Regulating epithelial-mesenchymal plasticity from 3D genome organization

Check for updates

Qing You Pang [1], Yi-Chia Chiu[2] & Ruby Yun-Ju Huang [2,3,4] ✉

Epithelial-mesenchymal transition (EMT) is a dynamic process enabling polarized epithelial cells to acquire mesenchymal features implicated in development and carcinoma progression. As our understanding evolves, it is clear the reversible execution of EMT arises from complex epigenomic regulation involving histone modifications and 3-dimensional (3D) genome structural changes, leading to a cascade of transcriptional events. This review summarizes current knowledge on chromatin organization in EMT, with a focus on hierarchical structures of the 3D genome and chromatin accessibility changes.

The paradigm shift of EMT from a binary to a spectral perspective

Epithelial-mesenchymal transition (EMT) is a cellular process enabling epithelial cells to lose their epithelial features and adopt distinct mesenchymal phenotypes facilitating cell invasion and migration. EMT occurs in both physiological conditions such as gastrulation and wound healing and pathological conditions such as tissue fibrosis and cancer progression. The conventional view of EMT emphasizes on the transition in between two broadly defined states[1,2]. The "epithelial" and the "mesenchymal" states are viewed from their cellular features such as apical-basal polarity, cell-cell adhesion junctions, cell-matrix interactions, and their functional consequences in the migratory and invasive behaviour. These two broadly defined traits have established the binary concept of EMT and the reversibility during transition. Core regulatory networks that encompass a repertoire of EMT transcription factors and its downstream effector have been known to govern the binary EMT process[3]. Recently, the concept of EMT has undergone a major paradigm shift to the notion of multiple intermediate states of hybrid phenotypes as an EMT Spectrum[4–6] with epithelial-mesenchymal plasticity (EMP) governing the convertibility among these states[5] (Fig. 1). These in-between states have more dynamic and complex 'mix-and-match' of the above-mentioned cellular features. The understanding of how plasticity is regulated along the entire EMT spectrum thus becomes more challenging. For instance, the detailed transcriptional programme guiding the plasticity in between the various states within the EMT spectrum is less clear. Recent data has suggested that there might be transcriptional regulatory network switches to trigger the sequential transition during carcinoma progression[7]. A hierarchical regulatory landscape between 46 (co)transcription factors and 13 miRNAs was identified to be critically required for EMT in the non-tumorigenic mouse mammary epithelial cells NMuMG[8]. But what has been of contention is whether these transcriptional repertoires are context specific[9] that there might not be universal master transcription regulators particularly in scenarios with known diverse heterogeneity such as tissue fibrosis and cancer metastasis.

Multi-scale regulations of EMT

With the context-dependent nature of EMT regulation in mind, there is a need to go beyond transcriptional programs and examine the regulatory landscape of EMT at the genomic level. Additionally, integrating different perspectives is crucial to enable a broad and in-depth approach for a more comprehensive understanding. In a recent updated guideline for EMT research[5], the official academic society for the EMT field, The EMT International Association (TEMTIA), recommended that EMT status cannot be assessed based solely on one or a small number of molecular markers and these molecular markers together with EMT driving transcription factors should be assessed in conjunction with changes in cellular characteristics to define EMT. The guideline concluded that a combinatorial approach should be adopted to identify reliable EMT markers and their regulatory mechanisms.

To understand the multi-scale regulations of EMT, one could refer to how regulations of cell fate determination have been studied. The EMT process is crucial during the gastrulation stage of early embryo development for neural crest cells to delaminate and embark on differentiation towards their derivatives. During neural crest migration, after progressing through a sequence of common transcriptional traits, a series of binary decisions are made to determine subsequent cell fates[10]. The intertwined execution of EMT and cell fate commitment during neural crest development implies that the determination of epithelial and mesenchymal traits follows the concept of cell fate determination as depicted by the Waddington landscape proposed by Conrad Waddington in 1957. The entire landscape is featured with ridges and valleys indicating the barriers between each lineage and state with different intermediates to feature a plasticity spectrum. Therefore,

[1]Neuro-Oncology Research Laboratory, National Neuroscience Institute, Singapore 308433, Singapore. [2]School of Medicine, College of Medicine, National Taiwan University, Taipei 10051, Taiwan. [3]Center for Advanced Computing and Imaging in Biomedicine, National Taiwan University, Taipei 10051, Taiwan. [4]Department of Obstetrics & Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119077, Singapore. ✉e-mail: rubyhuang@ntu.edu.tw

Fig. 1 | Cells at different phases of the EMT spectrum display varying combinations and degrees of cellular features associated with epithelial or mesenchymal states. Epithelial cells form connections with one another through different types of junctions, including adherens junctions, desmosomes, gap junctions, and tight junctions. The proper organization these cell-cell adhesion in epithelial cells is guided by apical-basal polarity. The epithelial cells are connected to the underlying basement membrane via hemidesmosomes, which contains integrin that facilitates binding to the basement membrane and is also linked to cytokeratins within the cell. Mesenchymal cells lack functional epithelial junctions and exhibit back-front polarity in their actin stress fibers. Vimentin-based intermediate filaments are present in mesenchymal cells, and they use integrin-containing focal adhesions to attach to the extracellular matrix. Some of the transcription factors and regulators known to regulate the different phases of the epithelial-to-mesenchymal transition are shown below the spectrum (epithelial regulators – purple, and mesenchymal regulators, orange).
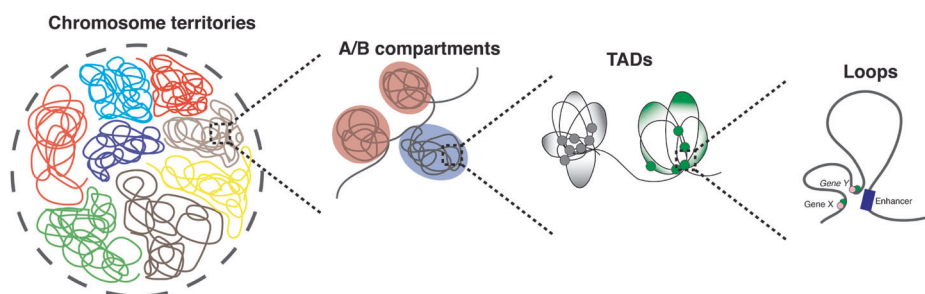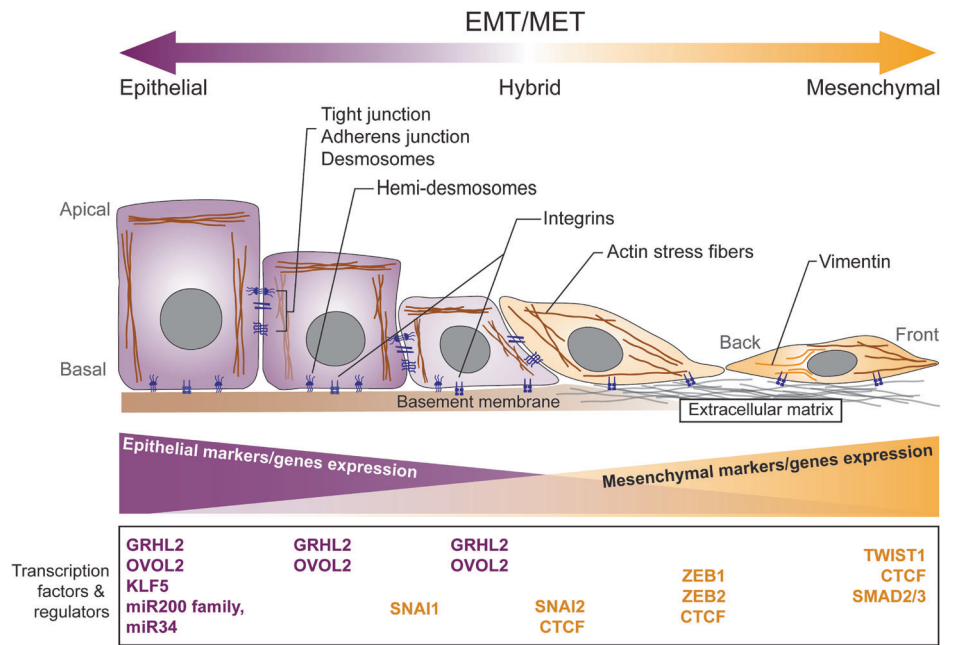

Fig. 2 | Hierarchical order of the 3D genome. The 3D genome in the nucleus is organized into hierarchical units of chromatin, with the formation of chromosomal territories (CTs) at the chromosomal resolution, where each chromosome occupies a distinct and non-overlapping area. Within CTs, there are chromatin compartments, comprising A compartments with open euchromatin enriched in highly expressed genes, and B compartments with closed heterochromatin associated with repressed transcription regions. Further organization within the compartments involves topologically associated domains (TADs), stable units guiding genome folding and long-range regulation. The finest level, often involving kilobases resolution, consists of chromatin loops which establish interactions between distant promoters and enhancers, playing a pivotal role in directly regulating gene expression.

reprogramming (direct or indirect) would require the system to overcome these barriers via the manipulation of various regulatory gatekeepers in multi-directionality[11]. Plasticity and fluidity of cell fate determination or EMT execution is thus dictated by the balance between different regulatory circuits governing the genome architecture, epigenetic modifications, transcription, translation, and post-translational modifications. Several reviews have nicely summarized the transcriptional[12], translational[13] control and post-translational modifications[14] in EMT and will not be repeated here. In this Mini-Review, we will discuss studies looking at the regulation of EMT from the lenses of 3D genome organization and epigenetic reprogramming to explain how the plasticity along the EMT Spectrum could be established.

3D Genome 101
The genome architecture in three dimensions
Following the emergence of chromatin conformation capture-based methods[15–19], the studies of the chromatin architecture has broadened our understandings on how the genome follows hierarchical organization inside the nucleus. Harnessing the ease of mapping the 3D genome with these technology breakthroughs, the reorganization of the 3D chromatin conformation has been shown to influence cell identity during lineage-differentiation, cell cycle progression and could be dysregulated in diseases[20–24].

The 3D genome is organized in the hierarchical orders from chromosome territories, active/inactive (A/B) chromosomal compartments, topological associating domains (TADs), and chromatin interactions involving long- and short-ranged DNA looping. Our genome thus follows certain origami-like patterns in coordination with a variety of components, such as transcription factors, architectural proteins, chromatin regulators, and non-coding RNAs to modulate transcription. At the largest resolution, chromosomes preferentially form topological territories with limited intermingling between chromosomes[25] (Fig. 2). The chromosome territories can be further sub-divided into 2 main compartments corresponding to different chromatin types[19], namely A (active) and B (inactive). Within the compartments, specific loci are arranged in spatial proximity to each other thereby forming TADs[22], architectural "stripes"[26] and loops[27].

Chromatin conformation capture-based techniques
The 3D genome can be studied in a genome-wide fashion using proximity ligation-based chromatin conformation capture methods such as Hi-C 1.0, HiC 2.0, HiC 3.0, Micro-C and Pore-C[27–32], which often requires deep sequencing to fully map out the chromatin structural features. Alternatively, protein- or gene-centric chromatin conformation capture could be applied to interrogate the 3D genome in a more efficient and cost-effective manner[18,33–39]. With the plethora of techniques made available to map the

3D genome architecture, it is essential that one understands the information output of each technique before applying it to an EMT model system.

In light of approaches that focuses on local chromatin conformation of a handful of genes, low throughput chromatin conformation capture methods could prove to be useful such as 3C, where the interaction frequency between two or more interacting regions could be quantified[24,40]. In addition to query specific one-one interactions between two genomic loci, circularised chromatin conformation capture (4C) to look at the chromatin interactions originating from one loci[41,42], or across many target loci along a continuous genomic region in chromosome conformation capture carbon copy (5C)[16].

Besides ligation-based techniques, there are other alternatives to capture the 3D genome architecture, which includes genome architecture mapping (GAM)[43], split-pool recognition of interactions by tag extension (SPRITE)[44] and chromatin interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop)[45]. In addition, there were two recent techniques that allowed the capture of chromatin conformation without the need for formaldehyde prefixing, namely DNA adenine methyltransferase (Dam)C[46] and chemical-crosslinking assisted proximity capture (CAP-C)[47]. In essence, there are several methods available to map the 3D genome, as nicely summarized by previous reviews[48–50], ranging from a high-throughput fashion such as HiC, SPRITE or GAM; to limited throughput techniques such as 3C, 4C and 5C which are capable of assessing specific enhancer-promoter interactions.

Reconstructing the 3D genome in EMT models
The regulation of EMT from the epigenomic perspective has been elucidated from several recent studies. It is now clear that the transitional phases of EMT occurs together with complex changes at the epigenomic landscape[7,51–54]. During development, the expression of EMT genes have been shown to be regulated via both histone modifications and the local enhancer-promoter looping[55]. Similar regulatory mechanisms to the EMT genes have also been shown to exist in cancer cells[51,52,56,57]. The 3D genome hierarchical orders, together with histone modifications and local DNA looping, establishes the architectural backbone of the genome in response to any functional ques during EMT or cell fate determination during development and in cancer cells[53,58].

Higher order organizations in chromatin domains
Genome-wide analysis of the 3D genome organization in EMT has been very limited among the vast expanding literature body on EMT. Two studies[59,60] have shed light on potential changes in the 3D genome organization during EMT, which could occur at the nuclear periphery. These studies underscores the significance of large organized chromatin K9 modifications[61] (LOCKs) and nuclear lamina-associated domains[62,63] (LADs) in playing a contributory role in the functional reorganization of the 3D genome within the nucleus. Notably, both LOCKs and LADs are typically transcriptionally repressed chromatin regions while often in proximity to the nuclear lamina[61,63]. The studies explore their involvement and events leading to the transcriptional activation of EMT-related genes during the EMT process. Using chromatin immunoprecipitation followed by genomic DNA microarray analysis (ChIP-chip) in TGF-β-induced AML12 mouse hepatocytes, the study by McDonald et al.[59] has implicated extensive chromatin reprograming during EMT at the LOCKs, which are large (100 kb–5 Mb), non-repetitive heterochromatin domains enriched in H3K9me2[59]. In this model, there is a global reduction in the heterochromatin mark H3 Lys9 dimethylation (H3K9me2), accompanied by the increase in euchromatin marks H3 Lys4 trimethylation (H3K4me3) and mark H3 Lys36 trimethylation (H3K36me3). The quantitative reduction in H3K9me2 is due to LSD1 demethylation the LOCKs, leading to the recruitment of H3K4me3 within the LOCKs and the enrichment of H3K36me3 at the boundaries of the LOCKs. This enrichment of H3K36me3 at the boundaries leads to the transcriptional activation of genes with EMT-related functions. Further studies would need to be conducted to assess whether these LOCK-specific H3K4me3 changes could trigger the

release from the nuclear lamina, prior to transcriptional activation. While lamin A/C's functional role in gene regulation is well-established, it predominantly associates with euchromatic regions, whereas lamin B is primarily linked to heterochromatin[64]. Interestingly, the presence of lamin B1 within euchromatic regions follows a dynamic nature. Utilizing a TGF-β induced EMT model in non-transformed mouse mammary gland epithelial cell line (NMuMG)[60], the authors discovered the existence of euchromatin LADs (eLADs), which is formed as lamin B1 engages with euchromatin regions. These eLADs exhibit transcriptional activity in the active compartments throughout the EMT process (Fig. 3). Intriguingly, as the cell becomes mesenchymal, there is a surge in transcriptionally inactive eLADs within the inactive compartments. This increase corresponds to newly formed eLADs, disputing the notion of eLADs merely shifting between active and inactive compartments. Moreover, the study identified a correlation between the enrichment of lamin B1 and the strength of TAD boundaries in active chromatin compartments, highlighting the possible role of lamin B1 as an architectural protein in mediating the establishment of new genome architecture during EMT.

These two studies support the notion that LADs and possibly the interface between the heterochromatin and euchromatin at the compartment level serve as the crucial chromatin structural organizing sites.

Dynamics of active and inactive compartments in EMT
The comprehension of compartment changes during EMT remains incomplete due to limited available studies. However, insights into the dynamics of 3D genome compartments during EMT can be inferred from observations in two different cell line models, namely (i) stem cell differentiation[65] and (ii) cancer EMT[53]. Both models employed the Hi-C technique to interrogate the 3D genome organization. One model focused on human embryonic stem cells and their derivative differentiated cell lines[65], while the other involved cancer cells representing various transitional phases of EMT[53]. In both the stem cell differentiation model (Fig. 4a) and the cancer EMT cell line model, the A/B compartments housing EMT genes displayed dynamic changes and conservation across different EMT transitional phases. This underscores the role of A/B compartments in regulating stem cell differentiation, as previously reported[65,66], which is also evident in the epithelial and mesenchymal compartments. However, it's important to note that the magnitude of change in the compartments doesn't necessarily correspond to a similar change in the expression of specific epithelial/mesenchymal genes (Fig. 4b). Only a subset of EMT genes exhibit such concordant dynamics between compartments and gene expression (Fig. 4c). For instance, classical epithelial and mesenchymal genes, including KRT19, LOX, COL5A2, and CDH2, exhibited concordant compartmental switch between EM states in both the stem cell differentiation (Fig. 4c) and cancer EMT cell line models[53]. Taken together, the two models collectively indicate distinct regulation patterns for epithelial and mesenchymal genes at the 3D genome level. This distinction was further elucidated in the cancer EMT model, where epithelial genes were progressively repressed along the EMT spectrum, characterized by increased H3K27me3 binding and cell-type specific changes in chromatin interactions. In contrast, the regulation of the mesenchymal genes was more closely associated with the changes in local chromatin conformation, where contact frequency within the mesenchymal TADs were shown to increase along the EMT spectrum[53]. The reprogramming of histone modifications within the TADs was also evident in an TGF-β induced EMT model in lung cancer cells, dynamic changes of specific histone H3 lysine 27 acetylation (H3K27ac)-marked enhancers mainly occur within pre-existing TADs in epithelial cells to activate mesenchymal genes[67]. This underscores the inherent and distinct regulation of epithelial and mesenchymal genes in EMT at the 3D genome level.

Promoter-enhancer interaction
The activation of gene transcription relies on the crucial physical contact between enhancers and promoters[68], often necessitating DNA looping, which can span distances ranging from 1 kb to several megabases[69–73]. Studies on chromatin conformation changes during EMT have employed
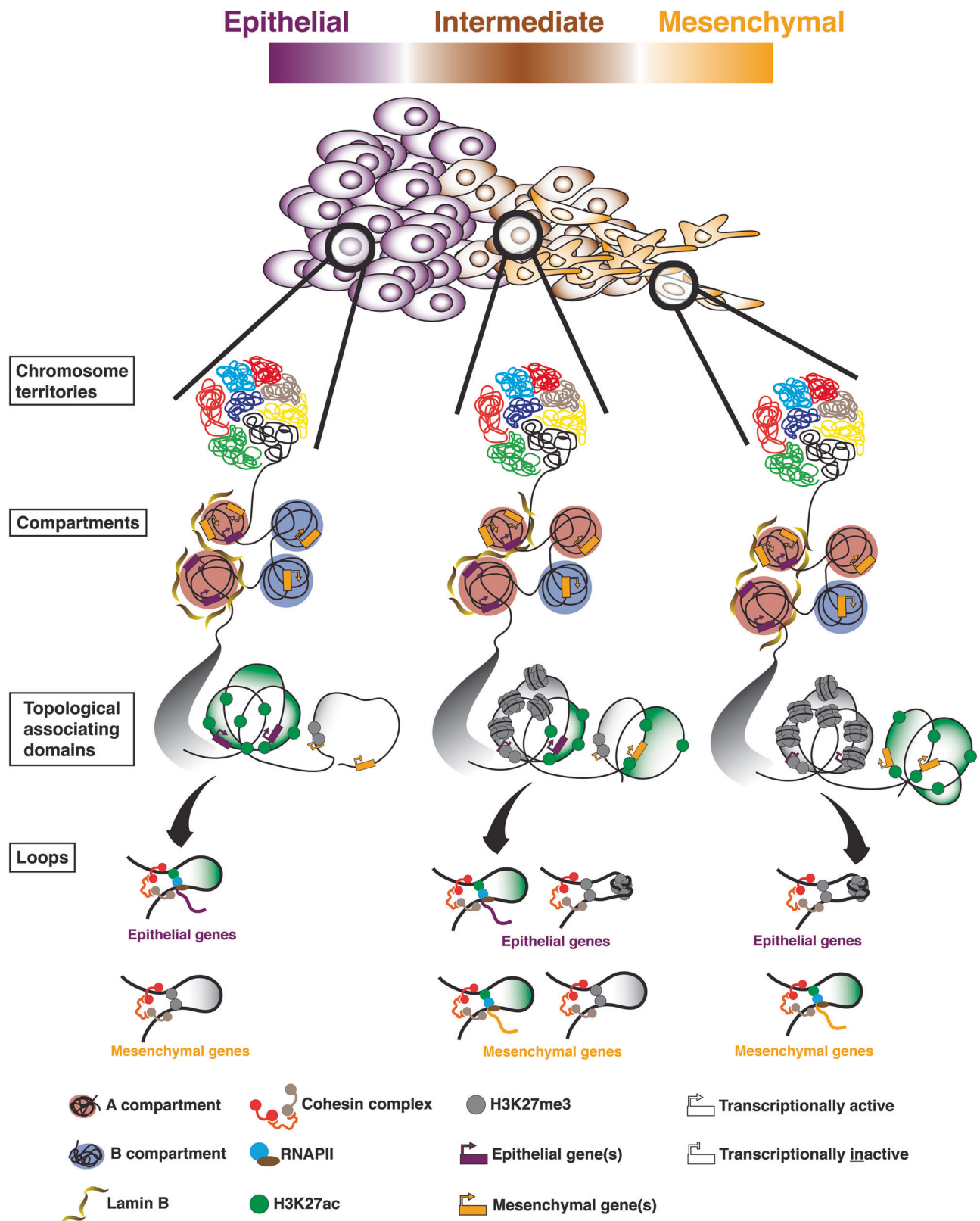
Fig. 3 | Illustration of the dynamic changes in the hierarchical organization of the 3D genome across different EMT transitional phases along the spectrum. The transition from an epithelial to mesenchymal state is depicted from left to right in the illustration. The hierarchical orders of the 3D genome organization, viewed from larger to smaller resolution (top to bottom) — chromosome territories, chromatin compartments, TADs and DNA loops —are indicated on the left. At the order of chromatin compartments, the eLADs are shown to be involved in the formation of new inactive chromatin compartments as the cells becomes mesenchymal. While the TAD borders remain relatively conserved between the EMT, the dynamic changes in histone modifications and chromatin conformation within the TADs are crucial in transcriptional regulation of EMT genes. Lastly, the formation of EMT-TF centric enhancer-promoter interactions, via loop extrusion model, within the TADs is required for activation of EMT gene transcription.
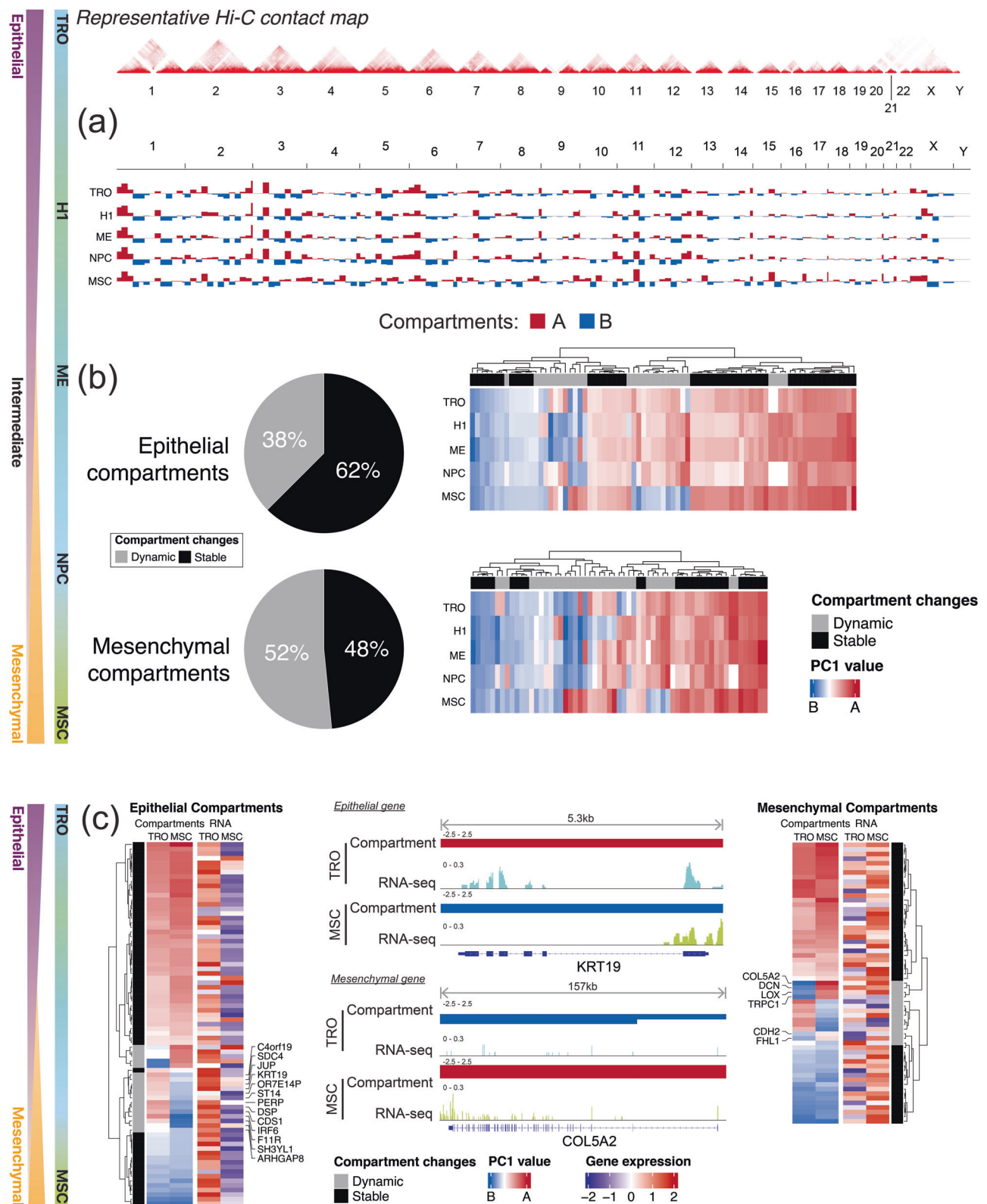
Fig. 4 | Illustration of 3D genome compartment dynamics of EMT in stem cell differentiation and cancer models. Stem cell model was derived from the differentiation of H1-hESC cells. H1 – H1 hESC, ME – mesendoderm, TRO – trophoblast-like cell, NPC – neural progenitor cell and MSC – mesenchymal stem cell. Cell lines are arranged by their EMT score. **a** Genome-wide view of A and B compartments, colored red and blue respectively, in H1-hESC and its differentiated cell lines. **b** Pie chart depicting the distribution of dynamic (gray) and stable (black) compartments across the cell lines at both epithelial and mesenchymal compartments. Heatmap representation of the compartment PC1 values are shown on the right of each pie chart. **c** Heatmap representation of PC1 and expression values of epithelial (right) and mesenchymal (left) genes in TRO (EMT score = −0.17) and MSC (EMT score = 0.62). IGV Snapshots of an epithelial (KRT19; 5.3 kilobase locus) and mesenchymal (COL5A2; 157 kilobase locus) gene demonstrating concordant compartment and expression dynamics is shown in the middle.

the 3C technique to investigate the regulation of local promoter and enhancer interactions for specific genes[55,56,74]. For instance, in the *Drosophila* embryo, Snail represses the transcription of 2 dorsal-ventral patterning genes *sog* and *brk* by blocking the formation of promoter-enhancer interactions when Snail is bound to the distal enhancer[75]. In the mouse epithelium, the epithelial-specific transcription factor (TF) GRHL2 regulates the looping between the promoter and the enhancer of *Cdh1* at intron 2 to maintain epithelial integrity[55]. During TGF-β induced-EMT in NMuMG cells, the expression of E-cadherin was mediated by DNA looping between the transcription start site (TSS) of *Cdh1* to the enhancers bound by GRHL2 and HNFα[74]. These findings confirm that short-range local promoter-enhancer interactions, regulated by the TF complex via DNA looping, play a crucial role at EMT gene loci.

Cohesin-mediated loop formation

The stability of promoter-enhancer looping relies on the cohesin complex[76,77]. The cohesin complex consists of SMC1, SMC3, RAD21 and SA1 or SA2 proteins[78]. The loading and unloading of cohesin from chromatin are further controlled by the complexes, NIPBL-MAU2 and WAPL-PDS5[79–81]. The formation of loops requires cohesin intrinsic ATPase activity to load and slide along DNA as proposed in a loop extrusion model. The idea of loop extrusion was first described in studies of axolotl lampbrush chromosomes in 1882 by Walther Flemming[82]. Subsequently, the loop extrusion hypothesis was further explored[83–85] and nicely summarized in a recent review[86]. Recent studies tested the loop extrusion model and this was suggested by ChIP-seq data in ATP-depleted cells where RAD21 binding was stronger at NIPBL loading sites while weaker at CTCF anchor sites[26]. Although ATP-independent loops are also formed in the absence of cohesin translocation along DNA, it was observed that cohesin-mediated loop formation tends to favor long-range functional interactions between enhancer and their cognate promoters[26]. Additionally, the loop extrusion model was also shown by cohesin depletion by Schwarzer et al. in a mouse model[87] and through real-time imaging the formation DNA loops by *Saccharomyces cerevisiae* condensin complex further validated the model[88].

In an EMT model of breast cancer cells, reduced binding of the cohesin complex component RAD21 at the loci of *TGFB1* and *ITGA5* leads to a loss of intrachromosomal interactions that allows recruitment of transcriptional machinery to the promoters of the mesenchymal genes[56]. In a SNAI1-inducible EMT model of ovarian cancer cells, RAD21 was found to be selectively depleted from the distal enhancer sites of the epithelial genes *ERBB3* and *PERP*. This selective depletion of RAD21 binding was correlated with the various states along the EMT Spectrum[51]. A recent study further suggested that another cohesin complex component SMC1A could bind to the *SNAI1* promoter to regulate its transcription[89] pointing to a possible feedback regulation between cohesin and the EMT TF machinery. However, it is unclear whether these cohesin complex binding events would have impact on the local DNA looping since there were no chromatin conformation data available. Nevertheless, these studies provided us with a zoomed in glance at the local DNA looping interactions of selected EMT gene loci.

The epigenomic accessibility landscape in EMT models

The epigenomic landscape in EMT models, intertwined with 3D genome architecture, significantly influences chromatin accessibility. The dynamic phase transition along the EMT spectrum mimics the energy gradient which could be segregated according to the epigenetic changes[90] to drive cellular plasticity. Many of the EMT TFs are known to recruit epigenetic modifiers affecting downstream targets[91], with pioneer-factor like TFs[52] demonstrating modulation of EMT plasticity through nucleosome remodeling. These EMT TFs themselves are often epigenetically regulated thus forming a reciprocal feedback control. Through a combined approach of pharmacological and CRISPR interference screens coupled with synthetic genetic tracing of EMT in lung cancer cells, several chromatin remodelers, writers, and readers were found to govern over phenotypic plasticity via functional antagonism[92]. Through exploiting the EMT plasticity facilitated by

Polycomb repressive complex 2 (PRC2) inhibition in lung cancer cells, Serresi et al. uncovered functional antagonism between several chromatin modulators occurring at active chromatin regions during EMT[92]. Pivotal to plasticity, the state of bivalent chromatin regulated by PRC2 is particularly crucial in governing the epigenomic landscape of EMT[93]. Bivalency in chromatin characterization was first coined by Bernstein and colleagues to describe the conflicting combination of active and inactive histone marks observed at same genomic regions[94]. Earlier studies have shown that PRC2 subunits co-orchestrate with EMT TFs during the epigenetic reprogramming of EMT[95,96]. In a breast cancer cell model, PRC2 methyltransferase enhancer of zeste homolog 2 (EZH2) cooperates with SOX4 in TGF-β-induced EMT regulating lysine 27 on histone H3 (H3K27me3) modifications at the promoters of EMT genes[41]. In another *TWIST1*-induced EMT model in immortalized human mammary epithelial cells (HMLE), EZH2 is required for EMT execution, and the maintenance of bivalent chromatin marked by H3K27me3[96]. In this study, switching in between active histone mark (H3K4me3) to the bivalent/repressive mark (H3K27me3) is found at the promoter of EMT genes. Following *TWIST1*-induced EMT, there are 75 genes presenting promoter switching from H3K4me3 to H3K27me3 and 144 genes presenting promoter switching from H3K27me3 to H3K4me3. In another model of ovarian cancer cells, the loss of the epithelial-specific TF *GRHL2* causes the switch of the chromatin states at the promoters of known EMT signature genes[97] during EMT. Approximately 5% of the active chromatin switched into poised/bivalent or PRC2-repressed chromatin at the GRHL2 binding sites[52]. These findings highlight the important role of the bivalent chromatin associated with the PRC2-related H3K27me3 modification during EMT.

What is intriguing is the fact that there are very few overlapping genes with a concordant H3K27me3 modification change following EMT induction between these studies. For example, the pioneer TF of B-cell specification and commitment gene *Ebf1* is the only gene showing concordance in these two studies with the loss of H3K27me3 mark at its promoter following EMT. Another overlapping gene *Lgr6* is upregulated with the loss of H3K27me3 mark in the Tiwari study but is found to harbour the promoter switch from H3K4me3 to H3K27me3 in the Malouf study. Among a selected panel of EMT signature genes showing the switch from the active state to the poised/bivalent or repressed states at the promoter in the Chung study, three genes (*ESRP1*, *GRHL2*, *LAD1*) are concordant with those switching from H3K4me3 to H3K27me3 at the promoter in the Malouf study. A recent study might be able to provide some clues for this discrepancy that these H3K27me3 enriched PRC2 sites are further classified according to different regulatory mechanisms[98]. Since PRC2 preferentially binds to the CpG islands (CGI) at the promoters, this study explores the genome-wide PRC2-occupied CGI (PRC2+-CGI) across 16 TCGA cancer types. Approximately 50% of H3K27me3 enriched PRC2+-CGI promoters in at least one cancer type are hypermethylated with down-regulation of transcription. There are about 35% of PRC2+-CGI promoters which are transcriptionally up-regulated with high chromatin accessibility in at least one cancer type. These upregulated PRC2+-CGI genes (not the hypermethylated GCIs) are enriched in the EMT pathway in many cancer types including basal breast cancer (Basal BRCA) and lung cancers. A small subset of mesenchymal genes (*CDH2*, *DKK1*, *SERPINE2*, *MATN3*, *APLP1*, *CXCL1*, *FOXC2*, *BMP1*) in Basal BRCA has been validated in this PRC2+-CGI pool with *SERPINE2* and *CXCL1* being the 2 most commonly overlapping genes found in 7 and 6 cancer types, respectively. This cancer-specific pattern of regulation at the PRC2 target sites is mediated by the cooperation with specific TFs binding to the distal enhancers to confer tissue specificity. This creates the promoter-enhancer interactions to allow cancer-specific TFs from distal enhancers to affect the chromatin accessibility at proximal promoters to regulate different sets of EMT genes. Findings from these studies reveal a versatile landscape of epigenomic regulation of EMT and the complexity of context-dependent regulation indicates that there might not exist a universal master regulator. This further provides a mechanistic ground as why the definition of EMT cannot rely on one or a small number of molecular markers[5].

Box 1 | Outstanding questions and challenges

- How do we map the diverse trajectory paths of master EMT-TFs occurring on a multi-omics scale – from epigenome, transcriptome, to proteome? EMT is a multistep process orchestrated by a plethora of EMT-TFs[4,5,127]. Evidence of transitional EMT states have been described in in vivo models[7,128] with different modes of regulation undertaken by EMT-TFs. Perturbational single cell omics could serve as a useful tool in interrogation of diverse EMT regulation networks[129].
  However, context is critical as cell-type specific EMT regulation needs to be considered.
  Hitherto, single-cell sequencing studies of EMT in cancer (Table S1) are mainly focused on the transcriptome or epigenome.
  The lack of simultaneous capture of epigenome, protein, and mRNA expression at single-cell depth in these studies may not adequately define the EMT states.
- How does the 3D genome contribute to EMT in a spatiotemporal manner, and to what extent?
  In the context of dynamic and preformed loops, what are the implications that might arise from such chromatin conformation models during the transition from epithelial to mesenchymal state?

Heterogeneity and dynamic genome organization of EMT

A recent study using live-cell imaging to trace single cells in a TGF-β induced EMT model in the lung cancer A549 cells further confirms that the transition dynamics proceeds through two parallel paths[99]. The challenges ahead in the EMT field would lie in the delineation of the different trajectories undertaken by each single cell that are at different phases of the EMT spectrum. Single cell transcriptomic profiling has shown that EMT is not a linear but highly variable process that branches into various trajectories upon induction[9,100]. The complexity is further amplified by the regulatory controls present at multidimensional levels. While single cell transcriptomic profiling has enabled us to look at EMT from an individualistic perspective, the chromatin regulation of the transcriptional activity of EMT genes in concert with TFs is still largely uncharted at a single cell resolution. By using Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) in single cells, germline stem cells (GSCs) have been clustered into epithelial-like and mesenchymal-like states that constitute into an EMT spectrum[101].

With advances in single cell multi-omics sequencing, simultaneous profiling of the transcriptome and epigenome paves the way to bridge the link between epigenetic controls leading to transcriptional events that constitute the heterogeneity of EMT. For instance, several techniques have been developed to capture both mRNA with chromatin accessibility from the same cell[102–104]. This could provide information to correlate between chromatin accessibility and gene expression in EMT models, as well as elucidate the possible TFs that orchestrate the accessibility changes at the EMT genes at different EMT states. Taken into account that EMT-TFs could also regulate their downstream genes through DNA methylation, a technique developed by Gu et al. that captures information on both DNA methylation and accessibility from a single cell would have potential to provide further insights on the regulatory mechanisms of EMT-TFs along the different transitional phases during EMT[105].

To decipher the role of 3D genome in the regulation of EMT at the single cell resolution, technical breakthroughs are crucial to accelerate discovery. The techniques available at the moment for generating single cell 3D genomes are either of low throughput[53] or lack the capability to simultaneously characterize the EMT states of each single cell via the expression of epithelial/mesenchymal markers at the high throughput setting[23,106–109].

In conjunction to scHi-C, computational methods which maximises for the sparseness observed in most scHi-C data are also crucial. Current computational methods to enhance scHi-C datasets include unsupervised embedding[110,111], random-walk-based imputation[112] or hypergraph representation learning of contact maps[113]. Despite the sparsity of scHi-C

contacts, it was shown that at the single-cell level, the chromosomal conformation at *VIM* was able to delineate single-cells by their EMT states, as determined by fluorescence measure of the EMT protein markers[53]. With the advent of more technological breakthroughs, sophisticated seqFISH[114,115] techniques such as DNA and intron seqFISH+ could enable us to elucidate the dynamics of nuclear organization and transcriptomics during EMT through spatial genomics in a temporal fashion, aspects that are largely obscured in current single-cell techniques (Box 1).

Future perspectives

Currently, we discussed the EMT spectrum in the context of transcriptional signature, there are other functional subtypes of the spectrum, including the ameboid state[116]. Although EMT occurs in a context-dependent manner, therein lies core regulatory networks present in both physiological and pathological conditions which are previously summarized in recent reviews[4,117], and the checkpoints at each transitional phase of the EMT spectrum could be governed by different EMT-TFs of the core regulatory networks (EMT-inducers or EMT-gatekeepers)[7,52,118,119]. Therefore, the use of chromatin conformation techniques in future EMT models, such mouse models[7,60] or cell lines[53,67,120], should be considered carefully and as recommended by the latest TEMTIA guidelines[5]. In a comprehensive approach to map the 3D genome changes between transitional phases along the EMT spectrum, HiChIP or ChIA-PET could be used to elucidate the EMT-TFs-centric chromatin interactions. By mapping the diverse EMT-TFs trajectory paths in a spatiotemporal manner, it would be possible to identify epigenomic events which predate the downstream activation of EMT transcriptional programs at various transitional phases of EMT. This could facilitate our comprehensive understanding of how the 3D genome and the EMT-TFs co-orchestrate the plasticity landscape. Given the transitional and continuous nature of the EMT Spectrum[7,97,121–126], it is crucial to delineate the regulatory landscape of subpopulations of different EMT states under different context. The heterogeneity and dynamics of EMT regulatory epigenome needs to be answered via single cell-omics platforms with close collaborations with multidisciplinary computational and mathematical modelling expertise. These efforts will help decipher how the plasticity occurs during EMT under different context to guide the future development of translational applications of the EMT concept in diseases.

References
1. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J. Clin. Investig.* **119**, 1420–1428 (2009).
2. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871–890 (2009).
3. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
4. Nieto, M. A., Huang, RubyY.-J., Jackson, RebeccaA. & Thiery, JeanP. EMT: 2016. *Cell* **166**, 21–45 (2016).
5. Yang, J. et al. Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341–352 (2020).
6. Huang, R. Y.-J., Guilford, P. & Thiery, J. P. Early events in cell adhesion and polarity during epithelial-mesenchymal transition. *J. Cell Sci.* **125**, 4417–4422 (2012).
7. Pastushenko, I. et al. Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463–468 (2018).
8. Meyer-Schaller, N. et al. A Hierarchical Regulatory Landscape during the Multiple Stages of EMT. *Developmental Cell* **48**, 539 (2019).
9. Cook, D. P. & Vanderhyden, B. C. Context specificity of the EMT transcriptional response. *Nat. Commun.* **11**, 2142 (2020).

10. Soldatov, R. et al. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364** https://doi.org/10.1126/science.aas9536 (2019).

11. Takahashi, K. & Yamanaka, S. A developmental framework for induced pluripotency. *Development* **142**, 3274–3285 (2015).

12. Lavin, D. P. & Tiwari, V. K. Unresolved Complexity in the Gene Regulatory Network Underlying EMT. *Front Oncol.* **10**, 554 (2020).

13. Bera, A. & Lewis, S. M. Regulation of Epithelial-to-Mesenchymal Transition by Alternative Translation Initiation Mechanisms and Its Implications for Cancer Metastasis. *Int. J. Mol. Sci.* **21** https://doi.org/10.3390/ijms21114075 (2020).

14. Kang, E., Seo, J., Yoon, H. & Cho, S. The Post-Translational Regulation of Epithelial-Mesenchymal Transition-Inducing Transcription Factors in Cancer Metastasis. *Int. J. Mol. Sci.* **22** https://doi.org/10.3390/ijms22073591 (2021).

15. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing Chromosome Conformation. *Science* **295**, 1306–1311 (2002).

16. Dostie, J. et al. Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).

17. Simonis, M. et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat. Genet.* **38**, 1348 https://www.nature.com/articles/ng1896#supplementary-information (2006).

18. Li, X. et al. Long-read ChIA-PET for base-pair resolution mapping of haplotype-specific chromatin interactions. *Nat. Protoc.* **12**, 899–915 (2017).

19. Lieberman-Aiden, E. et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289–293 (2009).

20. Lupianez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).

21. Tang, Z. et al. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* **163**, 1611–1627 (2015).

22. Dixon, J. et al. *Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions*. Vol. 485 (2012).

23. Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61 https://www.nature.com/articles/nature23001#supplementary-information (2017).

24. Naumova, N., Smith, E. M., Zhan, Y. & Dekker, J. Analysis of long-range chromatin interactions using Chromosome Conformation Capture. *Methods (San Diego, Calif.)* **58**, https://doi.org/10.1016/j.ymeth.2012.1007.1022 (2012).

25. Tavares-Cadete, F., Norouzi, D., Dekker, B., Liu, Y. & Dekker, J. Multi-contact 3C reveals that the human genome during interphase is largely not entangled. *Nat. Struct. Mol. Biol.* **27**, 1105–1114 (2020).

26. Vian, L. et al. The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* **173**, 1165–1178.e1120 (2018).

27. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).

28. Belaghzal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods (San. Diego, Calif.)* **123**, 56–65 (2017).

29. Hsieh, T. et al. Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* **162**, 108–119 (2015).

30. Deshpande, A. S. et al. Identifying synergistic high-order 3D chromatin conformations from genome-scale nanopore concatemer sequencing. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01289-z (2022).

31. Lafontaine, D. L., Yang, L., Dekker, J. & Gibcus, J. H. Hi-C 3.0: Improved Protocol for Genome-Wide Chromosome Conformation Capture. *Curr. Protoc.* **1**, e198 (2021).

32. Akgol Oksuz, B. et al. Systematic evaluation of chromosome conformation capture assays. *Nat. Methods* **18**, 1046–1055 (2021).

33. Li, G. et al. Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) sequencing technology and application. *BMC Genomics* **15**, S11 (2014).

34. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).

35. Ma, W. et al. Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat. methods* **12**, 71 (2015).

36. Martin, P. et al. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat. Commun.* **6**, 10069 https://www.nature.com/articles/ncomms10069#supplementary-information (2015).

37. Mifsud, B. et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598 https://www.nature.com/articles/ng.3286#supplementary-information (2015).

38. Sahlén, P. et al. Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol.* **16**, 156 (2015).

39. Schoenfelder, S. et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* **25**, 582–597 (2015).

40. Hagege, H. et al. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* **2**, 1722–1733 (2007).

41. Zhao, Z. et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347 (2006).

42. van de Werken, H. J. G. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969 https://www.nature.com/articles/nmeth.2173#supplementary-information (2012).

43. Beagrie, R. A. et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017).

44. Quinodoz, S. A. et al. Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell* **174**, 744–757.e724 (2018).

45. Zheng, M. et al. Multiplex chromatin interactions with single-molecule precision. *Nature* **566**, 558–562 (2019).

46. Redolfi, J. et al. DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat. Struct. Mol. Biol.* **26**, 471–480 (2019).

47. You, Q. et al. Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution. *Nat. Biotechnol.* **39**, 225–235 (2021).

48. Kempfer, R. & Pombo, A. Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* **21**, 207–226 (2020).

49. Davies, J. O., Oudelaar, A. M., Higgs, D. R. & Hughes, J. R. How best to identify chromosomal interactions: a comparison of approaches. *Nat. Methods* **14**, 125–134 (2017).

50. Jerkovic´, I. & Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nat. Rev. Mol. Cell Biol.* **22**, 511–528 (2021).

51. Sundararajan, V. et al. SNAI1-Driven Sequential EMT Changes Attributed by Selective Chromatin Enrichment of RAD21 and GRHL2. *Cancers (Basel)* **12** https://doi.org/10.3390/cancers12051140 (2020).

52. Chung, V. Y. et al. The role of GRHL2 and epigenetic remodeling in epithelial–mesenchymal plasticity in ovarian cancer cells. *Commun. Biol.* **2**, 272 (2019).

53. Pang, Q. Y. et al. 3D genome organization in the epithelial-mesenchymal transition spectrum. *Genome Biol.* **23**, 121 (2022).

54. Johnson, K. S. et al. CTCF Expression and Dynamic Motif Accessibility Modulates Epithelial-Mesenchymal Gene Expression. *Cancers (Basel)* **14** https://doi.org/10.3390/cancers14010209 (2022).

55. Werth, M. et al. The transcription factor grainyhead-like 2 regulates the molecular composition of the epithelial apical junctional complex. *Development* **137**, 3835–3845 (2010).

56. Yun, J. et al. Dynamic cohesin-mediated chromatin architecture controls epithelial–mesenchymal plasticity in cancer. *EMBO Rep.* **17**, 1343–1359 (2016).

57. Sun, Y. et al. PHF13 epigenetically activates TGFβ driven epithelial to mesenchymal transition. *Cell Death Dis.* **13**, 487 (2022).

58. Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**, 535–550 (2019).

59. McDonald, O. G., Wu, H., Timp, W., Doi, A. & Feinberg, A. P. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat. Struct. Mol. Biol.* **18**, 867–874 (2011).

60. Pascual-Reguant, L. et al. Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. *Nat. Commun.* **9**, 3420 (2018).

61. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet* **41**, 246–250 (2009).

62. Briand, N. & Collas, P. Lamina-associated domains: peripheral matters and internal affairs. *Genome Biol.* **21**, 85 (2020).

63. Guelen, L. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).

64. Gesson, K. et al. A-type lamins bind both hetero- and euchromatin, the latter being regulated by lamina-associated polypeptide 2 alpha. *Genome Res.* **26**, 462–473 (2016).

65. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).

66. Schmitt, A. D. et al. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep.* **17**, 2042–2059 (2016).

67. Qiao, Y. et al. Enhancer Reprogramming within Pre-existing Topologically Associated Domains Promotes TGF-beta-Induced EMT and Cancer Metastasis. *Mol. Ther.* **28**, 2083–2095 (2020).

68. Chen, H. et al. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* **50**, 1296–1303 (2018).

69. Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F. & de Laat, W. Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. *Mol. Cell* **10**, 1453–1465 (2002).

70. Lettice, L. A. et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).

71. Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning Human Gene Deserts for Long-Range Enhancers. *Science* **302**, 413–413 (2003).

72. Qin, Y. et al. Long-range activation of Sox9 in Odd Sex (Ods) mice. *Hum. Mol. Genet.* **13**, 1213–1218 (2004).

73. Javierre, B. M. et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* **167**, 1369–1384.e1319 (2016).

74. Alotaibi, H. et al. Enhancer cooperativity as a novel mechanism underlying the transcriptional regulation of E-cadherin during mesenchymal to epithelial transition. *Biochimica et. Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1849**, 731–742 (2015).

75. Chopra, V. S., Kong, N. & Levine, M. Transcriptional repression via antilooping in the Drosophila embryo. *Proc. Natl Acad. Sci. USA* **109**, 9460–9464 (2012).

76. Losada, A. Cohesin in cancer: chromosome segregation and beyond. *Nat. Rev. Cancer* **14**, 389–393 (2014).

77. O'Neil, N. J., van Pel, D. M. & Hieter, P. Synthetic lethality and cancer: cohesin and PARP at the replication fork. *Trends Genet.* **29**, 290–297 (2013).

78. Kim, Y. & Yu, H. Shaping of the 3D genome by the ATPase machine cohesin. *Exp. Mol. Med.* **52**, 1891–1897 (2020).

79. Bermudez, V. et al. In vitro loading of human cohesin on DNA by the human Scc2-Scc4 loader complex. *Proc. Natl Acad. Sci.* **109**, 9366–9371 (2012).

80. Arumugam, P. et al. ATP Hydrolysis Is Required for Cohesin's Association with Chromosomes. *Curr. Biol.* **13**, 1941–1953 (2003).

81. Watrin, E. et al. Human Scc4 Is Required for Cohesin Binding to Chromatin, Sister-Chromatid Cohesion, and Mitotic Progression. *Curr. Biol.* **16**, 863–874 (2006).

82. Flemming, W. *Zellsubstanz, kern und zelltheilung*. (Vogel, 1882).

83. Nasmyth, K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev. Genet.* **35**, 673–745 (2001).

84. Alipour, E. & Marko, J. F. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res* **40**, 11202–11212 (2012).

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

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

87. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).

88. Ganji, M. et al. Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).

89. Zhang, X. et al. SMC1A regulated by KIAA1429 in m6A-independent manner promotes EMT progress in breast cancer. *Mol. Ther. Nucleic Acids* **27**, 133–146 (2022).

90. Tam, W. L. & Weinberg, R. A. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.* **19**, 1438 (2013).

91. Skrypek, N., Goossens, S., De Smedt, E., Vandamme, N. & Berx, G. Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity. *Trends Genet* **33**, 943–959 (2017).

92. Serresi, M. et al. Functional antagonism of chromatin modulators regulates epithelial-mesenchymal transition. *Sci. Adv.* **7**, eabd7974 (2021).

93. Chaffer, C. L. et al. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74 (2013).

94. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).

95. Tiwari, N. et al. Sox4 Is a Master Regulator of Epithelial-Mesenchymal Transition by Controlling Ezh2 Expression and Epigenetic Reprogramming. *Cancer Cell* **23**, 768–783 (2013).

96. Malouf, G. G. et al. Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. *Genome Biol.* **14** https://doi.org/10.1186/gb-2013-14-12-r144 (2013).

97. Tan, T. Z. et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol. Med.* **6**, 1279–1293 (2014).

98. Zhang, Q. et al. PALI1 facilitates DNA and nucleosome binding by PRC2 and triggers an allosteric activation of catalysis. *Nat. Commun.* **12** https://doi.org/10.1038/s41467-021-24866-3 (2021).

99. Wang, W., Poe, D., Yang, Y., Hyatt, T. & Xing, J. Epithelial-to-mesenchymal transition proceeds through directional destabilization of multidimensional attractor. *Elife* **11** https://doi.org/10.7554/eLife.74866 (2022).

100. McFaline-Figueroa, J. L. et al. A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. *Nat. Genet* **51**, 1389–1398 (2019).

101. Liao, J. et al. scATAC-Seq reveals epigenetic heterogeneity associated with an EMT-like process in male germline stem cells and its regulation by G9a. 2020.2010.2012.336834 https://doi.org/10.1101/2020.10.12.336834 (2021).

102. Ma, S. et al. Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103–1116.e1120 (2020).

103. Cao, J. et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380 (2018).

104. Chen, S., Lake, B. B. & Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat. Biotechnol.* **37**, 1452–1457 (2019).

105. Gu, C., Liu, S., Wu, Q., Zhang, L. & Guo, F. Integrative single-cell analysis of transcriptome, DNA methylome and chromatin accessibility in mouse oocytes. *Cell Res.* **29**, 110–123 (2019).

106. Nagano, T. et al. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat. Protoc.* **10**, 1986–2003 (2015).

107. Stevens, T. J. et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017).

108. Ramani, V. et al. Sci-Hi-C: A single-cell Hi-C method for mapping 3D genome organization in large number of single cells. *Methods* **170**, 61–68 (2020).

109. Tan, L., Xing, D., Chang, C. H., Li, H. & Xie, X. S. Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018).

110. Liu, J., Lin, D., Yardimci, G. G. & Noble, W. S. Unsupervised embedding of single-cell Hi-C data. *Bioinforma. (Oxf., Engl.)* **34**, i96–i104 (2018).

111. Kim, H.-J. et al. Capturing cell type-specific chromatin compartment patterns by applying topic modeling to single-cell Hi-C data. *PLOS Comput. Biol.* **16**, e1008173 (2020).

112. Zhou, J. et al. Robust single-cell Hi-C clustering by convolution- and random-walk–based imputation. *Proc. Natl Acad. Sci.* **116**, 14011–14018 (2019).

113. Zhang, R., Zhou, T. & Ma, J. Multiscale and integrative single-cell Hi-C analysis with Higashi. *Nat. Biotechnol.* **40**, 254–261 (2022).

114. Shah, S. et al. Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. *Cell* **174**, 363–376.e316 (2018).

115. Takei, Y. et al. High-resolution spatial multi-omics reveals cell-type specific nuclear compartments. 2023.2005.2007.539762 https://doi.org/10.1101/2023.05.07.539762 (2023).

116. Graziani, V., Rodriguez-Hernandez, I., Maiques, O. & Sanz-Moreno, V. The amoeboid state as part of the epithelial-to-mesenchymal transition programme. *Trends Cell Biol.* **32**, 228–242 (2022).

117. Chakraborty, P., George, J. T., Tripathi, S., Levine, H. & Jolly, M. K. Comparative Study of Transcriptomics-Based Scoring Metrics for the Epithelial-Hybrid-Mesenchymal Spectrum. *Front. Bioeng. Biotechnol.* **8** https://doi.org/10.3389/fbioe.2020.00220 (2020).

118. Vasaikar, S. V. et al. EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br. J. Cancer* **124**, 259–269 (2021).

119. Sundararajan, V. et al. SNAI1 recruits HDAC1 to suppress SNAI2 transcription during epithelial to mesenchymal transition. *Sci. Rep.* **9**, 8295 (2019).

120. Deshmukh, A. P. et al. Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. *Proc. Natl. Acad. Sci. USA* **118** https://doi.org/10.1073/pnas.2102050118 (2021).

121. Huang, R. Y. J. et al. An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Amp* **4**, e915 https://www.nature.com/articles/cddis2013442#supplementary-information (2013).

122. Cook, D. P. & Vanderhyden, B. C. Transcriptional census of epithelial-mesenchymal plasticity in cancer. *Sci. Adv.* **8**, eabi7640 (2022).

123. Wang, H. et al. Single-Cell Analyses Reveal Mechanisms of Cancer Stem Cell Maintenance and Epithelial–Mesenchymal Transition in Recurrent Bladder Cancer. *Clin. Cancer Res.* **27**, 6265–6278 (2022).

124. LaFave, L. M. et al. Epigenomic State Transitions Characterize Tumor Progression in Mouse Lung Adenocarcinoma. *Cancer Cell* **38**, 212–228.e213 (2020).

125. Sha, Y., Wang, S., Bocci, F., Zhou, P. & Nie, Q. Inference of Intercellular Communications and Multilayer Gene-Regulations of Epithelial-Mesenchymal Transition From Single-Cell Transcriptomic Data. *Front Genet* **11**, 604585 (2020).

126. Puram, S. V. et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611–1624.e1624 (2017).

127. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69–84 (2019).

128. Aiello, N. M. et al. EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Developmental Cell* **45**, 681–695.e684 (2018).

129. Ji, Y., Lotfollahi, M., Wolf, F. A. & Theis, F. J. Machine learning for perturbational single-cell omics. *Cell Syst.* **12**, 522–537 (2021).

Acknowledgements

Author contributions

R.Y.H., P.Q.Y, and Y.C.C. contributed to the conceptualization and writing of the manuscript. Funding acquisition and administration of the project was provided by R.Y.H. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no conflict of interest. R.Y.H. is an Editorial Board Member for Communications Biology, but was not involved in the editorial review of, nor the decision to publish this article.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-06441-w.

Correspondence and requests for materials should be addressed to Ruby Yun-Ju Huang.

Peer review information *Communications Biology* thanks David Cook and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Eve Rogers and Christina Karlsson Rosenthal.

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

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