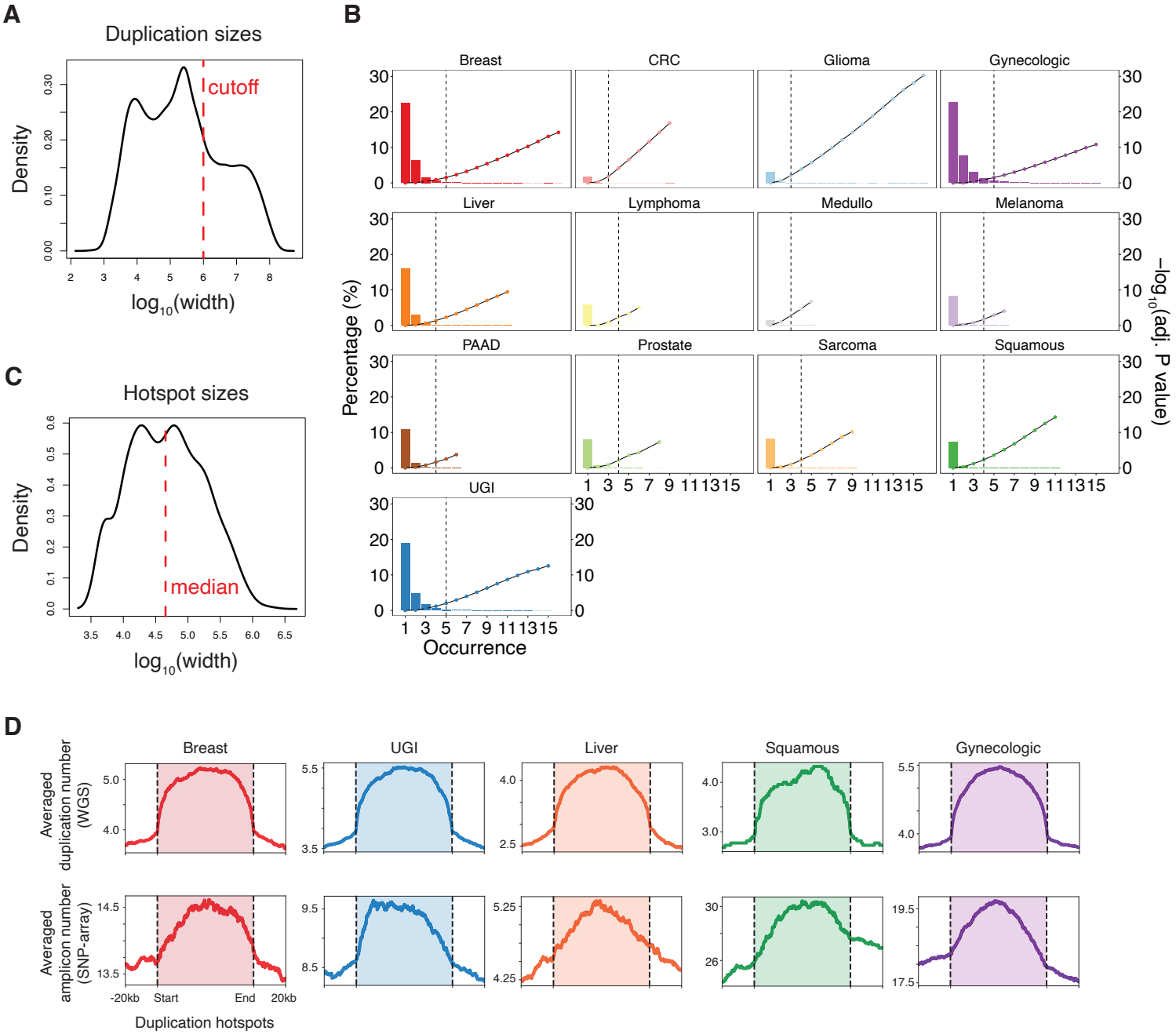
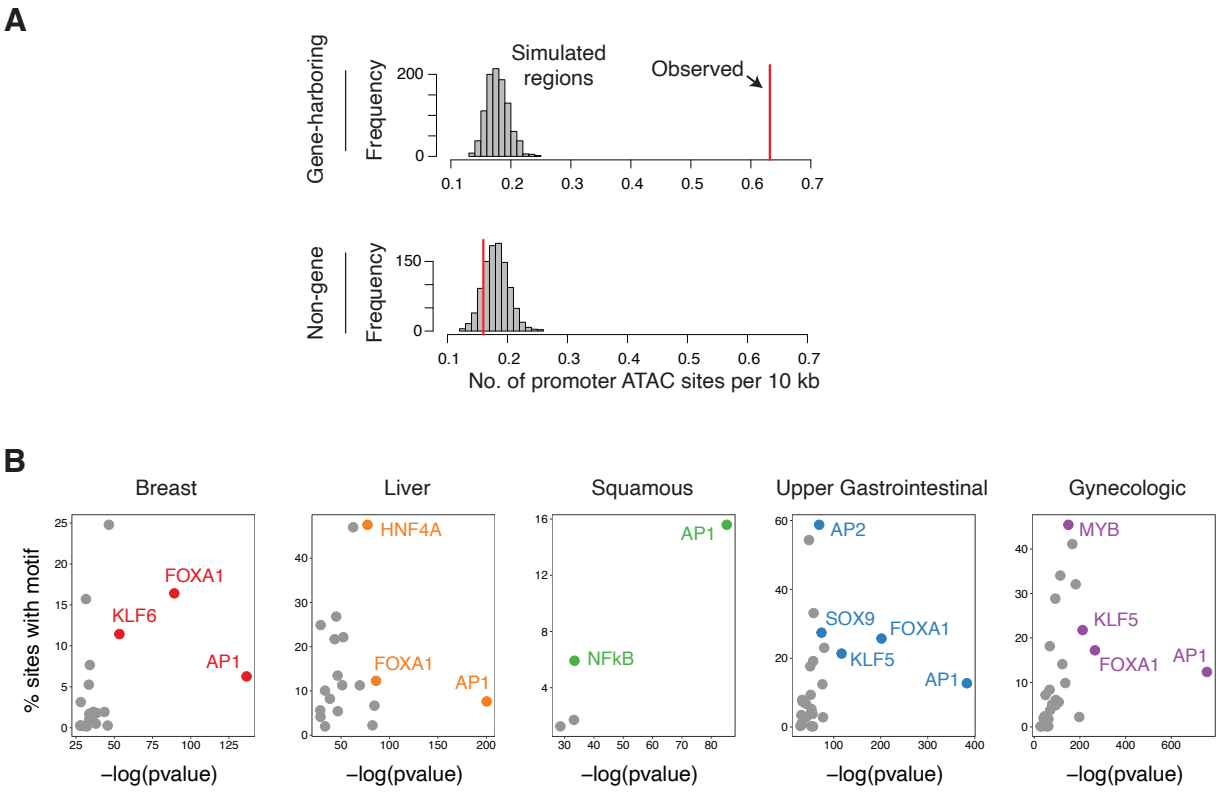
Supplementary Figure 1



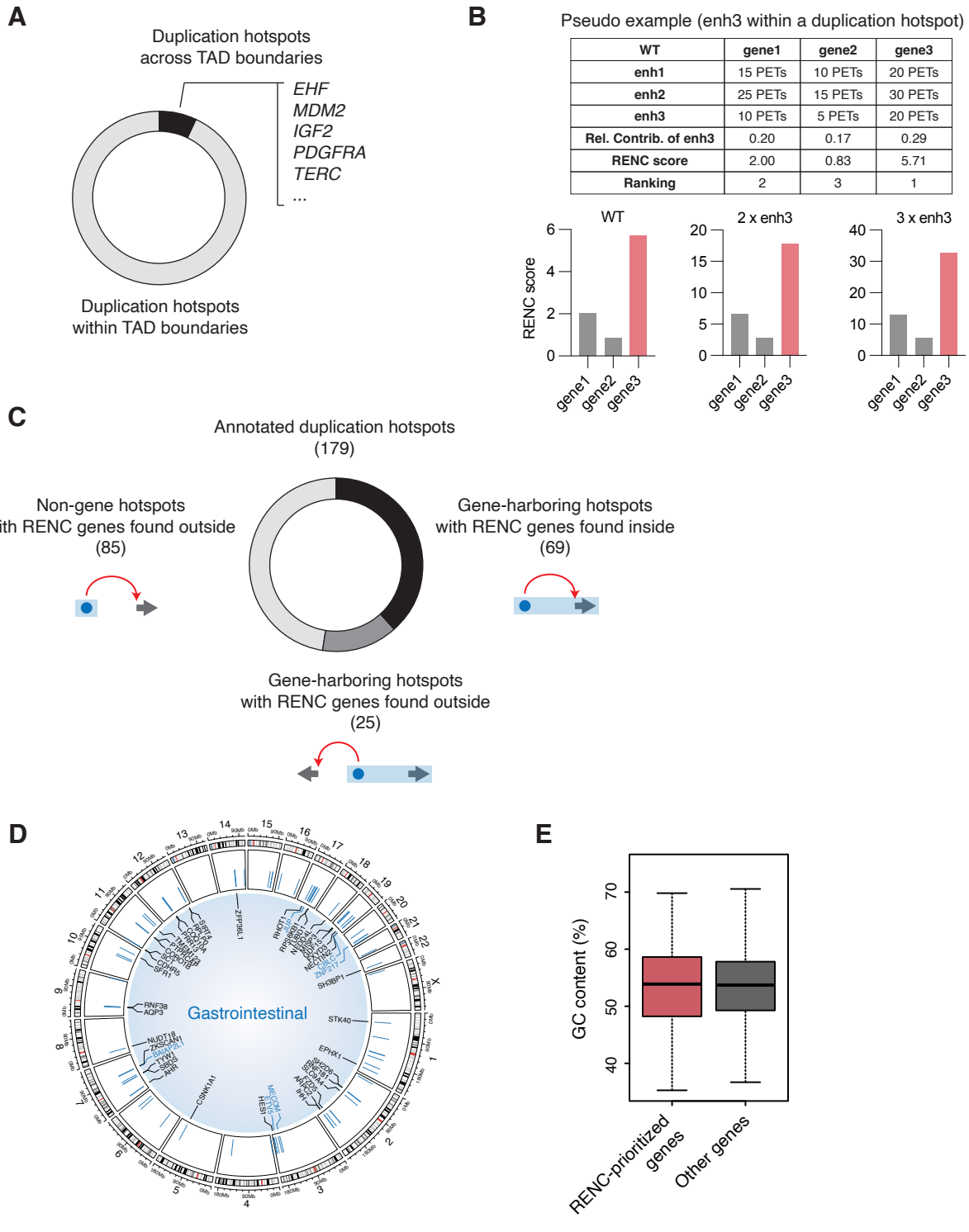
Supplementary Figure 1:
A. Size distribution of duplications called in the PCAWG project.
B. The occurrence of duplications at 5 kb bins (x axis), their frequency (bars, corresponding to the left y axis), and associated P values (curves, corresponding to the right y axis).
C. Size distribution of duplication hotspots identified in this study.
D. Averaged number of duplication events called from PCAWG WGS data (upper) and potential amplification events called from TCGA SNP-array data (mean segment score > 0.2) (bottom) at the identified duplication hotspots.

Supplementary Figure 2



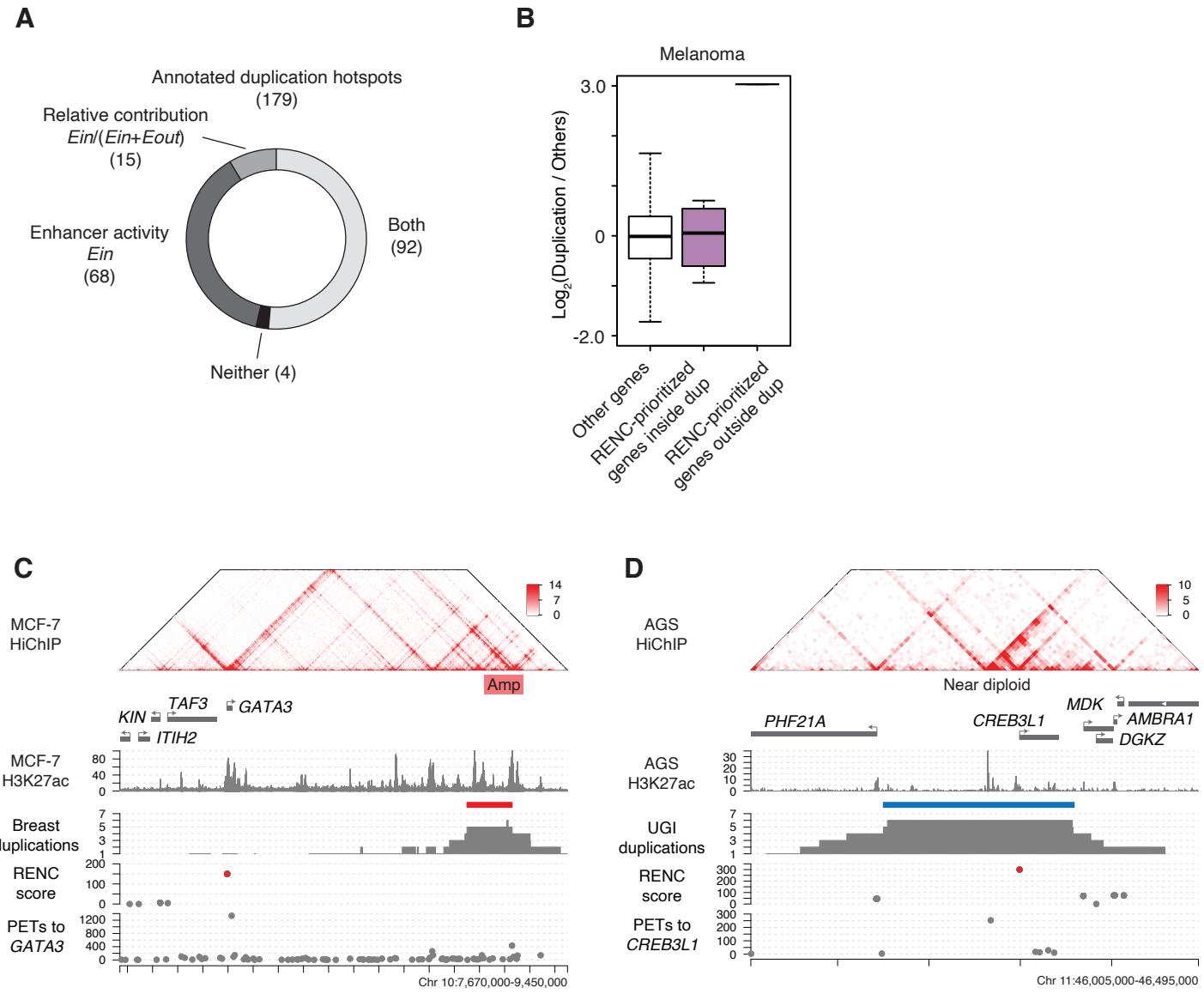
Supplementary Figure 2:
A. Number of promoter ATAC sites per 10 kb within the identified duplication hotspots or randomly shuffled regions matching the lengths of the hotspots.
B. HOMER analysis showing transcription factor motifs enriched in distal ATAC sites within the identified duplication hotspots. Highlighted are the ones that are relevant to the corresponding cancer types. The percentage of sites present with the motifs and the P values are calculated by the HOMER pipeline.

Supplementary Figure 3



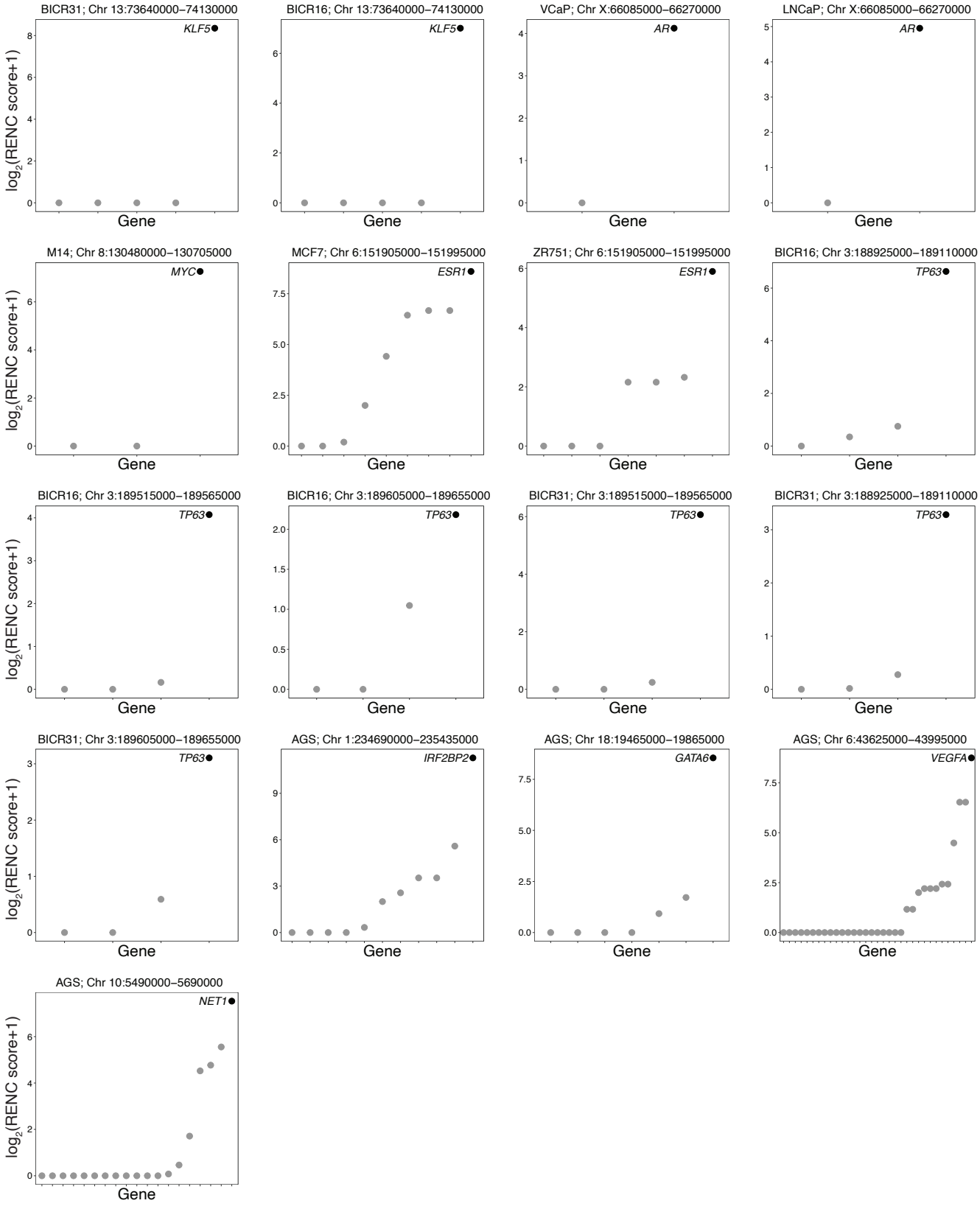
Supplementary Figure 3:
A. ~7% of the identified duplication hotspots span TAD boundaries.
B. A pseudo example showing the duplication status of an enhancer (enh3) affects the RENC scores but not the ranking of the target genes.
C. RENC-prioritized genes for the majority of the duplication hotspots are outside of the hotspot regions.
D. Circos plots presenting the genomic positions of the identified duplication hotspots (the ones not included in Figure 4B) in upper gastrointestinal cancer, their duplication occurrences (scale: 0 to >=10), and their associated target genes based on the RENC analysis.
E. The GC content (%) in promoters of RENC-prioritized genes versus the other genes.

Supplementary Figure 4



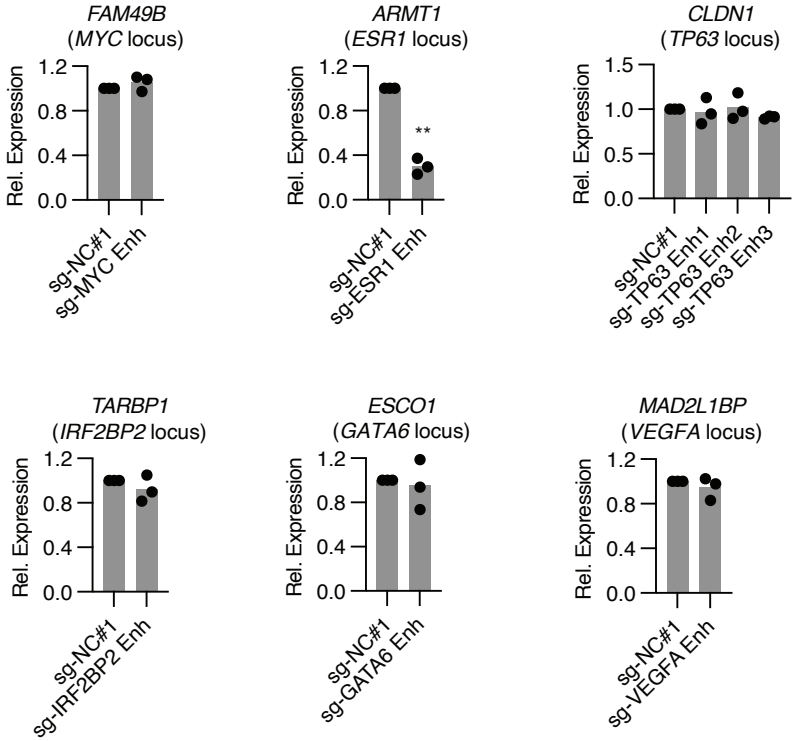
Supplementary Figure 4:
A. The annotated duplication hotspots are grouped based on if their RENC-prioritized genes are the same as the ones prioritized by the hotspot-delivered enhancer activity (*Ein*) or the relative contribution as compared to all the linked enhancers of each gene (*Ein*/(*Ein*+*Eout*)). "Both" indicate that both *Ein* and *Ein*/(*Ein*+*Eout*) prioritize the same genes as the ones prioritized by RENC for the hotspot.
B. Expression fold change (log_2-transformed) of RENC-prioritized target genes for the duplication hotspots in melanoma samples with the associated duplications versus the ones without duplications or deletions of the hotspot regions. RENC genes inside dup: RENC-prioritized target genes that are within the duplication hotspots; RENC genes outside dup: RENC-prioritized target genes that are outside of the duplication hotspots; Other genes: genes that are not prioritized for the duplication hotspots based on RENC analysis, which are used as negative controls. For the "other genes", expression fold changes between samples with duplications of any hotspot versus the ones without duplications or deletions of these hotspots were calculated.
C-D. Duplications of enhancers are linked to *GATA3* in breast cancer (C) and *CREB3L1* in upper gastrointestinal cancer (D). Presented tracks include H3K27ac HiChIP signal, H3K27ac ChIP-seq signal, positions of duplication hotspots, duplication events observed in the corresponding cancer types, RENC scores prioritizing gene promoters that are more likely to be activated by enhancers within the duplication hotspot, and the number of PETs connecting each enhancer to the RENC-prioritized gene promoter. The SNP-array-based copy number status of the hotspot in the corresponding cancer cell lines is indicated.

Supplementary Figure 5



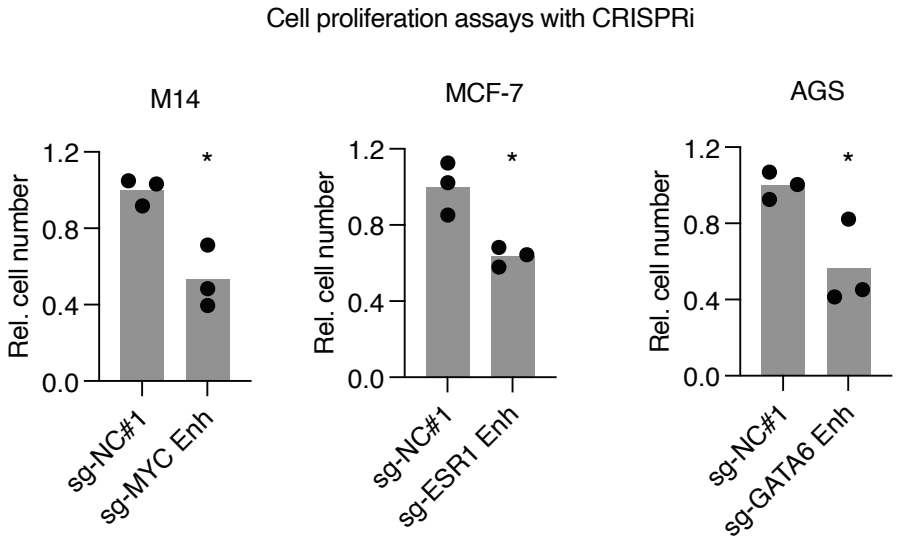
Supplementary Figure 5:
For each duplication hotspot highlighted in the main figures, we plotted the RENC scores of the RENC-prioritized genes, other genes linked to the hotspot via enhancer-promoter loops, and other genes within +/- 500 kb of the hotspot.

Supplementary Figure 6



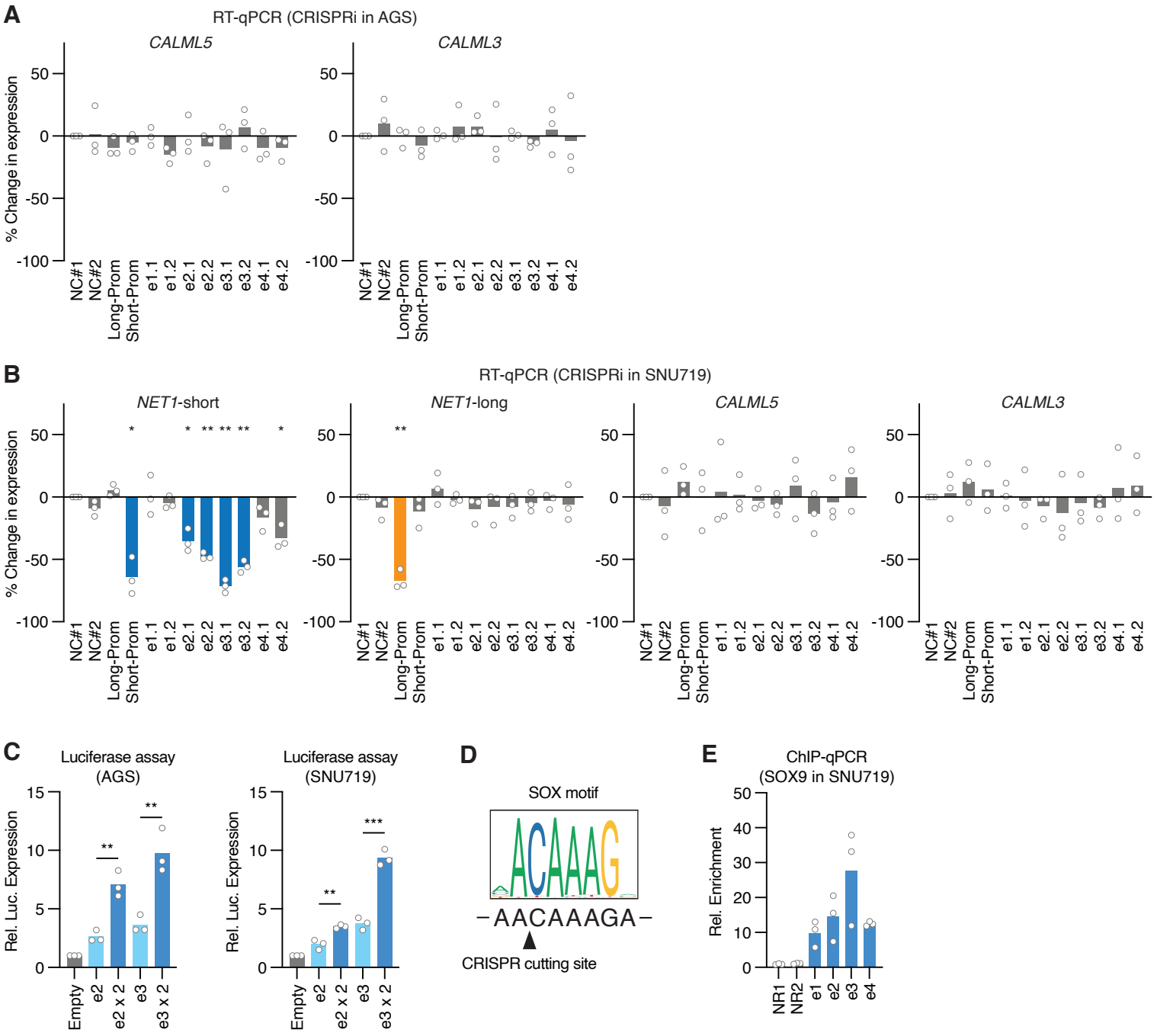
Supplementary Figure 6:
RT-qPCR of genes nearby the RENC-prioritized genes for six selected duplication hotspots in the corresponding cell lines with and without CRISPRi of the enhancers indicated in Figure 5. As described in the main text, these genes are either ranked second to the prioritized ones (*ARMT1* for the *ESR1* locus, *CLDN1* for the *TP63* locus, *TARBP1* for the *IRF2BP2* locus, *ESCO1* for the *GATA6* locus, and *MAD2L1BP* for the *VEGFA* locus) or adjacent to the duplication hotspot if no other genes are linked to the hotspot via enhancer-promoter loops (*FAM49B* for the *MYC* locus). Note that some of the genes such as *CLDN1* and *ESCO1* are outside of the windows presented in Figure 5. n = 3 biological replicates. P value was derived from a two-sided *t*-test: ** < 0.01.

Supplementary Figure 7



Supplementary Figure 7:
Cell proliferation results in M14, MCF-7, and AGS cells with and without CRISPRi of the enhancers indicated in Figure 5. The cell numbers (7 days post seeding) were normalized to those of the negative control sgRNA NC#1 (n = 3 biological replicates). P values were derived from two-sided *t*-tests: * < 0.05.

Supplementary Figure 8



Supplementary Figure 8:
A. RT-qPCR results showing expression changes of *CALML5* and *CALML3* after CRISPRi of the enhancers e1-e4 in AGS cells.
B. RT-qPCR results showing expression changes of *NET1*-short, *NET1*-long, *CALML5,* and *CALML3* after CRISPRi of the enhancers e1-e4 in SNU719 cells. P values were derived from two-sided *t*-tests: * < 0.05; ** < 0.01.
C. Luciferase assays measuring the activity of the e2 and e3 enhancers and their duplications (x 2). The signal was normalized to Renilla (co-transfected) and then to the Empty control. P values were derived from two-sided *t*-tests: ** < 0.01;*** < 0.001.
D. The presence of a SOX motif in the e3 enhancer. Indicated is the CRISPR cutting site that is used to disrupt the motif.
E. SOX9 ChIP-qPCR results showing the enrichment of SOX9 at the e1-e4 enhancers in SNU719 cells, with the strongest signal observed at e3. NR1 and NR2 are two negative control regions. The ChIP-qPCR signal from each region was normalized to the genomic input (sonicated DNA without ChIP) and then normalized to the averaged signal of NR1 and NR2.