







ARTICLE

# Chromosome compartmentalization alterations in prostate cancer cell lines model disease progression

Rebeca San Martín<sup>1</sup> , Priyjit Das<sup>2</sup>, Renata Dos Reis Marques<sup>1</sup> , Yang Xu<sup>2</sup> , Justin M. Roberts<sup>3</sup> , Jacob T. Sanders<sup>1</sup> , Rosela Gollosi<sup>1</sup>, and Rachel Patton McCord<sup>1</sup> 

Prostate cancer aggressiveness and metastatic potential are influenced by gene expression and genomic aberrations, features that can be influenced by the 3D structure of chromosomes inside the nucleus. Using chromosome conformation capture (Hi-C), we conducted a systematic genome architecture comparison on a cohort of cell lines that model prostate cancer progression, from normal epithelium to bone metastasis. We describe spatial compartment identity (A-open versus B-closed) changes with progression in these cell lines and their relation to gene expression changes in both cell lines and patient samples. In particular, 48 gene clusters switch from the B to the A compartment, including androgen receptor, *WNT5A*, and *CDK14*. These switches are accompanied by changes in the structure, size, and boundaries of topologically associating domains (TADs). Further, compartment changes in chromosome 21 are exacerbated with progression and may explain, in part, the genesis of the *TMPRSS2-ERG* translocation. These results suggest that discrete 3D genome structure changes play a deleterious role in prostate cancer progression.

## Introduction

Prostate cancer is the predominant new cancer diagnosis in males in the United States. It is also the second most common cause of male cancer-related deaths, second only to lung cancer (Siegel et al., 2020). Patients with late-stage prostate cancer present with a higher incidence of metastases to trabecular bone (Jacobs, 1983; Bubendorf et al., 2000; Hernandez et al., 2018).

The specific mechanisms that promote metastasis to bone are not understood. However, disseminated tumor cells can be detected in the blood of ~25% of prostate cancer patients with localized disease. The abundance of these circulating tumor cells positively correlates with metastatic occurrence (Moreno et al., 2005; Danila et al., 2007; Todenhofer et al., 2016).

The human genome's packaging into the nucleus's constrained space requires a systematic organization, from chromosomal territories (Meaburn and Misteli, 2007) to transcriptionally active and inactive chromatin compartments (Lieberman-Aiden et al., 2009). Within compartments, chromatin further organizes into topologically associating domains (TADs), which are segregated from each other by the insulator protein CTCF (Dixon et al., 2012, Nora et al., 2017), and finally into chromatin loops (Nuebler et al., 2018). With the advent of technologies such as genome-wide chromosome conformation capture (Hi-C), it has become clear

that these structures are essential for proper gene regulation, DNA replication, and repair (Hnisz et al., 2016; Dekker and Mirny, 2016; Pope et al., 2014; McCord and Balajee, 2018). Further, rearrangement of these domains can impact both the nuclei's ability to squeeze through tight spaces during metastatic migration and the expression patterns of oncogenes (Gerlitz and Bustin, 2010; Stephens et al., 2018; Hnisz et al., 2016; Barutcu et al., 2015). Nuclear atypia is a common diagnostic tool in prostate cancer (Verdone et al., 2015; Diamond et al., 1982), and there is evidence that the nuclear lamina content of prostate cells differs between normal epithelium, benign prostatic hyperplasia (BPH), and cancer (Partin et al., 1993). This suggests that genome architectural changes may occur in prostate cancer and might influence cancer-promoting gene expression profiles and the nuclear malleability necessary for cancer cells to metastasize.

Mutations, structural alterations, and gene regulatory changes at numerous genomic loci have been associated with a higher risk for prostate cancer (Ahmadiyeh et al., 2010; Helfand et al., 2015; Du et al., 2016). In particular, the *TMPRSS2-ERG* translocation in chromosome 21 has been found to associate with poor patient prognosis (Zhou et al., 2020; Demichelis et al., 2007; Tomlins et al., 2005). It has been shown that overexpression of

<sup>1</sup>Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, TN; <sup>2</sup>University of Tennessee - Oak Ridge National Lab (UT-ORNL) Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, TN; <sup>3</sup>Department of Genitourinary Medical Oncology and the David H. Koch Center for Applied Research of Genitourinary Cancer, University of Texas MD Anderson Cancer Center, Houston, TX.

Correspondence to Rachel Patton McCord: [rmccord@utk.edu](mailto:rmccord@utk.edu).

© 2021 San Martín et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).

ERG results in chromatin conformation changes (Rickman et al., 2012), and FISH experiments combined with irradiation have suggested that such induced chromosome rearrangements alongside DNA breaks can encourage the formation of the *TMPRSS2-ERG* translocation (Mani et al., 2009). Indeed, it is known that relative chromosome proximity can influence the pattern of translocations that occur (Zhang et al., 2012; Balajee et al., 2018), so this type of local rearrangement could result from chromosome compartmentalization changes that increase contact frequency among the different loci (Engreitz et al., 2012).

Recent studies of the 3D genome structure associated with prostate cancer (Rhie et al., 2019; Luo et al., 2017; Taberlay et al., 2016) have contributed some insight into regulatory chromatin loops, epigenetic alterations, and the influence that variable structures have in transcription. However, these studies did not address whether there are early changes in genome architecture that persist throughout progression. In this context, changes in compartment identity from transcriptionally repressed heterochromatin to euchromatin poised for transcription could identify genes required for early oncogenesis and those necessary for metastasis. Further, changes in TAD positioning or shifting of TAD boundaries could reveal altered interactions of neighboring promoter-enhancer regions.

In the present study, we use Hi-C to characterize the genome organization across a cohort of nine cell lines that model prostate cancer progression from the normal epithelium to bone metastasis, including two bone metastatic cell lines of African American lineage. We further assess the different hierarchical levels of genome organization: from large interchromosomal translocations to compartment identity to TAD location and TAD boundary shifting. We have identified a cohort of 386 genes that change compartment identity across prostate cancer progression. Interestingly, most of these genes switch compartments as proximal clusters. These compartment identity changes are accompanied by distinct structural features at higher resolution such as gain or loss of TAD structure, stalled transcriptional loops and structural deserts, and TAD boundary appearance, disappearance, or positional shifting. Finally, our results revealed several “genomic architecture hotspots” whose structural changes are persistent throughout the metastatic models; these include *WNT5A*, *CDK14*, androgen receptor (*AR*), and the *TMPRSS2-ERG* locus, among others. These results suggest that the 3D genome structure may serve as a prognostic marker for the progression of prostate cancer to bone metastasis.

## Results

### Hi-C reveals distinct changes in the 3D genome structure of a cohort of cell lines that model prostate cancer progression

To determine how the genome’s organization is affected by disease stage, we selected a cohort of nine cell lines that model the progression of prostate cancer, as follows: RWPE1 (Bello et al., 1997) was used to represent normal epithelium. This cell type is epithelial and nontumorigenic, and although it is an artificial cell line model, we verified as described below that its

genome structure properties are similar to primary prostate PrEC cells. LNCaP (Horoszewicz et al., 1983), originally isolated from lymph node metastasis, was used as an early adenocarcinoma model. VCaP (Korenchuk et al., 2001) and MDaPCa2a/b (Navone et al., 1997) were used as models for prototypical osteoblastic bone metastasis; these cells are of Caucasian and African American origins, respectively. 22RV1 (Sramkoski et al., 1999) and LNCaP C4-2B (Thalmann et al., 2000) were included in the study as cells that, although isolated from human sources, are models of murine bone metastasis. Atypical metastatic cell lines used in this study include PC3 (Kaighn et al., 1979; osteoclastic and androgen-independent) and Du145 (Stone et al., 1978; brain metastasis; Fig. 1).

Hi-C was performed on cells under normal cell culture conditions (see Materials and methods). Additionally, publicly available datasets for RWPE, PrEC, and LNCaP-C42B were used as comparisons (Rhie et al., 2019; Taberlay et al., 2016). Hi-C mapping and quality control statistics for all samples can be found in Table S1.

The frequency of chromosome contacts is represented in a heatmap where the XY axes are the chromosomal coordinates. The color intensity reflects the frequency with which two particular locations were found to be in contact. At a resolution of 1-Mb bins, the Hi-C heatmaps reveal chromosome territories. At a 250-kb resolution, it is possible to see a plaid pattern of interaction strength, which represents the spatial segregation of A and B compartments. We classify each genomic region as belonging to the A or B compartment using principal component analysis (PCA). Positive values of the first eigenvector (eigen1 or PC1) represent A compartment regions (typically open euchromatin) while negative values denote B compartment (heterochromatin). Finally, at 40-kb resolution, distinct TADs are evident, which are regions of enhanced contacts (which may promote interactions between promoters and enhancers) segregated by the insulator protein CTCF (Fig. 2 A).

Whole-genome contact maps for cells in our model reveal that the highest incidence of interactions occurs in cis: chromosomes primarily interacting within themselves (Fig. 2, B–E). Chromosomal translocations are evident as very strong interactions occurring in trans between different chromosomes. Comparing our Hi-C results with published spectral karyotyping (SKY) data (Pan et al., 1999; van Bokhoven et al., 2003; Roh et al., 2008) and karyotyping information from the American Type Culture Collection (ATCC), we found that our Hi-C data detects 84% of all previously reported translocations. Owing to the high resolution of Hi-C data, we also characterized several previously unreported, smaller translocation events (Fig. S1, Fig. S2, and Fig. S3). Interestingly, chromosomal territories remain well defined throughout progression, with subtle changes in intra-chromosomal interactions, as shown in 250-kb resolution heatmaps of each chromosome (refer to Fig. S2 for examples). From these contact maps, it is evident that some chromosomes in some cell lines are broken into multiple pieces (Fig. S4). However, when we perform compartment analysis on each of these broken pieces separately, we find that their underlying A/B compartmentalization is largely similar to cell lines with unbroken chromosomes.

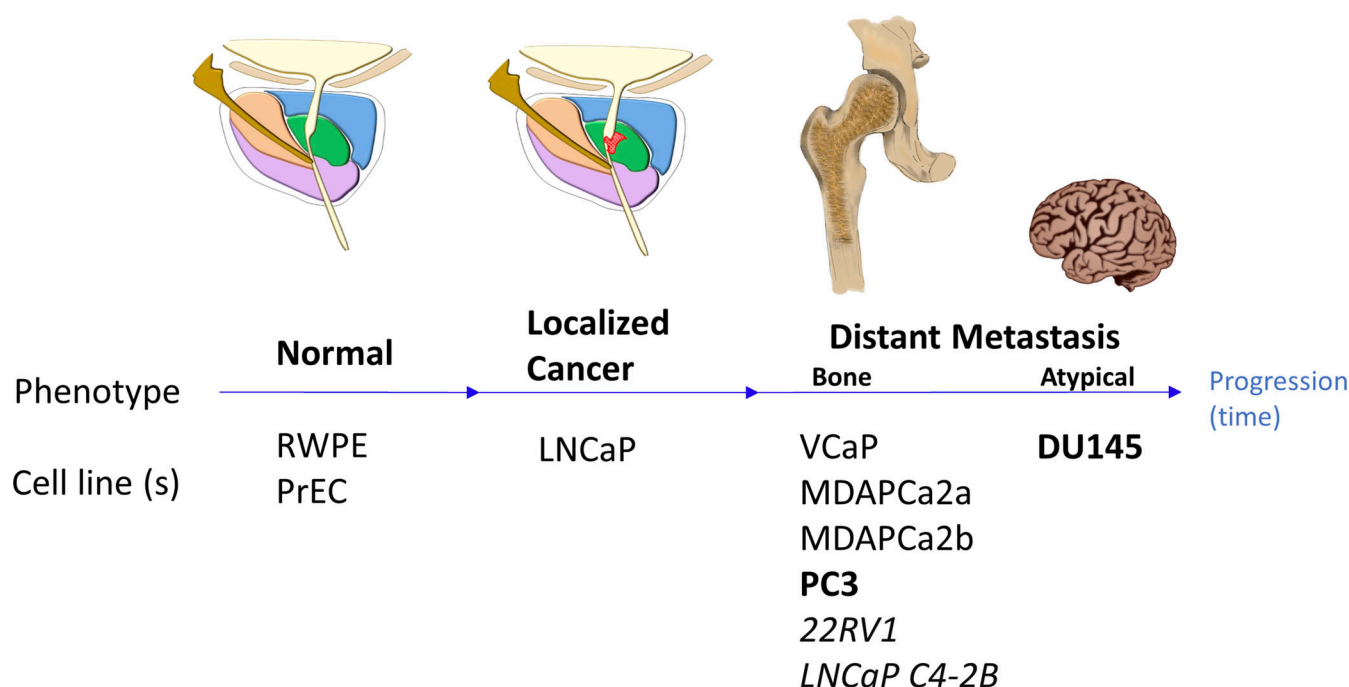


Figure 1. **A cell line-based model for the progression of prostate cancer.** Cell lines used in this study model different stages in the progression from normal epithelium (RWPE, PrEC) to localized disease (LNCaP) to prostate cancer that has metastasized to bone (VCaP, MDAPCa2a, MDAPCa2b, PC3) or to an atypical site (DU145). Italics: cell lines developed through mouse xenografts. Bold: androgen-independent cell lines.

### Genes important to prostate cancer progression switch chromatin compartment identity, and these changes are accompanied by transcription activation

Examination of higher resolution (250-kb) z-score correlation heatmaps of cis interactions for each chromosome shows that the characteristic plaid contact pattern, associated with spatial compartment identity, changes among cell lines (Fig. 3 A). These patterns can be observed throughout all chromosomes, and the changes in compartmentalization are specific to each cell line (Fig. S5). Motivated by an apparent compartment “erosion” in VCaP, we quantified the overall strength of interactions within A compartment regions and within B compartment regions in the different cell types. We find that all prostate cancer cell lines show a loss of A compartment interaction strength relative to RWPE, reflecting an increased intermixing of the A compartment with B compartment regions. Meanwhile, B compartment strength experiences less change in LNCaP and VCaP and even increases in MDAPCa cell lines. Overall, this leads to an imbalance of B and A compartment strength in all prostate cancer cell lines (Fig. S6, A and B).

Systematic analysis of the A/B compartment tracks, per chromosome, for all the cell lines in our model, revealed regions where the compartment identity remained the same and where it was changed (Fig. S5). We note that there is a general level of similarity between epithelial-derived cell lines (Fig. S6 C). By comparing compartment tracks for each cancer cell line to the nontransformed control (RWPE), we find that the changes among cell lines are specific and localized. For example, eigenvector tracks for chromosome 21 show three distinct hotspots of compartment identity switches between the RWPE and cancer

cell lines (Fig. 3 B). Interestingly, the rightmost region encompasses the TMPRSS2-ERG locus, a local translocation site that correlates with worse progression and metastasis in prostate cancer (Tomlins et al., 2005; Demichelis et al., 2007; Hägglöf et al., 2014).

A comparison of genome-wide compartment classifications for all cell lines by hierarchical clustering demonstrates that RWPE indeed clusters very closely with primary prostate epithelium PrEC cells (Fig. 3 C). We further demonstrate that RWPE and PrEC compartment profiles are consistent at important genomic locations of interest (Fig. S7). Because RWPE is a reproducible cell culture model, while PrEC cells vary by donor over time, we remain focused on RWPE for the remainder of the comparisons. When we examine the compartment patterns by nearest neighbor analysis (SPRING plot; Weinreb et al., 2018; see Materials and methods), we observe that LNCaP is a central node in a model axis to metastasis (Fig. 3 D, axis highlighted in yellow), connecting RWPE (nontransformed epithelium) to bone metastatic cell lines (VCaP and MDAPCa). Interestingly, the compartment patterns of MDAPCa2A and 2B, which were derived from an African American patient, is distinct from VCaP, which is of Caucasian lineage. This analysis also clusters atypical metastatic lines DU145 and PC3 together along a third axis radiating from the LNCaP central node.

To consistently mathematically classify regions that significantly changed compartments between cell lines, PC1 values per bin per cell line were normalized by subtracting the corresponding value derived from the RWPE cell line. Significant changes in compartment identity ( $|\bar{X} + 1.5\sigma|$ ) were identified and the corresponding genomic areas annotated. Interestingly,

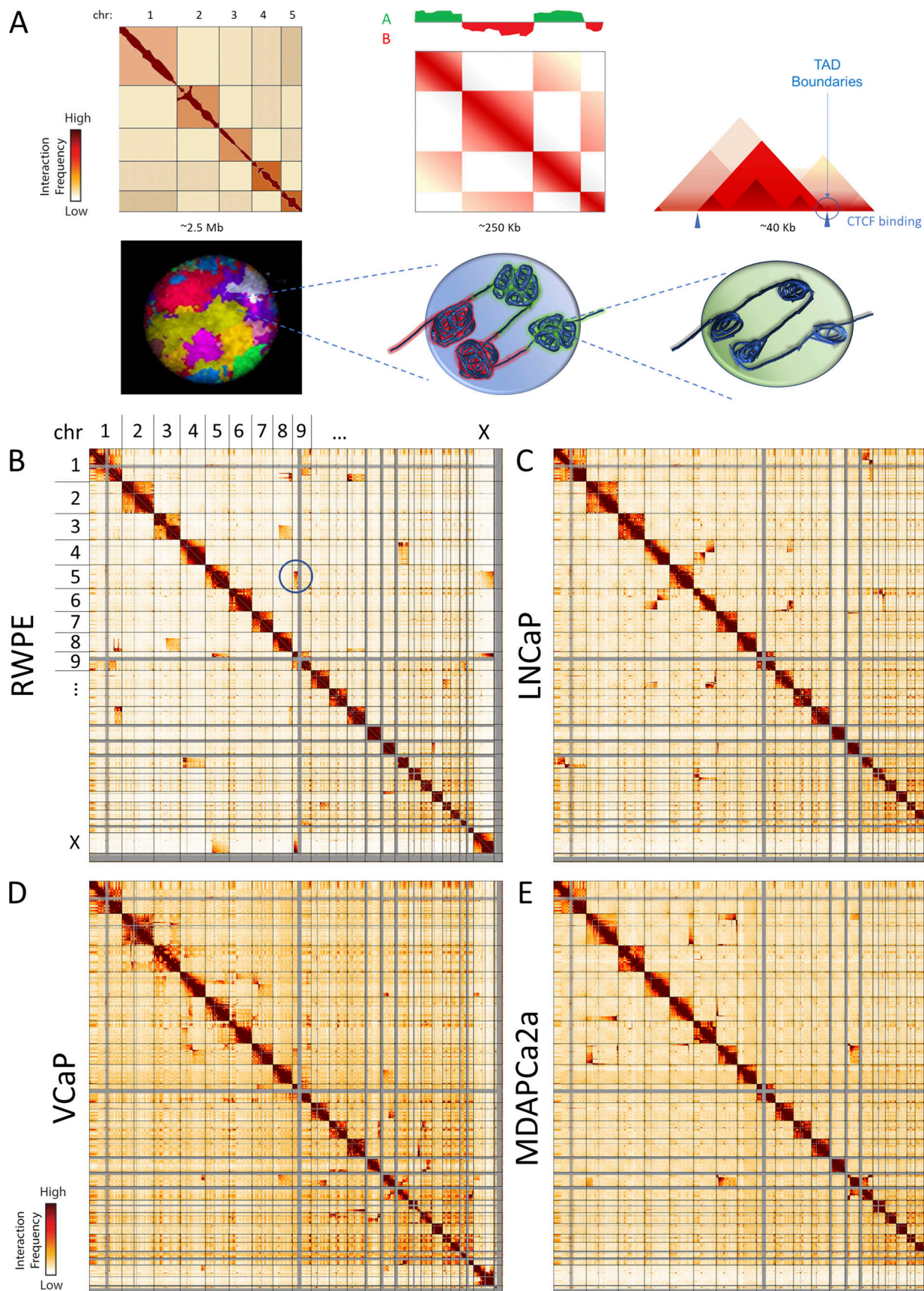


Figure 2. **Hi-C characterizes the hierarchical genome of each cell line.** (A) At 1-Mb resolution (left), Hi-C heatmaps reveal defined chromosome (chr) territories along the diagonal. At 250-kb resolution, a characteristic plaid pattern emerges: PCA of this matrix reveals A/B chromosome compartmentalization

(center). At 40-kb resolution, TADs, within compartments, become apparent (right). 2.5-Mb Hi-C heatmaps for RWPE (B), LNCaP (C), VCaP (D), and MDAPCa2a (E). Translocations between chromosomes appear as high interaction frequency areas away from the diagonal. Translocation between chromosomes 8 and 5 highlighted as an example in RWPE (circle). While these translocation events have been validated by SKY analysis, the high resolution of Hi-C data allows for the characterization of smaller events (Fig. S1).

genomic regions that switch compartment in the LNCaP-RWPE comparison tend to have a switched identity in bone metastatic cell lines as well (Fig. 3 E). These regions thus change compartment status early in progression and then maintain their new spatial localization rather than switching back or experiencing progressively more shift in compartment. Overall, we have identified 181 genomic bins (250 kb in size) whose compartmentalization changes from the B to the A compartment in either (1) six or more cancer cell lines compared with RWPE or (2) all cells in the progression axis, compared with RWPE. These genomic bins contain 300 genes that are therefore moving from the B to the A compartment (Table S2), suggesting that these genes become more likely to be transcribed. These include *AR*, *WNT5A*, *CDK14*, and genes located close to *TMPRSS2*, such as *BACE2*. The majority of these genes (256) are grouped in 48 proximal clusters (Table S3). We have also identified 86 genes whose genomic loci switch compartments from the A to B compartment, suggesting a genome structure rearrangement more likely to result in gene silencing (Table S2). Such is the case for certain cadherins, annexins, and mediators of inflammation. Of these genes, 64 are grouped in 16 clusters (Table S3).

A comparison between the bone metastatic cell line (VCaP) and the adenocarcinoma cell line (LNCaP) revealed that 345 more bins switch from the B to the A compartment, comprising 313 genes. Interestingly, these events appear to be VCaP-specific, suggesting that the MDA bone metastatic lineages undertake a different pathway to oncogenesis.

Since a switch from the B to A compartment could make a region permissive for transcriptional activation, we used the Clariom-S microarray to profile expression levels in the targets identified. Using the expression level of RWPE as a baseline, we have found that 47% of genes that switch from B to A compartment show a significant transcriptional induction in LNCaP (twofold or higher). In turn, 49% of those genes are even further up-regulated in VCaP compared with LNCaP, suggesting an exacerbation of this expression pattern with progression. These genes include, among others, *AR*, *TMPRSS2*, *CDK14*, and *WNT5a* (Fig. 4 A). In contrast, only 18% of the genes that show induction in LNCaP are induced further in the MDAPCa cell lines (Fig. 4 B).

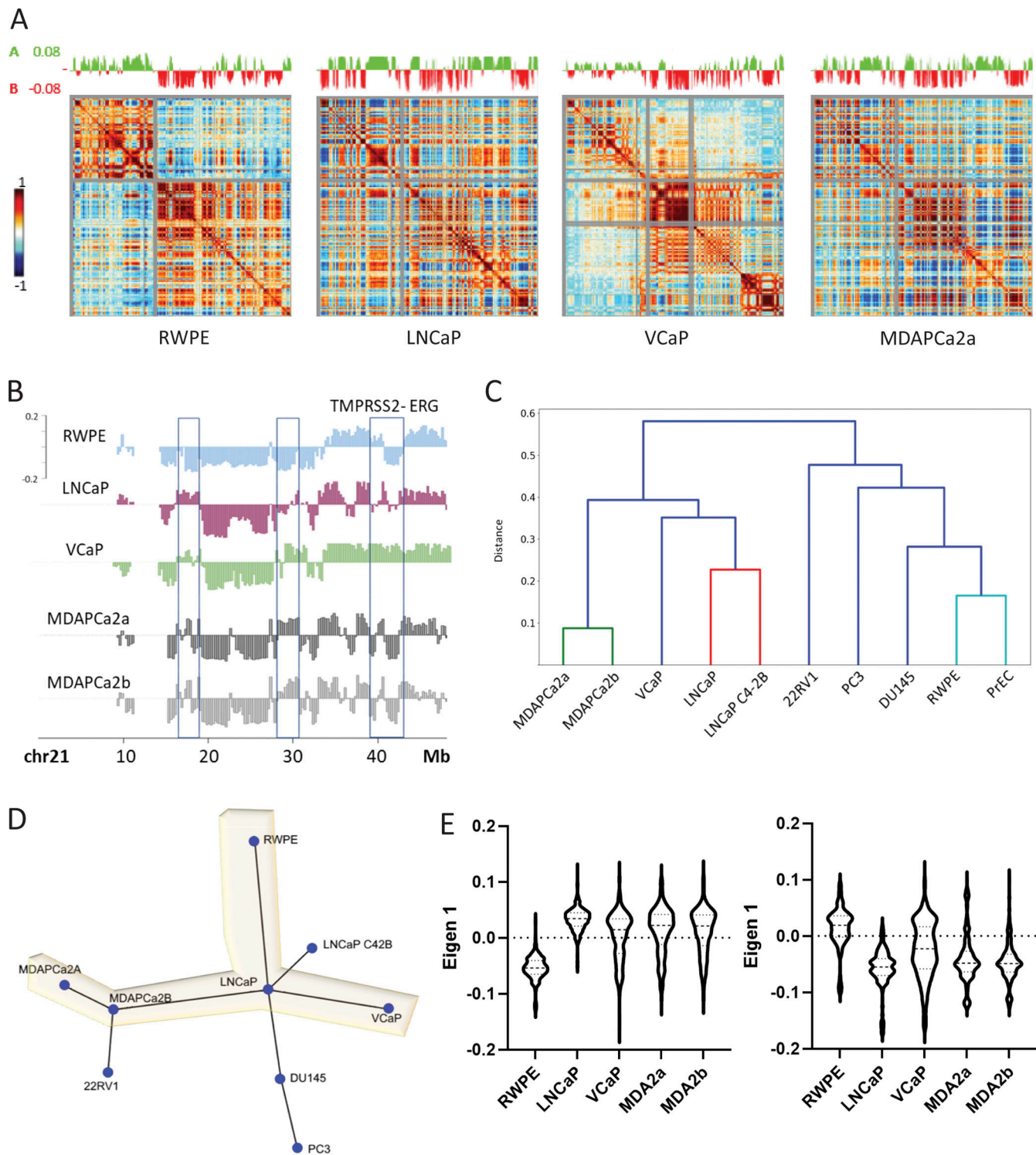
Finally, we found 25 genes underwent a B to A compartment switch, were up-regulated in LNCaP versus RWPE, and were further up-regulated in both MDAPCa cell lines and VCaP relative to LNCaP. This set includes *CDK14*, *WNT5a*, and *BACE2*, which is a close neighbor of *TMPRSS2* (Fig. 4 C). These data suggest that these transcriptional hotspots are shared in osteoblastic metastatic cells and might be necessary for colonization and survival in a secondary bone site. Interestingly, while amyloid precursor protein (*APP*), the proteolytic target of  $\beta$  secretase 2 (*BACE2*), does not change compartment between RWPE and LNCaP, it does switch both compartments from B to A and shows an expression increase when we compare VCaP to LNCaP.

### Compartment identity switches are accompanied by distinct structural changes at the TAD level, including boundary shifts

TADs contribute to the 3D architecture of the genome by sequestering enhancers and promoters with their target genes, with a low likelihood of interaction across boundaries (Lupiáñez et al., 2015; Nora et al., 2012; Dixon et al., 2012). The disappearance of TAD boundaries has been identified as a potential activator of oncogenes (Hnisz et al., 2016). To analyze whether there are changes in 3D genomic structure at the local level surrounding genes associated with compartment identity switches, we used higher-resolution Hi-C heatmaps (40 kb).

The *WNT5a-ERC2* locus (Fig. 5, Cluster 1) is in the B compartment in nontransformed control, switching to the A compartment in all four members of the model axis to progression: LNCaP, VCaP, and both MDAPCa cell lines. While the right TAD boundary location is relatively consistent, the location of the left boundary shifts positions: in LNCaP, DU145, MDAPCa2B, and PC3, *WNT5A* localizes in a TAD with *LRTM1* instead of *ERC2*. Interestingly, in VCaP, these three genes are clustered in the same TAD. *FZD1* and *CDK14*, a receptor and activator cyclin of noncanonical WNT signaling, and which have been associated with the function of *WNT5A*, also switch to the A compartment (Fig. 5, Cluster 2). In this case, the genes are separated by a TAD boundary that rests atop the gene body of *CDK14*, but increased loops are evident on the separate TADs in the cancer cell lines, compared with RWPE. Remarkably, *CDK14* is one of the genes whose transcription is consistently up-regulated in all metastatic cell lines in the progression axis. TAD boundary shifting, appearance, and disappearance were also observed in other clusters that switch compartment identity (Fig. S8).

A detailed survey of all compartment-switched loci revealed five possible categories of changes in local structure. We classified all compartment switch regions into these categories. Strikingly, the *AR* locus displays all five types of structure change across the different cell lines. First, a highly disorganized area (no TADs, uniformly distributed interactions across a region of the heatmap) becomes highly organized, or vice versa (Fig. 6 A). In *AR*, the gene is in the B compartment in RWPE, and the area around it is highly disorganized. With progression, the area becomes organized into TADs and sub-TADs for both LNCaP C4-2b and VCaP. The third type of structural conformation is the appearance of a high incidence of interactions along the edge the TAD (Fig. 6 B). We call these interactions “loops” or “stalled loops” because these have previously been associated with the phenomenon of cohesin becoming stalled as it extrudes loops but encounters RNA polymerase, CTCF, or other barriers. This increased loop formation is evident at the *AR* locus in LNCaP, MDAPCa2a, and MDAPCa2b. The fourth structural feature is a complete absence of structure (loss of contacts in the entire region) associated with the gene of interest. This “structural desert” is observed at the *AR* locus in the 22RV1 cell line (Fig. 6 C) and is reminiscent of previously observed



**Figure 3. Changes in A/B compartment identity are relevant to prostate cancer progression.** (A) 250-kb resolution, z-score correlation heatmaps of cis interactions along the X chromosome show a characteristic plaid pattern that denotes compartment identity (compartments classified by the first principal component, PC1, shown above the matrix: green = A, red = B). Alterations in this pattern are seen in each cancer cell line, with a particular disruption and weakening of this pattern observed in the bone metastatic cell line VCaP. (B) PC1 compartment tracks for chromosome 21 (chr21) show distinct hotspots of compartment identity switches (boxes) along the chromosome between the control (RWPE), adenocarcinoma, and bone metastatic cell lines. The rightmost region encompasses the *TMRSS2-ERG* locus. (C) Hierarchical clustering analysis based on spatial compartment identity for all the cells in this study reveal high degree of similarity between nontransformed epithelial cell lines (PrEC and RWPE) and between bone metastatic cell lines MDAPCa2a/b, which were derived from the same patient. (D) A comparison of genome-wide compartment tracks for all cell lines by nearest neighbor analysis reveals that LNCaP is a central node in a model axis to metastasis (highlighted yellow), connecting RWPE (nontransformed epithelium) to bone metastatic cell lines (VCaP and MDA cell lines). Atypical androgen independent lines cluster together along a different axis from LNCaP. (E) The distribution of eigen 1 values in all cell lines in the axis, for bins

that change from B to A (left,  $n = 624$ ) and from A to B (right,  $n = 273$ ) in LNCaP versus RWPE. Violin plots show the full distribution of values, while thick dashed line indicates median and thin dashed lines indicate 25th and 75th percentiles. Changes in distributions for all cells compared with RWPE are significant (Kruskal-Wallis,  $P < 0.0001$ ).

structural features at highly transcribed long genes (Heinz et al., 2018; Leidescher and Nübler, 2020 Preprint). Finally, the local structure can remain unchanged. For AR, this is the case in PC3 and DU145, where the gene remains in a disorganized region of the B compartment (Fig. 6 D).

Overall, for those genes that switch from the B to the A compartment, all types of changes happen with fairly even probability: 12.64% lose structure at the local level, 14.37% gain structure, 28.16% present stalled loops, 21.26% associate with structural deserts, and 23.56% do not change. However, genes located in the A to the B compartment switches predominantly

lose structure (39.06%) or do not change (21.88%). Only ~3% of A to B switch loci gain structure. The remaining loci are distributed evenly between stalled loop areas and structural desert change (Fig. 6 E). From this survey of prostate cancer cell lines, therefore, we also gain basic insight about the types of local structure change that most often accompany compartment level changes.

### The TMPRSS2-ERG locus shows an increase in local interactions in cell lines in the metastatic progression axis

As previously mentioned, the incidence of the *TMPRSS2-ERG* translocation has a positive correlation with prostate cancer

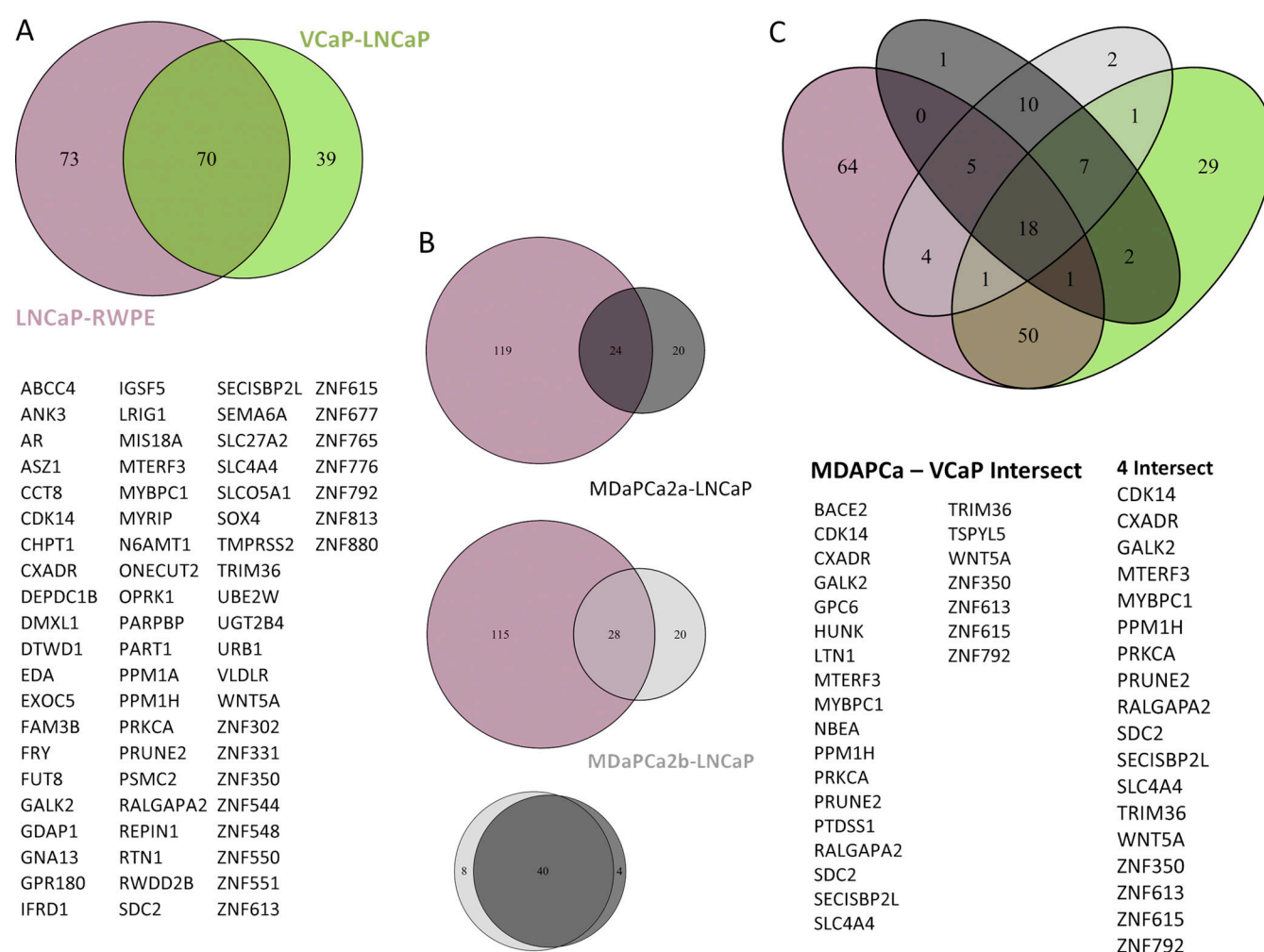


Figure 4. **Genes with expression increases that correlate with changes in B to A compartment identity.** (A) 143 genes identified as switching from the B to the A compartment show overexpression (twofold or greater increase) in LNCaP when compared with RWPE (purple). Of these, 70 are further overexpressed in VCaP (green) compared with LNCaP, denoting an exacerbation of this expression pattern with progression. This overlapping gene set includes known prostate cancer targets such as AR and *TMPRSS2*. (B) Of the genes up-regulated in LNCaP versus RWPE (purple), 24 and 28 genes are further overexpressed in MDAPCa2a (top) and MDAPCa2B (middle), respectively, when compared with LNCaP. Overlap of the genes overexpressed in the MDA cell lines compared with LNCaP shows a high concordance between these cell lines (bottom), which were derived from the same patient. (C) Genes that change compartment identity from B to A and are overexpressed in all bone metastatic cell lines, compared with LNCaP, which in turn are overexpressed in LNCaP compared with normal epithelium.

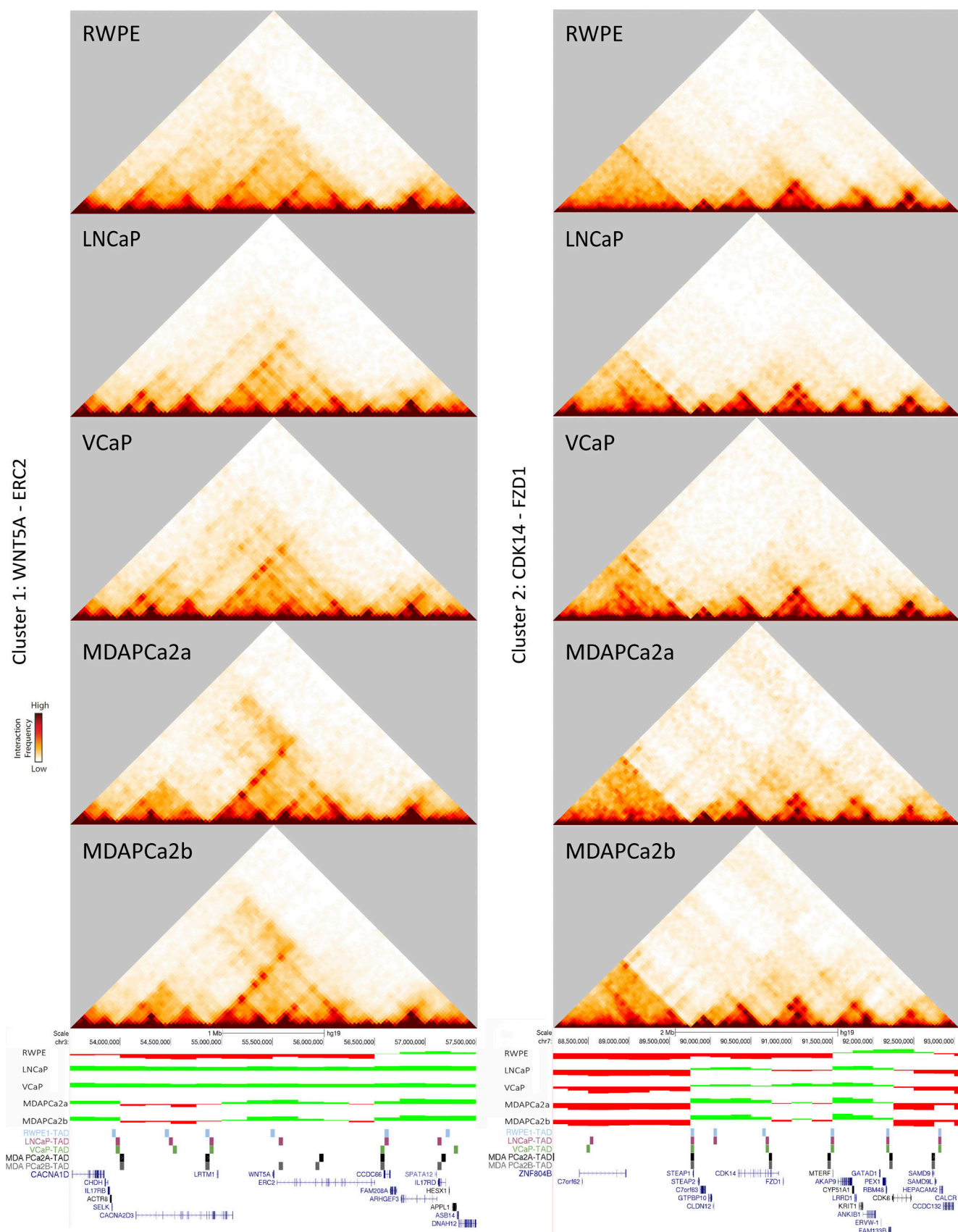


Figure 5. **Gene clusters that switch from the B to the A compartment include genes critical for prostate cancer progression.** 40-kb resolution heatmaps, compartment tracks, and TAD boundaries are shown for two representative clusters of genes that change compartment across the primary axis of

progression. Left: *WNT5A* switches from the B to the A compartment in cells of the primary progression axis. *WNT5A* sits at the location of a shifting TAD boundary, and strong interaction sites in the TAD surrounding this gene become more apparent in cancer progression as compared with RWPE. Right: *CDK14* switches from the B to the A compartment in most cancer cell lines. A TAD boundary localizes atop the gene body of *CDK14*.

progression and a poor prognosis. For the *TMPRSS2-ERG* locus, we observe that the region that encompasses *TMPRSS2*, *BACE2*, *PLAC4*, *MXI*, *MX2*, and *FAM3B* switches from the B compartment in RWPE to the A compartment in all cancer cell lines. In the model cell line for nontransformed epithelium RWPE, this genomic area is enclosed in a large TAD, including genes downstream of *TMPRSS2* (Fig. 7 A), as is also the case for LNCaP, 22RV1, MDAPCa2a, and PC3, with a slight shifting of the TAD boundary location. In LNCaP-C42B, an osteoblastic cell line derived from LNCaP, a sub-TAD appears within this locus, sequestering the *MXI* gene. Sub-TAD fragmentation also occurs in DU145, MDAPCa2b, and VCaP. In contrast, *ERG* is in the A compartment in RWPE, switching to the B compartment only in LNCaP and MDAPCa2b. The TAD boundary location around *ERG* does not change. Of note, the whole genomic area surrounding these two clusters is located in the A compartment in VCaP, a known carrier of the *TMPRSS2-ERG* translocation, evident as a high-interaction location in the Hi-C heatmap (Fig. 7 A, VCaP, rectangle).

In the bone metastatic cell line VCaP, the *TMPRSS2-ERG* translocation is apparent in the Hi-C heatmap of the long arm of chromosome 21 at a 250-kb resolution (Fig. 7 B). While a translocation is not evident in the MDAPCa cell lines, there is a distinct higher incidence of interactions close to that area: a log<sub>2</sub> ratio comparison against control (RWPE) shows that that chromosomal region is enriched for interactions in adenocarcinoma (LNCaP; Fig. 7 C), and that this phenotype is aggravated in both bone metastatic MDAPCa cell lines.

### Genes identified as switching compartments show altered expression in prostate cancer patients

To evaluate how the findings of our cell line-based model relate to patient data, we examined the expression data from the Prostate Adenocarcinoma cohort (PRAD) of The Cancer Genome Atlas (TCGA). A total of 549 patients were taken into consideration, comprising 52 normal and 497 tumor samples, stratified by Gleason score, an indication of cancer severity. First, we examined all genes in regions that switch from the A to the B compartment across our cell lines (as in Table S2). We observe a significant decrease in the pooled set of genes in tumor samples versus normal samples overall (Fig. 8 A). We find that 55% of genes in this set show significant decreases in tumor versus normal samples, many of which are relevant to prostate cancer (Fig. 8, B and C). This is a higher proportion of down-regulated genes than would be expected at random, as in 200 random gene sets of the same size, we never found any sets in which 55% or more of the genes were down-regulated in tumors (Fig. 8 C).

In contrast, when we consider the set of genes that switch from B to A (as in Table S2), we observe that they are actually not up-regulated more often than would be expected at random in PRAD tumor samples (Fig. 8 D). This is consistent with our model that movement into the A compartment is not

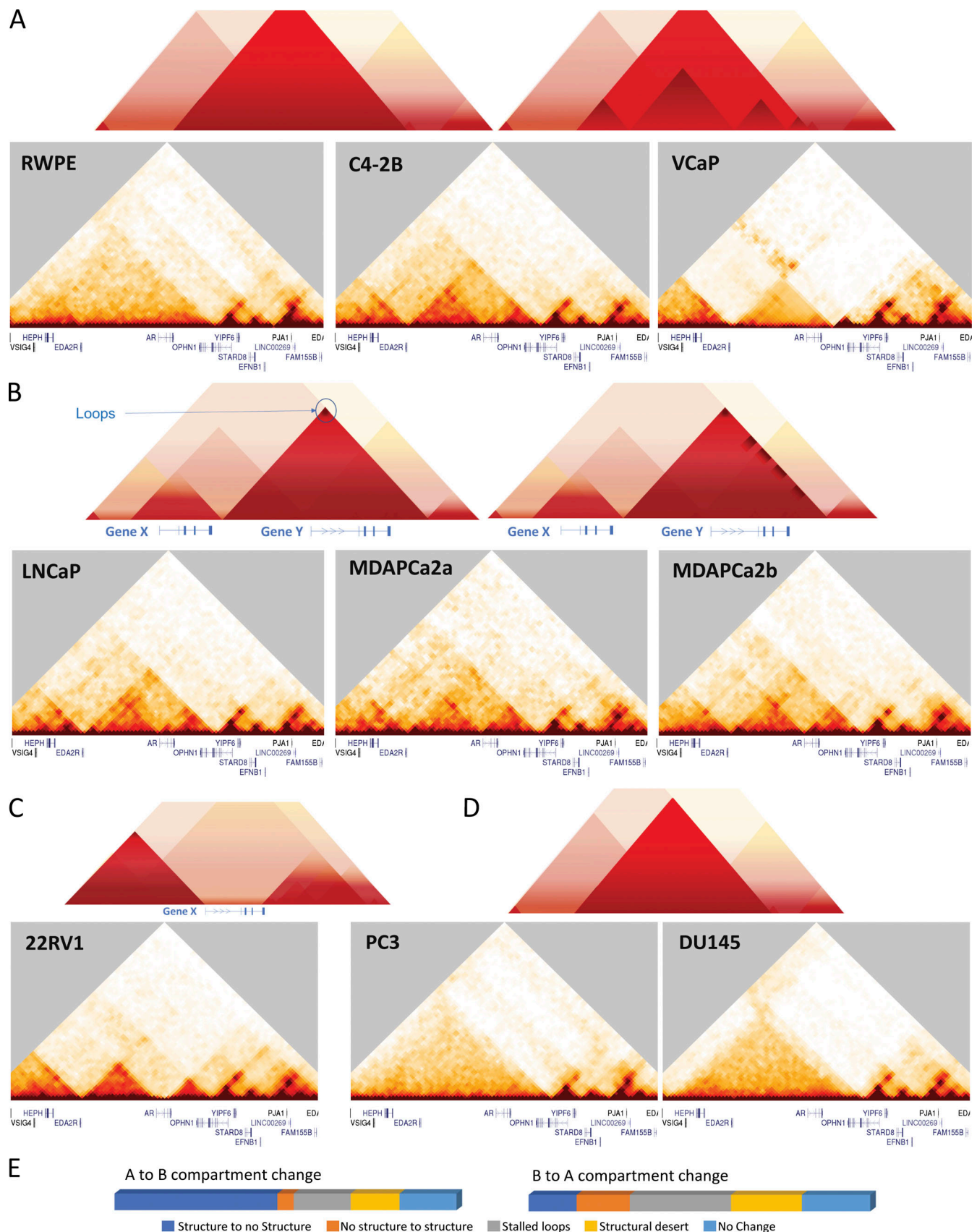
synonymous with gene activation, but instead only provides an opportunity for genes to be activated. But when we looked at the subset of genes which both change from the B to A compartment and change expression in our cell line models, we found that a more significant proportion of genes (47%) is up-regulated in patient tumor samples (Fig. 8, E and F). Only 6.5% of 200 random gene sets of the same size show this proportion of up-regulation. But some key genes in this set, such as *WNT5A* and *HUNK*, clearly do not show an increase in expression in patient tumor samples versus normal overall (Fig. 8 E). Instead, these genes show increasing expression with increasing Gleason score, a measure of more aggressive disease (Fig. 8 G). Indeed, most of the genes (75%) in this B to A switched subset that did not change expression in overall tumor versus normal instead show an increase in expression over cancer progression (Fig. 8 H). Thus, overall, a significant proportion of genes that switch from the B to the A compartment and changed expression in our cell line models either increase expression in patient tumors in general or increase their expression gradually with worsening disease in patients.

## Discussion

The deliberate, hierarchical organization of chromatin within the eukaryotic nucleus's constrained space is necessary for adequate DNA maintenance, repair, and gene transcription or silencing, all of which contribute to the cell's homeostasis. In this study, we have characterized the genomic architecture in a cohort of cells that model prostate cancer progression. At the genome-wide level, given the inherent high resolution of Hi-C data, we have identified small translocations that, to our knowledge, have not previously been reported in the literature.

Many regions across the genome are unchanged in their spatial compartmentalization across cell lines. This suggests that there are inherent 3D genome structure features of prostate epithelium that arise during initial differentiation and tissue patterning (Flyamer et al., 2017; Ke et al., 2017) and that these are persistent, regardless of malignancy status. These results suggest that genomic loci that switch compartment identity between the normal epithelium and cancer cells are associated with an oncogenic genomic architecture profile and that those features result from concerted biological events.

It is noteworthy that the majority of the compartment changes we identify involve a switch from the B to the A compartment. Correspondingly, we find in our compartment strength analyses that the A compartment becomes more intermixed, interacting more broadly, while the strongest B compartment regions remain more spatially segregated in the cancer cell lines. Both of these results point to a general shift in the prostate cancer lines toward a more open/poised-for-activation chromatin environment, which could lead to misactivation of oncogenes.



structure. TADs that appear C4-2B and VCaP are of different sizes and location. **(B)** Stalled loops. In addition to the appearance of a TAD structure, LNCaP, MDAPCa2a, and MDAPCa2b show enriched foci of interactions along the edges of TAD boundaries, consistent with previously documented cases of stalled loop extrusion. **(C)** Structural desert. In 22RV1, the entirety of the *AR* gene is in a region of depleted contacts. **(D)** No change. In PC3 and DU145, the *AR* locus remains disorganized as in RWPE. **(E)** Genome-wide distribution of above structural change categories across regions that change from A to B (left,  $n = 64$ ) or B to A (right,  $n = 174$ ) in either all four cell lines of the main axis or six or more of all cancer cell lines versus RWPE.

Our genome-wide compartment analysis revealed that compartment identity alone is enough to stratify prostate epithelium in a continuum of progression. LNCaP is a central node that connects the nontransformed epithelium (RWPE) to bone metastatic cell lines (MDAPCa2a/b and VCaP). These results provide further insight into the importance of cell line selection in progression studies: the atypical metastatic cell lines DU145 and PC3 cluster together away from the primary axis. This emphasizes that as researchers select cell lines for study, it is important to consider the differences in these cell lines and that not all will capture the most common pathways of metastasis.

It is important to consider that the compartment switching events observed here involve clusters of genes, which in some cases expand through several TADs. This observation echoes previous results in a plant system, where clusters of genes often changed their spatial compartmentalization and expression together (Nützmann et al., 2020). Interestingly, the clusters we observe include both genes previously identified in prostate cancer progression and others seemingly unrelated. Such is the case of the *WNT5a* locus, which is a known target of prostate tissue development, patterning, and cancer aggressiveness (Allgeier et al., 2008; Dai et al., 2008; Yamamoto et al., 2010; Huang et al., 2009). Our results show that in RWPE, *WNT5a* is sequestered within a TAD with the ELKS/RAB6-interacting/CAST family member 2 (*ERC2*). In contrast, TAD boundary shifting or eviction in the proximity of *WNT5a* in the bone metastatic cell lines represented in our primary axis for progression (VCaP and MDAPCa2a/b) results in TAD-limited interactions with leucine-rich repeats and transmembrane domains-containing protein 1 (*LRTM1*) instead. While this shift does not result in transcriptional induction of *LRTM1*, it could potentially lead to aberrant interactions between the promoters for both genes, ultimately resulting in the overexpression of *WNT5a*, as is also observed with increased Gleason score in patients. Since TAD structure is essential for proper gene regulation (Lupiañez et al., 2015; Rhie et al., 2019; Guo et al., 2018), this phenomenon requires further exploration. Still, it is attractive to consider TAD-targeted therapies that hold the potential to reverse the deleterious effects of TAD shifting. Interestingly, another target of noncanonical WNT signaling, *CDK14* (reviewed by Davidson and Niehrs, 2010), also switches from the B to the A compartment and is transcriptionally activated in our metastatic axis.

The clustering of different genes in the described compartment switches raises the interesting question of whether certain genes could act as a driver of compartment identity switches while neighboring genes act as “passengers.” For example, the transcriptional activation of one gene could influence a whole genomic region to switch to the A compartment. Indeed, previous work has shown that binding of transcriptional activators

can prefigure spatial compartment alterations (Stadhouders et al., 2018; Therizols et al., 2014). This compartment switch would then result in the switching of neighboring genes, which may increase their probability of later becoming activated as well. Recent work has shown that spatial reorganization of a chromosome region can make it more permissive for de-repression, even if the structural switch does not immediately change its expression level (Manjón et al., 2021 Preprint). Is it possible that earlier transcriptional events required for the cell to survive a particular insult trigger a full compartment shift? Is this, in turn, a potential trigger for oncogenic transcriptional activity? Such seems to be the case of the observed APP-*BACE2* axis along chromosome 21. APP presents with enhanced expression in the LNCaP-MDAPCa-VCaP axis compared with RWPE (Fig. S9). APP is also proximal to one of the persistent compartment switches in chromosome 21, but it does not change compartment itself: the A compartment identity atop the gene gets stronger with progression. Mounting evidence from the Alzheimer’s field, where abnormal amyloid processing results in aggregation and neurodegeneration, has shown that this protein and derived peptides serve a crucial role as antimicrobials and are necessary for mounting an appropriate host response to infection (thoroughly reviewed by Moir et al., 2018). If an environmental signal such as infection results in the prostate epithelium being exposed to excessive or chronic APP, could this trigger expression of its processing enzyme *BACE2*? Our evidence of the *BACE2* gene switching compartmentalization and experiencing higher levels of transcription in all three prototypical bone metastatic lines suggests this may be so. Critically, these events would imply the need for the compartment switch around the *TMPRSS2* locus.

Since its discovery, the *TMPRSS2-ERG* translocation in chromosome 21 has been recognized as an important indicator of poor prognosis and a higher risk of prostate cancer-related death (Tomlins et al., 2005; Perner et al., 2006; Demichelis et al., 2007; Hägglöf et al., 2014; Deplus et al., 2017). While a significant body of work has explored the role of *ERG* overexpression, its influence on *AR* signaling, and its association with chromatin structure (Li et al., 2020; Rickman et al., 2012), the genesis of this translocation and fusion remains poorly understood. Here, we implicate a 3D genome organization change of compartmentalization in this event. As mentioned before, *TMPRSS2* switches from the B to the A compartment consistently in all prostate cancer cell lines we queried. Meanwhile, *ERG* switches from the A to the B compartment in LNCaP and its nearest neighbor in our experimental metastatic axis, MDAPCa2b. Remarkably, *ERG* remains in the A compartment for VCaP and MDAPCa2a, suggesting a potential dual switching event during progression to more aggressive phenotypes. All these spatial rearrangements in such proximity result in the enhanced

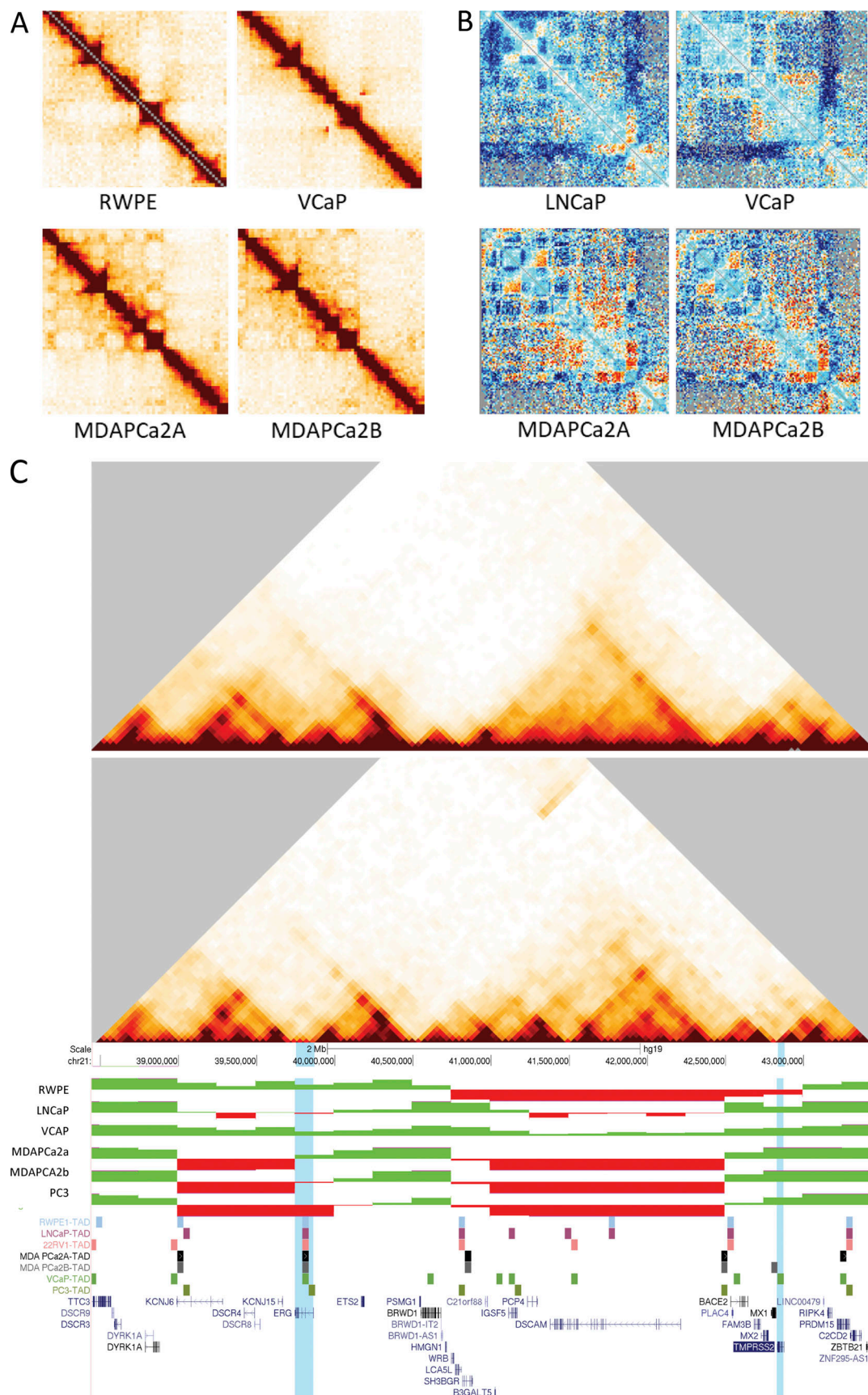


Figure 7. **Alterations in interactions around the *TMPPRSS2-ERG* locus in cell lines in the metastatic progression axis. (A)** Interaction heatmaps of the portion of chr21 containing *TMPPRSS2-ERG*: chromosome 21 (chr21): 30–45.2 Mb at 250-kb resolution. The *TMPPRSS2-ERG* translocation is visible in VCaP (blue

square) as a high interaction away from the diagonal. Interestingly, the other bone metastatic cell lines (MDAPCa2a and b) also show enhanced interactions close to that area. **(B)** Log2 ratio of interactions in each cell line compared with RWPE. Red areas have more interactions in the cell line compared with RWPE, while blue represents fewer interactions compared with RWPE. The area adjacent to the *TMPRSS2-ERG* translocation (red box) shows an increase in interactions in LNCaP and MDA cell lines. This increase of interactions in the region surrounding these genes is no longer evident in VCaP, where the *TMPRSS2-ERG* translocation has occurred. **(C)** *TMPRSS2* switches from the B to the A compartment in all cancer cell lines while *ERG* switches from A to B only on certain cell lines. The *TMPRSS2-ERG* translocation is easily visible in VCaP as a high interaction area away from the diagonal. Changes in TAD boundary locations around *TMPRSS2* are also observed.

contact incidence observed in all of these cell lines, even the early adenocarcinoma model LNCaP (as described in Fig. 7), and ultimately could lead to the translocation event, as it has been shown in other systems (McCord and Balajee, 2018, Zhang et al., 2012, Baca et al., 2013).

Recent efforts (Hawley et al., 2021 Preprint) have characterized 3D genomic profiles in prostate tumor cohorts. These studies recapitulate our findings that the 3D genome organization between malignant and benign prostate tissues remains largely consistent. We propose that prostate cancer progression is associated with specific changes in the 3D genome structure that arise early in the disease and facilitate an oncogenic expression phenotype. Based on these results, we can hypothesize that analyzing the 3D genome structure of patient-derived samples could be a prognostic marker for progression and bone metastasis.

## Materials and methods

### Cell lines

RWPE-1, LNCaP, DU145, 22RV1, VCaP, and PC3 cell lines were obtained from the Physical Sciences Oncology Network Bio-resource Core Facility, supported by ATCC (Manassas, VA).

All cell lines were cultured according to standard protocols, subculturing cells as they reached 80% confluency with the following media formulations: RWPE media was comprised of keratinocyte-specific media supplemented with EGF and bovine pituitary extract (Gibco; 17005042). RPMI (Gibco; 11835030) was supplemented to match the suggested formula by ATCC (4.5 g/liter glucose, 2.383 g/liter Hepes, and 0.11 g/liter sodium pyruvate) and 10% FBS (Corning; 35-010-CV). These media was used for LNCaP, 22RV1, and PC3. DU145 cells were cultured in DMEM supplemented with 10% FBS, and DMEM F12: Ham 1:1 (Gibco; 11-320-033) with 10% FBS was used for VCaP.

Cell lines MDAPCa2a and MDAPCa2b were a kind gift of Dr. Nora Navone (MD Anderson Cancer Center, Houston, TX). The identity of the received cell lines was verified by ATCC Human Cell STR testing services. These cell lines were grown in HPC1 media (Athena Enzyme Systems) supplemented with 10% FBS on standard cell culture T75 flasks coated with FCN Coating Mix (Athena Enzyme Systems) as per the manufacturer's instructions.

All media formulations were supplemented with 100 µg/ml penicillin-streptomycin (Gibco; 15-140-122).

### Hi-C

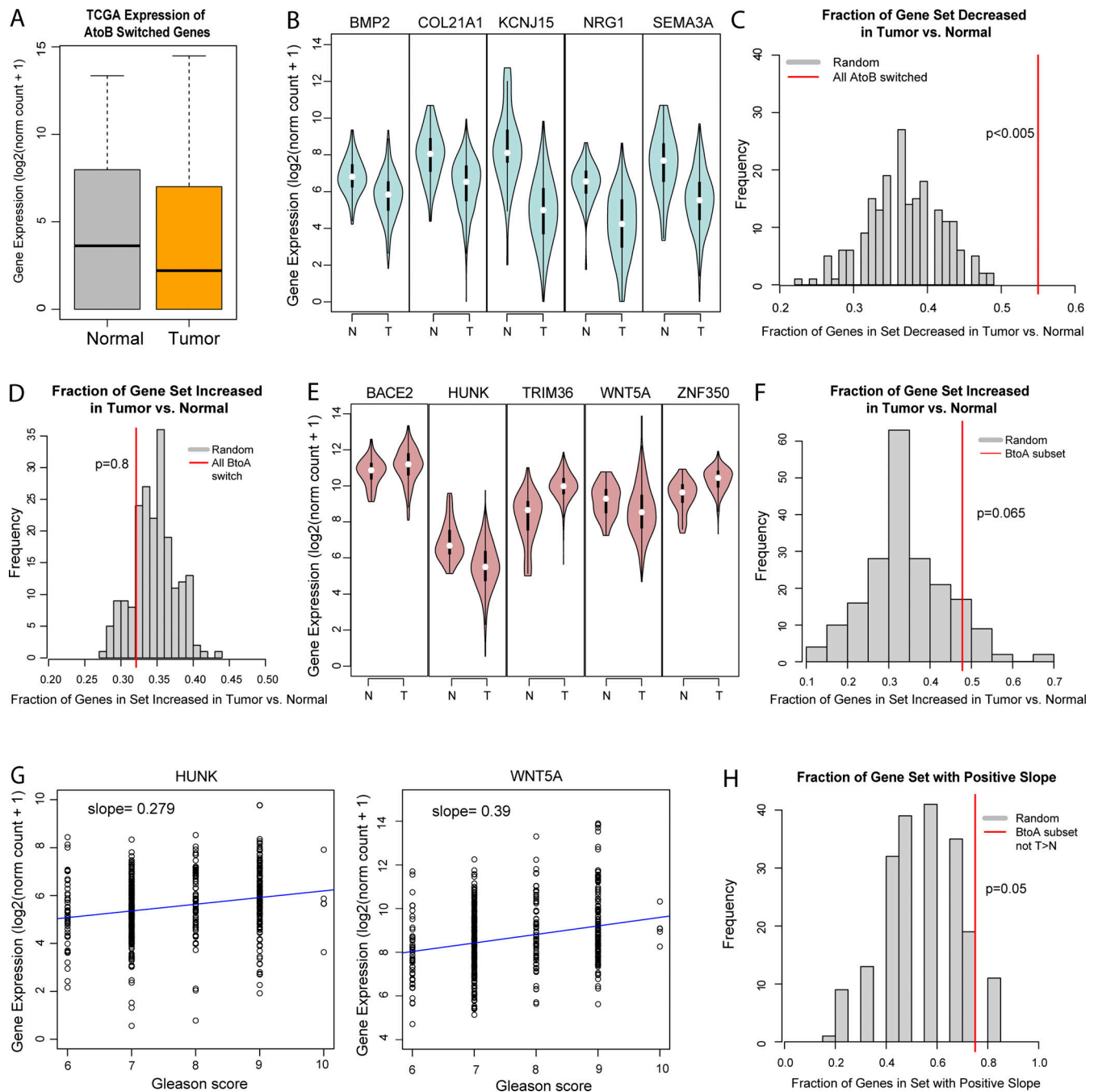
Cell pellets for Hi-C were prepared as previously described (Gollosi et al., 2018). Briefly, cells growing in monolayer in standard T75 flasks were quickly washed with 10 ml of HBSS

(Gibco; 14-025-134) at room temperature and cross-linked with 10 ml 1% formaldehyde (Fisher Bioreagents; BP531-25, in HBSS) for 10 min on a shaking platform. The cross-linking reaction was quenched by adding glycine to a final concentration of 0.14 M (MP Biomedicals; ICN19482591), followed by a 5-min incubation at room temperature, with shaking. After cooling down the plates on ice for 15 min, the formaldehyde solution was aspirated from the plate and substituted with 10 ml of ice-cold HBSS containing 1× Halt Protease Inhibitor cocktail (Thermo Fisher Scientific; PI78438). 5 million cell aliquots were collected by centrifugation and snap-frozen in liquid nitrogen.

VCaP, MDAPCa2a, and PC3 Hi-C was conducted as previously described (Gollosi et al., 2018). Briefly, cross-linked cells were lysed in cold hypotonic buffer, followed by dounce homogenization, and then nuclei were permeabilized by incubation in 0.1% SDS. After the permeabilization reaction was terminated by a final concentration of 1% Triton X-100, chromatin was digested overnight using 400 U DpnII (New England Biolabs) at 37 °C. Digested ends were filled in with biotin-dATP by Klenow DNA polymerase, and the blunt ends of interacting fragments were ligated together using T4 DNA ligase. DNA was then purified by phenol-chloroform extraction and ethanol precipitation. The ligated DNA was sonicated with a Covaris sonicator and then size-selected using Ampure XP beads to an average size of 200–400 bp. Biotinylated interacting fragments were pulled down using MyOne Streptavidin C1 beads (Invitrogen). For library preparation, the NEBNext Ultra II DNA Library prep kit (NEB) was used to carry out End Prep, Adaptor Ligation, and PCR amplification on streptavidin bead-bound DNA libraries.

Hi-C was also performed on LNCaP, DU145, VCaP, MDAPCa2a, MDAPCa2b, and PC3 using the Arima Hi-C (Arima Genomics) kit, following the manufacturer's protocol A160141 v01 for library amplification using the NEBNext Ultra II kit (NEB; E7645S). Sequencing was performed by Genewiz on either an Illumina NovaSeq or HiSeq platform with 50 or 150 bp paired-end reads. Sequenced reads were mapped to a reference human genome (hg19), binned, and iteratively corrected according to established pipelines (Imakaev et al., 2012) using the dekkerlab-cMapping tool available at <https://github.com/dekkerlab/cMapping>.

In addition, the same analysis above was performed on fastq files from previously published Hi-C datasets for RWPE, 22RV1, and LNCaP C4-2B: GSE118629 and GSE73782 (Rhie et al., 2019; Luo et al., 2017). See Table S1 for all data sources and statistics. Newly generated LNCaP results were checked for consistency with previously published LNCaP Hi-C results (ENCSR346DCU; Taberlay et al., 2016). Since our LNCaP data had a dramatically higher cis/trans ratio than these previously published datasets, comparisons were difficult, and so we proceeded with only our newly generated data.



**Figure 8. Genes that switch compartments show expression changes in TCGA patient samples. (A)** Genes in regions that switched from the A to B compartment from RWPE to cancer cell lines were queried in the TCGA-PRAD gene expression data for normal (norm;  $n = 52$ ) and tumor ( $n = 497$ ) samples. Black line indicates median while upper bound of box indicates 75th percentile of expression.  $P = 2.5 \times 10^{-15}$ , two-sided Wilcoxon test. **(B)** Expression of selected A to B switched genes in the PRAD cohort that shows decreases in tumor (T) compared with normal (N) samples. White dot indicates median expression, and thick black bar indicates range from 25th to 75th percentile. **(C)** Fraction of genes in the A to B switched set significantly down-regulated ( $P < 0.05$ ,  $t$  test) in tumor versus normal (red line, 0.55). 200 random gene sets were evaluated for significant down-regulation in the same way (gray histogram).  $P < 0.005$  represents chance of finding at least 55% of genes down-regulated in the random gene sets. **(D)** Fraction of genes significantly up-regulated in the B to A switched regions (red line) is not different from the fractions seen in 200 random gene sets (gray bars). **(E)** Expression of selected B to A switched genes in the PRAD cohort, comparing tumor (T) and normal (N) samples. **(F)** Examining only B to A switched genes also up-regulated in our bone metastatic cell lines, a larger fraction of genes are significantly up-regulated in tumor versus normal in PRAD cohort (red line), and this is less likely to be seen in 200 random gene sets of the same size (gray bars). **(G)** Example B to A switched genes that were not up-regulated in PRAD tumor versus normal overall but that show increasing expression (positive slope) when PRAD tumor samples are stratified by Gleason score (6–10). Linear fit shown in blue. **(H)** A large fraction (red line) of genes in the B to A switched set that were not up-regulated in PRAD tumor versus normal overall show a positive slope of expression over increasing Gleason score. This is a larger than the observed fraction of positive slope in most random gene sets of the same size.

## Analysis of Hi-C data

All Hi-C data analysis was performed using the existing cworld-dekker pipeline, available on github (<https://github.com/dekkerlab/cworld-dekker>) as follows.

Hi-C heatmaps were generated for genome-wide datasets at a 2.5-Mb resolution, and per chromosome at a 250-kb resolution, using the heatmap script.

Compartment analysis was performed via PCA on 250-kb binned matrices using the matrix2compartment script on each chromosome individually. Positive and negative PC1 values are initially arbitrary, so this script uses gene density to determine which set of bins (positive or negative) should be assigned as A (higher gene density) or B (lower gene density), respectively. PC1 profiles were visually inspected to ensure that they reflected a compartment-like pattern rather than a separation of chromosome arms or structural variant breaks. On some highly fragmented chromosomes, breaks between regions hindered proper compartment detection. In these cases, we detected breakpoints by finding extreme minima in insulation score profiles (Crane et al., 2015), creating separate matrices for each unbroken chunk of chromosome separately, and then running compartment analysis on each sub-region of the chromosome (Fig. S4).

The PC1 values per bin for the control cell line (RWPE) were subtracted from the values from each cancer cell line, resulting in a normalized distribution (referred to as  $\Delta$ PC1 analysis). Significant changes in compartment identity were defined as those bins whose subtracted value fell either under the mean minus 1.5 the SD or the mean plus 1.5 the SD, for at least six cell lines or all four cell lines in the primary axis, as defined by nearest neighbor analysis as described below (SPRING plot).

The genes contained in regions of interest determined from the  $\Delta$ PC1 analysis were annotated using the knownGene primary table as referenced in the University of California Santa Cruz (UCSC) Genome Browser (<https://genome.ucsc.edu/>).

To facilitate the visualization of chromatin compartmentalization, heatmaps for cis interactions were generated by first calculating the z-score of the interactions at 250-kb resolution compared with an expected interaction at each distance from the diagonal and then taking the Pearson correlation of each row and column of the heatmap (z-score correlation matrices). Multi-compartment track figures and overlays were constructed using the visualization tool Sushi (Phanstiel et al., 2014).

Genome-wide TAD boundaries were determined using the matrix2insulation script on 40-kb binned matrices, following the insulation score approach with an insulation square size of 500 kb (Crane et al., 2015).

## Calculation of A-A and B-B compartment interaction strengths

To calculate A-A and B-B compartment interaction strengths for each chromosome, distance corrected Hi-C intra-chromosomal interaction frequencies at 250-kb resolution were reordered according to their corresponding PC1 values (from strongest B to strongest A). Then, the reordered intra-chromosomal interaction matrix was smoothed at 500-kb resolution. Interactions were classified as A-A, B-B, and A-B and thresholded to include only the top 20% of interactions. The median value of each A-A, B-B, and A-B interaction is calculated.

Finally, the relative A-A and B-B compartment interaction strengths were obtained by subtracting the absolute A-B compartment interaction strength from the absolute A-A and B-B interaction compartment strengths, respectively.

For the average compartment interaction strength, the mean of the relative A-A and B-B interaction compartment strengths was calculated. A stronger A-A or B-B compartmentalization level compared with the A-B compartment intermixing would produce a higher positive value. On the other hand, a value close to zero suggests a weaker level of compartmentalization.

Saddle plots of compartment interaction strength were created by reordering a given chromosome matrix according to the PC1 values, from most negative (B compartment) to most positive (A compartment). Then, the interactions between each pair of bins were represented by the z-score, which compares the interaction frequency observed to the expected interactions between any two bins separated by that genomic distance. We subtract saddle plots from different cell types to observe the change in compartment interaction strength between them.

## Spring plot

To construct the spring plot (Weinreb et al., 2018; Miura et al., 2020) of the PCa cells based on the Hi-C compartmental data, the compartment profile of the cells at the 250-kb resolution was binarized. For example, the genomic regions where the compartment strength are  $>0$  were converted to 1 and the negative strengths to  $-1$ . The reason behind using that discretization step was to only consider the A/B compartment signature irrespective of the compartment strength. Once binarized, the data were then organized in a matrix format, where the rows are the genomic regions and columns are cells, and PCA was performed on that data. Since we were dealing with a small set of samples (cells) for our analysis, we kept all the principal components from the PCA transformation for further analysis. Then, a k-nearest neighbor graph with two nearest neighbors (includes the node itself) was constructed from the PCA-transformed data, and the network was visualized with a force-directed layout.

## Microarray

RNA was purified from 5 million cells at three different passages, per cell line, using the RNEasy mini kit (Qiagen; 74104) using QIAshredder (Qiagen; 79654) for homogenization. Purification was followed up by cleanup, and concentration using the RNase-Free DNase Set (Qiagen; 79254) and QIAquick PCR Purification Kit (Qiagen; 28104), respectively. RNA concentration was determined using a NanoDrop One (Thermo Fisher Scientific).

Clariom S microarray was performed through the Transcriptome Analysis Services (Transcriptome Profiling) from Thermo Fisher Scientific.

In addition, for cross-validation purposes, data for HG-U133 Plus2 microarray was also collected from the Encyclopedia of DNA Elements (ENCODE), as follows: RWPE (GSM966512, GSM966513, GSM966514), LNCaP (GSM2571978, GSM2571979, GSM2571980), DU145 (GSM1374469), PC3 (GSM1517530, GSM1517531, GSM1517532), LNCaP-C4-2B (GSM1565257, GSM1565258), and 22RV1 (GSM2571966, GSM2571967, GSM2571968).

Analysis of microarray data was performed using the Transcription Analysis Console from Applied Biosystems (Thermo Fisher Scientific), curating the log<sub>2</sub> fold up-regulated/down-regulated genes ( $P < 0.05$ ) with targets identified in the  $\Delta$ PC1 analysis.

### Analysis of TCGA data

Gene expression of all the genes identified as switching compartments in our model cell lines (Table S2) was analyzed on the TCGA-PRAD using the USCS Xena Browser interface (Goldman et al., 2020). Data for normalized gene counts for all samples were downloaded from <https://tcga-xena-hub.s3.us-east-1.amazonaws.com/download/TCGA.PRAD.sampleMap%2FHiSeqV2.gz>. Samples were annotated with Gleason score labels provided by TCGA. With custom R scripts, we statistically evaluated the expression change between tumor and normal samples for each individual gene (one-sided  $t$  test) and calculated the proportion of a given gene set with changes ( $P < 0.05$ ) in the direction we expected (that is, up-regulation for B to A switched gene sets and down-regulation for A to B switched gene sets). We then selected random gene sets of the same size as our query set for comparison. 200 independent sets of random genes were selected for each comparison (that is, if we were considering 87 genes that switched A to B, we would choose 200 sets of 87 randomly selected genes). The proportion of genes called significantly up or down-regulated ( $P < 0.05$  in one-sided  $t$  test) was calculated in each of the 200 random gene sets. The fraction of random gene sets containing the same or greater proportion of up or down-regulated genes was used to evaluate whether the amount of up- or down-regulation observed in the compartment switch gene sets was likely to be seen by chance. Note that by comparing proportions of genes that met a certain statistical threshold between real and random sets, we account for problems of multiple hypothesis testing. On the log count scale, most of the gene expression distributions are not highly skewed by outliers, so the mean expression was a representative metric for evaluation by the  $t$  test. We did not test for normality of the gene expression data distributions before using the  $t$  test because the  $t$  test does not require the underlying data to be normally distributed (Boneau, 1960; Norman, 2010). Instead, the  $t$  test assumes normality in the distribution of the sample means, a condition that is met, according to the central limit theorem, for sample sizes larger than 5–10, which we exceed in this case (Norman, 2010). Slopes of gene expression across Gleason score were calculated in R with a linear fit. The proportion of genes in a set with positive slope was evaluated and compared with the proportion of genes with positive slope in 200 sets of a matched number of genes chosen at random.

### Online supplemental material

Fig. S1, Fig. S2, and Fig. S3 show that Hi-C data at 2.5-Mb resolution recapitulate previously reported and novel translocations in the cell lines in this study. Fig. S4 shows how we modified our compartment analysis to consider heavily fragmented chromosomes. Fig. S5 shows a comparison of Hi-C heatmaps, cis interaction heatmaps, and compartment analysis for chromosome X for RWPE, LNCaP, and VCaP. Fig. S6

characterizes differences in genome-wide chromosome compartmentalization for all cells in this study. Fig. S7 validates the consistency of the genome-wide compartmentalization between the RWPE cell line (used as control for this study) and PrEC, a human primary prostate epithelial cell line. Fig. S8 includes examples of gene clusters that switch compartments in advanced prostate cancer. Fig. S9 highlights the APP-BACE2 region, located on chromosome 21, and its potential role as a driver in compartment switching of neighboring genes. Table S1 includes the database reference and quality control metrics for this study. Table S2 lists the genes identified as switching compartment identity. Table S3 denotes the gene clusters that switch compartment identity, as related to prostate cancer progression.

### Data availability

All Hi-C and microarray data are available on GEO at accession no. GSE172099. Other processed data figures and a UCSC Genome Browser Track Hub containing all data are available for browsing at <https://3dgenome.utk.edu/3d-genome-architecture-in-prostate-cancer-progression/>.

### Acknowledgments

We thank Jeremy Hughes (Web Communications Manager, University of Tennessee, Knoxville. Department of Arts & Sciences) for his help in constructing the accompanying website. We thank Nora Navone, Ph.D. (MD Anderson Cancer Center, Houston, TX) for providing the MDAPCa cell lines.

This work was supported, in part, by National Institutes of Health National Institute of General Medical Sciences grant R35GM133557 to R.P. McCord. R. San Martin was supported by a postdoctoral fellowship from the American Cancer Society (134060-PF-19-183-01-CSM).

The authors declare no competing financial interests.

Author contributions: Conceptualization: R. San Martin and R.P. McCord; investigation: R. San Martin and R. Gollosi; formal analysis: R. San Martin, R.P. McCord, P. Das, J.T. Sanders, and R. Dos Reis Marques; visualization: R. San Martin, R.P. McCord, P. Das, Y. Xu, and R. Dos Reis Marques; resources and validation: R. San Martin and J.M. Roberts; writing – original draft: R. San Martin and R.P. McCord; software: P. Das, Y. Xu, and R.P. McCord; writing – review and editing: all authors; supervision: R.P. McCord.

Submitted: 23 April 2021

Revised: 31 August 2021

Accepted: 17 November 2021

### References

- Ahmadiyeh, N., M.M. Pomerantz, C. Grisanzio, P. Herman, L. Jia, V. Al-mendro, H.H. He, M. Brown, X.S. Liu, M. Davis, et al. 2010. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc. Natl. Acad. Sci. USA*. 107:9742–9746. <https://doi.org/10.1073/pnas.0910668107>
- Allgeier, S.H., T.M. Lin, C.M. Vezina, R.W. Moore, W.A. Fritz, S.Y. Chiu, C. Zhang, and R.E. Peterson. 2008. WNT5A selectively inhibits mouse ventral prostate development. *Dev. Biol.* 324:10–17. <https://doi.org/10.1016/j.ydbio.2008.08.018>

- Baca, S.C., D. Prandi, M.S. Lawrence, J.M. Mosquera, A. Romanel, Y. Drier, K. Park, N. Kitabayashi, T.Y. MacDonald, M. Ghandi, et al. 2013. Punctuated evolution of prostate cancer genomes. *Cell*. 153:666–677. <https://doi.org/10.1016/j.cell.2013.03.021>
- Balajee, A.S., J.T. Sanders, R. Gollosi, I. Shuryak, R.P. McCord, and N. Dai-niak. 2018. Investigation of Spatial Organization of Chromosome Territories in Chromosome Exchange Aberrations After Ionizing Radiation Exposure. *Health Phys.* 115:77–89. <https://doi.org/10.1097/HP.0000000000000840>
- Barutcu, A.R., B.R. Lajoie, R.P. McCord, C.E. Tye, D. Hong, T.L. Messier, G. Browne, A.J. van Wijnen, J.B. Lian, J.L. Stein, et al. 2015. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol.* 16: 214. <https://doi.org/10.1186/s13059-015-0768-0>
- Bello, D., M.M. Webber, H.K. Kleinman, D.D. Waringer, and J.S. Rhim. 1997. Androgen responsive adult human prostatic epithelial cell lines immortalized by human papillomavirus 18. *Carcinogenesis*. 18:1215–1223. <https://doi.org/10.1093/carcin/18.6.1215>
- Boneau, C.A. 1960. The effects of violations of assumptions underlying the test. *Psychol. Bull.* 57:49–64. <https://doi.org/10.1037/h0041412>
- Bubendorf, L., A. Schöpfer, U. Wagner, G. Sauter, H. Moch, N. Willi, T.C. Gasser, and M.J. Mihatsch. 2000. Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Hum. Pathol.* 31:578–583. <https://doi.org/10.1053/hp.2000.6698>
- Crane, E., Q. Bian, R.P. McCord, B.R. Lajoie, B.S. Wheeler, E.J. Ralston, S. Uzawa, J. Dekker, and B.J. Meyer. 2015. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*. 523: 240–244. <https://doi.org/10.1038/nature14450>
- Dai, J., C.L. Hall, J. Escara-Wilke, A. Mizokami, J.M. Keller, and E.T. Keller. 2008. Prostate cancer induces bone metastasis through Wnt-induced bone morphogenetic protein-dependent and independent mechanisms. *Cancer Res.* 68:5785–5794. <https://doi.org/10.1158/0008-5472.CAN-07-6541>
- Danila, D.C., G. Heller, G.A. Gignac, R. Gonzalez-Espinoza, A. Anand, E. Tanaka, H. Lilja, L. Schwartz, S. Larson, M. Fleisher, and H.I. Scher. 2007. Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. *Clin. Cancer Res.* 13:7053–7058. <https://doi.org/10.1158/1078-0432.CCR-07-1506>
- Davidson, G., and C. Niehrs. 2010. Emerging links between CDK cell cycle regulators and Wnt signaling. *Trends Cell Biol.* 20:453–460. <https://doi.org/10.1016/j.tcb.2010.05.002>
- Dekker, J., and L. Mirny. 2016. The 3D Genome as Moderator of Chromosomal Communication. *Cell*. 164:1110–1121. <https://doi.org/10.1016/j.cell.2016.02.007>
- Demichelis, F., K. Fall, S. Perner, O. Andrén, F. Schmidt, S.R. Setlur, Y. Hoshida, J.M. Mosquera, Y. Pawitan, C. Lee, et al. 2007. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene*. 26:4596–4599. <https://doi.org/10.1038/sj.onc.1210237>
- Deplus, R., C. Delliaux, N. Marchand, A. Flourens, N. Vanpouille, X. Leroy, Y. de Launoit, and M. Dutertre-Coquillaud. 2017. TMPRSS2-ERG fusion promotes prostate cancer metastases in bone. *Oncotarget*. 8:11827–11840. <https://doi.org/10.18632/oncotarget.14399>
- Diamond, D.A., S.J. Berry, H.J. Jewett, J.C. Eggleston, and D.S. Coffey. 1982. A new method to assess metastatic potential of human prostate cancer: relative nuclear roundness. *J. Urol.* 128:729–734. [https://doi.org/10.1016/S0022-5347\(17\)53158-4](https://doi.org/10.1016/S0022-5347(17)53158-4)
- Dixon, J.R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J.S. Liu, and B. Ren. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485:376–380. <https://doi.org/10.1038/nature11082>
- Du, M., L. Tillmans, J. Gao, P. Gao, T. Yuan, R.L. Dittmar, W. Song, Y. Yang, N. Sahr, T. Wang, et al. 2016. Chromatin interactions and candidate genes at ten prostate cancer risk loci. *Sci. Rep.* 6:23202. <https://doi.org/10.1038/srep23202>
- Engreitz, J.M., V. Agarwala, and L.A. Mirny. 2012. Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. *PLoS One*. 7:e44196. <https://doi.org/10.1371/journal.pone.0044196>
- Flyamer, I.M., J. Gassler, M. Imakaev, H.B. Brandão, S.V. Ulianov, N. Abdennur, S.V. Razin, L.A. Mirny, and K. Tachibana-Konwalski. 2017. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*. 544:110–114. <https://doi.org/10.1038/nature21711>
- Gerlitz, G., and M. Bustin. 2010. Efficient cell migration requires global chromatin condensation. *J. Cell Sci.* 123:2207–2217. <https://doi.org/10.1242/jcs.058271>
- Goldman, M.J., B. Craft, M. Hastie, K. Repečka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A.N. Brooks, et al. 2020. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* 38:675–678. <https://doi.org/10.1038/s41587-020-0546-8>
- Gollosi, R., J.T. Sanders, and R.P. McCord. 2018. Iteratively improving Hi-C experiments one step at a time. *Methods*. 142:47–58. <https://doi.org/10.1016/j.jymeth.2018.04.033>
- Guo, Y., A.A. Perez, D.J. Hazelett, G.A. Coetzee, S.K. Rhie, and P.J. Farnham. 2018. CRISPR-mediated deletion of prostate cancer risk-associated CTCF loop anchors identifies repressive chromatin loops. *Genome Biol.* 19:160. <https://doi.org/10.1186/s13059-018-1531-0>
- Häggglöf, C., P. Hammarsten, K. Strömwall, L. Egevad, A. Josefsson, P. Stattin, T. Granfors, and A. Bergh. 2014. TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. *PLoS One*. 9:e86824. <https://doi.org/10.1371/journal.pone.0086824>
- Hawley, J.R., S. Zhou, C. Arlidge, G. Grillo, K. Kron, R. Hugh-White, T. Van Der Kwast, M. Fraser, P.C. Boutros, R.G. Bristow, et al. 2021. Cis-regulatory Element Hijacking Overshadows Topological Changes in Prostate Cancer. *bioRxiv*. 2021.01.05.425333 (Preprint posted April 27, 2021)
- Heinz, S., L. Texari, M.G.B. Hayes, M. Urbanowski, M.W. Chang, N. Givarkes, A. Rialdi, K.M. White, R.A. Albrecht, L. Pache, et al. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell*. 174:1522–1536.e22. <https://doi.org/10.1016/j.cell.2018.07.047>
- Helfand, B.T., K.A. Roehl, P.R. Cooper, B.B. McGuire, L.M. Fitzgerald, G. Cancel-Tassin, J.N. Cornu, S. Bauer, E.L. Van Blarigan, X. Chen, et al. 2015. Associations of prostate cancer risk variants with disease aggressiveness: results of the NCI-SPORE Genetics Working Group analysis of 18,343 cases. *Hum. Genet.* 134:439–450. <https://doi.org/10.1007/s00439-015-1534-9>
- Hernandez, R.K., S.W. Wade, A. Reich, M. Pirulli, A. Liede, and G.H. Lyman. 2018. Incidence of bone metastases in patients with solid tumors: analysis of oncology electronic medical records in the United States. *BMC Cancer*. 18:44. <https://doi.org/10.1186/s12885-017-3922-0>
- Hnisz, D., A.S. Weintraub, D.S. Day, A.L. Valton, R.O. Bak, C.H. Li, J. Goldmann, B.R. Lajoie, Z.P. Fan, A.A. Sigova, et al. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*. 351: 1454–1458. <https://doi.org/10.1126/science.1249024>
- Horoszewicz, J.S., S.S. Leong, E. Kawinski, J.P. Karr, H. Rosenthal, T.M. Chu, E.A. Mirand, and G.P. Murphy. 1983. LNCaP model of human prostatic carcinoma. *Cancer Res.* 43:1809–1818.
- Huang, L., Y. Pu, W.Y. Hu, L. Birch, D. Luccio-Camelo, T. Yamaguchi, and G.S. Prins. 2009. The role of Wnt5a in prostate gland development. *Dev. Biol.* 328:188–199. <https://doi.org/10.1016/j.ydbio.2009.01.003>
- Imakaev, M., G. Fudenberg, R.P. McCord, N. Naumova, A. Goloborodko, B.R. Lajoie, J. Dekker, and L.A. Mirny. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods*. 9: 999–1003. <https://doi.org/10.1038/nmeth.2148>
- Jacobs, S.C. 1983. Spread of prostatic cancer to bone. *Urology*. 21:337–344. [https://doi.org/10.1016/0090-4295\(83\)90147-4](https://doi.org/10.1016/0090-4295(83)90147-4)
- Kaighn, M.E., K.S. Narayan, Y. Ohnuki, J.F. Lechner, and L.W. Jones. 1979. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest. Urol.* 17:16–23.
- Ke, Y., Y. Xu, X. Chen, S. Feng, Z. Liu, Y. Sun, X. Yao, F. Li, W. Zhu, L. Gao, et al. 2017. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell*. 170: 367–381.e20. <https://doi.org/10.1016/j.cell.2017.06.029>
- Korenchuk, S., J.E. Lehr, L. McLean, Y.G. Lee, S. Whitney, R. Vessella, D.L. Lin, and K.J. Pienta. 2001. VCaP, a cell-based model system of human prostate cancer. *In Vivo*, 15, 163–8.
- Leidescher, S., and J. Nübner. Y. Feodorova, E. Hildebrand, S. Ullrich, S. Bultmann, S. Link, K. Thanisch, J. Dekker, H. Leonhardt, et al. 2020. Spatial organization of transcribed eukaryotic genes. *bioRxiv*. 2020.05.20.106591 (Preprint posted April 13, 2021)
- Li, F., Q. Yuan, W. Di, X. Xia, Z. Liu, N. Mao, L. Li, C. Li, J. He, Y. Li, et al. 2020. ERG orchestrates chromatin interactions to drive prostate cell fate reprogramming. *J. Clin. Invest.* 130:5924–5941. <https://doi.org/10.1172/JCI137967>
- Lieberman-Aiden, E., N.L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B.R. Lajoie, P.J. Sabo, M.O. Dorschner, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 326:289–293. <https://doi.org/10.1126/science.1181369>
- Luo, Z., S.K. Rhie, F.D. Lay, and P.J. Farnham. 2017. A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. *Cell Rep.* 21:1411–1417. <https://doi.org/10.1016/j.celrep.2017.10.048>

- Lupiáñez, D.G., K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, D. Horn, H. Kayserili, J.M. Opitz, R. Laxova, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. 161:1012–1025. <https://doi.org/10.1016/j.cell.2015.04.004>
- Mani, R.S., S.A. Tomlins, K. Callahan, A. Ghosh, M.K. Nyati, S. Varambally, N. Palanisamy, and A.M. Chinnaiyan. 2009. Induced chromosomal proximity and gene fusions in prostate cancer. *Science*. 326:1230. <https://doi.org/10.1126/science.1178124>
- Manjón, A.G., D.P. Hupkes, N.Q. Liu, A. Friskes, S. Joosten, H. Teunissen, M. Aarts, S. Prekovic, W. Zwart, E. De Wit, et al. 2021. Perturbations in 3D genome organization can promote acquired drug resistance. *bioRxiv*. 2021.02.02.429315 (Preprint posted February 2, 2021)
- McCord, R.P., and A. Balajee. 2018. 3D Genome Organization Influences the Chromosome Translocation Pattern. *Adv. Exp. Med. Biol.* 1044:113–133. [https://doi.org/10.1007/978-981-13-0593-1\\_8](https://doi.org/10.1007/978-981-13-0593-1_8)
- Meaburn, K.J., and T. Misteli. 2007. Cell biology: chromosome territories. *Nature*. 445:379–381. <https://doi.org/10.1038/445379a>
- Miura, H., S. Takahashi, T. Shibata, K. Nagao, C. Obuse, K. Okumura, M. Ogata, I. Hiratani, and S.-I. Takebayashi. 2020. Mapping replication timing domains genome wide in single mammalian cells with single-cell DNA replication sequencing. *Nat Protoc.* 15:4058–4100. <https://doi.org/10.1038/s41596-020-0378-5>
- Moir, R.D., R. Lathe, and R.E. Tanzi. 2018. The antimicrobial protection hypothesis of Alzheimer's disease. *Alzheimers Dement.* 14:1602–1614. <https://doi.org/10.1016/j.jalz.2018.06.3040>
- Moreno, J.G., M.C. Miller, S. Gross, W.J. Allard, L.G. Gomella, and L.W. Terstappen. 2005. Circulating tumor cells predict survival in patients with metastatic prostate cancer. *Urology*. 65:713–718. <https://doi.org/10.1016/j.urol.2004.11.006>
- Navone, N.M., M. Olive, M. Ozen, R. Davis, P. Troncoso, S.M. Tu, D. Johnston, A. Pollack, S. Pathak, A.C. von Eschenbach, and C.J. Logothetis. 1997. Establishment of two human prostate cancer cell lines derived from a single bone metastasis. *Clin. Cancer Res.* 3:2493–2500.
- Nora, E.P., B.R. Lajoie, E.G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N.L. van Berkum, J. Meisig, J. Sedat, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485:381–385. <https://doi.org/10.1038/nature11049>
- Nora, E.P., A. Goloborodko, A.L. Valton, J.H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L.A. Mirny, and B.G. Bruneau. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*. 169:930–944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>
- Norman, G. 2010. Likert scales, levels of measurement and the “laws” of statistics. *Adv. Health Sci. Educ. Theory Pract.* 15:625–632. <https://doi.org/10.1007/s10459-010-9222-y>
- Nuebler, J., G. Fudenberg, M. Imakaev, N. Abdennur, and L.A. Mirny. 2018. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. USA*. 115:E6697–E6706. <https://doi.org/10.1073/pnas.1717730115>
- Nützmann, H.W., D. Doerr, A. Ramírez-Colmenero, J.E. Sotelo-Fonseca, E. Wegel, M. Di Stefano, S.W. Wingett, P. Fraser, L. Hurst, S.L. Fernandez-Valverde, and A. Osbourn. 2020. Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *Proc. Natl. Acad. Sci. USA*. 117:13800–13809. <https://doi.org/10.1073/pnas.1920474117>
- Pan, Y., S. Kytölä, F. Farnebo, N. Wang, W.O. Lui, N. Nupponen, J. Isola, T. Visakorpi, U.S. Bergerheim, and C. Larsson. 1999. Characterization of chromosomal abnormalities in prostate cancer cell lines by spectral karyotyping. *Cytogenet. Cell Genet.* 87:225–232. <https://doi.org/10.1159/000015432>
- Partin, A.W., R.H. Getzenberg, M.J. CarMichael, D. Vindivich, J. Yoo, J.I. Epstein, and D.S. Coffey. 1993. Nuclear matrix protein patterns in human benign prostatic hyperplasia and prostate cancer. *Cancer Res.* 53:744–746.
- Perner, S., F. Demichelis, R. Beroukhi, F.H. Schmidt, J.M. Mosquera, S. Setlur, J. Tchinda, S.A. Tomlins, M.D. Hofer, K.G. Pienta, et al. 2006. TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res.* 66:8337–8341. <https://doi.org/10.1158/0008-5472.CAN-06-1482>
- Phanstiel, D.H., A.P. Boyle, C.L. Araya, and M.P. Snyder. 2014. SushiR: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. *Bioinformatics*. 30:2808–2810. <https://doi.org/10.1093/bioinformatics/btu379>
- Pope, B.D., T. Ryba, V. Dileep, F. Yue, W. Wu, O. Denas, D.L. Vera, Y. Wang, R.S. Hansen, T.K. Canfield, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature*. 515:402–405. <https://doi.org/10.1038/nature13986>
- Rhie, S.K., A.A. Perez, F.D. Lay, S. Schreiner, J. Shi, J. Polin, and P.J. Farnham. 2019. A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat. Commun.* 10:4154. <https://doi.org/10.1038/s41467-019-12079-8>
- Rickman, D.S., T.D. Soong, B. Moss, J.M. Mosquera, J. Dlabal, S. Terry, T.Y. MacDonald, J. Tripodi, K. Bunting, V. Najfeld, et al. 2012. Oncogene-mediated alterations in chromatin conformation. *Proc. Natl. Acad. Sci. USA*. 109:9083–9088. <https://doi.org/10.1073/pnas.112570109>
- Roh, M., O.E. Franco, S.W. Hayward, R. van der Meer, and S.A. Abdulkadir. 2008. A Role for Polyploidy in the Tumorigenicity of Pim-1-Expressing Human Prostate and Mammary Epithelial Cells. *PLoS ONE*. 3:e2572. <https://doi.org/10.1371/journal.pone.0002572>
- Siegel, R.L., K.D. Miller, and A. Jemal. 2020. Cancer statistics, 2020. *CA Cancer J. Clin.* 70:7–30. <https://doi.org/10.3322/caac.21590>
- Sramkoski, R.M., T.G. Pretlow II, J.M. Giacomia, T.P. Pretlow, S. Schwartz, M.S. Sy, S.R. Marengo, J.S. Rhim, D. Zhang, and J.W. Jacobberger. 1999. A new human prostate carcinoma cell line, 22Rv1. *In Vitro Cell Dev. Biol. Anim.* 35:403–409.
- Stadhouders, R., E. Vidal, F. Serra, B. Di Stefano, F. Le Dily, J. Quilez, A. Gomez, S. Collombet, C. Berenguer, Y. Cuartero, et al. 2018. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* 50:238–249. <https://doi.org/10.1038/s41588-017-0030-7>
- Stephens, A.D., P.Z. Liu, E.J. Banigan, L.M. Almashalha, V. Backman, S.A. Adam, R.D. Goldman, and J.F. Marko. 2018. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol. Biol. Cell*. 29:220–233. <https://doi.org/10.1091/mbc.E17-06-0410>
- Stone, K.R., D.D. Mickey, H. Wunderli, G.H. Mickey, and D.F. Paulson. 1978. Isolation of a human prostate carcinoma cell line (DU 145). *Int. J. Cancer*. 21:274–281. <https://doi.org/10.1002/ijc.2910210305>
- Taberlay, P.C., J. Achinger-Kawecka, A.T. Lun, F.A. Buske, K. Sabir, C.M. Gould, E. Zotenko, S.A. Bert, K.A. Giles, D.C. Bauer, et al. 2016. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* 26:719–731. <https://doi.org/10.1101/gr.201517.115>
- Thalman, G.N., R.A. Sikes, T.T. Wu, A. Degeorges, S.M. Chang, M. Ozen, S. Pathak, and L.W. Chung. 2000. LNCaP progression model of human prostate cancer: androgen-independence and osseous metastasis. *Prostate*. 44:91–103.
- Therizols, P., R.S. Illingworth, C. Courilleau, S. Boyle, A.J. Wood, and W.A. Bickmore. 2014. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science*. 346:1238–1242. <https://doi.org/10.1126/science.1259587>
- Todenhofer, T., E.S. Park, S. Duffy, X. Deng, C. Jin, H. Abdi, H. Ma, and P.C. Black. 2016. Microfluidic enrichment of circulating tumor cells in patients with clinically localized prostate cancer. *Urol. Oncol.* 34:483.e9–483.e16.
- Tomlins, S.A., D.R. Rhodes, S. Perner, S.M. Dhanasekaran, R. Mehra, X.W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer, et al. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 310:644–648. <https://doi.org/10.1126/science.1117679>
- van Bokhoven, A., A. Caires, M.D. Maria, A.P. Schulte, M.S. Lucia, S.K. Nordeen, G.J. Miller, and M. Varella-Garcia. 2003. Spectral karyotype (SKY) analysis of human prostate carcinoma cell lines. *Prostate*. 57:226–244. <https://doi.org/10.1002/pros.10291>
- Verdone, J.E., P. Parsana, R.W. Veltri, and K.J. Pienta. 2015. Epithelial-mesenchymal transition in prostate cancer is associated with quantifiable changes in nuclear structure. *Prostate*. 75:218–224. <https://doi.org/10.1002/pros.22908>
- Weinreb, Caleb, Samuel Wolock, and Allon M Klein. 2018. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics*. 34(7):1246–1248. <https://doi.org/10.1093/bioinformatics/btx792>
- Yamamoto, H., N. Oue, A. Sato, Y. Hasegawa, H. Yamamoto, A. Matsubara, W. Yasui, and A. Kikuchi. 2010. Wnt5a signaling is involved in the aggressiveness of prostate cancer and expression of metalloproteinase. *Oncogene*. 29:2036–2046. <https://doi.org/10.1038/ncr.2009.496>
- Zhang, Y., R.P. McCord, Y.J. Ho, B.R. Lajoie, D.G. Hildebrand, A.C. Simon, M.S. Becker, F.W. Alt, and J. Dekker. 2012. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*. 148:908–921. <https://doi.org/10.1016/j.cell.2012.02.002>
- Zhou, E., B. Zhang, K. Zhu, E. Schaafsma, R.D. Kumar, and C. Cheng. 2020. A TMPRSS2-ERG gene signature predicts prognosis of patients with prostate adenocarcinoma. *Clin. Transl. Med.* 10:e216. <https://doi.org/10.1002/ctm.2.216>

## Supplemental material

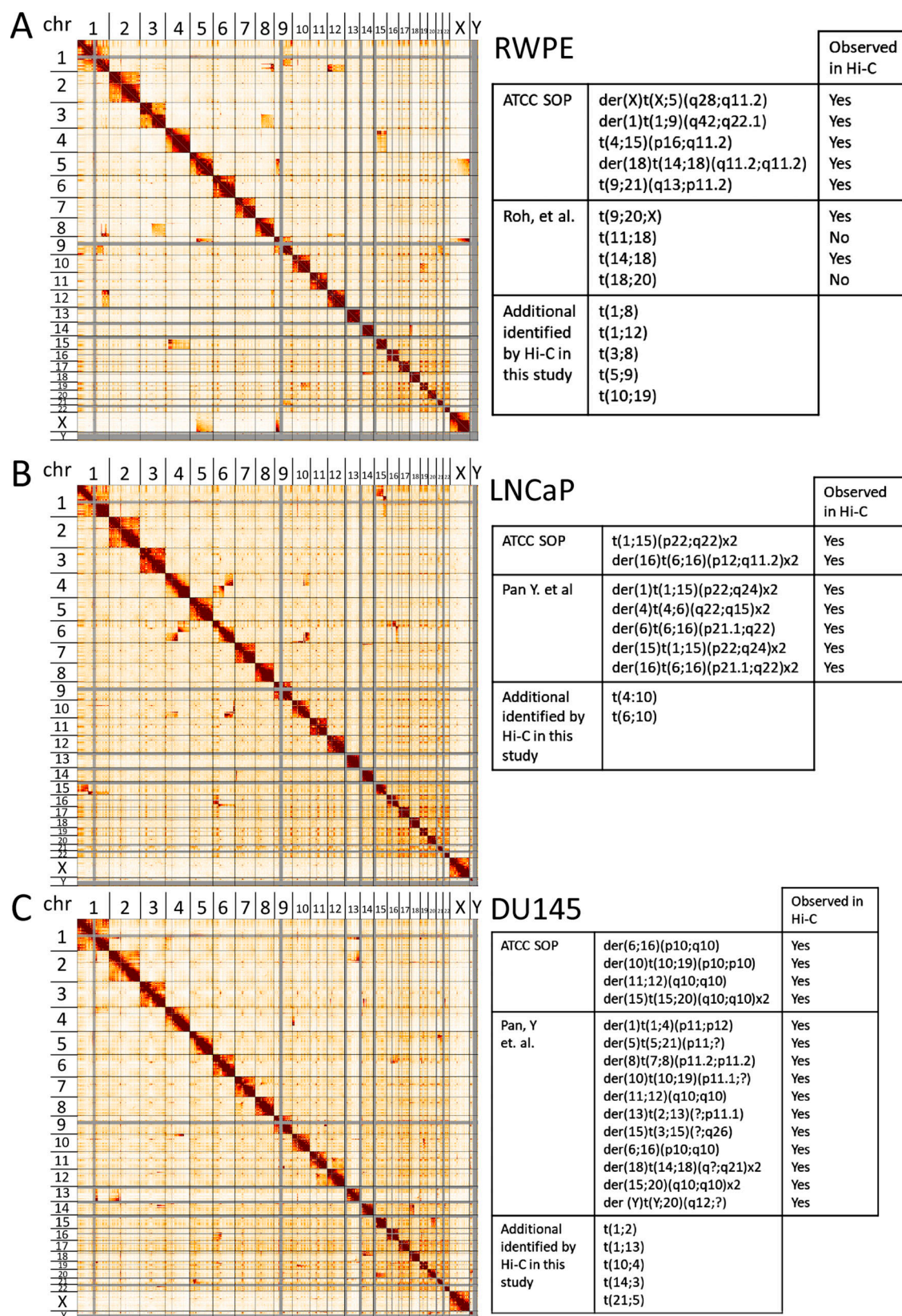


Figure S1. **Detection of known and novel translocations in RWPE, LNCaP, and DU145 cells using 2.5 Mb-resolution Hi-C data.** 2.5-Mb Hi-C heatmaps for RWPE (A), LNCaP (B), and DU145 (C). Translocations between chromosomes (chr) appear as high interaction frequency areas away from the diagonal. Translocations previously described by SKY analysis are listed on the right of each heatmap, along with those events identified in this study; our study has an 83% concordance with previously reported SKY data. The inherent higher resolution of the data obtained by Hi-C makes it possible to identify smaller translocation events.

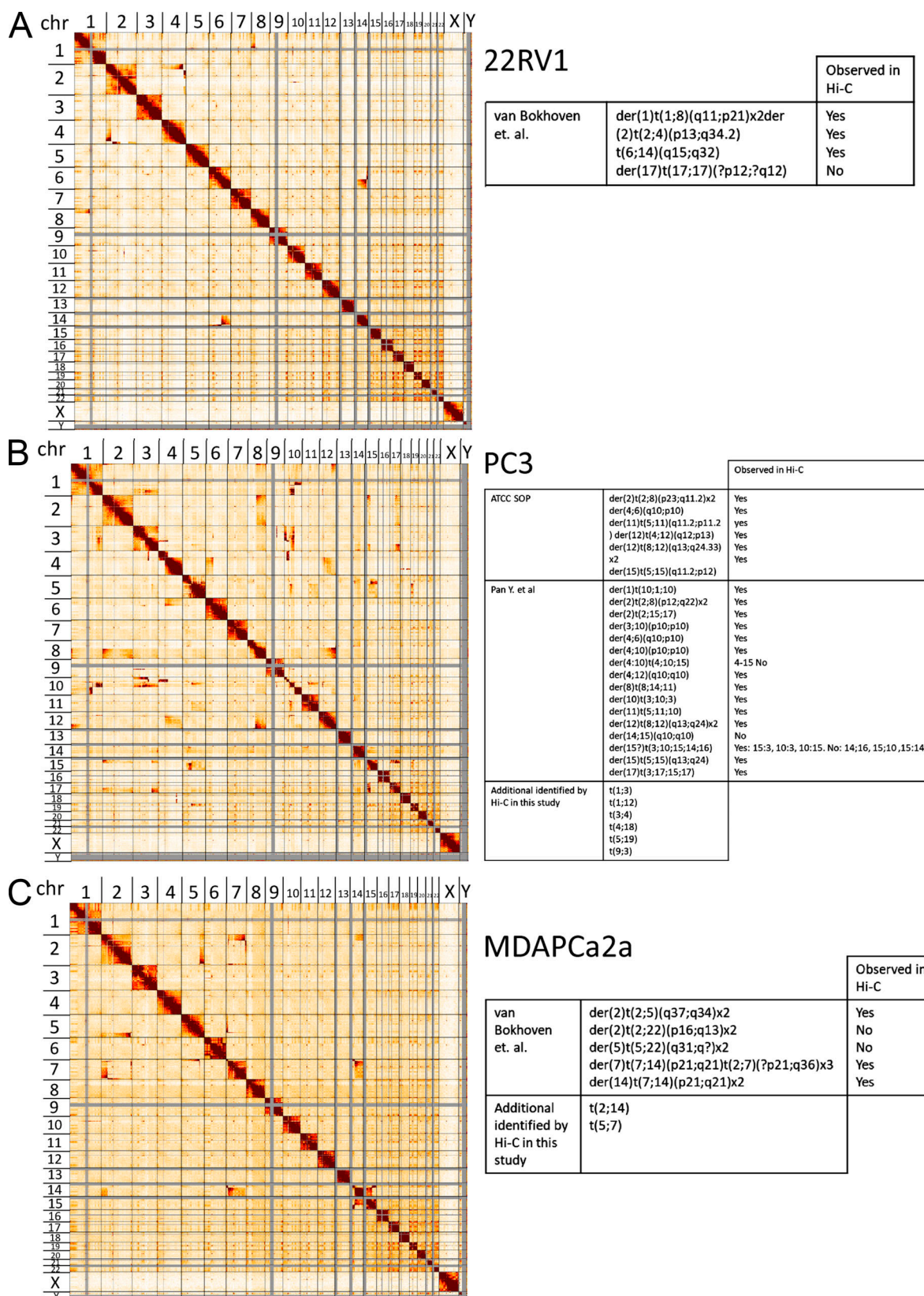
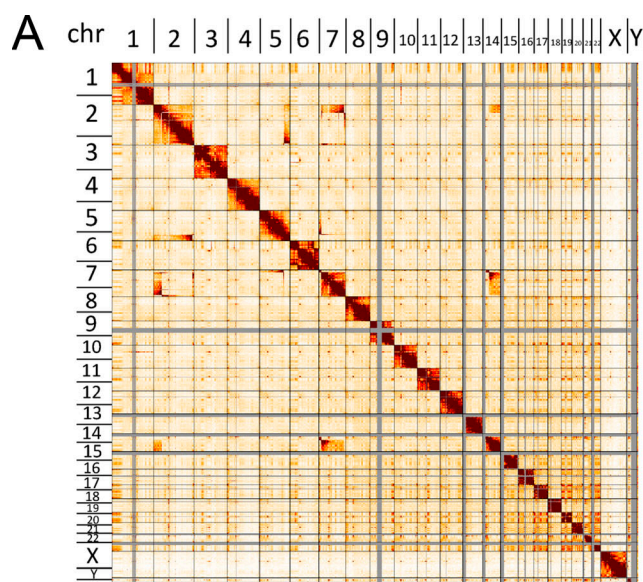
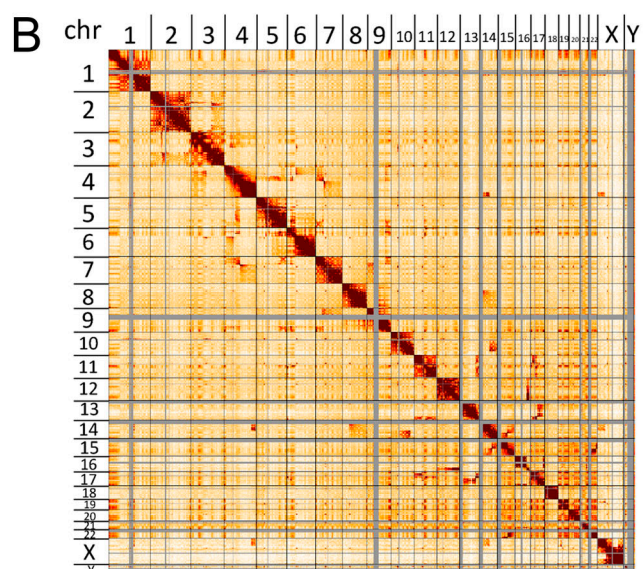


Figure S2. **Detection of known and novel translocations in 22RV1, PC3, and MDAPCa2a cells using 2.5-Mb resolution Hi-C data.** 2.5-Mb Hi-C heatmaps for 22RV1 (A), PC3 (B), and MDAPCa2a (C). Translocations between chromosomes appear as high interaction frequency areas away from the diagonal. Translocations previously described by SKY analysis are listed on the right of each heatmap, along with those events identified in this study; our study has an 83% concordance with previously reported SKY data. The inherent higher resolution of the data obtained by Hi-C makes it possible to identify smaller translocation events.



## MDAPCa2b

|   |   | Observed in Hi-C |
|---|---|------------------|
| van Bokhoven et. al.                        | der(2)t(2;5)(q37;q34)                         | Yes              |
|   | der(2)t(2;22)(p16;q13)                        | No               |
|   | der(5)t(5;22)(q31;q?)                         | No               |
|   | der(7)t(7;14)(p21;q21)t(2;7)(?p21;q36)t(14;7) | Yes              |
|   | inv(6)(p12q24)                                | Yes              |
| Additional identified by Hi-C in this study | t(2;14)                                       |                  |
|   | t(5;7)  |                  |



## VCaP

|   |   | Observed in Hi-C |
|---|---|------------------|
| van Bokhoven et. al.                        | der(2)t(2;3)(p23;q13.3)                       | Yes              |
|   | der(3)t(3;4)(q11.1;q12)                       | Yes              |
|   | der(4)t(4;17)(?;q11)                          | No               |
|   | der(5)t(4;5)(?;q35)                           | Yes              |
|   | der(6)t(4;6)(p15;p22)                         | Yes              |
|   | der(7)t(4;7)(q21;p11)                         | Yes              |
|   | der(7)t(7;9)(p21;p13)                         | Yes              |
|   | der(8)t(4;8)(p12;p11)                         | No               |
|   | der(8)t(8;9)(q24.3;q13)                       | Yes              |
|   | der(8;14)(p10;q10)                            | Yes              |
|   | der(10)t(10;14)(q23;q22)                      | Yes              |
|   | der(11)t(9;11)(?;p11)                         | No               |
|   | ins(12;16)(q24;?)x2                           | Yes              |
|   | der(13)t(13;17)(q32;q21)                      | Yes              |
|   | t(14;15)(q24;q22)                             | Yes              |
|   | der(14)t(1;14)(?;q22)t(X;1)(?;?)              | Yes              |
|   | der(16)t(6;16)(?;q22)x2                       | No               |
|   | der(17)t(11;17)(p12;p11.2)                    | Yes              |
|   | t(13;17)(?;q11.2)                             | Yes              |
|   | t(11;13)(?;?)t(11;13)(p11.2;?)                | Yes              |
|   | der(21)t(21;21)(p11;q11)                      | Yes              |
|   | der(22)t(15;22)(q11;q13)x2                    | Yes              |
|   | der(X)t(X;9)(p11.4;?)                         | No               |
|   | der(4)t(4;7)(q21;?)t(7;9)(?;q11)t(5;9)(?;q21) | Yes, Yes, No     |
| Additional identified by Hi-C in this study | t(4;9)  |                  |
|   | t(4;14)                                       |                  |
|   | t(4;X)  |                  |
|   | t(5;6)  |                  |
|   | t(4;7)  |                  |
|   | t(17;16)                                      |                  |
|   | t(17;9)                                       |                  |
|   | t(6;9)  |                  |

Figure S3. **Detection of known and novel translocations in MDAPCa2b and VCaP cells using 2.5-Mb resolution Hi-C data.** 2.5-Mb Hi-C heatmaps for MDAPCa2b (A) and VCaP (B). Translocations between chromosomes appear as high interaction frequency areas away from the diagonal. Translocations previously described by SKY analysis are listed on the right of each heatmap, along with those events identified in this study; our study has an 83% concordance with previously reported SKY data. The inherent higher resolution of the data obtained by Hi-C makes it possible to identify smaller translocation events.

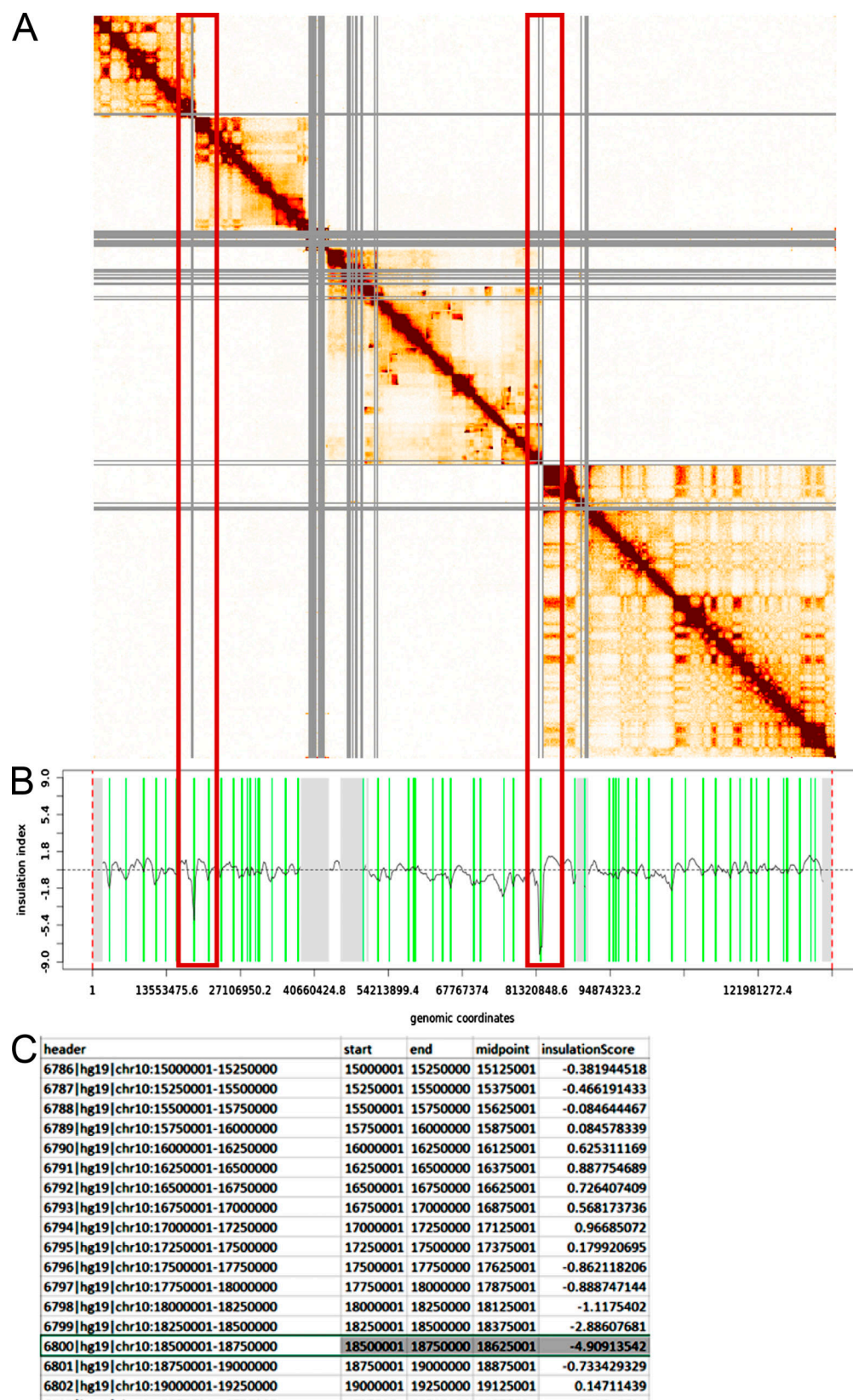


Figure S4. **Mathematical reconstruction of highly segmented genomic regions for compartment analysis.** (A) Hi-C Heatmap of a highly segmented chromosome (chr10, PC3). This type of fragmentation precludes compartment analysis. (B) By performing the insulation calculation on the 250 kb-binned data, using an insulation square size of 1.5 Mb, it is possible to identify the largest dips in insulation plot, which denote the breaks in the chromosome. (C) Compartment analysis is performed on the fragments defined by the bin location. For that particular chromosome, the compartment analysis for RWPE is done for the fragments to provide a fair comparison for subtraction.

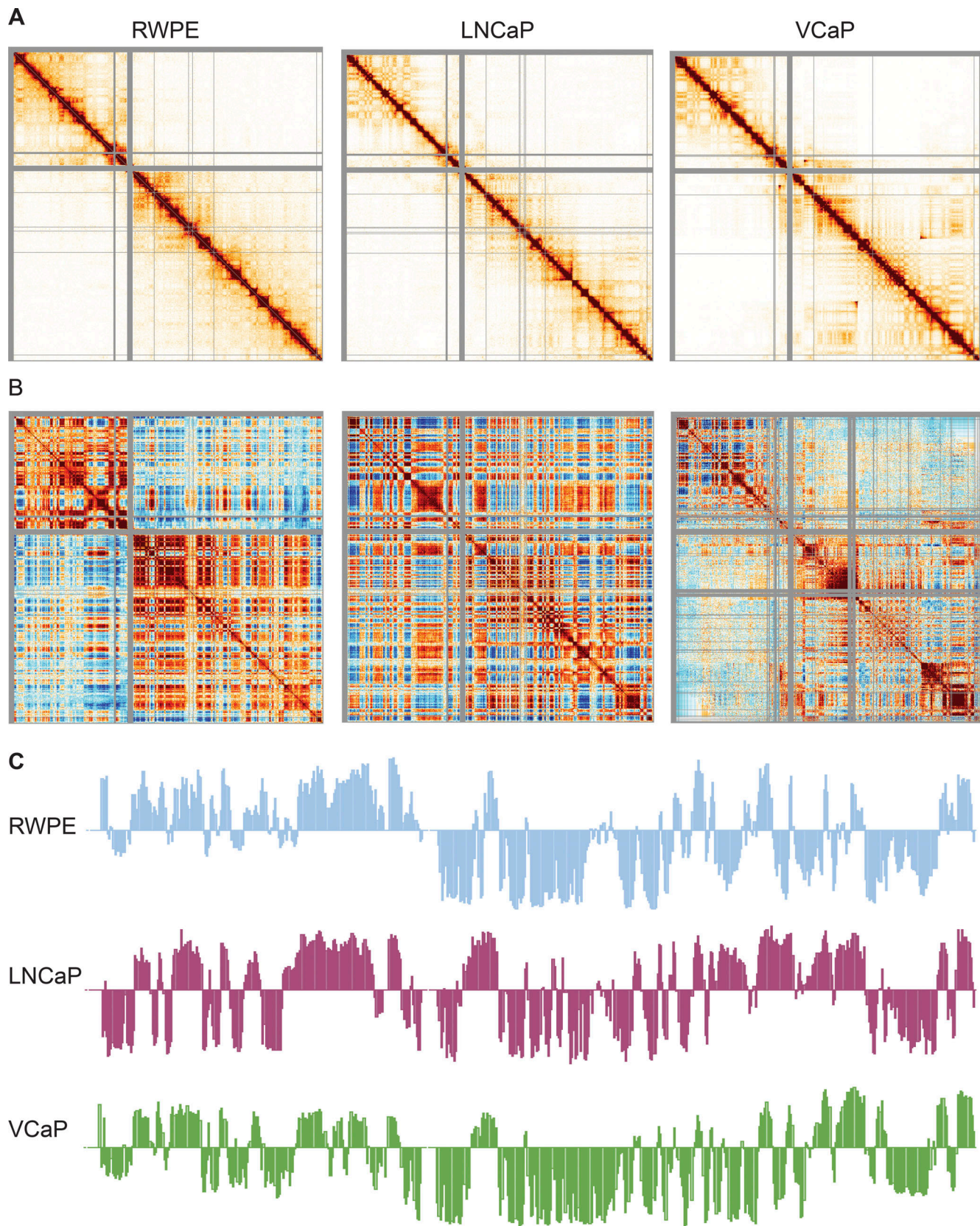
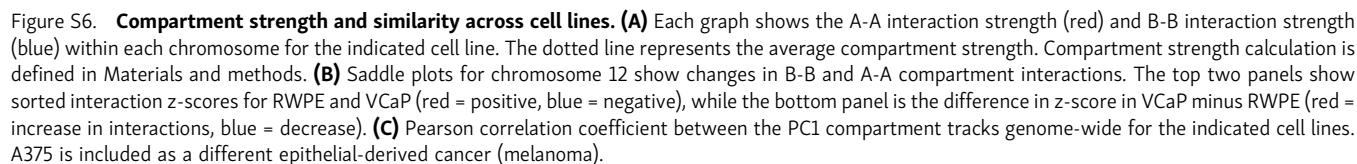


Figure S5. **Changes in chromosome compartmentalization are widespread in prostate cancer progression.** (A) Hi-C Heatmaps for chromosome X, at a 250-kb bin resolution for RWPE, LNCaP, and VCaP. (B) Cis Interaction heatmaps of 250-kb binned data for chromosome X for RWPE, LNCaP and VCaP. (C) Plots of the first eigenvector for chromosome X, obtained from PCA (PC1) of 250-kb binned data, for RWPE, LNCaP, and VCaP, revealing the compartment identity.



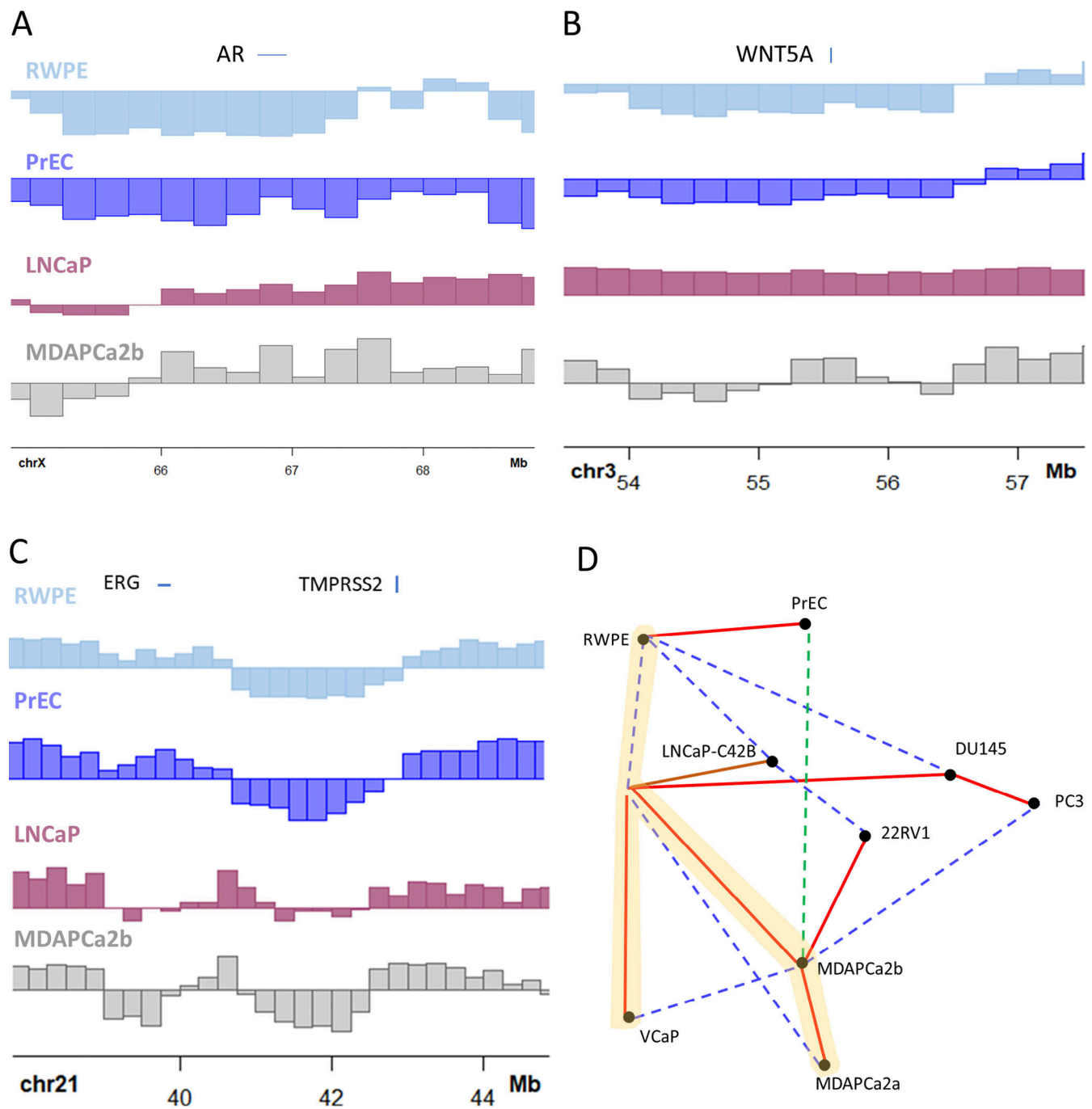
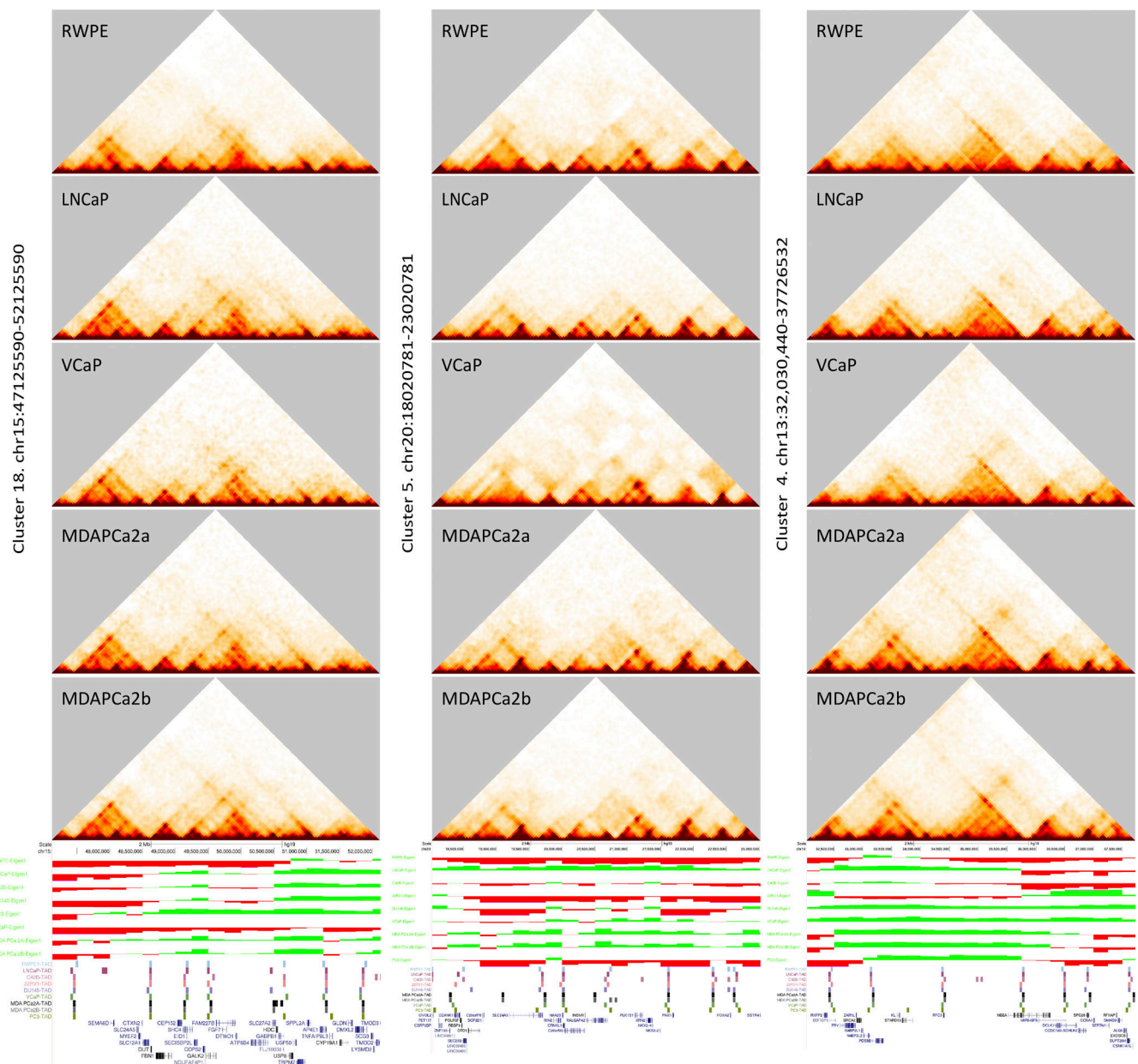
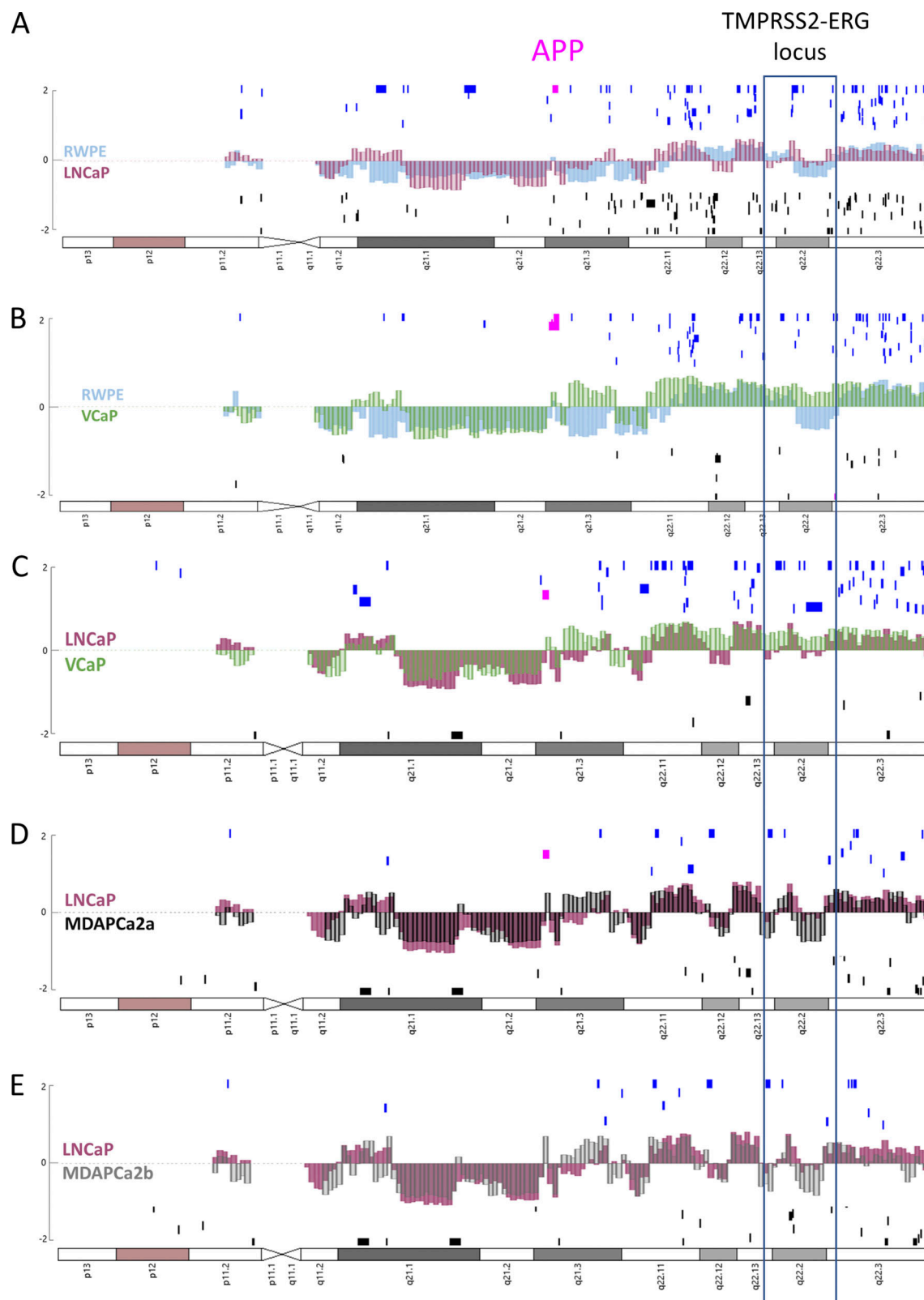


Figure S7. **The non-transformed epithelial lines RWPE and PrEC have similar compartment identities across the genome.** (A–C) PC1 compartment tracks for the AR, WNT5a, and TMPRSS2-ERG loci, showing a similar compartment identity between our chosen nontumorigenic control cell line (RWPE) and the primary prostate epithelium PrEC cells. PC1 compartment tracks for an adenocarcinoma model (LNCaP), and a bone metastatic cell line (MDAPCa2b), show compartment switching relative to both control cell lines around the regions of interest. The rightmost region encompasses the TMPRSS2-ERG locus. (D) A comparison of genome-wide compartment tracks for all cell lines by principal components nearest neighbor analysis reveals that the two nontransformed cell lines (RWPE and PrEC) show a high level of concordance in compartment identity. The cell line-based model axis to metastasis described in Fig. 3 is highlighted in yellow.



**Figure S8. Examples of gene clusters.** Examples of genes that switch compartments. Left: Cluster 18 includes the following genes: *FBN1*, *CEP152*, *SHC4*, *EID1*, *SECISBP2L*, *COPS2*, *NDUFA4P1*, *GALK2*, *FAM227B*, *FGF7*, *DTWD1*, *ATP8B4*, *SLC27A2*, and *HDC*. While in some cell lines the whole area switches compartments, in others the B compartment persists. Center: Cluster 5 includes the following genes: cluster with *SLC24A3*, *RIN2*, *NAA20*, *CRNKL1*, *CFAP61*, *INSM1*, *RALGAPA2*, *PLK1S1*, *XRN2*, *NKX2-4*, *NKX2-2*, *PAX1*, *FOXA2*, and *TTC6*. Changes to the A compartment in all cell lines in the main progression axis. Right: Cluster 4 changes compartments as a cluster (in DU145, VCaP and MDA lines) or in fragments. It is composed of the following genes: *FRY* (switches B to A), *BRCA2* (switches to the B compartment in 22RV1), and *N4BP2L1*, *RFC3*, *DCLK1*, *CCNA*, *SMAD9*, *ALG5*, *ZARIL*, *RFXP2*, and *EEF1DP3*, all of which switch from the B to the A compartment.



**Figure S9. The APP-BACE2 axis, located in chromosome 21, is a potential connection between transcriptional activity and compartment switching.** Compiled data aligning microarray expression (Log2) for genes along chromosome 21. Genes whose expression increases twofold or more or decreases twofold or less in the comparison, with a P value < 0.05 represented as vertical tick marks along the x axis (chromosome location). Chromosome compartment tracks overlay, superimposed along the microarray track. APP-specific probes highlighted in pink. An increased APP expression is observed in all comparisons, except for MDAPCa2b/LNCaP. **(A–C)** Data from HG-U133\_Plus\_2 microarray comparing LNCaP/RWPE, VCaP/RWPE, and VCaP/LNCaP, respectively. **(D and E)** Data from Clariom S microarray, comparing MDAPCa2a/LNCaP and MDAPCa2b/LNCaP, respectively.

Table S1, Table S2, and Table S3 are provided online as separate Excel files. Table S1 shows database references and valid pair counts for all datasets used in this study. Table S2 shows genes identified through  $\Delta$ Eigen 1 analysis (in at least six cancer cell lines or in all four cancer cell lines represented in the axis). Table S3 shows gene clusters that switch compartment identity with progression. Gene clusters that switch from the B to the A compartment are highlighted in green, those that switch from the A to the B compartment in red. 85% of all genes identified as switching from the B to the A compartment are included in 48 proximal clusters. Of the genes that switch from the A to the B compartment, 74% are included in 16 clusters.