scHiDiffEM: Reconstructing Sparse Single-Cell Hi-C Contact Maps with Diffusion Expectation–Maximization

Anonymous authors
Paper under double-blind review

Abstract

Single-cell Hi-C (scHi-C) profiling elucidates three-dimensional (3D) genome architecture at cellular resolution, yet individual contact maps are severely corrupted by extreme count sparsity, cell-to-cell sequencing depth heterogeneity, and power-law genomic distance decay. An observed zero conflates true biological non-contact with technical sampling dropout, making it difficult to reconstruct cell-specific topology without averaging away biological variation. We introduce scHiDiffEM (**S**ingle-**C**ell **Hi**-C **Diff**usion **E**xpectation–**M**aximization), which treats reconstruction as a population-informed inverse problem. A distance-aware low-rank decoder represents chromosome contact rates, and a library-size-normalized negative-binomial model relates these rates directly to raw counts. With clean single-cell targets unavailable, depth-normalized mini-pseudo-bulks from unsupervised subpopulations supply canonical seeds to initialize a population-conditioned diffusion prior. A generalized expectation–maximization (EM) framework then jointly trains the model: the E-step restores individual single-cell latents via raw-count data consistency and bounded diffusion proximal updates, while sequential M-steps iteratively optimize the observation likelihood parameters and refine the diffusion prior. Across 18 independent benchmark datasets spanning five biological collections, scHiDiffEM substantially improves overall performance across cell embeddings, contact-pattern imputation, and topological structure recovery without dataset-specific tuning. These results establish scHiDiffEM as a principled framework for scHi-C reconstruction, enabling faithful recovery of single-cell chromatin architecture from extreme sparsity without requiring unattainable ground truth.

1 Introduction

Three-dimensional genome organization shapes interactions between regulatory elements, closely associated with transcription, development, and cell identity. While conventional bulk Hi-C captures population-averaged contact frequencies (Lieberman-Aiden et al., 2009), single-cell Hi-C (scHi-C) resolves structural heterogeneity in chromatin conformations across individual cells (Dautle & Chen, 2025). Effectively analyzing scHi-C data requires complementary capabilities: learning low-dimensional cell embeddings that distinguish cell types and developmental states, imputing severely undersampled contact maps, and delineating higher-order chromatin structures such as topologically associating domains (TADs) along with their insulating boundaries (Dixon et al., 2012).

However, realizing this potential is complicated by severe sparsity, technical noise, and strong distance decay. Most locus pairs contain zero observed counts, where a true biological non-contact cannot be distinguished from a missed interaction due to shallow sequencing or dropouts. Furthermore, contact frequency decays steeply with genomic separation, allowing broad distance effects to overshadow subtle, biologically informative local structures (so-called band bias (Zheng et al., 2022)). Besides, the significant disparity in sequencing depth and resolution, as well as the high dimensionality further challenge computational scalability at fine resolutions. Effective reconstruction must therefore combine information across cells while accounting for sampling and distance biases without erasing cell-specific structure (Zheng et al., 2022).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Existing methods address these challenges through spatial smoothing (Zhou et al., 2019), hypergraph learning (Zhang et al., 2022a), tensor factorization (Zhang et al., 2022b), or probabilistic modeling (Zheng et al., 2022) (Section 2). While effective for clustering, heuristic smoothing can blur sharp domain boundaries and dampen cell-specific contacts, and high cluster separation in embedding space does not guarantee accurate contact-level imputation or the preservation of local structural minima such as insulation troughs. Moreover, fitting preprocessed or transformed targets often absorbs depth biases. While diffusion models provide expressive priors for underdetermined inverse problems (Chung et al., 2023) and have recently been applied to contact matrices via continuous Gaussian diffusion (Wang & Cheng, 2024), adapting them to scHi-C faces two domain obstacles: ground-truth clean single-cell maps do not exist for training, and standard matrix-space Gaussian noise fails to capture discrete, depth-dependent count sampling dropouts.

To address these challenges, we propose scHiDiffEM, which formulates scHi-C reconstruction as a population-informed statistical inverse problem and solves it through generalized expectation–maximization (EM), illustrated in Figure 1. First, to capture the underlying biology, a distance-aware geometric low-rank decoder separates cell-specific contact rates from global distance decay, while a library-size-normalized Negative Binomial (NB) likelihood models directly on raw count observations, supervising both zero and non-zero entries (Section 3.1). Second, to overcome the absence of clean single-cell training targets, we employ depth-normalized mini-pseudo-bulk seeds from unsupervised clustered subpopulations to initialize and train a population-conditioned diffusion prior via denoising score matching and simulated sparse observation recovery (Section 3.2). Third, our model is jointly trained via a generalized EM framework: the E-step restores individual single-cell latents by alternating raw-count data fidelity with bounded diffusion proximal corrections, while sequential M-steps iteratively optimize the observation parameters and refine the generative diffusion prior (Section 3.3). Finally, the restored latents yield common-depth rate maps for physical contact analysis and gauge-invariant spectral embeddings for cell typing (Section 3.4).

We summarize our main contributions as follows: (1) we formalize scHi-C reconstruction as a population-informed statistical inverse problem and solve it through generalized EM training with latent diffusion prior; (2) we propose geometric-aware raw-count reconstruction model and population-informed diffusion prior to achieve sparse-observation recovery without clean single-cell targets; and (3) we conduct comprehensive evaluations across 18 benchmark settings, demonstrating that scHiDiffEM achieves superior overall aggregate performance across cell representation, contact map imputation, and chromosome-structure recovery without per-dataset hyperparameter tuning.

2 Related Work

Single-cell Hi-C analysis. Methods address sparsity using within-map structure or cross-cell information: scHiCluster applies convolution and random walks (Zhou et al., 2019), whereas Higashi learns hypergraph representations (Zhang et al., 2022a) and Fast-Higashi uses scalable tensor decomposition (Zhang et al., 2022b). In parallel, BandNorm and scVI-3D address distance and depth biases through normalization and generative modeling (Zheng et al., 2022). Recently, HiCDiff (Wang & Cheng, 2024) applied continuous diffusion directly to contact matrices, but treating discrete counts as continuous images with additive Gaussian noise neglects discrete sampling physics and depth disparities. In contrast, scHiDiffEM diffuses over compact geometric latent coordinates and couples them to raw counts via a discrete observation likelihood. Although sequencing coverage and computational complexity strongly influence overall utility (Plummer et al., 2025), informative embeddings alone do not guarantee accurate contact recovery or the preservation of fine structural extrema. By contrast, we incorporate raw-count likelihood modeling directly with geometry-informed restoration, simultaneously evaluating cell representations, reconstructed maps, and structure identifications.

Diffusion and EM for inverse problems. Diffusion posterior sampling combines a pretrained prior with measurement constraints (Chung et al., 2023), leaving open how to learn that prior without clean training data. Recent computer vision frameworks address this through distinct routes: Ambient Diffusion (Daras et al., 2023) trains score models directly on corrupted measurements under known forward operators without latent imputation, whereas EMDiffusion (Bai et al., 2024) alternates posterior restoration and score network retraining. Single-cell Hi-C poses unique domain challenges: discrete count sampling, severe depth disparities, and non-identifiability between biological

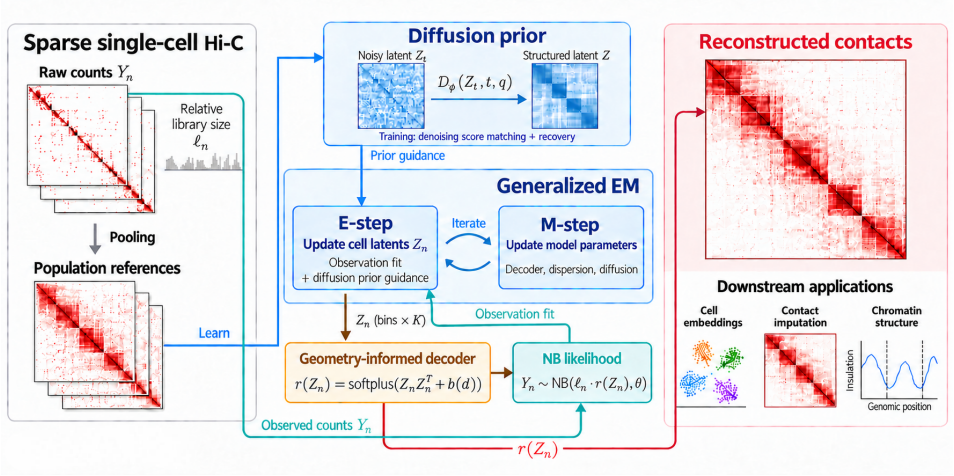

Figure 1: Overview of scHiDiffEM. Generalized EM couples a population-informed diffusion prior with a geometry-informed decoder and a library-size-aware negative-binomial likelihood, alternating latent reconstruction and model updates. The recovered representations support embedding, imputation, and structure analysis.

and sampling zeros. We therefore train restoration on population-derived references and simulated sparse observations rather than single-cell posterior factors, applying bounded MAP corrections. These approximate updates operate as a generalized EM procedure without classical monotonic convergence guarantees.

3 METHOD

3.1 GEOMETRIC RATE PARAMETERIZATION AND RAW-COUNT LIKELIHOOD

Geometry-informed contact rate parameterization. A contact map reflects two intertwined physical factors: locus-specific 3D spatial conformation, which varies across cells, and universal polymer scaling, which causes contact frequency to decay rapidly with 1D genomic separation. We model a chromosome discretized into M genomic bins (whose resolution depends on the sequencing method) within a diagonal band of width D, considering bin pairs $(i, i+d)$ with genomic distance $1 \le d \le D$. Restricting to this near-diagonal band is physically and experimentally principled: chromatin contact probability drops precipitously over genomic distance, such that ultra-distal entries are exceedingly sparse and predominantly contaminated by non-specific ligation noise, whereas functionally salient architectural domains (such as TADs and loops) reside within this local-to-intermediate range. To decouple locus-specific architecture from global distance decay, we represent cell n through a latent matrix $Z_n = [z_{n1}, \dots, z_{nM}]^\top \in \mathbb{R}^{M \times K}$ of rank K. The contact logit F_{nid} and corresponding common-depth interaction rate r_{nid} for bins $(i, i+d)$ are defined as

$$F_{nid} = z_{ni}^\top z_{n,i+d} + b_\phi(d), \qquad r_{nid} = \text{softplus}(F_{nid}), \tag{1}$$

where $b_\phi(d)$ is a multilayer perceptron (MLP) applied to $\log(1+d)$. The Gram inner product $z_{ni}^\top z_{n,i+d}$ captures low-rank geometric organization, while $b_\phi(d)$ absorbs the baseline distance decay shared across the cohort. Because $b_\phi(d)$ acts as a cohort-wide regularizing baseline, cell-state-dependent compaction shifts (e.g., mitotic condensation during the cell cycle; Nagano et al. 2017) are absorbed by the latent factors Z_n; as smooth distance-decay variations are low-frequency modes, rank-K Gram factors capture them without sacrificing capacity for local domain boundaries. Furthermore, as the Gram product $Z_n Z_n^\top$ is invariant under orthogonal gauge transformations $Z_n \mapsto Z_n Q$ for $Q \in \mathcal{O}(K)$, population seed factors are aligned to their consensus Fréchet barycenter via Generalized Procrustes Analysis (GPA; Gower 1975) prior to diffusion training, iteratively rotating factors against the consensus mean until convergence.

Library-size-aware observation likelihood. Rather than regressing on heuristic preprocessed inputs, we model raw contact counts directly. Let $s_n = \sum_{i,d} y_{nid}$ be the total in-band contacts for cell

n, with relative library size $\ell_n = \mathrm{clip}(s_n/\bar{s}, 0.05, 20)$ scaled by the cohort mean $\bar{s} = N^{-1}\sum_m s_m$. Conditional on Z_n and ℓ_n, raw counts follow a Negative Binomial distribution with quadratic variance (NB2): $y_{nid} \mid Z_n, \ell_n \sim \mathrm{NB2}(\mu_{nid}, \theta_d)$, where expected count $\mu_{nid} = \ell_n r_{nid}$, variance $\mathrm{Var}(y_{nid}) = \mu_{nid} + \mu_{nid}^2/\theta_d$, and θ_d is a distance-stratified dispersion parameter. To counteract stratum cardinality imbalance ($M - d$) and prioritize proximal architectural bands ($1 + 3\exp(-d/8)$), the observation objective is formulated as a weighted negative pseudo-log-likelihood

$$\mathcal{L}_{\mathrm{obs}} = -\frac{1}{N}\sum_{n,d,i} w_d \log p_{\mathrm{NB}}(y_{nid} \mid \ell_n r_{nid}, \theta_d), \qquad w_d = \frac{1 + 3\exp(-d/8)}{M - d}, \tag{2}$$

where $\log p_{\mathrm{NB}}(y \mid \mu, \theta)$ denotes the log-likelihood of count y under distribution $\mathrm{NB2}(\mu, \theta)$. Formulated as a pseudo-likelihood within generalized EM, it guides the MAP E-step toward structurally salient features. Crucially, all observed zero entries are retained in Equation 2, providing negative supervision against unsupported contacts. Compared to a Poisson likelihood, NB overdispersion accommodates sampling variability and reduces penalties on false-negative dropouts. We avoid a separate zero-inflation gate since gating parameters cannot be reliably identified without ground-truth zero labels.

3.2 Learning a generative diffusion prior without clean single-cell maps

Standard score-based diffusion requires clean training targets, which single-cell Hi-C does not provide. Furthermore, pooling all cells globally would wash out cell-type-specific chromatin conformations. We address this dilemma through a two-fold strategy: constructing clean-ish latent seeds from unsupervised subpopulation aggregates, and training a latent diffusion model on both continuous score matching and discrete corruption inversion.

Subpopulation conditioning and mini-pseudo-bulk seeds. Band-normalized (BandNorm) log-contacts $X_n^{\mathrm{aux}} = \log(1 + \mathrm{BandNorm}(Y_n))$ (Zheng et al., 2022) are employed to extract auxiliary features f_n across block-averaged strata. MiniBatchKMeans identifies 16–32 subpopulation clusters. For each center c_j, the soft population condition is $q_{nj} = \frac{\exp(-\delta_{nj}/\tau_q)}{\sum_k \exp(-\delta_{nk}/\tau_q)}$, with $\delta_{nj} = \|f_n - c_j\|_2^2 - \min_k \|f_n - c_k\|_2^2$ and temperature τ_q set from the median second-nearest-center distance. This condition guides the diffusion prior without averaging neighboring contact maps. Within each cluster, cells are partitioned into small groups G_g of at most 16 cells to form depth-normalized mini-pseudo-bulks $\bar{r}_{gid} = \frac{\sum_{n\in G_g} y_{nid}}{\sum_{n\in G_g} \ell_n}$ with group condition $q_g = \frac{1}{|G_g|}\sum_{n\in G_g} q_n$. Fitting each aggregated rate matrix via weighted Poisson regression yields population-informed mini-pseudo-bulk seeds $Z_g^\star = \arg\min_{Z_g} \sum_{i,d} w_d\,[r_{gid}(Z_g) - \bar{r}_{gid}\log(r_{gid}(Z_g) + \epsilon)]$. Aggregation substantially increases contact coverage, while small group sizes preserve biological heterogeneity.

Latent diffusion via denoising score matching and sparse observation recovery. Let normalized seeds be $\widetilde{Z}_0 = Z^\star/s_Z$, where s_Z is the empirical standard deviation of seed latents. Following the Gaussian diffusion formulation (Ho et al., 2020), a cosine schedule (Nichol & Dhariwal, 2021) defines $\widetilde{Z}_t = \sqrt{\bar{\alpha}_t}\widetilde{Z}_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon$, $\epsilon \sim \mathcal{N}(0, I)$. We employ a one-dimensional residual U-Net along genomic bins to parametrize $\epsilon_\psi(\widetilde{Z}_t, t, q)$ conditioned on time and population embeddings. The seed-only denoising score matching (DSM) loss (Song et al., 2021) uses minimum signal-to-noise ratio (Min-SNR) weighting (Hang et al., 2023)

$$\mathcal{L}_{\mathrm{DSM}} = \mathbb{E}_{g,t,\epsilon}\left[\omega_t \|\epsilon - \epsilon_\psi(\widetilde{Z}_{g,t}, t, q_g)\|_F^2\right], \quad \omega_t = \min\left(1, \frac{\gamma}{\mathrm{SNR}_t}\right), \quad \gamma = 5. \tag{3}$$

However, Gaussian score matching alone does not reproduce discrete sequencing counts with dropouts. To explicitly train the reverse mapping from sparse counts to clean structures, for each mini-pseudo-bulk seed with rate $\bar{R}_g$, we draw low-depth observations $Y_g^{\mathrm{corr}} \sim \mathrm{Poisson}(\ell \bar{R}_g)$ with sampled empirical relative library size ℓ. We fit corrupted factors Z_g^{corr} under the NB likelihood and add noise at intermediate timesteps $t \in [10, 250]$. The one-step clean prediction

$$\widehat{\widetilde{Z}}_g^{\mathrm{clean}} = \frac{\widetilde{Z}_{g,t}^{\mathrm{corr}} - \sqrt{1 - \bar{\alpha}_t}\,\epsilon_\psi(\widetilde{Z}_{g,t}^{\mathrm{corr}}, t, q_g)}{\sqrt{\bar{\alpha}_t}}, \tag{4}$$

is optimized via

$$\mathcal{L}_{\mathrm{rec}} = \mathrm{SmoothL1}_{\beta=0.1}\left(\widehat{\widetilde{Z}}_g^{\mathrm{clean}}, Z_g^\star / s_Z\right). \tag{5}$$

This objective trains the mapping from a low-depth observation to its corresponding population-reference factor.

3.3 Model training and single-cell restoration via generalized diffusion–EM

Given the trained prior and raw-count likelihood, single-cell restoration is formulated as a maximum a posteriori (MAP) estimation problem within a plug-and-play framework (Venkatakrishnan et al., 2013). Each single-cell latent is initialized from its subpopulation seed center. In the MAP E-step, each iteration executes $K_E = 20$ inner alternating updates. First, data consistency is enforced via a gradient step on the weighted observation pseudo-likelihood

$$Z_n^{\mathrm{data}} \leftarrow Z_n - \eta_E \nabla_{Z_n} \mathcal{L}_{\mathrm{obs}}(Y_n \mid Z_n, \ell_n). \tag{6}$$

To match the perturbed score marginals, the denoiser $\mathcal{D}_\psi$ injects forward noise $\widetilde{Z}_n^{(s)}(t) = \sqrt{\bar{\alpha}_t}(Z_n^{\mathrm{data}}/s_Z) + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}^{(s)}$ with $\boldsymbol{\epsilon}^{(s)} \sim \mathcal{N}(0, I)$ at annealed timestep t ($250 \to 10$), taking average over Tweedie estimates across $S = 2$ Monte Carlo draws

$$\mathcal{D}_\psi(Z_n^{\mathrm{data}}; t, q_n) = s_Z \cdot \frac{1}{S}\sum_{s=1}^{S} \mathrm{clip}\left(\frac{\widetilde{Z}_n^{(s)}(t) - \sqrt{1-\bar{\alpha}_t}\,\boldsymbol{\epsilon}_\psi(\widetilde{Z}_n^{(s)}(t), t, q_n)}{\sqrt{\bar{\alpha}_t}}, -C, C\right), \tag{7}$$

with clipping $C = 2.5$. The nominal prior step $\Delta Z_n^{\mathrm{prior}} = \lambda_m[\mathcal{D}_\psi(Z_n^{\mathrm{data}}; t, q_n) - Z_n^{\mathrm{data}}]$ with relaxation $\lambda_m \in [0.02, 0.10]$ is then bounded via trust-region projection for numerical stability and preventing over-regularization

$$Z_n \leftarrow Z_n^{\mathrm{data}} + \min\left(1, \frac{\rho_n}{\|\Delta Z_n^{\mathrm{prior}}\|_F}\right)\Delta Z_n^{\mathrm{prior}}, \quad \rho_n = 0.05\max\left(\|Z_n\|_F, 0.25 s_Z\sqrt{MK}\right). \tag{8}$$

The M-step then proceeds sequentially across two iterated stages: (1) with latent factors Z_n fixed, we optimize observation parameters (ϕ, θ_d) under $\mathcal{L}_{\mathrm{obs}}$; (2) with observation parameters fixed, we train the diffusion score network ψ under $\mathcal{L}_{\mathrm{DSM}} + \mathcal{L}_{\mathrm{rec}}$ using an exponential moving average (EMA). Importantly, anchoring score training on canonical seeds rather than noisy single-cell latents prevents feedback-loop error accumulation, while the EM coupling is mediated through the shared observation parameters (ϕ, θ_d) and iterative recalibration of the latent scale s_Z. Since the proximal update in Equation 8 is parameter-driven rather than an exact posterior maximizer, this alternating scheme operates as a generalized EM procedure without classical monotonic convergence guarantees.

3.4 Reconstructed contact maps and gauge-invariant embeddings

Common-depth contact rates are decoded via Equation 1, yielding reconstructed maps

$$\widehat{X}_{nid} = \max\{0, \log(1 + \mathrm{softplus}(\widehat{F}_{nid})) - \tau\}, \tag{9}$$

with a small background threshold $\tau = 0.01$. For representation learning, each chromosome's decoded log-rate band is reduced to 50 dimensions via uncentered truncated singular value decomposition (SVD). Chromosome representations are concatenated genome-wide, reduced to 30 dimensions, and row-normalized. Because SVD is applied directly to decoded Gram maps rather than unaligned latent matrices Z_n, the resulting embeddings are mathematically invariant to orthogonal gauge transformations.

4 Experiments and Results

4.1 Experimental setup

Benchmark datasets. We evaluate five biological collections at 18 combinations of dataset and resolution: cell-cycle dynamics in mouse embryonic stem cells (50 kb–1 Mb resolutions (Nagano

et al., 2017)), early mouse embryogenesis (100 kb–1 Mb (Collombet et al., 2020)), mouse brain profiling via HiRes (100–500 kb (Liu et al., 2023b)), adult mouse hippocampus (100–200 kb (Liu et al., 2023a)), and human prefrontal cortex (PFC, 50–500 kb (Lee et al., 2019; Luo et al., 2022)). Hi-C Data at different sequencing resolutions are measured independently on different cells and should be treated as distinct datasets. Hence, each resolution is analyzed independently without pooling across scales. Details are shown in Table 6 (Appendix A).

Evaluation. We evaluate methods across three complementary tasks: (1) **Cell embedding**: Inspired by a recent benchmark paper (Plummer et al., 2025), we construct 15-nearest-neighbor graphs, sweep Leiden resolutions from 0.1 to 2.0 to maximize normalized mutual information (NMI), and report NMI, adjusted Rand index (ARI), adjusted mutual information (AMI), and cell-type silhouette score. (2) **Split-half contact imputation**: Each cell type is randomly split into two halves before benchmark sampling. Predictions from query half are log-normalized, and compared with the opposite raw-count pseudo-bulk reference (ensuring query cells never contribute to the reference). We report stratum-adjusted correlation (SCC (Yang et al., 2017)), global Pearson and Spearman correlations, mean stratum-wise Pearson, and cell-type-specific residual metrics (Pearson correlation after global half-bulk subtraction, Top-1 type retrieval accuracy, and margin over the strongest competing type). (3) **Structure recovery**: on split-half maps, we compute insulation profiles and call domain boundaries, reporting insulation Pearson and Spearman correlations (Crane et al., 2015), boundary AUROC, AUPRC, and boundary F1 score with two-bin tolerance ($\mathrm{F1}_{\pm 2}$). Unweighted sums ($C_{\mathrm{emb}}, C_{\mathrm{imp}}, C_{\mathrm{str}}$) provide descriptive composite summaries (formal definitions in Appendix B).

Baselines. Embedding evaluation compares scHiDiffEM, scHiCluster (Zhou et al., 2019), Higashi (Zhang et al., 2022a), Fast-Higashi (Zhang et al., 2022b), and scVI-3D (Zheng et al., 2022). Map-based tasks evaluate the contact-reconstructing methods among them (scHiDiffEM, scHiCluster, Higashi, and scVI-3D).

Implementation. To assess scHiDiffEM as a general-purpose framework for scHi-C analysis, we used a shared set of model hyperparameters across all datasets and resolutions without dataset- or resolution-specific fine-tuning. We use latent rank $K = 10$, $D = 50$ off-diagonal bins, 1,000 cosine diffusion steps, Min-SNR weighting ($\gamma = 5$), and 50 generalized EM iterations (20 updates per E-step, 5 diffusion epochs per M-step). Details are shown in Appendix D.

4.2 Results

Across 18 independent dataset–resolution benchmarks spanning five diverse biological collections, scHiDiffEM achieves superior overall performance across cell-type clustering agreement, contact-pattern recovery, and insulation and boundary metrics (Tables 1 and 2). Notably, without dataset-specific hyperparameter tuning, scHiDiffEM consistently attains the highest overall aggregate score across all three core benchmark tasks (Table 1), maintaining robust performance advantages across diverse biological contexts, sequencing depths, and resolution scales (Table 2). Complete per-metric dataset results are provided in Appendix G.

Table 1: Mean performance across all 18 dataset–resolution settings for the three core benchmark tasks. Agg denotes task aggregate score ($C_{\mathrm{emb}}, C_{\mathrm{imp}}, C_{\mathrm{str}}$). Detailed metric definitions are in Appendix B.

	Cell Embedding (Task 1)					**Contact Imputation (Task 2)**								**Structure Recovery (Task 3)**					
Method	Agg	ARI	NMI	AMI	Silh	Agg	SCC	PG	SG	PS	TP	Top-1	Marg	Agg	Ins-P	Ins-S	AUROC	AUPRC	$\mathrm{F1}_{\pm 2}$
scHiDiffEM	**2.3329**	**0.7046**	**0.7325**	**0.7291**	**0.1668**	**3.9415**	0.6447	0.8382	0.7830	0.6244	**0.2451**	**0.7213**	**0.0847**	**3.5124**	**0.9141**	**0.8691**	0.6841	**0.2442**	0.8009
scHiCluster	1.9994	0.6647	0.6661	0.6628	0.0057	3.0466	0.5602	0.7984	0.7701	0.5484	0.1241	0.2915	-0.0461	2.9749	0.7530	0.6707	0.6549	0.2185	0.6777
Higashi	1.9173	0.5929	0.6276	0.6205	0.0763	2.8813	0.5607	0.7107	0.7184	0.5519	0.1190	0.3059	-0.0853	2.7853	0.7900	0.6392	0.5916	0.1829	0.5817
Fast-Higashi	2.1237	0.6706	0.6861	0.6830	0.0840	–	–	–	–	–	–	–	–	–	–	–	–	–	–
scVI-3D	1.7486	0.5655	0.5703	0.5597	0.0531	3.7048	**0.7109**	**0.8714**	**0.8220**	**0.6876**	0.1670	0.4577	-0.0119	3.5117	0.9077	0.8612	**0.6899**	0.2413	**0.8115**

Cell embeddings. scHiDiffEM achieved the highest mean clustering agreement and cell-type silhouette, supporting a representation that preserves annotated biological variation beyond a particular graph partition (Tables 1 and 2). Figure 2 demonstrates the uniform manifold approximation and projection (UMAP) visualization for all five methods on matched cells from PFC, mouse hippocampus, embryo, and cell-cycle data. Though performance varies across biological settings, a significant improvement in overall performance can be observed, especially for datasets with relatively fewer sample size (Figure 2c-d). Detailed dataset-level metrics appear in Appendix G.

Table 2: Task aggregate scores ($C_{\text{emb}}, C_{\text{imp}}, C_{\text{str}}$) across all 18 independent dataset–resolution benchmarks. Best scores among learned methods are in bold. Detailed metrics are shown in Appendix G.

	Cell Embedding (C_{emb})					**Contact Imputation (C_{imp})**				**Structure Recovery (C_{str})**			
Dataset	Ours	scHiClust	Higashi	F-Higashi	scVI-3D	Ours	scHiClust	Higashi	scVI-3D	Ours	scHiClust	Higashi	scVI-3D
Cell cycle 50 kb	1.9228	1.6021	**1.9498**	1.9158	1.3460	3.6334	2.6189	1.9561	**3.9516**	**3.3847**	3.0406	2.5896	3.2926
Cell cycle 100 kb	**2.0276**	1.6116	1.8114	1.6852	1.5895	4.0477	2.8741	2.2789	**4.0876**	**3.4680**	2.9949	2.5425	3.4125
Cell cycle 200 kb	**1.9942**	1.6909	1.8451	1.6527	1.7086	**4.3518**	2.9830	2.0749	4.1429	**3.5252**	2.9019	2.4445	3.5091
Cell cycle 500 kb	**1.9375**	1.3199	1.8732	1.6999	1.7285	**4.7473**	3.0012	3.4975	4.3122	3.3806	2.4790	2.7795	**3.4686**
Cell cycle 1 Mb	1.7904	0.8846	**1.8032**	1.6916	1.5152	**4.9967**	3.0192	3.7769	4.3019	3.2708	2.3582	2.8145	**3.4656**
Embryo 100 kb	2.0745	**2.3349**	1.6873	2.1299	2.0573	3.8881	3.0326	3.1220	**4.2567**	**3.4649**	2.7086	2.4813	3.3878
Embryo 200 kb	**2.6268**	2.2794	1.9539	2.2542	2.3269	3.8293	2.9420	3.4459	**3.9570**	**3.5336**	2.6505	2.6812	3.4656
Embryo 500 kb	**2.5988**	2.0954	1.8772	2.2315	2.3111	**3.9553**	2.9036	3.6721	3.5404	3.4467	2.6264	2.7792	**3.4586**
Embryo 1 Mb	**2.5187**	1.8964	2.3766	2.4974	2.2557	**3.8942**	2.7900	3.6832	3.3225	3.3068	2.6547	2.8383	**3.4684**
Hi-res brain 100 kb	**1.8316**	1.6654	0.6217	1.5060	0.0249	**2.9559**	2.5464	2.1871	2.6397	**3.4496**	3.0171	2.7816	3.3005
Hi-res brain 200 kb	**1.8155**	1.4877	0.7171	1.4434	0.0970	**3.4836**	2.8208	2.5903	2.8505	**3.5624**	3.1048	2.9543	3.4396
Hi-res brain 500 kb	**1.8003**	1.3850	0.7777	1.6859	0.2653	**3.5125**	2.7564	2.7095	2.9704	3.4518	3.0555	3.0188	**3.5503**
Hippocampus 100 kb	**2.5115**	2.3891	2.4666	2.4316	2.0626	**3.9757**	3.5639	2.5790	3.6860	**3.6009**	3.3555	2.8275	3.5721
Hippocampus 200 kb	**2.6075**	2.3760	2.4526	2.4702	2.3260	**4.5052**	3.4543	3.1756	4.1386	3.8022	3.5265	3.3044	**3.8750**
PFC 50 kb	**2.8435**	2.8150	2.3524	2.1668	1.8214	2.9249	2.8880	1.7289	**3.0261**	**3.4232**	3.1749	2.3997	3.3323
PFC 100 kb	**3.0985**	2.6902	2.8655	2.9683	2.6011	**3.8290**	3.3008	2.2811	3.6515	3.6406	3.2952	2.6849	**3.6667**
PFC 200 kb	**3.0869**	2.8239	2.7148	3.0033	2.7565	**4.1452**	3.7133	3.2561	3.8517	**3.7831**	3.3464	3.0959	3.7693
PFC 500 kb	**2.9062**	2.6409	2.3658	2.7922	2.6818	**4.2707**	3.6301	3.8476	3.9994	3.7283	3.2565	3.1187	**3.7754**
Mean Aggregate	**2.3329**	1.9994	1.9173	2.1237	1.7486	**3.9415**	3.0466	2.8813	3.7048	**3.5124**	2.9749	2.7853	3.5117

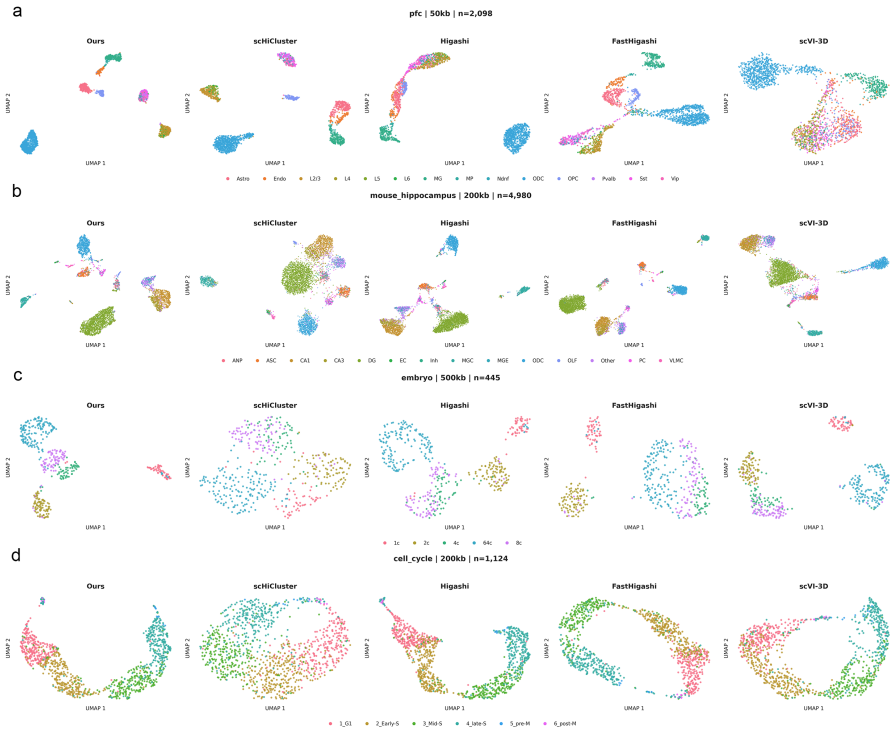

Figure 2: UMAP comparisons across five methods. Rows (a–d): PFC (50 kb), mouse hippocampus (200 kb), embryo (500 kb), and cell cycle (200 kb). Columns: scHiDiffEM, scHiCluster, Higashi, Fast-Higashi, and scVI-3D. Methods share matched cells and annotation colors within each row; coordinates are computed independently. Visualization of all datasets appears in Appendix E.

Contact maps imputation. Across all evaluated configurations, scHiDiffEM achieves the strongest overall contact recovery and uniquely excels at resolving cell-type-specific chromatin architectures (Tables 1 and 2). Crucially, after subtracting the shared half-bulk profile, scHiDiffEM maintains substantial advantages in type-specific correlation, top-1 retrieval accuracy, and discriminative margin over all baselines, whereas continuous generative baselines such as scVI-3D yield near-zero or negative retrieval margins. This indicates that while VAEs fit the smooth, generic distance-decay profile

(which yields high unstratified correlations), they risk over-smoothing subtype-defining chromatin contacts into an indistinct population average. In contrast, scHiDiffEM couples population diffusion priors with single-cell proximal likelihood updates, faithfully preserving idiosyncratic contacts. Furthermore, our background thresholding operator ($\tau = 0.01$ in Equation 9) deliberately eliminates near-zero background noise into exact zeros; while introducing rank ties that dampen global Spearman correlation, it sharpens biological contrast and structural fidelity where heuristic smoothers simply preserve artificial monotonic gradations.

From individual cells to population patterns. Figure 3 connects a single hippocampal medial ganglionic eminence (MGE)-annotated cell to the aggregate of the same type-half group, with PFC layer 5 (L5) cells providing a second biological setting. The individual prediction supplements sparse observed contacts, while the aggregate reveals local contact patterns that can be compared with an opposite-half reference. In these windows, scHiDiffEM preserves local variation that is more broadly smoothed by the comparison methods. The observed aggregate remains an important reference: pooling alone recovers substantial structure, so the figure illustrates the character of reconstruction rather than replacing the quantitative comparison.

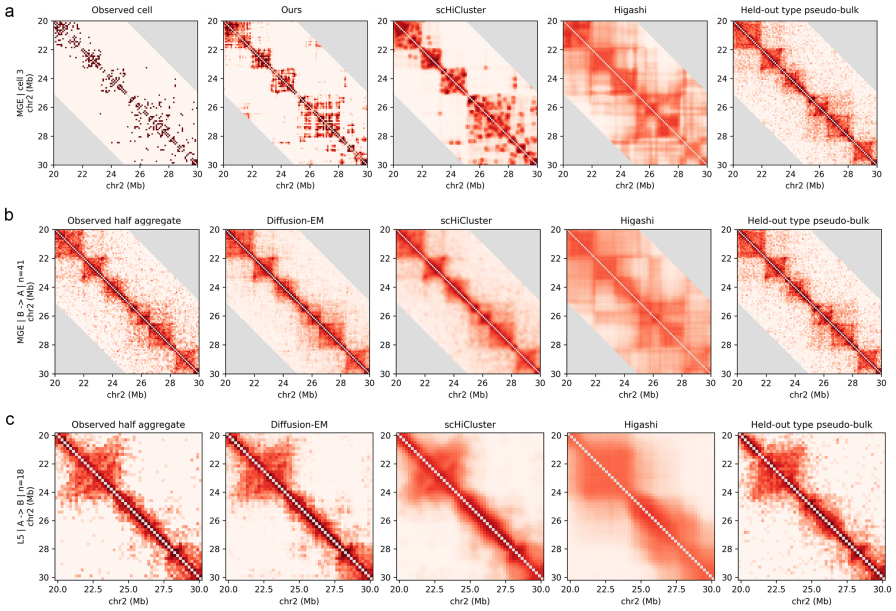

Figure 3: Contact maps on region 20–30 Mb, chromosome 2. (a) A mouse hippocampal MGE cell at 100 kb. (b) Its half-B MGE aggregate (41 cells), with a half-A reference. (c) PFC L5 at 200 kb (18 half-A cells), with a half-B reference. Methods share query cells and row-wise log1p color scales. Gray marks unmodeled entries. References are opposite-half pseudo-bulks, not single-cell ground truth. Full comparisons and colorbars appear in Appendix F.

Topological structure insulation and boundaries. The reconstructed single-cell contacts translate into superior downstream chromatin organization, yielding the highest overall structure recovery and the most faithful insulation profiles across benchmarks (Tables 1 and 2). By explicitly modeling raw-count sampling likelihoods rather than applying isotropic continuous smoothing, scHiDiffEM prevents boundary blurring and preserves sharp local insulation minima corresponding to true topological domain boundaries.

4.3 Ablation, Sensitivity Analysis and Cross-Seed Stability

Ablation. With fixed evaluation randomness, we removed either the diffusion proximal correction during inference or low-depth recovery training while retaining Gaussian denoising. Both ablations reduced SCC, stratum-wise Pearson, and type-specific Pearson on mouse embryo, cell-cycle, and human PFC data (Table 3), with larger losses after removing the proximal update. This supports the utility of learned restoration and count-corruption training within the model.

Table 3: Ablation results on contact map imputation. PS and TP denote stratum-wise and type-specific Pearson correlation.

Collection	Configuration	SCC	PS	TP
Embryo-200kb	Full	0.5395	0.5142	0.1790
Embryo-200kb	Without proximal update	0.5050	0.4851	0.1634
Embryo-200kb	Without explicit recovery	0.5215	0.4978	0.1699
Cell cycle-200kb	Full	0.7160	0.6727	0.3057
Cell cycle-200kb	Without proximal update	0.6616	0.6175	0.2512
Cell cycle-200kb	Without explicit recovery	0.6841	0.6386	0.2676
PFC-200kb	Full	0.7135	0.7101	0.3622
PFC-200kb	Without proximal update	0.6300	0.6294	0.2900
PFC-200kb	Without explicit recovery	0.6626	0.6591	0.3123

Table 4: Sensitivity analysis on hyperparameters. The results are shown as (SCC / type-specific Pearson).

Configuration	Embryo-200kb	Cell cycle-200kb	PFC-200kb
Full	0.5395 / 0.1790	0.7160 / 0.3057	0.7135 / 0.3622
Prior weight 0.05	0.5121 / 0.1659	0.6719 / 0.2553	0.6391 / 0.2933
Prior weight 0.20	0.5941 / 0.1753	0.7173 / 0.2141	0.7102 / 0.2515
Recovery mult 0.25	0.5293 / 0.1738	0.7008 / 0.2882	0.6880 / 0.3367
Recovery mult 2	0.5458 / 0.1818	0.7241 / 0.3145	0.7230 / 0.3691

Table 5: Cross-seed stability on embryo-200 kb dataset with 5 independent seeds.

Metric	Mean $\pm$ std
Imputation SCC	0.54096 ± 0.00085
Stratum Pearson	0.51543 ± 0.00077
Type Pearson	0.17967 ± 0.00040
Imputation agg	3.83485 ± 0.00283
Structure agg	3.53246 ± 0.00145
Embedding ARI	0.86593 ± 0.04267
Embedding NMI	0.85051 ± 0.02716

Prior strength. Varying final prior weight and recovery multiplier reveals expected regularization tradeoffs (Table 4). Specifically, an undersized prior weight under-regularizes sparse contact counts, whereas an overly heavy prior biases reconstructions toward a generic population prototype, compromising single-cell cell-type specificity despite stable global correlations.

Cross-seed stability. Across five training seeds on mouse embryo-200 kb dataset, SCC was 0.54096 ± 0.00085 (Table 5), showing the robustness and stability of scHiDiffEM, where variability substantially is much smaller than component-ablation drops.

5 Discussion

In this paper, we presented scHiDiffEM, a generative framework formulating single-cell Hi-C reconstruction as a statistical inverse problem. By coupling a geometric low-rank contact model and a raw-count Negative Binomial pseudo-likelihood with a population diffusion prior, generalized EM imputes sparse contacts and recovers sharp domain boundaries without heuristic smoothing or prototype collapse.

Despite the obvious advantages, there are still several limitations. First, restricting the contact model to a near-diagonal band ($D = 50$) omits ultra-distal and inter-chromosomal contacts, and subpopulation conditioning uses fixed initial clusterings rather than co-evolving with latents. Furthermore, while fixed $K = 10$ and $D = 50$ across 18 benchmarks demonstrate broad applicability without per-dataset tuning, adaptively scaling K and D with sequencing depth and resolution represents a natural avenue for further optimization. Moreover, extending scHiDiffEM to single-cell multimodal profiling by jointly modeling conformation with expression, chromatin accessibility, and DNA methylations (Park & Keleş, 2024; Zhou et al., 2024) will enrich regulatory discovery, while calibrated diffusion uncertainty can help distinguish topological heterogeneity from technical noise.

Despite these limitations, scHiDiffEM marks a substantial step forward for single-cell 3D genomics. By treating contact restoration as a solvable inverse problem rather than an empirical smoothing exercise, it demonstrates that cell-specific chromatin architecture can be faithfully recovered from ultra-sparse measurements without ground-truth maps. As single-cell conformation technologies continue to scale, such principled generative frameworks will be vital for transforming noisy contact counts into trustworthy biological discoveries.

AI USE STATEMENT

In this work, we used generative AI tools for designing or providing feedback on research methodology or experiments, implementing methods, and interpreting results. We have not used generative AI tools for generating or refining data sets, and developing theoretical models or writing proofs are not applicable to this work. Additionally, we used generative AI tools for creating or editing software code, creating or modifying scientific figures or images, and drafting parts of a research paper. We have reviewed all AI-assisted work, such as reviewing AI-generated codes and texts. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The source codes are submitted as supplementary materials. All the real data analyzed in the paper are publicly available via GEO collection GSE305523 (`https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE305523`). Comparison methods are implemented with official repositories provided by the original papers, available at `https://github.com/zhoujt1994/scHiCluster`, `https://github.com/ma-compbio/Higashi`, `https://github.com/ma-compbio/Fast-Higashi`, and `https://github.com/yezhengSTAT/scVI-3D`.

REFERENCES

Weimin Bai, Yifei Wang, Wenzheng Chen, and He Sun. An expectation-maximization algorithm for training clean diffusion models from corrupted observations. *Advances in Neural Information Processing Systems*, 37:19447–19471, 2024.

Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023.

Samuel Collombet, Noémie Ranisavljevic, Takashi Nagano, Csilla Varnai, Tarak Shisode, Wing Leung, Tristan Piolot, Rafael Galupa, Maud Borensztein, Nicolas Servant, et al. Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature*, 580(7801):142–146, 2020.

Emily Crane, Qian Bian, Rachel Patton McCord, Bryan R Lajoie, Bayly S Wheeler, Edward J Ralston, Satoru Uzawa, Job Dekker, and Barbara J Meyer. Condensin-driven remodelling of x chromosome topology during dosage compensation. *Nature*, 523(7559):240–244, 2015.

Giannis Daras, Kulin Shah, Yuval Dagan, Aravind Gollakota, Alex Dimakis, and Adam Klivans. Ambient diffusion: Learning clean distributions from corrupted data. *Advances in Neural Information Processing Systems*, 36:288–313, 2023.

Madison A Dautle and Yong Chen. Single-cell hi-c technologies and computational data analysis. *Advanced Science*, 12(9):2412232, 2025.

Jesse R Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380, 2012.

John C Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, 1975.

Tiankai Hang, Shuyang Gu, Chen Li, Jianmin Bao, Dong Chen, Han Hu, Xin Geng, and Baining Guo. Efficient diffusion training via min-snr weighting strategy. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7407–7417. IEEE, 2023.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Dong-Sung Lee, Chongyuan Luo, Jingtian Zhou, Sahaana Chandran, Angeline Rivkin, Anna Bartlett, Joseph R Nery, Conor Fitzpatrick, Carolyn O'Connor, Jesse R Dixon, et al. Simultaneous profiling of 3d genome structure and dna methylation in single human cells. *Nature Methods*, 16(10):999–1006, 2019.

Erez Lieberman-Aiden, Nynke L Van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R Lajoie, Peter J Sabo, Michael O Dorschner, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–293, 2009.

Hanqing Liu, Qiurui Zeng, Jingtian Zhou, Anna Bartlett, Bang-An Wang, Peter Berube, Wei Tian, Mia Kenworthy, Jordan Altshul, Joseph R Nery, et al. Single-cell dna methylome and 3d multiomic atlas of the adult mouse brain. *Nature*, 624(7991):366–377, 2023a.

Zhiyuan Liu, Yujie Chen, Qimin Xia, Menghan Liu, Heming Xu, Yi Chi, Yujing Deng, and Dong Xing. Linking genome structures to functions by simultaneous single-cell hi-c and rna-seq. *Science*, 380(6649):1070–1076, 2023b.

Chongyuan Luo, Hanqing Liu, Fangming Xie, Ethan J Armand, Kimberly Siletti, Trygve E Bakken, Rongxin Fang, Wayne I Doyle, Tim Stuart, Rebecca D Hodge, et al. Single nucleus multi-omics identifies human cortical cell regulatory genome diversity. *Cell Genomics*, 2(3), 2022.

Takashi Nagano, Yaniv Lubling, Csilla Várnai, Carmel Dudley, Wing Leung, Yael Baran, Netta Mendelson Cohen, Steven Wingett, Peter Fraser, and Amos Tanay. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*, 547(7661):61–67, 2017.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, pp. 8162–8171. PmLR, 2021.

Kwangmoon Park and Sündüz Keleş. Joint tensor modeling of single cell 3d genome and epigenetic data with muscle. *Journal of the American Statistical Association*, 119(548):2464–2477, 2024.

Dylan Plummer, Xiuyuan Lang, Shanshan Zhang, Yan Li, Jing Li, and Fulai Jin. A comprehensive benchmark of single-cell hi-c embedding tools. *Nature Communications*, 16(1):9119, 2025.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing*, pp. 945–948. IEEE, 2013.

Yanli Wang and Jianlin Cheng. Hicdiff: single-cell hi-c data denoising with diffusion models. *Briefings in Bioinformatics*, 25(4):bbae279, 2024.

Tao Yang, Feipeng Zhang, Galip Gürkan Yardımcı, Fan Song, Ross C Hardison, William Stafford Noble, Feng Yue, and Qunhua Li. Hicrep: assessing the reproducibility of hi-c data using a stratum-adjusted correlation coefficient. *Genome Research*, 27(11):1939–1949, 2017.

Ruochi Zhang, Tianming Zhou, and Jian Ma. Multiscale and integrative single-cell hi-c analysis with higashi. *Nature Biotechnology*, 40(2):254–261, 2022a.

Ruochi Zhang, Tianming Zhou, and Jian Ma. Ultrafast and interpretable single-cell 3d genome analysis with fast-higashi. *Cell Systems*, 13(10):798–807, 2022b.

Ye Zheng, Siqi Shen, and Sündüz Keleş. Normalization and de-noising of single-cell hi-c data with bandnorm and scvi-3d. *Genome Biology*, 23(1):222, 2022.

Jingtian Zhou, Jianzhu Ma, Yusi Chen, Chuankai Cheng, Bokan Bao, Jian Peng, Terrence J Sejnowski, Jesse R Dixon, and Joseph R Ecker. Robust single-cell hi-c clustering by convolution- and random-walk–based imputation. *Proceedings of the National Academy of Sciences*, 116(28): 14011–14018, 2019.

Yufan Zhou, Tian Li, Lavanya Choppavarapu, Kun Fang, Shili Lin, and Victor X Jin. Integration of schi-c and scrna-seq data defines distinct 3d-regulated and biological-context dependent cell subpopulations. *Nature Communications*, 15(1):8310, 2024.

A BENCHMARK COLLECTIONS AND DATA COHORTS

Table 6 summarizes the evaluation cohorts across the five biological collections after quality filtering and cell-name matching. Hi-C Data at different sequencing resolutions are measured independently on different cells and should be treated as distinct datasets. Cells are filtered globally across autosomes (excluding chrY, chrM, and chrEBV) using a total contact threshold: $\geq 50,000$ contacts for Cell Cycle, Embryo, HIRES Brain, and Mouse Hippocampus, and $\geq 5,000$ contacts for the ultra-sparse human PFC collection. To ensure fair and rigorous benchmarking, cohorts across all comparison methods are constructed via deterministic matching: we identify the common cell intersection successfully processed by all methods, retain cell types with at least 2 cells (enabling valid split-half evaluations), and balance cohorts with up to 500 cells per type. All the datasets are available at `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE305523`, and preprocessed by previous benchmark paper (Plummer et al., 2025). Comparison methods are implemented with official repositories provided by the original papers, available at `https://github.com/zhoujt1994/scHiCluster`, `https://github.com/ma-compbio/Higashi`, `https://github.com/ma-compbio/Fast-Higashi`, `https://github.com/yezhengSTAT/scVI-3D`.

Table 6: Benchmark biological collections. Cell counts indicate ranges when quality filtering yields slightly different cohorts across resolutions.

Collection	Tested Resolutions	Matched Cells	Cell Types / States
Cell cycle (Nagano et al., 2017)	50, 100, 200, 500 kb; 1 Mb	1,123–1,124	6
Mouse Embryo (Collombet et al., 2020)	100, 200, 500 kb; 1 Mb	445–453	5
Mouse brain (HIRES) (Liu et al., 2023b)	100, 200, 500 kb	398	7
Mouse hippocampus (Liu et al., 2023a)	100, 200 kb	4,924–4,980	14
Human Prefrontal cortex (PFC) (Luo et al., 2022)	50, 100, 200, 500 kb	1,634–2,098	14

B EVALUATION METRICS AND COMPOSITE SUMMARY MEASURES

For a compact, unweighted overview of each task, we sum performance metrics oriented such that larger values indicate superior performance:

$$C_{\text{emb}} = \text{ARI} + \text{NMI} + \text{AMI} + \text{Silhouette}, \tag{10}$$

$$C_{\text{imp}} = \text{SCC} + \rho_{\text{global}} + \rho^{\text{rank}}_{\text{global}} + \rho_{\text{strata}} + \rho_{\text{type}} + \text{Top1} + \text{Margin}, \tag{11}$$

$$C_{\text{str}} = \rho_{\text{ins}} + \rho^{\text{rank}}_{\text{ins}} + \text{AUROC} + \text{AUPRC} + \text{F1}_{\pm 2}. \tag{12}$$

Cell counts, cluster counts, boundary counts, and other support statistics are excluded. Individual metric scores across all 18 dataset–resolution settings are detailed in Appendix G.

C TRAINING AND INFERENCE PROCEDURE

1. **Data ingestion**: Read raw in-band counts y_{nid} and compute cell-level relative exposures ℓ_n.
2. **Subpopulation conditioning**: Compute BandNorm auxiliary features, fit MiniBatchK-Means ($J \in [16, 32]$ clusters with centers $c_1, \ldots, c_J$), and obtain soft population conditions $q_n \in \Delta^{J-1}$.
3. **Canonical seeds**: Form depth-normalized mini-pseudo-bulks (groups of ≤ 16 cells) and factorize seed latents $Z^\star_g$, aligned via Generalized Procrustes Analysis (GPA).
4. **Synthetic recovery pairs**: Generate thinned counts from seed rates at single-cell depths and fit corrupted recovery factors $\widetilde{Z}^{\text{corr}}_g$.
5. **Prior pretraining**: Pretrain the conditional diffusion score network $\boldsymbol{\epsilon}_\psi$ with Equations 3 and 5.
6. **Latent initialization**: Initialize each training single-cell latent from its subpopulation prototype center: $Z^{(0)}_n = \bar{Z}_{j(n)} + \boldsymbol{\epsilon}_{\text{init}}$, where $\bar{Z}_j = \frac{1}{|G_j|} \sum_{g \in G_j} Z^\star_g$ is the average of seeds belonging to cluster j, and $\boldsymbol{\epsilon}_{\text{init}} \sim \mathcal{N}(0, 0.02^2 I)$.

7. **Generalized EM optimization**: For 50 outer EM iterations, alternate:
   (a) **E-step**: Execute $K_E = 20$ inner alternating steps. In each step $k \in \{1, \ldots, K_E\}$, take an Adam gradient step on the observation pseudo-likelihood $\mathcal{L}_{\text{obs}}$ ($\eta_E = 10^{-2}$), followed by the conditional diffusion proximal correction at timestep $t_k = \text{round}(t_{\text{start}} + \frac{k-1}{K_E-1}(t_{\min} - t_{\text{start}}))$ with $S = 2$ Monte Carlo noise samples (Equation 7) and trust-region projection (Equation 8).
   (b) **M-step**: In the first stage, optimize observation parameters (ϕ, θ_d) under $\mathcal{L}_{\text{obs}}$ for 1 epoch (lr $= 10^{-3}$). In the second stage, update the score network $\boldsymbol{\epsilon}_\psi$ under $\mathcal{L}_{\text{DSM}} + \mathcal{L}_{\text{rec}}$ for 5 epochs (lr $= 10^{-4}$) with Min-SNR weighting and EMA decay 0.9999.
8. **Decoding and out-of-sample inference**: Decode common-depth interaction rates via Equation 1 and apply background thresholding $\tau = 0.01$ (Equation 9). For an unseen query cell, its auxiliary preprocessed band is projected to compute condition q_{query}; its latent is initialized from the nearest cluster prototype $Z^{(0)}_{\text{query}} = \bar{Z}_{j^*}$ ($j^* = \text{argmax}_j\, q_{\text{query},j}$) and restored via the MAP E-step using its observed raw counts.

D DEFAULT HYPERPARAMETERS

See Table 7. To assess scHiDiffEM as a general-purpose framework for scHi-C analysis, we used a shared set of model hyperparameters across all datasets and resolutions without dataset- or resolution-specific fine-tuning.

Table 7: Default scHiDiffEM hyperparameters (used in all main experiments).

Component	Parameter	Value
Decoder	latent rank K / band width D	10 / 50 bins
Diffusion	timesteps T / schedule	1,000 / cosine
Generalized EM	outer iterations	50
E-step	inner steps per iter (K_E) / learning rate (η_E)	20 / 10^{-2}
Observation M-step	epochs / learning rate	1 / 10^{-3}
Diffusion M-step	epochs / learning rate	5 / 10^{-4}
Diffusion training	Min-SNR γ / EMA decay	5 / 0.9999
MAP prior	relaxation schedule (λ_m)	$0.02 \rightarrow 0.10$
MAP denoising	timestep (t) / MC samples (S) / trust radius ($\rho_{\max}$)	$250 \rightarrow 10$ / 2 / 0.05
NB2	initial / bounded dispersion	5 / $[0.1, 100]$
Population seeds	group size / maximum seeds	16 / 320
Population model	automatic cluster range (J)	16–32
Recovery	latent-fit steps / loss weight	20 / 1
Output	background threshold cutoff (τ)	0.01

E ADDITIONAL UMAP VISUALIZATIONS

The following figures show resolution-specific method comparisons, with resolutions stacked vertically. Each row contains scHiDiffEM (Ours), scHiCluster, Higashi, Fast-Higashi, and scVI-3D from left to right, using the same matched cells and annotation colors within that setting. Points are colored by dataset annotations, with legends below each row.

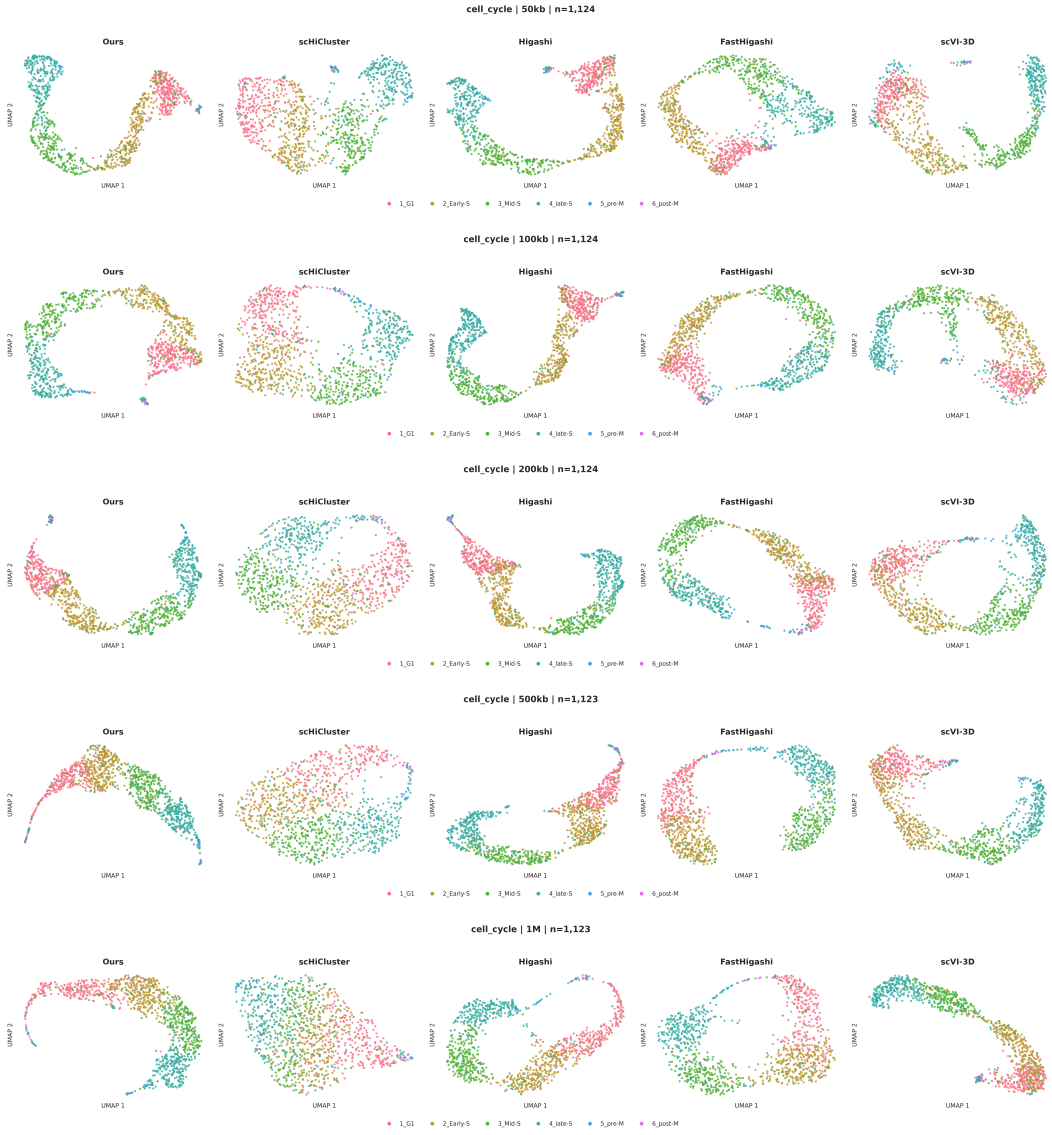

Figure 4: UMAP visualizations for the cell-cycle collection with sequencing resolutions 50 kb, 100 kb, 200 kb, 500 kb, and 1 Mb, from top to bottom.

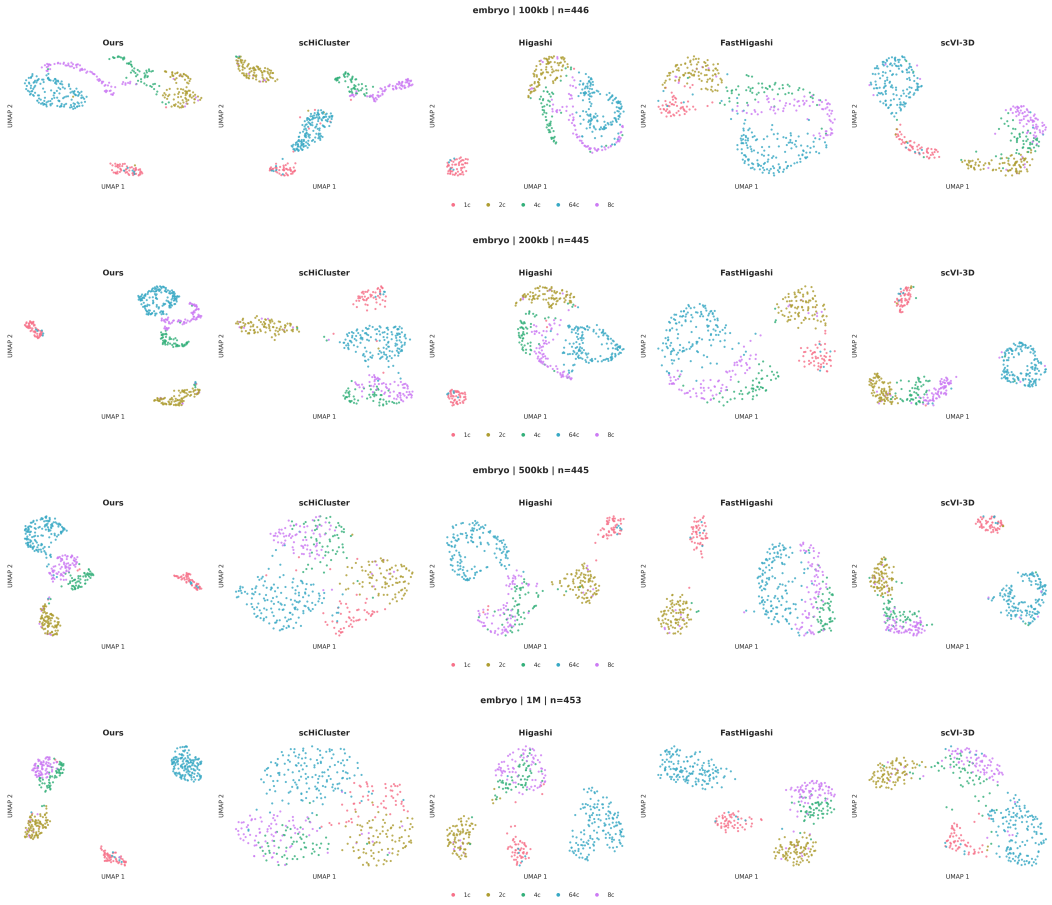

Figure 5: Full method comparisons for the embryo collection with sequencing resolutions 100 kb, 200 kb, 500 kb, and 1 Mb, from top to bottom.



Figure 6: UMAP visualizations for the HiRes mouse brain collection at 100 kb, 200 kb, and 500 kb, from top to bottom.

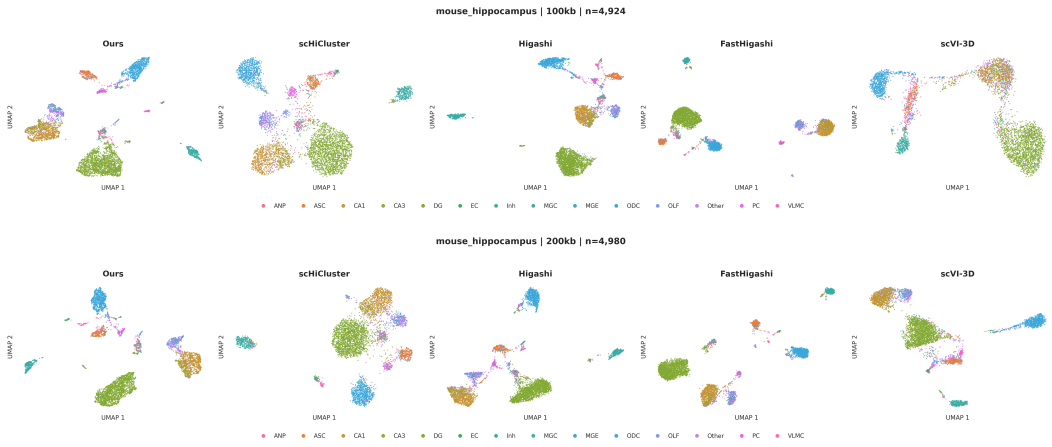

Figure 7: Full method comparisons for mouse hippocampus with sequencing resolutions 100 kb (top) and 200 kb (bottom).

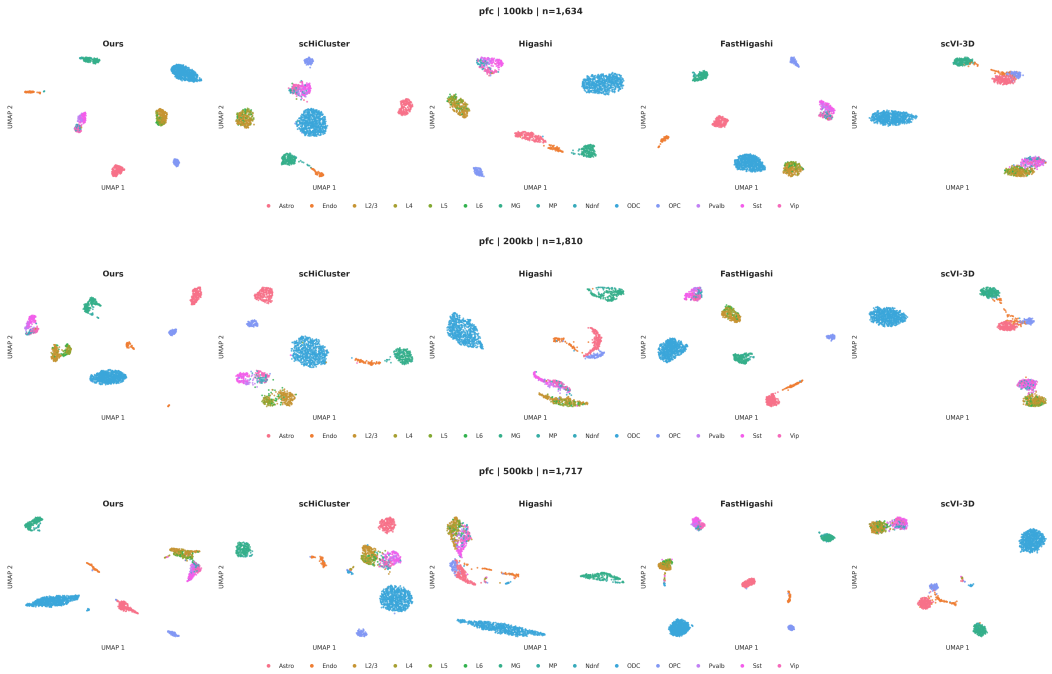

Figure 8: Additional UMAP visualizations for the prefrontal-cortex collection with sequencing resolutions 100 kb, 200 kb, and 500 kb, from top to bottom.

F ADDITIONAL CONTACT-MAP COMPARISONS

Split-half aggregates. The first four figures reproduce the complete original plots, with all sampled types and both directions. The query half contains benchmark cells; the reference is the opposite half of the full labeled population. All methods use the same query cells. These extend the main examples, not additional experiments or independent biological replicates.

Individual cells. The last two figures reproduce the original single-cell comparisons. Each row contains one cell, whereas each aggregate row above combines multiple cells. Both use an opposite-half population reference, which is not individual-cell ground truth. Original normalization, genomic coordinates, masks, and colorbars are retained throughout.

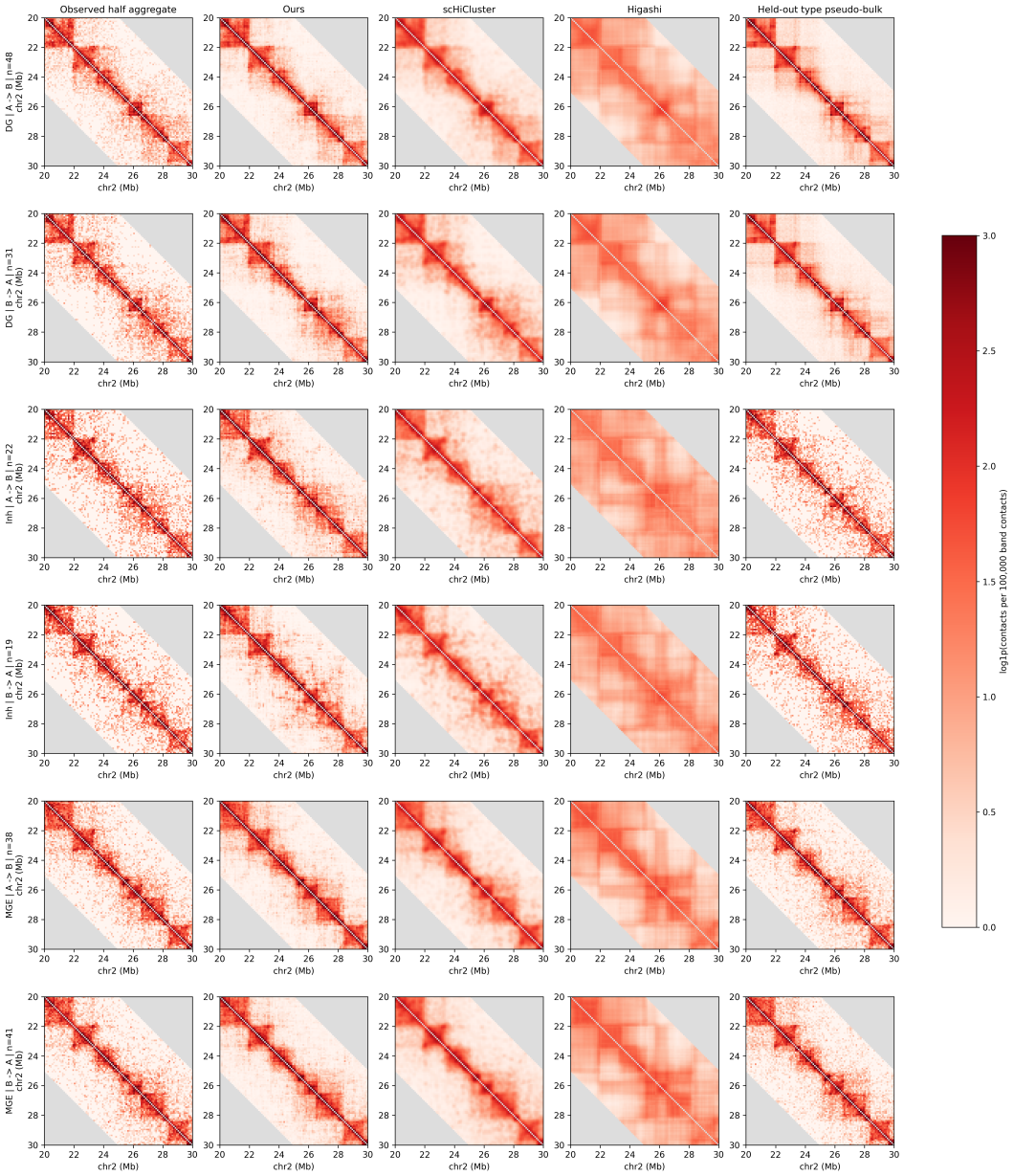

Figure 9: Complete split-half aggregates for mouse hippocampus at 100 kb. Dentate gyrus (DG), inhibitory (Inh), and MGE annotations are each shown in both directions; the MGE B-to-A row contains the single cell used in the main figure. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

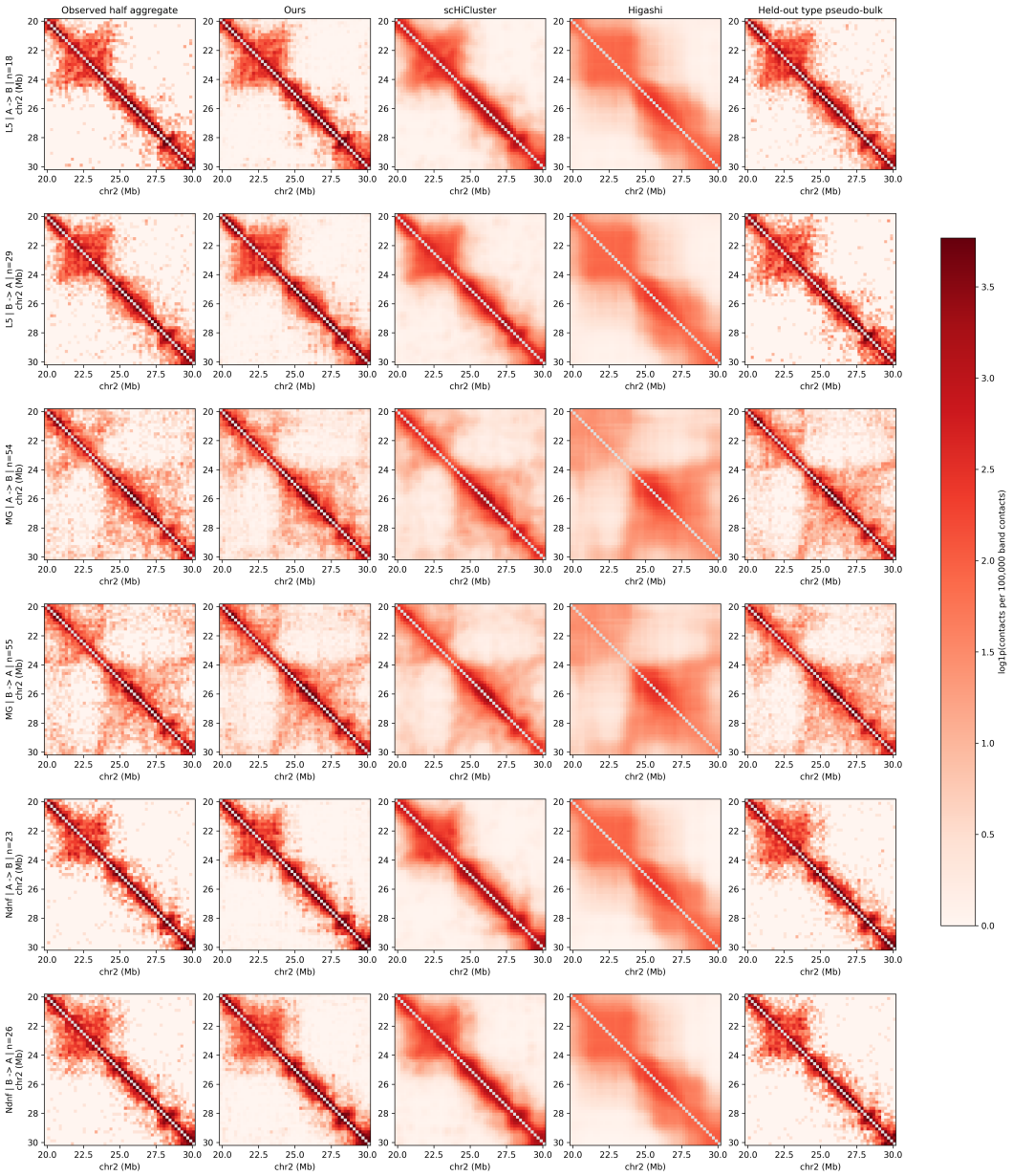

Figure 10: Split-half aggregates for PFC at 200 kb. L5, microglia (MG), and neuron-derived neurotrophic factor (Ndnf)-annotated cells are shown in both directions, extending the L5 A-to-B example in the main figure. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

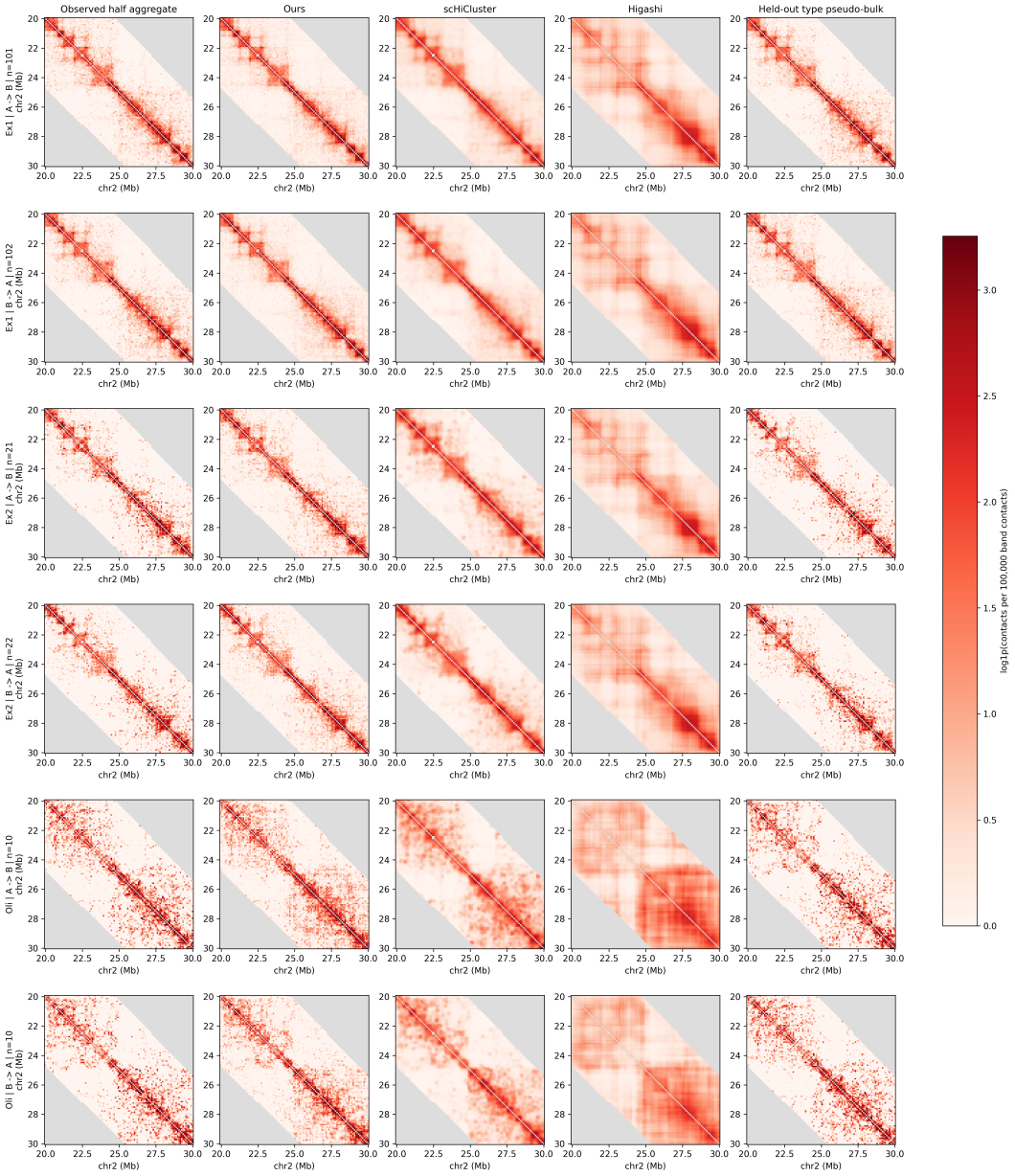

Figure 11: Split-half aggregates for HIRES brain at 100 kb, showing excitatory groups 1 and 2 (Ex1 and Ex2) and oligodendrocytes (Oli) in both directions. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

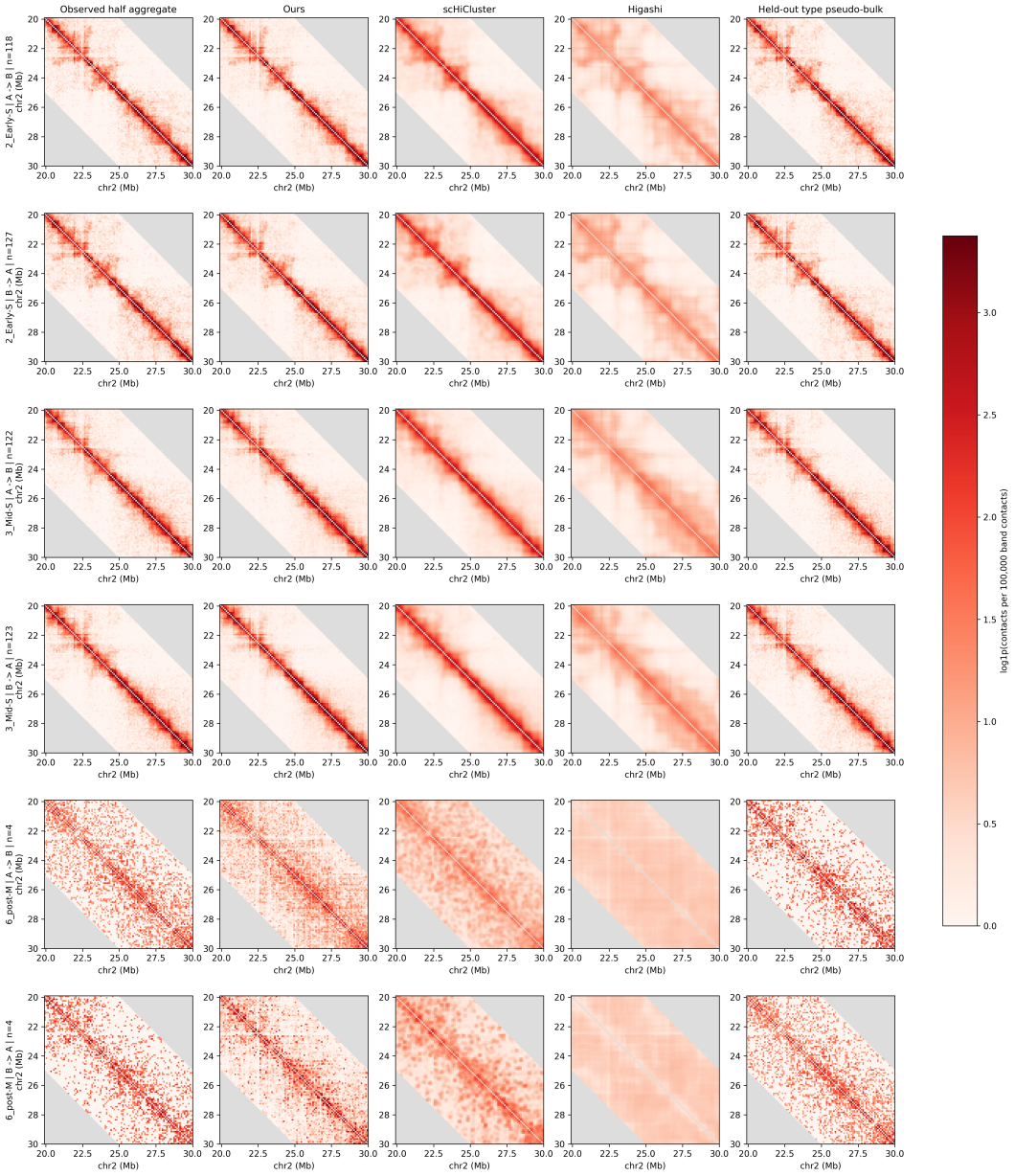

Figure 12: Split-half aggregates for the cell-cycle collection at 100 kb. Early-S, Mid-S, and post-M groups are shown in both directions. The post-M query groups contain only four cells each, limiting reference reliability. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

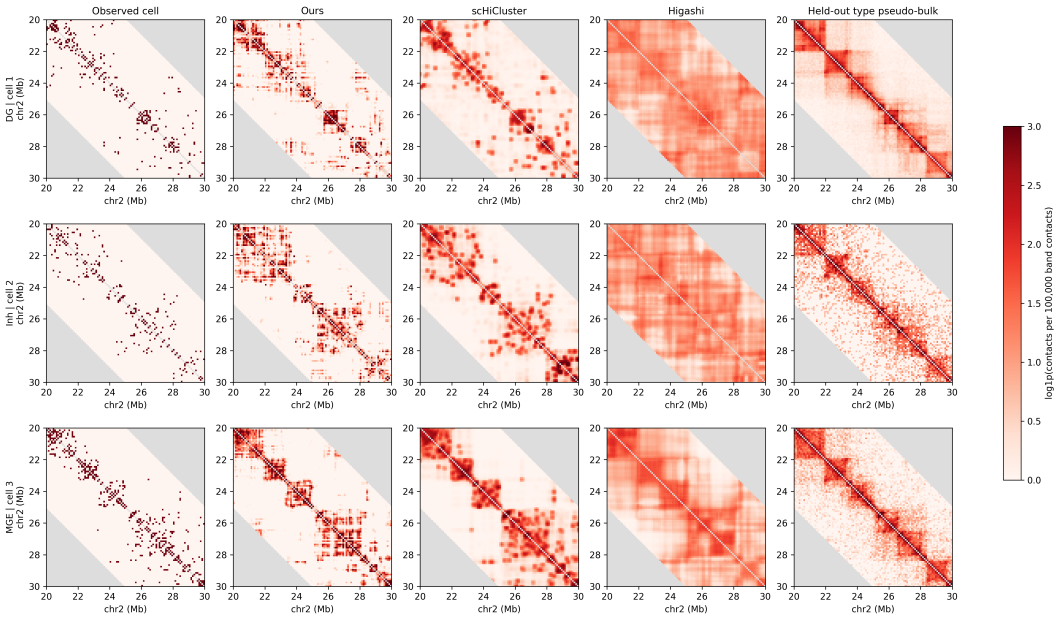

Figure 13: Individual-cell comparisons for mouse hippocampus at 100 kb: DG, Inh, and MGE. Each row shows one observed cell and its three method outputs against an opposite-half type reference. The MGE example is also shown in the main figure; the reference is not that cell's clean map. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

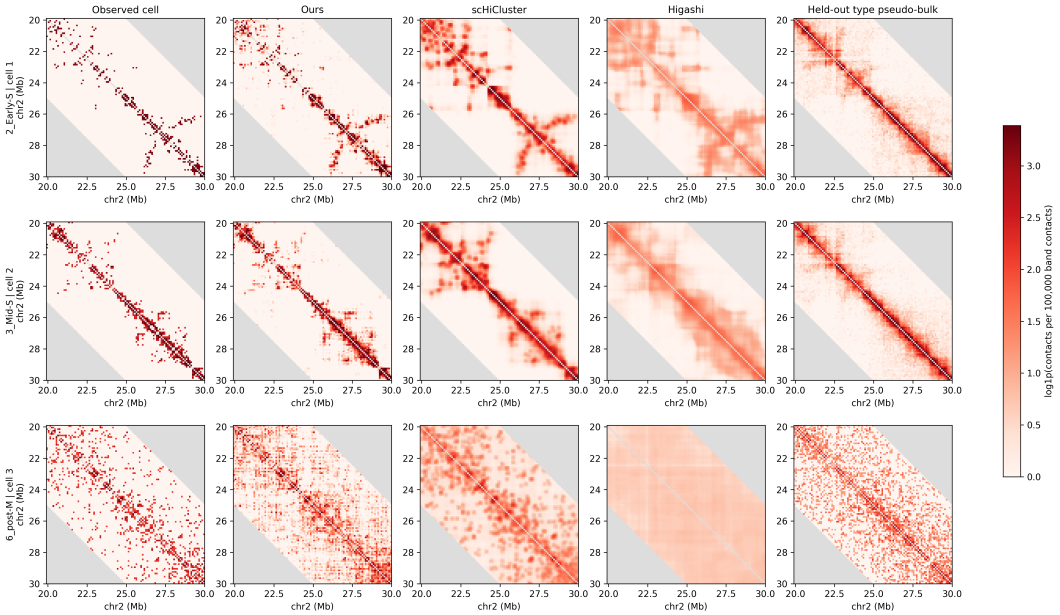

Figure 14: Individual-cell comparisons for the cell-cycle collection at 100 kb: Early-S, Mid-S, and post-M. These are single-cell outputs, not the aggregated maps in the preceding split-half figure. The pseudo-bulk reference describes the opposite-half population. Columns follow the original plot. All maps demonstrate the 20–30 Mb region of chromosome 2, with a shared color scale within the figure.

G DATASET-LEVEL METRICS

The following tables report the original performance metrics separately for every dataset–resolution setting and method, rather than sums of metrics. Embedding values are taken directly from the dataset-level evaluation. Contact-imputation and structure values use the saved chromosome-summary means, retaining the evaluator's aggregation without pooling different resolutions.

For imputation, PG and SG denote global Pearson and Spearman correlation, PS denotes stratum-wise Pearson, and TP denotes type-specific Pearson. Top-1 and Margin summarize type retrieval. For structure, Ins. P and Ins. S denote insulation Pearson and Spearman correlation; F1 uses a tolerance of two bins. Cell counts and other bookkeeping fields are not performance scores and are not summed into these tables.

Table 8: Embedding: original metrics for each dataset and resolution.

Method	ARI	NMI	AMI	Silhouette
cell cycle 50 kb				
scHiDiffEM	0.6503	0.6170	0.6150	0.0405
scHiCluster	0.6059	0.5777	0.5756	-0.1571
Higashi	0.6735	0.6287	0.6264	0.0212
Fast-Higashi	0.6482	0.6059	0.6039	0.0579
scVI-3D	0.4628	0.4819	0.4786	-0.0774
cell cycle 100 kb				
scHiDiffEM	0.6467	0.6331	0.6308	0.1171
scHiCluster	0.5637	0.5617	0.5594	-0.0731
Higashi	0.5716	0.5860	0.5833	0.0705
Fast-Higashi	0.5518	0.5517	0.5488	0.0329
scVI-3D	0.5579	0.5387	0.5357	-0.0429
cell cycle 200 kb				
scHiDiffEM	0.6282	0.6151	0.6126	0.1384
scHiCluster	0.6005	0.5745	0.5718	-0.0559
Higashi	0.5761	0.5831	0.5800	0.1058
Fast-Higashi	0.5382	0.5471	0.5448	0.0226
scVI-3D	0.5685	0.5651	0.5623	0.0127
cell cycle 500 kb				
scHiDiffEM	0.6094	0.6059	0.6034	0.1188
scHiCluster	0.4477	0.4745	0.4705	-0.0729
Higashi	0.5975	0.5857	0.5831	0.1068
Fast-Higashi	0.5553	0.5657	0.5630	0.0159
scVI-3D	0.5493	0.5724	0.5697	0.0371
cell cycle 1 Mb				
scHiDiffEM	0.5712	0.5692	0.5664	0.0836
scHiCluster	0.3067	0.3293	0.3250	-0.0763
Higashi	0.5751	0.5706	0.5674	0.0902
Fast-Higashi	0.5692	0.5646	0.5618	-0.0040
scVI-3D	0.4943	0.5128	0.5096	-0.0015
embryo 100 kb				
scHiDiffEM	0.5872	0.7316	0.7266	0.0291
scHiCluster	0.7699	0.7794	0.7772	0.0083
Higashi	0.5332	0.5978	0.5930	-0.0367
Fast-Higashi	0.7048	0.6936	0.6899	0.0415
scVI-3D	0.6468	0.6575	0.6542	0.0988
embryo 200 kb				
scHiDiffEM	0.8724	0.8476	0.8458	0.0610
scHiCluster	0.7554	0.7726	0.7704	-0.0191
Higashi	0.6563	0.6561	0.6527	-0.0112
Fast-Higashi	0.7133	0.7507	0.7483	0.0418
scVI-3D	0.7413	0.7259	0.7232	0.1365

Continued on next page

Method	ARI	NMI	AMI	Silhouette
embryo 500 kb				
scHiDiffEM	0.8418	0.8287	0.8266	0.1017
scHiCluster	0.7174	0.7114	0.7086	-0.0421
Higashi	0.5435	0.6724	0.6692	-0.0080
Fast-Higashi	0.7033	0.7492	0.7467	0.0323
scVI-3D	0.7434	0.7257	0.7230	0.1190
embryo 1 Mb				
scHiDiffEM	0.8034	0.7946	0.7916	0.1291
scHiCluster	0.6496	0.6473	0.6440	-0.0445
Higashi	0.7800	0.7928	0.7909	0.0129
Fast-Higashi	0.8060	0.8327	0.8311	0.0275
scVI-3D	0.7320	0.7168	0.7141	0.0929
hires brain 100 kb				
scHiDiffEM	0.4340	0.6437	0.6357	0.1182
scHiCluster	0.5597	0.5769	0.5692	-0.0404
Higashi	0.1104	0.3205	0.2922	-0.1015
Fast-Higashi	0.4116	0.5452	0.5384	0.0108
scVI-3D	0.0043	0.0805	0.0217	-0.0816
hires brain 200 kb				
scHiDiffEM	0.3875	0.6603	0.6488	0.1189
scHiCluster	0.4005	0.5427	0.5358	0.0088
Higashi	0.1298	0.3399	0.3070	-0.0596
Fast-Higashi	0.3985	0.5215	0.5144	0.0090
scVI-3D	0.0241	0.0932	0.0480	-0.0683
hires brain 500 kb				
scHiDiffEM	0.4535	0.6398	0.6319	0.0751
scHiCluster	0.3705	0.5124	0.5050	-0.0029
Higashi	0.1299	0.3465	0.3168	-0.0156
Fast-Higashi	0.5567	0.5648	0.5567	0.0077
scVI-3D	0.0364	0.1656	0.1178	-0.0545
mouse hippocampus 100 kb				
scHiDiffEM	0.8113	0.7399	0.7381	0.2222
scHiCluster	0.8305	0.7291	0.7271	0.1025
Higashi	0.8231	0.7328	0.7314	0.1792
Fast-Higashi	0.8369	0.7534	0.7518	0.0895
scVI-3D	0.6978	0.6340	0.6329	0.0980
mouse hippocampus 200 kb				
scHiDiffEM	0.8396	0.7522	0.7508	0.2650
scHiCluster	0.8318	0.7251	0.7233	0.0958
Higashi	0.8205	0.7238	0.7226	0.1856
Fast-Higashi	0.8378	0.7484	0.7468	0.1373
scVI-3D	0.7743	0.6977	0.6962	0.1578
pfc 50 kb				
scHiDiffEM	0.8907	0.8713	0.8700	0.2115
scHiCluster	0.8816	0.8565	0.8551	0.2217
Higashi	0.7070	0.7380	0.7366	0.1709
Fast-Higashi	0.5821	0.7279	0.7249	0.1319
scVI-3D	0.5765	0.6225	0.6193	0.0031
pfc 100 kb				
scHiDiffEM	0.8913	0.8838	0.8823	0.4411
scHiCluster	0.8885	0.8766	0.8750	0.0500
Higashi	0.8852	0.8732	0.8715	0.2356
Fast-Higashi	0.8890	0.8860	0.8845	0.3088
scVI-3D	0.8346	0.8085	0.8063	0.1518
pfc 200 kb				
scHiDiffEM	0.8946	0.8882	0.8869	0.4172
scHiCluster	0.9198	0.8962	0.8948	0.1130

Continued on next page

Method	ARI	NMI	AMI	Silhouette
Higashi	0.8397	0.8170	0.8152	0.2429
Fast-Higashi	0.8934	0.8831	0.8817	0.3450
scVI-3D	0.8780	0.8458	0.8439	0.1888
pfc 500 kb				
scHiDiffEM	0.8689	0.8628	0.8608	0.3137
scHiCluster	0.8655	0.8455	0.8432	0.0867
Higashi	0.7189	0.7321	0.7303	0.1844
Fast-Higashi	0.8743	0.8585	0.8565	0.2029
scVI-3D	0.8564	0.8213	0.8187	0.1854

Table 9: Cell-type contact imputation: original metrics for each dataset and resolution.

Method	SCC	PG	SG	PS	TP	Top-1	Margin
cell cycle 50 kb							
scHiDiffEM	0.5838	0.7226	0.6750	0.5345	0.1735	0.8333	0.1107
scHiCluster	0.5298	0.7063	0.6800	0.5067	0.1140	0.1667	-0.0845
Higashi	0.4544	0.4766	0.5077	0.4361	0.1165	0.1667	-0.2020
scVI-3D	0.6826	0.7968	0.7483	0.6372	0.1768	0.8220	0.0877
cell cycle 100 kb							
scHiDiffEM	0.6563	0.7913	0.7569	0.6097	0.2388	0.8333	0.1614
scHiCluster	0.5822	0.7601	0.7454	0.5614	0.1352	0.1667	-0.0769
Higashi	0.5255	0.5849	0.6019	0.4998	0.1252	0.1667	-0.2250
scVI-3D	0.7283	0.8374	0.8028	0.6825	0.1911	0.7722	0.0734
cell cycle 200 kb							
scHiDiffEM	0.7141	0.8472	0.8107	0.6708	0.3041	0.8333	0.1716
scHiCluster	0.6044	0.7963	0.7795	0.5832	0.1605	0.1667	-0.1075
Higashi	0.5093	0.4703	0.7498	0.4890	0.0767	0.1667	-0.3869
scVI-3D	0.7683	0.8725	0.8440	0.7224	0.1798	0.6951	0.0607
cell cycle 500 kb							
scHiDiffEM	0.7862	0.8949	0.8496	0.7624	0.3736	0.9199	0.1606
scHiCluster	0.6032	0.8055	0.7814	0.5890	0.1727	0.1929	-0.1434
Higashi	0.7342	0.8517	0.8322	0.7188	0.2364	0.2648	-0.1407
scVI-3D	0.8126	0.9006	0.8686	0.7836	0.1817	0.7196	0.0455
cell cycle 1 Mb							
scHiDiffEM	0.8275	0.9168	0.8787	0.8206	0.4286	0.9600	0.1644
scHiCluster	0.5574	0.7741	0.7673	0.5515	0.1704	0.3320	-0.1333
Higashi	0.7756	0.8625	0.8347	0.7698	0.2721	0.3573	-0.0951
scVI-3D	0.8371	0.9063	0.8645	0.8232	0.1905	0.6378	0.0425
embryo 100 kb							
scHiDiffEM	0.5224	0.8944	0.7885	0.4647	0.1654	0.9401	0.1126
scHiCluster	0.4585	0.8763	0.8154	0.4368	0.0703	0.3760	-0.0005
Higashi	0.4277	0.8422	0.7756	0.4011	0.0715	0.5642	0.0396
scVI-3D	0.6609	0.9305	0.8524	0.5999	0.1590	0.9632	0.0909
embryo 200 kb							
scHiDiffEM	0.5394	0.8988	0.7899	0.5139	0.1788	0.8000	0.1085
scHiCluster	0.4474	0.8676	0.7997	0.4336	0.0758	0.3323	-0.0144
Higashi	0.5049	0.8508	0.8139	0.4998	0.0831	0.6408	0.0526
scVI-3D	0.6507	0.9284	0.8457	0.6211	0.1377	0.7269	0.0464
embryo 500 kb							
scHiDiffEM	0.6031	0.8814	0.7906	0.6089	0.2035	0.7894	0.0785
scHiCluster	0.4445	0.8188	0.7534	0.4250	0.0897	0.4072	-0.0351
Higashi	0.5689	0.8032	0.7894	0.5775	0.1137	0.7621	0.0572
scVI-3D	0.6634	0.9018	0.8239	0.6640	0.1124	0.4104	-0.0355

Continued on next page

Method	SCC	PG	SG	PS	TP	Top-1	Margin
embryo 1 Mb							
scHiDiffEM	0.7418	0.8702	0.8091	0.7429	0.2002	0.5231	0.0069
scHiCluster	0.5495	0.7586	0.7195	0.5428	0.0749	0.2262	-0.0815
Higashi	0.6958	0.7869	0.7790	0.7021	0.1069	0.5690	0.0435
scVI-3D	0.7486	0.8757	0.7890	0.7359	0.0710	0.2488	-0.1464
hires brain 100 kb							
scHiDiffEM	0.4519	0.6795	0.6281	0.4355	0.1051	0.6423	0.0135
scHiCluster	0.4619	0.6943	0.6718	0.4567	0.1011	0.1892	-0.0287
Higashi	0.4282	0.6075	0.5898	0.4192	0.0811	0.1429	-0.0815
scVI-3D	0.5249	0.7430	0.7008	0.5145	0.0703	0.1516	-0.0655
hires brain 200 kb							
scHiDiffEM	0.5366	0.7712	0.7198	0.5254	0.1348	0.7640	0.0319
scHiCluster	0.5241	0.7632	0.7395	0.5196	0.0889	0.2114	-0.0259
Higashi	0.5151	0.7439	0.7016	0.4957	0.0855	0.1429	-0.0943
scVI-3D	0.5766	0.8044	0.7601	0.5688	0.0638	0.1599	-0.0831
hires brain 500 kb							
scHiDiffEM	0.5429	0.8047	0.7404	0.5327	0.1590	0.6953	0.0374
scHiCluster	0.4999	0.7806	0.7362	0.4941	0.0767	0.2037	-0.0348
Higashi	0.5297	0.7686	0.7528	0.5312	0.0609	0.1549	-0.0886
scVI-3D	0.5765	0.8247	0.7630	0.5649	0.0642	0.2180	-0.0410
mouse hippocampus 100 kb							
scHiDiffEM	0.7644	0.8740	0.8390	0.7272	0.2212	0.5109	0.0391
scHiCluster	0.7441	0.8522	0.8408	0.7224	0.1293	0.2990	-0.0239
Higashi	0.6146	0.6301	0.6462	0.6005	0.0612	0.1489	-0.1225
scVI-3D	0.8275	0.9056	0.8757	0.7918	0.1789	0.2281	-0.1216
mouse hippocampus 200 kb							
scHiDiffEM	0.8457	0.9222	0.9114	0.8406	0.3421	0.5764	0.0667
scHiCluster	0.7624	0.8579	0.8753	0.7621	0.1206	0.1679	-0.0919
Higashi	0.7250	0.8039	0.7973	0.7187	0.1329	0.1484	-0.1505
scVI-3D	0.8917	0.9433	0.9355	0.8831	0.2973	0.2848	-0.0970
pfc 50 kb							
scHiDiffEM	0.4069	0.6910	0.5880	0.3902	0.1086	0.7013	0.0390
scHiCluster	0.4657	0.7342	0.6687	0.4592	0.1180	0.4369	0.0053
Higashi	0.3387	0.4649	0.4830	0.3409	0.0412	0.1332	-0.0730
scVI-3D	0.5866	0.8067	0.7149	0.5654	0.0917	0.2652	-0.0045
pfc 100 kb							
scHiDiffEM	0.6304	0.8258	0.7848	0.6227	0.2310	0.6582	0.0760
scHiCluster	0.5973	0.8083	0.7999	0.5945	0.1148	0.3846	0.0014
Higashi	0.4564	0.5787	0.5973	0.4535	0.0424	0.2038	-0.0510
scVI-3D	0.7294	0.8780	0.8441	0.7174	0.2229	0.2874	-0.0276
pfc 200 kb							
scHiDiffEM	0.7128	0.8893	0.8590	0.7093	0.3614	0.5374	0.0759
scHiCluster	0.6442	0.8558	0.8461	0.6395	0.1895	0.5091	0.0290
Higashi	0.6333	0.8031	0.8243	0.6314	0.1096	0.2994	-0.0451
scVI-3D	0.7613	0.9086	0.8808	0.7526	0.2962	0.2844	-0.0322
pfc 500 kb							
scHiDiffEM	0.7383	0.9130	0.8746	0.7269	0.4827	0.4653	0.0699
scHiCluster	0.6067	0.8614	0.8426	0.5935	0.2316	0.4778	0.0166
Higashi	0.6549	0.8620	0.8555	0.6487	0.3252	0.4737	0.0276
scVI-3D	0.7696	0.9198	0.8817	0.7492	0.3216	0.3635	-0.0060

Table 10: Structure recovery: original metrics for each dataset and resolution.

Method	Ins. P	Ins. S	AUROC	AUPRC	$F1_{\pm 2}$
cell cycle 50 kb					
scHiDiffEM	0.9815	0.8319	0.6234	0.1407	0.8073
scHiCluster	0.9524	0.6335	0.6274	0.1432	0.6840
Higashi	0.9479	0.5001	0.5438	0.1137	0.4841
scVI-3D	0.9602	0.8366	0.6110	0.1248	0.7601
cell cycle 100 kb					
scHiDiffEM	0.9819	0.8609	0.6453	0.1712	0.8086
scHiCluster	0.9400	0.6003	0.6406	0.1665	0.6475
Higashi	0.9490	0.5089	0.5229	0.1204	0.4412
scVI-3D	0.9556	0.8461	0.6484	0.1617	0.8006
cell cycle 200 kb					
scHiDiffEM	0.9844	0.8714	0.6612	0.2121	0.7961
scHiCluster	0.8960	0.5775	0.6343	0.1899	0.6042
Higashi	0.9401	0.5981	0.5427	0.1448	0.2188
scVI-3D	0.9612	0.8567	0.6722	0.2066	0.8123
cell cycle 500 kb					
scHiDiffEM	0.8914	0.8435	0.6496	0.2612	0.7348
scHiCluster	0.5612	0.5009	0.6122	0.2205	0.5843
Higashi	0.7539	0.6744	0.5832	0.2066	0.5614
scVI-3D	0.8817	0.8405	0.6803	0.2749	0.7912
cell cycle 1 Mb					
scHiDiffEM	0.8510	0.8065	0.6535	0.2944	0.6654
scHiCluster	0.4787	0.4452	0.6090	0.2459	0.5794
Higashi	0.7060	0.6426	0.6174	0.2506	0.5979
scVI-3D	0.8507	0.8189	0.6985	0.3384	0.7591
embryo 100 kb					
scHiDiffEM	0.9007	0.8569	0.6840	0.2266	0.7967
scHiCluster	0.6569	0.5561	0.6483	0.2053	0.6421
Higashi	0.6532	0.5234	0.5876	0.1626	0.5545
scVI-3D	0.8916	0.8424	0.6672	0.2080	0.7787
embryo 200 kb					
scHiDiffEM	0.8978	0.8391	0.7025	0.2707	0.8234
scHiCluster	0.5909	0.5463	0.6460	0.2271	0.6402
Higashi	0.7214	0.5961	0.5950	0.1894	0.5792
scVI-3D	0.8867	0.8291	0.6909	0.2518	0.8071
embryo 500 kb					
scHiDiffEM	0.8737	0.8246	0.6832	0.2856	0.7795
scHiCluster	0.5742	0.5588	0.6303	0.2467	0.6165
Higashi	0.7414	0.6853	0.5867	0.2068	0.5589
scVI-3D	0.8702	0.8300	0.6839	0.2827	0.7917
embryo 1 Mb					
scHiDiffEM	0.8402	0.8103	0.6701	0.3042	0.6820
scHiCluster	0.5752	0.5658	0.6310	0.2645	0.6182
Higashi	0.7119	0.6751	0.6223	0.2562	0.5727
scVI-3D	0.8400	0.8096	0.7050	0.3444	0.7694
hires brain 100 kb					
scHiDiffEM	0.8848	0.8660	0.6737	0.2225	0.8027
scHiCluster	0.7355	0.7258	0.6592	0.2112	0.6854
Higashi	0.7509	0.7069	0.5763	0.1620	0.5855
scVI-3D	0.8718	0.8479	0.6470	0.1930	0.7408
hires brain 200 kb					
scHiDiffEM	0.8910	0.8731	0.7029	0.2664	0.8291
scHiCluster	0.7323	0.7413	0.6813	0.2425	0.7073
Higashi	0.7486	0.7029	0.6104	0.1927	0.6996
scVI-3D	0.8679	0.8459	0.6848	0.2365	0.8044

Continued on next page

Method	Ins. P	Ins. S	AUROC	AUPRC	$F1_{\pm 2}$
hires brain 500 kb					
scHiDiffEM	0.8444	0.8296	0.6820	0.2963	0.7995
scHiCluster	0.7185	0.7083	0.6567	0.2671	0.7048
Higashi	0.7567	0.7308	0.6153	0.2327	0.6834
scVI-3D	0.8407	0.8293	0.7138	0.3121	0.8544
mouse hippocampus 100 kb					
scHiDiffEM	0.9973	0.9490	0.6610	0.1415	0.8521
scHiCluster	0.9761	0.8447	0.6404	0.1343	0.7600
Higashi	0.9898	0.6550	0.5325	0.1015	0.5488
scVI-3D	0.9812	0.9359	0.6631	0.1307	0.8612
mouse hippocampus 200 kb					
scHiDiffEM	0.9598	0.9428	0.7271	0.2770	0.8956
scHiCluster	0.9448	0.8341	0.6865	0.2478	0.8133
Higashi	0.9740	0.7337	0.6146	0.2030	0.7791
scVI-3D	0.9799	0.9422	0.7456	0.2854	0.9219
pfc 50 kb					
scHiDiffEM	0.8996	0.8912	0.6895	0.1814	0.7614
scHiCluster	0.8110	0.8130	0.6657	0.1688	0.7165
Higashi	0.5720	0.5455	0.5984	0.1373	0.5466
scVI-3D	0.8907	0.8682	0.6684	0.1582	0.7468
pfc 100 kb					
scHiDiffEM	0.9097	0.9171	0.7363	0.2254	0.8522
scHiCluster	0.7809	0.7892	0.7407	0.2283	0.7562
Higashi	0.7634	0.5354	0.6082	0.1496	0.6284
scVI-3D	0.9587	0.9048	0.7383	0.2121	0.8529
pfc 200 kb					
scHiDiffEM	0.9378	0.9249	0.7508	0.2947	0.8748
scHiCluster	0.8222	0.8322	0.7079	0.2499	0.7342
Higashi	0.8087	0.7828	0.6461	0.2041	0.6542
scVI-3D	0.9280	0.9160	0.7598	0.2896	0.8760
pfc 500 kb					
scHiDiffEM	0.9273	0.9055	0.7174	0.3232	0.8549
scHiCluster	0.8075	0.8002	0.6704	0.2742	0.7042
Higashi	0.7313	0.7085	0.6459	0.2575	0.7754
scVI-3D	0.9214	0.9022	0.7405	0.3325	0.8788