

¹ RESEARCH

² **Modeling the cell-type specific mesoscale murine connectome with
3 anterograde tracing experiments**

⁴ **Samson Koelle**^{1,2}, **Dana Mastrovito**¹, **Jennifer D Whitesell**¹, **Karla E Hirokawa**¹, **Hongkui Zeng**¹, **Marina Meila**², **Julie A
5 Harris**¹, **Stefan Mihalas**¹

⁶ ¹ Allen Institute for Brain Science, Seattle, WA, USA

⁷ ²Department of Statistics, University of Washington, Seattle, WA, USA

⁸ **Keywords:** [Connectivity, Cell-type, Mouse]

ABSTRACT

⁹ The Allen Mouse Brain Connectivity Atlas consists of anterograde tracing experiments targeting
¹⁰ diverse structures and classes of projecting neurons. Beyond regional anterograde tracing done in
¹¹ C57BL/6 wild type mice, a large fraction of experiments are performed using transgenic Cre-lines.
¹² This allows access to cell-class specific whole brain connectivity information, with class defined by
¹³ the transgenic lines. However, even though the number of experiments is large, it does not come close
¹⁴ to covering all existing cell classes in every area where they exist. Here, we study how much we can fill
¹⁵ in these gaps and estimate the cell-class specific connectivity function given the simplifying
¹⁶ assumptions that nearby voxels have smoothly varying projections, but that these projection tensors
¹⁷ can change sharply depending on the region and class of the projecting cells.

¹⁸ This paper describes the conversion of Cre-line tracer experiments into class-specific connectivity
¹⁹ matrices representing the connection strengths between source and target structures. We introduce
²⁰ and validate a novel statistical model for creation of connectivity matrices. We extend the

21 Nadaraya-Watson kernel learning method which we previously used to fill in spatial gaps to also fill in
22 a gaps in cell-class connectivity information. To do this, we construct a "cell-class space" based on
23 class-specific averaged regionalized projections and combine smoothing in 3D space as well as in this
24 abstract space to share information between similar neuron classes. Using this method we construct a
25 set of connectivity matrices using multiple levels of resolution at which discontinuities in connectivity
26 are assumed. We show that the connectivities obtained from this model display expected cell-type
27 and structure specific connectivities. We also show that the wild type connectivity matrix can be
28 factored using a sparse set of factors, and analyze the informativeness of this latent variable model.

AUTHOR SUMMARY

29 Large-scale studies have described the connections between areas in multiple mammalian models in
30 ever expanding detail. Standard connectivity studies focus on the connection strength between areas.
31 However, when describing functions at a local circuit level, there is an increasing focus on cell types.
32 We have recently described the importance of connection types in the cortico-thalamic system, which
33 allows an unsupervised discovery of its hierarchical organization. In this study we focus on adding a
34 dimension of connection type for a brain-wide mesoscopic connectivity model. Even with our
35 relatively massive dataset, the data in the cell type direction for connectivity is quite sparse, and we
36 had to develop methods to more reliably extrapolate in such directions, and to estimate when such
37 extrapolations are impossible. This allows us to fill in such a connection type specific inter-areal
38 connectivity matrix to the extent our data allows. While analyzing this complex connectivity, we
39 observed that it can be described via a small set of factors. While not complete, this connectivity
40 matrix represents a a categorical and quantitative improvement in mouse mesoscale connectivity
41 models.

1 INTRODUCTION

42 The mammalian nervous system enables an extraordinary range of natural behaviors, and has
43 inspired much of modern artificial intelligence. Neural connections from one region to another form
44 the architecture underlying this capability. These connectivities vary by neuron type, as well as source
45 cell body location and target axonal projection structures. Thus, characterization of the relationship
46 between neuron type and source and target structure is important for understanding the overall
47 nervous system.

48 Viral tracing experiments - in which a viral vector expressing GFP is transduced into neural cells
49 through stereotaxic injection - are a useful tool for mapping these connections on the mesoscale
50 (Chamberlin, Du, de Lacalle, & Saper, 1998; Daigle et al., 2018; J. A. Harris, Oh, & Zeng, 2012). The GFP
51 protein moves into the axon of the projecting neurons. The long range connections between different
52 areas are generally formed by axons which travel from one region to another. Two-photon
53 tomography imaging can be used to determine the location and strength of the fluorescent signals in
54 two-dimensional slices. These locations can then be mapped back into three-dimensional space, and
55 the signal may then be integrated over area into cubic voxels to give a finely-quantized
56 three-dimensional fluorescence.

57 Several statistical models for the conversion of such experiment-specific signals into generalized
58 estimates of connectivity strength have been proposed (K. D. Harris, Mihalas, & Shea-Brown, 2016;
59 Knox et al., 2019; Oh et al., 2014; ?). Of these, Oh et al. (2014) and Knox et al. (2019) model **regionalized**
60 **connectivities**, which are voxel connectivities integrated by region. The value of these models is that
61 they provide some improvement over simply averaging the projection signals of injections in a given
62 region. However, these previous works only model connectivities observed in wild type mice in which
63 all neuron types were labeled, and so are poorly suited for extension to tracing experiments that
64 induce cell-type specific fluorescence (J. A. Harris et al., 2019). In particular, GFP expression is
65 induced by Cre-recombinase in cell-types specified by the combination of transgenic mouse strain
66 and transgene promoter. We refer to such sets of cells as **cell class**. For example, use of both wild-type
67 and transgenic mice would give rise to cell-class specific experiments, but with different albeit
68 perhaps overlapping classes of cells.

69 Thus, this paper introduces a class-specific statistical model to deal with the diverse set of **Cre-lines**
70 described in J. A. Harris et al. (2019), and expands this model to the entire mouse brain. Our model is a
71 to-our-knowledge novel estimator that takes into account both the spatial position of the labelled
72 source, as well as the categorical cell class. Like the previously state-of-the-art model in Knox et al.
73 (2019), this model predicts regionalized connectivity as an average over positions within the structure,
74 with nearby experiments given more weight. However, our model weighs class-specific behavior in a
75 particular structure against spatial position, so a nearby experiment specific to a similar cell-class is
76 relatively up-weighted, while a nearby experiment specific to a dissimilar class is down-weighted.
77 This model outperforms the model of Knox et al. (2019) based on its ability to predict held-out
78 experiments in leave-one-out cross-validation. We then use the trained model to estimate overall
79 connectivity matrices for each assayed cell class.

80 The resulting cell-type specific connectivity is a directed weighted multigraph which can be
81 represented as a tensor with missing values. We do not give an exhaustive analysis of this data, but do
82 establish a lower-limit of detection, verify several cell-type specific connectivity patterns found
83 elsewhere in the literature, and show that these cell-type specific signals are behaving in expected
84 ways. We also decompose the wild type connectivity matrix into factors representing archetypal
85 connectivity patterns. These components allow approximation of the regionalized connectivity using
86 a small set of latent components.

87 Section 2 gives information on the data and statistical methodology, and Section 3 presents our
88 results. These include connectivities, assessments of model fit, and subsequent analyses. Additional
89 information on our dataset, methods, and results are given in Supplemental Sections 5, 6, and 7,
90 respectively.

2 METHODS

91 We estimate and analyze cell class-specific connectivity functions using models trained on murine
92 brain viral tracing experiments. This section describes the data used to generate the model, the model
93 itself, the evaluation of the model against its alternatives, and the use of the model in creation of the
94 connectivity estimate matrices. It also includes background on the non-negative matrix factorization
95 method used for decomposing the wild type connectivity matrix into latent factors. Additional
96 information about our data and methods are given in Supplemental Sections 5 and 6, respectively.

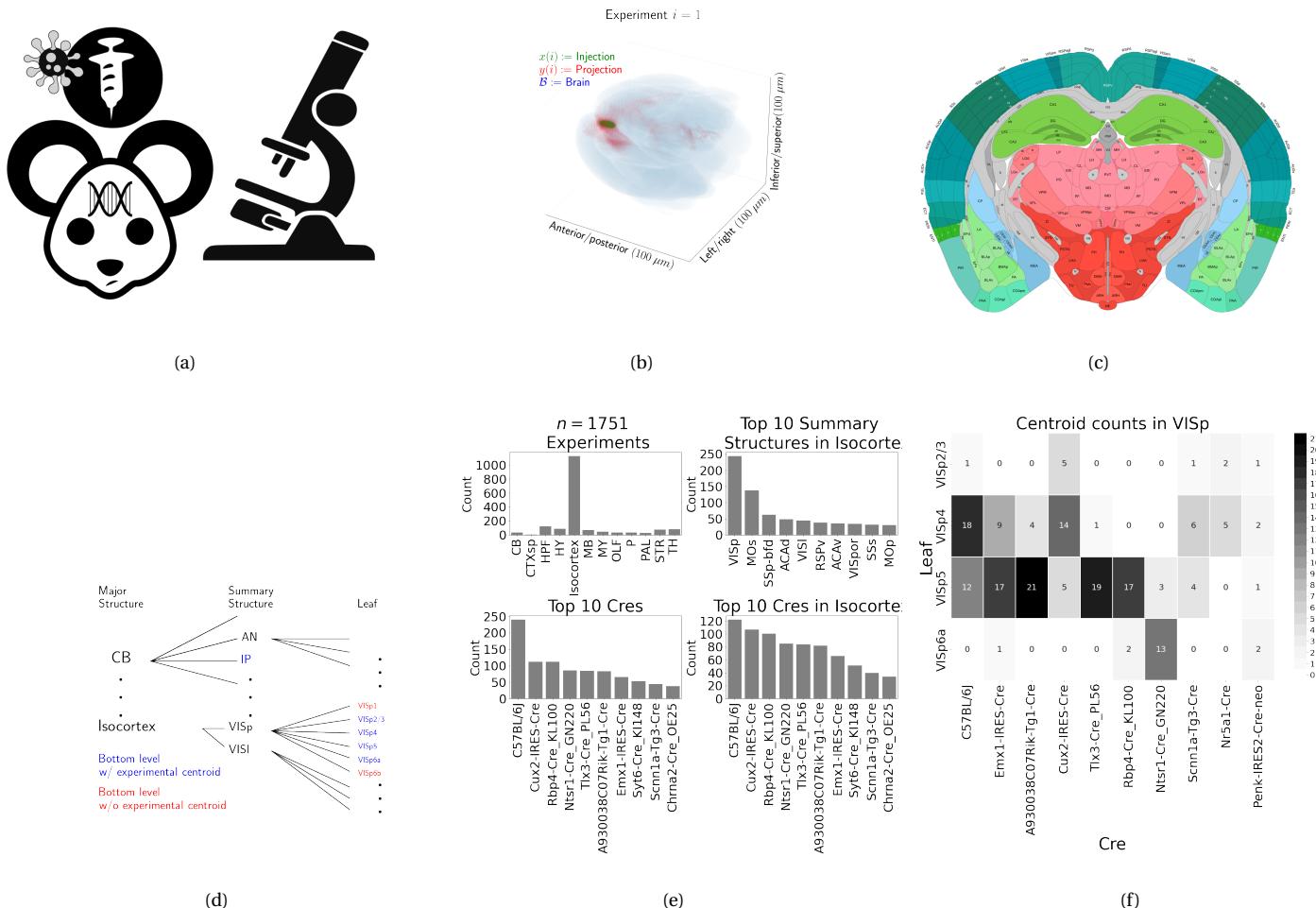


Figure 1: Experimental setting. 1a For each experiment, a Cre-dependent GFP-expressing transgene cassette is transduced by stereotaxic injection into a Cre-driver mouse, followed by serial two-photon tomography imaging. 1b An example of the segmentation of projection (targets) and injection (source) for a single experiment. Within each brain (blue), injection (green) and projection (red) areas are determined via histological analysis and alignment to the Allen Common Coordinate Framework (CCF). 1c Brain region parcellations within a coronal plane of CCFv3. 1d Explanation of nested structural ontology highlighting various levels of CCFv3 structure ontology. Lowest-level (leaf) structures are colored in blue, and structures without an injection centroid are colored in red. 1e Abundances of tracer experiments by Cre-line and region of injection. 1f Co-occurrence of layer-specific centroids and Cre-lines within VISp.

97 **Data**

98 Our dataset \mathcal{D} consists of $n = 1751$ publicly available murine brain viral tracing experiments from the
 99 Allen Mouse Brain Connectivity Atlas. Figure 1a summarizes the experimental process used to
 100 generate this data. In each experiment, a mouse is injected with an adeno-associated virus (AAV)
 101 encoding green fluorescent protein (GFP) into a single location in the brain. Location of fluorescence
 102 is mediated by the location of the injection, the characteristics of the transgene, and the genotype of
 103 the mouse. In particular, Cre-driver mice are engineered to express Cre under the control of a specific
 104 and single gene promoter. This localizes expression of Cre to regions with certain transcriptomic
 105 cell-types signatures. In such Cre-driver mice, we used a double-inverted floxed AAV to produce
 106 fluorescence that depends on Cre expression in infected cells. To account for the complex cell-type
 107 targeting induced by a particular combination of Cre-driver genotype and GFP promoter, we refer to
 108 the combinations of cell-types targeted by a particular combination of AAV and Cre-driver mice as
 109 cell-classes. For example, we include experiments from Cre-driver lines that selectively label cell
 110 classes located in distinct cortical layers or other nuclei across the whole brain. By our definition, wild
 111 type mice transduced with constitutively active GFP promoters induce fluorescence of a particularly
 112 broad cell class.

113 For each experiment, the fluorescent signal imaged after injection is aligned into the Allen
 114 Common Coordinate Framework (CCF) v3, a three-dimensional average template brain that is fully
 115 annotated with regional parcellations Wang et al. (2020). The whole brain imaging and registration
 116 procedures described in detail in Kuan et al. (2015); Oh et al. (2014) produce quantitative metrics of
 117 fluorescence discretized at the $100 \mu\text{m}$ voxel level. Given an experiment, this image was histologically
 118 segmented by an analyst into *injection* and *projection* areas corresponding to areas containing somas,
 119 dendrites and axons or exclusively axons of the transfected neurons. An example of a single
 120 experiment rendered in 3D is given in Figure 1b. Given an experiment i , we represent injections and
 121 projections as functions $x(i), y(i) : \mathcal{B} \rightarrow \mathbb{R}_{\geq 0}$, where $\mathcal{B} \subset [1 : 132] \times [1 : 80] \times [1 : 104]$ corresponds to the
 122 subset of the $(1.32 \times 0.8 \times 1.04)$ cm rectangular space occupied by the standard voxelized mouse brain.
 123 We also calculate injection centroids $c(i) \in \mathbb{R}^3$ and regionalized projections $y_{\mathcal{T}}(i) \in \mathbb{R}^T$ given by the
 124 sum of $y(i)$ in each region. A description of these steps is in Supplemental Section 6.

125 Our goal is the estimation of **regionalized connectivity** from one region to another. A visual
126 depiction of this region parcellation for a two-dimensional slice of the brain is given in Figure 1c. All
127 structures annotated in the CCF belong to a hierarchically ordered ontology, with different areas of the
128 brain are parcellated to differing finer depths within a hierarchical tree. We denote the main levels of
129 interest as major structures, summary structures, and layers. Not every summary structure has a layer
130 decomposition within this ontology, so we typically consider the finest possible regionalization - for
131 example, layer within the cortex, and summary structure within the thalamus, and denote these
132 structures as leafs. As indicated in Figure 1d, the dataset used to generate the connectivity model
133 reported in this paper contains certain combinations of region and cell class frequently, and others
134 not at all. A summary of the most frequently assayed cell classes and structures is given in Figures 1e
135 and 1f. Since users of the connectivity matrices may be interested in particular combinations, or
136 interested in the amount of data used to generate a particular connectivity estimate, we present this
137 information about all experiments in Supplemental Section 5.

138 ***Modeling Regionalized Connectivity***

We define voxelized cell-class specific connectivity $f : \mathcal{V} \times \mathcal{B} \times \mathbb{R}_{\geq 0}$ as giving the voxelized connectivity strength of a particular cell class from a source voxel to a target voxel. In contrast to Knox et al. (2019), which only uses wild type C57BL/6J mice, our dataset has experiments targeting $|\mathcal{V}| = 114$ different combinations of Cre-driver mice and Cre-regulated AAV transgenes jointly denoted as $\mathcal{V} := \{v\}$. As in Knox et al. (2019), we ultimately estimate an integrated regionalized connectivity defined with respect to a set of $S = 564$ source leafs $\mathcal{S} := \{s\}$ and $T = 1123$ target leafs $\mathcal{T} := \{t\}$, of which $1123 - 564 = 559$ are contralateral. That is, we define

$$\text{regionalized connectivity strength } \mathcal{C} : \mathcal{V} \times \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_{\geq 0} \text{ with } \mathcal{C}(v, s, t) = \sum_{l_j \in s} \sum_{l_{j'} \in t} f(v, l_j, l_{j'}),$$

$$\text{normalized regionalized connectivity strength } \mathcal{C}^N : \mathcal{V} \times \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_{\geq 0} \text{ with } \mathcal{C}^N(v, s, t) = \frac{1}{|s|} \mathcal{C}(v, l_j, l_{j'}),$$

$$\text{normalized regionalized projection density } \mathcal{C}^D : \mathcal{V} \times \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_{\geq 0} \text{ with } \mathcal{C}^D(v, s, t) = \frac{1}{|s||t|} \mathcal{C}(v, l_j, l_{j'})$$

139 where l_j and $l_{j'}$ are the locations of source and target voxels, and $|s|$ and $|v|$ are defined to be the
 140 number of voxels in the source and target structure, respectively. Since the normalized strength and
 141 densities are computable from the strength via a fixed normalization, our main statistical goal is to
 142 estimate $\mathcal{C}(v, s, t)$ for all v, s and t . In other words, we want to estimate matrices $\mathcal{C}_v \in \mathbb{R}_{\geq 0}^{S \times T}$. We call this
 143 estimator $\hat{\mathcal{C}}$.

Construction of such an estimator raises the questions of what data to use for estimating which connectivity, how to featurize the dataset, what statistical estimator to use, and how to reconstruct the connectivity using the chosen estimator. We represent these considerations as

$$\hat{\mathcal{C}}(v, s, t) = f^*(\hat{f}(f_*(\mathcal{D}(v, s))). \quad (1)$$

144 This makes explicit the data featurization f_* , statistical estimator \hat{f} , and any potential subsequent
 145 transformation f^* such as summing over the source and target regions. Denoting \mathcal{D} as a function of v
 146 and s reflects that we consider using different data to estimate connectivities for different cell-classes
 147 and source regions. Table 1 reviews estimators used for this data-type used in previous work, as well
 148 as our two main extensions: the Cre-NW and **Expected Loss** (EL) models. The main differences in our
 149 data featurization from (Knox et al., 2019) are that we regionalize our data at the leaf level where

150 available so that its layer-specific behavior is visible, and normalize our data by projection signal in
 151 order to account for differences between cell class. Additional model selection results are given in
 152 Supplemental Section 5 for alternative normalization strategies, and more detail on estimation is
 153 given in Supplemental Section 6.

Name	f^*	\hat{f}	f_*	$\mathcal{D}(v, s)$
NNLS (Oh et al., 2014)	$\hat{f}(1_s)$	NNLS(X,Y)	$X = x_{\mathcal{S}}, Y = y_{\mathcal{T}}$	I_m/I_m
NW (Knox et al., 2019)	$\sum_{l_s \in s} \hat{f}(l_s)$	NW(X,Y)	$X = l_s, Y = y_{\mathcal{T}}$	I_m/I_m
Cre-NW	$\sum_{l_s \in s} \hat{f}(l_s)$	NW(X,Y)	$X = l_s, Y = y_{\mathcal{T}}$	$(I_l \cap I_v)/I_m$
Expected Loss (EL)	$\sum_{l_s \in s} \hat{f}(s)$	EL(X, Y, v)	$X = l_s, Y = y_{\mathcal{T}}, v$	I_l/I_m

Table 1: Estimation of \mathcal{C} using connectivity data. The regionalization, estimation, and featurization steps are denoted by f^* , \hat{f} , and f_* , respectively. The training data used to fit the model is given by index set I . We denote experiments with centroids in particular major brain divisions and leafs as I_m and I_l , respectively. Data I_l/I_m means that, given a location $l_s \in s \in m$, the model \hat{f} is trained on all of I_m , but only uses I_l for prediction. The non-negative least squares estimator (NNLS) fits a linear model that predicts regionalized projection signal $y_{\mathcal{T}}$ as a function of regionalized injection signal $x_{\mathcal{S}}$. Thus, the regionalization step for a region s is given by applying the learned matrix \hat{f} to the s -th indicator vector. In contrast, the Nadaraya-Watson model (NW) is a local smoothing model that generates a prediction for each voxel within the source structure that are then averaged to create estimate the structure-specific connectivity.

154 Our contributions - the Cre-NW and Expected Loss (EL) models - have several differences from the
 155 previous methods. In contrast to the non-negative least squares (Oh et al., 2014) and
 156 Nadaraya-Watson (Knox et al., 2019) estimators that account only for source region s , our new
 157 estimators account cell class v , The Cre-NW estimator only uses experiments from a particular class
 158 to predict connectivity for that class, while the EL estimator shares information between classes
 159 within a structure. Both of these estimator take into account both the cell-class and the centroid
 160 position of the experimental injection. Like the NW and Cre-NW estimator, the EL estimator generates

¹⁶¹ predictions for each voxel in a structure, and then sums them together to get the overall connectivity.
¹⁶² However, in contrast to the NW approaches, the EL estimate of the projection vector for a cell-class at
¹⁶³ a location weights the average projection of that cell-class in the region containing the location
¹⁶⁴ against the relative locations of all experimental centroids in the region regardless of class. That is,
¹⁶⁵ cell-class and source region combinations with similar average projection vectors will be upweighted
¹⁶⁶ when estimating \hat{f} . Thus, all experiments that are nearby in three-dimensional space can help
¹⁶⁷ generate the prediction, even when there are few nearby experiments for the cell-class in question. A
¹⁶⁸ detailed mathematical description of our new estimator is given in Supplemental Section 6.

169 ***Model evaluation***

170 We select optimum functions from within and between our estimator classes using **leave-one-out**
 171 **cross validation**, in which the accuracy of the model is assessed by its ability to predict projection
 172 vectors experiments excluded from the training data on the basis of their cell class and experimental
 173 centroid. Equation 1 includes a deterministic step f^* included without input by the data. The
 174 performance of $\hat{\mathcal{C}}(v, s, t)$ is thus determined by performance of $\hat{f}(f_*(\mathcal{D}(v, s)))$. Thus, we evaluate
 175 prediction of $f_{\mathcal{T}} : \mathbb{R}^3 \rightarrow \mathbb{R}_{\geq 0}^T$ - the regionalized connection strength at a given location.

176 Another question is what combinations of v , s , and t to generate a prediction for. Our EL and
 177 Cre-NW models are leaf specific. They only generate predictions for cell-classes in leafs where at least
 178 one experiment with a Cre-line targeting that class has a centroid. To accurately compare our new
 179 estimators with less-restrictive models such as used in Knox et al. (2019), we restrict our evaluation set
 180 to Cre driver/leaf combinations that are present at least twice. The sizes of these evaluation sets are
 181 given in Supplemental Section 5.

We use weighted l_2 -loss to evaluate these predictions.

$$\text{l2-loss } \ell(y_{\mathcal{T}}(i)), \widehat{y_{\mathcal{T}}(i)}) := \|y_{\mathcal{T}}(i)) - \widehat{y_{\mathcal{T}}(i)}\|_2^2.$$

$$\text{weighted l2-loss } \mathcal{L}(\widehat{f}(f_*)) := \frac{1}{|\{\mathcal{S}, \mathcal{V}\}|} \sum_{s, v \in \{\mathcal{S}, \mathcal{V}\}} \frac{1}{|I_s \cap I_v|} \sum_{i \in (I_s \cap I_v)} \ell(y_{\mathcal{T}}(i)), \hat{f}_{\mathcal{T}}(f_*(\mathcal{D}(v, s) \setminus i)).$$

182 I_s refers to the set of experiments with centroid in structure s , and I_v refers to the set of experiments
 183 with Cre-line v , so $|I_s \cap I_v|$ is the number of experiments of Cre-line v with injection centroid in
 184 structure s . This is a somewhat different loss from Knox et al. (2019) because of the increased
 185 weighting of rarer combinations of s and v implicit in the $\frac{1}{|I_s \cap I_v|}$ term in the loss. The establishment of
 186 a lower limit of detection and the extra cross-validation step used in the EL model to establish the
 187 relative importance of regionally averaged cell-class projection and injection centroid position are
 188 covered in Supplemental Section 6.

189 ***Connectivity analyses***

190 We examine latent structure underlying our estimated connectome using hierarchical clustering and
 191 non-negative matrix factorization. First, we use agglomerative hierarchical clustering with Ward's
 192 criterion to compare outputs from connectivities from different Cre-lines. Details of this approach are
 193 given in ?? . Second, we use non-negative matrix factorization (NMF) to factor the wild-type
 194 connectivity matrix into a small set of underlying components. Inspired by ?, we refer to these latent
 195 coordinates as **connectivity archetypes** since they represent underlying patterns from which we can
 196 reconstruct a broad range of observed connectivities, although we note that the genomic archetypal
 197 analysis in that paper is slightly methodologically distinct.

198 Our application of NMF to decompose the estimated long-range connectivity into latent
 199 coordinates that linearly combine to reproduce the observed connectivity is some independent
 200 interest, since we censor short range connections due to their clear biological derivation from
 201 diffusion and their high mathematical rank. NMF refers to a collection of **dictionary-learning**
 202 algorithms for decomposing a non-negatively-valued matrix such as \mathcal{C} into positively-valued
 203 matrices called, by convention, weights $W \in \mathbb{R}_{\geq 0}^{S \times q}$ and hidden units $H \in \mathbb{R}_{\geq 0}^{q \times T}$. NMF assumes a simple
 204 linear statistical model: that the observed matrix is composed of linear combinations of latent
 205 coordinates (?). Unlike PCA, NMF specifically accounts for the fact that data are all in the positive
 206 orthant, and it is more stable and interpretable in assays of complex biological systems than
 207 hierarchical clustering (Brunet, Tamayo, Golub, & Mesirov, 2004) The matrix H is typically used to
 208 identify latent structures with interpretable biological meaning, and the choice of matrix factorization
 209 method reflects particular scientific subquestions and probabilistic interpretations.

210 Our NMF algorithm solves the following optimization problem

$$\text{NMF}(\mathcal{C}, \lambda, q) := \arg \min_{W \in \mathbb{R}_{\geq 0}^{S \times q}, H \in \mathbb{R}_{\geq 0}^{q \times T}} \frac{1}{2} \| \mathbf{1}_{d(s,t) > 1500\mu m} \odot \mathcal{C} - WH \|_2^2 + \lambda (\| H \|_1 + \| W \|_1).$$

211 For this decomposition we ignore connections between source and target regions less than $1500\mu m$
 212 apart. This is because short-range projections resulting from diffusion dominate the matrices $\hat{\mathcal{C}}$, and
 213 represent a less-interesting type of biological structure. We set $\lambda = 0.002$ to encourage sparser and
 214 therefore more interpretable components. We use unsupervised cross-validation to determine an

215 optimum q , and show the top 15 stable components (Perry, 2009). Stability analysis accounts for the
216 difficult-to-optimize NMF program by clustering the resultant H from multiple replicates. Since the
217 NMF objective is difficult to optimize and sensitive to initialization, we follow up with a stability
218 analysis. The medians of the component clusters appearing frequently across NMF replicates are
219 selected as **connectivity archetypes**. Details of these approaches are given in Supplementary Sections
220 6 and 7.

3 RESULTS

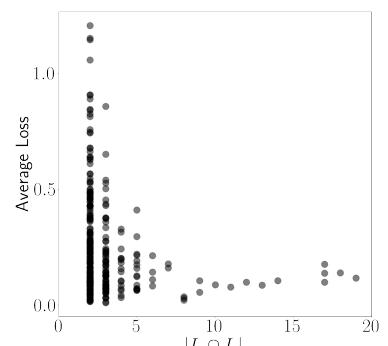
²²¹ We main result of this paper is the collection of cell-type specific connectivity estimates from the Allen
²²² Mouse Connectivity Atlas (MCA) experiments. We first establish that our novel expected-loss (EL)
²²³ estimator performs best in validation assays for estimating wild-type and cell-type specific
²²⁴ connectivities. We then show that Cre-specific connectivity matrices generated using this model are
²²⁵ consistent with known biology. Finally, we factor some of these connectivity matrices to show how
²²⁶ connectivity arises from underlying statistical patterns, and that these may be associated with
²²⁷ cell-types.

²²⁸ ***Model evaluation***

²²⁹ In general, our EL model performs better than the other estimators that we consider. Table 2a
²³⁰ contains weighted losses from leave-one-out cross-validation of candidate models, such as the NW
²³¹ Major-WT model from Knox et al. (2019). The EL model combines the good performance of
²³² class-specific models like NW Leaf-Cre in regions like Isocortex with the good performance of
²³³ class-agnostic models in regions like Thalamus. Additional information on model evaluation,
²³⁴ including class and structure specific performance, is given in Appendix 5. In particular,
²³⁵ Supplementary Table 4 contains the sizes of these evaluation sets in each major structure, and
²³⁶ Supplementary Section 7 contains the structure- and class specific losses.

\hat{f}	Mean Leaf-Cre Mean $I_c \cap I_L$	NW Major-Cre NW $I_c \cap I_M$	NW Leaf-Cre NW $I_c \cap I_L$	NW Leaf NW I_L	NW Major-WT NW $I_{wt} \cap I_M$	NW Major NW I_M	EL EL I_L
Isocortex	0.239	0.252	0.234	0.279	0.274	0.274	0.228
OLF	0.193	0.233	0.191	0.135	0.179	0.179	0.138
HPF	0.175	0.332	0.170	0.205	0.228	0.228	0.153
CTXsp	0.621	0.621	0.621	0.621	0.621	0.621	0.621
STR	0.131	0.121	0.128	0.169	0.232	0.232	0.124
PAL	0.203	0.205	0.203	0.295	0.291	0.291	0.188
TH	0.673	0.664	0.673	0.358	0.379	0.379	0.369
HY	0.360	0.382	0.353	0.337	0.317	0.317	0.311
MB	0.168	0.191	0.160	0.199	0.202	0.202	0.159
P	0.292	0.292	0.292	0.299	0.299	0.299	0.287
MY	0.268	0.347	0.268	0.190	0.189	0.189	0.204
CB	0.062	0.062	0.062	0.068	0.112	0.112	0.068

(a)



(b)

Table 2: 2a Losses from leave-one-out cross-validation of candidate models. **Bold** numbers are best for their major structure. 2b Empirical performance of selected EL model by data abundance. The model is more accurate in Cre-leaf combinations where it draws on more data.

237 **Connectivities**

238 Our main result is the estimation of matrices $\hat{\mathcal{C}}_v \in \mathbb{R}_{\geq 0}^{S \times T}$ representing connections of source structures
 239 to target structures for particular Cre-lines v . We confirm the detection of several well-established
 240 connectivities within our tensor, although we expect additional interesting biological processes to be
 241 identifiable. The connectivity tensor and code to reproduce it are available at
 242 https://github.com/AllenInstitute/mouse_connectivity_models/tree/2020.

243 *Overall connectivity* Several known biological projection patterns are evident in the wild-type
 244 connectivity matrix \mathcal{C}_{wt} shown in Figure 2a. This matrix shows connectivity from leaf sources to leaf
 245 targets. Large-scale patterns like intraareal connectivities and ipsilateral connections between cortex
 246 and thalamus are clear, as in previous estimates in J. A. Harris et al. (2019); Knox et al. (2019); Oh et al.
 247 (2014).

248 The layer-specific targeting of the different Cre-lines enables our estimated wild-type connectivities
 249 to resolve at the layer level. This contrasts the model in Knox et al. (2019), which is denominated as
 250 the NW Major-WT model whose accuracy is evaluated in Table 2a. Layer-specificity is a major

251 advantage of including distinct Cre-lines. For comparison, we also plot averages over component
252 layers weighted by layer size projections between summary-structure sources and targets in the cortex
253 in Figure 2b. Our connectivities therefore exhibit a finer differentiation between regions than those in
254 Knox et al. (2019) This also enables us to lower our limit of detection and examine a broader range of
255 connectivity strengths. Importantly, as shown in Table 2a this finer spatial resolution corresponds to
256 the increased accuracy of our EL model over the NW Major-WT model.

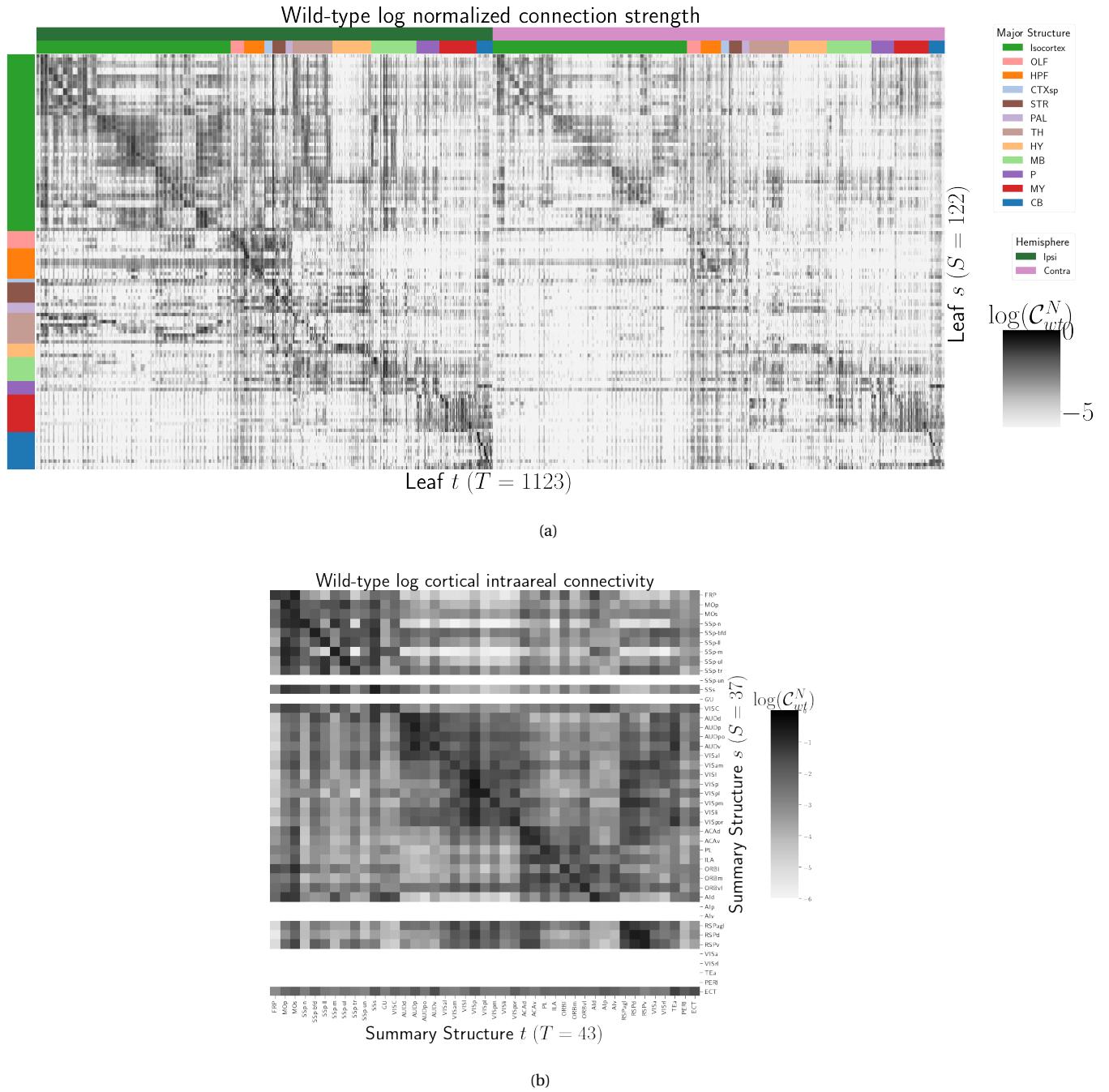


Figure 2: Wild-type connectivities. 2a Log wild-type connectivity matrix $\log \mathcal{C}(s, t, v_{wt})$. 2b Log wild-type intracortical connectivity matrix at the summary structure level.

257 *Class-specific connectivities* Since source and cell-type combinations which project similarly indicate
 258 the network structure underpinning cognition, we investigate the presence of several biological
 259 processes within our connectivity estimates. Although there is a rich anatomical literature using
 260 anterograde tracing data to describe projection patterns from subcortical sources to a small set of
 261 targets of interest, much of the accessible whole brain projection data is from the MCA project used
 262 here to generate the connectome models. Thus, to validate our results while avoiding a circular
 263 validation of the data used to generate the model weights, we confirm that these class-specific
 264 connectivities exhibit certain known behaviors. The cell types and source areas with extensive
 265 previous anatomical descriptions of projections using both bulk tracer methods with cell type
 266 specificity and single cell reconstructions that we investigate are 1) thalamic-projecting neurons in
 267 the visual and motor cortical regions, 2) cholinergic neurons in the medial septum and nucleus of the
 268 diagonal band (MS/NDB); and 3) serotonergic neurons of the dorsal raphe nucleus (DR). Our
 269 estimated connections are in agreement with literature on these cell types.

270 DEPENDENCE OF THALAMIC CONNECTION ON CORTICAL LAYER. Visual cortical areas VISp and VISl and
 271 cortical motor areas MOp and MOs have established layer-specific projection patterns that can be
 272 labeled with the layer-specific Cre-lines from the Allen datasets and others J. A. Harris et al. (2019);
 273 Jeong et al. (2016). Figure 3a shows that in VISp, the Ntsr1-Cre line strongly targets the core part of the
 274 thalamic LGd nucleus while in VISl, it a strong projection to the LP nucleus. In VISp, the Rbp4-Cre line
 275 strongly targets LP as well. Rbp4-Cre and Ntsr1-Cre injections target layers 5 and 6 respectively. Since
 276 we only generate connectivity estimates for structures with at least one injection centroid, this is
 277 shown by the position of non-zero rows in Figure 3a. To fill these gaps, as a heuristic alternative
 278 model, we also display an average connectivity matrix over all Cre-lines.

279 MS AND NDB PROJECTIONS IN THE CHAT-IRES-CRE-NEO MODEL. Cholinergic neurons in the MS and
 280 NDB are well-known to strongly innervate the hippocampus, olfactory bulb, piriform cortex,
 281 entorhinal cortex, and lateral hypothalamus (Watson, Paxinos, & Puelles, 2012; Zaborszky et al., 2015).
 282 In the Allen MCA, cholinergic neurons were labeled by injections into Chat-IRES-Cre-neo mice.
 283 Figure 3b checks the estimated connectome weights to targets in these major brain divisions from MS
 284 and NDB. We observed that all these expected divisions were represented above the 90th percentile of
 285 weights from these source structures.

286 We also compared our Chat-IRES-Cre connectome model data for MS and NDB with the targets
 287 identified by Li et al. (2018). This single cell whole brain mapping project using Chat-Cre mice fully
 288 reconstructed n=50 cells to reveal these same major targets and also naming additional targets from
 289 MS/NDB (Li et al., 2018). We identified 150 targets at the fine leaf structure level among the top decile
 290 of estimated weights. To directly compare our data across studies, we merged structures as needed to
 291 get to the same ontology level, and remove ipsilateral and contralateral information. After formatting
 292 our data, we found 51 targets in the top 10%; Li et al. reported 47 targets across the 50 cells. There was
 293 good consistency overall between the target sets; 35 targets were shared, 12 were unique to the single
 294 cell dataset, and 16 unique to our model data. We checked whether targets missing from our dataset
 295 were because of the threshold level. Indeed, lowering the threshold to the 75 th percentile confirmed 6
 296 more targets-in-common, and all but 2 targets from Li et al. (2018) were above the 50th percentile
 297 weights in our model. Of note, the absence of a target in the single cell dataset that was identified in
 298 our model data is most likely due to the sparse sampling of all possible projections from only n=50
 299 MS/NDB cells.

300 DR PROJECTIONS IN THE SLC6A4-CRE'ET33 MODEL. Serotonergic projections from cells in the
 301 dorsal raphe (DR) are widely distributed and innervate many forebrain structures including isocortex
 302 and amygdala. In the Allen MCA, serotonergic neurons were labeled using Slc6a4-Cre_ET33 and
 303 Slc6a4-CreERT2_EZ13 mice. This small nucleus appears to contain a complex mix of molecularly
 304 distinct serotonergic neuron subtypes with some hints of subtype-specific projection patterns (Huang
 305 et al., 2019; Ren et al., 2018, 2019). We expect that the Cre-lines used here in the Allen MCA, which use
 306 the serotonin transporter promoter (Slc6a4-Cre and -CreERT2), will lead to expression of tracer in all
 307 the serotonergic subtypes recently described in an unbiased way, but this assumption has not been
 308 tested directly. We compared our model data to a single cell reconstruction dataset consisting of n=50
 309 serotonergic cells with somas in the DR that also had bulk tracer validation 3c. After processing our
 310 data to match the target structure ontology level across studies, we identified 37 targets from the DR
 311 with weights above the 90th percentile, whereas Ren et al. (2019) listed 55 targets across the single cell
 312 reconstructions. Twenty seven of these targets where shared.

313 Overall there was good consistency between targets in olfactory areas, cortical subplate, CP, ACB
 314 and amygdala areas, as well in palidum and midbrain, while the two major brain divisions with the

315 least number of matches are the isocortex and thalamus. There are a few likely reasons for these
316 observations. First, in the isocortex, there is known to be significant variation in the density of
317 projections across different locations, with the strongest innervation in lateral and frontal orbital
318 cortices Ren et al. (2019); ?. Indeed, when we lower the threshold and check for weights of the targets
319 outside of the 90%, we see all but one of these regions (PTLp, parietal cortex which is not frontal or
320 lateral) has a weight assigned in the top half of all targets. In the thalamus, our model predicted strong
321 connections to several medial thalamic nuclei (i.e., MD, SMT) that were not targeted by the single
322 cells. This discrepancy may be at least partially explained by the complex topographical organization
323 of the DR that, like the molecular subtypes, is not yet completely understood. A previous bulk tracer
324 study that specifically targeted injections to the central, lateral wings, and dorsal subregions of the DR
325 reported semi-quantitative differences in projection patterns (Muzerelle, Scotto-Lomassese, Bernard,
326 Soiza-Reilly, & Gaspar, 2016). Notably, Muzerelle et al. (2016) report that cells in the ventral region of
327 DR project more strongly to medial thalamic nuclei, whereas the lateral and dorsal DR cells innervate
328 more lateral regions (e.g., LGd). Thus, it is possible that the single cell somas did not adequately
329 sample the entire DR.

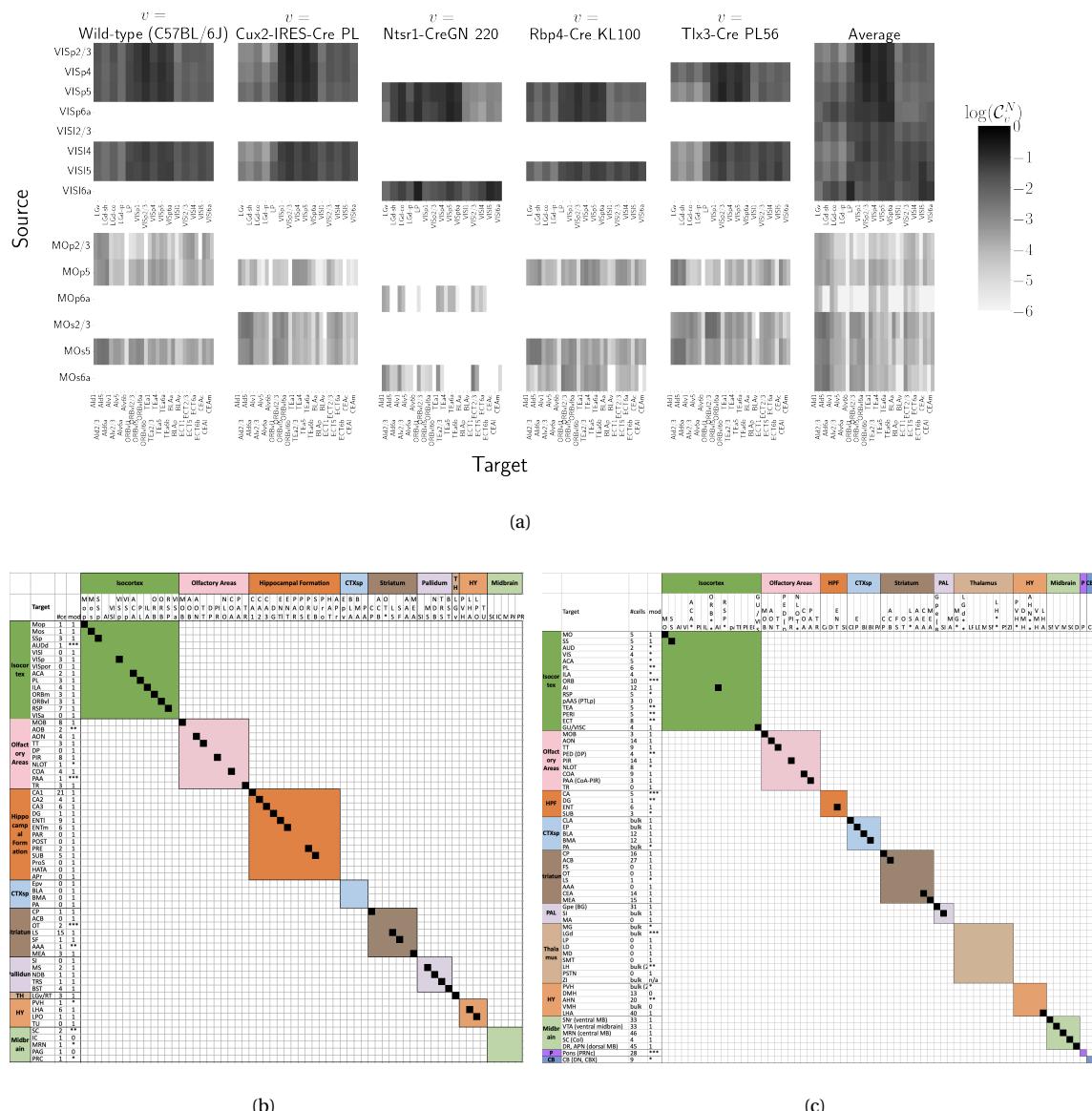


Figure 3: Cell-class specificity. 3a Selected cell-class and layer specific connectivities from two visual and two motor areas. Sources without an injection in the Cre driver line are not estimated due to lack of data for that Cre-line in that structure. 3b Targets reported for cholinergic cells in MS/NDB in Li et al. (2018) and above the 90th percentile in our Chat-IRES-Cre-neo model. A black box along the diagonal indicates the target was identified in both studies ($n = 35$). The number of single cells (out of 50) with a projection to each target is shown in the first column after the target acronym and the next column shows whether a target appeared in the 90% thresholded model weights (1=present, 0=absent). Some of these targets only appeared at lower thresholds as indicated by asterisks, * * * > 85th%, ** > 75th%, * > 50th%. 3c Targets reported for serotonergic cells in DR in Ren et al. (2018, 2019) and above the 90th percentile in our Slc6a4-Cre_ET33 model.

330 Connectivity Analyses

331 While the manual analysis in Section 3 is valuable for validation, scaling our interpretation of our
 332 connectivity estimates motivated us to apply dimension reduction methods to our connectivity
 333 estimates and understand whether the learned structure was consonant with the expected biology.
 334 For example, Supplemental Figure ?? shows a collection of connectivity strengths generated using
 335 Cre-specific models for wild-type, Cux2, Ntsr1, Rbp4, and Tlx3 cre-lines from VIS areas at leaf level in
 336 the cortex to cortical and thalamic nuclei. Sorting source and target structure/cell-class combinations
 337 hierarchical clustering shows, for example, that layer 6-specific Ntsr1 Cre-lines source regions cluster
 338 together. This makes sense, since layer 6-specific Ntsr1 Cre-line distinctly projects to thalamic nuclei,
 339 regardless of source summary structure, and in contrast with the tendency of other cell-classes to
 340 project to nearby regions within the cortex.

341 The low-dimensional coordinates returned by NMF highlight important features within the
 342 connectivity matrix in a data-driven way. The learned projection archetypes H and model weights W
 343 are plotted in Figure 4. Intrastructural connections such as MB-MB MY-MY are visible in the 7th and
 344 11th archetypes, but there is no obvious layer-6 specific signal. These factors may be used to generate
 345 a reconstructed connectivity matrix using the implied statistical model. Despite the relatively small
 346 number of learned additive factors, comparing with 2a shows that this reconstructed matrix has a
 347 relatively globally plausible structure. Supplemental Sections 6 and 7 contain quantitative
 348 performance of this model, and assessment of factorization stability to ensure the decomposition is
 349 reliable across computational replicates. These indicate that though the displayed coordinate are
 350 fewer than the optimum number, this optimum is still an order of magnitude smaller than the
 351 dimension of the matrix itself. The supplement also quantitatively shows differential association of
 352 projection archetypes in this model with projection vectors of sources from the Cux2, Ntsr1, Rbp4,
 353 and Tlx3 Cre-lines.

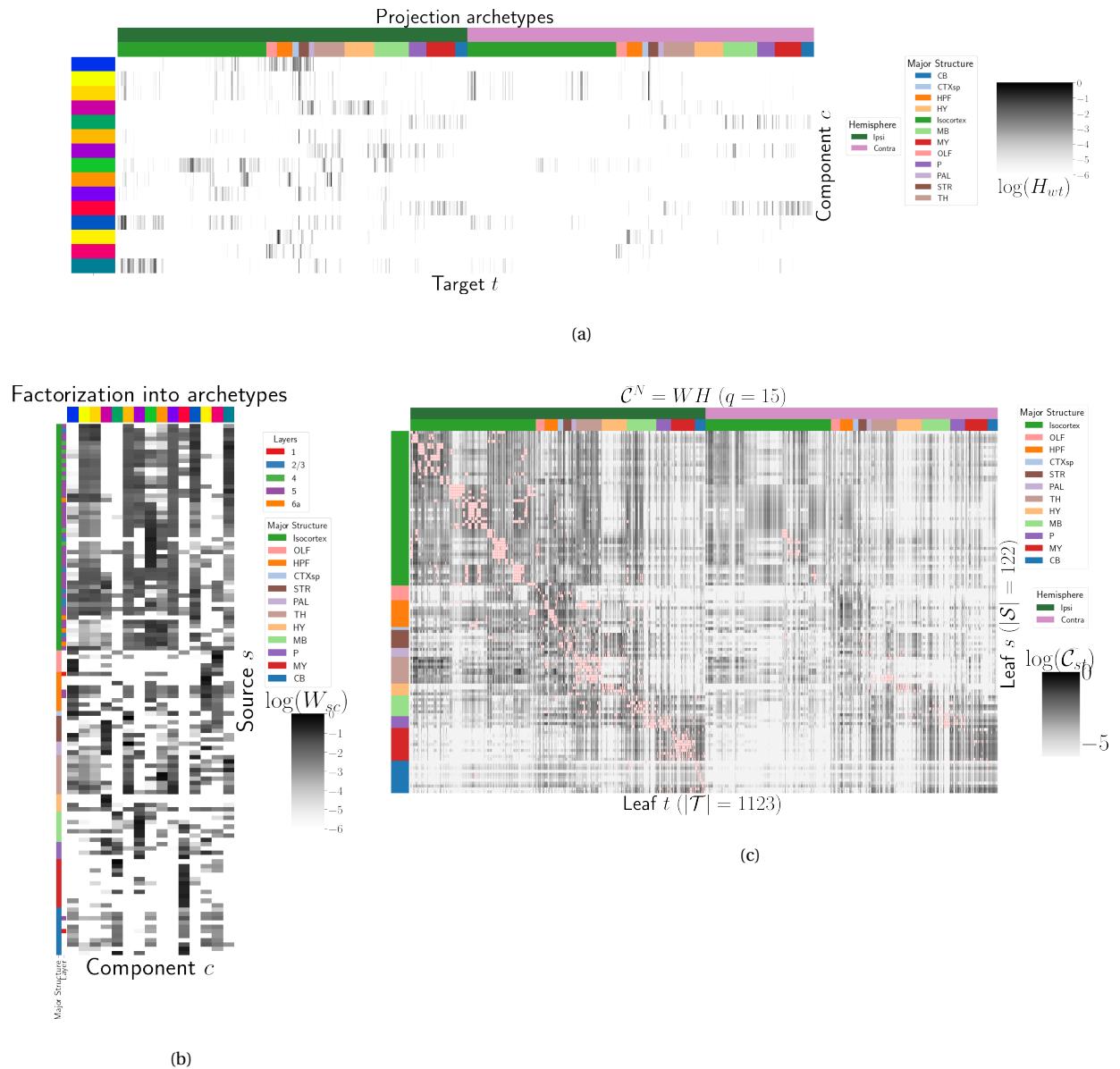


Figure 4: Non-negative matrix factorization results $\mathcal{C}_{wt}^N = W H$ for $q = 15$ components. **4a** Latent space coordinates H of \mathcal{C} . Target major structure and hemisphere are plotted. **4b** Loading matrix W . Source major structure and layer are plotted. **4c** Reconstruction of the normalized distal connectivity strength using the top 15 archetypes. Areas less than $1500\mu m$ apart are not modeled, and therefore shown in pink.

4 DISCUSSION

354 The model presented here among the first cell-type specific whole brain projectome models for a
 355 mammalian species, and it opens the door for a large number of models linking brain structure to
 356 computational architectures. Overall, we find expected targets, based on our anatomical expertise
 357 and published reports, but underscore that the core utility of this bulk connectivity analysis is not
 358 only in validation of existing connection patterns, but also in identification of new ones. We note that
 359 although the concordance appeared stronger for the cholinergic cells than the serotonergic cells, any
 360 differences might still be explained by the lack of high quality “ground-truth” datasets to validate
 361 these Cre connectome models. Larger numbers of single cell reconstructions, that saturate all
 362 possible projection types, would be a better gold standard than the small number of cells reported
 363 here. Perhaps future iterations of the connectome model may also take into account some single cell
 364 axon projection data.

365 The Nadaraya-Watson estimator presented here is novel. Beyond using a Nadaraya-Watson kernel
 366 regression defined in physical space, we define a cell-type space based on similarities of projections,
 367 and theoretically justify the use of an intermediate shape-constrained estimator. While methods like
 368 non-negative least squares can also account for covariates, the centroid method from Knox et al.
 369 (2019) was shown that the more precise notion of injection location than the non-negative least
 370 squares in Oh et al. (2014). Furthermore, our sample size seems too low to utilize a fixed or mixed
 371 effect, particularly since the impact of the virus depend on the particular injection region. In a sense
 372 both the NNLS and NW models can be thought of as improvements over the structure-specific
 373 average, and so is also possible that a yet undeveloped residual-based data-driven blend of these
 374 models could provide improved performance.

375 We see several other opportunities for improving on our model. Ours is certainly not the first
 376 cross-validation based model averaging method Gao, Zhang, Wang, and Zou (2016). However, our use
 377 of shape-constrained estimator in target-encoded feature space is novel and fundamentally different
 378 from Nadaraya-Watson estimators that use an optimization method for selecting the weights (Saul &
 379 Roweis, 2003). The properties of this estimator, as well as its relation to estimators fit using an
 380 optimization algorithm, are a possible future avenue of research (??) A deep model such as Lotfollahi,
 381 Naghipourfar, Theis, and Alexander Wolf (2019) could be appropriate, provided enough data was

382 available. Finally, a Wasserstein-based measure of injection similarity per structure would combine
383 both the physical simplicity of the centroid model while also incorporating the full distribution of the
384 injection signal.

385 The factorization of the connectivity matrix could also be improved and better used. From a
386 statistical perspective, stability-based method for establishing archetypal connectivities in NMF is
387 similar to those applied to genomic data Kotliar et al. (2019); Wu et al. (2016). However, non-linear
388 data transformations or matrix decompositions, or tensor factorizations that account for correlations
389 between cell-types could better capture the true nature of latent neural connections (K. D. Harris et
390 al., 2016). An interesting goal for improved decompositions of the connectivity matrices here would be
391 to identify layer-specific signals from within the wild-type matrix. Regardless of statistical approach,
392 latent low-dimensional organization in connectivity inspires search for similarly parsimonious
393 biological correlates.

ACKNOWLEDGMENTS

³⁹⁴ We thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement,
³⁹⁵ and support.

396 This supplement is divided into information about our dataset, supplemental methods, and
397 supplemental results. However, certain topics are revisited between sections. Thus, if a reader is
398 interested in, say, non-negative matrix factorization, they may find relevant information in both
399 methods and results.

5 SUPPLEMENTAL INFORMATION

400 Our supplementary information consists of abundances of leaf/Cre-line combinations, information
401 about distances between structures, and the size of our restricted evaluation dataset.

402 *Cre/structure combinations in \mathcal{D}*

403 This section describes the abundances of structure and Cre-line combinations in our dataset. That is,
404 it indicates how many experiments in our dataset with a particular Cre-line have an injection centroid
405 in a particular structure. Users of the connectivity matrices who are interested in a particular Cre-line
406 or structure can see the quantity and type of data used to compute and evaluate that connectivity.

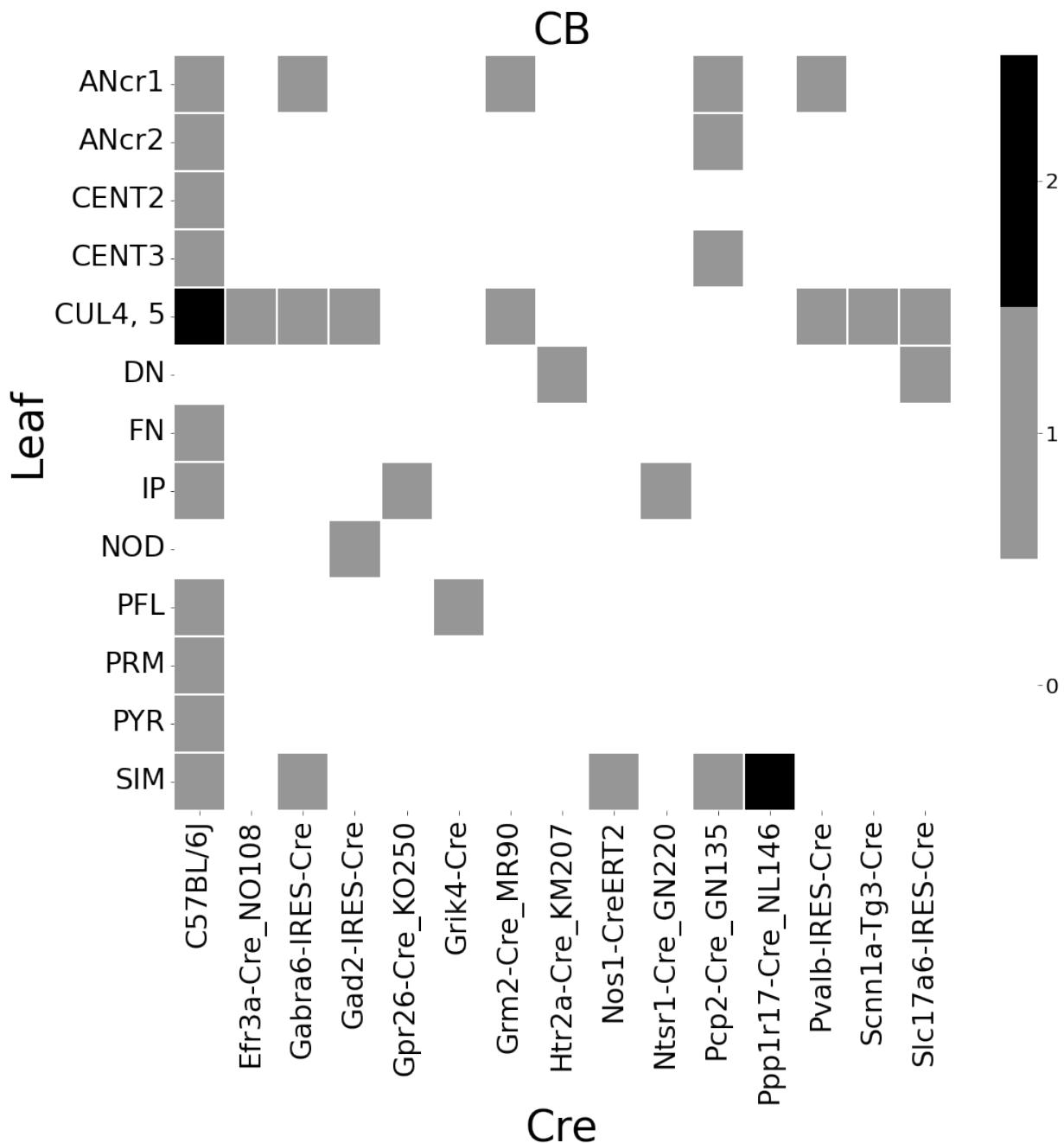


Figure 5: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

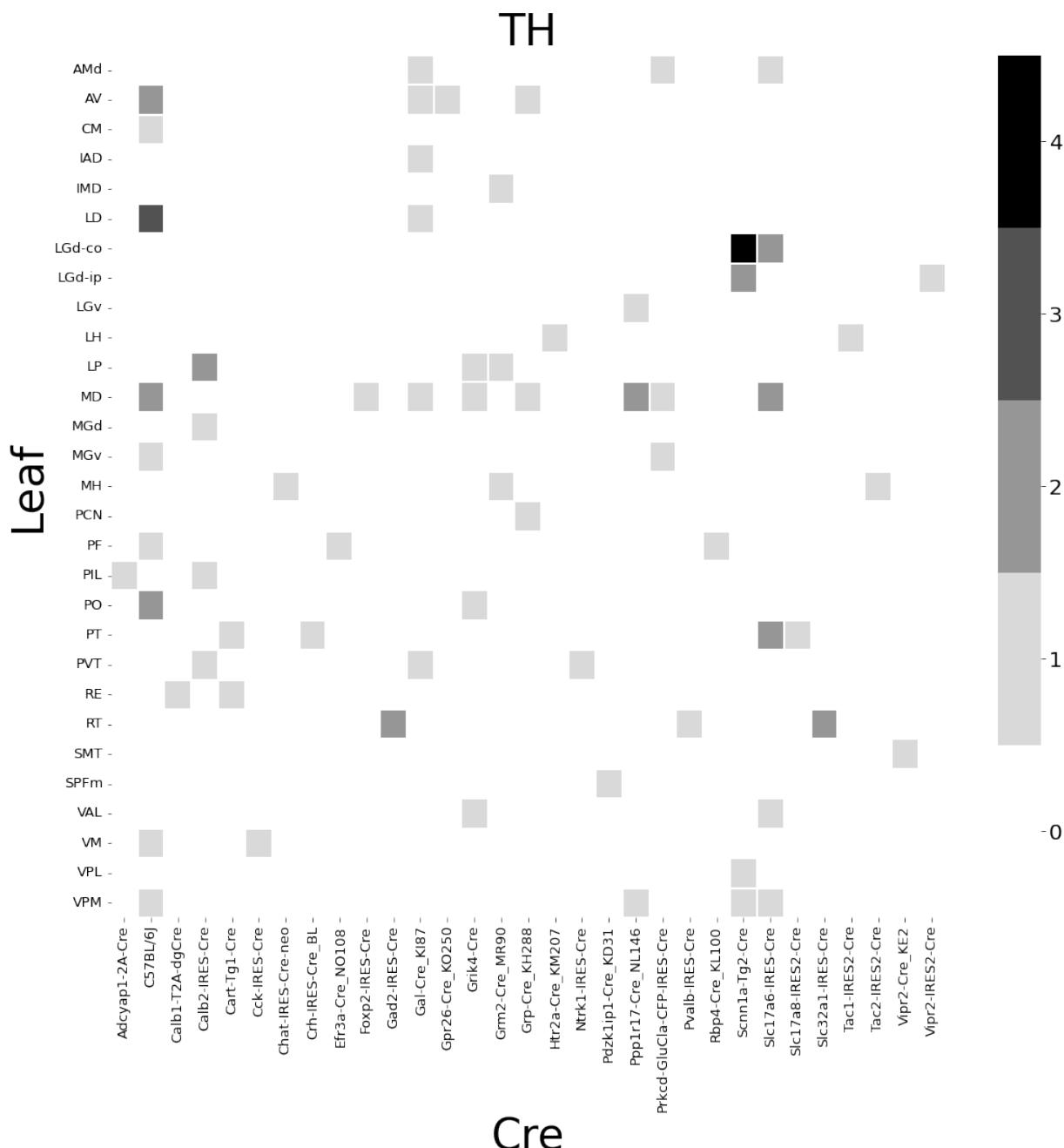


Figure 6: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

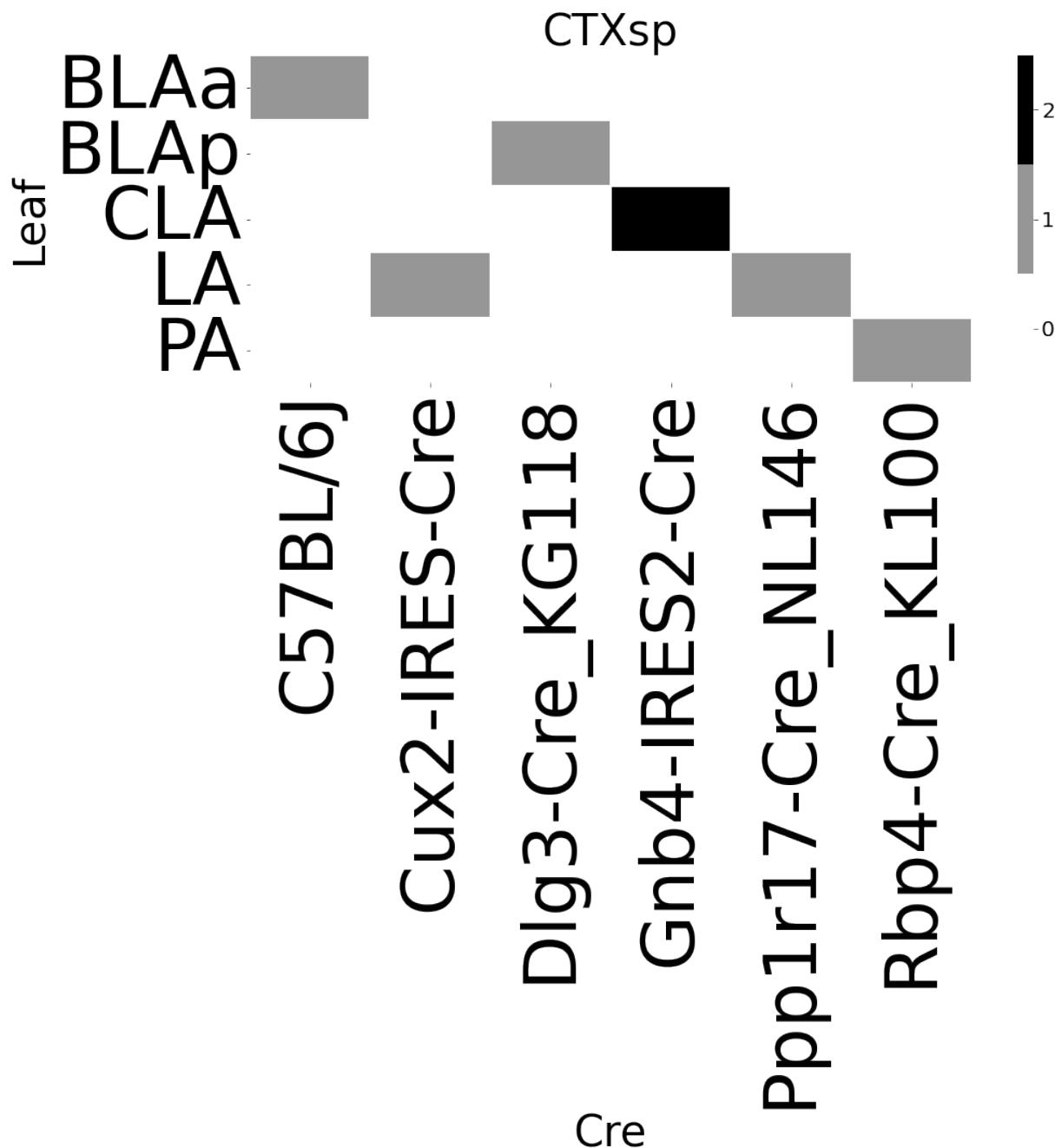


Figure 7: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

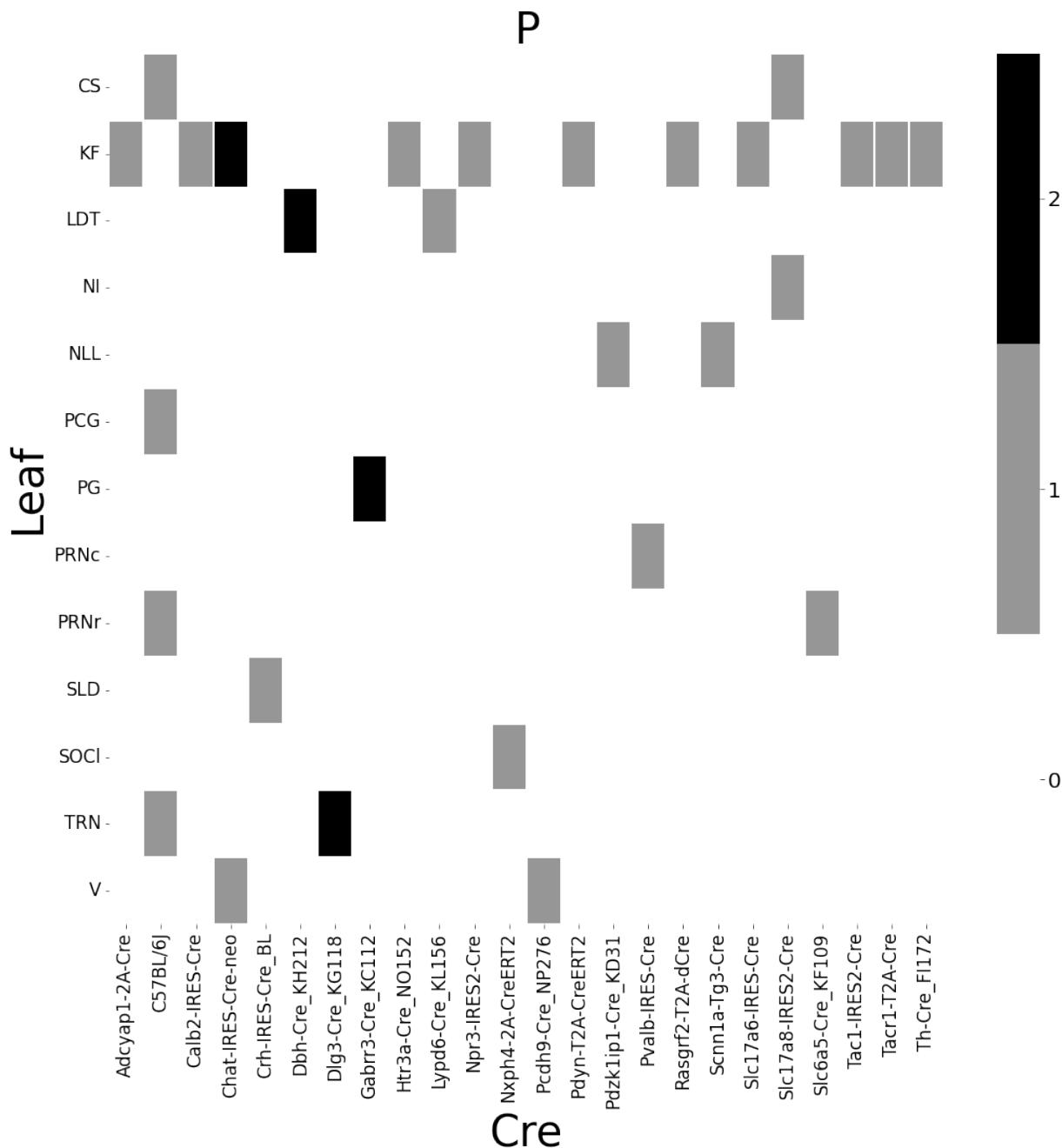


Figure 8: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

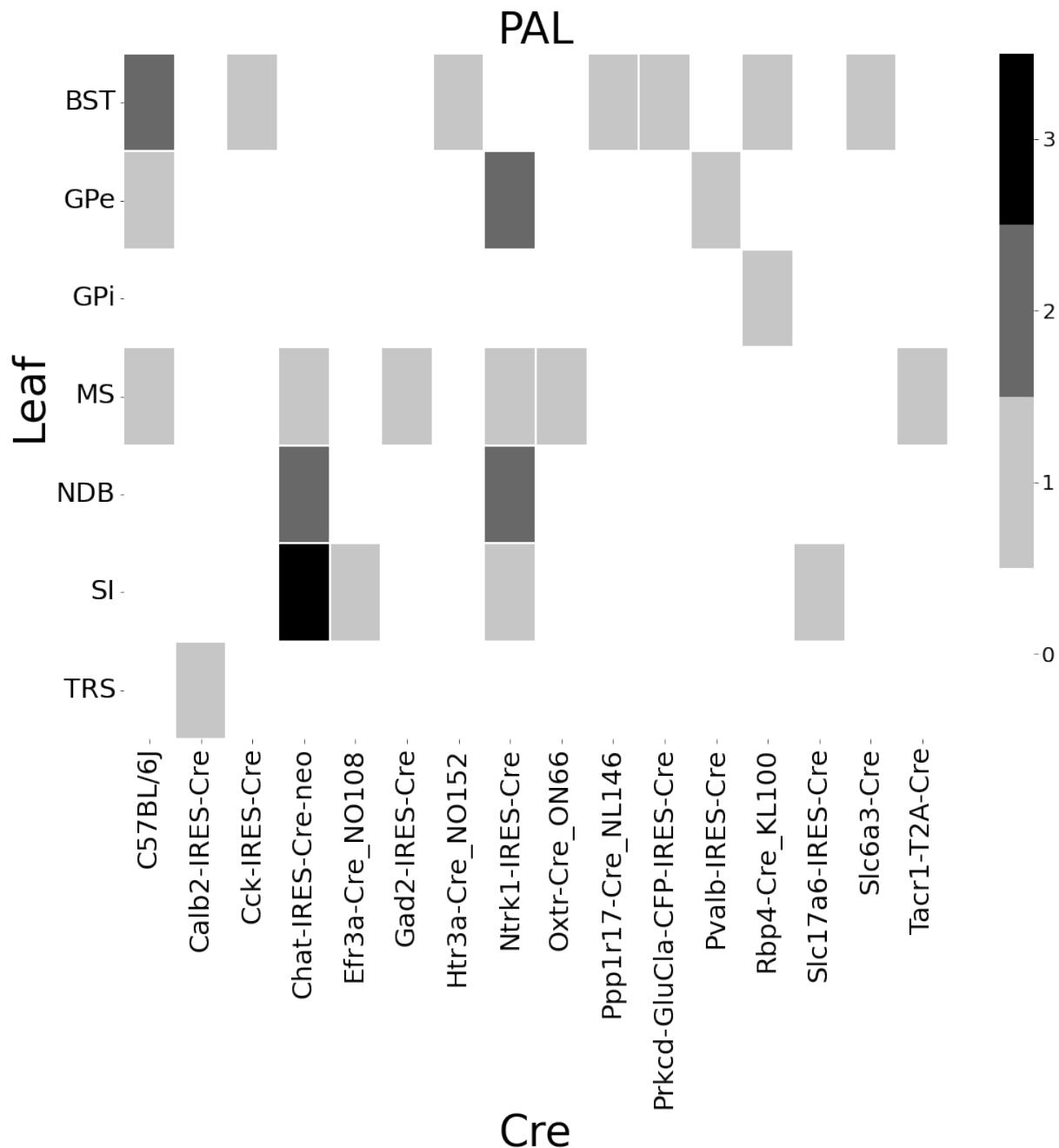


Figure 9: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

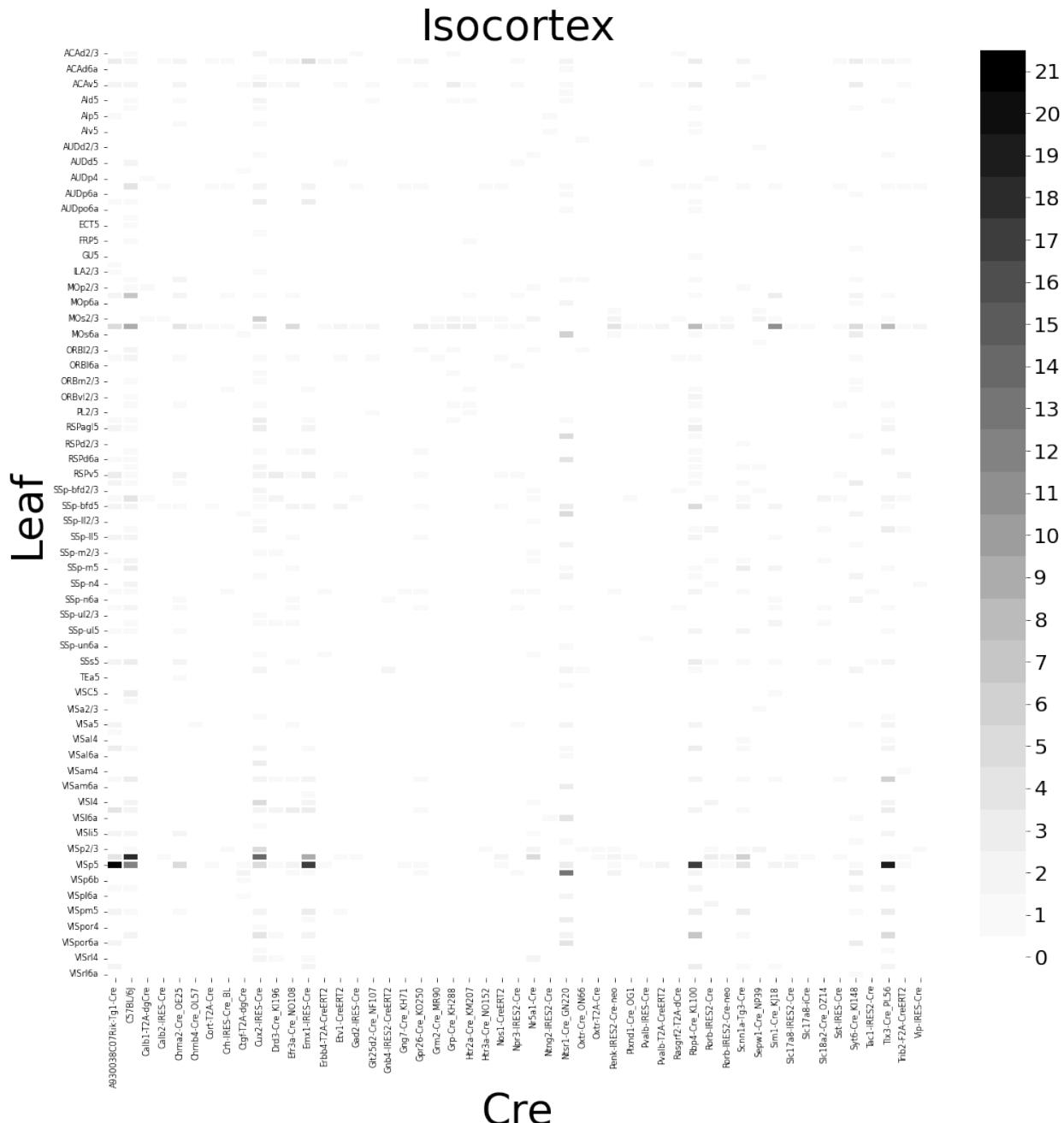


Figure 10: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

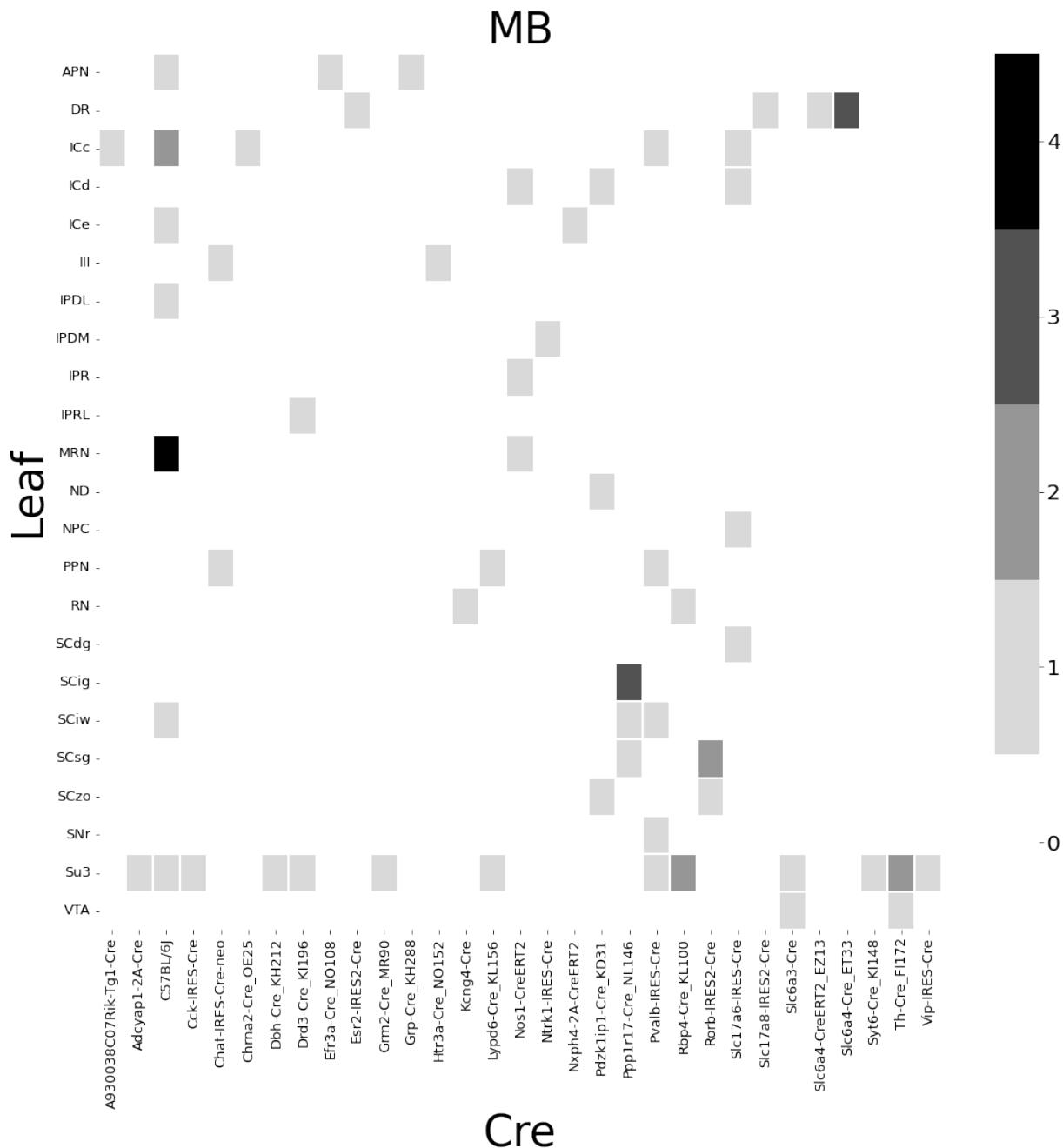


Figure 11: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

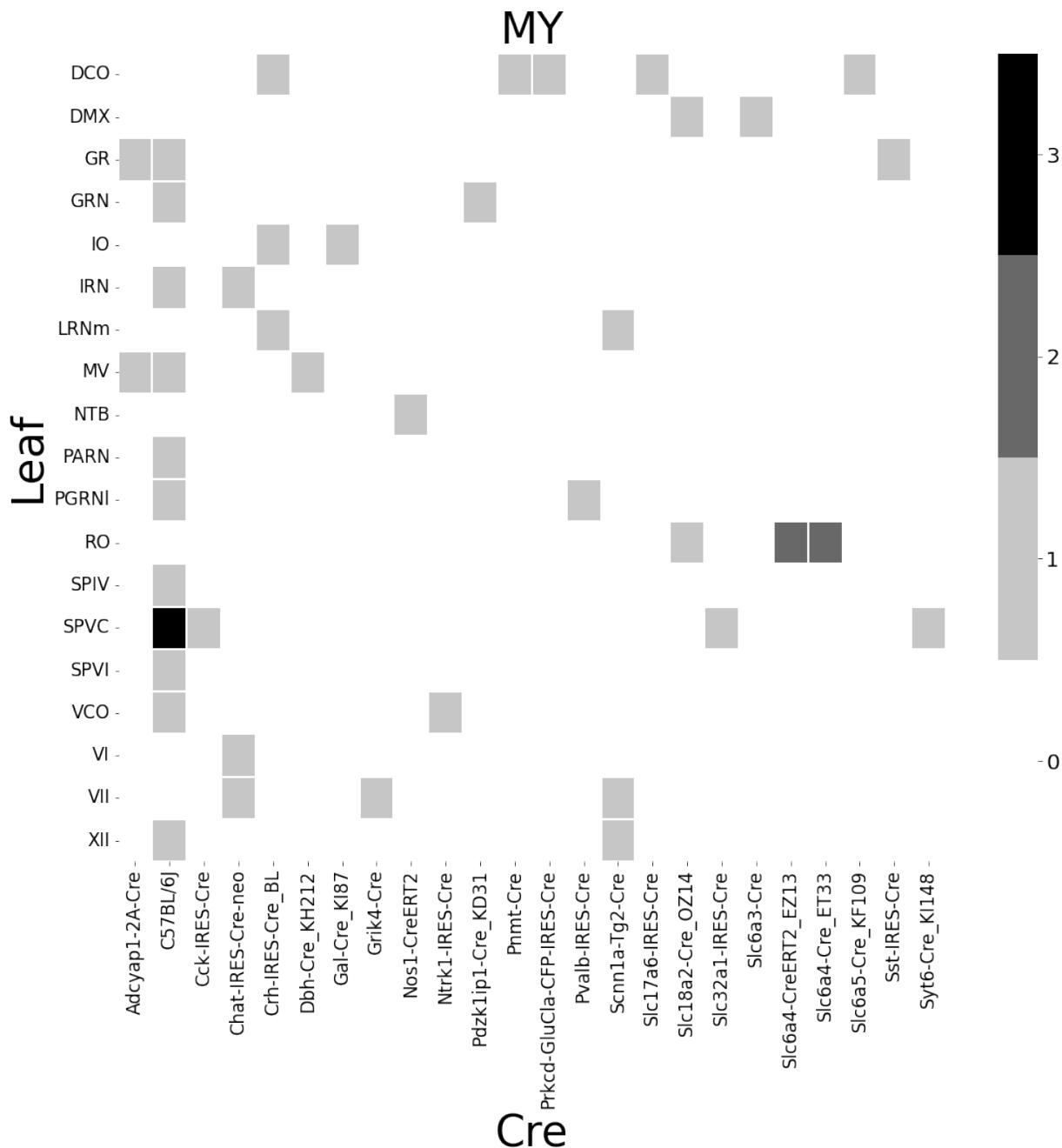


Figure 12: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

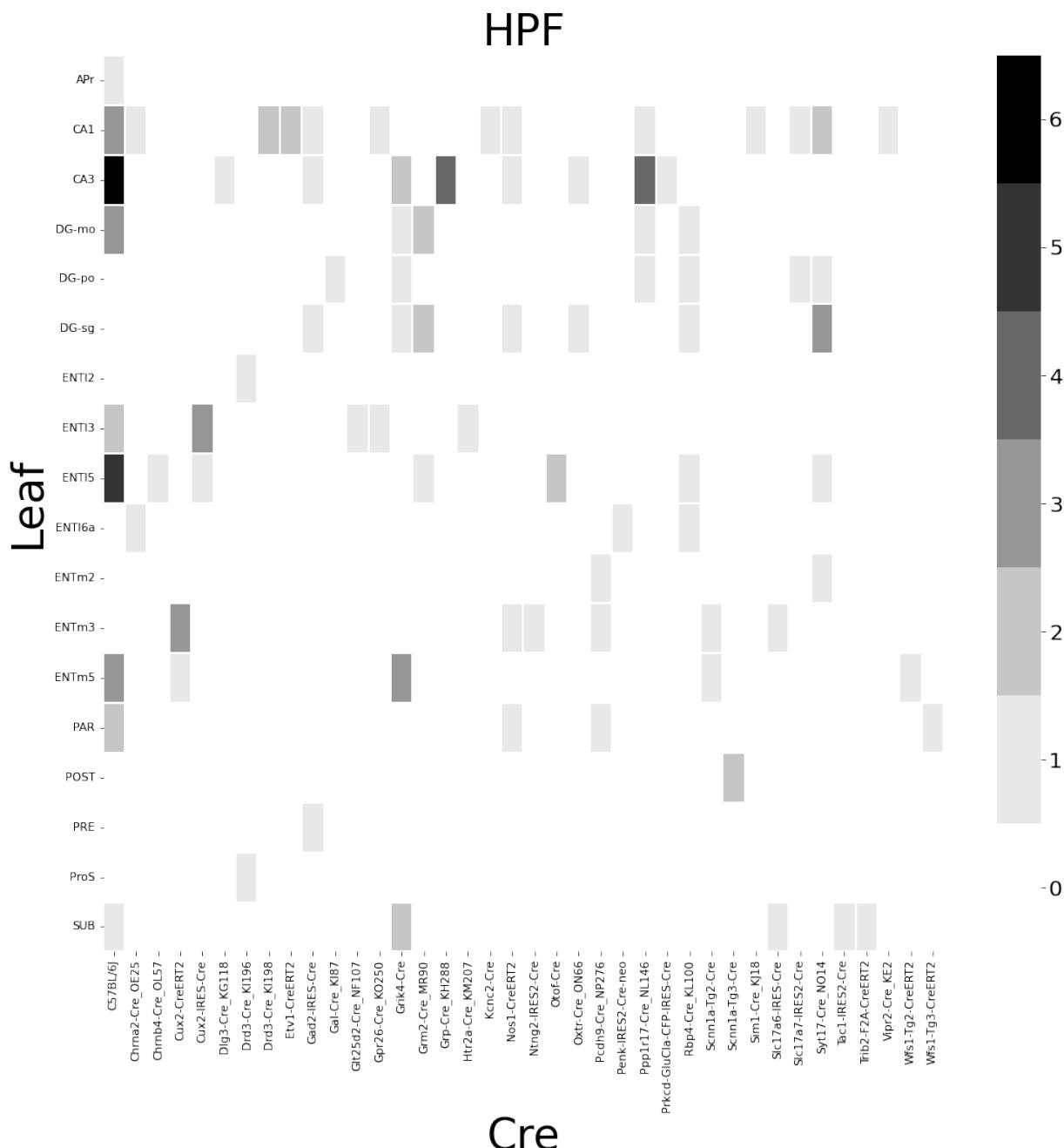


Figure 13: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

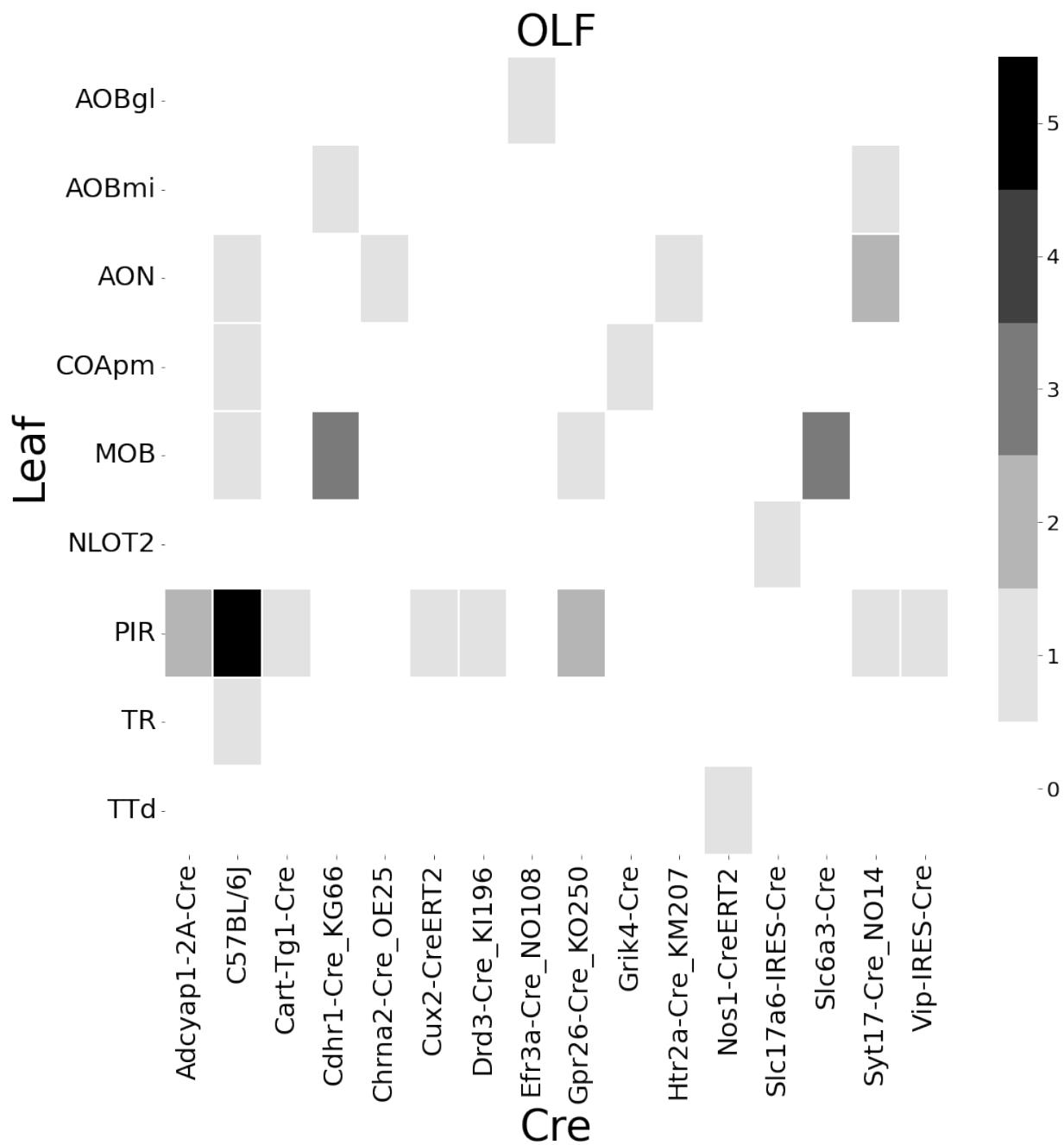


Figure 14: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

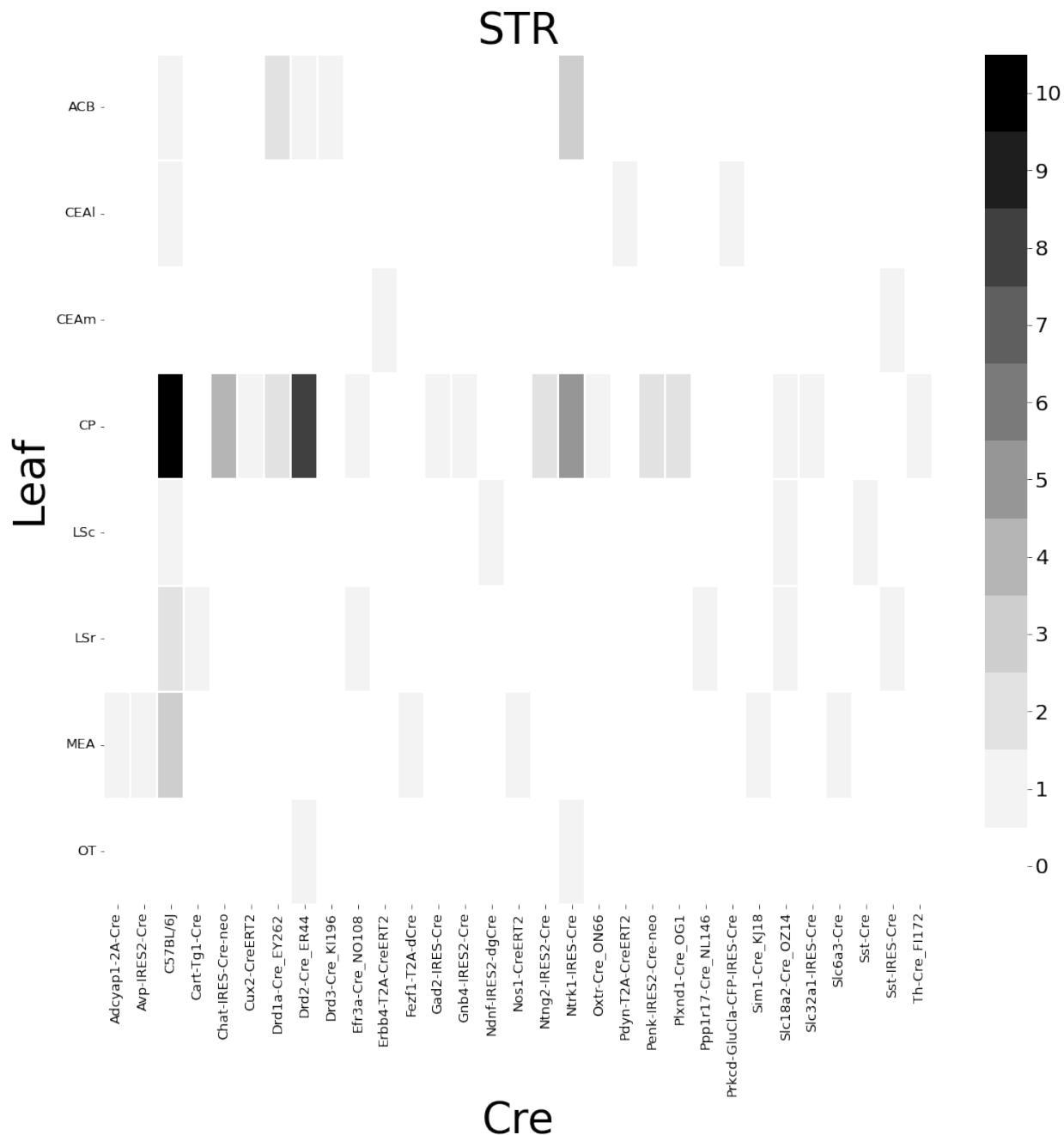


Figure 15: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

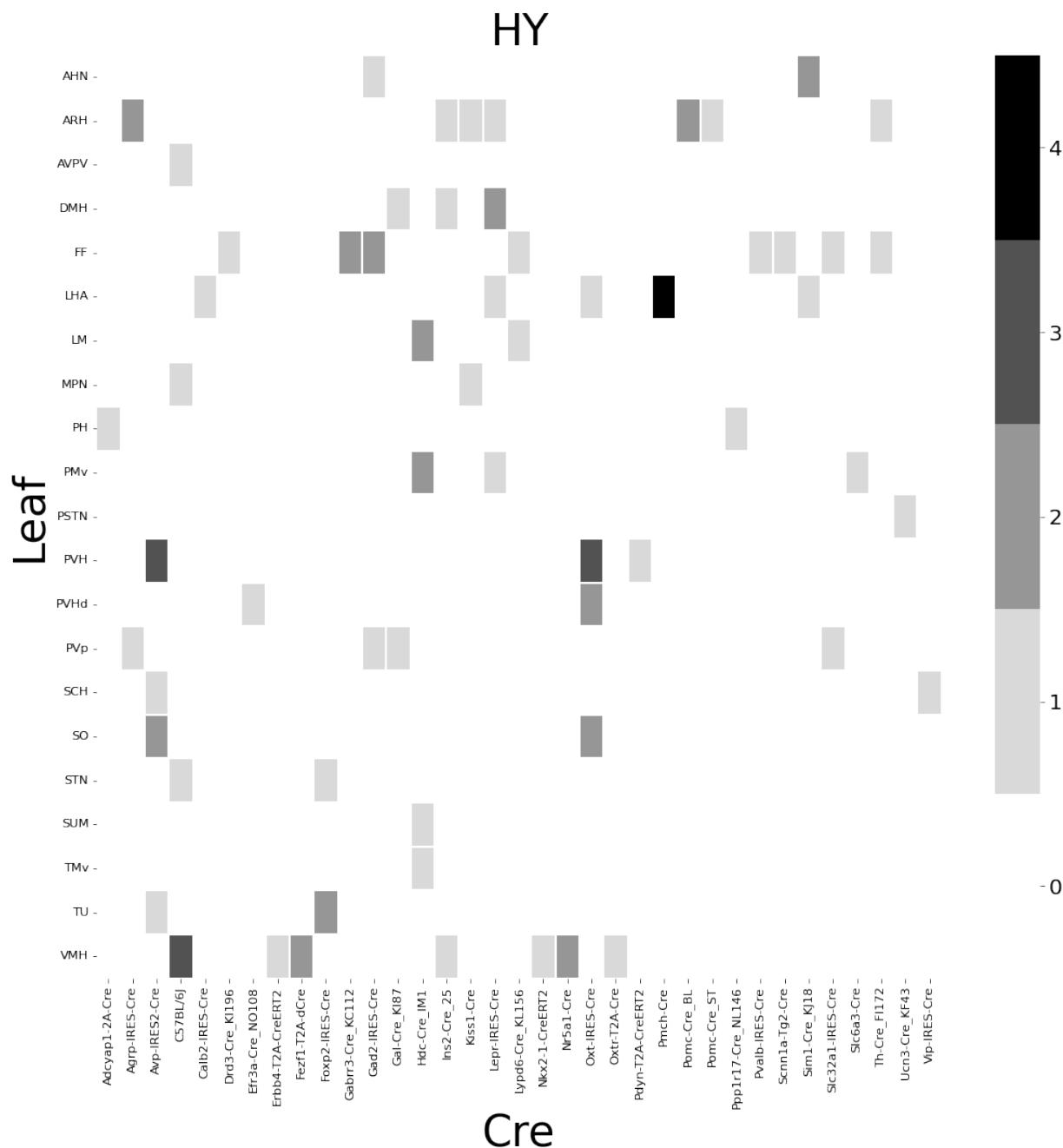


Figure 16: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

407 ***Distances between structures***

408 The distance between structures has a strong effect on the connectivity (Knox et al., 2019). For
 409 reference, we show these distances here. Short range distances are not used in our matrix
 410 factorization approach. This masking is methodologically novel.

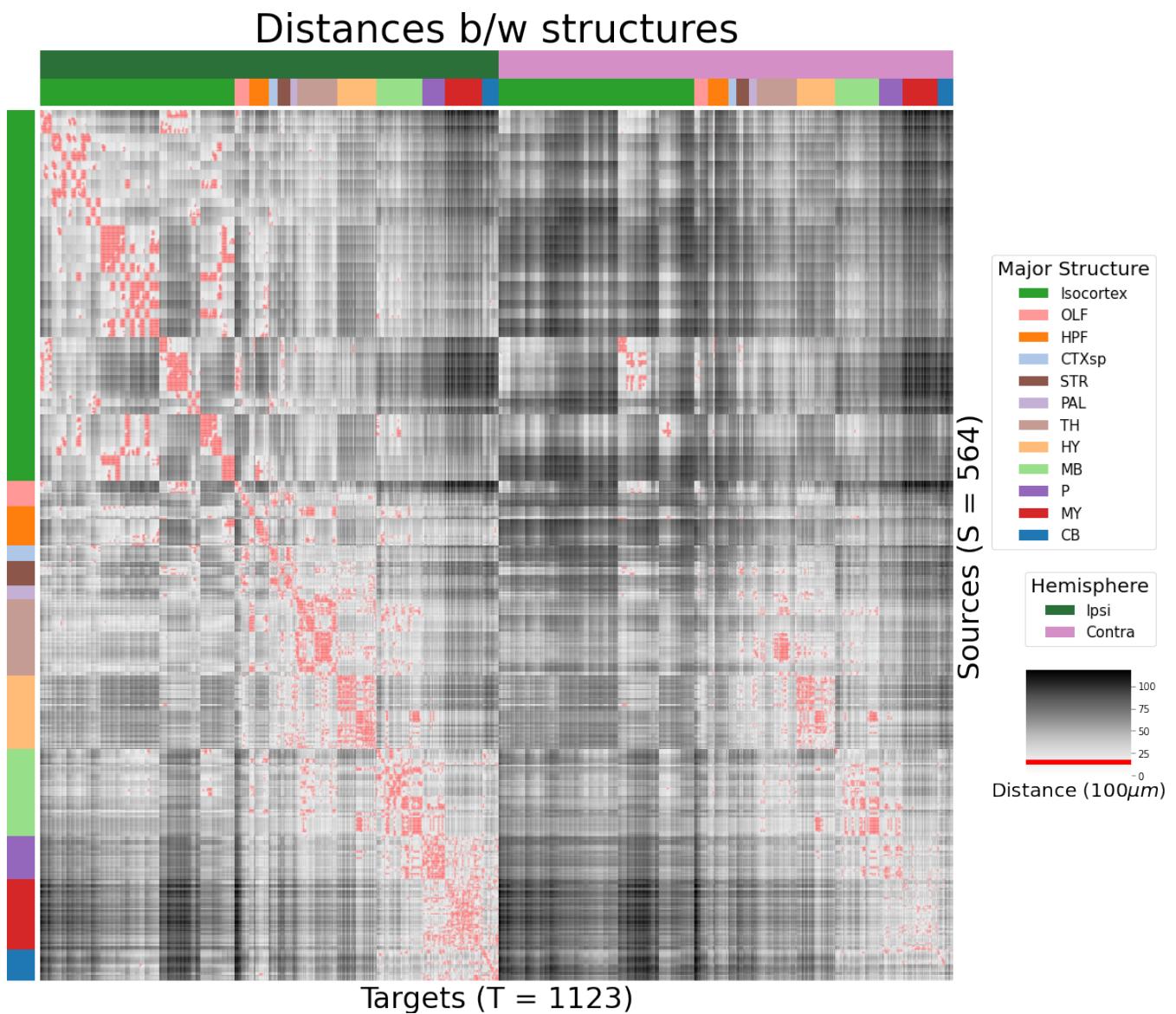


Figure 17: Distance between structures. Short-range connections are masked in red.

⁴¹² We give the sizes of our evaluation sets in leave-one-out cross-validation and additional losses using
⁴¹³ the injection-based normalization scheme from Knox et al. (2019).
⁴¹¹ ***Model evaluation***

⁴¹⁴ NUMBER OF EXPERIMENTS IN EVALUATION SETS In order to compare between methods, we restrict to
⁴¹⁵ the smallest set of evaluation indices. That is, the set of experiments used to validate our models
⁴¹⁶ combinations that are those whose combination of Cre-line and injection centroid leaf are present at
⁴¹⁷ least twice, since one experiment at least must be held out. This means that our evaluation set is
⁴¹⁸ smaller in size than our overall list of experiments.

	Total	Evaluation set	Injection-thresholded evaluation
Isocortex	1128	732	709
OLF	35	17	17
HPF	122	62	59
CTXsp	7	2	2
STR	78	45	44
PAL	30	11	11
TH	83	29	29
HY	85	41	38
MB	68	18	18
P	33	8	8
MY	46	7	7
CB	36	4	4

Table 3: Number of experiments available to evaluate models in leave-one-out cross validation. The size of the evaluation set is lower than the total number of experiments since models that rely on a finer granularity of modeling have less data available to validate with, and we restrict all models to the smallest evaluation set necessitated by any of the modes. In this case, the Expected Loss and Cre-NW models require at least two experiments to be present with a combination of injection centroid structure and Cre-line for leave-one-out cross-validation. We also include results with a slightly smaller evaluation set that removes experiments without a sufficiently strong injection signal, as in Knox et al. (2019)

419 INJECTION-NORMALIZED LOSSES To compare with the injection-normalization procedure from Knox
 420 et al. (2019), we also remove experiments with small injection, and here give results for this slightly
 421 reduced set using injection-normalization. That is, instead of dividing the projection signal of each
 422 experiment by its l_1 norm, we divide by the l_1 norm of the corresponding injection signal. We find
 423 that setting a summed injection-signal of threshold of 1 is sufficient for evading pathological edge
 424 cases in this normalization, while still retaining a large evaluation set.

\hat{f}	Mean	NW	EL				
\mathcal{D}	$I_c \cap I_L$	$I_c \cap I_M$	$I_c \cap I_L$	I_L	$I_{wt} \cap I_M$	I_M	I_L
Isocortex	0.413	0.453	0.408	0.538	0.528	0.528	0.396
OLF	0.499	0.504	0.494	0.441	0.543	0.543	0.437
HPF	0.336	0.483	0.332	0.444	0.501	0.501	0.321
CTXsp	0.497	0.497	0.497	0.497	0.497	0.497	0.497
STR	0.359	0.386	0.359	0.364	0.433	0.433	0.322
PAL	0.519	0.497	0.519	0.436	0.459	0.459	0.434
TH	0.769	0.767	0.769	0.514	0.539	0.539	0.556
HY	0.414	0.439	0.414	0.441	0.452	0.452	0.399
MB	0.459	0.396	0.397	0.358	0.324	0.324	0.403
P	0.562	0.562	0.562	0.758	0.764	0.764	0.562
MY	0.699	0.552	0.621	0.439	0.578	0.578	0.439
CB	0.849	0.689	0.849	0.500	0.615	0.615	0.495

Table 4: Losses from leave-one-out cross-validation of candidate for injection-normalized regionalized connectivity on injection-thresholded evaluation set. **Bold** numbers are best for their major structure.

425 PROJECTION-NORMALIZED LOSSES ON THRESHOLDED SET We also give results for the
 426 projection-normalization procedure from the main text on this reduced subset.

\hat{f}	Mean	NW	EL				
\mathcal{D}	$I_c \cap I_L$	$I_c \cap I_M$	$I_c \cap I_L$	I_L	$I_{wt} \cap I_M$	I_M	I_L
Isocortex	0.229	0.248	0.224	0.274	0.269	0.269	0.217
OLF	0.193	0.233	0.191	0.135	0.179	0.179	0.138
HPF	0.178	0.342	0.172	0.212	0.235	0.235	0.172
CTXsp	0.621	0.621	0.621	0.621	0.621	0.621	0.621
STR	0.128	0.117	0.124	0.171	0.234	0.234	0.125
PAL	0.203	0.205	0.203	0.295	0.291	0.291	0.188
TH	0.673	0.664	0.673	0.358	0.379	0.379	0.417
HY	0.358	0.378	0.351	0.331	0.312	0.312	0.314
MB	0.168	0.191	0.160	0.199	0.202	0.202	0.160
P	0.292	0.292	0.292	0.299	0.299	0.299	0.287
MY	0.268	0.347	0.268	0.167	0.189	0.189	0.196
CB	0.062	0.062	0.062	0.068	0.108	0.108	0.061

Table 5: Losses from leave-one-out cross-validation of candidate for normalized regionalized connectivity on injection-thresholded evaluation set. **Bold** numbers are best for their major structure.

6 SUPPLEMENTAL METHODS

⁴²⁷ This section consists of additional information on preprocessing of the neural connectivity data,
⁴²⁸ estimation of connectivity, and matrix factorization.

⁴²⁹ ***Data preprocessing***

⁴³⁰ Several data preprocessing steps take place prior to evaluations of the connectivity matrices. These
⁴³¹ steps are described in Algorithm PREPROCESS. The arguments of this normalization process - injection
⁴³² signals $x(i)$, projection signals $y(i)$, injection fraction $F(i)$, and data quality mask $q(i)$ - were
⁴³³ downloaded using the Allen SDK, a programmatic interface to the brain connectivity data. The
⁴³⁴ injections and projection signals $\mathcal{B} \rightarrow [0, 1]$ were segmented manually in histological analysis. The
⁴³⁵ projection signal gives the proportion of pixels within the voxel displaying fluorescence, and the
⁴³⁶ injection signal gives the proportion of pixels within the histologically-selected injection subset
⁴³⁷ displaying fluorescence. The injection fraction $F(i) : \mathcal{B} \rightarrow [0, 1]$ gives the proportion of pixels within
⁴³⁸ each voxel in the injection subset. Finally, the data quality mask $q(i) : \mathcal{B} \rightarrow \{0, 1\}$ gives the voxels that
⁴³⁹ have valid data.

⁴⁴⁰ Our preprocessing makes use of the above ingredients, as well as several other essential steps. First,
⁴⁴¹ we compute the weighted injection centroid

$$c(i) = \sum_{l \in \mathcal{B}} x(i)|_l$$

⁴⁴² where $x(i)|_l$ is the injection density at location $l \in \mathbb{R}^3$. Given a regionalization \mathcal{R} from the Allen SDK,
⁴⁴³ we can also access regionalization map $R : \mathcal{B} \rightarrow \mathcal{R}$. This induces a functional of connectivities from
⁴⁴⁴ the space of maps $\{\mathcal{X} = x : \mathcal{B} \rightarrow [0, 1]$

$$1_{\mathcal{R}} : \mathcal{X} \rightarrow \mathcal{R} \times \mathbb{R}_{\geq 0}$$

$$x \mapsto \sum_{l \in r} x(l) \text{ for } r \in \mathcal{R}.$$

⁴⁴⁵ We also can restrict a signal to a individual structure as

$$1|_S : \mathcal{X} \rightarrow \mathcal{X}$$

$$x(l) = \begin{cases} x(l) & \text{if } l \in S \\ 0 & \text{otherwise.} \end{cases}$$

⁴⁴⁶ Finally, given a vector or array $a \in \mathbb{R}^T$, we have the $l1$ normalization map

$$n : a \mapsto \frac{a}{\sum_{j=1}^T a_j}.$$

⁴⁴⁷ Denote m as the major structure containing an experiment, and define \odot for maps $\mathcal{B} \rightarrow [0, 1]$ by e.g.

⁴⁴⁸ $(y(i) \odot q(i))|_l := (y(i)|_l)(q(i)|_l)$. We then can write the preprocessing algorithm.

PREPROCESS 1 Input Injection x , Projection y , Injection centroid $c \in \mathbb{R}^3$, Injection fraction F , data quality mask q

Injection fraction $x_F \leftarrow x \odot F$

Data-quality censor $y_q \leftarrow y \odot q, x_q \leftarrow x_F \odot q$

Restrict injection $x_m = 1|_m x_q$.

Compute centroid c from x_m

Regionalize $\tilde{y}_{\mathcal{T}} \leftarrow 1_{\mathcal{T}}(y_q)$

Normalize $y_{\mathcal{T}} \leftarrow n(\tilde{y}_{\mathcal{T}})$

Output $\tilde{y}_{\mathcal{T}}, c$

449 **Estimators**

450 As mentioned previously, we can consider our estimators as modeling a connectivity vector
 451 $f_{\mathcal{T}}(\nu, s) \in \mathbb{R}_{\geq 0}^T$. Thus, for the remainder of this section, we will discuss only $f(\nu, s)$. We review the
 452 Nadaraya-Watson estimator from Knox et al. (2019), and describe its conversion into our cell-class
 453 specific Expected Loss estimator.

454 *Centroid-based Nadaraya-Watson* In the Nadaraya-Watson approach of Knox et al. (2019), the injection
 455 is considered only through its centroid $c(i)$, and the projection is considered regionalized. That is,

$$f_*(i) = \{c(i), y_{\mathcal{T}}(i)\}.$$

456 Since the injection is considered only by its centroid, this model only generates predictions for
 457 particular locations l , and the prediction for a structure s is given by integrating over locations within
 458 the structure

$$f^*(\hat{f}(f_*(\mathcal{D})))(\nu, s) = \sum_{l \in s} \hat{f}(f_*(\mathcal{D}(I)))(\nu, l).$$

459 Here, I is the training data, and \hat{f} is the Nadaraya-Watson estimator

$$\hat{f}_{NW}(c(I), y_{\mathcal{T}}(I))(l) := \sum_{i \in I} \frac{\omega_{il}}{\sum_{i \in I} \omega_{il}} y_{\mathcal{T}}(i)$$

460 where $\omega_{il} := \exp(-\gamma d(l, c(i))^2)$ and d is the Euclidean distance between centroid $c(i)$ and voxel with
 461 position l .

462 Several facets of the estimator are visible here. A smaller γ corresponds to a greater amount of
 463 smoothing, and the index set $I \subseteq \{1 : n\}$ generally depends on s and ν . Varying γ bridges between
 464 1-nearest neighbor prediction and averaging of all experiments in I . In Knox et al. (2019), I consisted
 465 of experiments sharing the same brain division, i.e. $I = I_m$, while restricting of index set to only
 466 include experiments with the same cell class gives the class-specific Cre-NW model. Despite this
 467 restriction, we fit γ by leave-one-out cross-validation for each m rather than a smaller subset like s or
 468 ν . That is,

$$\hat{\gamma}_m = \arg \min_{\gamma \in \mathbb{R}_{\geq 0}} \frac{1}{|\{s, \nu\}|} \sum_{s, \nu \in \{m, \mathcal{V}\}} \frac{1}{|I_s \cap I_\nu|} \sum_{i \in (I_s \cap I_\nu)} \ell(y_{\mathcal{T}}(i)), \hat{f}_{\mathcal{T}}(f_*(\mathcal{D}(\nu, s) \setminus i)). \quad (2)$$

⁴⁶⁹ *The Expected-Loss estimator* Besides location of the injection centroid, cell class also influences
⁴⁷⁰ projection. Thus, we introduce method for estimating the effect of Cre-distance, which we define as
⁴⁷¹ the distance between the projections of the mean experiment of one (Cre,leaf) pair with another.
⁴⁷² Equivalently, relatively small Cre-distance defines what we call similar cell classes. This method
⁴⁷³ assigns a predictive weight to each pair of training points that depends both on their
⁴⁷⁴ centroid-distance and Cre-distance. This weight is determined by the expected prediction error of
⁴⁷⁵ each of the two feature types

⁴⁷⁶ We define Cre-line behavior as the average regionalized projection of a Cre-line in a given structure
⁴⁷⁷ (i.e. leaf). The vectorization of categorical information is known as **target encoding**

$$\bar{y}_{\mathcal{T},s,v} := \frac{1}{|I_s \cap I_v|} \sum_{i \in (I_s \cap I_v)} y_{\mathcal{T}}(i)$$

⁴⁷⁸ The Cre-distance is then defined as a **Cre-distance** in a leaf to be the distance between the target-encoded
⁴⁷⁹ projections of two Cre-lines. The relative predictive accuracy of Cre-distance and centroid distance is
⁴⁸⁰ determined by fitting a surface of projection distance as a function of Cre-distance and centroid
⁴⁸¹ distance. When we use shape-constrained B-splines to estimate this weight, the weights then may be
⁴⁸² said to be used in a Nadaraya-Watson estimator. For this reason, we call this the Expected Loss
⁴⁸³ Estimator. The resulting weights are then utilized in a Nadaraya-Watson estimator in a final
⁴⁸⁴ prediction step.

⁴⁸⁵ In mathematical terms, our full feature set consists of the centroid coordinates and the
⁴⁸⁶ target-encoded means of the combinations of virus type and injection-centroid structure. That is,

$$f_*(\mathcal{D}_i) = \{c(i), \{\bar{y}_{\mathcal{T},s,v} \forall v\}, y_{\mathcal{T}}(i)\}.$$

⁴⁸⁷ f^* is defined as in (2). The expected loss estimator is then

$$\hat{f}_{EL}(c(I), y_{\mathcal{T}}(I))(l, v) := \sum_{i \in I} \frac{v_{ilv}}{\sum_{i \in I} v_{ilv}} y_{\mathcal{T}}(i)$$

⁴⁸⁸ where

$$v_{ilv} := \exp(-\gamma g(d(l, c(i))^2, d(\bar{y}_{\mathcal{T},s,v}, \bar{y}_{\mathcal{T},s,v(i)})^2))$$

⁴⁸⁹ and s is the structure containing l .

490 The key step therefore is finding a suitable function g with which to weight the positional and
 491 (Cre,leaf) information. Note that g must be a concave, non-decreasing function of its arguments with
 492 with $g(0, 0) = 0$. Then, g defines a metric on the product of the metric spaces defined by experiment
 493 centroid and target-encoded cre-line, and \hat{f}_{EL} is a Nadaraya-Watson estimator. A derivation of this
 494 fact is given later in this section.

495 We therefore use a linear generalized additive model of shape-constrained B-splines to estimate g
 496 (Eilers & Marx, 1996). This is a method for generating a predictive model g that minimizes the loss of

$$\sum_{i, i' \in S} \| \|y_{\mathcal{T}}(i) - y_{\mathcal{T}}(i')\|_2 - \sum_{q=1}^Q \rho_q B_q(\|c(i') - c(i)\|_2, \|\bar{y}_{\mathcal{T}, s, v} - \bar{y}_{\mathcal{T}, s, v}(i)\|) \|_2$$

497 given the constraints on g . That is, given all pairs of experiments with injection centroid in the same
 498 structure, g gives a prediction of the distance between their projections made using the distance
 499 between the average behavior of their Cre-lines given their injection centroid, and the distance
 500 between their injection centroids. In particular, g is the empirically best such function within the
 501 class of *B*-splines, which Similarly to the Nadaraya-Watson model, we make the decision to fit a g
 502 separately for each major brain division, and select γ as in 2. We set $Q = 10$ and leave validation of this
 503 parameter, as well as the precise nature of the polynomial *B*-spline terms B_q out of the scope of this
 504 paper. Empirically this leads to a smooth surface using the pyGAM Python package (Servén D., n.d.).

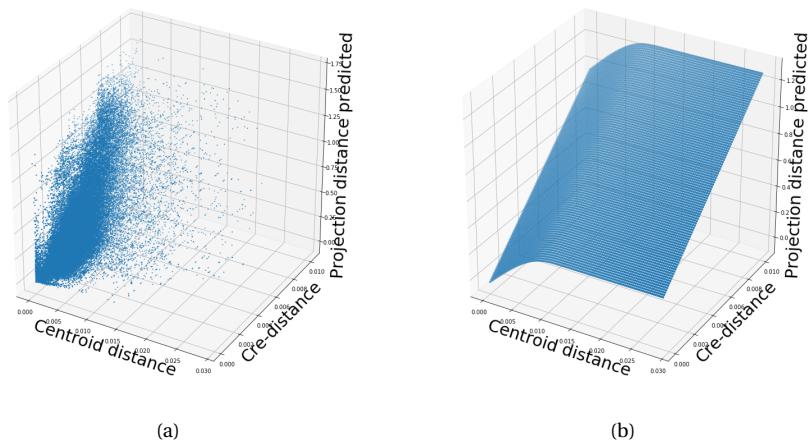


Figure 18: Fitting g . 18a Distribution of projection errors against centroid distance and cre-distance in Isocortex. 18b estimated \hat{g} using B-splines.

505 JUSTIFICATION OF SHAPE CONSTRAINT The shape-constrained expected-loss estimator introduced
 506 in this paper is, to our knowledge, novel. It should be considered an alternative method to the classic
 507 weighted kernel method (??). While we do not attempt a detailed theoretical study of this estimator,
 508 we do establish the need for the shape constraint in our spline estimator. Though this fact is probably
 509 well known, we prove a (slightly stronger) version here for completeness.

510 **Proposition 1.** *Given a collection of metric spaces X_1, \dots, X_n with metrics d_1, \dots, d_n (e.g. $d_{centroid}, d_{cre}$),
 511 and a function $f : (X_1 \times X_1) \dots \times (X_n \times X_n) = g(d_1(X_1 \times X_1), \dots, d_n(X_n \times X_n))$, then f is a metric if g is
 512 concave, non-decreasing and $g(d) = 0 \iff d = 0$.*

513 *Proof.* We show g satisfying the above properties implies that f is a metric.

- 514 ▪ The first property of a metric is that $f(x, x') = 0 \iff x = x'$. The left implication:
 515 $x = x' \implies f(x_1, x'_1, \dots, x_n, x'_n) = g(0, \dots, 0)$, since d are metrics. Then, since $g(0) = 0$, we have that
 516 $f(x, x') = 0$. The right implication: $f(x, x') = 0 \implies d = 0 \implies x = x'$ since d are metrics.
- 517 ▪ The second property of a metric is that $f(x, x') = f(x', x)$. This follows immediately from the
 518 symmetry of the d_i , i.e. $f(x, x') = f(x_1, x'_1, \dots, x_n, x'_n) = g(d_1(x_1, x'_1), \dots, d_n(x_n, x'_n)) =$
 519 $g(d_1(x'_1, x_1), \dots, d_n(x'_n, x_n)) = f(x'_1, x_1, \dots, x'_n, x_n) = f(x', x)$.
- 520 ▪ The third property of a metric is the triangle inequality: $f(x, x') \leq f(x, x^*) + f(x^*, x')$. To show this
 521 is satisfied for such a g , we first note that $f(x, x') = g(d(x, x')) \leq g(d(x, x^*) + d(x^*, x'))$ since g is
 522 non-decreasing and by the triangle inequality of d . Then, since g is concave,
 523 $g(d(x, x^*) + d(x^*, x')) \leq g(d(x, x^*)) + g(d(x^*, x')) = f(x, x^*) + f(x^*, x')$.

525 ***Setting a lower detection threshold***

526 The lower detection threshold of our approach is a complicated consequence of our experimental and
 527 analytical protocols. For example, the Nadaraya-Watson estimator is likely to generate many small
 528 false positive connections, since the projection of even a single experiment within the source region
 529 to a target will cause a non-zero connectivity in the Nadaraya-Watson weighted average. On the other
 530 hand, the complexities of the experimental protocol itself and the image analysis and alignment can
 531 also cause spurious signals. Therefore, it is of interest to establish a lower-detection threshold below
 532 which we have very little power-to-predict, and set estimated connectivities below this threshold to
 533 zero.

534 We set this threshold with respect to the sum of Type 1 and Type 2 errors

$$\iota = \sum_{i \in \mathcal{E}} 1_{y_{\mathcal{T}}^T(i)=0}^T 1_{\hat{f}_{\mathcal{T}}(\nu(i), c(i)) > \tau} + 1_{y_{\mathcal{T}}^T(i) > 0}^T 1_{\hat{f}_{\mathcal{T}}(\nu(i), c(i)) < \tau}.$$

535 We then select the τ that minimizes ι . Results for this approach are given in Supplemental Section 7.

536 ***Decomposing the connectivity matrix***

537 We utilize non-negative matrix factorization (NMF) to analyze the principal signals in our
 538 connectivity matrix. Here, we review this approach as applied to decomposition of the distal elements
 539 of the estimated connectivity matrix $\hat{\mathcal{C}}$ to identify q connectivity archetypes. Aside from the NMF
 540 program itself, the key elements are selection of the number of archetypes q and stabilization of the
 541 tendency of NMF to give random results over different initializations.

542 *Non-negative matrix factorization* As discussed in Knox et al. (2019), one of the most basic processes
 543 underlying the observed connectivity is the tendency of each source region to predominantly project
 544 to proximal regions. For example, the heatmap in Supplemental Figure 17 shows that the pattern of
 545 infrastructure distances resembles the connectivity matrix in 2. These connections are biologically
 546 meaningful, but also unsurprising, and their relative strength biases learned latent coordinate
 547 representations away from long-range structures. For this reason, we establish a $1500\mu\text{m}$ 'distal'
 548 threshold within which to exclude connections for our analysis.

549 Given a matrix $X \in \mathbb{R}_{\geq 0}^{a \times b}$ and a desired latent space dimension q , the non-negative matrix
 550 factorization is thus

$$\text{NMF}(\mathcal{C}, \lambda, q, \mathbf{1}_M) = \arg \min_{W \in \mathbb{R}_{\geq 0}^{S \times q}, H \in \mathbb{R}_{\geq 0}^{q \times T}} \frac{1}{2} \|\mathbf{1}_M \odot \mathcal{C} - WH\|_2^2 + \lambda(\|H\|_1 + \|W\|_1).$$

551 The mask $\mathbf{1}_M$ specifies this objective for detecting patterns in long-range connections. We note the
 552 existence of NMF with alternative norms for certain marginal distributions, but leave utilization of
 553 this approach for future work (Brunet et al., 2004).

554 The mask $\mathbf{1}_M \in \{0, 1\}^{S \times T}$ serves two purposes. First, it enables computation of the NMF objective
 555 while excluding self and nearby connections. These connections are both strong and linearly
 556 independent, and so would unduly influence the *NMF* reconstruction error over more biologically
 557 interesting or cell-type dependent long-range connections. Second, it enables cross-validation based
 558 selection of the number of retained components.

559 *Cross-validating NMF* We review cross-validation for NMF following (Perry, 2009). In summary, a
 560 NMF model is first fit on a reduced data set, and an evaluation set is held out. After random masking
 561 of the evaluation set, the loss of the learned model is then evaluated on the basis of successful
 562 reconstruction of the held-out values. This procedure is performed repeatedly, with replicates of
 563 random masks at each tested dimensionality q . This determines the point past which additional
 564 hidden units provide no additional value for reconstructing the original signal.

565 The differentiating feature of cross-validation for NMF compared with supervised learning is the
 566 randomness of the masking matrix 1_M . Cross-validation for supervised learning generally leaves out
 567 entire observations, but this is insufficient for our situation. This is because, given W , our H is the
 568 solution of a regularized non-negative least squares optimization problem

$$H := \hat{e}_W(1_M \odot \mathcal{C}) = \arg \min_{\beta \in \mathbb{R}_{\geq 0}^{q \times T}} \|1_M \odot \mathcal{C} - W\beta\|_2^2 + \|\beta\|_1. \quad (3)$$

569 The negative effects of an overfit model can therefore be optimized away from on the evaluation set.

We therefore generate uniformly random masks $1_{M(p)} \in \mathbb{R}^{S \times T}$ where

$$1_{M(p)}(s, t) \sim \text{Bernoulli}(p).$$

NMF is then performed using the mask $1_{M(p)}$ to get W . The cross-validation error is then

$$\epsilon_q = \frac{1}{R} \sum_{r=1}^R (\|1_{M(p)_r^c} \odot X - W(\hat{e}_W(1_{M(p)_r^c} \odot X))\|_2^2$$

where $1_{M(p)_r^c}$ is the binary complement of $1_{M(p)_r}$ and R is a number of replicates. Theoretically, the optimum number of components is then

$$\hat{q} = \operatorname{arg\,min}_q \epsilon_q.$$

570 *Stabilizing NMF* The NMF program is non-convex, and, empirically, individual replicates will not
 571 converge to the same optima. One solution therefore is to run multiple replicates of the NMF
 572 algorithm and cluster the resulting vectors. This approach raises the questions of how many clusters
 573 to use, and how to deal with stochasticity in the clustering algorithm itself. We address this issue
 574 through the notion of clustering stability (von Luxburg, 2010a).

The clustering stability approach is to generate L replicas of k-cluster partitions $\{C_{kl} : l \in 1 \dots L\}$ and then compute the average dissimilarity between clusterings

$$\xi_k = \frac{2}{L(L-1)} \sum_{l=1}^L \sum_{l'=1}^l d(C_{kl}, C_{kl'}).$$

Then, the optimum number of clusters is

$$\hat{k} = \arg \min_k \xi_k.$$

575 A review of this approach is found in von Luxburg (2010b). Intuitively, archetype vectors that cluster
 576 together frequently over clustering replicates indicate the presence of a stable clustering. For d , we
 577 utilize the adjusted Rand Index - a simple dissimilarity measure between clusterings. Note that we
 578 expect to select slightly more than the q components suggested by cross-validation, since archetype
 579 vectors which appear in one NMF replicate generally should appear in others. We then select the q
 580 clusters with the most archetype vectors - the most stable NMF results - and take the median of each
 581 cluster to create a sparse representative archetype Kotliar et al. (2019); Wu et al. (2016). We then find
 582 the according H using Program 3. Experimental results for these cross-validation and stability
 583 selection approaches are given in Supplemental Section 7.

7 SUPPLEMENTAL EXPERIMENTS

584 Setting detection threshold τ

585 We give results on the false detection rate at different limits of detection. These conclusively show that
 586 10^{-6} is the good threshold for our normalized data.

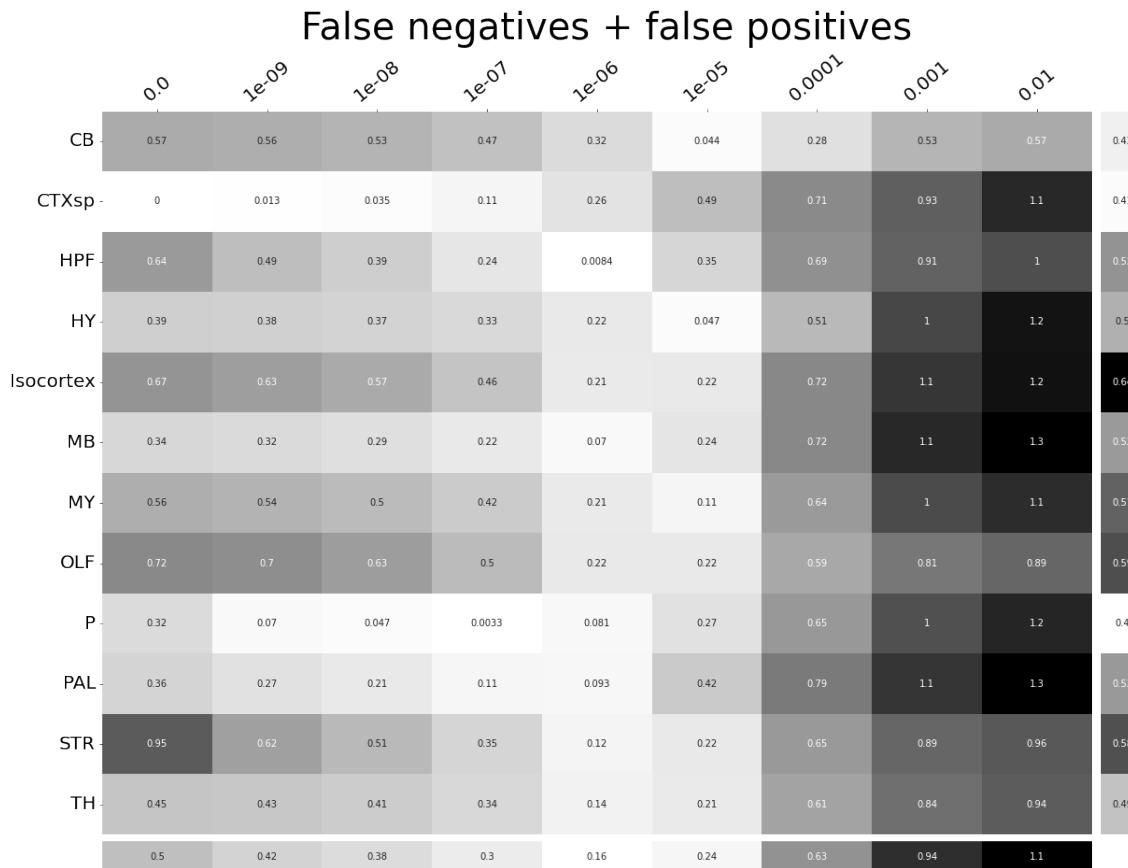


Figure 19: τ at different limits of detection in different major structures. 10^{-6} is the optimal detection threshold.

587 ***Loss subsets***

588 We report model accuracies for our *EL* model by neuron class and structure. These expand upon the
 589 results in Table 5 and give more specific information about the quality of our estimates. CTXsp is
 590 omitted due to the small evaluation set.



Figure 20: Weighted loss for Cre-leaf combinations in CB. Missing values are omitted. Row and column averages are also plotted.

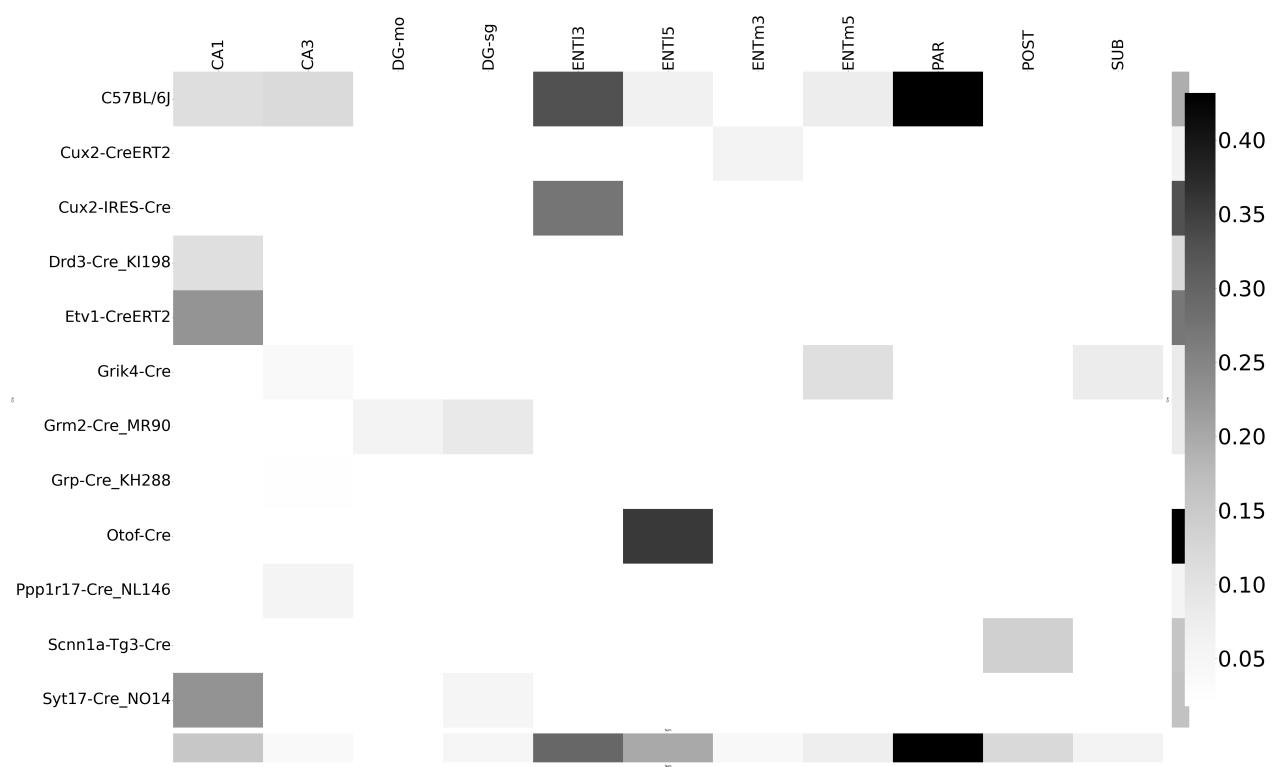


Figure 21: Weighted loss for Cre-leaf combinations in HPF. Missing values are omitted. Row and column averages are also plotted.

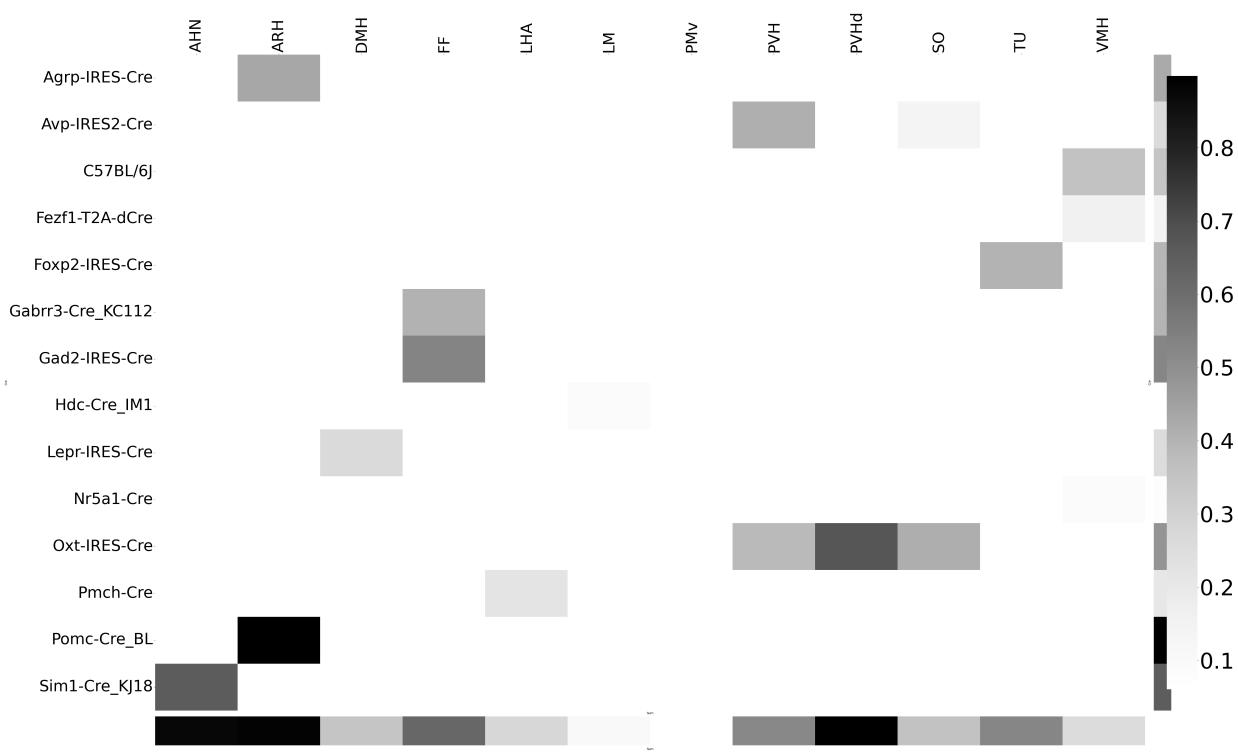


Figure 22: Weighted loss for Cre-leaf combinations in HY. Missing values are omitted. Row and column averages are also plotted.

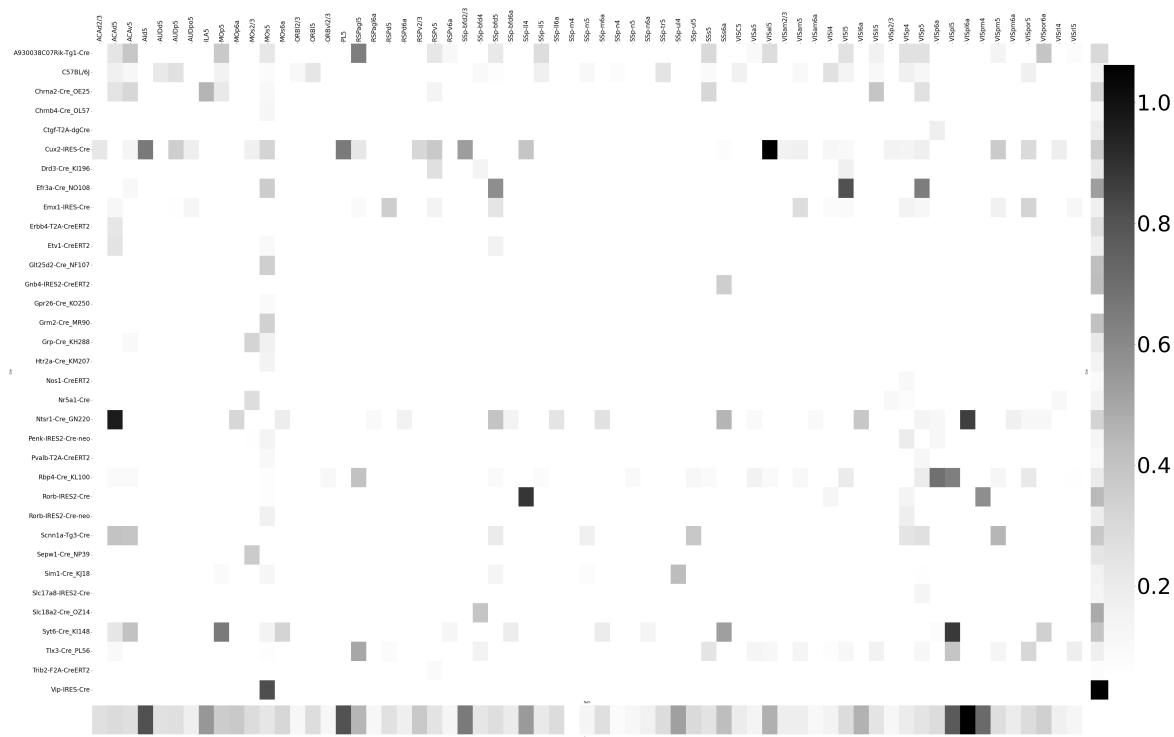


Figure 23: Weighted loss for Cre-leaf combinations in Isocortex. Missing values are omitted. Row and column averages are also plotted.



Figure 24: Weighted loss for Cre-leaf combinations in MB. Missing values are omitted. Row and column averages are also plotted.

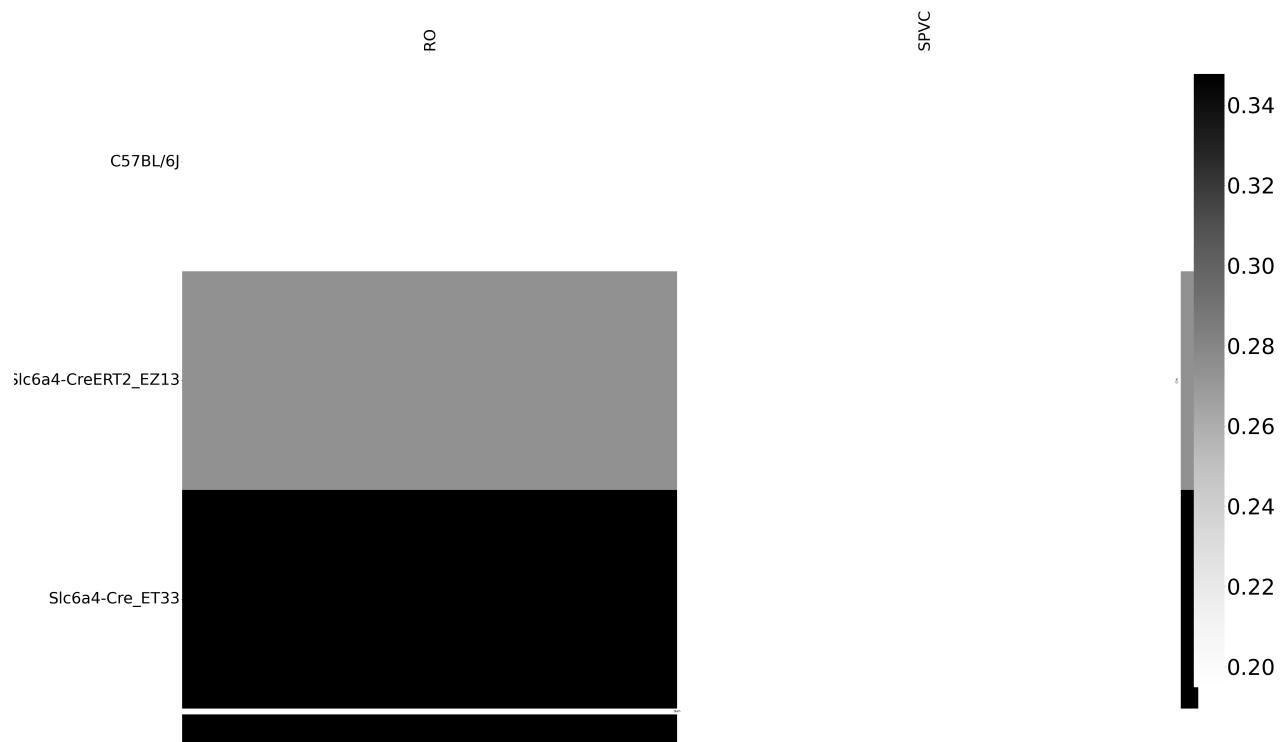


Figure 25: Weighted loss for Cre-leaf combinations in MY. Missing values are omitted. Row and column averages are also plotted.

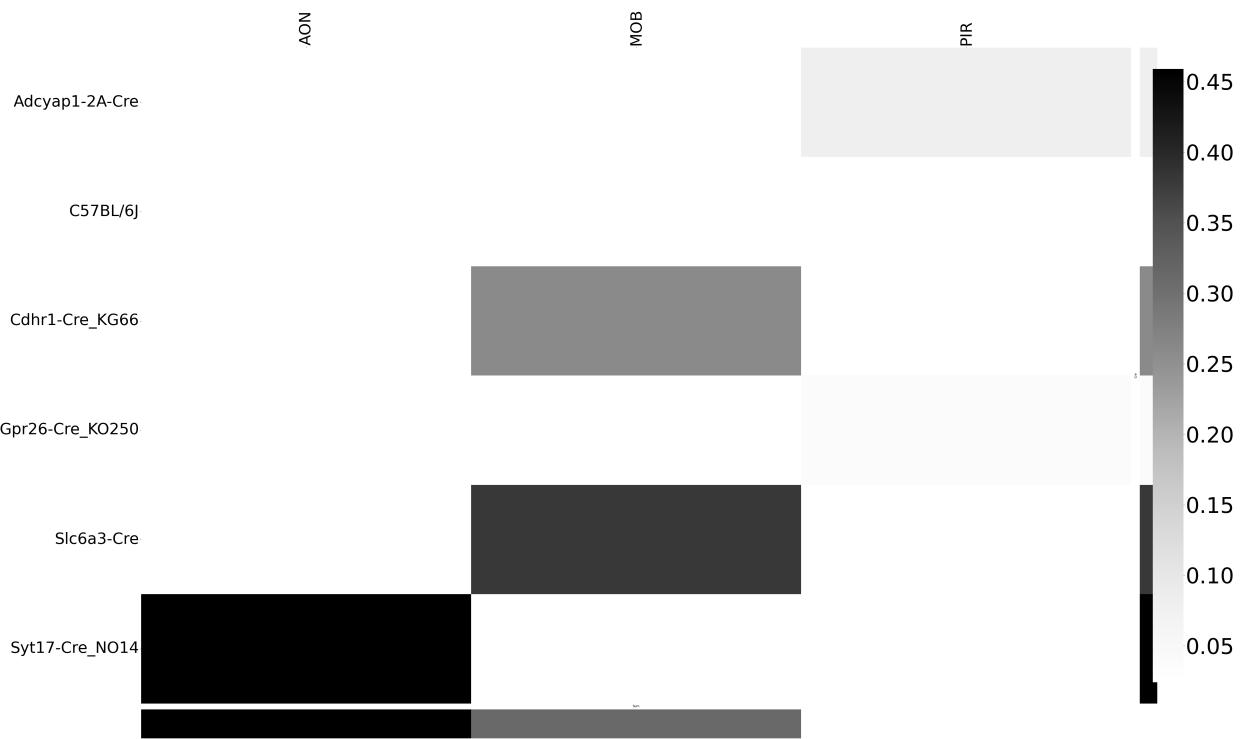


Figure 26: Weighted loss for Cre-leaf combinations in OLF. Missing values are omitted. Row and column averages are also plotted.

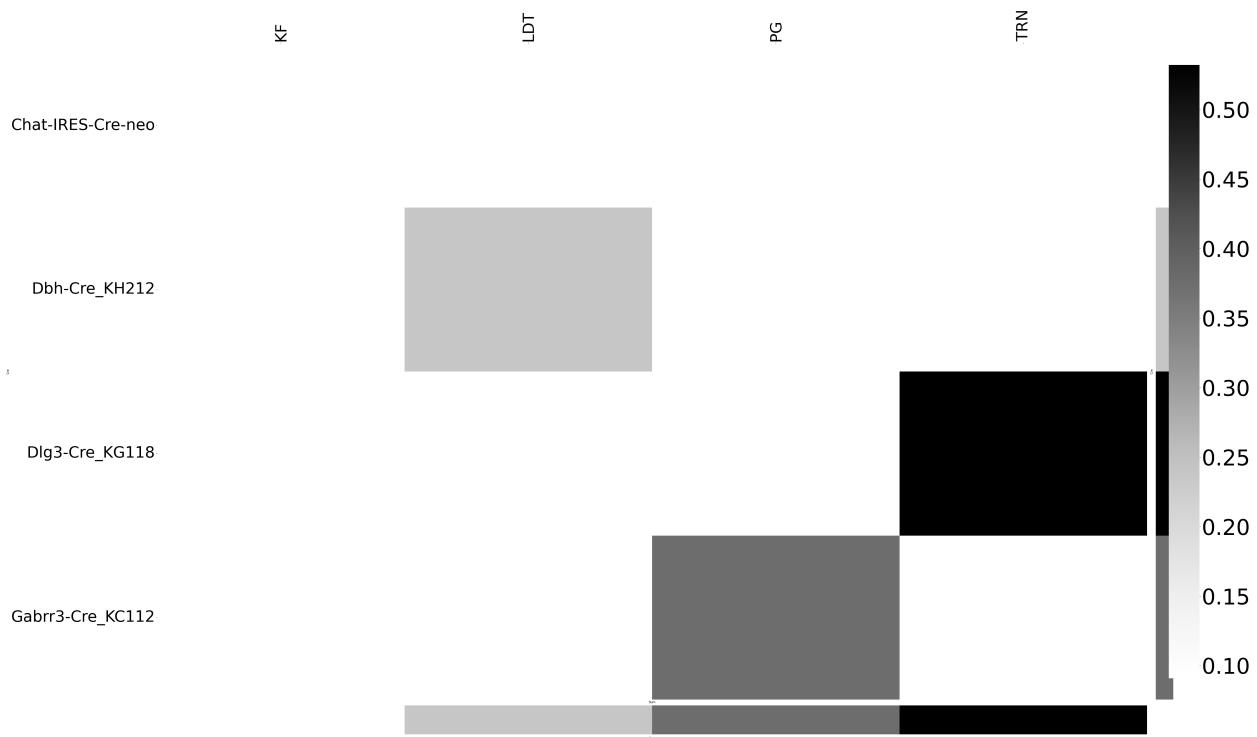


Figure 27: Weighted loss for Cre-leaf combinations in P. Missing values are omitted. Row and column averages are also plotted.

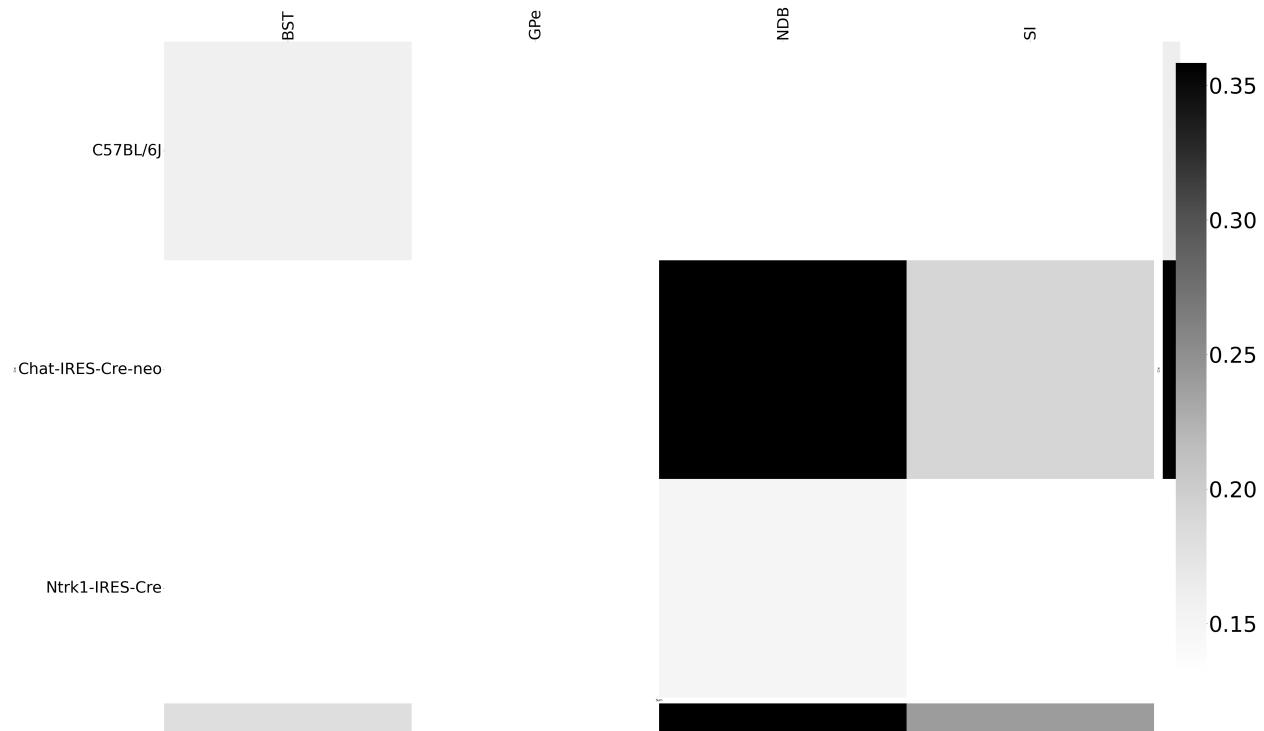


Figure 28: Weighted loss for Cre-leaf combinations in PAL. Missing values are omitted. Row and column averages are also plotted.

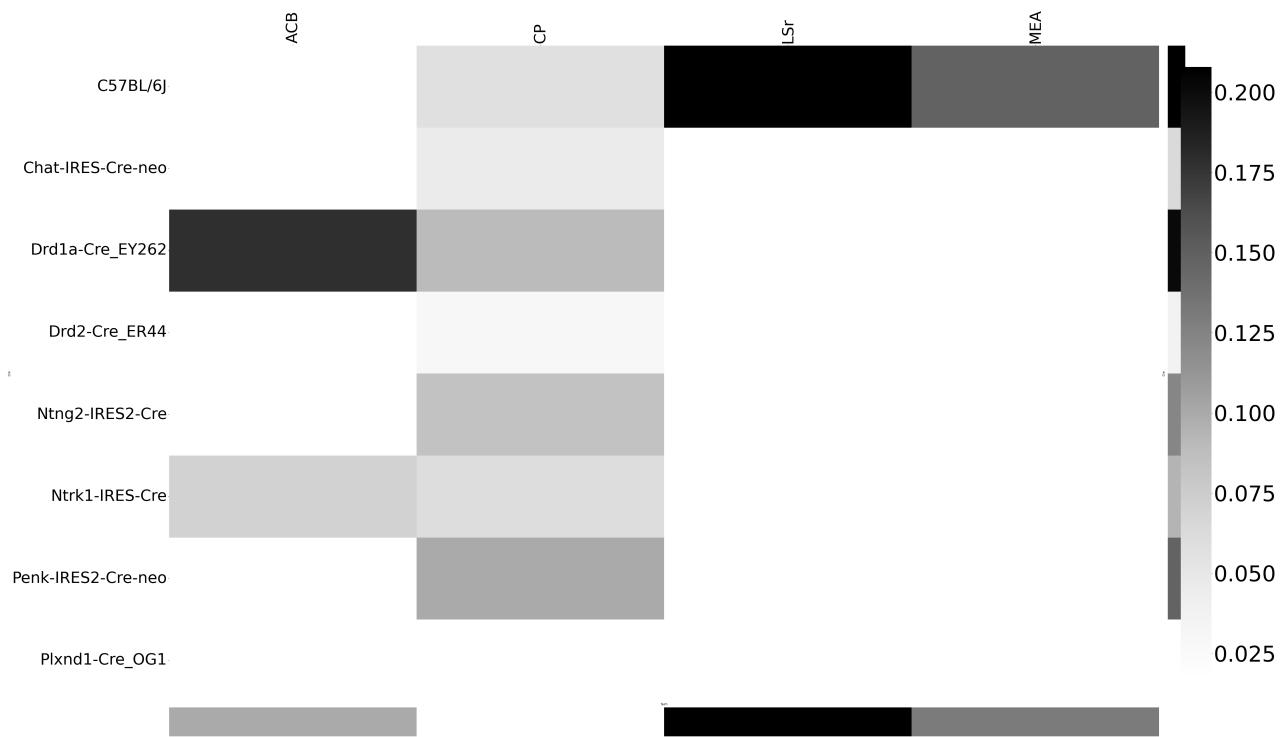


Figure 29: Weighted loss for Cre-leaf combinations in STR. Missing values are omitted. Row and column averages are also plotted.

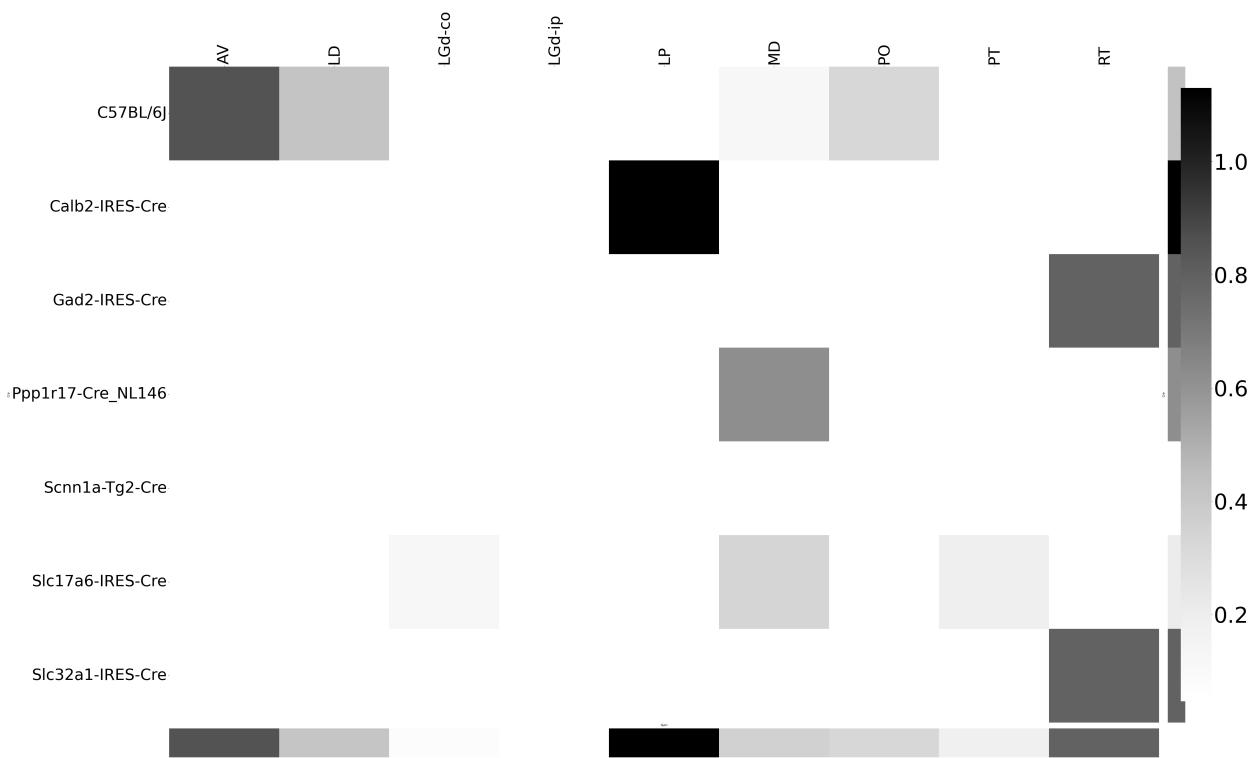


Figure 30: Weighted loss for Cre-leaf combinations in TH. Missing values are omitted. Row and column averages are also plotted.

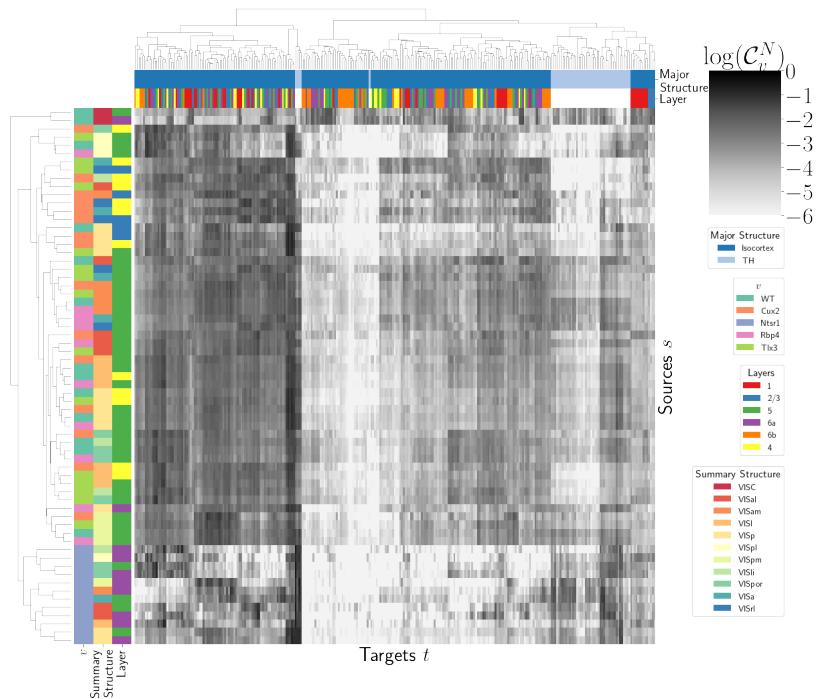


Figure 31: ?? Hierarchical clustering of connectivity strengths from visual cortex cell-types to cortical and thalamic targets. Cre-line, summary structure, and layer are labelled on the sources. Major brain division and layer are labelled on the targets.

591 **Cell-type specificity**

592 We performed agglomerative heirarchical clustering using the default method in Seaborn (?) to
 593 investigate shared projection patterns across Cre-lines. This showed clustering of Ntsr1 projections to
 594 Thalymic nucelii.

595 **Matrix Factorization**

596 We give additional results on the generation of the archetypal connectome latent variables. These
 597 consist of cross-validation selection of q , the number of latent components, stability analysis, and
 598 visualization of the reconstructed wild-type connectivity.

599 *Cross-validation* We set $\alpha = 0.002$ and run Program 2 on \mathcal{C}_{wt} . We use a random mask with $p = .3$ to
 600 evaluate prediction accuracy of models trained on the unmasked data on the masked data. To
 601 account for stochasticity in the NMF algorithm, we run $R = 8$ replicates at each potential dimension q .
 602 This selects $\hat{q} = 60$. However, the low decrease in error at higher values of q and need for brevity in our
 603 figures motivated us to choose $q = 15$ for the purposes of display.

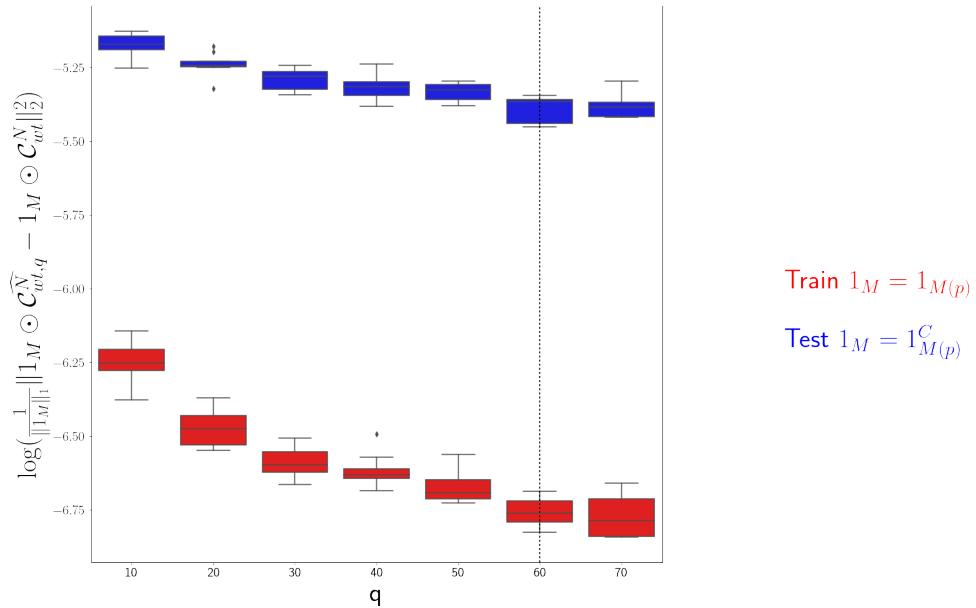


Figure 32: Train and test error using NMF decomposition.

604 *Stability* To address the instability of the NMF algorithm in identifying components, we k -means
 605 cluster components over $R = 10$ replicates with $k \in \{10, 15, 20, 25, 30\}$. Since the clustering is itself
 606 unstable, we repeat the clustering 25 times and select the k with the largest Rand index, a standard
 607 method of clustering stability (??).

q	10.000000	20.000000	30.000000	40.000000	50.000000
Rand index	0.772544	0.844981	0.932957	0.929827	0.885862

609 Since k -means is most stable at $k = 30$, we cluster the $qR = 150$ components into 30 clusters and
 610 select the 15 clusters appearing in the most replicates.

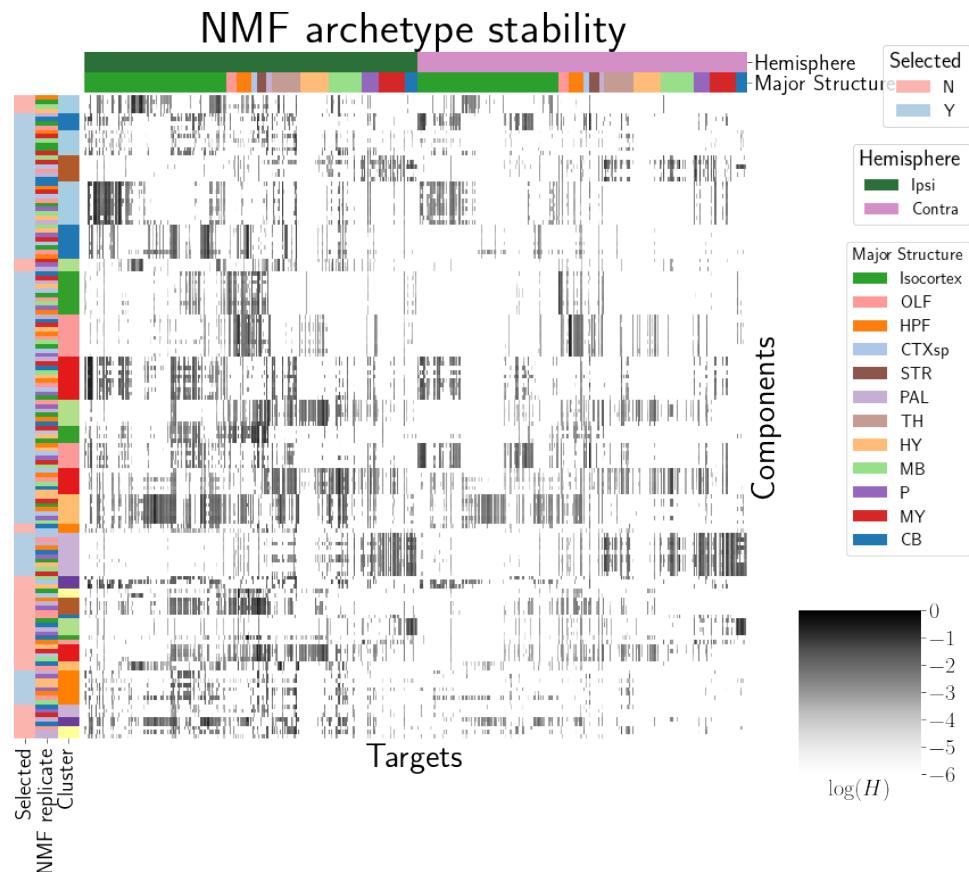


Figure 33: Stability of NMF results across replicates. Replicate and NMF component are shown on rows. Components that are in the top 15 are also indicated.

611 We plot the medians of these components in Figure 4a and in the main text. These are the
612 connectivity archetypes. We then fit a non-negative least squares (the second step in the standard
613 NMF optimization algorithm) to determine W (?)

614 *Association with Cre-line* Finally, we show the association of our learned archetypes with projections
 615 from sources with injection centroids from the Ntsr1, Cux2, Rbp4, and Tlx3 Cre-lines. While we make
 616 no statistical claims on these associations, the distribution of cosine similarities of sources from each
 617 of the Cre-line lines shows an association of learned archetypes with Cre-line.

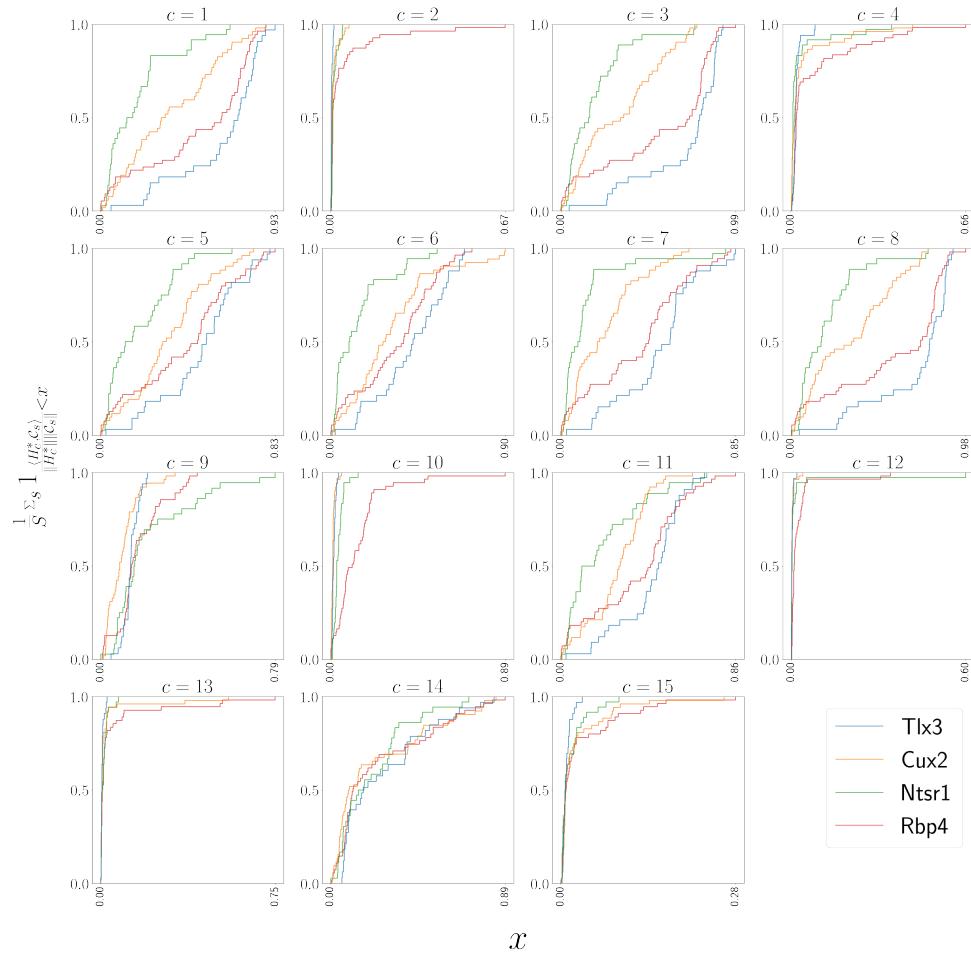


Figure 34: Empirical cumulative distributions of cosine similarities between source structures and connectivity components for four different Cre-lines.

8 COMPETING INTERESTS

⁶¹⁸ This is an optional section. If you declared a conflict of interest when you submitted your manuscript,
⁶¹⁹ please use this space to provide details about this conflict.

620

621 **REFERENCES**

622

623 Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix
 624 factorization. *Proc. Natl. Acad. Sci. U. S. A.*, 101(12), 4164–4169.

625 Chamberlin, N. L., Du, B., de Lacalle, S., & Saper, C. B. (1998). Recombinant adeno-associated virus vector: use for
 626 transgene expression and anterograde tract tracing in the CNS. *Brain Res.*, 793(1-2), 169–175.

627 Daigle, T. L., Madisen, L., Hage, T. A., Valley, M. T., Knoblich, U., Larsen, R. S., ... Zeng, H. (2018). A suite of transgenic
 628 driver and reporter mouse lines with enhanced Brain-Cell-Type targeting and functionality. *Cell*, 174(2), 465–480.e22.

629 Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *SSO Schweiz. Monatsschr.
 630 Zahnheilkd.*, 11(2), 89–121.

631 Gao, Y., Zhang, X., Wang, S., & Zou, G. (2016). Model averaging based on leave-subject-out cross-validation. *J. Econom.*,
 632 192(1), 139–151.

633 Harris, J. A., Mihalas, S., Hirokawa, K. E., Whitesell, J. D., Choi, H., Bernard, A., ... Zeng, H. (2019). Hierarchical
 634 organization of cortical and thalamic connectivity. *Nature*, 575(7781), 195–202.

635 Harris, J. A., Oh, S. W., & Zeng, H. (2012). Adeno-associated viral vectors for anterograde axonal tracing with fluorescent
 636 proteins in nontransgenic and cre driver mice. *Curr. Protoc. Neurosci., Chapter 1, Unit 1.20.1–18*.

637 Harris, K. D., Mihalas, S., & Shea-Brown, E. (2016). Nonnegative spline regression of incomplete tracing data reveals high
 638 resolution neural connectivity.

639 Huang, K. W., Ochandarena, N. E., Philson, A. C., Hyun, M., Birnbaum, J. E., Cicconet, M., & Sabatini, B. L. (2019).
 640 Molecular and anatomical organization of the dorsal raphe nucleus. *Elife*, 8.

641 Jeong, M., Kim, Y., Kim, J., Ferrante, D. D., Mitra, P. P., Osten, P., & Kim, D. (2016). Comparative three-dimensional
 642 connectome map of motor cortical projections in the mouse brain. *Sci. Rep.*, 6, 20072.

643 Knox, J. E., Harris, K. D., Graddis, N., Whitesell, J. D., Zeng, H., Harris, J. A., ... Mihalas, S. (2019). High-resolution
 644 data-driven model of the mouse connectome. *Netw Neurosci*, 3(1), 217–236.

- 645 Kotliar, D., Veres, A., Nagy, M. A., Tabrizi, S., Hodis, E., Melton, D. A., & Sabeti, P. C. (2019). Identifying gene expression
646 programs of cell-type identity and cellular activity with single-cell RNA-Seq. *Elife*, 8.
- 647 Kuan, L., Li, Y., Lau, C., Feng, D., Bernard, A., Sunkin, S. M., ... Ng, L. (2015). Neuroinformatics of the allen mouse brain
648 connectivity atlas. *Methods*, 73, 4–17.
- 649 Li, X., Yu, B., Sun, Q., Zhang, Y., Ren, M., Zhang, X., ... Qiu, Z. (2018). Generation of a whole-brain atlas for the
650 cholinergic system and mesoscopic projectome analysis of basal forebrain cholinergic neurons. *Proc. Natl. Acad. Sci.
651 U. S. A.*, 115(2), 415–420.
- 652 Lotfollahi, M., Naghipourfar, M., Theis, F. J., & Alexander Wolf, F. (2019). Conditional out-of-sample generation for
653 unpaired data using trVAE.
- 654 Muzerelle, A., Scotto-Lomassese, S., Bernard, J. F., Soiza-Reilly, M., & Gaspar, P. (2016). Conditional anterograde tracing
655 reveals distinct targeting of individual serotonin cell groups (B5-B9) to the forebrain and brainstem. *Brain Struct.
656 Funct.*, 221(1), 535–561.
- 657 Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., ... Zeng, H. (2014). A mesoscale connectome of the
658 mouse brain. *Nature*, 508(7495), 207–214.
- 659 Perry, P. O. (2009). Cross-Validation for unsupervised learning.
- 660 Ren, J., Friedmann, D., Xiong, J., Liu, C. D., Ferguson, B. R., Weerakkody, T., ... Luo, L. (2018). Anatomically defined and
661 functionally distinct dorsal raphe serotonin sub-systems. *Cell*, 175(2), 472–487.e20.
- 662 Ren, J., Isakova, A., Friedmann, D., Zeng, J., Grutzner, S. M., Pun, A., ... Luo, L. (2019). Single-cell transcriptomes and
663 whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *Elife*, 8.
- 664 Saul, L. K., & Roweis, S. T. (2003). Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *J.
665 Mach. Learn. Res.*, 4(Jun), 119–155.
- 666 Servén D., B. C. (n.d.). *pygam: Generalized additive models in python*.
- 667 von Luxburg, U. (2010a). Clustering stability: An overview.
- 668 von Luxburg, U. (2010b). Clustering stability: An overview.

- 669 Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., ... Ng, L. (2020). The allen mouse brain common coordinate
670 framework: A 3D reference atlas. *Cell*, 181(4), 936–953.e20.
- 671 Watson, C., Paxinos, G., & Puelles, L. (2012). The mouse nervous system..
- 672 Wu, S., Joseph, A., Hammonds, A. S., Celtniker, S. E., Yu, B., & Frise, E. (2016). Stability-driven nonnegative matrix
673 factorization to interpret spatial gene expression and build local gene networks. *Proc. Natl. Acad. Sci. U. S. A.*, 113(16),
674 4290–4295.
- 675 Zaborszky, L., Csordas, A., Mosca, K., Kim, J., Gielow, M. R., Vadász, C., & Nadasdy, Z. (2015). Neurons in the basal
676 forebrain project to the cortex in a complex topographic organization that reflects corticocortical connectivity
677 patterns: an experimental study based on retrograde tracing and 3D reconstruction. *Cereb. Cortex*, 25(1), 118–137.

9 TECHNICAL TERMS

678 **Technical Term** a key term that is mentioned in an NETN article and whose usage and definition may
679 not be familiar across the broad readership of the journal.

680 **Cre-line** Refers to the combination of cre-recombinase expression in transgenic mouse and
681 cre-induced promotion in the vector that induces labelling of cell-class specific projection.

682 **Cell class** The projecting neurons targeted by a particular cre-line

683 **Structural connectivities** connectivity between structures

684 **Voxel** A $100\mu m$ cube of brain.

685 **Structural connection tensor** Connectivities between structures given a neuron class

686 **dictionary-learning** A family of algorithms for finding low-dimensional data representations.

687 **Shape constrained estimator** A statistical estimator that fits a function of a particular shape (e.g.
688 monotonic increasing, convex).

689 **Nadaraya-Watson** A simple smoothing estimator.

690 **Connectivity archetypes** Typical connectivity patterns

691 **Expected loss** Our new estimator that weights different features by their estimated predictive
692 power.