

# Transcriptional programs of cell identity and p53-induced stress responses are associated with distinctive features of spatial genome organization

Gony Shanel<sup>1,†</sup>, Tsung-Han S. Hsieh<sup>2,†</sup>, Claudia Cattoglio<sup>2,3</sup>, Hadar Amira Haham<sup>4</sup>, Hsin-Jung Chou<sup>2</sup>, Jack Z. Li<sup>2,5,6</sup>, Ron Shamir<sup>4</sup>, Xavier Darzacq<sup>2</sup>, Ran Elkon<sup>1,\*</sup>

<sup>1</sup>Department of Human Molecular Genetics and Biochemistry, Gray Faculty of Medical & Health Sciences, Tel Aviv University, Tel Aviv 69978, Israel

<sup>2</sup>Department of Molecular and Cell Biology, Li Ka Shing Center for Biomedical and Health Sciences, CIRM Center of Excellence, University of California, Berkeley 94720, United States

<sup>3</sup>Howard Hughes Medical Institute, Berkeley 94720, United States

<sup>4</sup>The Blavatnik School of Computer Science and Artificial Intelligence, Tel Aviv University, Tel Aviv 69978, Israel

<sup>5</sup>Department of Dermatology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, United States

<sup>6</sup>Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, United States

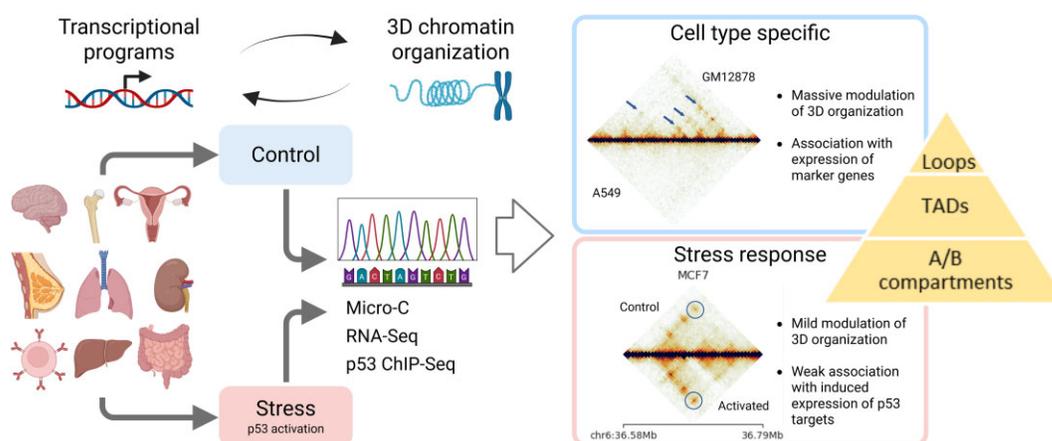
\*To whom correspondence should be addressed. Email: ranel@tauex.tau.ac.il

†The first two authors should be regarded as Joint First Authors.

## Abstract

To date, most studies explored changes in 3D-genome organization between different tissues or during differentiation, which involve massive reprogramming of transcriptional programs. Much fewer studies examined alterations in genome organization in response to cellular stress, which involves less pervasive transcriptional modulation. Here, we examined associations between spatial chromatin organization and gene expression in two different biological contexts: transcriptional programs determining cell identity and transcriptional responses to stress, using p53 activation as a model. We selected 10 cell lines of diverse tissues, and in each performed micro-C, RNA-seq, and p53 ChIP-seq, before and after p53 induction. In the comparison between cell types, we delineated marked correlations between gene expression and spatial genome organization and identified hundreds of active enhancer–promoter loops associated with the expression of cell-type marker genes. In contrast, within each cell type, no such links were observed for expression changes induced by p53 activation, even for enhancers and promoters activated by p53 binding. Our analysis points to a fundamental difference between chromatin interactions that define cell identity and those that are established in response to cellular stress. Our results on p53-induced transcriptional responses support the recently proposed TF activity gradient model, which speculated a contact-independent mechanism for enhancer–promoter communication.

## Graphical abstract



Received: January 10, 2025. Revised: May 14, 2025. Editorial Decision: June 3, 2025. Accepted: June 13, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

## Introduction

The relationship between genome structure and function is a longstanding question [1]. Transcriptional programs that determine cell fate or cellular responses to environmental stresses are controlled by hundreds of thousands regulatory elements embedded in our genome, which exert their functions in the context of the dynamic 3D chromatin organization [2–4]. The development and refinement of the Hi-C technique [5–7] and its enhanced derivatives, like the micro-C method [8, 9], which measure interaction frequencies between any two mappable segments in the genome, revolutionized our understanding of chromatin spatial organization in the nucleus and its association with transcriptional regulation. Yet, despite recent advancements in the field, key aspects of cell-type specific finer-scale loops and chromatin structures induced by transient stress responses remain unsolved.

A three-layer 3D organization hierarchy has emerged from Hi-C studies [10]: at the top of the hierarchy, the genome is spatially divided into two major compartments referred to as A and B compartments, which roughly correspond to active and inactive transcriptional regions, respectively [11]. Markedly, a study that explored the 3D genome organization in 21 primary human tissues and cell types observed a substantial A/B compartment switching across tissues, finding that ~60% of the genome showed a change in A/B compartmentalization over the analyzed tissue panel [12].

In the next organization layer are the topologically associating domains (TADs), defined by the preferential interaction of chromatin segments located within a domain and relative depletion of interactions between genomic segments located in different domains. TAD boundaries are significantly enriched with CTCF-binding sites, mostly set in a convergent orientation [13, 14]. TADs are relatively conserved across different tissues [12]. Many studies applying perturbations of TAD boundary elements indicated that TADs form insulated domains for transcriptional regulation, which determine specificity of enhancer–promoter (E–P) interactions by considerably confining them to contacts between elements residing within the same TAD [15–18]. However, overall, the insulation imposed by TADs is generally low, as intra-TAD interaction frequency is only ~2-fold higher than interactions between distance-matched inter-TAD loci [19]. Yet, even mild TAD insulation can have a marked regulatory impact, as the disruption of a weak TAD boundary can result in a substantial (>10-fold) difference in gene expression [20].

However, interestingly, a severe loss of TAD structures across the genome, induced by acute degradation of the architectural proteins CTCF or cohesin, resulted in an unexpectedly limited effect on gene expression [21]. Therefore, it seems that TAD boundaries with very strong roles in gene regulation are not common or that their regulatory effect may be highly context dependent [22, 23]. Moreover, both recent single-cell Hi-C studies and chromatin tracing experiments coupled with super-resolution microscopy show that intra-TAD contacts vary significantly from cell to cell and that the globular structure of TADs only emerges when the entire cell ensemble is considered collectively [24, 25]. Thus, the perspective that TADs represent population-level statistical patterns of dynamic chromatin polymer movements that mainly occur within confined domains is gaining increasing support [26–28].

As for the third layer in the hierarchy of the 3D genome organization, TADs are often subdivided into smaller, nested micro-compartments [29]. Some of the strongest intra-TAD contacts are between active enhancers and promoters (referred to as “E–P loops”). Many such E–P loops are formed during cell differentiation and correlate with the activation of cell-type specific genes [4, 30, 31]. Imaging studies showed that these contacts are dynamic and transient rather than stable structures and that even for the strongest E–P loops detected by Hi-C, the enhancers and promoters were in spatial proximity in only 10%–30% of cells [27, 32, 33]. These observations bring into focus key open questions, which require a higher-resolution 3D genome mapping, including whether E–P proximity is a general rule for gene activation and if such proximity is a prerequisite for transcription activation of cell-identity or stress-induced genes.

Many studies explored changes in 3D genome organization between different tissues or during differentiation, which involves massive reprogramming of cellular transcriptional programs [34–37]. Much fewer studies examined alterations in spatial genome organization in response to cellular stress, which involves less pervasive and more transient transcriptional modulation [38, 39]. In this study, we used a panel of 10 cell lines to examine and contrast associations between spatial chromatin organization and gene expression in transcriptional programs that define cell identity on one hand and transcriptional responses to p53 activation (by Nutlin-3a treatment [40]), which we used as a model for stress responses, on the other.

## Materials and methods

### Micro-C experiments and data generation

Cells were treated with 20  $\mu$ M Nutlin-3a (dissolved in ethanol) or the same volume of pure ethanol (mock treatment) for 4 h. Cells were cross-linked for 10 min at room temperature with 1% formaldehyde-containing medium; cross-linking was stopped by Tris-glycine. Cells were washed twice with PBS and cross-linked for 45 min at room temperature with 3 mM disuccinimidyl glutarate (DSG) in phosphate-buffered saline (PBS). Cross-linking was halted by Tris-glycine, and cells were washed twice in PBS. Micro-C protocol was performed as the following steps: (i) digest cross-linked chromatin by MNase; (ii) repair fragment ends with biotin-dNTP; (iii) proximal ligation and purge unligated ends; and (iv) purify ligated dinucleosomal DNA. Nucleosomal fragment size distribution at various steps is shown in [Supplementary Fig. S1](#). Purified DNA with biotin-dNTPs was captured by Dynabeads® MyOne™ Streptavidin C1 [41]. Micro-C libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB E7645) according to manufacturer instructions with a few modifications. The sequencing library was amplified by Kapa HiFi PCR enzyme with the lowest possible cycles to reduce polymerase chain reaction (PCR) duplicates. Library concentration, quality, and fragment size were assessed by Qubit fluorometric quantification (Qubit™ dsDNA HS Assay Kit, Invitrogen™ Q32851), quantitative polymerase chain reaction (qPCR), and Fragment Analyzer™. Twenty multiplexed libraries were pooled and sequenced in six lanes on the Illumina NovaSeq sequencing platform (100 bp, paired-end reads).

### Micro-C data analysis

We used HiC-Pro (2.10 version) [42] to generate allValidPairs from 40 libraries total, two biological replicates per condition. mcool files were converted from allValidPairs by the Cooler package [43]. Mapping and QC statistics are provided in [Supplementary Table S1](#). All analyses were conducted using human genome version hg38.

### A/B compartments correlation with gene expression

Principal Component Analysis (PCA) was applied to the contact matrix of each sample (after merging duplicates) using cooltools [44] with 100 kb resolution contact maps. Genes were assigned to the A/B compartment based on the sign of PC1 at the location of the gene's transcription start site (TSS). The intersection between compartments and TSS locations was done using bedtools [45]. Gene expression level differed significantly between the two compartments. The "A" label was assigned to the compartment with higher levels of gene expression.

### TADs analysis

TADs were called in each sample using the arrowhead algorithm implemented in the Juicer package [46] with 10 kb resolution contact maps and default parameters. TADs longer than 1 Mb were filtered out.

### Chromatin loops

Chromatin loops were identified using Mustache [47] with 5 kb resolution contact maps and using False Discovery Rate (FDR) < 0.1. To create a pooled set of all the loops detected in all samples, we merged individual loop lists using pgltools [48]. Loop interaction frequency ("loop intensity") was extracted from .cool files using cooler [43]. Finally, pooled loop intensity data were normalized across all samples using quantile normalization.

### Promoter interactions

Promoter loops (P loops) were identified by intersecting chromatin loops and genes' TSS coordinates after adding  $\pm 2.5$  kb flanks around the TSSs.

### RNA-seq experiments and data generation

Cells were treated with 20  $\mu$ M Nutlin-3a (dissolved in ethanol) or the same volume of pure ethanol (mock treatment) for 4 h. Total RNA was extracted with TRIzol reagent followed by ribosomal RNA (rRNA) depletion [NEBNext® rRNA Depletion Kit (Human/Mouse/Rat) with RNA Sample Purification Beads; cat. E6350L]. RNA-seq libraries were prepared with the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (cat. E7760S). Library concentration, quality, and fragment size were assessed by Qubit fluorometric quantification (Qubit™ dsDNA HS Assay Kit, Invitrogen™ Q32851), qPCR, and Fragment Analyzer™. Ten multiplexed libraries were pooled and sequenced in one lane on the Illumina HiSeq4000 sequencing platform (100 bp, paired-end reads). RNA-seq reads (40 libraries total, two biological replicates per condition) were quantified with Kallisto [49]. Mapping statistics are provided in [Supplementary Table S1](#). Differentially expressed genes (between different cell

lines and within a cell line in response to Nutlin-3a treatment) were called using DESeq2 [50].

### Removal of copy number aberration effects

Cancer cell lines are characterized by high level of genomic aberration (large deletion and amplification events) leading to changes in copy number of chromosomal segments (CNV). CNV may affect both interaction frequencies measured by micro-C and gene expression data measured by RNA-seq. To inspect the effect of CNVs on our data, we used the CaSpER tool [51], which identifies gross chromosomal deletions and amplifications from RNA-seq data. CaSpER divides the genome in each cell line into five CNV states: (i) homozygotic deletion, (ii) heterozygotic deletion, (iii) diploid, (iv) triploid amplification, and (v) amplification. We then assigned each gene to a CNV state according to the state at the location of its TSS. As expected, in all cell lines, CNV state was significantly correlated with raw interaction frequency and gene expression levels ([Supplementary Fig. S2](#)). ICE balancing [52] effectively removed CNV impact on interaction frequency ([Supplementary Fig. S2A](#)). As for gene expression, to remove CNV effects, we fitted, for each cell line, a linear model (expression  $\sim$  CNV state) ([Supplementary Fig. S2B](#)), and in subsequent analyses, used the residual expression levels, which represent the component that is independent of CNV.

### p53 ChIP-seq experiments and data generation

ChIP was performed as described with few modifications [53]. Cells were treated with 20  $\mu$ M Nutlin-3a (dissolved in ethanol) or the same volume of pure ethanol (mock treatment) for 4 h and cross-linked for 6 min at room temperature with 1% formaldehyde-containing, serum-free medium. Cross-linking was stopped by PBS-glycine (0.125 M final). Cells were washed twice with ice-cold PBS, scraped, centrifuged for 10 min, and pellets were flash-frozen. Cell pellets were thawed and resuspended in cell lysis buffer (5 mM PIPES, pH 8.0, 85 mM KCl, and 0.5% NP-40, 1 ml/15 cm plate) with protease inhibitors and incubated for 10 min on ice. Lysates were centrifuged for 10 min at 4000 rpm, and nuclear pellets resuspended in six volumes of sonication buffer [50 mM Tris-HCl, pH 8.1, 10 mM ethylenediaminetetraacetic acid (EDTA), 0.1% sodium dodecyl sulphate (SDS)] with protease inhibitors, incubated on ice for 10 min, and sonicated to obtain DNA fragments below 2000 bp in length (Covaris S220 sonicator, 20% Duty factor, 200 cycles/burst, 150 peak incident power, 7–16 cycles 30 s on and 30 s off). Sonicated lysates were cleared by centrifugation, and chromatin (100–800  $\mu$ g per antibody) was diluted in RIPA buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl) with protease inhibitors to a final concentration of 0.8  $\mu$ g/ $\mu$ l, precleared with Protein G sepharose (GE Healthcare) for 2 h at 4°C, and immunoprecipitated overnight with 1  $\mu$ g of anti-p53 antibody (Cell Signaling, #9282) per 100  $\mu$ g of chromatin. About 4% of the precleared chromatin was saved as input. Immunoprecipitated DNA was purified with the Qiagen QIAquick PCR Purification Kit and eluted in 45  $\mu$ l of 0.1 $\times$  TE (1 mM Tris-HCl, pH 8.0, 0.01 mM EDTA). ChIP-seq libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB E7645) according to manufacturer instructions with a few modifications. Twenty nanograms of ChIP input DNA (as measured by Nan-

odrop) and 25  $\mu$ l of the immunoprecipitated DNA were used as a starting material. The recommended reagents' volumes were cut in half. The NEBNext Adaptor for Illumina was diluted 1:5 to 1:10 in Tris/NaCl, pH 8.0 (10 mM Tris-HCl, pH 8.0, 10 mM NaCl), and the ligation step extended to 30 min. After ligation, a single purification step with 0.9 $\times$  volumes of Agencourt AMPure XP PCR purification beads (Beckman Coulter A63880) was performed, eluting DNA in 22  $\mu$ l of 10 mM Tris-HCl (pH 8.0). Twenty microliters of the eluted DNA were used for the library enrichment step, performed with the KAPA HotStart PCR kit (Roche Diagnostics KK2502) in 50  $\mu$ l of total reaction volume (10  $\mu$ l 5 $\times$  KAPA buffer, 1.5  $\mu$ l 10 mM dNTPs, 0.5  $\mu$ l 10 uM NEB Universal PCR primer, 0.5  $\mu$ l 10  $\mu$ M NEB index primer, 1  $\mu$ l KAPA polymerase, 16.5  $\mu$ l nuclease-free water, and 20  $\mu$ l sample). Samples were enriched with 10–12 PCR cycles (98°C, 45 s; [98°C, 15 s; 60°C, 10 s]  $\times$  9; 72°C, 1 min; 4°C, hold), purified with 0.9 volumes of AMPure XP PCR purification beads, and eluted with 33  $\mu$ l of 10 mM Tris-HCl, pH 8.0. Library concentration, quality and fragment size were assessed by Qubit fluorometric quantification (Qubit™ dsDNA HS Assay Kit, Invitrogen™ Q32851) qPCR and Fragment Analyzer™. Ten multiplexed libraries were pooled and sequenced in one lane on the Illumina HiSeq4000 sequencing platform (50 bp, single-end reads).

ChIP-seq raw reads from ethanol- or Nutlin-3a-treated cells (80 libraries total, two biological replicates per condition) were quality-checked with FastQC, trimmed with cutadapt and aligned onto the human genome (hg38 assembly) using bowtie2 [54]. Mapping statistics are provided in [Supplementary Table S1](#). Biological replicates were pooled and p53 binding events (“peaks”) induced by Nutlin-3a were detected using MACS2 [55], comparing p53 IP samples in the treated versus untreated cells. MACS2 bedGraph output files were converted to bigwig files with bedGraphToBigWig [56]. p53 ChIP-seq peaks were tested for transcription factor binding motif enrichment using HOMER (Hypergeometric Optimization of Motif Enrichment) [57].

### Signal of epigenetic markers over genomic intervals

All analyses of signal of epigenetic markers over genomic intervals were done by calculating the mean fold-change-over-control signal over the specified intervals using publicly available data from ENCODE ([Supplementary Table S6](#)) using ENCODE's bigWigAverageOverBed tool [58].

## Results

We selected 10 cell lines of diverse tissue of origin (Fig. 1): GM12878 and IMR90 (normal cell lines); HEK293 (immortalized); and MCF7, HCT116, HepG2, HeLa, A549, SK-N-SH, and U2OS (cancer cell lines). All these cell lines are included in the ENCODE project (and therefore have ample publicly available epigenomic profiles). As a model for stress responses, we focused on p53 activation (using the potent p53 activator Nutlin-3a [40]). All the selected cell lines except two (HeLa and HEK293) have functional wild-type p53. To allow comprehensive and systematic investigation of genome organization in 3D and gene expression under these conditions, for each cell line we performed micro-C, RNA-seq, and p53 ChIP-seq experiments, both in basal conditions and after Nutlin-3a treatment. In the first part of this study, we focus on differences

between the cell lines in basal conditions and analyze correlations between cell-type specific gene expression programs and differential 3D genome organization. In the second part, we turn to analyze alterations in gene expression and 3D genome organization upon induction of p53.

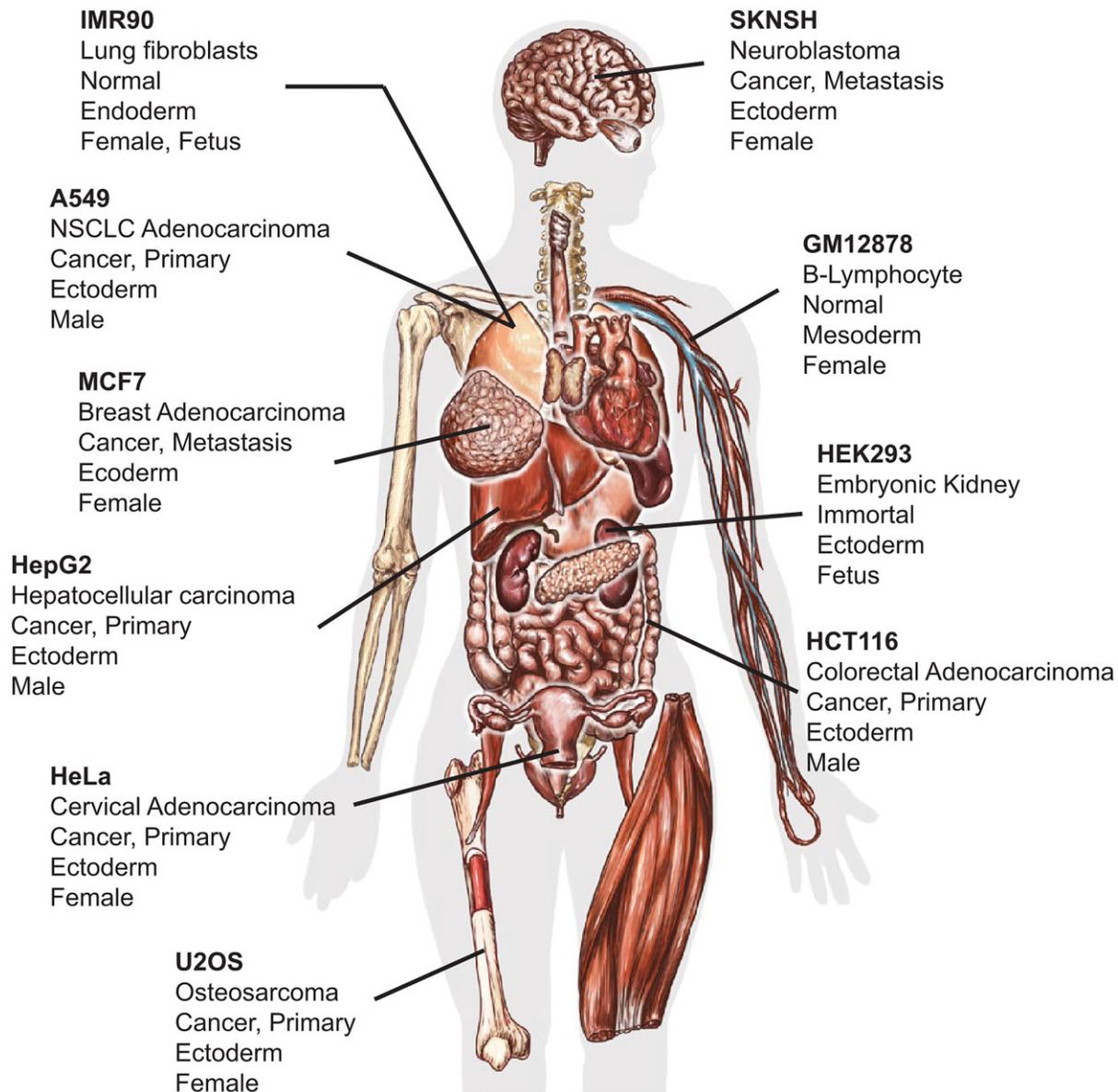
Seven cell lines in our panel are tumor-derived, carrying unbalanced karyotypes with frequent amplifications and deletions. As expected, we observed significant correlations between copy number aberrations (CNAs) and chromatin interactions and gene expression ([Supplementary Fig. S2](#)). As our aim was to analyze cell-type-intrinsic regulatory links between chromatin organization and gene expression (rather than characterizing the impact of chromosomal amplifications and deletions on expression levels), we first aimed to remove the impact of CNA. In line with previous reports [59], we found that balancing the interaction frequency matrix using the iterative correction and eigenvector decomposition algorithm [52] largely cancelled the CNA bias ([Supplementary Fig. S2A](#)). For the RNA-seq data, we applied linear regression to cancel the link between CNA and expression level (see the “Materials and methods” section; [Supplementary Fig. S2B](#)). The subsequent analyses were performed on these normalized datasets.

In the next sections, we present the analyses in accordance with their order in the hierarchy of the spatial organization of the genome: (i) A/B compartments, (ii) TADs, and (iii) intra-TAD loops.

### Association between chromatin compartmentalization and differential gene expression between cell lines

For each micro-C sample, A/B compartments were called using cooltools [44] with a resolution of 100k bp (see the “Materials and methods” section). As expected, in all cell lines there was a marked difference in expression levels between genes assigned to the two compartments, and we labeled by “A” the compartment with the higher expression ([Supplementary Fig. S3A](#)). Next, for each pair of cell lines, we assigned each gene to one of four groups: AA, BB, AB, or BA, according to the gene's compartment in the two cell lines. In all 45 pairwise analyses, compartment switch was strongly associated with a corresponding change in gene expression (Fig. 2A and B). To further quantify the association between A/B compartmentalization and differential expression between cell lines, we calculated, for each pairwise comparison, the percentage of “concordant genes”—that is, genes that showed expression difference in the expected direction: genes switching from A to B compartment between cell lines 1 and 2 have higher expression in cell line 1 than cell line 2, and vice versa. Over the 45 comparisons, the average percentage of such concordant genes was 64% (with SD of 5%) ([Supplementary Fig. S3B](#)).

Following these results, we next tested if compartmentalization can be represented as a continuous feature (using the quantitative value of the eigenvector as a *compartment score*) rather than a binary one (A/B determined by the sign of the eigenvector) [60]. This finer analysis showed a significant correlation between changes in compartment scores and differences in gene expression in all 45 pairwise comparisons (Fig. 2C and D). While the correlation was highly statistically significant, its magnitude was mostly low ([Supplementary Fig. S3C](#); mean = 0.18, SD = 0.09), indicating that compartment-



**Figure 1.** The cell lines used in our study and their tissue of origin.

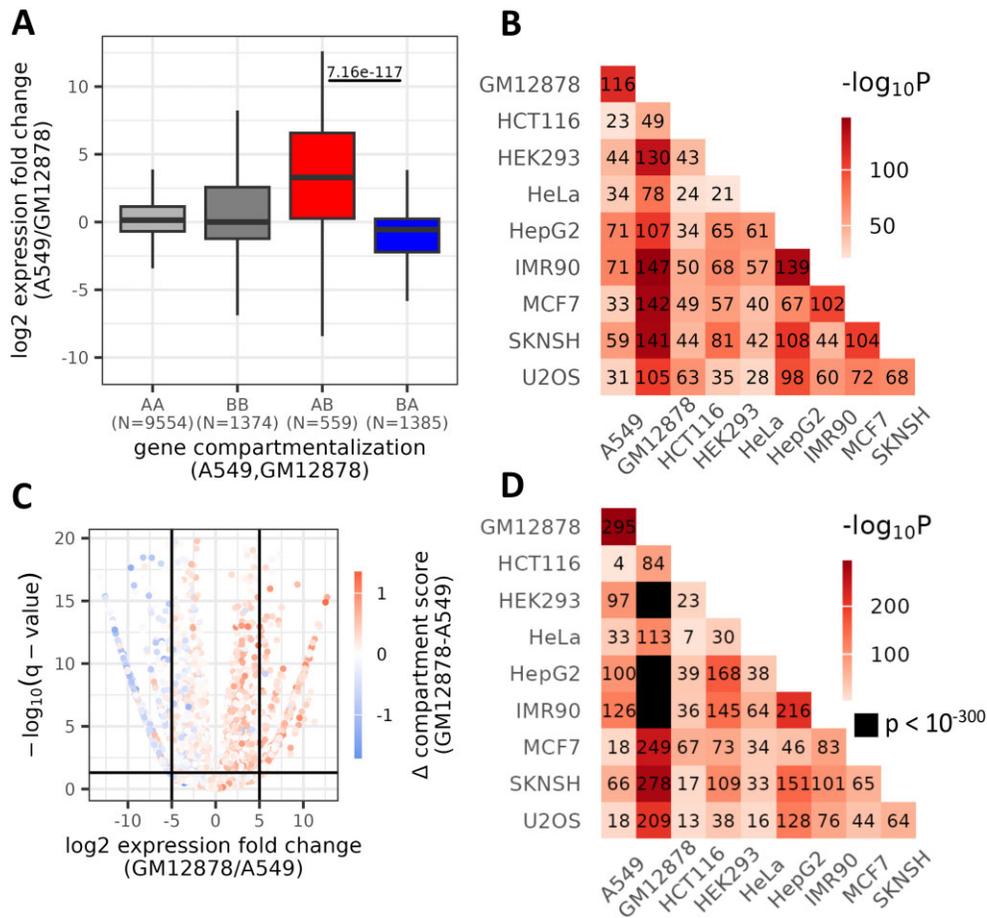
talization is only one of many factors that affect changes in gene expression between different cell lines.

### Association between TADs and differential gene expression between cell lines

TADs were called in each cell line using the arrowhead algorithm implemented in Juicer [46] with contact maps of 10k bp resolution. The average number of TADs detected per cell line was 4836 (min = 3717, max = 6598; Fig. 3A), the average TAD length was ~308k bp (median ~230k bp) (Fig. 3B), and the average number of genes per TAD was 2.36 (SD = 3) (Fig. 3C). TADs are considered largely cell-type invariants [12]. To investigate this in our dataset, we examined the overlap between TADs detected in different cell lines and compared it to the overlap between TADs detected in samples from the same cell line. As the contact maps we obtained from individual samples were relatively sparse, hindering robust detection of TADs, we calculated the overlap between TADs de-

tected in the same cell line in basal state and after p53 activation (merging replicate samples in each biological condition), in addition to calculating TAD overlaps between individual replicate samples. Conservatively assuming that p53 activation has only a mild effect on TADs, the comparison between the basal and Nutlin-treated samples provides us with a lower bound estimate for TADs overlap expected for replicate samples using our experimental protocol and sequencing depth. While TAD overlap between samples from the same cell line was ~60%, the overlap observed between TADs called in different cell lines was markedly lower (~40%) (Fig. 3D). This indicates that in addition to a substantial core set of TADs that are cell-type invariant, there is also a non-negligible portion of dynamic TAD structures.

Previous studies established that CTCF plays a key role in constructing TAD boundaries. Accordingly, the boundaries of the TADs called in our cell lines were significantly enriched for CTCF binding signal (Fig. 3E). We next examined overlaps between TAD boundaries and gene promoters. Of the

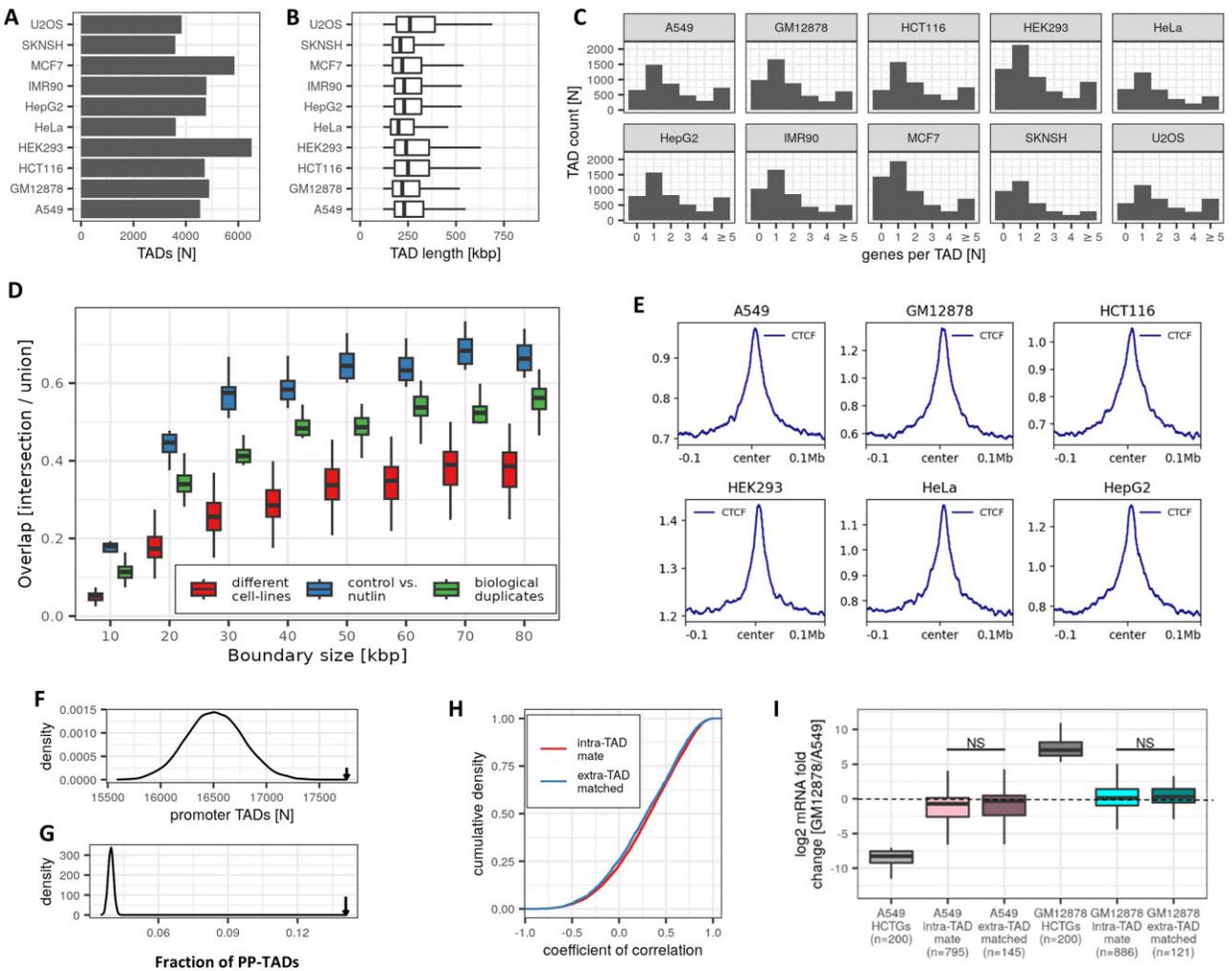


**Figure 2.** Association between chromatin compartmentalization and differential gene expression between cell lines. **(A)** Association between switch in chromatin compartmentalization and changes in gene expression level. As an example, shown here are results for the comparison between A549 and GM12878. Genes were assigned to one of the four groups: AA, BB, AB, and BA, according to their compartment in A549 and GM12878, respectively.  $P$ -value for the difference in the distribution of fold-change (FC) of expression between the genes assigned to the AB and BA groups was calculated using Wilcoxon's test. **(B)** Results of the analysis described in panel (A) applied to all 45 pairwise comparisons. For all pairs, difference in the relative expression of genes assigned to the AB and BA groups were highly statistically significant. **(C)** Volcano plot for differential gene expression between GM12878 and A549. Genes are colored according to the difference in their compartment score between the two cell lines. The horizontal line indicates  $q$ -value = 0.05 for the differential expression test (using DESeq). **(D)** The statistical significance of the correlation observed between expression FC and differential compartmentalization score in each of the 45 pairwise comparisons.  $P$ -values calculated using Spearman's correlation test.

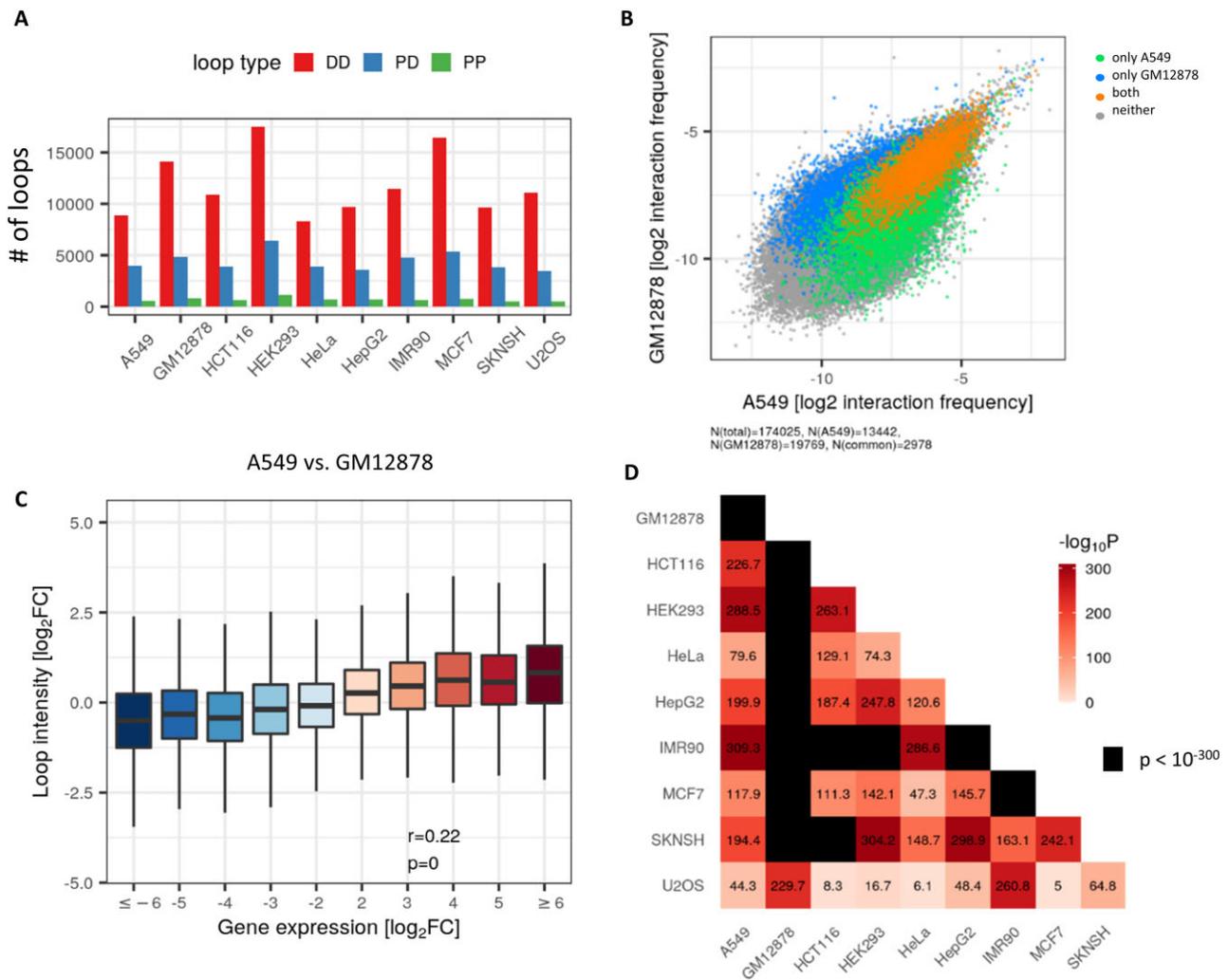
58 520 TADs identified across all cell lines, 22% overlap a promoter at one of their boundaries (we named such TADs “P-TADs”). Permutation tests showed that this overlap is highly significant (Fig. 3F). TADs detected in only a single cell line (“unique TADs”) showed markedly lower overlap with promoters (13%), compared to 25% of the TADs detected in at least seven cell lines (“invariant TADs”). Examination of the binding of CTCF and RAD21 (cohesin-complex subunit) at the boundaries of TADs showed these signals were significantly stronger at the invariant TADs' boundaries than in the unique TADs' boundaries (Supplementary Fig. S4A). Similarly, we observed that CTCF signal was stronger at TAD boundaries overlapping promoters (Supplementary Fig. S4B). Next, we sought TADs where both boundaries overlapped promoters and named them “PP-TADs.” We found that ~14% of the P-TADs were PP-TADs. Permutation tests showed that this fraction of PP-TADs is much higher than expected by chance (Fig. 3G).

TADs are often regarded as insulated regulatory units of transcriptional control. Previous studies suggested that genes residing in the same TAD show a greater level of co-

regulation compared to genes located in adjacent TADs [61]. Seeking evidence for such roles for TADs, we first examined whether genes located within the same TAD (intra-TAD genes) show higher expression correlation over the 10 cell lines in our panel compared to matched control pairs of genes located at comparable distance but not within the same TAD. Somewhat unexpectedly, the intra-TAD genes did not show significantly higher correlated expression patterns than the matched controls (Fig. 3H and Supplementary Fig. S4C). Seeking further support for the notion of TADs as insulated regulatory territories, we next focused on highly cell-type specific genes (“HCTS genes”) that share their TADs with additional genes. Under the assumption that genes within a TAD show higher level of co-regulation, we expected that genes sharing TADs with HCTS genes would show a concordant cell-type specific expression pattern. However, compared to distance-matched extra-TAD genes, we did not find evidence for higher level of co-regulation between HCTS genes and their intra-TAD companion genes: While the intra-TAD mates of HCTS genes did show some degree of concordant elevated expression in the same cell line as their HCTS genes,



**Figure 3.** Association between TADs and differential gene expression across cell lines. **(A)** Number of TADs detected in each cell line (mean = 4836, SD = 933). **(B)** Distribution of TAD lengths in each cell line (Md = 230 kb, SD = 166.6 kb). **(C)** Distribution of the number of genes per TAD in each cell line (average 2.36, SD = 3). **(D)** Overlap between TADs detected in different cell lines (red), in the same cell line without and after Nutlin-3a treatment (blue), and between individual replicate samples (green), using different flanks around TAD boundaries. For each pair of samples, TAD overlap was calculated using the Jaccard similarity (defined as the ratio between the number of overlapping TADs and the number of TADs in the union set). (The overlap obtained between individual replicates was lower than the one obtained for the same cell line without and after Nutlin treatment, as merging replicates doubles depth and results in a more robust TAD detection.) **(E)** Profiles of CTCF ChIP-seq signal (ENCODE data) around TAD boundaries. **(F)** A permutation test for enrichment of TAD boundaries overlapping promoters (P-TADs). We created random shifts of TAD boundaries, which preserve the length of the original ones, by shifting, in each chromosome, the original boundaries' coordinates together by a constant shift that was randomly selected from the [10K, 1M] interval. We repeated these shifts 10 000 times, and in each iteration, counted the number of boundaries that overlapped promoters. This created a Null distribution for the number of P-TADs. The number of P-TADs observed in the real data (black arrow) was significantly higher. **(G)** A similar randomization test as in panel (F) to examine the enrichment of P-TADs for PP-TADs. Here, we took record of the lengths of the P-TADs and then randomly matched each promoter boundary with a distal boundary in a way that preserved the original length distribution of the P-TADs, and then counted the number of PP-TADs that occurred in this random setting. We repeated this random shuffling 10 000 times to create a Null distribution for the proportion of P-TADs that are PP-TADs. The observed proportion in the real dataset (black arrow) is significantly higher than the expected under the Null. **(H)** Correlation between expression profiles over the 10 cell lines for intra-TAD gene pairs (red) and distance-matched control gene pairs (“extra-TADs”; blue). **(I)** Comparing GM12878 and A549, we focused on the top 400 differentially expressed genes (“highly cell-type specific genes”—HCTGs) that have additional genes within their TADs (taking the top 200 HCTGs in each cell line compared to the other). Then, for each intra-TAD mate gene of an HCTS gene, we selected a matched control gene that is located in an adjacent TAD or outside any TAD but is as close to the HCTS gene as the intra-TAD mate [that is, located at a distance from the HCTS gene that is not larger than the intra-TAD gene (Supplementary Fig. S4C)]. The distribution of expression FC (between GM12878 and A549) did not differ between the intra-TAD mate and the matched control genes (*t*-test; ns = not significant). Similar results were obtained for pairwise comparisons between all cell lines.



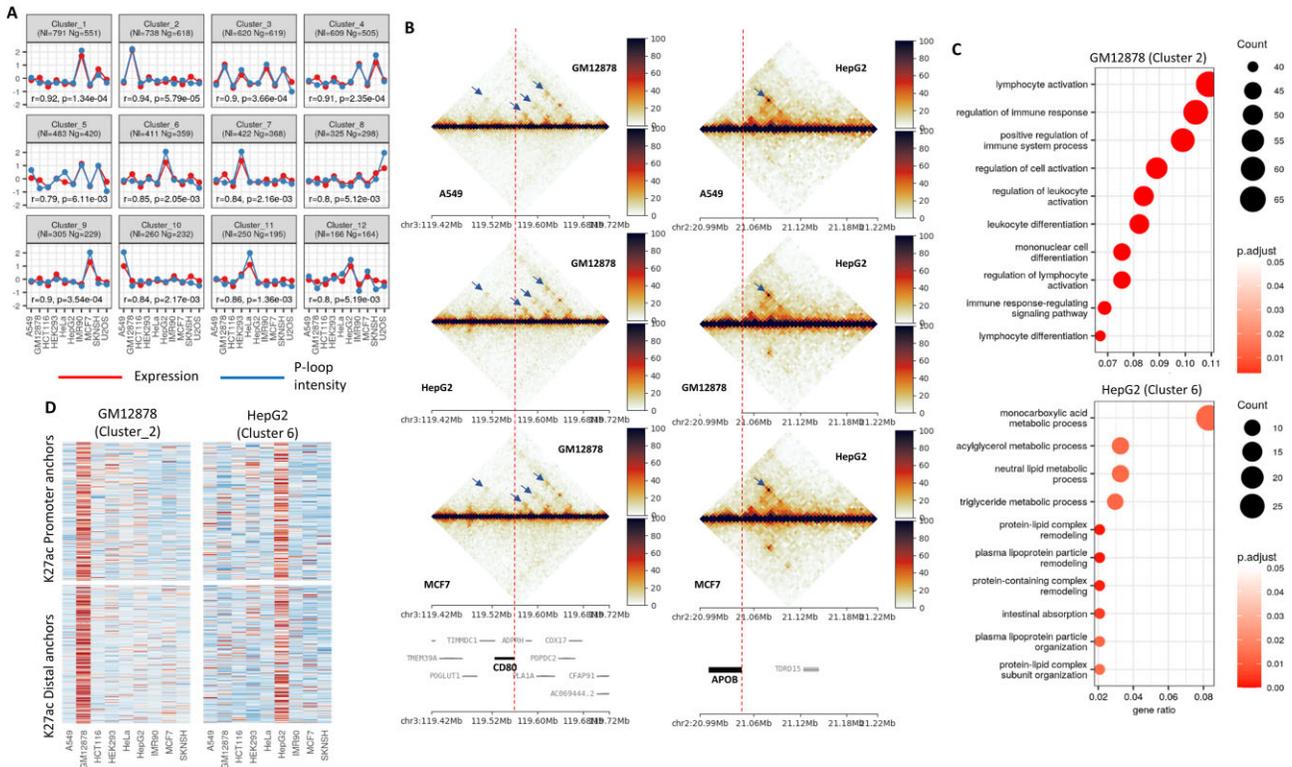
**Figure 4.** Association between chromatin loops and differential gene expression across cell lines. **(A)** Number of PP, PD, and DD loops called in each cell line. **(B)** Interaction intensities measured in A549 and in GM12878 cell lines for all the loops in the pooled set. Green: loops called only in A549; blue: called only in GM12878; orange: called in both; and gray: called in none of these two cell lines. **(C)** Correlation between changes in gene expression and in promoter interaction intensity. Comparing A549 and GM12878, for each gene we calculated the FC in expression level and the FC in interaction intensity of all the loops that have an anchor in the gene’s promoter. We divided the genes into 10 bins according to FC in expression and calculated the distribution of FC in P-loops interaction intensities in each bin. The correlation between expression FC and mean interaction intensity over the bins was calculated using Pearson coefficient. **(D)** Results of the correlation test in panel (C) applied to all 45 pairwise comparisons. In all cases, the correlation between differential gene expression and changes in P-loop intensity was highly statistically significant.

the distance-matched control genes showed a similar trend (Fig. 3I).

### Association between chromatin loops and differential gene expression across cell lines

Chromatin loops in each sample were identified using Mustache [47] with 5k bp resolution contact maps. We identified an average of 17 920 loops per sample, with a total of 174 025 unique chromatin loops across all samples. We categorized the loops into three groups: PP (promoter–promoter,  $N = 6031$ ), PD (promoter–distal,  $N = 41\,861$ ), and DD (distal–distal,  $N = 126\,133$ ) loops—indicating loops with two, one, and no anchors overlapping a promoter region, respectively (Fig. 4A). Permutation tests showed that the loops detected in our dataset are highly enriched for P loops (Supplementary Fig. S5A) and that the P loops themselves are highly enriched for PP loops (Supplementary Fig. S5B). Interestingly, CTCF and Cohesin (RAD21 subunit) binding was

enriched much more strongly at anchors of cell-line invariant loops (“constitutive loops,” detected in more than eight cell lines) than in anchors of cell-type specific loops (detected in only one cell line) (Supplementary Fig. S5C). Next, seeking to handle chromatin loops as quantitative, rather than binary (present/absent) entities, we analyzed the pooled set of loops from all cell lines and extracted the normalized interaction frequencies of each loop (“loop intensities”) in each of the 10 cell lines (Fig. 4B). Importantly, in all pairwise comparisons, changes in gene expression were significantly correlated with changes in intensities of the corresponding P loops (Fig. 4C and D). This observation remained valid when PD and, to a lesser extent, PP loops were considered separately (Supplementary Fig. S5D and E). Notably, in line with a previous observation [20, 62], the dynamic range of changes in gene expression levels was an order of magnitude larger than the range of changes in interaction frequencies. This implies that mild alterations in P-loop interaction intensity could be associated with strong changes in gene expression.



**Figure 5.** Candidate functional P loops. **(A)** Clustering of the candidate functional P loops. Each cluster shows the mean activity pattern of the P loops (blue) and the mean expression profile (red) of the genes assigned to the cluster. The numbers of P loops ( $N_p$ ) and genes ( $N_g$ ) assigned to each cluster are indicated in its title. Clustering was done using the CLICK algorithm from the EXPANDER package [77]. Interaction intensities and gene expression levels were standardized across the 10 samples (mean = 0, SD = 1) before clustering. **(B)** Examples of contact maps in the vicinity of two highly cell-type specific genes: CD80 (GM12878-specific) and APOB (HepG2-specific). Cell-type specific functional P loops linked to these genes are indicated by arrows. Dashed red line marks the location of the gene's TSS. (The plot was generated using FANC [78]). **(C)** Enriched Gene Ontology (GO) categories (FDR < 5%) detected in cluster 2 (P loops specifically active in GM12878) and cluster 6 (P loops specifically active in HepG2). GO enrichment analysis was done using the clusterProfiler R package [79] (The set of all genes associated with any P loop was used as the background set in these tests). **(D)** Heatmaps of H3K27ac signal (based on ENCODE ChIP-seq data) in the promoter and distal anchors of the P loops assigned to clusters 2 (GM12878) and 6 (HepG2).

Notwithstanding the significant association between changes in P-loop intensities and differential gene expression, the magnitude of the correlation was modest (<0.25 for all pairwise analyses), indicating that, in general, alterations in intensity of promoter chromatin interactions account for only a small part of gene expression modulation. Therefore, we next sought to identify candidate P loops that play a principal role in controlling the transcription of their target genes. To this end, we linked each gene with its P loops, and then for each of the 61 138 gene-P-loop pairs in our dataset, we calculated the correlation between the intensity of the P loop and the expression level of its linked gene across all samples in the dataset. While the distribution of correlations was significantly skewed to positive values, the correlation magnitude was generally low (mean = 0.11) (Supplementary Fig. S6A). Yet, this analysis identified thousands of P loops whose intensities are highly correlated with the expression of their target genes (Supplementary Fig. S6A; for example, 6897 P loops associated with 3132 genes showed  $r > 0.5$ , ~7-fold higher than the number showing the opposite trend of  $r < -0.5$ ). A permutation test showed that our dataset is markedly enriched for P loops whose intensities are highly correlated ( $r > 0.5$ ) with the expression of their linked genes (Supplementary Fig. S6B). As these P loops are candidate regulatory interactions that play a marked role in

the transcriptional regulation of their targets, hereafter we refer to them as “functional P loops.”

The functional P loops identified in our dataset display many different intensity profiles over the panel of 10 cell lines. Thus, we next applied clustering analysis to systematically delineate sets of loops with similar profiles. This analysis identified 12 major clusters, each with >150 functional P loops. Each cluster is characterized by a particular activity profile shared by the P loops assigned to it and a highly correlated expression profile of their associated target genes (Fig. 5A). Some of these clusters represent P loops and target genes that are active in only one specific cell line (e.g. cluster 2—GM12878; cluster 6—HepG2). Thus, these clusters represent cell-type specific transcriptional programs associated with their cell-type specific chromatin loops. Two prominent examples of cell-type specific genes and the 3D genome organization in their vicinity are shown in Fig. 5B: CD80, specifically expressed in GM12878, and APOB, specifically expressed in HepG2.

GO enrichment analysis showed that clusters with cell-type specific profiles are enriched for genes that carry out fundamental biological processes related to that cell type. For example, cluster 2, which shows an activity profile that is highly specific to GM12878 (a lymphoblastoid cell line), is enriched for genes that function in various immune processes. Clus-

ter 6, which shows an activity profile that is highly specific to HepG2 (a liver cancer cell line), is enriched for genes that function in main liver processes (Fig. 5C and Supplementary Fig. S6C). Therefore, our analysis delineated hundreds of P loops that are associated with cell-type specific gene regulation (Supplementary Table S2). We expected that many of these loops represent E–P loops. To test this expectation, we examined the signal for H3K27ac and H3K4me1, key markers of active regulatory elements, at both the promoter and distal anchors of the cell-type specific functional P loops (by construction, one of the anchors of P loops is located at a promoter). Collectively, for most cell-type specific clusters, we indeed found an enrichment for H3K27ac and H3K4me1 signals specifically in the cell line where the P loops and the target genes are active (Fig. 5D and Supplementary Fig. S6D), suggesting that many of the functional P loops we detected are genuine active E–P loops.

### Associations between spatial chromatin organization and differential gene expression upon p53 activation

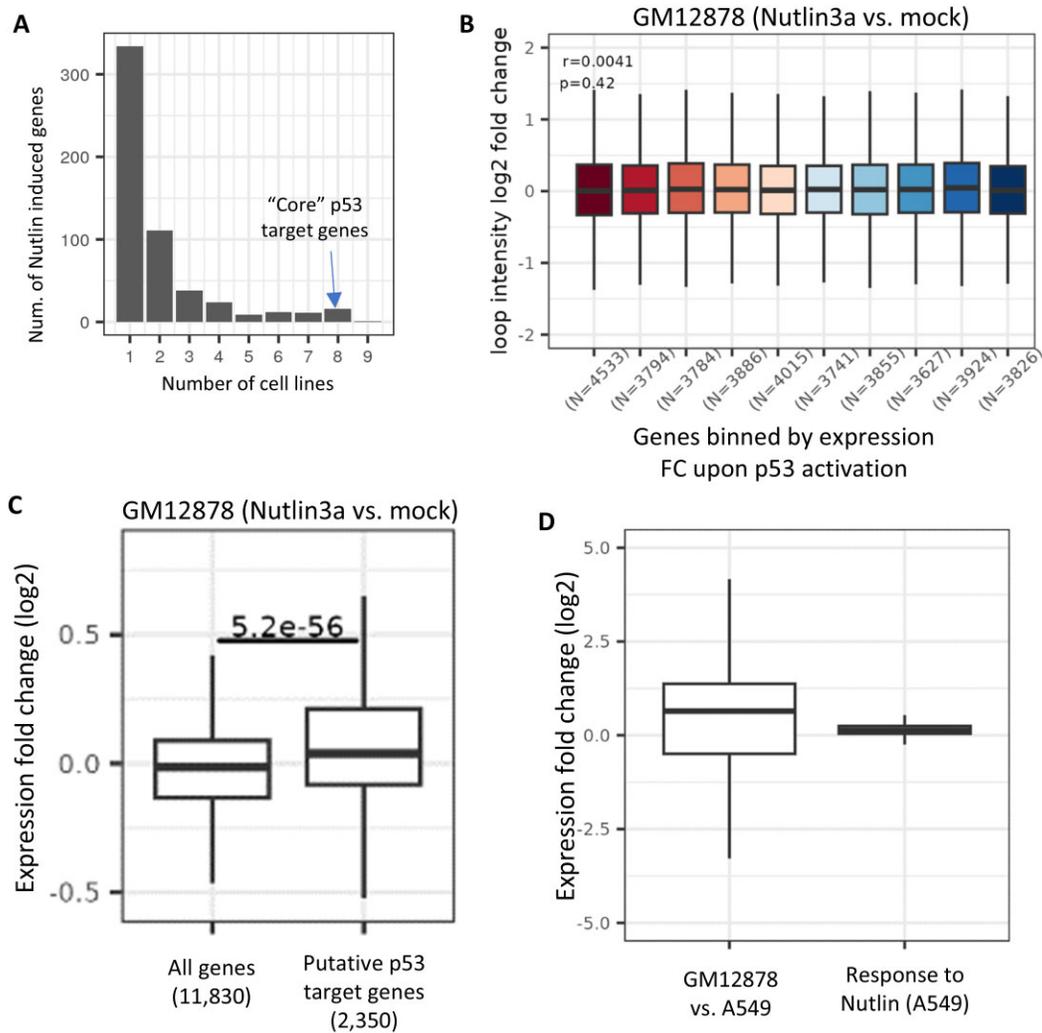
The preceding analyses focused on differences between cell lines of highly diverse tissue of origin and demonstrated a strong link between spatial genome organization and transcriptional programs that determine cell identity. Next, we turned to analyze milder transcriptional differences—namely, those that are induced by cellular stress. As a model system, we used transcriptional responses to p53 activation. We treated each of the panel's cell lines with the potent p53 activator Nutlin-3a, which acts by inhibiting MDM2, a protein that directs p53 for degradation [40]. In each cell line, we recorded the spatial genome organization 4 h after Nutlin-3a treatment using micro-C. In parallel, we performed RNA-seq and p53 ChIP-seq in the same biological conditions to obtain comprehensive snapshots of gene expression levels and p53-chromatin binding profiles, both without and after Nutlin-3a treatment. All the cell lines in our panel except two, HeLa and HEK293, have functional p53. The number of induced genes differed by more than threefold between the cell lines with functional p53 (min = 45; max = 199; as expected, HeLa and HEK293 showed a minimal response, Supplementary Fig. S7A; Supplementary Table S3), and most of the responsive genes were cell-type specific (Fig. 6A). The core responding genes, induced in all eight cell lines with functional p53, included well-documented canonical p53 target genes as *CDKN1A*, *MDM2*, *SESN1-2*, *BBC3*, *GDF15*, and *PPM1D*.

We next turned to correlate the transcriptional response induced by p53 activation with changes in 3D chromatin organization. First, we found that in all cell lines, there were minimal events of A/B compartment switching upon Nutlin-3a treatment (<10 in all cell lines; Supplementary Fig. S7B). Second, hypothesizing that many of the induced genes are driven by E–P loops that are established or stabilized upon p53 activation, we sought correlations between changes in gene expression and alterations in P-loop intensity induced in response to Nutlin-3a treatment. However, in contrast to the significant association that we observed when comparing different cell lines (Fig. 4C and D), changes in gene expression upon p53 activation were not correlated with corresponding alterations in P-loop intensities (Fig. 6B). Aiming to enhance the detection of such correlations, we con-

finned this analysis to the set of genes that responded most strongly to p53 activation, contrasting, in each cell line, the induced and repressed genes. However, this analysis too did not detect significant association between alterations in P-loop intensities and changes in the expression of the responding genes (Supplementary Fig. S7C). Nevertheless, visual inspection of the contact plots near p53 canonical target genes did detect some changes in chromatin organization that were induced upon Nutlin-3a treatment (Supplementary Fig. S7D). But these changes were subtle and did not pass strict statistical tests for differential loop intensity.

Next, we used our p53 ChIP-seq data to define sets of putative direct target genes of p53 in each cell line. p53 ChIP-seq analysis detected thousands of p53-chromatin binding events that were induced upon Nutlin-3a treatment (Supplementary Tables S3 and S4). Reassuringly, in all cell lines, the genomic sequences at these p53 binding locations (“p53 peaks”) were highly enriched for the known p53 binding motif (Supplementary Table S5). Intersecting p53 peaks with the pooled set of loops detected in our micro-C dataset identified chromatin loops with p53 binding at one of their anchors (we call such loops “p53 loops”). Unexpectedly, when considered collectively, we did not find an increase in the intensity of these loops upon Nutlin-3a treatment and the ensuing p53 binding (Supplementary Fig. S7E). However, examining the expression of genes linked to p53 P loops (that is, genes linked to loops with one of their anchors at the gene's promoter and the other at a Nutlin-3a induced p53 binding site) did show modest and nominally significant ( $P < .05$ ) induction upon Nutlin-3a treatment in seven of the eight cell lines with functional p53 (Supplementary Fig. S7F). In addition, we created in each cell line a set of putative p53 direct target genes by merely linking p53 peaks to their nearest gene (up to a distance of 50 kb). In line with previous studies [63], most of these genes were not induced in response to p53 activation (Supplementary Table S6). Yet, when considered collectively, we found in all eight cell lines that these sets of genes showed significant up-regulation in response to Nutlin-3a treatment (Fig. 6C, Supplementary Fig. S7G and H). The marked induction of these genes indicates a regulatory role for the binding of p53 at the respective regulatory elements detected by the ChIP-seq analysis. However, in most of these cases, our micro-C data did not detect chromatin looping between the p53 binding sites and the promoters of the induced genes. Taken together, despite the significant correlation we observed between the induction of p53 binding to its regulatory sites and the transcriptional responses to p53 activation, our micro-C data did not detect a correlated modulation of the spatial chromatin organization.

Last, seeking some determinants of the highly cell-type specific transcriptional responses to p53 activation (Fig. 6A), we examined if cell-type specific p53 binding sites (as detected by our ChIP-seq analysis) were primed for activation before the exposure of cells to Nutlin-3a treatment. To this end, we analyzed public DHS-seq data (from ENCODE), which were measured under basal conditions in the cell lines used in our study. We found that, in general, cell-type specific p53 binding sites were associated with an open-chromatin state in the untreated condition specifically in the cell line in which these sites were bound by p53 in response to the treatment (Supplementary Fig. S8A–C). Furthermore, in few cell lines, motif analysis found that the p53-bound regions were enriched for motifs of cell-fate determining TFs of the corre-



**Figure 6.** Associations between spatial chromatin organization and differential gene expression upon p53 activation. **(A)** Most of the transcriptional response to Nutlin-3a treatment is cell-type specific. The histogram shows the distribution of the number of cell lines in which the activated genes were induced. Seventeen p53 core target genes were induced in all eight cell lines with functional p53. **(B)** Genes expressed in GM12878 (CPM  $\geq 1$ ,  $N = 11\,069$ ) were binned into 10 equally sized bins according to their expression FC in response to Nutlin-3a treatment (781 genes per bin). The distribution of loop intensity FC was calculated for the P loops associated with the genes in each bin (numbers of P loops in each bin are indicated). No correlation was found between changes in expression and promoter interaction intensity. Similar results were obtained for all cell lines. **(C)** p53 peaks induced by Nutlin-3a in GM12878 were linked to their nearest promoters, up to a distance of 50 kb, and the expression FC of the corresponding genes ("putative p53 targets") was compared to the FC of all the other genes in the dataset.  $P$ -value calculated using Wilcoxon's test. **(D)** Distribution of expression FC, considering all genes in the data, in the comparison between different cell lines (GM12878 versus A549) and within a cell line in response to Nutlin-3a treatment (A549).

sponding cell line lineages (in addition to the p53 motif). These TFs included IRFs, NF $\kappa$ B, and SPI [64] in GM12878 lymphoblastoid cell line, HNF4 [65] in HepG2 liver cancer cell line, GRHLs [66] in MCF7 breast cancer cell line and NFIX [67] in SKNSH neuroblastoma cell line. This result suggests that in each cell type, lineage-determining TFs play a role in outlining cell-type specific p53 responses by shaping the chromatin landscape that is available for binding of p53 upon its activation.

## Discussion

In the first part of our study, we focused on characterizing associations between the spatial organization of the genome and gene expression profiles in a diverse set of cell lines. At the layer of A/B compartmentalization, our results reaffirm

the substantial correlation between differential expression and compartment switching, suggesting dynamic localization of mega-base-pair chromatin compartments to the nucleus' periphery or center in a way that is linked to transcriptional activity (Fig. 2). As for TAD structures, while our results indicate a substantial core set of cell-type invariant TADs, they also suggest a non-negligible portion of TADs that are dynamic (Fig. 3D). Focusing on highly cell-type specific genes (HCTGs), whose expression is supposedly controlled by potent cell-type specific enhancers, we did not observe evidence supporting strong and global insulating activity for TADs. We found that while genes that share TADs with HCTGs show consistent, albeit attenuated, preferential expression in the same cell type, suggesting some extent of enhancer sharing between genes in the same TAD, genes located outside the TAD but matched with respect to their distance to the

HCTGs showed similar level of preferential expression as the mate TAD genes (Fig. 3I).

As for intra-TAD chromatin loops, first, we observed a significant enrichment for P loops and P–P loops (Supplementary Fig. S5A and B), further supporting the concept that active promoters form higher-order hub-like structures in which multiple active promoters co-localize [68–70]. Second, we show the quantitative nature of the micro-C measurements (Fig. 4B) and demonstrate that across cell types, differential promoter transcriptional activity is strongly correlated with differential engagement of the promoters in chromatin looping (Fig. 4C and D). In line with a previous report [62], we observed that changes in promoter interaction intensity are an order of magnitude lower than the correlated changes in gene expression (Fig. 4C), indicating that even mild changes in contact frequency between an active enhancer and its target promoter could be associated with substantial induction of gene expression. Next, we found hundreds of chromatin P loops that are highly correlated with the expression profile of their target genes, which play key roles in biological processes that carry out fundamental functions of the respective cell types/tissues. Thus, we have systematically delineated multiple networks of E–P loops and their associated gene expression programs that determine cell identity (Fig. 5).

Last, we characterized correlations in a much subtler biological context: cellular response to stress, using p53 activation as a model system, which elicits transcriptional programs that are weaker than those establishing cell identity. Accordingly, p53 activation, which typically resulted in the induction of a few dozen genes (in contrast to thousands of differentially expressed genes in comparisons between different cell lines), was associated with very few, if any, events of A/B compartment switching. Additionally, we did not observe a global correlation between gene induction and elevated chromatin interaction intensity at the promoters of the induced genes (Fig. 6B). However, the genes that are closest to p53-bound regulatory elements in each cell line did show significant induction of expression upon p53 activation (Fig. 6C and Supplementary Fig. S7G). Yet, our micro-C data detected only very mild 3D genome reorganization near the promoters of p53 canonical target genes (Supplementary Fig. S7D) and only subtle links between “p53 loops” and gene activation upon Nutlin-3a treatment (Supplementary Fig. S7F). These results may represent the much weaker and more transient propensity of E–P loops established upon p53 activation in comparison to E–P loops that control cell-identity genes. Congruently, alteration in gene expression driven by p53 activation (max. induction is typically 2–4-fold) is markedly narrower than the level of differential expression exhibited by cell-identity genes (Fig. 6D). As an alternative explanation for the lack of correlation between change in expression and interaction intensity at the promoters of the induced genes, our results are in line with the recently proposed TF activity gradient (TAG) model [71], which suggested a contact-independent mechanism for E–P communication. According to the TAG model, TFs are first recruited to enhancers where they are activated (e.g. by p300 acetylation). Next, activated TFs disengage from the enhancers and communicate with target promoters through diffusion. Rapid deactivation (deacetylation) of TFs ensures that the spatial range of their effect is limited. Thus, under this model, the requirement for proximity between the distal enhancer and its target promoter(s) is met without necessitating a direct contact between the two elements [71]. No-

tably, p53 particularly fits the TAG model, as its activation by p300 acetylation upon its binding to enhancers was demonstrated [72, 73].

One puzzle that remains is the high cell-type specificity of the transcriptional response to p53 activation (Fig. 6A), reflected by the high cell-type specific landscape of p53-chromatin interactions (Supplementary Fig. S8B). Interestingly, we found that regulatory elements that are bound by p53 in a specific cell line are in an open-chromatin state already in the basal condition, specifically in that cell line (Supplementary Fig. S8C). Furthermore, for several cell lines, we found that these genomic regions are enriched for binding motifs of key TFs that determine cell fate of the corresponding cell lineage. This suggests that key transcriptional regulators in each cell type shape the genomic space that is available for binding by stress-induced TFs, and consequently, play a role in the determination of cell-type specific transcriptional responses to stress.

Our study has a few limitations. First, we used cell lines as our experimental systems. It will be interesting to examine how our observations generalize to *in vivo* tissues and primary cells. Second, although we used the micro-C technique to delineate the spatial genome organization, the highest resolution of contact maps we used in our analyses was 5 kb. We also performed chromatin loop analysis using contact matrices with lower bin sizes, but these matrices were too sparse, and the loops detected were less robust. Yet, the advantages of using Micro-C in our study are that MNase digestion introduces lower sequence biases compared to restriction digestion and that the use of MNase allows for milder detergent pretreatment of cross-linked cells, thereby improving the preservation of E–P contacts, which are more sensitive to experimental conditions than CTCF-based TAD interactions. [74, 75]. Advanced variation of micro-C (Micro-Capture-C) can determine E–P contacts in specified genomic regions at base-pair resolution [76]. Application of such techniques to p53-bound enhancers would further examine the validity of the contact-independent TAG model for the activation of p53 target genes. Last, our study is based on correlative patterns, and as such, the set of putative functional P loops identified by our analysis requires an experimental validation.

Our micro-C data, together with matched RNA-seq and p53 ChIP-seq data, provide an extensive resource for the genomics research community. Our results further establish links between chromatin organization and transcriptional regulation. Collectively, we have delineated hundreds of candidate functional cell-type specific E–P loops and highlighted the contrast between chromatin interactions associated with genes involved in cell-fate determination and those involved in cellular responses to p53 stress.

## Acknowledgements

*Author contributions:* Gony Shanel (Conceptualization [supporting], Formal analysis [lead], Writing—original draft [lead]), Tsung-Han S. Hsieh (Conceptualization [lead], Data curation [lead], Formal analysis [supporting], Methodology [lead], Writing—original draft [supporting]), Claudia Cattoglio (Data curation [supporting], Formal analysis [lead], Methodology [lead], Writing—original draft [lead]), Hadar Amira Haham (Formal analysis [supporting]), Hsin-Jung Chou (Formal analysis [supporting]), Jack Li (Formal analysis [supporting]), Ron Shamir (Supervision [supporting],

Writing—review & editing [supporting]), Xavier Darzacq (Conceptualization [lead], Funding acquisition [lead], Resources [lead], Supervision [supporting], Writing—review & editing [supporting]), Ran Elkon (Conceptualization [lead], Funding acquisition [lead], Methodology [lead], Resources [lead], Supervision [lead], Writing—original draft [lead]).

## Supplementary data

Supplementary data is available at NAR online.

## Conflict of interest

None declared.

## Funding

This study was supported by grants from the Koret-UC Berkeley-Tel Aviv University Initiative in Computational Biology and Bioinformatics to R.E., R.S. (EJSCB, TAU), and X.D. (UCB), and by Tel Aviv University Center for AI and Data Science (TAD) to R.E. and R.S. R.E. is a Faculty Fellow of the Edmond J. Safra Center for Bioinformatics at Tel Aviv University. G.S. and H.A.H. were partially supported by a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel Aviv University. Funding to pay the Open Access publication charges for this article was provided by TAU Research funds.

## Data availability

The Micro-C, ChIP-seq, and RNA-seq data generated in this publication are available in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) through accession no. GSE279021. ENCODE accession numbers for published data reused in this manuscript are listed in Supplementary Table S7. The code used to generate the main figures, together with the relevant input files, is available at <https://github.com/ElkonLab/micro-C> and <https://doi.org/10.5281/zenodo.15600163>.

## References

- Oudelaar AM, Higgs DR. The relationship between genome structure and function. *Nat Rev Genet* 2021;22:154–68. <https://doi.org/10.1038/s41576-020-00303-x>
- Andersson R, Gebhard C, Miguel-Escalada I *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* 2014;507:455–61. <https://doi.org/10.1038/nature12787>
- Long HK, Prescott SL, Wysocka J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* 2016;167:1170–87. <https://doi.org/10.1016/j.cell.2016.09.018>
- Furlong EEM, Levine M. Developmental enhancers and chromosome topology. *Science* 2018;361:1341–5. <https://doi.org/10.1126/science.aau0320>
- Lieberman-Aiden E, van Berkum NL, Williams L *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009;326:289–93. <https://doi.org/10.1126/science.1181369>
- Kempfer R, Pombo A. Methods for mapping 3D chromosome architecture. *Nat Rev Genet* 2020;21:207–26. <https://doi.org/10.1038/s41576-019-0195-2>
- Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 2013;14:390–403. <https://doi.org/10.1038/nrg3454>
- Hsieh TH, Weiner A, Lajoie B *et al.* Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* 2015;162:108–19. <https://doi.org/10.1016/j.cell.2015.05.048>
- Hsieh TS, Fudenberg G, Goloborodko A *et al.* Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. *Nat Methods* 2016;13:1009–11. <https://doi.org/10.1038/nmeth.4025>
- Gibcus JH, Dekker J. The hierarchy of the 3D genome. *Mol Cell* 2013;49:773–82. <https://doi.org/10.1016/j.molcel.2013.02.011>
- Rowley MJ, Corces VG. Organizational principles of 3D genome architecture. *Nat Rev Genet* 2018;19:789–800. <https://doi.org/10.1038/s41576-018-0060-8>
- Schmitt AD, Hu M, Jung I *et al.* A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep* 2016;17:2042–59. <https://doi.org/10.1016/j.celrep.2016.10.061> []
- Rao SS, Huntley MH, Durand NC *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159:1665–80. <https://doi.org/10.1016/j.cell.2014.11.021>
- Dixon JR, Selvaraj S, Yue F *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;485:376–80. <https://doi.org/10.1038/nature11082>
- Downen JM, Fan ZP, Hnisz D *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 2014;159:374–87. <https://doi.org/10.1016/j.cell.2014.09.030>
- Hanssen LLP, Kassouf MT, Oudelaar AM *et al.* Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function *in vivo*. *Nat Cell Biol* 2017;19:952–61. <https://doi.org/10.1038/ncb3573>
- Lupianez DG, Kraft K, Heinrich V *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 2015;161:1012–25. <https://doi.org/10.1016/j.cell.2015.04.004>
- Dixon JR, Gorkin DU, Ren B. Chromatin domains: the unit of chromosome organization. *Mol Cell* 2016;62:668–80. <https://doi.org/10.1016/j.molcel.2016.05.018>
- Chang LH, Ghosh S, Noordermeer D. TADs and their borders: free movement or building a wall? *J Mol Biol* 2020;432:643–52. <https://doi.org/10.1016/j.jmb.2019.11.025>
- Xiao JY, Hafner A, Boettiger AN. How subtle changes in 3D structure can create large changes in transcription. *eLife* 2021;10:e64320. <https://doi.org/10.7554/eLife.64320>
- Rao SSP, Huang SC, Glenn St Hilaire B *et al.* Cohesin loss eliminates all loop domains. *Cell* 2017;171:305–20.e24. <https://doi.org/10.1016/j.cell.2017.09.026>
- Cavalheiro GR, Pollex T, Furlong EE. To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? *Curr Opin Genet Dev* 2021;67:119–29. <https://doi.org/10.1016/j.gde.2020.12.015>
- Ghavi-Helm Y, Jankowski A, Meiers S *et al.* Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat Genet* 2019;51:1272–82. <https://doi.org/10.1038/s41588-019-0462-3>
- Hafner A, Boettiger A. The spatial organization of transcriptional control. *Nat Rev Genet* 2023;24:53–68. <https://doi.org/10.1038/s41576-022-00526-0>
- Nagano T, Lubling Y, Varnai C *et al.* Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* 2017;547:61–7. <https://doi.org/10.1038/nature23001>
- Espinola SM, Gotz M, Bellec M *et al.* Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early *Drosophila* development. *Nat Genet* 2021;53:477–86. <https://doi.org/10.1038/s41588-021-00816-z>
- Bintu B, Mateo LJ, Su JH *et al.* Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* 2018;362:eau1783. <https://doi.org/10.1126/science.aau1783>

28. Mir M, Bickmore W, Furlong EEM *et al.* Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? *Development* 2019;146:dev182766. <https://doi.org/10.1242/dev.182766>
29. Goel VY, Huseyin MK, Hansen AS. Region Capture Micro-C reveals coalescence of enhancers and promoters into nested microcompartments. *Nat Genet* 2023;55:1048–56. <https://doi.org/10.1038/s41588-023-01391-1>
30. Schoenfelder S, Fraser P. Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet* 2019;20:437–55. <https://doi.org/10.1038/s41576-019-0128-0>
31. Heinz S, Romanoski CE, Benner C *et al.* The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* 2015;16:144–54. <https://doi.org/10.1038/nrm3949>
32. Hansen AS, Cattoglio C, Darzacq X *et al.* Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* 2018;9:20–32. <https://doi.org/10.1080/19491034.2017.1389365>
33. Conte M, Fiorillo L, Bianco S *et al.* Polymer physics indicates chromatin folding variability across single-cells results from state degeneracy in phase separation. *Nat Commun* 2020;11:3289. <https://doi.org/10.1038/s41467-020-17141-4>
34. Novo CL, Javierre BM, Cairns J *et al.* Long-range enhancer interactions are prevalent in mouse embryonic stem cells and are reorganized upon pluripotent State transition. *Cell Rep* 2018;22:2615–27. <https://doi.org/10.1016/j.celrep.2018.02.040>
35. Phanstiel DH, Van Bortle K, Spacek D *et al.* Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol Cell* 2017;67:1037–48. <https://doi.org/10.1016/j.molcel.2017.08.006>
36. Bonev B, Mendelson Cohen N, Szabo Q *et al.* Multiscale 3D genome rewiring during mouse neural development. *Cell* 2017;171:557–72. <https://doi.org/10.1016/j.cell.2017.09.043>
37. Batut PJ, Bing XY, Sisco Z *et al.* Genome organization controls transcriptional dynamics during development. *Science* 2022;375:566–70. <https://doi.org/10.1126/science.abi7178>
38. Amat R, Bottcher R, Le Dily F *et al.* Rapid reversible changes in compartments and local chromatin organization revealed by hyperosmotic shock. *Genome Res* 2019;29:18–28. <https://doi.org/10.1101/gr.238527.118>
39. Ray J, Munn PR, Vihervaara A *et al.* Chromatin conformation remains stable upon extensive transcriptional changes driven by heat shock. *Proc Natl Acad Sci USA* 2019;116:19431–9. <https://doi.org/10.1073/pnas.1901244116>
40. Wang B, Fang L, Zhao H *et al.* MDM2 inhibitor Nutlin-3a suppresses proliferation and promotes apoptosis in osteosarcoma cells. *ABBS* 2012;44:685–91. <https://doi.org/10.1093/abbs/gms053>
41. Slobodyanyuk E, Cattoglio C, Hsieh TS. Mapping mammalian 3D genomes by micro-C. *Methods Mol Biol* 2022;2532:51–71.
42. Servant N, Varoquaux N, Lajoie BR *et al.* HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol* 2015;16:259. <https://doi.org/10.1186/s13059-015-0831-x>
43. Abdennur N, Mirny LA. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* 2020;36:311–6. <https://doi.org/10.1093/bioinformatics/btz540>
44. Open2C, Abdennur N, Sameer A *et al.* . Cooltools: enabling high-resolution Hi-C analysis in Python. *PLoS Comput Biol* 2024;20:e1012067. <https://doi.org/10.1371/journal.pcbi.1012067>
45. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2. <https://doi.org/10.1093/bioinformatics/btq033>
46. Durand NC, Shamim MS, Machol I *et al.* Juicer provides a one-click system for analyzing loop-resolution hi-C experiments. *Cell Syst* 2016;3:95–8. <https://doi.org/10.1016/j.cels.2016.07.002>
47. Roayaei Ardakany A, Gezer HT, Lonardi S *et al.* Mustache: multi-scale detection of chromatin loops from hi-C and Micro-C maps using scale-space representation. *Genome Biol* 2020;21:256. <https://doi.org/10.1186/s13059-020-02167-0>
48. Greenwald WW, Li H, Smith EN *et al.* Pgltools: a genomic arithmetic tool suite for manipulation of hi-C peak and other chromatin interaction data. *BMC Bioinformatics* 2017;18:207. <https://doi.org/10.1186/s12859-017-1621-0>
49. Bray NL, Pimentel H, Melsted P *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. <https://doi.org/10.1038/nbt.3519>
50. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>
51. Serin Harmanci A, Harmanci AO, Zhou X. CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat Commun* 2020;11:89. <https://doi.org/10.1038/s41467-019-13779-x>
52. Imakaev M, Fudenberg G, McCord RP *et al.* Iterative correction of hi-C data reveals hallmarks of chromosome organization. *Nat Methods* 2012;9:999–1003. <https://doi.org/10.1038/nmeth.2148>
53. Testa A, Donati G, Yan P *et al.* Chromatin immunoprecipitation (ChIP) on chip experiments uncover a widespread distribution of NF-Y binding CCAAT sites outside of core promoters. *J Biol Chem* 2005;280:13606–15. <https://doi.org/10.1074/jbc.M414039200>
54. Langmead B, Trapnell C, Pop M *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>
55. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-seq (MACS). *Genome Biol* 2008;9:R137. <https://doi.org/10.1186/gb-2008-9-9-r137>
56. Kent WJ, Zweig AS, Barber G *et al.* BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 2010;26:2204–7. <https://doi.org/10.1093/bioinformatics/btq351>
57. Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;38:576–89. <https://doi.org/10.1016/j.molcel.2010.05.004>
58. Luo Y, Hitz BC, Gabdank I *et al.* New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res* 2020;48:D882–9. <https://doi.org/10.1093/nar/gkz1062>
59. Wu P, Li T, Li R *et al.* 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. *Nat Commun* 2017;8:1937. <https://doi.org/10.1038/s41467-017-01793-w>
60. Krietenstein N, Abraham S, Venev SV *et al.* Ultrastructural details of mammalian chromosome architecture. *Mol Cell* 2020;78:554–65. <https://doi.org/10.1016/j.molcel.2020.03.003>
61. Nora EP, Lajoie BR, Schulz EG *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 2012;485:381–5. <https://doi.org/10.1038/nature11049>
62. Greenwald WW, Li H, Benaglio P *et al.* Subtle changes in chromatin loop contact propensity are associated with differential gene regulation and expression. *Nat Commun* 2019;10:1054. <https://doi.org/10.1038/s41467-019-08940-5>
63. Rashi-Elkeles S, Warnatz HJ, Elkon R *et al.* Parallel profiling of the transcriptome, cistrome, and epigenome in the cellular response to ionizing radiation. *Sci Signal* 2014;7:rs3. <https://doi.org/10.1126/scisignal.2005032>
64. Iwanaszko M, Kimmel M. NF-kappaB and IRF pathways: cross-regulation on target genes promoter level. *BMC Genomics* 2015;16:307. <https://doi.org/10.1186/s12864-015-1511-7>
65. Babeu JP, Boudreau F. Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks. *World J Gastroenterol* 2014;20:22–30. <https://doi.org/10.3748/wjg.v20.i1.22>
66. Wang Z, Coban B, Wu H *et al.* GRHL2-controlled gene expression networks in luminal breast cancer. *Cell Commun Signal* 2023;21:15. <https://doi.org/10.1186/s12964-022-01029-5>
67. Campbell CE, Piper M, Plachez C *et al.* The transcription factor Nfix is essential for normal brain development. *BMC Dev Biol* 2008;8:52. <https://doi.org/10.1186/1471-213X-8-52>

68. Allahyar A, Vermeulen C, Bouwman BAM *et al.* Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet* 2018;50:1151–60. <https://doi.org/10.1038/s41588-018-0161-5>
69. Uyehara CM, Apostolou E. 3D enhancer–promoter interactions and multi-connected hubs: organizational principles and functional roles. *Cell Rep* 2023;42:112068. <https://doi.org/10.1016/j.celrep.2023.112068>
70. Oudelaar AM, Davies JOJ, Hanssen LLP *et al.* Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat Genet* 2018;50:1744–51. <https://doi.org/10.1038/s41588-018-0253-2>
71. Karr JP, Ferrie JJ, Tjian R *et al.* The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer–promoter communication. *Genes Dev* 2022;36:7–16. <https://doi.org/10.1101/gad.349160.121>
72. Dornan D, Shimizu H, Perkins ND *et al.* DNA-dependent acetylation of p53 by the transcription coactivator p300. *J Biol Chem* 2003;278:13431–41. <https://doi.org/10.1074/jbc.M211460200>
73. Ceskova P, Chichger H, Wallace M *et al.* On the mechanism of sequence-specific DNA-dependent acetylation of p53: the acetylation motif is exposed upon DNA binding. *J Mol Biol* 2006;357:442–56. <https://doi.org/10.1016/j.jmb.2005.12.026>
74. Golov AK, Gavrillov AA, Kaplan N *et al.* A genome-wide nucleosome-resolution map of promoter-centered interactions in human cells corroborates the enhancer–promoter looping model. *eLife* 2024;12:RP91596. <https://doi.org/10.7554/eLife.91596.3>
75. Hsieh TS, Cattoglio C, Slobodyanyuk E *et al.* Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol Cell* 2020;78:539–53. <https://doi.org/10.1016/j.molcel.2020.03.002>
76. Hua P, Badat M, Hanssen LLP *et al.* Defining genome architecture at base-pair resolution. *Nature* 2021;595:125–9. <https://doi.org/10.1038/s41586-021-03639-4>
77. Hait TA, Maron-Katz A, Sagir D *et al.* The EXPANDER integrated Platform for transcriptome analysis. *J Mol Biol* 2019;431:2398–406. <https://doi.org/10.1016/j.jmb.2019.05.013>
78. Kruse K, Hug CB, Vaquerizas JM. FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biol* 2020;21:303. <https://doi.org/10.1186/s13059-020-02215-9>
79. Yu G, Wang LG, Han Y *et al.* clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284–7. <https://doi.org/10.1089/omi.2011.0118>