Article

# TP63 mediates the generation of tumour-specific chromatin loops that underlie MYC activation in radiation-induced tumorigenesis

Qiaowei Liu[1,2,7], Shuai Jiang[3,7], Xiang Xu[4,7], Kang Xu[4], Yawen Luo[4], Zongyuan Yu[4], Meida Xiang[4], Zhouyang Xu [2], Lijie Wang[1], Sujie Zhang[1], Haitao Tao[1], Xuhui Yang[1], Chongbo Jiang[1], Xiaoguang Qi[1], Hao Li [2,4], Yiming Gao[1,5], Yao Li[1,5], Shihui Fu [6], Pingkun Zhou[2], Xiaochen Bo [4], Hebing Chen [4] ✉, Xiaohua Chen[2] ✉ & Yi Hu [1,5] ✉

Alterations in 3D chromatin conformation may disrupt the interplay between promoters and distal enhancers. How gene regulatory circuits are reshaped during ionizing radiation-induced tumorigenesis remains unclear, and little is known about the mediators that drive these processes. To decipher the chromatin alterations in radiation-induced lung cancer, we performed ATAC-seq, RNA-seq and Hi-C analyses of human bronchial epithelial cells and corresponding radiation-induced malignantly transformed cell lines. We found that this malignant transformation is accompanied by chromatin switching from the inactive B compartment to the active A compartment, an increased number of TADs and gained ATAC-seq peaks that mediate new distal chromatin contacts. We identified tumour protein 63 (TP63) as a mediator of new chromatin-accessible sites that anchor tumour-specific chromatin contacts in radiation-induced tumour cells. A TP63-mediated accessible chromatin site anchors a tumour-specific TAD boundary and multiple tumour-specific chromatin loops, which might underlie *MYC* oncogene activation during malignant transformation.

Tumour cells undergo considerable rewiring of regulatory networks during tumorigenesis, which leads to the activation of oncogenes and the repression of tumour suppressors[1-4]. The new networks include silent genes that might be antagonistic to the starting cell type and thus reside in chromatin that is inaccessible to many transcription factors (TFs) and the transcriptional machinery[5]. Chromatin is the essential medium through which transcription factors, signalling pathways, and related cues alter gene activity and cellular phenotypes[2]. Alterations in the multilevel chromatin architecture have been characterized as hallmarks of tumorigenesis[6,7], and these alterations might occur when cells are exposed to ionizing radiation, which is a well-recognized cause of lung cancer[8]. Chromatin alteration, the direct consequence of the erroneous processing of DNA double-strand breaks (DSBs) induced by ionizing radiation, results in gene network

[1]Department of Oncology, the Fifth Medical Center, Chinese PLA General Hospital, Beijing 100071, China. [2]Beijing Institute of Radiation Medicine, Beijing 100850, P. R. China. [3]Nanhu Laboratory, State Key Laboratory of Biomedical Analysis (SKLBA, formerly known as National Center of Biomedical Analysis, NCBA), Beijing 100039, P.R. China. [4]Academy of Military Medical Sciences, Beijing 100850, P. R. China. [5]Medical School of Chinese PLA, Beijing 100853, P. R. China. [6]Department of Cardiology, Hainan Hospital of Chinese People's Liberation Army General Hospital, Sanya 572013, P. R. China. [7]These authors contributed equally: Qiaowei Liu, Shuai Jiang, Xiang Xu. ✉e-mail: chb-1012@163.com; chenxh0676@sina.com; huyi301zlxb@sina.com

transitions and might lead to tumorigenesis[9,10]. However, the mechanism through which local chromatin is opened to enable cooperative events for new genetic networks during tumorigenesis induced by ionizing radiation remains unclear, and little is known about the mediators that drive these processes and their functional impacts on tumorigenesis.

The human genome is organized in three-dimensional (3D) space in the nucleus and exhibits a multilevel 3D chromatin architecture that plays an important role in the transcriptional control of genes[11,12]. In recent years, the high-throughput sequencing of chromosome conformation capture (Hi-C) approach has been introduced as a method for identifying higher-order chromatin interactions throughout the genome[13]. Interphase chromatin, which rapidly decondenses into structures that allow regulated access to functional elements and thus enables genome functions such as transcription, gene silencing, and DNA replication[14], is linearly partitioned into megabase-long regions spatially segregated into active 'A' compartments and inactive 'B' compartments. The 'A' compartments correspond to gene-rich regions and are enriched in active chromatin marks, whereas the 'B' compartments are gene-poor regions that display features of heterochromatin; Switching between the 'A' and 'B' compartments is related to alterations in gene transcriptional activity during development and in various diseases[15–17]. Identification of compartments using principal component analysis (PCA) and high-resolution Hi-C data binned at 10-50 kb suggests that compartments can be as small as a single active or inactive locus[18]. Recently, the mammalian genome has been divided into at least six subcompartments with various combinations of repressive and active chromatin modifications[19]. At the submegabase scale, consecutive regions are spatially organized into self-interacting domains called topologically associating domains (TADs), which are both structural and functional units of the human genome that are frequently delineated by CTCF-binding insulators that can block the function of enhancers or the spread of gene silencing[13,20]. Genes are often coexpressed within each TAD[21,22], and concordant increases and decreases in enhancer activities are characteristics of TADs in senescent cells[23]. In mammals, an additional level of TAD organization involves dynamic cohesin-dependent loops between CTCF-binding sites at convergent orientations[21]. Recent single-cell Hi-C results and Oligopaint experiments coupled with superresolution microscopy have confirmed that TAD structures are present in single cells[24–26]. Studies focused on developmental processes have revealed a convincing link between TAD boundary disruption, ectopic enhancer-promoter interactions and alterations in gene expression levels[27–29]. At the finest scale, chromatin loops emerge between linearly distant genomic regions with unusually high spatial contact frequencies[19]. Chromatin loops often connect distal regulatory elements such as enhancers and promoters and facilitate enhancer selection via their target genes[30–32]. The formation and maintenance of hierarchical chromatin structure in the nucleus depend on but are not limited to loop extrusion and compartmentalization[33]. Because the spatial folding of chromatin orchestrates the gene expression network, genome conformation is an integral part of the chromatin landscape that TFs must navigate to perform their gene regulatory functions[34].

This multilevel 3D chromatin architecture is under rigorous spatial and temporal control during embryonic development and lineage commitment[15,20,35,36], and perturbation of this architecture might lead to aberrant gene regulation and cell dysfunction[37]. Disruption of the 3D chromatin architecture has been identified in various diseases, such as cancers, including prostate cancer[38], childhood acute lymphoblastic leukaemia (ALL)[39], colorectal cancer[37], and acinic cell carcinoma[40]. TADs are genomic regulatory units with a high degree of internal stability that can be remodelled by structural genomic variations[41]. Widespread structural variation events in cancer genomes result in novel three-dimensional chromatin structural domains through alteration of the boundaries[42]. Deletion or inactivation of a TAD

boundary can lead to inappropriate oncogene expression[43], and the juxtaposition of active regulatory sequences to known oncogenes can also contribute to tumorigenesis[42]. Each cell expresses a specific subset of all genes, and the genome conformation is partly cell-type-specific and plays an important role in cell fate decisions[34,43]. A subset of fate-changing TFs termed pioneer factors can scan and target DNA sequences and mediate cooperative events that can open local chromatin and allow the binding of other TFs to induce cell-type-specific gene signatures[5,44]. Therefore, the specific pattern of chromatin conformation in radiation-induced tumour cells needs to be identified, and the underlying mediators need to be determined to further understand the mechanisms through which ionizing radiation leads to lung cancer.

For the current study, we employed Hi-C, ATAC-seq and RNA-seq to decipher the chromatin alterations in radiation-induced lung cancer and determine the underlying mediators. We developed a high-resolution map of 3D chromatin structures by using Hi-C in BEP2D cells, a human bronchial epithelial cell line, and their radiation-induced tumour cell counterparts to reveal alterations in chromatin structure at different scales, ranging from megabase-scale A/B compartments to the finest-scale chromatin loops anchored at specific genes. ATAC-seq was used to map genome-wide chromatin accessibility, revealing that chromatin contacts are enhanced and reinforced at tumour-specific open chromatin sites. By examining the predicted binding sites of 746 TFs, we identified tumour protein 63 (TP63) as a mediator that underlies the reinforcement of functional chromatin contacts in radiation-induced tumour cells. Via combined analysis of the putative TP63 binding sites with publicly available TP63 ChIP-seq datasets of lung squamous cell carcinoma (LUSC)[45], we showed that TP63 binds to multiple tumour-specific open chromatin sites that collectively anchor a tumour-specific hub of chromatin loops that underlie *MYC* oncogene activation in radiation-induced tumour cells. By comprehensive profiling of the transcriptome and TF-binding events in wild-type BEP2D cells, BEP2D cells overexpressing TP63 and malignantly transformed BERP35T1 cells using RNA-seq, ATAC-seq and CUT&Tag targeting TP63, CTCF, and STAG2, we demonstrated that overexpression of TP63 in BEP2D cells led to the formation of novel accessible chromatin sites and CTCF-binding sites and increased cohesin levels at TP63-binding sites. Micro-C profiling of these cells further demonstrated that these epigenomic alterations induced by TP63 binding led to enhanced chromatin contacts and elevated expression levels of corresponding anchor genes functionally enriched in biological processes related to tumorigenesis. We conclude that TP63 might act as a pioneer factor mediating chromatin structure remodelling in radiation-induced lung cancer and activating the *MYC* oncogene by orchestrating a tumour-specific gene expression network.

## Results

### Ionizing radiation-induced tumour cells exhibit aberrant gene expression and pervasive accumulation of accessible chromatin sites

To elucidate the interplay between chromatin accessibility and chromatin structure and their regulatory role in radiation-induced lung cancer, we established two malignantly transformed cell lines, BERP35T1 and BERP35T4, which were derived from BEP2D cells (a previously established human papillomavirus (HPV18)-immortalized human bronchial epithelial cell line[46]) irradiated with 1.5 Gy of α-particles emitted from a $^{238}$Pu source as previously described in detail[47].

We subsequently performed RNA-seq and ATAC-seq to obtain atlases of the transcriptome and epigenome, respectively, of BEP2D, BERP35T1 and BERP35T4 cells. First, we identified alterations in gene transcriptional activity between normal and tumour cells with RNA-seq. Specifically, we identified 414 significantly upregulated and 204 significantly downregulated genes in BERP35T1 cells and

396 significantly upregulated and 284 significantly downregulated genes in BERP35T4 cells compared with BEP2D cells (Fig. 1A). Most of the differentially expressed genes (70.52% of all upregulated genes and 55.41% of all downregulated genes) in BERP35T1 and BERP35T4 cells were shared (Fig. S1A, B). These results reveal a similarity in the transcriptional alterations between the two radiation-induced tumour cell lines.

To investigate the biological functions of the differentially expressed genes in ionizing radiation-induced malignantly transformed cell lines, we performed Gene Ontology (GO) pathway analysis of the upregulated genes in BERP35T1 and BERP35T4 cells (Fig. 1B). We noted that the upregulated genes in both cell lines were enriched in the angiogenesis pathway. Previous studies have shown that angiogenesis plays a crucial role in tumour growth, invasion, and metastasis[48]. This characteristic, although observed in cells cultured

in vitro, indicates that malignantly transformed cells may have a potential role in promoting angiogenesis. Cell adhesion is essential for the development and structural integrity of multicellular tissues and is a key attribute of tumour cells[49,50]. The maintenance of epithelial cell–cell adhesion is vital for both organ function and tumour formation[51]. The epidermal development pathway shares numerous genes with tumorigenesis[52], and any disruption in the balance between cellular renewal and differentiation can result in uncontrolled growth and tumour initiation[53,54]. These findings suggest that ionizing radiation-induced malignantly transformed cell lines might share similar pathways with cancer cells. Our previous study has highlighted significant morphological differences between BEP2D cells and ionizing radiation-induced malignantly transformed cell lines[55]. Our findings indicate that the cell adhesion and cell-cell adhesion properties of malignantly transformed cell lines are
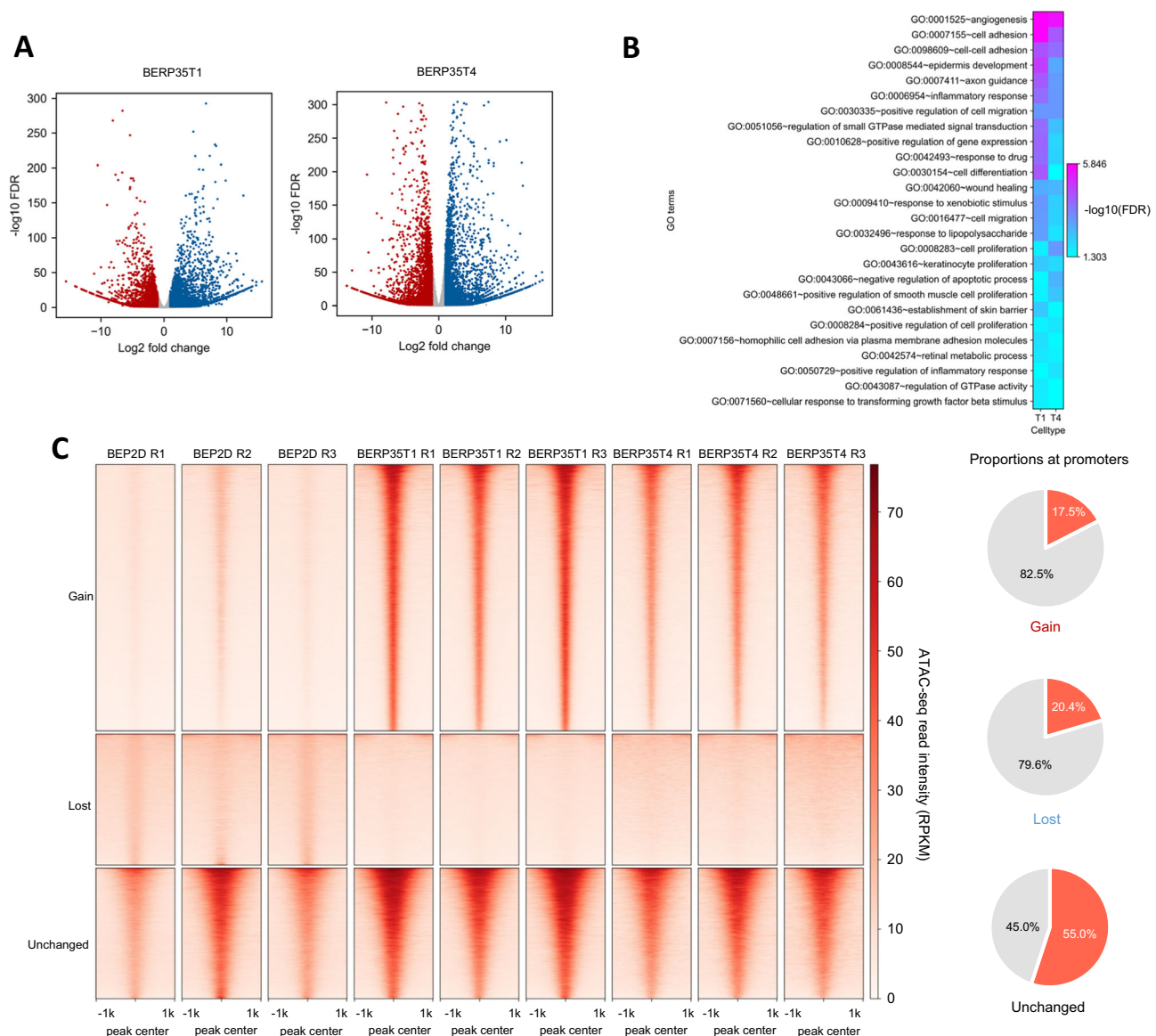


**Fig. 1 | Ionizing radiation-induced tumour cells exhibit aberrant gene expression and pervasive accumulation of accessible chromatin sites. A** Volcano plots showing genes that are up- and downregulated in BERP35T1 and BERP35T4 tumour cells compared to BEP2D cells. **B** Heatmap showing significantly enriched GO terms of up-regulated genes in BERP35T1 and BERP35T4 tumour cells. **C** Heatmaps showing the read intensities of 2-kb regions centered on ATAC-seq peaks in normal BEP2D cells and tumour cells. All ATAC-seq peaks were categorized into 3 types: gained in tumour cells, lost in tumour cells, and unchanged (from top to bottom). All replicates were shown in one heatmap with the same order of all genes. The pie charts show the proportions of ATAC-seq peaks located at gene promoters of each type of ATAC-seq peak (orange: ATAC-seq peaks at promoters, ATAC-seq peaks not at promoters).

considerably stronger than those of normal cells. This is evidenced by the notably prolonged digestion time of malignant cells.

To identify the regulatory basis of aberrant gene expression in tumour cells, we analysed the chromatin accessibility profiles of these 3 cell types with ATAC-seq. We identified 14220 and 5360 gained ATAC-seq peaks in BERP35T1 and BERP35T4 cells, respectively, compared with BEP2D cells. However, we identified only 832 and 1043 ATAC-seq peaks in BEP2D cells that were lost in BERP35T1 and BERP35T4 cells, respectively (Fig. 1C). Interestingly, only 2.3% of the gained ATAC-seq peaks in BERP35T1 cells and 2.9% of the gained ATAC-seq peaks in BERP35T4 cells were located in promoter regions, whereas 22.6% and 29.9% of ATAC-seq peaks that were unchanged between normal cells and tumour cells colocalized with promoters in BERP35T1 and BERP35T4 cells, respectively (Fig. 1C). Together, these results indicate an overall increase in the number of accessible chromatin sites in the genomes of radiation-induced BEP2D cell-derived tumour cells and suggest that, compared with genomic sites with stable accessibility, genomic sites with increased accessibility in these tumour cells are located primarily in nonpromoter regions.

## A/B compartment switching and alterations in intra-TAD structures are associated with dysregulated gene expression in ionizing radiation-induced tumour cells

Accessible chromatin sites in nonpromoter regions delineate distal regulatory elements such as enhancers, and their interplay with the 3D chromatin architecture is the primary factor underlying their regulatory function[38,56–58]. To identify potential alterations in the chromatin architecture of radiation-induced tumour cells at different scales, we performed a Hi-C analysis of BEP2D cells and two radiation-induced tumour cell lines (BERP35T1 and BERP35T4). At the largest scale, interphase chromatin is spatially segregated into the 'A' and 'B' compartments, and we comprehensively analysed the Hi-C data of these cell lines to characterize potential alterations in the A/B compartments in the interphase chromatin architecture after the radiation-induced malignant transformation of BEP2D cells. The 'A' and 'B' compartments were annotated in consecutive 200-kb bins by principal component analysis (PCA). To avoid the potential influence of translocations on A/B compartment annotation, we used the PC1 value generated by PCA to exclude chromatin that harboured large-scale translocations and that showed weak correlations between the PC1 value and gene density in downstream analyses of the 'A' and 'B' compartments. We adopted a matrix balance approach to minimize technical bias in the data processing pipeline to avoid the potential influence on our analysis. Our results revealed that 9.7% of the analysed bins corresponded to B-to-A compartment switching during the transformation of normal BEP2D cells to BERP35T1 tumour cells, whereas only 2.4% of the analysed bins corresponded to A-to-B compartment switching (Fig. 2A, left panel). Similar results were observed with BERP35T4 cells: 7.5% and 2.5% of the bins corresponded to B-to-A and A-to-B compartment switching, respectively (Fig. 2A, right panel). And we also identified 665 consistent B-to-A bins in BERP35T1 and BERP35T4 (Fig. 2B). Consistent with previous reports[59], genes located in genomic regions corresponding to B-to-A compartment switching tended to be upregulated ($P$ value = 9.5×10$^{-22}$ for BERP35T1, $P$ value = 6.3×10$^{-18}$ for BERP35T4), whereas genes in regions corresponding to A-to-B compartment switching tended to be downregulated ($P$ value = 1.2×10$^{-4}$ for BERP35T1, $P$ value = 0.004 for BERP35T4, Fig. 2C). These results suggest an overall imbalance in compartment switching biased towards the active 'A' compartments associated with increased transcriptional activity in radiation-induced tumour cells.

To explore the biological significance of genes involved in the B-to-A switching, we performed GO pathway analysis based on these genes (Fig. S2A). Notably, we noticed that genes located in B-to-A switching regions are enriched in keratinization-related pathways. It is reported that keratinization is one of the hallmark histopathology features of non-small cell lung cancer and keratinization of lung

squamous cell carcinoma is associated with poor clinical outcome[60]. We also observed cellular phenotype changes in BERP35T1 and BERP35T4 cell lines[55] which are both lung squamous cell carcinoma cell lines. This suggested that genes located in B-to-A switching regions are associated with keratinization.

Although the positions of TADs are generally conserved in different cell types[61], numerous studies have also shown that the intra-TAD structure can be significantly changed in cells with alterations in enhancer activity and gene expression[62–64]. Additionally, alterations in TAD structures have been identified in various cancers, but their functional impact remains largely uncharacterized. To investigate alterations in TADs and their potential regulatory functions during the radiation-induced malignant transformation of BEP2D cells, we first identified 3348 TADs in BEP2D cells and 3974 and 3746 TADs in radiation-induced BERP35T1 and BERP35T4 tumour cells, respectively, at 20-kb resolution (Fig. S2B). These results are consistent with previous studies that identified more TADs in tumour cells than in normal cells by Hi-C analysis[16,65]. Despite the increased number of TADs in tumour cells, the genomic positions of most TAD boundaries were conserved between BEP2D cells and both radiation-induced tumour cell lines (Fig. S2C). To investigate whether conserved TADs undergo structural and functional alterations during radiation-induced tumorigenesis, we compared the TAD boundaries in BEP2D, BERP35T1 and BERP35T4 cells and identified conserved TAD boundaries that were not shifted by more than 1 bin in either BERP35T1 or BERP35T4 tumour cells compared with BEP2D cells. A TAD was identified as a conserved TAD among all three cell types when both TAD boundaries were conserved. We identified 1874 TADs that were conserved among all three cell types (Supplementary Data 1). We previously showed that the spatial density of open chromatin (SDOC) is a key regulatory property associated with the TAD level[63]. In brief, SDOC quantitatively measures the level of spatial aggregation of active regulatory elements and reflects the regulatory properties of TADs. A higher level of SDOC indicates a more active regulatory state within the TAD, whereas a lower level of SDOC indicates a repressive regulatory state. To investigate the impact of TAD alterations on gene expression levels within these TADs, we calculated the SDOC level for the conserved TADs. Indeed, we identified 296 TADs with increased SDOC and 316 TADs with decreased SDOC in BERP35T1 tumour cells compared with BEP2D cells and identified 276 TADs with increased SDOC and 258 TADs with decreased SDOC in BERP35T4 tumour cells compared with BEP2D cells (Supplementary Data 2). In addition, genes located within TADs with increased SDOC tended to be upregulated (Fig. 2D), whereas the transcriptional activity of genes located within TADs with decreased SDOC tended to be repressed, during the radiation-induced malignant transformation of BEP2D cells (Fig. 2D). We also found that TADs with increased SDOC were more enriched in regions undergoing B-to-A compartment switching in radiation-induced tumour cells (Fig. 2E), and both were associated with increased gene transcriptional activity. The SDOC alteration in conserved TADs indicates functional changes in the chromatin state and structure of TADs and might underlie the dysregulation of gene transcription within these TADs in radiation-induced tumour cells. We demonstrated that genomic regions with compartment switching are associated with alterations in chromatin accessibility and transcriptional activity (Fig. 2F, S2D).

Together, these results indicate that the chromatin architecture is altered at both the compartment level and the TAD level and that the changes at both scales are associated with alterations in gene expression in radiation-induced tumour cells.

## Dynamic changes in chromatin contact anchor gene expression in SDOC-increased TADs

The significant association between alterations in SDOC and gene expression inspired us to further investigate the characteristics and potential regulatory impact of chromatin alterations in TADs with
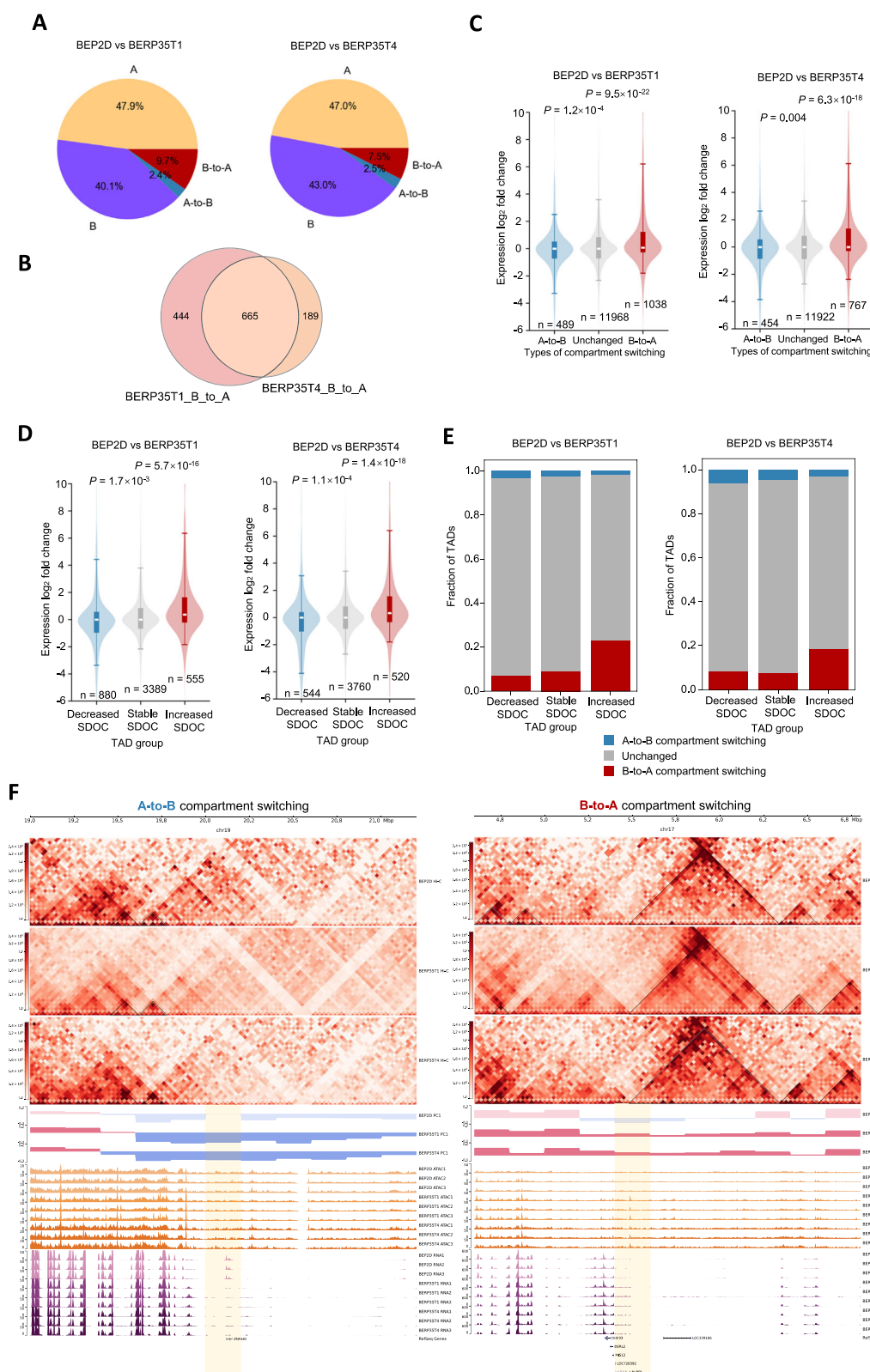
altered SDOC levels during radiation-induced malignant transformation. We first assessed the level of alterations in contact frequencies within the 3 groups of TADs classified according to the types of SDOC alterations during tumorigenesis.

We found that alterations in contact frequency are significantly greater in TADs with increased levels of SDOC than in TADs with stable or decreased levels of SDOC. This observation is consistent when

comparing both BERP35T1 cells with BEP2D cells and BERP35T4 cells with BEP2D cells (Fig. 3A). As an increased SDOC value indicates a greater spatial chromatin contact frequency, these results indicate that increased contact frequencies in a subset of TADs may play a key role in establishing regulatory circuits within these TADs, shaping a more active regulatory state during radiation-induced malignant transformation of BEP2D cells.

**Fig. 2 | A/B compartments switching and alterations in intra-TAD structures are associated with dysregulated gene expression in the ionizing radiation-induced tumour cells. A** Pie charts showing the proportions of each type of locus (stable compartment **A**, **A**-to-**B** switching, **B**-to-**A** switching and stable compartment **B**) comparing BEP2D with BERP35T1 and BERP35T4 respectively. **B** Overlap of the compartments from **B**-to-**A** in BERP35T1 and BERP35T4. **C** Violin plots showing the log2 fold changes in the expression of genes located within each type of locus (**A**-to-**B** compartment switching, unchanged compartment, and **B**-to-**A** compartment switching). *P*-values from a two-sided t-test between groups are shown. Box limits represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Source data are provided as a Source Data file.

**D** Boxplots showing the log2 fold changes in the expression of genes located within each type of TAD (from left to right: decreased SDOC in tumour, unchanged SDOC in tumour, increased SDOC in tumour). *P*-values from a two-sided t-test between groups are shown. Box limits represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Source data are provided as a Source Data file. **E** Stacked bar plots showing the fractions of the 3 groups of TADs in regions with an **A**-to-**B** compartment switching and regions with a **B**-to-**A** compartment switching. **F** Examples of **A/B** compartment switching with chromatin accessibility and gene expression tracks, where the shadings correspond to genomic regions with compartment switching that associates with alterations in chromatin accessibility and transcriptional activity.
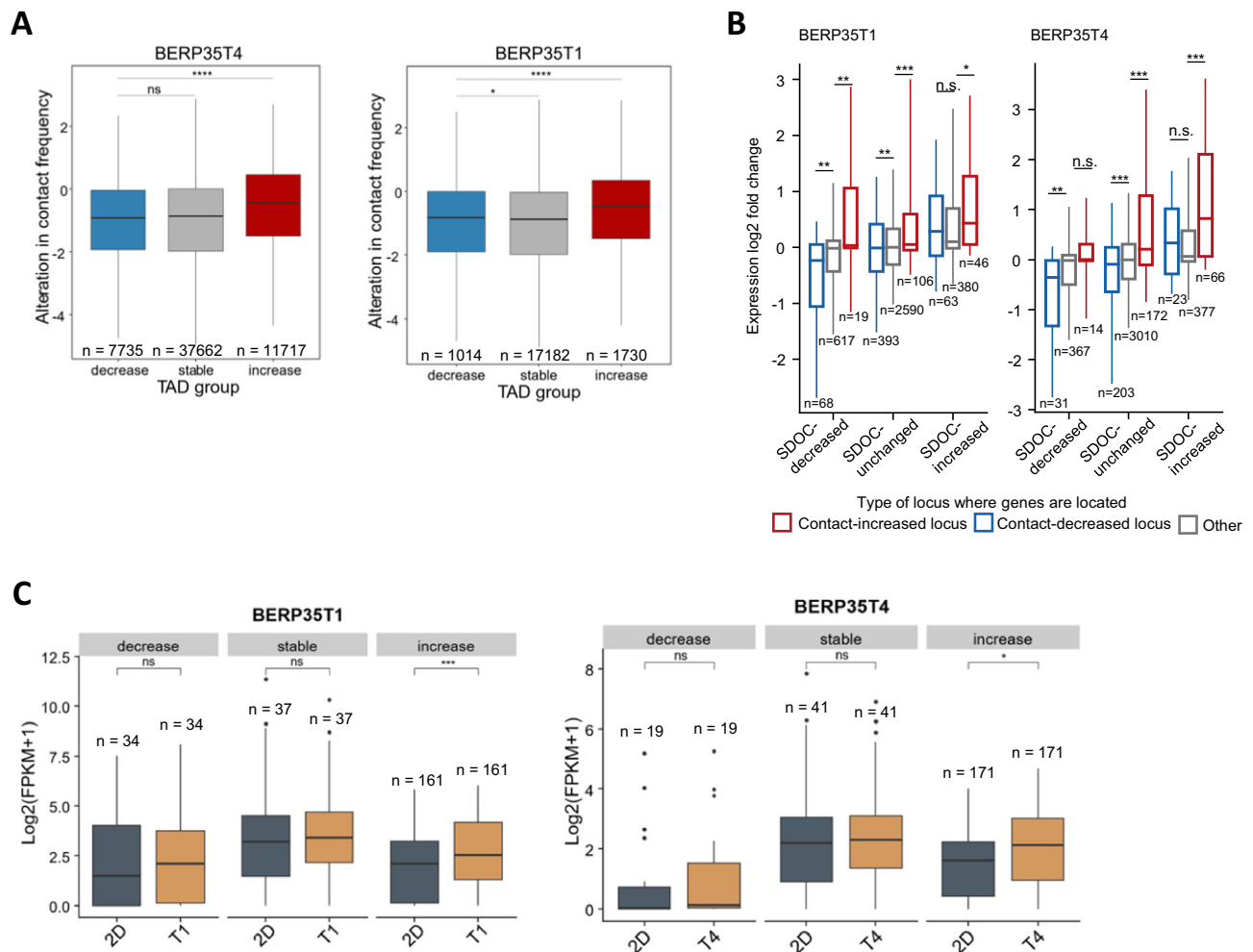


**Fig. 3 | Dynamic changes in different SDOC during the process of radiation-induced tumor formation. A** The contact frequency alterations within SDOC-decreased, SDOC-stable, and SDOC-increased TADs in BERP35T4 (left) and BERP35T1 (right) cells. Box limits represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Significance: n.s. not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-sided t-test. Source data are provided as a Source Data file. **B** Boxplots showing the log2 fold changes in the expression of genes located at the 3 types of loci (contact-increased, contact-decreased and other) within the 3 groups of TADs (SDOC increased in tumour cells, SDOC unchanged in tumour cells and SDOC decreased in tumour cells). Box limits

represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Significance: n.s. not significant, *P < 0.05, ***P < 0.001, two-sided t-test. Source data are provided as a Source Data file. **C** Boxplots showing the log2(FPKM + 1) in the expression of oncogenes located at the 3 groups of TADs (SDOC increased in tumour cells, SDOC unchanged in tumour cells and SDOC decreased in tumour cells). Box limits represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Significance: n.s. not significant, *P < 0.05, **P < 0.01, ***P < 0.001, two-sided t-test. Source data are provided as a Source Data file.

We then sought to further investigate whether increased chromatin contact frequencies might contribute to aberrant gene activation in the BERP35T1 and BERP35T4 malignantly transformed cell lines. To this end, we assigned each gene to a locus pair if the transcription start site (TSS) of the gene overlapped with either anchor of the locus

pair and defined them as "anchor genes". All anchor genes were then categorized into nine groups on the basis of the SDOC alteration of their corresponding TADs and alterations in the contact frequency of the locus pair. Finally, alterations in the expression of anchor genes were compared between groups. Overall, genes located at an anchor of

enhanced distal contact, which was defined as a locus pair with a differential contact frequency > 1, were generally upregulated regardless of whether the SDOC value of the corresponding TAD was altered, and genes located at an anchor of weakened distal contact, which was defined as a locus pair with a differential contact frequency < -1, were generally downregulated regardless of whether the SDOC level of the corresponding TAD was altered (Fig. 3B). The statistically non-significant result for TADs with decreased SDOC in BERP35T4 cells was likely due to the small sample size. Therefore, alterations in the transcriptional activity of anchor genes might be associated primarily with alterations in the contact frequencies where they are located, where alterations in the overall TAD regulatory context may play a less significant role.

Then we checked whether the SDOC increase events are consistent between the BERP35T1 and BERP35T4 cell lines (Fig. S3A). We observed 202 overlapping SDOC increase events between BERP35T1 and BERP35T4. Further GO pathway analysis showed that genes located in SDOC increase regions were enriched in desmosome-related pathways (Fig. S3B). Recent studies using mouse genetic approaches have uncovered a role for desmosomes in tumour suppression, demonstrating that desmosome downregulation occurs before that of adherens junctions to drive tumour development and early invasion[66].

On the basis of the above results, we speculated that during radiation-induced malignant transformation of BEP2D cells, alterations in TADs as well as intra-TAD chromatin contacts may contribute to the abnormal expression of oncogenes. Therefore, we next investigated the contribution of SDOC alterations to oncogene expression. Briefly, we compiled the oncogene list[67–71] and combined the gene expression levels by calculating the mean expression within each TAD to enable comparison of TAD-level gene expression in different groups of TADs. Our results revealed that TADs with increased SDOC correspond to significantly increased expression of oncogenes during tumorigenesis (two-sided t-test, Fig. 3C), suggesting that SDOC alteration may indeed contribute to oncogene activation during radiation-mediated malignant transformation of BEP2D cells.

Together, our results suggest that the dynamic changes in SDOC values during radiation-induced tumorigenesis may be closely related to the dysregulation of gene expression, which may be accompanied by specific binding of transcription factors.

## TP63 is a mediator of tumour-specific chromatin contact

As a variety of TFs have been shown to contribute to the establishment of cell-type-specific chromatin contacts that regulate the expression of cell identity genes[34,72,73], we hypothesize that a key regulatory TF may drive alterations in SDOC levels within TADs and mediate altered gene transcription in radiation-mediated malignant transformation. We first downloaded the JASPAR CORE Position Frequency Matrix (PFM) of Vertebrates from the JASPAR database, which represents 746 non-redundant TF binding motifs[74]. We scanned all accessible chromatin regions in each cell line with these PFMs to identify putative TF-binding sites (TFBSs) using the FIMO package from the MEME suite[75]. To rule out potential false-positive "bystander" TFs that are not functionally significant but "happen" to be located at accessible chromatin sites, we assessed the gene expression level of each assessed TF using RNA-seq data to exclude those that presented decreased expression levels in tumour cells. We then calculated the Pearson correlation coefficients between the alterations in the number of binding sites for each TF in each TAD and the changes in SDOC levels in each TAD, as a higher correlation indicates a stronger association between the alteration in TF binding and the alteration in the SDOC level. Our results revealed that the top candidate TFs that had the most significantly increased expression levels and were positively correlated with SDOC alterations in both BERP35T1 cells and BERP35T4 cells were IRF6, STAT3, FOSL1, and TP63 (Fig. 4A), suggesting that these transcription factors are most likely to be associated with changes in SDOC during the radiation-

mediated malignant transformation of BEP2D cells. Notably, among these TFs, TP63 was most significantly activated in both BERP35T1 and BERP35T4 cells, suggesting its potential impact in reshaping the structural and regulatory states of TADs during malignant transformation of BEP2D cells.

On the basis of the observation of increased ATAC-seq peaks in both BERP35T1 and BERP35T4 malignant cells compared with BEP2D cells, we further hypothesized that certain TFs may be involved in mediating tumour-specific chromatin contacts via the generation of novel chromatin sites. We then identified potential TFs responsible for the generation of novel accessible chromatin sites in radiation-induced tumour cells by comparing the chromatin accessibility levels in normal BEP2D cells and tumour cells at all putative binding sites for each TF and measured the level of increase in chromatin accessibility by calculating the fold change in the mean ATAC-seq read density between normal cells and tumour cells. We also took the expression level of each TF into consideration to rule out potential false-positive results. Overall, multiple TFs presented high increases in chromatin accessibility ( > 3-fold) in BERP35T1 and relatively weak but consistent increases in BERP35T4. Among the analysed TFs, TP63 presented the greatest increase in chromatin accessibility in both BERP35T1 and BERP35T4 radiation-induced tumour cells compared with BEP2D cells, and it was also the only TF that presented a more than 3-fold increase in chromatin accessibility in both tumour cell lines (Fig. 4A). Additionally, TP63 was only weakly expressed in normal BEP2D cells but was significantly upregulated in both tumour cell lines. Notably, TP63 presented the highest degree of transcriptional activation among all the TF genes in both tumour cell lines (Fig. 4B). Although gained ATAC-seq peaks in tumour cells were enriched in loci that anchor regions of increased spatial chromatin contact (contact-increased loci), a minority of gained ATAC-seq peaks were located within anchor loci exhibiting weakened chromatin contact (contact-decreased loci) in tumour cells. Interestingly, the TFBSs of TP63 were preferentially located at gained ATAC-seq peaks within contact-increased loci rather than at gained ATAC-seq peaks within contact-decreased loci (Fig. 4C), indicating that TP63 preferentially binds to anchors with increased spatial chromatin contact in addition to gained ATAC-seq peaks in tumour cells. We also analysed the binding sites of other TFs with increased accessibility in tumour cells (SOX15, SOX2, FOS, FOSL1, JUNB, FOXA2, FOXE1 and BARX2), and most of these sites did not exhibit preferential localization at gained ATAC-seq peaks within contact-increased loci compared with those within contact-decreased loci (Fig. S4A). Therefore, TP63 is a tumour-specific TF that specifically binds to tumour-specific ATAC-seq peaks that anchor tumour-specific chromatin contacts in BERP35T1 and BERP35T4 radiation-induced tumour cells.

Histopathological analysis has indicated that BERP35T1 and BERP35T4 radiation-induced tumour cells originate from squamous cell carcinoma[47]. To further support the role of TP63 as a potential mediator of tumour-specific chromatin contact, we used a public TP63 ChIP-seq dataset of an LUSC cell line to investigate the preferential binding of TP63 at loci with increased contact frequencies in tumour cells. The analysis of a public ChIP-seq dataset for TP63 in LUSC cell lines[45] revealed that TP63-occupied ATAC-seq peaks were preferentially located within loci that exhibited increased contact frequency with multiple other distal loci (Fig. 4D). Because the mean ATAC-seq read density of SOX2 TFBSs in BERP35T1 tumour cells was more than 3-fold greater than that in BEP2D cells (Fig. 4B), we analysed the SOX2 ChIP-seq dataset provided in the abovementioned study[23]. Unlike for TP63, we found no preferential localization in contact-increased loci for SOX2 (Fig. S4B).

Together, these results show that tumour-specific opening of chromatin regions mediated by TP63 might underlie the formation and reinforcement of distal chromatin contacts in BERP35T1 and BERP35T4 tumour cells.
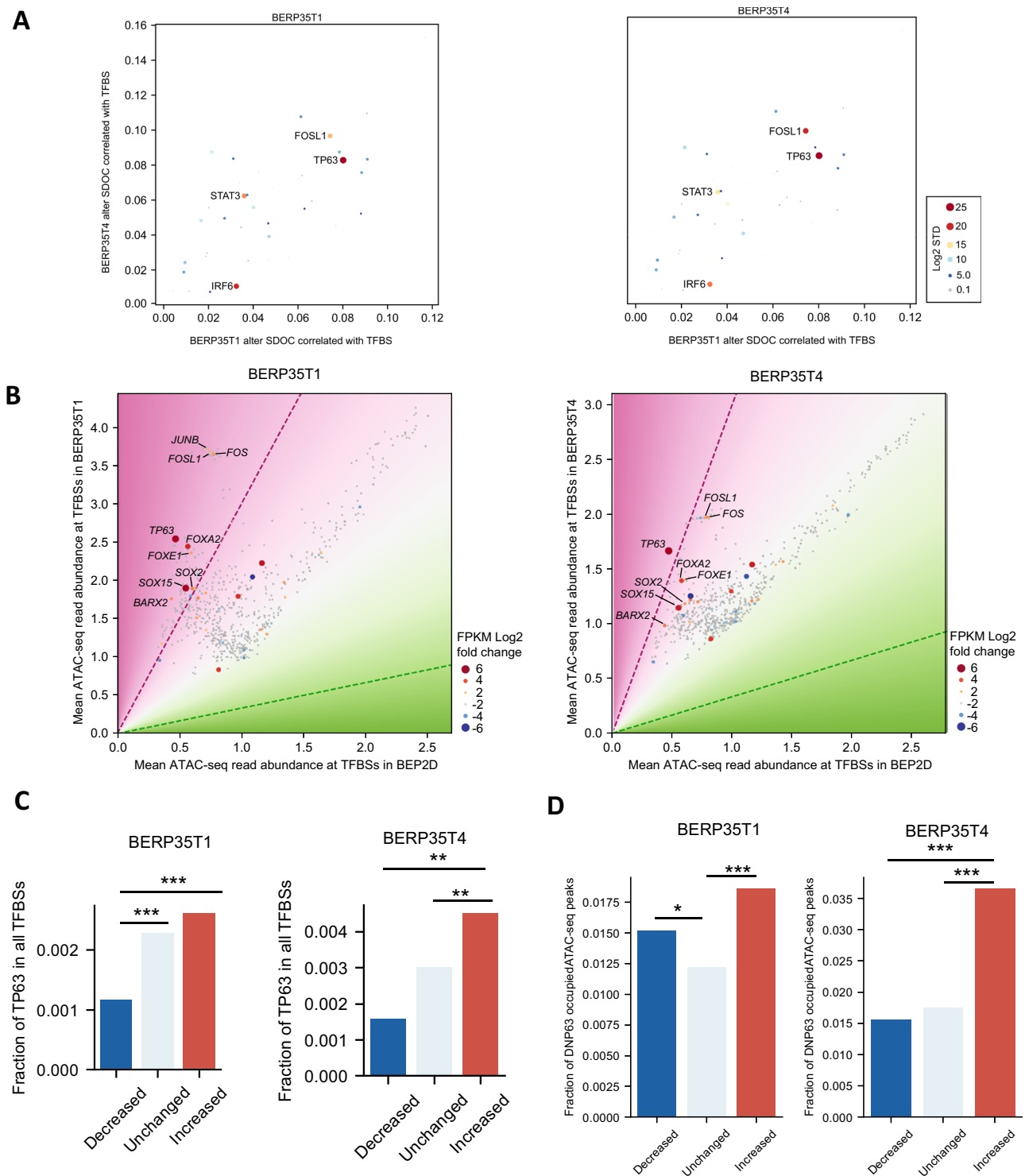
**Fig. 4 | TP63 is a mediator of tumour-specific chromatin contact. A** Scatter diagram of SDOC alteration correlated with TFBS in BERP35T1 and BERP35T4. The color and size represent log2 STD of gene expression alteration in BERP35T1 (left) and BERP35T4 (right). **B** Average chromatin accessibility levels at binding sites of each TF in normal BEP2D cells (x-axis) and tumour cells (y-axis) measured as reads per kilobase of transcript per million reads mapped (RPKM). The pink and green dashed lines correspond to 3-fold increases and 3-fold decreases, respectively, in ATAC-seq read abundances in tumor cells. The names of TFs that were significantly upregulated and showed an increase in the ATAC-seq read abundance of more than 3-fold in either tumour cell line are indicated. FPKM log2 fold change of each differentially expressed TF were marked by different sizes and color. **C** Bar plots showing the fractions of putative TP63 binding sites in all putative TFBSs located

within the 3 groups of gained-in-tumour ATAC-seq peaks (Left bar in each panel: ATAC-seq peaks gained in tumour cells and located at contact-decreased loci; middle bar in each panel: ATAC-seq peaks gained in tumour cells and located at contact-unchanged loci; right bar in each panel: ATAC-seq peaks gained in tumour cells and located at contact-increased loci). Significance: **$P < 0.01$, ***$P < 0.001$, two-sided Fisher's exact test. **D** Bar plots showing the fractions of ATAC-seq peaks colocalized with P63 ChIP-seq peaks in the 3 groups of loci. Loci were grouped by the number of other distal loci that showed concomitantly increased, unchanged or decreased contact frequencies in tumour cells. ChIP-seq data is from Sato et al. (Cancer Research, 79(24), 6084–6100.). Significance: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, two-sided Fisher's exact test.

## TP63 anchors multiple tumour-specific chromatin loops that instruct *MYC* oncogene activation in radiation-induced tumour cells

To gain further insight into the functional role of the reinforced chromatin contact anchored by gained ATAC-seq peaks in radiation-induced tumour cells, we performed a genome-wide search for differentially expressed genes anchored by increased distal chromatin contact in both the BERP35T1 and BERP35T4 tumour cell lines. We identified 48 genes, of which 46 were upregulated in both tumour cell lines, which is consistent with our findings that locus pairs with an increased contact frequency in tumour cells are associated with increased transcriptional activities of anchor genes. The 46 upregulated anchor genes, including the known oncogene *MYC* (Supplementary Data 3), whose expression was 5-fold greater in both BERP35T1 and BERP35T4 tumour cells than in BEP2D cells (Fig. 5A), are located at a tumour-specific hub of chromatin interactions anchored by multiple upstream genomic regions that harbour ATAC-seq peaks gained in tumour cells (Fig. 5B, C). One of these gained ATAC-seq peaks is located within a tumour-specific TAD boundary ~340 kb upstream of the adjacent TAD boundary that contains the *MYC* promoter. Taken together, these two adjacent TAD boundaries delineate a tumour-specific TAD and form a tumour-specific chromatin loop at the apex of the TAD in both BERP35T1 and BERP35T4 cells. In addition, both anchor loci of this tumour-specific chromatin loop contain a tumour-specific ATAC-seq peak that colocalizes with a TP63 ChIP-seq peak in the LK2 LUSC cell line (Fig. 5C). In this example, the upregulation of the oncogene *MYC* was associated with the construction of a tumour-specific chromatin contact via TP63-occupied tumour-specific open chromatin sites, and this result is consistent with the findings from our genome-wide analysis, which suggested that TP63 mediates the formation of open chromatin regions that underlie functional chromatin contacts in radiation-induced BERP35T1 and BERP35T4 tumour cells.

## Validation of TP63-mediated chromatin contacts and transcriptional regulatory roles

To validate the role of TP63 in mediating novel chromatin contacts, we constructed a TP63-overexpressing BEP2D cell line to investigate the functional impact of TP63 (Methods). For comprehensive investigation of the impact of the chromatin architecture, epigenome and transcriptome of TP63, we performed Micro-C on WT BEP2D (BEP2DNC) cells, TP63-overexpressing BEP2D (BEP2DOE) cells and BERP35T1 cells and probed the transcriptome, chromatin accessibility, TP63 and cohesin binding and chromatin structure by profiling these cell lines using RNA-seq, ATAC-seq, and CUT&Tag targeting TP63, STAG2 and CTCF. The RNA-seq data revealed significantly greater expression of the MYC gene in BEP2DOE cells than in BEP2DNC cells ($p$ value = $2.2 \times 10^{-43}$; Fig. 6A). The increased expression of MYC by TP63 overexpression in BEP2D cells was further verified using qPCR and Western blot (WB) analysis (Fig. 6B, S5A). These results revealed a causal link between TP63 overexpression and increased expression of the oncogene MYC.

We also found that the profile of gene expression alterations in TP63-overexpressing BEP2D cells compared with BEP2DNC cells significantly correlated with those observed in BERP35T1 cells compared with BEP2DNC cells ($p$ value < $10^{-300}$, Fig. S5B), indicating that the overexpression of TP63 resulted in a significant shift in the transcriptome towards that of malignant BERP35T1 cells. Moreover, there was a significantly greater proportion of overlapping upregulated genes compared with the rate by chance, and there was a significantly greater proportion of overlapping downregulated genes compared with the rate by chance, although to a lesser extent than upregulated genes (Fisher's exact test, p = 8.9e-23, odds ratio = 10.35; Fig. 6C). Importantly, GO analysis revealed that the overlapping upregulated genes were most significantly enriched in biological processes related to negative regulation of apoptotic processes, positive regulation of

mitotic cell division and positive regulation of cell proliferation, suggesting positive regulatory effects in tumorigenesis (Fig. S5C). Similarly, the most significantly enriched biological process among the overlapping downregulated genes was the apoptotic process (Fig. S5D). These results showed that the transcriptional alteration induced by TP63 overexpression represents a significant part of the aberrant transcriptome alteration in transformed BERP35T1 cells and is functionally related to malignant transformation.

To validate the role of TP63 in mediating chromatin reorganization and chromatin contact in the malignant transformation of BEP2D cells, we performed integrative analysis of our CUT&Tag and ATAC-seq data and assessed the potential impact of TP63 binding on the level of chromatin accessibility as well as the binding of CTCF and cohesin at the corresponding TP63 binding sites. We first noted that the number of TP63 CUT&Tag peaks (hereafter referred to as TP63 peaks) with increased chromatin accessibility was greater than that with decreased chromatin accessibility, suggesting an overall increase in the chromatin accessibility level at TP63 peaks, but the level of chromatin accessibility alteration varied at different TP63 peaks. To enable a detailed and context-dependent assessment of the regulatory impact of TP63 chromatin binding events, we further classified TP63 peaks into four categories according to alterations in chromatin accessibility with/without TP63 binding by comparing TP63-overexpressing BEP2D cells with WT BEP2D cells. Specifically, we classified the TP63 peaks into 4 groups: strongly increased (fold change≥4), slightly increased (2≤fold change<4), not significantly changed (0.5<fold change<2) and decreased (fold change≤0.5). We observed that the "strongly increased" group corresponded to the strongest TP63 peaks (fold change = 39.23), indicating that stronger and more robust TP63 binding results in a greater increase in chromatin accessibility. These results further support TP63 as a pioneer factor mediating novel accessible chromatin sites during the malignant transformation of BEP2D cells. We next explored how the four types of TP63 peaks affect CTCF and cohesin binding (Fig. 6D, S5E). As expected, the increase in chromatin accessibility at TP63 peaks is associated with increased CTCF and cohesin binding. This suggested that TP63 is involved in the establishment of novel CTCF bindings by creating accessible chromatin sites. Moreover, these novel CTCF peaks might be capable of blocking cohesin[76], resulting in the enrichment of cohesin, and favouring the establishment of chromatin contacts on these sites.

On the basis of the above observations, we further explored the potential of TP63 binding to mediate chromatin contacts. We first studied the expression alterations of genes located in loop anchors (Fig. S5F). The expression of genes located at TP63-bound loop anchors was significantly greater than that of TP63-bound loop anchor genes but not loop anchor genes (p = 0.029) or other genes (p = 0.011), indicating the regulatory potential of TP63-anchored chromatin contacts. Next, we investigated the impact of TP63 on chromatin interactions across the whole genome, which was not limited to loop anchors. As weaker TP63 peaks only presented moderate alterations in chromatin accessibility, CTCF binding intensity and cohesin level, we focused on the most significantly enriched TP63 peaks (n = 4000). As TP63 peaks with a greater degree of increase in chromatin accessibility are more likely to recruit CTCF and cohesin (Fig. 6D, S5E), we inferred that these TP63 binding sites are more likely to mediate enhanced chromatin interactions. Therefore, on the basis of the extent of increased chromatin accessibility at novel TP63 peaks, we divided these peaks into two classes: those with a fold change in chromatin accessibility greater than or equal to 3 were defined as class 1 TP63 peaks, and those with a less prominent increase in chromatin accessibility were defined as class 2 TP63 peaks as internal controls. We then compared the intensities of chromatin contacts mediated by at least one class 1 TP63 peak with the intensities of those mediated by class 2 TP63 peaks alone in the BEP2DNC, BEP2DOE and BERP35T1 cell lines
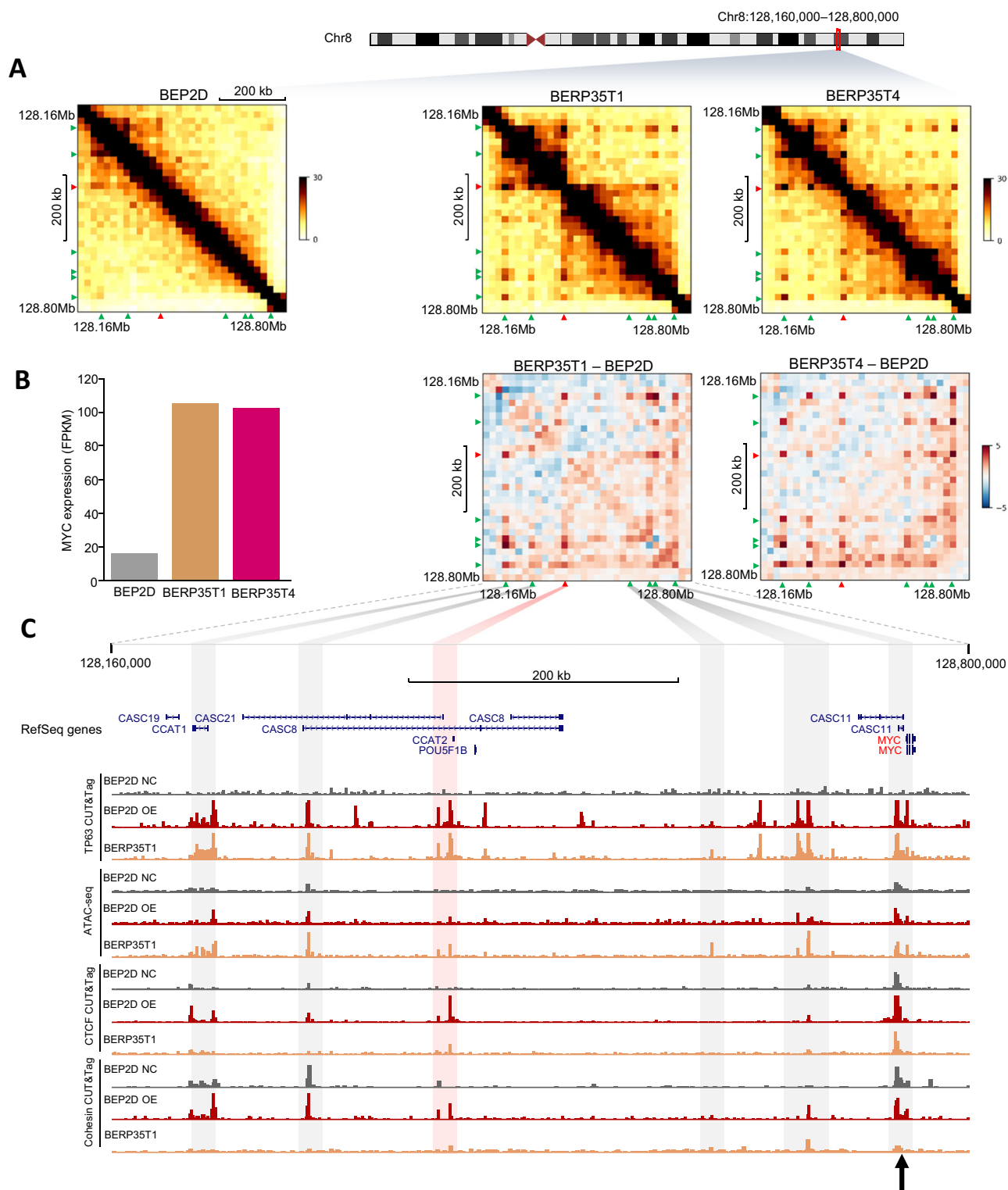
**Fig. 5 | TP63 anchors multiple tumour-specific chromatin loops that instruct MYC oncogene activation in radiation-induced tumour cells. A** Hi-C contact maps of the MYC locus in BEP2D, BERP35T1 and BERP35T4 cells. Differential contact maps between each type of tumour cells and normal BEP2D cells are shown below the Hi-C contact maps of the corresponding tumour types. The red arrows mark the locus where a new TAD boundary in each radiation-induced tumor cells was established. The green arrows mark the anchors of the tumour-specific hub of chromatin loops. **B** Expression level of MYC in BEP2D, BERP35T1 and BERP35T4 cells. **C** Genome-browser like plot of the MYC locus showing RefSeq genes, ATAC-seq, TP63, CTCF, Cohesin CUT&Tag tracks of BEP2DNC, BEP2DOE and BERP35T1 cell lines.

(Fig. 6E, S5G). We observed that class 1 TP63 peaks mediated interactions that were significantly stronger in BEP2DOE and BERP35T1 cell lines (APA score = 1.91 in BEP2DOE cells; APA score = 1.90 in BERP35T1 cells) than in BEP2DNC cells (APA score = 1.83), whereas class

2 TP63 peaks mediated interactions that showed unchanged intensities (APA score = 2.12, 2.04 and 2.18 in BEP2DNC, BEP2DOE and BERP35T1 cells, respectively). These results suggest that TP63 binding sites that mediate increased chromatin accessibility are likely to
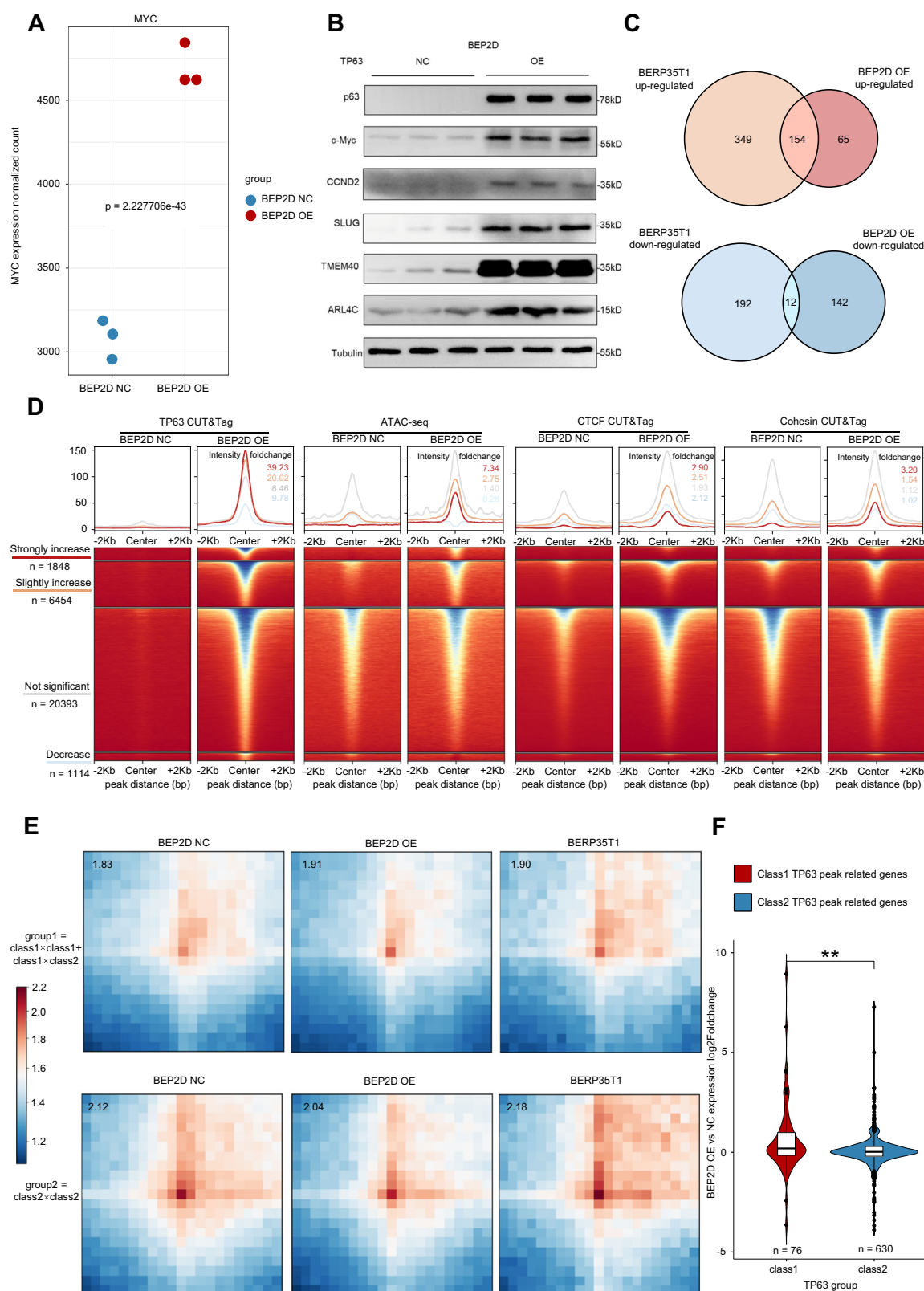
mediate enhanced chromatin interactions. This finding is consistent with the observation that TP63 peaks with increased chromatin accessibility presented increased CTCF binding.

Next, we explored whether increased interaction intensities led to the activation of gene expression. We compared the gene expression fold changes at class 1 TP63 peaks and class 2 TP63 peaks (Fig. 6F, S5H). As expected, genes located at class 1 TP63 peaks presented

significantly greater increases in expression than did genes located at class 2 TP63 peaks.

Compared with those in BEP2DNC cells, we observed novel chromatin contacts at similar genomic sites in BEP2DOE cells and BERP35T1 cells upstream of MYC gene (Fig. S6, S7). These novel chromatin contacts were anchored by chromatin sites with clear TP63 CUT&Tag peaks (Fig. S6). Notably, we observed novel CTCF peaks that

**Fig. 6 | Effects of TP63 overexpression on gene expression, chromatin accessibility and chromatin structure. A** MYC expression difference before and after TP63 overexpression. pValue is calculated using DEseq2. **B** Western blot of MYC before and after TP63 overexpression. The experiment was repeated three times independently with similar results. Uncropped and unprocessed scans of all blots are provided as a Source Data file. **C** Venn diagram shows the overlap between BERP35T1 up-regulated genes and BEP2D OE up-regulated genes. **D** TP63 CUT&Tag, ATAC-seq, CTCF CUT&Tag and Cohesin CUT&Tag heatmap before and after TP63 overexpression. According to ATCA-seq signal foldchange, TP63 peaks were divided into 4 groups, including strongly increase (foldchange≥4), slightly increase (2≤foldchange<4), not significant (0.5<foldchange<2) and decrease (foldchange≤0.5). **E** APA of contact mediated by different group of TP63 peaks. **F** Gene expression foldchange mediated by different group of TP63 peaks. Box limits represent 25th percentile, median and 75th percentile. Whiskers represent 5th and 95th percentile. Points represent outliers beyond whiskers. Significance: \*\**P* < 0.01, two-sided Wilcox-test. Source data are provided as a Source Data file.

colocalized with novel cohesin peaks in a subset of these clear TP63 peaks (n = 2, Fig. S6). One of the novel CTCF peaks was located ~340 kb upstream of the MYC promoter and corresponded to a novel TAD boundary in both BEP2DOE cells and BERP35T1 cells but not in BEP2DNC cells. Importantly, we also observed that the corner of this novel TAD corresponds to a chromatin loop that connects the novel TAD boundary and the MYC gene promoter. Importantly, this observation is consistent with the novel TAD boundary shown in our Hi-C results (Fig. 5), suggesting that this novel TAD boundary is mediated by TP63 expression during radiation-mediated malignant transformation of BEP2D cells, resulting in activated MYC expression.

Together, these results validated the role of TP63 in mediating the opening of chromatin at binding sites, establishing novel CTCF binding sites, thereby potentially mediating chromatin contacts via elevated levels of cohesin blockade, favouring chromatin loop extrusion and resulting in activated gene expression. These epigenome and 3D genome alterations mediated by overexpressed TP63 contributed to the aberrant transcriptome and the activation of pathways related to the malignant transformation of BEP2D cells.

## Discussion

Although alterations in the multilevel chromatin architecture have been characterized as hallmarks of tumorigenesis[6,7], little is known about the mediators that drive these processes and their functional impacts on tumorigenesis. To fill this research gap and reveal the mechanism of radiation carcinogenesis, we performed Hi-C, ATAC-seq, and RNA-seq analyses of the α particle-induced carcinogenesis of human bronchial epithelial cells and then performed a multi-omic integrative analysis.

First, we showed that the chromatin accessibility landscape undergoes pervasive alterations before and after tumorigenesis. The landscape of chromatin accessibility broadly reflects the regulatory capacity and is a critical determinant of chromatin organization and function[77]. Previous work on cutaneous T-cell lymphoma (CTCL) revealed the accessibility signature of oncogenic dysregulation, with implications for therapeutic intervention[78]. In our radiation carcinogenesis model, we found a general tendency towards increased chromatin accessibility at both open and closed chromatin sites in BERP35T1 and BERP35T4 tumour cells. Unlike chromatin sites with stable accessibility levels, which were enriched in gene promoters, very few chromatin sites with increased accessibility in tumour cells were colocalized with promoter regions, which indicated that most remodelling events occur at promoter-distal sites.

Second, we also showed that the remodelling of chromatin accessibility is closely related to finer-scale alterations in the 3D chromatin structure. Chromatin accessibility is determined by the linear structure of chromatin at the nucleosomal length scale, and higher-order nucleosome organization might also impact access to DNA[77]; however, few studies have investigated the relationship between chromatin accessibility and chromatin structure in tumorigenesis. We found that radiation-induced tumorigenesis significantly increased the frequency of chromatin contacts between genomic sites with increased accessibility and significantly decreased the frequency of chromatin contacts between sites with decreased accessibility. The greatest alteration in the chromatin contact frequency was observed

between two distal sites with simultaneously increased or decreased chromatin accessibility. An increased contact frequency between distal loci often corresponded to increased transcriptional activity of genes located at either of the anchor loci, which indicated a regulatory role of chromatin structure alterations in radiation-induced tumorigenesis.

Third, we identified TP63 as a mediator of functional chromatin contact in BERP35T1 and BERP35T4 tumour cells. We propose that TP63 functions as a pioneer factor that creates accessible chromatin sites that recruit other TFs, including known insulators such as CTCF and YY1[30], thereby blocking chromatin loop extrusion and creating new functional chromatin loops. We showed that TP63 might participate in the formation of new chromatin interaction hubs that anchor the *MYC* oncogene to multiple distal tumour-specific accessible regions, which might underlie the activation of *MYC* in BERP35T1 and BERP35T4 tumour cells. Previous studies have shown that tumour cells harbour more TADs than do cells of normal origin[16,38]. Chromatin loop extrusion has been shown to underlie TAD formation[79]. In accordance with previous studies, we detected an overwhelmingly greater number of gained ATAC-seq peaks than lost peaks in radiation-induced tumour cells. These tumour-specific open chromatin regions might awaken insulators that are repressed in normal cells and thus result in the creation of new chromatin loops and TADs in tumour cells. In support of this hypothesis, we identified a CTCF binding motif ~340 kb upstream of the *MYC* promoter that colocalized with a gained ATAC-seq peak in both BERP35T1 and BERP35T4 tumour cells; in addition, a TP63 binding motif was found in this region. As TP63 was activated in both BERP35T1 and BERP35T4 tumour cells, we concluded that it might function as a tumour-specific pioneer factor that creates a functional insulator at this site, which in turn underlies the formation of a new TAD boundary in tumour cells. Indeed, TP63 has been shown to establish a squamous subtype-specific enhancer landscape in squamous cell carcinomas (SCCs) and has been shown to regulate *MYC* expression[80]. TP63 has also been shown to be a pioneer factor in zebrafish[44]. To further support these conclusions via bioinformatic analysis of putative TFBS predictions, we generated a TP63-overexpressing BEP2D cell line and conducted a series of sequencing experiments to generate a comprehensive map of WT BEP2D cells, TP63-overexpressing BEP2D cells and BERP35T1 malignant cells, including transcriptome, epigenome and 3D genome datasets. We performed integrative analysis and provided robust results showing that TP63 binding indeed leads to genome-wide alterations in chromatin accessibility, CTCF binding, and chromatin contacts at corresponding sites. Using high-resolution Micro-C data, we also replicated the observation of Hi-C data on the oncogene *MYC* locus. Specifically, we proved that in both TP63-overexpressing BEP2D cells and BERP35T1 cells, a novel TP63 binding site ~320 kb upstream of the *MYC* promoter mediates a new accessible chromatin site and results in the establishment of a colocalized CTCF binding site. This TP63-mediated CTCF binding site mediates a novel TAD boundary as well as a newly formed chromatin loop that anchors the *MYC* promoter on the other side, leading to elevated *MYC* expression. Thus, our study further extends our understanding of the regulatory function of chromatin architecture alterations and the underlying role of TP63 in radiation-induced lung tumorigenesis.

Radiation has been reported to have a limited effect on 3D chromatin organization, with the main effect being the strengthening of TAD boundaries in repair-proficient cells[81]. They also reported that repair-deficient cells undergo little change, indicating the presence of a mechanism that protects 3D genome structure integrity during the process of DNA damage repair. In contrast to previous work, our study demonstrated alterations in multiple chromatin structural levels, including compartments, TADs and intra-TAD chromatin contacts, in radiation-induced malignantly transformed BEP2D cells (BERP35T1 and BERP35T4 cells) compared with normal BEP2D cells. The potential reasons for the differences in the observations could be multifactorial. First, the difference in cell type may play an important role. While a previous study investigated fibroblasts and lymphoblastoid cells, our study focused on epithelial cells, which might have different chromatin states and susceptibilities to radiation-induced alterations that facilitate malignant transformation. Second, the focus of the previous study was mainly on the immediate responses to radiation, while our work suggests that long-term chromatin remodelling may emerge beyond the immediate timeframe. Finally, previous analyses of cells revealed little evidence of malignant transformation upon exposure to radiation, where alterations in chromatin contact may reflect mainly the process of DNA damage repair. Our work focused on malignantly transformed cells, where alterations in chromatin contact have been shown to facilitate transcriptome alterations related to tumorigenesis.

Our work revealed that the role of TP63 in mediating chromatin contacts is functionally related to the regulation of oncogene expression. As TP63 is specifically activated in lung squamous cell carcinoma (LUSC), our findings may provide new insight into the potential regulatory mechanisms underlying the development of LUSC. Furthermore, the comprehensive multi-omic dataset provided in our study could be a valuable resource for studying the regulation and function of 3D chromatin structure in the process of malignant transformation.

In summary, this study, by integrating genome-wide high-resolution Hi-C data, chromatin accessibility profiling data and gene expression data, identified TP63 as a pioneer factor that facilitates remodelling of the 3D genomic architecture and might activate *MYC* expression by mediating tumour-specific chromatin interaction hubs in a model of radiation-induced lung cancer. These findings increase our understanding of the radiation-induced malignant transformation of bronchial epithelial cells, which might aid in the clinical treatment of and drug development for lung cancer. Our results are based on a cell model and thus should be validated via replication in animal models and patients in future studies.

## Methods

### Cell culture
The cell lines used in this study were described in detail in a previous paper[47,55]. Briefly, the malignantly transformed cell lines BERP35T1 and BERP35T4 were derived from BEP2D cells irradiated with 1.5 Gy of α-particles emitted from a $^{238}$Pu source and were described in detail in a previous paper[47]. The cells were cultured in serum-free LHC-8 medium (Gibco, USA) at 37 °C in an atmosphere containing 95% air and 5% $CO_2$. The culture medium was changed every 3 days. When the cells reached 90% confluency, they were detached with trypsin (Sigma–Aldrich, USA) for 5 min, and the suspension was separated by centrifugation to collect the cells. The cells were then subjected to subculturing. For this study, cells at passages 48-49 were used.

### Lentivirus transfection
NP63-overexpressing lentivirus (HBLV-h-TP63-3xflag-ZsGreen-PURO) and the corresponding control lentivirus (HBLV-ZsGreen-PURO) were purchased from Hanbio Biotechnology Co. Ltd. (Shanghai). For lentivirus transfection, $2 \times 10^5$ BEP2D cells were seeded in 12-well plates in LHC-8 medium overnight and infected with either a lentivirus expressing the target gene or the control lentivirus. After 16 h, the medium was replaced with fresh medium. After 48 h, LHC-8 medium supplemented with 1 µg/ml puromycin was used to select puromycin-resistant cells. After 2 weeks, the protein expression level of TP63 in the cells was determined by Western blot analysis.

### Western blot
Western blot analysis was performed via a standard technique. An anti-GAPDH antibody (#5174S, 1:2000, Cell Signaling Technology) was used as a control for whole-cell lysates. Immunoblotting was performed with anti-NP63 (#67825S, 1:1000, Cell Signaling Technology), anti-c-Myc (#18583S, 1:1000, Cell Signaling Technology), anti-SLUG (#9585S, 1:1000, Cell Signaling Technology), anti-CCND2 (#10934-1-AP, 1:1000, Proteintech), anti-TMEM40 (#sc-393601, 1:1000, Santa Cruz), anti-ARL4C (#10934-1-AP, 1:1000, Proteintech), 10202-1-AP Anti-rabbit IgG, and HRP-linked antibodies (#7074, 1:5000, Cell Signaling Technology).

### RNA-seq
Total RNA was extracted from $1 \times 10^6$ cells with RNAiso Reagent (TaKaRa, China) following the manufacturer's instructions. RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was checked using a NanoPhotometer® spectrophotometer (IMPLEN, USA). The RNA concentration was measured using a Qubit® RNA Assay Kit with a Qubit® 2.0 fluorometer (Life Technologies, USA). RNA integrity was assessed using an RNA Nano 6000 Assay Kit in a 2100 Bioanalyzer system (Agilent Technologies, USA). A total of 1 µg of RNA per sample was used as input material for RNA sample preparation. The sequencing libraries were generated using an NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations, and in-house index codes (Supplementary Data 4) were added to attribute sequences to each sample. In brief, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations at an elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5×). First-strand cDNA was synthesized using random hexamer primers and M-MuLV Reverse Transcriptase (RNase H-). Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. The remaining overhangs were converted to blunt ends via exonuclease/polymerase activity. After adenylation of the 3′ ends of the DNA fragments, NEBNext adaptors with hairpin loop structures were ligated in preparation for hybridization. To preferentially select cDNA fragments with lengths of 150 - 200 base pairs (bp), the library fragments were purified with the AMPure XP system (Beckman Coulter, USA). Then, 3 µl of USER Enzyme (NEB, USA) was incubated with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. Then, PCR was performed with Phusion High-Fidelity DNA polymerase, universal PCR primers and the Index (X) Primer. Finally, the PCR products were purified (AMPure XP system), and the library quality was assessed with the Agilent 2100 Bioanalyzer system (3 nM). Clustering of the index-coded samples was performed in a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were pooled together and sequenced on the Illumina HiSeq X Ten platform, and 2×150 bp paired-end reads were generated. Library construction and sequencing were accomplished with the help of Annoroad Gene Technology (Beijing, China).

### ATAC-seq
ATAC-seq was performed as previously described[82]. Fifty thousand cells were collected and centrifuged. The cell pellets were resuspended in lysis buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.5% NP-40) and immediately centrifuged at $500 \times g$ and 4 °C for 10 min. The supernatants were removed, and the pellets (nuclear fraction) were resuspended in a solution from a TruePrep™ DNA Library Prep Kit V2 for Illumina® (Vazyme, China) for 30 min at 37 °C.

DNA was purified using VAHTS DNA Clean Beads (Vazyme, China). Following the manufacturer's protocol, the libraries were amplified with a TruePrep™ Index Kit V2 for Illumina® (Vazyme, China). Libraries were analysed for size distribution with the Agilent 2100 Bioanalyzer system and quantified by real-time PCR. Finally, the high-quality DNA libraries were pooled together and sequenced on an Illumina NovaSeq platform, and 2×150 bp paired-end reads were generated. Library sequencing was accomplished by Berry Genomics Corporation (Beijing, China).

### Cleavage under targets and tagmentation (CUT&Tag) assay

CUT&Tag was performed according to the instructions of the Hyperactive® Universal CUT&Tag Assay Kit for Illumina (TD903, Vazyme). Briefly, 100,000 sorted cells were mixed with Concanavalin A-coated magnetic beads and permeabilized with digitonin. The cell-bead complexes were incubated with the primary antibody at 4 °C overnight, followed by incubation with the secondary antibody for 1 h at room temperature. After washing, the cells were incubated at room temperature for one hour with pAG-Tn5. The nuclei were collected and subjected to a transposition reaction at 37 °C for 1 h with a Bio-Rad T100 Thermal Cycler. Then, the transposed fragments were extracted, purified and amplified by using the reagents provided by the kit. Libraries were analysed for size distribution with the Agilent 2100 Bioanalyzer system. Finally, the libraries were pooled together and sequenced on an Illumina NovaSeq 6000 platform, and 2×150 bp paired-end reads were generated.

The antibodies used in this study for CUT&Tag were specific for NP63 (#67825S, 1:50, Cell Signaling Technology), STAG2 (#5882S, 1:50, Cell Signaling Technology), CTCF (#3418S, 1:50, Cell Signaling Technology), Goat Anti-Rabbit IgG H&L (#Ab207, 1:100, Vazyme).

### Hi-C

Approximately $2 \times 10^7$ cells were chemically crosslinked by the addition of a 1/36 volume of fresh 37% formaldehyde solution to the medium and incubation for 10 min at room temperature with gentle shaking. Crosslinking was stopped by adding 2.5 M glycine at a final concentration of 0.2 M and incubating for 10 min at room temperature. After being rinsed twice with PBS, the cells were harvested in a 1.5 ml tube by scraping and centrifugation and stored at −80 °C until use. The fixed cells were resuspended in 1 ml of lysis buffer (10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 0.2% Igepal CA-630, and 1/10 vol. of proteinase inhibitor cocktail (Sigma, USA)) and then incubated on ice for 20 min. The nuclei were pelleted by centrifugation at 4 °C and 600× g for 5 min, washed with 1 ml of lysis buffer and subjected to another centrifugation step under similar conditions. After washing twice with restriction enzyme buffer, the nuclei were resuspended in 400 μl of restriction enzyme buffer and transferred to a safe-lock tube. Next, the chromatin was solubilized with diluted SDS and incubated at 65 °C for 10 min. After the SDS was quenched with Triton X-100, the chromatin was digested overnight with a 4-cutter restriction enzyme (400 units of MboI) at 37 °C on a rocking platform. The subsequent steps were conducted as previously described[19]; the methods used were specific to Hi-C and included labelling of the DNA ends with biotin-14-dCTP and performing blunt-end ligation of crosslinked fragments. The proximal chromatin DNA was religated with a T4 DNA ligase (Thermo Fisher, USA). The nuclear complexes were reverse crosslinked by incubation with proteinase K at 65 °C. DNA was purified via phenol–chloroform extraction. Biotin-C was removed from nonligated fragment ends using T4 DNA polymerase. Fragments were sheared to a size of 200–600 bp by sonication with a Bioruptor for 90 seconds and subjected to end repair with a mixture of T4 DNA polymerase, T4 polynucleotide kinase and Klenow DNA polymerase. Biotin-labelled Hi-C samples were specifically enriched using streptavidin C1 magnetic beads (Thermo Fisher, USA). Fragment ends were subjected to A-tailing with Klenow fragments (exo-), and Illumina paired-end

sequencing adaptors (NEBNext Ultra DNA Library Prep Kit for Illumina) were then added to the ligation mixture. Finally, the Hi-C libraries were amplified via 12–14 cycles of PCR by using universal PCR primers according to the instructions of Illumina and sequenced on the Illumina HiSeq platform. Sequence interaction patterns were obtained on the Illumina HiSeq X Ten platform with 2×150 bp reads. Library construction and sequencing were accomplished with the help of Annoroad Gene Technology (Beijing, China).

### Micro-C-XL

Micro-C-XL libraries were constructed according to previously established methods[83,84]. Briefly, 107 cells were cross-linked for 15 min with 1% FA at room temperature (RT) and then quenched with 0.125 M glycine for 5 min. FA-fixed nuclei were subjected to additional cross-linking with freshly prepared 3 mM EGS for 40 min at RT. FA + EGS dual cross-linked cells were quenched with 0.4 M glycine for 10 min and then washed once with PBS supplemented with 0.05% BSA. Next, the cells were lysed, and the chromatin was fragmented with the appropriate amount of MNase for 10 min at 37 °C. The digestion was terminated by the addition of 1.5 mM EGTA at 65 °C for 10 min. The 3′ ends of the DNA were dephosphorylated, and the 5′ ends were phosphorylated using 20 units of T4 polynucleotide kinase (NEB, #M0201L) at 37 °C for 15 min. The DNA overhangs were filled with 40 units of DNA Polymerase I, Large Klenow Fragment (NEB, #M0210L) with biotin-dCTP and biotin-dATP (Invitrogen) at 25 °C for 45 min. The reaction was terminated by incubation at 65 °C for 20 min with 0.03 M EDTA. Chromatin was collected and ligated using 50 units of T4 DNA Ligase (NEB, #M0202L) at room temperature for 3 h. After proximity ligation, the biotin from the unligated ends was removed using 200 units of exonuclease III (NEB, #M0206S) at 37 °C for 5 min. After the cross-linking was reversed by reaction with proteinase K at 65 °C for 2 h, the ligated DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, #69506) in accordance with the manufacturer's instructions. Then, 250–400-bp DNA fragments were selected and purified using 0.7X + 0.3X Ampure XP beads, which were then subjected to blunt-end repair, poly-adenylation, and adaptor addition using the VAHTS Universal Plus DNA Library Prep Kit for MGI (Vazyme, #NDM617). The fragments were subsequently subjected to streptavidin C1 bead (Invitrogen, #65001)-mediated pull-down and polymerase chain reaction (PCR) amplification (95 °C for 3 min; 10–12 cycles of 98 °C for 20 s, 60 °C for 15 s, and 72 °C for 30 s; 72 °C for 5 min; and a 4 °C hold). The amplified products were purified with 0.8X AMPure XP beads. Finally, Micro-C-XL libraries were quantified and sequenced on the MGI-seq platform (BGI, China).

### RNA-seq data processing

Sequence alignment to the hg19 reference genome was performed using HISAT2 v2.2.1. Reads with MAPQ < 30 were filtered out. Gene expression levels were determined using featureCounts v2.0.1. To assess the statistical significance of differential expression for each gene, the $P$ value and corresponding adjusted $p$ value were calculated using DESeq2 v1.38.3. Differentially expressed genes were identified on the basis of both $\log_2$-fold changes in expression and q values. Specifically, genes with an adjusted $p$ value < 0.05 and a $\log_2$-fold change in expression > 2 were identified as significantly upregulated genes, and genes with an adjusted $p$ value < 0.05 and a $\log_2$-fold change in expression < -2 were identified as significantly downregulated genes.

### ATAC-seq data processing

Sequence alignment to the hg19 reference genome was performed using Bowtie2[85]. ATAC-seq peak calling was performed with MACS2 using the parameter "-- nomodel" and default parameters otherwise[86].

We compared the ATAC-seq peak calling results for normal BEP2D cells with those for BERP35T1 and BERP35T4 tumour cells. First, the

peak calling results (narrowPeak files) of all replicates were merged using the "merge" command in BEDtools software (v2.29.1). To rigorously identify gained or lost ATAC-seq peaks in tumour cells, we first excluded peaks that were identified in both the normal and tumour cell lines. We subsequently calculated the reads per kilobase of transcript per million reads mapped (RPKM) value of each ATAC-seq peak in tumour cells and calculated the RPKM value of the corresponding region in normal BEP2D cells. We then calculated the read intensity ratio between the tumour cells and normal cells using the following equation:

$$R_i = \frac{(RPKM_{ti} + Pseudovalue)}{(RPKM_{ni} + Pseudovalue)} \quad (1)$$

where $RPKM_{Ti}$ is the RPKM value of the ith peak in the tumour cells, and $RPKM_{Ni}$ is the RPKM value of the corresponding region in the normal BEP2D cells. We used 1 as the pseudovalue to prevent division by 0 and to eliminate the effect of extremely low $RPKM_{ni}$ values that would otherwise increase $R_i$ and generate false positives. We retained only peaks with $R_i > 2$ as ATAC-seq peaks obtained from tumour cells. ATAC-seq peaks lost in tumour cells were identified via a similar procedure. In brief, ATAC-seq peaks identified in both tumour cells and normal cells were removed, $R_i$ was calculated for each peak, and only peaks with $R_i < 0.5$ were retained.

## Hi-C data processing

Sequence alignment to the hg19 reference genome was performed using Bowtie2. The HOMER[87] package was used to process the aligned data. First, the "makeTagDirectory" command was used to remove all PCR duplicates and to create a HOMER tag directory with the parameter "-tbp 1", with all other parameters set to default values. The "AnalyzeHiC" command was used to generate normalized Hi-C contact matrices using the following parameters: "-res 20000 -window 20000". "-balance" parameter in the HOMER software was used to perform Hi-C matrix normalization, which normalize the sum of each bin of the Hi-C contact matrix to the same scale. Default settings were used for the other parameters. The "runHiCpca.pl" script was used for PCA and the calculation of the first eigenvector (PC1) for A/B compartment assignment. The parameters "-res 200000 -window 200000" were used, with default settings used for all other parameters. The "findTADsAndLoops.pl" script was used to detect TADs and loops with the parameters "-res 20000 -window 20000", with default settings used for all other parameters.

## Micro-C data processing

Sequence alignment to the hg19 reference genome was performed using BWA[88] software. Pairtools[89] was used to process the aligned data. First, the "pairtools parse" command was used to record valid ligation events with the following parameters: "--min-mapq 40 --walks-policy 5unique --max-inter-align-gap 30 --nproc in 8 --nproc-out 8 --chroms-path hg19.genome". Next, the parsed pairs are then sorted using "pairtools sort". The "pairtools dedup" command was subsequently used to detect and remove PCR duplicates. The "pairtools split" command was further used to split the final files into two files:.pairs and bam files. The bam files were sorted and indexed using SAMtools[90] with the "samtools sort" and "samtools index" commands. Finally, the pairs files were converted to hic files and visualized using Juicer[91]. Aggregate Peak Analysis (APA) was performed using Coolpup.py[92] software with parameters: --ignore_diags 2 --maxdist 3000000. And APA scores were also calculated and visualized using Coolpup.py with plotpup.py command. To give more quantitative information about the level of enrichment of interactions in the center of the pileup, the average value of the few central pixels of the heatmap was labelled.

## CUT&Tag data processing

Sequence alignment to the hg19 reference genome was performed using Bowtie2[85]. SAMtools[90] was used to sort the bam files and create an index. The "bamCoverage" command of deeptools[93] was used to compute the coverage of reads, and bigWig files were output for visualization and analysis. We used the following parameters: "--binSize 20 --normalizeUsing RPKM --ignoreForNormalization chrX". MACS2 was used to call peaks with the following parameters: "-q 0.05 -f BAM -g hs".

## SDOC calculation

First, conserved TADs among all three datasets were rigorously identified. To this end, we used TAD datasets generated by HOMER software and compared the two boundaries of each TAD in BEP2D cells to those of the overlapping TADs in both BERP35T1 and BERP35T4 cells. A TAD was identified as conserved among all 3 datasets if neither of the TAD boundaries in the BEP2D cell dataset were shifted by more than 1 bin in either tumour cell dataset.

Then, we calculated the SDOC in the conserved TADs using the "get_SDOC.py" python script in the SDOC GitHub repository (https://github.com/birmjiangs/SDOC). Specifically, three-dimensional coordinates of all loci in individual TADs were calculated using Pastis-PM2[94]. These coordinates represented simulated chromatin structures within the TAD and were used to construct the convex hull of each TAD. The raw SDOC was calculated as the ratio of the number of open chromatin peaks (the total number of ATAC-seq peaks in each TAD region) to the raw TAD volume. Quantile normalization was performed on the raw SDOC using the preprocessCore package of R to normalize the SDOC value of each dataset to a Gaussian distribution.

## Calculation of the differential chromatin contact frequency

First, z score normalization was performed on all Hi-C contact matrices in each dataset. For each data point $x$ in a dataset, the corresponding z score $z$ was calculated using the following equation:

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

where $\mu$ is the mean value of the dataset, and $\sigma$ is the standard deviation of the dataset. Specifically, all cells in intrachromatin Hi-C contact matrices that share the same genomic distances, which were first set to the bin size of the Hi-C contact maps, were z score normalized. This procedure was subsequently repeated by traversing each genomic distance up to 2 Mb. This z score normalization procedure eliminated the effects of both sequencing depth and genomic distance. The differential chromatin contact frequency of a certain cell was subsequently calculated by subtracting the corresponding z score-normalized contact frequency in the BEP2D cell dataset from that in each tumour cell dataset.

## Generation of differential contact maps between normal and tumour cells

First, average z score-normalized contact maps centred on each group of locus pairs in normal and tumour cells were generated for BEP2D, BERP35T1 and BERP35T4 cells. Then, for each group of locus pairs, the differential contact map between the BEP2D cell dataset and a tumour cell dataset was generated by subtracting the average z score-normalized contact map of the normal BEP2D cell dataset from that of each tumour cell dataset.

## Comparison of chromatin accessibility levels of TFBSs between normal and tumour cells

Putative TFBSs were predicted using the MEME suite (v5.1.0). TF binding motif datasets were downloaded from the JASPAR database (http://jaspar.genereg.net/). The open chromatin regions in BEP2D,

BERP35T1 and BERP35T4 cells were merged using the "merge" command in BEDtools software. The output.bed file containing the union set of open chromatin sites was used for TFBS searching. The "fimo" command in the MEME suite was used for identification of putative TFBSs with the parameter "--thresh 1e-5" and default parameters otherwise.

The chromatin accessibility level of each TFBS was assigned as the RPKM-normalized read intensity of the ATAC-seq peak containing the TFBS. The chromatin accessibility level of each TF was calculated as the mean of all the chromatin accessibility levels of all the TFBSs in the dataset.

## Statistical analysis and reproducibility

The RNA-seq experiments for the BEP2D, BERP35T1 and BERP35T4 cell lines were repeated three times. The Hi-C experiments for the BEP2D, BERP35T1 and BERP35T4 cell lines were repeated three times. The ATAC-seq experiments for the BEP2D, BERP35T1, and BERP35T4 cell lines were repeated three times, and two replications were applied to the BEP2DNC, BEP2DPOE, and BERP35T1 cell lines. The Micro-C experiments for the BEP2DNC, BEP2DOE and BERP35T1 cell lines were repeated four times.

A t-test was used to assess the statistical significance of the differences in the mean log2 expression fold changes between different gene groups, as shown in Figs. 2C, D and 3E. Fisher's exact test was used to assess the statistical significance of the differences in the fraction of overlapping epigenomic features between different groups, as shown in Fig. 4B, C. The levels of significance were calculated using the SciPy library (version 1.5.2) of Python (version 3.7.0).

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The sequencing data generated in this study have been deposited in the GEO database (accession number: GSE175857, GSE175858, GSE175859, GSE175860, GSE272817, GSE272819, GSE272820, GSE272821 and GSE272822). The source data underlying Figs. 2, 3 and 6 are provided as a Source Data file. Uncropped and unprocessed scans of all blots are provided as a Source Data file. Source data are provided with this paper.

## References

1. Knoechel, B. et al. An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nat. Genet.* **46**, 364–370 (2014).
2. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science (New York, N.Y.)* **357** (2017).
3. Chatterjee, A., Rodger, E. J. & Eccles, M. R. Epigenetic drivers of tumourigenesis and cancer metastasis. *Semin. Cancer Biol.* **51**, 149–159 (2018).
4. Zhu, H. et al. Candidate cancer driver mutations in distal regulatory elements and long-range chromatin interaction networks. *Mol. Cell* **77**, 1307–1321.e10 (2020).
5. Iwafuchi, M. et al. Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones. *Nat. Genet.* **52**, 418–427 (2020).
6. Akdemir, K. C. et al. Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nat. Genet.* **52**, 294–305 (2020).
7. Corces, M. R. & Corces, V. G. The three-dimensional cancer genome. *Curr. Opin. Genet. Dev.* **36**, 1–7 (2016).
8. Tirmarche, M. Cancer risk following alpha-emitter exposure. *Ann. ICRP* **47**, 115–125 (2018).
9. Iliakis, G., Mladenov, E. & Mladenova, V. Necessities in the processing of DNA double strand breaks and their effects on genomic instability and cancer. *Cancers* **11**, 1671 (2019).
10. Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.* **27**, 247–254 (2001).
11. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
12. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Sci. (N. Y., N. Y.)* **326**, 289–293 (2009).
13. Jr, D. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
14. Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
15. Hu, G. et al. transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T Cells. *Immunity* **48**, 227–242.e8 (2018).
16. Wu, P. et al. 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. *Nat. Commun.* **8**, 1937 (2017).
17. Johnstone, S. E. et al. Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **182**, 1474–1489.e23 (2020).
18. Rowley, M. J. et al. Evolutionarily conserved principles predict 3D chromatin organization. *Mol. Cell* **67**, 837–852.e7 (2017).
19. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
20. Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**, 535–550 (2019).
21. Szabo, Q. et al. TADs are 3D structural units of higher-order chromosome organization in Drosophila. *Sci. Adv.* **4**, eaar8082 (2018).
22. Szabo, Q. et al. Regulation of single-cell genome organization into TADs and chromatin nanodomains. *Nat. Genet.* **52**, 1151–1157 (2020).
23. Guan, Y. et al. Senescence-activated enhancer landscape orchestrates the senescence-associated secretory phenotype in murine fibroblasts. *Nucleic Acids Res.* **48**, 10909–10923 (2020).
24. Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
25. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
26. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362** (2018).
27. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
28. Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
29. van Bemmel, J. G. et al. The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat. Genet* **51**, 1024–1034 (2019).
30. Weintraub, A. S. et al. YY1 is a structural regulator of enhancer-promoter loops. *Cell* **171**, 1573–1588.e28 (2017).
31. Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
32. Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
33. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet* **52**, 8–16 (2020).

34. Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).

35. Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381.e20 (2017).

36. Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).

37. Weischenfeldt, J. et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2017).

38. Taberlay, P. C. et al. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731 (2016).

39. Yang, M. et al. Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nat. Commun.* **10**, 1519 (2019).

40. Haller, F. et al. Enhancer hijacking activates oncogenic transcription factor NR4A3 in acinic cell carcinomas of the salivary glands. *Nat. Commun.* **10**, 368 (2019).

41. Franke, M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).

42. Dixon, J. R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).

43. Valton, A.-L. & Dekker, J. TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* **36**, 34–40 (2016).

44. Santos-Pereira, J. M., Gallardo-Fuentes, L., Neto, A., Acemel, R. D. & Tena, J. J. Pioneer and repressive functions of p63 during zebrafish embryonic ectoderm specification. *Nat. Commun.* **10**, 3049 (2019).

45. Sato, T. et al. Epigenomic profiling discovers trans-lineage sox2 partnerships driving tumor heterogeneity in lung squamous cell carcinoma. *Cancer Res.* **79**, 6084–6100 (2019).

46. Willey, J. C. et al. Immortalization of normal human bronchial epithelial cells by human papillomaviruses 16 or 18. *Cancer Res.* **51**, 5370–5377 (1991).

47. Sun, J. F. et al. Decreased efficiency of gamma-ray-induced DNA double-strand break rejoining in malignant transformants of human bronchial epithelial cells generated by alpha-particle exposure. *Int. J. Radiat. Biol.* **78**, 773–780 (2002).

48. Liu, Z. L., Chen, H. H., Zheng, L. L., Sun, L. P. & Shi, L. Angiogenic signaling pathways and anti-angiogenic therapy for cancer. *Signal Transduct. Target Ther.* **8**, 198 (2023).

49. Choi, H. J. et al. A conserved phosphorylation switch controls the interaction between cadherin and β-catenin in vitro and in vivo. *Dev. Cell* **33**, 82–93 (2015).

50. Bhat, S. S., Anand, D. & Khanday, F. A. p66Shc as a switch in bringing about contrasting responses in cell growth: implications on cell proliferation and apoptosis. *Mol. Cancer* **14**, 76 (2015).

51. Schipper, K. et al. Rebalancing of actomyosin contractility enables mammary tumor formation upon loss of E-cadherin. *Nat. Commun.* **10**, 3800 (2019).

52. Kern, F., Niault, T. & Baccarini, M. Ras and Raf pathways in epidermis development and carcinogenesis. *Br. J. Cancer* **104**, 229–234 (2011).

53. Bansaccal, N. et al. The extracellular matrix dictates regional competence for tumour initiation. *Nature* **623**, 828–835 (2023).

54. Ge, Y. et al. Stem cell lineage infidelity drives wound repair and cancer. *Cell* **169**, 636–650.e14 (2017).

55. Liu, Q. et al. Genome-wide identification and analysis of A-to-I RNA editing events in the malignantly transformed cell lines from bronchial epithelial cell line induced by α-particles radiation. *PLOS ONE* **14**, e0213047 (2019).

56. Hoang, P. H. et al. Whole-genome sequencing of multiple myeloma reveals oncogenic pathways are targeted somatically through multiple mechanisms. *Leukemia* **32**, 2459–2470 (2018).

57. Krijger, P. H. & de Laat, W. Regulation of disease-associated gene expression in the 3D genome. *Nat. Rev. Mol. Cell Biol.* **17**, 771–782 (2016).

58. Li, R. et al. 3D genome and its disorganization in diseases. *Cell Biol. Toxicol.* **34**, 351–365 (2018).

59. Chronis, C. et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459.e20 (2017).

60. Park, H. J. et al. Keratinization of lung squamous cell carcinoma is associated with poor clinical outcome. *Tuberc. Respir. Dis. (Seoul.)* **80**, 179–186 (2017).

61. Schmitt, A. D. et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep.* **17**, 2042–2059 (2016).

62. Li, Y. et al. Alterations of specific chromatin conformation affect ATRA-induced leukemia cell differentiation. *Cell Death Dis.* **9**, 200 (2018).

63. Jiang, S. et al. Spatial density of open chromatin: an effective metric for the functional characterization of topologically associated domains. *Briefings in Bioinformatics* (2020).

64. Barrington, C. et al. Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nat. Commun.* **10**, 2908 (2019).

65. Achinger-Kawecka, J., Taberlay, P. C. & Clark, S. J. Alterations in three-dimensional organization of the cancer genome and epigenome. *Cold Spring Harb. Symposia Quant. Biol.* **81**, 41–51 (2016).

66. Dusek, R. L. & Attardi, L. D. Desmosomes: new perpetrators in tumour suppression. *Nat. Rev. Cancer* **11**, 317–323 (2011).

67. Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).

68. Lawrence, M. S. et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).

69. Kandoth, C. et al. Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).

70. Schroeder, M. P., Rubio-Perez, C., Tamborero, D., Gonzalez-Perez, A. & Lopez-Bigas, N. OncodriveROLE classifies cancer driver genes in loss of function and activating mode of action. *Bioinformatics* **30**, i549–i555 (2014).

71. Zehir, A. et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med* **23**, 703–713 (2017).

72. Zhou, Q. et al. ZNF143 mediates CTCF-bound promoter-enhancer loops required for murine hematopoietic stem and progenitor cell function. *Nat. Commun.* **12**, 43 (2021).

73. Wang, R. et al. MyoD is a 3D genome structure organizer for muscle cell identity. *Nat. Commun.* **13**, 205 (2022).

74. Castro-Mondragon, J. A. et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **50**, D165–d173 (2022).

75. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res* **43**, W39–W49 (2015).

76. Davidson, I. F. & Peters, J. M. Genome folding through loop extrusion by SMC complexes. *Nat. Rev. Mol. Cell Biol.* **22**, 445–464 (2021).

77. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220 (2019).

78. Qu, K. et al. Chromatin Accessibility Landscape of Cutaneous T Cell Lymphoma and Dynamic Response to HDAC Inhibitors. *Cancer Cell* **32**, 27–41.e4 (2017).

79. Fudenberg, G. et al. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* **15**, 2038–2049 (2016).

80. Qu, J., Yi, G. & Zhou, H. p63 cooperates with CTCF to modulate chromatin architecture in skin keratinocytes. *Epigenetics Chromatin* **12**, 31 (2019).

81. Sanders, J. T. et al. Radiation-induced DNA damage and repair effects on 3D genome organization. *Nat. Commun.* **11**, 6178 (2020).

82. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

83. Hsieh, T. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553.e8 (2020).

84. Hsieh, T. S., Fudenberg, G., Goloborodko, A. & Rando, O. J. Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. *Nat. Methods* **13**, 1009–1011 (2016).

85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

86. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).

87. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

88. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

89. Abdennur, N. et al. Pairtools: From sequencing data to chromosome contacts. *PLoS Comput Biol.* **20**, e1012164 (2024).

90. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

91. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).

92. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980–2985 (2020).

93. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187–W191 (2014).

94. Varoquaux, N., Ay, F., Noble, W. S. & Vert, J. P. A statistical approach for inferring the 3D structure of the genome. *Bioinformatics* **30**, i26–i33 (2014).

## Acknowledgements

## Author contributions

Q.L., S.J., and X.X. conceived the study and wrote the manuscript with input from all coauthors. Q.L. performed the molecular biology, RNA-seq, ATAC-seq and Hi-C experiments with help from L.W., S.Z., H.T., X.Y., C.J., X.Q., Y.G., S.F. and Y.L. S.J., X.X., Z.X., K.X., Y.L., Z.Y., M.X. and H.L. performed the bioinformatic analyses and integrated and visualized the data. X.B. and P.Z. contributed reagents and materials. H.C. and X.C. provided valuable advice, and Y.H. supervised the research.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-63754-y.

**Correspondence** and requests for materials should be addressed to Hebing Chen, Xiaohua Chen or Yi Hu.

**Peer review information** *Nature Communications* thanks Ilya Flyamer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.