Supplementary Material to
"Spatial Cellular Networks from omics data with SpaCeNet"

1 Simulation studies

Data simulation Data were simulated from the full probability density (main article, Eq. (1)) with potentials $\rho_{ij}(r_{ab}) = \Delta\rho_{ij} \exp(-\phi_{ij} r_{ab})$, using the following procedure:

1. Initialize empty $p \times p$ matrices for $\mathbf{\Omega}$ and $\mathbf{\Delta\boldsymbol{\rho}}$.
2. Randomly draw a set of symmetric edges (we chose 5% of all possible edges) for $\mathbf{\Delta\boldsymbol{\rho}}$ from Unif$(-1, 1)$ and draw corresponding range parameters $\phi_{ij} = \phi_{ji}$ for given positive ϕ_{min} and ϕ_{max} from Unif(ϕ_{min}, ϕ_{max}).
3. Randomly draw a set of symmetric edges (we chose 10% of all possible edges) for $\mathbf{\Omega}$ from Unif$(-1, 1)$.
4. Draw $s = 1, \ldots, S$ samples of spatial coordinates, uniformly sampled from 3d space with an average density of 100 cells per volume unit, each with n cells and calculate the pairwise distances $\mathbf{R}^{(s)}$ for each sample.
5. Based on $\mathbf{\Omega}$, $\mathbf{\Delta\boldsymbol{\rho}}$, ϕ_{ij} and $\mathbf{R}^{(s)}$ construct $\mathbf{\Lambda}^{(s)}$ for all s.
6. Calculate the row-wise sum of the absolute values of all $\mathbf{\Lambda}^{(s)}$. For each of the p variables, select the maximum value from the corresponding $n \cdot S$ sums, add a small constant (we chose 10^{-7}) and fill the respective diagonal element of $\mathbf{\Omega}$ with it to ensure that Λ is positive definite.
7. Update all $\mathbf{\Lambda}^{(s)}$, sample $\boldsymbol{\xi}^{(s)}$ from $\mathbf{\Lambda}^{(s)}$ by means of a Cholesky decomposition and add a random $\boldsymbol{\mu}$ if desired.

With this setup, we simulated 24 scenarios with different parameter settings for n, S and ϕ_{ij} with 20 independently seeded replicates for each. All simulations used $p = 20$ variables. Note that the data are simulated from the full joint probability Eq. (1) of the main manuscript with a precision matrix of size $pn \times pn$, making data simulation computationally expensive for large p and n.

In our studies, ϕ_{ij} was chosen between 5 and 20. The rationale behind this choice is motivated as follows. Given a unit density of η, k cells on average occupy a volume $V_k = k/\eta$, which corresponds to a sphere of radius $r_k = (3k/(4\pi\eta))^{1/3}$. We used r_k as an estimate for the average distance between a cell and its k nearest neighbors. For $\eta = 100$, this yields $r_1 \approx 0.13$, $r_8 = 2r_1 \approx 0.27$, and $r_{27} = 3r_1 \approx 0.40$. The range of the potentials $\exp(-\phi_{ij} r)$ can be quantified by $1/\phi_{ij}$. With the average nearest-neighbor distance r_1 as reference point, $\phi_{ij} = 5$ therefore corresponds to a long-range and $\phi_{ij} = 20$ to a short-range potential.

SpaCeNet model selection We set the expansion order to $L = 3$ and chose r_0 to equal the minimal observed distance between two cells. Note that there are reasonable alternative choices for r_0, e.g., the average nearest-neighbor distance provides a length scale that might be more appropriate for larger numbers of cells n. We then performed a grid search on all data sets with 4 values for $\alpha \in [10^{-5}, 10^{-3}, 10^{-1}, 10]$ and $\beta \in [10^{-5}, 10^{-3}, 10^{-1}, 10]$ each. The grid was then successively refined 6 times such that about 100 different hyper-parameter combinations were evaluated in total. The best set of hyper-parameters was chosen based on the maximum pseudo-log-likelihood of test data. To this end, the full data set was split 70:30 into a training and a test set. If more than one sample was available ($S > 1$), the split was performed between different samples and otherwise ($S = 1$) all observations in 30 % of the spatial volume were used for testing. We initialized the optimization with a step size of 10^{-6}. If overflows were encountered, we successively reduced the step size by a factor of 10. The convergence threshold for the proximal gradient algorithm was set to 10^{-5} and training was terminated after a maximum of 3,088 optimization steps. The AUROC and AUPRC based evaluation methods rely on a threshold-based classification of estimated parameters into positives and negatives. When considering the spatial association parameters we classify an association between two variables i and j to be positive if $|\Delta\rho_{ij}^{(l)}|$ is greater than the threshold for at least one considered order l in the expansion.

Reconstruction of interaction potentials Data for the reconstruction of the interaction potentials (main article Figure 2) was generated in line with previous procedure, but using $p = 5$ and only a single spatial edge connecting two of the variables with the potentials given in the caption of main article Figure 2A to D, respectively. We simulated data for $n = 10$ and $S = 1000$, and the potentials were fitted with $r_0 = 0.1$.

2 Estimates from conditional densities

By construction, we assume a spatially constant mean vector $\hat{\mathbf{x}} = \boldsymbol{\mu}$ for all cells in the full density distribution of SpaCeNet. Then, we directly obtain from the conditional densities that

$$\hat{x}_j^a = \mu_j - \frac{1}{\omega_{jj}} \left[\sum_{k \neq j}^{p} \omega_{jk}(x_k^a - \mu_k) + \sum_{b=1}^{n} \sum_{l=1}^{L} \sum_{k=1}^{p} (\Theta_{ab}^{(l)} \Delta\rho_{jk}^{(l)})(x_k^b - \mu_k) \right] , \tag{1}$$

where the mean of the normal distribution is shifted. Eq. (1) can be used as an estimate for the variable x_j^a provided all other variables are known.

This method was used to calculate the residuals for the mouse visual cortex data (main article Fig. 4), where $\boldsymbol{\Delta\rho}^{(\cdot)}$ was either set to 0 or the estimated parameters. In this case, information $\mathbf{x}_{\setminus j}^a$ is required, i.e., all other variables of cell a have to be known already.

Similarly, the joint conditional density for all variables of one cell a is given by

$$\begin{aligned} f(\mathbf{x}^a | \mathbf{R}, \mathbf{X}^{\setminus a}) &= \frac{\sqrt{|\boldsymbol{\Omega}|}}{(2\pi)^{p/2}} \exp\left\{ -\frac{1}{2} (\mathbf{x}^a - \boldsymbol{\mu} + \boldsymbol{\Omega}^{-1}\boldsymbol{\rho}^a)^\top \boldsymbol{\Omega} (\mathbf{x}^a - \boldsymbol{\mu} + \boldsymbol{\Omega}^{-1}\boldsymbol{\rho}^a) \right\} \\ &= \frac{\sqrt{|\boldsymbol{\Omega}|}}{(2\pi)^{p/2}} \exp\left\{ -\frac{1}{2} (\mathbf{x}^a - \hat{\mathbf{x}}^a)^\top \boldsymbol{\Omega} (\mathbf{x}^a - \hat{\mathbf{x}}^a) \right\} \end{aligned}$$

with

$$\begin{aligned}
\hat{\mathbf{x}}^a &= \boldsymbol{\mu} - \boldsymbol{\Omega}^{-1}\boldsymbol{\rho}^a \, , \\
\rho_j^a &= \sum_{l=1}^{L}\sum_{k=1}^{p}\sum_{b=1}^{n} \Theta_{ab}^{(l)}(x_k^b - \mu_k)\Delta\rho_{jk}^{(l)} \, .
\end{aligned}$$

The mean vector $\hat{\mathbf{x}}^a$ can be used as an estimate for the variables of cell a, provided that all variables of all other cells are known. Note that $\boldsymbol{\rho}^a$ only depends on other cells $b \neq a$ since $\Theta_{aa}^{(l)} = 0$ for all l. A comparison of the results obtained with the two methods is shown in main article Fig. 6.

3 SpaCeNet analysis of high-throughput spatial transcriptomics from Chen et al. [2022]

For both datasets Mosta A and B, the top hits predominantly showed a positive spatial association ($\Delta\rho_{ij} < 0$). For instance, among the top 5 hits in both analyses, we observed the gene pairs *Mt1-Mt2* and *Mobp-Mbp*. The melatonin receptors Mt1 and Mt2 are G protein-coupled receptors that are activated by the release of melatonin in response to darkness from the pineal gland, which is tightly controlled by circadian pacemaker cells in the suprachiasmatic nucleus (SCN) of the hypothalamus through a multi-synaptic pathway. The clock rhythm is entrained by daily and seasonal changes in the environmental light-dark cycle, which are sensed by melanopsin in a subset of retinal ganglion cells that innervate the SCN [Berson et al., 2002]. The human MT1 and MT2 receptors are 350 and 362 amino acids long, respectively, with molecular weights of 39–40 kDa and 55% amino acid homology overall (70% within the transmembrane domains), while the respective receptors in mouse are 353 and 365 amino acids long. Activation of MT1 receptors promotes sleep and has been mainly implicated in the regulation of rapid eye movement (REM) sleep, whereas MT2 receptors selectively increase non-REM (NREM) sleep and phase shift circadian rhythms of activity and neuronal firing [Dubocovich, 2007]. Aside from these well-known functions, melatonin receptors are involved in various other physiological activities such as blood pressure regulation and immunomodulation [Pandi-Perumal et al., 2008]. Spatial distribution of MT1 and MT2 in mammalian brains has been investigated by various techniques, including receptor autoradiography with [125I]iodomelatonin, nested RT-PCR, Western blotting, immunohistochemistry, and genetically modified mouse models expressing fluorescent protein or a marker enzyme under control of the endogenous *MT1* and *MT2* promoters. *MT1* and *MT2* mRNA has been amplified from human cerebellum, cortex, thalamus, hippocampus, suprachiasmatic nucleus, and retina. Expression of MT1/2 in the latter two structures is consistent with their important role in biological rhythmicity [Dubocovich et al., 2010]. Expression of *MT1* and *MT2* mRNA was also found in rat SCN tissue and confirmed in protein extracts by Western blot using polyclonal rabbit anti-human MT1 and MT2 antiserum [Rivera-Bermúdez et al., 2004]. Protein expression of MT1 receptor in rat SCN varies across the circadian cycle, compared to non-SCN areas, and is the highest at dusk [Waly and Hallworth, 2015]. In contrast, MT2 receptor labeling could not be detected in the rat SCN, but was clearly visible in the paraventricular nucleus and retina. These findings could be corroborated by an immunohistochemical study conducted on the entire adult rat brain using also polyclonal anti-MT1 and anti-MT2 antibodies, which revealed selective MT1 and MT2 localization on neuronal cell bodies and dendrites in numerous regions of the rat telencephalon, mesencephalon, and diencephalon [Lacoste et al., 2015]. This study reported the highest density for MT2 labeling in the hippocampus, thalamus and hypothalamus, except for the SCN. Density of MT1 labeling, in contrast, was the highest in the cerebral cortex and basal forebrain, the epithalamus, the *pars tuberalis* of the pituitary gland, and the SCN. The most comprehensive mapping of the distribution of melatonin receptors in mouse brain, to date, has been accomplished by a knock-in strategy replacing the *Mt1* and *Mt2* coding sequences, respectively, with the LacZ reporter enzyme without modifying the endogenous upstream and downstream regulatory elements [Klosen et al., 2019]. Expression during the light phase, when melatonin binding and expression of melatonin receptor mRNA are the highest, was measured in mouse brain sections by both, LacZ histoenzymological staining of beta-galactosidase activity and immunohistochemistry using a chicken anti-LacZ antibody. This strategy showed striking differences in the distribution of Mt1 and Mt2 receptors. Histoenzymological staining of Mt1-LacZ yielded high density labeling in very few structures, notably the SCN, the *pars tuberalis*, as well as the paratenial and paraventricular nuclei of the thalamus, which perfectly coincided with the distribution of [125I]iodomelatonin binding, which exhibits a higher affinity for Mt1 than Mt2. However, the findings are completely at odds with RT-PCR studies, which have detected expression of *Mt1* mRNA in the caudate-putamen, substantia nigra, ventral tegmental area, olfactory tubule, and hippocampus of the mouse brain [Uz et al., 2005, Jenwitheesuk et al., 2017], which may reflect differences in detection sensitivity or indicate that the presence of mRNA may not necessarily represent protein expression, as some mRNAs are translationally controlled. Expression of Mt2, on the other hand, was far more widespread and particularly dense in the forebrain, the paraventricular nucleus of the hypothalamus, and the amygdaloid complex. Co-expression of both melatonin receptors was rare, with the exception of the SCN that showed high and medium density staining for Mt1 and Mt2, respectively. Regardless of the published ambiguity over the differential expression of Mt1 and Mt2, no study has ever reported the extent of co-expression of the two melatonin receptors in adult mouse brain as displayed in the mouse organogenesis spatiotemporal transcriptomic atlas (Suppl. Figure S18), raising concerns over the ability of the Stereo-seq chip to capture selectively *Mt1* and *Mt2* mRNA. Thus, the strong spatial association computed by SpaCeNet based on the MOSTA data, which implies that high expression of *Mt1* is associated with high expression of *Mt2* in neighboring cells, is not supported by the extensive body of literature available on melatonin receptor localization and function. Expression of myelin-associated oligodendrocytic basic protein (Mobp) and myelin basic protein (Mbp), which yielded the second highest positive spatial assoiation, is in contrast to the neuro-specific expression of Mt1 and Mt2 specific to oligodendrocytes. They are major components of the myelin sheath, which is formed from cellular processes that extend from oligodendrocytes and wrap in a spiral fashion multiple times around the nerve axons. The major function of Mbp is intracellular compaction of the myelin, whereby it binds to the two opposing cytoplasmic leaflets of the myelin bilayer and brings them closely together via its polymerization into a fibrous network, giving rise to the so-called major dense line in electron microscope images [Bakhti et al., 2014]. Mobp, on the other hand, reinforces apposition of the compacted cytoplasmic leaflets and contributes to the normal arrangement of the so-called radial component, which ensures extracellular apposition between myelin membranes [Yamamoto et al., 1999]. A single oligodendrocyte myelinates between 40 and 60 different axons but only one segment per axon, which ranges from 20 - 200 µm in length in the CNS [Stadelmann et al., 2019]. Thus, each axon in the CNS is myelinated by multiple oligodendrocytes. Both Mobp and Mbp are highly basic polypeptides. To minimize interference with other proteins, the mRNAs of Mbp and Mobp are packed into granules in the perikaryon and transported along the peripheral processes to the myelin compartments for localized protein synthesis [Holz et al., 1996, Ainger et al., 1997]. Thus, given the branching structure of oligodendrocytes, which support numerous myelin ensheathments [Osanai et al., 2022], Mbp and Mobp mRNAs localizing to the various myelin compartments of an oligodendrocyte over an extended area may be interpreted as belonging to different cells, *i.e.* axons, using an array-based method for spatial transcriptomics. This may explain the positive spatial association observed for these two genes, which may have been further confounded by the low sequencing depth (Suppl. Figure S17). Finally, we observed a strong negative spatial association between the genes *Ptgds* and *Apoe*, which encode prostaglandin-H2 D-isomerase and apolipoprotein E, respectively. The former is expressed in oligodendrocytes and the leptomeninges [Urade et al., 1993] and catalyzes the conversion of prostaglandin H2 (PGH2) to prostaglandin D2 (PGD2), which is unstable and readily undergoes non-enzymatic dehydration to yield biologically active prostaglandins of the J series, including PGJ2, delta12,14-PGJ2, and 15-deoxy-delta12,14-PGJ2 (15d-PGJ2). The latter is a potent agonist of the nuclear transcription factor peroxisome proliferator–activated receptor-gamma (PPARgamma) [Waku et al., 2009] that has been shown to decrease transcription of Apoe, whereas PPARgamma knock-down results in a strong increase in expression of Apoe-mRNA [Subramanian et al., 2017]. Apolipoprotein E is mainly produced by astrocytes to facilitate transport of lipids to other brain cells such as neurons [Konings et al., 2021]. In summary, even in scenarios where SpaCeNet's model assumptions are challenged, it may yield promising findings for further analyses. However, spatial transcriptomic data should be subjected to careful scrutiny before drawing any conclusions.

4 Tables and Figures

Table S1: SpaCeNet results of simulation studies.

				AUROC $\boldsymbol{\Omega}$		AUPRC $\boldsymbol{\Omega}$		AUROC $\boldsymbol{\Delta\rho}$		AUPRC $\boldsymbol{\Delta\rho}$	
				mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\phi_{\min}$	$\phi_{\max}$	$n \cdot S$	S								
5.0	5.0	1000	1	0.89	0.04	0.85	0.04	0.55	0.07	0.28	0.21
			10	0.91	0.04	0.88	0.04	0.63	0.11	0.27	0.17
			100	0.96	0.02	0.94	0.02	0.93	0.07	0.73	0.12
		10000	10	0.95	0.02	0.94	0.02	0.87	0.10	0.64	0.16
			100	0.97	0.02	0.96	0.02	0.94	0.06	0.83	0.11
		100000	100	0.98	0.02	0.98	0.02	0.98	0.03	0.92	0.12
	20.0	1000	1	0.96	0.02	0.94	0.03	0.59	0.10	0.25	0.17
			10	0.97	0.02	0.95	0.02	0.62	0.09	0.26	0.17
			100	0.97	0.01	0.96	0.02	0.78	0.10	0.40	0.12
		10000	10	0.98	0.02	0.98	0.02	0.84	0.13	0.64	0.15
			100	0.99	0.01	0.98	0.01	0.90	0.06	0.75	0.14
		100000	100	0.99	0.01	0.99	0.01	0.96	0.06	0.88	0.13
10.0	10.0	1000	1	0.95	0.02	0.93	0.02	0.62	0.10	0.24	0.15
			10	0.96	0.02	0.94	0.03	0.63	0.09	0.25	0.16
			100	0.97	0.01	0.96	0.02	0.84	0.12	0.48	0.13
		10000	10	0.98	0.01	0.98	0.02	0.88	0.08	0.72	0.13
			100	0.99	0.01	0.98	0.01	0.93	0.07	0.81	0.12
		100000	100	0.99	0.01	0.99	0.01	0.98	0.02	0.92	0.11
20.0	20.0	1000	1	0.98	0.02	0.97	0.02	0.57	0.08	0.20	0.17
			10	0.98	0.02	0.97	0.02	0.55	0.06	0.22	0.20
			100	0.98	0.01	0.97	0.02	0.67	0.10	0.26	0.13
		10000	10	0.99	0.01	0.99	0.01	0.76	0.15	0.49	0.16
			100	1.00	0.01	0.99	0.01	0.88	0.09	0.65	0.14
		100000	100	1.00	0.01	1.00	0.01	0.94	0.07	0.84	0.14

Table S2: Spatial edge recovery in terms of AUPRC for the correlation baseline for a varying number of nearest neighbours considered for the spatial environment.

				1knn $\boldsymbol{\Delta\rho}$		5knn $\boldsymbol{\Delta\rho}$		10knn $\boldsymbol{\Delta\rho}$		20knn $\boldsymbol{\Delta\rho}$		50knn $\boldsymbol{\Delta\rho}$	
				mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
ϕ_{min}	ϕ_{max}	$n \cdot S$	S										
5.0	5.0	1000	1	0.07	0.05	0.08	0.06	0.08	0.05	0.07	0.03	0.07	0.04
			10	0.06	0.03	0.07	0.05	0.08	0.05	0.07	0.04	0.07	0.04
			100	0.18	0.10	0.33	0.14	0.33	0.14	0.33	0.14	0.33	0.14
		10000	10	0.10	0.07	0.16	0.09	0.24	0.12	0.23	0.14	0.25	0.12
			100	0.22	0.12	0.48	0.17	0.55	0.20	0.57	0.20	0.43	0.15
		100000	100	0.62	0.16	0.81	0.10	0.85	0.11	0.88	0.10	0.85	0.10
	20.0	1000	1	0.12	0.08	0.14	0.08	0.10	0.05	0.08	0.04	0.07	0.04
			10	0.07	0.04	0.08	0.06	0.09	0.05	0.07	0.03	0.06	0.03
			100	0.15	0.11	0.11	0.07	0.11	0.07	0.11	0.07	0.11	0.07
		10000	10	0.39	0.19	0.39	0.15	0.39	0.19	0.25	0.16	0.18	0.11
			100	0.42	0.19	0.53	0.19	0.48	0.19	0.39	0.16	0.18	0.10
		100000	100	0.84	0.13	0.84	0.10	0.80	0.12	0.76	0.12	0.64	0.12
10.0	10.0	1000	1	0.13	0.08	0.16	0.09	0.11	0.05	0.10	0.04	0.07	0.04
			10	0.07	0.04	0.09	0.06	0.09	0.05	0.07	0.04	0.06	0.03
			100	0.16	0.12	0.12	0.07	0.12	0.07	0.12	0.07	0.12	0.07
		10000	10	0.41	0.19	0.48	0.16	0.48	0.20	0.33	0.17	0.24	0.13
			100	0.46	0.18	0.60	0.20	0.55	0.20	0.48	0.20	0.21	0.11
		100000	100	0.88	0.11	0.89	0.09	0.89	0.10	0.88	0.10	0.81	0.11
20.0	20.0	1000	1	0.11	0.07	0.09	0.06	0.07	0.05	0.06	0.03	0.06	0.03
			10	0.07	0.04	0.07	0.05	0.07	0.04	0.06	0.03	0.06	0.03
			100	0.11	0.09	0.06	0.04	0.06	0.04	0.06	0.04	0.06	0.04
		10000	10	0.38	0.18	0.21	0.13	0.16	0.10	0.10	0.11	0.08	0.05
			100	0.37	0.17	0.29	0.16	0.22	0.15	0.14	0.11	0.06	0.03
		100000	100	0.81	0.12	0.78	0.13	0.69	0.17	0.55	0.17	0.31	0.18

Table S3: Spatial edge recovery in terms of AUPRC for the partial correlation baseline for a varying number of nearest neighbours considered for the spatial environment.

				1knn $\boldsymbol{\Delta\rho}$		5knn $\boldsymbol{\Delta\rho}$		10knn $\boldsymbol{\Delta\rho}$		20knn $\boldsymbol{\Delta\rho}$		50knn $\boldsymbol{\Delta\rho}$	
				mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\phi_{\min}$	$\phi_{\max}$	$n \cdot S =$	S										
5.0	5.0	1000	1	0.07	0.05	0.08	0.05	0.08	0.05	0.08	0.05	0.07	0.04
			10	0.06	0.04	0.07	0.04	0.08	0.04	0.07	0.03	0.06	0.04
			100	0.18	0.13	0.32	0.17	0.32	0.17	0.32	0.17	0.32	0.17
		10000	10	0.11	0.09	0.18	0.11	0.25	0.12	0.23	0.13	0.24	0.11
			100	0.20	0.11	0.46	0.17	0.53	0.20	0.53	0.18	0.42	0.15
		100000	100	0.59	0.16	0.75	0.12	0.79	0.13	0.83	0.11	0.82	0.10
	20.0	1000	1	0.11	0.08	0.13	0.08	0.11	0.06	0.08	0.05	0.08	0.06
			10	0.07	0.05	0.08	0.05	0.08	0.05	0.06	0.03	0.06	0.04
			100	0.14	0.11	0.11	0.08	0.11	0.08	0.11	0.08	0.11	0.08
		10000	10	0.40	0.16	0.37	0.16	0.37	0.16	0.24	0.13	0.16	0.09
			100	0.40	0.18	0.51	0.18	0.47	0.18	0.35	0.15	0.16	0.10
		100000	100	0.78	0.15	0.77	0.13	0.75	0.15	0.73	0.16	0.64	0.19
10.0	10.0	1000	1	0.12	0.07	0.15	0.09	0.13	0.07	0.10	0.06	0.07	0.04
			10	0.07	0.05	0.09	0.06	0.09	0.05	0.07	0.04	0.06	0.04
			100	0.15	0.13	0.13	0.09	0.13	0.09	0.13	0.09	0.13	0.09
		10000	10	0.41	0.16	0.46	0.16	0.46	0.16	0.33	0.15	0.21	0.11
			100	0.43	0.17	0.57	0.18	0.54	0.20	0.45	0.18	0.19	0.12
		100000	100	0.81	0.13	0.82	0.11	0.82	0.12	0.83	0.11	0.79	0.11
20.0	20.0	1000	1	0.11	0.08	0.08	0.05	0.06	0.04	0.07	0.05	0.07	0.04
			10	0.07	0.05	0.06	0.04	0.06	0.03	0.06	0.03	0.06	0.04
			100	0.11	0.10	0.05	0.03	0.05	0.03	0.05	0.03	0.05	0.03
		10000	10	0.36	0.16	0.20	0.14	0.15	0.10	0.08	0.07	0.07	0.04
			100	0.35	0.16	0.29	0.14	0.21	0.11	0.11	0.06	0.07	0.04
		100000	100	0.75	0.13	0.71	0.14	0.64	0.17	0.54	0.17	0.31	0.19

Table S4: Performance in terms of AUROC to recover intra- (left) and intercellular (right) interactions from *in silico* tissues generated via mechanistic modeling for MISTy and SpaCeNet.

AUROC (intracellular)	MISTy	SpaCeNet
Tissue 1	0.66	0.81
Tissue 2	0.66	0.81
Tissue 1+2	0.66	0.80

AUROC (intercellular)	MISTy	SpaCeNet
Tissue 1	0.60	0.62
Tissue 2	0.62	0.64
Tissue 1+2	0.60	0.65

Table S5: Top 10 spatial interactions discovered by SpaCeNet on the StarMap data on natural scale (left), on log scale (middle) and for two copies of gene *Mbp* denoted as _Mbp*_ and _Mbp**_ on a natural scale (right).

Recovered edges (natural scale)			Recovered edges (log scale)			Artificial _Mbp*_ colinear to _Mbp**_ (natural scale)		
Gene 1	Gene 2	**$\Delta\rho$**	Gene 1	Gene 2	**$\Delta\rho$**			
Mbp	*Flt1*	0.327	*Sst*	*Pvalb*	0.178	_Mbp*_	_Mbp**_	-0.173
Ctgf	*Gja1*	0.102	*Mbp*	*Flt1*	0.150	_Mbp*_	*Flt1*	0.165
Ctgf	*Pcp4*	0.083	*Ctgf*	*Pcp4*	0.115	*Flt1*	_Mbp**_	0.165
Sst	*Reln*	0.077	*Ctgf*	*Gja1*	0.114	*Ctgf*	*Gja1*	0.110
Slc17a7	*Egr1*	0.075	*Npy*	*Sst*	0.111	*Ctgf*	*Pcp4*	0.083
Cux2	*Gja1*	0.060	*Sst*	*Vip*	0.107	*Sst*	*Reln*	0.078
Cux2	*Egr1*	0.055	*Cux2*	*Pcp4*	-0.104	*Slc17a7*	*Egr1*	0.077
Npy	*Vip*	0.054	*Cux2*	*Plcxd2*	0.098	*Cux2*	*Gja1*	0.062
Egr1	*Egr2*	0.052	*Mgp*	*Cck*	0.098	*Cux2*	*Egr1*	0.057
Cux2	*Pcp4*	-0.048	*Slc17a7*	*Pcp4*	0.095	*Npy*	*Vip*	0.055

Table S6: Top 10 spatial interactions discovered by SpaCeNet on the Drosophila data obtained on natural scale (left) and on log-scale (right).

Natural scale			Log scale		
Gene 1	Gene 2	**$\Delta\rho$**	Gene 1	Gene 2	**$\Delta\rho$**
sna	*twi*	-0.0360	*sna*	*twi*	-0.0316
ems	*noc*	0.0301	*dan*	*danr*	-0.0316
Dfd	*lok*	0.0210	*ems*	*noc*	0.0312
Ance	*CG10479*	0.0205	*Ance*	*CG10479*	0.0215
dan	*danr*	-0.0193	*Dfd*	*lok*	0.0194
cnc	*erm*	0.0175	*apt*	*tll*	-0.0159
apt	*tll*	-0.0171	*cnc*	*erm*	0.0150
Blimp-1	*brk*	0.0156	*Blimp-1*	*brk*	0.0129
CG14427	*kni*	0.0122	*kni*	*Nek2*	0.0126
cnc	*kni*	0.0117	*cnc*	*kni*	0.0118

Table S7: Top 10 absolute highest values of $\Delta\rho$ for genes appearing in at least 30% of all cells (left) and genes appearing in 10% of cells (right).

	30%			10%	
Gene 1	Gene 2	$\boldsymbol{\Delta\rho}$	Gene 1	Gene 2	$\boldsymbol{\Delta\rho}$
Mt2	*Mt1*	−0.363	*Mobp*	*Mbp*	−0.448
Mobp	*Mbp*	−0.355	*Sst*	*Npy*	0.374
Ptgds	*Apoe*	0.319	*Fth1*	*Mbp*	−0.356
Il31ra	*Camk1d*	−0.287	*Mt2*	*Mt1*	−0.347
Bc1	*Hps5*	−0.287	*Il31ra*	*Camk1d*	−0.331
Gm42418	*Il31ra*	0.260	*Bc1*	*Hps5*	−0.318
Fth1	*Mbp*	−0.218	*Mbp*	*Plekhb1*	−0.308
Bc1	*Ppm1e*	−0.208	*Nefl*	*Nefm*	−0.296
Olfm1	*Camk2n1*	−0.206	*Nrsn1*	*Rgs4*	−0.257
Plp1	*Cd81*	−0.200	*Bc1*	*Ppm1e*	−0.249

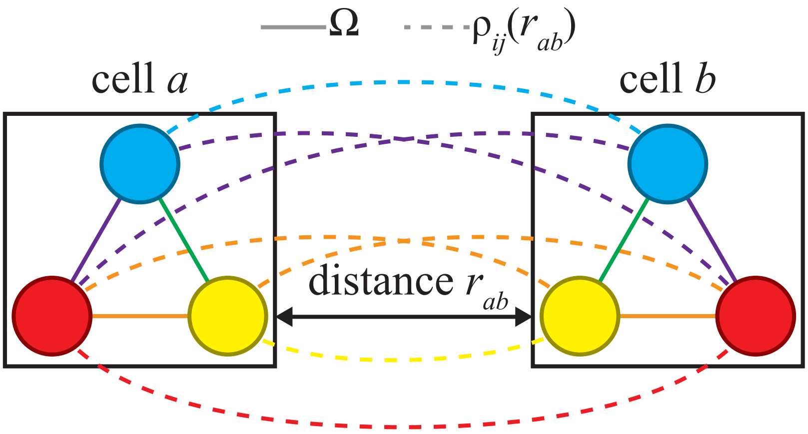

Figure S1: **Associations estimated by SpaCeNet:** SpaCeNet estimates intracellular and intercellular Spatial Conditional Independence (SCI) relationships reflected by network edges. Two cells a (left) and b (right) are shown. The circles of the same color represent the same molecular variable in each of the two cells. The solid lines represent edges of intracellular networks, indicating a direct association between molecular variables observed within the same cell. The dashed lines represent the intercellular network edges between cells a and b, indicating a direct spatial association between respective molecular variables. The potentials $\rho_{ij}(r_{ab})$ parameterize the intercellular interaction strength between variables i and j between cell a and cell b at distance r_{ab}. For instance, the missing dashed green lines between the blue and yellow variable imply SCI between the two ($\rho_{ij}(r_{ab}) = 0$), while SCI does not apply between the yellow and red variable ($\rho_{ij}(r_{ab}) \neq 0$).

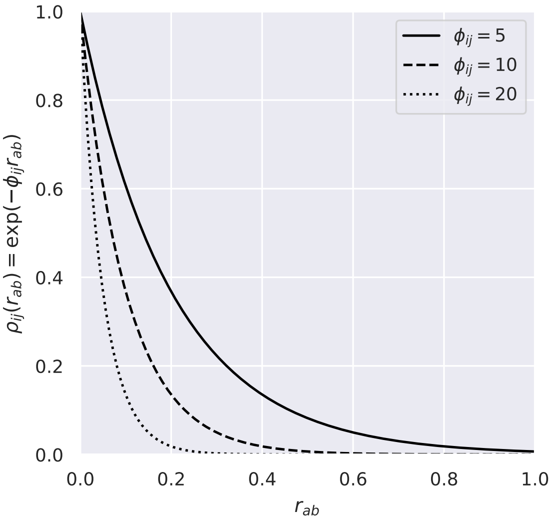

Figure S2: Illustration of different exponential potentials $\rho_{ij}(r_{ab}) = \exp(-\phi_{ij} r_{ab})$ used in the simulation studies, where $\phi_{ij} \sim \text{Unif}(5, 20)$ controls the interaction range. Small values correspond to long range associations ($\phi_{ij} = 5$, solid line), large values to short range associations ($\phi_{ij} = 20$, dotted line) and values in between to medium range associations ($\phi_{ij} = 10$, dashed line).

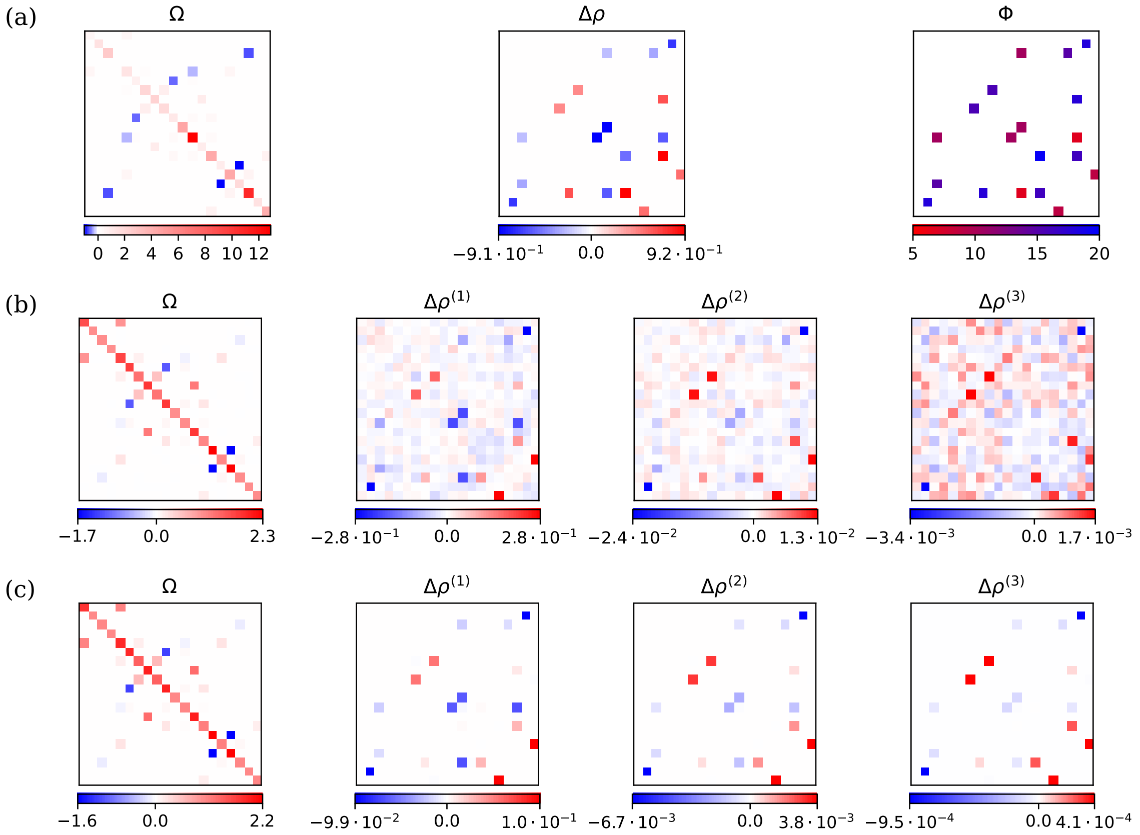

Figure S3: Parameters for a simulated data set with $n = 10^3,\ S = 100,\ \phi_{ij} \in [5, 20]$. (a) True parameters that have been used for sampling. (b) Estimated parameters, where the hyper-parameters were selected based on test-set loss, and (c) estimated parameters, where the hyper-parameters were manually chosen.

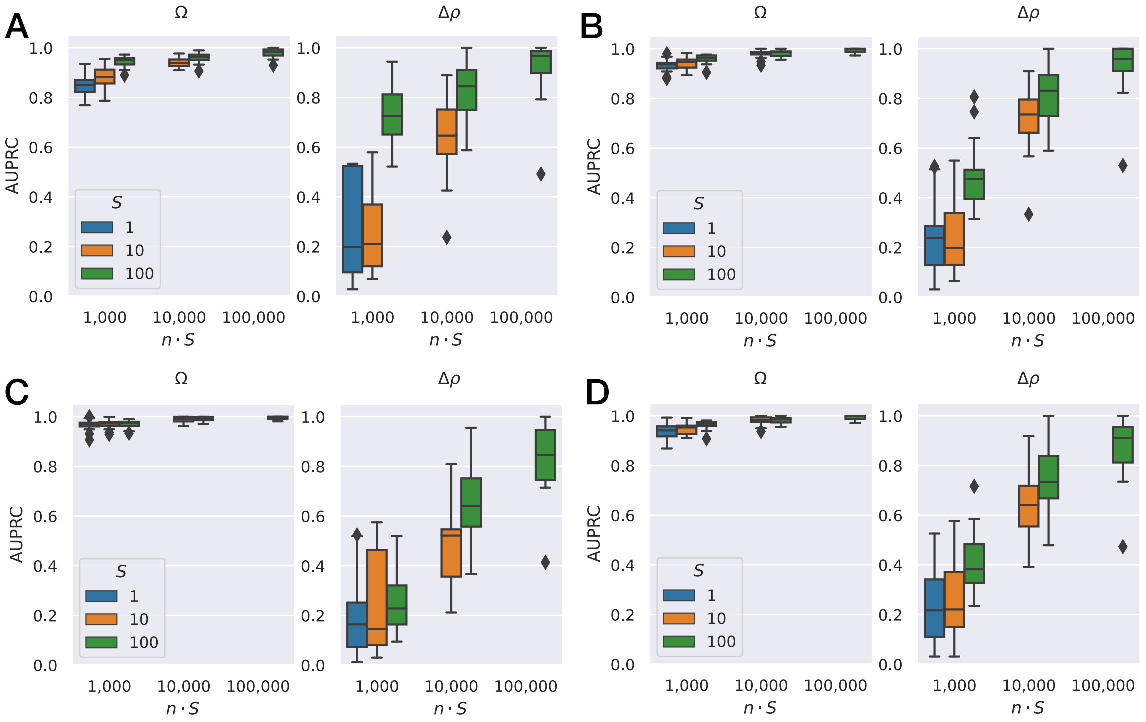

Figure S4: Edge recovery by SpaCeNet assessed in a simulation study using (A-C) fixed, radially decreasing potentials $\rho(r_{ab}) = \Delta\rho_{ij} \exp(-\phi_{ij} r_{ab})$ for all cell-cell interactions with constant range parameters $\phi_{ij} \in \{5, 10, 20\}$, respectively, and (D) flexible potentials $\rho_{ij}(r_{ab}) = \Delta\rho_{ij} \exp(-\phi_{ij} r_{ab})$ with $\phi_{ij} \sim \text{Unif}(5, 20)$ that mediate the interaction between molecular variables i and j. The y-axes give the performance in terms of the area under the precision recall curve (AUPRC). Left figures correspond to the inner-cellular networks (the intracellular precision matrix $\boldsymbol{\Omega}$) and the right figures to the extracellular networks (the cell-cell interaction parameters $\boldsymbol{\Delta\rho}$). The x-axis stratifies the analysis with respect to total cell numbers $n \cdot S$, where n is the number of cells within a measurement and S the number of measurements. Here, $S = 1$ is shown in blue, $S = 10$ in orange, and $S = 100$ in green.

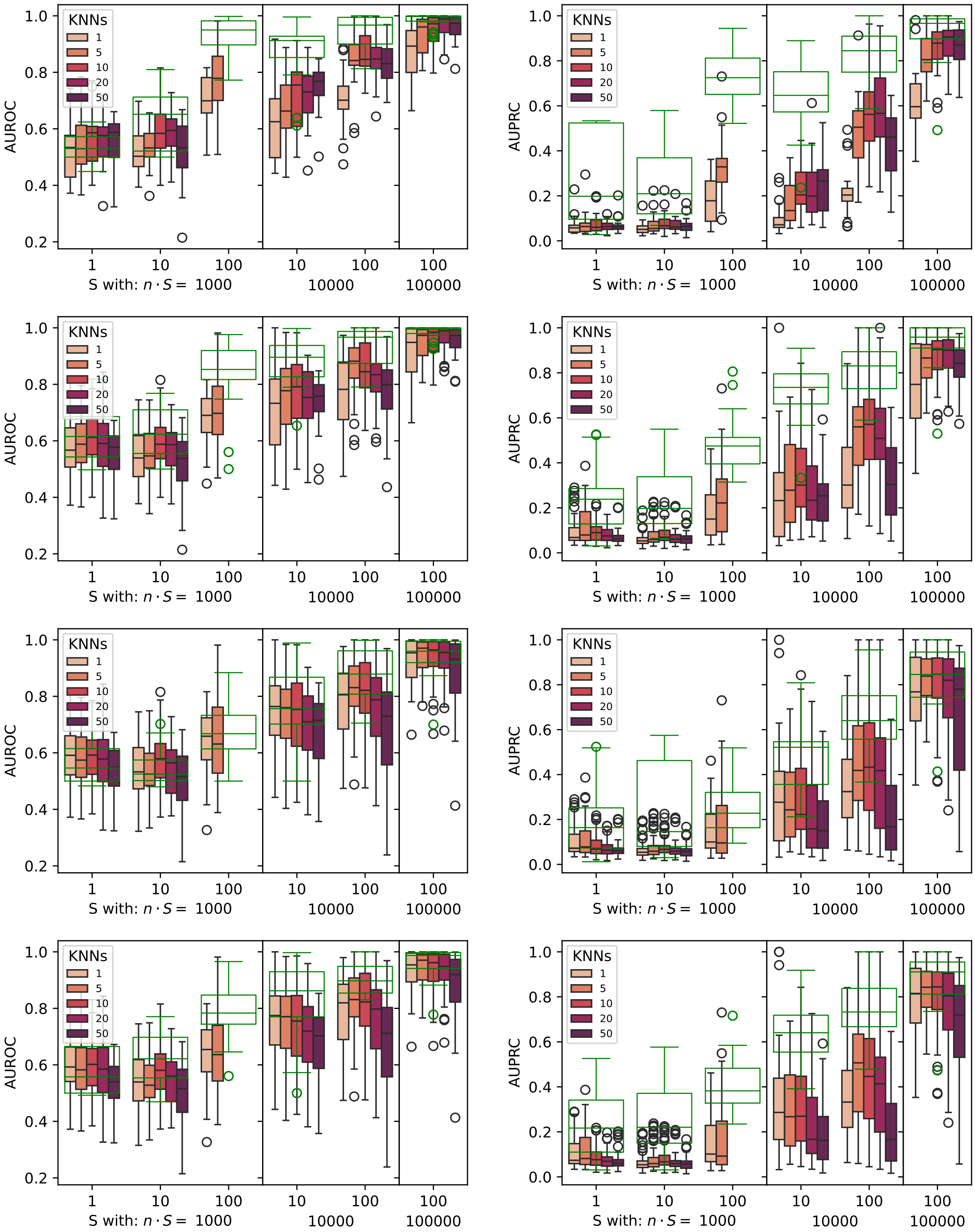

Figure S5: Edge recovery for intercellular associations $\Delta\rho$ of the correlation baseline considering a spatial environment given by the K nearest neighbours. Row 1 to 4 correspond to simulation settings A-D, respectively (see Fig. S4). Results from SpaCeNet are shown in green.

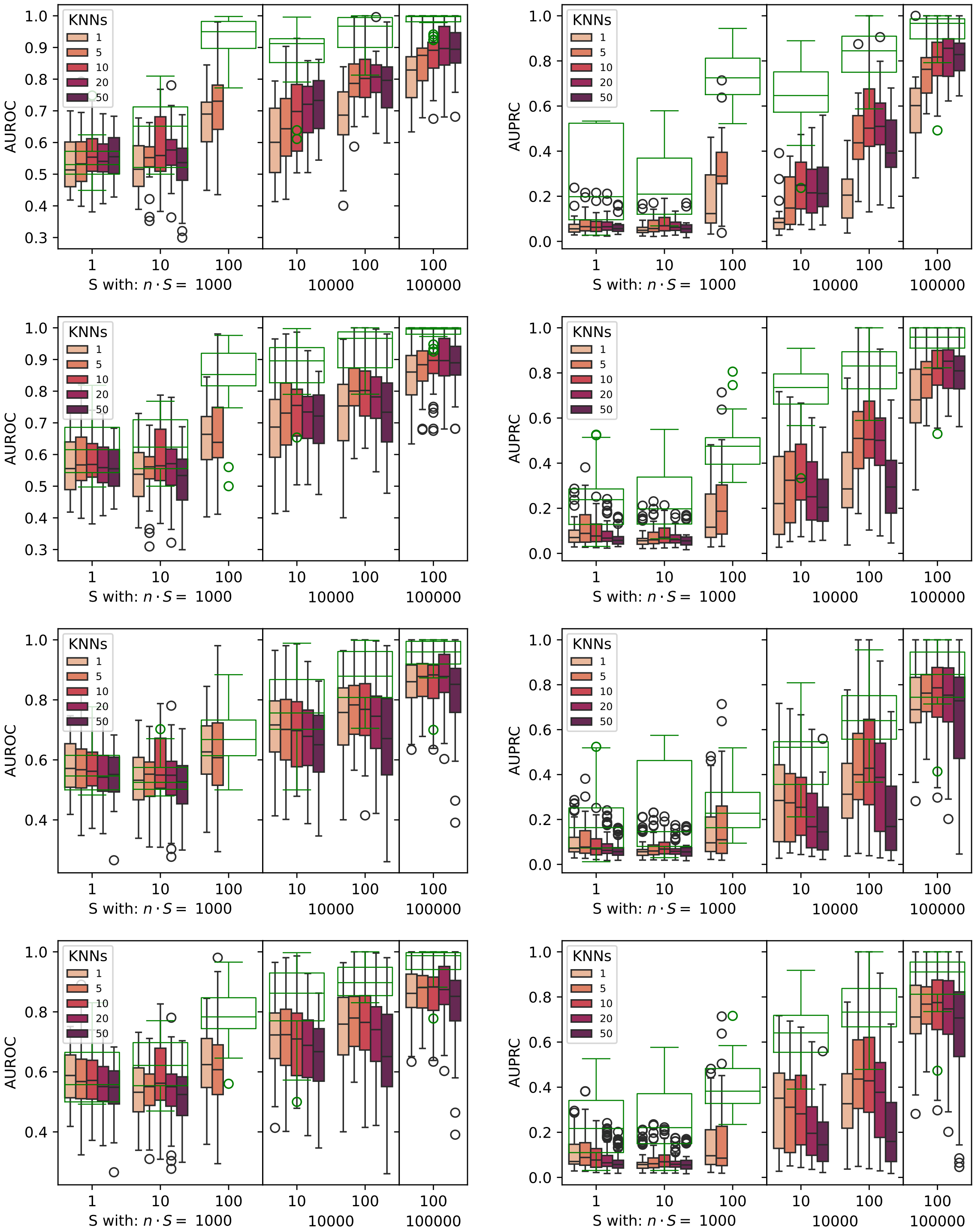

Figure S6: Edge recovery for intercellular associations $\Delta\rho$ of the partial correlation baseline considering a spatial environment given by the K nearest neighbours. Row 1 to 4 correspond to Simulation settings A-D, respectively (see Fig. S4). Results from SpaCeNet are shown in green.

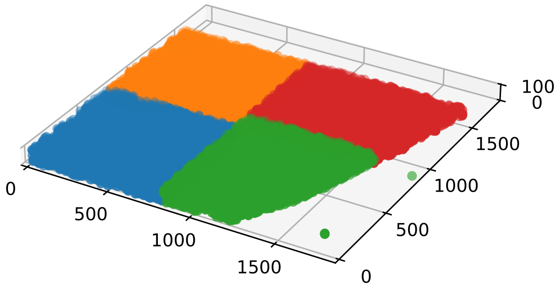

Figure S7: Cellular positions in the mouse visual cortex. Colors indicate the training (blue, orange, green) and validation (red) batches used for hyper-parameter screening and model development.

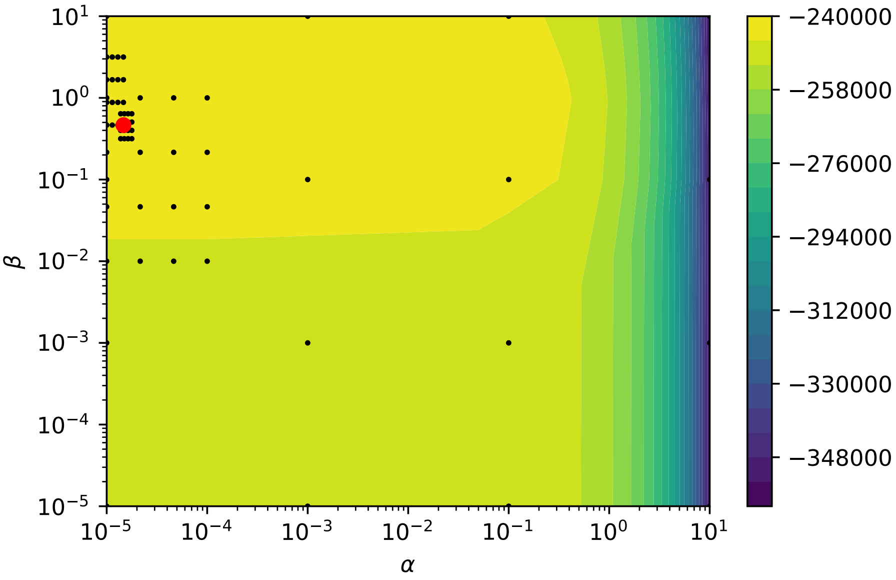

Figure S8: Evaluated hyper-parameter space and corresponding validation pseudo-log-likelihoods colored in blue (low values) to yellow (high values) based on the mouse visual cortex data provided by [Wang et al., 2018]. Black dots are tested hyper-parameters and the red dot corresponds to the optimal set of hyper-parameters in the grid search.

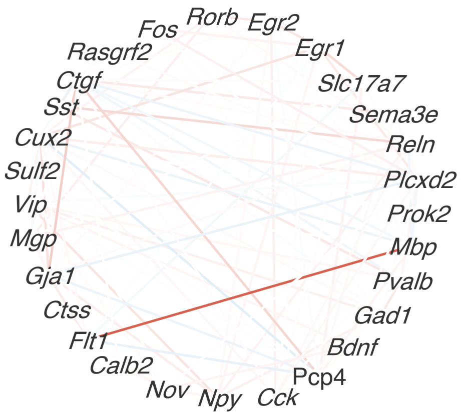

Figure S9: Network of spatial interactions between molecular variables estimated by SpaCeNet using data of the mouse visual cortex provided by [Wang et al., 2018]. Blue edges correspond to positive associations (negative entries of $\boldsymbol{\Delta\rho}^{(\cdot)}$) and red edges to negative associations (positive entries of $\boldsymbol{\Delta\rho}^{(\cdot)}$).

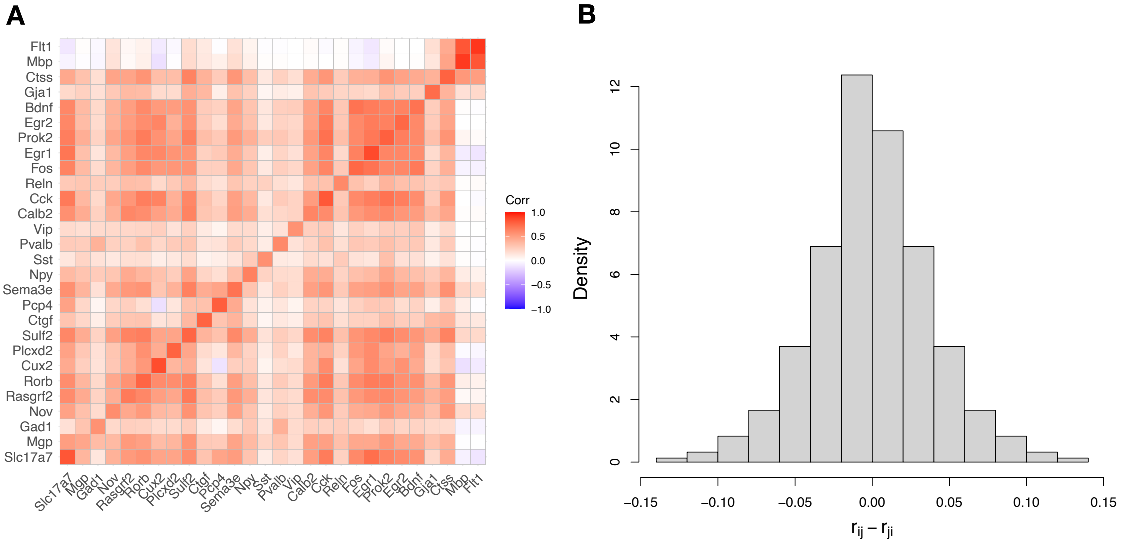

Figure S10: Estimated pair-wise spatial correlations (Figure A). The shown matrix is not symmetric by definition, although empirically approximately symmetric. This finding is supported by the histogram of differences $r_{ij} - r_{ji}$ (Figure B).

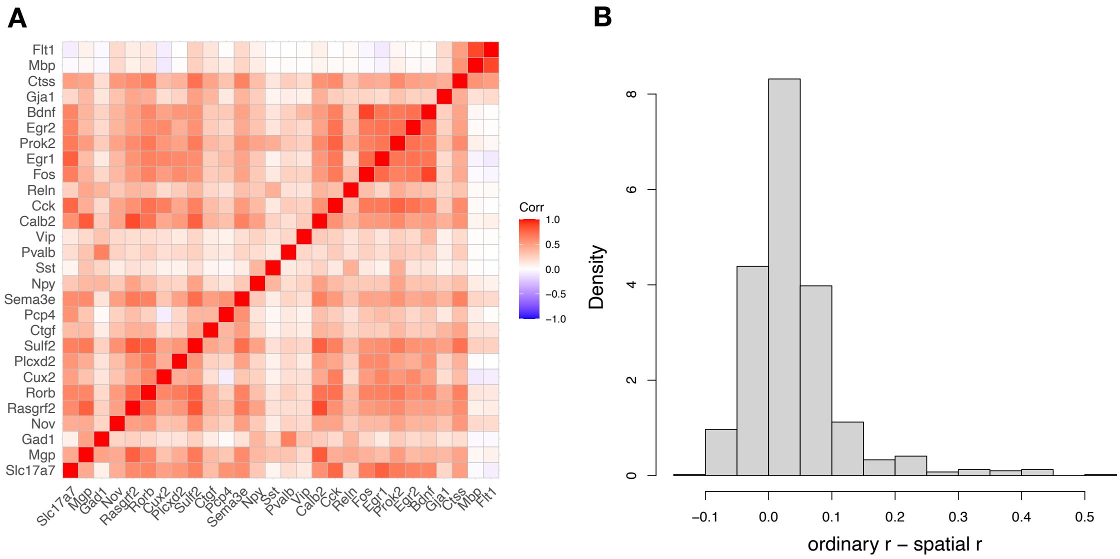

Figure S11: Ordinary correlations calculated between gene expression levels based on the individual cells' molecular profiles (Figure A). Figure B shows a histogram of the differences between the ordinary and spatial correlations.

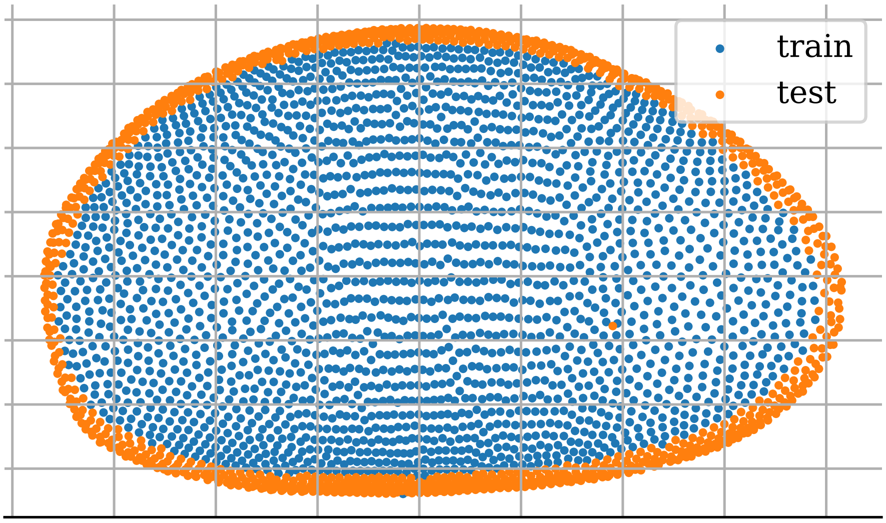

Figure S12: Spatial coordinates (mapped from 3D to 2D via a principal component analysis) of the virtual *Drosophila* embryo [Fowlkes et al., 2008] and the corresponding training/test splitting used for the SpaCeNet hyper-parameter screening shown in blue/orange, respectively.

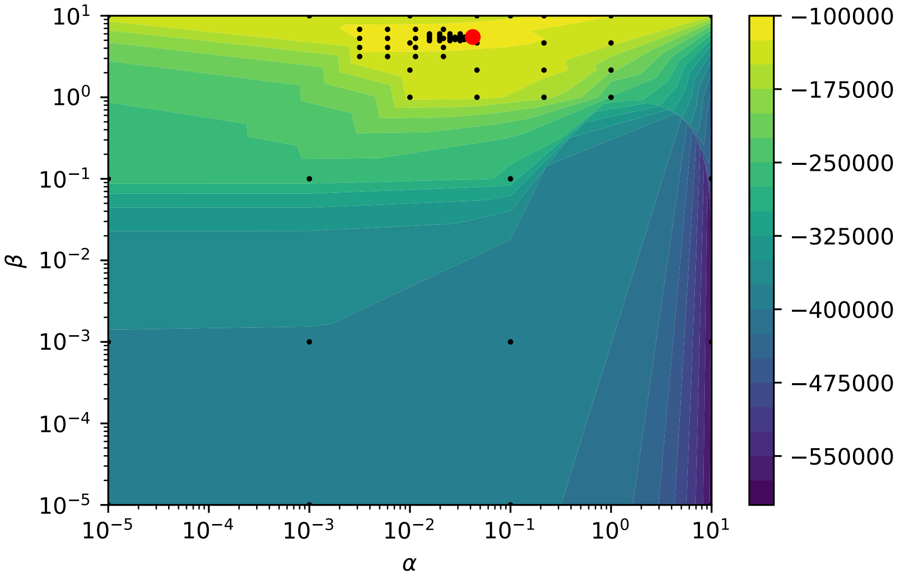

Figure S13: Evaluated hyper-parameter space and corresponding validation pseudo-log-likelihoods colored in blue (low values) to yellow (high values) based on the virtual *Drosophila* embryo data provided by [Fowlkes et al., 2008]. Black dots are tested hyper-parameters and the red dot corresponds to the optimal set of hyper-parameters in the grid search.

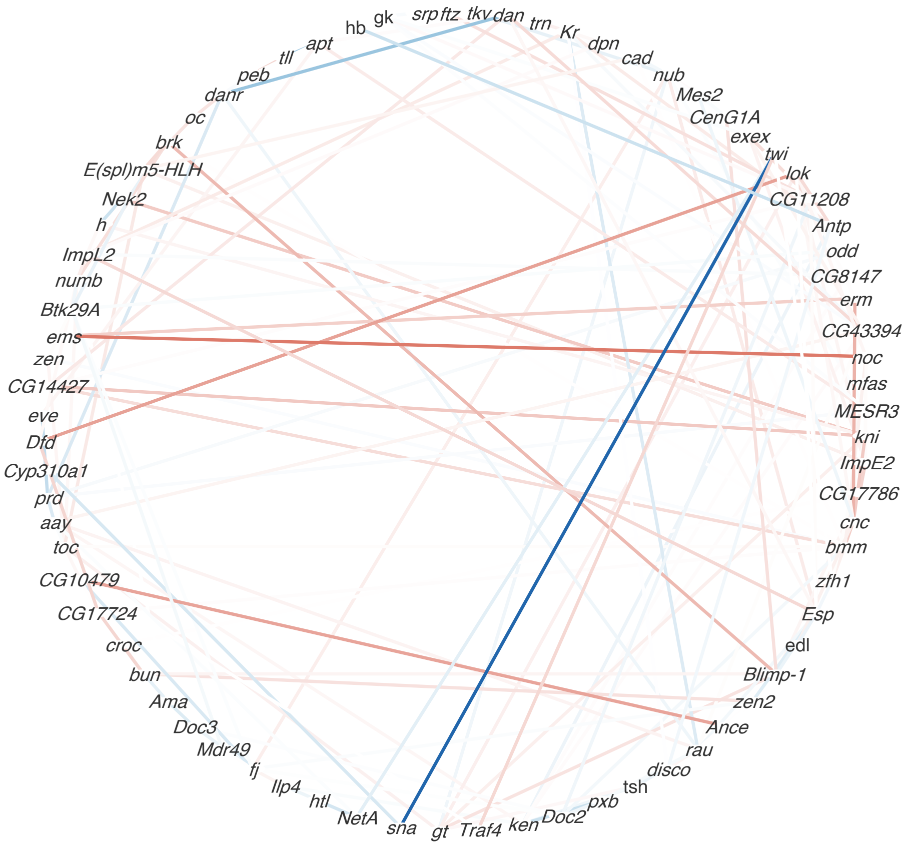

Figure S14: Network of spatial interactions between molecular variables estimated by SpaCeNet using data of the virtual *Drosophila* embryo provided by [Fowlkes et al., 2008]. Blue edges correspond to positive associations (negative entries of $\boldsymbol{\Delta\rho}^{(\cdot)}$) and red edges to negative associations (positive entries of $\boldsymbol{\Delta\rho}^{(\cdot)}$).

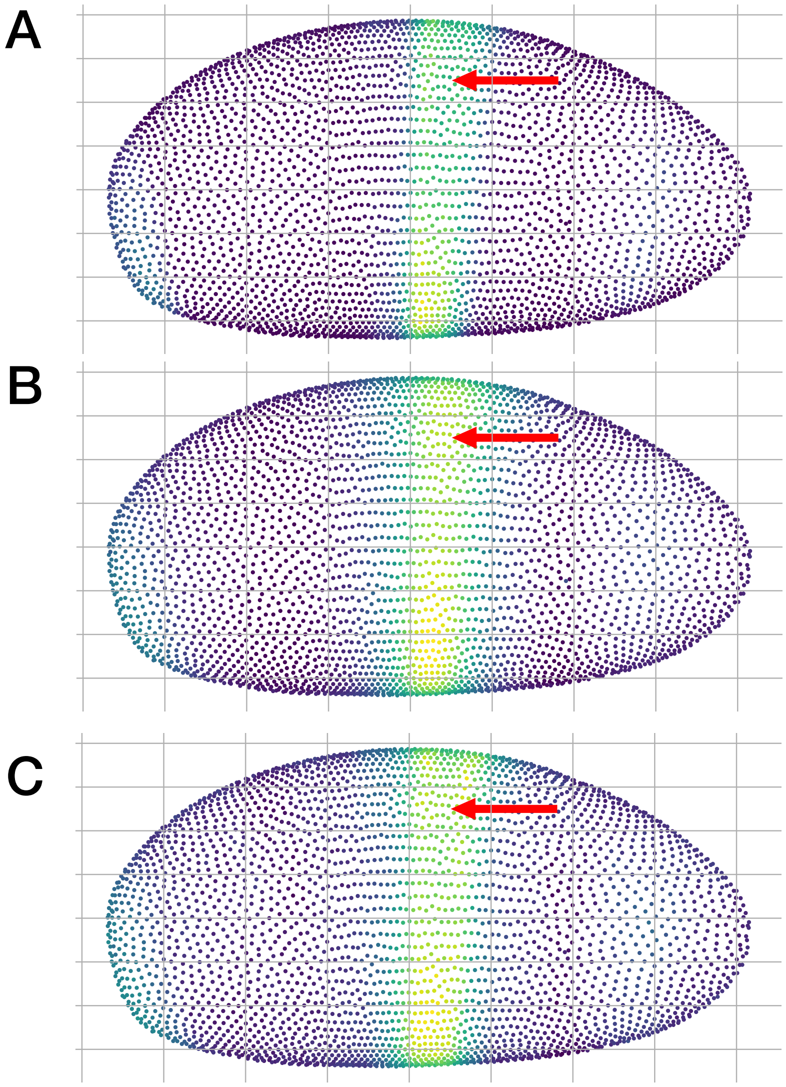

Figure S15: **SpaCeNet is an inferential tool which can predict gene expression from cellular context.** **A**, expression of the Krüppel protein gene *Kr* based on the *Drosophila* blastoderm data of [Fowlkes et al., 2008]. **B**, corresponding SpaCeNet prediction, where each cell's expression was predicted in a leave-one-cell-out approach (here, the density $f(\mathbf{x}^a|\mathbf{R}, \mathbf{X}^{\setminus a})$ was used). **C**, corresponding predictions based on cellular context and the expression of the remaining genes in the predicted cell (using density $f(x_j^a|\mathbf{R}, \mathbf{X}^{\setminus a}, \mathbf{x}^a_{\setminus j}))$, meaning that the expression $\mathbf{x}_j^a$ is predicted using the expression levels of all other cells $\mathbf{x}^b$ with $b \neq a$ and the expression levels of cell a except j. The red arrow highlights an area where the predictions of B and C differ most.

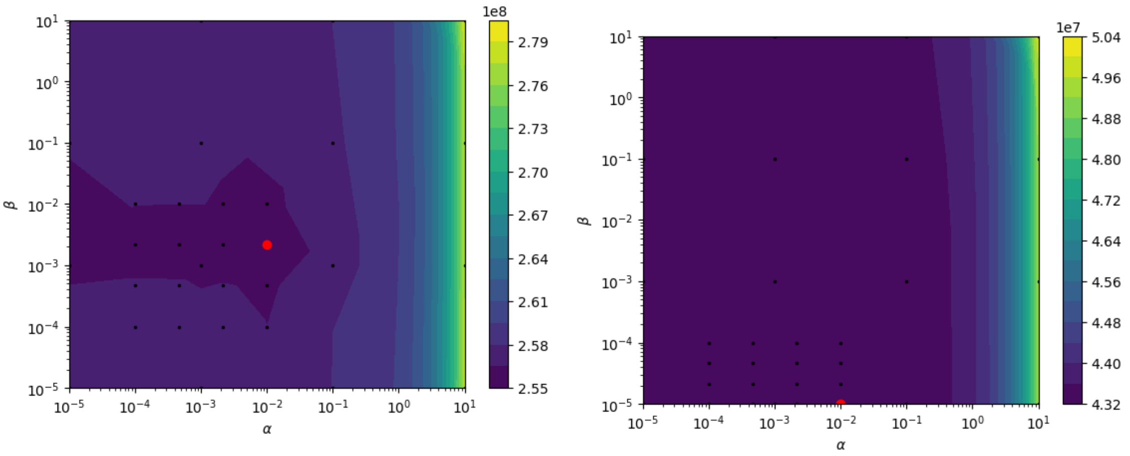

Figure S16: Evaluated hyper-parameter space and corresponding validation AIC colored in blue (low values) to yellow (high values) based on the MOSTA data, for genes appearing in at least 30% of all cells (left) and genes appearing in 10% of cells (right). Black dots are tested hyper-parameters and the red dot corresponds to the optimal set of hyper-parameters in the grid search.

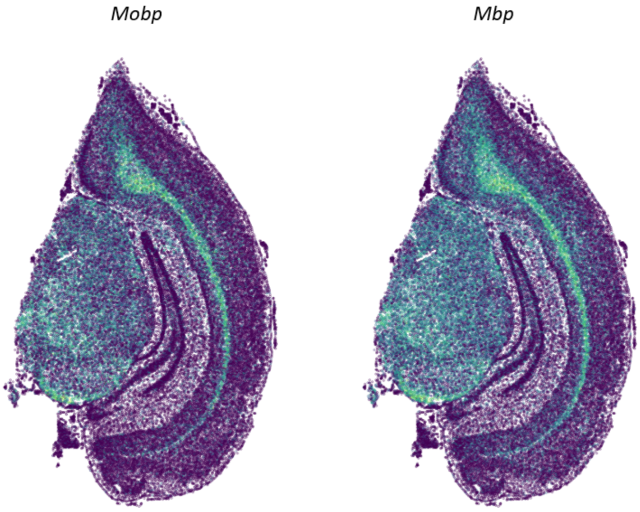

Figure S17: Scatter plot of the *Mobp* and *Mbp* genes in the MOSTA mouse adult brain colored in blue (low UMI count) to yellow (high UMI count).

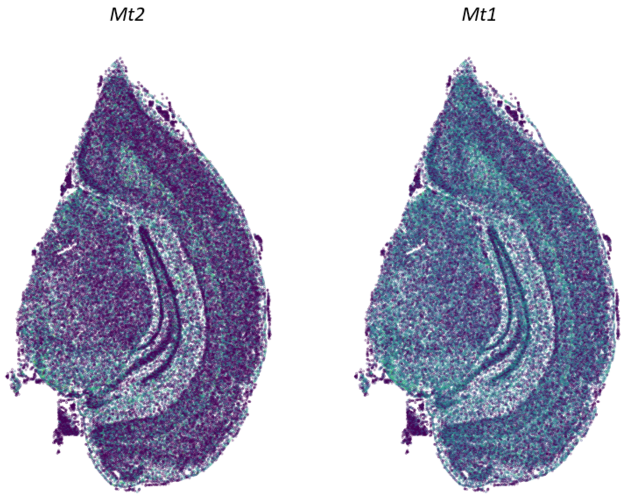

Figure S18: Scatter plot of the *Mt2* and *Mt1* genes in the MOSTA mouse adult brain colored in blue (low UMI count) to yellow (high UMI count).

References

Ao Chen, Sha Liao, Mengnan Cheng, Kailong Ma, Liang Wu, Yiwei Lai, Xiaojie Qiu, Jin Yang, Jiangshan Xu, Shijie Hao, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 185(10):1777–1792, 2022.

David M Berson, Felice A Dunn, and Motoharu Takao. Phototransduction by retinal ganglion cells that set the circadian clock. *Science*, 295(5557):1070–1073, 2002.

Margarita L Dubocovich. Melatonin receptors: role on sleep and circadian rhythm regulation. *Sleep medicine*, 8:34–42, 2007.

Seithikurippu R Pandi-Perumal, Ilya Trakht, Venkataramanujan Srinivasan, D Warren Spence, Georges JM Maestroni, Nava Zisapel, and Daniel P Cardinali. Physiological effects of melatonin: role of melatonin receptors and signal transduction pathways. *Progress in neurobiology*, 85 (3):335–353, 2008.

Margarita L Dubocovich, Philippe Delagrange, Diana N Krause, David Sugden, Daniel P Cardinali, and James Olcese. International Union of Basic and Clinical Pharmacology. LXXV. Nomenclature, classification, and pharmacology of G protein-coupled melatonin receptors. *Pharmacological reviews*, 62(3):343–380, 2010.

Moisés A Rivera-Bermúdez, Monica I Masana, Gregory M Brown, David J Earnest, and Margarita L Dubocovich. Immortalized cells from the rat suprachiasmatic nucleus express functional melatonin receptors. *Brain research*, 1002(1-2):21–27, 2004.

Nermien E Waly and Richard Hallworth. Circadian pattern of melatonin MT1 and MT2 receptor localization in the rat suprachiasmatic nucleus. *Journal of circadian rhythms*, 13, 2015.

Baptiste Lacoste, Debora Angeloni, Sergio Dominguez-Lopez, Sara Calderoni, Alessandro Mauro, Franco Fraschini, Laurent Descarries, and Gabriella Gobbi. Anatomical and cellular localization of melatonin MT1 and MT2 receptors in the adult rat brain. *Journal of Pineal Research*, 58(4): 397–417, 2015.

Paul Klosen, Sarawut Lapmanee, Carole Schuster, Beatrice Guardiola, David Hicks, Paul Pevet, and Marie Paule Felder-Schmittbuhl. MT1 and MT2 melatonin receptors are expressed in nonoverlapping neuronal populations. *Journal of pineal research*, 67(1):e12575, 2019.

Tolga Uz, Ahmet D Arslan, Murat Kurtuncu, Marta Imbesi, Mustafa Akhisaroglu, Yogesh Dwivedi, Ghanshyam N Pandey, and Hari Manev. The regional and cellular expression profile of the melatonin receptor MT1 in the central dopaminergic system. *Molecular brain research*, 136(1-2): 45–53, 2005.

Anorut Jenwitheesuk, Parichart Boontem, Prapimpun Wongchitrat, Jiraporn Tocharus, Sujira Mukda, and Piyarat Govitrapong. Melatonin regulates the aging mouse hippocampal homeostasis via the sirtuin1-FOXO1 pathway. *EXCLI journal*, 16:340, 2017.

Mostafa Bakhti, Shweta Aggarwal, and Mikael Simons. Myelin architecture: zippering membranes tightly together. *Cellular and molecular life sciences*, 71:1265–1277, 2014.

Yoichi Yamamoto, Hiroo Yoshikawa, Seiichi Nagano, Gen Kondoh, Shigeki Sadahiro, Takahiro Gotow, Takehiko Yanagihara, and Saburo Sakoda. Myelin-associated oligodendrocytic basic protein is essential for normal arrangement of the radial component in central nervous system myelin. *European Journal of Neuroscience*, 11(3):847–855, 1999.

Christine Stadelmann, Sebastian Timmler, Alonso Barrantes-Freer, and Mikael Simons. Myelin in the central nervous system: structure, function, and pathology. *Physiological reviews*, 99(3): 1381–1431, 2019.

Andreas Holz, Nicole Schaeren-Wiemers, Christoph Schaefer, Uwe Pott, Raymond J Colello, and Martin E Schwab. Molecular and developmental characterization of novel cDNAs of the myelin-associated/oligodendrocytic basic protein. *The Journal of neuroscience*, 16(2):467, 1996.

Kevin Ainger, Daniela Avossa, Amy S Diana, Christopher Barry, Elisa Barbarese, and John H Carson. Transport and localization elements in myelin basic protein mRNA. *The Journal of cell biology*, 138(5):1077–1087, 1997.

Yasuyuki Osanai, Reiji Yamazaki, Yoshiaki Shinohara, and Nobuhiko Ohno. Heterogeneity and regulation of oligodendrocyte morphology. *Frontiers in Cell and Developmental Biology*, 10:1030486, 2022.

Yoshihiro Urade, Kunio Kitahama, Hitoshi Ohishi, Takeshi Kaneko, Noboru Mizuno, and Osamu Hayaishi. Dominant expression of mRNA for prostaglandin D synthase in leptomeninges, choroid plexus, and oligodendrocytes of the adult rat brain. *Proceedings of the National Academy of Sciences*, 90(19):9070–9074, 1993.

Tsuyoshi Waku, Takuma Shiraki, Takuji Oyama, Yoshito Fujimoto, Kanako Maebara, Narutoshi Kamiya, Hisato Jingami, and Kosuke Morikawa. Structural insight into PPARγ activation through covalent modification with endogenous fatty acids. *Journal of molecular biology*, 385(1):188–199, 2009.

Shobana Subramanian, William K Gottschalk, So Young Kim, Allen D Roses, and Ornit Chiba-Falek. The effects of PPARγ on the regulation of the TOMM40-APOE-C1 genes cluster. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1863(3):810–816, 2017.

Sabine C Konings, Laura Torres-Garcia, Isak Martinsson, and Gunnar K Gouras. Astrocytic and neuronal apolipoprotein E isoforms differentially affect neuronal excitability. *Frontiers in Neuroscience*, 15:734001, 2021.

Xiao Wang, William E Allen, Matthew A Wright, Emily L Sylwestrak, Nikolay Samusik, Sam Vesuna, Kathryn Evans, Cindy Liu, Charu Ramakrishnan, Jia Liu, et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, 361(6400):eaat5691, 2018.

Charless C Fowlkes, Cris L Luengo Hendriks, Soile VE Keränen, Gunther H Weber, Oliver Rübel, Min-Yu Huang, Sohail Chatoor, Angela H DePace, Lisa Simirenko, Clara Henriquez, et al. A quantitative spatiotemporal atlas of gene expression in the drosophila blastoderm. *Cell*, 133(2): 364–374, 2008.