

RESEARCH ARTICLE

Deciphering the molecular mechanism of the cancer formation by chromosome structural dynamics

Xiakun Chu¹, Jin Wang^{1,2*}

1 Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York, United States of America, **2** Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York, United States of America

* jin.wang.1@stonybrook.edu

Abstract

Cancer reflects the dysregulation of the underlying gene network, which is strongly related to the 3D genome organization. Numerous efforts have been spent on experimental characterizations of the structural alterations in cancer genomes. However, there is still a lack of genomic structural-level understanding of the temporal dynamics for cancer initiation and progression. Here, we use a landscape-switching model to investigate the chromosome structural transition during the cancerization and reversion processes. We find that the chromosome undergoes a non-monotonic structural shape-changing pathway with initial expansion followed by compaction during both of these processes. Furthermore, our analysis reveals that the chromosome with a more expanding structure than those at both the normal and cancer cell during cancerization exhibits a sparse contact pattern, which shows significant structural similarity to the one at the embryonic stem cell in many aspects, including the trend of contact probability declining with the genomic distance, the global structural shape geometry and the spatial distribution of loci on the chromosome. In light of the intimate structure-function relationship at the chromosomal level, we further describe the cell state transition processes by the chromosome structural changes, suggesting an elevated cell stemness during the formation of the cancer cells. We show that cell cancerization and reversion are highly irreversible processes in terms of the chromosome structural transition pathways, spatial repositioning of chromosomal loci and hysteresis loop of contact evolution analysis. Our model draws a molecular-scale picture of cell cancerization from the chromosome structural perspective. The process contains initial reprogramming towards the stem cell followed by the differentiation towards the cancer cell, accompanied by an initial increase and subsequent decrease of the cell stemness.

OPEN ACCESS

Citation: Chu X, Wang J (2021) Deciphering the molecular mechanism of the cancer formation by chromosome structural dynamics. *PLoS Comput Biol* 17(11): e1009596. <https://doi.org/10.1371/journal.pcbi.1009596>

Editor: Yaakov Koby Levy, Weizmann Institute of Science, ISRAEL

Received: July 2, 2021

Accepted: October 28, 2021

Published: November 9, 2021

Copyright: © 2021 Chu, Wang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The necessary files for setting up Gromacs (version 4.5.7 with PLUMED version 2.5.0) simulations and analysis programs/scripts are publicly available at <https://osf.io/cd9fk/>.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Cancer is among the leading causes of human death. Cancer is regulated by the underlying regulatory network of gene expressions, which are in intimate relation to the 3D chromosome architectures. Numerous efforts have been spent on elucidating the chromosome

structural variants in tumorigenesis, while the dynamical picture of how chromosomes structurally evolve during cancer formation is still missing. Here we integrate the Hi-C data into the polymer simulations to build the chromosome structural ensembles in the normal and cancer cells. Then we use a nonequilibrium landscape-switching model to simulate the chromosome structural dynamics during the cancerization and reversion processes. With quantified pathways, we show that the chromosomes at transient intermediate states in the cancer formation possess a significant degree of structural similarity to those at the stem cell. Our findings indicate the formation of stem-like states during cancer formation from the chromosome structural perspective. We draw a molecular-scale picture of the cancer formation, which contains initial reprogramming towards the stem cell followed by differentiation towards the cancer cell.

Introduction

Currently, our knowledge of how cancer cells form and proliferate is still quite limited. In general, the development of cancer is controlled by the underlying gene regulatory network [1], which relies on the molecular interactions between the spatially organized genome and a broad class of proteins, including the transcription factors and chromatin remodelers [2]. As the structural scaffold for the genome function, the 3D genome architecture has been recognized to play an important role in regulating the gene expression [3–5], leading to the intimate structure-function relationships at the genomic level. Several studies have shown that alterations in chromosome structures and interactions can make significant contributions to the dysregulation of the gene expressions forming the specific cancer signatures [6–9]. These findings have fostered a view that the 3D context of the genome is the major player in the development and progression of cancer [10]. Therefore, investigating the disorganization of the 3D genome structure in cancer cells can provide the key to understanding the pathogenesis of cancer.

Although the chromosome structural variant, which is the major form of the genome instability, has been regarded as a hallmark of almost all human cancers [11–13], determining the genome structure has been a long-term challenge until the emergence of chromosome conformation capture (3C) techniques nearly two decades ago [14]. As an advanced derivative of the 3C, Hi-C measures the spatial proximity of the chromosomal loci across the entire genome in terms of the contact frequency map and often provides an ensemble description of the genome organization within a large number of cells [15, 16]. Recently, Hi-C techniques were applied to characterize the disorganization of the cancer genomes and determine their functional consequences [17]. Taberlay et al. observed the smaller-sized topologically associated domains (TADs) formed in prostate cancer cells than those in normal cells due to the emergence of additional domain boundaries. This further leads to alterations of TP53 tumor suppressor locus [18]. In addition, the effects of the structural disruptions within TADs on leading to various cancer types were established [19–21]. At the higher hierarchical level, Barutcu et al. found that the significant A/B compartment switching between the normal and breast cancer cell is associated with gene expression changes [22]. These studies have shown that the structural variants in the cancer genome occur throughout the genome sequence, and they have important impacts on the gene expression alterations for inducing cancer formation.

Increasing evidence indicates that cancer development is orchestrated by a small subpopulation of the cancer cells, namely the cancer stem (CS) cells. [23–26]. CS cells, which possess stem-like properties and functions, are capable of performing self-renewal, proliferation and

differentiation. Thus they provide the driving forces for cancer progression [27, 28]. However, experimental identifications on the CS cell and the associated dynamics are non-trivial due to its rare population in the cancer cells [29]. The previous theoretical studies using gene regulatory work, which includes the cancer and developmental genetic markers as well as the interactions between them, provided a landscape view of how the cancer cell develops and CS cell forms [30, 31]. However, the work focused on simplified core gene regulatory networks to describe the cancer systems, so the results were restricted in the form of regulation pathways among several marker genes and this description is at the gene network level. Therefore, there is still a lack of a complete molecular chromosomal-structural level understanding of cancer development and CS cell formation. On the other hand, the strategies on inducing the cancer cell to normal cell by reversion [32] or transdifferentiation processes [33–35] are rapidly developing. These approaches have opened new avenues for cancer treatments, but the underlying mechanism of the cancer reversion is still unclear [36].

Despite the significant achievements made by the Hi-C technique on elucidating the spatial genome disorganization in the cancer cells [17], the temporal dynamic rearrangement of the genome structure during cancer development is still not available. The picture is fundamentally important as it may provide a molecular-level understanding of the cancer mechanism. Nevertheless, measuring the chromosome structural dynamics during the cancer cell developmental process is extremely challenging due to the spatial and temporal resolution limits of the current experiments. Owing to the rapid development of the 4D Nucleome program [37], in particular the time-course Hi-C technique, it has been possible to characterize the spatiotemporal organizations of the genome during different cellular processes [38–40]. Recently, the data probed at the discrete time-points along the biological processes in the experiments were used to model the structures of the chromosomes, and further implemented to construct the gradual and smooth trajectories through simulations [41, 42]. These data-driven approaches have provided useful information to interpret the experiments and moved further to characterize the relationship between the structure and function of the chromosome. However, the time-course Hi-C experiments are often laborious and expensive, and the measurements do not provide the real-time dynamics but in the form of ensemble-based data, which inevitably contain the cell state heterogeneity, thus impeding our understanding of the intermediate states during the transition [43, 44].

Here, we developed molecular dynamics models and performed associated simulations to investigate the dynamic cancer-related chromosome structural evolution, which provides the microscopic description of the cancer process. First, we applied the maximum entropy principle to incorporate the Hi-C data of the normal and cancer cells into two independent sets of simulations. These simulations generated two potentials for describing the chromosome structural ensembles that reproduced the Hi-C data, in the normal and cancer cells, respectively (S1 and S2 Figs) [45]. In the cell nucleus, the chromosome constantly experiences the nonequilibrium effects even at one cell state, leading to a nonequilibrium system. Theoretically, it has been demonstrated that the dynamics of the nonequilibrium system can be described using the concept of an effective landscape in some circumstances [46, 47]. Recent studies revealed that under an effective equilibrium landscape, the chromosome dynamics reproduces many aspects of kinetic behaviors as observed in experiments, including anomalous diffusion, viscoelasticity, and spatially coherent dynamics [48, 49]. Here, to examine whether the potentials accounting for chromosome structures obtained from maximum entropy principle simulations can also be used for describing the chromosome dynamics, we performed additional simulations under these two potentials and calculated the diffusion behaviors of the chromosome motions in the normal and cancer cells, respectively (S3 Fig). We observed sub-diffusivity of chromosome dynamics with the scaling exponents of the mean square displacement in good

agreement with multiple experiments [50, 51]. Our results have suggested that these two potentials generated by the maximum entropy principle simulations for the chromosome in the normal and cancer cells can be further regarded as the effective energy landscapes that govern the chromosome conformational dynamics within the normal and cancer cells, respectively. Although each state can be described by an effective equilibrium landscape, the inter-basin dynamics of switching between the normal and cancer states often requires a significant amount of energy input [52] and therefore is a nonequilibrium process, which cannot be treated in an equilibrium way. This is the motivation and target of our current study. In this regard, the cancer-related chromosome structural transitions are further described by the connection between these two effective landscapes during the cell cancerization and reversion processes.

To bridge these two landscapes for describing transitions between the normal and cancer cells, we then used a landscape-switching model, which was developed to simulate the chromosome structural transition during the cell cycle [53] and the cell developmental processes [45]. Briefly, the chromosome structural dynamics at either normal or cancer cell was initially described by an effective equilibrium landscape generated by the data-driven maximum entropy principle simulation. Then, the cell state transition process was triggered by an instantaneous energy excitation that switches the system from one landscape to the other. This energy excitation implementation provides the extra energy for the inter-landscape switching, thus it breaks the detailed balance of the system, leading to the nonequilibrium process. Finally, the relaxation dynamics of the chromosome structure occur on the post-switching effective equilibrium landscape (see “Materials and methods”). The rationales for approximating the transitions between the normal and cancer cells to simple switches are based on the following two facts.

1. The cancer process often exhibits switch-like behavior between two steady cell states, in accordance with the landscape-switching model. The cell state transition system, including cancer, usually shows bistability at the initial and final cell states [54, 55]. The normal and cancer cells, deemed as the attractors on the cell developmental landscape, dictate the transition processes [31, 56, 57]. Increasing experimental evidence suggested that the cell state transitions likely undergo switching between the two stable cell states [58–61]. To accommodate these features, the model used a landscape-switching implementation to trigger cell cancerization and reversion. For practical correspondence, the switch mimics the roles of the genetic mutations and epigenetic modifications in initiating cancer processes [62, 63]. Although the individual genetic mutations and epigenetic modifications during the cancer processes may occur frequently, the changes of the cell phenotypes can only be steadily realized by the cooperative changes of the chromatin states at a global scale, leading to abrupt switches between the two distinct cell states [64, 65].

2. From the physical perspective, cell cancerization and reversion can be approximately described by the nonequilibrium nonadiabatic processes, quantified using the landscape-switching model. It has been recognized that the degree of the adiabaticity is determined by the interplay of the timescales for intra- and inter-landscape dynamics [66, 67]. A faster (slower) intra-landscape motion than inter-landscape hopping leads to a nonadiabatic (adiabatic) process. In reality, the normal and cancer cell states are stable and the transitions between them are impossible to occur spontaneously. In contrast, numerous internal and external factors associated with a significant amount of energy input work collaboratively to achieve the transition processes between these two cell states with distinct phenotypes [68]. These features lead to a significantly slower timescale for the cell waiting for the state transition (inter-landscape dynamics) than the cell relaxing within one stable state (intra-landscape dynamics). In analogous to the “surface hopping” method [69], we separated the simulations

of the chromosome dynamics at one cell state and the nonequilibrium nonadiabatic inter-state switching dynamics, giving rise to the landscape-switching model.

We used the landscape-switching model to study the chromosome structural dynamics during the transitions between the human normal and cancer lung cells. We predicted that the chromosome can form more expanding structures than those at the normal and cancer cells during both of these processes but with different structural characteristics. Further analyses revealed that the chromosome at the transient intermediate state with a more expanding structure than those at the normal and cancer cells during the cancerization process share significant structural similarity to the one at the embryonic stem (ES) cell. This feature implies forming a cell state with the features of stemness or the CS cell during the cancerization process from the chromosome structural perspective. We observed the high irreversibility for cancerization and reversion through the quantified chromosome structural transition pathways, the spatial repositioning of chromosomal loci and the hysteresis loop of establishing the chromosomal contacts. These findings underlined distinct mechanisms for these two cancer cellular processes. The prediction of forming a stem-like intermediate cell state at the chromosomal-structural level during cancerization can be tested by the future time-course Hi-C experiments designed for the cancer developmental processes.

Results

Chromosome structural transitions during cancerization and reversion

We used the landscape-switching model to investigate the chromosome dynamics during the cell cancerization and reversion processes. The chromosome used in this study is the long arm of human chromosome 14 (20.5–106.1 Mb). We focused on the terminally differentiated human lung fibroblast cell (IMR90) and lung cancer cell (A549), and investigated the chromosome structural changes during the transitions between these two cell states. We note that there are no chromosome abnormalities, such as the chromosome deletion, duplication, inversion, substitution and translocation, occurring on the chromosome 14 in the A549 cell [17, 70]. This feature indicates that the structural differences of chromosome 14 between the IMR90 and A549 cells are all related to the intra-chromosome structural rearrangements and can be studied by molecular dynamics simulations.

The implementation of the landscape-switching model in molecular dynamics simulation is briefly summarized as follows. First, we iteratively fit a generic polymer model to reproduce the experimental Hi-C data through the maximum entropy principle simulation for the IMR90 and A549 cell, separately [71]. Previous studies have shown that the resulting potential not only captures the thermodynamics of the chromosome (i. e., Hi-C) [72–74], but also describes the correct kinetic properties of the chromosomal loci diffusion within one cell state (one phase in cell cycle or one cell state in cell differentiation/reprogramming) [48, 49]. Here, we termed the potential $V(\mathbf{r}|S)$ as the effective energy landscape for describing the chromosome dynamics in the normal or cancer cell, where \mathbf{r} is the coordinate of the system at the cell state S (IMR90 or A549). Next, the simulation was set up with the chromosome exploring the structural dynamics under either the energy landscape of the normal ($V(\mathbf{r}|IMR90)$) or cancer ($V(\mathbf{r}|A549)$) cell. Then, the energy landscape underwent a switch from normal to cancer cell ($V(\mathbf{r}|IMR90) \rightarrow V(\mathbf{r}|A549)$) or cancer to normal cell ($V(\mathbf{r}|A549) \rightarrow V(\mathbf{r}|IMR90)$) to trigger the chromosome structural transition during cancerization or reversion, respectively. Finally, the chromosome dynamics was governed by the post-switching energy landscape of the cancer cell ($V(\mathbf{r}|A549)$) or normal cell ($V(\mathbf{r}|IMR90)$). The model allowed us to observe the chromosome transformation during the cancerization and reversion processes with affordable computational resources (Fig 1).

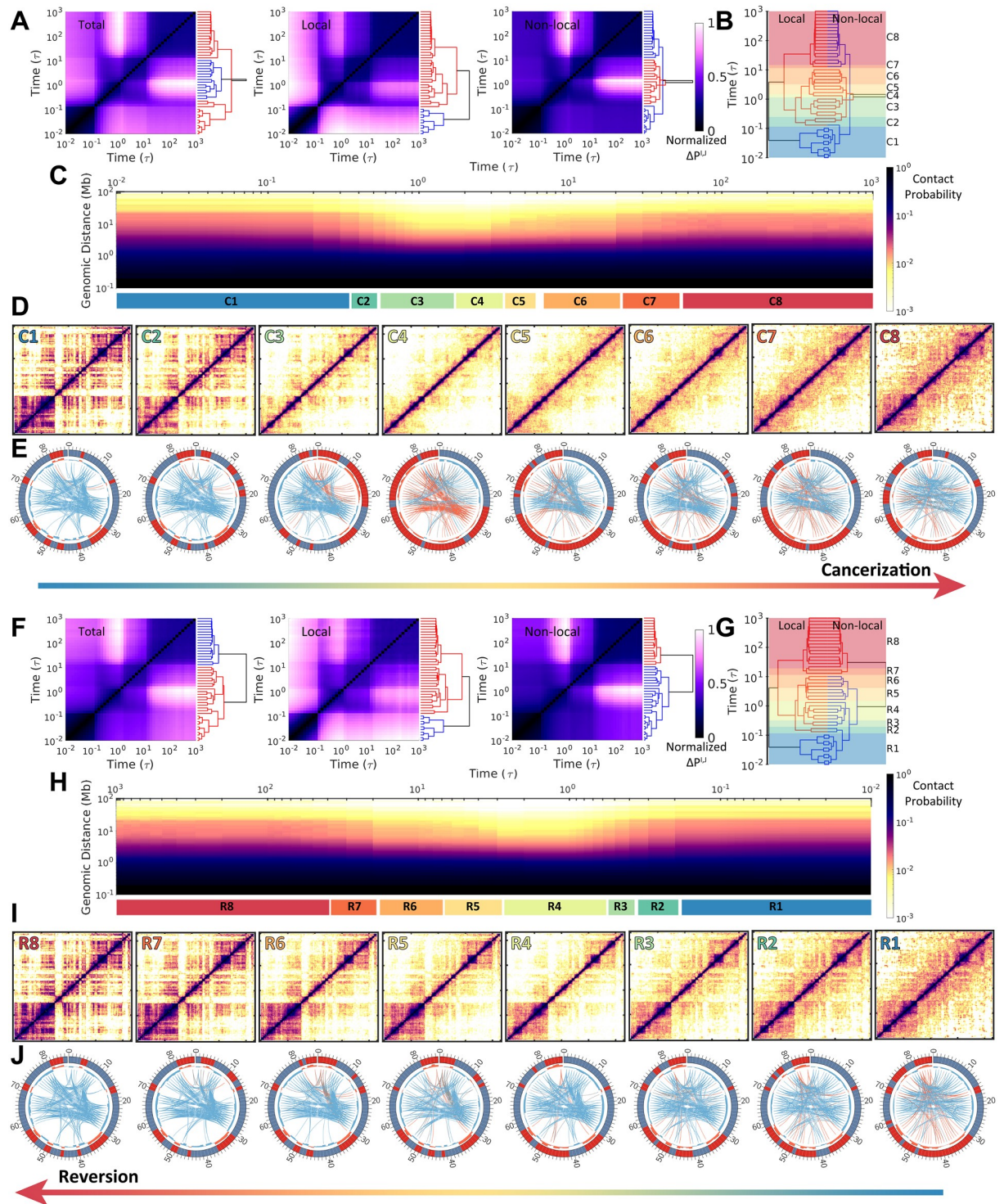


Fig 1. Chromosome structural transitions during cancerization of IMR90 (normal cell) to A549 (cancer cell) (Upper) and reversion of A549 to IMR90 (Lower). (A) Hierarchical clustering of the chromosomal contact probability among each time frame $t = I, J$ during cancerization varied by total, local ($< 2\text{Mb}$), and non-local ($\geq 2\text{Mb}$) contact ranges. The clustering was performed on the contact map difference $\Delta P^{I,J}$ (see “Materials and Models”). (B) Reduced 8 stages (“C1”-“C8”) for the cancerization process based on the combination and comparison of the dendrograms of local and non-local $\Delta P^{I,J}$ established in (A). Thus the chromosomes within one stage possess relatively similar contact probability maps. (C) The change of contact probability $P(l)$ versus genomic distance l in the chromosome during cancerization with the 8 stages indicated at the bottom. (D) Hi-C heat (contact probability) maps of the chromosome for the 8 stages during cancerization. (E) The circle plots of chromosomes for the 8 stages during cancerization [75]. The red and blue bands indicate the loci in compartments A and

B, respectively. The compartment profiles are further shown as histograms near to the band plots. The connections show the long-range (> 5Mb) interactions which are identified by the P_{obs}/P_{exp} , where P_{obs} and P_{exp} are observed and expected contact probability, respectively [15]. The red, blue, and grey lines indicate the interactions between chromosomal loci within compartment A, within compartment B, and between compartment A and B. The line widths correspond to the logarithmic scale of the P_{obs}/P_{exp} and only the top 200 weighted contacts are shown for better visualization. (F-J) are similar with (A-E) but for the reversion process of A549 to IMR90. Another reduced 8 stages during the reversion (“R1”–“R8”) are determined in (G).

<https://doi.org/10.1371/journal.pcbi.1009596.g001>

We performed hundreds of independent landscape-switching simulations starting from different chromosome structures in the normal (and cancer) cell state to investigate the chromosome structural transitions during the cancerization (and reversion) process (See “[Materials and methods](#)”). By comparing the differences of the contact probability formed through the pairwise chromosomal loci between time-series frames during the cancerization process (denoted as $\Delta P^{I,J}$, where P is contact probability and I and J are the time points), the transition can be reduced into two clusters according to the hierarchical clustering of $\Delta P^{I,J}$ along with the processing time (Fig 1A). However, the clustering dendrograms for the local and non-local chromosomal contacts are very different, indicating distinct behaviors of structural rearrangements at the local and non-local ranges. The separation of the two clusters at the local contacts occurs early at $\sim 0.1\tau$, where τ is the time unit in the simulation. In contrast, the non-local chromosomal contacts at the beginning of the transition are relatively similar to those at the late stages, and they are separated by the second cluster at $\sim 0.5-10\tau$. This is also reflected by the contact probability $P(l)$ versus the genomic distance l between a pair of chromosomal loci (Fig 1C), where an apparent decrease of $P(l)$ is observed at time $\sim 0.5-10\tau$.

By examining the clustering dendrogram patterns of the local and non-local contact formations in the chromosome, we further reduced the cancerization process into 8 stages (Fig 1B). As a result, chromosomes within one stage are structurally similar in terms of the local and non-local contacts. The contact maps of the 8 stages show a picture of how chromosomes perform the structural rearrangements in the cancer formation process (Fig 1D). Interestingly, we found that as the cancerization proceeds, the chromosome heat maps show lower probabilities (more sparse contacts in heat maps) than those in both normal and cancer cells. The result implies that the chromosome structure is expanding during cancerization. Besides, we observed a complex manner of the chromosome in organizing the compartment change during cancerization shown by the colored bands in the circle plots (Fig 1E). From the stage “C1” to “C3”, the populations of the chromosomal loci in compartment A increase associated with an increasing quantity of contacts (lines in each circle plot) formed within compartment A. At the stage “C4”, the chromosomal loci are roughly assigned to compartment A and B by sequence and the long-range contacts are rare. This implies the extensive breaking of the contacts formed in the chromosome at the normal cell. From the stage “C5” to “C8”, the chromosomal compartment profiles and contact patterns gradually adapt to those at the cancer cell. Finally, we see that the chromosome in the cancer cell has more chromosomal loci in compartment A and forms more sparse contacts than the one in the normal cell. In the ES cells, chromosomes possess high populations of euchromatin, which maps with compartment A, to benefit the cell pluripotency and differentiation [76–78]. Our observation of the chromosome with populated expressive loci and open structure in the cancer cell resonates with the experimental evidence that the cancer cells possess certain contents of the “stemness” in favor of the self-renewal, differentiation, and proliferation [23, 79, 80].

Then we applied similar analyses into the reversion process. We found that the chromosome structural transition during the reversion process is not the simple reversal of the cancerization. From the chromosomal contact formation perspective, the reversion process can also be grouped into two clusters (Fig 1F). However, these two clusters are separated by the time in

sequence from the dendrogram plots. In detail, the time for separating the two clusters at local contact ($\sim 0.1\tau$) occurs much earlier than that at non-local contacts ($\sim 10\tau$). This indicates that chromosome structurally evolves faster at local than non-local range during reversion. Interestingly, the time evolution of $P(l) \sim l$ also shows a chromosome structural expansion at the time $\sim 0.5-10\tau$ (Fig 1H). However, the expansion appears to be less significant than that observed during cancerization. The chromosomal contact maps show distinct evolving ways deviated from the reversal of that in cancerization (Fig 1I). In particular, there is no significant compartment-switching during the reversion process, as the compartment undergoes gradual and moderate change, associated with reducing the contacts formed within compartment A (Fig 1J). Overall, we found that the chromosome structural transitions undergo initial expansions followed by compactations during both the cancerization and the reversion processes but through irreversible paths.

Chromosome structural expansions during cancerization and reversion

To examine the chromosome structural expansion during cancerization and reversion, we monitored the contact probability $P(l)$ versus the genomic distance l . $P(l)$ provides the information on how the chromosome organizes its structure along the genomic sequence distance, thus it dictates the polymer state of the chromosome [81, 82]. Chromosomes in both the normal and cancer cells show similar $P(l)$ trends with higher curves, corresponding to more compacted structures, than the ones in the ES cell [78] (Fig 2A). This can also be inferred by the slope in the logarithmic relation of $P(l) \sim l$ (Fig 2B). As the normal cell converts to the cancer cell, the profile of $P(l)$ (colored by time) initially goes down and approaches the one at the ES cell (blue line), then increases to the cancer cell (purple line). The result draws a picture of the chromosome expanding its structure followed by compaction during cancerization. To quantitatively measure this trend, we calculated the slope in the logarithmic relation of $P(l) \sim l$. We see that the slope at $\sim 2\tau$ has the lowest value and approximates to the value found at the ES cell. This implies that the chromosome with the maximum structural expansion possesses a similar contact formation scaling versus the genomic sequence distance with the one at the ES cell. In other words, the chromosome may form the chromosome structural characteristics of the ES cell during cancerization. Besides, it is worth noting that the chromosome in the cancer cell is more expanding than it is in the normal cell, as the $P(l)$ plot for the cancer cell is slightly below the one for the normal cell.

To measure the similarity of the chromosomal contact formation with that in the ES cell during cancerization, we calculated the difference of the contact probability between the processing state at time point t during cancerization and the normal, cancer and ES cells at the local and non-local ranges (denoted as $P^{t,S}$ in Fig 2C and 2D, see “Materials and methods”). We found that the chromosome at $\sim 0.5-10\tau$, when the chromosome appears to be more expanding than those at the normal and cancer cells, as indicated in Fig 2B, $P^{t,S}$ of the normal and cancer increase significantly while $P^{t,S}$ of the ES cell decreases. This implies that the contact formation in the chromosome with a more expanding structure than those at the normal and cancer cells, becomes similar to the one at the ES cell. The results indicate that the chromosome may go through stem-like structures to accomplish the cancerization process.

In contrast, the chromosome expansion in the reversion process from the cancer to normal cell is not significant (Fig 2E). When the chromosome has the most significant structural expansion with the lowest $P(l)$, the profile of $P(l)$ still visually deviates from the one in the ES cell. The steepest slope in $P(l) \sim l$ shows a slight decrease from that in the cancer cell and occurs at the time $\sim 0.5-3\tau$. Besides, we can still observe that the similarity of the chromosomal contact formation between the processing state and the ES cell increases, but it is not as

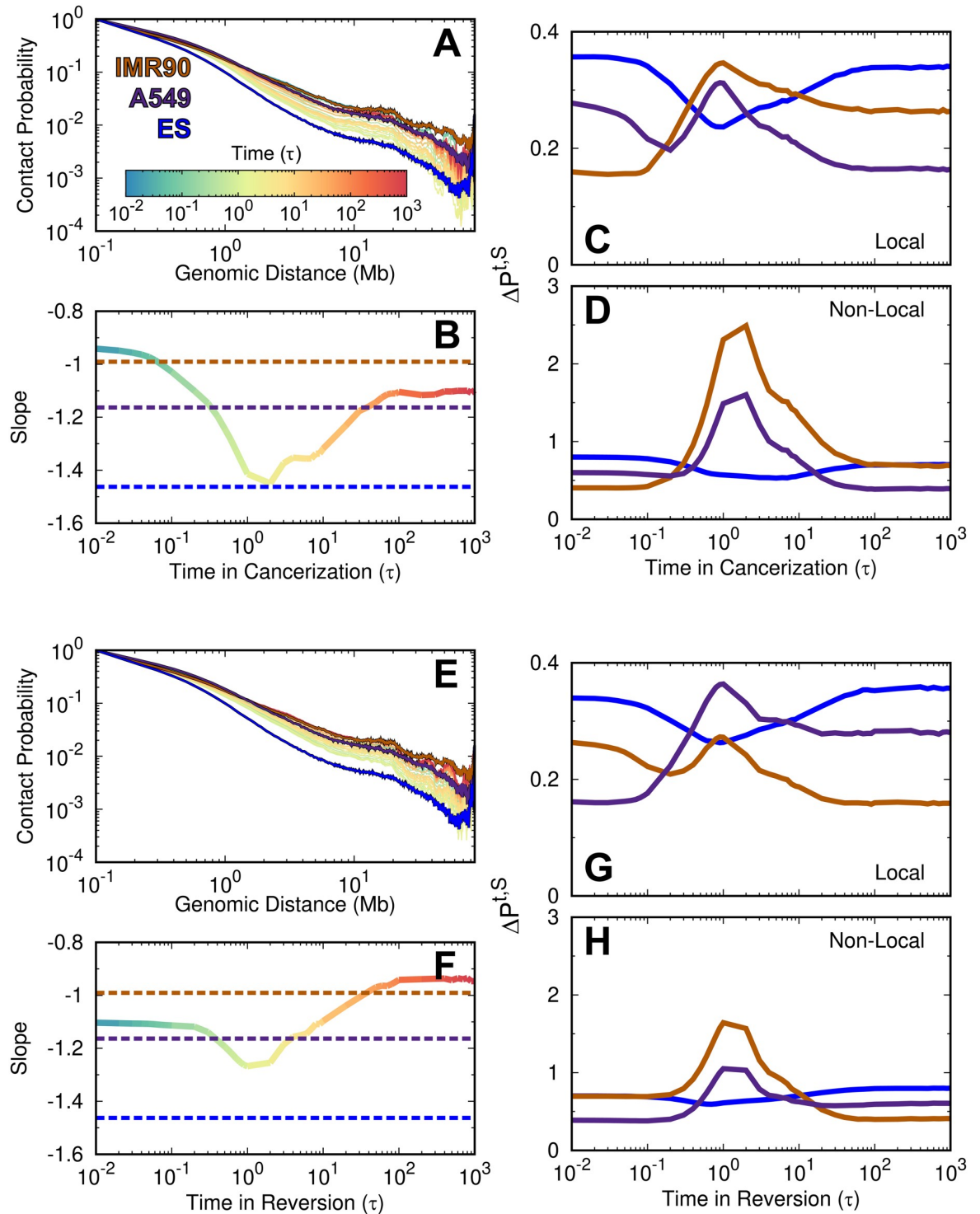


Fig 2. The time evolution of the chromosomal contact probability during cell cancerization and reversion. (A) The time evolution of contact probability $P(l)$ versus genomic distance l during cancerization. The Hi-C data of the IMR90, A549, and ES cells are also plotted. (B) The time evolution of the slope in the logarithmic relation of $P(l) \sim l$ at the range of 0.5–7 Mb during cancerization. The differences of contact probability map during the cancerization process to the Hi-C data of IMR90, A549 and ES cells at the (C) local and (D) non-local ranges. (E–H) are the same with (A–D) but for the reversion process.

<https://doi.org/10.1371/journal.pcbi.1009596.g002>

much as that observed during cancerization (Fig 2G and 2H). Therefore, the chromosome during the reversion process may possess less structural properties at the ES cell than during cancerization.

Quantified chromosome structural transition pathways during cancerization and reversion

In order to obtain a quantitative picture of how the chromosome structural transitions occur during the cancerization and reversion processes, we projected all the transition trajectories as well as the averages onto several order parameters, which describe the shape of the chromosome structure and contact formation at various ranges (Fig 3 and S4 and S5 Figs). We first used the extension lengths of the chromosome structure along the longest and shortest principal axes (PA1 and PA3) [73] (Fig 3A). We found that the chromosomes at the normal and cancer cells show very similar extension lengths. However, changing the shapes of the chromosomes does not follow the direct straight line connecting the initial and final states during both the cancerization and reversion processes. In contrast, there is an increase followed by the decrease of shape extension on the chromosome structure along with both PAs during both the cancerization and reversion processes. In addition, the individual pathways show high stochasticity, reminiscent of the stochastic dynamics in cell development [83]. From the averaged pathways, we see that the transitions of these two processes do not follow the same route. This is also observed by projecting the pathways onto the radius of gyration (R_g) and aspheric quantity (Δ). Δ measures the asphericity of the chromosome structure, and the deviation of Δ from 0 measures the deviation from the perfect sphere [84]. Besides, we can see that the shape of the chromosome with the most significant expanding structure is geometrically different from the one in the ES cell. However, the chromosome at the stage “C4” in cancerization is closer to the ES cell than the stage “R4” in reversion.

The chromosome structural expansion during the cancerization and reversion can also be reflected by the local and non-local contact formations (Fig 3C and 3D and S4 Fig). Here we used the fraction of “native” contact Q , which was widely applied in protein folding [85], with references being the averages of the pairwise distances between loci in chromosome ensembles at the normal and cancer cells. There are different pathways for local chromosome structural formations during these two processes, which appear to deviate significantly from the chromosome in the ES cell. For non-local contacts, the pathways become overlapped but still show differences at the most expanding chromosome structure. Further projections of the trajectories onto the contact probability formed by different genomic distances show simultaneous adaptations of the contacts at both local and non-local ranges (S5 Fig). These results imply that the highly irreversible chromosome structural transitions occur universally. The chromosomes undergo isotropic structural expansions followed by the compaction with adapting the shapes and the contact formations during the cancerization and reversion.

Spatial rearrangements of chromosomal loci during cancerization and reversion

It has been recognized that the spatial distribution of the chromosomal loci strongly influences the transcriptional activity [86, 87]. To see how the chromosome rearranges the spatial distribution of the chromosomal loci along with the compartment formation during cancerization and reversion, we calculated the radial density $\rho(r)$ of chromosomal loci and further classified it into the compartment A and B. We found that the normal cell, which is a terminally differentiated cell, tends to locate the chromosomal loci of compartment A and B towards the chromosome’s surface and interior, respectively (Fig 4A). The finding is in line with previous

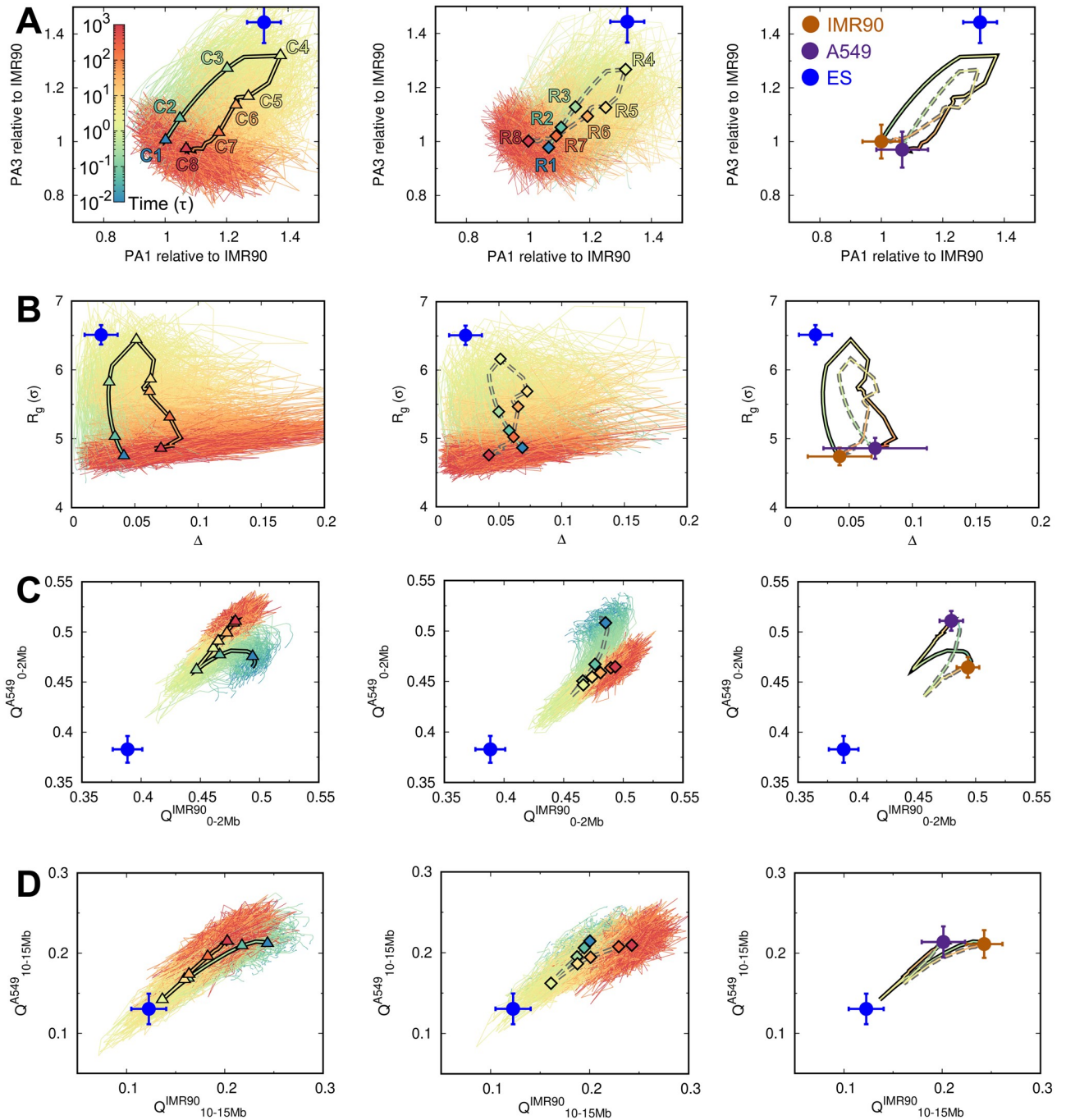


Fig 3. The pathways of chromosome structural transitions during cell cancerization and reversion. The pathways are projected onto the order parameters of the chromosome with all trajectories presented for cancerization (*Left*) and reversion (*Middle*). The 8 stages are indicated by triangles (“C1”–“C8” in cancerization, *Left*) and diamonds (“R1”–“R8” in reversion, *Middle*) on the averaged pathways of the cancerization (solid line) and reversion (dashed line), respectively. The averaged pathways are additionally shown and colored by time (*Right*). The quantities of the chromosome in the IMR90, A549, and ES cells are correspondingly placed as brown, purple and blue points, respectively. The pathways are projected onto (A) the extensions of the longest and the shortest principal axes (PA1 and PA3), (B) the radius of gyration (R_g) and the aspheric quantity (Δ), contact similarity in terms of the fraction of native contact Q to the IMR90 and A549 at (C) local range (0–2Mb), and (D) long-range (10–15Mb). The data of the ES cell are obtained from our previous work [45].

<https://doi.org/10.1371/journal.pcbi.1009596.g003>

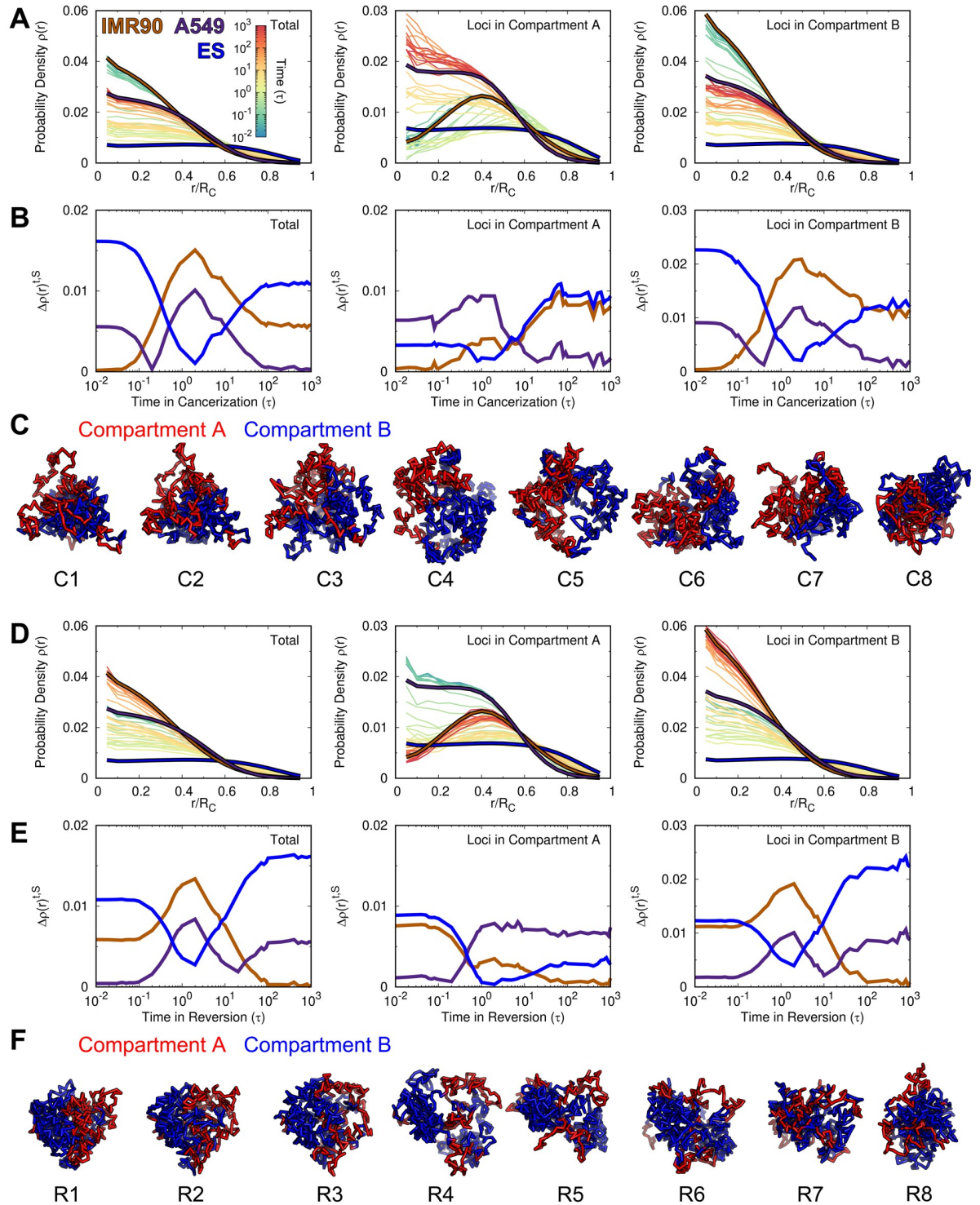


Fig 4. The change of the radial density in the chromosome during cell cancerization and reversion. (A) The change of the radial density profile $\rho(r)$ for the whole loci (Left), loci in compartment A (Middle) and loci in compartment B (Right) in the chromosome during cancerization. The profiles of ρ at the IMR90, A549 and ES cells are colored brown, purple and blue, respectively. (B) The difference of the radial density, calculated by $\Delta\rho(r)^{t,S} = \int_0^{R_C} |\rho(r)^t - \rho(r)^S| dr/R_C$, from the processing time (t) to the reference cell (S). (C) Structural illustrations of the chromosome with loci colored by compartment states during cancerization. (D-F) are the same with (A-C) but for the reversion process.

<https://doi.org/10.1371/journal.pcbi.1009596.g004>

simulations [49, 88, 89] and experiment [90]. In contrast, the chromosome in the ES cell exhibits a roughly uniform distribution of the chromosomal loci regardless of the compartment status. For the cancer cell, $\rho(r)$ appears to be the intermediate between those at the normal and ES cells. We further calculated the radial density similarity $\Delta\rho(r)^{t,S}$ between the processing state at time point t and the normal, cancer and ES cells for the total loci and the loci in compartment A and B (Fig 4B). When the normal cell converts to the cancer cell, $\Delta\rho(r)^{t,S}$ of the ES cell decreases and reaches the minimum at time $\sim 0.9-3\tau$, which are varied by the total loci and the loci in the compartment A and B (Fig 4B). We note that this period corresponds to the stage when chromosome forms more expanding structure than those at the normal and cancer cell (Fig 2B), so the results indicate that the spatial distribution of the loci in chromosome during this period is similar to that in the ES cell. Further proceeding cancerization decreases $\Delta\rho(r)^{t,S}$ of the cancer cell. Overall, we found a non-monotonic spatial repositioning of the chromosomal loci during cancerization. In the beginning, the chromosome segregates the active loci located in compartment A and inactive loci located in compartment B towards the surface and interior, respectively (Fig 4C). Then, the compartment segregation of the chromosome is still effective, but it forms a stem-like pattern with a uniform radial distribution of genomic loci regardless of the compartment states. Finally, the spatial distribution of the genomic loci reaches the one at the cancer cell.

Interestingly, we observed a similar trend of spatial repositioning of the chromosomal loci during the reversion process (Fig 4D, 4E and 4F). In other words, $\rho(r)$ during reversion also goes through processing states sharing similarity with that in the ES cell. In detail, we found that $\Delta\rho(r)^{t,S}$ of the ES cell decreases more for the loci in compartment A than the ones in compartment B at time $\sim 1-3\tau$ (Fig 4E), corresponding to the stage when chromosome forms more expanding structures than those at the normal and cancer cells (Fig 2F). This is different from that observed in the cancerization process, where spatial repositioning of the chromosomal loci in compartment B is dominant to be the stem-like, as decreasing $\Delta\rho(r)^{t,S}$ of the ES is more significant for the loci in compartment B than the ones in compartment A.

In order to see how the spatial distribution of the genes on the chromosome changes during the cancerization and reversion processes, we further calculated the radial density of genes $\rho(r)^{gene}$ during the transition (S6 Fig). We found that $\rho(r)^{gene}$ at the normal, cancer and ES cells show similar distribution behaviors with $\rho(r)$ of the chromosomal loci with the compartment A status at the corresponding cells. Changes in $\rho(r)^{gene}$ also follow similar trends of the changes in $\rho(r)$ of chromosomal loci with the compartment A status during both cancerization and reversion processes. The findings are likely due to the fact that the chromosomal loci in compartment A has a much higher gene density than the ones in compartment B. The combined results suggest that during both cancerization and reversion processes, the chromosome dynamically rearranges the spatial positions of the chromosomal loci and genes, and the distribution can resemble the one at the ES cell.

Stemness and irreversibility of chromosome structural transitions during cancerization and reversion

Our results have implied that the chromosome during the cancerization and reversion may adopt the chromosome structure in the ES cell. To quantitatively assess this, we divided the contacts by different ranges and then performed the principal component analysis (PCA) on the contact probability evolving trajectories. Since the contacts at long-ranges are usually formed with very low probabilities, it may introduce large uncertainty and imprecision when directly applying to PCA. To resolve this issue, we instead used the matrix P_{obs}/P_{exp} , which captures the essence of the compartment formation [15]. We also note that the boundaries of

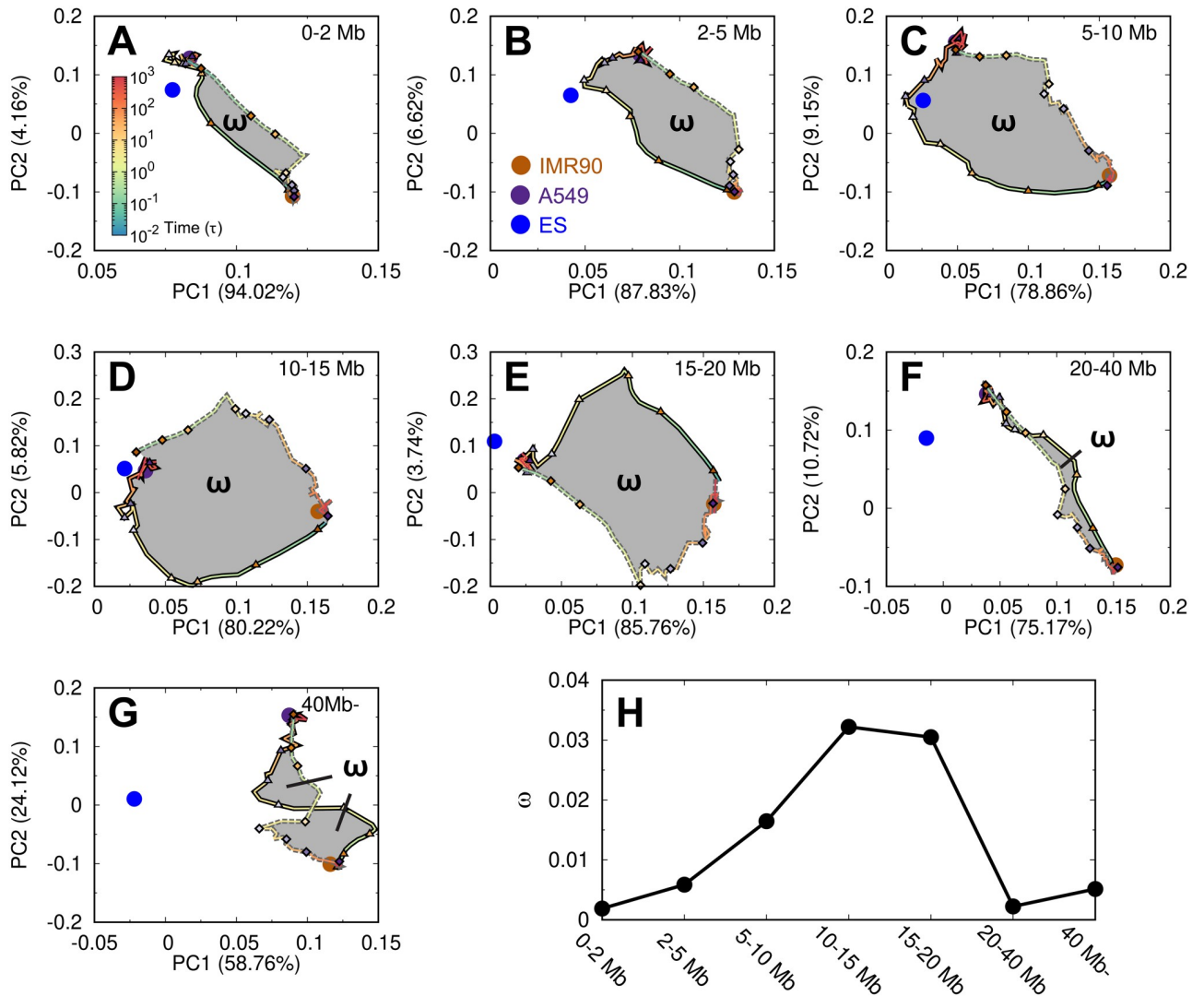


Fig 5. The pathways of the chromosome structural transition during cell cancerization and reversion. The PCA plots on the contact matrix P_{obs}/P_{exp} evolving with the time during the processes projected on the first two PCs. The plots are further divided into different contact ranges (A-G). The reference points from the IMR90, A549 and ES cells are plotted as brown, purple and blue points, respectively. The 8 stages during the cancerization paths (solid lines) and reversion paths (dashed lines) are indicated by the triangles and diamonds, respectively. The hysteresis loop (ω) is colored grey in (A-G), and the area is calculated for different contact ranges (H).

<https://doi.org/10.1371/journal.pcbi.1009596.g005>

TADs are mostly conserved during the transitions (S7 Fig). This indicates that the chromosome structural arrangements during cancerization and reversion should mostly rely on the contacts at the long ranges.

We performed the PCA plots on the transition trajectories with the first and second most weighted PCs (Fig 5). In this respect, at the local range (< 2 Mb), we see that the ES cell is not on either of the two pathways. As the contact range increases (2–5 and 5–10 Mb), the location of the ES cell becomes close to the cancerization pathway but is far from the reversion pathway. Further increasing the contact range to 10–15 and 15–20 Mb shows that the ES cell is close to the cancer cell. When the contact range is very long (20–40 Mb and > 40 Mb), the ES cell locates far from both of these two pathways. Our results showed that the chromosome during the cancerization process may explore the structures formed in the ES cell at the moderate

contact range (2–20 Mb). In contrast, the chromosome during the reversion process seems to be structurally different from the one in the ES cell, though the chromosome is prone to adopt many of the structural properties in the ES cell, such as the contact probability scaling, structural shape geometry and spatial distribution of the chromosomal loci.

Interestingly, we found that the two pathways do not overlap, resulting in the hysteresis phenomenon. Hysteresis indicates a toggle-like switching behavior and was often described in the context of ferromagnetism. In biology, hysteresis was found in the cell cycle and demonstrated as the driving force for the irreversible cell cycle transitions [91, 92]. In this respect, the hysteresis loop area provides a quantitative measure of the degree of the irreversibility, thus it was calculated here (Fig 5H). We found that the largest hysteresis area, which corresponds to the most significant irreversibility of the chromosome structural transitions during cancerization and reversion, is at the moderate contact range. Combining the results, we showed that the chromosome structural transitions during the cancerization and reversion processes are highly irreversible, and the chromosome tends to form stem-like structures during the cancerization.

Discussion

We used the landscape-switching model to explore the chromosome structural transitions during the cell cancerization and reversion processes. More than two hundred independent simulations (226 simulations for cancerization and 246 simulations for reversion, see “[Materials and methods](#)”) were performed for a long time under the potential designated for describing the chromosome dynamics in the normal (or cancer) cell state. We found that these simulations with the initial chromosome structures chosen based on the clustering method recapitulate the structural ensembles in terms of the contact maps, TADs and compartments (S8 Fig). The feature suggests that the selected chromosome structures can represent the chromosome structural distributions at the original cell states and are statistically sufficient to initialize the landscape-switching simulations. The Langevin stochastic simulations were performed to study the chromosome structural dynamics with friction and noisy terms implicitly representing the hydrodynamic interactions from the surrounding environments and fluctuations from the numerous biological activities, respectively [49]. We further described the cell state transition processes between the normal and cancer cells from the chromosome structural perspective, based on the intimate relationship between the gene expression and the chromosome structure. By projecting the transition trajectories onto different order parameters, we observed highly fluctuating, stochastic and diverse pathways. The results are reminiscent of recent increasing experimental evidence that chromosome structures exhibit high cell-to-cell variability [43, 93–96]. At the same time, we found that the averaged pathways showed clear routes for the chromosomes to accomplish the structural transitions, suggesting the deterministic dynamics in organizing the chromosome structures during the cancerization and reversion. This also resonates with the single-cell Hi-C measurements on the cell cycle process, where the cyclic chromosome structural dynamics are heterogeneous at the single-cell level but show deterministic trends collectively [38]. Therefore, our simulations indicate that the chromosome structural dynamics during the cancerization and reversion are a combination of stochastic and deterministic dynamics.

In combination with the results on the contact probability decaying with the genomic distance, the global structural geometry, the contact formation and the spatial distribution of chromosomal loci, we showed that the chromosomes with the highest degree of structural expansion during cell cancerization and reversion share significant structural similarities to the ones at the ES cell. However, there are differences between the chromosomes with

expanding structures during these two processes. As shown by the enhanced contact probability evolving trajectories (Fig 5), the chromosome during cancerization can form a stem-like chromosomal contact pattern up to the range of 20 Mb. In contrast, this is not observed during reversion. The results indicate a potential role of the cell state with stemness during cancerization in guiding the process. The chromosome focused in our study is the long arm of chromosome 14. From the curves of the contact probability $P(l)$ versus genomic distance l for other chromosomes in the IMR90 and A549 cells (S9 Fig), we found similar trends of $P(l)$ for the chromosomes 1 to 14 as well as their changes between the IMR90 and A549 cells. As $P(l)$ reflects the strengths of the chromosome interactions at different genomic distances, this feature may suggest that the results obtained in our simulations are applicable to the segments in chromosomes 1 to 14 as long as there are no chromosome abnormalities.

In order to see how the transcriptional activity changes between the normal and cancer cells, we analyzed the RNA-seq data of the IMR90, A549 and ES cells. We found that the gene expression levels of the loci across the chromosome segment focused in this study are overall quite similar for these three cells (S10A Fig). However, there are more loci with low gene expression levels in the IMR90 cell than those in the A549 and ES cell (S10B Fig), indicating that a large number of genes switch from low to high levels during the cancerization. Interestingly, we found that there is a correlation between the gene expression level changes and the compartment status switching for the cell state transition between the IMR90 and A549 cells (S10C Fig). In general, the compartment switching from B to A leads to a more significant increase of the gene expression level than the compartment switching from A to B. This is a demonstration of how the chromosome reorganizes its structure in favor of gene expression (function). In our simulations, we have shown that chromosome dynamically changes the compartment status and the spatial distribution of the chromosomal loci with different compartment status during the cancerization and reversion. Thus the dynamical adaptation of transcriptional activity along the chromosome should also be expected during the processes.

Based on our simulation results, we can propose a pictorial Waddington's landscape from the chromosome structural perspective to understand the cancerization and reversion (Fig 6). In the context of Waddington's epigenetic landscape [97], the cell, metaphorically referred to as a ball, rolls down from the ES cell at the top, which possesses the highest degree of stemness, to the terminally differentiated cell at the basin of the landscape to accomplish cell differentiation. As the reverse process, cell reprogramming transforms the differentiated cell to the ES cell by gaining the stemness [98, 99]. Hi-C data revealed that the chromosome at the cancer cell has more loci in compartment A than compartment B (Fig 1), corresponding to more populated active, open euchromatin than that of inactive, closed heterochromatin. The analyses on the gene expression levels at the IMR90, A549 and ES cells show an apparent increase in the transcriptional activity from the normal to cancer cells (S10 Fig). In particular, the increased gene expression levels for the chromosomal loci associated with compartment switching from B to A are more significant than those associated with compartment switching from A to B when the cell state changes from the normal to the cancer cell. These features suggest that the cancer cell, which exhibits the characteristics of the stem cell, should be located at a landscape layer, higher than the normal cell. The normal cell is terminally differentiated with a large proportion of the chromosomal loci in the repressive compartment B associated with low gene expression levels. Transforming the normal cell to the cancer cell increases the stemness. However, our simulations showed that cancerization is a non-monotonic process, during which the chromosome with the highest degree of structural expansion exhibits significant structural similarity with the one at the ES cell. In the landscape view, the cell at the first stage of cancerization climbs up on the landscape approaching the ES cell, reminiscent of cell reprogramming. The second stage of cell cancerization corresponds to a rolling down process

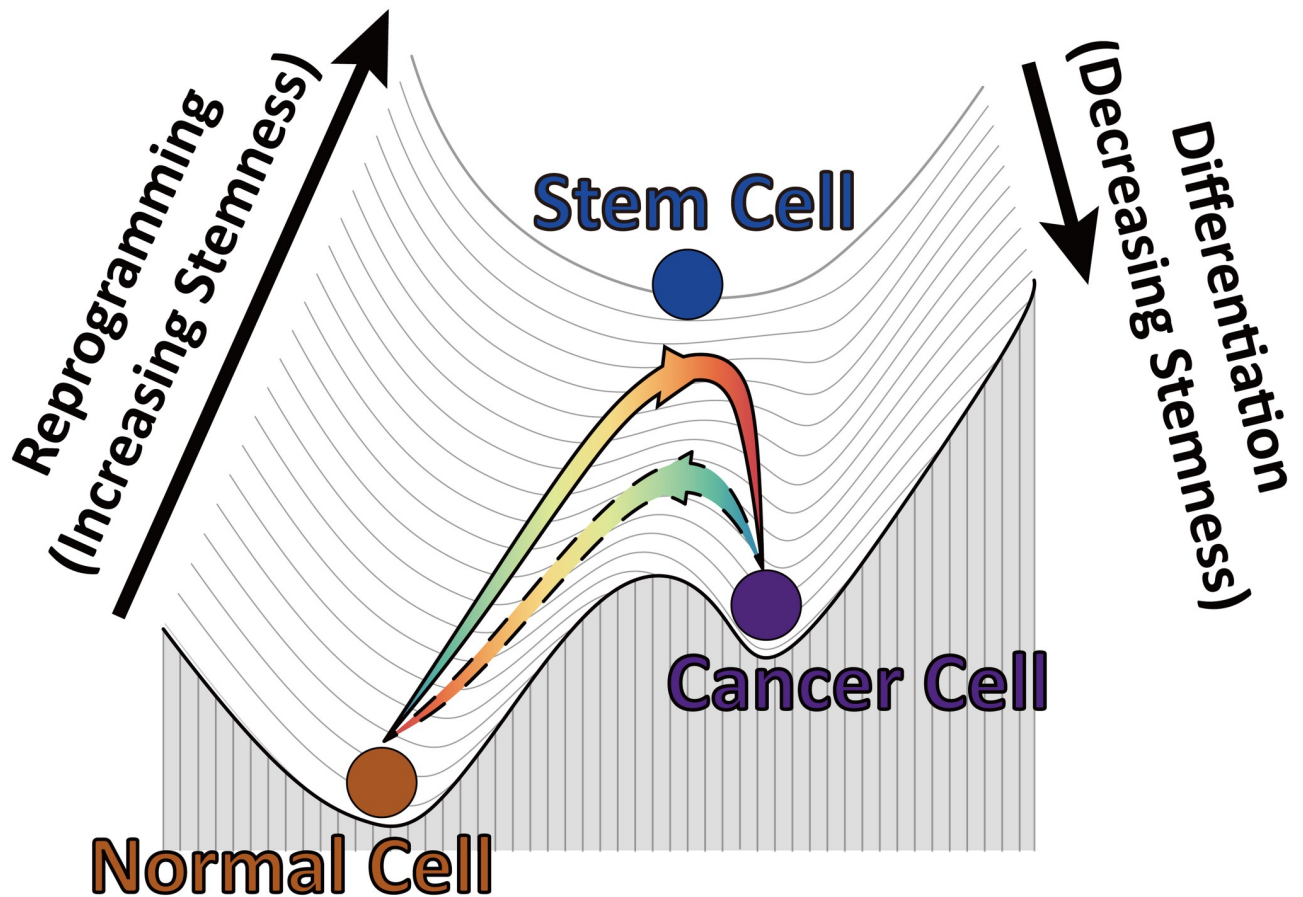


Fig 6. The Waddington landscape of the cell cancerization (solid line) and reversion (dashed line) processes from the chromosome structural transition perspective.

<https://doi.org/10.1371/journal.pcbi.1009596.g006>

toward the cancer cell, reminiscent of cell differentiation. On the other hand, the chromosome during the reversion process also expands and reorganizes its structure towards the one at the ES as an implication of gaining the stemness. However, the structural difference of the chromosome to the one at the ES during reversion is more significant than that during the cancerization. This leads to a relatively weak reprogramming process at the first stage of the reversion process. Thus the reversion pathway should appear to be under the cancerization pathway (Fig 6).

Recently, the non-monotonic chromosome structural transition was observed in a short-lived intermediate state during the mitotic exit process by the time-course Hi-C experiments [44]. The chromosome in this intermediate state was further characterized to be unbound with the structural maintenance of chromosomes (SMC) protein complexes, corresponding to the over-expanding chromosome observed in the simulations [53]. The advantage of the chromosome being over-expanding is likely because the large accessible surface in the loosely formed chromosome structure can facilitate the subsequent proteins binding. Besides, using an integrative approach combined with the measurements on gene expressions and epigenetic information, Cacchiarelli et al. found that reprogramming of the fibroblast cell to the stem cell is a non-monotonic process, which contains a reverse of differentiation process to the pre-implantation-like cell state followed by a transition to the post-implantation-like stem cell state at the late stage of reprogramming [100]. In light of the intimate relations among the gene

expression, epigenetics and 3D genome architecture [3–5], the non-monotonic transitions in chromosome structure should also be expected during the reprogramming. With a landscape-switching model, we observed the over-expanding chromosome structures, which are more open than those at the fibroblast and ES cells, during the reprogramming, resonating with the experimental evidence [45]. Here, we found that the chromosome during cell cancerization opens its structure and partially forms the structure, similar to the one formed in the ES cell. The ES cells that are richer in the less compact euchromatin than the highly condensed heterochromatin [78], were found to form the globally open chromosome structures [101, 102]. Such open chromosome structures in the ES cell may promote the chromatin-protein binding, which controls the pluripotent differentiation [103, 104]. Our findings echo with a recent experiment using super-resolution imaging techniques [105], where Xu et al. observed that the chromosome tends to open its structure with decompaction of heterochromatin in the early stages of carcinogenesis. Furthermore, it was found that the disruption of the compact heterochromatin prior to tumor formation occurs ubiquitously across multiple cancer types and shares the similar trait of the chromosome organization with the stem cells [106].

We can propose a possible explanation for the non-monotonic chromosome structural transition during the cell state transitions observed in our model. It is worth noting that the non-monotonic chromosome structures were only observed during the transitions, where the chromosomes are more compacted in the initial state than in the final state. Compared to the less compacted chromosome, the more compacted chromosome is formed by more chromosomal contacts, associated with a smaller configurational space of the structural ensemble. During the cell state transition involving the chromosome decompaction, the contacts have to be broken to overcome the entropic barrier with the aid of the extra energy input. The breaking of the contacts led by the extra energy input can be excessive to induce over-expanding chromosome structures, which are in favor of the entropy. This provides opportunities for the system to escape the trapped state and enough configurational space associated with high energy to search for the new state. In contrast, during the cell state transitions involving the chromosome compaction, the chromosome may not have to break a significant amount of contacts. In this respect, the extra energy input tends to mainly guide the chromosome to form and stabilize the contacts to compact the chromosome, so the over-expanding chromosome structures were not observed. When it comes to the cancer processes, where the chromosomes have a similar degree of compaction at the initial and final states, the breaking and formation of contacts can occur simultaneously at comparable degrees, so the over-expanding chromosome structures can still be observed.

Our simulation showed that cancerization and reversion are irreversible processes. Previously, it was assumed that the forward and reverse processes of cell state transitions may pass through the same intermediate cell states, referred to as the metastable attractors on the Waddington's landscape [107, 108]. Theoretical studies at the gene network level have shown that irreversibility is a universal feature for cell state transition due to the nonequilibrium effects [56, 109]. Our simulation results provided a molecular-level description that the irreversibility of cell cancerization and reversion can be reflected by the chromosome structural dynamics. Besides, we proposed a quantitative way to measure the irreversibility in cancerization and reversion using the hysteresis loop analysis of the chromosomal contact formation. Our findings also resonate with a recent experiment, which used a combination of Hi-C and replication timing analyses to characterize the irreversible non-overlapped pathways between the cell state transition, namely the differentiation and reprogramming [110].

The stemness-related genetic markers and interactions in cells are controlled by the underlying gene regulatory network [111], which uses chromosomes as the structural scaffold for functionalities. Besides, the CS cells were found to have similar marker expression profiles

with normal stem cells [112]. These features support that our approach for inferring the stemness and connecting it to the partial formation of the CS cell is reasonable and based on the intimate structure-function relationship at the genomic level [4, 5, 86]. Therefore, we speculate that the stem-like intermediate cell state during cell cancerization may contribute to the formation of the CS cells, of which the terminally differentiated somatic cell has been suggested as one potential origin [113]. Due to the lack of the Hi-C data at the CS cells, the precise and direct comparisons of the intermediate state to the CS cells were not performed. A strict assessment of our statement can be made when the experimental data are available.

Our predictions of the transient intermediate state with stemness during cell cancerization may have implications for cancer treatments. The increase of stemness during cancerization may promote the cell self-renewal and differentiation processes, similar to what the CS cells do. Thus, decreasing the stemness of the cell by deviating the chromosome structures from the one at the ES cell can help to repress the cancer initiation and progression. Further efforts can be focused on developing effective genome structure engineering methods for modulating and disrupting the specific chromosomal interactions similar to those formed in the ES cells [114–118], serving as a potential therapeutic approach.

Our predictions can be tested by future experiments. In this regard, the time-course Hi-C experiments appear to be promising and competent in monitoring the chromosome structural evolution during the cancerization and reversion processes [37]. It is expected that the Hi-C data at the high temporal resolution can provide a dynamical picture of how the chromosome structure rearranges and offer useful clues as to the formation of the stem-like intermediate state during cell state transition using advanced data-analysis methods [44, 119]. Meanwhile, mounting theoretical and experimental studies confirm the existence of intermediates during the epithelial-mesenchymal transition (EMT) [120–123], which contributes to cancer metastasis and tumor relapse. The EMT intermediates often exhibit a gain of stemness compared with the epithelial and mesenchymal cells [124–127], leading to a non-monotonic increasing-followed-by-decreasing trend in stemness during the EMT [128]. These properties of the EMT process are similar to what we have observed in our simulations of the transition from the normal cell to the cancer cell. Therefore, the fruitful approaches developed for uncovering the EMT intermediate and studying the EMT pathways can be adapted to investigate the cancerization processes [129]. Ultimately, the combined efforts on the direct cancerization and the EMT processes will help us to complete the understanding of how cancer initializes and metastasizes.

Our landscape-switching model can be improved from the following three aspects. First, the model used a simplified regime at the nonadiabatic limit to describe the cancerization and reversion processes. However, the ratio between the rates of the intra- and inter-landscape dynamics are not infinite, thus there is a certain degree of adiabaticity in the cell state transitions processes. To take into account the adiabaticity, more switches of the landscape can be implemented, which can result in an elevated rate for the inter-landscape hopping [66, 67]. In this regard, the experimental determinations of the timescales for cancer initiation and progression are prerequisites to include the adiabaticity in the model. Noteworthy, as an extreme case where multiple switches of the landscape occur frequently, corresponding to a much faster inter-landscape hopping than intra-landscape motion, the process approaches the adiabatic limit, which can be described by an averaged classic intra-landscape barrier crossing on a single effective landscape [130]. In practice, cell cancerization and reversion should occur between two extreme cases: the nonadiabatic landscape-switching process and the effective intra-landscape barrier-crossing process. The precise determination of the adiabaticity of the process relies on future experiments focusing on measuring the timescales of the chromosome dynamics relaxation within one cell state and hopping between the two cell states. Second, the

model considers a bistable system with abrupt switches for describing the cell state transitions. However, the cell state transitions should occur continuously with finite rates in reality. This indicates that further improvements on the connections between these two landscapes are needed. Nevertheless, there is still a lack of clear guidance to determine the form of the connections that can best capture the essence of the processes. One may build a “double-basin” landscape widely used in studying the protein conformational changes, which usually occur between the two distinct structural states [131–133]. However, this landscape leads to an equilibrium process with an often very high potential barrier, which makes it almost impossible to switch the state. This contradicts the nonequilibrium nature of chromosome structural changes during the cell state transitions, which require significant energy input to help overcome the barrier. The landscape-switching model used a simple way to establish the connections between two landscapes with considering the nonequilibrium effects. Further improvements on the model can be made by refining the form of the connections between the two landscapes with the experiments. It is worth noting that the (meta)stable intermediate cell states may exist during the cancerization and reversion [134]. Thus, we expect that including the intermediate states through stepwise switches among the multiple (meta)stable cell states in the system can improve the accuracy of the connections between the normal and cancer cells. This can be implemented in future studies when the experimental Hi-C data on the intermediate states are available. Third, the model did not take into account the cell cycle process. The cell state transitions between the normal and cancer cells often undergo multiple rounds of cell cycle processes. In general, the interphase occupies a significant amount of time during the cell cycle. In this regard, we focused on the slow cell state transitions in the interphase during the cell cancerization and reversion processes while treating the fast cell cycle dynamics as an averaged background. In reality, the cell development and cell cycle are controlled by two different regulatory networks [135]. The quantitative characterization of the interplay between these two networks in future studies can help the further improvements on the model by incorporating the cell cycle processes.

In summary, we presented a chromosome structural-level study of illuminating the molecular mechanisms of cancer formation and its reversion processes. The targeted cells belong to the human lung organism and should be regarded as one example of cancer. Up to now, more than a hundred distinct types of cancer have been uncovered. Different types of cancer may vary substantially in their behaviors. However, all types of cancer are caused by the common mechanism of uncontrolled cell division, reflecting the dysregulation of cell self-renewal. Therefore, we speculate that our observation of the stem-like intermediate state during cancer formation may likely be observed by the other types of cancer. The open chromosome structure in the stem-like intermediate state may contribute to the increased genomic instability and active transcription for cells to gain plasticity that has been demonstrated as the universal key to facilitate cancer progression [136]. Future studies on elucidating the cancer formation pathways and the stem-like intermediate state can help to decipher the fundamental mechanisms that underlie the different types of cancer.

Materials and methods

Chromosome polymer model

We built a beads-on-a-string model at the 100 kb resolution to represent the chromosome segment (chr14: 20.5–106.1 Mb), resulting in a total of 857 beads in the system. The basic background of our polymer model is a generic polymer model with homogeneously weighted

bonded and nonbonded potentials:

$$V_{\text{Polymer}} = V_{\text{Bond}} + V_{\text{Nonbond}}.$$

The bonded potential contains the finitely extensible nonlinear elastic (FENE) bond stretching interactions, linear-promoting angle bending interactions, and neighboring beads non-overlapped interactions with the expression as follows [137]:

$$V_{\text{Bond}} = V_{\text{FENE}} + V_{\text{Angle}} + V_{\text{hc}},$$

where the FENE potential for the bond ($r_{i,i+1}$) between neighboring beads ($i, i+1$) is expressed as:

$$V_{\text{FENE}} = -0.5K_b R_0^2 \ln\left[1 - \left(\frac{r_{i,i+1}}{R_0}\right)^2\right],$$

and the non-overlapped interaction between ($i, i+1$) is described by a hard-core Lennard-Jones potential (V_{LJ}):

$$V_{\text{hc}} = \begin{cases} V_{LJ}(r_{i,i+1}), & r_{i,i+1} \leq \sigma 2^{1/6}, \\ 0, & r_{i,i+1} > \sigma 2^{1/6}, \end{cases}$$

where

$$V_{LJ}(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^6 \right] + \epsilon.$$

To consider the stiffness of the chain at the local range, we applied an angle potential acting on the three adjacent beads ($i - 1, i, i + 1$) with the following form, which was used in the previous studies [72, 137]:

$$V_{\text{Angle}} = K_a [1 - \cos(\theta_i - \pi)].$$

The angle term describes the local stiffness of the chain, thus it favors the linear placement of the three adjacent beads with $\theta_i = \pi$.

The nonbonded potential for every nonbonded pair (i,j) was set as the soft-core interactions that allow the chain-crossing to mimic the effects of a large number of topoisomerases in the cell nucleus on efficiently unknotting the DNA molecules *in vivo* [138, 139].

$$V_{\text{Nonbond}} = \begin{cases} 2(1 + \tanh[0.5V_{LJ}(r_{ij}) - 1]), & r \leq r_0, \\ V_{LJ}(r_{ij}), & r_0 < r_{ij} \leq \sigma 2^{1/6}, \\ 0, & r_{ij} > \sigma 2^{1/6}. \end{cases}$$

In addition, a spherical confinement with the radius R_C was used to mimic the volume fraction of the chromosome in the cell nucleus at 10%, same as used previously [72, 137]. With only the generic polymer potential, the model generates an ensemble that resembles the equilibrium globule [45].

Reduced units were used throughout simulations. The energy unit was $\epsilon = 1.0$. The bond length σ was set to be the length unit. In FENE potential, $R_0 = 1.5\sigma$ allows the bonds to stretch flexibly and $K_b = 30.0/\sigma^2$. The angle potential has a moderate strength of $K_a = 2.0$. We note that a very large K_a would lead to a high degree of chain stiffness at the local range, contradictory to

the highly flexible nature of the chromatin. In contrast, a very weak angle term would eliminate the effects of the stiffness. In the soft-core potential, we set $r_0 = \sigma / ((1 + \sqrt{2})/2)^{1/6}$, so $V_{LJ}(r_0) = 2.0$ [72]. The radius of the spherical confinement R_C was set to be 9.7σ . The temperature in all simulations was set to 1.0 in energy units by multiplying by the Boltzmann constant. Langevin stochastic dynamics was applied with a time step of 0.0005τ and a friction coefficient of $10.0\tau^{-1}$, where τ is the reduced time unit.

Growing lines of evidence showed that the activities of topoisomerases are effective throughout the cell cycle [140–142]. For simplicity, we kept the soft-core potential constant throughout the simulations. The strength of V_{Nonbond} appears to be moderate with a maximum of 4.0 when the two beads are fully overlapped. A very strong soft-core potential, which resembles the hard-core potential, would reduce the possibility of chain-crossing [143]. In contrast, a very weak soft-core potential could lead to an unrealistic biological picture, where the chromosomal loci do not have excluded volumes and the reaction rates of the topoisomerases on DNA topology are infinite.

Maximum entropy principle simulation

The potential $V(\mathbf{r})$ in the maximum entropy principle simulations is made up of the generic polymer potential V_{Polymer} and the Hi-C restraint potential $V_{\text{Hi-C}}$:

$$V(\mathbf{r}) = V_{\text{Polymer}} + V_{\text{Hi-C}}.$$

The entropy of system under $V(\mathbf{r})$ relative to a given prior distribution $\varrho_0(\mathbf{r})$ under polymer potential V_{Polymer} is defined as [71]:

$$\mathbb{S}[\varrho|\varrho_0] = - \int d\mathbf{r} \varrho(\mathbf{r}) \ln \frac{\varrho(\mathbf{r})}{\varrho_0(\mathbf{r})}.$$

This entropy should be maximized subject to the following constraints in order to be compatible with observations (contact probability $P_{i,j}$ and experimental Hi-C data $f_{i,j}$):

$$\left\{ \begin{array}{l} \int d\mathbf{r} \varrho(\mathbf{r}) P_{i,j} = \langle P_{i,j} \rangle = f_{i,j}, \\ \int d\mathbf{r} \varrho(\mathbf{r}) = 1. \end{array} \right.$$

The maximization of the entropy can be obtained using the method of Lagrangian multipliers by searching for the stationary points of the Lagrange function:

$$\mathcal{L} = \mathbb{S}[\varrho|\varrho_0] + \sum_{i,j} \alpha_{i,j} (\int d\mathbf{r} \varrho(\mathbf{r}) P_{i,j} - f_{i,j}) + \beta (\int d\mathbf{r} \varrho(\mathbf{r}) - 1),$$

where $\alpha_{i,j}$ and β are Lagrangian multipliers. By setting $\frac{\delta \mathcal{L}}{\delta \varrho(\mathbf{r})} = 0$ and neglecting the normalization factor, the posterior distribution has the following expression:

$$\varrho(\mathbf{r}) \propto e^{-\sum_{i,j} \alpha_{i,j} P_{i,j}} \cdot \varrho_0(\mathbf{r}).$$

Therefore, the potential $V(\mathbf{r})$ with maximum entropy principle, is expressed as follows:

$$V(\mathbf{r}|S) = V_{\text{Polymer}} + \sum_{i,j} \alpha_{i,j} P_{i,j},$$

where $V_{\text{Hi-C}}$ is in a linear form of contact probabilities and S represents the cell state (IMR90 or A549). In practice, $P_{i,j}$ is the calculated contact probability between the chromosomal loci i

and j using step function expressed as:

$$P_{i,j} = \frac{1}{2} [1 + \tanh(\mu(R_0 - r_{i,j}))],$$

where we set $\mu = 3.0/\sigma$. $\alpha_{i,j}$ is the corresponding contact strength, which is iteratively adapted by simulations. $V_{\text{Hi-C}}$ acts on the non-neighboring beads with $|i - j| > 1$, which includes the pair $(i - 1, i + 1)$ involved in the angle term. During the iteration process, the contact probability $P_{i,j}$ is restrained by the Hi-C data $f_{i,j}$. In the end, the chromosome under the potential $V(\mathbf{r}|S)$ generates an ensemble that can reproduce the Hi-C maps at the normal and cancer cell, separately (S1 Fig). Details of the maximum entropy principle simulation can be found in previous studies [72, 73]. The maximum entropy principles simulations for the ES and IMR90 cells were done in our previous work [45].

Landscape-switching model

We used the resulting potentials $V(\mathbf{r}|S)$ of the maximum entropy principle simulations for the normal and cancer cells to represent the effective energy landscapes. We performed the hierarchical clustering on the chromosome ensembles generated by the maximum entropy principle simulations. Two chromosome structures in each cluster, which has a population higher than 0.3%, were picked out as the initial structures for performing the landscape-switching simulations (S2 Fig). The purpose of choosing the chromosome structures in the populated clusters rather than using the whole sets of the structures in the ensembles is to generate a limited number of structures that can sufficiently represent the ensemble in order to significantly reduce the computational expenses. This led to 226 and 246 trajectories for the cancerization and reversion processes, respectively.

The landscape-switching simulations were performed as follows. First, the simulations were run under the potential $V(\mathbf{r}|IMR90)$ (cancerization) or $V(\mathbf{r}|A549)$ (reversion) for 5000τ , where τ is the time unit of the simulations. Then a sudden switch of potential in terms of $V(\mathbf{r}|IMR90) \rightarrow V(\mathbf{r}|A549)$ (cancerization) or $V(\mathbf{r}|A549) \rightarrow V(\mathbf{r}|IMR90)$ (reversion) was implemented. Finally, the simulations were under the new potential $V(\mathbf{r}|A549)$ (cancerization) or $V(\mathbf{r}|IMR90)$ (reversion) for 10000τ . All the trajectories were collected and combined for the analyses to generate the results present in this study.

Trajectory analysis

To cluster the processing states during the transition based on the chromosome structural similarity, we calculated the contact map difference $\Delta P^{I,J}$ between the time point $t = I$ and $t = J$ by:

$$\Delta P^{I,J} = \sum_{i,j} |P_{i,j}^{t=I} - P_{i,j}^{t=J}| / \sum_{i,j} P_{i,j}^{t=I},$$

where $P_{i,j}^{t=I(J)}$ is the contact probability between the chromosomal loci i and j at the time $t = I$ or J during the transition. $\Delta P^{I,J}$ was then normalized from 0 to 1. We performed the hierarchical clustering on $\Delta P^{I,J}$ and presented the dendrogram on the relationships between the processing states during the transition. This was done based on the total contact pairs, local contact pairs ($|i - j| < 2Mb$) and non-local contact pairs ($|i - j| \geq 2Mb$). In order to see how the chromosome structures evolve with respect to the ones of the IMR90, A549 and ES cells during the transitions, we replaced the contact probability map at the time point $t = J$ with the ones at the IMR90, A549 and ES cells. Therefore, we obtained the differences of contact maps at the

processing states relative to the Hi-C data of the IMR90, A549 and ES cells:

$$\Delta P^{t,S} = \sum_{ij} |P_{ij}^t - P_{ij}^S| / \sum_{ij} P_{ij}^t,$$

where S stands for the IMR90, A549 and ES cells, respectively.

The fraction of native contacts Q was used to compare the structural similarity of the chromosomes to ones at the IMR90 or A549 cells. The references of the pairwise distances in Q were set to be the averages of the pairwise distances in the chromosome ensemble at the IMR90 or A549 cell obtained from the maximum entropy principle simulations. The fraction of native contacts Q_{ij} for the chromosomal loci i and j is expressed as:

$$Q_{ij}^{IMR90(A549)} = \frac{1}{N_{sum}} \sum_{ij} \exp\left[-\frac{(r_{ij} - \langle r_{ij}^{IMR90(A549)} \rangle)^2}{2(0.5\sigma)^2}\right],$$

where N_{sum} is the number of the summed pairs and $\langle r_{ij}^{IMR90(A549)} \rangle$ is the averaged distance between the chromosomal loci i and j of the structural ensemble at the IMR90 (A549) cell. To see how the chromosome structures form at different genomic sequence distances, Q can be calculated based on the genomic distance l :

$$Q_l = \frac{1}{N_{sum}} \sum_{ij}^{|i-j| \in l} Q_{ij}.$$

The radial density $\rho(r)$ was used to describe the spatial distribution of chromosomal loci, expressed as:

$$\rho(r) = \frac{1}{N} \frac{n(r)}{4\pi r^2 \Delta r},$$

where N is the number of the beads in the simulation, r is the distance to the center of the spherical confinement, which has a radius of R_C , $n(r)$ is the number of chromosomal loci found in the spherical shell of $r, r + \Delta r$. Here, we set $\Delta r = R_C/20$. We further calculated $\rho(r)$ based on the loci from different compartment status and gene density, which gives rise to the radial gene density $\rho(r)^{gene}$.

We calculated the enhanced contact matrix P_{obs}/P_{exp} , which is the ratio between the observed contact probability P_{obs} and expected contact probability P_{exp} [15]. The matrix was then divided into different categories based on different genomic distances l between the interacting loci. We performed the PCA plots of the time evolution of the P_{obs}/P_{exp} at different genomic distances l during the transitions. In practice, a PCA plot shows clusters of samples based on their similarity and reduces the number of dimensions by constructing PCs. The first two PCs were shown to describe the trajectories of P_{obs}/P_{exp} evolution.

Supporting information

S1 Text. Additional materials and methods.

(PDF)

S1 Fig. Comparisons of the maximum entropy principle simulation results and the experimental Hi-C data in the A549 cells. (A) Hi-C contact maps of the chromosome ensemble obtained from the simulations and the Hi-C data at the global (*Left*) and local (*Right*) scales. (B) Contact probability versus genomic distance in the chromosome for the simulations and the Hi-C data with a slope of -1.0 in the logarithmic scale at 0.5–7 Mb. (C) Insulation score of

the chromosome obtained by the simulations and the Hi-C data. (D) Compartment profiles of the chromosome obtained by the simulations and the Hi-C data. (E) Correlations of insulation score (*Left*) and compartment profiles (*Right*) between the simulations and the Hi-C data.

(TIF)

S2 Fig. Chromosome ensemble in the A549 cells. (A) Contact distance (d) versus genomic distance (l) in the chromosome. (B) The probability distributions of aspheric parameters of the chromosome. Δ and S_{shape} were calculated using the inertia tensor [84]. Deviation of Δ from 0 (the value corresponding to a sphere) gives an indication of the extent of anisotropy. Negative values of S_{shape} correspond to oblate shapes and positive values of S_{shape} to prolate shapes. (C) The probability distribution of the configurational extension on the three principal axes of the chromosome. (D) The time evolution of the average root mean square distance (d_{rms}) between every genomic pair in chromosome at the time t relative to its initial value: $\delta(t) = \sum_{i,j} d_{rms}(i, j, t)/N_{sum}$, where N_{sum} is the number of summed pairs. The maximum of $\delta(t)$ is close to and below 3σ . (E) The hierarchical clustering of the chromosome structural ensemble shown as a dendrogram (*Top*) and the populations of the cluster (*Bottom*). Cut-off distance 3σ was applied. (F) The top 3 most populated chromosome clusters. Each is shown with a mixed contact map (*Left*), which contains 5 structures within the cluster, and one representative structure (*Right*).

(TIF)

S3 Fig. Chromosome diffusion dynamics in the IMR90 and A549 cells. (A) Fitting of MSD to the power-law function ($MSD \sim \alpha x^\beta$) in the IMR90 cell. MSD was calculated by averaging the MSD from 5 independent simulations with the potential from the maximum entropy principle simulation $V(r|IMR90)$. (B) MSD of all the individual chromosomal loci in the IMR90 cell state obtained from one simulation with average shown as the black line. (C) and (D) are same as (A) and (B) but for the A549 cell.

(TIF)

S4 Fig. Pathways of chromosome structural transitions projected on Q varied by different contact ranges during cancerization (*Left*) and reversion (*Middle*), as well as their averages (*Right*).

(TIF)

S5 Fig. Pathways of chromosome structural transitions projected on contact probability varied by different contact ranges during cancerization (*Left*) and reversion (*Middle*), as well as their averages (*Right*).

(TIF)

S6 Fig. Gene density and radial density distribution of the genes in the chromosome during cell cancerization and reversion. (A) Gene density along the chromosome and the compartment status of chromosomal loci at the IMR90, A549 and ES cells. Compartment A and B are colored red and blue, respectively. (B) The change of the radial gene density profile $\rho(r)^{gene}$ in the chromosome during cancerization. The profiles of the IMR90, A549 and ES cells are colored brown, purple and blue, respectively. (C) The difference of the radial density from the processing time to the reference cell (the IMR90, A549 or ES cell) during cancerization. (D) and (E) are the same with (B) and (C) but for the reversion process. The plotting details of (B-E) are the same with Fig 4, but for the radial density distributions of the genes.

(TIF)

S7 Fig. TAD structural formation. (A-C) The correlation between the insulation score among the IMR90, A549 and ES cells. The correlation coefficient of insulation score of the processing state during (D) cancerization and (E) reversion with those of the IMR90, A549 and ES cells. (F) An illustration of TAD boundary overlapping between the Hi-C data (IMR90 or A549) and simulation processing state during cancerization and reversion. The numbers of TAD boundaries detected by insulation score for experimental Hi-C (IMR90 or A549), simulation processing state, and the overlap between the Hi-C (IMR90 or A549) and processing state are denoted as $N_{Hi-C}^{IMR90(A549)}$, N_{PS} , and $N_{ol}^{IMR90(A549)}$, respectively. (G) The change of TAD boundary overlapping population from the simulation processing state to the Hi-C of the IMR90 and A549 (calculated by $N_{ol}^{IMR90}/N_{Hi-C}^{IMR90}$ and $N_{ol}^{A549}/N_{Hi-C}^{A549}$) and from the Hi-C of the IMR90 and A549 to the simulation processing state (indicated by superscript “+” and calculated by N_{ol}^{IMR90}/N_{PS} and N_{ol}^{A549}/N_{PS}) during cancerization and reversion. (TIF)

S8 Fig. The comparisons of the chromosome structures between the clusters for initializing the landscape-switching simulations and the ensemble generated by the maximum entropy principle simulations. (A) The IMR90 cell. (B) The A549 cell. (TIF)

S9 Fig. Contact probability versus genomic distance in Hi-C data for chromosomes 1-22 at the IMR90 (brown) and A549 cells (purple). The dashed lines indicate the chromosome segment used in this study (chr14: 20.5-106.1 Mb). (TIF)

S10 Fig. Gene expression at the IMR90, A549 and ES cells. (A) Gene expression level along the chromosomal loci. Gene expression data were measured by the RNA-seq. The expression level for each bead in our model, which represents a DNA segment of 100 Kb in length, was determined as the sum of the reads in this 100 kb DNA segment. The value was further scaled by 10^6 to indicate the reads of “per million” and was represented in the logarithmic scale (\log_2). The calculation of the gene expression level is similar to the Reads Per Kilobase Million (RPKM), widely used in RNA-seq analysis. (B) Distributions of gene expression levels in the IMR90, A549 and ES cells. (C) Distributions of the changes in the gene expression levels for the genes that change the compartment status (“A to B” or “B to A”) or that remain the same (“stable”) when comparing the IMR90 cell to the A549 and ES cells, respectively. The red lines and green diamonds in the box plots indicate the median and mean values of the distributions, respectively. (TIF)

S11 Fig. The correlations between the contact probability from the maximum entropy principle simulations and the Hi-C data for the IMR90, A549 and ES cells. Pearson’s correlation coefficients were calculated based on the absolute values and logarithmic scaled values of contact probability. The calculations based on the logarithmic scaled values of contact probability were performed after removing the contacts with the probability equal to 0 in the Hi-C data. The linear fit of the simulated contact probability and the Hi-C data is shown. (TIF)

S12 Fig. Polymer ensemble from the simulations with the generic polymer model under the potential $V_{polymer}$ and spherical confinement. (A) The contact maps of the polymer at global (*top*) and local (*bottom*) scales. (B) Insulation score. (TIF)

Acknowledgments

The authors would like to thank Stony Brook Research Computing and Cyberinfrastructure, and the Institute for Advanced Computational Science at Stony Brook University for access to the high-performance SeaWulf computing system.

Author Contributions

Conceptualization: Xiakun Chu, Jin Wang.

Data curation: Xiakun Chu.

Formal analysis: Xiakun Chu, Jin Wang.

Investigation: Xiakun Chu, Jin Wang.

Methodology: Xiakun Chu.

Project administration: Jin Wang.

Resources: Xiakun Chu.

Software: Xiakun Chu.

Supervision: Jin Wang.

Validation: Xiakun Chu, Jin Wang.

Visualization: Xiakun Chu, Jin Wang.

Writing – original draft: Xiakun Chu, Jin Wang.

Writing – review & editing: Xiakun Chu, Jin Wang.

References

1. Repana D, Nulsen J, Dressler L, Bortolomeazzi M, Venkata SK, Tourna A, et al. The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome biology*. 2019; 20(1):1. <https://doi.org/10.1186/s13059-018-1612-0> PMID: 30606230
2. Croce CM. Oncogenes and cancer. *New England journal of medicine*. 2008; 358(5):502–511. <https://doi.org/10.1056/NEJMra072367> PMID: 18234754
3. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013; 503(7475):290–294. <https://doi.org/10.1038/nature12644> PMID: 24141950
4. Cavalli G, Misteli T. Functional implications of genome topology. *Nature structural & molecular biology*. 2013; 20(3):290. <https://doi.org/10.1038/nsmb.2474> PMID: 23463314
5. Bonev B, Cavalli G. Organization and function of the 3D genome. *Nature Reviews Genetics*. 2016; 17(11):661. <https://doi.org/10.1038/nrg.2016.112> PMID: 27739532
6. Akhtar-Zaidi B, Cowper-Sal R, Corradin O, Saiakhova A, Bartels CF, Balasubramanian D, et al. Epigenomic enhancer profiling defines a signature of colon cancer. *Science*. 2012; 336(6082):736–739. <https://doi.org/10.1126/science.1217277> PMID: 22499810
7. Kron KJ, Bailey SD, Lupien M. Enhancer alterations in cancer: a source for a cell identity crisis. *Genome medicine*. 2014; 6(9):77. <https://doi.org/10.1186/s13073-014-0077-3> PMID: 25473436
8. Ooi WF, Xing M, Xu C, Yao X, Ramlee MK, Lim MC, et al. Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nature communications*. 2016; 7(1):1–17. <https://doi.org/10.1038/ncomms12983> PMID: 27677335
9. Flavahan WA, Drier Y, Liao BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. 2016; 529(7584):110–114. <https://doi.org/10.1038/nature16490> PMID: 26700815
10. Corces MR, Corces VG. The three-dimensional cancer genome. *Current opinion in genetics & development*. 2016; 36:1–7. <https://doi.org/10.1016/j.gde.2016.01.002> PMID: 26855137

11. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. *Nature reviews cancer*. 2004; 4(3):177–183. <https://doi.org/10.1038/nrc1299> PMID: 14993899
12. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability—an evolving hallmark of cancer. *Nature reviews Molecular cell biology*. 2010; 11(3):220–228. <https://doi.org/10.1038/nrm2858> PMID: 20177397
13. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *cell*. 2011; 144(5):646–674. <https://doi.org/10.1016/j.cell.2011.02.013> PMID: 21376230
14. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *science*. 2002; 295(5558):1306–1311. <https://doi.org/10.1126/science.1067799> PMID: 11847345
15. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009; 326(5950):289–293. <https://doi.org/10.1126/science.1181369> PMID: 19815776
16. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014; 159(7):1665–1680. <https://doi.org/10.1016/j.cell.2014.11.021> PMID: 25497547
17. Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, et al. Integrative detection and analysis of structural variation in cancer genomes. *Nature genetics*. 2018; 50(10):1388. <https://doi.org/10.1038/s41588-018-0195-8> PMID: 30202056
18. Taberlay PC, Achinger-Kawecka J, Lun AT, Buske FA, Sabir K, Gould CM, et al. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome research*. 2016; 26(6):719–731. <https://doi.org/10.1101/gr.201517.115> PMID: 27053337
19. Li L, Barth NK, Pilarsky C, Taher L. Cancer is associated with alterations in the three-dimensional organization of the genome. *Cancers*. 2019; 11(12):1886. <https://doi.org/10.3390/cancers11121886> PMID: 31783642
20. Akdemir KC, Le VT, Chandran S, Li Y, Verhaak RG, Beroukhim R, et al. Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature genetics*. 2020; 52(3):294–305. <https://doi.org/10.1038/s41588-019-0564-y> PMID: 32024999
21. Akdemir KC, Le VT, Kim JM, Killcoyne S, King DA, Lin YP, et al. Somatic mutation distributions in cancer genomes vary with three-dimensional chromatin structure. *Nature Genetics*. 2020; p. 1–11. <https://doi.org/10.1038/s41588-020-0708-0> PMID: 33020667
22. Barutcu AR, Lajoie BR, McCord RP, Tye CE, Hong D, Messier TL, et al. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome biology*. 2015; 16(1):214. <https://doi.org/10.1186/s13059-015-0768-0> PMID: 26415882
23. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *nature*. 2001; 414(6859):105–111. <https://doi.org/10.1038/35102167> PMID: 11689955
24. Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, et al. Identification of a cancer stem cell in human brain tumors. *Cancer research*. 2003; 63(18):5821–5828. PMID: 14522905
25. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, et al. Identification of human brain tumour initiating cells. *nature*. 2004; 432(7015):396–401. <https://doi.org/10.1038/nature03128> PMID: 15549107
26. Galli R, Binda E, Orfanelli U, Cipelletti B, Gritti A, De Vitis S, et al. Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer research*. 2004; 64(19):7011–7021. <https://doi.org/10.1158/0008-5472.CAN-04-1364> PMID: 15466194
27. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. *Journal of biomedical science*. 2018; 25(1):1–18. <https://doi.org/10.1186/s12929-018-0426-4> PMID: 29506506
28. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature reviews cancer*. 2008; 8(10):755–768. <https://doi.org/10.1038/nrc2499> PMID: 18784658
29. Schatton T, Frank NY, Frank MH. Identification and targeting of cancer stem cells. *Bioessays*. 2009; 31(10):1038–1049. <https://doi.org/10.1002/bies.200900058> PMID: 19708024
30. Li C, Wang J. Quantifying the underlying landscape and paths of cancer. *Journal of The Royal Society Interface*. 2014; 11(100):20140774. <https://doi.org/10.1098/rsif.2014.0774> PMID: 25232051
31. Li C, Wang J. Quantifying the landscape for development and cancer from a core cancer stem cell circuit. *Cancer research*. 2015; 75(13):2607–2618. <https://doi.org/10.1158/0008-5472.CAN-15-0079> PMID: 25972342
32. Powers S, Pollack RE. Inducing stable reversion to achieve cancer control. *Nature Reviews Cancer*. 2016; 16(4):266. <https://doi.org/10.1038/nrc.2016.12> PMID: 27458638

33. Rapino F, Robles EF, Richter-Larrea JA, Kallin EM, Martinez-Climent JA, Graf T. C/EBP α induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell reports*. 2013; 3(4):1153–1163. <https://doi.org/10.1016/j.celrep.2013.03.003> PMID: 23545498
34. Ishay-Ronen D, Diepenbruck M, Kalathur RKR, Sugiyama N, Tiede S, Ivanek R, et al. Gain fat—lose metastasis: converting invasive breast cancer cells into adipocytes inhibits cancer metastasis. *Cancer Cell*. 2019; 35(1):17–32. <https://doi.org/10.1016/j.ccell.2018.12.002> PMID: 30645973
35. Cheng Z, He Z, Cai Y, Zhang C, Fu G, Li H, et al. Conversion of hepatoma cells to hepatocyte-like cells by defined hepatocyte nuclear factors. *Cell research*. 2019; 29(2):124–135. <https://doi.org/10.1038/s41422-018-0111-x> PMID: 30560924
36. Cho KH, Lee S, Kim D, Shin D, Joo JI, Park SM. Cancer reversion, a renewed challenge in systems biology. *Current Opinion in Systems Biology*. 2017; 2:49–58. <https://doi.org/10.1016/j.coisb.2017.01.005>
37. Dekker J, Belmont AS, Guttman M, Leshyk VO, Lis JT, Lomvardas S, et al. The 4D nucleome project. *Nature*. 2017; 549(7671):219–226. <https://doi.org/10.1038/nature23884> PMID: 28905911
38. Nagano T, Lubling Y, Várnai C, Dudley C, Leung W, Baran Y, et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*. 2017; 547(7661):61. <https://doi.org/10.1038/nature23001> PMID: 28682332
39. Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, et al. A pathway for mitotic chromosome formation. *Science*. 2018; 359(6376):eaao6135. <https://doi.org/10.1126/science.aao6135> PMID: 29348367
40. Stadhouders R, Vidal E, Serra F, Di Stefano B, Le Dily F, Quilez J, et al. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nature genetics*. 2018; 50(2):238–249. <https://doi.org/10.1038/s41588-017-0030-7> PMID: 29335546
41. Di Stefano M, Stadhouders R, Farabella I, Castillo D, Serra F, Graf T, et al. Transcriptional activation during cell reprogramming correlates with the formation of 3D open chromatin hubs. *Nature communications*. 2020; 11(1):1–12. <https://doi.org/10.1038/s41467-020-16396-1> PMID: 32444798
42. Shinkai S, Nakagawa M, Sugawara T, Togashi Y, Ochiai H, Nakato R, et al. PHI-C: deciphering Hi-C data into polymer dynamics. *NAR genomics and bioinformatics*. 2020; 2(2):lqaa020. <https://doi.org/10.1093/nargab/lqaa020> PMID: 33575580
43. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 2013; 502(7469):59. <https://doi.org/10.1038/nature12593> PMID: 24067610
44. Abramo K, Valton AL, Venev SV, Ozadam H, Fox AN, Dekker J. A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *NATURE CELL BIOLOGY*. 2019; 21(11):1393–1402. <https://doi.org/10.1038/s41556-019-0406-2> PMID: 31685986
45. Chu X, Wang J. Microscopic chromosomal structural and dynamical origin of cell differentiation and reprogramming. *Advanced Science*. 2020; 7(20):2001572. <https://doi.org/10.1002/advs.202001572> PMID: 33101859
46. Wang S, Wolynes PG. Communication: Effective temperature and glassy dynamics of active matter. *Journal of Chemical Physics*. 2011; 135(5):051101–051101. <https://doi.org/10.1063/1.3624753> PMID: 21823683
47. Wang S, Wolynes P. Tensegrity and motor-driven effective interactions in a model cytoskeleton. *The Journal of chemical physics*. 2012; 136(14):145102. <https://doi.org/10.1063/1.3702583> PMID: 22502548
48. Di Pierro M, Potoyan DA, Wolynes PG, Onuchic JN. Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. *Proceedings of the National Academy of Sciences*. 2018; 115(30):7753–7758. <https://doi.org/10.1073/pnas.1806297115> PMID: 29987017
49. Liu L, Shi G, Thirumalai D, Hyeon C. Chain organization of human interphase chromosome determines the spatiotemporal dynamics of chromatin loci. *PLoS computational biology*. 2018; 14(12):e1006617. <https://doi.org/10.1371/journal.pcbi.1006617> PMID: 30507936
50. Bronstein I, Israel Y, Kepten E, Mai S, Shav-Tal Y, Barkai E, et al. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical review letters*. 2009; 103(1):018102. <https://doi.org/10.1103/PhysRevLett.103.018102> PMID: 19659180
51. Shinkai S, Nozaki T, Maeshima K, Togashi Y. Dynamic nucleosome movement provides structural information of topological chromatin domains in living human cells. *PLoS computational biology*. 2016; 12(10):e1005136. <https://doi.org/10.1371/journal.pcbi.1005136> PMID: 27764097
52. DeBerardinis RJ, Chandel NS. Fundamentals of cancer metabolism. *Science advances*. 2016; 2(5):e1600200. <https://doi.org/10.1126/sciadv.1600200> PMID: 27386546

53. Chu X, Wang J. Conformational state switching and pathways of chromosome dynamics in cell cycle. *Applied Physics Reviews*. 2020; 7(3):031403. <https://doi.org/10.1063/5.0007316> PMID: 32884608
54. Kim D, Rath O, Kolch W, Cho KH. A hidden oncogenic positive feedback loop caused by crosstalk between Wnt and ERK pathways. *Oncogene*. 2007; 26(31):4571–4579. <https://doi.org/10.1038/sj.onc.1210230> PMID: 17237813
55. Wu M, Liu L, Chan C. Identification of novel targets for breast cancer by exploring gene switches on a genome scale. *BMC genomics*. 2011; 12(1):547. <https://doi.org/10.1186/1471-2164-12-547> PMID: 22053771
56. Wang J, Zhang K, Xu L, Wang E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences*. 2011; 108(20):8257–8262. <https://doi.org/10.1073/pnas.1017017108> PMID: 21536909
57. Jia D, Jolly MK, Kulkarni P, Levine H. Phenotypic plasticity and cell fate decisions in cancer: insights from dynamical systems theory. *Cancers*. 2017; 9(7):70. <https://doi.org/10.3390/cancers9070070> PMID: 28640191
58. Ferrell JE, Machleder EM. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science*. 1998; 280(5365):895–898. <https://doi.org/10.1126/science.280.5365.895> PMID: 9572732
59. Xiong W, Ferrell JE. A positive-feedback-based bistable ‘memory module’ that governs a cell fate decision. *Nature*. 2003; 426(6965):460–465. <https://doi.org/10.1038/nature02089> PMID: 14647386
60. Fang X, Liu Q, Bohrer C, Hensel Z, Han W, Wang J, et al. Cell fate potentials and switching kinetics uncovered in a classic bistable genetic switch. *Nature communications*. 2018; 9(1):1–9. <https://doi.org/10.1038/s41467-018-05071-1>
61. Jiang Z, Tian L, Fang X, Zhang K, Liu Q, Dong Q, et al. The emergence of the two cell fates and their associated switching for a negative auto-regulating gene. *BMC biology*. 2019; 17(1):1–14. <https://doi.org/10.1186/s12915-019-0666-0> PMID: 31202264
62. Kanwal R, Gupta S. Epigenetic modifications in cancer. *Clinical genetics*. 2012; 81(4):303–311. <https://doi.org/10.1111/j.1399-0004.2011.01809.x> PMID: 22082348
63. Ellis L, Atadja PW, Johnstone RW. Epigenetics in cancer: targeting chromatin modifications. *Molecular cancer therapeutics*. 2009; 8(6):1409–1420. <https://doi.org/10.1158/1535-7163.MCT-08-0860> PMID: 19509247
64. Sasai M, Kawabata Y, Makishi K, Itoh K, Terada TP. Time scales in epigenetic dynamics and phenotypic heterogeneity of embryonic stem cells. *PLoS computational biology*. 2013; 9(12):e1003380. <https://doi.org/10.1371/journal.pcbi.1003380> PMID: 24348228
65. Ashwin S, Sasai M. Effects of collective histone state dynamics on epigenetic landscape and kinetics of cell reprogramming. *Scientific reports*. 2015; 5(1):1–12. <https://doi.org/10.1038/srep16746> PMID: 26581803
66. Zhang K, Sasai M, Wang J. Eddy current and coupled landscapes for nonadiabatic and nonequilibrium complex system dynamics. *Proceedings of the National Academy of Sciences*. 2013; 110(37):14930–14935. <https://doi.org/10.1073/pnas.1305604110> PMID: 23980160
67. Wang J. Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Advances in Physics*. 2015; 64(1):1–137. <https://doi.org/10.1080/00018732.2015.1037068>
68. Brock A, Krause S, Ingber DE. Control of cancer formation by intrinsic genetic noise and microenvironmental cues. *Nature Reviews Cancer*. 2015; 15(8):499–509. <https://doi.org/10.1038/nrc3959> PMID: 26156637
69. Tully JC. Molecular dynamics with electronic transitions. *The Journal of Chemical Physics*. 1990; 93(2):1061–1071. <https://doi.org/10.1063/1.459170>
70. Chakraborty A, Ay F. Identification of copy number variations and translocations in cancer cells from Hi-C data. *Bioinformatics*. 2018; 34(2):338–345. <https://doi.org/10.1093/bioinformatics/btx664> PMID: 29048467
71. Cesari A, Reißer S, Bussi G. Using the maximum entropy principle to combine simulations and solution experiments. *Computation*. 2018; 6(1):15. <https://doi.org/10.3390/computation6010015>
72. Zhang B, Wolynes PG. Topology, structures, and energy landscapes of human chromosomes. *Proceedings of the National Academy of Sciences*. 2015; 112(19):6062–6067. <https://doi.org/10.1073/pnas.1506257112> PMID: 25918364
73. Zhang B, Wolynes PG. Shape transitions and chiral symmetry breaking in the energy landscape of the mitotic chromosome. *Physical review letters*. 2016; 116(24):248101. <https://doi.org/10.1103/PhysRevLett.116.248101> PMID: 27367409
74. Lin X, Qi Y, Latham AP, Zhang B. Multiscale modeling of genome organization with maximum entropy optimization. *The Journal of Chemical Physics*. 2021; 155(1):010901. <https://doi.org/10.1063/5.0044150> PMID: 34241389

75. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome research*. 2009; 19(9):1639–1645. <https://doi.org/10.1101/gr.092759.109> PMID: 19541911
76. Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell stem cell*. 2010; 6(5):479–491. <https://doi.org/10.1016/j.stem.2010.03.018> PMID: 20452322
77. Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*. 2013; 153(5):1134–1148. <https://doi.org/10.1016/j.cell.2013.04.022> PMID: 23664764
78. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. 2015; 518(7539):331. <https://doi.org/10.1038/nature14222> PMID: 25693564
79. Clarke MF, Hass AT. Cancer stem cells. *Reviews in Cell Biology and Molecular Medicine*. 2006.
80. Nguyen LV, Vanner R, Dirks P, Eaves CJ. Cancer stem cells: an evolving concept. *Nature Reviews Cancer*. 2012; 12(2):133–143. <https://doi.org/10.1038/nrc3184> PMID: 22237392
81. Mirny LA. The fractal globule as a model of chromatin architecture in the cell. *Chromosome research*. 2011; 19(1):37–51. <https://doi.org/10.1007/s10577-010-9177-0> PMID: 21274616
82. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of chromosomal domains by loop extrusion. *Cell reports*. 2016; 15(9):2038–2049. <https://doi.org/10.1016/j.celrep.2016.04.085> PMID: 27210764
83. Losick R, Desplan C. Stochasticity and cell fate. *science*. 2008; 320(5872):65–68. <https://doi.org/10.1126/science.1147888> PMID: 18388284
84. Dima RI, Thirumalai D. Asymmetry in the shapes of folded and denatured states of proteins. *The Journal of Physical Chemistry B*. 2004; 108(21):6564–6570. <https://doi.org/10.1021/jp037128y>
85. Cho SS, Levy Y, Wolynes PG. P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes. *Proceedings of the National Academy of Sciences*. 2006; 103(3):586–591. <https://doi.org/10.1073/pnas.0509768103> PMID: 16407126
86. Misteli T. Beyond the sequence: cellular organization of genome function. *Cell*. 2007; 128(4):787–800. <https://doi.org/10.1016/j.cell.2007.01.028> PMID: 17320514
87. Bickmore WA, van Steensel B. Genome architecture: domain organization of interphase chromosomes. *Cell*. 2013; 152(6):1270–1284. <https://doi.org/10.1016/j.cell.2013.02.001> PMID: 23498936
88. Di Pierro M, Zhang B, Aiden EL, Wolynes PG, Onuchic JN. Transferable model for chromosome architecture. *Proceedings of the National Academy of Sciences*. 2016; 113(43):12168–12173. <https://doi.org/10.1073/pnas.1613607113> PMID: 27688758
89. Qi Y, Zhang B. Predicting three-dimensional genome organization with chromatin states. *PLoS computational biology*. 2019; 15(6):e1007024. <https://doi.org/10.1371/journal.pcbi.1007024> PMID: 31181064
90. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*. 2016; 529(7586):418. <https://doi.org/10.1038/nature16496> PMID: 26760202
91. Sha W, Moore J, Chen K, Lassaletta AD, Yi CS, Tyson JJ, et al. Hysteresis drives cell-cycle transitions in *Xenopus laevis* egg extracts. *Proceedings of the National Academy of Sciences*. 2003; 100(3):975–980. <https://doi.org/10.1073/pnas.0235349100> PMID: 12509509
92. Pomerening JR, Sontag ED, Ferrell JE. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nature cell biology*. 2003; 5(4):346–351. <https://doi.org/10.1038/ncb954> PMID: 12629549
93. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*. 2017; 544(7648):59. <https://doi.org/10.1038/nature21429> PMID: 28289288
94. Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*. 2017; 544(7648):110–114. <https://doi.org/10.1038/nature21711> PMID: 28355183
95. Bintu B, Mateo LJ, Su JH, Sinnott-Armstrong NA, Parker M, Kinrot S, et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*. 2018; 362(6413). <https://doi.org/10.1126/science.aau1783> PMID: 30361340
96. Xie WJ, Qi Y, Zhang B. Characterizing chromatin folding coordinate and landscape with deep learning. *PLoS computational biology*. 2020; 16(9):e1008262. <https://doi.org/10.1371/journal.pcbi.1008262> PMID: 32986691

97. Waddington CH. The strategy of the genes. Routledge; 1957.
98. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *cell*. 2006; 126(4):663–676. <https://doi.org/10.1016/j.cell.2006.07.024> PMID: 16904174
99. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *cell*. 2007; 131(5):861–872. <https://doi.org/10.1016/j.cell.2007.11.019> PMID: 18035408
100. Cacchiarelli D, Trapnell C, Ziller MJ, Soumillon M, Cesana M, Karnik R, et al. Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell*. 2015; 162(2):412–424. <https://doi.org/10.1016/j.cell.2015.06.016> PMID: 26186193
101. Francastel C, Schübeler D, Martin DI, Groudine M. Nuclear compartmentalization and gene activity. *Nature reviews Molecular cell biology*. 2000; 1(2):137–143. <https://doi.org/10.1038/35040083> PMID: 11253366
102. Arney KL, Fisher AG. Epigenetic aspects of differentiation. *Journal of Cell Science*. 2004; 117(19):4355–4363. <https://doi.org/10.1242/jcs.01390> PMID: 15331660
103. Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental cell*. 2006; 10(1):105–116. <https://doi.org/10.1016/j.devcel.2005.10.017> PMID: 16399082
104. Meshorer E, Misteli T. Chromatin in pluripotent embryonic stem cells and differentiation. *Nature reviews Molecular cell biology*. 2006; 7(7):540–546. <https://doi.org/10.1038/nrm1938> PMID: 16723974
105. Xu J, Ma H, Ma H, Jiang W, Mela CA, Duan M, et al. Super-resolution imaging reveals the evolution of higher-order chromatin folding in early carcinogenesis. *Nature communications*. 2020; 11(1):1–17. <https://doi.org/10.1038/s41467-020-15718-7> PMID: 32313005
106. Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M, Cosma MP. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*. 2015; 160(6):1145–1158. <https://doi.org/10.1016/j.cell.2015.01.054> PMID: 25768910
107. Takahashi K. Cellular reprogramming—lowering gravity on Waddington’s epigenetic landscape. *Journal of cell science*. 2012; 125(11):2553–2560. PMID: 22736045
108. Crompton JG, Clever D, Vizcardo R, Rao M, Restifo NP. Reprogramming antitumor immunity. *Trends in immunology*. 2014; 35(4):178–185. <https://doi.org/10.1016/j.it.2014.02.003> PMID: 24661777
109. Fang X, Wang J. Nonequilibrium Thermodynamics in Cell Biology: Extending Equilibrium Formalism to Cover Living Systems. *Annual review of biophysics*. 2020; 49:227–246. <https://doi.org/10.1146/annurev-biophys-121219-081656> PMID: 32375020
110. Miura H, Takahashi S, Poonperm R, Tanigawa A, Takebayashi Si, Hiratani I. Single-cell DNA replication profiling identifies spatiotemporal developmental dynamics of chromosome organization. *Nature genetics*. 2019; 51(9):1356–1368. <https://doi.org/10.1038/s41588-019-0474-z> PMID: 31406346
111. Klonisch T, Wiechec E, Hombach-Klonisch S, Ande SR, Wesselborg S, Schulze-Osthoff K, et al. Cancer stem cell markers in common cancers—therapeutic implications. *Trends in molecular medicine*. 2008; 14(10):450–460. <https://doi.org/10.1016/j.molmed.2008.08.003> PMID: 18775674
112. Shackleton M; Elsevier. Normal stem cells and cancer stem cells: similar and different. *Seminars in cancer biology*. 2010; 20(2):85–92. <https://doi.org/10.1016/j.semcancer.2010.04.002> PMID: 20435143
113. Bjerkvig R, Tysnes BB, Aboody KS, Najbauer J, Terzis A. The origin of the cancer stem cell: current controversies and new insights. *Nature Reviews Cancer*. 2005; 5(11):899–904. <https://doi.org/10.1038/nrc1740> PMID: 16327766
114. Kostriken R, Wedeen CJ. Engineered interphase chromosome loops guide intrachromosomal recombination. *The EMBO journal*. 2001; 20(11):2907–2913. <https://doi.org/10.1093/emboj/20.11.2907> PMID: 11387223
115. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*. 2012; 149(6):1233–1244. <https://doi.org/10.1016/j.cell.2012.03.051> PMID: 22682246
116. Morgan SL, Mariano NC, Bermudez A, Arruda NL, Wu F, Luo Y, et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nature communications*. 2017; 8(1):1–9. <https://doi.org/10.1038/ncomms15993> PMID: 28703221
117. Kim JH, Rege M, Valeri J, Dunagin MC, Metzger A, Titus KR, et al. LADL: light-activated dynamic looping for endogenous gene expression control. *Nature methods*. 2019; 16(7):633. <https://doi.org/10.1038/s41592-019-0436-5> PMID: 31235883
118. Du M, Zou F, Yan Y, Bai L. Chemically Induced Chromosomal Interaction (CICI): A New Tool to Study Chromosome Dynamics and Its Biological Roles. *bioRxiv*. 2020.

119. Liu NQ, de Wit E. A transient absence of SMC-mediated loops after mitosis. *Nature cell biology*. 2019; 21(11):1303–1304. <https://doi.org/10.1038/s41556-019-0421-3> PMID: 31685985
120. Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, et al. Identification of the tumour transition states occurring during EMT. *Nature*. 2018; 556(7702):463–468. <https://doi.org/10.1038/s41586-018-0040-3> PMID: 29670281
121. Nieto MA, Huang RYJ, Jackson RA, Thiery JP. EMT: 2016. *Cell*. 2016; 166(1):21–45. <https://doi.org/10.1016/j.cell.2016.06.028> PMID: 27368099
122. Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E. MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proceedings of the National Academy of Sciences*. 2013; 110(45):18144–18149. <https://doi.org/10.1073/pnas.1318192110> PMID: 24154725
123. Bocci F, Jolly MK, George JT, Levine H, Onuchic JN. A mechanism-based computational model to capture the interconnections among epithelial–mesenchymal transition, cancer stem cells and Notch–Jagged signaling. *Oncotarget*. 2018; 9(52):29906. <https://doi.org/10.18632/oncotarget.25692>
124. Strauss R, Li ZY, Liu Y, Beyer I, Persson J, Sova P, et al. Analysis of epithelial and mesenchymal markers in ovarian cancer reveals phenotypic heterogeneity and plasticity. *PloS one*. 2011; 6(1): e16186. <https://doi.org/10.1371/journal.pone.0016186> PMID: 21264259
125. Jolly MK, Huang B, Lu M, Mani SA, Levine H, Ben-Jacob E. Towards elucidating the connection between epithelial–mesenchymal transitions and stemness. *Journal of the Royal Society Interface*. 2014; 11(101):20140962. <https://doi.org/10.1098/rsif.2014.0962> PMID: 25339690
126. Tripathi S, Jolly MK, Woodward WA, Levine H, Deem MW. Analysis of hierarchical organization in gene expression networks reveals underlying principles of collective tumor cell dissemination and metastatic aggressiveness of inflammatory breast cancer. *Frontiers in oncology*. 2018; 8:244. <https://doi.org/10.3389/fonc.2018.00244> PMID: 30023340
127. Ruscetti M, Quach B, Dadashian EL, Mulholland DJ, Wu H. Tracking and functional characterization of epithelial–mesenchymal transition and mesenchymal tumor cells during prostate cancer metastasis. *Cancer research*. 2015; 75(13):2749–2759. <https://doi.org/10.1158/0008-5472.CAN-14-3476> PMID: 25948589
128. Ombrato L, Malanchi I. The EMT universe: space between cancer cell dissemination and metastasis initiation. *Critical Reviews in Oncogenesis*. 2014; 19(5). PMID: 25404150
129. Pastushenko I, Blanpain C. EMT transition states during tumor progression and metastasis. *Trends in cell biology*. 2019; 29(3):212–226. <https://doi.org/10.1016/j.tcb.2018.12.001> PMID: 30594349
130. Onuchic JN, Wolynes PG. Classical and quantum pictures of reaction dynamics in condensed matter: Resonances, dephasing, and all that. *The Journal of Physical Chemistry*. 1988; 92(23):6495–6503. <https://doi.org/10.1021/j100334a007>
131. Whitford PC, Miyashita O, Levy Y, Onuchic JN. Conformational transitions of adenylate kinase: switching by cracking. *Journal of molecular biology*. 2007; 366(5):1661–1671. <https://doi.org/10.1016/j.jmb.2006.11.085> PMID: 17217965
132. Wang Y, Gan L, Wang E, Wang J. Exploring the dynamic functional landscape of adenylate kinase modulated by substrates. *Journal of chemical theory and computation*. 2012; 9(1):84–95. <https://doi.org/10.1021/ct300720s> PMID: 26589012
133. Chu X, Liu F, Maxwell BA, Wang Y, Suo Z, Wang H, et al. Dynamic conformational change regulates the protein–DNA recognition: an investigation on binding of a Y-family polymerase to its target DNA. *PLoS computational biology*. 2014; 10(9):e1003804. <https://doi.org/10.1371/journal.pcbi.1003804> PMID: 25188490
134. Chen C, Wang J. A physical mechanism of cancer heterogeneity. *Scientific reports*. 2016; 6(1):1–9. <https://doi.org/10.1038/srep20679> PMID: 26854017
135. Zhang K, Wang J. Exploring the underlying mechanisms of the coupling between cell differentiation and cell cycle. *The Journal of Physical Chemistry B*. 2019; 123(16):3490–3498. <https://doi.org/10.1021/acs.jpcc.9b00509> PMID: 30933510
136. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature Reviews Genetics*. 2016; 17(5):284–299. <https://doi.org/10.1038/nrg.2016.13> PMID: 26972587
137. Rosa A, Everaers R. Structure and dynamics of interphase chromosomes. *PLoS computational biology*. 2008; 4(8):e1000153. <https://doi.org/10.1371/journal.pcbi.1000153> PMID: 18725929
138. Wang JC. Recent studies of DNA topoisomerases. *Biochimica et Biophysica Acta (BBA)—Gene Structure and Expression*. 1987; 909(1):1–9. [https://doi.org/10.1016/0167-4781\(87\)90040-6](https://doi.org/10.1016/0167-4781(87)90040-6)
139. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, et al. Organization of the mitotic chromosome. *Science*. 2013; 342(6161):948–953. <https://doi.org/10.1126/science.1236083> PMID: 24200812

140. Swedlow JR, Sedat JW, Agard DA. Multiple chromosomal populations of topoisomerase II detected in vivo by time-lapse, three-dimensional wide-field microscopy. *Cell*. 1993; 73(1):97–108. [https://doi.org/10.1016/0092-8674\(93\)90163-K](https://doi.org/10.1016/0092-8674(93)90163-K) PMID: 8384932
141. Poljak L, Käs E. Resolving the role of topoisomerase II in chromatin structure and function. *Trends in cell biology*. 1995; 5(9):348–354. [https://doi.org/10.1016/S0962-8924\(00\)89068-6](https://doi.org/10.1016/S0962-8924(00)89068-6) PMID: 14732077
142. Lee JH, Berger JM. Cell cycle-dependent control and roles of DNA topoisomerase II. *Genes*. 2019; 10(11):859. <https://doi.org/10.3390/genes10110859> PMID: 31671531
143. Racko D, Benedetti F, Goundaroulis D, Stasiak A. Chromatin loop extrusion and chromatin unknotting. *Polymers*. 2018; 10(10):1126. <https://doi.org/10.3390/polym10101126> PMID: 30961051