Article

Dynamics and Pathways of Chromosome Structural Organizations during Cell Transdifferentiation

Xiakun Chu and Jin Wang*

Read Online

| ACCESS | Metrics & More | Article Recommendations | Supporting Information |

ABSTRACT: Direct conversion of one differentiated cell type into another is defined as cell transdifferentiation. In avoidance of forming pluripotency, cell transdifferentiation can reduce the potential risk of tumorigenicity, thus offering significant advantages over cell reprogramming in clinical applications. Until now, the mechanism of cell transdifferentiation is still largely unknown. It has been well recognized that cell transdifferentiation is determined by the underlying gene expression regulation, which relies on the accurate adaptation of the chromosome structure. To dissect the transdifferentiation at the molecular level, we develop a non-equilibrium landscape-switching model to investigate the chromosome structural dynamics during the state transitions between the human fibroblast and neuron cells. We uncover the high irreversibility of the transdifferentiation at the local chromosome structural ranges, where the topologically associating domains form. In contrast, the pathways in the two opposite directions of the transdifferentiation projected onto the chromosome compartment profiles are highly overlapped, indicating that the reversibility vanishes at the long-range chromosome structures. By calculating the contact strengths in the chromosome at the states along the paths, we observe strengthening contacts in compartment A concomitant with weakening contacts in compartment B at the early stages of the transdifferentiation. This further leads to adapting contacts toward the ones at the embryonic stem cell. In light of the intimate structure−function relationship at the chromosomal level, we suggest an increase of "stemness" during the transdifferentiation. In addition, we find that the neuron progenitor cell (NPC), a cell developmental state, is located on the transdifferentiation pathways projected onto the long-range chromosome contacts. The findings are consistent with the previous single-cell RNA sequencing experiment, where the NPC-like cell states were observed during the direct conversion of the fibroblast to neuron cells. Thus, we offer a promising microscopic and physical approach to study the cell transdifferentiation mechanism from the chromosome structural perspective.

KEYWORDS: Cell-state transition, Chromosome dynamics, Energy landscape, Nonequilibrium model, Cell Stemness
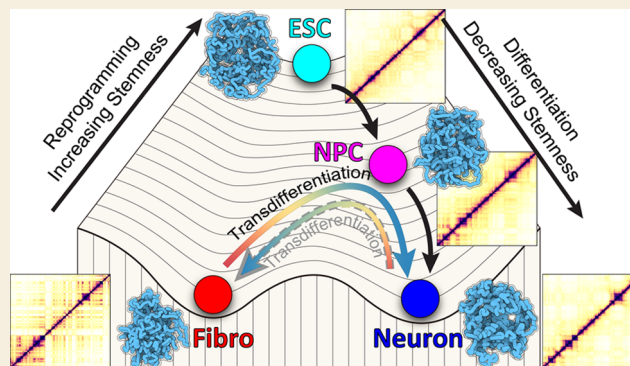
Cell transdifferentiation corresponds to the direct conversion of one differentiated cell type into another. Similar to cell reprogramming, where the differentiated cell is induced to the pluripotent cell by a defined cocktail of transcription factors,[1,2] cell transdifferentiation can be realized by a combination of the lineage-specific transcription factors.[3−5] These transcription factors activate the genes needed for the cell fate determination, thus dictating the cell transdifferentiation. The faithful switching of the cellular phenotype is further established by the epigenetic modifications, which regulate the patterns of the gene expressions through the modulations of the DNA accessibility and the chromatin structure.[6,7] It has been acknowledged that transdifferentiation is a safer process for a differentiated cell to gain a new function than reprogramming in the potential clinical applications, as the cell can reduce the potential risk of cancer development by avoiding the pluripotency during the state transition process.[8] Despite the numerous successes in

implementing the transdifferentiation for various cell types,[9] the mechanistic understanding of the process is still quite limited.

Recent experimental studies used the single-cell RNA-sequencing (RNA-seq) technologies to dissect the cell transdifferentiation.[10−12] The transcriptome data of the states on the transdifferentiation path can be subsequently compared with the ones found along the cell developmental path in order to explore the relationship between transdifferentiation and differentiation.[10] The results provided useful information to
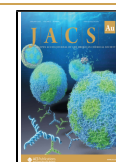
characterize the transdifferentiation pathways and identify the intermediate states from the transcriptome perspective.[13] It has been well recognized that the transcriptional regulation is strongly dependent on the three-dimensional chromosome organization.[14−19] In this respect, cell transdifferentiation, which is determined by the underlying gene regulation network, is closely related to the chromosome structural changes during the transition. To quantitatively understand the cell transdifferentiation, it is critical to obtain the molecular-level picture of how the chromosome dynamically arranges its structure toward that in the destined cell. However, this is a very challenging task as the current experiments are intrinsically limited by the temporal or spatial scale and resolution.[20]

Nowadays, the high-resolution chromosome structural determination relies on the Hi-C technique,[21,22] which measures the probability of the spatial contact formed by the loci throughout the genome. The analysis of the Hi-C contact map revealed that the chromosome is hierarchically organized. At the submegabase scales, the chromosome forms the topologically associating domains (TADs),[23−25] which correspond to the square blocks of elevated interaction probability centered along the diagonal of the contact map. Within a TAD, the loci interact with each other more strongly than with the loci outside the TAD. At large scale (>5 Mb), the chromosome segregates into two spatial regions with specific preferential long-range interactions, leading to the plaid pattern on the Hi-C contact map.[21,22,26] These two mutually excluded regions, referred to as compartment A and B, correspond to the active euchromatin and inactive heterochromatin, respectively.[21,22,26] The Hi-C data on the human embryonic stem cell (ESC) and its derived multilineage cells at the early embryo developmental stages indicate the extensive switching of compartment A to B, accompanied by the changes of epigenetic modifications upon the differentiation.[27,28] Recent studies on different cell state transitions established the intimate relationship between the chromosome compartment switching and the transcriptional changes.[29,30] These findings suggest that the large-scale chromosome structural changes are indispensable and needed for the cell state transitions.

The Hi-C measurements on the initial and final states do not provide the dynamics of the chromosome structural changes during the cell state transition. Recently, the time-series 4D Hi-C experiments studied the mechanisms of transitions by measuring the contact maps at discrete time points during the processes.[31−36] However, there is still a lack of connections between the neighboring data, so the transition pathway inference strongly depends on the temporal resolution of the data. Meanwhile, given the highly stochastic nature of the cell dynamics, the data at each time point inevitably contain the temporal heterogeneity of cells at different transition stages. This impedes the precise characterization of chromosome structures at the intermediate states during the cell transitions.[37,38]

Here, we used a nonequilibrium landscape-switching simulation approach to study the chromosome structural transition between the human fibroblast (Fibro) and Neuron cells. The transdifferentiation from the Fibro to Neuron cells has been extensively realized in experiments.[3−5,39,40] However, the mechanisms of the transitions remain elusive. We quantified the transdifferentiation pathways by projecting the trajectories onto different chromosome structural characteristics, including TADs, compartments, and contacts at different types and ranges. We identified the nonoverlapped pathways

for the two transitions from the Fibro to Neuron cells, and from the Neuron to the Fibro cells, leading to the irreversibility of the transdifferentiation. The degree of the irreversibility in transdifferentiation is negatively correlated with the ranges of the chromosome structures. However, because of the conserved boundaries of TADs during cell transdifferentiation, we speculate that the high irreversibility of structural reorganization at the TAD level may have little effects on distinguishing the mechanistic differences between the forward and reverse transitions. We observed that the strengths of chromosome contacts at the intermediate states during both of the transitions adapt toward the ones at the ESC. This is an implication of increasing "stemness" during the transdifferentiation. We found that the NPC is located at the transdifferentiation pathways projected onto the long-range contacts. The observations are consistent with the previous experiment,[10] where the NPC-like intermediate cell states were uncovered during the transition from the Fibro to Neuron cells. We launched a promising theoretical framework to study the chromosome structural dynamics during the transdifferentiation and provided extensive predictions that can be assessed by future experiments.

■ RESULTS

Chromosome Structural Transitions during the Transdifferentiation between the Fibro and Neuron Cells

We used molecular dynamics (MD) simulations to study the chromosome structural dynamics during cell transdifferentiation between the Fibro and Neuron cells. It has been demonstrated that an effective equilibrium energy landscape can accurately describe the nonequilibrium chromosome system at one cell state in favorable circumstances.[41−45] Thus, the chromosome dynamical transitions during cell transdifferentiation, which corresponds to the transitions between two cell states, are determined by the connections between two effective equilibrium energy landscapes. As the cell state transition processes are governed by the principles of nonequilibrium dynamics facilitated often by the energy pumping processes such as ATP hydrolysis, the interlandscape connections should inherently have a nonequilibrium nature. In our previous work,[46] we developed a nonequilibrium landscape-switching model that is capable of simulating the slow and large-scale chromosome structural transitions during various cell processes, including the cell cycle,[46] cell differentiation and reprogramming,[47] as well as the cancer cell formation.[48,49] The landscape-switching model utilizes an approximation to connect the two landscapes at the initial and final cell states through an instantaneous energy excitation, which triggers the transition from one cell state to the other. As an external energy input, the energy excitation breaks the detailed balance of the system and drives the system out of equilibrium. Therefore, the model not only captures the essence of the nonequilibrium nature of the cell-state transition dynamics, but also significantly expedites the cell transformation processes, which are often too slow to be investigated by the conventional MD approaches.

The rationale of the landscape-switching model can be explained by the following two facts. First, from the biological perspective, the cell state transition process exhibits the switching dynamics. Various experimental evidence suggested that the cell developmental system shows multistability and the transition between the multiple stable cell states undergoes like
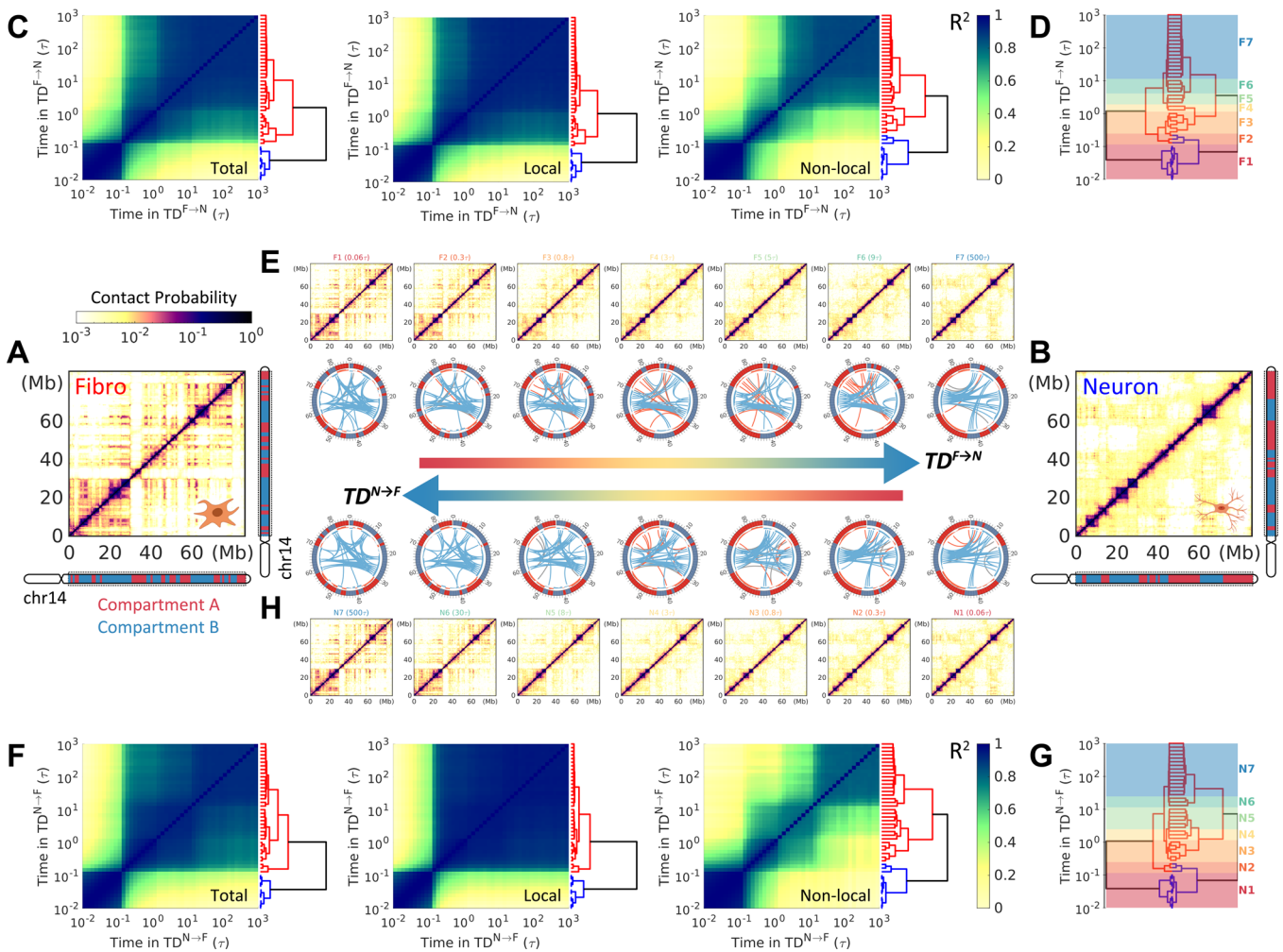
Figure 1. Chromosome structural dynamics during cell transdifferentiation from the Fibro to Neuron cells (upper) and from the Neuron to Fibro cells (lower). The experimental Hi-C contact maps and the ideograms of chromosome 14 (20.5−106.1 Mb) annotated by compartment status for (A) the Fibro cell and (B) the Neuron cell. (C) Hierarchical clustering of the chromosome contact maps between the transdifferentiation time $t = I, J$ for total contacts (left), contacts in local range ($\leqslant$2 Mb, middle), and contacts in nonlocal range (>2 Mb, right). The comparison between the contact maps at $t = I$ and J is made by calculating the coefficient of determination $R^2(I, J)$ between the contact probability $P_{i,j}$ formed by the chromosomal loci "i" and "j" at time $t = I$ and J: $R^2(I, J) = 1 - \sum_{i,j} [P_{i,j}(t = I) - P_{i,j}(t = J)]^2$ $\bigg/ \sqrt{\sum_{i,j} [P_{i,j}(t = I) - \langle P_{i,j}(t) \rangle]^2 \sum_{i,j} [P_{i,j}(t = J) - \langle P_{i,j}(t) \rangle]^2}$. $R^2(I, J)$, which measures the similarity of the two contact maps at time $t = I$ and J ($R^2(I, J) = 1$ corresponds to the identical Hi-C contact maps with $P_{i,j}(t=I) = P_{i,j}(t=J)$ and the deviation of $R^2(I, J)$ from 1 indicates the degree of difference between these two contact maps), is shown by 2D plots. (D) The 7 states ("F1" to "F7") that represent the typical chromosome contact formation during the transdifferentiation from the Fibro to Neuron cells, based on the comparison of the local and nonlocal dendrogram cluster trees. (E) The chromosome structural transition of the 7 states denoted in (D) during the transdifferentiation from the Fibro to Neuron cells shown in terms of the contact maps (upper) and circle plots (lower). In circle plots, the red and blue bands indicate the chromosomal loci in the compartment A and B, respectively. The compartment profiles are further shown as histograms near to the band plots. The connections show the long-range (>5 Mb) interactions. The interactions are identified by the enhanced contact probability P_{obs}/P_{exp}, where P_{obs} and P_{exp} are the observed and expected contact probability, respectively.[21] The red, blue, and gray lines indicate the interactions between the chromosomal loci within compartment A, within compartment B, and between compartments A and B. Only the top 50 weighted contacts are shown for better visualization. (F−H) The same as (C−E), but for the transdifferentiation from the Neuron to Fibro cells. The representative 7 states during the transdifferentiation from the Neuron to Fibro cells are denoted as "N1" to "N7".

a switch.[50−53] Theoretical studies showed that the simple circuitry model of switching between two distinct gene states is capable of capturing many characteristics of the cell state transition.[54−57] Second, from the physical perspective, the cell state transition process should be described by the nonadiabatic nonequilibrium dynamics. We have demonstrated that the degree of the adiabaticity of the system depends on the time scales of the processes involved and a faster (slower) intralandscape motion than the interlandscape hopping leads

to a nonadiabatic (adiabatic) process.[58,59] In reality, a cell can stably reside in the terminally differentiated cell state and the transdifferentiation cannot occur spontaneously. This indicates a slower time scale for the cell waiting for transdifferentiation (interlandscape dynamics) than the cell relaxing within one stable state (intralandscape dynamics), resulting in the nonadiabatic process.

Here, we applied the landscape-switching model to study the chromosome structural dynamics during cell transdifferentia-
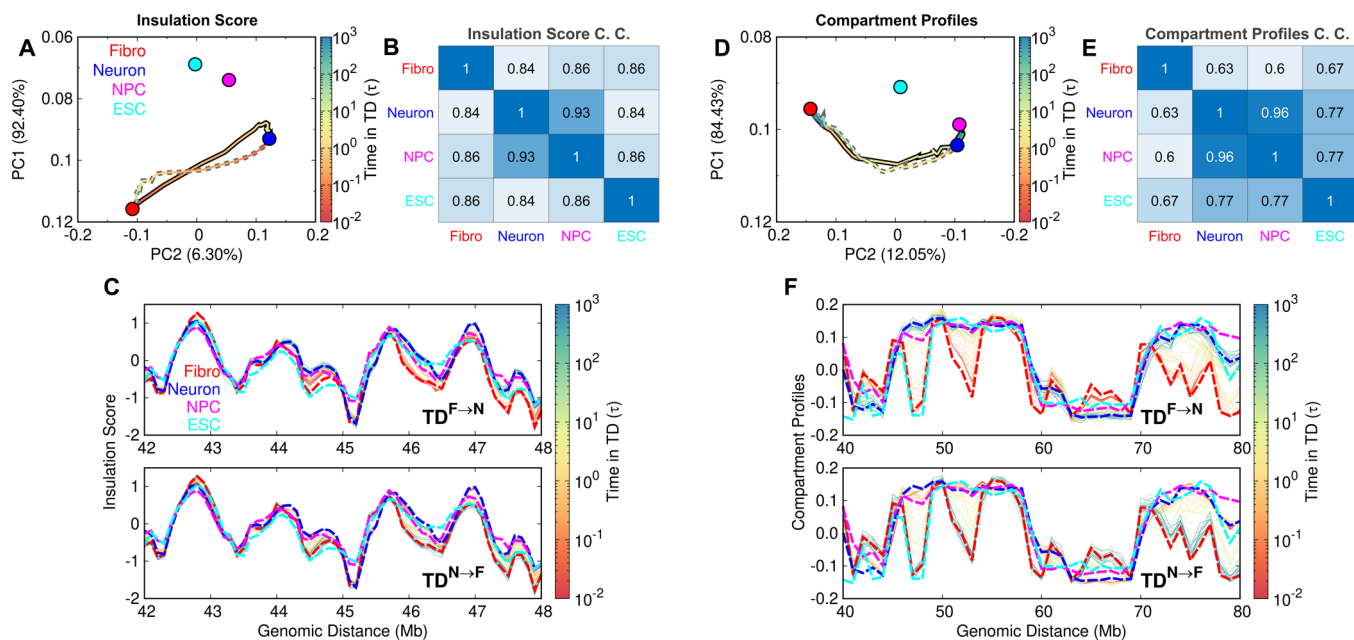
Figure 2. Chromosome structural transition pathways in terms of the TADs and compartments during cell transdifferentiation. (A) PCA of the insulation score profiles during the transdifferentiation. The solid and dashed lines are the PCA trajectories of the transdifferentiation projected onto the first two PCs from the Fibro to Neuron cells, and from the Neuron to Fibro cells, respectively. The Hi-C data of the Fibro cell, Neuron cell, NPC, and ESC are plotted as circles. (B) Correlation coefficients of the insulation score profiles between any pair among the Fibro cell, Neuron cell, NPC and ESC. (C) Evolution of the insulation score profiles during the transdifferentiation. Hi-C data at the Fibro cell, Neuron cell, NPC, and ESC are shown in dashed lines. (D−F) Same as (A−C) but for the compartment profiles.

tion. First, we used the Hi-C data as restraints to obtain the landscapes for the chromosome dynamics at the Fibro and Neuron cells, separately. In practice, we performed data-driven MD simulations guided by the maximum entropy principle to obtain the simulated chromosome contact maps that closely resemble the experimental Hi-C data (Figures S1−S4).[60] The prominent outcomes of the restrained MD simulations are two potentials. Further analyses on the kinetics of the chromosome systems under these potentials at one cell state show good agreements with numerous experiments in many faces, including spatial coherence, viscoelasticity, and the subdiffusive behavior of the motion in the chromosome.[61,62] These features indicate that these potentials can reproduce both the thermodynamic and kinetic properties of the chromosomes in one cell state; thus, they correspond to the energy landscapes of the chromosome dynamics at the Fibro and Neuron cells, respectively. Then, the landscape-switching model was implemented to establish the connections between these two landscapes, leading to the nonequilibrium cell-state transitions between the Fibro and Neuron cells, namely, the cell transdifferentiation (see Materials and Methods).

We focused on the long arm of human chromosome 14 with a range of 20.5−106.1 Mb. The Hi-C maps of the chromosomes at the Fibro and Neuron cells show significant differences (Figure 1A and B), in particular, at the regions far from the diagonal. This implies that there are large-scale chromosome structural changes during the transdifferentiation between these two cell states. We first studied the transition process from the Fibro to Neuron cells (denoted as $TD^{F \to N}$) and obtained the Hi-C-like contact probability maps along with the transition time. We compared the similarity of the contact probability maps between any of two time points during the transition (measured by the coefficient of determination $R^2(I, J)$, where I and J correspond to two time

points $t = I$ and J) for different chromosomal loci interacting ranges (Figure 1C). We found that the local and nonlocal chromosome contacts organize with different time scales, as the $R^2(I, J)$ plots show different patterns. Stepwise comparison of the local and nonlocal dendrogram clustering trees can further simplify the transition process into 7 representative states ("F1" to "F7") (Figure 1D). The chromosomes within one state show high structural similarity at both local and nonlocal ranges. The contact maps of these 7 states indicate the chromosome structural adaptation during the transdifferentiation, where the changes of long-range contacts are more significant than short-range contacts (Figure 1E, upper). The changes of long-range contacts further lead to the compartment state switching associated with the rearrangements of the interactions based on the compartment state during the transition (Figure 1E, lower). It is worth noting that although the contacts within the compartment B at both the Fibro and Neuron cells are highly weighted (blue lines in Figure 1E, lower), the chromosomes at the states on the transition path tend to increase the weights of contacts within the compartment A (red lines in Figure 1E, lower). The results indicate that the chromosome strengthens the long-range interactions within the active compartment A regions, followed by weakening them during the transdifferentiation.

We then studied the chromosome structural dynamics during the transdifferentiation from the Neuron to Fibro cells (denoted as $TD^{N \to F}$). Overall, the chromosome structural evolution during the $TD^{N \to F}$ shows the same behavior as that during the $TD^{F \to N}$, where the local and nonlocal contacts organize the changes through different time scales (Figure 1F). Similarly, we used 7 representative states ("N1" to "N7") to simplify the transition process (Figure 1G). The representative states during the $TD^{N \to F}$, except the initial and final states, show different contact maps with the states during the $TD^{F \to N}$
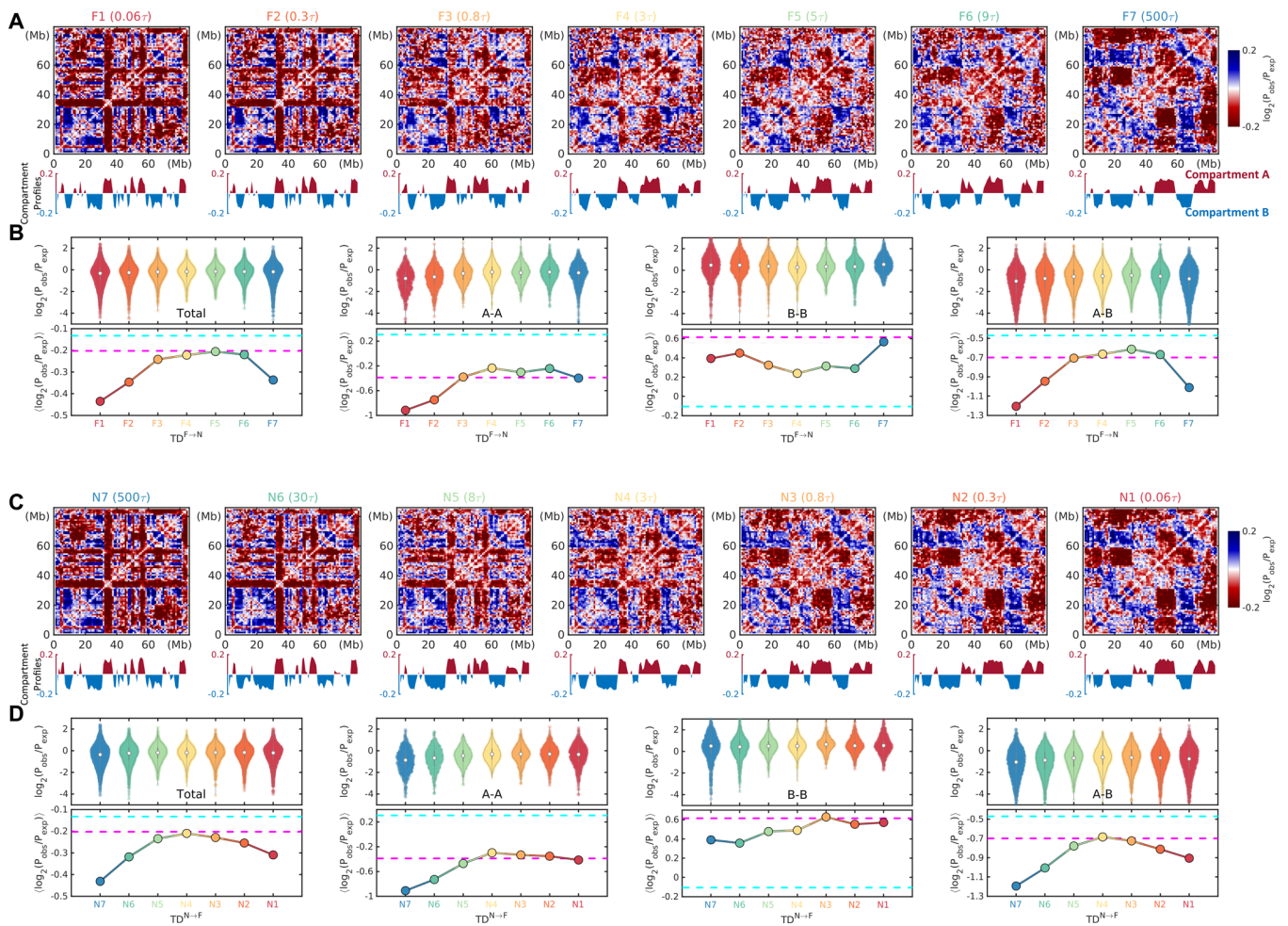
Figure 3. Evolution of the chromosome enhanced contact probability during cell transdifferentiation. (A) Chromosome enhanced contact maps (upper) and compartment profiles (lower) of the 7 states during the transdifferentiation from the Fibro to Neuron cells. (B) Statistics of enhanced contact probability of the 7 states during the transdifferentiation from the Fibro to Neuron cells. The violin plots (upper) and average (lower) of the contact probability are based on the interaction loci pairs from (left to right) total, within compartment A, within compartment B, and between compartments A and B. (C, D) The same as (A) and (B), but for the transdifferentiation from the Neuron to Fibro cells.

(Figure 1H). This indicates that these two transitions in the opposite directions do not undergo the same route, leading to the irreversibility of the transdifferentiation. Despite the different contact maps between the representative states during these two transitions, the long-range contacts within compartment A in the chromosome are enhanced at the intermediates during the $TD^{N \to F}$, similar to the observations during the $TD^{F \to N}$. The results suggest that strengthening the contacts within the active compartment A regions in the chromosome is prevalent during the transdifferentiation regardless of the transition direction.

Dynamical Changes of TADs and Compartments during the Transdifferentiation between the Fibro and Neuron Cells

Since TADs have been deemed as the structural units of the chromosome,[15,63,64] we focused on how the TADs structurally change during the transdifferentiation. We used the insulation score, which was previously introduced by Crane et al.,[65] to quantify the structure of the chromosome at the local ranges, where TADs form. The insulation score is calculated based on the chromosome contact map and presented in terms of a one-dimensional profile. Every locus in the chromosome has an insulation score that reflects the aggregate of the interactions

occurring across this locus at a local range. Therefore, the local minima of the insulation score profile denote the loci of high insulation that can potentially be classified as the TAD boundaries. The insulation score provides a quantitative structural measurement of the TADs and the associated boundaries. It was used here to capture the TAD formation. We performed the principal component analysis (PCA) on the evolution of the insulation score profiles along with the transition time during the transdifferentiation (Figure S5) and projected the trajectories onto the first two principal components (PCs) (Figure 2A). The two pathways that show the dynamical changes of the local chromosome structures during the forward and reverse transitions do not overlap. This indicates a high degree of irreversibility in TAD structural dynamics during the transdifferentiation. Furthermore, we observed that the ESC and NPC are located far away from both pathways. The findings suggest that TADs do not form the structures in either ESC or NPC during the transdifferentiation between the Fibro and Neuron cells.

Extensive analyses on Hi-C data have revealed that the structures of TADs are highly conserved across different cell types.[23,28,66] We also found that there are strong correlations among the insulation score profiles of the Fibro cell, Neuron
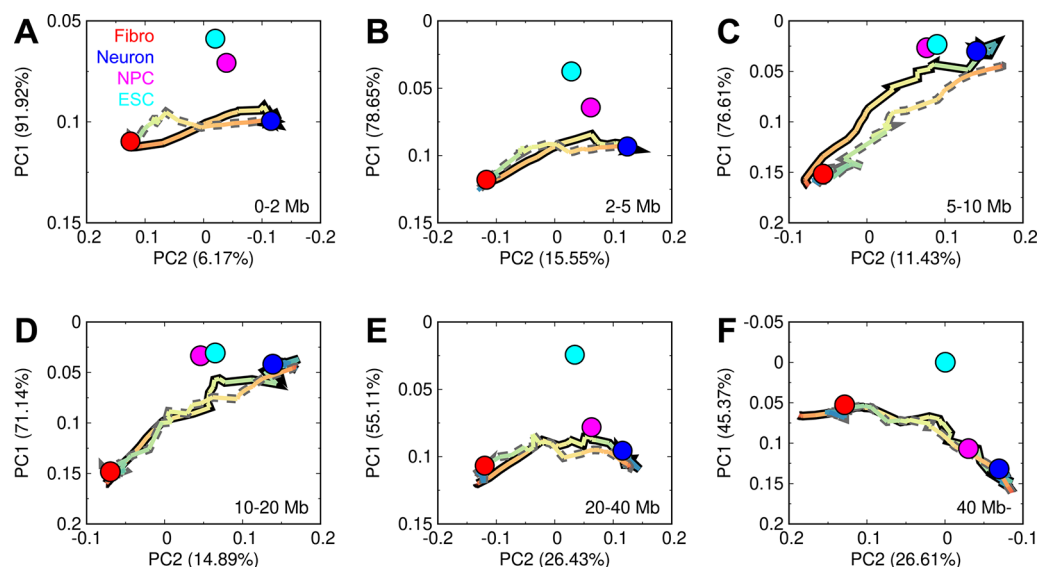
Figure 4. Chromosome structural transition pathways in terms of the chromosome contact probability. The PCA plots of the enhanced contact probability $\log_2(P_{\text{obs}}/P_{\text{exp}})$ evolving with the transition time are projected onto the first two PCs. The pathways are divided into different contact ranges based on the contact probability as (A) 0−2 Mb, (B) 2−5 Mb, (C) 5−10 Mb, (D) 10−20 Mb, (E) 20−40 Mb, and (F) 40 Mb−.

cell, NPC, and ESC (Figure 2B). This implies that the changes of TAD structures during the transdifferentiation may not be significant. We collected all the insulation score profiles during the transitions and compared them with the ones from the experimental Hi-C data of the Fibro cell, Neuron cell, NPC, and ESC (Figure 2C). All the insulation score profiles show very similar trends during the transitions and do not deviate significantly from the ones obtained through the Hi-C data of the 4 cell states. The results suggest that the large-scale chromosome structural changes during the transdifferentiation should mainly occur at the ranges beyond the TADs.

We then studied the changes of the compartment, which is a long-range structural motif in the chromosome, during the transdifferentiation. We used the compartment profile to quantify the structural compartmentalization in the chromosome. The compartment profile is the first PC (PC1) of the enhanced contact probability and its direction is assigned based on the gene density. A positive (negative) value in the compartment profile dictates the compartment A (B) chromosome region, which further correlates with the euchromatin (heterochromatin) state.[21,22] The compartment A was found to be associated with open chromatin and the compartment B with closed chromatin. Interactions are largely constrained to occur between chromosome regions belonging to the same compartment status, which is denoted by the compartment profile. Therefore, the compartment profile provides the quantitative measurement of the strength of the chromosome compartmentalization and was used here to capture the compartment formation. We performed the PCA of the compartment profiles evolving with the transition time and projected the trajectories onto the first two PCs (Figure 2D). Interestingly, we observed that the pathways of the compartment changes share similar routes for the two transitions. The ESC is still located far away from the pathways, indicating that the ESC does not lay in the transdifferentiation processes from the compartment perspective. On the other hand, we found that the NPC is very close to the Neuron cell, due to the strong correlation of compartment profiles between the NPC and Neuron cell (Figure 2E). Thus,

it cannot be determined whether the NPC is on the transition pathways based on the compartment changes. Different from the TADs, the trajectories of the compartment profiles show significant changes with compartment state switching between A and B during the transition (Figure 2F). Therefore, our results show that the chromosome structural changes during transdifferentiation are mainly associated with the compartment changes, which undergo reversible pathways during the forward and reverse transitions.

Chromosome Structural Transitions during the Transdifferentiation between the Fibro and Neuron Cells at Different Contact Types and Ranges

As shown in Figure 1, the irreversibility of the transdifferentiation can be observed at the adaptation of the long-range chromosome contacts. This indicates that the compartment profile, which is the one-dimensional projection (PC1) of the enhanced contact map, is insufficient to accurately describe the chromosome structural dynamics during the transdifferentiation. Therefore, we present the 2D maps of the enhanced contact probability at the representative 7 states during the $TD^{N\rightarrow F}$ and $TD^{F\rightarrow N}$ (Figure 3A and C). It is shown that the contact maps of the states, except the initial and final states, at the forward direction have clear differences from those at the reverse direction (Figure S7A), though they have highly similar compartment profiles (Figure S7B).

To see how the chromosome contacts rearrange during the transdifferentiation, we categorized the contacts based on the compartment state of the involved loci. For the whole contact sets ("Total" type), the contact probability shows an initial increase followed by a decrease during both transitions (Figure 3B and D). This nonmonotonic change of contact probability during the transdifferentiation was also observed for the interacting loci within the compartment A ("A-A" type) and between the compartment A and B ("A-B" type). Interestingly, the changes of contact probability involved by the loci within the compartment B ("B-B" type) show distinct behaviors during the $TD^{N\rightarrow F}$ and $TD^{F\rightarrow N}$. During the $TD^{F\rightarrow N}$, the probability of the "B-B" contact type decreases first toward the value at the ESC, followed by a sharp increase from the state

"F6" to "F7". In contrast, the probability of the "B-B" contact type during the $TD^{N \to F}$ undergoes a monotonic decrease. It is worth noting that the chromosome in the ESC shows the highest probabilities of the "A-A" and "A-B" contact types and the lowest probability of the "B-B" contact types among the 4 cell states (Figure S8). Therefore, the initial rearrangements of the chromosome structures with strengthening the contacts of the "A-A" and "A-B" types and weakening the contacts of the "B-B" type during the transdifferentiation potentially adapt the chromosome contacts toward the ones formed at the ESC.

To see how the contacts in the chromosome organize at different ranges during the transdifferentiation, we projected the trajectories of the enhanced contact probability onto the first two PCs at the different contact ranges during the two transitions (Figure 4). We found that the pathways of the forward and reverse transitions are more likely to be irreversible at short contact ranges, compared to the ones at long contact ranges. At the local contact range within 2 Mb, where TADs form (Figure 4A), the ESC and NPC are located far away from these two nonoverlapped pathways, echoing with the analysis based on the insulation score (Figure 2A). Increasing the contact ranges from 2 Mb to 20 Mb progressively moves the ESC and NPC close to the pathways (Figure 4B,C,D). It is worthing noting that the NPC and ESC are very close to each other at the contact ranges of 5−10 Mb and 10−20 Mb, indicating that the chromosome interactions at these ranges are very similar at the ESC and NPC. Interestingly, when the contact ranges exceed 20 Mb (Figure 4E and F), the NPC appears to be on the two pathways, while the ESC is located far away from the NPC and these two pathways. The results indicate that the chromosome structures at the very long ranges (>20 Mb) can go through the ones formed at the NPC during both transitions in the transdifferentiation. Together, our findings suggest that the chromosomes during the transdifferentiation between the Fibro and Neuron cells can partially form the structures in the NPC at the long ranges.

■ DISCUSSION AND CONCLUSIONS

In this work, we studied the chromosome structural dynamics during the transdifferentiation between the Fibro and Neuron cells with a data-driven model followed by a nonequilibrium MD simulation approach. We analyzed the formations of TADs, compartments, and contacts, and quantified the chromosome structural transition pathways during cell transdifferentiation. We found that the degree of the irreversibility in cell transdifferentiation varies by contact range and type. In particular, we observed that the local contacts and the contacts within compartment B are formed irreversibly during cell transdifferentiation. Based on the intimate relationship between the structure and function of the chromosome, we described the cell transdifferentiation process from the chromosome structural perspective. The irreversibility, which reflects the nonequilibrium nature of the cell state transition process, underlines the differences in the mechanisms of the transitions from the Fibro to Neuron cells and from the Neuron to Fibro cells.

Our results indicate that the two transdifferentiation pathways in the opposite directions at the local structural level are nonoverlapped, indicating a significant degree of irreversibility for forming TADs. In contrast, these two pathways, when projected on the compartment profile, reflecting the chromosome structure at the long range, appear to be almost reversible. These distinct features of the chromosome structural transition pathways at the TAD and the compartment level may be correlated with the distinct mechanisms for forming the TADs and the compartments in the chromosome. For TADs, the structures are assumed to be formed by the active extrusion of the chromatin loops.[67,68] The active loop extrusion is undertaken by the structural maintenance of chromosome (SMC) protein complexes,[69] and the process is highly ATP-dependent.[70] In other words, the structural changes of the TADs in the chromosome are led by the active, motor-driven, nonequilibrium dynamics. From the physical perspective, the nonequilibrium dynamics is the origin of the irreversibility. Therefore, it is expected that there is a high degree of irreversibility in the chromosome structural transition pathways of cell transdifferentiation at the short-range where the TADs form. On the contrary, the exact segregation mechanism of the chromosome compartmentalization and its molecular players have not been fully identified. Nevertheless, compartments are deemed to rely in part on the association of the loci according to the histone modifications,[22,71] which leads to the different binding affinities across different loci and the potential recruitment of the heterochromatin protein 1 (HP1) or the other proteins.[72,73] Although our current limited knowledge on the mechanism of the compartmental organization does not rule out the possible active events involved in the process, it has been demonstrated that the spatially segregated compartment can be well explained by the equilibrium microphase segregation,[74,75] which can be spontaneously induced by the specific association of the chromosomal loci from the same compartment status with or without the participation of the proteins. Therefore, the nonequilibrium effects on the structural formations of the compartments in the chromosome should be much weaker, compared to the TADs, leading to largely overlapped transdifferentiation pathways of the forward and reverse transitions from the compartment perspective.

The nonoverlapped pathways at the TAD level indicate distinct behaviors of the chromosome reorganizing its structure during the forward and reverse processes. However, we speculate that the irreversibility at the TAD level may have negligible effects on distinguishing the two transitions due to the following two facts. First, we found that the insulation score profiles preserve similar trends during the transdifferentiation, so the boundaries of TADs are conserved. The conserved boundaries of TADs are supposed to maintain the same functional roles as spatial constraints for facilitating the specific enhancer-promoter interactions.[14] Second, a recent experiment found that the disruptions of the TAD organizations through CTCF depletion do not affect the transdifferentiation of the B cells into macrophages.[29] Besides, extensive studies showed that the structural changes on TADs in various species only caused minor changes in gene expressions.[76−80] Therefore, the different routes for TAD reorganizations during the two transdifferentiation processes should not lead to mechanistic differences in the gene regulation at the TAD level.

The irreversibility of transitions was further observed in terms of the contact adaptations within B compartment regions during the transdifferentiation. A previous study uncovered that the chromosomes in the mouse Neuron cell and NPC have weaker interactions within the active compartment regions accompanied by stronger interactions within the inactive compartment B than the chromosomes in the mouse

ESC.[81] We found a similar trend in calculating the contact strengths with different types in the chromosomes at the human Fibro cell, Neuron cell, NPC, and ESC (Figure S8). This suggests a common mechanism that the strong aggregation in the chromosome is more likely to occur at the euchromatin regions than the heterochromatin regions in the pluripotent ESC, and vice versa for the partially and terminally differentiated cells. These delicate arrangements of the compartment segregation strengths may be used to characterize the different stages of the cell developmental processes.[27,28,82,83] In this respect, we claim that the initial increase of the "A-A" contact strengths during the transdifferentiation potentially adapts the chromosome interactions within the compartment A regions toward the ones formed at the ESC. At the same time, the decrease of the "B-B" contact strengths during the two transitions also promotes the formation of chromosome interactions within the compartment B regions toward that at the ESC. However, it leads to irreversibility as the strengths of the "B-B" contacts at the late stages in the $TD^{F \to N}$ have to increase to reach the values at the Neuron cell (Figure 3B and D). Therefore, the chromosome in these two transitions has a similar propensity to organize its structure toward that formed in the ESC at the early stages of the transdifferentiation. The irreversibility between these two transitions originates from the different strengthens of the interactions formed within the B compartment. During the transition from the Fibro to Neuron cell, the chromosome can reach a weaker interaction strength in the inactive compartment B regions, possessing more degrees of ESC-like structures than during the transition from the Neuron to Fibro cells.

Combining all the findings present in this study allows us to draw a schematic Waddington's landscape for cell transdifferentiation between the Fibro and Neuron cell from the chromosome structural perspective.[84] In the pictorial Waddington's epigenetic landscape, the cell, which is metaphorically described as a ball, rolls down from the top to the bottom of the landscape during cell differentiation. The ESC, which possesses the highest degree of pluripotency, locates at the top of the landscape, and the terminally differentiated cells (e.g., the Fibro and Neuron cells) are placed at the minima of the landscape. During the ESC differentiation, the cell forms the multipotent NPC prior to the arrival at the Neuron cell. The cell transdifferentiation processes correspond to the transitions that occur directly between the two minima on the landscape (Figure 5). Our calculations on the strengths of the different contact types suggest that an initial climb up the landscape toward the ESC is a prerequisite for cell transdifferentiation (Figure 3B and D). The observation that the ESC is far away from the quantified pathways projected at the structural characteristics of the TADs and compartments (Figure 2A and D) indicates that the transdifferentiation does not undergo a full reprogramming process,[1,2] which corresponds to a complete rise of the cell state from the bottom to the top of the landscape. The long-range contact evolution pathways show that the chromosomes during the transdifferentiation form the structures, which largely resemble those in the NPC (Figure 4C−F). This suggests that the NPC, as a cell developmental state, lies close to the transdifferentiation paths. Our theoretical results resonate with a recent experimental work, where Treutlein et al. used the single-cell RNA-seq technique to reveal NPC-like intermediate cell states during the direct conversion of the mouse Fibro to the Neuron cell.[10] Furthermore, the transcriptomic data of the inter-
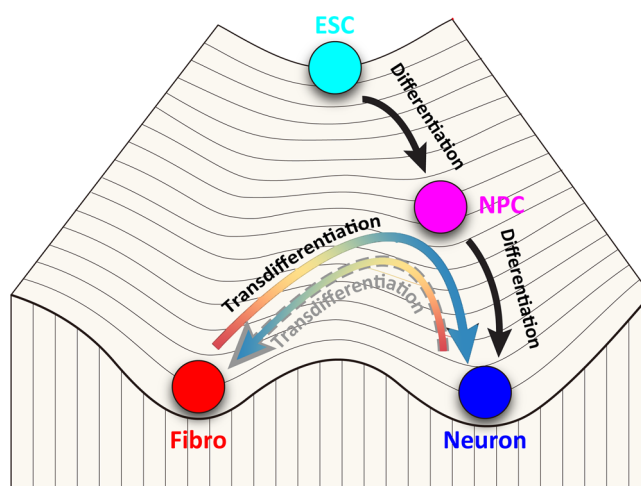
Figure 5. Scheme illustrating Waddington's landscape for cell transdifferentiation between the Fibro and Neuron cells from the chromosome structural dynamics perspective.

mediate states show certain degrees of deviation from the ones of the NPC.[10] This is also in line with our simulation that the NPC does not always locate on the transition pathways. The findings suggest the slightly divergent pathways for the late stages of the transdifferentiation from the Fibro to Neuron cell and the differentiation from the NPC to the Neuron cell, leading to nonoverlapped paths on Waddington's landscape (Figure 5). Finally, the two paths of transdifferentiation with opposite directions should be separated on the landscape to reflect the irreversibility of the transitions. However, the separation of these two paths should be minor, as the two transitions share many features in organizing the chromosome contacts, as indicated by our simulations.

It is noteworthy that there are two prominent limitations in our current model that can be improved in the future. First, the model with a focus on the chromosome structural changes in the interphase during cell transdifferentiation does not take into account the cell cycle process. Currently, the experimental Hi-C data for the cell cycle during cell development are still lacking, and the relationship between the cell cycle and cell transdifferentiation remains elusive. In this regard, we treated the cell cycle dynamics as an averaged background and focused on the slow cell developmental process. Thus, our results indicate the chromosome structural dynamics at the interphase of the cell during transdifferentiation. The model can be improved by implementing the cell-cycle dynamics using the Hi-C data at different cell-cycle phases measured from future experiments. Second, the model that simplifies cell transdifferentiation into a bistable switch does not take into account the possible existence of the intermediate state. The intermediate state, which corresponds to the metastable state on Waddington's landscape, should contribute to dictate the developmental pathway. Further improvement on the model can be made by applying multiple stepwise switching between the cell states during the developmental process when the Hi-C data at the intermediates state are available.

In summary, we quantified the chromosome structural dynamics during the transdifferentiation and uncovered the molecular mechanisms of the cell state transitions at the chromosomal level. Our predictions can be assessed by future experiments, in particular, with the rapid development of the 4D nucleome techniques.[85,86] On the other hand, further

improvements on our model can be made by introducing the cell-cycle dynamics and the intermediate states into the landscape-switching model, when the relevant Hi-C data are available. We anticipate that our approach can be extended to study the chromosome structure dynamics during different cell state transition processes and quantify the essential pathways and mechanisms that are difficult to characterize in the current experiments.

■ MATERIALS AND METHODS

Hi-C Data Processing

The Hi-C data of the Fibro cell and ESC were downloaded from the Gene Expression Omnibus database with accession number GSE63525[22] and GSE35156,[23] respectively. The Hi-C data of the Neuron cell and NPC are accessible through an open source platform (synapse, ID: syn12979101).[87] The analyses of Hi-C data were performed by the Hi-C Pro software through the standard pipeline.[88] The contact matrices were generated at a resolution of 100 kb and further normalized by the iterative correction and eigenvector decomposition (ICE) method.[89] We focused on the long arm of chromosome 14 with a range of 20.5−106.1 Mb, so there are 857 beads in our simulations. In order to perform the simulations, we further normalized the contact frequency to contact probability $f_{i,j}$, based on the reasonable assumptions that the neighboring beads are always in the contacts with probability $f_{i,j} \equiv 1$.[43−45,90]

Chromosome Simulation Model

In our simulations, the chromosome is described by a beads-on-a-string model, where the neighboring chromosomal loci are connected by the pseudo bonds.[91] Soft-core repulsive interactions are implemented on any pair of the beads to allow the chain-crossing, mimicking the effects of topoisomerase.[43,44,92] These bonded and nonbonded interactions with an additional spheric confinement potential, which mimics the effects of the nuclear membrane, make up a homopolymer potential $V_{Polymer}$. In our previous work, we showed that the chromosome under $V_{Polymer}$ results in an equilibrium globular ensemble, which has no biasing to specific structures.[47]

To accurately reproduce the experimental observations with the MD simulations, one can introduce additional restraints based on the experimental data. Here, we used the maximum entropy principle to implement the experimental observations, i.e., Hi-C contact maps, into the MD simulations. As requested by the maximum entropy principle, the potential from experimental restraints should be in the linear form of the contact probabilities.[60,93] Therefore, the potential of chromosome system at one cell state V_{Cell}, where the subscript "Cell" corresponds to "Fibro" or "Neuron", is expressed as the following:

$$V_{Cell} = V_{Polymer} + \sum_{i,j} \alpha_{i,j} P_{i,j}$$

where $P_{i,j}$ is the contact probability between the chromosomal loci "i" and "j" with prefactor $\alpha_{i,j}$ controlling the strength. $\alpha_{i,j}$ is determined by the maximum entropy principle through multiple rounds of iterations to ultimately match $P_{i,j}$ from the simulated chromosome ensembles to the $f_{i,j}$, which is the experimental Hi-C contact probability. Details of the model can be found in the Supporting Information and the previous studies.[43−45]

Landscape-Switching Model

We performed the landscape-switching model, which was developed in our previous work,[46−49] to study the chromosome structural dynamics during cell transdifferentiation. We briefly describe the model here. First, the chromosomes are simulated under the potential obtained by the maximum entropy principle simulations for the Fibro (V_{Fibro}) or Neuron (V_{Neuron}) cell. Then, a switch of the potential from the Fibro cell to Neuron cell ($V_{Fibro} \rightarrow V_{Neuron}$), or from the Neuron cell to Fibro cell ($V_{Neuron} \rightarrow V_{Fibro}$), is implemented. The switching implementation results in an instantaneous energy excitation to transfer the system from the equilibrium energy minimum on the preswitching landscape to an excited state on the postswitching landscape. Finally, the system relaxes to the energy minimum at the postswitching landscape, and the dynamics of the relaxation correspond to the transitions between the two cell states.

MD Simulations

We used Gromacs (version 4.5.7) MD software[94] with PLUMED (version 2.5.0)[95] to undertake the simulation tasks. Langevin dynamics was used with a friction coefficient $10\tau^{-1}$, where τ is the reduced time unit. Temperature is in the energy unit by multiplying the Boltzmann constant and ϵ is the reduced energy unit. The temperature in our model does not have a direct connection to the real one. Instead, it reflects the environmental scale in affecting the chromosome dynamics under the potential energy. Time step was set to be 0.0005τ. In each round of maximum entropy principle simulations, 100 independent MD simulations starting at different chromosome structures were performed with a length of 1000τ. A simulated annealing technique was applied in the individual simulations, where the temperature was gradually reduced from 4ϵ to ϵ during the first 250τ time. Then, the temperature was then kept at ϵ, and the second half of the trajectory from 500τ to 1000τ was collected to calculate the simulated contact probability $P_{i,j}$.

We performed the hierarchical clustering on the chromosome structural ensembles obtained by the maximum entropy principle simulations at the Fibro and Neuron cells (Figures S1 and S2). Two chromosome structures with the closest distances to the center of each cluster, which has populations more than 0.3%, were picked out for initializing the landscape-switching model simulations. These led to 220 and 270 chromosome structures to represent the ensembles at the Fibro and Neuron cells, respectively, and to perform the chromosome dynamics transition from the Fibro to Neuron cells and from the Neuron to Fibro cells, respectively. The trajectories at the landscape-switching model simulations with either V_{Fibro} or V_{Neuron} potential were performed with a length of $1 \times 10^4\tau$, within which the systems are able to reach the equilibrium (Figures S1 and S2).

Identifications of TADs and Compartments

We used the insulation score to identify the signals of the TADs. The insulation score was introduced by Crane et al.[65] Here, we applied the same size of sliding space (500 kb ×500 kb) suggested previously,[65] to calculate the insulation score of each chromosomal locus from the contact probability map. A minimum of the insulation score profile indicates a strong local insulation tendency to form the TAD boundary. TADs are usually megabase-sized structural domains, thus consisting of about 10 beads in our polymer model, which has a resolution of 100 kb. In our previous study at the same resolution,[47] we performed the simulations under only potential $V_{Polymer}$ and found that no TAD formed with the insulation scores equal to 0 across the whole chromosome segment. In this regard, TADs can be reliably established by forming and enhancing the block-sized contacts between the nonbonded beads with including the experimental restraint from the Hi-C data.

We used the enhanced contact probability P_{obs}/P_{exp} at the 1 Mb resolution to calculate the compartment profile. We performed PCA on the enhanced contact map after the ICE normalization and extracted the first PC to represent the compartment profile. Since the direction of the PC1 is arbitrary, we further determined the status of the compartment based on the association with the gene density, so the positive (compartment A) and negative (compartment B) values were assigned to the gene-rich and gene-poor regions (Figure S9). Finally, the direction of the PC1 values of states during the transdifferentiation was determined by calculating the correlation coefficient to the compartment profile (PC1) of the Fibro cell. The PCA plots of the trajectories of the enhanced contact probability were divided based on different contact ranges, which are determined by different genomic distances between the interacting loci. In practice, we first performed the PCA on the time evolution of the enhanced contact probability at different genomic distances, respectively. The trajectories were then projected onto the first two PCs, leading to the quantified pathways in terms of the enhanced contact probability.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.1c00416.

> Details of chromosome simulation model; figures of data-driven simulation results (Figures S1–S4), time-evolving trajectories of TADs and compartments (Figure S5 and S6), correlation between contact probability and compartment during transdifferentiation (Figure S7), contact probability varied by contact type (Figure S8), and gene density (Figures S9) (PDF)

AUTHOR INFORMATION

Corresponding Author

Jin Wang − *Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794, United States; Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York 11794, United States;* ● orcid.org/0000-0002-2841-4913; Email: jin.wang.1@stonybrook.edu

Author

Xiakun Chu − *Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794, United States;* ● orcid.org/0000-0003-3166-7070

Complete contact information is available at:
https://pubs.acs.org/10.1021/jacsau.1c00416

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676.

(2) Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**, *131*, 861–872.

(3) Vierbuchen, T.; Ostermeier, A.; Pang, Z. P.; Kokubu, Y.; Südhof, T. C.; Wernig, M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **2010**, *463*, 1035–1041.

(4) Pang, Z. P.; Yang, N.; Vierbuchen, T.; Ostermeier, A.; Fuentes, D. R.; Yang, T. Q.; Citri, A.; Sebastiano, V.; Marro, S.; Südhof, T. C.; Wernig, M. Induction of human neuronal cells by defined transcription factors. *Nature* **2011**, *476*, 220–223.

(5) Pfisterer, U.; Kirkeby, A.; Torper, O.; Wood, J.; Nelander, J.; Dufour, A.; Björklund, A.; Lindvall, O.; Jakobsson, J.; Parmar, M. Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10343–10348.

(6) Jenuwein, T.; Allis, C. D. Translating the histone code. *Science* **2001**, *293*, 1074–1080.

(7) Bernstein, B. E.; Meissner, A.; Lander, E. S. The mammalian epigenome. *Cell* **2007**, *128*, 669–681.

(8) Cieślar-Pobuda, A.; Knoflach, V.; Ringh, M. V.; Stark, J.; Likus, W.; Siemianowicz, K.; Ghavami, S.; Hudecki, A.; Green, J. L.; Łos, M. J. Transdifferentiation and reprogramming: overview of the processes, their similarities and differences. *Biochim. Biophys. Acta, Mol. Cell Res.* **2017**, *1864*, 1359–1369.

(9) Jopling, C.; Boue, S.; Belmonte, J. C. I. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 79–89.

(10) Treutlein, B.; Lee, Q. Y.; Camp, J. G.; Mall, M.; Koh, W.; Shariati, S. A. M.; Sim, S.; Neff, N. F.; Skotheim, J. M.; Wernig, M.; Quake, S. R. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **2016**, *534*, 391–395.

(11) Liu, Z.; Wang, L.; Welch, J. D.; Ma, H.; Zhou, Y.; Vaseghi, H. R.; Yu, S.; Wall, J. B.; Alimohamadi, S.; Zheng, M.; Yin, C.; Shen, W.; Prins, J. F.; Liu, J.; Qian, L. Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* **2017**, *551*, 100–104.

(12) Biddy, B. A.; Kong, W.; Kamimoto, K.; Guo, C.; Waye, S. E.; Sun, T.; Morris, S. A. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **2018**, *564*, 219–224.

(13) Reid, A.; Tursun, B. Transdifferentiation: do transition states lie on the path of development? *Current opinion in systems biology* **2018**, *11*, 18–23.

(14) De Laat, W.; Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **2013**, *502*, 499–506.

(15) Gorkin, D. U.; Leung, D.; Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell stem cell* **2014**, *14*, 762–775.

(16) Dekker, J.; Mirny, L. The 3D genome as moderator of chromosomal communication. *Cell* **2016**, *164*, 1110–1121.

(17) Spielmann, M.; Lupiáñez, D. G.; Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **2018**, *19*, 453–467.

(18) Furlong, E. E.; Levine, M. Developmental enhancers and chromosome topology. *Science* **2018**, *361*, 1341–1345.

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

(20) Di Stefano, M.; Paulsen, J.; Jost, D.; Marti-Renom, M. A. 4D nucleome modeling. *Curr. Opin. Genet. Dev.* **2021**, *67*, 25–32.

(21) Lieberman-Aiden, E.; et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **2009**, *326*, 289–293.

(22) Rao, S. S.; Huntley, M. H.; Durand, N. C.; Stamenova, E. K.; Bochkov, I. D.; Robinson, J. T.; Sanborn, A. L.; Machol, I.; Omer, A. D.; Lander, E. S.; Aiden, E. L. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **2014**, *159*, 1665–1680.

(23) Dixon, J. R.; Selvaraj, S.; Yue, F.; Kim, A.; Li, Y.; Shen, Y.; Hu, M.; Liu, J. S.; Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **2012**, *485*, 376.

(24) Nora, E. P.; Lajoie, B. R.; Schulz, E. G.; Giorgetti, L.; Okamoto, I.; Servant, N.; Piolot, T.; van Berkum, N. L.; Meisig, J.; Sedat, J.; Gribnau, J.; Barillot, E.; Blüthgen, N.; Dekker, J.; Heard, E. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **2012**, *485*, 381.

(25) Sexton, T.; Yaffe, E.; Kenigsberg, E.; Bantignies, F.; Leblanc, B.; Hoichman, M.; Parrinello, H.; Tanay, A.; Cavalli, G. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **2012**, *148*, 458–472.

(26) Boettiger, A. N.; Bintu, B.; Moffitt, J. R.; Wang, S.; Beliveau, B. J.; Fudenberg, G.; Imakaev, M.; Mirny, L. A.; Wu, C.-t.; Zhuang, X. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **2016**, *529*, 418.

(27) Xie, W.; et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **2013**, *153*, 1134–1148.

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

(29) Stik, G.; et al. CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response. *Nat. Genet.* **2020**, *52*, 655−661.

(30) Vilarrasa-Blasi, R.; et al. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. *Nat. Commun.* **2021**, *12*, 1−18.

(31) Nagano, T.; Lubling, Y.; Várnai, C.; Dudley, C.; Leung, W.; Baran, Y.; Cohen, N. M.; Wingett, S.; Fraser, P.; Tanay, A. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **2017**, *547*, 61.

(32) Stadhouders, R.; Vidal, E.; Serra, F.; Di Stefano, B.; Le Dily, F.; Quilez, J.; Gomez, A.; Collombet, S.; Berenguer, C.; Cuartero, Y.; Hecht, J.; Filion, G. J.; Beato, M.; Marti-Renom, M. A.; Graf, T. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **2018**, *50*, 238−249.

(33) Gibcus, J. H.; Samejima, K.; Goloborodko, A.; Samejima, I.; Naumova, N.; Nuebler, J.; Kanemaki, M. T.; Xie, L.; Paulson, J. R.; Earnshaw, W. C.; Mirny, L. A.; Dekker, J. A pathway for mitotic chromosome formation. *Science* **2018**, *359*, eaao6135.

(34) Paulsen, J.; Ali, T. M. L.; Nekrasov, M.; Delbarre, E.; Baudement, M.-O.; Kurscheid, S.; Tremethick, D.; Collas, P. Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nat. Genet.* **2019**, *51*, 835−843.

(35) Miura, H.; Takahashi, S.; Poonperm, R.; Tanigawa, A.; Takebayashi, S.-i.; Hiratani, I. Single-cell DNA replication profiling identifies spatiotemporal developmental dynamics of chromosome organization. *Nat. Genet.* **2019**, *51*, 1356−1368.

(36) Sati, S.; et al. 4D genome rewiring during oncogene-induced and replicative senescence. *Mol. Cell* **2020**, *78*, 522−538.

(37) Abramo, K.; Valton, A.-L.; Venev, S. V.; Ozadam, H.; Fox, A. N.; Dekker, J. A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **2019**, *21*, 1393−1402.

(38) Liu, N. Q.; de Wit, E. A transient absence of SMC-mediated loops after mitosis. *Nat. Cell Biol.* **2019**, *21*, 1303−1304.

(39) Caiazzo, M.; et al. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* **2011**, *476*, 224−227.

(40) Yoo, A. S.; Sun, A. X.; Li, L.; Shcheglovitov, A.; Portmann, T.; Li, Y.; Lee-Messer, C.; Dolmetsch, R. E.; Tsien, R. W.; Crabtree, G. R. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* **2011**, *476*, 228−231.

(41) Wang, S.; Wolynes, P. G. Communication: Effective temperature and glassy dynamics of active matter. *J. Chem. Phys.* **2011**, *135*, 051101−051101.

(42) Wang, S.; Wolynes, P. Tensegrity and motor-driven effective interactions in a model cytoskeleton. *J. Chem. Phys.* **2012**, *136*, 145102.

(43) Zhang, B.; Wolynes, P. G. Topology, structures, and energy landscapes of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 6062−6067.

(44) Zhang, B.; Wolynes, P. G. Shape transitions and chiral symmetry breaking in the energy landscape of the mitotic chromosome. *Phys. Rev. Lett.* **2016**, *116*, 248101.

(45) Qi, Y.; Reyes, A.; Johnstone, S. E.; Aryee, M. J.; Bernstein, B. E.; Zhang, B. Data-driven polymer model for mechanistic exploration of diploid genome organization. *Biophys. J.* **2020**, *119*, 1905−1916.

(46) Chu, X.; Wang, J. Conformational state switching and pathways of chromosome dynamics in cell cycle. *Appl. Phys. Rev.* **2020**, *7*, 031403.

(47) Chu, X.; Wang, J. Microscopic Chromosomal Structural and Dynamical Origin of Cell Differentiation and Reprogramming. *Advanced Science* **2020**, *7*, 2001572.

(48) Chu, X.; Wang, J. Deciphering the molecular mechanism of the cancer formation by chromosome structural dynamics. *PLoS Comput. Biol.* **2021**, *17*, e1009596.

(49) Chu, X.; Wang, J. Insights into the molecular mechanisms of cell fate decision making processes from chromosome structural dynamics. *bioRxiv* **2021**, 1 DOI: 10.1101/2021.05.09.443292.

(50) Ferrell, J. E.; Machleder, E. M. The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science* **1998**, *280*, 895−898.

(51) Xiong, W.; Ferrell, J. E. A positive-feedback-based bistable 'memory module'that governs a cell fate decision. *Nature* **2003**, *426*, 460−465.

(52) Fang, X.; Liu, Q.; Bohrer, C.; Hensel, Z.; Han, W.; Wang, J.; Xiao, J. Cell fate potentials and switching kinetics uncovered in a classic bistable genetic switch. *Nat. Commun.* **2018**, *9*, 1−9.

(53) Jiang, Z.; Tian, L.; Fang, X.; Zhang, K.; Liu, Q.; Dong, Q.; Wang, E.; Wang, J. The emergence of the two cell fates and their associated switching for a negative auto-regulating gene. *BMC Biol.* **2019**, *17*, 1−14.

(54) Sasai, M.; Wolynes, P. G. Stochastic gene expression as a many-body problem. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2374−2379.

(55) Hornos, J. E.; Schultz, D.; Innocentini, G. C.; Wang, J.; Walczak, A. M.; Onuchic, J. N.; Wolynes, P. G. Self-regulating gene: an exact solution. *Phys. Rev. E* **2005**, *72*, 051907.

(56) Wang, J.; Zhang, K.; Xu, L.; Wang, E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8257−8262.

(57) Li, C.; Wang, J. Quantifying Waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *J. R. Soc., Interface* **2013**, *10*, 20130787.

(58) Zhang, K.; Sasai, M.; Wang, J. Eddy current and coupled landscapes for nonadiabatic and nonequilibrium complex system dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14930−14935.

(59) Wang, J. Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Adv. Phys.* **2015**, *64*, 1−137.

(60) Cesari, A.; Reißer, S.; Bussi, G. Using the maximum entropy principle to combine simulations and solution experiments. *Computation* **2018**, *6*, 15.

(61) Liu, L.; Shi, G.; Thirumalai, D.; Hyeon, C. Chain organization of human interphase chromosome determines the spatiotemporal dynamics of chromatin loci. *PLoS Comput. Biol.* **2018**, *14*, e1006617.

(62) Di Pierro, M.; Potoyan, D. A.; Wolynes, P. G.; Onuchic, J. N. Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 7753−7758.

(63) Pope, B. D.; et al. Topologically associating domains are stable units of replication-timing regulation. *Nature* **2014**, *515*, 402.

(64) Dixon, J. R.; Gorkin, D. U.; Ren, B. Chromatin domains: the unit of chromosome organization. *Mol. Cell* **2016**, *62*, 668−680.

(65) Crane, E.; Bian, Q.; McCord, R. P.; Lajoie, B. R.; Wheeler, B. S.; Ralston, E. J.; Uzawa, S.; Dekker, J.; Meyer, B. J. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **2015**, *523*, 240.

(66) Schmitt, A.; Hu, M.; Jung, I.; Xu, Z.; Qiu, Y.; Tan, C.; Li, Y.; Lin, S.; Lin, Y.; Barr, C.; Ren, B. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep.* **2016**, *17*, 2042−2059.

(67) Sanborn, A. L.; et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E6456−E6465.

(68) Fudenberg, G.; Imakaev, M.; Lu, C.; Goloborodko, A.; Abdennur, N.; Mirny, L. A. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **2016**, *15*, 2038−2049.

(69) Uhlmann, F. SMC complexes: from DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 399.

(70) Ganji, M.; Shaltiel, I. A.; Bisht, S.; Kim, E.; Kalichava, A.; Haering, C. H.; Dekker, C. Real-time imaging of DNA loop extrusion by condensin. *Science* **2018**, *360*, 102−105.

(71) Filion, G. J.; van Bemmel, J. G.; Braunschweig, U.; Talhout, W.; Kind, J.; Ward, L. D.; Brugman, W.; de Castro, I. J.; Kerkhoven, R. M.; Bussemaker, H. J.; van Steensel, B. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **2010**, *143*, 212−224.

(72) Strom, A. R.; Emelyanov, A. V.; Mir, M.; Fyodorov, D. V.; Darzacq, X.; Karpen, G. H. Phase separation drives heterochromatin domain formation. *Nature* **2017**, *547*, 241−245.

(73) Larson, A. G.; Elnatan, D.; Keenen, M. M.; Trnka, M. J.; Johnston, J. B.; Burlingame, A. L.; Agard, D. A.; Redding, S.; Narlikar, G. J. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **2017**, *547*, 236−240.

(74) Jost, D.; Carrivain, P.; Cavalli, G.; Vaillant, C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* **2014**, *42*, 9553−9561.

(75) Di Pierro, M.; Zhang, B.; Aiden, E. L.; Wolynes, P. G.; Onuchic, J. N. Transferable model for chromosome architecture. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 12168−12173.

(76) Rao, S. S.; et al. Cohesin loss eliminates all loop domains. *Cell* **2017**, *171*, 305−320.

(77) Nora, E. P.; Goloborodko, A.; Valton, A.-L.; Gibcus, J. H.; Uebersohn, A.; Abdennur, N.; Dekker, J.; Mirny, L. A.; Bruneau, B. G. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **2017**, *169*, 930−944.

(78) Despang, A.; Schöpflin, R.; Franke, M.; Ali, S.; Jerković, I.; Paliou, C.; Chan, W.-L.; Timmermann, B.; Wittler, L.; Vingron, M.; Mundlos, S.; Ibrahim, D. M. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **2019**, *51*, 1263−1271.

(79) Ghavi-Helm, Y.; Jankowski, A.; Meiers, S.; Viales, R. R.; Korbel, J. O.; Furlong, E. E. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* **2019**, *51*, 1272−1282.

(80) Williamson, I.; Kane, L.; Devenney, P. S.; Flyamer, I. M.; Anderson, E.; Kilanowski, F.; Hill, R. E.; Bickmore, W. A.; Lettice, L. A. Developmentally regulated Shh expression is robust to TAD perturbations. *Development* **2019**, *146*, dev179523.

(81) Bonev, B.; Cohen, N. M.; Szabo, Q.; Fritsch, L.; Papadopoulos, G. L.; Lubling, Y.; Xu, X.; Lv, X.; Hugnot, J.-P.; Tanay, A.; Cavalli, G. Multiscale 3D genome rewiring during mouse neural development. *Cell* **2017**, *171*, 557−572.

(82) Hawkins, R. D.; et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell stem cell* **2010**, *6*, 479−491.

(83) Solovei, I.; et al. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **2013**, *152*, 584−598.

(84) Waddington, C. H. *The strategy of the genes*; Routledge, 1957.

(85) Dekker, J.; Belmont, A. S.; Guttman, M.; Leshyk, V. O.; Lis, J. T.; Lomvardas, S.; Mirny, L. A.; O'shea, C. C.; Park, P. J.; Ren, B.; Politz, J. C. R.; Shendure, J.; Zhong, S. the 4D Nucleome Network, The 4D nucleome project. *Nature* **2017**, *549*, 219−226.

(86) Marti-Renom, M. A.; Almouzni, G.; Bickmore, W. A.; Bystricky, K.; Cavalli, G.; Fraser, P.; Gasser, S. M.; Giorgetti, L.; Heard, E.; Nicodemi, M.; Nollmann, M.; Orozco, M.; Pombo, A.; Torres-Padilla, M.-E. Challenges and guidelines toward 4D nucleome data and model standards. *Nat. Genet.* **2018**, *50*, 1352−1358.

(87) Rajarajan, P.; et al. Neuron-specific signatures in the chromosomal connectome associated with schizophrenia risk. *Science* **2018**, *362*, eaat4311.

(88) Servant, N.; Varoquaux, N.; Lajoie, B. R.; Viara, E.; Chen, C.-J.; Vert, J.-P.; Heard, E.; Dekker, J.; Barillot, E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **2015**, *16*, 259.

(89) Imakaev, M.; Fudenberg, G.; McCord, R. P.; Naumova, N.; Goloborodko, A.; Lajoie, B. R.; Dekker, J.; Mirny, L. A. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **2012**, *9*, 999.

(90) Qi, Y.; Zhang, B. Predicting three-dimensional genome organization with chromatin states. *PLoS Comput. Biol.* **2019**, *15*, e1007024.

(91) Rosa, A.; Everaers, R. Structure and dynamics of interphase chromosomes. *PLoS Comput. Biol.* **2008**, *4*, e1000153.

(92) Naumova, N.; Imakaev, M.; Fudenberg, G.; Zhan, Y.; Lajoie, B. R.; Mirny, L. A.; Dekker, J. Organization of the mitotic chromosome. *Science* **2013**, *342*, 948−953.

(93) Lin, X.; Qi, Y.; Latham, A. P.; Zhang, B. Multiscale modeling of genome organization with maximum entropy optimization. *J. Chem. Phys.* **2021**, *155*, 010901.

(94) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(95) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.