Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

Characterization of network hierarchy reflects cell state specificity in genome organization

Jingyao Wang, Yue Xue, Yueying He, et al.

Genome Res. 2023 33: 247-260 originally published online February 24, 2023
Access the most recent version at doi:10.1101/gr.277206.122

References This article cites 77 articles, 9 of which can be accessed free at:
http://genome.cshlp.org/content/33/2/247.full.html#ref-list-1

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see https://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**


CRISPR and RNAi Genetic Screening.
Your new superpower.
LEARN MORE
CELLECTA

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

© 2023 Wang et al.; Published by Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

Characterization of network hierarchy reflects cell state specificity in genome organization

Jingyao Wang,[1,4] Yue Xue,[1,4] Yueying He,[1] Hui Quan,[1] Jun Zhang,[3] and Yi Qin Gao[1,2,3]

[1] *Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing, 100871, China;* [2] *Biomedical Pioneering Innovation Center (BIOPIC), Peking University, Beijing, 100871, China;* [3] *Changping Laboratory, Beijing, 102206, China*

Dynamic chromatin structure acts as the regulator of transcription program in crucial processes including cancer and cell development, but a unified framework for characterizing chromatin structural evolution remains to be established. Here, we performed graph inferences on Hi-C data sets and derived the chromatin contact networks. We discovered significant decreases in information transmission efficiencies in chromatin of colorectal cancer (CRC) and T-cell acute lymphoblastic leukemia (T-ALL) compared to corresponding normal controls through graph statistics. Using network embedding in the Poincaré disk, the hierarchy depths of chromatin from CRC and T-ALL patients were found to be significantly shallower compared to their normal controls. A reverse trend of change in chromatin structure was observed during early embryo development. We found tissue-specific conservation of hierarchy order in chromatin contact networks. Our findings reveal the top-down hierarchy of chromatin organization, which is significantly attenuated in cancer.

[Supplemental material is available for this article.]

Genome architecture plays an important role in transcription regulation (Liu et al. 2018; Deng et al. 2022; He et al. 2022). Instead of maintaining a static configuration, chromatin structure is dynamic (Barth et al. 2020). A large and diverse set of gene regulatory elements in chromatin are found to regulate transcription through long-range interactions. For example, promoter–enhancer interactions regulate the dynamic activation of genes (Cloney 2016), whereas insulators and repressors play the opposite roles (Yang and Corces 2012). Graph theory has been widely used in deciphering biological networks including protein interaction and gene expression networks (Koutrouli et al. 2020). Dynamic information transmission in networks is vital for maintaining functionalities (Eisenberg et al. 2000; Brazhnik et al. 2002).

Chromatin structure evolves during crucial life processes including cancer and cell development. Although extensive studies have attributed oncogenesis to gene mutations (Fujimoto et al. 1992; Brekelmans et al. 2001; Davies et al. 2002; Spruck et al. 2002), targeting chromatin structure for diagnosis and therapy has increasingly gained attention (Kaur et al. 2019). Conventional microscopic observation of aberrant chromatin was used for cancer diagnosis (Zink et al. 2004). The recently developed Hi-C technology allows detailed inspection of aberrant chromatin structure in cancer. Pathological cancer tissues were observed to be accompanied by disturbances in compartments and TAD structures in Hi-C (Johnstone et al. 2020; Iyyanki et al. 2021; Yang et al. 2021). On the other side, the similarity between embryonic stem cells and cancer cells has also attracted attention (Kim and Orkin 2011; Malta et al. 2018; Feng et al. 2021). The characteristics shared by embryonic stem cells and cancer cells include unlimited proliferation and high cellular plasticity (Seftor et al. 2002; Hendrix et al. 2007). However, the similarity in chromatin structure between embryo tissues and pathological cancer tissues has not been extensively studied.

Although a number of methods have been reported to delineate the hierarchy of Hi-C data, they are confined to the notion of nested hierarchy, which describes a containment relationship of units. For example, Soler-Vila et al. introduced TADpole to identify the hierarchical subdivisions of TADs in intrachromosomal Hi-C (Soler-Vila et al. 2020). These TAD prediction methods help decode the modularity of chromatin structure. However, it remains to be investigated how genomes assemble beyond the scope of discrete structural units like TADs.

Chromatin contacts converge into a scale-free hierarchical network, where hub loci are especially important for chromatin stability and the proper functioning of cells (Sandhu et al. 2012; Ding et al. 2021). Considering the important role of chromatin interactions in regulating transcription, it is useful to identify the key elements in genome structure organization. This study is devoted to delineating the hierarchy describing the top-down organization of chromatin. We considered chromatin segments (loci) forming contacts with extensive genomic regions as top-level nodes in the hierarchy, whereas loci with local contacting patterns are bottom-level nodes. We focused on the roles of chromatin hierarchy in important processes including cancer and cell development.

Result

Graph statistics reflect changes in chromatin contact networks in cancer and cell differentiation

We examined changes in information transmission properties of chromatin contact networks (CCNs) from normal to cancer cells

---

[4]**These authors contributed equally to this work.**
Corresponding author: gaoyq@pku.edu.cn

Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.277206.122.

© 2023 Wang et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see https://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

using published Hi-C data sets of colorectal cancer (CRC) and T-cell acute lymphoblastic leukemia (T-ALL) (Johnstone et al. 2020; Yang et al. 2021). We first performed a graph inference to each intrachromosomal Hi-C matrix (Fig. 1A, see in Methods). We then characterized the traffic fluxes of edges in CCNs using betweenness centrality (Freeman 1977) (BC, see in Methods). BC of an edge measures its average probability to constitute the shortest path between any node pair, thus edges with high BC are essential for information transmission in a functional network (Girvan and Newman 2002). The expected BC is calculated as the average BC of edges at each sequential distance. We found a positive relationship between expected BC and the sequential distance it strides across, which underlines the bridge-like characteristics of long-range contacts (Fig. 1B). We next identified the normal- and cancer-specific edges in CRC, T-ALL, and corresponding normal controls. For instance, an edge existing in more than 75% normal CCNs and concurrently absent in more than 75% cancer CCNs was defined as normal specific for a certain cell type. In both cases of CRC and T-ALL, we found a significant decrease in linear sequential distances for cancer-specific edges compared with normal-specific edges (Fig. 1C, $P = 1.6 \times 10^{-12}$ for CRC and $P = 2.7 \times 10^{-154}$ for T-ALL). The analogous pattern of

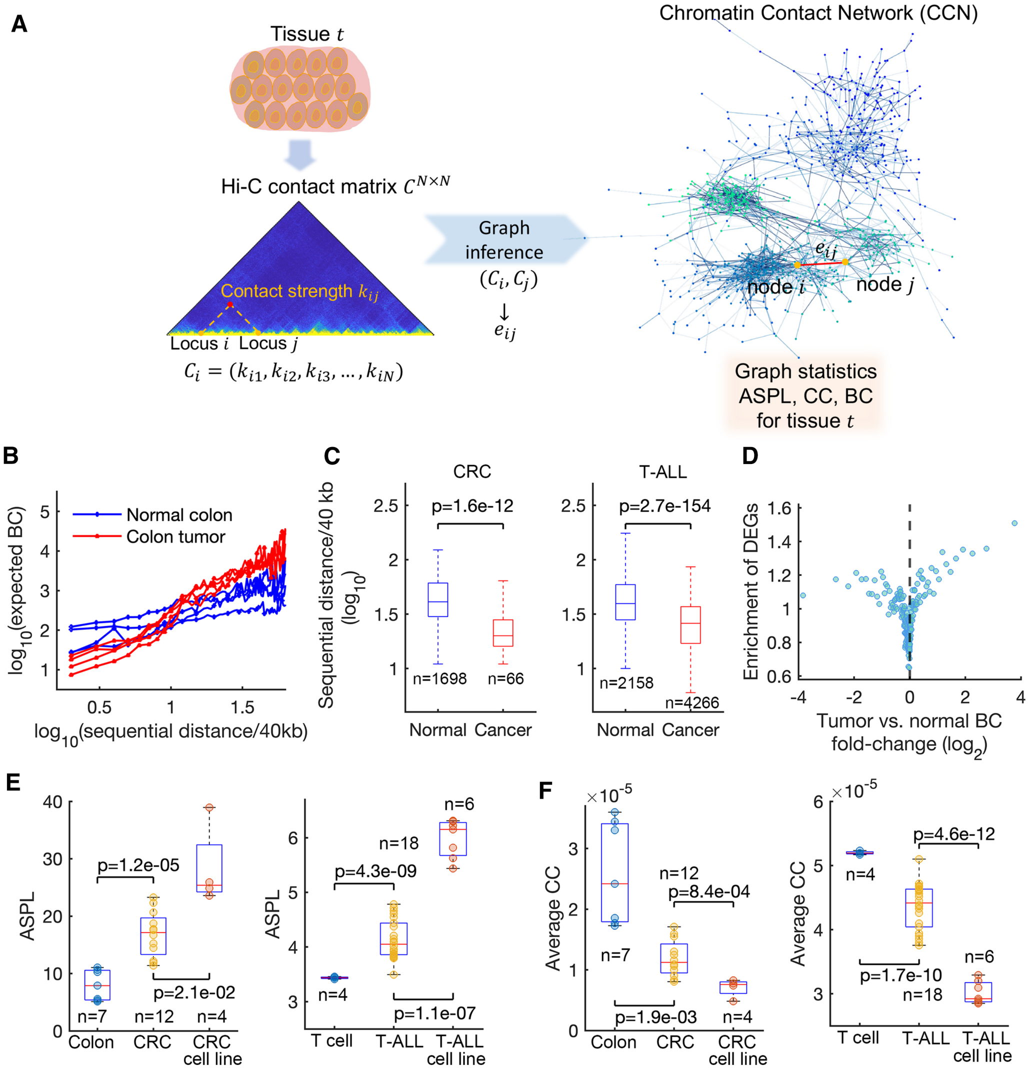

Figure 1. Graph statistics reflect decrease in information transmission efficiencies in chromatin contact networks (CCNs) in cancer. (*A*) We performed network inferences on Hi-C matrices (see in Methods). The element k_{ij} is the contact strength between the *i*th locus and the *j*th locus in sequential order. Then we characterized graph statistics including betweenness centrality (BC), average shortest path length (ASPL), and closeness centrality (CC). (*B*) Logarithm plots of expected BC at various sequential distances for normal colon (blue) and tumor (red) samples from four individuals. (*C*) Boxplots showing logarithms of sequential distances of cancer- or normal-specific edges in colorectal cancer (CRC) and T-cell acute lymphoblastic leukemia (T-ALL). (*D*) Differentially expressed gene (DEG) enrichment in edges with different levels of BC fold change (FC_b) for CRC. The enrichment was quantified as the observed/expected value of DEG ratio in edges. Edges were stratified into 200 groups according to FC_b rank. The *x* and *y* coordinates of points are group averages of FC_b and DEG enrichment, respectively. (*E*) Boxplots showing ASPL in cancer, cancer cell lines, and corresponding normal controls for CRC (*left*) and T-ALL (*right*). (*F*) Boxplots showing average CC in cancer, cancer cell lines, and corresponding normal controls for CRC (*left*) and T-ALL (*right*). Identification of normal- and cancer-specific edges and graph statistics including BC, ASPL, and CC were performed on the intrachromosomal CCN of Chromosome 1. In each boxplot, the *center* line, the *bottom* hinge, the *top* hinge, and the whiskers indicated the median, the 25th percentile, the 75th percentile, and 1.5 times the interquartile range (IQR) of each data set, respectively.

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

edge redistribution shared by tumor and leukemia indicates a generality of cancers, which represents a loss of bridge-like long-range contacts.

We further investigated the possible influences of bridge-reorganization in CCNs in CRC. We stratified edges into groups according to their BC fold changes (FC_b) from normal colons to tumors. FC_b of an edge e is defined as the adjusted fold change of BC,

$$FC_b(e) = \log_2\left(\frac{1 + BC^t(e)/\overline{BC^t}}{1 + BC^n(e)/\overline{BC^n}}\right),$$

where $BC^n(e)$ and $BC^t(e)$ stand for BC of the edge e in a normal or a tumor CCN, and the overbar denotes averaging over all edges in the normal or tumor CCN, respectively. Through adding the pseudocount 1, the adjusted definition of fold change preserves edges with largely different BC from normal to tumor. We downloaded gene expression profiles of normal colon and CRC samples from the TCGA database and identified differentially expressed genes (DEGs, see in Methods). Edges with high absolute FC_b tend to show enrichment in DEGs (Fig. 1D), which was defined as the observed/expected ratio of DEGs in either anchor of edges. The expected DEG number is calculated as the product of the DEG ratio in all loci and the number of loci in a divided group.

It is known that bridge edges in networks are important for efficient routing (Krioukov et al. 2010). Considering the loss of bridge-like long-range contacts in cancer (Fig. 1B,C), we hypothesized that the global information transmission efficiencies in CCNs should decrease. The inverse of the average shortest path length (ASPL) measures the global information transmission efficiency of networks (Latora and Marchiori 2001; Mao and Zhang 2013; Shimizu and Mori 2016). We characterized the ASPL of CCN for each Hi-C sample (see in Methods). We found that cancer CCNs have significantly longer ASPL than normal CCNs in CRC and T-ALL (Fig. 1E, $P = 1.2 \times 10^{-5}$ for CRC and $P = 4.3 \times 10^{-9}$ for T-ALL), indicating a significant decrease of global information transmission efficiency in cancer CCNs. CCNs of cancer cell lines have even longer ASPL than those of cancer samples, underlining a distinction between cancer and cancer cell lines (Fig. 1E, $P = 2.1 \times 10^{-2}$ for CRC and $P = 1.1 \times 10^{-7}$ for T-ALL).

Aside from characterizing the global information transmission efficiency with ASPL, we further quantified the information transmission efficiency of each node by CC (Golbeck 2013) (see in Methods). The average CC values are significantly decreased in CRC and T-ALL compared to their normal controls (Fig. 1F, $P = 1.9 \times 10^{-3}$ for CRC and $P = 1.7 \times 10^{-10}$ for T-ALL). Such a tendency continues from cancer to cancer cell lines ($P = 8.4 \times 10^{-4}$ for CRC and $P = 4.6 \times 10^{-12}$ for T-ALL). The global loss of CC indicates a localization tendency of the network (Ahmad et al. 2020). Nodes with the highest 10% CC in the normal colon contain genes enriched in the regulation of transcription from RNA polymerase II promoter and cell cycle (Supplemental Fig. S1A; Supplemental Table S3). We normalized CC along the sequence into its Z-score vector and identified nodes with significant decrease ($P < 0.05$) in CRC. These nodes contain genes enriched in cell differentiation and negative regulation of cell proliferation (Supplemental Fig. S1B; Supplemental Table S4).

For each CCN, we generated a random network with the same number of nodes and edges and calculated their ASPL and average CCs). With randomized network structures, normal and cancer samples no longer show significant changes in these properties (Supplemental Fig. S2).

Poincaré embedding of CCNs reveals cell state–related changes in hierarchy

To visualize the hierarchical structure of chromatin contact networks, we next embedded the Hi-C of each sample in the Poincaré disk. Graph embedding is a dimensionality reduction strategy which transforms graph connectivity into distances in the embedding space (Perozzi et al. 2014; Grover and Leskovec 2016). Because hyperbolic geometry underlies a hierarchical network with a scale-free structure, Poincaré embedding is a promising and suitable method for representing chromatin contact networks while revealing the latent hierarchy. We embedded each intrachromosomal CCN into the Poincaré disk adapting the algorithms published by Nickel and Kiela (Nickel and Kiela 2017) (see in Methods). We quantified the hierarchy statuses of individual nodes using Poincaré radii (R_P), because nodes with smaller R_P are closer to all other nodes and have higher hierarchy statuses (Krioukov et al. 2010; Nickel and Kiela 2017).

Discriminations between normal and cancer CCNs are clearly seen in the Poincaré disk for both CRC and T-ALL (Fig. 2A,B). From normal to cancer, the nodes in the embedded structure show a tendency of shifting from the disk center to the periphery. Similar tendencies were observed in CRC and T-ALL, and continue in the change from cancer to cancer cell lines (Fig. 2A,B). We defined the hierarchy depth of CCN as the negative of the median Poincaré radii (R_P). Generally, the median R_P shows an ascending trend from normal tissues to cancer, as well as from cancer to cancer cell lines (Fig. 2C). Analogous changes occurred on all individual chromosomes (Supplemental Fig. S3), indicating that the decrease of hierarchy depth occurred genome-wide in cancer. The median R_P of normal colon cell line is between those of normal colons and CRC cell lines (Fig. 2C), probably implying a malignancy-related abnormality in chromatin structure of normal cell lines. We generated random networks having the same number of nodes and edges with a pair of normal colon and CRC samples. We performed similar Poincaré embedding procedures for control. We derived randomized embedded structures as shown in Supplemental Figure S4.

Intrigued by the stemness in cancer cells (Seftor et al. 2002; Weeraratna et al. 2002; Balint et al. 2005; Kim and Orkin 2011; Malta et al. 2018), we also looked into the cell development process, during which chromatin structure undergoes significant changes (Dixon et al. 2015; Vilarrasa-Blasi et al. 2021). We first explored changes in CCNs during early embryonic development using published Hi-C data of human (Chen et al. 2019). Compared to the changes from normal to cancer, CCNs change in an opposite way during early embryonic development. As shown in Figure 2D, an initial establishment of hierarchy depth is manifested by a decrease in the median R_P before blastocyst. In cancer, the hierarchical chromatin structure collapsed, forming a less structured conformation largely dictated by the 1-D linear DNA sequence (Fig. 2A,B,D).

To investigate chromatin hierarchy changes in further differentiation from embryo stem cells, we performed Poincaré embeddings on published Hi-C data sets of human embryonic stem cells (hESCs) and three hESC-derived lineages (Dixon et al. 2015). Of note, we observed a significant increase in the hierarchy depth after differentiation from hESCs, indicated by the decrease in the median R_P (Fig. 3A,B). Whereas mesendoderm cells (MES) give rise to all cell types derived from mesoderm and endoderm, trophoblast-like cells (TBL) and neural progenitor cells (NPCs) differentiate into extraembryonic structures and neural cells, respectively. With more limited differential potential, TBL and NPCs have their hierarchy depth of

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

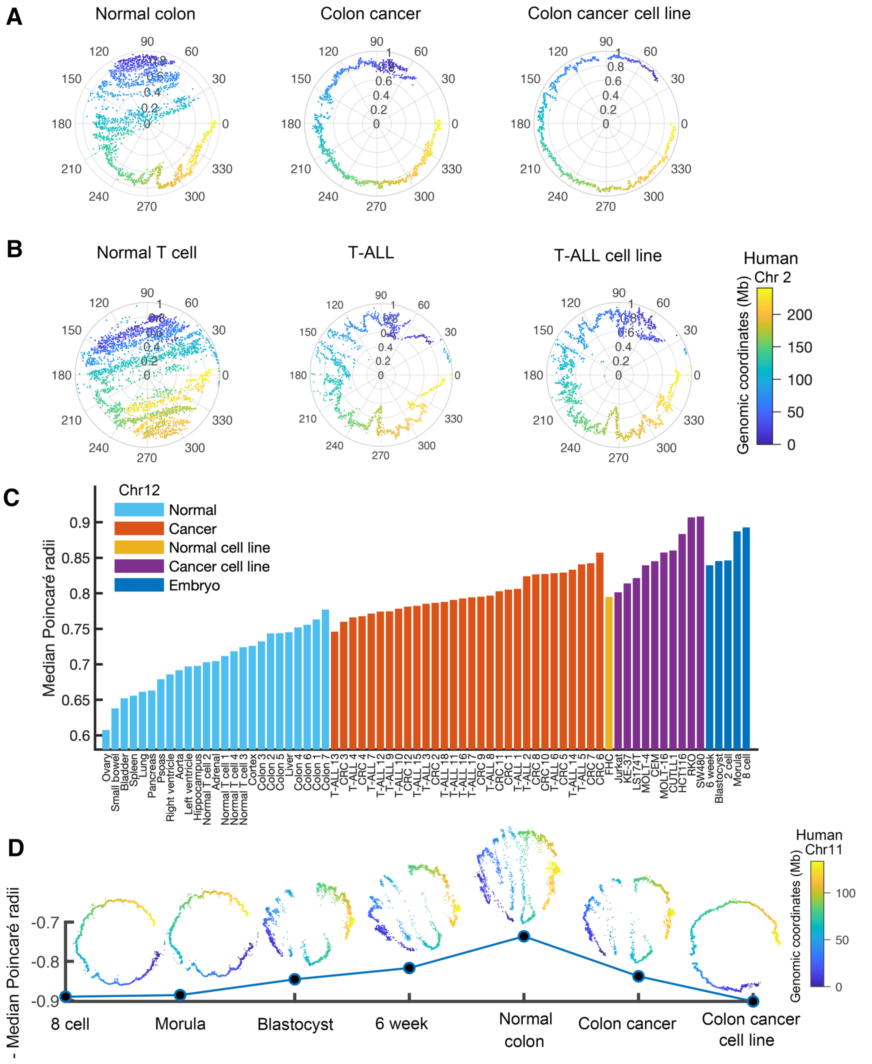

Figure 2. Poincaré embedding shows cell state–related hierarchy changes of CCNs in cancer and early embryo development. (*A*,*B*) Embedded structures of CCNs in the Poincaré disk for normal controls (*left*), cancer (*middle*), and cancer cell lines (*right*) of CRC and T-ALL. The color distribution of embedded nodes represents DNA sequential order. (*C*) Bars showing the median Poincaré radii (R_P) of each sample with cell types varying from primary tissues, cancer, cancer cell lines, and early embryos. (*D*) Changes in the negative median Poincaré radii during human early embryo development and in cancer.

CCNs more significantly increased than MES. Early embryo development initiates the cell differentiation process, and primary tissues are at the other end of the cell differentiation trajectory (Peng et al. 2017; Theilgaard-Mönch et al. 2022). We investigated changes in chromatin hierarchy in cell differentiation by comparing the primary tissues to human early embryos (from two-cell to blastocyst) using published Hi-C data sets (Schmitt et al. 2016; Chen et al. 2019). From human early embryos to differentiated tissues, a significant decrease in the median R_p was observed genome widely (Supplemental Fig. S3), reflecting the establishment of chromatin hierarchy in cell differentiation.

We performed a similar network embedding analysis after normalizing the contact frequencies based on sequence length. Because the normalization operation heightened the long-range contacts, hierarchy depth of CCNs increased generally after normalization (Supplemental Fig. S5). However, under normalization, the cancer CCN still displayed a shallower hierarchy than its normal counterpart, indicating increased sparsity in long-range contacts in cancer. The biased loss of long-range contacts may be resultant from a larger nucleus size, which has been observed in cancer (Denais and Lammerding 2014). We used the GRCh37 reference genome for analyses. We also transformed the coordinates of each 40-kb bin center from GRCh37 to GRCh38 using UCSC liftOver (Kuhn et al. 2013) and performed identical Poincaré embedding analyses. As shown in Supplemental Figure S6, with the transformed Hi-C matrices as inputs, the trends of hierarchy establishment during cell development and the hierarchy collapse in cancer were reproduced.

CCN hierarchy disorder in cancer

From the foregoing analysis, hierarchy depth is an important descriptor of a chromatin contact network. On the other side, we hypothesized that the hierarchy order of nodes in CCNs is also vital and related to cell states. To explore the ability of the hierarchy order of nodes in characterizing the cell states, we transformed the R_P vector into its Z-score vector along sequence. Such an operation was performed to evaluate each node's relative hierarchy status among all nodes. We next performed a principal component analysis (PCA) on normalized R_P vectors for all samples, based on which each sample was represented by a low-dimensional vector in the principal component (PC) space. In both cases of CRC and T-ALL, samples are departed in PC1 according to their belonging clusters of normal, cancer, and cancer cell line (Fig. 4A). The PCA is conducted on R_P values of all 40-kb bins in the entire genome. Cell types can also be separated by PCA of R_P on individual chromosomes (Supplemental Fig. S7). Such a result inspired us to further study the association between cell identity and the hierarchy order. We evaluated the similarity in hierarchy order using the Spearman's correlation coefficients between R_P vectors. Conservation of intrachromosomal hierarchy order among normal colons was observed, which was weakened among colon tumors (Fig. 4B), reflecting increased cell heterogeneity in cancer. Distinctions in hierarchy order between the normal colons and colon tumors indicate a hierarchy disorder in CRC (Fig. 4C). On the other hand, each normal primary tissue has a unique hierarchy order, except for two pairs (cortex and hippocampus, left ventricle and right ventricle), which have high (>0.9) Spearman's correlation coefficients (Fig. 4C). This result further indicates that CCN hierarchy order not only distinguishes between normal and cancer tissues but also reveals tissue specificity.

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

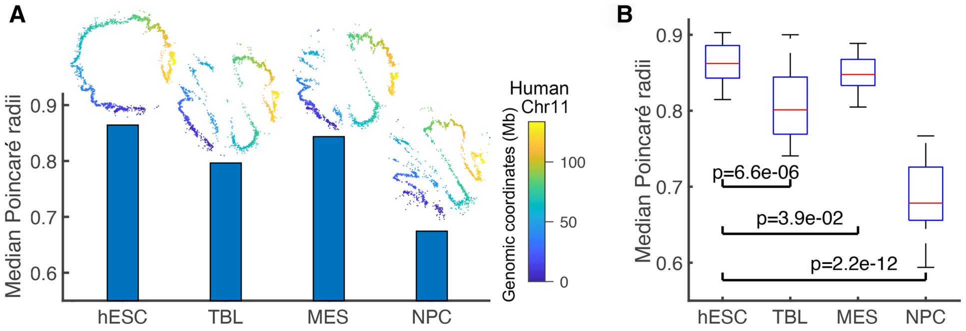

Figure 3. Changes in hierarchy depth of CCN from human embryonic stem cells (hESCs) to three hESC-derived lineages. We defined the hierarchy depth of CCN as the negative of the median Poincaré radii (R_P). (*A*) Bars showing the median Poincaré radii of Chr 11 for human embryonic stem cell (hESC), trophoblast-like cells (TBL), mesendoderm cells (MES), and neural progenitor cell (NPC). Corresponding Poincaré embedding structures are shown upon each bar. (*B*) Boxplot showing the median Poincaré radii of all intrachromosomal CCNs in each sample.

The foregoing analysis revealed hierarchy disorder in intrachromosomal CCNs in cancer. We further investigated changes in the context of whole chromatin during the same process. Due to existence of chromatin territories, intrachromosomal contact networks are relatively independent components of the whole-chromatin CCN. We embedded the whole-chromatin CCNs in the Poincaré disk at 100-kb resolution (see in Methods). As shown in Figure 4D, small and gene-rich Chromosomes 16, 17, 19, 20, 21, and 22 are located near the disk center in normal colon, whereas they are relocated towards the disk periphery in colon tumor. We reasoned that chromosomes possess different hierarchy statuses owing to an ordered radial distribution in the normal nucleus. In fact, Chromosomes 16, 17, 19, 20, 21, and 22 preferentially interact with each other in a karyotypically normal human lymphoblastoid cell line (Lieberman-Aiden et al. 2009). We defined the interchromosomal closeness matrix (see in Methods) to evaluate the relative distance between each pair of chromosomes. As shown in the closeness matrix for the normal colon, long chromosomes are far from each other whereas short chromosomes are close to each other (Fig. 4E). The small and gene-poor Chromosome 18 behaves analogously to long chromosomes. Actually, Chromosome 18 is preferentially located near the nuclear periphery (Ramani et al. 2016), whereas small and gene-rich chromosomes are frequently observed in the nucleus interior (Boyle et al. 2001; Tanabe et al. 2002). In colon tumor, the distinction between contacting patterns of short and long chromosomes was attenuated (Fig. 4E), implying a disorder in the radial distribution of chromosomes in the nucleus. In fact, long chromosomes are enriched with lamina associated domains (LADs) (Kind et al. 2015). In cancer, lamina associated heterochromatin was observed to be relocated into the nuclear interior (Johnstone et al. 2020).

Hierarchical chromatin organization contributes to cell state maintenance

Through Poincaré embedding, we quantified the hierarchy statuses of nodes in a CCN with R_P. We identified conservative loci in each hierarchy class in normal colon. Specifically, for each sample, nodes in an intrachromosomal CCN were stratified into five classes according to their entries in the R_P vector (i.e., sections of 1st to 20th, 20th to 40th, …, 80th to 100th quantiles). For each hierarchy class, the intersection of loci among normal samples ($n = 7$ for colon) was identified as the conservative loci of the hierarchy class. We quantified the conservativeness of each hierarchy class by Jaccard Index (JI), which was calculated as the size of intersection divided by the size of union. Loci in the highest hierarchy class (H-loci) and loci in the lowest hierarchy class (L-loci) have larger JI compared to the middle classes (Fig. 5A), indicating their stable hierarchy statuses among samples. The global-contacting mode of the H-loci was manifested by the largest number of contacts and the highest proportion of long-range (>4 Mb) contacts (Fig. 5B).

Exemplary contact patterns of genes in each hierarchy class in the normal colon are shown in Figure 6. For instance, *OTX1* in the top hierarchy encodes a transcription factor important for brain development (Levantini et al. 2003), it forms more long-range contacts than genes with lower hierarchy (Fig. 6A). Housekeeping genes are the most enriched in the bottom hierarchy (Fig. 5A). One such example is *ODC1*, which has a local contacting pattern characterized by sparse long-range contacts (Fig. 6E). With moderate amounts of long-range contacts, genes in the middle hierarchy classes have intermediate contacting patterns between the top and the bottom hierarchy (Fig. 6B–D).

We further investigated the regulatory-element related histone modifications in each hierarchy class. Enhancers and silencers, decorated with high H3K27ac and H3K27me3 signals respectively, are known to function as transcription activators and repressors via chromatin interactions with promoters (Creyghton et al. 2010; Cai et al. 2021; Xu et al. 2022). We found that H3K27ac peaks (enhancer-like) are enriched in both the H-loci and the L-loci (Fig. 5A). Such an observation indicated the existence of distinct enhancers with local or global contacting patterns, respectively. Loci in the middle hierarchy classes (class 2, 3, and 4) have medium or low H3K27ac signals on average. SEs are also enriched in both the H-loci and the L-loci (Fig. 5A). The enrichment of SEs in L-loci is more significant than that in H-loci. Exemplary tracks of SEs from the top and the bottom hierarchy are displayed in Figure 5, C and D, respectively. The SE in the L-loci is situated in a more H3K27ac-rich region than that in the H-loci. As shown in the virtual 4C tracks, the SE in the H-loci has more long-range contacts than that in the L-loci.

Associations between tissue specificities and spatial-temporal heterogeneity of SEs have been reported (Huang et al. 2018; Kai et al. 2021). We further investigated the relationship between the tissue specificity of SEs and their structural hierarchy. We downloaded the genomic locations of SEs from the SEdb database (Supplemental Table S5). We measured the tissue specificity of a SE by the number of tissues it is shared by; for example, a SE exclusive

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

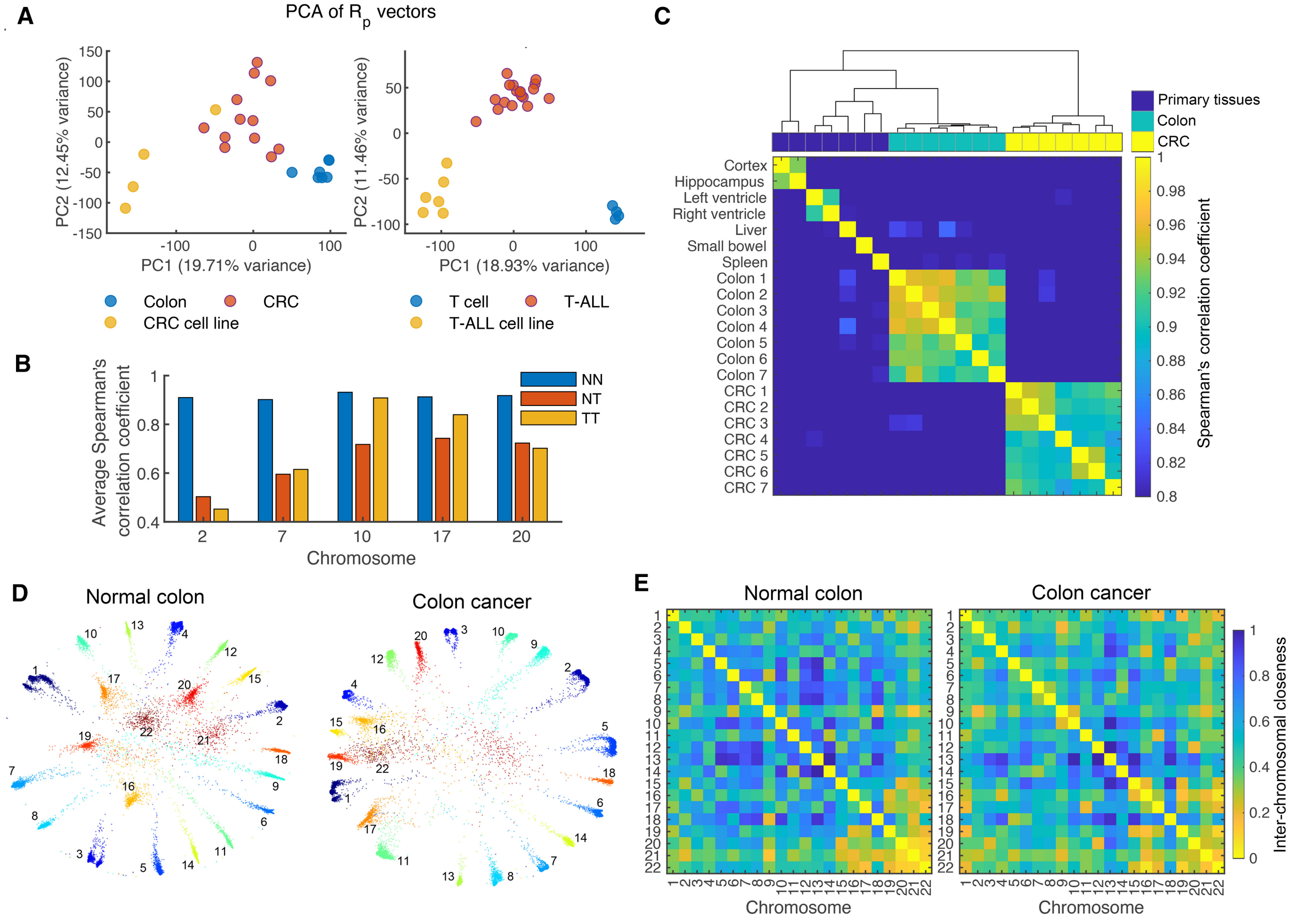

Figure 4. CCN hierarchy disorder in cancer. (*A*) Principal component analysis (PCA) for R_P vectors of samples in CRC and T-ALL data sets. The PCA is conducted on R_P values of all 40-kb bins in the entire genome. Each point represents a sample in PC space defined by the first two principal components. (*B*) Bar graphs showing the average Spearman's correlation coefficients of R_P between all pairs of normal colons (NN), between all pairs of colon tumors (TT), and between each normal colon and each colon tumor (NT). (*C*) Rank similarity in R_P among samples including primary tissues, normal colons, and colon tumors. Each unit color in the heatmap represents the value of Spearman's correlation coefficient between R_P vectors of corresponding sample pair. Samples were clustered using the Ward's hierarchical clustering method. The heatmap was derived from the intrachromosomal CCNs of Chromosome 10, therein the normal colon and tumor sets both have conserved hierarchy order as shown in *B*. (*D*) Embedded structures of whole chromatin contact networks for a pair of normal colon (*left*) and colon tumor (*right*) of one individual. Chromosome numbers were annotated alongside corresponding chromosomes in embedded structures. (*E*) Interchromosomal closeness matrices corresponding to (*D*). Each heatmap entry ranging from 0 to 1 indicates the closeness between each pair of chromosomes.

to lung has the highest specificity for lung. We quantified the relative hierarchy status of each SE with its rank score in Poincaré radii (R_P), a smaller rank score means a higher relative hierarchy status. For each tissue, SEs were ranked in ascending order according to corresponding R_P. The rank score of a SE was calculated as its rank in R_P divided by the number of entries in R_P. Then each SE was classified into groups by the number of tissues sharing it. We found a positive correlation between the tissue specificity of SEs and their relative hierarchy statuses (Fig. 5E). Such a result implies that SEs with stronger tissue specificity are involved in more advanced structural regulation, whereas general SEs tend to function locally. We assumed the structural hierarchy of SEs also reflects the order of their temporal employment in cell development, because specific cell functions emerge late.

H-loci also show enrichment for colon complementary genes (see in Methods) and have the highest level of average H3K27me3 signal (Fig. 5A), which is the marker of gene silencers. As shown in Figure 6A, the gene *OTX1* in the top hierarchy is decorated with strong H3K27me3 and weak H3K27ac signals around its promoter. We investigated contacts involving loci containing H3K27ac peaks (enhancer-like) in the highest hierarchy class (HEs) and H3K27me3-rich loci (silencer-like) in the highest hierarchy class (HSs). We observed contacting preferences of HEs for colon TSGs and HSs for colon complementary genes, respectively. Specifically, HEs had significantly more observed/expected (O/E) contacts with TSGs than with complementary genes, whereas the situation is reversed for HSs (Fig. 5F). The O/E contacts were calculated as the contact strength divided by the average contact strength at corresponding sequential distances. Moreover, contact preferences existed among HEs in normal cells, compared to those among loci containing H3K27ac peaks (enhancer-like) in the lowest hierarchy class (LEs) (Fig. 5G).

H-loci were involved in more significantly decreased long-range contacts in CRC compared to loci in lower hierarchy (Fig. 5H). Similar results were observed for the subsets of contacts between promoter loci and loci containing H3K27ac-peaks (E-Ps) and between promoter loci and H3K27me3-rich loci (S-Ps) (Supplemental Fig. S8). Although HEs and LEs have comparable

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

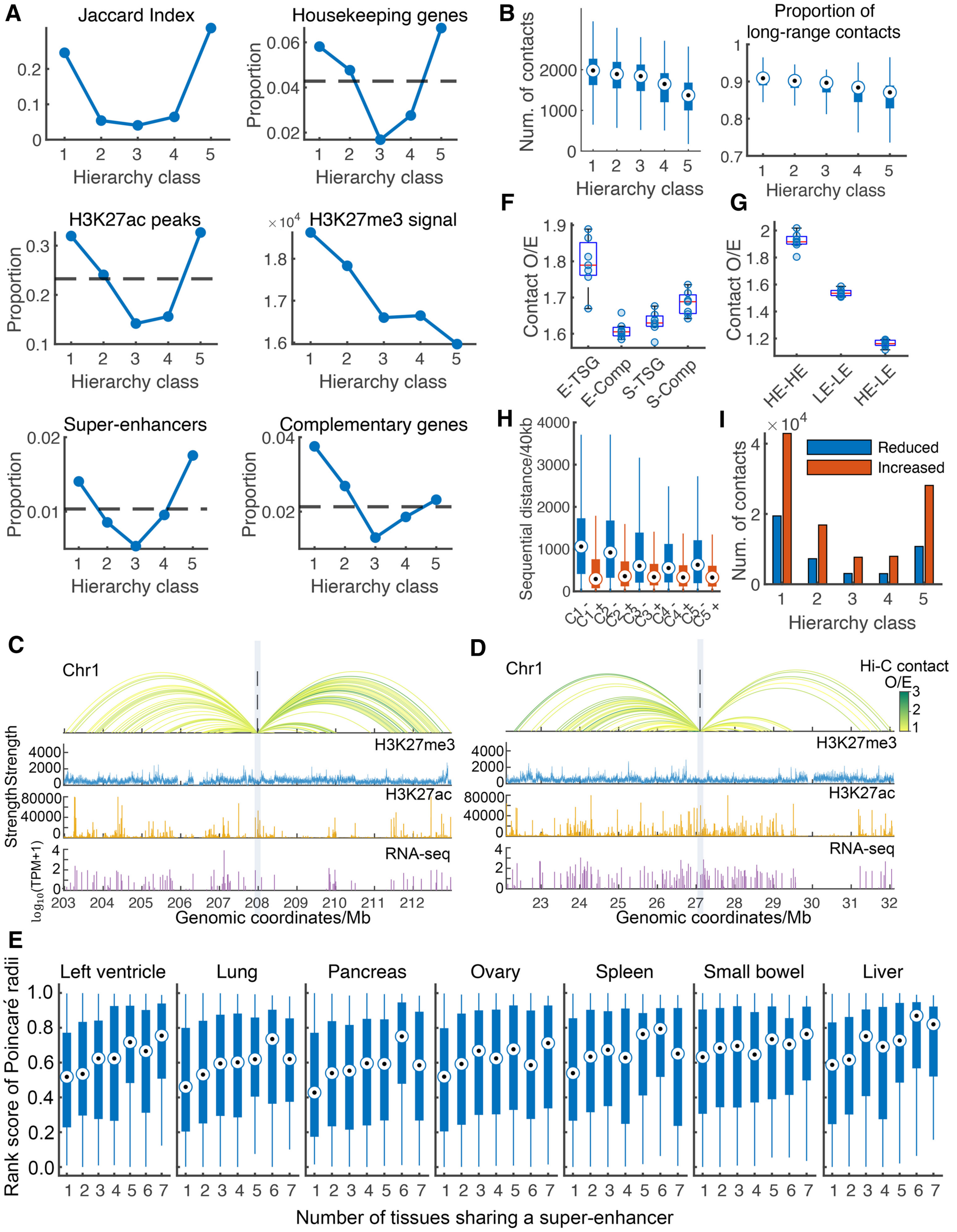

Figure 5. Properties of chromatin segments in each hierarchy class. (*A*) Proportion of loci (40 kb segments) containing H3K27ac peaks, super-enhancers, colon complementary genes, or housekeeping genes in each hierarchy class, as well as the Jaccard index (JI) and the average H3K27me3 signal of each hierarchy class. Numbers *beneath* the *x*-axis denote the hierarchy classes for normal colon (e.g., class 1 is the highest hierarchy class). The horizontal dashed lines indicate the expected proportion of loci containing regulatory elements/genes in all loci. (*B*) Boxplots showing the number of formed contacts and the proportion of long-range (>4 Mb) contacts for loci in each hierarchy class for normal colon. Genomic tracks of exemplary super-enhancers (SEs) from the hierarchy class 1 (*C*) and class 5 (*D*) in the normal colon. The *top* panels are virtual 4C traits at the SE center baits; Hi-C contacts with observed/expected (O/E) strength larger than 1 are displayed. The O/E contacts were calculated as the contact strength divided by the average contact strength at corresponding sequential distances. The colors of 4C traits represent the O/E contact strength. The *second* and the *third* panels are H3K27me3 and H3K27ac signals, respectively. The *bottom* panels are RNA-seq data, the y coordinates are $\log_{10}$(TPM + 1). (*E*) Boxplots showing rank scores by R_P for SEs in each tissue. For one primary tissue, each SE was classified into groups by the number of tissues sharing it. The rank score of a SE was calculated as its rank in R_P divided by the number of entries in R_P. A smaller rank score means a higher relative hierarchy status for a SE. (*F*) Boxplots showing the observed/expected (O/E) contact frequencies between colon tissue-specific genes (TSGs) or colon complementary genes and potential regulatory elements in the highest hierarchy class in normal colon. "E," "S," and "Comp" denote loci containing H3K27ac peaks (enhancer-like), with rich H3K27me3 signals (silencer-like) and containing colon complementary genes, respectively. (*G*) Boxplot showing O/E contact frequencies within the HEs, within the LEs, between HEs and LEs in normal colon. HE and LE denote loci containing H3K27ac peaks (enhancer-like) in the highest hierarchy class and those in the lowest hierarchy class, respectively. (*H*) Boxplots showing sequential distances of significantly changed contacts from normal colon to CRC for each hierarchy class on Chromosome 1. For instance, "C1−" denotes the decreased contacts involving loci in the highest hierarchy class, whereas "C5+" denotes the increased contacts involving loci in the lowest hierarchy class. (*I*) Bars showing numbers of significantly changed E-P and S-P contacts in each hierarchy class from normal colon to CRC. E-P and S-P denote contacts between promoter loci and loci containing H3K27ac-peaks (enhancer-like) and between promoter loci and H3K27me3-rich (silencer-like) loci, respectively.

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

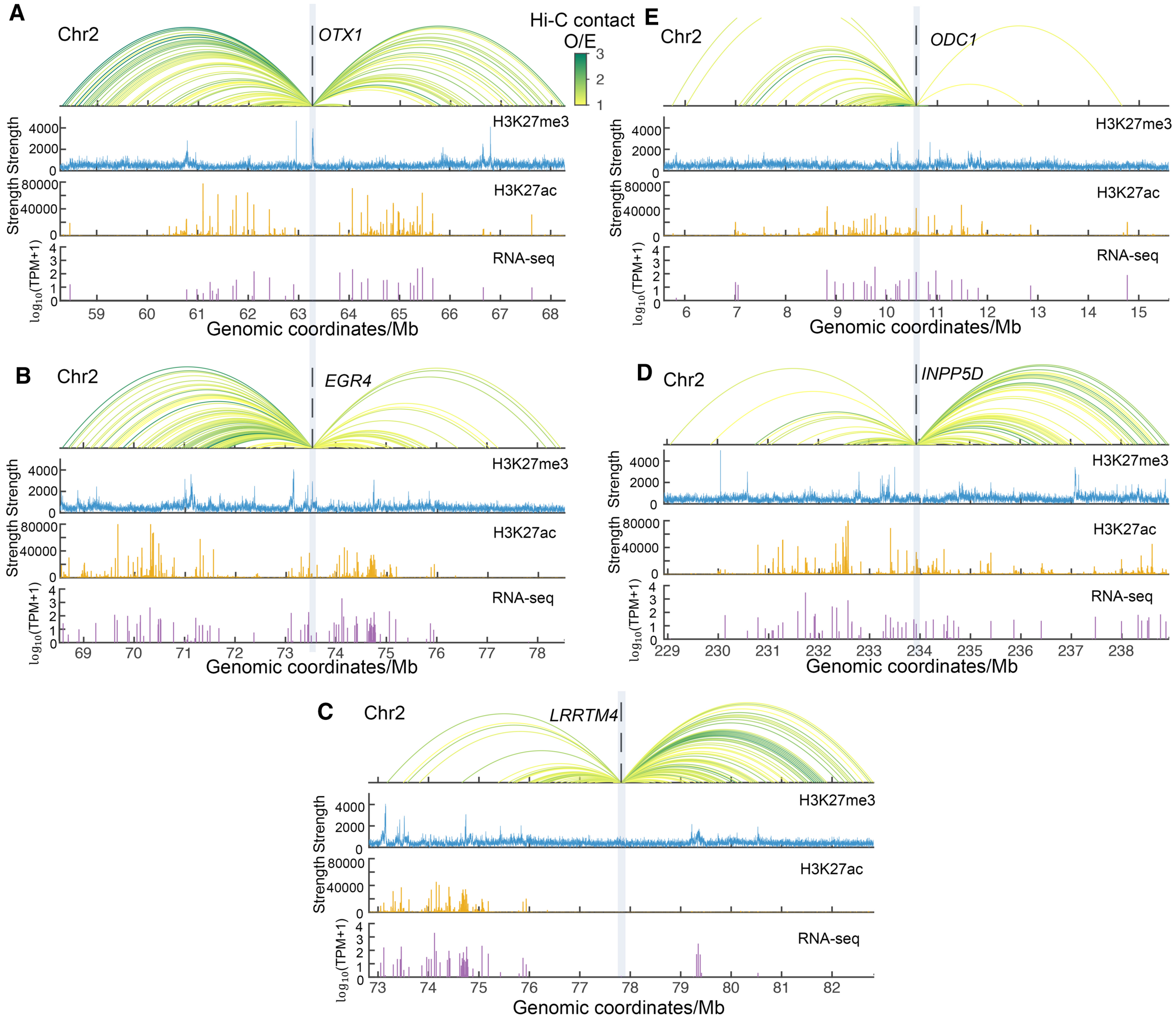

Figure 6. Genomic tracks of exemplary genes from each hierarchy class. Genomic tracks of genes *OTX1*, *EGR4*, *LRRTM4*, *INPP5D*, and *ODC1* from the hierarchy class 1, 2, 3, 4, and 5 in the normal colon are shown in panels *A*, *B*, *C*, *D*, and *E*, respectively. The *top* panels are virtual 4C traits at the gene promoter baits; Hi-C contacts with observed/expected (O/E) strength larger than 1 are displayed. The O/E contacts were calculated as the contact strength divided by the average contact strength at corresponding sequential distances. The colors of 4C traits represent the O/E contact strength. The *second* and the *third* panels are H3K27me3 and H3K27ac signals, respectively. The *bottom* panels are RNA-seq data; the y coordinates are $\log_{10}$(TPM + 1). TPM is averaged in the normal colons from the TCGA database.

amounts, more contacts between active promoters and HEs underwent significant changes in cancer, compared with those between active promoters and LEs (Fig. 5I).

We also found the potency that regulatory elements in H-loci had in altering gene transcription through rewired contacts in cancer. We observed stronger impacts of reshuffled E-P contacts on transcription for HEs, compared with those for LEs. Specifically, by identifying significantly changed E-Ps, we found a positive correlation between the number of HEs that an active promoter lost contact with and its transcription down-regulation (Fig. 7A). In addition, increased contacts between promoters and HEs were accompanied by the up-regulation of genes in cancer. In comparison, the relevance of de novo increased contacts to transcription up-regulation is less evident for LE-Ps (Fig. 7A), implying more potent roles of HEs in gene activation. On the other hand, the significantly changed S-P contacts are negatively correlated with transcription alteration (Fig. 7A), validating the repressive impacts of H3K27me3-rich loci via physical interactions. To control for variability related to copy number variants (CNVs), we identified significantly changed contacts for genomic regions with rare CNV occurrence (see in Methods).

IGSF9 is in the top hierarchy in the normal colon and is significantly down-regulated in expression in CRC ($FC_{TPM} = -1.67$, $P = 6.76 \times 10^{-19}$). As shown in Figure 7, B and C, promoter of *IGSF9* lost contacts with distant strong H3K27ac peaks and gained contacts with nearer H3K27ac-poor regions in CRC. Because long-range contacts have high expected BC (Fig. 1B), losing them may bring about severely decreased BC for corresponding edges. For *IGSF9*, the lost contacts with distant H3K27ac peaks are manifested as BC fold-changes near −1, whereas gained short-range contacts are manifested as BC fold-changes with small absolute values. Changes in BC between gene promoters and potential regulators reflect the rewired Hi-C contacts, which may lead to transcription misregulation in cancer. For instance, *IGSF9* inhibits breast cancer

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

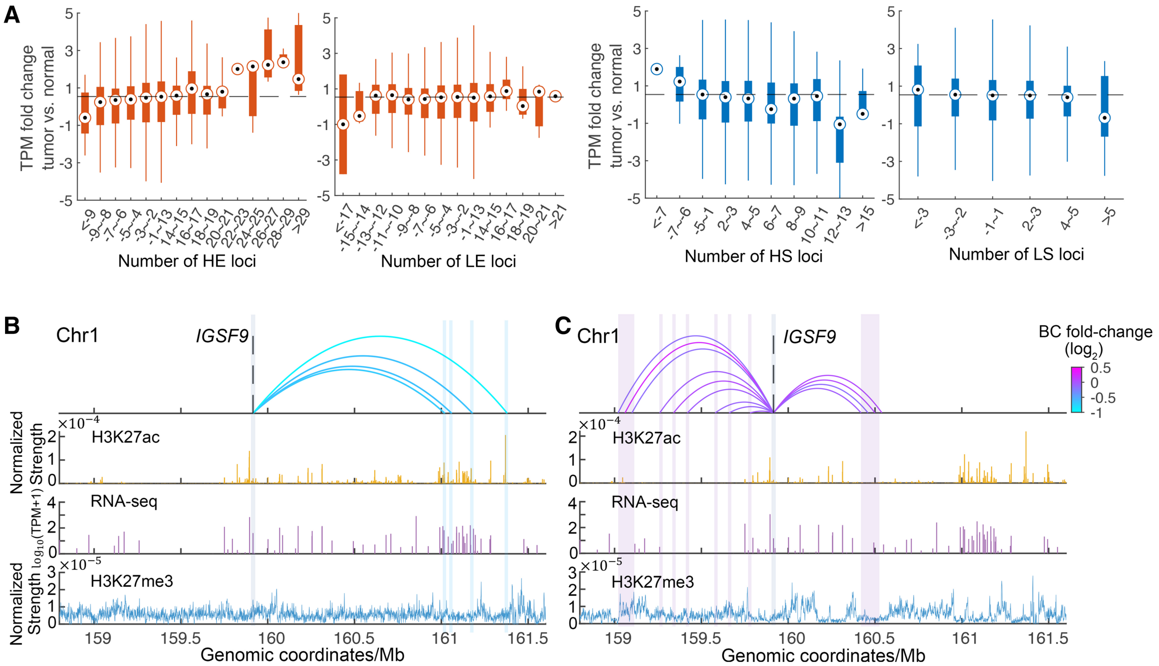

Figure 7. Correlation between gene expression changes and rewired contacts in cancer. (*A*) In each box, the x coordinate is the number of loci that a gene significantly ($P < 0.05$) lost or gained contacts with from normal colon to CRC; the y coordinate is TPM fold change ($\log_2$). A negative or a positive x coordinate means losing or gaining contacts with loci of corresponding type. HE and LE denote loci containing H3K27ac-peaks (enhancer-like) in the highest hierarchy class and those in the lowest hierarchy class, whereas HS and LS denote H3K27me3-rich (silencer-like) loci in the highest hierarchy class and those in the lowest hierarchy class, respectively. Changes of genomic tracks from normal colon (*B*) to colorectal cancer (*C*) for the gene *IGSF9*. Significantly ($P <$ 0.05) changed Hi-C contacts from normal colon to colorectal cancer at the gene promoter bait are shown in the *top* panel, where decreased and increased contacts are shown in panels *B* and *C*, respectively. The colors of contact traits represent betweenness centrality (BC) fold-change from normal colon to colorectal cancer. The *second* and the *third* panels are H3K27ac signals and RNA-seq data, respectively. TPM is averaged in the normal colons and the CRC samples from the TCGA database. The *bottom* panels are H3K27me3 signals. The ChIP-seq signals were normalized by corresponding sequencing depth.

metastasis and is down-regulated in breast cancer, melanoma, and colorectal polyposis (Smith et al. 2005; Bowden et al. 2007; Li et al. 2022).

Loci with the highest hierarchy statuses (in the central region of the Poincaré disk) underwent the most significant hierarchy degradation in CRC (Fig. 8A). Genes with the highest hierarchy statuses in normal cells contain the largest ratio of down-regulated genes in CRC, whereas genes with the lowest hierarchy statuses contain the largest ratio of up-regulated genes (Fig. 8B). We derived the gene sets in the highest hierarchy class in normal colon and CRC, respectively. Active genes in the highest hierarchy class in normal cells are enriched in pathways including intracellular signal transduction (e.g., *NPR1* and *RPS6KA2*) and apoptotic process (e.g., *PTEN* and *ERCC3*) (Fig. 8C; Supplemental Table S6). Genes constituting the most enriched pathways were generally down-regulated in cancer (Fig. 8C). Active genes in the highest hierarchy class in CRC are enriched in pathways including replication fork processing (e.g., *MMS22L* and *NUCKS1*) and DNA repair (e.g., *UBE2T* and *DNA2*), which were generally up-regulated in cancer (Fig. 8D; Supplemental Table S7).

Discussion

Decreased global information transmission efficiency and hierarchy depth in cancer chromatin

In a CCN, a physical contact may transmit latent genetic signals (e.g., enhancer–promoter interactions). Bridge edges in a network guarantee the efficient routing of signals (Krioukov et al. 2010). In chromatin, long-range interactions are bridge-like edges with high information flow (Fig. 1B). Loss of long-range contacts was observed in cancer cell lines compared with normal cells (Xue et al. 2022). In this study, we observed loss of long-range contacts and gain of short-range contacts in cancer tissues (Fig. 1C). Such redistribution of contacts may lead to a hindered global communication among chromatin loci. The reduction in global information transmission efficiency was manifested by a decrease in CC as well as an increase in ASPL.

Centralities and path length are both important network properties related to the network hierarchy (Krioukov et al. 2010; Mones et al. 2012; Rajeh et al. 2020). We next performed Poincaré embedding to reveal the latent hierarchy in CCNs. Chromatin segments with high hierarchy statuses in normal cells underwent severe hierarchy down-gradation in cancer, reflected by increased Poincaré radii (R_P). It is known that targeted attacks on high-hierarchy nodes bring about severe loss of network efficiency and the global hierarchy (Zamani et al. 2018). The attenuated top-down hierarchical characteristic in cancer resulted in a more homogeneous distribution of node statuses. Specifically, with decreased long-range contacts and increased short-range contacts in cancer, linearly adjacent loci were less distinguished and tended to locate near in the embedding space. Resultantly, the embedded structures were delineated by linear configurations lacking wrinkled characteristics.

We observed rewired E-P and S-P contacts related to gene transcription alteration in CRC (Fig. 7A). Promoters losing or gaining contacts with a large number of HEs coincided with significant changes in transcription. Such quantitative effect may reflect the

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

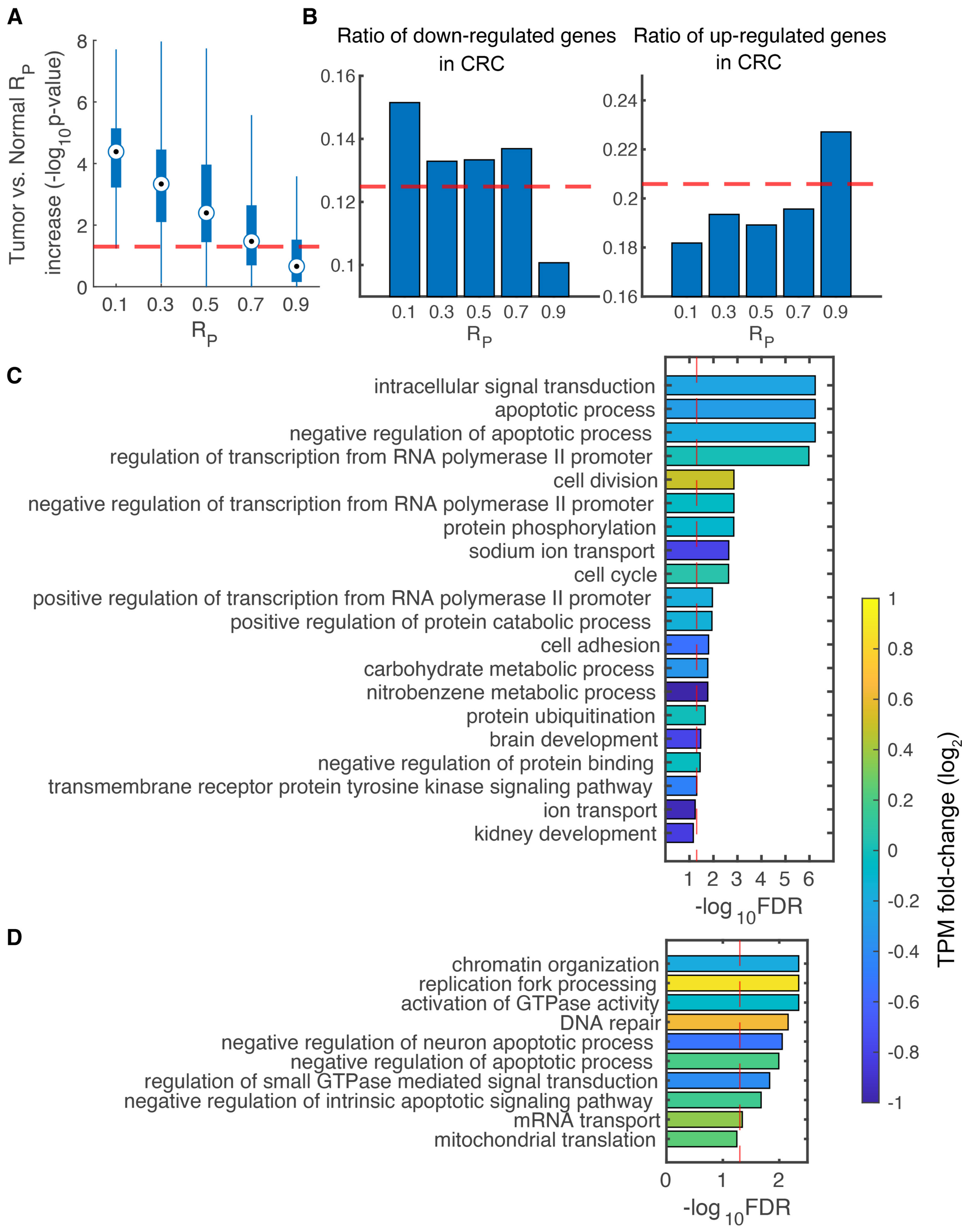

Figure 8. Correlation between differential gene expression and CCN hierarchy in cancer. (*A*) Boxplot showing significance [$-\log_{10}$(*P*-value)] of R_P increase from normal colon to CRC. Loci were grouped according to R_P of normal colon. For instance, 0.1 in the *x* axis means the annular area defined by a radius ranging from 0 to 0.2 in the Poincaré disk. The horizontal dashed line indicates $-\log_{10}(0.05)$. (*B*) Bars showing the ratio of down-regulated genes (*left*) and that of up-regulated genes (*right*) in CRC for genes in each group. The horizontal dashed lines indicate the proportion of corresponding differentially expressed genes in all genes. Enriched functions for colon-specific (*C*) and CRC-specific genes (*D*) in the highest hierarchy class. For each bar representing a pathway, the average TPM fold change ($\log_2$) of involved genes from normal colon to CRC is manifested by the bar color. The vertical dashed lines indicate FDR-adjusted *P*-value equaling 0.05.

coordination of multiple enhancers in activating a single promoter. In fact, deletion of individual enhancers not necessarily affect the transcription of target genes (Osterwalder et al. 2018; Lin et al. 2022). Moreover, contact preferences exist among HEs in normal cells, compared with those among LEs (Fig. 5G). These results may indicate that HEs form an active transcription environment via spatial contacting in normal cells. In fact, enhancers and active promoters can form clusters in 3D space and activate each other (Zhu et al. 2021). In cancer, with severe loss of long-range contacts (Fig. 5H), it is possible that HEs disassembled and genes therein were deactivated, resulting in down-regulation of genes (Fig. 8B).

In this work, the used Hi-C data sets were from bulk tissues, which reflected population-averaged chromatin contacts. However, it is unknown whether the transcription of certain genes and the mediating contacts exist in each individual cell at one time point. To examine this, we interrogated the published single-cell

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

imaging data with single-cell RNA-seq of IMR-90 cell lines (Su et al. 2020). We found several genes having discrepant distributions of physical distances with other loci under transcription on and off states. One example is *LCA5L*, which tends to approach a distant locus under active transcription states (Supplemental Fig. S9). Our observation reflects that the transcription state of certain genes stratifies individual cells into subsets with distinct average contacting profiles. In this aspect, population-averaged chromatin contacts may better correspond to transcription levels of bulk tissues. We used Hi-C data sets for graph analyses and embedding. However, our methods can be extended to multidirectional interactions by decomposing multidirectional interactions into combinations of pairwise interactions.

Changes in CCN hierarchy in cancer and cell development

We found that Poincaré embedding can sensitively reveal variations in chromatin structure especially in processes during which the CCN hierarchy depth significantly changes, including canceration and cell development. The CCN hierarchy of blastocyst embryo is initially built up but only become completely established upon further differentiation into downstream monofunctional cells. In cancer, CCN hierarchy collapses to resemble stem cells.

The shallow CCN hierarchy of cancer and pluripotent embryos is reminiscent of the correlation between network stability and hierarchy. The more hierarchical a network is, the more resistance and resilience it has against disturbance (Liu et al. 2012; Zamani et al. 2018). However, a network with deep hierarchy can jump to a state with shallow hierarchy under strong disturbance (Zamani and Vicsek 2017). Because network instability leads to plasticity (Prinz 2008; Mermillod et al. 2013), we speculate that the shallow hierarchy or quasi-1D-like characteristics of chromatin in embryos are related to their proficiency to differentiate. For cancer cells, structural instability in chromatin may favor the occurrence of tumor heterogeneity. Based on partial similarity in chromatin structure and substantial differences in cell nature between cancer and stem cells, it would be interesting to understand their genetic differences.

In this study, we systematically performed graph analyses and Poincaré embedding from Hi-C data sets to investigate the role of chromatin structure hierarchy in cancer and cell development. We found that chromatin segments with high hierarchy statuses contribute to cell state maintenance. The top-down hierarchy of chromatin structure establishes during cell development and collapses in cancer. The significant changes in graph statistics of CCNs suggest potential usage in cancer risk stratification.

Methods

Hi-C normalization

For each sample, normalization was performed on the raw Hi-C matrix using the iterative correction and eigenvector decomposition (ICE) method (Imakaev et al. 2012), yielding the normalized Hi-C matrix. The normalization was performed on intrachromosomal Hi-C matrices with 40 kb resolution and the whole-chromatin Hi-C matrices with 100 kb resolution.

Closeness centrality (CC), average shortest path length (ASPL) and betweenness centrality (BC) in chromatin contact networks (CCNs)

We constructed the CCNs through a network inference procedure from Hi-C. The input of the procedure is an ICE-normalized intrachromosomal Hi-C matrix of 40 kb resolution. We first calculated a correlation matrix by computing Pearson correlation coefficients between rows of the Hi-C matrix. Pearson correlation method was used in inferring effective connectivity strength between nodes in biological networks (Sun et al. 2014; Tijms et al. 2014). Then we assigned edges with the top 5% highest Pearson correlation coefficients as the effective links in a CCN, where each node is a 40-kb genomic segment. In this way, an unweighted CCN was derived and fed into following calculation of network metrics including CC, ASPL, and BC.

CC was calculated as the inverse sum of the shortest paths from a node to all other nodes in the graph (Golbeck 2013),

$$CC(i) = \left(\frac{A_i}{N-1}\right)^2 \frac{1}{C_i}.$$

$\boldsymbol{A_i}$ is the number of reachable nodes for node i (excluding i), $\boldsymbol{N}$ is the number of nodes in G, and $\boldsymbol{C_i}$ is the sum of shortest paths from node i to all reachable nodes. The shortest paths of all node pairs were calculated with breadth-first search algorithm on unweighted graphs (Franciosa et al. 1997).

The BC of an edge e is the sum of the fraction of all pair-wise shortest paths that pass through e (Freeman 1977):

$$BC(e) = \sum_{s,t \in V} \frac{\sigma(s,t|e)}{\sigma(s,t)},$$

where V is the set of nodes, $\sigma(s, t)$ is the number of shortest (s, t)-paths, and $\sigma(s, t|e)$ is the number of these paths passing through e.

Gene expression analysis

We downloaded all available RNA-seq data of CRC and normal colons from the TCGA database, which were then converted into TPM format. TPM fold change of each gene from normal colons to colon tumors was calculated as follows:

$$\mathrm{FC_{TPM}} = \log_2\left(\frac{\mathrm{TPM_{cancer}} + 1}{\mathrm{TPM_{normal}} + 1}\right).$$

$\mathrm{TPM_{cancer}}$ and $\mathrm{TPM_{normal}}$ of a gene were calculated by averaging TPM in all cancer and normal samples, respectively. We performed differential gene expression analysis with the R/Bioconductor package "DESeq2" (Love et al. 2014). A DEG was defined as a gene with an absolute $\mathrm{FC_{TPM}}$ larger than 1 and a P-value lower than 0.05.

Gene function analysis

In this study, Gene Ontology (GO) enrichment analysis of interested genes were performed using the R package clusterProfiler (Yu et al. 2012).

Embedding chromatin contact networks in the Poincaré disk

We used the ICE-normalized Hi-C matrices as the input of network embedding. An intrachromosomal Hi-C matrix with 40 kb resolution was used in embedding an intrachromosomal CCN, where each node is a 40-kb genomic segment. We first calculated the pairwise contact probabilities among genomic segments from the Hi-C matrix. For the largest 25% nonzero entries in a Hi-C matrix, their corresponding contact probabilities were assigned as 1. The rest of the nonzero entries were divided by the 75th quantile of the Hi-C matrix and assigned as corresponding contact probabilities. The probability matrix was the adjacency matrix of a probabilistic graph, where we used a Monte Carlo strategy to embed into the Poincaré disk. At each training epoch, we sampled an unweighted adjacency matrix based on the probability matrix. In embedding CCNs of the whole chromatin, we used Hi-C matrices of 100-kb

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

resolution. For entries ranging from the 85th to 95th quantile in the Hi-C matrix, their corresponding edges were assigned as effective in the contact network. The unweighted adjacency matrix was used through all training epochs.

At each training epoch, the unweighted graph was embedded in the Poincaré disk based on the Riemannian optimization (Nickel and Kiela 2017). Loss function was optimized by encouraging connected nodes to be close in the Poincaré disk:

$$\mathcal{L}(\Theta) = \sum_{(u,v)\in\mathcal{D}} \log \frac{e^{-d(\boldsymbol{u},\boldsymbol{v})}}{\sum_{\boldsymbol{v}'\in\mathcal{N}(u)} e^{-d(\boldsymbol{u},\boldsymbol{v}')}},$$

where $\mathcal{N}(u) = \{v|(u, v) \notin \mathcal{D}\} \cup \{u\}$ is the set of unconnected nodes for the node u. Θ is the set of embedded coordinates of all nodes in the Poincaré disk. The hyperbolic distance $d(\boldsymbol{u}, \boldsymbol{v})$ was calculated as follows:

$$d(\boldsymbol{u}, \boldsymbol{v}) = \operatorname{arcosh}\left(1 + 2\frac{\| \boldsymbol{u} - \boldsymbol{v} \|^2}{(1 - \| \boldsymbol{u} \|^2)(1 - \| \boldsymbol{v} \|^2)}\right),$$

where $\boldsymbol{u}$, $\boldsymbol{v}$ are the embedded coordinates of node u, v. The same training protocol and experimental setting was applied to all individual Hi-C samples.

Interchromosomal closeness matrix

We first calculated the average interchromosomal Poincaré distances for each pair of autosomes for each embedded structure,

$$\overline{d_{mn}} = \sum_{i\in V_m, j\in V_n} d_{ij}/N_m N_n,$$

where V_m and V_n are sets of nodes belonging to Chromosomes m and n, respectively, d_{ij} is the Poincaré distance between node i and j, N_m and N_n are the size of set V_m and V_n, respectively. Each element in the interchromosomal distance matrix (22 × 22) were then divided by the geometric average of the corresponding pair of average intrachromosomal distances, to evaluate the closeness between each chromosome pair:

$$d_{mn}^{clo} = \overline{d_{mn}}/\sqrt{\overline{d_{mm}} \times \overline{d_{nn}}}.$$

Each element in the closeness matrix was further normalized by subtraction of 1 and division by the maximum element, so that matrix values range from 0 to 1 and the intrachromosomal closeness equals 0. The matrix entry ranging from 0 to 1 represents the closeness between each pair of chromosomes.

ChIP-seq analysis

The H3K27ac and H3K27me3 ChIP-seq data of normal colon and CRC were downloaded from the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/), accession number GSE133928. The enhancer-associated mark H3K27ac is punctate peaks, whereas the repressive modification H3K27me3 is in broader regions. H3K27ac peaks were called using MACS2 (Zhang et al. 2008) with default parameters. H3K27me3 signals were merged for each 40kb loci. The H3K27me3-rich loci were defined as loci with the top 10% strongest H3K27me3 signals.

Definition of tissue-specific genes and complementary tissue-specific genes

The normalized RNA-seq data of the GTEx project (The GTEx Consortium 2013) was downloaded from Sonawane et al. (2017). The tissue specificity of gene i in tissue t was defined as

$$s_i^t = \frac{\varepsilon_i^t - \mu_i^{all}}{\mu_i^{all}},$$

where ε_i^t and μ_i^{all} are the mean expression level of gene i in tissue t and all tissues examined, respectively. A gene with a tissue specificity value greater than 2 was defined as a TSG. All tissue specific genes except for those in tissue t are defined as complementary TSGs for tissue t.

Controlling for copy number variants (CNVs)

We downloaded the inferred CNVs for 26 CRC samples published by Johnstone et al. (Johnstone et al. 2020). For each 40-kb locus, its CNV frequency was calculated as the number of CRC samples in which the locus was overlapped with CNV regions. The loci with rare CNV occurrences were defined as CNV probability < 0.2, that is, loci that were overlapped with CNV in less than 5 CRC samples in the 26 CRC samples.

Statistics

The significance levels of inequality in graph metrics and expression levels between cancer/embryo samples and their corresponding normal controls were calculated under the one-sided Welch's *t*-test.

Visualization methods

Statistical graphs and visualization of embedded CCNs in the Poincaré disk were drawn by MATLAB codes. Visualization of networks in Figure 1A was performed by Python package NetworkX.

Data sets

The data that support the findings of this study were derived from the following resources available in the public domain: GEO, TCGA, and SEdb. The detailed data accession for Hi-C data, RNA-seq and super-enhancer location can be found in Supplemental Tables S1, S2, and S5, respectively.

Software availability

The scripts and source codes for graph analyses and Poincaré embedding based on Hi-C data sets are available as Supplemental Code and at GitHub (https://github.com/wjychem/GGL).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (92053202, 22050003, 21821004). The results shown here are partly based on data generated by the TCGA Research Network (http://cancergenome.nih.gov).

Author contributions: Conceptualization, Y.Q.G.; data curation, J.Y.W., Y.X., Y.Y.H., H.Q., and J.Z.; formal analysis, J.Y.W., Y.X., and Y.Y.H.; funding acquisition, Y.Q.G.; investigation, J.Y.W., Y.X., Y.Y.H., H.Q., and J.Z.; supervision, Y.Q.G.; visualization, J.Y.W.; writing—original draft, J.Y.W., Y.X., YY.H., H.Q., and J.Z.; writing—review and editing, J.Y.W., Y.X., and Y.Q.G. All authors have read and agreed to the published version of the manuscript.

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

References

Ahmad T, Li XJ, Seet B-C, Cano J-C. 2020. Social network analysis based localization technique with clustered closeness centrality for 3D wireless sensor networks. *Electronics (Basel)* **9:** 738. doi:10.3390/electronics9050738

Balint K, Xiao M, Pinnix CC, Soma A, Veres I, Juhasz I, Brown EJ, Capobianco AJ, Herlyn M, Liu ZJ. 2005. Activation of Notch1 signaling is required for β-catenin-mediated human primary melanoma progression. *J Clin Invest* **115:** 3166–3176. doi:10.1172/JCI25001

Barth R, Bystricky K, Shaban HA. 2020. Coupling chromatin structure and dynamics by live super-resolution imaging. *Sci Adv* **6:** eaaz2196. doi:10.1126/sciadv.aaz2196

Bowden NA, Croft A, Scott RJ. 2007. Gene expression profiling in familial adenomatous polyposis adenomas and desmoid disease. *Hered Cancer Clin Pract* **5:** 79–96. doi:10.1186/1897-4287-5-2-79

Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet* **10:** 211–219. doi:10.1093/hmg/10.3.211

Brazhnik P, de la Fuente A, Mendes P. 2002. Gene networks: how to put the function in genomics. *Trends Biotechnol* **20:** 467–472. doi:10.1016/S0167-7799(02)02053-X

Brekelmans C, Seynaeve C, Bartels C, Tilanus-Linthorst M, Meijers-Heijboer E, Crepin C, van Geel A, Menke M, Verhoog L, van Den Ouweland A, et al. 2001. Effectiveness of breast cancer surveillance in *BRCA1/2* gene mutation carriers and women with high familial risk. *J Clin Oncol* **19:** 924–930. doi:10.1200/JCO.2001.19.4.924

Cai Y, Zhang Y, Loh YP, Tng JQ, Lim MC, Cao Z, Raju A, Lieberman Aiden E, Li S, Manikandan L, et al. 2021. H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions. *Nat Commun* **12:** 719. doi:10.1038/s41467-021-20940-y

Chen X, Ke Y, Wu K, Zhao H, Sun Y, Gao L, Liu Z, Zhang J, Tao W, Hou Z, et al. 2019. Key role for CTCF in establishing chromatin structure in human embryos. *Nature* **576:** 306–310. doi:10.1038/s41586-019-1812-0

Cloney R. 2016. Dynamic enhancer–promoter interactions for transcriptional bursting. *Nat Rev Genet* **17:** 437–437. doi:10.1038/nrg.2016.81

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci* **107:** 21931–21936. doi:10.1073/pnas.1016071107

Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, et al. 2002. Mutations of the *BRAF* gene in human cancer. *Nature* **417:** 949–954. doi:10.1038/nature00766

Denais C, Lammerding J. 2014. Nuclear mechanics in cancer. In *Cancer biology and the nuclear envelope: Recent advances may elucidate past paradoxes* (ed. Schirmer EC, de las Heras JI), pp. 435–470. Springer, New York. doi:10.1007/978-1-4899-8032-8_20

Deng S, Feng Y, Pauklin S. 2022. 3D chromatin architecture and transcription regulation in cancer. *J Hematol Oncol* **15:** 49. doi:10.1186/s13045-022-01271-x

Ding B, Liu Y, Liu Z, Zheng L, Xu P, Chen Z, Wu P, Zhao Y, Pan Q, Guo Y, et al. 2021. Noncoding loci without epigenomic signals can be essential for maintaining global chromatin organization and cell viability. *Sci Adv* **7:** eabi6020. doi:10.1126/sciadv.abi6020

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518:** 331–336. doi:10.1038/nature14222

Eisenberg D, Marcotte EM, Xenarios I, Yeates TO. 2000. Protein function in the post-genomic era. *Nature* **405:** 823–826. doi:10.1038/35015694

Feng Y, Liu X, Pauklin S. 2021. 3D chromatin architecture and epigenetic regulation in cancer stem cells. *Protein Cell* **12:** 440–454. doi:10.1007/s13238-020-00819-2

Franciosa PG, Frigioni D, Giaccio R. 1997. Semi-dynamic shortest paths and breadth-first search in digraphs. In *STACS 97* (ed. Reischuk R, Morvan M), pp. 33–46. Springer, Berlin, Heidelberg.

Freeman LC. 1977. A set of measures of centrality based on betweenness. *Sociometry* **40:** 35–41. doi:10.2307/3033543

Fujimoto K, Yamada Y, Okajima E, Kakizoe T, Sasaki H, Sugimura T, Terada M. 1992. Frequent association of p53 gene mutation in invasive bladder cancer. *Cancer Res* **52:** 1393–1398.

Girvan M, Newman MEJ. 2002. Community structure in social and biological networks. *Proc Natl Acad Sci* **99:** 7821–7826. doi:10.1073/pnas.122653799

Golbeck J. 2013. Chapter 2 - nodes, edges, and network measures. In *Analyzing the social Web* (ed. Golbeck J), pp. 9–23. Morgan Kaufmann, Boston. doi:10.1016/B978-0-12-405531-5.00002-X

Grover A, Leskovec J. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, San Francisco, pp. 855–864.

The GTEx Consortium. 2013. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45:** 580–585. doi:10.1038/ng.2653

He Y, Xue Y, Wang J, Huang Y, Liu L, Huang Y, Gao YQ. 2022. Computational enhanced Hi-C data reveals the function of structural geometry in genomic regulation. bioRxiv doi:10.1101/2022.07.12.499232

Hendrix MJ, Seftor EA, Seftor RE, Kasemeier-Kulesa J, Kulesa PM, Postovit LM. 2007. Reprogramming metastatic tumour cells with embryonic microenvironments. *Nat Rev Cancer* **7:** 246–255. doi:10.1038/nrc2108

Huang J, Li K, Cai W, Liu X, Zhang Y, Orkin SH, Xu J, Yuan G-C. 2018. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun* **9:** 943. doi:10.1038/s41467-018-03279-9

Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9:** 999–1003 doi:10.1038/nmeth.2148

Iyyanki T, Zhang B, Wang Q, Hou Y, Jin Q, Xu J, Yang H, Liu T, Wang X, Song F, et al. 2021. Subtype-associated epigenomic landscape and 3D genome structure in bladder cancer. *Genome Biol* **22:** 105. doi:10.1186/s13059-021-02325-y

Johnstone SE, Reyes A, Qi Y, Adriaens C, Hegazi E, Pelka K, Chen JH, Zou LS, Drier Y, Hecht V, et al. 2020. Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **182:** 1474–1489.e23. doi:10.1016/j.cell.2020.07.030

Kai Y, Li BE, Zhu M, Li GY, Chen F, Han Y, Cha HJ, Orkin SH, Cai W, Huang J, et al. 2021. Mapping the evolving landscape of super-enhancers during cell differentiation. *Genome Biol* **22:** 269. doi:10.1186/s13059-021-02485-x

Kaur J, Daoud A, Eblen ST. 2019. Targeting chromatin remodeling for cancer therapy. *Curr Mol Pharmacol* **12:** 215–229. doi:10.2174/1874467212666190215112915

Kim J, Orkin SH. 2011. Embryonic stem cell-specific signatures in cancer: insights into genomic regulatory networks and implications for medicine. *Genome Med* **3:** 75. doi:10.1186/gm291

Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M, et al. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* **163:** 134–147. doi:10.1016/j.cell.2015.08.040

Koutrouli M, Karatzas E, Paez-Espino D, Pavlopoulos GA. 2020. A guide to conquer the biological network era using graph theory. *Front Bioeng Biotechnol* **8:** 34. doi:10.3389/fbioe.2020.00034

Krioukov D, Papadopoulos F, Kitsak M, Vahdat A, Boguñá M. 2010. Hyperbolic geometry of complex networks. *Phys Rev E* **82:** 036106. doi:10.1103/PhysRevE.82.036106

Kuhn RM, Haussler D, Kent WJ. 2013. The UCSC genome browser and associated tools. *Brief Bioinformatics* **14:** 144–161. doi:10.1093/bib/bbs038

Latora V, Marchiori M. 2001. Efficient behavior of small-world networks. *Phys Rev Lett* **87:** 198701. doi:10.1103/PhysRevLett.87.198701

Levantini E, Giorgetti A, Cerisoli F, Traggiai E, Guidi A, Martin R, Acampora D, Aplan PD, Keller G, Simeone A, et al. 2003. Unsuspected role of the brain morphogenetic gene *Otx1* in hematopoiesis. *Proc Natl Acad Sci* **100:** 10299–10303. doi:10.1073/pnas.1734071100

Li Y, Deng Y, Zhao Y, Zhang W, Zhang S, Zhang L, Wang B, Xu Y, Chen S. 2022. Immunoglobulin superfamily 9 (IGSF9) is *trans*-activated by p53, inhibits breast cancer metastasis via FAK. *Oncogene* **41:** 4658–4672. doi:10.1038/s41388-022-02459-8

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326:** 289–293. doi:10.1126/science.1181369

Lin X, Liu Y, Liu S, Zhu X, Wu L, Zhu Y, Zhao D, Xu X, Chemparathy A, Wang H, et al. 2022. Nested epistasis enhancer networks for robust genome regulation. *Science* **77:** 1077–1085. doi:10.1126/science.abk3512

Liu Y-Y, Slotine J-J, Barabási A-L. 2012. Control centrality and hierarchical structure in complex networks. *PLoS One* **7:** e44459. doi:10.1371/journal.pone.0044459

Liu S, Chen H, Ronquist S, Seaman L, Ceglia N, Meixner W, Chen P-Y, Higgins G, Baldi P, Smale S, et al. 2018. Genome architecture mediates transcriptional control of human myogenic reprogramming. *iScience* **6:** 232–246. doi:10.1016/j.isci.2018.08.002

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15:** 550. doi:10.1186/s13059-014-0550-8

Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, et al. 2018. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* **173:** 338–354.e15. doi:10.1016/j.cell.2018.03.034

Downloaded from genome.cshlp.org on June 22, 2026 . Published by Cold Spring Harbor Laboratory Press

Mao G, Zhang N. 2013. Analysis of average shortest-path length of scale-free network. *J Appl Math* **2013:** 865643. doi:10.1155/2013/865643

Mermillod M, Bugaiska A, Bonin P. 2013. The stability-plasticity dilemma: investigating the continuum from catastrophic forgetting to age-limited learning effects. *Front Psychol* **4:** 504. doi:10.3389/fpsyg.2013.00504

Mones E, Vicsek L, Vicsek T. 2012. Hierarchy measure for complex networks. *PLoS One* **7:** e33799. doi:10.1371/journal.pone.0033799

Nickel M, Kiela D. 2017. Poincaré embeddings for learning hierarchical representations. In *Advances in neural information processing systems* (ed. Guyon I, et al.), Vol. 30. Curran Associates, Inc., Red Hook, NY.

Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, et al. 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554:** 239–243. doi:10.1038/nature25461

Peng G, Tam PPL, Jing N. 2017. Lineage specification of early embryos and embryonic stem cells at the dawn of enabling technologies. *Natl Sci Rev* **4:** 533–542. doi:10.1093/nsr/nwx093

Perozzi B, Al-Rfou R, Skiena S. 2014. DeepWalk: online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, pp. 701–710.

Prinz AA. 2008. 16 - Stability and plasticity in neuronal and network dynamics. In *Computational neuroscience in epilepsy* (ed. Soltesz I, Staley K), pp. 247–258, VIII. Academic Press, San Diego. doi:10.1016/B978-012373649-9.50019-3

Rajeh S, Savonnet M, Leclercq E, Cherifi H. 2020. Interplay between hierarchy and centrality in complex networks. *IEEE Access* **8:** 129717–129742. doi:10.1109/ACCESS.2020.3009525

Ramani V, Shendure J, Duan Z. 2016. Understanding spatial genome organization: methods and insights. *Genomics Proteomics Bioinformatics* **14:** 7–20. doi:10.1016/j.gpb.2016.01.002

Sandhu KS, Li G, Poh Huay M, Quek Yu Ling K, Sia Yee Y, Peh Su Q, Mulawadi Fabianus H, Lim J, Sikic M, Menghi F, et al. 2012. Large-scale functional organization of long-range chromatin interaction networks. *Cell Rep* **2:** 1207–1219. doi:10.1016/j.celrep.2012.09.022

Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan C, Li Y, Lin S, Lin Y, Barr CL, et al. 2016. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep* **17:** 2042–2059. doi:10.1016/j.celrep.2016.10.061

Seftor EA, Meltzer PS, Schatteman GC, Gruman LM, Hess AR, Kirschmann DA, Seftor RE, Hendrix MJ. 2002. Expression of multiple molecular phenotypes by aggressive melanoma tumor cells: role in vasculogenic mimicry. *Crit Rev Oncol Hematol* **44:** 17–27. doi:10.1016/S1040-8428(01)00199-8

Shimizu N, Mori R. 2016. Average shortest path length of graphs of diameter 3. In *2016 Tenth IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*, Nara, Japan, pp. 1–6. doi:10.1109/NOCS.2016.7579335

Smith AP, Hoek K, Becker D. 2005. Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular differences between nevi/melanoma in situ and advanced-stage melanomas. *Cancer Biol Ther* **4:** 1018–1029. doi:10.4161/cbt.4.9.2165

Soler-Vila P, Cuscó P, Farabella I, Di Stefano M, Marti-Renom Marc A. 2020. Hierarchical chromatin organization detected by TADpole. *Nucleic Acids Res* **48:** e39. doi:10.1093/nar/gkaa087

Sonawane AR, Platig J, Fagny M, Chen C-Y, Paulson JN, Lopes-Ramos CM, DeMeo DL, Quackenbush J, Glass K, Kuijjer ML. 2017. Understanding tissue-specific gene regulation. *Cell Rep* **21:** 1077–1088. doi:10.1016/j.celrep.2017.10.001

Spruck CH, Strohmaier H, Sangfelt O, Müller HM, Hubalek M, Müller-Holzner E, Marth C, Widschwendter M, Reed SI. 2002. hCDC4 gene mutations in endometrial cancer. *Cancer Res* **62:** 4535–4539.

Su J-H, Zheng P, Kinrot SS, Bintu B, Zhuang X. 2020. Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* **182:** 1641–1659.e1626. doi:10.1016/j.cell.2020.07.032

Sun Y, Yin Q, Fang R, Yan X, Wang Y, Bezerianos A, Tang H, Miao F, Sun J. 2014. Disrupted functional brain connectivity and its association to structural connectivity in amnestic mild cognitive impairment and Alzheimer's disease. *PLoS One* **9:** e96505. doi:10.1371/journal.pone.0096505

Tanabe H, Habermann FA, Solovei I, Cremer M, Cremer T. 2002. Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutat Res* **504:** 37–45. doi:10.1016/S0027-5107(02)00077-5

Theilgaard-Mönch K, Pundhir S, Reckzeh K, Su J, Tapia M, Furtwängler B, Jendholm J, Jakobsen JS, Hasemann MS, Knudsen KJ, et al. 2022. Transcription factor-driven coordination of cell cycle exit and lineage-specification in vivo during granulocytic differentiation. *Nat Commun* **13:** 3595. doi:10.1038/s41467-022-31332-1

Tijms BM, Yeung HM, Sikkes SA, Möller C, Smits LL, Stam CJ, Scheltens P, van der Flier WM, Barkhof F. 2014. Single-subject gray matter graph properties and their relationship with cognitive impairment in early- and late-onset Alzheimer's disease. *Brain Connect* **4:** 337–346. doi:10.1089/brain.2013.0209

Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, Russiñol N, Di Stefano M, Chapaprieta V, Clot G, Farabella I, Cuscó P, Kulis M, et al. 2021. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. *Nat Commun* **12:** 651. doi:10.1038/s41467-020-20849-y

Weeraratna AT, Jiang Y, Hostetter G, Rosenblatt K, Duray P, Bittner M, Trent JM. 2002. Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell* **1:** 279–288. doi:10.1016/S1535-6108(02)00045-4

Xu J, Song F, Lyu H, Kobayashi M, Zhang B, Zhao Z, Hou Y, Wang X, Luan Y, Jia B, et al. 2022. Subtype-specific 3D genome alteration in acute myeloid leukaemia. *Nature* **611:** 387–398. doi:10.1038/s41586-022-05365-x

Xue Y, Yang Y, Tian H, Quan H, Liu S, Zhang L, Yang L, Zhu H, Wu H, Gao YQ. 2022. Computational characterization of domain-segregated 3D chromatin structure and segmented DNA methylation status in carcinogenesis. *Mol Oncol* **16:** 699–716. doi:10.1002/1878-0261.13127

Yang J, Corces VG. 2012. Insulators, long-range interactions, and genome function. *Curr Opin Genet Dev* **22:** 86–92. doi:10.1016/j.gde.2011.12.007

Yang L, Chen F, Zhu H, Chen Y, Dong B, Shi M, Wang W, Jiang Q, Zhang L, Huang X, et al. 2021. 3D genome alterations associated with dysregulated *HOXA13* expression in high-risk T-lineage acute lymphoblastic leukemia. *Nat Commun* **12:** 3708. doi:10.1038/s41467-021-24044-5

Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16:** 284–287. doi:10.1089/omi.2011.0118

Zamani M, Vicsek T. 2017. Glassy nature of hierarchical organizations. *Sci Rep* **7:** 1382. doi:10.1038/s41598-017-01503-y

Zamani M, Camargo-Forero L, Vicsek T. 2018. Stability of glassy hierarchical networks. *New J Phys* **20:** 023025. doi:10.1088/1367-2630/aaa8ca

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137. doi:10.1186/gb-2008-9-9-r137

Zhu I, Song W, Ovcharenko I, Landsman D. 2021. A model of active transcription hubs that unifies the roles of active promoters and enhancers. *Nucleic Acids Res* **49:** 4493–4505. doi:10.1093/nar/gkab235

Zink D, Fischer AH, Nickerson JA. 2004. Nuclear structure in cancer cells. *Nat Rev Cancer* **4:** 677–687. doi:10.1038/nrc1430

Received August 15, 2022; accepted in revised form January 31, 2023.