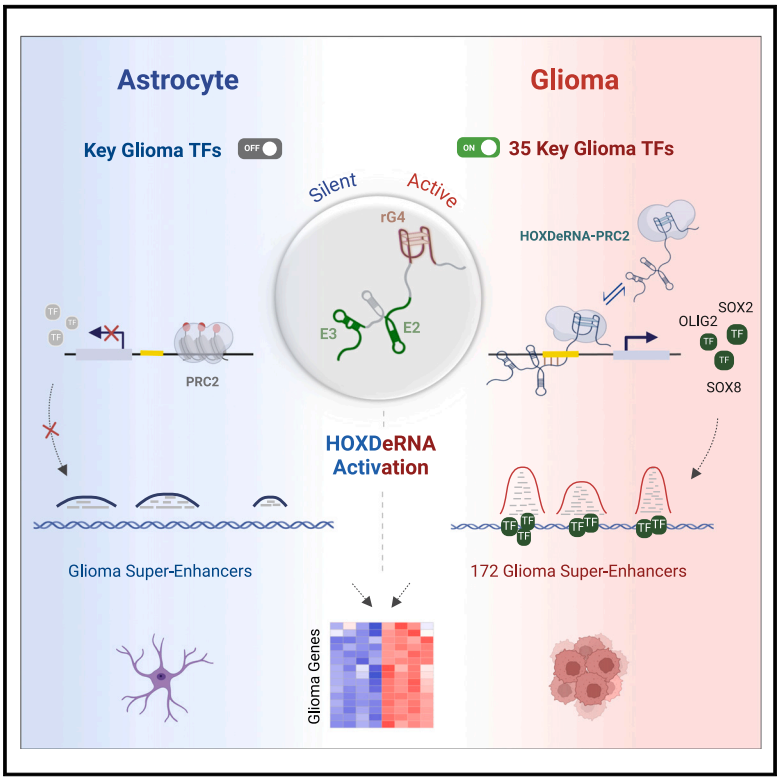
Molecular Cell

Article

HOXDeRNA activates a cancerous transcription program and super enhancers via genome-wide binding

Graphical abstract



Highlights

- HOXDeRNA transforms astrocytes into glioma-like stem cells
- HOXDeRNA directly binds to PRC2-covered CpG islands in the promoters of glioma driver genes
- HOXDeRNA removes EZH2 through its RNA quadruplex sequences

Authors

Evgeny Deforzh, Prakash Kharel, Yanhong Zhang, Anton Karelin, Abdellatif El Khayari, Pavel Ivanov, Anna M. Krichevsky

Correspondence

akrichevsky@bwh.harvard.edu

In brief

Deforzh et al. demonstrate the ability of the lncRNA HOXDeRNA to transform astrocytes into glioma-like stem cells by sequestering and protecting key glioma genes from Polycomb2 repression in an RNA quadruplex-dependent manner.

Deforzh et al., 2024, Molecular Cell *84*, 3950–3966
October 17, 2024
https://doi.org/10.1016/j.molcel.2024.09.018

CellPress

CellPress

Molecular Cell

Article

HOXDeRNA activates a cancerous transcription program and super enhancers via genome-wide binding

Evgeny Deforzh,[1] Prakash Kharel,[2] Yanhong Zhang,[1] Anton Karelin,[1] Abdellatif El Khayari,[1,3] Pavel Ivanov,[2] and Anna M. Krichevsky[1,4,*]

[1]Department of Neurology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA
[2]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA
[3]Institute of Biological Sciences (ISSB-P), UM6P Faculty of Medical Sciences, Mohammed VI Polytechnic University, Ben-Guerir 43150, Morocco
[4]Lead contact
*Correspondence: akrichevsky@bwh.harvard.edu
https://doi.org/10.1016/j.molcel.2024.09.018

SUMMARY

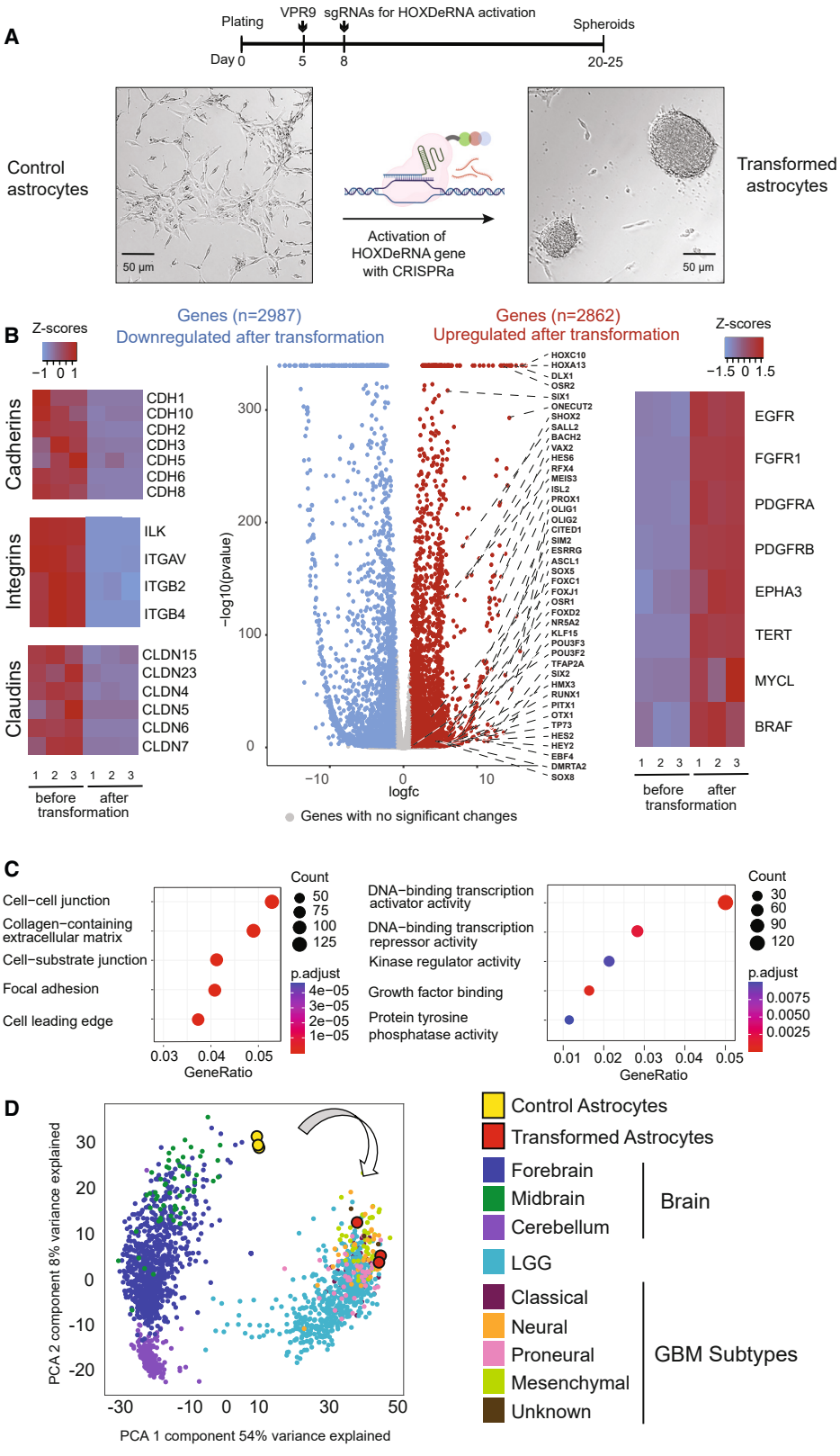
The role of long non-coding RNAs (lncRNAs) in malignant cell transformation remains elusive. We previously identified an enhancer-associated lncRNA, LINC01116 (named HOXDeRNA), as a transformative factor converting human astrocytes into glioma-like cells. Employing a combination of CRISPR editing, chromatin isolation by RNA purification coupled with sequencing (ChIRP-seq), *in situ* mapping RNA-genome interactions (iMARGI), chromatin immunoprecipitation sequencing (ChIP-seq), HiC, and RNA/DNA FISH, we found that HOXDeRNA directly binds to CpG islands within the promoters of 35 glioma-specific transcription factors (TFs) distributed throughout the genome, including key stem cell TFs SOX2, OLIG2, POU3F2, and ASCL1, liberating them from PRC2 repression. This process requires a distinct RNA quadruplex structure and other segments of HOXDeRNA, interacting with EZH2 and CpGs, respectively. Subsequent transformation activates multiple oncogenes (e.g., EGFR, miR-21, and WEE1), driven by the SOX2- and OLIG2-dependent glioma-specific super enhancers. These results help reconstruct the sequence of events underlying the process of astrocyte transformation, highlighting HOXDeRNA's central genome-wide activity and suggesting a shared RNA-dependent mechanism in otherwise heterogeneous and multifactorial gliomagenesis.

INTRODUCTION

Glioblastoma (GBM, or grade IV astrocytoma) is the most prevalent malignant primary tumor of the central nervous system in adults, with a median survival of 15 months. Despite substantial research, the molecular events underlying the transformation of normal cells to glioma-initiating cells and the development of this highly heterogeneous disease are poorly understood. Traditionally, these events have been examined within the context of mutational landscapes. Various combinations of mutations can transform terminally differentiated GFAP$^+$ astrocytes, leading to GBM formation *in vivo*. For example, loss-of-function mutations inactivating RB family,[1–3] CDKN2A,[4–6] NF1,[7,8] PTEN,[1,2,7,9] or TRP53[1,4,7–10] tumor suppressor genes, and activating mutations in EGFR,[5,11] HRAS,[10,12] KRAS,[6,13] or AKT1[13,14] play pivotal roles in this process.

On the other hand, it has been observed that multiple key glioma transcription factors (TFs) and other genes responsible for glioma cell identity are predominantly silenced in brain astrocytes, the cells of origin for glioma, and become activated in gliomagenesis.[15] The silent status of these genes depends on the Polycomb repressive complex 2 (PRC2) component EZH2, which catalyzes the tri-methylation of lysine 27 on histone H3 (H3K27Me3).[16] Gain and loss of PRC2 and H3K27Me3 marks on chromatin are associated, respectively, with gene repression and activation during carcinogenesis. In glioma, inhibition of PRC2 components reactivates PRC2 target genes, promoting tumor progression.[17] PRC2 is recruited to chromatin through CpG islands (CGIs) located mostly at the promoters,[18] and insertion of CGIs into the genome could be sufficient to induce PRC2 recruitment.[19–21]

PRC2 is also modulated by RNA.[22,23] EZH2 can bind multiple RNA species *in vitro* and *in vivo*.[24–27] Although the selectivity of EZH2-RNA binding is debatable,[28] essential regulatory functions of several long non-coding RNAs (lncRNAs), such as Xist and HOTAIR, were linked to their PRC2-binding capacity.[29,30] In line with the promiscuous RNA binding of PRC2, recent reports also suggested EZH2's high affinity and specificity for G-rich sequences and RNA quadruplex (rG4) structures formed by the stacking of two or more guanine quartets (rG4).[31,32] However,



(legend on next page)

the biological significance of lncRNA-EZH2 interactions, and particularly of rG4-EZH2 binding, has not been elucidated.

We have previously identified an enhancer-associated lncRNA, LINC01116 (named HOXDeRNA here), silenced in the normal cells of the brain and commonly activated in glioma.[33] High levels of this eRNA are strongly associated with glioma progression and shortened survival of GBM patients.[33–36] Targeted activation of this eRNA, but not of its neighboring LINC01117, induces astrocyte transformation accompanied by notable morphological changes. Our previous work described the local role of HOXDeRNA in controlling transcription of HOXD/miR-10b genes located within a 500 kb region, involving the regulation of CTCF/cohesin-dependent looping.[33] We now demonstrate the global chromatin remodeling and transcriptional reprogramming triggered by this eRNA. Using advanced techniques, we show that HOXDeRNA derepresses key glioma genes not only in *cis* but also in *trans* in a genome-wide manner. We describe a multifactorial molecular mechanism, which involves rG4-dependent recruitment of HOXDeRNA to PRC2-covered CGIs near transcription start sites (TSSs) of glioma driver genes, including key TFs, followed by the removal of PRC2 repression from the gene bodies and further activation of glioma super enhancers (SEs) and oncogenes. Targeted base editing of a specific rG4 structure in HOXDeRNA abrogates HOXDeRNA recruitment to chromatin and the removal of PRC2, repressing the glioma signature genes and transformation. Hence, the reported mechanism offers a perspective on carcinogenesis, presenting it as a pathology driven by regulatory RNA. This process entails epigenetic and transcriptional reprogramming, ultimately leading to neoplastic transformation, even in the absence of genetic alterations.

RESULTS

Activation of HOXDeRNA in astrocytes induces glioma transcriptional programs

We expanded prior findings[33] that HOXDeRNA (Ensembl: ENST00000295549) was not expressed in normal brain tissues, astrocytes, oligodendrocytes, and neural stem cells but was commonly transcribed in glioma and GBM tumor tissues, cells, and glioma stem cells (GSCs) using multiple datasets (Figures S1A and S1B). To investigate HOXDeRNA role in glioma biology, we employed human immortalized astrocytes[37–39] whose transcriptome mirrors that of primary cortical astrocytes (Figure S1C). As we showed previously, activation of HOXDeRNA in astrocytes using the CRISPR activation system promoted their transformation into glioma-like spheroids (Figure 1A).[33] Quantification of unspliced HOXDeRNA (pre-HOXDeRNA, 8,223 nucleotides) and mature spliced HOXDeRNA (1,058 nucleotides) using droplet digital PCR in these transformed astrocytes, similar to glioma cells and 3D GSC spheroid cultures, demonstrated the prevalence of the predominantly nuclear pre-HOXDeRNA form over the cytosolic mature form (Figure S1D; Table S1). The median copy number of pre-HOXDeRNA transcripts, measured with smFISH in different glioma cells and transformed astrocytes, ranges between 45 and 90 copies/cell, while absent in naive astrocytes (Figures S1E and S1F). Differential gene expression analysis of control astrocytes and the astrocytes transformed by HOXDeRNA activation revealed 2,862 upregulated and 2,987 downregulated genes in the transformed spheroids (FC > 2, $p < 0.05$) (Figures 1A and 1B). Gene Ontology (GO) analysis demonstrated that cell-cell junction, focal, and cell adhesion genes, including cadherins, claudins, and integrins, were overrepresented among the downregulated genes, suggesting a molecular basis for the observed phenotypic transition from adherent cells to spheroids (Figure 1C). Notably, the genes upregulated in transformed cells were related to transcriptional regulation, growth factors, and tyrosine kinases, which were known to be associated with glioma biology. Among the genes upregulated by HOXDeRNA activation were those encoding 35 key glioma TFs[15] (Figure 1B, central). RT-qPCR verified strong 500- to 1,000-fold upregulation of many representative TFs (Figure S1G). The list of TFs includes glioma-specific SOX2 and OLIG2 (Figure S2A).[15,40] Analysis of publicly available RNA sequencing (RNA-seq) and H3K27Me3 chromatin immunoprecipitation sequencing (ChIP-seq) datasets demonstrated that most of the 35 TFs controlled by HOXDeRNA are elevated in GBM compared with normal brains and also epigenetically silenced in normal brains, brain organoids, and astrocytes compared with GBM cells of different molecular subtypes (classical, proneural, and mesenchymal; Figures S2B and S2C). In addition, multiple key factors frequently mutated and/or activated in gliomas, including the established therapeutic targets such as EGFR, PDGFRA, BRAF, and TERT, were upregulated by HOXDeRNA (Figure 1B). These groups of genes define transcriptional programs and glioma-like identity of the transformed astrocytes. To explore whether the identified differentially expressed genes can distinguish GBM from normal brain tissues, we integrated transcriptomic profiles of normal and HOXDeRNA-transformed astrocytes with TCGA GBM datasets. A gradient boosting model was trained on

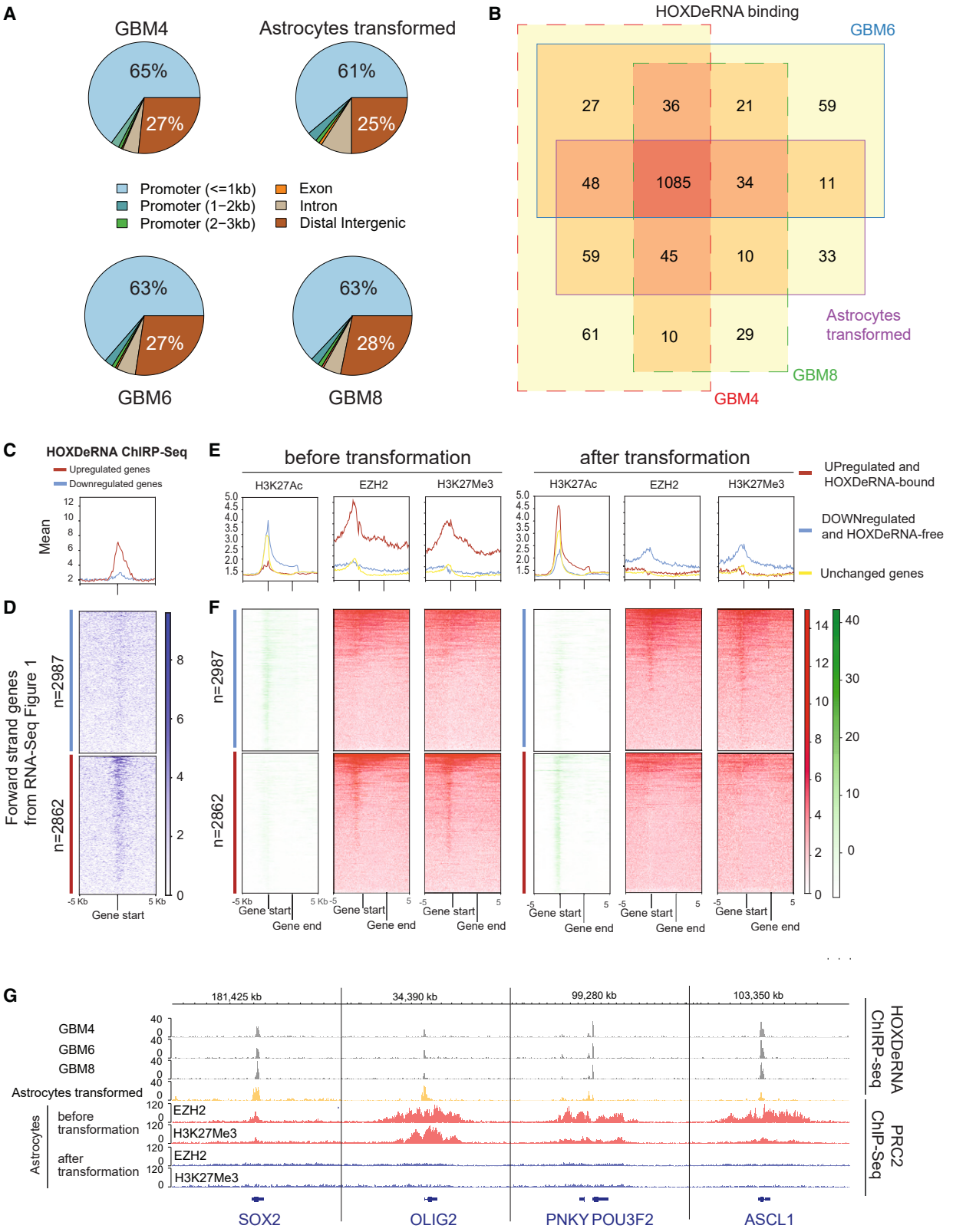
Figure 1. Targeted activation of the HOXDeRNA leads to astrocyte transformation with phenotypic and transcriptomic switch to glioma

(A) Timeline for transduction of astrocytes with the CRISPR activation system leading to transformation (top). Representative images of astrocytes transduced with non-targeting sgRNA (control astrocytes) and HOXDeRNA-activating sgRNA (HOXDeRNA-activated astrocytes) are shown.

(B) Volcano plot (middle) showing differentially expressed genes (DEGs) between the non-transformed and transformed astrocytes. The genes upregulated in transformed astrocytes (red), downregulated in transformed astrocytes (blue) (fold change > 2 and adjusted $p = 0.05$), and 10% of non-significantly changed genes (gray) are shown. The 35 glioma-specific TFs that are upregulated after astrocyte transformation are indicated. The heatmaps exhibit Z scores for cell junction and cell adhesion genes downregulated in transformed astrocytes (left, $n = 3$) and major glioma-associated genes upregulated in transformed astrocytes (right, $n = 3$).

(C) The top 5 categories of GO gene sets downregulated (left) or upregulated (right) in transformed astrocytes shown for DEG (FC > 2, $p < 0.01$).

(D) Control and transformed astrocytes are associated with "normal brain" and "glioma" expression signatures, respectively. Machine learning model trained on forebrain ($n = 857$ samples), cerebellum ($n = 214$), midbrain ($n = 57$), low-grade glioma (LGG, $n = 522$), mesenchymal ($n = 54$), classical ($n = 41$), proneural ($n = 39$), neural ($n = 28$), and unknown subtype ($n = 4$) GBM samples from TCGA and GTEX datasets classifies control and transformed astrocytes as normal brain and glioma, correspondingly, based on our RNA-seq data. PCA visualization is shown, and every dot represents an individual sample (see STAR Methods for details).



(legend on next page)

differentially expressed genes from our RNA-seq data (see STAR Methods), and the results were visualized as principal-component analysis (PCA) maps. They demonstrated the respective association of control and transformed astrocytes with the normal brain and glioma (Figure 1D), confirming that HOXDeRNA drives astrocytes into a glioma state.

HOXDeRNA binds and evicts PRC2 from genes encoding key glioma-specific TFs in transformed astrocytes

The RNA-seq data showing that HOXDeRNA activation changed the expression of thousands of genes suggested a genome-wide regulatory activity of HOXDeRNA via chromatin binding. To investigate this scenario, we employed the chromatin isolation by RNA purification coupled with sequencing (ChIRP-seq) technique, which captures interactions between the RNA of interest and chromatin, using the probes for the unspliced pre-HOXDeRNA. The binding peaks were called using input control for each cell line, and their quality was assessed with negative controls (see STAR Methods). The peaks of pre-HOXDeRNA binding in transformed astrocytes and GSCs were highly consistent between biological replicates, with Pearson correlation higher than 0.8 and at least 70% common peaks (Table S2). We detected 1,085 common chromatin binding sites for HOXDeRNA in transformed astrocytes and three GSC lines (Figures 2A and 2B). Genomic distributions of HOXDeRNA binding sites in transformed astrocytes and GSCs were similar, with 61%–65% binding events observed in gene promoters. They were mapped to all chromosomes (Figure S3A), suggesting a global role for HOXDeRNA in gene transcription. Notably, HOXDeRNA bound almost exclusively to the promoters of genes that were upregulated in transformed astrocytes but not the genes whose expression was downregulated (Figures 2C, 2D, S3B, and S3C). The corresponding gene lists (Table S3) were obtained from the differential gene expression analysis shown in Figure 1B (volcano plot). HOXDeRNA binding to the promoters of representative TFs was verified by ChIRP-qPCR (Figure S3F). Of note, HOXDeRNA ChIRP-seq also revealed substantial binding in intergenic regions. This phenomenon, consistently observed in other ChIRP-seq datasets and frequently exceeding 30% (see Table S4), may be attributed to the off-target binding of multiple probes utilized in ChIRP experiments or other sources of technical bias.

We next investigated the epigenetic status of HOXDeRNA-occupied genes in astrocytes, using H3K27Ac as the mark of active chromatin state and H3K27Me3 and EZH2 (PRC2) as the marks of repressed chromatin. Notably, the genes bound by HOXDeRNA and upregulated after astrocyte transformation were occupied by EZH2/H3K27Me3 and depleted of the H3K27Ac in naive NHA cells. The PRC2 coverage of these genes decreased, and H3K27Ac coverage increased after HOXDeRNA-induced transformation (Figures 2E, 2F, S3D, and S3E). This contrasted with the unchanged and downregulated genes not bound by HOXDeRNA and not exhibiting such a trend (Figures 2E, 2F, S3D, and S3E). Overall, the differential EZH2 binding analysis revealed 762 and 388 regions enriched in naive and transformed astrocytes, respectively (Table S5). Importantly, genes encoding the 33 key glioma-specific TFs were bound by HOXDeRNA and lost their PRC2 marks after cell transformation (Figures 2G, S3G–S3J, and S4). We highlight this effect using representative key TFs SOX2, OLIG2, POU3F2, ASCL1, and others, which genes were bound by HOXDeRNA in three GSCs and transformed astrocytes (Figures 2G and S4: gray and yellow tracks), and HOXDeRNA binding in transformed astrocytes was associated with the reduced PRC2 coverage (Figures 2G and S4: red and blue tracks, respectively; Figure S3K). Altogether, these data suggested that HOXDeRNA bound to glioma-specific genes and promoted their transcription by titrating out the PRC2 silencing complex.

HOXDeRNA selectively engages with CGIs of genes upregulated in transformed astrocytes

To further investigate HOXDeRNA interactions with chromatin in transformed astrocytes and GSCs, we performed *in situ* mapping RNA-genome interactions (iMARGI).[41] This technique, designed to discover chromatin-associated RNA interactions at single-nucleotide resolution, was further optimized to detect direct RNA-DNA binding by decrosslinking chromatin *in situ* immediately after RNA and DNA fragmentation before adapter ligation (see STAR Methods). We filtered out all reads containing multi-mapped RNA or DNA segments, ensuring that only uniquely mapped read segments, each corresponding to a single specific gene, were used in the subsequent analysis. HOXDeRNA-interacting genomic DNA segments were extracted, and a list of common segments for all samples was generated. Similarly to ChIRP, iMARGI revealed genome-wide HOXDeRNA binding (Figures S5A and S5B), with exclusive interactions detected between its exons 2 and 3 and genes upregulated after astrocyte transformation (Figure 3A; Table S6).

Figure 2. Genome-wide binding of HOXDeRNA is associated with exclusive PRC2 removal from transformation-induced genes

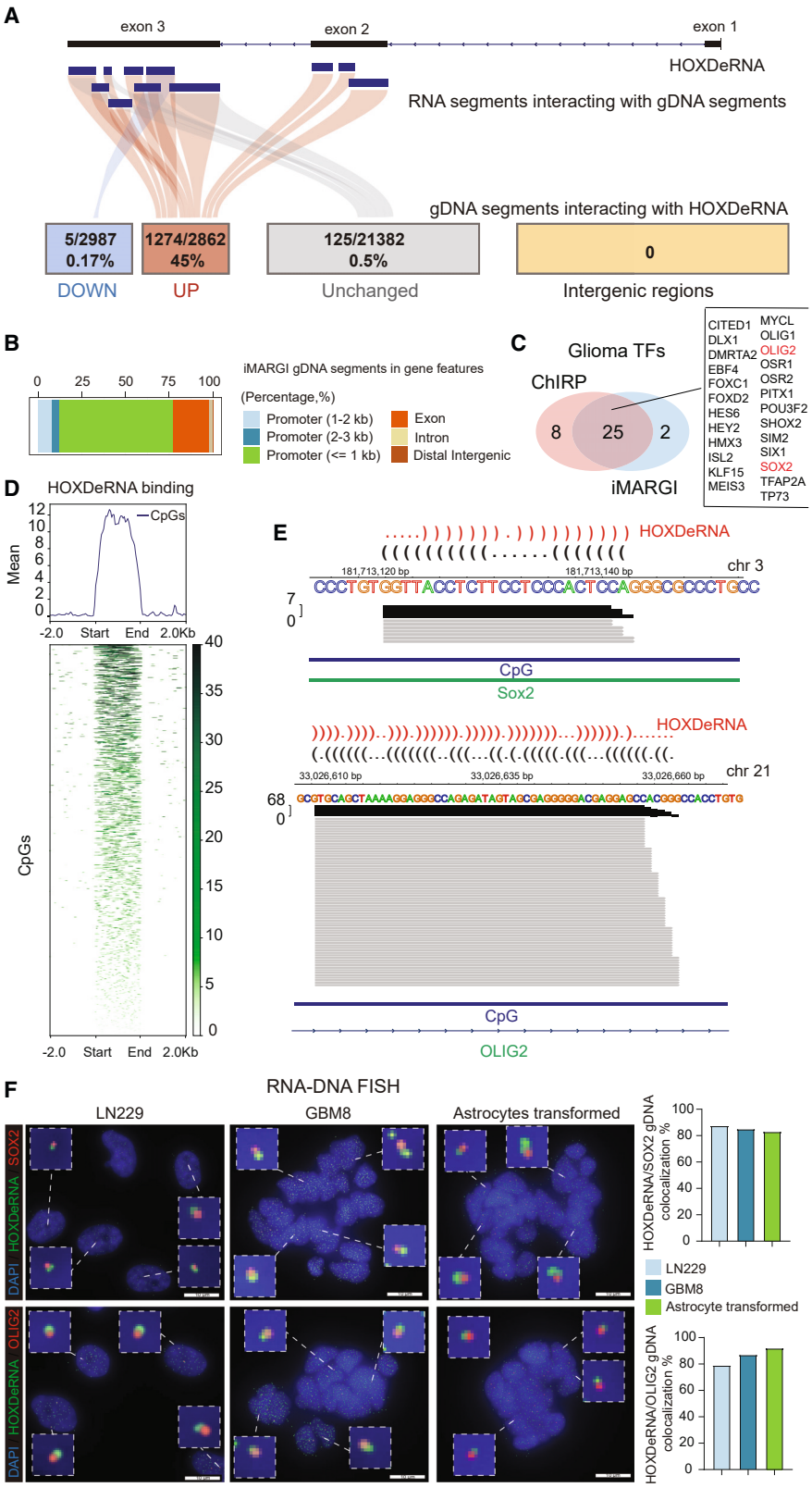
(A) ChIRP-seq analysis demonstrates that transformed astrocytes and three GSC lines (GBM4, GBM6, and GBM8) exhibit similar distribution of HOXDeRNA peaks across the genome, with most peaks mapped to gene promoters.

(B) HOXDeRNA binds to the same gene promoters in transformed astrocytes and GSCs. HOXDeRNA ChIRP peaks were annotated to the nearest genes, and the gene lists produced for the four cell types were intersected and visualized as a Venn diagram (see STAR Methods for details).

(C and D) ChIRP-seq raw read coverage signal, representing the HOXDeRNA binding at the TSS (±5 kb) of the forward strand of genes either downregulated or upregulated after astrocyte transformation, is visualized as an average value for each group (C) or for individual genes (D). The heatmap is accompanied by a color scheme representing the value of the raw read counts. A similar coverage for reverse strand genes is shown in Figures S3B and S3C.

(E and F) Epigenetic status of genes upregulated or downregulated after astrocyte transformation. H3K27Ac, H3K27Me3, and EZH2 ChIP-seq raw signals covering gene bodies of the positive strand (±5 kb) were normalized to gene length and counted, followed by visualization of the average signal for 3 groups: genes upregulated after astrocyte transformation and bound by HOXDeRNA, downregulated after astrocyte transformation and HOXDeRNA-free, and unchanged genes (E). Individual gene body coverage values were visualized as heatmaps (F) (see reverse strand gene coverage in Figures S3D and S3E).

(G) HOXDeRNA ChIRP-seq tracks in GSCs and transformed astrocytes, aligned with PRC2 (H3K27Me3, EZH2) ChIP-seq coverage, before and after astrocyte transformation, visualized for selected key glioma TF genes.



(legend on next page)

Minimal HOXDeRNA interactions were detected with downregulated or unchanged genes, as well as with other genomic regions. HOXDeRNA-interacting genomic segments were highly enriched among promoters (~80%) and exonic regions (~15%) of protein-coding genes (Figures 3B and S5C). Both ChIRP and iMARGI demonstrated HOXDeRNA binding to hundreds of genes (Table S7), including those encoding 25 glioma-specific TFs, such as SOX2 and OLIG2, whose promoters directly interact with HOXDeRNA (Figure 3C). Furthermore, iMARGI showed that 95% of HOXDeRNA-interacting DNA segments intersected with CGIs of upregulated genes after astrocyte transformation (Figure 3D). Analysis of these interactions suggested partial HOXDeRNA-DNA complementarity (Table S6), as depicted for SOX2 and OLIG2 promoters in Figure 3E. Further validating these interactions, combined RNA-DNA FISH imaging detected frequent colocalization of HOXDeRNA at SOX2 and OLIG2 promoters (Figure 3F).

Astrocyte transformation involves activation of key GSC SEs

Previous reports suggested that TFs, including SOX2 and OLIG2, may regulate SEs.[42,43] We next analyzed the H3K27Ac enhancer chromatin mark in glioma primary tumors (n = 35) and GSCs (n = 43)[42] and compiled lists of typical glioma enhancers (TEs) and SE. The astrocyte transformation induced by HOXDeRNA resulted in a significant increase in H3K27Ac coverage of the SEs, signifying their activation (Figure 4A, left), while the coverage of TEs remained constant (Figure 4A, right). SEs activated after transformation were associated with critical protein-coding and non-coding oncogenes such as SOX2, WEE1, EGFR, and miR-21 (Figure 4B). As SEs can be activated by TF binding (reviewed by Grosveld et al.[44] and Bruter et al.[45]), we investigated TF motifs in the H3K27Ac-covered regions of the activated SEs. The binding motifs for OLIG2 and SOX2, two TF marks of GBM subtypes induced by HOXDeRNA and upregulated after transformation, were highly enriched (Figure 4C), suggesting their regulation of numerous SEs. A comprehensive analysis of SOX2 and OLIG2 ChIP-seq datasets from GSCs (see STAR Methods) revealed that 82% of the SEs were occupied by SOX2, OLIG2, or both (Figure 4D). Confirming this, the double knockdown (KD) of SOX2 and OLIG2 following astrocyte transformation substantially reduced the activity of glioma SEs, indicating that SEs are regulated by these TFs (Figures 4E and 4F). Moreover, KD of SOX2 and OLIG2 in astrocytes prior to HOXDeRNA activation reduced spheroid formation, demonstrating that these TFs are essential for SE activation and astrocyte transformation (Figure 4G).

We hypothesized that activation of glioma SEs mediated many of the HOXDeRNA downstream effects on gene expression. To identify the genes regulated by SEs, we called the loops from available glioma HiC datasets,[46] merged them, and annotated the loops' anchors. Multiple genes (including those encoding TFs) within the loops of the activated SEs exhibited reduced mRNA expression after SOX2 and OLIG2 KD in transformed astrocytes, reverting to the levels observed in naive astrocytes (Figure 4H). To further investigate the effects of SE activation on the coordinated transcriptional programs in glioma, we integrated our RNA-seq and H3K27Ac datasets with the CTCF ChIP-seq and HiC datasets. We hypothesized that SEs might concordantly activate clustered genes located within large topologically associated domains (TADs). We searched for such clusters exhibiting concordant transcriptional activation after astrocyte transformation and located within the same TADs in glioma. We found that HOXDeRNA transformation-induced protocadherin (PCDH) genes were characterized by genomic interactions with glioma-specific SEs (see the legend for Figure S6 for the list of tracks and accession numbers). Multiple PCDHB genes located in the same glioma TAD with HOXDeRNA-activated SEs were concordantly upregulated after transformation (Figure S6, dotted rectangles mark specific HiC interactions enriched in GSCs compared with astrocytes, and glioma SEs are depicted in green). Moreover, these interactions appeared dependent on a cluster of CTCF sites bound by CTCF proteins in glioma cells but not in astrocytes (CTCF tracks, regions highlighted in yellow). The mRNA expression of all PCDH genes located in the indicated TAD depends on SOX2/OLIG2 double KD (Figure 4H).

Of note, ChIRP and iMARGI data analysis indicated that none of the glioma SEs was directly bound by HOXDeRNA, reinforcing that their activation is downstream of the TFs such as SOX2 and OLIG2. Our findings align with the paradigm of a self-regulatory transcriptional network represented by key TFs and SEs (Figure 4I) and define those critical factors in GBM.

Figure 3. HOXDeRNA interacts directly and exclusively with CGIs of genes upregulated after astrocyte transformation

(A) HOXDeRNA displays exclusive interactions with genes that are upregulated following astrocyte transformation, as revealed by iMARGI (see STAR Methods). The schematic representation of HOXDeRNA transcript is depicted at the top (in black). The Sankey plot illustrates interactions between various HOXDeRNA segments (blue bars) in exons 2 and 3, and groups of genes either downregulated, upregulated, or unchanged in transformed astrocytes, or intergenic regions. The numerators indicate the numbers of HOXDeRNA-interacting genes, and the denominators indicate a total number of genes in each respective category.

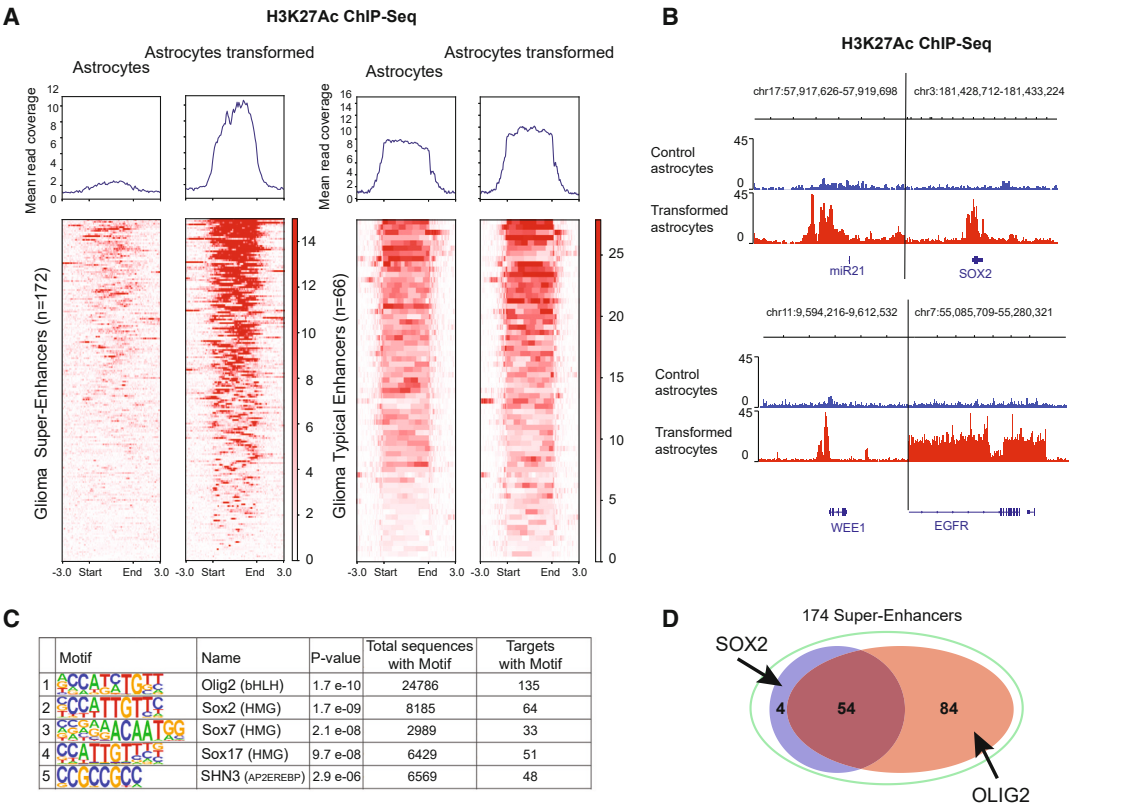
(B) HOXDeRNA predominantly interacts with gene promoters (~80%) and exons (~15%) of the upregulated genes. Genomic DNA segments interacting with HOXDeRNA were categorized based on gene features and visualized as percentage using a stacked bar graph.

(C) Venn diagram demonstrates 25 out of 35 glioma TF genes identified by both ChIRP-seq and iMARGI-seq.

(D) HOXDeRNA exhibits exclusive and direct interactions with CGIs of genes upregulated with astrocyte transformation. Genomic DNA segment coverage within CGIs and surrounding sequences (±2 kb of CGIs) was computed as an average (top line graph) and individual values (bottom heatmap).

(E) Visualization of HOXDeRNA-interacting DNA segments in the promoters of SOX2 and OLIG2 genes. The tracks, from top to bottom, include HOXDeRNA (depicted in red) and gDNA (in black) interactions presented in dot-bracket notation; genomic coordinates of the visualization window; reference gDNA sequence of the visualization window; gDNA nucleotide coverage (in black) with values on the y axis; individual gDNA segments (in gray); CGI (depicted as a blue bar); and a portion of the corresponding gene (in green).

(F) Combined RNA-DNA FISH shows HOXDeRNA (green) co-localizing with OLIG2 (bottom, red) on Chr.21 and SOX2 (top, red) on Chr.3, in glioma cells and transformed astrocytes, with quantification of SOX2 and OLIG2 loci colocalization with HOXDeRNA.



(legend on next page)

HOXDeRNA is globally recruited to chromatin by EZH2

Since genome-wide binding of HOXDeRNA was associated with the removal of PRC2 from essential glioma genes upon astrocyte transformation, we conducted a ChIRP-western blot to investigate whether HOXDeRNA directly binds to EZH2. The results demonstrated a direct binding between HOXDeRNA and EZH2 (Figure 5A). Simultaneously, crosslinking and immunoprecipitation (CLIP) experiments showed that EZH2 exhibited high enrichment for pre-HOXDeRNA but not for mature HOXDeRNA (Figure 5B). Other members of PRC2 family (EED, SUZ12, and RbAp46) did not bind HOXDeRNA (Figure S7A). We, thus, hypothesized that EZH2 might recruit nuclear HOXDeRNA to chromatin. We identified 1,453 EZH2-covered and 166 EZH2-free genomic regions in astrocytes that gain binding of HOXDeRNA after transformation. We then knocked down EZH2 by small interfering RNA (siRNA) (Figure 5C), verified its effective removal from chromatin (Figure S7B), and examined HOXDeRNA binding in these regions using ChIRP-seq. EZH2 KD reduced HOXDeRNA binding at EZH2-covered genomic regions but did not affect HOXDeRNA binding at EZH2-free regions (Figure 5D, left). Notably, the binding of HOXDeRNA to the TSS of key glioma TFs, such as SOX2, OLIG2, and others (all repressed by PRC2 in astrocytes) disappeared following the EZH2 KD in GSCs and transformed astrocytes (Figure 5E). Conversely, overexpression of EZH2 in transformed astrocytes increased its genome-wide chromatin occupancy (Figure S7C) and HOXDeRNA-DNA binding (Figure 5D, right). These data indicate that EZH2 plays a critical role in recruiting HOXDeRNA to its chromatin binding sites, suggesting that the HOXDeRNA-EZH2 balance may determine the activation status of target genes.

We further monitored the kinetics of HOXDeRNA-induced EZH2 displacement and, in parallel, transcriptional activation of target genes, specifically focusing on the representative key glioma TFs SOX2, OLIG2, and POU3F2. Through a combination of targeted EZH2 ChIP, HOXDeRNA ChIRP, and RNA-seq at 8–24 h post-HOXDeRNA CRISPRa in astrocytes, we observed a gradual rise in HOXDeRNA binding, accompanied by transcriptional activation, and closely associated with the removal of EZH2, initiated at the 8-h mark (Figures 5F, S8A, and S8B). The siRNA-mediated KD of pre-HOXDeRNA restored EZH2 binding and gene repression. By contrast, the KD of the mature HOXDeRNA under the same conditions failed to rescue these effects (Figures 5F, S8A, and S8B).

Binding of HOXDeRNA to EZH2 depends on the rG4 structure

To explore whether EZH2 interacted with HOXDeRNA through an rG4 structure, as reported for its interaction with other RNAs,[31,32] we initially examined the ability of HOXDeRNA to form rG4s structures. The analysis using the QGRS mapper tool (https://bioinformatics.ramapo.edu/QGRS/index.php) yielded 5 putative rG4-forming sequences (Figures S9A and S9B). We then employed circular dichroism (CD) spectroscopy to assess the ability of these predicted G-rich HOXDeRNA motifs to fold into rG4 structures. An increased CD peak of the RNA oligonucleotides at 263 nm in the rG4-favoring K^+ environment and a reduction in the corresponding peak intensity in the rG4-unfavoring Li^+ environment suggest the formation of an rG4.[47] Among the tested sequences, three oligonucleotides (rG4-1, rG4-2, and rG4-4) showed a typical rG4 CD behavior, suggesting the highly probable formation of the rG4 structures in the cells (Figure S9C). To further check the stability of these rG4 structures, we used CD melting approach,[48] which identified rG4-1 as the most stable structure with a Tm of 57.5°C (Figure S9D). We then tested the ability of potential HOXDeRNA rG4 structures to bind to EZH2 in glioma cells and found that rG4-1 exhibited the highest affinity for EZH2 (Figure 6A). These data suggested that EZH2 binding to HOXDeRNA was mediated by the rG4-1 structure.

To explore the role of rG4-1 in HOXDeRNA function, one of the two G-nucleotide stretches forming the rG4-1 structure was edited using CRISPR base editing[49] (Figure 6B). We tested whether these G-to-A substitutions, which abolished rG4-1 within HOXDeRNA, impacted its binding to EZH2 and its transformative effects on astrocytes. Indeed, the ablation of rG4-1 reduced HOXDeRNA binding to EZH2 about 20-fold (Figure 6C) without affecting its expression (Figure S9E). As a downstream readout, we selected a panel of marker genes repressed by PRC2 in astrocytes but bound by HOXDeRNA and transcriptionally activated after transformation. The results indicated that the ablation of rG4-1 eliminated

Figure 4. Induction of HOXDeRNA activates GSC-specific SEs

(A) Astrocyte transformation induced by HOXDeRNA activates SEs but not TEs. Monitoring raw H3K27Ac coverage at glioma-specific SEs (on the left) and TEs (on the right) in naive and HOXDeRNA-activated astrocytes. Enhancers were divided into an identical number of bins for length normalization. The line graphs (top) or heatmaps (bottom) represent the average read coverage value per bin across all enhancers or the individual value per bin, respectively. The enhancers' margins are defined by the "start" and "end" marks.

(B) H3K27Ac ChIP-seq signals are shown for representative glioma SEs in control and transformed astrocytes.

(C) The top 5 TF binding motifs enriched in the SE-associated H3K27Ac peaks are displayed. TF enrichment analysis was conducted using Homer software.

(D) ChIP-seq demonstrates that SOX2, OLIG2, or both bind to 82% of GBM SEs. The Venn diagram visualizes the intersection of the lists of SEs bound by SOX2 or OLIG2.

(E) Knockdown of SOX2 and OLIG2 results in the loss of glioma SEs activity marked by H3K27Ac. Transfected transformed astrocytes with control siRNAs or a combination of SOX2 and OLIG2 siRNAs underwent H3K27Ac ChIP-seq. The visualized H3K27Ac read coverage is presented as an average (line graph) or individual values (heatmap) at glioma SEs. The enhancers' margins are defined by the "start" and "end" marks.

(F) H3K27Ac ChIP-seq signals display representative glioma SEs in transformed astrocytes transfected with control or a mix of SOX2 and OLIG2 siRNAs (red and blue tracks, respectively).

(G) Double KD of SOX2 and OLIG2 48 h prior to HOXDeRNA activation reduces astrocyte transformation. The numbers of spheroids were quantified in control (n = 3, blue) and siRNA KD (n = 3, red) conditions (mean + SD).

(H) KD of SOX2 and OLIG2 reduces the expression of genes associated with glioma SEs. Naive astrocytes and transformed astrocytes transfected with control or a mix of SOX2 and OLIG2 siRNAs underwent RNA-seq. The count matrix was MRN-scaled and transformed to Z scores by rows for heatmap visualization.

(I) A diagram illustrates the glioma self-regulatory transcriptional network represented by key TFs and SEs.

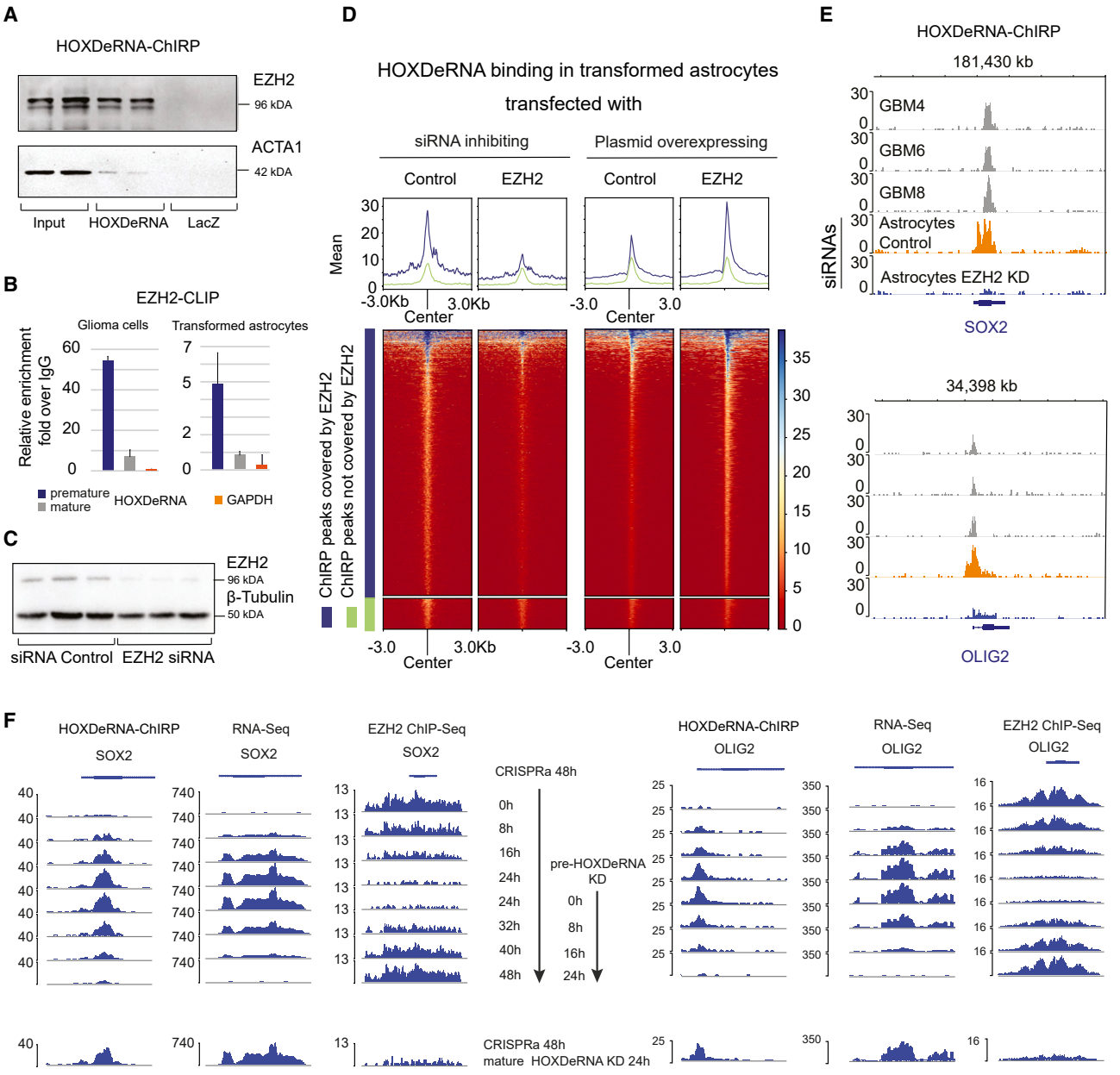


Figure 5. HOXDeRNA genome-wide binding depends on EZH2

(A) ChIRP with probes for HOXDeRNA and LacZ (negative control) was conducted on transformed astrocytes, followed by western blots with antibodies recognizing EZH2 and ACTA1, visualized with 1% input. Two representative biological replicates per group are shown.

(B) CLIP with EZH2 and IgG antibodies, followed by RT-qPCR detection of pre-HOXDeRNA, spliced HOXDeRNA, and GAPDH mRNA, was performed in glioma LN229 cells and transformed astrocytes. EZH2/IgG ratios are demonstrated (n = 3, mean + SD).

(C) Western blot validating EZH2 inhibition in transformed astrocytes at 48 h post-transfection with EZH2 siRNAs (n = 3).

(D) HOXDeRNA ChIRP signals were measured in transformed astrocytes transfected with either EZH2-inhibiting siRNAs (left) or EZH2-overexpressing plasmid (right). These signals were visualized as an average (line graph, top) or individual values (heatmap, bottom) at HOXDeRNA peaks (center ± 3 kb). HOXDeRNA binding was analyzed separately for the peaks covered (blue) or not covered by EZH2 (green) in control astrocytes.

(E) HOXDeRNA binding at the promoters of representative glioma TF genes is shown for GSCs (top three tracks) and transformed astrocytes transfected with either control siRNAs or EZH2 siRNAs (two bottom tracks).

(F) The kinetics of HOXDeRNA and EZH2 binding to the SOX2 and OLIG2 genes are tightly associated with the genes' expression in transformed astrocytes. After HOXDeRNA activation with the CRISPRa system, NHA cells were collected at different time points, followed by targeted HOXDeRNA ChIRP-seq, RNA-seq, and EZH2 ChIP-seq. At 24 h after CRISPRa, the cells were transfected with specific siRNAs against either pre-HOXDeRNA or mature HOXDeRNA and analyzed in similar ways.

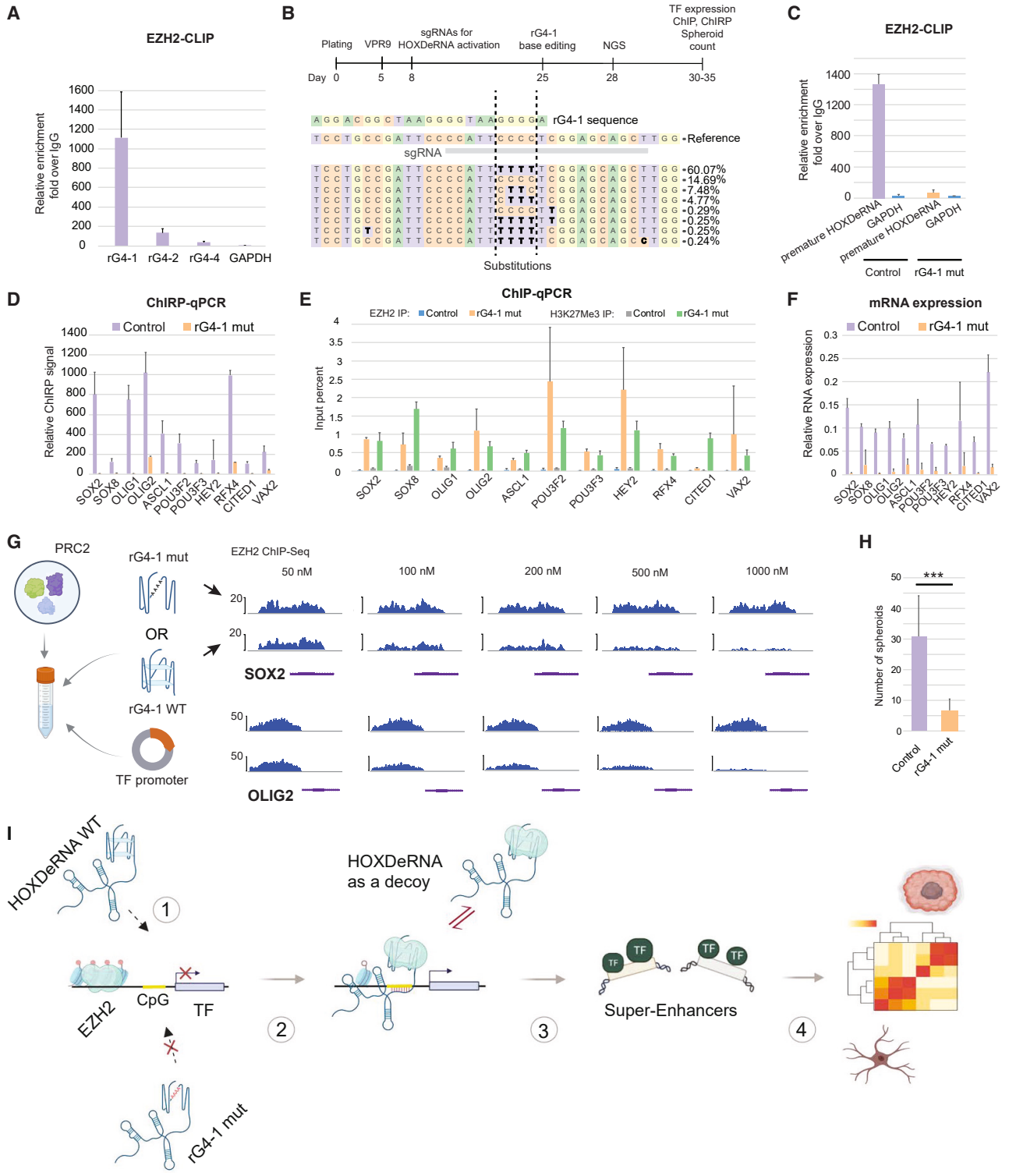


Figure 6. A specific RNA quadruplex rG4-1 element in HOXDeRNA mediates its EZH2 binding and PRC2 removal, global regulatory activity, and transformation capacity

(A) CLIP with EZH2 antibody in glioma LN229 cells, followed by RT-qPCR detection of three putative HOXDeRNA rG4-containing regions, using GAPDH mRNA as a negative control (mean + SD, *n* = 3).

(legend continued on next page)

HOXDeRNA binding to the gene promoters and led to the corresponding increase in EZH2 and H3K27Me3 marks, ultimately hindering the expression of the genes (Figures 6D–6F).

Finally, we used an *in vitro* titration system to test the effect of HOXDeRNA rG4-1 on PRC2 displacement from gene promoters directly. Recombinant PRC2 complex (composed of EZH2, EED, and Suz12 subunits) was mixed with plasmids containing the promoter regions of SOX2 and OLIG2, along with increasing concentrations of either HOXDeRNA wild-type (WT) rG4-1 or rG4-1-"mutated" oligonucleotides. Our findings revealed a concentration-dependent eviction of EZH2 by the WT rG4-1 but not its mutated variant (Figure 6G). Consistent with these observations, astrocytes deficient in HOXDeRNA rG4-1 exhibited a reduced transformation capacity, as evidenced by matrix-independent spheroid growth (Figure 6H). Notably, the activation of HOXDeRNA in astrocytes, followed by the selective KD of either mature HOXDeRNA or pre-HOXDeRNA, resulted in distinct phenotypes. While the KD of pre-HOXDeRNA reduced spheroid formation, the KD of the mature cytosolic form had no impact (Figure S8C), further validating the essential role of nuclear, chromatin-associated HOXDeRNA in the transformation of human astrocytes.

We propose a model in which HOXDeRNA is recruited to crucial glioma genes in an rG4-1/EZH2 and CpG-dependent manner, maintaining their activity by displacing PRC2 (Figure 6I). The specificity of HOXDeRNA binding to these genes is determined by two factors: PRC2 coverage and the presence of CGIs. HOXDeRNA regulates a dynamic balance between activating and repressive molecular mechanisms, preventing the accumulation of PRC2 in transformed astrocytes and glioma cells to levels critical for gene silencing. This process triggers the activation of multiple oncogenes and SEs, establishing and reinforcing glioma transcriptional program.

DISCUSSION

lncRNAs and particularly eRNAs are associated with chromatin-modifying complexes and may affect gene expression.[50,51] Nevertheless, only a few lncRNAs have been functionally implicated in human physiology and pathology.[52–62] Here, we report a broad genome-wide function of a specific poorly conserved eRNA, HOXDeRNA, that upon its activation in astroglia, binds to and derepresses promoters of genes encoding key glioma TFs scattered throughout the genome. These TFs in turn activate multiple glioma-specific SEs and their associated target genes, including protein-coding, miRNAs, and lncRNAs, involved in the malignant phenotype of glioma. Among them are the major glioma factors such as EGFR, PDGFR, TERT, miR-21, miR-10b, and HOXD-AS2. This study, therefore, provides a detailed description of how lncRNA is recruited to its gene targets genome-wide through an rG4/EZH2 and CpG-dependent mechanism and drives a cancer-specific transcriptional program by alleviating PRC2 repression from critical components of glial regulatory circuitry.

Key TFs, SEs, and their gene targets dictate cell fates in normal and disease states. Overexpression of a few TFs of pluripotency induces significant transcriptomic alterations, cell reprogramming, and transdifferentiation.[63,64] A small number of core GSC TFs (SOX2, OLIG2, POU3F2, and SALL2) appears sufficient for the reprogramming of differentiated glioma cells into a cell population recapitulating tumor-propagating patient-derived GSCs.[40] Remarkably, our data indicate that activation of HOXDeRNA in astrocytes leads to its direct binding to promoters of four core and 31 additional glioma-specific TFs,[15] all silenced in astrocytes by PRC2, and derepresses their expression. There are two distinct groups of TFs (i.e., OLIG1/2 and RUNX1/2/TFAP2A) that define proneural and mesenchymal GBM subtypes, respectively.[42,65] Both are directly activated by HOXDeRNA, suggesting its involvement in both proneural and mesenchymal transcriptional programs. Although HOXDeRNA exhibits local activity on HOXD TFs[33] and global genome-wide function, the comparative kinetics of HOXD and SOX2, OLIG2, and POU3F2 genes' upregulation indicate that local effects do not precede the global ones (Figures 5F, S8B, and S10). Furthermore, OLIG2 and SOX2 binding motifs are prevalent in HOXDeRNA-activated SEs, many of which control additional

(B) Schematic timeline for rG4-1 base editing experiments in transformed astrocytes (top). DNA analysis confirming efficient C-to-T editing in the HOXDeRNA rG4-1 genomic region, corresponding to the G-to-A editing in the HOXDeRNA. Alleles with a substitution rate of >0.1% are visualized. Substituted nucleotides are shown in bold.

(C) CLIP analysis of WT and rG4-1-edited transformed astrocytes with EZH2 antibody, followed by RT-qPCR for HOXDeRNA and GAPDH mRNA (mean + SD, $n = 3$).

(D) rG4-1 base editing abolishes the binding of HOXDeRNA to its targets in transformed astrocytes. ChIRP-qPCR analysis of HOXDeRNA/gene promoter binding was performed in control ($n = 3$) and base-edited transformed astrocytes ($n = 3$). The data were normalized to the GAPDH gene and presented as bars (mean + SD).

(E) rG4-1 base editing prevents the removal of PRC2 from HOXDeRNA targets after its activation. ChIP-qPCR reactions with EZH2 and H3K27Me3 antibodies on control ($n = 3$) and edited clones ($n = 3$) were normalized for input and shown as bars (mean + SD).

(F) rG4-1 base editing disrupts the derepression of HOXDeRNA target genes after astrocyte transformation. RT-qPCR analysis of the corresponding set of HOXDeRNA-induced target mRNAs was performed in both control ($n = 3$) and base-edited ($n = 3$) transformed astrocytes and normalized to GAPDH mRNA levels. The data are shown as bars (mean + SD).

(G) PRC2 complex, composed of recombinant EZH2, EED, and Suz12 proteins, was mixed with the SOX2 and OLIG2 reporter plasmids and increasing concentrations of either WT rG4-1 or mutated rG4-1 oligonucleotides and incubated for 1 h. EZH2 coverage was analyzed at the promoters of SOX2 and OLIG2 by ChIP-seq.

(H) rG4-1 base editing inhibits astrocyte transformation. The numbers of spheroids were quantified in both control ($n = 3$) and base-edited ($n = 3$) transformed astrocytes (mean + SD).

(I) A model of genome-wide function of HOXDeRNA. rG4-dependent HOXDeRNA binding to EZH2 and direct recruitment to PRC2-silenced CGIs of promoters in astrocytes (1) leads to reduced PRC2 repression in the corresponding chromatin regions, gene derepression, and active state of key glioma TFs (2), followed by SE activation (3) and further transcriptional reprogramming (4).

glioma TFs and oncogenes (e.g., EGFR, BRD4, and miR-21). These data support the idea of the reciprocal relationship between TFs and SEs, where TF can be both a regulator and a target of the SE, and indicate that a compact glioma-specific auto-regulatory transcriptional network anchored on TFs and SEs is epigenetically controlled by an oncogenic eRNA that serves as a PRC2 modulator.

PRC2 silences genes by chemically depositing tri-methylation marks on histone H3 at lysine 27 with its catalytic subunit EZH2. Our CLIP analysis shows that EZH2 but not other members of PRC2 complex (SUZ12, EED, and RbAp46) binds pre-HOXDeRNA. Recent EZH2-CLIP studies highlighted the enrichment of 8-mer G-tract sequences at crosslinking sites *in vivo*.[24–27] Moreover, tethering G-tract or G-rich RNAs to the 5′ end of genes removes PRC2 components and H3K27Me3 from chromatin.[32] Consistent with these reports, we demonstrate that HOXDeRNA binds to EZH2, and their interaction is mediated by the rG4 structure. Using CRISPR base editing, which does not affect the activation and transcription of HOXDeRNA, in combination with *in vivo* kinetic and *in vitro* titration experiments, we show rG4-EZH2 binding and its global impact on gene expression. Furthermore, our study demonstrates the direct binding of HOXDeRNA to CGIs of genes upregulated in glioma. Hence, the binding specificity is defined by both PRC2 and the availability of CGIs. According to our model, HOXDeRNA is recruited to essential glioma-driving genes during the process of astrocyte transformation in a rG4/EZH2- and CpG-dependent manner, evicts PRC2, and counteracts its accumulation on these genes in glioma, regulating the dynamic balance between PRC2 and transcriptional machineries.

Two different methods demonstrated direct interactions between HOXDeRNA and chromatin: ChIRP-seq and iMARGI. The primary difference between these techniques is that ChIRP libraries are enriched for gDNA fragments interacting with a specific RNA of interest. By contrast, iMARGI detects all pairs of RNA-gDNA interactions by converting proximal RNA-DNA complexes into chimera sequences. Overall, since iMARGI is not a targeted technique, it is anticipated to detect fewer HOXDeRNA-specific interactions. Nevertheless, both ChIRP-seq and iMARGI demonstrate almost exclusive enrichment of HOXDeRNA binding to promoters of the genes upregulated after astrocyte transformation. The combined list of the promoters identified by these techniques covered all 35 key glioma TFs, with 25 of them detected by both methods. Of note, most prior studies utilizing iMARGI and similar techniques (RD-SPRITE, GRID-seq, and others) reported mainly in *cis* ("neighborhood") interactions, and those were largely attributed to highly abundant classes of RNA such as rRNA, small nucleolar RNA (snoRNA), and small nuclear RNA (snRNA).[66–68] By contrast, our analysis revealed that HOXDeRNA binds genome-wide, primarily to protein-coding genes (~83%–85% of *cis* and ~85% of *trans* interactions), with *trans* interactions exceeding *cis* interactions by 10-fold. Remarkably, 12%–14% of all interchromosomal RNA-gDNA interactions, measured across all RNAs in glioma and transformed astrocytes, were attributed to HOXDeRNA.

Finally, GBM is a highly heterogeneous disease characterized by a diverse array of mutations. However, mutations in the HOXDeRNA/HOXD region are rare and not linked to HOXDeRNA activation,[33] indicating that neoplastic transformation (i.e., HOXDeRNA-mediated) can occur without genetic alterations. Beyond TFs, HOXDeRNA upregulates multiple cancer drivers, including TERT, EGFR, BRAF, and PDGFR, which are established therapeutic targets. These findings enhance our understanding of gliomagenesis, traditionally viewed through mutational landscapes. They also suggest HOXDeRNA, a glioma-enriched RNA absent in normal neuroglial cells, as a unique therapeutic target. KD of HOXDeRNA significantly impairs glioma growth,[33–36,69,70] highlighting its potential for clinical translation. By controlling both "druggable" (e.g., EGFR, PDGFR, and BRAF) and "poorly druggable" elements (e.g., GSC tumor-initiating TFs and oncogenic miRNAs like miR-10b and miR-21) essential for glioma viability and therapy resistance, HOXDeRNA emerges as a promising therapeutic target.

Limitations of the study

This study has several limitations that open avenues for future research. First, while HOXDeRNA is expressed in some extracranial cells and dysregulated in various cancers, its role in normal cells remains largely unknown (reviewed in Xu et al.[71]). The proposed mechanisms in other cancers often focus on its cytosolic miRNA-sponging activity.[34,69,72] Additionally, HOXDeRNA binding to EZH2 has been reported in other cancers,[73,74] but the extent to which the global epigenetic mechanism described here applies to non-neuroglial cells remains unclear. Additional ChIRP-seq and iMARGI-seq experiments on other cell types would provide controlling and contrasting datasets. Second, the applications of bulk ChIRP-seq, ChIP-seq, and iMARGI-seq crosslinking-based techniques do not capture molecular dynamics and heterogeneity observed in GBM.[75,76] Additional studies enabling the analysis of clinical samples and mouse models at single-cell resolution will be necessary to investigate the details of HOXDeRNA activity across numerous promoters. Such studies will also further validate the physiological relevance of this newly identified mechanism in glioma formation *in vivo*. Third, although nuclear HOXDeRNA contains long introns, the segments interacting with target genes and EZH2 appear exclusively exonic. The prevalence of exonic over intronic reads, often observed with total non-polyA-enriched RNA in NGS techniques, could indicate a technical bias. In summary, the rules governing HOXDeRNA binding and chromatin modulation, as well as their applicability to other lncRNAs, remain exciting topics for future research.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact Dr. Anna M. Krichevsky, Brigham and Women's Hospital and Harvard Medical School, 60 Fenwood Road, Boston, MA 02115, USA. E-mail: akrichevsky@bwh.harvard.edu.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- NGS data generated in this study have been deposited at GEO with accession number GSE227805. Microscopy data and immunoblots were deposited to Mendeley (Mendeley Data: https://doi.org/10.17632/vzkdk7s4d5.1). The data are publicly available as of the date of publication.
- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the R21 NS098051, AG060019, R01 CA215072, and R01 NS113929 grants to A.M.K. and R01GM146997 and R01GM126150 grants to P.I. We thank the members of the Krichevsky laboratory and Dr. Leonid Mirny for helpful discussions and valuable insights. This manuscript was edited at Life Science Editors.

AUTHOR CONTRIBUTIONS

A.M.K. and E.D. conceived and designed the study. E.D. performed most experiments, data analysis, and visualization. P.K. performed CD and melting curves experiments. Y.Z. performed cell fractionation. A.K. assisted with computational analysis. P.I. contributed reagents and advice. A.M.K. supervised the work and acquired funding. E.D. and A.M.K. wrote the manuscript, and all authors revised and approved the manuscript.

DECLARATION OF INTERESTS

A patent application related to this work, listing Drs. Deforzh and Krichevsky as co-inventors, has been filed by Brigham and Women's Hospital.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell cultures and transfections
- METHOD DETAILS
  - Cell Fractionation and Droplet digital PCR
  - CRISPR activation of HOXDeRNA
  - CRISPR base editing: plasmids, cell transfection, and data analysis
  - Double RNA-DNA FISH
  - Chromatin Immunoprecipitation followed by DNA Sequencing (ChIP-Seq) and data analysis
  - Chromatin Isolation by RNA purification followed by DNA Sequencing (ChIRP-Seq) and data analysis
  - iMARGI and data analysis
  - mRNA-Seq, qRT-PCR, and data analysis
  - Cross-linking and immunoprecipitation (CLIP)
  - Circular dichroism (CD)
  - CD melting
  - HOXDeRNA-PRC2 titration
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2024.09.018.

Received: January 9, 2024
Revised: June 24, 2024
Accepted: September 13, 2024
Published: October 8, 2024

REFERENCES

1. Chow, L.M.L., Endersby, R., Zhu, X., Rankin, S., Qu, C., Zhang, J., Broniscer, A., Ellison, D.W., and Baker, S.J. (2011). Cooperativity within and between Pten, p53 and Rb pathways induces high-grade astrocytoma in adult brain. Cancer Cell *19*, 305–316. https://doi.org/10.1016/j.ccr.2011.01.039.
2. Song, Y., Zhang, Q., Kutlu, B., Difilippantonio, S., Bash, R., Gilbert, D., Yin, C., O'Sullivan, T.N., Yang, C., Kozlov, S., et al. (2013). Evolutionary etiology of high-grade astrocytomas. Proc. Natl. Acad. Sci. USA *110*, 17933–17938. https://doi.org/10.1073/pnas.1317026110.
3. Xiao, A., Wu, H., Pandolfi, P.P., Louis, D.N., and Van Dyke, T. (2002). Astrocyte inactivation of the pRb pathway predisposes mice to malignant astrocytoma development that is accelerated by PTEN mutation. Cancer Cell *1*, 157–168. https://doi.org/10.1016/s1535-6108(02)00029-6.
4. Hambardzumyan, D., Amankulor, N.M., Helmy, K.Y., Becher, O.J., and Holland, E.C. (2009). Modeling adult gliomas using RCAS/t-va technology. Transl. Oncol. *2*, 89–95.
5. Holland, E.C., Hively, W.P., DePinho, R.A., and Varmus, H.E. (1998). A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. Genes Dev. *12*, 3675–3685. https://doi.org/10.1101/gad.12.23.3675.
6. Uhrbom, L., Dai, C., Celestino, J.C., Rosenblum, M.K., Fuller, G.N., and Holland, E.C. (2002). Ink4a-Arf loss cooperates with KRas activation in astrocytes and neural progenitors to generate glioblastomas of various morphologies depending on activated Akt. Cancer Res. *62*, 5551–5558.
7. Kwon, C.-H., Zhao, D., Chen, J., Alcantara, S., Li, Y., Burns, D.K., Mason, R.P., Lee, E.Y.-H.P., Wu, H., and Parada, L.F. (2008). Pten haploinsufficiency accelerates formation of high-grade astrocytomas. Cancer Res. *68*, 3286–3294. https://doi.org/10.1158/0008-5472.CAN-07-6867.
8. Zhu, Y., Guignard, F., Zhao, D., Liu, L., Burns, D.K., Mason, R.P., Messing, A., and Parada, L.F. (2005). Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. Cancer Cell *8*, 119–130. https://doi.org/10.1016/j.ccr.2005.07.004.
9. Zheng, H., Ying, H., Yan, H., Kimmelman, A.C., Hiller, D.J., Chen, A.-J., Perry, S.R., Tonon, G., Chu, G.C., Ding, Z., et al. (2008). p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. Nature *455*, 1129–1133. https://doi.org/10.1038/nature07443.
10. Friedmann-Morvinski, D., Bushong, E.A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M.H., and Verma, I.M. (2012). Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. Science *338*, 1080–1084. https://doi.org/10.1126/science.1226929.
11. Bachoo, R.M., Maher, E.A., Ligon, K.L., Sharpless, N.E., Chan, S.S., You, M.J., Tang, Y., DeFrances, J., Stover, E., Weissleder, R., et al. (2002). Epidermal growth factor receptor and Ink4a/Arf: convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. Cancer Cell *1*, 269–277. https://doi.org/10.1016/s1535-6108(02)00046-6.
12. Ding, H., Roncari, L., Shannon, P., Wu, X., Lau, N., Karaskova, J., Gutmann, D.H., Squire, J.A., Nagy, A., and Guha, A. (2001). Astrocyte-specific expression of activated p21-ras results in malignant astrocytoma formation in a transgenic mouse model of human gliomas. Cancer Res. *61*, 3826–3836.
13. Holland, E.C., Celestino, J., Dai, C., Schaefer, L., Sawaya, R.E., and Fuller, G.N. (2000). Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice. Nat. Genet. *25*, 55–57. https://doi.org/10.1038/75596.
14. Le, D.M., Besson, A., Fogg, D.K., Choi, K.-S., Waisman, D.M., Goodyer, C.G., Rewcastle, B., and Yong, V.W. (2003). Exploitation of astrocytes by glioma cells to facilitate invasiveness: a mechanism involving matrix metalloproteinase-2 and the urokinase-type plasminogen activator-plasmin cascade. J. Neurosci. *23*, 4034–4043. https://doi.org/10.1523/JNEUROSCI.23-10-04034.2003.
15. Rheinbay, E., Suvà, M.L., Gillespie, S.M., Wakimoto, H., Patel, A.P., Shahid, M., Oksuz, O., Rabkin, S.D., Martuza, R.L., Rivera, M.N., et al. (2013). An aberrant transcription factor network essential for Wnt signaling and stem cell maintenance in glioblastoma. Cell Rep. *3*, 1567–1579.

16. Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr. Opin. Genet. Dev. *14*, 155–164. https://doi.org/10.1016/j.gde.2004.02.001.

17. de Vries, N.A., Hulsman, D., Akhtar, W., de Jong, J., Miles, D.C., Blom, M., van Tellingen, O., Jonkers, J., and van Lohuizen, M. (2015). Prolonged Ezh2 depletion in glioblastoma causes a robust switch in cell fate resulting in tumor progression. Cell Rep. *10*, 383–397. https://doi.org/10.1016/j.celrep.2014.12.028.

18. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022. https://doi.org/10.1101/gad.2037511.

19. Mendenhall, E.M., Koche, R.P., Truong, T., Zhou, V.W., Issac, B., Chi, A.S., Ku, M., and Bernstein, B.E. (2010). GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet. *6*, e1001244. https://doi.org/10.1371/journal.pgen.1001244.

20. Jermann, P., Hoerner, L., Burger, L., and Schübeler, D. (2014). Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proc. Natl. Acad. Sci. USA *111*, E3415–E3421. https://doi.org/10.1073/pnas.1400672111.

21. Lynch, M.D., Smith, A.J.H., De Gobbi, M., Flenley, M., Hughes, J.R., Vernimmen, D., Ayyub, H., Sharpe, J.A., Sloane-Stanley, J.A., Sutherland, L., et al. (2012). An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. EMBO J. *31*, 317–329. https://doi.org/10.1038/emboj.2011.399.

22. Song, J., Gooding, A.R., Hemphill, W.O., Love, B.D., Robertson, A., Yao, L., Zon, L.I., North, T.E., Kasinath, V., and Cech, T.R. (2023). Structural basis for inactivation of PRC2 by G-quadruplex RNA. Science *381*, 1331–1337. https://doi.org/10.1126/science.adh0059.

23. Gail, E.H., Healy, E., Flanigan, S.F., Jones, N., Ng, X.H., Uckelmann, M., Levina, V., Zhang, Q., and Davidovich, C. (2024). Inseparable RNA binding and chromatin modification activities of a nucleosome-interacting surface in EZH2. Nat. Genet. *56*, 1193–1202. https://doi.org/10.1038/s41588-024-01740-8.

24. Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1258–1264.

25. Davidovich, C., Zheng, L., Goodrich, K.J., and Cech, T.R. (2013). Promiscuous RNA binding by Polycomb repressive complex 2. Nat. Struct. Mol. Biol. *20*, 1250–1257.

26. Betancur, J.G., and Tomari, Y. (2015). Cryptic RNA-binding by PRC2 components EZH2 and SUZ12. RNA Biol. *12*, 959–965.

27. Long, Y., Bolanos, B., Gong, L., Liu, W., Goodrich, K.J., Yang, X., Chen, S., Gooding, A.R., Maegley, K.A., Gajiwala, K.S., et al. (2017). Conserved RNA-binding specificity of polycomb repressive complex 2 is achieved by dispersed amino acid patches in EZH2. eLife *6*, e31558.

28. Guo, J.K., Blanco, M.R., Walkup, W.G., Bonesteele, G., Urbinati, C.R., Banerjee, A.K., Chow, A., Ettlin, O., Strehle, M., Peyda, P., et al. (2024). Denaturing purifications demonstrate that PRC2 and other widely reported chromatin proteins do not appear to bind directly to RNA in vivo. Mol. Cell *84*, 1271–1289.e12. https://doi.org/10.1016/j.molcel.2024.01.026.

29. Wu, L., Murat, P., Matak-Vinkovic, D., Murrell, A., and Balasubramanian, S. (2013). Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. Biochemistry *52*, 9519–9527.

30. Sarma, K., Cifuentes-Rojas, C., Ergun, A., Del Rosario, A., Jeon, Y., White, F., Sadreyev, R., and Lee, J.T. (2014). ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. Cell *159*, 869–883.

31. Wang, X., Goodrich, K.J., Gooding, A.R., Naeem, H., Archer, S., Paucek, R.D., Youmans, D.T., Cech, T.R., and Davidovich, C. (2017). Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. Mol. Cell *65*, 1056–1067.e5.

32. Beltran, M., Tavares, M., Justin, N., Khandelwal, G., Ambrose, J., Foster, B.M., Worlock, K.B., Tvardovskiy, A., Kunzelmann, S., Herrero, J., et al. (2019). G-tract RNA removes Polycomb repressive complex 2 from genes. Nat. Struct. Mol. Biol. *26*, 899–909.

33. Deforzh, E., Uhlmann, E.J., Das, E., Galitsyna, A., Arora, R., Saravanan, H., Rabinovsky, R., Wirawan, A.D., Teplyuk, N.M., El Fatimy, R., et al. (2022). Promoter and enhancer RNAs regulate chromatin reorganization and activation of miR-10b/HOXD locus, and neoplastic transformation in glioma. Mol. Cell *82*, 1894–1908.e5. https://doi.org/10.1016/j.molcel.2022.03.018.

34. Jiang, L., Cheng, C., Ji, W., Wang, H., Du, Q., Dong, X., Shao, J., and Yu, W. (2021). LINC01116 promotes the proliferation and invasion of glioma by regulating the microRNA-744-5p-MDM2-p53 axis. Mol. Med. Rep. *23*, 366. https://doi.org/10.3892/mmr.2021.12005.

35. Wang, T., Cao, L., Dong, X., Wu, F., De, W., Huang, L., and Wan, Q. (2020). LINC01116 promotes tumor proliferation and neutrophil recruitment via DDX5-mediated regulation of IL-1β in glioma cell. Cell Death Dis. *11*, 302.

36. Ye, J., Zhu, J., Chen, H., Qian, J., Zhang, L., Wan, Z., Chen, F., Sun, S., Li, W., and Luo, C. (2020). A novel lncRNA-LINC01116 regulates tumorigenesis of glioma by targeting VEGFA. Int. J. Cancer *146*, 248–261. https://doi.org/10.1002/ijc.32483.

37. Sonoda, Y., Ozawa, T., Hirose, Y., Aldape, K.D., McMahon, M., Berger, M.S., and Pieper, R.O. (2001). Formation of intracranial tumors by genetically modified human astrocytes defines four pathways critical in the development of human anaplastic astrocytoma. Cancer Res. *61*, 4956–4960.

38. Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474–478. https://doi.org/10.1038/nature10860.

39. Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W.M., Lu, C., Ward, P.S., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479–483. https://doi.org/10.1038/nature10866.

40. Suvà, M.L., Rheinbay, E., Gillespie, S.M., Patel, A.P., Wakimoto, H., Rabkin, S.D., Riggi, N., Chi, A.S., Cahill, D.P., Nahed, B.V., et al. (2014). Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell *157*, 580–594.

41. Wu, W., Yan, Z., Nguyen, T.C., Bouman Chen, Z., Chien, S., and Zhong, S. (2019). Mapping RNA–chromatin interactions by sequencing with iMARGI. Nat. Protoc. *14*, 3243–3272. https://doi.org/10.1038/s41596-019-0229-4.

42. Mack, S.C., Singh, I., Wang, X., Hirsch, R., Wu, Q., Villagomez, R., Bernatchez, J.A., Zhu, Z., Gimple, R.C., Kim, L.J.Y., et al. (2019). Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. J. Exp. Med. *216*, 1071–1090. https://doi.org/10.1084/jem.20190196.

43. Yoo, W., Song, Y.W., Kim, J., Ahn, J., Kim, J., Shin, Y., Ryu, J.-K., and Kim, K.K. (2023). Molecular basis for SOX2-dependent regulation of super-enhancer activity. Nucleic Acids Res. *51*, 11999–12019. https://doi.org/10.1093/nar/gkad908.

44. Grosveld, F., van Staalduinen, J., and Stadhouders, R. (2021). Transcriptional regulation by (super) enhancers: from discovery to mechanisms. Annu. Rev. Genomics Hum. Genet. *22*, 127–146.

45. Bruter, A.V., Rodionova, M.D., Varlamova, E.A., and Shtil, A.A. (2021). Super-enhancers in the regulation of gene transcription: general aspects and antitumor targets. Acta Nat. *13*, 4–15.

46. Johnston, M.J., Nikolic, A., Ninkovic, N., Guilhamon, P., Cavalli, F.M.G., Seaman, S., Zemp, F.J., Lee, J., Abdelkareem, A., Ellestad, K., et al. (2019). High-resolution structural genomics reveals new therapeutic vulnerabilities in glioblastoma. Genome Res. *29*, 1211–1222. https://doi.org/10.1101/gr.246520.118.

47. Guiset Miserachs, H., Donghi, D., Börner, R., Johannsen, S., and Sigel, R.K.O. (2016). Distinct differences in metal ion specificity of RNA and DNA G-quadruplexes. J. Biol. Inorg. Chem. *21*, 975–986. https://doi.org/10.1007/s00775-016-1393-4.

48. Morris, M.J., and Basu, S. (2009). An unusually stable G-quadruplex within the 5′-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. Biochemistry *48*, 5313–5319. https://doi.org/10.1021/bi900498z.

49. Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. *36*, 843–846.

50. Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., et al. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc. Natl. Acad. Sci. USA *106*, 11667–11672.

51. Wei, J.-W., Huang, K., Yang, C., and Kang, C.-S. (2017). Non-coding RNAs as regulators in epigenetics (Review). Oncol. Rep. *37*, 3–9.

52. Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., and Shen, F. (2019). Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. Int. J. Mol. Sci. *20*, 5573.

53. 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.

54. Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. Cell *146*, 119–133.

55. McHugh, C.A., Chen, C.-K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232–236.

56. Colognori, D., Sunwoo, H., Kriz, A.J., Wang, C.-Y., and Lee, J.T. (2019). Xist deletional analysis reveals an interdependency between Xist RNA and polycomb complexes for spreading along the inactive X. Mol. Mol. Cell *74*, 101–117.e10.

57. Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.-C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071–1076. https://doi.org/10.1038/nature08975.

58. Marín-Béjar, O., Mas, A.M., González, J., Martinez, D., Athie, A., Morales, X., Galduroz, M., Raimondi, I., Grossi, E., and Guo, S. (2017). The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. Genome Biol. *18*, 1–15.

59. Yang, L., Lin, C., Jin, C., Yang, J.C., Tanasa, B., Li, W., Merkurjev, D., Ohgi, K.A., Meng, D., Zhang, J., et al. (2013). lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature *500*, 598–602.

60. Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-DiNardo, D., and Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol. Cell *32*, 232–246.

61. Schertzer, M.D., Braceros, K.C., Starmer, J., Cherney, R.E., Lee, D.M., Salazar, G., Justice, M., Bischoff, S.R., Cowley, D.O., and Ariel, P. (2019). lncRNA-induced spread of polycomb controlled by genome architecture, RNA abundance, and CpG island DNA. Mol. Cell *75*, 523–537.e10.

62. Jain, A.K., Xi, Y., McCarthy, R., Allton, K., Akdemir, K.C., Patel, L.R., Aronow, B., Lin, C., Li, W., Yang, L., et al. (2016). LncPRESS1 is a p53-regulated LncRNA that safeguards pluripotency by disrupting SIRT6-mediated de-acetylation of histone H3K56. Mol. Cell *64*, 967–981. https://doi.org/10.1016/j.molcel.2016.10.039.

63. Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

64. Barzilay, R., Melamed, E., and Offen, D. (2009). Introducing transcription factors to multipotent mesenchymal stem cells: making transdifferentiation possible. Stem Cells *27*, 2509–2515.

65. Zhao, K., Cui, X., Wang, Q., Fang, C., Tan, Y., Wang, Y., Yi, K., Yang, C., You, H., Shang, R., et al. (2019). RUNX1 contributes to the mesenchymal subtype of glioblastoma in a TGFβ pathway-dependent manner. Cell Death Dis. *10*, 877.

66. Bell, J.C., Jukam, D., Teran, N.A., Risca, V.I., Smith, O.K., Johnson, W.L., Skotheim, J.M., Greenleaf, W.J., and Straight, A.F. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. eLife *7*, e27024. https://doi.org/10.7554/eLife.27024.

67. Li, X., Zhou, B., Chen, L., Gou, L.-T., Li, H., and Fu, X.-D. (2017). GRID-seq reveals the global RNA-chromatin interactome. Nat. Biotechnol. *35*, 940–950. https://doi.org/10.1038/nbt.3968.

68. Quinodoz, S.A., Jachowicz, J.W., Bhat, P., Ollikainen, N., Banerjee, A.K., Goronzy, I.N., Blanco, M.R., Chovanec, P., Chow, A., Markaki, Y., et al. (2021). RNA promotes the formation of spatial compartments in the nucleus. Cell *184*, 5775–5790.e30. https://doi.org/10.1016/j.cell.2021.10.014.

69. Brodie, S., Lee, H.K., Jiang, W., Cazacu, S., Xiang, C., Poisson, L.M., Datta, I., Kalkanis, S., Ginsberg, D., and Brodie, C. (2017). The novel long non-coding RNA TALNEC2, regulates tumor cell growth and the stemness and radiation response of glioma stem cells. Oncotarget *8*, 31785–31801.

70. Zhang, Z.F., Xu, H.H., Hu, W.H., Hu, T.Y., and Wang, X.B. (2019). LINC01116 promotes proliferation, invasion and migration of osteosarcoma cells by silencing p53 and EZH2. Eur. Rev. Med. Pharmacol. Sci. *23*, 6813–6823. https://doi.org/10.26355/eurrev_201908_18720.

71. Xu, Y., Yu, X., Zhang, M., Zheng, Q., Sun, Z., He, Y., and Guo, W. (2021). Promising advances in LINC01116 related to cancer. Front. Cell Dev. Biol. *9*, 736927. https://doi.org/10.3389/fcell.2021.736927.

72. Lou, J., Wang, P., Chang, K., Wang, G., Geng, X., Wu, Y., Zhang, W., and Niu, G. (2021). Knocking down LINC01116 can inhibit the regulation of TGF-β through miR-774-5p axis and inhibit the occurrence and development of glioma. Am. J. Transl. Res. *13*, 5702–5719.

73. Liu, W., Liang, F., Yang, G., and Xian, L. (2021). LncRNA LINC01116 sponges miR-93-5p to promote cell invasion and migration in small cell lung cancer. BMC Pulm. Med. *21*, 50. https://doi.org/10.1186/s12890-020-01369-3.

74. Liang, W., Wu, J., and Qiu, X. (2021). LINC01116 facilitates colorectal cancer cell proliferation and angiogenesis through targeting EZH2-regulated TPM1. J. Transl. Med. *19*, 45. https://doi.org/10.1186/s12967-021-02707-7.

75. Couturier, C.P., Ayyadhury, S., Le, P.U., Nadaf, J., Monlong, J., Riva, G., Allache, R., Baig, S., Yan, X., Bourgey, M., et al. (2020). Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. Nat. Commun. *11*, 3406. https://doi.org/10.1038/s41467-020-17186-5.

76. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An integrative model of cellular states, plasticity, and genetics for glioblastoma. Cell *178*, 835–849.e21. https://doi.org/10.1016/j.cell.2019.06.024.

77. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

78. Liu, T. (2014). Use model-based analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods Mol. Biol. *1150*, 81–95. https://doi.org/10.1007/978-1-4939-0512-6_4.

79. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

80. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

81. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

82. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383. https://doi.org/10.1093/bioinformatics/btv145.

83. Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. https://doi.org/10.1145/2939672.2939785.

84. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4.

85. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

86. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

87. Wakimoto, H., Kesari, S., Farrell, C.J., Curry, W.T., Jr., Zaupa, C., Aghi, M., Kuroda, T., Stemmer-Rachamimov, A., Shah, K., Liu, T.-C., et al. (2009). Human glioblastoma–derived cancer stem cells: establishment of invasive glioma models and treatment with oncolytic herpes simplex virus vectors. Cancer Res. *69*, 3472–3481.

88. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192. https://doi.org/10.1093/bib/bbs017.

89. Chu, C., Quinn, J., and Chang, H.Y. (2012). Chromatin isolation by RNA purification (ChIRP). J. Vis. Exp. https://doi.org/10.3791/3912.

90. Khan, A., and Mathelier, A. (2017). Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. BMC Bioinformatics *18*, 287. https://doi.org/10.1186/s12859-017-1708-7.

91. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

92. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

93. Videm, P., Kumar, A., Zharkov, O., Grüning, B.A., and Backofen, R. (2021). ChiRA: an integrated framework for chimeric read analysis from RNA-RNA interactome and RNA structurome data. GigaScience *10*, giaa158. https://doi.org/10.1093/gigascience/giaa158.

94. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

95. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

96. Kucukural, A., Yukselen, O., Ozata, D.M., Moore, M.J., and Garber, M. (2019). DEBrowser: interactive differential expression analysis and visualization tool for count data. BMC Genomics *20*, 6. https://doi.org/10.1186/s12864-018-5362-x.

97. Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., et al. (2009). R package version 2.10.1. CRAN. R-project.org/package=gplots.

98. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

99. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527. https://doi.org/10.1038/nbt.3519.

100. Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.I. (2020). From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. *2*, 56–67. https://doi.org/10.1038/s42256-019-0138-9.

101. Jensen, K.B., and Darnell, R.B. (2008). CLIP: crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins. Methods Mol. Biol. *488*, 85–98. https://doi.org/10.1007/978-1-60327-475-3_6.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
EZH2	Active Motif	RRID:AB_2614956; Cat# 39902
H3K27Ac	Cell Signaling Technology	RRID:AB_10545273; Cat# 4353
H3K27Me3	Cell Signaling Technology	RRID:AB_2616029; Cat# 9733
ACTA1	GenScript	Cat# A00885-40
EED Monoclonal Antibody (41D)	Active Motif	RRID:AB_2615071; Cat# 50-199-3233
Suz12 Monoclonal Antibody (2AO9)	Active Motif	RRID:AB_2793377; Cat# 39877
RbAp46/48 Polyclonal Antibody	Active Motif	RRID:AB_2615007; Cat# 39199
Critical commercial assays		
NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®	NEB	Cat# E7546S
SimpleChIP® Enzymatic Chromatin IP Kit	Cell Signaling	Cat# 9003
Monarch PCR and DNA Clean Up Kit	NEB	Cat# T1030L
Droplet Digital PCR System	MOgene	RRID:SCR_019707; Cat# QX200
Deposited data		
Raw and analyzed NGS data	This Study	GEO: GSE227805
Imaging data and Immunoblots	This Study	Mendeley: https://doi.org/10.17632/vzkdk7s4d5.1
Experimental models: Cell lines		
Human: GBM4	Hiroaki Wakimoto laboratory	N/A
Human: GBM6	Hiroaki Wakimoto laboratory	N/A
Human: GBM8	Hiroaki Wakimoto laboratory	N/A
Human: LN229	ATCC	RRID: CVCL-0393; CRL-2611
Human: Normal human astrocytes	Dr. Yukihiko Sonoda laboratory	N/A
Oligonucleotides		
Stellaris® Probes for HOXDeRNA ChIRP and combined RNA-DNA FISH. Table S6.	Biosearch Technologies	N/A
Primers for RNA detection by qRT-PCR, CLIP-qPCR and gDNA detection by ChIRP-qPCR. Related to Figures 5B, 6D–6F, S1E, S2F, S6A, S7D, and S8 and Table S7	This study	N/A
Recombinant DNA		
pCMV_BE4max_P2A_GFP plasmid	Addgene	RRID: Addgene_112094
pLV[gRNA]-EGFP:T2A:Puro-U6	VectorBuilder	N/A
pRP[Exp]-{SOX2 promoter}>Luc2	VectorBuilder	VB230906-1377enp
pRP[Exp]-{OLIG2 promoter}>Luc2	VectorBuilder	VB230906-1375sgq
CRISPRmod CRISPRa lentiviral dCas9-VPR	Dharmacon	Cat# VCAS11918
Edit-R lentiviral HOXDeRNA sgRNAs	Dharmacon	N/A
Software and algorithms		
Bowtie2	Langmead and Salzberg[77] (version 2.3.4.1)	(RRID: SCR_016368) http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
MACS2	Liu[78] (version 2.2.7.1)	RRID: SCR_013291
Samtools	Li et al.[79] (version 1.13)	RRID: SCR_002105
Deeptools	Ramírez et al.[80] (version 3.5.0)	RRID: SCR_016366

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BEDtools	Quinlan et al.[81]	RRID: SCR_006646
ChIPseeker	Yu et al.[82] (version 1.34.1)	RRID: SCR_021322
Xgboost	Chen and Guestrin[83]	RRID: SCR_021361
HISAT2	Kim et al.[84] (Version 2.2.1)	RRID: SCR_015530
Trimmomatic	Bolger et al.[85] (version 0.39)	RRID: SCR_011848
ImageJ	Schneider et al.[86]	RRID: SCR_003070 https://imagej.nih.gov/ij/
R 3.6.3	R Core Team	RRID: SCR_001905 https://www.r-project.org/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell cultures and transfections

Human cells were used in accordance with institutional review board guidelines at Brigham and Women's Hospital. Low-passage human GBM stem cells (GBM4, GBM6, and GBM8) were a generous gift from Dr Hiroaki Wakimoto, MGH. The tumorigenic, genetic, and molecular properties of these cells have been described previously.[87] Cells were maintained in serum-free neurobasal media supplemented with N-2 and B-27 Plus Supplements (Gibco™), 3 mM Gibco® GlutaMAX™ Supplement, 50 units/ml penicillin and 50 units/ml streptomycin (Gibco™), 2 μg/ml heparin (Sigma-Aldrich), 20 ng/ml FGF2 (Sigma-Aldrich), and 20 ng/ml EGF (Sigma-Aldrich). Cells were passaged by dissociation using the Neurocult Stem Cells chemical dissociation kit (Stem Cells Technologies). Normal human astrocytes immortalized by E6/E7/hTERT (a generous gift from Dr. Yukihiko Sonoda[37] were maintained in Neurobasal medium similarly to GSCs. Human primary astrocytes were cultured as previously described.[33] Cell transfections with EZH2, OLIG2 or SOX2 siRNAs (50nM) have been performed as previously described.[33] Cell cultures were regularly tested for mycoplasma and cell lines were authenticated.

METHOD DETAILS

Cell Fractionation and Droplet digital PCR

A total of 1×10^7 GSCs were dissociated into single cells (StemCell Technologies, Catalog # 05707). Subsequently, cell fractionation was performed using PARIS Kit (Invitrogen, AM1921), following the manufacturer's prescribed protocol. Briefly, single cells were gently re-suspended in 500 μl of cold Cell Fractionation Buffer through pipetting and then incubated on ice for 5 minutes. The sample was subjected to centrifugation at 4°C and 500g for 5 minutes, then the supernatant, known as the cytoplasmic fraction, was carefully collected. An additional 500 μl of Cell Fractionation Buffer was used to wash the nuclear pellet, effectively eliminating any residual cytoplasmic constituents. The nuclear fraction was vigorously extracted from the nuclear pellet using 500 μl Cell Disruption Buffer by vortex. RNA was isolated from both the cytoplasmic and nuclear samples using a total RNA purification kit (Norgen, Catalog # 17270), and genomic DNA was eliminated with a DNase I kit (Norgen, Catalog # 25720). For detecting U6 expression, we performed reverse transcription of total RNA using the miRCURY LNA RT kit (QIAGEN, 339340) and employed the miRCURY LNA SYBR Green PCR kit (QIAGEN, 339347) with specific miRCURY LNA™ primers (YP00203907 for U6). To assess β-actin and GAPDH expression, 200 ng of total RNA underwent reverse transcription to cDNA using BIO-RAD (Catalog #1708841). Primers are listed in Table S9. Fold change was calculated as follows: delta Ct for each technical replicate = Cytoplasm or Nucleus Ct – Mean total RNA Ct, followed by 2^-delta Ct.

Isolated RNA from nuclear and cytoplasmic fractions was processed by MOgene using QX200 Droplet Digital PCR System with Taqman primers specific for premature and mature HOXDeRNA (Table S1).

CRISPR activation of HOXDeRNA

Lentiviral plasmids to generate stable dCas9-VPR nuclease-expressing cell populations (CRISPRa dCas9-VPR), Edit-R CRISPRa lentiviral sgRNA non-targeting control, and two custom-made Edit-R CRISPRa human HOXDeRNA lentiviral sgRNAs, all from Dharmacon, were employed as previously described.[33] sgRNAs used for HOXDeRNA gene activation were: AAGGCGCAGGCTG GCAGTTC, CCAGCCTGCGCCTTTGCAGC. Edit-R CRISPRa lentiviral sgRNA non-targeting control (cat. GSGC11913) was used as control sgRNA. NHA cells were sequentially transduced with dCas9-VPR followed by sgRNAs according to the Dharmacon technical manual.

CRISPR base editing: plasmids, cell transfection, and data analysis

pCMV_BE4max_P2A_GFP plasmid was a gift from Dr. David Liu (Addgene plasmid # 112094; http://n2t.net/addgene:112094; RRID: Addgene_112094). Plasmids encoding sgRNA for HOXDeRNA rG4-1 base editing (5'CCATTCCCCTCGGAGCAGCT) and control

sgRNA were purchased from VectorBuilder (pLV[gRNA]-EGFP:T2A:Puro-U6, VB230203-1324ybu). Six-well plates of transformed astrocytes in Neurobasal medium were transfected with 6 μl Lipofectamine 2000, 750 ng pCMV_BE4max_P2A_GFP, and 250 ng sgRNA per well. Cell culture medium was changed 4 hours later, and Puromycin (100 μg/mL) was used for cell selection 24 hours post-transfection. Genomic DNA was extracted using the genomic DNA purification kit (Dneasy Blood and Tissue kit, Cat. 69504, Qiagen), the specific fragment containing rG4-1 sequence was amplified using Phire Green Hot Start II PCR Master Mix (F126S) with the forward primer 5'TCCGCCTGGAAAAGAAGTCC and reverse primer 5'GAGGCAAGACTTTGGTGGGA, and the purified amplicon was sequenced at MGH DNA core facility.

Double RNA-DNA FISH

RNA FISH experiments employing Stellaris (Biosearch Technologies) probes were performed following the manufacturer's protocol for adherent cells. Briefly, cells were seeded onto coverslips one day before fixation. The fixation process involved treating cells with 4% paraformaldehyde (PFA) for 10 minutes, followed by washing in phosphate-buffered saline (PBS), and subsequent incubation in 70% ethanol at 4°C overnight for permeabilization. The next day, cells underwent a 5-minute incubation at room temperature in a wash buffer A (comprising 10% formamide in 2× saline sodium citrate (SSC)). Subsequently, the coverslips were inverted onto a drop of 200 μl hybridization buffer (containing 10% formamide, 2× SSC, and 10% Dextran sulfate) with 125 nM probe. This assembly was placed in a humidified chamber sealed with parafilm and incubated in darkness at 37°C overnight. Following hybridization, the coverslips were transferred to a fresh well and subjected to two 30-minute washes with Buffer A at 37°C, along with an additional brief wash in Buffer B. DAPI was employed as a DNA counterstain, and cells were mounted using Vectashield Antifade (ref H-1200) mounting medium. The probes used for HOXDeRNA detection are described in Table S8.

Further, DNA probes complementary to 'template' genomic DNA strands of SOX2 and OLIG2 promoters were designed with Stellaris software and manufactured by Biosearch Corporation. Following RNA visualization, coverslips were washed with PBS and incubated with RNAse A for 30 minutes at 37°C. DNA was denatured with probes in Hybridization buffer for 10 minutes at 80°C, followed by incubation in a humidity chamber at 37°C overnight. Washes after hybridization were performed in 0.4× SSC for 5 minutes at 72–74°C, then in 4× SSC with 0.1% Tween20 for 2 minutes at room temperature, and PBS for 10 minutes at room temperature. Mounting medium was applied with subsequent visualization. The images were obtained using a Dmi8 Leica fluorescent microscope, and pictures were saved in TIFF format. For Figures 3F (left) and S1E, all layers of the Z-stack were collapsed into one layer using maximum projection function of ImageJ software.[86]

Chromatin Immunoprecipitation followed by DNA Sequencing (ChIP-Seq) and data analysis

ChIP-Seq was performed using the SimpleChIP® Enzymatic Chromatin IP Kit (Magnetic Beads) (#9003, Cell Signaling). Briefly, 10 million cells were cross-linked with 1% formaldehyde and washed twice with ice-cold PBS. The collected pellet was resuspended in 2 ml RIPA buffer (BP -115X, Boston BioProducts) with protease inhibitors and fragmented to ~300 bp using MISONIX S-4000 Sonicator (amplitude 30%, 30 sec ON / 30 sec OFF, 30 min). 20 μg of chromatin in 1 ml of IP dilution buffer (16.7 mM Tris-HCl pH 8, 0.01% SDS, 1% Triton X-100, 167 mM NaCl, 1.2 mM EDTA) per ChIP was mixed with 10 μg of the following antibodies: H3K27Ac (#4353), H3K27Me3 (#9733) (all from Cell Signaling), EZH2 (Cat. 39002, Active Motif). The IP mixes were incubated overnight under rotation at 4°C, followed by an additional 4 hours with Dynabeads™ Protein G (30 μl per sample). IPs were washed twice in a low salt buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 150 mM NaCl, 1% Triton X-100 in distilled water), once with a high salt buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 500 mM NaCl, 1% Triton X-100 in distilled water) and once in TE buffer, eluted and de-crosslinked according to the instructions (SimpleChIP® Enzymatic Chromatin IP Kit). DNA was purified using the Monarch PCR and DNA Clean Up Kit (#T1030L, NEB) and ChIP-seq libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (E7546S, NEB) and sequenced on the Illumina HiSeq 2500 platform configured for 50-bp single-end reads.

For data analysis, Fastq files were aligned to the human reference genome (GRCh37/hg19) using Bowtie2 with default parameters. PCR duplicates were removed using samtools rmdup version 1.13.[79] The resulting aligned Bam files were transformed to Bigwig format without scaling or normalization using deeptools version 3.5.0.[80] Bed files for peaks were created using MACS2 (version 2.2.7.1).[78] The IGV web application[88] was used for visualization.

For Figure 4D, SOX2 (GSM1306360_MGG8TPC.SOX2, n=2) and OLIG2 (GSM1306365_MGG8TPC.OLIG2, n=2) BED files (peaks) were downloaded from GEO. All overlapping intervals were merged with MergeBED function (BEDtools[81]) after replicates were concatenated tail-to-head. "Intersect intervals" function (BEDtools) was used to generate lists of SOX2-bound and OLIG2-bound super enhancers. Resulting lists were intersected and visualized as Venn diagram.

ChIP-qPCR was performed with primers indicated in Table S9. Data were analysed in 4 steps: 1) 1% input Ct values were adjusted to 100% by subtracting 6.64; 2) Ct (IP) was subtracted from the adjusted input Ct to obtain delta Ct; 3) calculate 100*2^delta CT for each IP; 4) calculate a mean with SD for each condition.

Chromatin Isolation by RNA purification followed by DNA Sequencing (ChIRP-Seq) and data analysis

ChIRP was performed based on the previously published protocol.[89] Briefly, 48 biotinylated oligonucleotide probes specific for nascent HOXDeRNA RNA were designed and manufactured using ChIRP probe designer from LGC Biosearch Technologies (https://www.biosearchtech.com/support/tools/design-software/stellaris-probe-designer). Probes for LacZ mRNA (Sigma-Aldrich, 03-307) were used as negative control. Single cell suspensions were prepared from 100 million adherent or suspension cells followed

by cross-linking with 2% glutaraldehyde (Sigma, G5882-50), and rotation at 20 rpm RT for 20 minutes. The reaction was quenched with glycine solution (7005S, Cell Signaling), and the cells were washed twice with ice-cold PBS. Sonication was performed in 15 ml falcon tubes with RIPA buffer and protease inhibitors (cOmplete Mini EDTA free), and chromatin was fragmented to ~200 bp using MISONIX S-4000 sonicator (amplitude 30%, 30 sec ON / 30 sec OFF, 20 min). The chromatin from each cell line was mixed with the protease and RNAse inhibitors (SuperaseIn RNAse Inhibitor, AM2696, ThermoFisher Scientific), split into three tubes (for input, LacZ, and HOXDeRNA binding), incubated for 4 h at 37C 1000 rpm with respective sets of biotinylated probes, and followed by another 2 h incubation with Dynabeads MyOne Streptavidine C1 (62001, ThermoFisher). The beads were washed twice with low and high salt buffers at 37C, 1000 rpm, followed by de-crosslinking with Proteinase K for 4 hours at 65C. For ChIRP-WB analysis, the beads were loaded into NuPAGE Bis-Tris protein gels (Invitrogen, NP0335PK2). EZH2 (Cat. 39002, Active Motif, dilution 1:1000) and ACTA1 (Catalog # A00885-40, GenScript, dilution 1:500) antibodies were used for protein detection. For sequencing, DNA was purified using the Monarch PCR and DNA Clean Up Kit (#T1030L, NEB), Illumina NGS libraries were prepared using the NEBNext Ultra II DNA Kit (NEB, E7645S), and sequenced for 50-nt SE reads. For data analysis, Fastqsanger files were mapped to hg19 genome using bowtie2[77] (default settings), and the resulting aligned Bam files were transformed to Bigwig format using deeptools version 3.5.0.[80] Raw reads were quantified for each 50-nucleotide bin for each Bigwig file. The Bigwigs signal tracks were visualized with IGV browser. Peaks were called using MACS version 2.0.10[78] (default parameters). The resulting bed files were intersected with other interval data. For Figure 2A, ChIRP peaks were annotated (using Homo_sapiens.GRCh37.75.gtf (hg19)) with ChIPseeker (version 1.34.1, [82]) to the nearest genomic feature. The number of peaks corresponding to various genomic features was plotted as a Venn diagram. For Figure 2B, ChIRP peaks were annotated to the nearest gene for each cell line using ChIPseeker, followed by visualization of intersected gene lists with a web tool 'Intervene'[90]. Average profile plots and heatmaps for Figures 2C–2F, 3A, 4C, and S2 were generated using the deeptools functions plotProfile and plotHeatmap.[80]

The following negative controls were included for the assessment of the peaks' quality: HOXDeRNA ChIRP-seq on naive NHA cells not expressing HOXDeRNA, LacZ ChIRP-seq on all tested cell lines, and ChIRP-seq with probes for the genomic template (antisense) strand of HOXDeRNA for all cell lines. The relative number of binding events in intergenic regions were measured with ChIPseeker in all ChIRP-Seq GEO datasets with available BED files: GSE208012, GSE112602, GSE198808, GSE201567, GSE156121, GSE114981, GSE128369, GSE76418, GSE47804, GSE31332. These control experiments produced no specific signals, validating the specificity of HOXDeRNA binding to indicated DNA regions.

The ChIRP-qPCR data were analyzed using 2^-delta Ct formula, where delta Ct= Ct (IP) -Ct (GAPDH). Primers used in the reactions are listed in Table S9.

iMARGI and data analysis

iMARGI was conducted following the described protocol with slight modifications to enhance the detection of direct, non-protein-mediated RNA-gDNA interactions.[41] In brief, 20 million cells were crosslinked using 1% formaldehyde, and the nuclei were isolated. Nuclear RNA and DNA were fragmented using RNase I and AluI restriction enzyme. Crosslinks were reversed by incubating nuclei with proteinase K for 2 hours at 65°C. Nuclei integrity was monitored under the microscope. A specially designed linker was introduced to ligate with the fragmented RNA and subsequently ligate with spatially proximal DNA. Nucleic acids were purified, and exonucleases were used to remove any linker sequences that did not successfully ligate with both RNA and DNA. The resulting ligation products, in the form of RNA–linker–DNA, were captured with streptavidin beads. The RNA part of the captured sequence was reverse transcribed into cDNA, generating a complementary strand of (5′) DNA–linker–cDNA (3′). Single-stranded DNA–linker–cDNA was released from streptavidin beads, circularized, and re-linearized, producing single-stranded DNA in the form of left.half.Linker–cDNA–DNA–right.half.Linker. The two halves of the linker (left.half.Linker and right.half.Linker) served as templates for PCR amplification. The linearized DNA was then amplified with NEBNext PCR primers for Illumina, size-selected, and subjected to 150 cycles of pair-end sequencing with an Illumina Hiseq 2500.

Fastq forward and reverse reads for each sample were merged using FLASH, version 1.2.11[91]), and read duplicates were collapsed into a fasta with the calculation duplicates counts per read (fastx_toolkit (Version 0.0.14)). Reads were aligned to reference fasta file containing transcript sequences (hg38_Homo_sapiens.GRCh38.100.fasta) using BWA-MEM[92] in local alignment mode. The alignment considered both strands with the following parameters: Seed length for 1st mapping iteration = 12, Seed length for 2nd mapping iteration =12, Minimum alignment score for 1st mapping iteration = 18, Minimum alignment score for 2nd mapping iteration = 16, Matching score for 1st mapping iteration = 1, Mismatch penalty for 1st mapping iteration = 4, Matching score for 2nd mapping iteration = 1, Mismatch penalty for 2nd mapping iteration = 6, Gap opening penalty for 1st mapping iteration = 6, Gap extension penalty for 1st mapping iteration = 1, Gap opening penalty for 2nd mapping iteration = 100, Gap extension penalty for 2nd mapping iteration = 100, Maximum number of allowed multi hits per read in 1st iteration = 50, Maximum number of allowed multi hits per read in 2nd iteration = 100. Multi-mapped reads were filtered out. Unmapped reads were then mapped to the reference genome (hg38_Homo_sapiens.GRCh38.dna.toplevel.fa) in order to detect intergenic binding for Figure 3A. Significantly overlapping aligned positions were merged based on the reference mapping locations provided in GTF format (hg38_Homo_sapiens.GRCh38.100.gff) to define the expressed loci, which are regions on the reference built from overlapping alignments. Common Read Loci, a set of expressed loci sharing most of their multi-mapped reads, were determined and quantified. After obtaining information about the loci expression and Common Read Loci, the final step involved extracting only the best scoring interacting partners for each read.[93] The resulting file with chimeric reads, having HOXDeRNA as the RNA arm and genomic DNA as the second arm, along with column definitions, can be found in a Table S6.

mRNA-Seq, qRT-PCR, and data analysis

Total RNA was isolated with Norgen Biotek kit followed by treatment with DNAse I (Promega) at 37°C for 30 minutes. qRT-PCR reactions were performed as previously described,[33] with primers listed in the Table S9. The data were analysed in three steps: 1) Ct (gene of interest) was subtracted from GAPDH Ct (normalization control) to obtain delta Ct; 2) 2^-delta Ct was calculated for each sample; 3) means and SD for each group were calculated and plotted as bars. Alternatively, cDNA libraries were prepared, and deep sequencing was performed by Novogene. For the analysis of RNAseq data, Fastq files were trimmed to remove Illumina adapters with Trimmomatic (version 0.39[85]) and aligned to the GRCh37/hg19 genome using HISAT 2.2.1[84] with the annotation Homo_sapiens.GRCh37.75.gtf (hg19) containing exon-exon splice junctions. Raw counts were measured with featureCounts (part of the subread 2.0.0 package[94]). Differential expression analysis was performed using Deseq2 (version 1.39.4[95]). PCA plot and GO analysis for Figure S1A were produced and visualized using Debrowser (version 1.24.1.[96] Heatmaps for Figures 1B and S3 were generated with heatmap.2 function of the R gplots package, with z-scores calculated for each gene across all samples (version: 3.1.3).[97]

Principal Component Analysis (PCA) based on machine learning selected features for Figure 1D was performed as following. We used the TCGA TARGET GTEx dataset from the UCSC RNA-seq compendium, where TCGA, TARGET, and GTEx samples were processed using the same bioinformatic pipeline, including alignment to the hg38 genome, and gene expression calling with RSEM[98] and Kallisto[99] methods. Uniform processing eliminated computational batch effects in this dataset. The dataset can be downloaded at: https://xenabrowser.net/datapages/?cohort=TCGA%20TARGET%20GTEx&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443. To make our RNA-Seq data directly comparable to the selected datasets, they were processed through the same pipeline.

Feature selection and training for generating Figure 1D was performed as following. To define DEG between the groups, we used a significance threshold of $p < 0.05$, followed by a selection of the top first 500 genes sorted by fold change from our RNA-seq data – in naïve astrocytes vs. transformed astrocytes. Feature selection was then applied to these 500 genes using the SHAP package.[100] We calculated the importance of the genes using the Shapley values and used this information to improve the performance of the model by removing the less important genes. In the final step, we selected 30 genes with the highest scores for model prediction. To account for possible batch effects of different datasets, we transformed the data for the selected genes into a rank order before training the model. We used an xgboost model[83] to predict whether the samples were associated with normal brain categories (Forebrain (n=857), Cerebellum (n=214), Midbrain (n=57)), LGG (n=522), or GBM categories (Mesenchymal (n=54), Classical (n=41), Proneural (n=39), and Neural (n=28)).

Cross-linking and immunoprecipitation (CLIP)

CLIP experiments were performed based on previously published protocol.[101] Briefly, to covalently cross-link proteins to nucleic acids, 2×10^7 cells were exposed to UV irradiation (200 mJ/cm^2) for 2 minutes. The cells were then lysed with RIPA buffer containing Protease Inhibition Cocktail and RNase Inhibitor. Immunoglobin-coated magnetic protein beads (Thermofisher Scientific) were incubated with EZH2, EED, SUZ12, RbAp46 (39002, 50-199-3233, 50-199-2923, 50-199-3951, all from Active Motif) or IgG (Catalog # 2729S, Cell Signaling) antibodies. The complexes were washed with RIPA buffer and the samples were treated with DNAse I (Promega) for 30 minutes at 37°C, followed by Proteinase K (10% SDS and 10 mg/ ml Proteinase K in RIPA buffer) treatment for 30 minutes at 37°C with shaking. RNA was further isolated with a phenol:chloroform:isoamyl alcohol (25:24:1) solution, precipitated with isopropanol, and resuspended in RNase-free water. qRT-PCR was performed for nascent HOXDeRNA and GAPDH mRNA. Primers are listed in Table S9. Fold enrichment over IgG was calculated in two steps: 1) DDCt = (Ct IP) - (Ct IgG), and 2) fold enrichment = 2^-DDCt.

Circular dichroism (CD)

The oligonucleotides were dissolved in 150 mM K^+ in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA). 200 μL of 10 μM solution was placed in quartz cuvettes (1 mm path length) and the spectra were collected in the range between 200 and 320 nm at 20 °C from three scans, and a buffer baseline was subtracted from each spectrum. Increased peak intensity of the oligo under the K^+ environment (compared to the Li^+ environment) at 260-265 nm and a trough at around 240 nm suggests the formation of an rG4. JASCO J815 spectropolarimeter was used to collect CD spectra. The data was plotted using GraphPad Prism (Smooth: 10 neighbors on each side, second order smoothing polynomial).

CD melting

CD melting experiments were performed in the same spectrophotometer and similar solution conditions as described above. The CD ellipticity of 10 μM folded solution of rG4s was tracked at 263 nm using a variable temperature measurement method. The thermal data was collected between 20 °C to 95 °C sampling per °C at a rate of 1°C/ min. The data was plotted in a XY format using GraphPad Prism (Smooth: 6 neighbors in each side, fourth order smoothing polynomial) and first derivative of the data was plotted to calculate the melting temperature (T_m) of the structure, which is defined as a temperature at which the sequence is equally populated in the folded and unfolded states.

HOXDeRNA-PRC2 titration

Recombinant PRC2 complex (composed of EZH2, EED, and Suz12 subunits, Active Motif Catalog No 31337) was mixed with 25 μg of plasmids containing the promoter regions of SOX2 (VB230906-1377enp vector, chr3:181711153-181712186, hg38) or OLIG2

(VB230906-1375sgq vector, chr21:33024853-33026163, hg38) and with increasing concentrations of either HOXDeRNA WT rG4-1 or rG4-1-mutated oligonucleotides in 50 μl reaction mix (20 mM phosphate pH 7.4, 0.05% Tween-20, 40 μM S-adenosylmethionine, and 1000 ng enzyme complex) and incubated for 1h RT. The mix was adjusted to 1 ml with RIPA, sonicated and incubated with EZH2 antibody (Cat. 39002, Active Motif, dilution 1:1000) overnight at 4°C. EZH2 coverage was analysed at promoters of SOX2 and OLIG2 by ChIP-Seq.

QUANTIFICATION AND STATISTICAL ANALYSIS

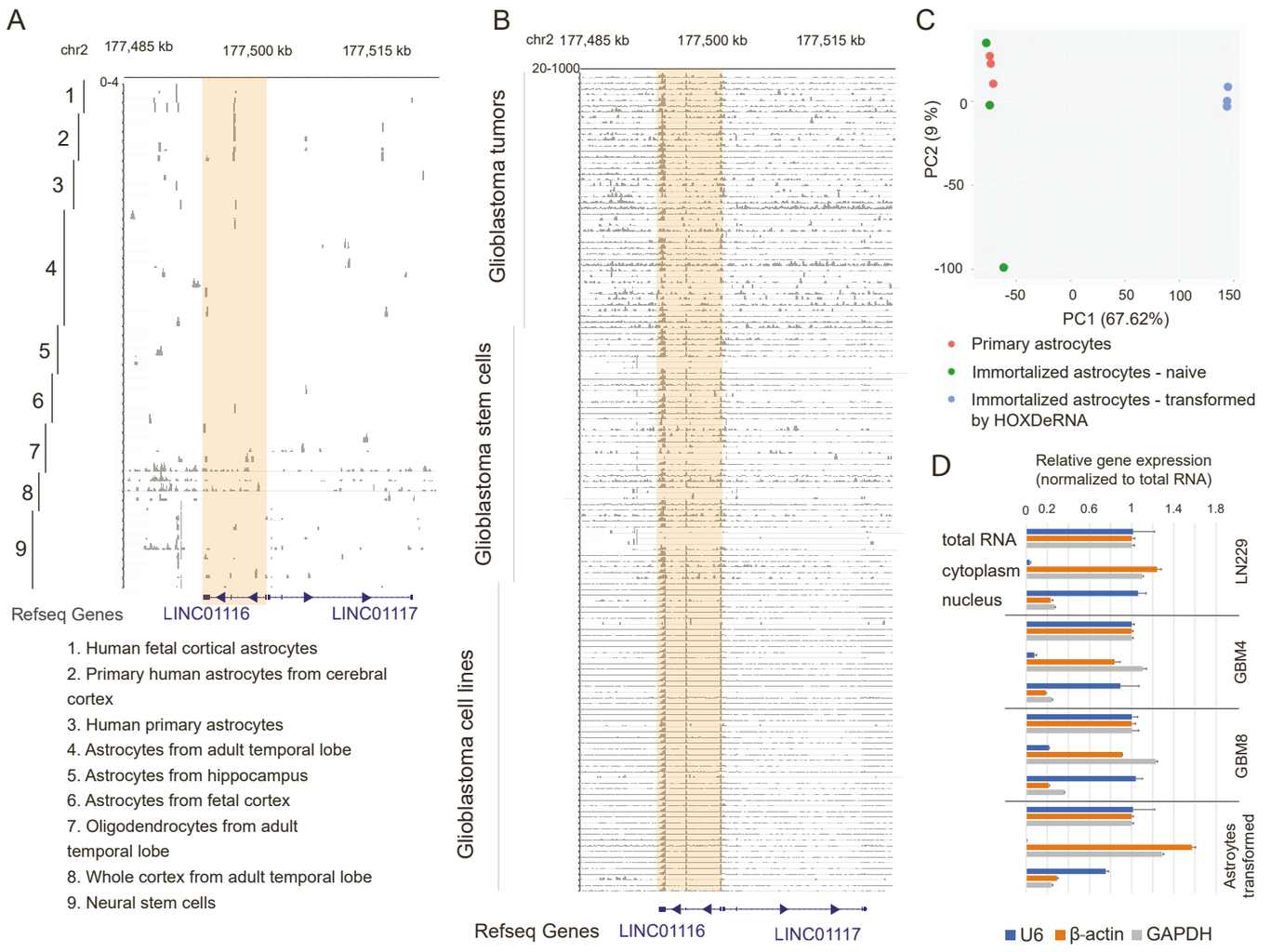
All statistical analyses were performed using Prism 8 GraphPad Software. Statistical significance of the differences between groups was measured with two-tailed unpaired t-test. In figure legends: *p< 0.05, **p<0.01, ***P<0.001, ****p<0.0001. Non-significant differences are marked with NS.

Molecular Cell, Volume *84*

Supplemental information

HOXDeRNA activates a cancerous transcription program and super enhancers via genome-wide binding

Evgeny Deforzh, Prakash Kharel, Yanhong Zhang, Anton Karelin, Abdellatif El Khayari, Pavel Ivanov, and Anna M. Krichevsky


A
chr2 177,485 kb
177,500 kb
177,515 kb
0-4
1
2
3
4
5
6
7
8
9
Refseq Genes
LINC01116
LINC01117
1. Human fetal cortical astrocytes
2. Primary human astrocytes from cerebral cortex
3. Human primary astrocytes
4. Astrocytes from adult temporal lobe
5. Astrocytes from hippocampus
6. Astrocytes from fetal cortex
7. Oligodendrocytes from adult temporal lobe
8. Whole cortex from adult temporal lobe
9. Neural stem cells
B
chr2 177,485 kb
177,500 kb
177,515 kb
20-1000
Glioblastoma tumors
Glioblastoma stem cells
Glioblastoma cell lines
Refseq Genes
LINC01116
LINC01117
C
PC2 (9 %)
PC1 (67.62%)
Primary astrocytes
Immortalized astrocytes - naive
Immortalized astrocytes - transformed by HOXDeRNA
D
Relative gene expression (normalized to total RNA)
total RNA
cytoplasm
nucleus
LN229
GBM4
GBM8
Astroytes transformed
U6
β-actin
GAPDH
E
HOXDeRNA / DAPI
Zoom in
GBM8
Astrocytes
10 µm
F
Number of premature HOXDeRNA copies/cell
LN229
GBM4
GBM8
Astrocytes transformed
G
qRT-PCR validation of RNA-Seq
Astrocyte control
Astrocytes transformed
Relative gene expression (normalized to GAPDH)
SOX2
SOX8
OLIG1
OLIG2
ASCL1
POU3F2
POU3F3
HEY2
RFX4
CITED1
VAX2

Supplemental Figure 1. HOXDeRNA is not expressed in normal neuroglial cells, in contrast to glioma cells, and its activation globally alters transcriptional programs. Related to Figure 1.

A-B. HOXDeRNA expression monitored by RNA-Seq in normal brain cells (GSE122701, GSE157461, GSE166847, GSE73721 and GSE119834 datasets) (A) and GBM tissues, GSCs (GSE119834), and glioma Cancer Cell Line Encyclopaedia 78 (B). HOXDeRNA transcript, composed of 3 exons, is expressed in all glioma samples but not in the normal cells of the brain.

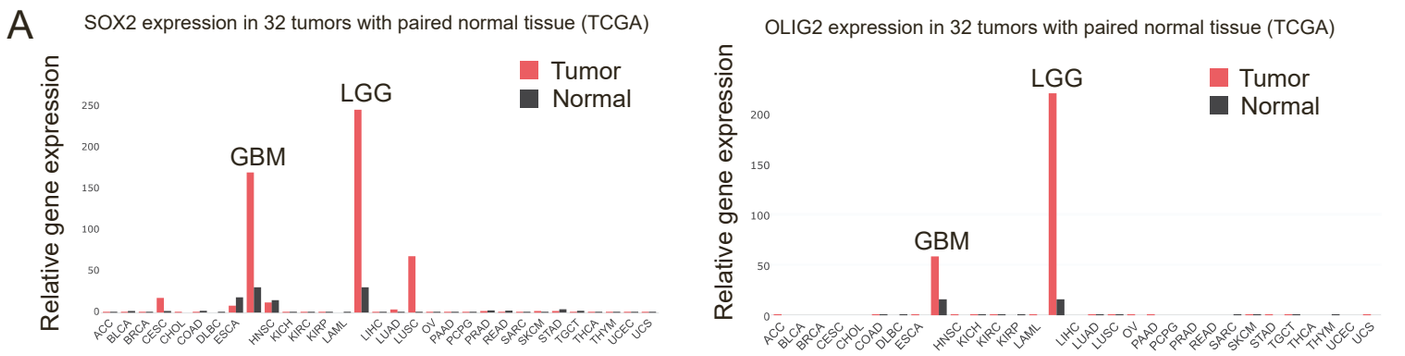
C. Transcriptomic differences among primary astrocytes (n=3) and immortalized astrocytes naive or transformed by HOXDeRNA (n=3), visualized as PCA plot.

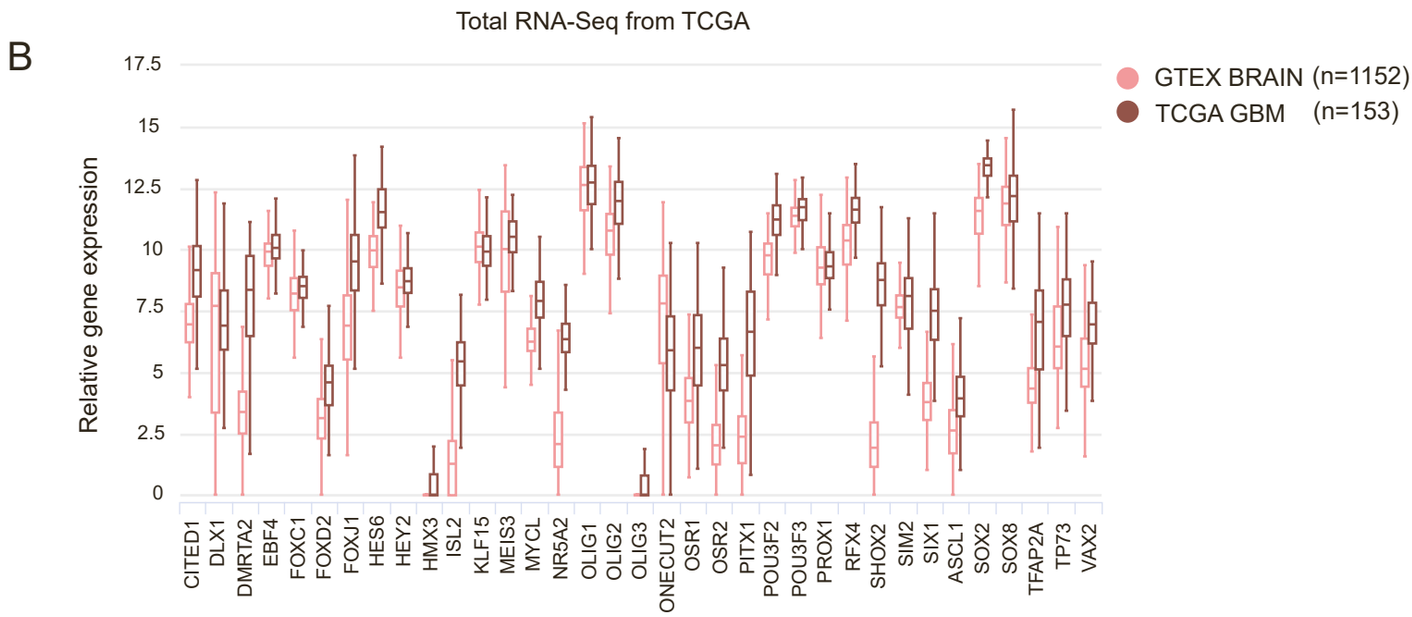
D. The purity of cytoplasmic and nuclear fractions for HOXDeRNA quantification was assessed in LN229 glioma cells, GSCs, and transformed astrocytes. SnRNA U6, β-actin, and GAPDH served as markers for nuclear and cytoplasmic fractions, respectively. After normalization to their levels in total RNA fractions, Ct values were visualized as bars (mean of 3 technical replicates +SD).

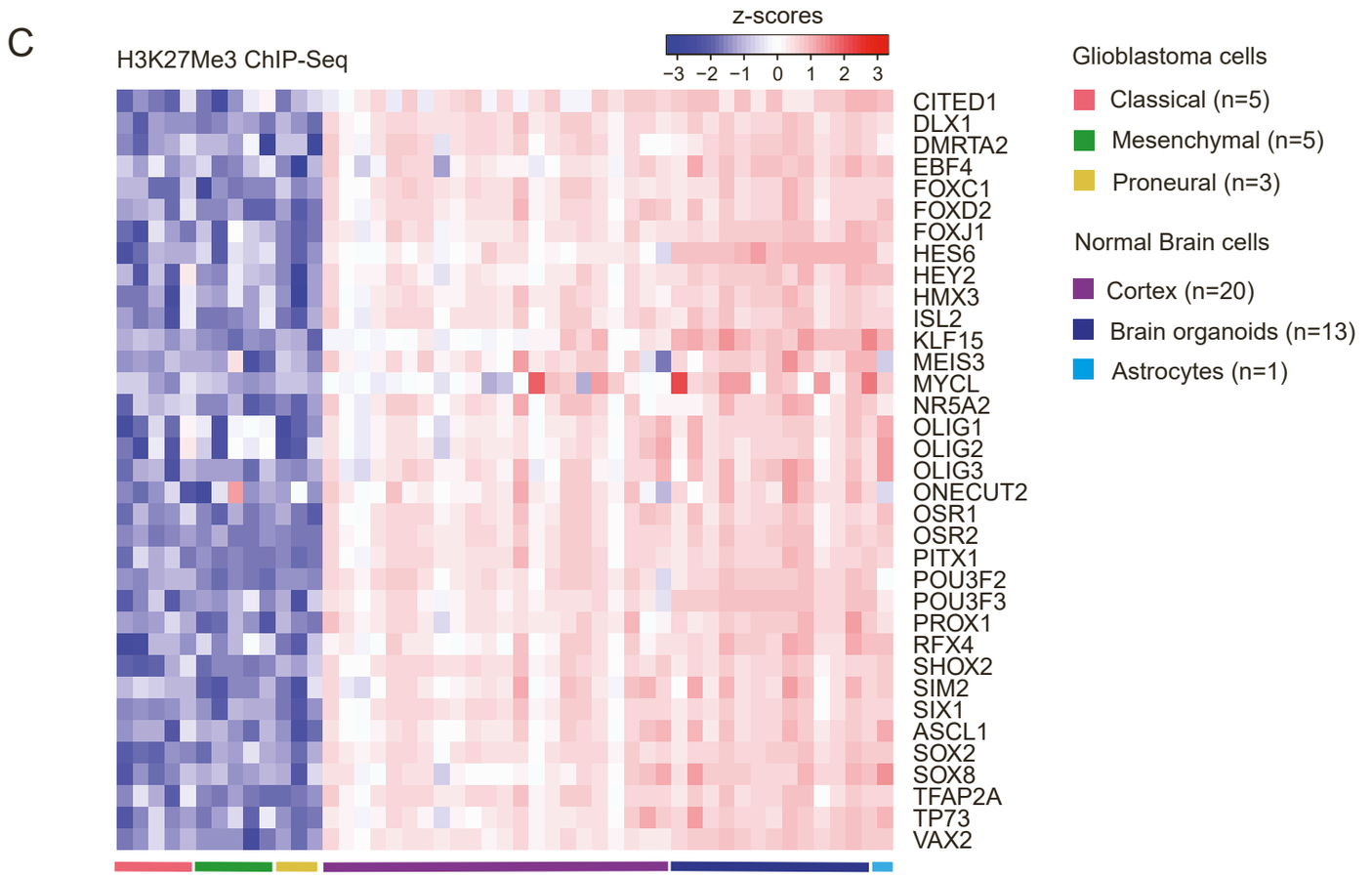
E. smFISH probes specifically detect pre-HOXDeRNA in GBM8 cells (top), with no signal observed in naïve normal human astrocytes (bottom).

F. Pre-HOXDeRNA was detected by smFISH and quantified in glioma cells, GSCs, and transformed astrocytes, with levels ranging from 45 to 90 copies per nucleus. The center line shows the median expression, box limits indicate the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values.

G. Representative key glioma TFs are upregulated after HOXDeRNA-mediated astrocyte transformation. Relative gene expression was measured by qRT-PCR, normalized to GAPDH and visualized as bar graph (mean +SD, n=3).


A
SOX2 expression in 32 tumors with paired normal tissue (TCGA)
Relative gene expression
Tumor
Normal
GBM
LGG
OLIG2 expression in 32 tumors with paired normal tissue (TCGA)
Relative gene expression
Tumor
Normal
GBM
LGG


B
Total RNA-Seq from TCGA
Relative gene expression
GTEX BRAIN (n=1152)
TCGA GBM (n=153)
CITED1
DLX1
DMRTA2
EBF4
FOXC1
FOXD2
FOXJ1
HES6
HEY2
HMX3
ISL2
KLF15
MEIS3
MYCL
NR5A2
OLIG1
OLIG2
OLIG3
ONECUT2
OSR1
OSR2
PITX1
POU3F2
POU3F3
PROX1
RFX4
SHOX2
SIM2
SIX1
ASCL1
SOX2
SOX8
TFAP2A
TP73
VAX2

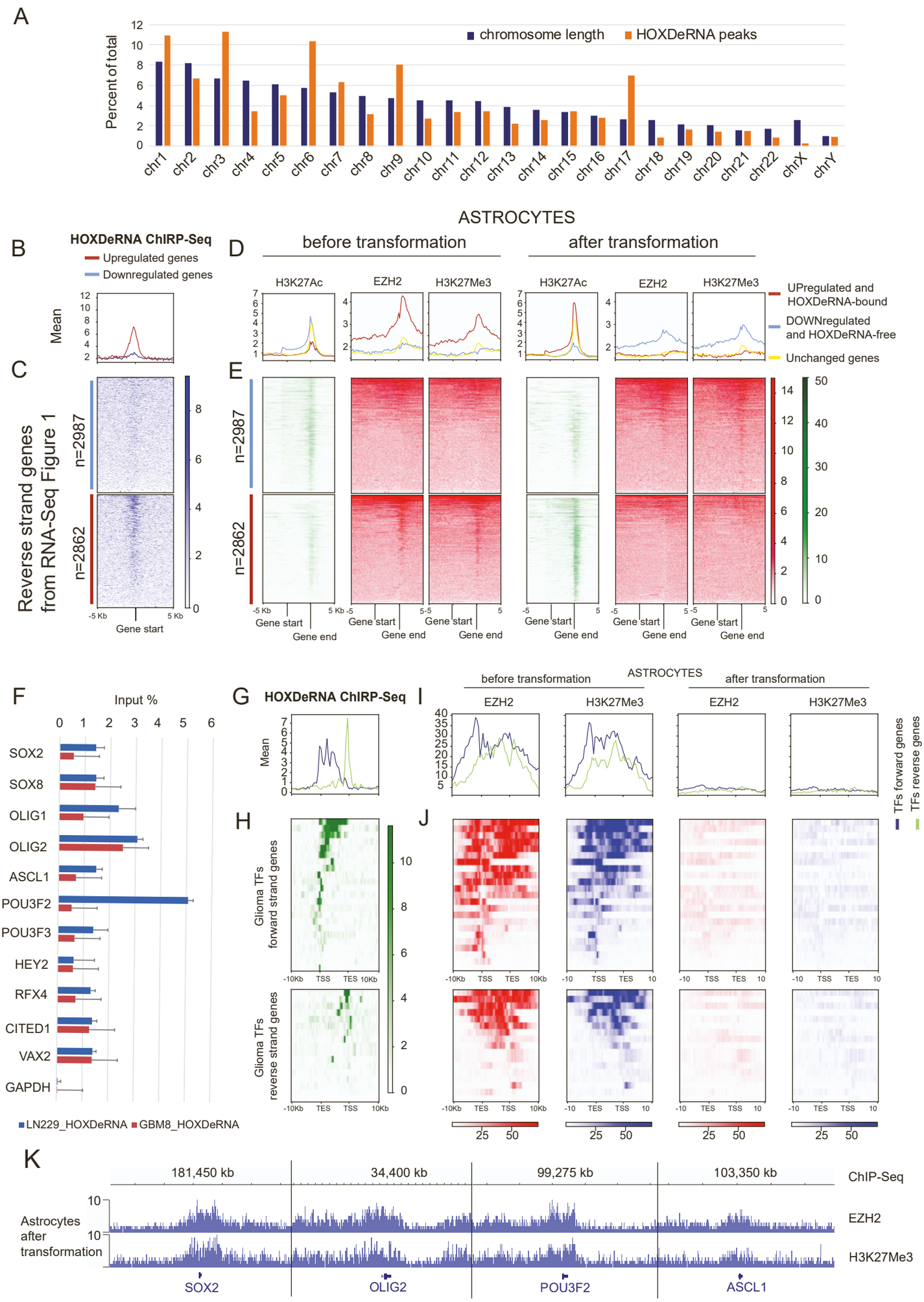

C
H3K27Me3 ChIP-Seq
z-scores
Glioblastoma cells
Classical (n=5)
Mesenchymal (n=5)
Proneural (n=3)
Normal Brain cells
Cortex (n=20)
Brain organoids (n=13)
Astrocytes (n=1)

Supplemental Figure 2. Glioma-specific TFs are repressed in the normal brain and activated in GBM. Related to Figure 1.

A. SOX2 and OLIG2 are glioma-specific TFs. mRNA levels of SOX2 and OLIG2 were visualized across 32 tumors and corresponding normal tissues (RNAseq, TCGA database) using the GEPIA2 website. Bar heights represent mean TPM values for each tumor or normal tissue.

B. mRNA expression of 35 glioma TFs in the human brain (n=1152) and GBM (n=153). TCGA TARGET GTEx RNAseq dataset visualized as boxplots; unit: log2(norm_count+1). Center line shows the median expression, box limits indicate the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values.

C. The 35 glioma TFs are epigenetically silenced in normal brain cells compared to GBM, based on TCGA and ENCODE dataset analysis. Total reads between raw samples were equalized, log2 transformed, and computed Z-scores are plotted.


A
Percent of total
chromosome length
HOXDeRNA peaks
chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18 chr19 chr20 chr21 chr22 chrX chrY
B
HOXDeRNA ChIRP-Seq
Upregulated genes
Downregulated genes
Mean
C
Reverse strand genes from RNA-Seq Figure 1
n=2987
n=2862
-5 Kb
5 Kb
Gene start
D
ASTROCYTES
before transformation
after transformation
H3K27Ac
EZH2
H3K27Me3
UPregulated and HOXDeRNA-bound
DOWNregulated and HOXDeRNA-free
Unchanged genes
E
Gene start
Gene end
F
Input %
SOX2
SOX8
OLIG1
OLIG2
ASCL1
POU3F2
POU3F3
HEY2
RFX4
CITED1
VAX2
GAPDH
LN229_HOXDeRNA
GBM8_HOXDeRNA
G
HOXDeRNA ChIRP-Seq
Mean
H
Glioma TFs forward strand genes
Glioma TFs reverse strand genes
-10Kb
TSS
TES
10Kb
I
ASTROCYTES
before transformation
after transformation
EZH2
H3K27Me3
TFs forward genes
TFs reverse genes
J
K
181,450 kb
34,400 kb
99,275 kb
103,350 kb
ChIP-Seq
Astrocytes after transformation
EZH2
H3K27Me3
SOX2
OLIG2
POU3F2
ASCL1

Supplemental Figure 3. Genome-wide binding of HOXDeRNA is associated with exclusive PRC2 removal from the transformation-induced genes. Related to Figure 2.

A. HOXDeRNA binds to all chromosomes. Relative chromosome length and number of HOXDeRNA peaks were calculated for each chromosome as percent of total.

B, C. ChIRP-seq raw read coverage signal, representing the HOXDeRNA binding at the TSS (+/- 5Kb) of the reverse strand of genes downregulated or upregulated after astrocyte transformation, is visualized as an average value for each group (B) or for each individual gene (C). The heatmap is accompanied by a colour scheme indicating the value of the raw read counts.

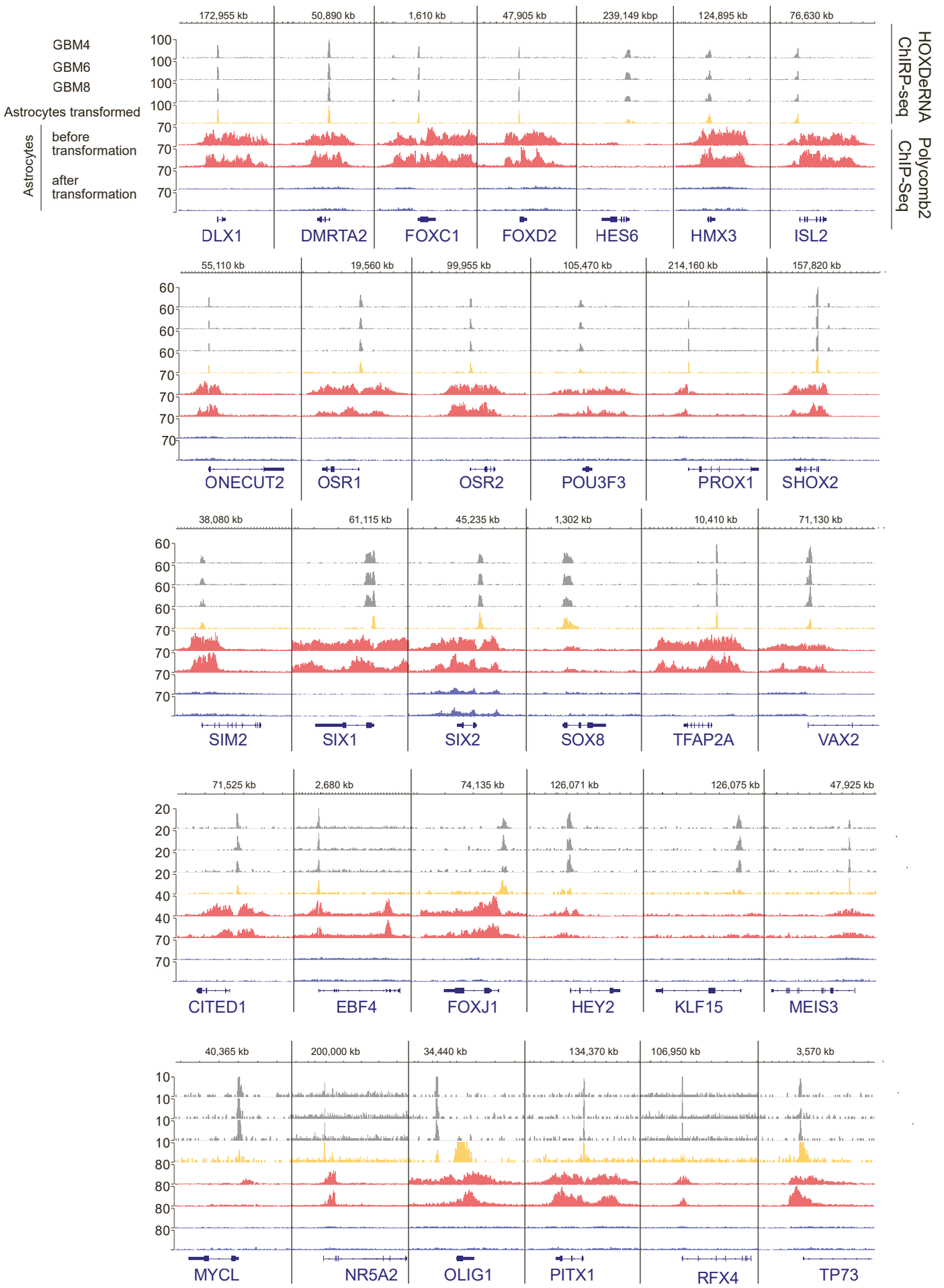
D, E. Epigenetic status of genes upregulated or downregulated in transformed astrocytes. H3K27Ac, H3K27Me3, and EZH2 ChIP-seq raw signals covering gene bodies of the negative strand (+/- 5Kb) were normalized to gene length and counted, followed by visualization of the average signal for 3 groups: genes upregulated after astrocytes transformation and bound by HOXDeRNA, downregulated after astrocyte transformation and HOXDeRNA-free, unchanged genes (D). Read coverage values for every individual gene body were visualized as heatmaps for negative strand genes, which are enriched in control or transformed astrocytes (E).

F. HOXDeRNA ChIRP-seq was validated by ChIRP-qPCR with primers specific for representative glioma TFs in glioma cells. The HOXDeRNA binding was visualized as Input % (mean, +SD, n=3).

G, H. Genes encoding key glioma TFs are bound by HOXDeRNA after astrocyte transformation. HOXDeRNA ChIRP-Seq raw read signal was normalized to gene length and visualized as mean value (G) or individual values (H) for each TF gene body (+/- 10Kb) on forward and reverse strands. The colour schemes represent the values of the raw read counts.

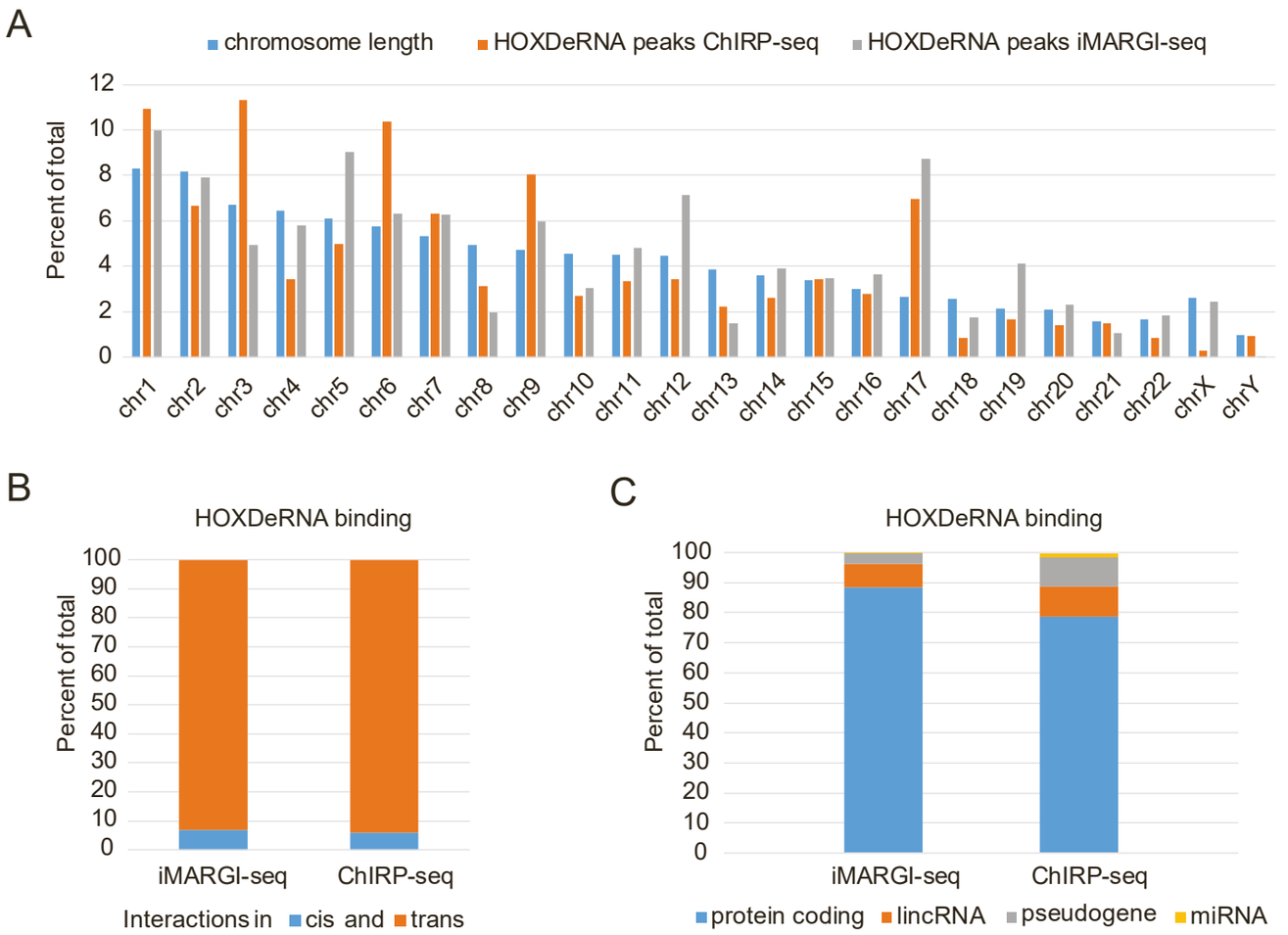
I, J. PRC2 is removed from genes encoding key glioma TFs after astrocyte transformation. The EZH2 and H3K27Me3 raw read coverage signal was normalized to gene length and visualized as mean value (I) or individual values (J) for each gene body (+/- 10Kb) on both forward and reverse strands. The colour scheme represents the values of the raw read counts.

K. H3K27Me3 and EZH2 ChIP-seq tracks in transformed astrocytes visualized for selected key glioma TF genes indicate incomplete removal of PRC2 activity in bulk chromatin after transformation. A different scale is presented in Fig. 2G bottom tracks to visualize relative levels in both non-transformed and transformed astrocytes.


GBM4
GBM6
GBM8
Astrocytes transformed
Astrocytes
before transformation
after transformation
HOXDeRNA ChIRP-seq
Polycomb2 ChIP-Seq
172,955 kb
50,890 kb
1,610 kb
47,905 kb
239,149 kbp
124,895 kb
76,630 kb
DLX1
DMRTA2
FOXC1
FOXD2
HES6
HMX3
ISL2
55,110 kb
19,560 kb
99,955 kb
105,470 kb
214,160 kb
157,820 kb
ONECUT2
OSR1
OSR2
POU3F3
PROX1
SHOX2
38,080 kb
61,115 kb
45,235 kb
1,302 kb
10,410 kb
71,130 kb
SIM2
SIX1
SIX2
SOX8
TFAP2A
VAX2
71,525 kb
2,680 kb
74,135 kb
126,071 kb
126,075 kb
47,925 kb
CITED1
EBF4
FOXJ1
HEY2
KLF15
MEIS3
40,365 kb
200,000 kb
34,440 kb
134,370 kb
106,950 kb
3,570 kb
MYCL
NR5A2
OLIG1
PITX1
RFX4
TP73

Supplemental Figure 4. HOXDeRNA binds to and removes PRC2 repression from multiple glioma TFs, including key TFs. Related to Figure 2.

HOXDeRNA ChIRP-seq tracks in GSCs and transformed astrocytes, aligned with PRC2 (H3K27Me3, EZH2) ChIP-seq coverage, before and after astrocyte transformation, and visualized for representative key glioma TF genes.
.

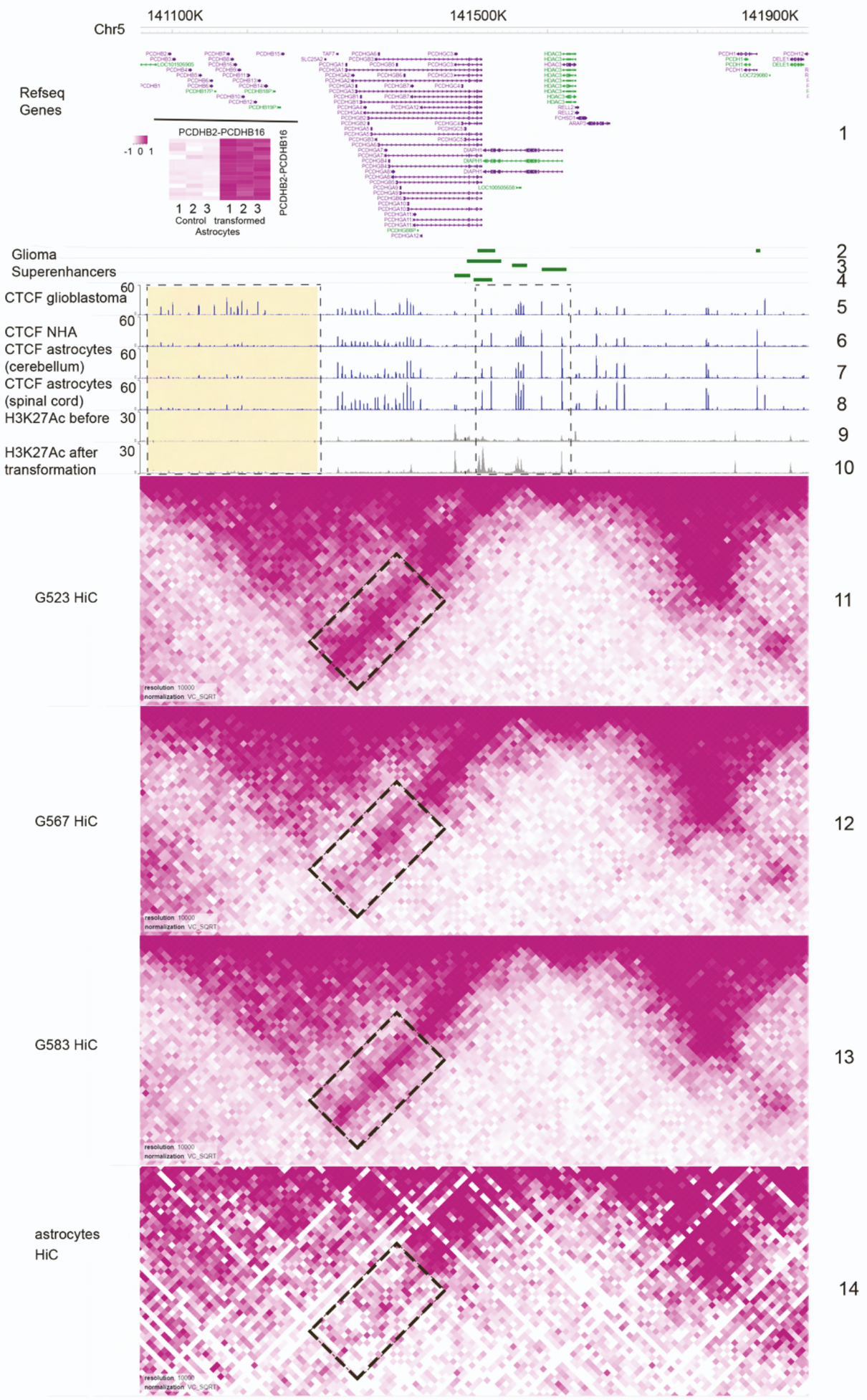


Supplemental Figure 5. Comparison of HOXDeRNA-bound regions detected by ChIRP-seq and iMARGI-seq. Related to Figure 3.

A. Distribution of HOXDeRNA-bound gDNA regions across all chromosomes measured by ChIRP-seq and iMARGI-seq in transformed astrocytes, along with chromosome lengths.

B. The percentage of cis-interactions on chromosome 2 (where HOXDeRNA gene is located) and trans-interactions (on other chromosomes), detected by ChIRP-seq and iMARGI-seq in transformed astrocytes.

C. Classes of HOXDeRNA-bound genes detected using ChIRP-seq or iMARGI-seq in transformed astrocytes. The percentage of total for each class is visualized.


Chr5
141100K
141500K
141900K
Refseq Genes
PCDHB2-PCDHB16
-1 0 1
1 2 3 1 2 3
Control transformed
Astrocytes
PCDHB2-PCDHB16
1
Glioma
2
Superenhancers
3
4
CTCF glioblastoma
60
5
CTCF NHA
60
6
CTCF astrocytes (cerebellum)
60
7
CTCF astrocytes (spinal cord)
60
8
H3K27Ac before
30
9
H3K27Ac after transformation
30
10
G523 HiC
11
resolution 10000
normalization VC_SQRT
G567 HiC
12
resolution 10000
normalization VC_SQRT
G583 HiC
13
resolution 10000
normalization VC_SQRT
astrocytes HiC
14
resolution 10000
normalization VC_SQRT

Supplemental Figure 6. Integration of astrocytes RNA-seq and ChIP-seq data with ENCODE and GEO datasets suggests a mechanism of Protocadherin (PCDHB) family upregulation in GBM. Related to Figure 4.

List of tracks from top to bottom:

1) protocadherin genes annotated with Refseq. A heatmap visualizing PCDHB2-16 RNAseq-based expression in control and transformed astrocytes, transformed to z-scores;

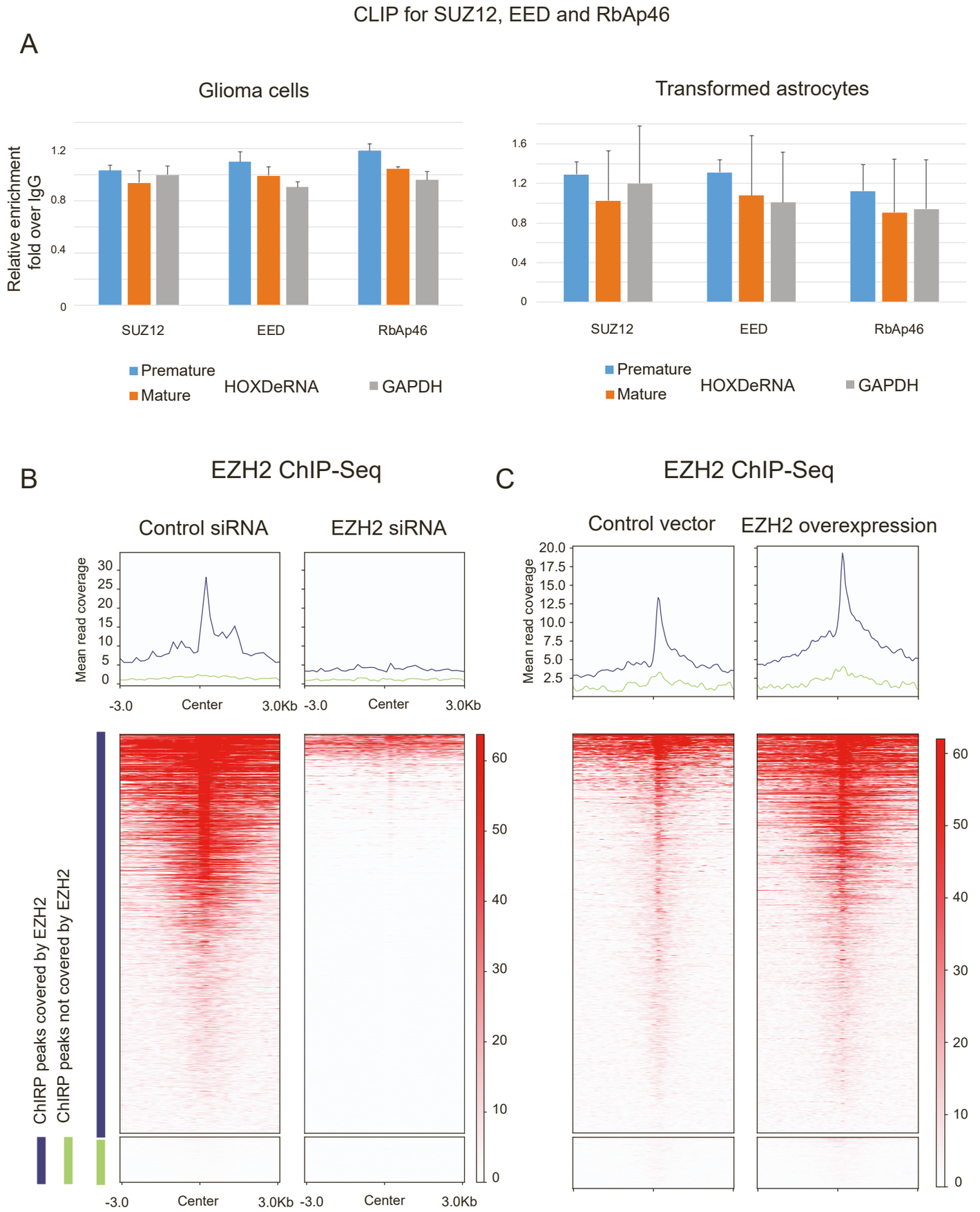
2-4) SE tracks in three GCSs (GSE121601);

5-8) CTCF ChIP-seq tracks from ENCODE database for GBM (GSM822303) and three types of astrocytes (GSM733765, GSM1022662, GSM749696). Genomic region with differential CTCF binding is highlighted in yellow;

9-10) H3K27Ac coverage before and after astrocyte transformation shows activation of GBM specific SEs (dotted lines);

11-14) 4 HiC maps for GSCs: G583, G567, G523 from EGAS00001003493, available for download at (https://wangftp.wustl.edu/hubs/johnston_gallo/), and astrocytes (GSE105194). A subset of HiC contacts representing SEs-protocadherins interactions are highlighted with dotted rectangles.

.

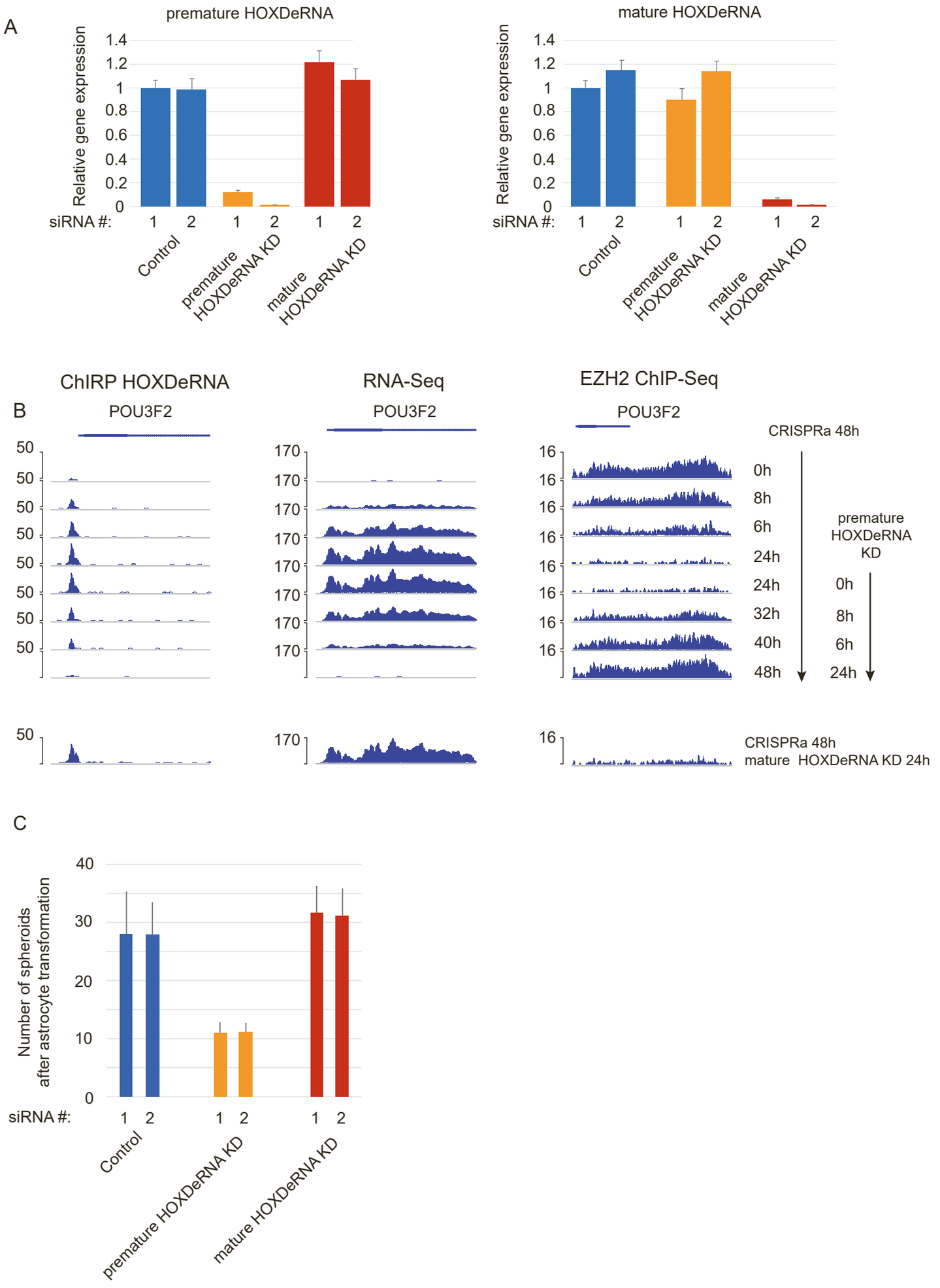

CLIP for SUZ12, EED and RbAp46
A
Glioma cells
Transformed astrocytes
Relative enrichment fold over IgG
1.2
0.8
0.4
0
1.6
1.2
0.8
0.4
0
SUZ12
EED
RbAp46
Premature
Mature
HOXDeRNA
GAPDH
B
EZH2 ChIP-Seq
Control siRNA
EZH2 siRNA
Mean read coverage
30
25
20
15
10
5
0
-3.0
Center
3.0Kb
ChIRP peaks covered by EZH2
ChIRP peaks not covered by EZH2
60
50
40
30
20
10
0
C
EZH2 ChIP-Seq
Control vector
EZH2 overexpression
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5

Supplemental Figure 7. Confirmation of EZH2 modulation on chromatin upon its siRNA-mediated KD and overexpression. Related to Figure 5D.

A. CLIP with EED, SUZ12, and RbAp46 antibodies, followed by qRT-PCR detection of pre-HOXDeRNA, spliced HOXDeRNA, and GAPDH mRNA, was performed in glioma LN229 cells and transformed astrocytes. EZH2/IgG ratios are demonstrated (n=3, mean+ SD).

B. EZH2 KD in transformed astrocytes leads to its disappearance from HOXDeRNA-bound chromatin. Transformed astrocytes were transfected with siRNAs (control or EZH2-targeting), followed by EZH2 ChIP-Seq. EZH2 reads were visualized as average (line graph) and individual values (heatmap) at HOXDeRNA ChIRP peaks (center +/- 3 Kb), either overlapping (blue) or non-overlapping (green) with EZH2.

C. EZH2 overexpression in transformed astrocytes increases EZH2 coverage at HOXDeRNA peaks. EZH2 reads were visualized as average (line graph) and individual values (heatmap) at HOXDeRNA ChIRP peaks (center +/- 3 Kb), either overlapping (blue) or non-overlapping (green) with EZH2.


A
premature HOXDeRNA
Relative gene expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
siRNA #:
1 2 1 2 1 2
Control
premature HOXDeRNA KD
mature HOXDeRNA KD
mature HOXDeRNA
B
ChIRP HOXDeRNA
RNA-Seq
EZH2 ChIP-Seq
POU3F2
CRISPRa 48h
0h
8h
6h
24h
24h
32h
40h
48h
premature HOXDeRNA KD
0h
8h
6h
24h
CRISPRa 48h
mature HOXDeRNA KD 24h
C
Number of spheroids after astrocyte transformation
40
30
20
10
0
siRNA #:
1 2 1 2 1 2
Control
premature HOXDeRNA KD
mature HOXDeRNA KD

Supplemental Figure 8. Pre-HOXDeRNA is responsible for TF derepression and spheroid formation. Related to Figure 5F.

A. The siRNA-mediated KD of either mature HOXDeRNA or pre-HOXDeRNA strongly and specifically decreases expression of the respective cognate targets in transformed astrocytes. Cells were transfected with two different siRNAs against control, premature, or mature HOXDeRNA. Relative RNA levels were measured at 24-48 hours post-transfection by qRT-PCR and normalized to GAPDH (mean +SD, n=3). At 72 hours post-transfection, pre-HOXDeRNA-targeting siRNA led to the reduced levels of mature HOXDeRNA as well (data not shown).

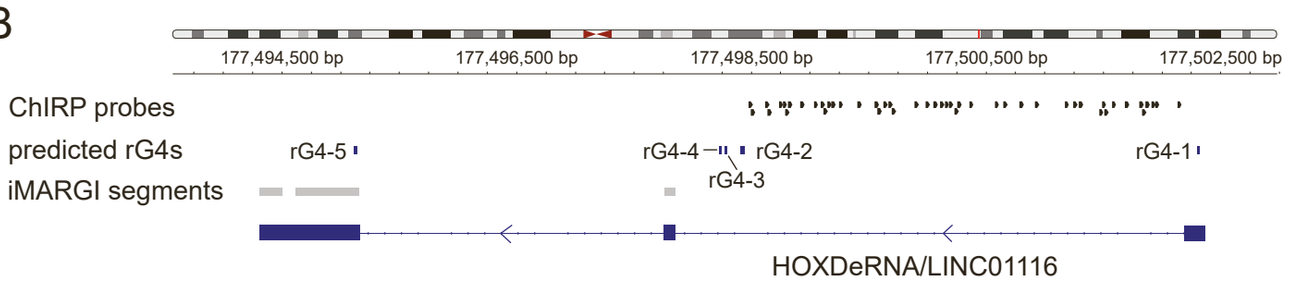
B. The kinetics of HOXDeRNA and EZH2 binding to POU3F2 gene are tightly associated with the genes' expression in transformed astrocytes. After HOXDeRNA activation with CRISPRa system, NHA cells were collected at different timepoints, followed by targeted HOXDeRNA ChIRP-Seq, RNA-Seq, and EZH2 ChIP-Seq. At 24h after CRISPRa, the cells were transfected with specific siRNAs against either premature HOXDeRNA or mature HOXDeRNA and analysed in similar ways.

C. KD of premature HOXDeRNA, but not mature HOXDeRNA, inhibits spheroid formation of HOXDeRNA-activated NHA. Following astrocyte transformation by HOXDeRNA, cells were transfected with two different siRNAs against Control, premature, or mature HOXDeRNA, and spheroids counted, with their numbers visualized as bar graphs (mean +SD, n=3).

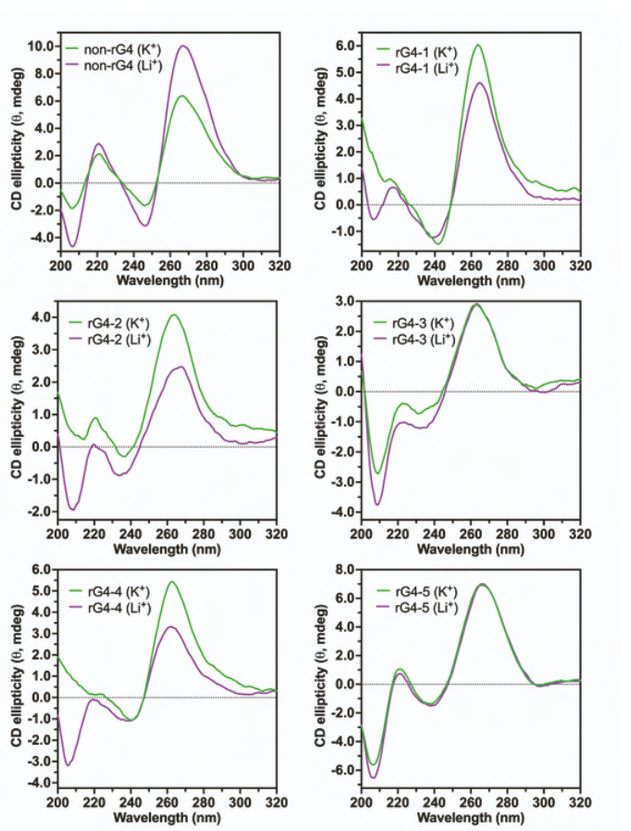
A

Position	Putative G4 forming sequences	G-Score	Name
62	GAGGGGAATGGGGAATCGGCAGGA	33	rG4-1
7241	CAGGCTGGAGTGTGGTGGCA	32	rG4-2
3937	CGGGCCGGGCGCGGTGGCT	34	rG4-3
3977	TGGGAGGCCGAGGCGGGCGGAT	34	rG4-4
4111	GGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGG	34	rG4-5
N/A	GAGAUGAATGAUGAATCGACAAGA	N/A	Control

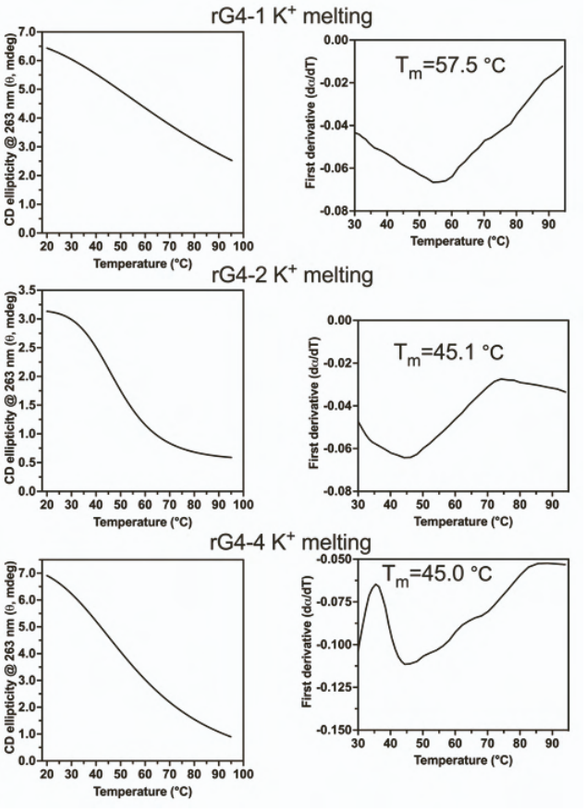
B



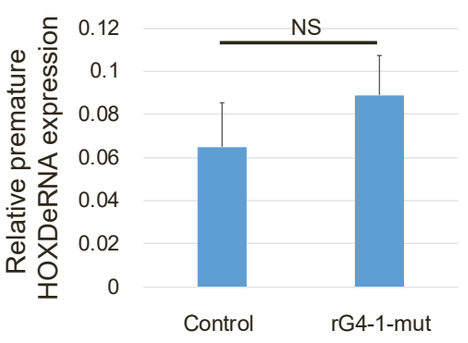
C



D



E



Supplemental Figure 9. Characterization of rG4 candidates. Related to Figure 6.

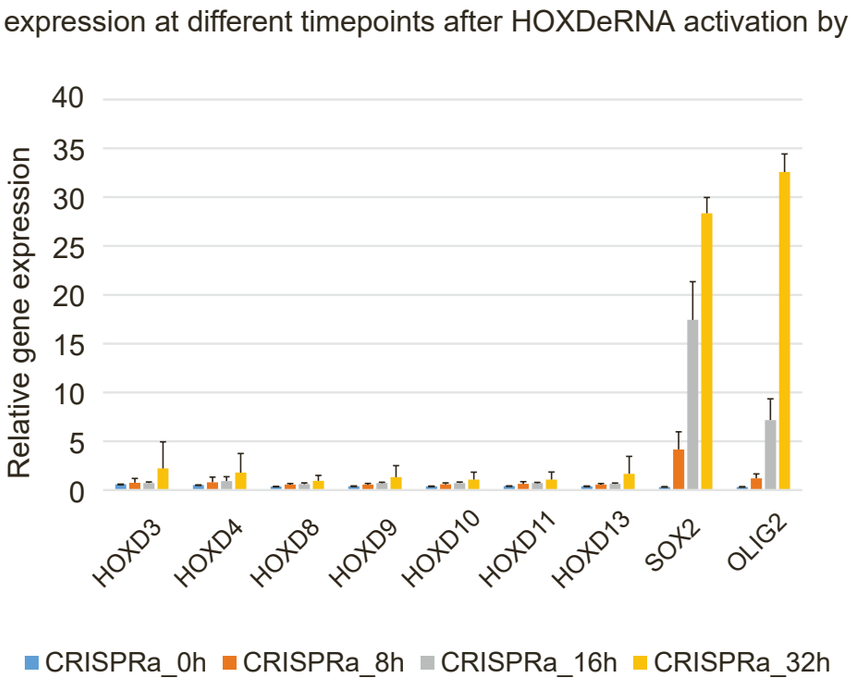
A. Five putative rG4 forming sequences detected with QGRS mapper tool. Search parameters: QGRS max length: 45, min G-Group size: 2, loop size: from 1 to 36.

B. Schematics of pre-HOXDeRNA with the following regions indicated: ChIRP probes (black), predicted G-quadruplexes (blue), and iMARGI-detected fragments bound to the upregulated genes (grey).

C. CD spectra of five putative rG4 forming sequences (rG4-1, rG4-2, rG4-3, rG4-4, and rG4-5), along with one non-rG4 sequence. rG4-1, rG4-2, and rG4-4 exhibit common rG4 CD features (a higher intensity peak at 263 nm under K+ environment).

D. CD-melting curves indicating the stability of predicted rG4 structures. rG4-1 showed the highest melting temperature of 57.5 °C.

E. Base editing of rG4-1 does not alter premature HOXDeRNA expression levels in transformed astrocytes. qRT-PCR analysis, normalized to GAPDH mRNA levels (mean + SD, n=3).



Supplemental Figure 10. Transformation driven SOX2 and OLIG2 genes are derepressed earlier than HOXD genes. Related to Figure 5F.

HOXD genes are barely derepressed at 32 hours post-HOXDeRNA induction in astrocytes, whereas SOX2 and OLIG2 are derepressed starting from 8 hours. mRNA levels were measured by qRT-PCR at different timepoints after HOXDeRNA activation. Ct values were normalized to GAPDH and visualized as bar graphs (mean +SD, n=3).