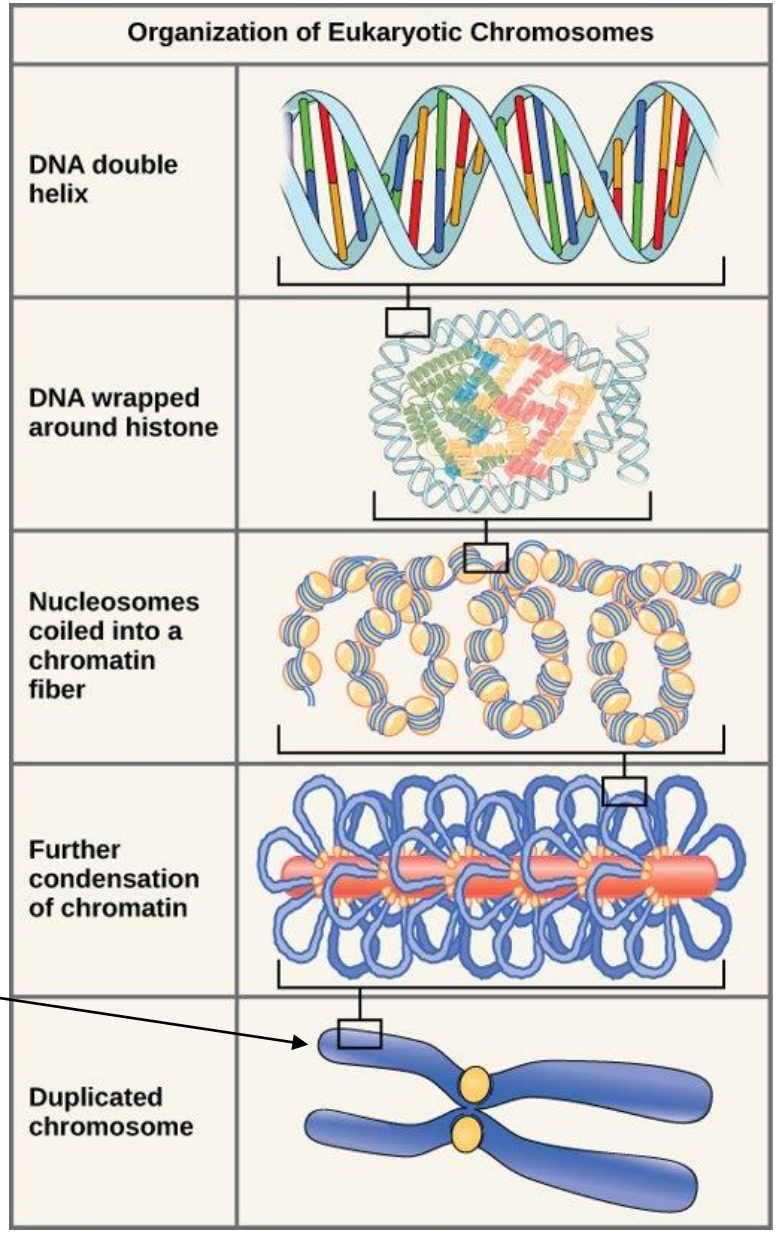
Disruption of chromatin folding domains in human cancer

TYLER CAMP

2020/03/10

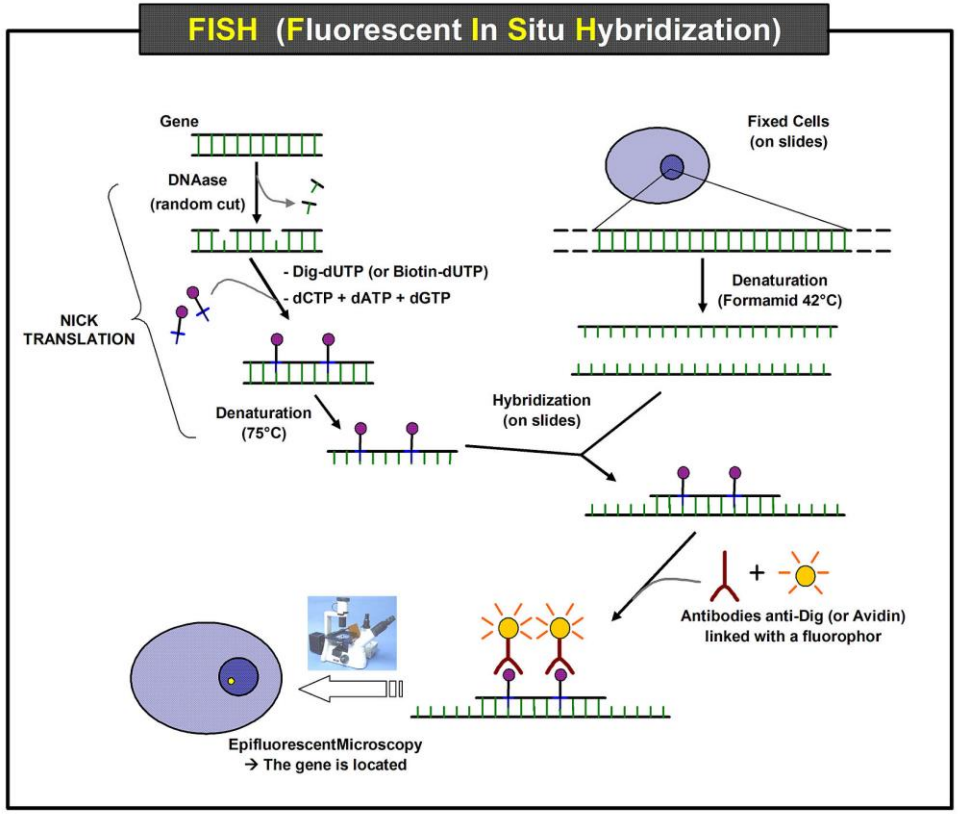
Eukaryotic DNA Organization



These molecules only visible after replication

https://courses.lumenlearning.com/wm-biology1/chapter/reading-dna-packaging-in-eukaryotes-and-prokaryotes/

Identifying DNA-DNA Interactions



https://en.wikipedia.org/wiki/Fluorescence_in_situ_hybridization

Chromosome Conformational Capture

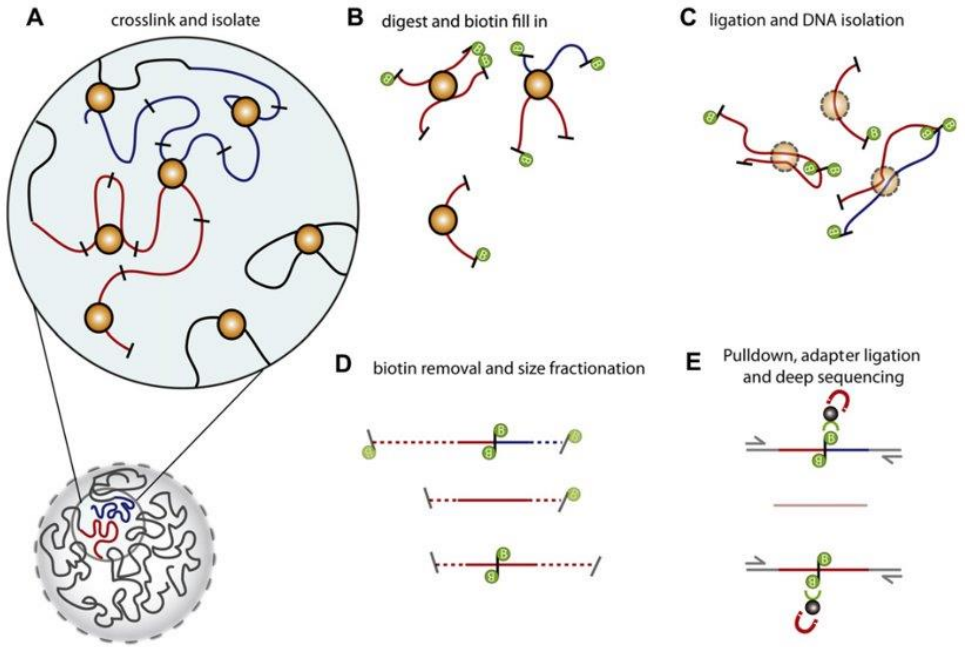
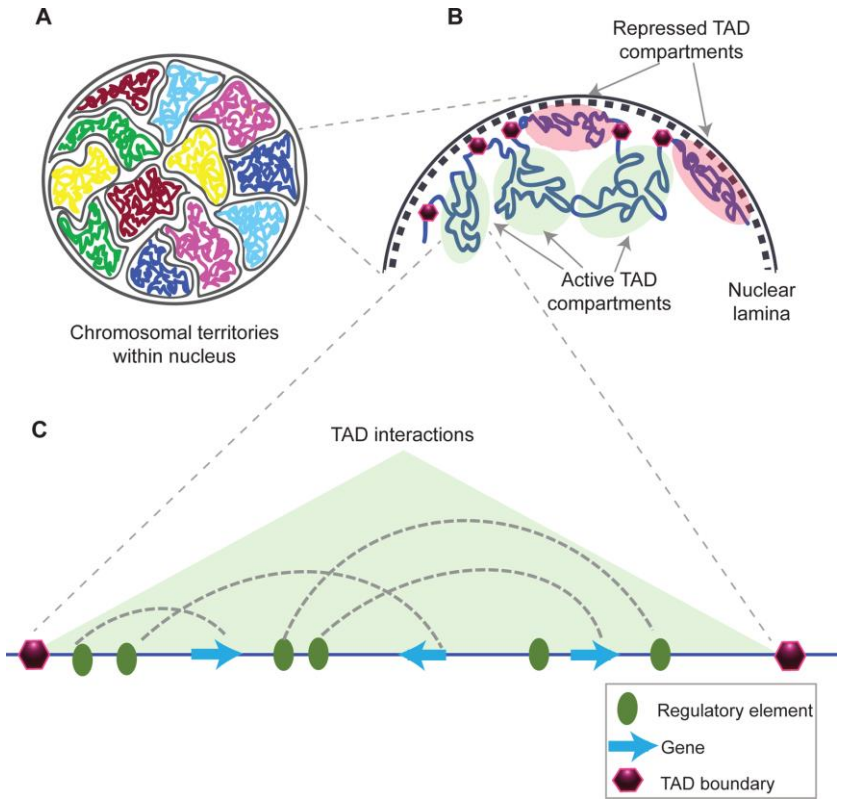


Fig. 1. Overview of Hi–C technology. (A) Hi–C detects chromatin interactions both within and between chromosomes by covalently crosslinking protein/DNA complexes with formaldehyde. (B) The chromatin is digested with a restriction enzyme and the ends are marked with a biotinylated nucleotide. (C) The DNA in the crosslinked complexes is ligated to form chimeric DNA molecules. (D) Biotin is removed from unligated ends of linear fragments and the molecules are fragmented to reduce their overall size. (E) Molecules with internal biotin incorporation are pulled down with streptavidin coated magnetic beads and modified for deep sequencing. Quantitation of chromatin interactions is achieved through massively parallel deep sequencing.

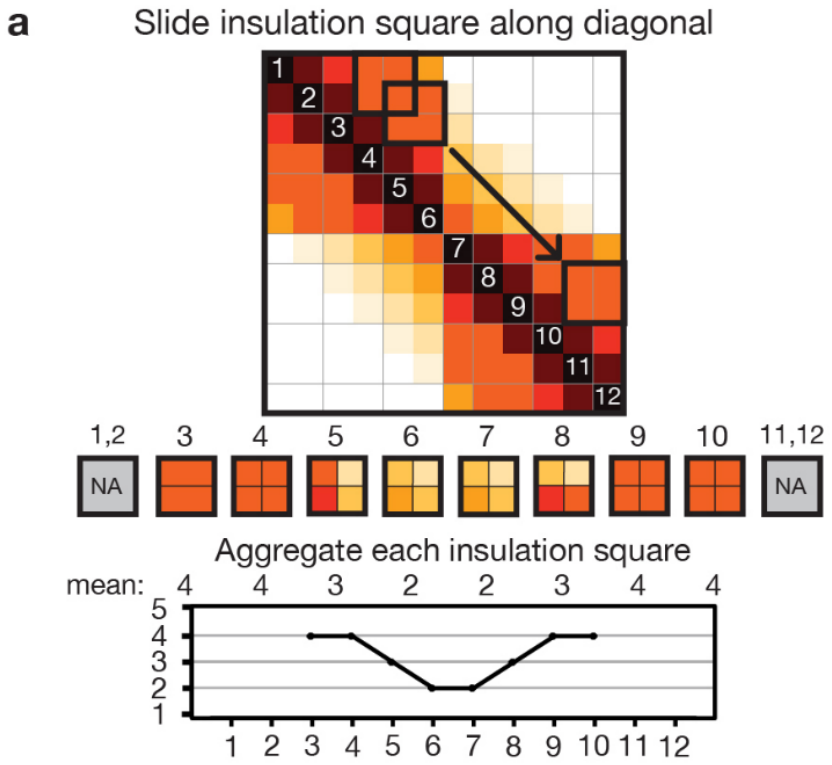
Belton, J. M., McCord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y., & Dekker, J. (2012). Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*, *58*(3), 268–276. https://doi.org/10.1016/j.ymeth.2012.05.001

Topologically Associated Domains (TADs)



https://en.wikipedia.org/wiki/Topologically_associating_domain

Identifying TADs



Extended Data Figure 2 | Insulation profile calculation parameters and boundary calling. **a**, Cartoon shows approach for calculating the insulation profile. A square is slid along each diagonal bin of the interaction matrix to aggregate the amount of interactions that occur across each bin (up to a specified distance upstream and downstream of the bin). Bins with a high insulation effect (for example, at a TAD boundary) have a low insulation score (as measured by the insulation square). Bins with low insulation or boundary activity (for example, in the middle of a TAD) have a high insulation score. Minima along the insulation profile are potential TAD boundaries. **b**, **c**, Heatmaps of chromosome X and

Crane, E., Bian, Q., McCord, R. P., Lajoie, B. R., Wheeler, B. S., Ralston, E. J., ... Meyer, B. J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*, *523*(7559), 240–244. https://doi.org/10.1038/nature14450

TAD Boundaries in Cancer

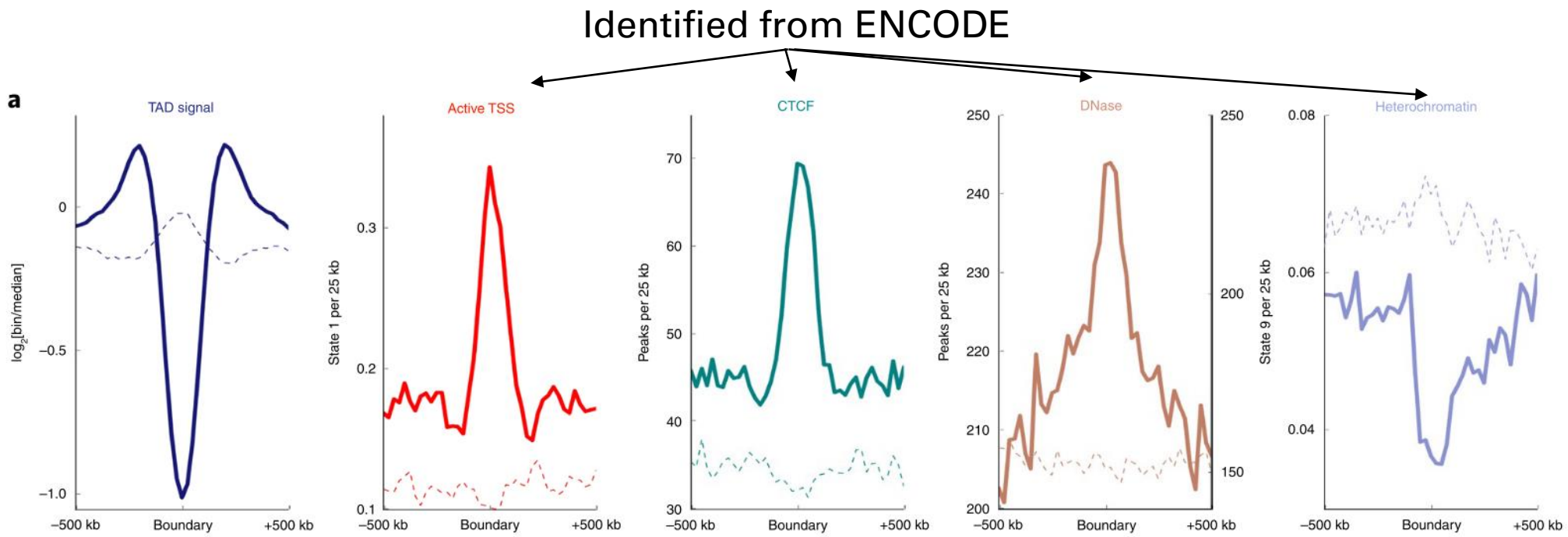
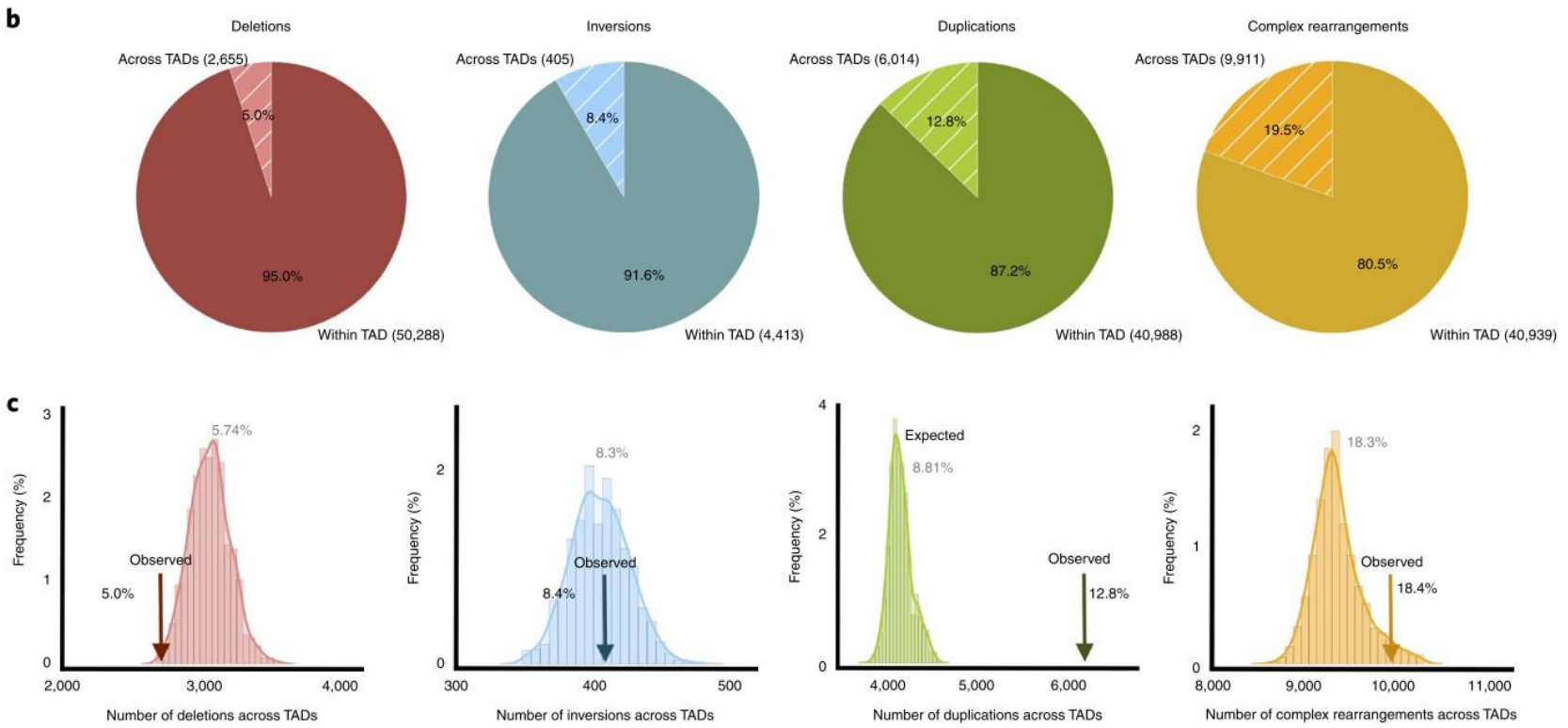


Fig. 1 | TAD boundaries are affected by different types of somatic SVs in cancer genomes. a, Profiles of TAD signals, active transcription start sites (TSS), CTCF peaks, DNase I hypersensitivity and heterochromatic states around common TAD boundaries. Dashed lines represent enrichment levels for the shuffled boundaries. **b**, The percentage of short-range SVs (length ≤ 2 Mb) across TADs (shaded) and within TADs (solid) for different SV types. SVs that

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

TAD Boundaries in Cancer



shuffled boundaries. **b**, The percentage of short-range SVs (length ≤ 2 Mb) across TADs (shaded) and within TADs (solid) for different SV types. SVs that occur across TADs are referred to as BA-SVs. **c**, Observed (arrows) and expected distribution (histograms) of BA-SVs. The expected distribution is based on randomly shuffled boundary data. **d**, Number of affected boundaries (*x* axis) per short-range SV length cut-off (*y* axis). The size of the circles indicates

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

Boundary Altering Structural Variants

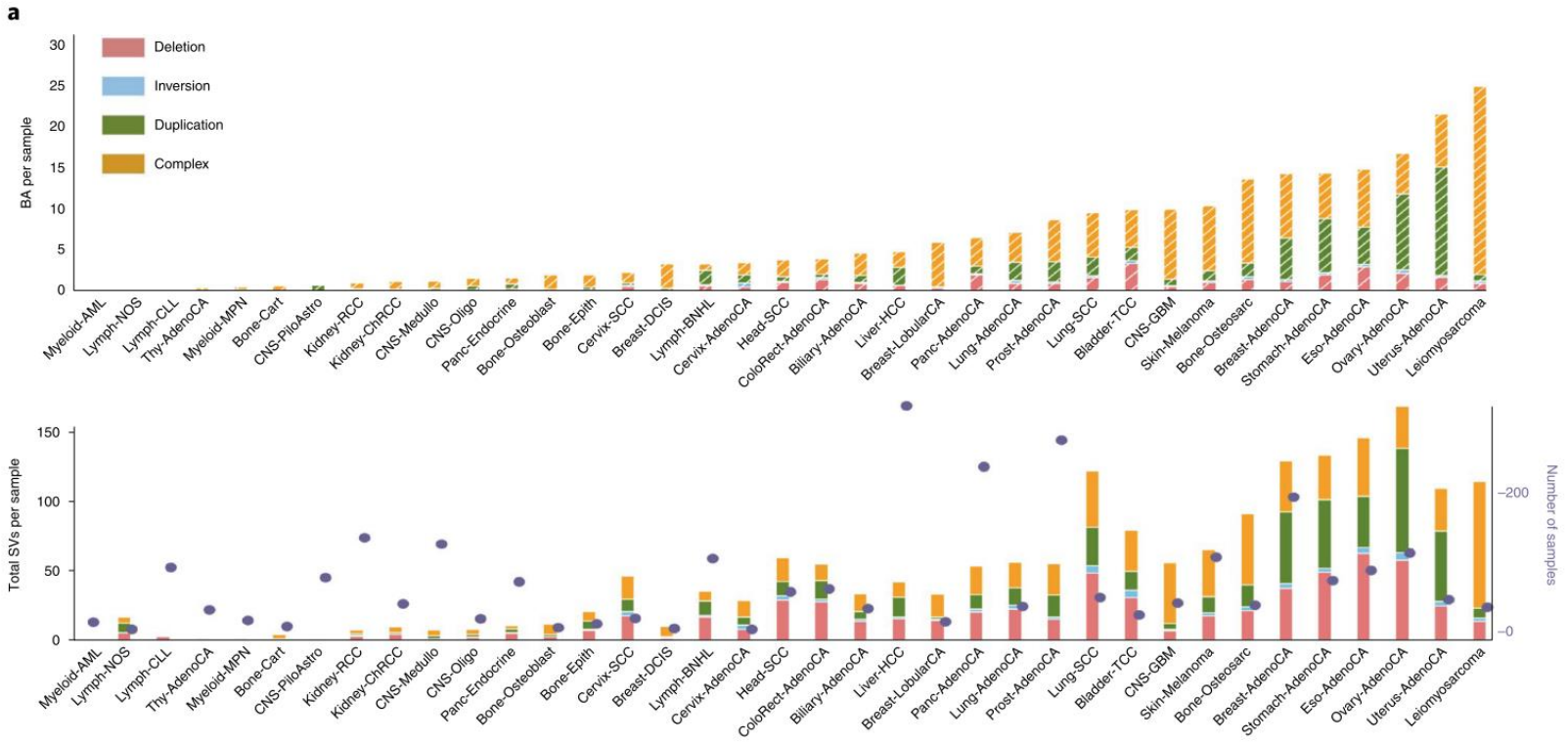


Fig. 2 | Chromatin folding disruptions are specific to histological subtypes. a, Top, the distribution of the average number of BA-SVs per sample for each histological type[22]. **Bottom**, The distribution of the average number of SVs observed in each histological type. Purple dots represent patient numbers for each histological type. Deletions, inversions, tandem duplications and complex rearrangements are shown in red, cyan, green and orange, respectively.

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

Boundaries and Driver Genes

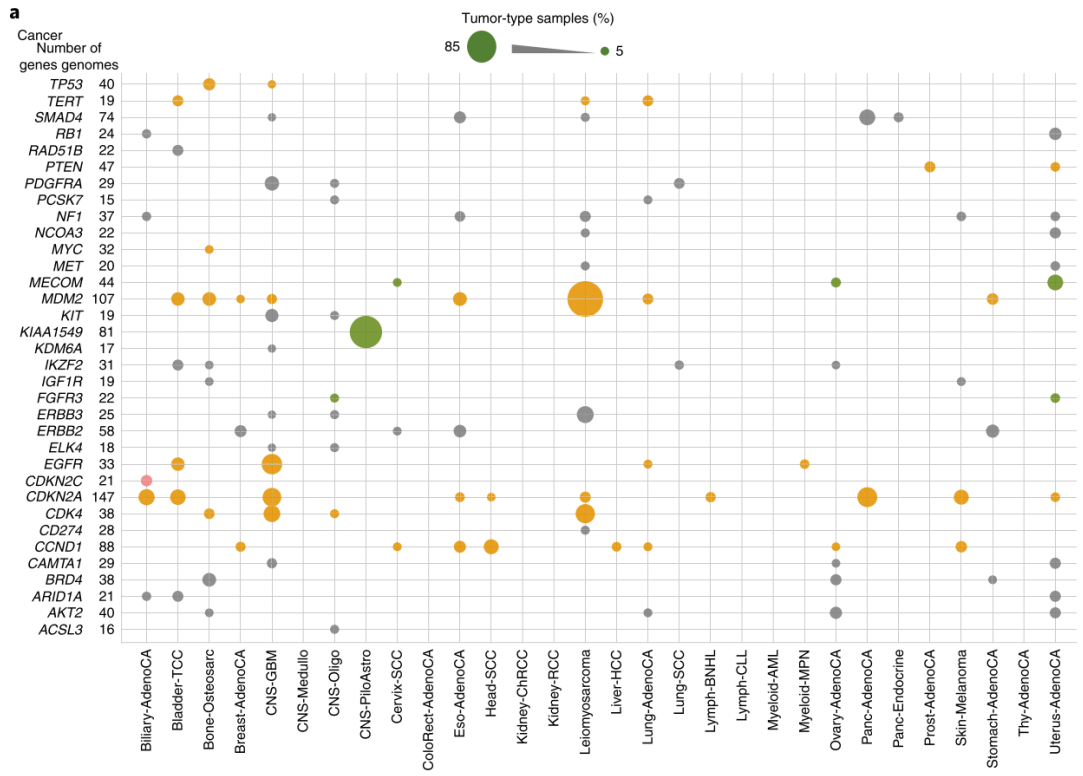


Fig. 3 | Recurrently affected boundaries in specific cancer types. a, Recurrently affected boundaries near known cancer driver genes (*y* axis) per histological type (*x* axis). The size of the circles indicates the portion of samples harboring a BA-SV event in a specific histological type. The color of the circles demonstrates the most common SV type (red, deletion; orange, complex; green, duplication; gray, different SV types were observed) in each histological type. **b**, Top, recurrently duplicated TAD boundaries in pilocytic astrocytoma. Columns of the heat map are TAD boundaries and rows

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

TAD Classification

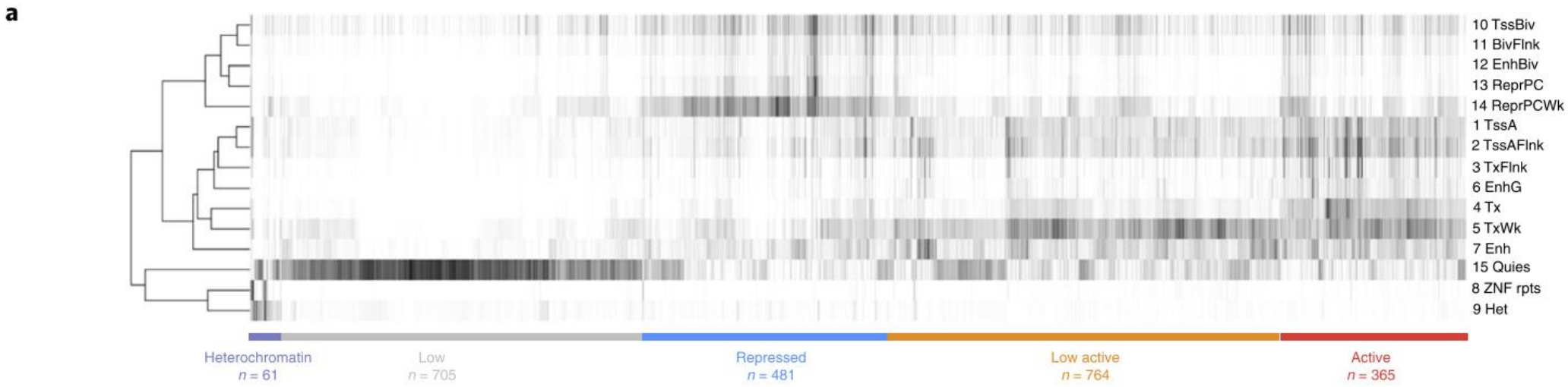
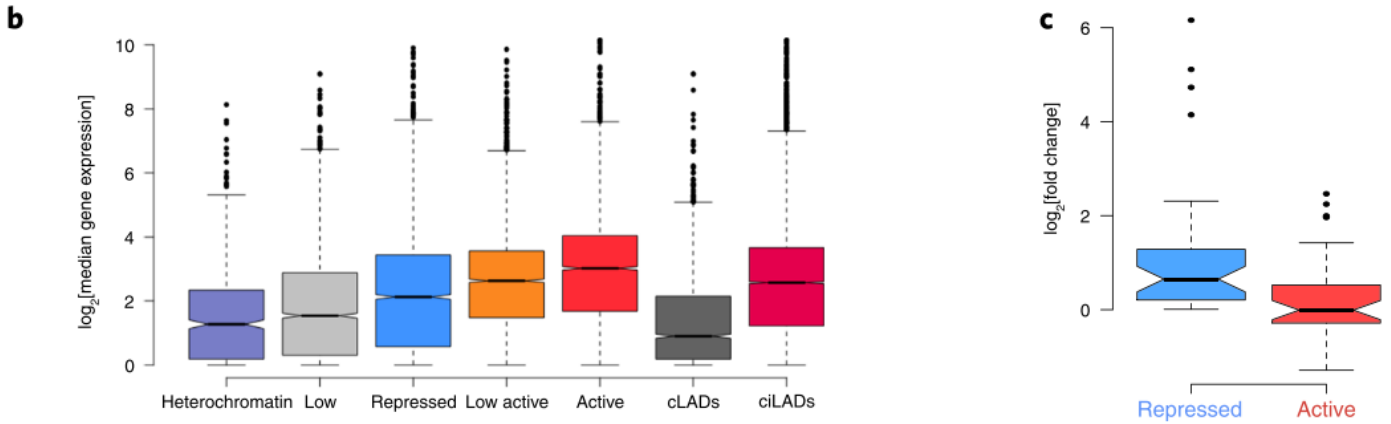


Fig. 4 | Most domain disruptions do not result in marked gene-expression changes. a, Classification of TADs based on chromatin state coverage. The heat map shows domain-length normalized coverage of each chromatin state (rows) in each domain (columns). Domains are classified into five groups according to chromatin state combinations: heterochromatin (purple), low/quiescent (gray), repressed (blue), low active (orange) and active (red).

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

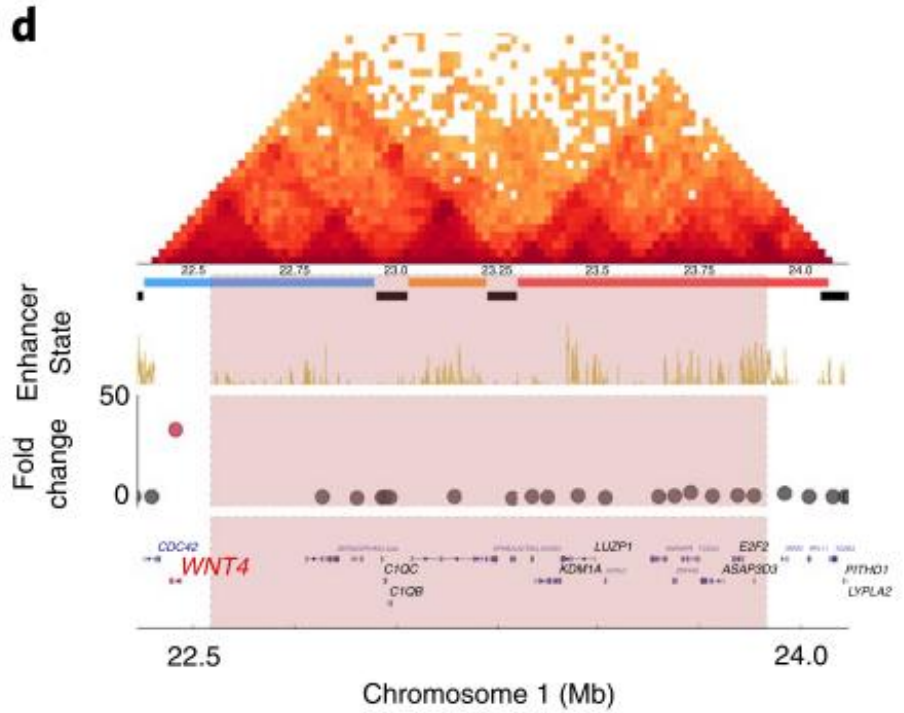
TADs and Gene Expression



b, Median expression levels (log_2) in all PCAWG samples are shown for genes located in each TAD, constitutive LADs (cLADs) and inter-LADs (ciLADs). The number of genes in each annotation group: heterochromatin, 624; low, 2,874; repressed, 3,690; low active, 4,319; active: 4,578; constitutive LAD, 2,384; constitutive inter-LAD, 14,430. In these and all other box plots, the center line is the median; box limits are the upper and lower quantiles; whiskers represent 1.5× the interquartile range. **c**, The log_2-transformed fold change in expression is shown for genes that are nearest to BA-deletion break-ends between repressed and active TADs ($n = 43$). **d**–**f**, Examples of BA-SV-harboring samples. Triangle heat maps represent chromatin contact frequency

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

TADs and Gene Expression



between repressed and active TADs ($n = 43$). **d**–**f**, Examples of BA-SV-harboring samples. Triangle heat maps represent chromatin contact frequency (log_2) in NHEK cells. BA-SV regions are shaded. Colored tiles represent domains and black bars denote TAD boundaries. Roadmap Epigenome enhancer-state frequencies are shown as a yellow histogram. Dots show fold changes in expression in a lymphoma sample harboring a BA-deletion near *WNT4* (**d**), a breast adenocarcinoma sample harboring a BA-deletion near *SLC22A2* (**e**) and an ovarian adenocarcinoma sample harboring a BA-deletion near *SLC2A10*

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y

Chromothripsis

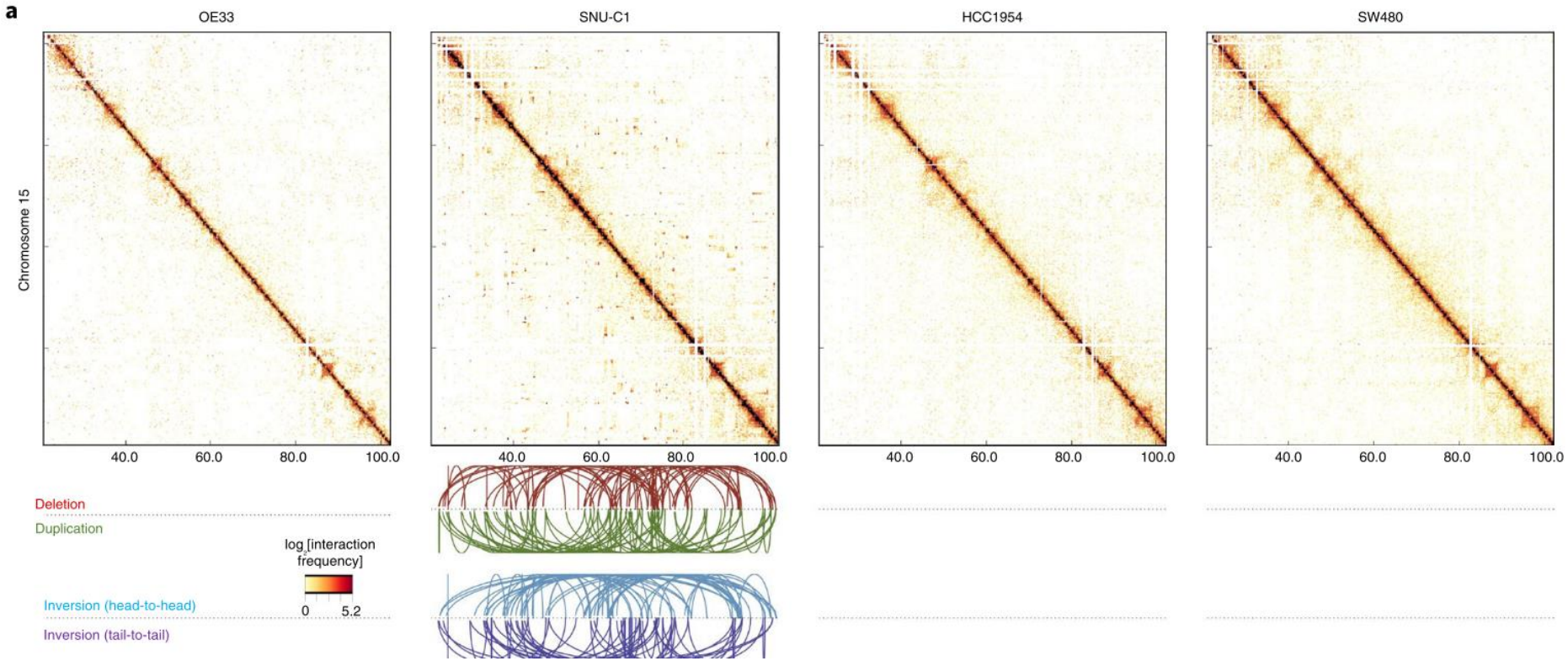


Fig. 6 | Complex rearrangements markedly change chromatin folding maps in the cancer genomes. a–d, The effects of complex rearrangements on chromatin folding domains. Contact frequencies (log_2) of each cell type, plotted with a 40-kb window size. Bottom arcs represent SV breakpoint locations with rearrangements coded by color. Green, tandem duplication; red, deletion; cyan and purple, inversion. **a**, SNU-C1 cells harbor a chromothripsis event that affects chromosome 15. **b**, HCC1954 cells contain a complex rearrangement on chromosome 21. **c**, The *MYC* locus contains regional complex

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., ... Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*(3), 294–305. https://doi.org/10.1038/s41588-019-0564-y