SCIENCE ADVANCES | RESEARCH ARTICLE

CANCER

Spatial 3D genome organization reveals intratumor heterogeneity in primary glioblastoma samples

Qixuan Wang[1]†, Juan Wang[1]†, Radhika Mathur[2], Mark W. Youngblood[1,3], Qiushi Jin[1], Ye Hou[1,4], Lena Ann Stasiak[1], Yu Luan[1], Hengqiang Zhao[1], Stephanie Hilz[2,5], Chibo Hong[2], Susan M. Chang[2], Janine M. Lupo[2], Joanna J. Phillips[2], Joseph F. Costello[2]*‡, Feng Yue[1]*‡

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Glioblastoma (GBM) is the most prevalent malignant brain tumor with poor prognosis. Although chromatin intratumoral heterogeneity is a characteristic feature of GBM, most current studies are conducted at a single tumor site. To investigate the GBM-specific 3D genome organization and its heterogeneity, we conducted Hi-C experiments in 21 GBM samples from nine patients, along with three normal brain samples. We identified genome subcompartmentalization and chromatin interactions specific to GBM, as well as extensive intertumoral and intratumoral heterogeneity at these levels. We identified copy number variants (CNVs) and structural variations (SVs) and demonstrated how they disrupted 3D genome structures. SVs could not only induce enhancer hijacking but also cause the loss of enhancers to the same gene, both of which contributed to gene dysregulation. Our findings provide insights into the GBM-specific 3D genome organization and the intratumoral heterogeneity of this organization and open avenues for understanding this devastating disease.

INTRODUCTION

Glioblastoma multiforme (GBM) is an aggressive and highly lethal brain tumor, comprising approximately 57% of all gliomas (*1*–*3*). Intratumoral heterogeneity (ITH), a well-established feature in GBM, denotes the presence of diverse tumor cell populations within a single tumor sample and constitutes a major contributor to therapeutic resistance and treatment failure (*4*–*7*). ITH of chromatin is a characteristic of GBM that has long been observed under the microscope, but the molecular underpinnings and consequences on gene expression are poorly understood.

The three-dimensional (3D) organization of the genome, which refers to the spatial organization and interactions of chromatin within the nucleus of a cell, plays a crucial role in transcriptional regulation through multiple levels of chromatin folding, including chromosome territories (*8*), A/B compartments (*9*), topologically associating domains (TADs) (*10*, *11*), and chromatin loops. Genome-wide chromatin folding structures can be captured using a variety of 3C-based technologies such as Hi-C (*9*), Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) (*12*), Hi-C chromatin immunoprecipitation (HiChIP) (*13*), or micro-C (*14*). When comparing 3D genome organization in cancer samples with normal samples, the widespread of copy number variations (CNVs) and structure variations (SVs) in cancer genomes added another layer of complexity. In particular, SVs—such as deletion, inversion, and translocations—can juxtapose enhancers or silencers from distance or even from another chromosome in proximity to near potent cancer-related genes and influence their expression, termed enhancer or silencer hijacking (*15*–*21*). In a recent work, Papantonis and colleagues (*22*) generated the Hi-C maps and studied enhancer hijacking in a variety of GBM cell lines, but so far, the study of 3D genome organization in GBM using primary patient samples has been limited.

Therefore, in this work, we conducted Hi-C experiments in a cohort of 21 tumor samples from nine primary GBM patients and three normal brain samples, generating the most comprehensive and highest resolution 3D genome dataset in GBM to date. To the best of our knowledge, this work represents the first large-scale 3D genome profiling in different regions of the same tumor in any cancer type. We sought to investigate GBM-specific 3D genome organization and the heterogeneity of the organization both between and within tumors.

RESULTS

Hi-C data generation in primary GBM samples

To comprehensively study 3D genome interactions of GBM patients, we performed 24 Hi-C experiments in nine patients with IDH wild-type GBM and three normal brain samples (Fig. 1A). By sampling multiple regions within the same patients, we sought to investigate intratumoral diversity, in addition to comparisons between patients. For patients 530, 524, and 529, we selected nine, three, and three samples, respectively, using a 3D spatial sampling method (details in Materials and Methods). The location and relative distance from centroid [defined as distance to tumor centroid/(distance to tumor centroid + distance to tumor periphery)] of the selected samples in these three patients are shown in the lower panel of Fig. 1 (A and B). Specimen locations were selected to optimize diversity in tumor evolution. For example, the tumor samples of P530 belong to two distinct lobes of the brain: temporal (sample #2, #4, and #5) and frontal (sample #10, #13, #14, #15, #18, and #19). Sample 4 and sample 18 of P530 correspond to regions that are physically closer to the tumor centroid, whereas the other samples of P530 were obtained from regions closer to the tumor periphery (Fig. 1B). In P524, all the three samples are located near the tumor periphery. For P529, sample 8 and sample 1 are closer to the tumor centroid and periphery, respectively, while sample 6 is located in between. In the remaining cohort of six patients with GBM, we

---

[1] Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [2] Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA. [3] Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [4] Institutes of Biomedical Sciences, College of Life Sciences, Inner Mongolia University, Hohhot, China. [5] Genentech Inc., 1 DNA Way, South San Francisco, CA, USA.

*Corresponding author. Email: joseph.costello@ucsf.edu (J.F.C.); yue@northwestern.edu (F.Y.)

†These authors contributed equally to this work.

‡These authors contributed equally to this work.

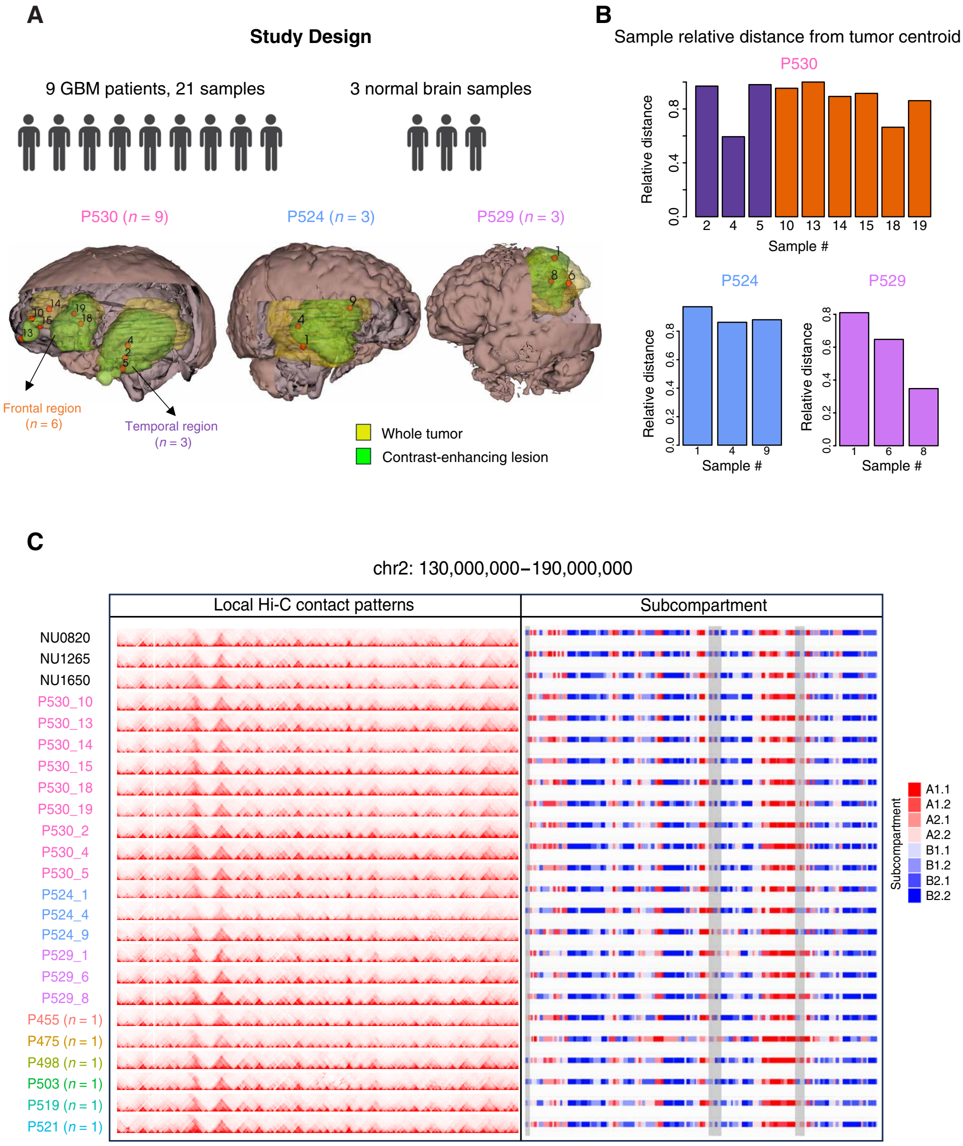

Fig. 1. Study design. (**A**) Study design including the number of patients and samples in this study and location of each spatially mapped samples within the brain. The yellow region is the whole tumor (T2-weighted hyperintense region) and the green region is the contrast-enhancing lesion (T1-weighted post-contrast). People icon was generated from BioRender. (**B**) Relative distance from centroid of each sample in P530, P524, and P529. Relative distance from centroid is defined as distance to tumor centroid/(distance to tumor centroid + distance to tumor periphery). (**C**) Local Hi-C contact patterns and corresponding subcompartment for the 24 samples on a representative region of chr2: 130,000,000–190,000,000. Example regions with variable subcompartments were highlighted in gray.

generated a single Hi-C library for each case and used the data for intertumoral comparisons. On average, we generated ~700 million paired-end reads for each Hi-C library (table S1). We show an example region of Hi-C data across all the 24 samples in Fig. 1C. For 12 of the Hi-C samples, we also have the RNA sequencing (RNA-seq) and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) in the same region and thus can study the effect of 3D genome organization on gene regulation for these spatially mapped samples.

Impact of extensive SVs on PC-based A/B compartment accuracy

The genome can be partitioned into A and B compartments, where A compartment is more associated with open chromatin regions, while B compartment is related to condensed regions (*9*, *23*). Recently, it has been shown that there are cancer- and subtype-specific compartment changes that can distinguish distinct pathologies, such as in colon cancer (*24*), prostate cancer (*25*), or leukemia (*18*). To profile the genome compartmentalization of GBM, we first derived the first principal component (PC1) value of the Hi-C matrices at 100-kb resolution for each sample and then assigned each bin into A (PC1 value > 0) or B compartments (PC1 value < 0).

However, we observed that complex SV events would disrupt PC1 value calculation (the second track in fig. S1, A and B). For example, there are extensive SVs in sample P475, and the PC1 values for these regions are mostly missing (fig. S1A). Similarly, in sample P524_9, the derived PC1 values are also missing at SV-affected regions (fig.

S1B), potentially leading to missed or wrong assigned A/B compartments. The main reason is that when PC1 values are derived (*9*), the approach does not include consideration for CNVs and SVs, but this alteration of the genome might alter the correlations step in PC calculation and lead to the inaccurate PC1 inference.

Identification of subcompartment with CALDER

Therefore, we investigated whether subcompartments (*26*, *27*) were less affected by SVs. In particular, we run the CALDER software (*27*) to identify subcompartments at 100-kb resolution (Fig. 2A and fig. S1, A and B). First, we noticed that CALDER were able to define subcompartments in the regions where PC1 failed to provide annotations (third and fourth tracks in fig. S1, A and B). In general, CALDER first calculates a correlation matrix based on a compressed Hi-C map and then identifies domains using TopDom based on the correlation matrix. The domains would then be clustered and assigned into eight subcompartments: A1.1, A1.2, A2.1, A2.2, B1.1, B1.2, B2.1, and B2.2. We found that the Hi-C map compressing, domain identification, and clustering steps largely mitigated the influences from SVs, confirming what the authors suggested in their paper (*27*). Specifically, we observed that at the complex SVs regions, CALDER was able to call meaningful subcompartments at most of the regions (fig. S1, A and B). To evaluate the accuracy of subcompartment, we further compared its states with both gene expression (Fig. 2B and fig. S1, C to E) and ATAC-seq signals (Fig. 2C and fig. S1, F to H). Both gene expression and ATAC-seq read counts exhibited patterns consistent with the subcompartments, confirming the reliability of subcompartment identification in each sample. In addition, we checked whether subcompartments in normal sample were associated with histone modifications using public normal brain histone marks ChIP-seq data (*28*). We observed that subcomartment A1.1, A1.2, A2.1, and A2.2 were more enriched with active enhancers (H3K27ac and H3K4me1), active promoters (H3K4me3 and H3K9ac), and transcribed gene markers (H3K36me3), while B1.1, B1.2, B2.1, and B2.2 were more enriched with polycomb repressed (H3K27me3) and heterochromatin markers (H3K9me3) (fig. S1I).

Compartment dynamics across different patients

First, we performed genome-wide classification based on subcompartments using Uniform Manifold Approximation and Projection (UMAP) (Fig. 2D), which revealed six clusters of these 24 samples and notably suggested unique patterns within individual tumors. We observed that all temporal region samples (#2, #4, and #5) of patent P530 were in the same cluster, while the frontal region samples were further divided into two different clusters. Sample 1 of P524 was grouped in the same cluster with the three normal brain samples (NU0820, NU1650, and NU1265), potentially due to its low tumor purity [~0.04, predicted by whole exome sequencing (WES) data]. We also performed clustering analysis of the Hi-C data using t-distributed stochastic neighbor embedding (t-SNE), and the result was largely similar (fig. S2A). To investigate whether Hi-C and RNA-seq generate similar clustering results, we used gene expression to cluster 18 samples in which we have matching RNA-seq data using both UMAP (fig. S2B) and t-SNE (fig. S2C). We observed that UMAP and t-SNE with RNA-seq data generated identical clusters. When we compared the clusters based on Hi-C versus RNA-seq, the results were similar except that P524_4 was in different clusters. This suggests that subcompartments are robust features that can be used to reflect sample specificity.

Next, we investigated the compartment dynamics across different samples. We defined the top, middle, and bottom variable bins across all 24 samples according to their subcompartment annotation. Unexpectedly, we found that the top 10% most variable bins were mostly in sub–B compartment (Fig. 2E, left), while the top 10% most invariable bins were in sub–A compartment (Fig. 2E, right). To validate this observation, we examined the expression of genes within these regions and observed that genes in the most stable bins had higher expression than those located in the most variable bins (fig. S3B). Similarly, the fraction of ATAC-seq signals in the most stable bins was also higher than that in the most variable bins (fig. S3C), supporting the observation made by subcompartment analysis. To evaluate the impact of resolution on compartment analysis, we performed the same analysis at 50- and 250-kb resolutions (fig. S3, D and E) and found the most stable regions were still mostly located in sub–A compartment at these resolutions. The most variable regions were still predominantly located in sub–B compartments at 50-kb resolution but were more evenly distributed between sub–A and sub–B compartments at the 250-kb resolution. Together, these results indicate that the transcriptionally more active regions are more stable, while the transcriptionally more repressive regions are more dynamic across different samples.

GBM sample specific sub–A/B compartments

To investigate the GBM sample–specific sub–A/B compartment, we compared the Hi-C data in our GBM samples with normal sample. We defined a region as A-to-B switch when this bin was annotated as any of the four A subcompartments in normal samples but was annotated as any of the four B subcompartments in a cancer sample. On average, we observed 1733 (~5.78% of the genome) A-to-B–switched regions and 2035 (~6.78% of the genome) B-to-A–switched regions (Fig. 2F). When we further dissected the sub–A to sub–B switch based on the eight subcompartment states, we noticed that most of the subcompartment switches were between neighboring states (Fig. 2G). In addition, more of the sub–A/B switches happened between A2 and B1 subcompartments. For example, when comparing normal sample NU0820 versus P524_9, 28.5% of A2 regions switched to sub–B compartments, and 35% of B1 regions switched to sub–A compartments, whereas 5.6% of A1 regions switched to sub–B compartments and 7.3% of B2 regions to sub–A switches (fig. S4, A to D). Also, the expression of genes located in these switched regions showed concordant pattern (fig. S4, E and F).

Next, we investigated whether there are recurrent A/B switches between GBM samples and normal samples and identified 360 recurrent A-to-B–switched regions and 624 recurrent B-to-A regions (Fig. 2H). All samples were separated into three groups, two GBM groups and one normal group, and group 1 GBM samples in general have higher tumor purity scores than group 2 samples (table S2). The gene expression levels were consistent with the subcompartment change pattern (Fig. 2I). Genes were either up- or down-regulated when they were in B-to-A or A-to-B subcompartment switching regions. For example, *CHI3L2*, a gene associated with poor prognosis and immune infiltration in gliomas (*29*), its expression level was much lower when it was in B compartment in normal samples than its expression in GBM samples when it was in A compartment (Fig. 2J, top). On the contrary, tumor suppressor genes *CAVIN2* and *TMEFF2* were down-regulated in tumor samples versus normal samples, and the expression patterns were consistent with their A/B compartment annotations (Fig. 2J, bottom).

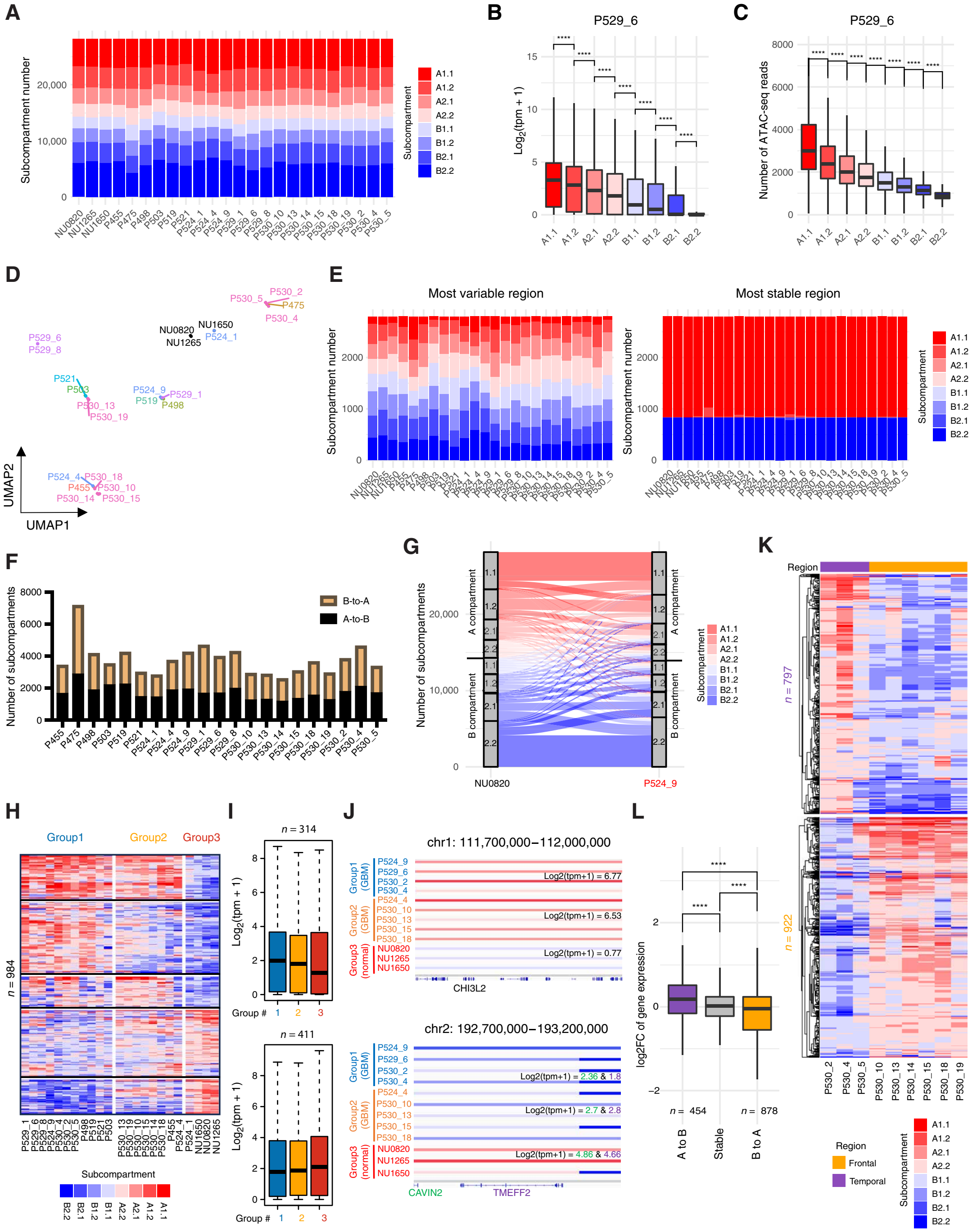

Fig. 2. Subcompartment analysis. (**A**) Barplot of number of each subcompartment in each sample. (**B** and **C**) Boxplot of gene expression (B) and ATAC-seq read counts (C) in each subcompartment in P529_6. The center line denotes the median, and the top and bottom of the boxes denote the first and third quartiles, respectively. *P* values calculated using two-sided Wilcoxon rank-sum test. (**D**) UMAP clustering based on subcompartments. (**E**). Barplots of number of each subcompartment at most variable regions (left) and most stable regions (right) in each sample at 100-kb resolutions. (**F**) Barplot showing number of A/B switched subcompartments between each GBM sample and normal sample. (**G**) Sankey diagram showing subcompartments changes from NU0820 to P524_9. (**H**) Heatmap of recurrent A/B-switched regions between GBM samples and normal samples. (**I**) Boxplots showing the average gene expression in each group for the top three clusters (top boxplot) and the bottom two clusters (bottom boxplot) from the heatmap in (H). The numbers of genes were labeled on top of the boxes. (**J**) Subcompartment tracks showing examples of B-to-A (top) and A-to-B (bottom) compartment switches between normal samples (group 3) and GBM samples (groups 1 and 2) in (H). The average expression of *CHI3L2*, *CAVIN2*, and *TMEFF2* for three groups was labeled in black, green, and purple. (**K**) Heatmap of A/B-switched regions between P530 temporal and frontal region samples. The numbers of A/B-switched subcompartments were labeled on the left side. (**L**) Boxplot of $\log_2$ fold change ($\log_2$FC) of gene expression in A-to-B, stable, and B-to-A regions between P530 temporal and frontal region samples. Log_2FC is defined as average gene expression in P530 temporal region samples/frontal region samples. The numbers of genes located in A-to-B– and B-to-A–switched regions were labeled at the bottom of the box. *P* values calculated using two-sided Wilcoxon rank-sum test.

Intratumor heterogeneity in sub–A/B compartments

Next, we focused on the three patients (P524, P529, and P530) with spatially mapped samples to further investigate the intratumoral heterogeneity in sub–A/B compartmentalization. Within the same patient, genome-wide sub–A/B switch also occurred (fig. S5, A to E), and genes located in these compartment-switching regions showed concordant expression changes (fig. S5F). Genes located in A-to-B regions were more associated with immune process regulation, while genes located in B-to-A regions were more related to synapse organization, transmission, and neuron development and differentiation (fig. S5G).

Previous studies have revealed that GBMs are located in both frontal and temporal regions in the brain (*30*); however, the difference of epigenetic regulatory programs between these two regions has not been well characterized. Taking advantage of our spatially mapped samples from the temporal and frontal lobes in P530 (three samples from temporal and six from frontal lobe), we investigated whether there were region-specific 3D genome and epigenomic features. In this section, we focused on the subcompartment analysis. We identified 797 sub–A to sub–B and 922 sub–B to sub–A compartment switch bins among these lobes (Fig. 2K). Genes located in the sub–A/B switching regions showed consistent changes of expression patterns with the subcompartment switches (Fig. 2L). In particular, we noticed that oncogene *ETV1* was located in P530 temporal-specific A regions and exhibited higher expression levels in temporal region than in frontal region (fig. S6A). We also observed frontal-specific A compartment genes, such as *NAV3* (fig. S6A). Gene Ontology shows that genes located in A-to-B–switched regions between temporal and frontal region samples were more associated with regulation of synapse structure or activity and dendrite development, whereas genes located in B-to-A–switched regions were more related to leukocyte migration and immune response (fig. S6B), suggesting a potentially higher tumor infiltration in the frontal regions from patient P530.

Next, we asked whether the within patient sample differences would be larger or smaller than the between patient sample differences. To do this, we used samples from P529, P524, and P530 as three within patient groups and the rest of the samples as between patient group. We calculated the adjusted stratum-adjusted correlation coefficient (SCC) score using HiCRep software (*31*) for samples in each group. We observed that the SCC scores in between patients' group were generally lower than those in the other three groups (within patient) (fig. S7), indicating potential larger differences between different patients.

GBM-specific chromatin interactions

To further investigate the chromatin loop level features in GBM, we identified chromatin loops at 10-kb resolution for each sample using the Peakachu software (*32*). On average, we identified 22,721 chromatin loops from GBM samples (Fig. 3A and table S3). For samples with ATAC-seq data, we stratified loops into enhancer-promoter (E-P) loops, promoter-promoter (P-P) loops, and promoter-none (P-N) loops. We found that genes with only E-P loops, only P-P loops, or both E-P and P-P loops showed significant higher expression comparing to genes with P-N loops (Fig. 3B and fig. S8A). In addition, we found that the number of enhancers is positively correlated with gene expression (Fig. 3C and fig. S8B). These results indicate that E-P and P-P loops can both contribute to gene expression activation, and more enhancers in the E-P loops would increase the activation effect.

We further identified GBM-specific chromatin loops as loops that were in each GBM sample but not in all three normal samples (fig. S9, A and B, and table S4). On average, we identified 726 GBM-specific loops per sample. For the samples with ATAC-seq data, we further identified GBM sample–specific E-P loops (on average, 138 loops) (Fig. 3, D and E). Genes with GBM-specific E-P loops showed higher expression in GBM sample than in normal samples (Fig. 3F). Gene Ontology showed that these genes were related to nervous system development, immune response, and axon development and guidance (fig. S9C).

Intratumor heterogeneity in chromatin interactions at ultrahigh resolution

To study region specific features, we aggregated all temporal specimens (sample #2, #4, and #5) into a single "temporal region combined" sample and combined all frontal specimens (sample #10, #13, #14, #15, #18, and #19) into a "frontal region combined" sample. Collectively, this aggregated more than 6 billion reads into two regional samples (4.46 billion in frontal versus 2.58 billion in temporal), enabling an ultrahigh-resolution view of 3D genome structures with a 2-kb resolution (Fig. 3G).

On the basis of this ultrahigh-resolution 3D genome datasets, we identified 1282 P530 temporal region–specific and 1311 frontal region–specific chromatin loops (table S5) with concordant gene expression patterns (Fig. 3H and fig. S10A). We noticed a P530 temporal region–specific chromatin loop near *API5*, with higher expression in temporal region samples than in frontal region samples (Fig. 3I, left). High expression of *API5* has been associated with worse GBM patient survival (*33*). In addition, a unique loop near *SELL* was observed in the frontal region samples, aligning with increased expression of this gene in frontal region samples in comparison to the temporal region samples (Fig. 3I, right). These results indicate intratumoral heterogeneity of chromatin loops between the temporal and frontal regions.

Intratumoral heterogeneity in chromatin interactions were also evident in other patients. We identified 602 P524_9-specific loops as validated by the aggregate peak analysis (APA) plots in fig. S10B. Figure S10C showed an example of sample P524_9-specific loop near the *ETV1* and *AGMO* genes, whose expression were higher in P524_9 than in both P524_1 and normal samples (fig. S10D). The genes located in P524_9-specific E-P loop anchors also showed higher expression in P524_9 than in both P524_1 and normal samples (fig. S10E). These results demonstrated the intratumoral heterogeneity of chromatin interactions in patients with GBM and indicated its potential roles in leading to heterogeneity in transcription regulation.

TAD boundary disruption by CNVs and SVs

To explore 3D genome features at TADs level, we used cooltools (*34*) software to compute the insulation scores at 40-kb resolution and defined TAD boundaries. For each sample, we identified ~3000 TADs (Fig. 4A). Next, we investigated whether CNV disrupted TAD structures. First, we derived CNV profiles using the Hi-C data for all 21 GBM samples with the NeoLoopFinder software (*35*), and their profiles were depicted in Fig. 4B and fig. S11. We observed the signature CNVs of GBM in most of the samples, such as chr7 amplification or the loss of chr10. To study the impact of CNVs on TADs, we compared the CNVs called from cancer samples with the TADs

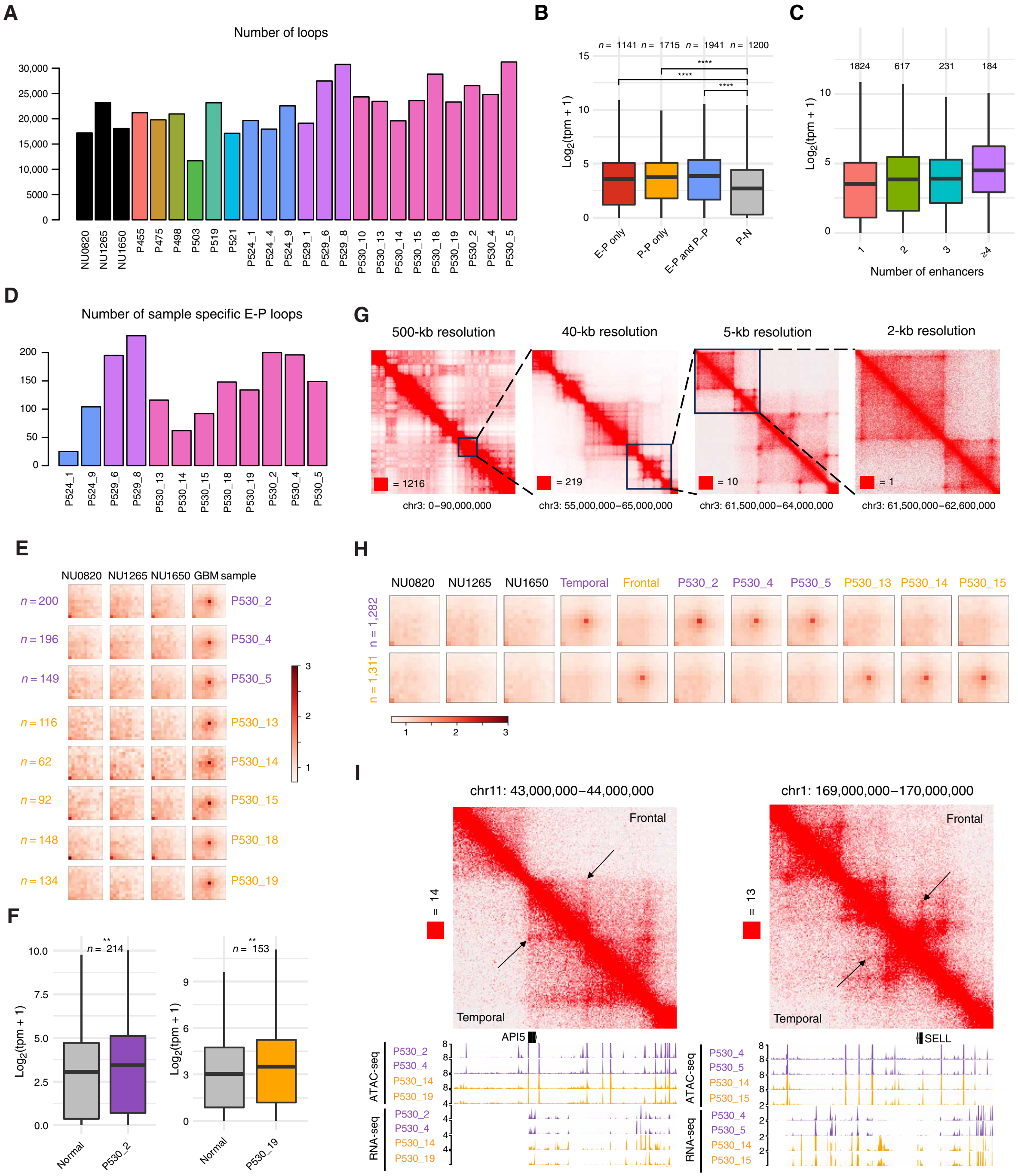

Fig. 3. GBM-specific chromatin interactions and intratumoral heterogeneity in chromatin interactions. (**A**) Barplot showing number of predicted chromatin loops in each sample. (**B**) Boxplot of expression of genes located in E-P only, P-P only, E-P, and P-P overlap and P-N loop anchors in P524_9. The numbers of genes in each category were labeled on top. *P* values calculated using two-sided paired Wilcoxon rank-sum test. (**C**) Boxplot of expression of genes with different numbers of enhancers (1, 2, 3, and ≥4) in P524_9. The numbers of genes in each category were labeled on top. (**D**) Barplot showing number of GBM sample–specific E-P loops in each sample. (**E**) Aggregate peak analysis (APA) showing interaction frequency of P530 sample–specific E-P loops at three normal samples and each P530 sample at 10-kb resolution. The numbers of specific E-P loops were labeled on the left side. (**F**) Boxplot showing expression of genes located in P530_2 (left)– and P530_19 (right)–specific E-P loop anchors in normal sample (average of three normal samples) and P530_2 and P530_19. The number of genes was labeled on top. *P* values calculated using two-sided paired Wilcoxon rank-sum test. (**G**) Multiresolution Knight-Ruiz (KR) matrix-balance normalized Hi-C maps of example regions on chr3 of P530_temporal region combined sample. (**H**) APA showing interaction frequency of P530 temporal and frontal region combined specific loops at three normal samples, P530 temporal and frontal region combined sample, and each individual P530 sample, respectively, at 10-kb resolution. The numbers of specific loops were labeled on the left side. (**I**) KR balanced Hi-C map at chr11: 43,000,000 to 44,000,000 (left) and chr1: 169,000,000 to 170,000,000 (right) showing P530 temporal region– and frontal region–specific chromatin interactions near *API5* and *SELL*. ATAC-seq and RNA-seq signal in temporal (purple) and frontal region samples (orange) for these regions were shown in the bottom panel.

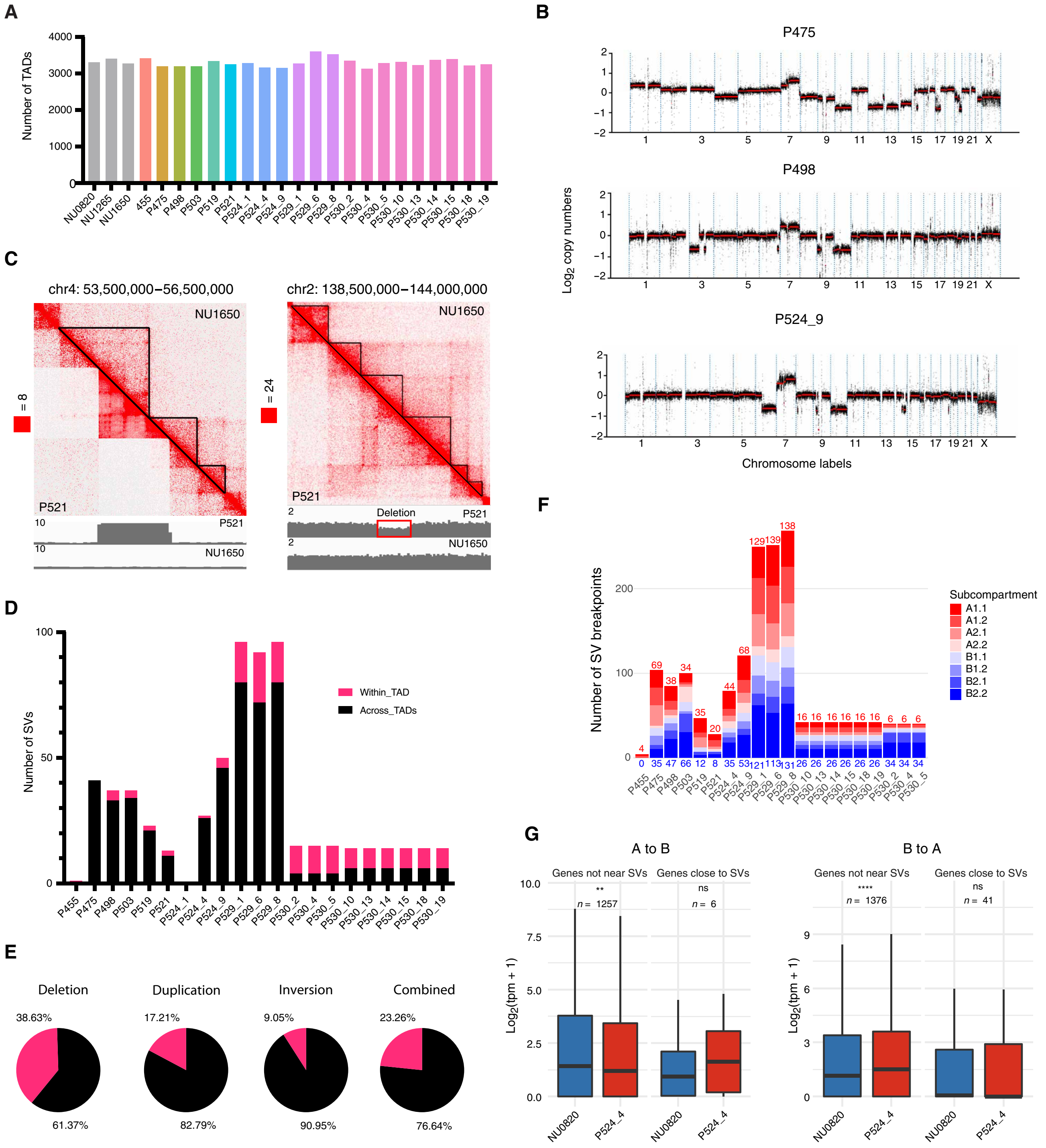

Fig. 4. The impact of SVs and CNVs on 3D genome. (**A**) Barplot showing number of TADs in each sample. (**B**) Hi-C data inferred CNV profiles for P475, P498, and P524_9. (**C**) KR matrix-balanced Hi-C map of chr4: 53,500,000 to 56,500,000 (left) and chr2: 138,500,000 to 144,000,000 (right) in P521 (lower triangle) and NU1650 (upper triangle) from JuiceBox. TADs in normal sample were highlighted by black triangles. CNV tracks for the same region were presented below the Hi-C map. (**D**) Barplot showing number of SVs whose two breakpoints are in different TADs (Across_TADs) or within the same TADs (Within_TAD) in each GBM sample. (**E**) Percentage of SVs whose two breakpoints are in different TADs (Across_TADs) or within the same TADs (Within_TAD) in each SV category (deletion, duplication, and inversion) and all three categories combined. SV events in this figure are the combined SV events from all 21 GBM samples. (**F**) Barplot of number of SV breakpoints located in each subcompartment of normal sample NU0820 in each GBM sample. Numbers of SV breakpoints that were located in A/B compartment of NU0820 were labeled in red/blue. (**G**) Boxplot of expression of genes close to SV breakpoints and genes not near SV breakpoints in NU0820 (normal) versus P524_4 A-to-B–switched (left) and B-to-A–switched (right) regions. The center line denotes the median, and the top and bottom of the boxes denote the first and third quartiles, respectively. The numbers of genes close to SV breakpoints and genes not near SV breakpoints were labeled on top of the box. *P* values calculated using two-sided paired Wilcoxon rank-sum test.

identified from normal samples. Majority of the two breakpoints of the same CNVs were located in different TADs in normal sample (fig. S12A), suggesting that they disrupted the TAD boundaries. We showed three such examples in Fig. 4C and fig. S12B, where CNVs disrupted the TAD boundaries defined in NU1650. The scenarios included both amplifications (Fig. 4C, left and fig. S12B) and a deletion (Fig. 4C, right).

We performed a similar analysis to investigate how SVs affected TAD boundaries. SVs have been shown to play an important role in cancer progression (*36*). To detect SV events in patients with GBM, we used our recently developed method EagleC (*37*). Most of the SVs (including deletion, duplication, and inversion) expanded across different TADs (Fig. 4D), suggesting that they disrupted TAD boundaries. Next, we dissect how SVs disrupted TAD boundaries according to their type: 61.37% of the deletions, 82.79% of the duplications, and 90.95% of the inversions potentially disrupted TAD boundaries (Fig. 4E and fig. S12C). We also wanted to note that this analysis was based on SVs identified by Hi-C, which tends to capture large scale SVs.

SVs, subcompartment, and gene expression

Next, we explored whether there is a higher occurrence of SVs within A or B compartments by comparing all SV breakpoints in each GBM sample with the subcompartment annotations from normal samples. SVs were more enriched in B compartment for three samples in P530, but not in the other six samples (Fig. 4F and fig. S13A). Similarly, SVs were more enriched for A compartment in P475, P519, P521, and P524_4, but not in other samples. Overall, we did not observe any clear pattern (Fig. 4F and fig. S13A).

Next, we asked whether the relationship between compartment switches and gene expression was influenced by SVs. We first identified A/B-switched regions between the normal sample and each GBM sample. We then categorized genes located in these regions into two groups based on whether they were within 500 kb of SV breakpoints (genes close to SVs) or outside 500 kb (genes not near SVs) in each sample. We found that the change of expression of genes away from SVs was correlated with A/B compartment switches (Fig. 4G and fig. S13B), consistent with previous reports. However, for genes close to SVs, the relationship between compartment switch and gene expression change was not clear (Fig. 4G and fig. S13B), suggesting potential influences by local chromatin state and a more complex relationship.

On the basis of the CNVs and SVs derived from Hi-C data, we were able to construct phylogenetic tree for three patients (P524, P529, and P530) with multiple Hi-C samples (Fig. 5A and fig. S14). The results based on CNVs (Fig. 5A) or SVs (fig. S14) were highly similar. Samples of P530 were separated into two subclones, one subclone contains three samples from temporal region and the other one contains six samples from frontal region (Fig. 5A). Similarly, we identified two clones in P529 and P524. The phylogenetic reconstructions based on CNVs and SVs resembled those based on mutations in Mathur *et al.* (*38*). These results demonstrated that we could use Hi-C data to perform phylogenetic analysis.

As we have multiple samples from P530, and we further explored the P530 donor-specific, region-specific, and sample-specific SVs. To define P530 donor-specific SVs, we used the SVs found in all nine samples from P530, but not in any other samples. In total, we identified 3 P530 donor-specific SVs, 12 P530 temporal region-specific SVs, and 13 P530 frontal region-specific SVs (table S6). We did not find any P530 sample-specific SVs, which is consistent with the phylogenetic tree analysis (fig. S14).

Enhancer hijacking and amputation in patients with GBM

To further investigate the influences of SVs on gene expression, we predicted SV-induced neo-loops for each sample using NeoLoopFinder (fig. S15A) (*35*). Then, we checked whether there were cancer-related genes located in the neo-loop anchors and observed a variable percentage (ranging from 10.81 to 75%) across different samples (fig. S15, B and C). In total, we identified 41 recurrent cancer-related genes (present in at least two samples) within neo-loop anchors, 4 of which were present across different patients (fig. S15D). Figure S15 (E and F) showed that oncogene *MELK* was located at neo-loop anchors resulting from inversion and deletion events in P524_9 and P498.

Next, we further investigated the impact of enhancer hijacking on gene expression. Unexpectedly, we find many cases where the same targe gene was involved in both enhancer hijacking (gain of enhancer) and enhancer amputation (the loss of its previous enhancers). We described such scenarios in Fig. 5 (B and C): (i) The E1 enhancer was deleted due to a deletion event (Fig. 5B) and (ii) due to inversion or translocation, the E1 enhancer was moved away from its original target gene G1 and no longer control its expression (Fig. 5C). In both cases, gene 1 gained another enhancer (E2) either due to either deletion or translocation.

As shown in Fig. 5 (D and E), P524_9 has a ~ 7-Mb deletion on chromosome 14 (Fig. 5D, black circle) near the *ZFP36L1* gene, which has been shown to play an important role in glioma cell proliferation regulation (*39*, *40*). We examined Hi-C data in P524_1 and did not observe the deletion in this sample (fig. S16, A and B), suggesting that this was a sample-specific deletion. We reconstructed the Hi-C map of P524_9 and observed a neo-TAD formation in P524_9 (Fig. 5E), which enabled the linkage between the *ZFP36L1* gene and a gained block of enhancers. In this scenario, the expression of *ZFP36L1* was increased compared with P524_1 {6.87 versus 5.81, $\log_2$[transcripts per million (tpm) + 1] transformed}.

In another example, there was a region-specific deletion on chr11 in P530_5 (fig. S16C, black circle) near *DLG2*, a known tumor suppressor in osteosarcoma and neuroblastoma (*41*–*44*). This deletion was not observed in sample P530_13. The deleted region contained several enhancers (marked gray rectangular in fig. S16D), likely linked to *DLG2*. After reconstructing the Hi-C map in P530_5, we detected a potential enhancer hijacking event in P530_5 (fig. S16E). However, despite this gained linked enhancer, the expression of *DLG2* in P530_5 was still decreased compared with P530_13 [2.56 versus 3.8, $\log_2$(tpm + 1) transformed], potentially due to the loss of multiple enhancers in the original genomic loci.

Last, there was a translocation between chr1 and chr11 in P530_18, but not in P530_4. We reconstructed Hi-C map in P530_18 and detected a potential enhancer-hijacking event involved with *ZBED6*, which has been associated with breast cancer susceptibility and cell growth of human colorectal cancer cells (*45*–*47*). This translocation led to potential gain of several enhancers for *ZBED6* from chr11 (Fig. 5F, gray rectangular shade). Again, despite the newly formed enhancer-hijacking events, the expression of *ZBED6* was decreased when compared with P530_4 where there was no SV in this region or enhancer hijacking [1.17 versus 2.87, log2(tpm + 1) transformed]. When we checked the upstream of the translocation breakpoint on chr1 in P530_4, we found that the SV also led to

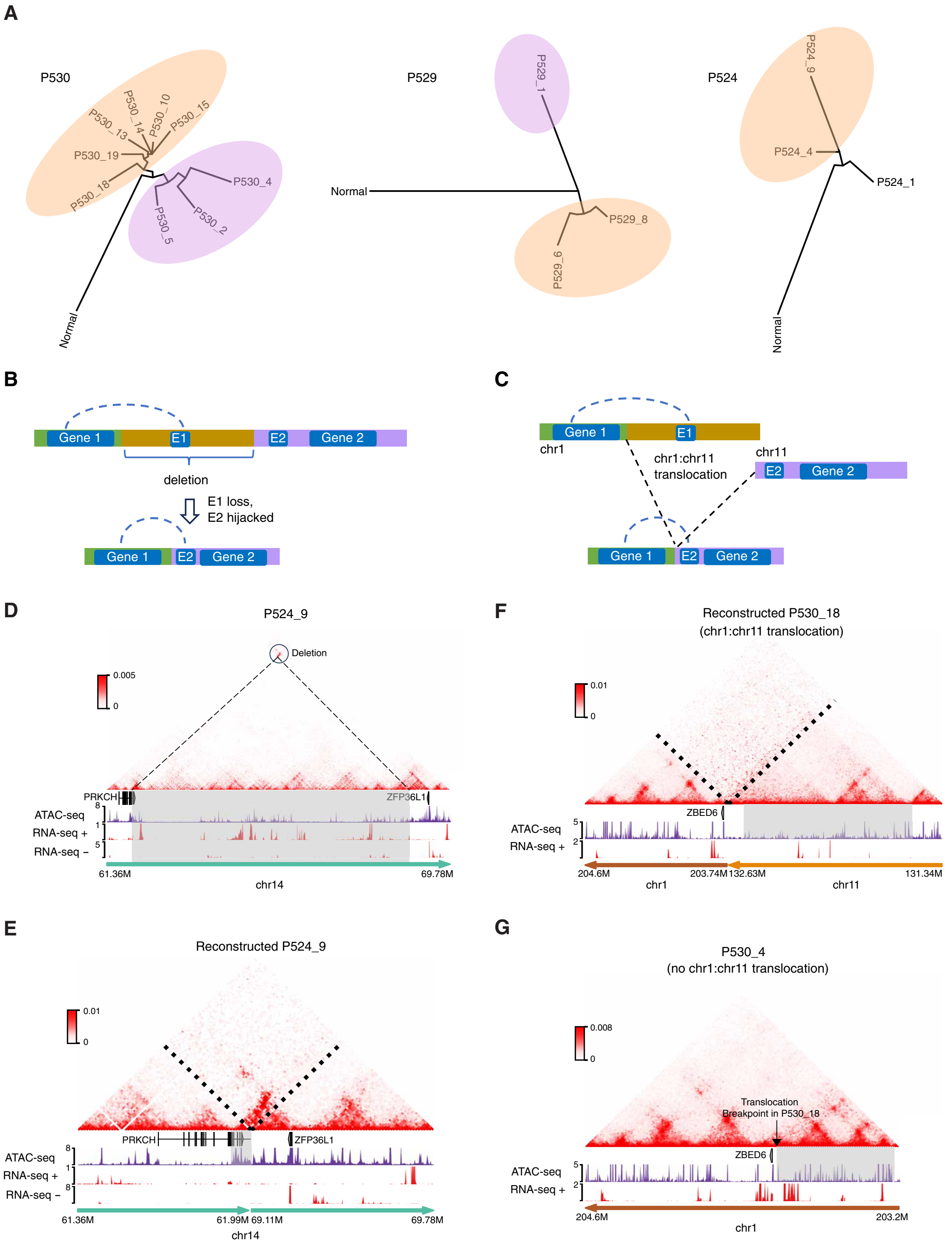

Fig. 5. Trade-off between enhancer hijacking and enhancer amputation. (**A**) Phylogenetic trees based on CNVs inferred from Hi-C data. (**B** and **C**) Schematic figure showing enhancer hijacking and enhancer amputation due to deletion event (B) and translocation event (C). (**D**) Hi-C map (top), ATAC-seq, and RNA-seq track (bottom) of chr14: 61,360,000 to 69,780,000 region in P524_9. Deletion marker was highlighted by black circle, and deletion regions were highlighted by black dashed line and gray rectangular. (**E**) Reconstructed Hi-C map (top), ATAC-seq, and RNA-seq track (bottom) of chr14: 61,360,000 to 69,780,000 region after removing deletion region (chr14: 61,990,000 to 69,110,000) in P524_9. Deletion breakpoints and region were marked by black dashed line. Potential gained enhancers of *ZFP36L1* were highlighted in gray rectangular. (**F**) Hi-C map (top), ATAC-seq, and RNA-seq track (bottom) of chr1: 203,740,000 to 204,600,000 and chr11: 131,340,000 to 132,630,000 region in frontal region sample P530_18. Translocation breakpoints were marked by black dashed line. Potential gained enhancers for *ZBED6* were highlighted by gray rectangular. (**G**) Hi-C map (top), ATAC-seq, and RNA-seq track (bottom) of chr1: 203,200,000 to 204,600,000 region in temporal region sample P530_4. Translocation breakpoint within this region in P530 frontal region samples was marked by black arrow. Potential enhancer loss for *ZBED6* in frontal region samples was highlighted in gray rectangular.

potential loss of several enhancers for *ZBED6* (Fig. 5G, gray rectangular). In this case, the trade-off between enhancers gain (hijacking) and enhancer loss (amputation) resulted in the decrease expression of *ZBED6*.

DISCUSSION

In this study, we explored GBM-specific, intertumor and intratumor heterogeneity at multiple 3D genome levels across 24 Hi-C samples from nine patients with GBM and three normal brain samples. We also generated the ultradeep sequencing Hi-C dataset (more than 6 billion reads) in primary GBM, which enabled us to explore 3D genomic interactions at frontal and temporal brain regions at kilobase resolution. In addition, we systematically identified the CNV and SV profiles across all the patient samples and evaluated their impacts on the 3D genome and transcriptional regulation. We identified both enhancer hijacking and amputation events and the complex relationship between two mechanisms in gene dysregulations.

We also investigated whether the estimated tumor purities could compound bulk Hi-C analysis. To do so, we performed deconvolution analysis using the bulk RNA-seq data with BayesPrism (*48*), and the results showed variable cell-type compositions in each sample (fig. S17). Specifically, we noticed that in the P530 frontal regions, there are also relatively large populations of myeloid (7.97%) and oligo cells (13.16%). Hence, we acknowledge that variations in the composition of cell populations between two samples might contribute to sample-specific 3D genome features. To fully address this concern, further studies could either adopt single-cell Hi-C experiments or design smart algorithms to take advantage of bulk Hi-C and single RNA/ATAC-seq data from the same sample to mitigate this potential bias.

We identified structural variations using Hi-C data and predicted both enhancer hijacking (gain of enhancers) and amputation events (loss of enhancers) in patients with GBM. While prior research mainly focused on the impact of gaining regulatory elements, such as enhancer or silencer hijacking (*15*–*21*), our study introduces an alternative perspective. Systematically cataloging and investigating the impact of SV-induce loss of enhancers would lead to important breakthroughs in cancer studies.

MATERIALS AND METHODS

Ethics approval and consent to participate

GBM sample use was approved by the University of California San Francisco's Committee on Human Research and the University of California, San Francisco's (UCSF's) Institutional Review Board (IRB). All patients provided informed written consent. Samples were also obtained from Northwestern University. All patients participating in this study underwent informed consent for molecular analysis of collected surgical specimens, and procedures were approved by the Northwestern University Institutional Review Board (IRB number: STU00095863). An honest broker was used to de-identify subjects before analysis to maintain patient privacy.

Sample acquisition and cataloging

Spatially distinct GBM samples were acquired during surgical resection at the discretion of each patient's treating neurosurgeon. Locations were selected to maximize diversity across each tumor and were recorded on the basis of stereotactic registration to a pre-operative magnetic resonance imaging (MRI) (BrainLAB). All specimens (tumor and blood) were collected and preserved by biorepository staff at UCSF.

When adequate material was available, flash-frozen tissue was divided for DNA/RNA extraction, ATAC-seq, and Hi-C. Sample information and additional methods regarding 3D MRI modeling and spatial analysis are provided in our companion paper (*38*).

Hi-C experiments

Hi-C experiments were conducted using the Arima Hi-C Kit with minor modifications. Approximately 10 to 20 mg of frozen tumor or brain tissue was submerged in liquid nitrogen and then pulverized using a mortar and pestle until the sample resembled a fine powder. Samples were then cross-linked using 2% formaldehyde, followed by quenching with glycine. The remaining steps were completed according to the manufacturer's protocol (Arima Genomics). Quality control was assessed before deep sequencing, which targeted 500 to 1200 million reads per sample on the Illumina platform.

Hi-C data processing

We first used Trim Galore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) for quality and adapter trimming of the raw Hi-C fastq files and then mapped them to hg19 human reference genome using runHiC pipeline (*49*). In specific, the raw sequencing reads were first mapped to hg19 using Burrows-Wheeler Aligner (bwa) (*50*). Then, it eliminated noise at both read level and fragment level. Contact matrices were then generated after binning the reads at 5-kb resolution. We then used the run-cool2multirescool script from 4DN consortium to generate the multiresolution cool files, spanning resolutions of 5 kb, 10 kb, 25 kb, 40 kb, 50 kb, 100 kb, 250 kb, 500 kb, 1 Mb, 2.5 Mb, 5 Mb, and 10 Mb, and concurrently perform iterative correction and eigenvector decomposition (ICE) normalization. We also used Juicer tools (*51*) to generate multiresolution .hic files, which can be visualized in Juicebox directly.

We combined the fastq files of all temporal region samples (#2, #4, and #5) into a single temporal region combined fastq file to achieve an ultradeep resolution in analyzing the 3D genome of the temporal region. The fastq files of all frontal region samples (#10, #13, #14, #15, #18, and #19) were also combined into a single frontal region combined fastq file for the same purpose. The data processing for these combined regional fastq files is similar to the processing of above individual samples, with the exception that, for the combined regional samples, an additional 2-kb resolution matrix is stored in the .hic and .mcool files.

Compartment, subcompartment, TAD, and loop calling

We used cooltools (*34*) to compute PC1 values from the Hi-C matrices for each sample at 100-kb resolution, subsequently assigning each bin into A/B compartments based on their respective PC1 values. We then used CALDER (*27*) to call subcompartments (A1.1, A1.2, A2.1, A2.2, B1.1, B1.2, B2.1, and B2.2) for each sample at 100-kb resolution. To check the most variable and stable regions across all samples at other resolutions, we also called subcompartment at 50- and 250-kb resolutions.

We used cooltools software to call diamond insulation scores using a 400-kb window size and a 40-kb bin size. We then defined TAD boundaries as the 4DN domain calling pipelines do: Local minima of the chromosome-wide topographic prominence track for $\log_2$(insulation score) exceeds a 0.2 threshold.

We predicted chromatin loops for each sample using a Random Forest classification framework called Peakachu (*32*). On the basis of the number of total intrachromosomal reads, we selected the appropriate high-confidence model at 10 kb for each of the 24 individual samples and at 2 kb for the two region combined samples of P530. We then used predicted probability >0.9 as cutoff to obtain high-confidence chromatin loops for each sample.

For GBM sample–specific chromatin loops, we selected loops in each GBM sample with a predicted probability below 0.3 in all three normal samples: NU0820, NU1265, and NU1650. In addition, we required that the predicted probability within a two-bin region surrounding each loop was also below 0.3 in all normal samples. For P530 region-specific chromatin loops, we selected loops in P530 temporal/frontal regions whose predicted probability is less than 0.5 in the other region and all normal samples. We also required that the predicted probability within a one-bin region around each loop was also below 0.5 in the other region and all normal samples.

SV prediction and neo-loop prediction

We first used NeoLoopFinder (*35*) to infer the CNV profile from Hi-C data for each sample and then eliminated the CNV effects. Contact matrices at three resolutions, including 5, 10, and 50 kb were used to predict SVs in each sample with "--balance-type CNV" parameters using EagleC (*37*). Results from the three resolutions were then integrated in a nonredundant fashion. The SVs were then compared with the Database of Genomic Variants (DGV) to filter out germline SV calls. The filtered SV events were then used as input for NeoLoopFinder to predict neo-loops and generate the reconstructed Hi-C map.

Phylogenetic tree construction using SV or CNV

We first generated an SV binary matrix and a CNV matrix for each patient. In the SV matrix, each row is an SV event, and each column is a sample. The value in the matrix is 0 or 1 depending on whether this sample has this specific SV event. For the CNV matrix, each row is a 50-kb genomic region, and each column is a sample. The value in this matrix is the copy number of this region in each sample. We then calculated the Manhattan distance from the CNV matrix and SV binary matrix and then used ape R package (*52*) and ggtree R package (*53*) to generate the neighbor-joining tree for each patient.

Cell-type deconvolution using RNA-seq data

To identify the cell-type composition in each sample, we analyzed our RNA-seq data in 15 GBM samples with a Bayesian method BayesPrism (*48*). We used single-cell RNA-seq data of eight patients with high-grade gliomas (*54*) as reference to deconvolute our bulk RNA-seq data.

Supplementary Materials

The PDF file includes:

Figs. S1 to S17

Legends for tables S1 to S6

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S6

REFERENCES AND NOTES

1. Q. T. Ostrom, H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, J. S. Barnholtz-Sloan, CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2011-2015. *Neuro Oncol.* **20**, iv1–iv86 (2018).
2. A. C. Tan, D. M. Ashley, G. Y. Lopez, M. Malinzak, H. S. Friedman, M. Khasraw, Management of glioblastoma: State of the art and future directions. *CA Cancer J. Clin.* **70**, 299–312 (2020).
3. F. Hanif, K. Muzaffar, K. Perveen, S. M. Malhi, S. U. Simjee, Glioblastoma multiforme: A review of its epidemiology and pathogenesis through clinical presentation and treatment. *Asian Pac. J. Cancer Prev.* **18**, 3–9 (2017).
4. M. Jamal-Hanjani, S. A. Quezada, J. Larkin, C. Swanton, Translational implications of tumor heterogeneity. *Clin. Cancer Res.* **21**, 1258–1266 (2015).
5. S. R. Y. Cajal, M. Sesé, C. Capdevila, T. Aasen, L. De Mattos-Arruda, S. J. Diaz-Cano, J. Hernández-Losa, J. Castellví, Clinical implications of intratumor heterogeneity: Challenges and opportunities. *J. Mol. Med. (Berl)* **98**, 161–177 (2020).
6. A. Sottoriva, I. Spiteri, S. G. M. Piccirillo, A. Touloumis, V. P. Collins, J. C. Marioni, C. Curtis, C. Watts, S. Tavaré, Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4009–4014 (2013).
7. M. Snuderl, L. Fazlollahi, L. P. Le, M. Nitta, B. H. Zhelyazkova, C. J. Davidson, S. Akhavanfard, D. P. Cahill, K. D. Aldape, R. A. Betensky, D. N. Louis, A. J. Iafrate, Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* **20**, 810–817 (2011).
8. T. Cremer, M. Cremer, Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010).
9. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
10. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
11. E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Bluthgen, J. Dekker, E. Heard, Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
12. M. J. Fullwood, M. H. Liu, Y. F. Pan, J. Liu, H. Xu, Y. B. Mohamed, Y. L. Orlov, S. Velkov, A. Ho, P. H. Mei, E. G. Y. Chew, P. Y. H. Huang, W.-J. Welboren, Y. Han, H. S. Ooi, P. N. Ariyaratne, V. B. Vega, Y. Luo, P. Y. Tan, P. Y. Choy, K. D. S. A. Wansa, B. Zhao, K. S. Lim, S. C. Leow, J. S. Yow, R. Joseph, H. Li, K. V. Desai, J. S. Thomsen, Y. K. Lee, R. K. M. Karuturi, T. Herve, G. Bourque, H. G. Stunnenberg, X. Ruan, V. Cacheux-Rataboul, W.-K. Sung, E. T. Liu, C.-L. Wei, E. Cheung, Y. Ruan, An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* **462**, 58–64 (2009).
13. M. R. Mumbach, A. J. Rubin, R. A. Flynn, C. Dai, P. A. Khavari, W. J. Greenleaf, H. Y. Chang, HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
14. T.-H. S. Hsieh, A. Weiner, B. Lajoie, J. Dekker, N. Friedman, O. J. Rando, Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162**, 108–119 (2015).
15. J. R. Dixon, J. Xu, V. Dileep, Y. Zhan, F. Song, V. T. Le, G. G. Yardımcı, A. Chakraborty, D. V. Bann, Y. Wang, R. Clark, L. Zhang, H. Yang, T. Liu, S. Iyyanki, L. An, C. Pool, T. Sasaki, J. C. Rivera-Mulia, H. Ozadam, B. R. Lajoie, R. Kaul, M. Buckley, K. Lee, M. Diegel, D. Pezic, C. Ernst, S. Hadjur, D. T. Odom, J. A. Stamatoyannopoulos, J. R. Broach, R. C. Hardison, F. Ay, W. S. Noble, J. Dekker, D. M. Gilbert, F. Yue, Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).
16. S. Gröschel, M. A. Sanders, R. Hoogenboezem, E. de Wit, B. A. M. Bouwman, C. Erpelinck, V. H. J. van der Velden, M. Havermans, R. Avellino, K. van Lom, E. J. Rombouts, M. van Duin, K. Döhner, H. B. Beverloo, J. E. Bradner, H. Döhner, B. Löwenberg, P. J. M. Valk, E. M. J. Bindels, W. de Laat, R. Delwel, A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* **157**, 369–381 (2014).
17. P. A. Northcott, C. Lee, T. Zichner, A. M. Stütz, S. Erkek, D. Kawauchi, D. J. H. Shih, V. Hovestadt, M. Zapatka, D. Sturm, D. T. Jones, M. Kool, M. Remke, F. M. G. Cavalli, S. Zuyderduyn, G. D. Bader, S. VandenBerg, L. A. Esparza, M. Ryzhova, W. Wang, A. Wittmann, S. Stark, L. Sieber, H. Seker-Cin, L. Linke, F. Kratochwil, N. Jäger, I. Buchhalter, C. D. Imbusch, G. Zipprich, B. Raeder, S. Schmidt, N. Diessl, S. Wolf, S. Wiemann, B. Brors, C. Lawerenz, J. Eils, H.-J. Warnatz, T. Risch, M.-L. Yaspo, U. D. Weber, C. C. Bartholomae, C. von Kalle, E. Turanyi, P. Hauser, E. Sanden, A. Darabi, P. Siesjö, J. Sterba, K. Zitterbart, D. Sumerauer, P. van Sluis, R. Versteeg, R. Volckmann, J. Koster, M. U. Schuhmann, M. Ebinger, H. L. Grimes, G. W. Robinson, A. Gajjar, M. Mynarek, K. von Hoff, S. Rutkowski, T. Pietsch, W. Scheurlen, J. Felsberg, G. Reifenberger, A. E. Kulozik, A. von Deimling, O. Witt, R. Eils, R. J. Gilbertson, A. Korshunov, M. D. Taylor, P. Lichter, J. O. Korbel, R. J. Wechsler-Reya, S. M. Pfister, Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* **511**, 428–434 (2014).
18. J. Xu, F. Song, H. Lyu, M. Kobayashi, B. Zhang, Z. Zhao, Y. Hou, X. Wang, Y. Luan, B. Jia, L. Stasiak, J. H. Wong, Q. Wang, Q. Jin, Q. Jin, Y. Fu, H. Yang, R. C. Hardison, S. Dovat, L. C. Platanias, Y. Diao, Y. Yang, T. Yamada, A. D. Viny, R. L. Levine, D. Claxton, J. R. Broach, H. Zheng, F. Yue, Subtype-specific 3D genome alteration in acute myeloid leukaemia. *Nature* **611**, 387–398 (2022).

19. J. Wang, T. Y.-T. Huang, Y. Hou, E. Bartom, X. Lu, A. Shilatifard, F. Yue, A. Saratsis, Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Sci. Adv.* **7**, eabg4126 (2021).
20. T. Liu, J. Wang, H. Yang, Q. Jin, X. Wang, Y. Fu, Y. Luan, Q. Wang, M. W. Youngblood, X. Lu, L. Casadei, R. Pollock, F. Yue, Enhancer coamplification and hijacking promote oncogene expression in liposarcoma. *Cancer Res.* **83**, 1517–1530 (2023).
21. J. Weischenfeldt, T. Dubash, A. P. Drainas, B. R. Mardin, Y. Chen, A. M. Stütz, S. M. Waszak, G. Bosco, A. R. Halvorsen, B. Raeder, T. Efthymiopoulos, S. Erkek, C. Siegl, H. Brenner, O. T. Brustugun, S. M. Dieter, P. A. Northcott, I. Petersen, S. M. Pfister, M. Schneider, S. K. Solberg, E. Thunissen, W. Weichert, T. Zichner, R. Thomas, M. Peifer, A. Helland, C. R. Ball, M. Jechlinger, R. Sotillo, H. Glimm, J. O. Korbel, Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2017).
22. T. Xie, A. Danieli-Mackay, M. Buccarelli, M. Barbieri, I. Papadionysiou, Q. G. D'Alessandris, C. Robens, N. Übelmesser, O. S. Vinchure, L. Lauretti, G. Fotia, R. F. Schwarz, X. Wang, L. Ricci-Vitiani, J. Gopalakrishnan, R. Pallini, A. Papantonis, Pervasive structural heterogeneity rewires glioblastoma chromosomes to sustain patient-specific transcriptional programs. *Nat. Commun.* **15**, 3905 (2024).
23. T. Ryba, I. Hiratani, J. Lu, M. Itoh, M. Kulik, J. Zhang, T. C. Schulz, A. J. Robins, S. Dalton, D. M. Gilbert, Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761–770 (2010).
24. S. E. Johnstone, A. Reyes, Y. Qi, C. Adriaens, E. Hegazi, K. Pelka, J. H. Chen, L. S. Zou, Y. Drier, V. Hecht, N. Shoresh, M. K. Selig, C. A. Lareau, S. Iyer, S. C. Nguyen, E. F. Joyce, N. Hacohen, R. A. Irizarry, B. Zhang, M. J. Aryee, B. E. Bernstein, Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **182**, 1474–1489.e23 (2020).
25. J. R. Hawley, S. Zhou, C. Arlidge, G. Grillo, K. J. Kron, R. Hugh-White, T. H. van der Kwast, M. Fraser, P. C. Boutros, R. G. Bristow, M. Lupien, Reorganization of the 3D genome pinpoints noncoding drivers of primary prostate tumors. *Cancer Res.* **81**, 5833–5848 (2021).
26. K. Xiong, J. Ma, Revealing Hi-C subcompartments by imputing inter-chromosomal chromatin interactions. *Nat. Commun.* **10**, 5069 (2019).
27. Y. Liu, L. Nanni, S. Sungalee, M. Zufferey, D. Tavernari, M. Mina, S. Ceri, E. Oricchio, G. Ciriello, Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. *Nat Commun* **12**, 2439 (2021).
28. Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A.-R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
29. L. Liu, Y. Yang, H. Duan, J. He, L. Sun, W. Hu, J. Zeng, CHI3L2 is a novel prognostic biomarker and correlated with immune infiltrates in gliomas. *Front. Oncol.* **11**, 611038 (2021).
30. M. Nakada, D. Kita, T. Watanabe, Y. Hayashi, L. Teng, I. V. Pyko, J.-I. Hamada, Aberrant signaling pathways in glioma. *Cancers (Basel)* **3**, 3242–3278 (2011).
31. T. Yang, F. Zhang, G. G. Yardimci, F. Song, R. C. Hardison, W. S. Noble, F. Yue, Q. Li, HiCRep: Assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res.* **27**, 1939–1949 (2017).
32. T. J. Salameh, X. Wang, F. Song, B. Zhang, S. M. Wright, C. Khunsriraksakul, Y. Ruan, F. Yue, A supervised learning framework for chromatin loop detection in genome-wide contact maps. *Nat. Commun.* **11**, 3428 (2020).
33. S. Upraity, S. Kazi, V. Padul, N. V. Shirsat, MiR-224 expression increases radiation sensitivity of glioblastoma cells. *Biochem. Biophys. Res. Commun.* **448**, 225–230 (2014).
34. Open2C, N. Abdennur, S. Abraham, G. Fudenberg, I. M. Flyamer, A. A. Galitsyna, A. Goloborodko, M. Imakaev, B. A. Oksuz, S. V. Venev, Y. Xiao, Cooltools: Enabling high-resolution Hi-C analysis in Python. *PLOS Comput. Biol.* **20**, e1012067 (2024).
35. X. Wang, J. Xu, B. Zhang, Y. Hou, F. Song, H. Lyu, F. Yue, Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. *Nat. Methods* **18**, 661–668 (2021).
36. Y. Li, N. D. Roberts, J. A. Wala, O. Shapira, S. E. Schumacher, K. Kumar, E. Khurana, S. Waszak, J. O. Korbel, J. E. Haber, M. Imielinski, PCAWG Structural Variation Working Group, J. Weischenfeldt, R. Beroukhim, P. J. Campbell, PCAWG Consortium, Patterns of somatic structural variation in human cancer genomes. *Nature* **578**, 112–121 (2020).
37. X. Wang, Y. Luan, F. Yue, EagleC: A deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. *Sci. Adv.* **8**, eabn9215 (2022).
38. R. Mathur, Q. Wang, P. G. Schupp, A. Nikolic, S. Hilz, C. Hong, N. R. Grishanina, D. Kwok, N. O. Stevers, Q. Jin, M. W. Youngblood, L. A. Stasiak, Y. Hou, J. Wang, T. N. Yamaguchi, M. Lafontaine, A. Shai, I. V. Smirnov, D. A. Solomon, S. M. Chang, S. L. Hervey-Jumper, M. S. Berger, J. M. Lupo, H. Okada, J. J. Phillips, P. C. Boutros, M. Gallo, M. C. Oldham, F. Yue, J. F. Costello, Glioblastoma evolution and heterogeneity from a 3D whole-tumor perspective. *Cell* **187**, 446–463.e16 (2024).
39. Q. Weng, J. Wang, J. Wang, D. He, Z. Cheng, F. Zhang, R. Verma, L. Xu, X. Dong, Y. Liao, X. He, A. Potter, L. Zhang, C. Zhao, M. Xin, Q. Zhou, B. J. Aronow, P. J. Blackshear, J. N. Rich, Q. He, W. Zhou, M. L. Suva, R. R. Waclaw, S. S. Potter, G. Yu, Q. R. Lu, Single-cell transcriptomics uncovers glial progenitor diversity and cell fate determinants during development and gliomagenesis. *Cell Stem Cell* **24**, 707–723.e8 (2019).
40. X. Guo, H. Piao, Y. Zhang, P. Sun, B. Yao, Overexpression of microRNA-129-5p in glioblastoma inhibits cell proliferation, migration, and colony-forming ability by targeting ZFP36L1. *Bosn. J. Basic Med. Sci.* **20**, 459–470 (2020).
41. Y. W. Shao, G. A. Wood, J. Lu, Q.-L. Tang, J. Liu, S. Molyneux, Y. Chen, H. Fang, H. Adissu, T. McKee, P. Waterhouse, R. Khokha, Cross-species genomics identifies DLG2 as a tumor suppressor in osteosarcoma. *Oncogene* **38**, 291–298 (2019).
42. S. Keane, M. Herring, P. Rolny, Y. Wettergren, K. Ejeskär, Inflammation suppresses DLG2 expression decreasing inflammasome formation. *J. Cancer Res. Clin. Oncol.* **148**, 2295–2311 (2022).
43. J. T. Siaw, N. Javanmardi, J. Van den Eynden, D. E. Lind, S. Fransson, A. Martinez-Monleon, A. Djos, R. M. Sjöberg, M. Östensson, H. Caren, G. Trøen, K. Beiske, A. P. Berbegall, R. Noguera, W.-Y. Lai, P. Kogner, R. H. Palmer, B. Hallberg, T. Martinsson, 11q deletion or ALK activity curbs DLG2 expression to maintain an undifferentiated state in neuroblastoma. *Cell Rep.* **32**, 108171 (2020).
44. S. Keane, S. Améen, A. Lindlöf, K. Ejeskär, Low DLG2 gene expression, a link between 11q-deleted and MYCN-amplified neuroblastoma, causes forced cell cycle progression, and predicts poor patient survival. *Cell Commun. Signal.* **18**, 65 (2020).
45. E. Markljung, L. Jiang, J. D. Jaffe, T. S. Mikkelsen, O. Wallerman, M. Larhammar, X. Zhang, L. Wang, V. Saenz-Vash, A. Gnirke, A. M. Lindroth, R. Barrés, J. Yan, S. Strömberg, S. De, F. Pontén, E. S. Lander, S. A. Carr, J. R. Zierath, K. Kullander, C. Wadelius, K. Lindblad-Toh, G. Andersson, G. Hjälm, L. Andersson, ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth. *PLOS Biol.* **7**, e1000256 (2009).
46. Q. Cai, B. Zhang, H. Sung, S. K. Low, S.-S. Kweon, W. Lu, J. Shi, J. Long, W. Wen, J.-Y. Choi, D.-Y. Noh, C.-Y. Shen, K. Matsuo, S.-H. Teo, M. K. Kim, U. S. Khoo, M. Iwasaki, M. Hartman, A. Takahashi, K. Ashikawa, K. Matsuda, M.-H. Shin, M. H. Park, Y. Zheng, Y.-B. Xiang, B.-T. Ji, S. K. Park, P.-E. Wu, C.-N. Hsiung, H. Ito, Y. Kasuga, P. Kang, S. Mariapun, S. H. Ahn, H. S. Kang, K. Y. K. Chan, E. P. S. Man, H. Iwata, S. Tsugane, H. Miao, J. Liao, Y. Nakamura, M. Kubo, DRIVE GAME-ON Consortium, R. J. Delahanty, Y. Zhang, B. Li, C. Li, Y.-T. Gao, X.-O. Shu, D. Kang, W. Zheng, Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat. Genet.* **46**, 886–890 (2014).
47. M. Akhtar Ali, S. Younis, O. Wallerman, R. Gupta, L. Andersson, T. Sjöblom, Transcriptional modulator ZBED6 affects cell cycle and growth of human colorectal cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7743–7748 (2015).
48. T. Chu, Z. Wang, D. Pe'er, C. G. Danko, Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nat. Cancer* **3**, 505–517 (2022).
49. X. Wang, runHiC: A user-friendly Hi-C data processing software based on hiclib (Zenodo, 2016); https://doi.org/10.5281/zenodo.55324.
50. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
51. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
52. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
53. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T.-Y. Lam, ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
54. J. Yuan, H. M. Levitin, V. Frattini, E. C. Bush, D. M. Boyett, J. Samanamud, M. Ceccarelli, A. Dovas, G. Zanazzi, P. Canoll, J. N. Bruce, A. Lasorella, A. Iavarone, P. A. Sims, Single-cell transcriptome analysis of lineage diversity in high-grade glioma. *Genome Med.* **10**, 57 (2018).

Acknowledgments: We thank the staff of the University of California, San Francisco Brain Tumor Center Tissue Bank for sample acquisition. **Funding:** Support provided by the Hana Jabsheh Research Initiative, the Panattoni Project, National Institutes of Health Grants F32

1F32CA239472-01 (to R.M.), T32 T32CA151022 (to R.M. and S.H.), R01 CA169316 (to J.F.C.), P01 CA118816-06 (to J.F.C., J.M.L, J.J.P., and S.M.C), P50 CA097257 (to J.F.C.), and U01 CA229345 (to J.J.P.). F.Y. was supported by NIH grants R35GM124820, 1R01HG009906, and R01HG011207.

Author contributions: Q.W.: Writing (original draft), conceptualization, investigation, writing (review and editing), resources, data curation, validation, formal analysis, software, and visualization. J.W.: Writing (original draft), conceptualization, investigation, writing (review and editing), resources, data curation, validation, formal analysis, software, and visualization. R.M.: Conceptualization, investigation, writing (review and editing), methodology, resources, funding acquisition, validation, project administration, and visualization. M.W.Y.: Investigation and writing (review and editing). Q.J.: Writing (original draft), conceptualization, investigation, writing (review and editing), methodology, resources, and validation. Y.H.: Investigation and validation. L.A.S.: Investigation. Y.L.: Formal analysis and software. H.Z.: Formal analysis. S.H.: Resources and data curation. C.H.: Investigation, methodology, resources, validation, and project administration. S.M.C.: Conceptualization, investigation, writing (review and editing), methodology, resources, funding acquisition, validation, and visualization. J.M.L.: Methodology, resources, and software. J.J.P.: Investigation, writing (review and editing), and resources. J.F.C.: Conceptualization, writing (review and editing), methodology, resources, funding acquisition, data curation, validation, supervision, and project administration. F.Y.: Writing (original draft), conceptualization, investigation, writing (review and editing), methodology, resources, funding acquisition, data curation, validation, supervision, formal analysis, software, project administration, and visualization. **Competing interests:** S.H. is a current stockholder of Roche stocks. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw data of all RNA-seq, ATAC-seq, and Hi-C datasets have been deposited in the European Genome-Phenome Archive (EGA) database under accession numbers EGAD00001010311, EGAD00001010312, and EGAD00001005222.

Submitted 3 December 2023
Accepted 5 February 2025
Published 12 March 2025
10.1126/sciadv.adn2830