Supplemental Methods

Overall design of C_TG

The Hi-C contact map depicts a proximity network $G(V,E)$, where the vertices $V=\{v_1,v_2,\ldots v_n\}$ denote the non-overlapping genomic regions and the edges $E = \{e_{i,j}\}$ denote the contact strength between pairwise connected genomic regions. Similar to diffusion-based methods for network denoising (1, 2), a Markov prosses (3) is used to describe the diffusion process on this network. $D_{i,i} = \sum_{j=1}^{n} e_{i,j}$, is the element of the diagonal degree matrix D for the network. The vector $P_i^{(1)} = \{P_{i,1}^{(1)}, P_{i,2}^{(1)}, \ldots P_{i,n}^{(1)}\}$ is the conditional transition probability transiting from vertex v_i to $V=\{v_1,v_2,\ldots v_n\}$ in one single step. Likewise, $P_i^{(k)} = \{P_{i,1}^{(k)}, P_{i,2}^{(k)}, \ldots P_{i,n}^{(k)}\}$ is the conditional transition probability in k steps and $P_{i,j}^{(k)} = \sum_{p=1}^{n} P_{i,p}^{(k-1)} P_{p,j}^{(k-1)}$. With increasing k, the transition probability from v_i to v_j gradually integrates neighbor information and expand the inclusion of edges, since v_i and v_j may not be connected in one step but they can be connected in some finite steps as the network G is a connected graph. Taking k=2 and $P_{i,j}^{(2)} = \sum_{p=1}^{n} P_{i,p}^{(1)} P_{p,j}^{(1)}$ as an example, when the two pairs of vertices (v_i and v_p, v_j and v_p) are pairwise neighbors, which means $P_{i,p}^{(1)} \neq 0$ and $P_{p,j}^{(1)} \neq 0$, v_p contributes to $P_{i,j}^{(2)}$. $P_i^{(k)}$ converges to an invariant distribution for connected graph and the difference between $P_i^{(k-1)}$ and $P_i^{(k)}$ decreases.

It is thus appropriate to use the integrated information on $\{P_i^{(1)}, P_i^{(2)}, \ldots P_i^{(k)}\}$ to describe the diffusion manner of vertex v_i within some given number of k steps, which can be infinite. In practice, we found that $P_i^{(k)}$ converges rapidly and therefore used the exponential decay to fit the convergence. $S_i^{(k)}$ is defined as the weighted summation of $P_i^{(t)}$ $(1 \leq t \leq k)$:

$$S_i^{(k)} = \sum_{t=1}^{k} \exp(-\alpha t)\, P_i^{(t)}$$

When k reaches infinity, $S_i^{(k)}$ converges to S_i (Supplementary note). As the weighted summation of $P_i^{(t)}$, S_i naturally integrates neighbor information of the connected graph and therefore alleviates in a physics-based manner the problems caused by the Hi-C data sparsity. On the other hand, the exponential decay ensures that the integration does not eliminate the distinction of each vertex, taking the rapid convergence of $P_i^{(k)}$ into consideration.

The physical succession of the genomic structure suggests that the proximal genomic regions should share similar diffusion manners. The similarity between pairwise vertices v_i and v_j is quantified by L1 distance between S_i and S_j. L1 distance is used as a measure since it mitigates the impact of outliers caused by distance matrices of higher-order terms. A C_TG distance matrix is then constructed based on the Hi-C contact map.

Proof 1

Eigenvalues Λ of the P are within the range of [-1,1].

For any eigenvector X of P:

$$PX = \lambda X$$

The maximum element of X is denoted as x_{max}, and the minimum element of X is denoted as x_{min}. As the row summation of P is normalized to 1, and P is positive,

$$x_{min} \leq \lambda x_{min} \leq x_{max}$$
$$x_{min} \leq \lambda x_{max} \leq x_{max}$$

Therefore,

$$-1 \leq \lambda x_{max} \leq 1$$

Proof 2

When n approaches infinity, the transition propensity matrix $M^{(n)}$ is convergent.

P is diagonalizable:

$$P(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m) = (\lambda_1 \vec{v}_1, \lambda_2 \vec{v}_2, \dots, \lambda_m \vec{v}_m) = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m) \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{bmatrix}$$

$$P^{(1)} = U^{-1} \Lambda U$$

$P^{(k)}$ can be written as:

$$P^{(k)} = P^k = U^{-1} \Lambda U$$

$S^{(n)}$ is the weighted summation of $P^{(k)}$:

$S^{(n)} = \sum_{k=1}^{n} \exp(-\alpha k)\, U^{-1} \Lambda^k\, U = \sum_{k=1}^{n} U^{-1} [\exp(-\alpha k)\, \Lambda^k]\, U$

According to the associative law of multiplication:

$$S^{(n)} = U^{-1} \sum_{k=1}^{n} [\exp(-\alpha k)\, \Lambda^k]\, U = U^{-1} [\sum_{k=1}^{n} \exp(-\alpha k)\, \Lambda^k]\, U$$

When n approaches infinity, we have

$$S = U^{-1} [\lim_{n \to \infty} \sum_{k=1}^{n} \exp(-\alpha k)\, \Lambda^k] U$$

In the above equation, $\exp(-\alpha k)\,\Lambda^k$ is a geometric progression, and

$$\lim_{n\to\infty} \exp(-\alpha k)\,\Lambda^k \to 0$$

Therefore, the summation over $\exp(-\alpha k)\Lambda^k$ is convergent when

$$\rho(P) < \exp(a), \rho(P) = \max|\lambda_i|$$

As $\rho(P) < 1$ and $\exp(a) > \exp(0) > 1$:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \exp(-\alpha k)\,\Lambda^k = \Lambda\,[\exp(\alpha) I - \Lambda]^{-1}$$

I denotes the identity matrix.

S is then also convergent and

$$S = U^{-1}\,\Lambda\,[\exp(\alpha) I - \Lambda]^{-1} U$$

1. Wang,B., Pourshafeie,A., Zitnik,M., Zhu,J., Bustamante,C.D., Batzoglou,S. and Leskovec,J. (2018) Network enhancement as a general method to denoise weighted biological networks. *Nat. Commun.*, **9**.
2. Cao,M., Zhang,H., Park,J., Daniels,N.M., Crovella,M.E., Cowen,L.J. and Hescott,B. (2013) Going the Distance for Protein Function Prediction: A New Distance Metric for Protein Interaction Networks. *PLoS One*, **8**, e76339.
3. Stochastic Processes in Physics and Chemistry (2007) Elsevier.
4. Halverson,J.D., Smrek,J., Kremer,K. and Grosberg,A.Y. (2014) From a melt of rings to chromosome territories: The role of topological constraints in genome folding. *Reports Prog. Phys.*, **77**.
5. Lieberman-Aiden,E., van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O., *et al.* (2009) Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, **326**, 289–293.

PCR Program

STEP	TEMP	TIME
Pre-extension	72°C	5 minutes
Initial Denaturation	98°C	30 seconds
10 Cycles	98°C	10 seconds
	65°C	30 seconds
	72°C	90 seconds
Final Extension	72°C	2 minutes
Hold	4°C	∞