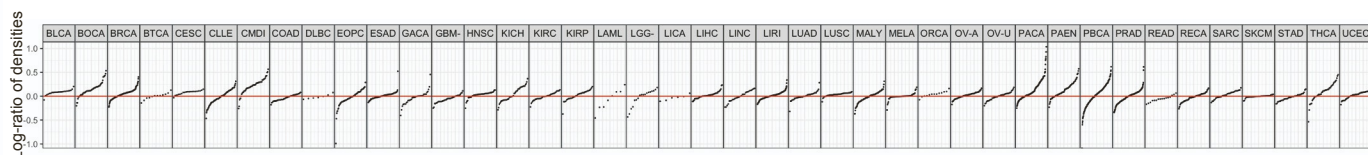
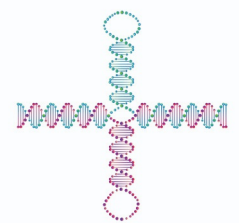


Supplemental information

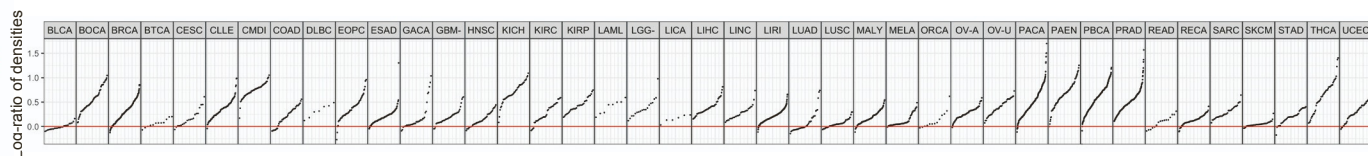
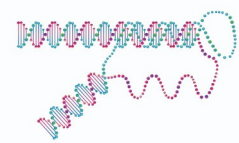
**APOBEC mutagenesis is low in most
types of non-B DNA structures**

Gennady V. Ponomarev, Bulat Fatykhov, Vladimir A. Nazarov, Ruslan Abasov, Evgeny Shvarov, Nina-Vicky Landik, Alexandra A. Denisova, Almira A. Chervova, Mikhail S. Gelfand, and Marat D. Kazanov

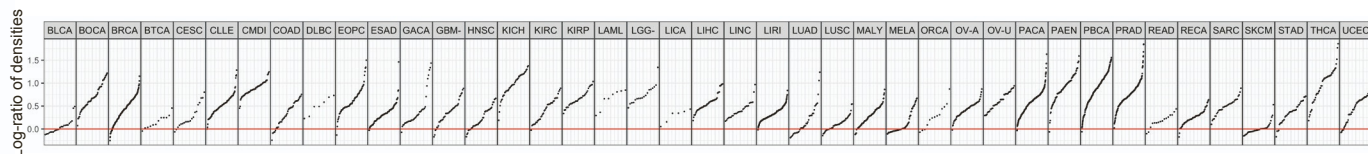
Inverted repeats



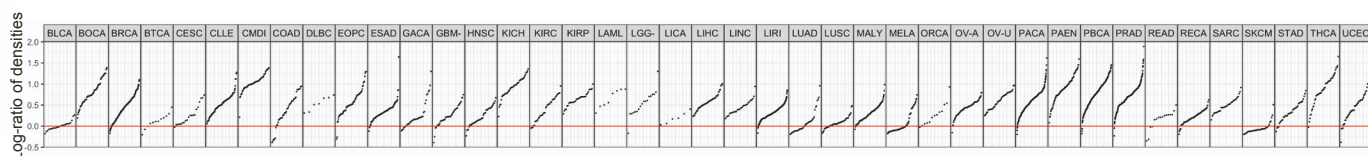
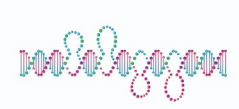
Mirror repeats



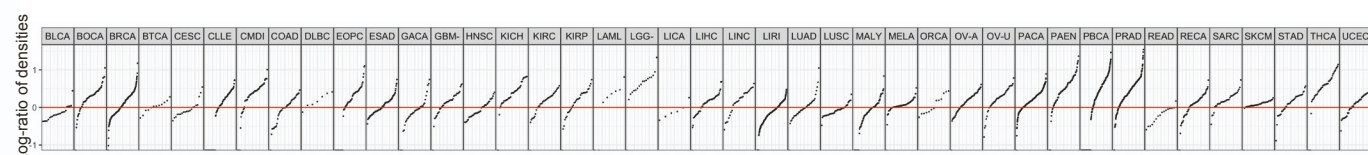
Direct repeats



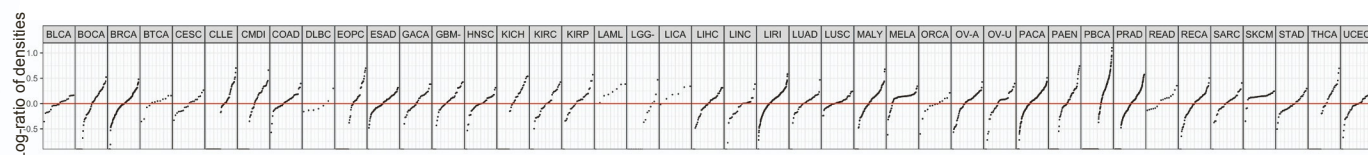
Short tandem repeats



G-quadruplex



A-phased repeats



Z-DNA

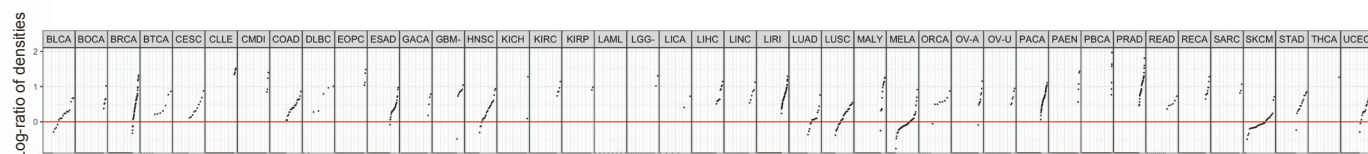
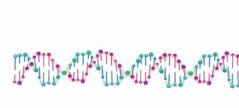


Figure S1. Enrichment of APOBEC-signature (TpC) somatic single-base substitutions (SBS) observed in various types of non-canonical DNA regions and different cancers, Related to Figure 1.

Single dot corresponds to a particular cancer sample, with the vertical position indicating the log-ratio of the mutational density in non-B DNA structure genome regions to the mutational density in B-DNA genome regions.

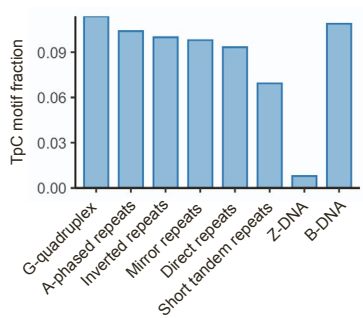
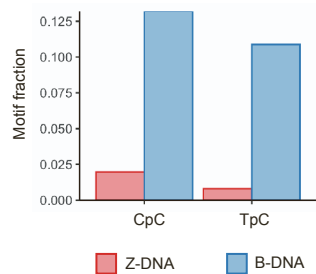
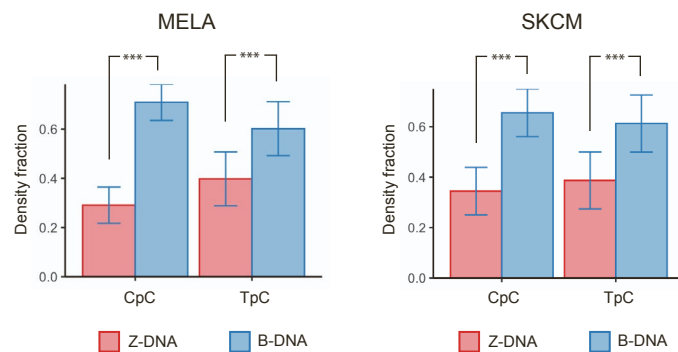
A**B****C**

Figure S2. Fraction of the APOBEC-signature motif in non-B DNA structures, Related to STAR Methods.

(a) Fraction of the APOBEC-signature motif, TpC, in each type of non-B DNA structure. (b) Fraction of the UV-signature motifs, CpC and TpC, in Z- and B-DNA. (c) Densities of UV-induced mutations in TpC and CpC di-nucleotides in Z- and B-DNA. Wilcoxon-Mann-Whitney test notation: *** – P -value < 0.001. Data are represented as mean \pm SEM.

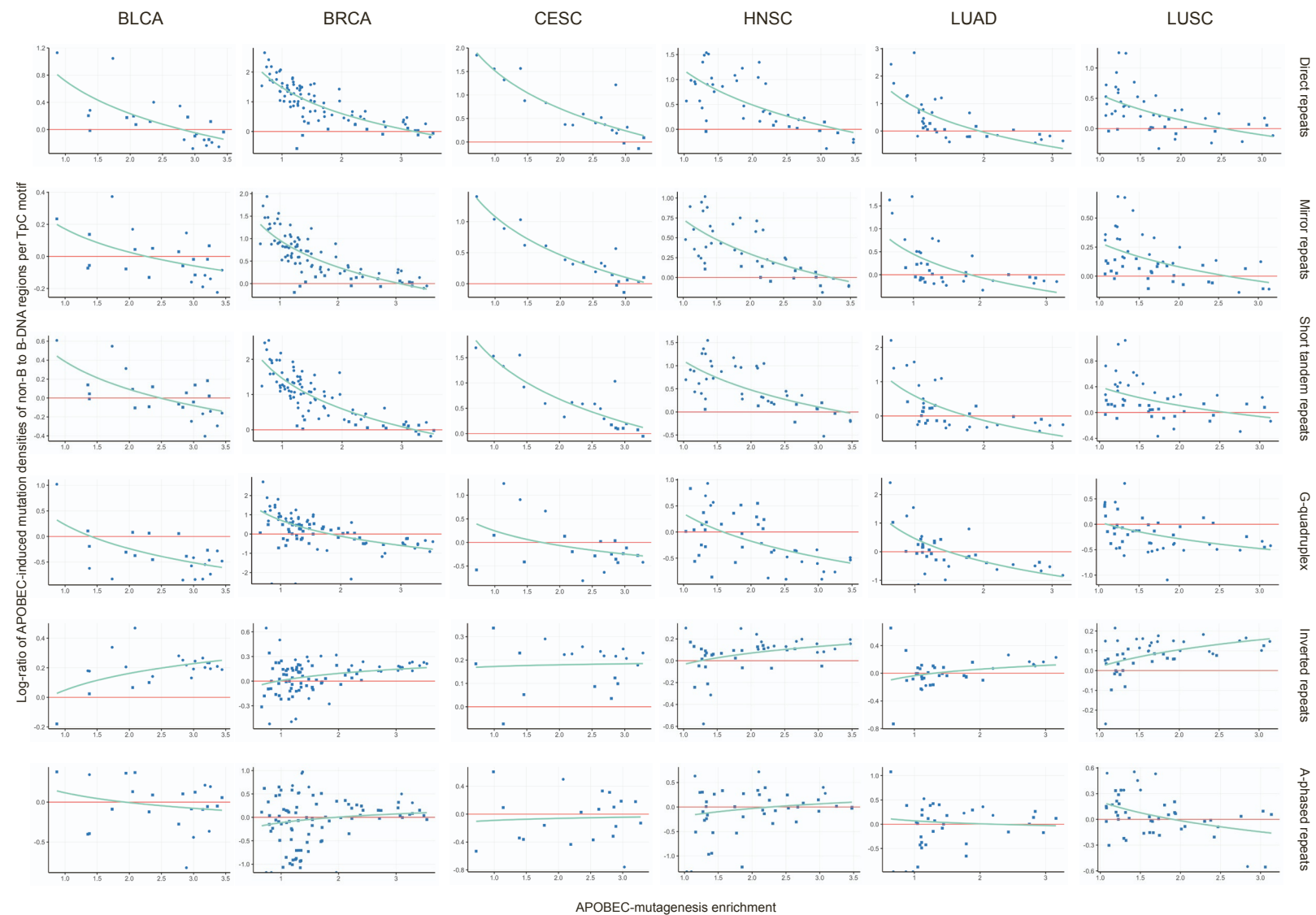


Figure S3. Dependence of the log-ratio of APOBEC-induced mutation densities in non-B DNA to B-DNA genome regions on the activity of APOBEC mutagenesis in cancer samples, Related to Figure 2.
 Point shape – round or square – corresponds to significant or insignificant statistical differences between two densities in a particular cancer sample, respectively.

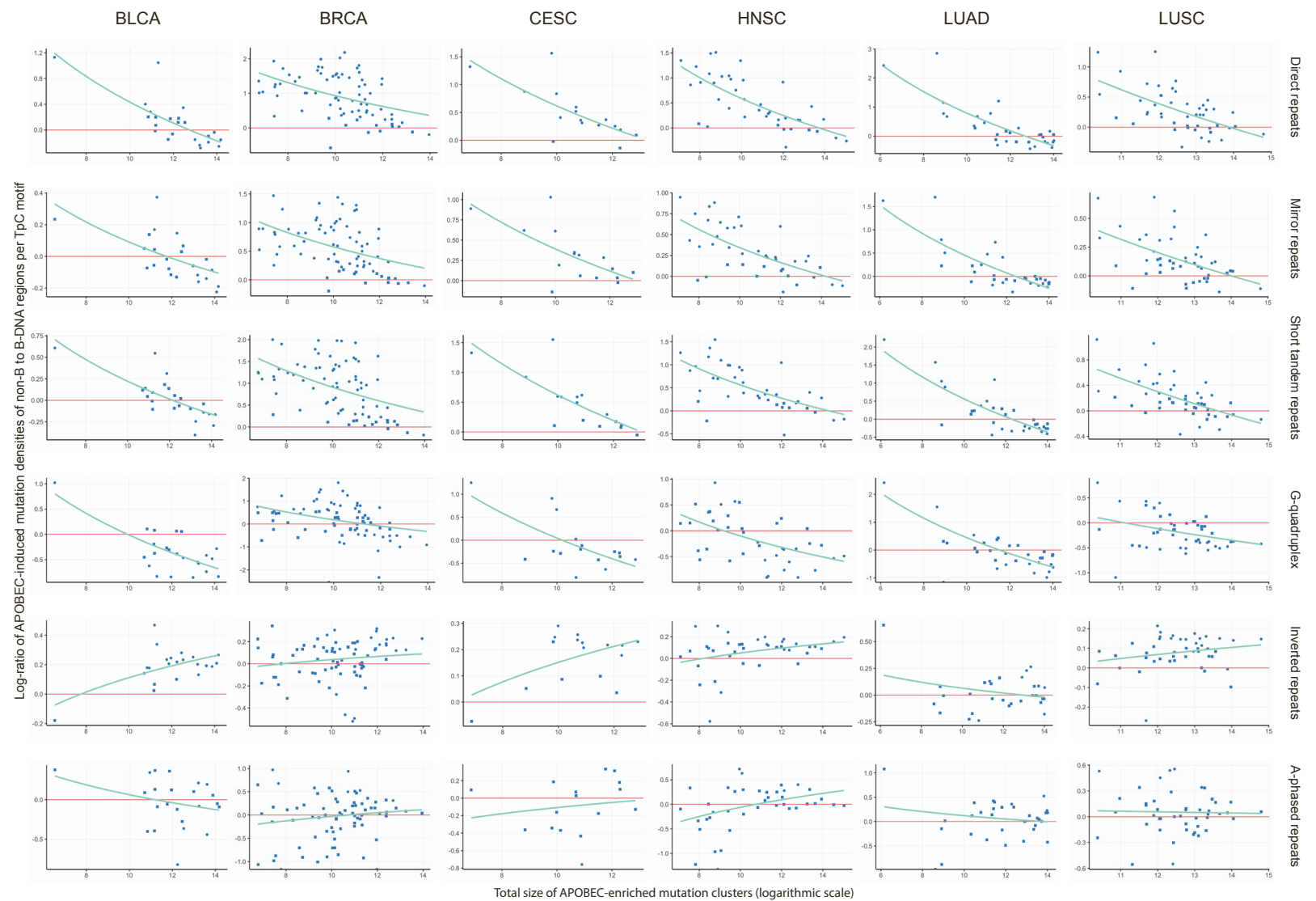
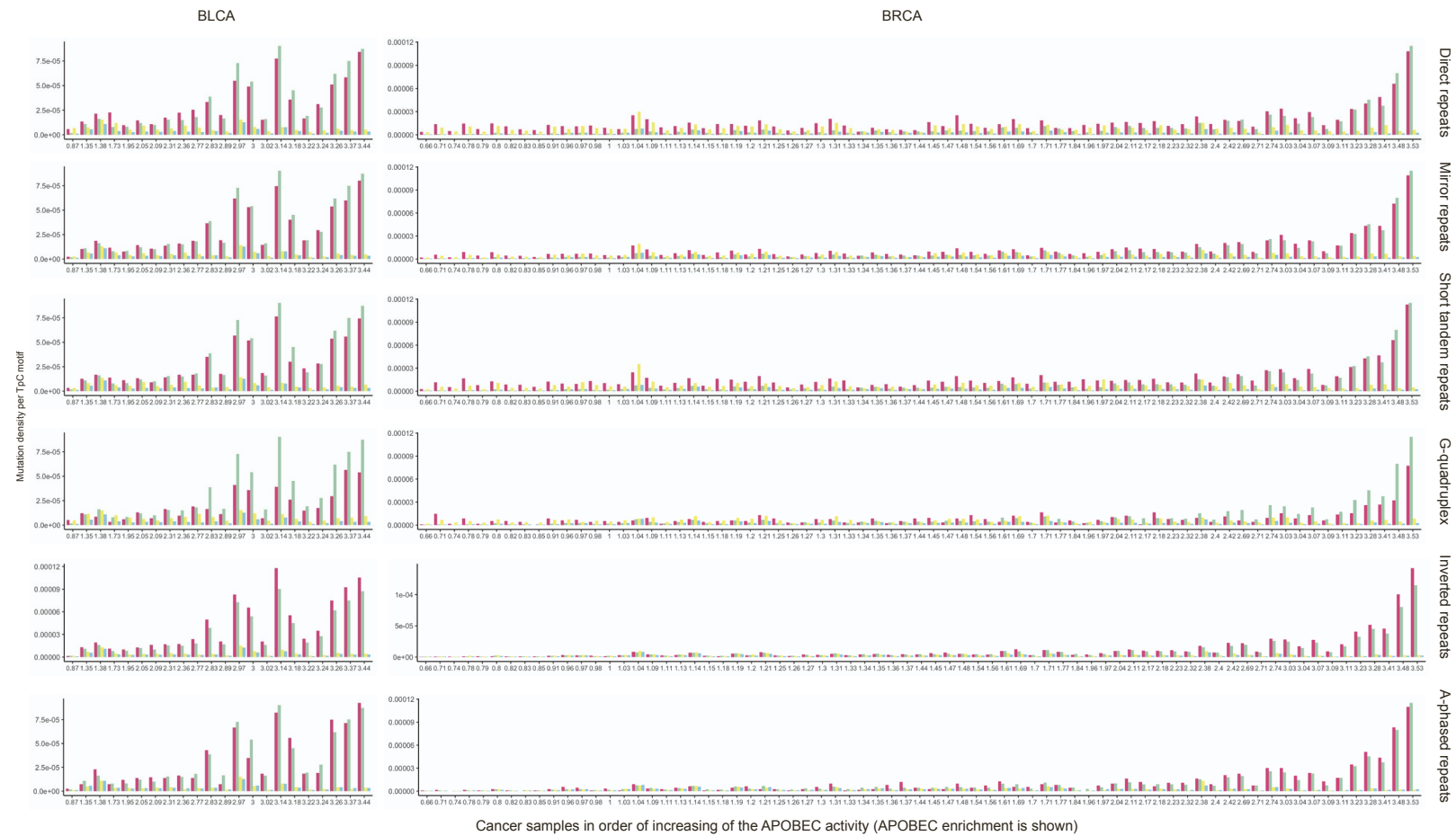
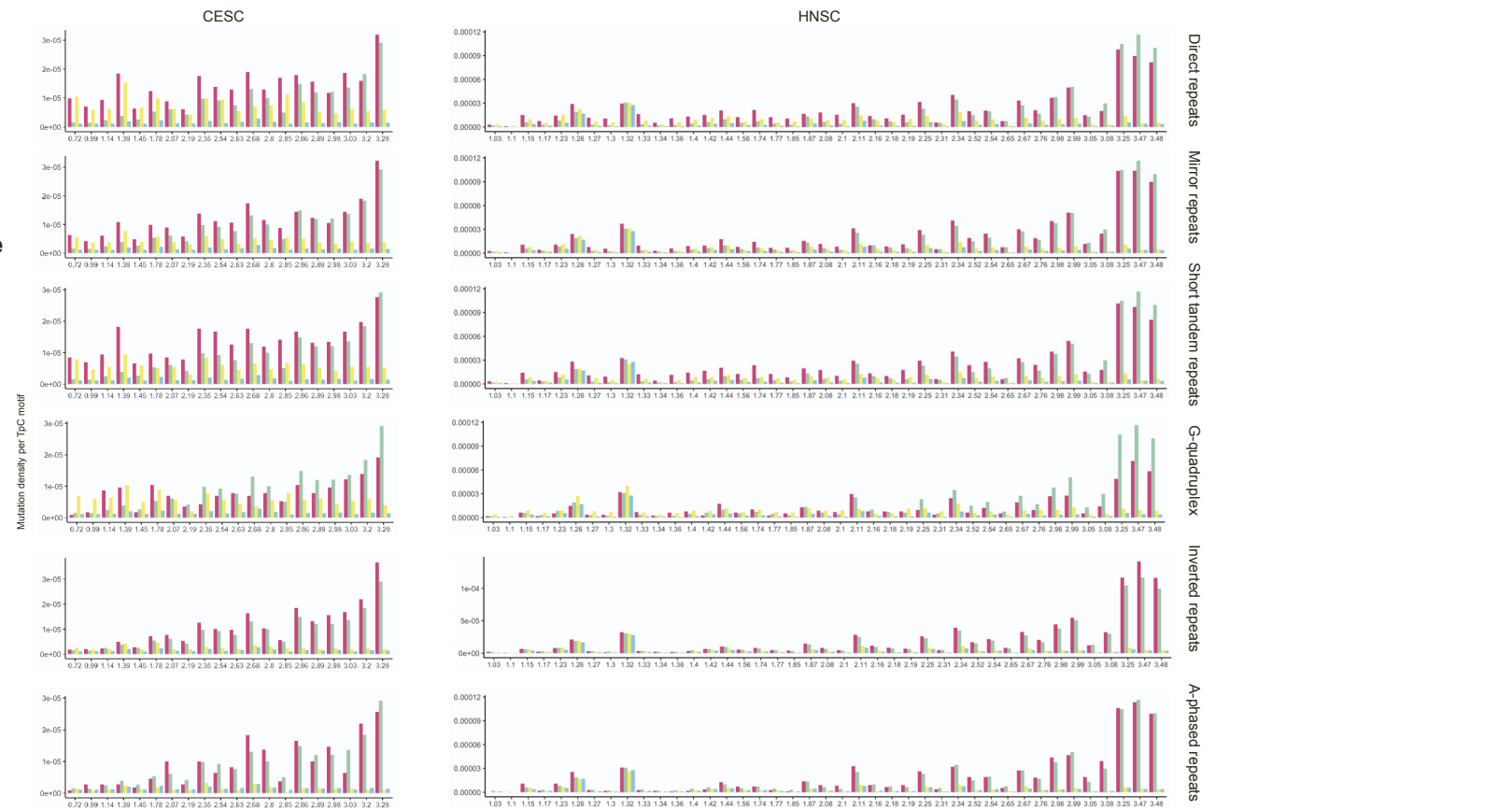


Figure S4. Dependence of the log-ratio of APOBEC-induced mutation densities in non-B DNA to B-DNA genome regions on the total size of APOBEC-enriched mutation clusters in cancer samples, Related to Figure 2.
 Notation as in Fig. S3.



Cancer samples in order of increasing of the APOBEC activity (APOBEC enrichment is shown)



Cancer samples in order of increasing of the APOBEC activity (APOBEC enrichment is shown)

■ APOBEC-signature mutations in non-B DNA regions
 ■ Non-APOBEC signature mutations in non-B DNA regions
■ APOBEC-signature mutations in B-DNA regions
 ■ Non-APOBEC signature mutations in B-DNA regions

Figure S5. APOBEC-induced and other mutation densities in non-B and B-DNA genome regions of BRCA, BLCA, CESC, and HNSC cancers, Related to Figure 2.

LUAD

LUSC

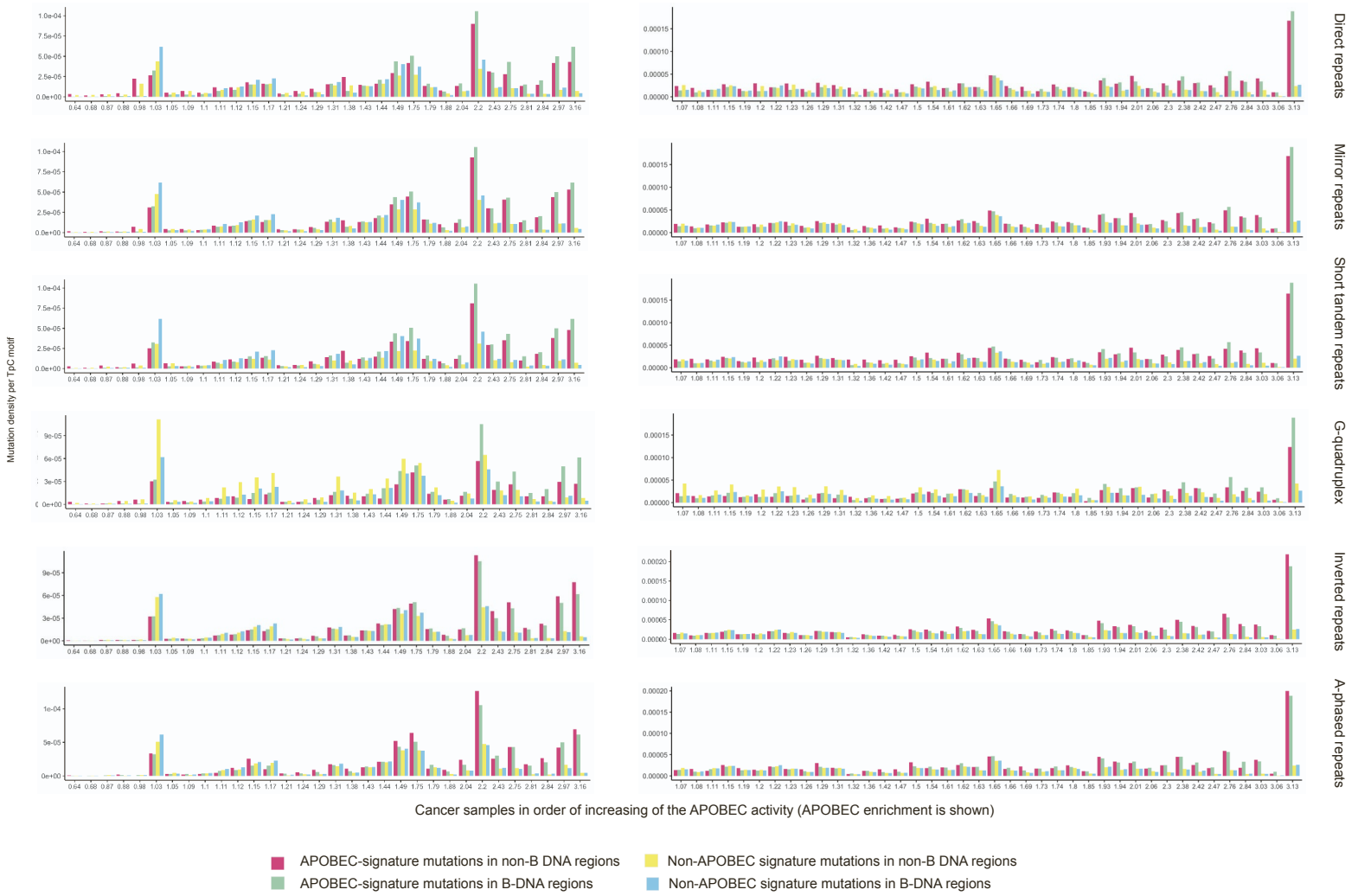


Figure S6. APOBEC-induced and other mutation densities in non-B and B-DNA genome regions of LUAD and LUSC cancers, Related to Figure 2.

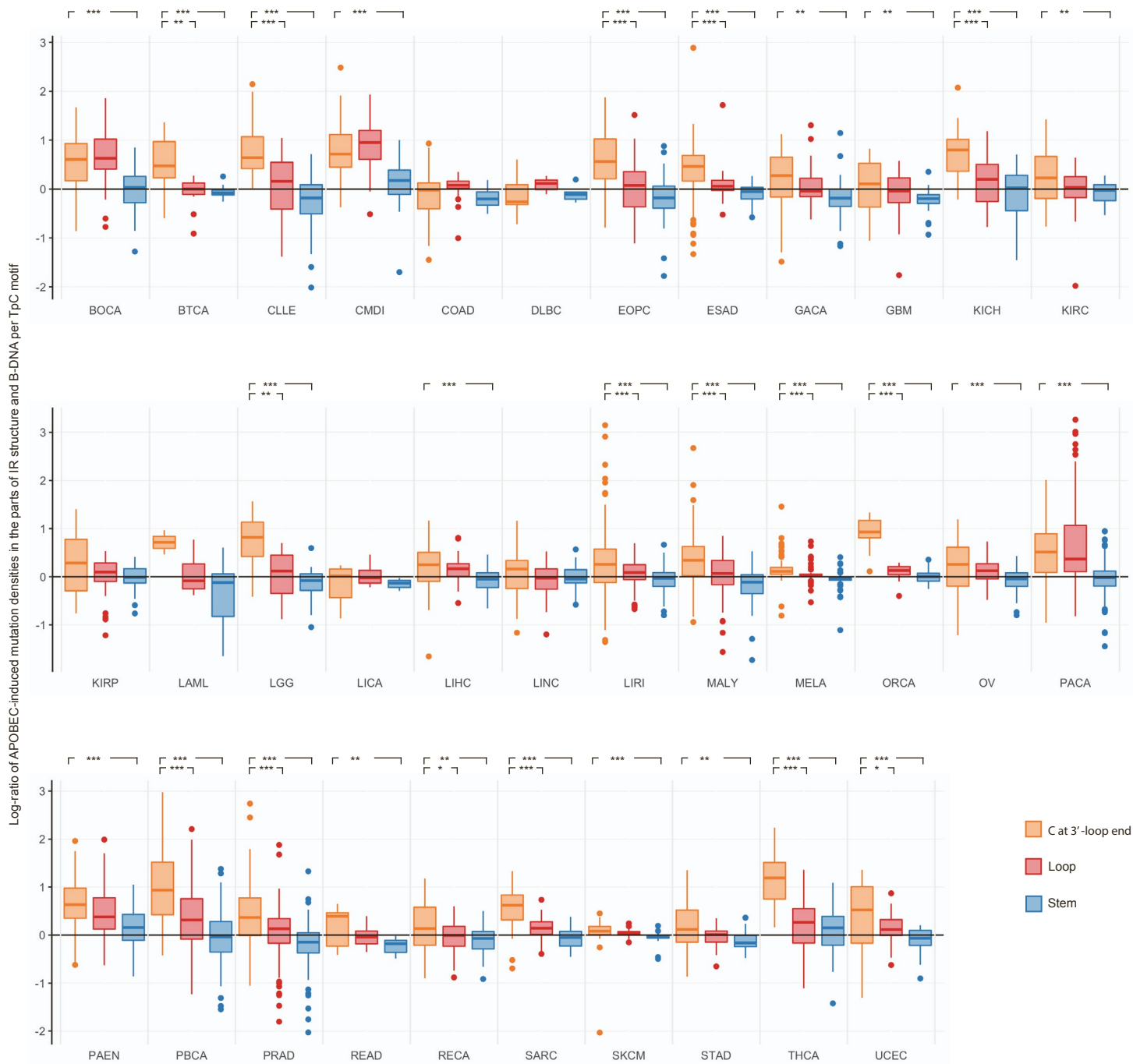


Figure S7. Distribution of APOBEC-signature mutational densities at TpC motifs in different parts of the IR secondary structure in APOBEC-negative cancers, Related to Figure 4.

Wilcoxon-Mann-Whitney test notation: *** – P -value < 0.001, ** – P -value < 0.01, * – P -value < 0.05. Data are represented as box plots displaying minimum, first quartile, median, third quartile, and maximum values.

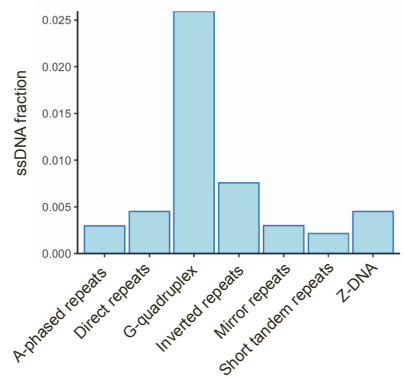


Figure S8. Fraction of ssDNA in non-B DNA structures, Related to STAR Methods.

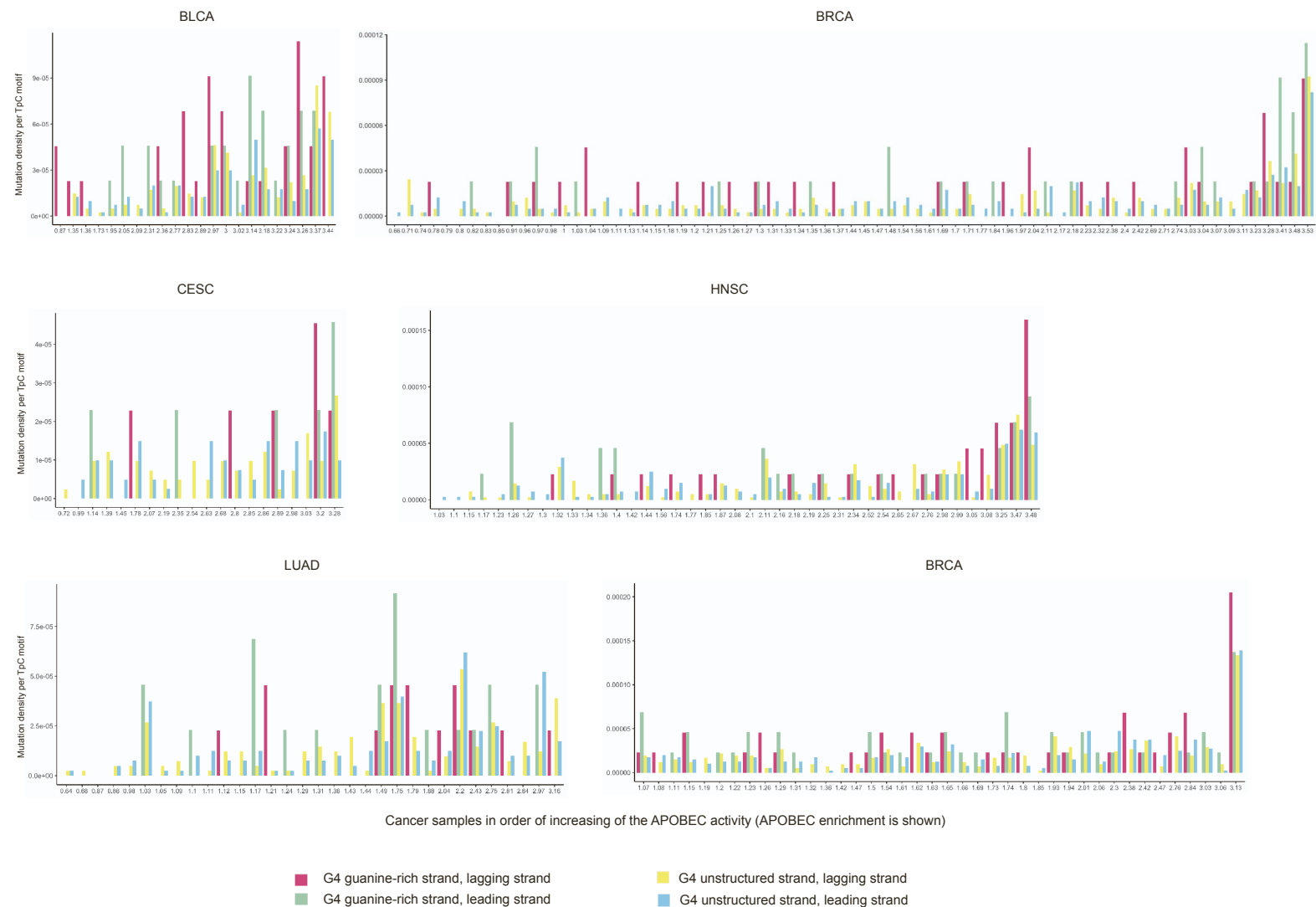


Figure S9. APOBEC-induced mutation density in the G-quadruplex guanine-rich and unstructured strands, divided into regions of lagging and leading strands of DNA replication, Related to Figure 5.

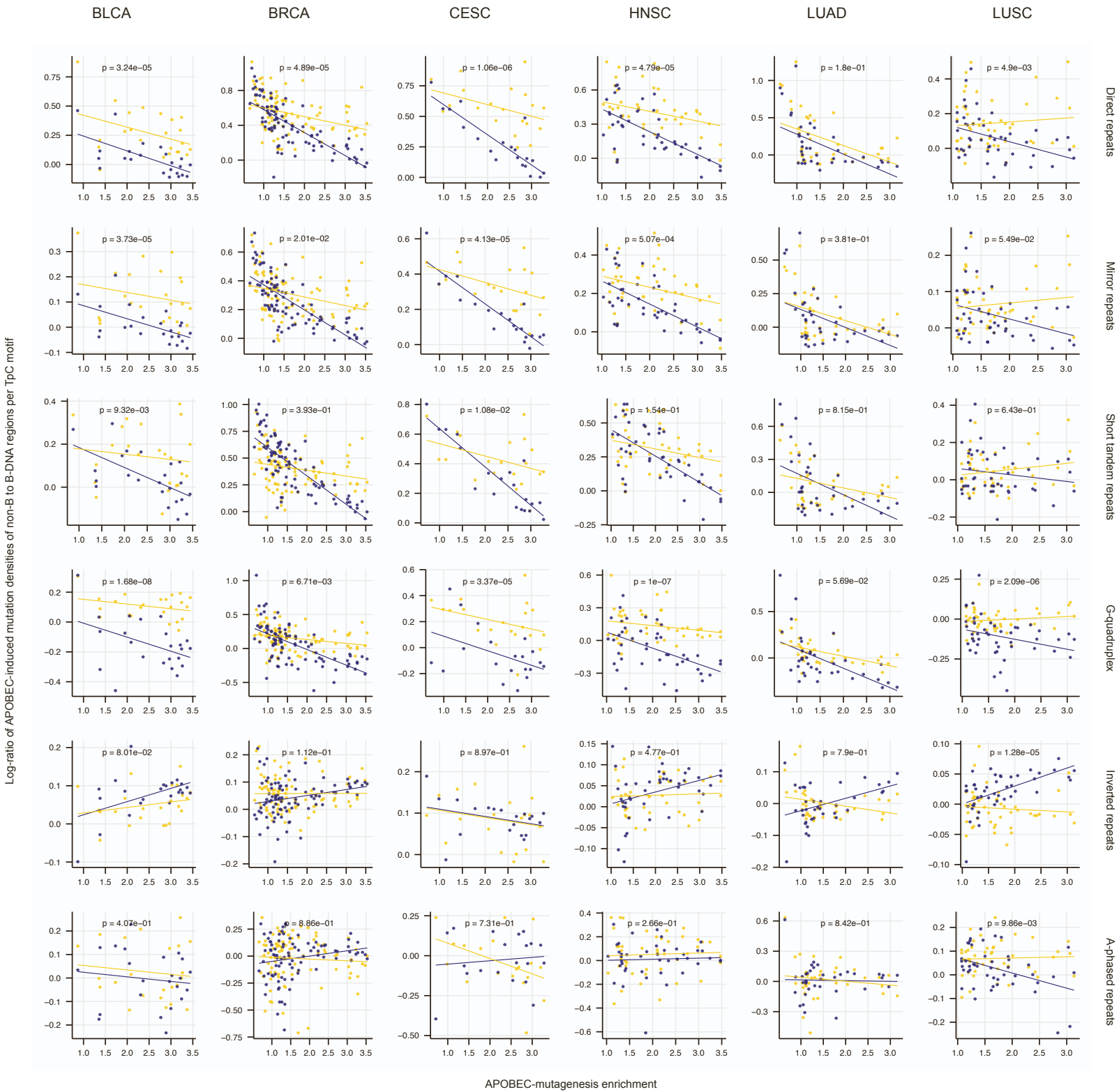


Figure S10. Dependence of the log-ratio of C->T and C->G mutation densities in non-B DNA to B-DNA genome regions in TpC (violet) and VpC (yellow) motifs on the activity of APOBEC mutagenesis in cancer samples, Related to Figure 2. The presented p-values estimate statistical significance of differences between slopes of two regression lines by combining both datasets and introducing the interaction term.

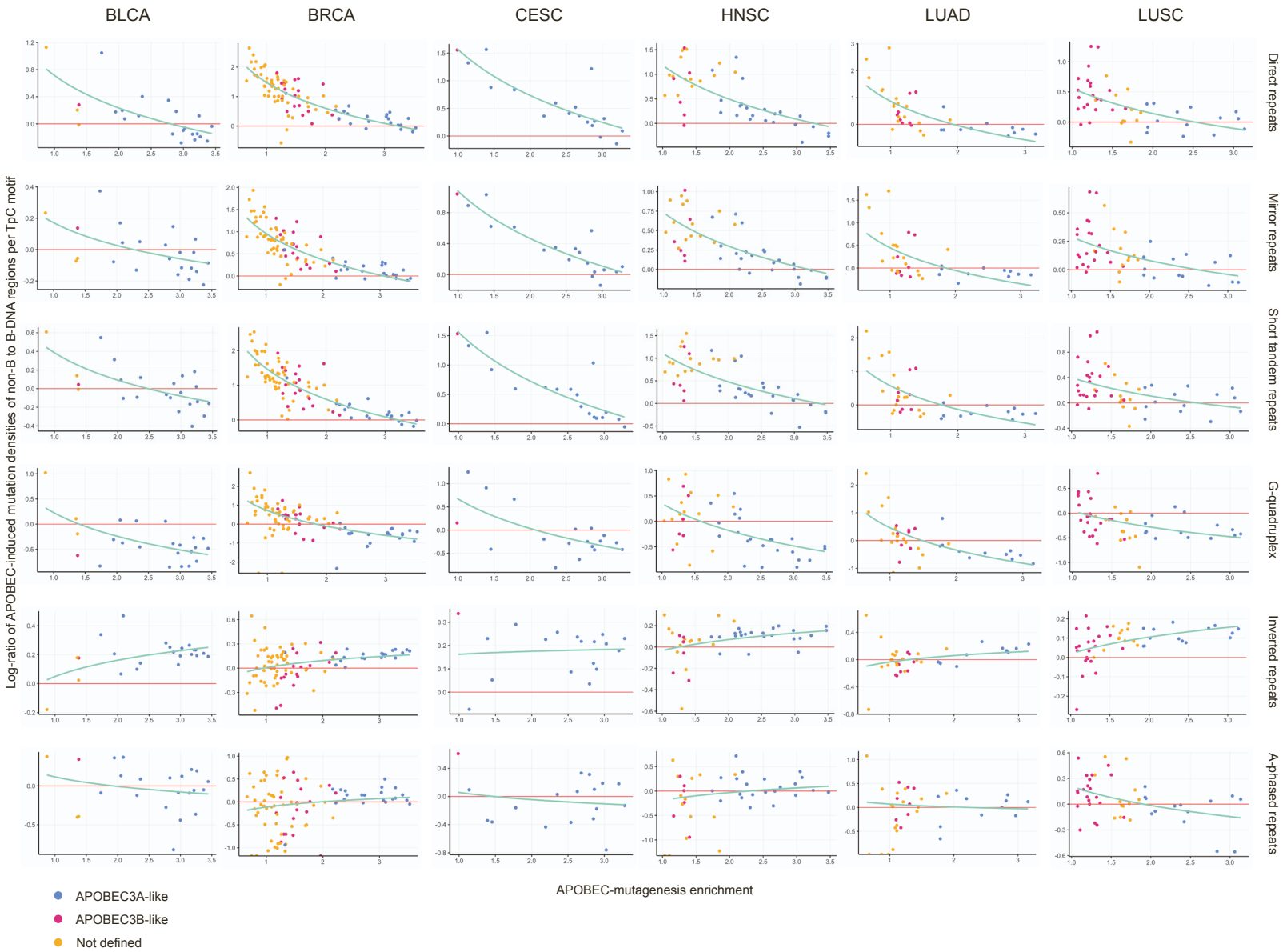


Figure S11. Dependence of the log-ratio of APOBEC-induced mutation densities in non-B DNA to B-DNA genome regions on the activity of APOBEC mutagenesis in cancer samples, Related to Figure 2.
 Color code of samples reflects the prevalence of APOBEC3A-like or APOBEC3B-like mutations.



Figure S12. Comparison of the AID-signature mutation density in non-B and B-DNA genome regions for each sequence component of the WRCY motif, Related to STAR Methods.

An increased AID-signature mutation density can be observed in the direct repeats for the whole WRCY motif and the short tandem repeats for the TACY motif. Statistical significance of the difference in the mutation density between non-B and B-DNA genome regions across cancer types was estimated using the paired Wilcoxon signed-rank test.

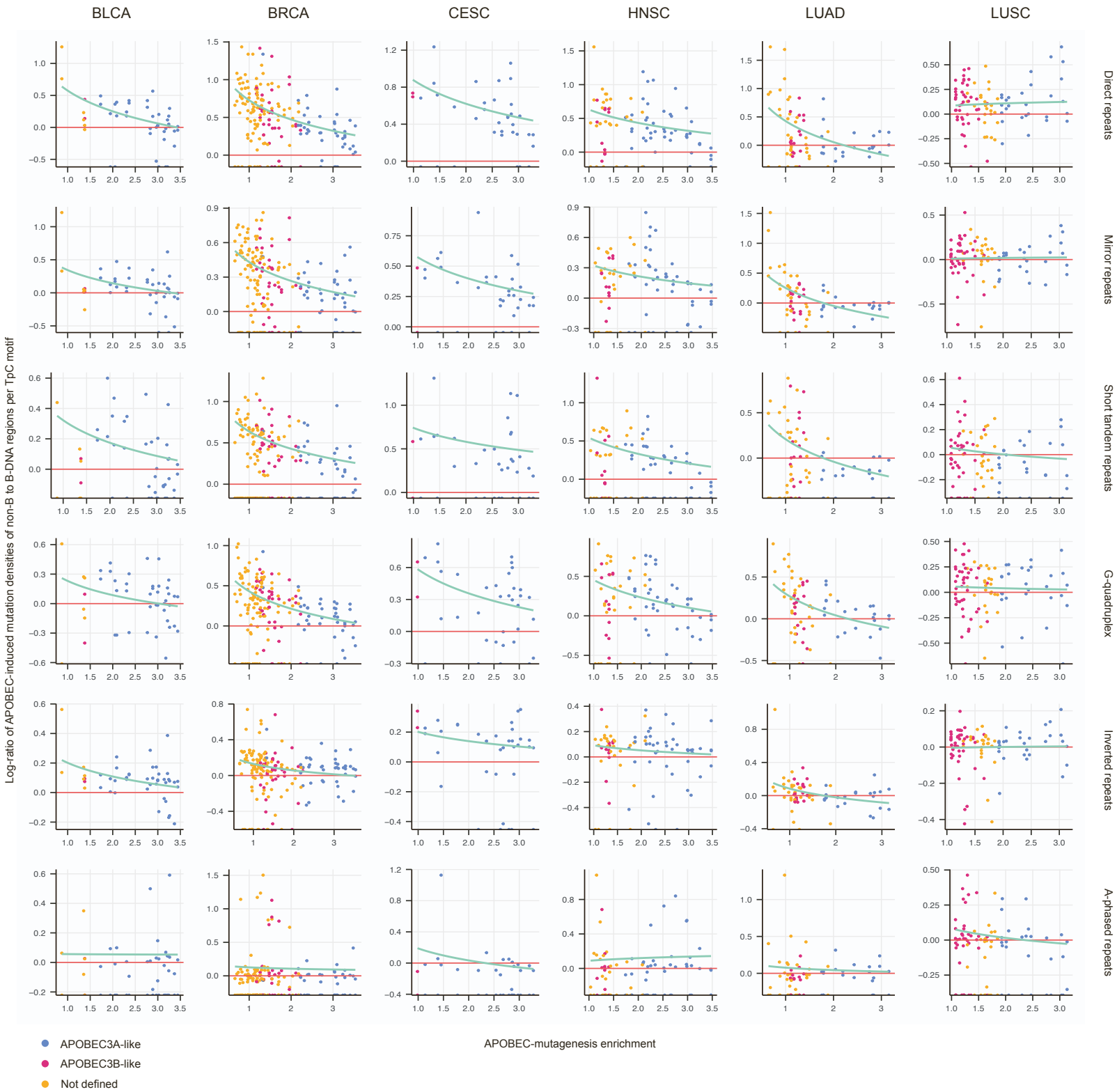


Figure S13. Dependence of the log-ratio of presumed APOBEC3A mutation densities (C->G and C->T substitutions in the CCCA motif) in non-B DNA to B-DNA genome regions on the activity of APOBEC mutagenesis in cancer samples, Related to STAR Methods.

Color code of samples reflects the prevalence of APOBEC3A-like or APOBEC3B-like mutations.

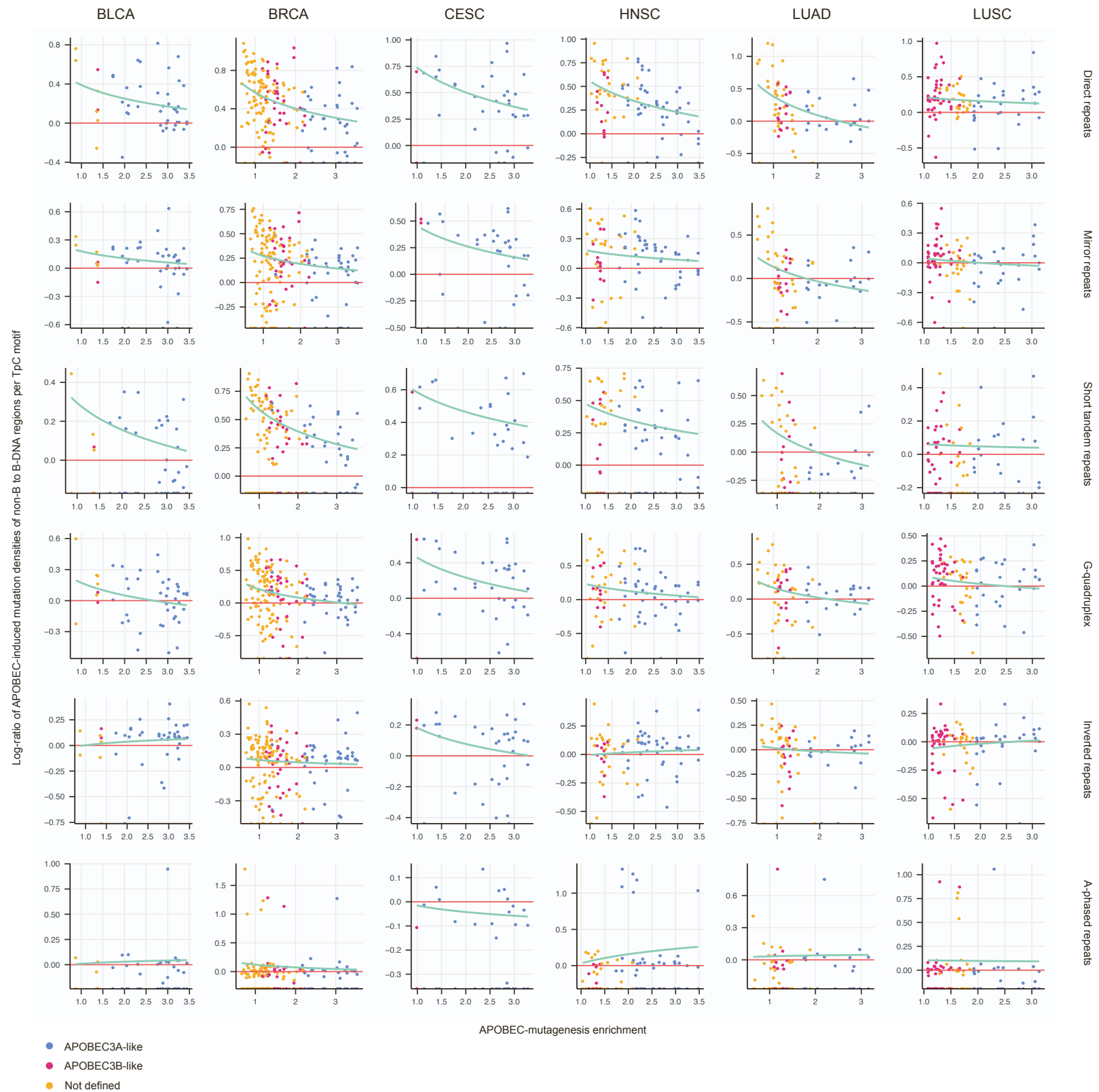


Figure S14. Dependence of the log-ratio of presumed APOBEC3A mutation densities (C→G and C→T substitutions in the CCCG motif) in non-B DNA to B-DNA genome regions on the activity of APOBEC mutagenesis in cancer samples, Related to STAR Methods.

Color code of samples reflects the prevalence of APOBEC3A-like or APOBEC3B-like mutations.