

Rearrangement of 3D genome organization in breast cancer epithelial - mesenchymal transition and metastasis organotropism

Priyjit Das, Rebeca San Martin, Tian Hong, Rachel Patton McCord 

UT-ORNL Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, USA •
Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, USA

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

 Copyright information

eLife Assessment

This **valuable** study explores the role of spatial genome organization in oncogenic transformation, addressing an ambitious and significant topic. The authors have assembled comprehensive datasets from various subtypes of localized and lung-metastatic breast cancer cells, as well as from healthy and cancerous lung cells. They identified switching patterns in the 3D genome organization of lung-metastatic breast cancer cells, revealing a reconfiguration of genome architecture that resembles that of lung cells. If validated, this study could be critical for the field; however, at this stage, it is **incomplete**, as the main claims are only partially substantiated.

<https://doi.org/10.7554/eLife.103697.1.sa3>

Abstract

Breast cancer cells exhibit organotropism during metastasis, showing preferential homing to certain organs such as bone, lung, liver, and brain. One potential explanation for this organotropic behavior is that cancer cells gain properties that enable thriving in certain microenvironments. Such specific metastatic traits may arise from gene regulation at the primary tumor site. Spatial genome organization plays a crucial role in oncogenic transformation and progression, but the extent to which chromosome architecture contributes to organ-specific metastatic traits is unclear. This work characterizes chromosome architecture changes associated with organotropic metastatic traits. By comparing a collection of genomic data from different subtypes of localized and lung metastatic breast cancer cells with both normal and cancerous lung cells, we find important trends of genomic reorganization. The most striking differences in 3D genome compartments segregate cell types according to their epithelial vs. mesenchymal status. This EMT compartment signature occurs at genomic regions distinct from transcription-defined EMT signatures, suggesting a separate layer of regulation. Specifically querying organotropism, we find 3D genome changes consistent with adaptations needed to survive in a new

microenvironment, with lung metastatic breast cells exhibiting compartment switch signatures that shift the genome architecture to a lung cell-like conformation and brain metastatic prostate cancer cells showing compartment shifts toward a brain-like state. TCGA patient data reveals gene expression changes concordant with these organ-permissive compartment changes. These results suggest that genome architecture provides an additional level of cell fate specification informing organotropism and enabling survival at the metastatic site.

Significance

Computational analysis of a cohort of cancer cell lines reveals 3D genome spatial compartment changes are associated with transitions in cancer cell state that favor metastasis (EMT) and enable survival in a new organ context.

Introduction

Breast cancer is the leading malignant disease in women worldwide. It is highly heterogeneous; based on molecular characteristics it can be divided into four broad subgroups: luminal A (estrogen receptor ER+ progesterone receptor PR+/-human epidermal growth factor receptor 2 HER2-), luminal B (ER+ PR+/-HER2+), HER2-enriched (ER-PR-HER2+), and triple-negative/basal (ER-PR-HER2-) (1 [↗](#)). About 20 to 30% of early-stage breast cancer patients have a chance to ultimately develop distant metastasis, which accounts for 90% of all breast cancer-related deaths (2 [↗](#)). Overall, metastasis implies the dissemination of cancer cells from the primary tumor site into the circulation system and later extravasation to a new site which leads to the formation of a secondary tumor. However, the organs affected by breast cancer metastasis vary remarkably, depending on the disease subtype, molecular features, primary tumor location, and secondary site microenvironment. This phenomenon is known as ‘metastasis organotropism’ (3 [↗](#)). For example, breast cancer cells can metastasize to bone, lung, liver, and brain. However, bone metastasis accounts for a major percentage of metastatic cases compared to other organs (4 [↗](#)). Moreover, luminal A and basal subtypes of breast cancer show a higher incidence of bone and lung metastasis, respectively, whereas luminal B preferentially metastasizes to both organs (5 [↗](#)). Organotropic behavior is congruent with the ‘seed and soil’ hypothesis, according to which a cancer cell (seed) having specific intrinsic abilities will only be able to survive in a host local microenvironment (soil) that is favorable to it (6 [↗](#)).

Secondary organ colonization by circulating tumor cells is a complex process, ruled by dynamic and evolving interactions between cancer cells and the host microenvironment. In most situations, the microenvironment is unfavorable, and the disseminated cells fail to survive in that ‘soil’. Despite the low survival rate of cancer cells during the metastatic colonization process, cells with the intrinsic capabilities to interact with the host microenvironment can survive and proliferate in the secondary site (7 [↗](#)). Alternatively, cells might arrive at the secondary site and remain dormant under the influence of the local microenvironment (8 [↗](#)-11 [↗](#)) slowly acquiring characteristics that would later allow their proliferation. Importantly, the microenvironment surrounding primary breast tumors seems to condition the cells to specific metastatic profiles. For example, if the primary tumor stroma is enriched in transforming growth factor beta (TGFB), lung metastasis prevails through the induction of angiopoietin-like 4 (12 [↗](#)). In contrast, a tumor stroma enriched with mesenchymal cytokines like C-X-C motif chemokine 12 (CXCL12) and insulin-like growth factor 1 (IGF1) predisposes primary cancer cell population to home into the bone marrow (13 [↗](#)). How these different reactive microenvironments condition the origin of specific metastatic traits still poses an outstanding question.

The current view in the cancer biology field is that transcriptional gene signatures associated with organ-specific metastatic traits are regulated in part by cues from the reactive microenvironment that accompanies primary tumorigenesis (12–15). However, whether regulating those genes involves changes in 3D genome organization is still an open question. Recent advancements in imaging and molecular biology techniques have shown that the genome organizes in a non-random multi-layered fashion inside the nucleus, and each layer of organization contributes both independently and synergistically to gene expression regulation (16). At a large scale, the reorganization of genome compartments, in which euchromatic and heterochromatic genomic regions are spatially segregated, is also found to be associated with cancer progression (17–19). It has been shown that as breast cancer becomes endocrine-resistant, the change in compartment organization is accompanied by gene expression changes and differential estrogen receptor binding (17). Other studies have also found that an unintended consequence of various drug treatments and therapies is the rearrangement of chromatin architecture at different levels (17,20,21). Though these types of studies are essential for elucidating the association between the reorganization of the genome in three dimensions and oncogenic transformation, they do not consider the genome organization changes experienced by the metastatic cell that occur as a result of the interaction with the microenvironment of the secondary site. Genome spatial compartmentalization is strongly associated with cell type definition: genomic regions segregated into the A (euchromatic) or B (heterochromatic) compartments differ among cell types and tissue origin (22). These compartmentalization profiles help encode stable cell fate gene expression (23–25). Therefore, we hypothesize that similar to gene expression, metastatic cancer cells experience 3D genome structure rearrangements that mimic the compartment signatures of their new host microenvironment. To test this hypothesis, here, we focused on the breast cancer lung metastasis system. We compared genome organization data from different types of breast epithelial cell lines (normal, localized cancer, and metastatic) with normal and cancerous lung cells. Our analysis revealed that metastatic breast cancer cells have a higher proportion of lung-specific spatial compartment changes than localized breast cancer cells. We find similar brain microenvironment-specific compartment changes in prostate cancer cells that have the potential to metastasize to brain. These distal-site concordant compartment shifts often correspond to changes in gene expression in both cell lines and TCGA patient samples. However, we find that 3D genome changes do not always correlate with gene expression changes. Instead, particularly across the epithelial mesenchymal transition (EMT) spectrum, we find a set of genes that change expression without needing spatial reorganization and then a distinct set of epithelial and mesenchymal phenotype genes that change their spatial compartment without altered gene expression. We hypothesize that spatial genome organization facilitates immediate changes in gene expression in certain loci but can also poise certain regions for future activation in a way that favors EMT and adaptation to metastatic organ microenvironments.

Results

Modeling the breast cancer lung metastasis system

To investigate whether breast cancer spatial genome organization showed rearrangements specific to a distant organ metastatic site microenvironment, we needed to assemble genome spatial compartment data from different types of normal and cancerous cells for both breast and a secondary organ. Since most of breast cancer subtypes show a preference for lung metastasis, we narrowed our field to this metastatic site (26). Breast and lung normal and cancerous cell types for which publicly available genome organization (Hi-C) are available were selected for this work (Supplementary Table 1). A total of 15 cell types including primary cells and cancer cell lines from both organs were obtained. A classification of the cells used in this study based on their malignancy status and other molecular signatures is shown in **Fig. 1a**. Briefly, for the breast cancer system we considered both HMEC and MCF10A as normal (non-tumorigenic) breast epithelial cells and placed them at the starting node of the model. Cancer cell lines were then

divided into three subcategories based on breast cancer molecular signatures: luminal (ZR751, BT474, ZR7530, MCF7 and T47D), HER2-enriched (HCC1954 and SKBR3) and triple-negative (HCC70, BT549 and MDA-MB-231). Further, within the luminal subtype, ZR751 and BT474 were designated as localized cancers and the remaining three cell lines, ZR7530, MCF7, and T47D were used as models for ascites, lung, and lung metastatic breast cancers respectively. Similarly, for the HER2-enriched subcategory, HCC1954 and SKBR3 were considered localized and lung metastatic cancers, respectively. Finally, in the case of the triple-negative subtype, HCC70 and BT549 were used to model localized cancer whereas MDA-MB-231 represents lung metastatic cancer. For the lung portion of the system, HTBE primary cells represent normal lung epithelium and both the A549 and H460 cell lines represent localized lung cancer.

Subtype and epithelial-mesenchymal state-specific signatures are present in breast cancer compartmental organization

We compared the supervised ordering of the cell lines described above with the ordering derived from unsupervised analysis of genome-wide A/B compartment organization data. Compartment identity strengths were calculated using eigenvector decomposition on the 250 kb binned Hi-C data for each cell line (Methods). In this analysis a positive value signifies a euchromatic A compartment region, while a negative value represents a heterochromatic B compartment association.

When we applied hierarchical clustering on the genome-wide compartment data, we observed a grouping of cells that was congruent with the cell ordering constructed based on molecular and clinical features (**Fig. 1b**). For example, all the normal cell types, including both the breast and lung systems, were clustered together. Similarly, cells from the luminal subtype and the lung carcinoma cells were grouped in their respective clusters. However, we also observed a few interesting exceptions from the clustering result. MDA-MB-231 cells were grouped with the lung carcinoma cells. This implies that the genome organization of this cell line is closer to that of lung cells than to other breast epithelial cell lines. We also found that HCC70 and HCC1954, which are both cells from localized cancer, clustered relatively close to the normal epithelial cells. Collectively, this unsupervised analysis shows that the different subtypes of cancer exhibit distinctive 3D genome organization features at the compartment level (**Fig. 1b**), as previous studies suggested (18).

Performing principal component analysis (PCA) on the genome-wide compartment data revealed similar trends (**Fig. 1c**). Interestingly, the ordering of cell types along the first principal component (PC1) captured a transition between epithelial and mesenchymal-like cells. Cells characterized as highly epithelial were found on the left extreme of the PC1 axis, and as we move towards the right along PC1, the cells acquire a mesenchymal phenotype. For example, MDA-MB-231 cells are derived from healthy cells of epithelial origin but have undergone an epithelial-mesenchymal transition and show a mesenchymal-like phenotype (27,28). In our Hi-C principal components analysis, this cell line has a highly positive PC1 value. In contrast, MCF7 cells maintain a more epithelial morphology and are found at the opposite end of the PC1 spectrum from MDA-MB-231. The organization of cell lines along PC1 from spatial compartmentalization matches quite well with the previously characterized epithelial-mesenchymal classification of these cell lines based on transcriptomic data (Fig. S1) (29,30). The PC2 axis segregates the cell types depending on the presence or absence of the HER2, suggesting that in breast cancer cell lines, the presence or absence of this receptor and its downstream signaling network has important effects on the genome organization at the compartment scale. To further characterize the genomic regions that showed distinctive compartment signatures across cell lines, we extracted lists of all genes that fell into genomic bins with the 100 most positive and negative eigenvalues for PC1. Gene ontology enrichment results for these gene lists coincide with the idea that PC1 represents the EMT axis: PC1 positive genes have functions related to cell adhesion

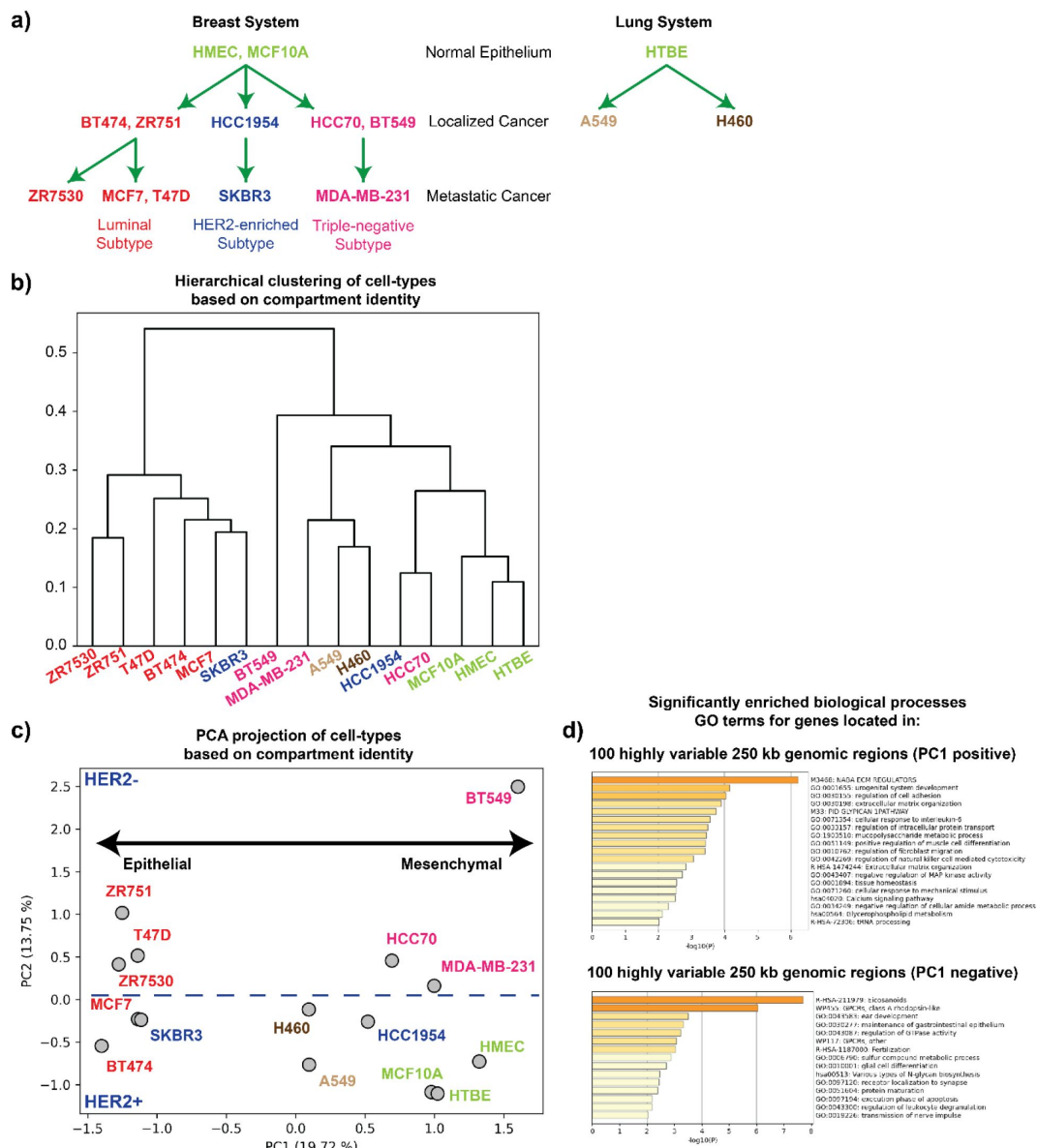


Fig 1.

Spatial compartment identity segregates cells by subtype and epithelial-mesenchymal status.

a) Primary cells and cell lines used in this study are arranged depending on their molecular subtypes and malignancy status based on prior available characteristics. Green: normal breast and lung epithelium, red: luminal subtype, blue: HER2-enriched subtype, pink: triple-negative subtype, brown: localized lung cancer. **b)** Hierarchical clustering of genome-wide compartment identity data for all breast and lung cell types at 250 kb resolution; colors as in **1a**. **c)** Principal component analysis of genome-wide compartment identity data for breast and lung cell types at 250 kb resolution; colors as in **1a**. The PC2 axis is divided into two subspaces based on HER2 status (blue thick dotted line). **d)** GO term enrichment (biological processes) of the genes in genomic regions corresponding to the top 100 positive and 100 negative elements of the first eigenvector of genome-wide compartment identity PCA.

(tenascin C, JAG1, ACTA2), extracellular matrix organization (APP, LAMA2, FGF2, BMP4) and positive regulation of cell migration, consistent with a mesenchymal phenotype, while PC1 negative genes show GO terms related to epithelial maturation ([Fig. 1d](#)). Overall, our analyses show that the genome organization is not only able to distinguish different subtypes of cancers but also exhibits changes along the epithelial-mesenchymal transition.

Breast cancer transcriptome profiles also exhibit epithelial-mesenchymal state-based segregation

We next analyzed the gene expression signatures that emerged across different subtypes of breast cancer (data sources described in Supplementary Table 2) and compared those with the patterns obtained from the compartmental analysis. We first performed PCA on the transcriptome profile of all the breast and lung system cells considered in this study ([Fig. 2a](#)). From the low dimensional PC projection result, we observed that similar breast cancer subtype cells were arranged closely in the PC space based on their transcriptome profile. As in the compartment signature, MDA-MB-231 cells localize close to lung cancer cells A549 and H460, suggesting that this lung metastatic cell line has transcriptional similarities to lung cancer. We also noticed that along the largest PC axis (PC1), cells are again ordered based on their epithelial-mesenchymal transition state. Performing GO term enrichment analysis using the genes with the top 100 most positive and 100 most negative eigenvalues of the first eigenvector, we find that genes associated with PC1 positive values relate to epithelial organization, differentiation and negative regulation of cell migration, denoting a more epithelial phenotype ([Fig. 2b](#)). In contrast, PC1 negative genes relate to the mesenchymal functions with enriched terms for wound healing, hemostasis, cell movement, and ECM remodeling ([Fig. 2b](#)). Along with the unsupervised PCA, we also performed gene expression analysis in a supervised fashion using two different methods – gene set variation analysis (GSVA) and non-negative principal component analysis (nnPCA) ([Fig. 2c](#) and [2d](#), see Methods).

Briefly, for each cell type, an E (epithelial) score and an M (mesenchymal) score were calculated based on the expression levels of a curated set of 232 epithelial and 193 mesenchymal genes ([30](#), [31](#)). A higher E score represents more epithelial characteristics whereas a higher M score corresponds to a mesenchymal phenotype. The GSVA and nnPCA transcriptome analyses recapitulate a similar ordering of cell lines according to epithelial-mesenchymal characteristics.

Distinct EMT gene sets are affected by spatial compartment changes vs. transcription changes, revealing multiple layers of EMT signature

Numerous studies have already demonstrated that cancer cells segregate along an EMT axis by transcriptomics ([29](#)–[34](#)). Given that spatial compartmentalization is often correlated with gene expression ([22](#)), a simple possible explanation for the EMT segregation of cells by compartment signatures could be that genes which change expression also change compartment. An alternative hypothesis would be that transcription and 3D genome organization change represent two independent aspects of the EMT transition. To evaluate these possibilities, we compared the genes obtained from the compartmental PCA analysis, the transcriptome PCA analysis, and the curated epithelial-mesenchymal gene set ([Fig. 3a](#)). We found that the genes in regions that change spatial compartmentalization have very little overlap with the sets of genes that most varied in transcription across these cell lines or were curated as differently expressed EMT genes ([Fig. 3a](#)). To further investigate this result, we examined the transcription status of genes that changed compartment across the EMT spectrum and, conversely, the compartment status of genes that changed transcription ([Fig. 3b](#), [c](#), and [d](#)). As expected, hierarchical clustering of the detected highly varying compartments ([Fig 3b](#), top) and highly varying transcripts ([Fig 3c](#) and [d](#), bottom) resulted in clear segregation of the cell types into similar groupings (i.e., HMEC and MCF10A together, BT474 and SKBR3 together with an opposing pattern).

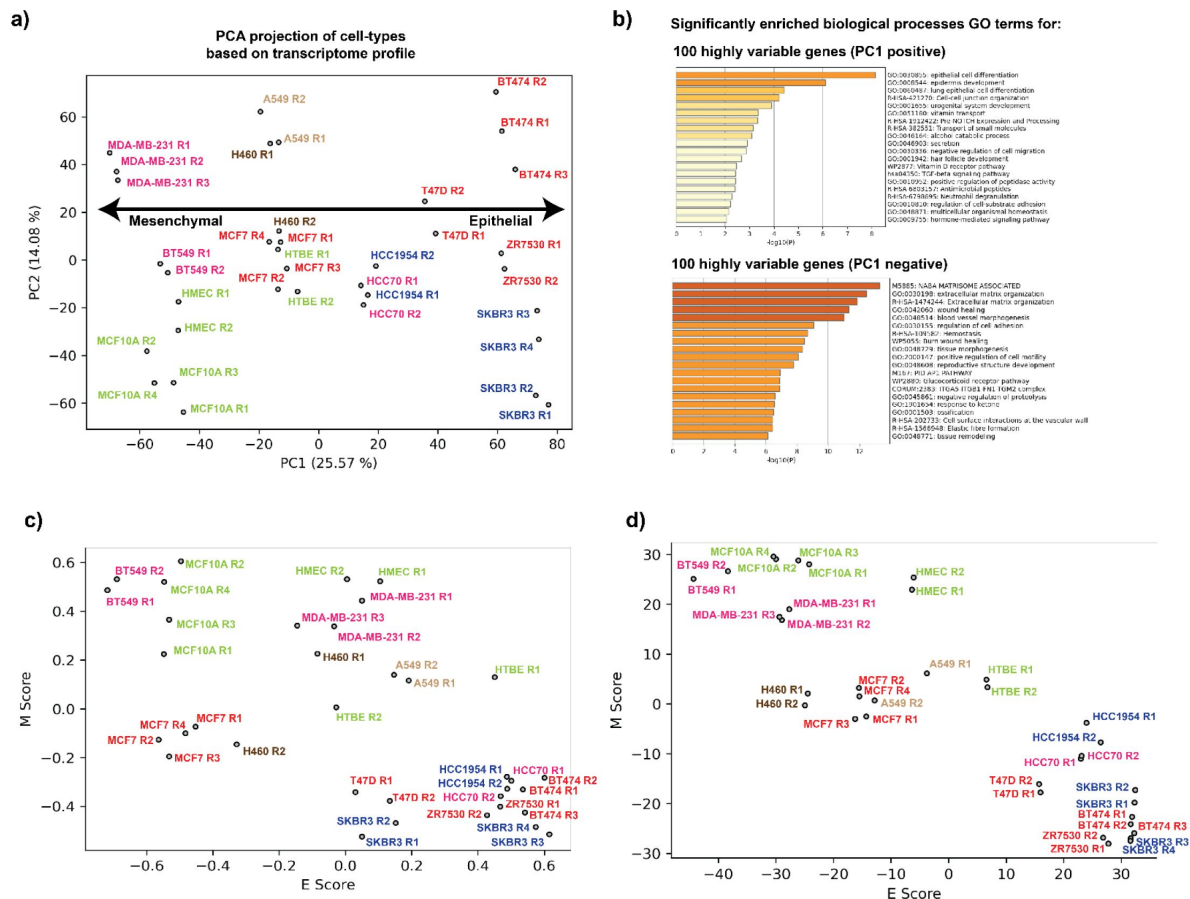


Fig 2.

Gene expression differences also segregate breast cancer cell lines along an EMT axis.

a) Principal component analysis of transcriptome data for breast and lung cell types in this study. Cell types are colored based on their molecular subtypes as represented in [Fig. 1a](#). **b)** GO term enrichment (biological processes) of the genes corresponding to top 100 positive and 100 negative elements of the first eigenvector of the transcriptome profile PCA. **c, d)** Projection of breast and lung cell types on a curated epithelial and mesenchymal axis based on their gene expression using gene set variation analysis (GSVA) (**2c**) and non-negative principal component analysis (nnPCA) (**2d**) methods.

But, the genes which strongly changed compartments between cell lines showed no consistent patterns of correlated expression change (**Fig. 3b**, bottom). Likewise, very few compartment differences existed across cell lines at genes whose transcription segregated cells by EMT status (**Fig. 3c** and **d**, top). Notably, for both the transcriptome PCA analysis and the curated EMT gene set, the majority of the differentially expressed genes are found in the active A compartment. Overall, our results favor the hypothesis that compartmentalization and gene expression signatures are capturing two independent factors in EMT transition. The genes which differ in expression across these cell lines are largely already accessible (A compartment) and therefore their expression is likely not controlled by spatial compartmentalization. Meanwhile, the additional EMT relevant genes that change compartment, but not transcription, may be now spatially poised for future gene regulation given additional environmental cues, as has been shown in other contexts (35–38).

Lung metastatic breast cancer cell lines acquire lung-like genome architecture

To survive a metastatic environment, a cancer cell must express a repertoire of binding factors, extracellular matrix proteins, and ligands that positively interact with the secondary tissue microenvironment and foster colonization and growth. We hypothesized that for that to happen, the genome organization of the metastatic cell should change to resemble that of the cells at the secondary site, as a prelude of transcriptional activation. To evaluate this hypothesis, we compared the 3D genome organization of the metastatic and localized breast cancer cells with lung epithelial vs. lung cancer cells (**Fig. 4a, 4b**). As a baseline, we first compared the breast cancer compartment identity data with that of the normal breast epithelial cells. For each comparison, this resulted in four possible compartment identity combinations: AA (A compartment in both normal breast and breast cancer), AB (A compartment in normal breast and B compartment in breast cancer), BA (B compartment in normal breast and A compartment in breast cancer), and BB (B compartment in both normal breast and breast cancer). Similarly, the compartment identity of lung cancer cell lines was also compared with that of the normal lung cell. We then performed a cross-comparison between the breast and lung cancer comparisons, leading to a total of 16 compartment identity combinations (**Fig. 4b**). Of these, we were particularly interested in two specific combinations: AB_BB (Genomic regions which are in the B compartment in both normal lung and localized lung cancer, and which change from A to B in the transition from normal breast to breast cancer) and BA_AA (Genomic regions which switch from B to A in breast cancer compared to normal breast epithelium, and which are in the A compartment in both normal lung and lung cancer). These combinations were termed “lung permissive genome organization changes”, as the compartment switches in breast cancer make the cells more like the lung compartment pattern. These combinations allow us to distinguish changes that could foster lung-specific-survival changes from cancer-specific traits such as AB_AB and BA_BA (compartments that change in the same direction from normal to cancer cells in both breast and lung). We then evaluated whether lung-permissive 3D genome changes occurred more frequently in lung metastatic breast cancer than in localized breast cancer.

To quantify the frequency of lung permissive changes, we calculated what fraction of the compartment changes in breast cancer vs. normal breast tissue were also counted as lung permissive. (e.g., for AB_BB breast cancer lung permissive changes, we calculate: # of AB_BB bins / # of breast cancer AB bins. Similarly, for BA_AA: # of BA_AA bins / # of breast cancer BA bins). We divided cells according to breast cancer subtype for this comparison. For the subtypes where several cell lines were available, we considered compartment identity combinations that were consistent between members of the group. We found that for luminal and triple-negative subtypes, metastatic breast cancer exhibits a higher level of lung permissive changes than localized cancer (**Fig. 4c**).

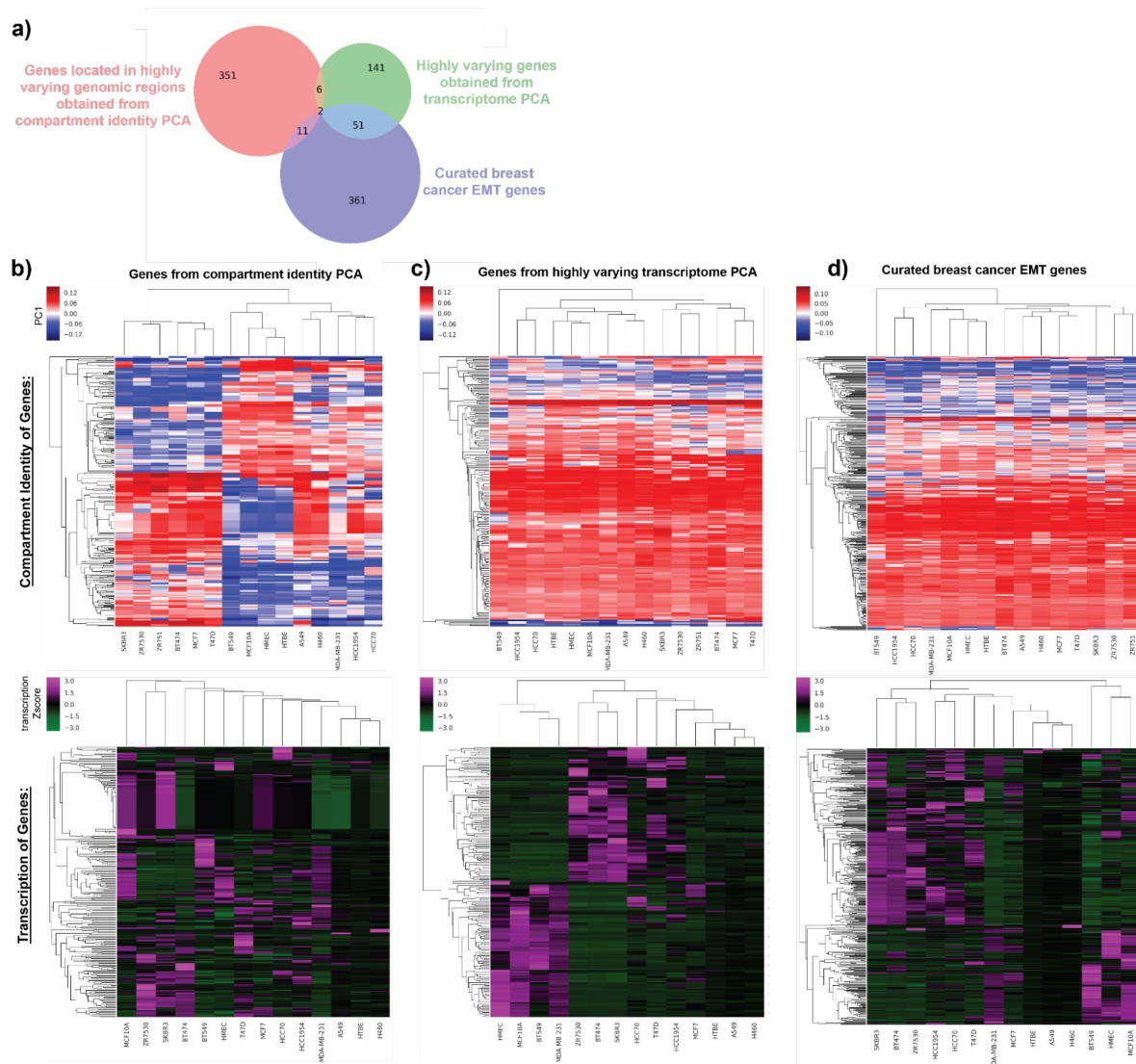


Fig 3.

Transcription and compartment changes capture distinct sets of EMT-related genomic regions.

a) Overlap analysis of the genes obtained from the compartmental analysis PC1 (red), the transcriptome analysis PC1 (green), and the curated breast cancer epithelial-mesenchymal gene set (blue). **b-d)** Hierarchical clustering of compartmental profiles (top row) and gene expression (bottom row) of the genes obtained from either **(b)** the compartmental analysis PC1, **(c)** transcriptome analysis PC1, or **(d)** the curated set of breast cancer epithelial and mesenchymal genes.

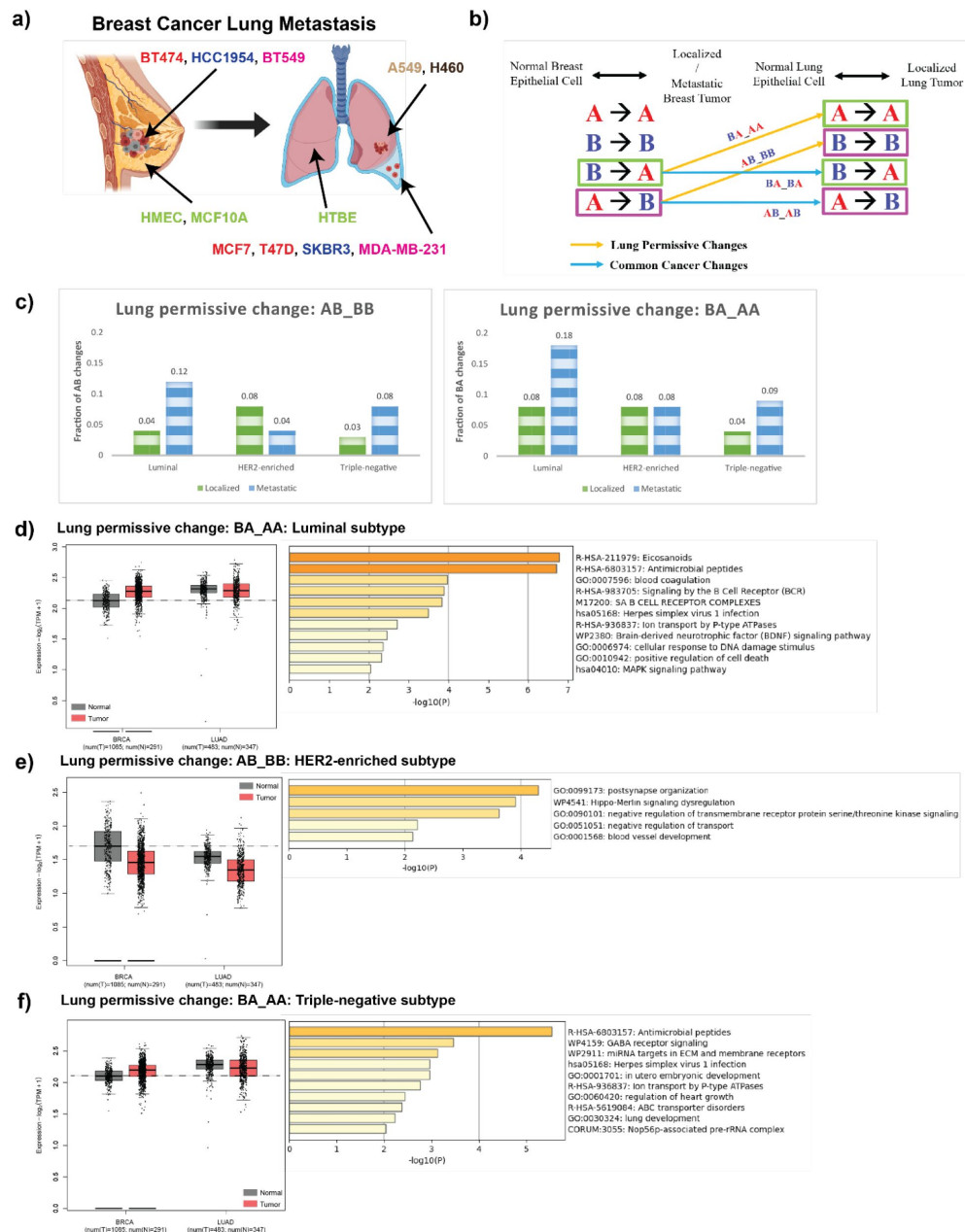


Fig 4.

a) Normal breast epithelium (HMEC and MCF10A), localized breast cancer (BT474, HCC1954, and BT549), and lung metastatic breast cancer (MCF7, T47D, SKBR3, and MDA-MB-231) cells are used to model different stages of breast cancer progression. To model diseased lung, normal lung epithelium (HTBE) and localized lung cancer (A549 and H460) cells are used. **b)** Schematic diagram representing breast cancer lung permissive changes calculation. First, for both breast and lung, cancer cell compartment identity is compared with the normal epithelial cell. This results in four possible compartment identity combinations: AA (A compartment in both normal and cancer), AB (A compartment in normal and B compartment in cancer), etc. Then, a cross-comparison between the breast and lung systems leads to 16 compartment identity combinations. Two specific combinations: AB_BB and BA_AA are defined as lung permissive changes (yellow arrows). **c)** Fraction of lung permissive changes shown by localized and metastatic cancers from different breast cancer subtypes. **d-f)** TCGA patient gene expression (left) and GO term enrichment (biological processes) (right) for genes from regions exhibiting **(d)** BA_AA lung permissive changes in the luminal metastatic breast cancer subtype, **(e)** AB_BB lung permissive changes in case of HER2-enriched metastatic breast cancer subtype, or **(f)** BA_AA lung permissive changes in case of triple-negative metastatic breast cancer subtype. TCGA data from BRCA and LUAD tumor vs. normal sets used and sample size indicated.

Genes affected by lung permissive compartment changes show lung-like expression levels cancer patients

After selecting the genomic regions that exhibit lung permissive changes in metastatic breast cancer cell lines, we extracted the underlying genes and used the GEPIA2 tool (39) to analyze the expression of these gene sets in TCGA breast and lung cancer patient data (Fig. 4d, 4e, and 4f; left panels). We found cases in which the gene expression levels change concordantly with the compartment. The results show that for the regions exhibiting lung permissive changes, the switch in compartmental identity is associated with the gene expression change from normal to tumor. That is, for BA_AA regions, the breast tumor expression of these genes is higher than normal breast tissue expression, coinciding with higher expression in the lung (Fig. 4d and f). Conversely, for AB_BB regions, the breast tumor expression of genes in these regions is lower than for normal tissue, corresponding to lower expression in the lung (Fig. 4e).

Organotropic 3D genome changes match target organ more than an unrelated organ

To verify that “lung permissive changes” truly occurred more than expected in lung metastatic cells, we used a different organ system—brain and brain cancer—as a negative control. Breast cancer can metastasize to brain (40,41), but the specific metastatic cell lines for which we have data are not metastatic to brain. Therefore, we would expect fewer “brain like” compartment changes in metastatic breast cancer that metastasizes to lung. Before comparing percent permissive changes between organs, we developed an adjusted version of the calculation to account for differing baseline probabilities of compartment change at the distant organ site (SFig. 2, Methods). The adjusted calculation still results in a higher level of lung permissive changes in lung metastatic breast cancer cell lines as compared to localized (SFig. 3d). For the brain comparison, we used the NHA cell line (normal human glia) and the SF9427 cell line (glioblastoma) as our reference models (SFig. 3a and 3b). We then calculated the adjusted level of brain permissive changes shown by localized and metastatic breast cancers (SFig. 3c). We found that the metastatic breast cancer cells used in this study showed lung-permissive changes more frequently than brain-permissive compartment changes (SFig. 3c and d). This is congruent with the fact that these metastatic breast cancer cell lines were all isolated from lung pleural effusions. We note that Hi-C data for brain metastatic breast cancer cell lines is not currently available for comparison.

Brain metastatic prostate cancer cell lines exhibit organotropic 3D genome changes toward brain

The analysis above suggests that lung metastatic breast cancer cell lines show more lung permissive than brain permissive genome organization changes. Next, we asked whether a cancer type that metastasizes to the brain exhibits brain-permissive compartment changes. For this purpose, we chose prostate cancer cell lines derived from localized (lymph node) and brain metastatic sites (LNCaP and DU145), and compared their genome architecture to the non-tumorigenic prostate epithelial cell line RWPE (SFig. 4a, b). We found that the brain metastatic cells showed more brain-permissive changes than the one derived from a lymph node (SFig. 4c). Also, particularly in the B to A direction, there were greater brain permissive changes in this prostate cancer system (Fig. 4c) than in the breast cancer system (SFig. 3c). This confirms that a cancer type that metastasizes to the brain matches the brain compartment profile more often than a cancer type that does not metastasize to the brain. We further found that the genes in the switched compartment regions have functions that could relate to leaving the prostate epithelium and adapting to the brain environment (SFig. 4d). For example, the DU145 (brain metastatic) brain-permissive B to A switched genes were enriched for functions in regulation of nervous system development (including semaphorin genes SEMA5A, SEMA3E, SEMA3A) and regulation of

canonical WNT signaling pathway (such as WNT2, FZD4). Conversely, the regions that switch from A to B to match the brain B compartment and are enriched for keratinization and cell adhesion, suggesting a downregulation of prostatic epithelial properties.

Discussion

During metastasis, breast cancer cells spread preferentially to certain organs, a series of events widely known as metastasis organotropism. This phenomenon, also shown by other types of cancer, may be explained, to some extent, by the ‘seed and soil’ hypothesis which states that cancer cells will be able to colonize a distant site if the cells have intrinsic ability to survive in the microenvironment of the secondary site. Though the process of metastasis is complex and involves other factors, the primary tumor microenvironment conveys survival capabilities to cancer cells that can potentially favor survival in other organs; for this to occur, the cancer cell must engage specific gene expression signatures. However, there are many open questions related to the rise of organ-specific survival traits. Are these gene signatures a result or consequence of genome organization changes? What is the role of genome rearrangements in metastasis organotropism for which there are no significant gene expression alteration? Does the genome organization of the metastatic cell become similar to that of the cells at the distant site? In this work, we have made a step toward answering these questions by analyzing compartmental organization data of breast cancer cell lines derived from lung metastasis in comparison to lung cell lines.

Unsupervised clustering of compartment organization data showed that compartment organization can distinguish cancer cell lines by subtype, providing the first indication that lung metastatic cell lines may show similarities to lung cell genome organization. For example, the triple-negative invasive metastatic breast cancer cell MDA-MB-231, clusters close to the lung carcinoma cells (A549 and H460). The clustering result supports the conclusion that the breast cancer subtypes are not only quite different at the gene expression level, but also at the level of genome organization. Furthermore, it also shows that the genome organization is highly consistent among luminal subtype cancers, irrespective of the tissue from which the cells were collected, while the other subtypes such as triple negative, were more heterogeneous in their structure. This is congruent with studies involving gene expression which have shown that the triple-negative breast cancer category is highly heterogeneous, encompassing six subtypes (42 [43](#)).

When we compared the compartmental organization changes shown by the localized and lung metastatic breast cancers to the compartment state of the lung, we found that, for luminal and triple-negative subtypes, metastatic cells exhibit higher percentages of lung permissive changes compared to localized cancers. Given that these cancer cell lines were derived from lung pleural effusions, we cannot tell whether these metastatic cells acquired a lung-like compartmental organization at the primary tumor site or later while adapting to the metastatic site. To answer this, we would need temporal genome organization data of metastatic progression, which is not currently available. For the lung permissive chromatin conformation changes, we find that shifts toward the A compartment are favored compared to the B compartment. It may be possible that shifts toward the A compartment are predominant because both activation and gene repression are possible once a region is in the A compartment while a shift to the B compartment is not necessary to accomplish gene repression (35 [43](#), 44 [45](#)). Alternatively, this preference for shifts toward an A compartment could also be related to the general loss of heterochromatin previously observed in cancer (45 [46](#)). Indeed, we found that genes in these spatially reorganized regions, in patient samples, tend to change expression concordantly with their compartment changes. Overall, our results are consistent with the idea that genome compartmentalization can play a role in facilitating organ-specific metastasis.

While our gene expression and compartment analyses echo the current dogma that 3D genome compartments correlate with gene expression changes, further analysis revealed that gene expression and compartment signatures across breast cancer types surprisingly capture two orthogonal sets of genomic regions. We found that compartment profiles group breast cancer cell lines according to their epithelial vs. mesenchymal states (46 [↗](#)). This observation becomes much more visible in the PCA transformed data as the cells are arranged along an epithelial-mesenchymal spectrum. While epithelial vs. mesenchymal states have been clearly associated with gene expression, it is also becoming clear that these states exhibit characteristic genome architectures as well. Indeed, a recent study shows that genome organization changes during epithelial-mesenchymal transition in ovarian cancer system, by considering some representative epithelial and mesenchymal cell lines (47 [↗](#)). Here, we extend this conclusion for another major cancer, the breast cancer system by including a spectrum of breast and lung cancer and normal cell lines. Across this cohort, we find not only that both 3D genome architecture and gene expression changes occur during EMT, but also that these changes happen in different regions. The 3D genome changes that most strongly capture the differences across the epithelial and mesenchymal spectrum occur at regions where gene expression is not notably distinct between epithelial and mesenchymal cells. Likewise, the signature gene expression changes occur in regions that tend to be in the A compartment across all cell types. This suggests that many genes key to EMT are regulated by local factors in the A compartment, while major spatial shifts in chromosome regions may relate more to changing the likelihood of future inducibility of genes or alterations in the physical/mechanical properties of the cell during EMT, which can affect chromatin state (48 [↗](#), 49 [↗](#)). Overall, our results suggest that genome spatial compartment changes can help encode a cell state that favors metastasis (EMT) as well as enabling survival in a new organ context (lung or brain).

Methods

Hi-C data processing

To process Hi-C sequencing reads, we used the HiC-Pro (<https://github.com/nservant/HiC-Pro> [↗](#)) pipeline with default parameters (50 [↗](#)). The reads were aligned to the hg19 reference genome. Also, for each cell type, all the available reads from all replicates were merged together to improve the quality of contact data for further downstream processing. After obtaining the valid pair interactions, we binned the interactions at 250 kb resolution for each individual chromosome and normalized the matrices using the ICE technique (51 [↗](#)). Following that, the number of normalized contacts of each chromosome was scaled to 1M using ‘scaleMatrix.pl’ script from the cworld-dekker (<https://github.com/dekkerlab/cworld-dekker> [↗](#)) tool suite. And finally, in order to get the compartmental organization for each individual chromosome, principal components analysis was performed on the normalized and scaled matrices using the ‘matrix2compartment.pl’ script from the cworld-dekker tool suite. The compartment analysis assigns A/B compartment identity to genomic regions based on the sign of the elements of the eigenvector (A-positive and B-negative).

RNA-seq data processing

The RNA-seq files were processed as per the steps described in the previous publication (35 [↗](#)). Briefly, the BBDuk (<https://github.com/kbaseapps/BBTools> [↗](#)) tool was used first for adapter and quality trimming (52 [↗](#)). After that, the reads were mapped to hg19 human reference genome using STAR aligner (RRID:SCR_004463) (53 [↗](#)). Once aligned reads were generated, HTSeq-Counts (<https://github.com/simon-anders/htseq> [↗](#)) tool was used to perform gene level feature count (54 [↗](#)). As the RNA-seq files were obtained from various sources and generated using different experimental techniques, they suffer from the batch effect which masks the biological signal. To remove the batch effect, ComBat-seq (<https://github.com/zhangyuqing/ComBat-seq> [↗](#)) tool was applied on the raw gene counts data (55 [↗](#)).

Hierarchical clustering of genome-wide compartmental organization

To cluster different cell types based on their compartmental organization, hierarchical clustering was performed on the genome-wide compartment identity data. First, the pairwise similarity score between all the cells were calculated using Pearson's correlation. The similarity data was then further converted to distance data by subtracting from 1. And finally, hierarchical clustering was performed on the distance data using 'ward' linkage.

Gene set variation analysis (GSVA)

GSVA was performed to compute the enrichment scores for individual cell lines using RNA-seq data (56 [↗](#), 57 [↗](#)). A list of 232 epithelial signature genes, and a list of 193 mesenchymal signature genes were used to compute the E enrichment score (E score) and M enrichment score (M score) respectively (30 [↗](#), 31 [↗](#), 58 [↗](#)).

Non-negative principal component analysis (nnPCA)

We performed nnPCA using the gene sets mentioned above. nnPCA provided higher resolutions for large number of samples in the EMT spectrum compared to GSVA according to previous studies (31 [↗](#), 59 [↗](#)). nnPCA determines the approximately orthogonal axes with non-negative coefficients (loadings) for features (genes). Variances of projections of data points (cell lines) onto these axes are maximized via an optimization method (60 [↗](#)).

Correction for comparing organ permissive changes across different organs

It is possible that while comparing the host organ permissive changes across different secondary organs, the compartment switch ratio of different secondary localized cancers might include inherent bias to the comparison. To eliminate the bias, for the secondary organ systems, we first calculate the probability of a genomic region to be in a specific compartment in localized cancer given that the region belongs to the same compartment in normal cell (SFig. 3b [↗](#) and SFig. 4d [↗](#)). For the A compartment, $p(TA|NA)$ represents the probability of a region in A compartment in localized cancer cell given that region also belongs to A compartment in normal cell. Similarly, $p(TB|NB)$ for the B compartment. If we have more than one cell line representing the secondary site localized cancer, the minimum probability across all the cells is considered. Once we have those, the probabilities are subtracted from one and the resultant values are multiplied with corresponding secondary organ specific percentages. For example, $(1-p(TA|NA))$ is multiplied to the percentage of BA_AA and $(1-p(TB|NB))$ to the AB_BB.

Data availability

Datasets used for this project were obtained from NCBI GEO, EMBL ENA, 4DN Data Portal and ENCODE repositories and their accession IDs are given in Supplementary Tables 1 and 2.

Acknowledgements

The authors would like to thank all the members of McCord lab for their fruitful suggestions in developing the research project. This work was supported by the National Institutes of Health NIGMS grant R35GM133557 to R.P.M and American Cancer Society postdoctoral fellowship 134060-

PF-19-183-01-CSM to R.S.M. T.H. is supported by the National Institutes of Health NIGMS grant R35GM149531.

Additional information

Disclosure statement

The authors declare no competing interests exist.

Author contributions

P.D. conceived the study in collaboration with R.P.M and R.S.M., performed data analyses, and prepared figures. R.S.M. performed the grouping and ordering of cancer cell lines into biological subtypes and interpreted the biological enrichment analysis. T.H. performed EMT gene signature analyses. R.P.M. supervised and guided the work. P.D. and R.P.M. wrote the manuscript with input from all authors.

References

1. Fragomeni SM, Sciallis A, Jeruss JS (2018) **Molecular subtypes and local-regional control of breast cancer** *Surgical Oncology Clinics* **27**:95–120
2. Chaffer CL, Weinberg RA (2011) **A perspective on cancer cell metastasis** *science* **331**:1559–64
3. Chen W, Hoffmann AD, Liu H, Liu X. (2018) **Organotropism: new insights into molecular mechanisms of breast cancer metastasis** *NPJ precision oncology* **2**:1–12
4. Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ (2016) **A novel gene expression signature for bone metastasis in breast carcinomas** *Breast cancer research and treatment* **156**:249–59
5. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, et al. (2008) **Subtypes of breast cancer show preferential site of relapse** *Cancer research* **68**:3108–14
6. Langley RR, Fidler IJ (2011) **The seed and soil hypothesis revisited—The role of tumor-stroma interactions in metastasis to different organs** *International journal of cancer* **128**:2527–35
7. Obenauf AC, Massagué J. (2015) **Surviving at a distance: organ-specific metastasis** *Trends in cancer* **1**:76–91
8. Sharma S, Xing F, Liu Y, Wu K, Said N, Pochampally R, et al. (2016) **Secreted Protein Acidic and Rich in Cysteine (SPARC) Mediates Metastatic Dormancy of Prostate Cancer in Bone** *J Biol Chem* **291**:19351–63
9. Taichman RS, Patel LR, Bedenis R, Wang J, Weidner S, Schumann T, et al. (2013) **GAS6 receptor status is associated with dormancy and bone metastatic tumor formation** *PLoS One* **8**
10. Shiozawa Y, Berry JE, Eber MR, Jung Y, Yumoto K, Cackowski FC, et al. (2016) **The marrow niche controls the cancer stem cell phenotype of disseminated prostate cancer** *Oncotarget* **7**:41217–32
11. Dalla E, Sreekumar A, Aguirre-Ghiso JA, Chodosh LA (2023) **Dormancy in Breast Cancer** *Cold Spring Harb Perspect Med* **13**
12. Padua D, Zhang XH-F, Wang Q, Nadal C, Gerald WL, Gomis RR, et al. (2008) **TGF β primes breast tumors for lung metastasis seeding through angiopoietin-like 4** *Cell* **133**:66–77
13. Zhang XH-F, Jin X, Malladi S, Zou Y, Wen YH, Brogi E, et al. (2013) **Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma** *Cell* **154**:1060–73
14. Sceneay J, Chow MT, Chen A, Halse HM, Wong CS, Andrews DM, et al. (2012) **Primary tumor hypoxia recruits CD11b⁺/Ly6C^{med}/Ly6G⁺ immune suppressor cells and compromises NK cell cytotoxicity in the premetastatic niche** *Cancer research* **72**:3906–11
15. Gao Y, Bado I, Wang H, Zhang W, Rosen JM, Zhang XH-F. (2019) **Metastasis organotropism: redefining the congenial soil** *Developmental cell* **49**:375–91

16. McCord RP, Kaplan N, Giorgetti L. (2020) **Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function** *Mol Cell* **77**
17. Achinger-Kawecka J, Valdes-Mora F, Luu P-L, Giles KA, Caldon CE, Qu W, et al. (2020) **Epigenetic reprogramming at estrogen-receptor binding sites alters 3D chromatin landscape in endocrine-resistant breast cancer** *Nature communications* **11**:1–17
18. Kim T, Han S, Chun Y, Yang H, Min H, Jeon SY, et al. (2022) **Comparative characterization of 3D chromatin organization in triple-negative breast cancers** *Experimental & Molecular Medicine* :1–16
19. San Martin R, Das P, Dos Reis Marques R, Xu Y, Roberts JM, Sanders JT, et al. (2021) **Chromosome compartmentalization alterations in prostate cancer cell lines model disease progression** *Journal of Cell Biology* **221**
20. Barutcu AR, Hong D, Lajoie BR, McCord RP, van Wijnen AJ, Lian JB, et al. (2016) **RUNX1 contributes to higher-order chromatin organization and gene regulation in breast cancer cells** *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1859**:1389–97
21. Yang Y, Choppavarapu L, Fang K, Naeini AS, Nosirov B, Li J, et al. (2020) **The 3D genomic landscape of differential response to EGFR/HER2 inhibition in endocrine-resistant breast cancer cells** *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1863**
22. Li H, Playter C, Das P, McCord RP (2024) **Chromosome compartmentalization: causes, changes, consequences, and conundrums** *Trends Cell Biol*
23. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, et al. (2016) **A compendium of chromatin contact maps reveals spatially active regions in the human genome** *Cell reports* **17**:2042–59
24. Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, Russiñol N, Di Stefano M, Chapaprieta V, et al. (2021) **Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation** *Nature communications* **12**
25. Owen JA, Osmanović D, Mirny LA (2022) **Design principles of 3D epigenetic memory systems** *bioRxiv*
26. Bennett C, Carroll C, Wright C, Awad B, Park JM, Farmer M, et al. (2022) **Breast Cancer Genomics: Primary and Most Common Metastases** *Cancers* **14**
27. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adélaïde J, Cervera N, et al. (2006) **Gene expression profiling of breast cell lines identifies potential new basal markers** *Oncogene* **25**:2273–84
28. Yin KB (2011) **The Mesenchymal-Like Phenotype of the MDA-MB-231 Cell Line** *IntechOpen*
29. Le AV-P, Szaumkessel M, Tan TZ, Thierry J-P, Thompson EW, Dobrovic A. (2018) **DNA methylation profiling of breast cancer cell lines along the epithelial mesenchymal spectrum—Implications for the choice of circulating tumour DNA methylation markers** *International journal of molecular sciences* **19**
30. Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RYJ, et al. (2014) **Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients** *EMBO molecular medicine* **6**:1279–93

31. Panchy N, Watanabe K, Takahashi M, Willems A, Hong T. (2022) **Comparative single-cell transcriptomes of dose and time dependent epithelial–mesenchymal spectrums** *NAR Genomics and Bioinformatics* **4**
32. Jung AR, Jung C-H, Noh JK, Lee YC, Eun Y-G. (2020) **Epithelial-mesenchymal transition gene signature is associated with prognosis and tumor microenvironment in head and neck squamous cell carcinoma** *Scientific Reports* **10**
33. Cook DP, Vanderhyden BC (2022) **Transcriptional census of epithelial-mesenchymal plasticity in cancer** *Science advances* **8**
34. Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, et al. (2018) **Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity** *Nature communications* **9**
35. Das P, San Martin R, McCord RP (2023) **Differential contributions of nuclear lamina association and genome compartmentalization to gene regulation** *Nucleus* **14**
36. Stik G, Vidal E, Barrero M, Cuartero S, Vila-Casadesús M, Mendieta-Esteban J, et al. (2020) **CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response** *Nature genetics* **52**:655–61
37. Leemans C, van der Zwalm MC, Brueckner L, Comoglio F, van Schaik T, Pagie L, et al. (2019) **Promoter-intrinsic and local chromatin features determine gene repression in LADs** *Cell* **177**:852–64
38. Honda A, Hoeksema MA, Sakai M, Lund SJ, Lakhdari O, Butcher LD, et al. (2022) **The Lung Microenvironment Instructs Gene Transcription in Neonatal and Adult Alveolar Macrophages** *The Journal of Immunology*
39. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. (2017) **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses** *Nucleic acids research* **45**:W98–W102
40. Bachmann C, Schmidt S, Staebler A, Fehm T, Fend F, Schittenhelm J, et al. (2014) **CNS metastases in breast cancer patients: prognostic implications of tumor subtype** *Medical Oncology* **32**
41. Yoneda T, Williams PJ, Hiraga T, Niewolna M, Nishimura R. (2001) **A Bone-Seeking Clone Exhibits Different Biological Properties from the MDA-MB-231 Parental Human Breast Cancer Cells and a Brain-Seeking Clone In Vivo and In Vitro** *Journal of Bone and Mineral Research* **16**:1486–95
42. Wang D-Y, Jiang Z, Ben-David Y, Woodgett JR, Zacksenhaus E. (2019) **Molecular stratification within triple-negative breast cancer subtypes** *Scientific reports* **9**:1–10
43. Hakim O, Sung M-H, Voss TC, Splinter E, John S, Sabo PJ, et al. (2011) **Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements** *Genome research* **21**:697–706
44. Vieux-Rochas M, Fabre PJ, Leleu M, Duboule D, Noordermeer D. (2015) **Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain** *Proceedings of the National Academy of Sciences* **112**:4672–7

45. Carone DM, Lawrence JB (2013) **Heterochromatin instability in cancer: from the Barr body to satellites and the nuclear periphery** Elsevier :99–108
46. D'Amato NC, Ostrander JH, Bowie ML, Sistrunk C, Borowsky A, Cardiff RD, et al. (2012) **Evidence for phenotypic plasticity in aggressive triple-negative breast cancer: human biology is recapitulated by a novel model system** *PLOS ONE* **7**
47. Pang QY, Tan TZ, Sundararajan V, Chiu Y-C, Chee EYW, Chung VY, et al. (2022) **3D genome organization in the epithelial-mesenchymal transition spectrum** *Genome Biology* **23**:1–31
48. Wang N, Zhang M, Chang Y, Niu N, Guan Y, Ye M, et al. (2019) **Directly observing alterations of morphology and mechanical properties of living cancer cells with atomic force microscopy** *Talanta* **191**:461–8
49. Scott AK, Rafuse M, Neu CP (2023) **Mechanically induced alterations in chromatin architecture guide the balance between cell plasticity and mechanical memory** *Frontiers in Cell and Developmental Biology* **11**
50. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. (2015) **HiC-Pro: an optimized and flexible pipeline for Hi-C data processing** *Genome biology* **16**:1–11
51. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. (2012) **Iterative correction of Hi-C data reveals hallmarks of chromosome organization** *Nat Methods* **9**
52. Bushnell B, Rood J, Singer E. (2017) **BBMerge—accurate paired shotgun read merging via overlap** *PloS one* **12**
53. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2013) **STAR: ultrafast universal RNA-seq aligner** *Bioinformatics* **29**:15–21
54. Anders S, Pyl PT, Huber W. (2015) **HTSeq—a Python framework to work with high-throughput sequencing data** *bioinformatics* **31**:166–9
55. Zhang Y, Parmigiani G, Johnson WE (2020) **ComBat-seq: batch effect adjustment for RNA-seq count data** *NAR genomics and bioinformatics* **2**
56. Deshmukh AP, Vasaikar SV, Tomczak K, Tripathi S, Den Hollander P, Arslan E, et al. (2021) **Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing** *Proceedings of the National Academy of Sciences* **118**
57. Hänzelmann S, Castelo R, Guinney J. (2013) **GSEA: gene set variation analysis for microarray and RNA-seq data** *BMC bioinformatics* **14**:1–15
58. Cursons J, Pillman KA, Scheer KG, Gregory PA, Foroutan M, Hediye-Zadeh S, et al. (2018) **Combinatorial targeting by microRNAs co-ordinates post-transcriptional control of EMT** *Cell systems* **7**:77–91
59. Panchy N, Watanabe K, Hong T. (2021) **Interpretable, scalable, and transferrable functional projection of large-scale transcriptome data using constrained matrix decomposition** *Frontiers in Genetics* **12**
60. Sigg CD, Buhmann JM (2008) **Expectation-maximization for sparse and non-negative PCA** *Proceedings of the 25th International Conference on Machine Learning* :960–7

Author information

Priyojit Das[†]

UT-ORNL Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, USA

ORCID iD: [0000-0002-6774-6718](https://orcid.org/0000-0002-6774-6718)

[†]Current address: Department of Molecular Biology, Massachusetts General Hospital, Boston, USA; Department of Genetics, Harvard Medical School, Boston, USA

Rebeca San Martin[‡]

Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, USA

ORCID iD: [0000-0001-7249-3922](https://orcid.org/0000-0001-7249-3922)

[‡]Current address: Department of Cell Biology, Ruth L and David S. Gottesman Institute for Stem Cell Research and Regenerative Medicine, Albert Einstein College of Medicine. Bronx, USA.

Tian Hong⁺

UT-ORNL Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, USA, Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, USA

ORCID iD: [0000-0002-8212-7050](https://orcid.org/0000-0002-8212-7050)

⁺Current address: Department of Biological Sciences, The University of Texas at Dallas, Richardson, USA. Cancer Dormancy Institute, Department of Medicine (Oncology). Montefiore Einstein Comprehensive Cancer Center Bronx, USA

Rachel Patton McCord

UT-ORNL Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, USA, Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, USA

ORCID iD: [0000-0003-0010-5323](https://orcid.org/0000-0003-0010-5323)

For correspondence: rmccord@utk.edu

Editors

Reviewing Editor

Anna Panchenko

Queen's University, Kingston, Canada

Senior Editor

Aleksandra Walczak

École Normale Supérieure - PSL, Paris, France

Reviewer #1 (Public review):**Summary:**

This study utilized publicly available Hi-C data to ensemble a comprehensive set of breast cancer cell lines (luminal, Her2+, TNBC) with varying metastatic features to answer whether breast cancer cells would acquire organ-specific features at the 3D genome level to metastasize to that specific organ. The authors focused on lung metastasis and included several controls as the comparison including normal mammary lines, normal lung epithelial lines, and lung cancer cell lines. Due to the lower resolution at 250KB binning size, the authors only addressed the compartments (A for active compartment and B for inactive compartment) not the other 3D organization of the genome. They started by performing clustering and PCA analysis for the compartment identity and discovered that this panel of cell lines could be well separated based on Her2 and epithelial-mesenchymal features according to the compartment identity. While correlating with the transcriptomic changes, the authors noticed the existence of concordance and divergence between the compartment changes and transcriptomic changes. The authors then switched gears to tackle the core question of metastatic organotropism to the lung. They discovered a set of "lung permissive compartment changes" and concluded that "lung metastatic breast cancer cell lines acquire lung-like genome architecture" and "organotropic 3D genome changes match target organ more than an unrelated organ". To prove the latter point, the authors enlisted an additional non-breast cancer cell line (prostate cancer) in the setting of brain metastasis. This is a piece of pure dry computational work without wet bench experiments.

Strengths:

The authors embarked on an ambitious journey to seek the answer regarding 3D genome changes predisposing to metastatic organotropism. The authors succeeded in the assembly of a comprehensive panel of breast cancer cell lines and the aggregation of the 3D genome structure data to conduct a hypothesis-driven computation analysis. The authors also achieved in including proper controls representing normal non-cancerous epithelium and the end organ of interest. The authors did well in the citation of relevant references in 3D genome organization and EMT.

Weaknesses:

(1) The authors should clearly indicate how they determine the patterns of spread of the breast cancer cell lines being utilized in this manuscript. How did the authors arrive at the conclusion that certain cell lines would be determined as "localized spread" and "metastatic tropism to the lung"? This definition is crucial, and I will explain why.

Todd Golub's team from the Broad Institute of MIT and Harvard published "A metastasis map of human cancer cell lines" to exhaustively create a first-generation metastasis map (MetMap) that reveals organ-specific patterns of metastasis. (By the way, this work was not cited in the reference in this manuscript.) The MetMap Explorer (<https://depmap.org/metmap/vis-app/index.html>) is a public resource that could be openly accessed to visualize the metastatic potential of each cell line as determined by the in vivo barcoding approach as described in the MetMap paper in the format of petal plots. 5 organs were tested in the MetMap paper, including brain, lung, liver, kidney, and bone. The authors would discover that some of the organ-specific metastasis patterns defined in the MetMap Explorer would be different from the authors' classification. For example, the authors defined MCF7 as a line as lung metastatic, and rightly so the MetMap charted a signal towards lung with low penetrance and low metastatic potential. The authors defined ZR751 as a line with localized spread, however, the MetMap charted a signal towards the kidney with low penetrance and low metastatic potential, the signal strength similar to the lung metastasis in MCF7. A similar argument could be made for T47D. The TNBC line MDA-MB-231 is indeed highly metastatic,

however, in MetMap data, its metastasis is not only specific to the lung but towards all 5 organs with high penetrance and metastatic potential. The 2 lung cancer cell lines mentioned in this study, A549 and H460, the authors defined them as localized spread to the lung. However, the MetMap data clearly indicated that A549 and H460 are highly metastatic to all 5 organs with high penetrance and high metastatic potential.

Since results will vary among different experimental models testing metastatic organotropism, (intra-cardiac injection was the metastasis model being adopted in the MetMap), the authors should state more clearly which experimental model system served as the basis for their definition of organ-specific metastasis. In my opinion, this is the most crucial first step for this entire study to be sound and solid.

(2) Figure 1b: The authors found that "MDA-MB-231 cells were grouped with the lung carcinoma cells. This implies that the genome organization of this cell line is closer to that of lung cells than to other breast epithelial cell lines.". In fact, another TNBC line BT549 was also clustered under the same clade. So this clade consisted of normal-like and highly metastatic lines. Therefore, the authors should be mindful of the fact that the compartment features might not directly link to metastasis (or even metastatic organotropism).

(3) Figure 3: In the text, the authors stated, "To further investigate this result, we examined the transcription status of genes that changed compartment across the EMT spectrum and, conversely, the compartment status of genes that changed transcription (Fig. 3b, c, and d)". However, it was not apparent in the figure that the cell lines were arranged according to an EMT spectrum. Also, the clustering heatmaps did not provide sufficient information regarding the genes with concordant/divergent compartments vs transcription changes. It would be more informative if the authors could spend more effort in annotating these genes/pathways.

(4) Figure 4: The title of the subheading of this section was 'Lung metastatic breast cancer cell lines acquire lung-like genome architecture'. Echoing my comments in point 1, I am a bit hesitant to term it as "lung metastatic" but rather "metastatic" in general since cell lines such as MDA-MB-231 do metastasize to other organs as well. However, I do get the point that the definition of "lung metastasis" is derived from the common metastasis features among the cell lines here (MCF7, T47D, SKBR3, MDA-MB-231).

There might be another argument about whether the "lung" carcinoma cell lines can be considered "localized" since they are also capable of metastasizing to other organs. In a way, what the authors probably were trying to leverage here is the "tissue" identity of that organ. Having said this, in addition to showing the "lung permissive changes", the authors should show the "breast identity conservation" as well. Because this section started to deal with the concept of "tissue/lineage identity", the authors should also clarify whether these breast cancer cell lines capable of making lung metastasis are also preserving their original tissue identity from the compartment features (which would most likely be the case).

(5) Rest of the sections: The authors started to claim that the organ-specific metastasis permissive compartmental features mimic the designated end organ. The authors utilized additional non-breast cancer cell lines (prostate cancer cell lines LNCaP as localized and DU145 as brain metastatic) in brain metastasis to strengthen this claim. (DU145 in MetMap again is highly metastatic to lung, brain, and kidney). However, this makes one wonder that for cell lines that are capable of metastasizing to multiple organ sites (eg. MDA-MB-231, DU145, A549, H460), does it mean that they all acquire the permissive features for all these organs? This scenario is clinically relevant in Stage 4 patients who often present with not only one metastatic lesion in one single organ but multiple metastatic lesions in more than one organ (eg. concomitant liver and lung metastasis). Do the authors think that there might be different clones having different tropism-permissive 3D genome features or there might be evolutionary trajectory in this?

In my opinion, to further prove this point, the authors might need to consider doing in vivo experiments to collect paired primary and organ-specific metastatic samples to look at the 3D genome changes.

(6) Technically, the study utilized public Hi-C data without generating new Hi-C data. The resolution of the Hi-C data for compartments was set at 250KB as the binning size indicating that the Hi-C data was at lower resolution so it might not be ideal to address other 3D genome architecture changes such as TADs or long-range loops. It is therefore unknown whether there might be permissive TAD/loop changes associated with organotropism and this is the limitation of this study.

(7) In the final sentence of the discussion the authors stated "Overall, our results suggest that genome spatial compartment changes can help encode a cell state that favors metastasis (EMT)". The "metastasis (EMT)" was in fact not clearly linked inside the manuscript. The authors did not provide a strong link between metastasis and EMT in their result description. It is also unclear whether the EMT-associated compartment identity would also correlate with the organotropic compartment identity.

<https://doi.org/10.7554/eLife.103697.1.sa2>

Reviewer #2 (Public review):

Summary:

This work addresses an important question of chromosome architecture changes associated with organotopic metastatic traits, showing important trends in genome reorganization. The most important observation is that 3D genome changes consistent with adaptations for new microenvironments, including lung metastatic breast cells exhibiting signatures of the genome architecture typical to a lung cell-like conformation and brain metastatic prostate cancer cells showing compartment shifts toward a brain-like state.

Strengths:

This work presents interesting original results, which will be important for future studies and biomedical implications of epigenetic regulation in norm and pathology.

Weaknesses:

The authors used publicly available data for 15 cell types. They should show how many different sources the data were obtained from and demonstrate that obtained results are consistent if the data from different sources were used.

<https://doi.org/10.7554/eLife.103697.1.sa1>

Author response:

We appreciate the constructive feedback from the reviewers and will work to address many of these concerns in a revised version. Here, we provide initial responses to a few key points that the reviewers raised:

(1) The reviewers rightly pointed out that it is very important to clearly define and explain what qualifies as metastatic potential to particular organs in our system. We acknowledge the valuable contributions of animal models in metastatic cancer studies, but here we intentionally limited our scope to metastasis that had occurred within the human system only. For example, we use data from cancer cells that model human organotropism from the

breast to the lung, since the cells originated from infiltrative ductal carcinoma (human breast) but were collected from pleural effusions (human lung). We propose that in this case a comparison with a human lung cancer-derived cell line that was itself purified from a pleural effusion could reveal factors essential for lung metastasis, without adding the confounder of an animal microenvironment. The MetMap Explorer contains valuable information, but the “metastatic potential of each cell line” is measured in a mouse environment. Knowing that a particular cell line, which originated from a human lung metastasis, can further metastasize to other organs in a mouse does not necessarily mean that those cells could do so in humans. The microenvironment responses to metastatic colonization can differ among species. Further, the changes a cell needs to make to adapt to a new organ system in a mouse could be confounded by the changes needed to adapt to mouse conditions in general. Finally, migration from a site of ectopic injection may not mimic migration from an initial tumor site. We agree that the very best data would come from matched primary and metastatic tumors in the same human patient, but those data do not currently exist and generating them would require future work beyond the scope of this study. In our revision, we will ensure that we more clearly explain how and why we chose the cell lines we did and what the advantages and limitations of this choice are.

(2) The reviewers are correct that our unsupervised Principal component analysis (PCA) does not precisely stratify cells according to epithelial-mesenchymal status. In a high dimensional, complex system, it is expected that an unsupervised analysis such as this will not capture just one biological feature in the first principal component. Therefore, when we performed PCA on the compartmental organization profiles of different healthy and cancerous cell lines, instead of finding the largest variation (PC1) following exactly EMT state, it captured an ordering that includes influences from epithelial-mesenchymal state, disease condition, nuclear geometry, and other cellular properties. However, it was striking that this completely unsupervised analysis did match previous annotations of EMT state so well (as seen in supp fig 1b). Therefore, we conclude that the most prominent variations in A/B compartment signature strongly relate to EMT state. In the revision, we will more clearly present the caveats of this interpretation.

(3) Our decision to focus on A/B compartmentalization rather than TAD or loop structure in this analysis was intentional and biologically motivated, rather than solely being a reflection of data resolution. Both compartments and topologically associated domains (TADs) are key parts of genome organization and disruption of these structures has the potential to alter downstream gene regulation, as shown by numerous studies. But, compartments have been found, more so than TADs, to be strongly associated with cell type and cell fate. Therefore, in this manuscript, we decided to focus only on the compartment organization changes between different healthy and cancerous cells as they are more likely to represent the stable alterations of the genome organization malignant transformations.

<https://doi.org/10.7554/eLife.103697.1.sa0>