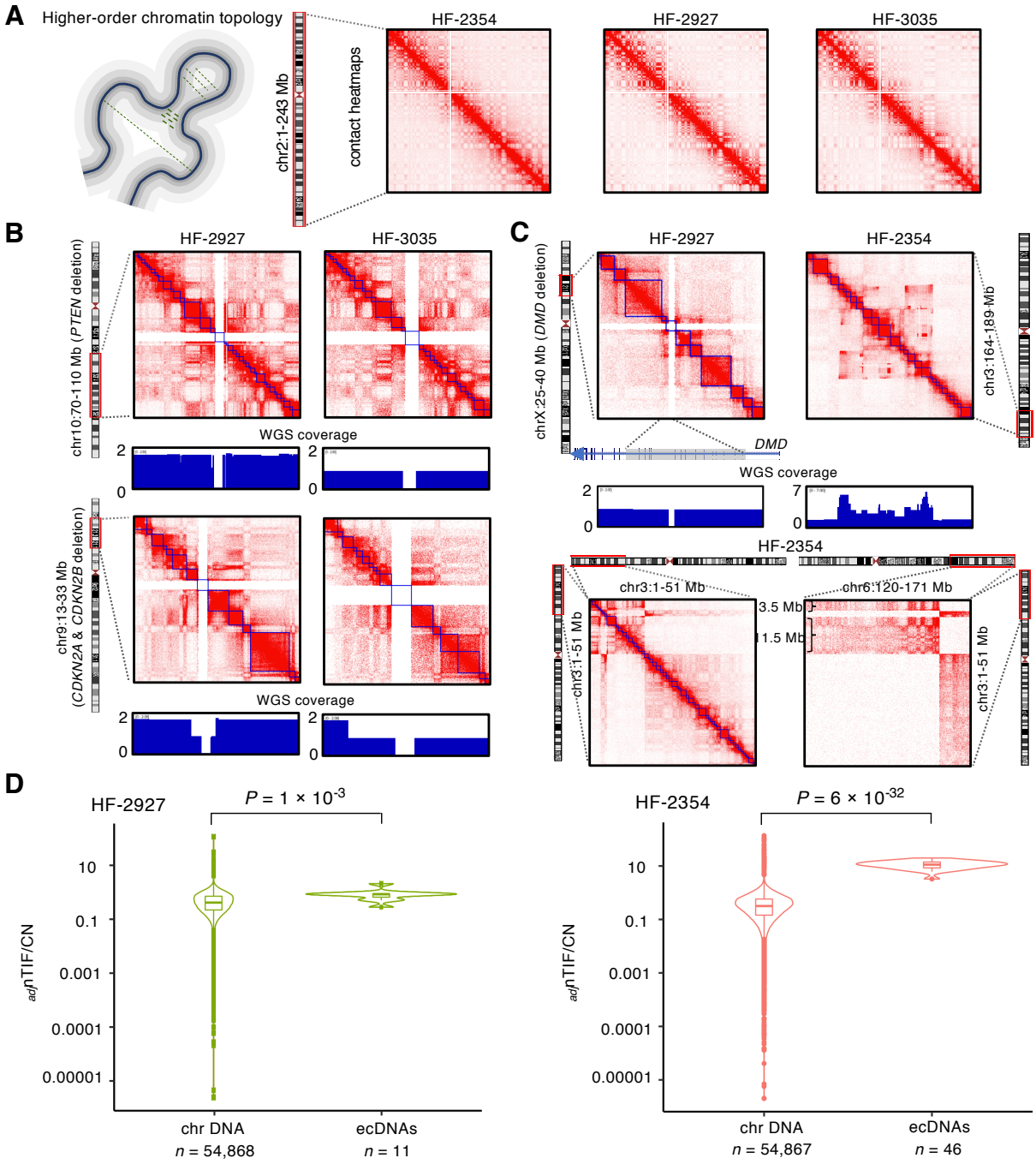


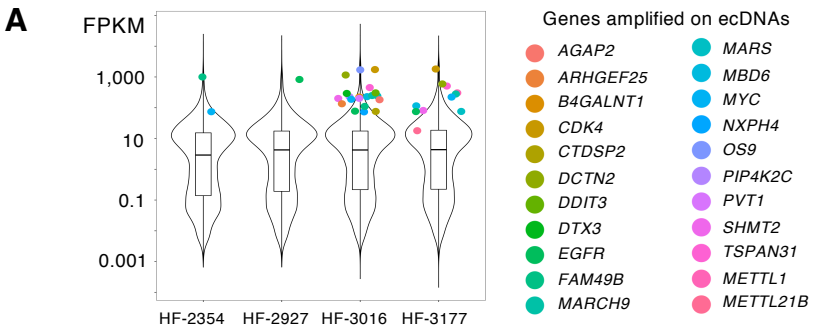
Figure S1. Hi-C assay detects chromatin topology changes caused by the genome structural variants and extensive ecDNA mediated chromatin interactions, Related to Figure 1.

(A) The spatial chromatin topology measured by the general chromatin contacts in Hi-C can be visualized by 2D contact heatmaps. Chromosome 2 heatmaps from all three GBM patient-derived neurosphere cell lines are shown.

(B) 2D contact heatmaps of genomic regions where *PTEN* (upper) and *CDKN2A* & *CDKN2B* (lower) deletions in HF-2927 and HF-3035, respectively are shown. TADs are demarcated by the blue lines. Deletion of the loci is represented by the loss of chromatin contacts. Coverage tracks from whole-genome sequencing in the corresponding regions are shown in blue.

(C) Deletion in *DMD* gene, a complex rearrangement in chromosome 3 and a double-translocations t(3;6) are visualized by the 2D heatmaps as unique and aberrant contact patterns.

(D) Box plots display $_{adj}$nTIF/CN between the known ecDNA regions and chromosomal DNA at 50-Kb resolution from the Hi-C analyses. *P* values determined for higher $_{adj}$nTIF/CN in ecDNA regions are shown (one-sided Wilcoxon Rank-Sum Test). See Table S2 for statistics summary. For all boxplots, center line, median; boxes, first and third quartiles; whiskers, 1.5 × the interquartile range (IQR); points, outliers.



B

	HF-2354	HF-2927	HF-3035	HF-3016	HF-3177
Protein binding	RNAPII Binding peaks				
	34,801	23,469	24,492	26,816	41,009
	H3K27ac peaks				
	36,577	34,395	45,581	36,175	51,894
Long-range chromatin interactions with RNAPII binding	Genome-wide *cis*-interactions				
	42,208	32,785	43,719	16,868	28,155
	Intra-ecDNA				
	189	64		945	686
	ecDNA-chromosomal				
	220	271		587	455

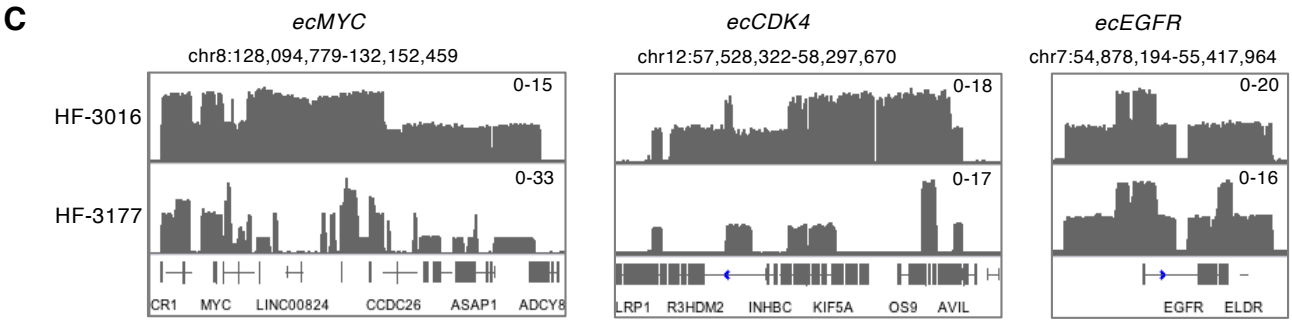


Figure S2. RNAPII ChIA-PET and ChIP-seq analyses in five GBM-derived neurosphere cell lines, Related to Figure 2.

(A) RNA abundance of genes amplified on the ecDNAs from ecDNA (+) GBM derived cell lines.

(B) Summary of the number of ChIP-seq peaks and long-range chromatin interactions detected in each of the five cell lines. The interactions are reported in three categories, the genome-wide significant *cis*-interactions, interactions within the ecDNA regions (intra-ecDNA interactions), and between ecDNA regions and the 23 chromosomes (ecDNA-chromosomal). Only interactions with RNAPII binding sites detected at both anchors are reported. Both interactions originated from the ecDNA regions and the corresponding native chromosomal loci contribute to the observed interactions. Because the copy numbers of ecDNA far exceed the copies of chromosomal alleles, majority of the interactions reported are expected to derive from the ecDNA regions.

(C) ecDNA regions defined by whole genome sequencing coverages in HF-3016 and HF-3177. Genes annotated in the defined regions are labelled.

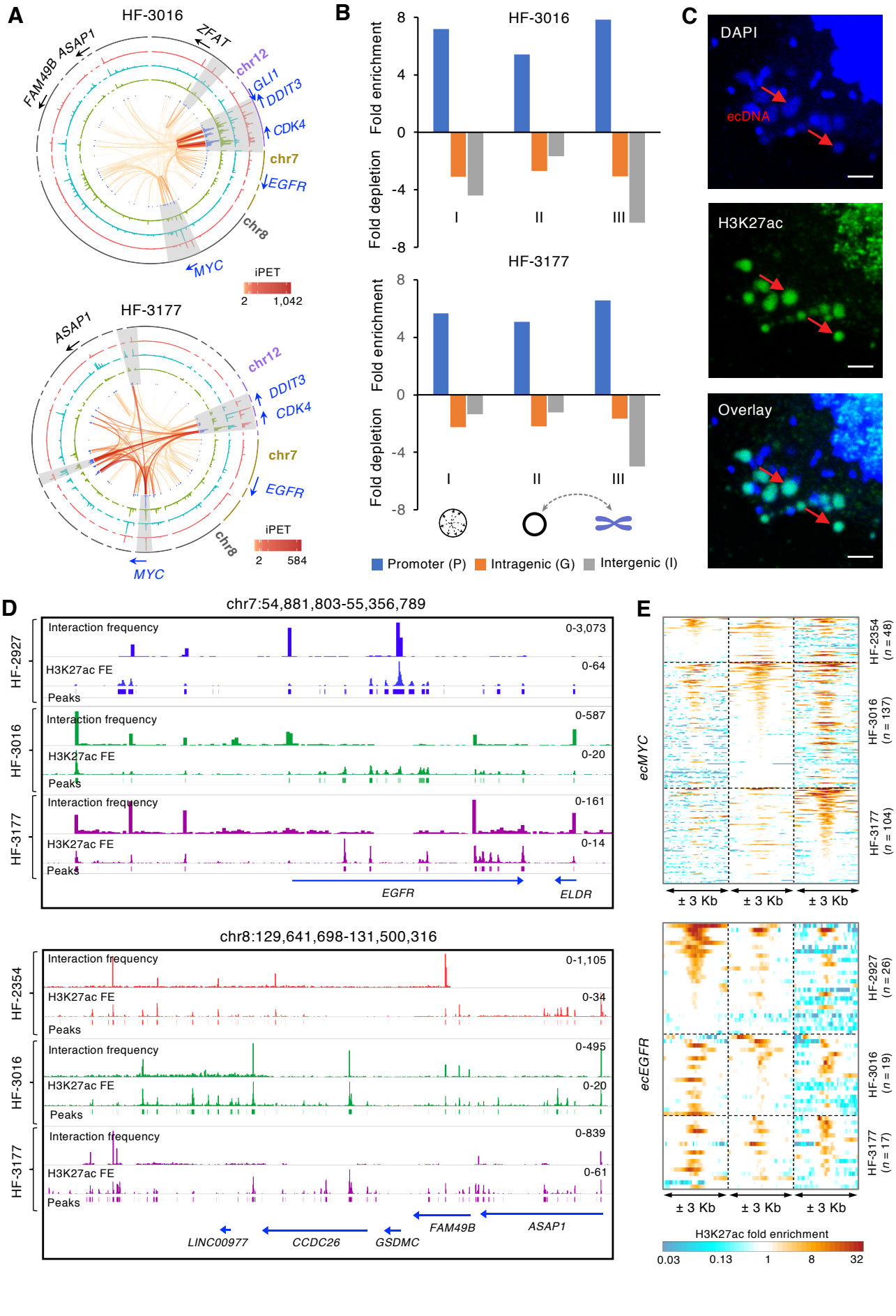

A
HF-3016
FAM49B
ASAP1
ZFAT
chr12
GLI1
DDIT3
CDK4
chr7
EGFR
chr8
MYC
iPET
2
1,042
HF-3177
ASAP1
chr12
DDIT3
CDK4
chr7
EGFR
chr8
MYC
iPET
2
584
B
HF-3016
Fold enrichment
Fold depletion
8
4
0
-4
-8
I
II
III
HF-3177
Promoter (P)
Intragenic (G)
Intergenic (I)
C
DAPI
ecDNA
H3K27ac
Overlay
D
chr7:54,881,803-55,356,789
HF-2927
HF-3016
HF-3177
Interaction frequency
H3K27ac FE
Peaks
0-3,073
0-64
0-587
0-20
0-161
0-14
EGFR
ELDR
chr8:129,641,698-131,500,316
HF-2354
HF-3016
HF-3177
0-1,105
0-34
0-495
0-20
0-839
0-61
LINC00977
CCDC26
GSDMC
FAM49B
ASAP1
E
ecMYC
ecEGFR
HF-2354 (n = 48)
HF-3016 (n = 137)
HF-3177 (n = 104)
HF-2927 (n = 26)
HF-3016 (n = 19)
HF-3177 (n = 17)
± 3 Kb
H3K27ac fold enrichment
0.03
0.13
1
8
32

Figure S3. EcDNA regions display intense *cis*- and *trans*-interactions with strong H3K27ac enrichment, Related to Figure 3.

(A) Circos plots of ecDNA regions defined in HF-3016 (upper) and HF-3177 (lower) ecDNA (+) cell lines. From inner to outer circles: intra-ecDNA interaction loops, blue: distribution of intra-ecDNA interaction frequency, green: distribution of ecDNA-chromosomal interaction frequency, cyan: distribution of H3K27ac fold enrichment, red: RNAPII binding enrichment intensity. The signal tracks are at 1-Kb resolution. Examples of high concordance regions between high H3K27ac enrichment and high interaction frequency are highlighted in grey.

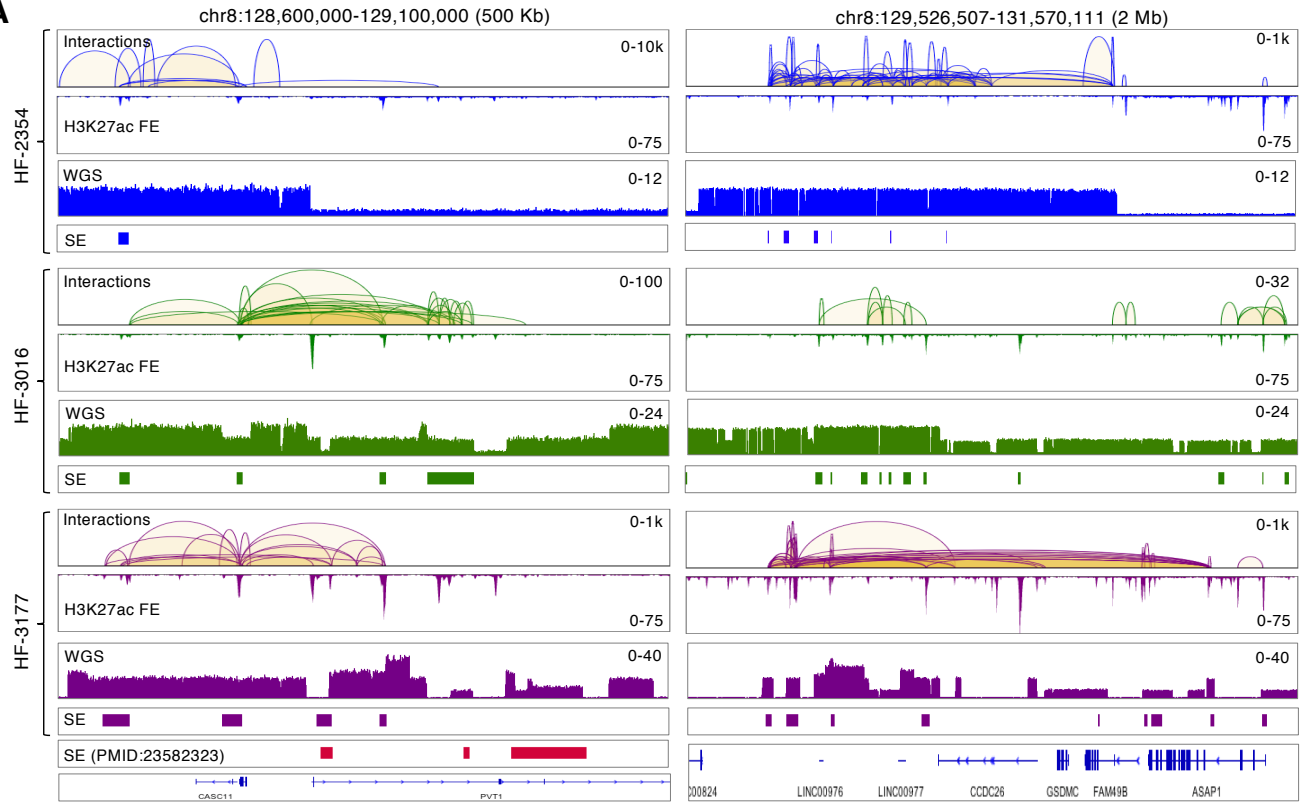
(B) Enrichment or depletion of promoter, intergenic and intragenic regions associated with ecDNA interaction anchors compared to genomic background on ecDNA regions. The interaction anchors from (I) intra-ecDNA, ecDNAs mediated trans chromosomal interaction (II) and their corresponding chromosomal contacts (III) in HF-3016 and HF-3177 are shown.

(C) H3K27ac immunostaining on metaphase chromosomes from HF-2927 cells. Distinct DAPI (blue) positive extrachromosomal spots overlap with H3K27ac staining spots (green). Scale bars = 2 μm.

(D) Concordance between interaction frequency and H3K27ac signals. *ecMYC* and *ecEGFR* regions in three cell lines are shown with H3K27ac fold enrichment, peaks and contact frequency tracks.

(E) Heatmaps of H3K27ac fold enrichment within ± 3 Kb of H3K27ac peaks located on *ecMYC* (top) or *ecEGFR* (bottom) across three cell lines. Scale bar is log2 scaled.

A



B

Amplicon classes	Total amplicons	Oncogene & Enhancer	Oncogene & No-Enhancer	Protein-coding Gene & Enhancer	Protein-coding Gene & No-Enhancer	No-gene & Enhancer	No-gene & No-Enhancer
Circular	579	402	0	555	1	22	1
BFB	503	375	0	497	1	4	1
Heavily-rearranged	814	553	0	799	5	9	1
Linear	3,895	1077	5	3,204	128	465	98

Figure S4. Enrichment of enhancer signals on ecDNAs and their variation among different ecDNA (+) cell lines, Related to Figure 4.

(A) Super-enhancers (SEs), H3K27ac fold enrichment (FE) and intra-ecDNA interactions within *ecMYC* regions found in their respective cell lines. For comparison, SEs defined in Lovén et al (PMID: 23582323) are shown.

(B) Summary of amplicon versus gene/enhancer content by amplicon class. Amplicons were classified as: (1) Circular, representing amplicons residing extrachromosomally; (2) 'BFB' if they bore a signature of having been created by a breakage-fusion-bridge (BFB) mechanism; (3) as Heavily-rearranged, for noncircular amplicons containing pieces of DNA (DNA segments) from different chromosomes or regions that were very far apart on chromosomes (>1 Mb); or (4) Linear for linear amplifications. There is no duplication of genes between the "Oncogene" and "Protein-coding Gene" classes.

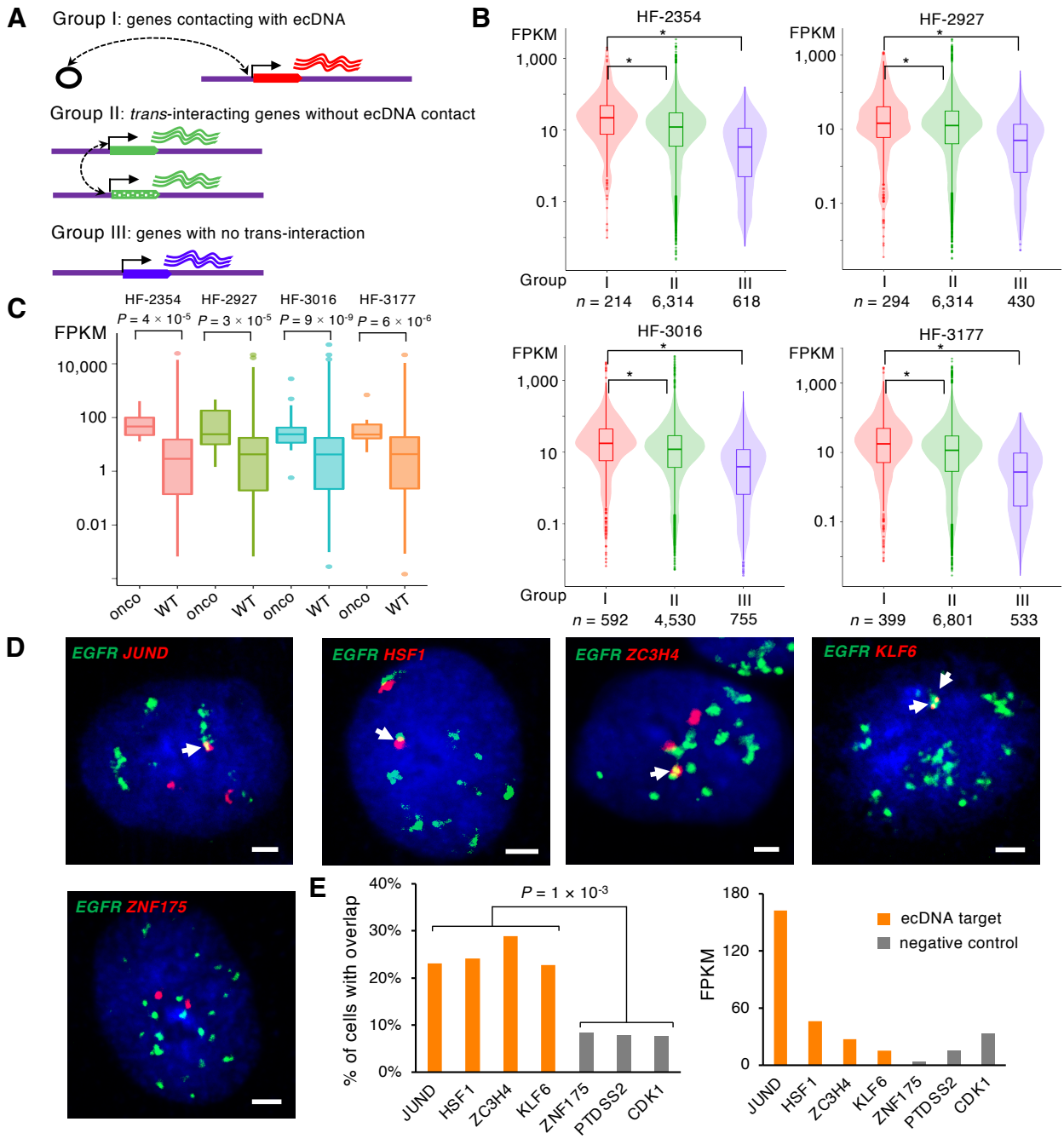


Figure S5. Chromosomal genes specific connected by ecDNAs are actively transcribed, Related to Figure 5.

(A) Based on their ecDNA connectivity status, genes are classified in three categories; Group I: genes with promoters connecting to ecDNAs, Group II: genes with promoters having *trans*-interactions but not with ecDNAs, and Group III: genes with no *trans*-interactions. Genes classified in each of the Group I, II and III and their expression are listed in Table S3.

(B) Box plots of steady-state RNA expression (FPKM) of genes from Groups I, II and III in HF-2354, HF-2927, HF-3016 and HF-3177 ecDNA (+) cell lines. * represents significant *P* values based on one-sided Wilcoxon Rank-Sum Test.

(C) Box plot shows the expression level of ecDNA-connected oncogenes (n = 9, 16, 34 and 18) vs. whole transcriptome (WT) (n = 21,170, 18,965, 19,150 and 19,140). Statistical analyses by one-sided Wilcoxon Rank-Sum Test. For all boxplots, center line, median; boxes, first and third quartiles; whiskers, 1.5 × the interquartile range (IQR); points, outliers. See Table S3 for (B) & (C) statistics summary.

(D) Multi-color FISH validation of interactions between *ecEGFR* and its chromosomal targets in HF-2927. Examples of FISH images from four *ecEGFR* targets (*JUND*, *HSF1*, *ZC3H4*, and *KLF6*) and one negative control (*ZNF175*) were shown. *EGFR* was in green, targets and controls were in red, DAPI staining of the nuclei was in blue. Overlapped signals were indicated by white arrows. Scale bar is 2 μm.

(E) Left panel shows the percentage of cells with overlapped signals in each of the probe pairs. *P* value was calculated by *t* test. Right panel shows the FPKM of the selected genes.

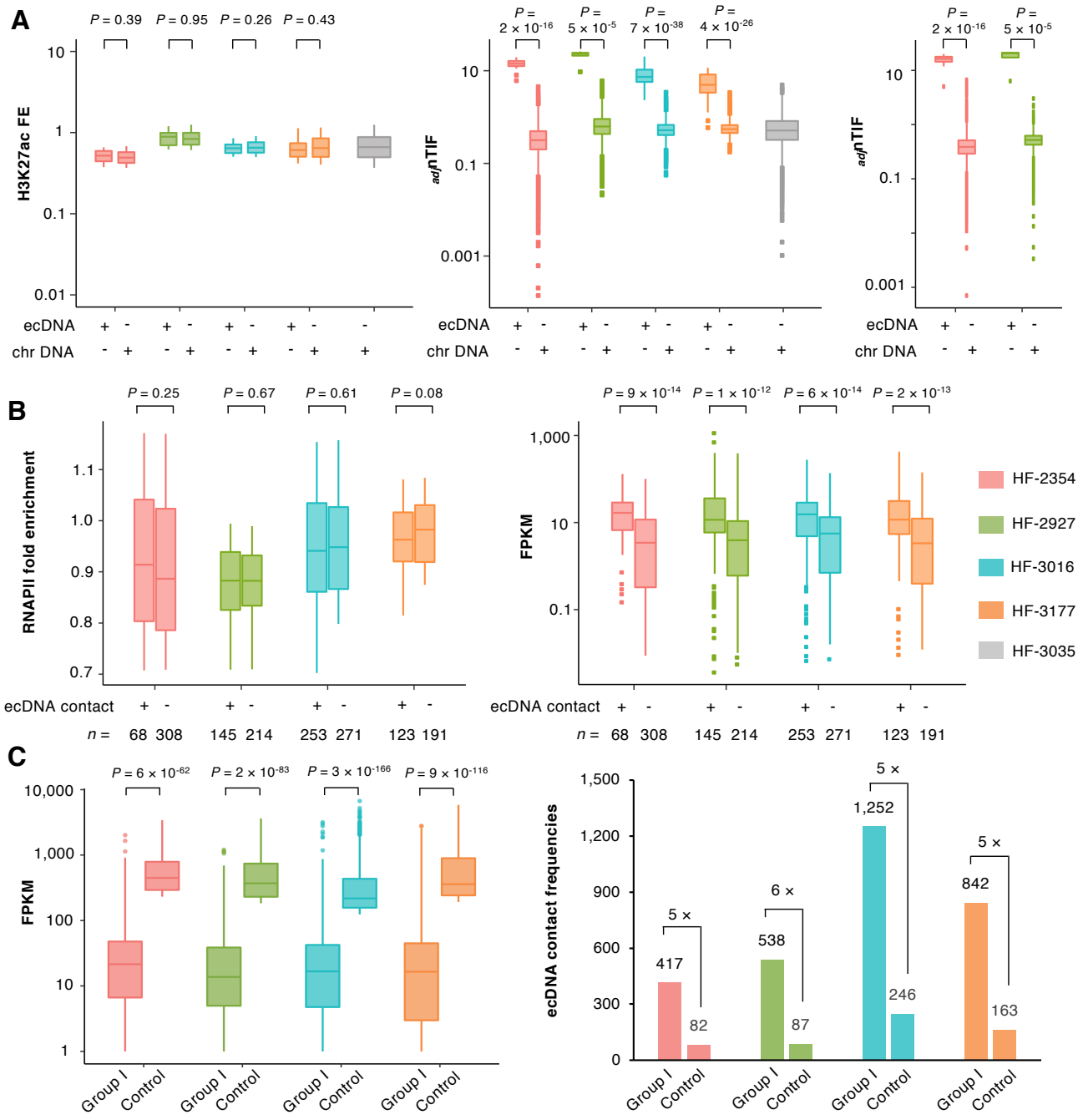


Figure S6. EcDNA-chromosomal contacts are specific, not resulted from random aggregation between regions of active transcription, Related to Figure 4 and 5.

(A) Subsampled genomic bins with matching H3K27ac signal enrichment at 50-Kb resolution from ecDNA and chromosomal regions were selected (Left). For ecDNA regions, $n = 22$, 5, 55 and 39; for chromosomal regions, $n =$ 12,525, 8,337, 10,882, 14,104, 17,538 in each line. Distributions of $_{adj}$nTIFs in the bins determined by ChIA-PET (Middle) and Hi-C (Right) are shown.

(B) Comparison of RNAPII signal (Left) and expression level (Right) between selected chromosomal genes *trans*-interacting with ecDNA and selected genes with no *trans*-chromosomal interactions from each of the four ecDNA (+) cell lines.

(C) Comparison of ecDNA contact frequencies between ecDNA targets (Group I genes) and genes of highest expression abundance in GBM lines. Boxplot shows the FPKM from selected control genes (highest FPKM genes in each of the four cell lines) and the ecDNA targeted Group I genes ($n = 214$, 294, 592, and 399 in HF-2354, HF-2927, HF-3016, and HF-3177, respectively). Right: contact frequencies (sums of total PET counts) between ecDNA to targeted Group I genes and to genes with highest FPKM.

P values of pair-wise comparison by one-sided Wilcoxon Rank-Sum Test were shown except two-sided for A and B Left plots. For all boxplots, center line, median; boxes, first and third quartiles; whiskers, 1.5 × the interquartile range (IQR); points, outliers.

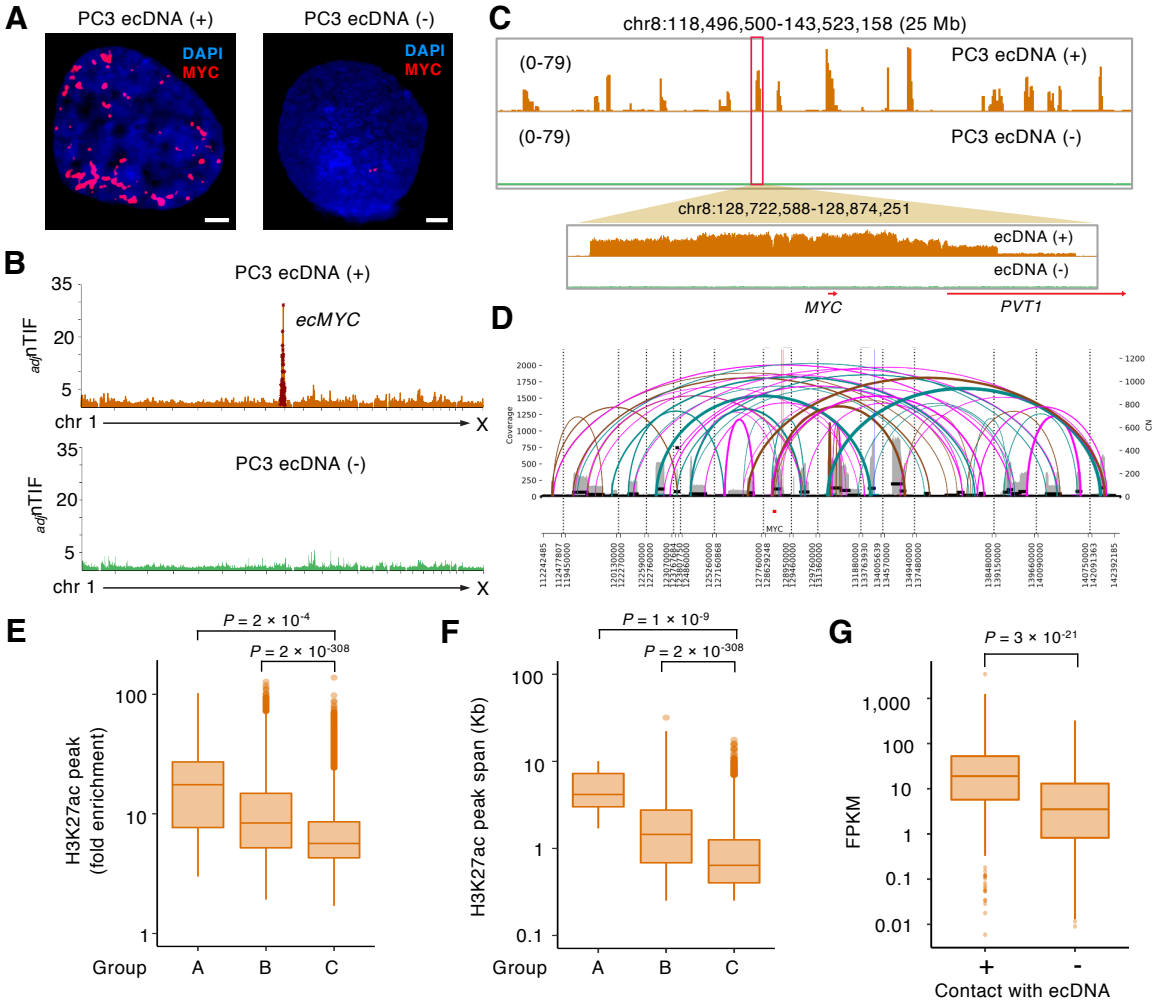


Figure S7. Characterization of *ecMYC* status in PC3 ecDNA (+) and ecDNA (-) cell lines, Related to Figure 5.
(A) FISH analysis with *MYC* DNA probe (red) superimposed with the DAPI staining (blue) in the two PC3 cell lines. Scale bar = 1.5 μm.
(B) Distribution of genome-wide $_{adj}$nTIF at 50-Kb resolution in the PC3 ecDNA (+) and ecDNA (-) cell lines. Elevated $_{adj}$nTIFs are observed on the chromosome 8 regions known to be amplified on *ecMYC*.
(C) Whole-genome sequencing coverage tracks in *ecMYC* region from PC3 ecDNA (+) (upper) and ecDNA (-) (lower) cell lines.
(D) Multiple circular structures of *ecMYC* predicted by AmpliconArchitect analysis.
(E-F) Box plots show the fold enrichment (E) and span size (F) distributions of H3K27ac peaks on *ecMYC* high interaction frequency foci (Group A, n = 14), their corresponding *trans*-interaction chromosomal partners (Group B, n = 8,263) and the genome-wide non-ecDNA and non-ecDNA target peaks (Group C, n = 38,218) from PC3 ecDNA (+) line. All H3K27ac signals in E-F were copy-number normalized. Y-axis are in log10 scales.
(G) The distributions of RNA expression (FPKM) between chromosomal genes *trans*-interacted with ecDNA (n = 238) and genes with no *trans*-chromosomal interactions (n = 514) from PC3 ecDNA (+) line. Y-axis are in log10 scales.
P values are calculated by one-sided Wilcoxon Rank-Sum Test. For all boxplots, center line, median; boxes, first and third quartiles; whiskers, 1.5 × the interquartile range (IQR); points, outliers.

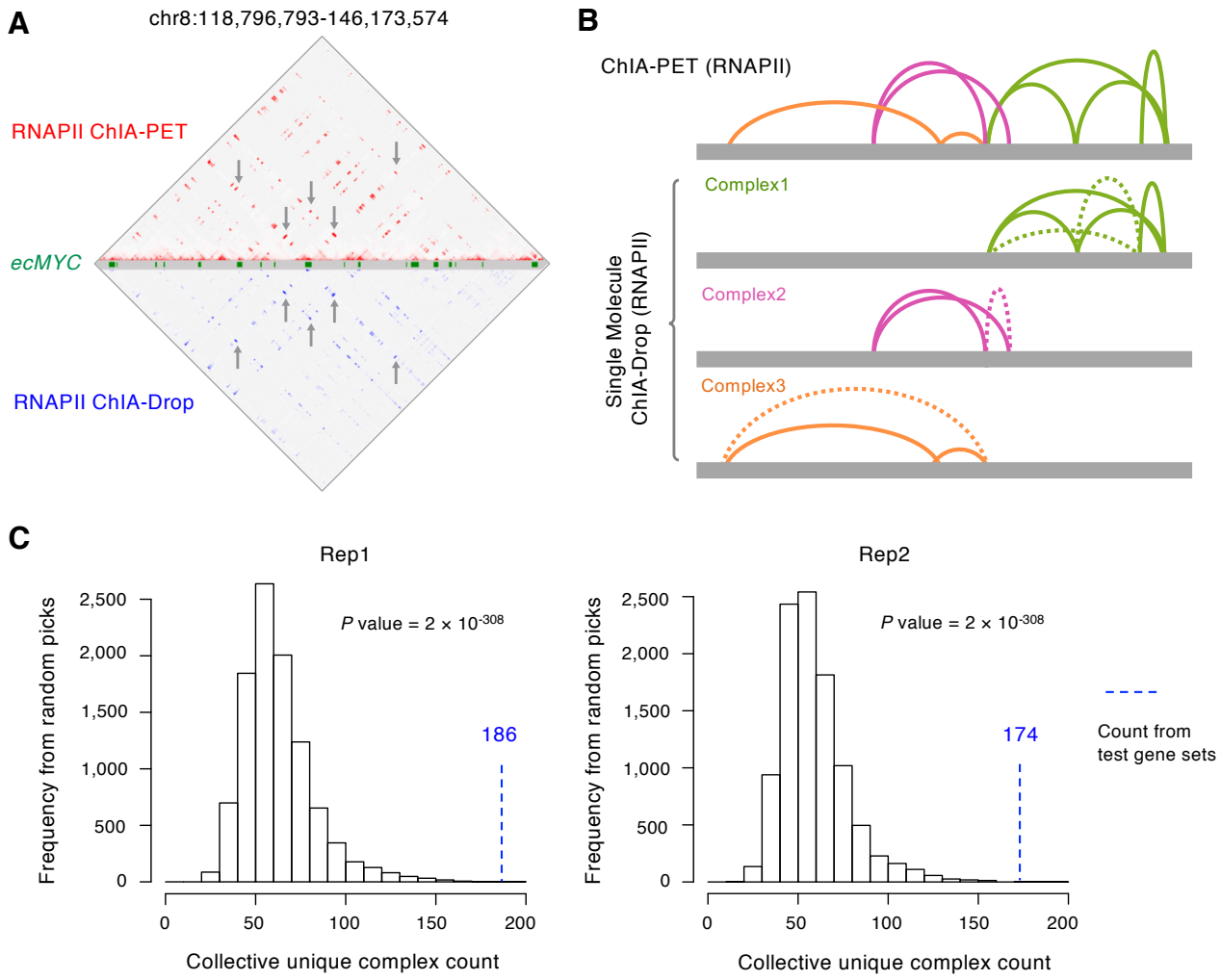


Figure S8. RNAPII ChIA-Drop faithfully captures the multiplex ecDNA-mediated chromatin interactions, Related to Figure 6.

(A) 2D chromatin contact maps of interactions detected by RNAPII ChIA-PET (top) and ChIA-Drop (bottom) in *ecMYC* region. Examples of concordant interaction foci are highlighted by arrows.

(B) Top panel: Loops captured by ChIA-PET color-coded based on the loops captured in individual ChIA-Drop complexes for the genome position chr8:112,200,001-146,364,022. Bottom panel: Interactions detected in three separate ecDNA-associated ChIA-Drop chromatin complexes. Solid circular blocks indicate aligned fragments. Fragments from the same complex are connected by horizontal line. Loops connect these fragments to indicate proximity in 3D genome space. Solid loops are detected by ChIA-PET. Dashed loops are proximity information inferred only in the ChIA-Drop data.

(C) The distribution of chromatin complex counts from randomly (randomized 10,000X) selected gene sets with comparable expression level from two ChIA-Drop replicates. The numbers of complexes detected from the upregulated genes were indicated in blue. *P* values determined for higher complex counts in the up-regulated genes are shown (one-sided one sample Wilcoxon Test).

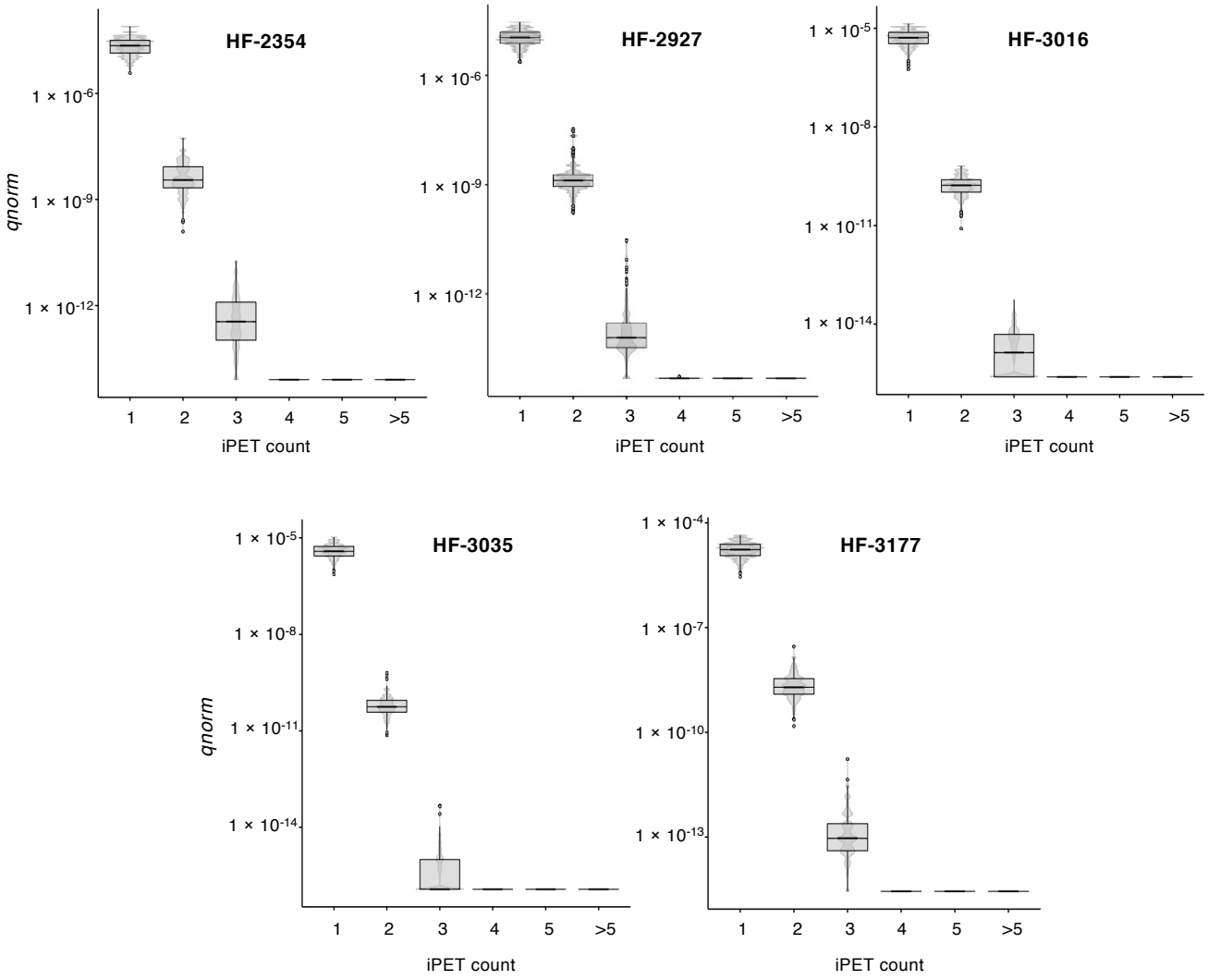


Figure S9. Distribution of normalized q values (*qnorm*) (Y-axis) for *trans*-chromosomal interaction as a function of iPET counts (X-axis) across five GBM derived cell lines, Related to STAR Methods.