

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 (MCA) consists of anterograde tracing experiments
10 targeting diverse structures and classes of projecting neurons. Beyond regional anterograde tracing
11 done in C57BL/6 wild type mice, a large fraction of experiments are performed using transgenic
12 Cre-lines. This allows access to cell-class specific whole brain connectivity information, with class
13 defined by the transgenic lines. However, even though the number of experiments is large, it does not
14 come close to covering all existing cell classes in every area where they exist. Here, we study how
15 much we can fill in these gaps and estimate the cell-class specific connectivity function given the
16 simplifying assumptions that nearby voxels have smoothly varying projections, but that these
17 projection tensors 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 including those from one region
44 to another form the architecture underlying this capability. These connectivities vary by neuron type,
45 as well as source (cell body) location and target (axonal projection) structures. Thus, characterization
46 of the relationship between neuron type and source and target structure is important for
47 understanding the overall 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 (Gămănuț et al., 2018; K. D. Harris, Mihalas, &
59 Shea-Brown, 2016; Knox et al., 2019; Oh et al., 2014). Of these, Oh et al. (2014) and Knox et al. (2019)
60 provide a model for **regionalized connectivities**, which are voxel connectivities integrated by region.
61 The value of these models is that they provide some improvement over simply averaging the
62 projection signals of injections in a given region. However, these previous works only model
63 connectivities observed in wild type mice which are suboptimally suited to assessment of cell-type
64 specific connectivity compared with fluorescence from Cre-recombinase induced eGFP expression in
65 cell-types specified by the combination of transgenic mouse strain and transgene promoter
66 (J. A. Harris et al., 2019). We generally refer to sets of targeted eGFP-expressing cells in tracing
67 experiments as a **cell class** since they may contain multiple types. For example, use of both wild-type
68 and transgenic mice would give rise to cell-class specific experiments, albeit with different yet
69 perhaps overlapping classes of cells.

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

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

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

2 METHODS

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

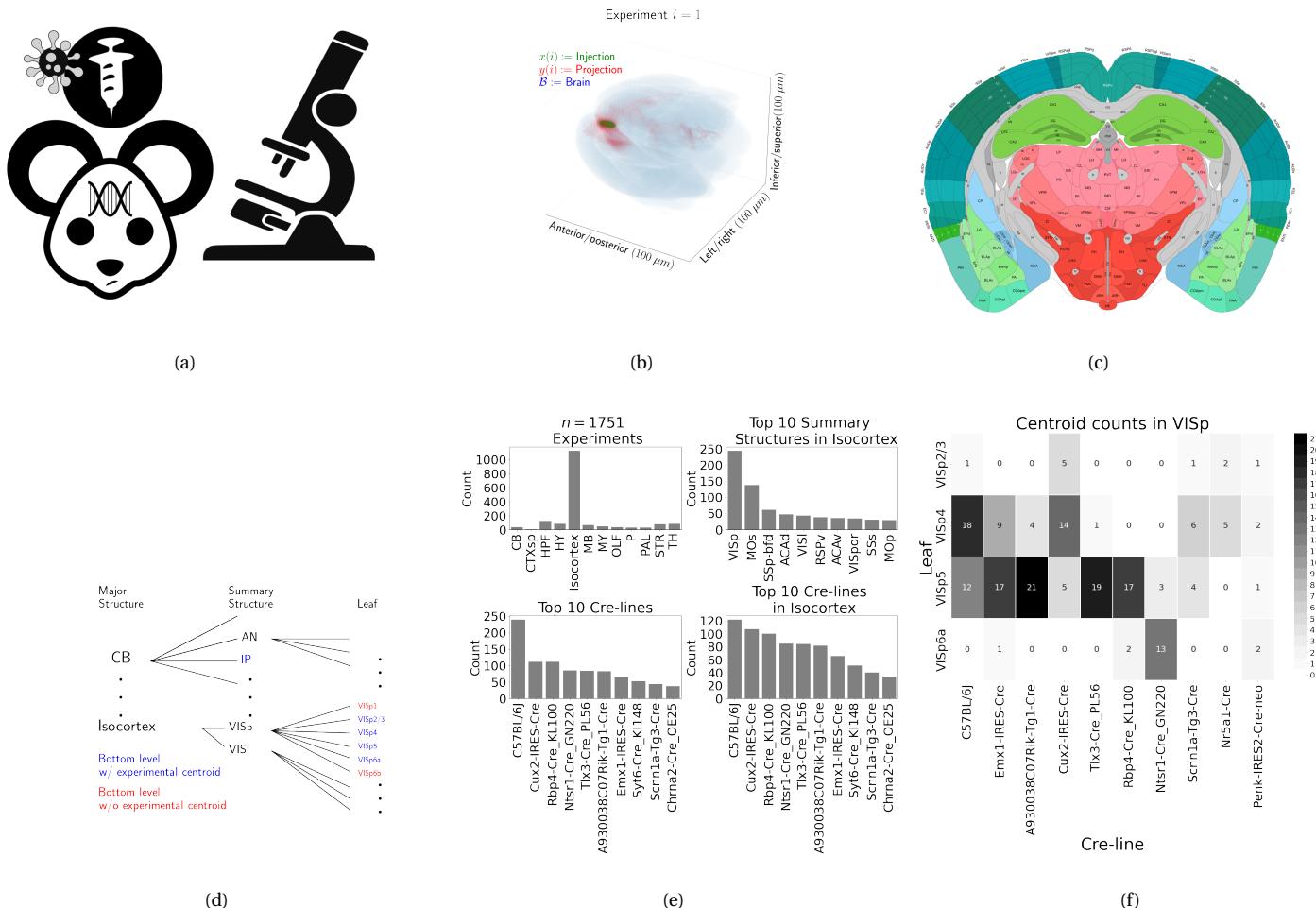


Figure 1: Experimental setting. 1a For each experiment, a Cre-dependent eGFP-expressing transgene cassette is transduced by stereotaxic injection into a Cre-driver mouse, followed by serial two-photon tomography imaging. 1b An example of the segmentation of projection (targets) and injection (source) 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.

99 Data

100 Our dataset \mathcal{D} consists of $n = 1751$ publicly available murine brain viral tracing experiments from the
 101 Allen Mouse Brain Connectivity Atlas. Figure 1a summarizes the experimental process used to
 102 generate this data. In each experiment, a mouse is injected with an adeno-associated virus (AAV)
 103 encoding green fluorescent protein (GFP) into a single location in the brain. Location of fluorescence
 104 is mediated by the location of the injection, the characteristics of the transgene, and the genotype of
 105 the mouse. In particular, Cre-driver or, equivalently, Cre-line mice are engineered to express Cre
 106 under the control of a specific and single gene promoter. This localizes expression of Cre to regions
 107 with certain transcriptomic cell-types signatures. In such Cre-driver mice, we used a double-inverted
 108 floxed AAV to produce eGFP fluorescence that depends on Cre expression in infected cells. To account
 109 for the complex cell-type targeting induced by a particular combination of Cre-driver genotype and
 110 GFP promoter, we refer to the combinations of cell-types targeted by a particular combination of AAV
 111 and Cre-driver mice as cell-classes. For example, we include experiments from Cre-driver lines that
 112 selectively label cell classes located in distinct cortical layers or other nuclei across the whole brain.
 113 For injections in the wild type mice, we used the Synapsin I promoter (Jackson, Dayton, Deverman, &
 114 Klein, 2016; Kügler, Kilic, & Bähr, 2003). For injections into Cre mice, we used the CAG promoter with
 115 a Flex cassette for Cre-mediated recombination control (Saunders, Johnson, & Sabatini, 2012).
 116 Additional details on are given in J. A. Harris et al. (2019).

117 For each experiment, the fluorescent signal imaged after injection is aligned into the Allen
 118 Common Coordinate Framework (CCF) v3, a three-dimensional average template brain that is fully
 119 annotated with regional parcellations Wang et al. (2020). The whole brain imaging and registration
 120 procedures described in detail in Kuan et al. (2015); Oh et al. (2014) produce quantitative metrics of
 121 fluorescence discretized at the $100 \mu\text{m}$ voxel level. Given an experiment, this image was histologically
 122 segmented by an analyst into *injection* and *projection* areas corresponding to areas containing somas,
 123 dendrites and axons or exclusively axons of the transfected neurons. An example of a single
 124 experiment rendered in 3D is given in Figure 1b. Given an experiment i , we represent injections and
 125 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
 126 subset of the $(1.32 \times 0.8 \times 1.04)$ cm rectangular space occupied by the standard voxelized mouse brain.

¹²⁷ 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
¹²⁸ sum of $y(i)$ in each region. A description of these steps is in Supplemental Section 6.

¹²⁹ Our goal is the estimation of **regionalized connectivity** from one region to another. A visual
¹³⁰ depiction of this region parcellation for a two-dimensional slice of the brain is given in Figure 1c. All
¹³¹ structures annotated in the CCF belong to a hierarchically ordered ontology, with different areas of the
¹³² brain are parcellated to differing finer depths within a hierarchical tree. We denote the main levels of
¹³³ interest as major structures, summary structures, and layers. Not every summary structure has a layer
¹³⁴ decomposition within this ontology, so we typically consider the finest possible regionalization - for
¹³⁵ example, layer within the cortex, and summary structure within the thalamus, and denote these
¹³⁶ structures as leafs. As indicated in Figure 1d, the dataset used to generate the connectivity model
¹³⁷ reported in this paper contains certain combinations of region and cell class frequently, and others
¹³⁸ not at all. A summary of the most frequently assayed cell classes and structures is given in Figures 1e
¹³⁹ and 1f. Since users of the connectivity matrices may be interested in particular combinations, or
¹⁴⁰ interested in the amount of data used to generate a particular connectivity estimate, we present this
¹⁴¹ information about all experiments in Supplemental Section 5.

142 ***Modeling Regionalized Connectivity***

We define voxelized cell-class specific connectivity $f : \mathcal{V} \times \mathcal{B} \times \rightarrow \mathbb{R}_{\geq 0}$ 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

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

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

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

143 where l_j and $l_{j'}$ are the locations of source and target voxels, and $|s|$ and $|v|$ are defined to be the
 144 number of voxels in the source and target structure, respectively. Since the normalized strength and
 145 densities are computable from the strength via a fixed normalization, our main statistical goal is to
 146 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
 147 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

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

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

154 available so that its layer-specific behavior is visible, and normalize our data by projection signal in
 155 order to account for differences between cell class. Additional model selection results are given in
 156 Supplemental Section 5 for alternative normalization strategies, and more detail on estimation is
 157 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.

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

¹⁶⁵ predictions for each voxel in a structure, and then sums them together to get the overall connectivity.
¹⁶⁶ However, in contrast to the NW approaches, the EL estimate of the projection vector for a cell-class at
¹⁶⁷ a location weights the average projection of that cell-class in the region containing the location
¹⁶⁸ against the relative locations of all experimental centroids in the region regardless of class. That is,
¹⁶⁹ cell-class and source region combinations with similar average projection vectors will be upweighted
¹⁷⁰ when estimating \hat{f} . Thus, all experiments that are nearby in three-dimensional space can help
¹⁷¹ generate the prediction, even when there are few nearby experiments for the cell-class in question. A
¹⁷² detailed mathematical description of our new estimator is given in Supplemental Section 6.

173 ***Model evaluation***

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

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

We use weighted l_2 -loss to evaluate these predictions.

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

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

186 I_s refers to the set of experiments with centroid in structure s , and I_v refers to the set of experiments
 187 with Cre-line v , so $|I_s \cap I_v|$ is the number of experiments of Cre-line v with injection centroid in
 188 structure s . This is a somewhat different loss from Knox et al. (2019) because of the increased
 189 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
 190 a lower limit of detection and the extra cross-validation step used in the EL model to establish the
 191 relative importance of regionally averaged cell-class projection and injection centroid position are
 192 covered in Supplemental Section 6.

193 ***Connectivity analyses***

194 We examine our connectome estimates with both comparisons to known biology and statistical
 195 decompositions. As an exploratory analysis, we use hierarchical clustering to compare outputs from
 196 connectivities from different Cre-lines. Details of and results from this approach are given in Section
 197 7. We then use non-negative matrix factorization (NMF) to factor the wild-type connectivity matrix
 198 into a small set of underlying components that can be linearly combined to reproduce the observed
 199 long-range wild-type connectivity. Inspired by Mohammadi, Ravindra, Gleich, and Grama (2018), we
 200 refer to these latent coordinates as **connectivity archetypes** since they represent underlying patterns
 201 from which we can reconstruct a broad range of observed connectivities, although we note that the
 202 genomic archetypal analysis in that paper is slightly methodologically distinct.

203 NMF refers to a collection of **dictionary-learning** algorithms for decomposing a
 204 non-negatively-valued matrix such as \mathcal{C} into positively-valued matrices called, by convention,
 205 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 linear statistical model: that the
 206 observed matrix is composed of linear combinations of latent coordinates (Devarajan, 2008). Unlike
 207 PCA, NMF specifically accounts for the fact that data are all in the positive orthant, and it is more
 208 stable and interpretable in assays of complex biological systems than hierarchical clustering (Brunet,
 209 Tamayo, Golub, & Mesirov, 2004). The choice of matrix factorization method reflects particular
 210 scientific subquestions and probabilistic interpretations, and the matrix H may be used to identify latent
 211 structures with interpretable biological meaning.

212 Our application of NMF to decompose the estimated long-range connectivity is some independent
 213 interest, since we ignore connections between source and target regions less than $1500 \mu m$ apart. This
 214 is because short-range projections resulting from diffusion and traveling fibers dominate the matrices
 215 $\hat{\mathcal{C}}$. Our NMF algorithm solves the following optimization problem

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

216 We set $\lambda = 0.002$ to encourage sparser and therefore more interpretable components. We use
 217 unsupervised cross-validation to determine an optimum q , and show the top 15 stable components
 218 (Perry, 2009). Since the NMF objective is difficult to optimize and sensitive to initialization, we follow

219 up with a stability analysis via clustering the resultant H from multiple replicates. The medians of the
220 component clusters appearing frequently across NMF replicates are selected as **connectivity**
221 **archetypes**. Details of these approaches are given in Supplementary Sections 6 and 7.

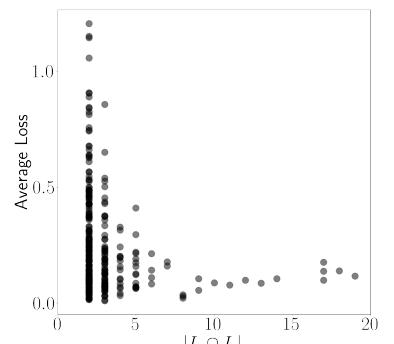
3 RESULTS

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

229 ***Model evaluation***

230 Our EL model generally performs better than the other estimators that we consider. Table 2a contains
231 weighted losses from leave-one-out cross-validation of candidate models, such as the NW Major-WT
232 model from Knox et al. (2019). The EL model combines the good performance of class-specific
233 models like NW Leaf-Cre in regions like Isocortex with the good performance of class-agnostic models
234 in regions like Thalamus. Additional information on model evaluation, including class and structure-
235 specific performance, is given in Appendix 5. In particular, Supplementary Table 4 contains the sizes
236 of these evaluation sets in each major structure, and Supplementary Section 7 contains the structure-
237 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. The dataset variable \mathcal{D} indicates the set of experiments used to model a given connectivity. For example, $I_c \cap I_L$ means only experiments with a given Cre-line in a given leaf are used to model connectivity for the corresponding cell-class in that leaf, while I_L means that all experiments in that leaf are used. $I_{wt} \cap I_M$ means all wild-type experiments in the major structure are used - this was the model in Knox et al. (2019)

238 **Connectivities**

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

244 *Overall connectivity* Several known biological projection patterns are evident in the wild-type
 245 connectivity matrix \mathcal{C}_{wt} shown in Figure 2a. This matrix shows connectivity from leaf sources to leaf
 246 targets. Large-scale patterns like intraareal connectivities and ipsilateral connections between cortex

247 and thalamus are clear, as in previous estimates in J. A. Harris et al. (2019); Knox et al. (2019); Oh et al.
248 (2014). However, the layer-specific targeting of the different Cre-lines enables our estimated wild-type
249 connectivities to display heterogeneity at the layer level. This contrasts the model in Knox et al.
250 (2019), which is denominated as the NW Major-WT model whose accuracy is evaluated in Table 2a.
251 For comparison, we also plot averages over component layers weighted by layer size projections
252 between summary-structure sources and targets in the cortex in Figure 2b. Importantly, as shown in
253 Table 2a this finer spatial resolution corresponds to the increased accuracy of our EL model over the
254 NW Major-WT model.

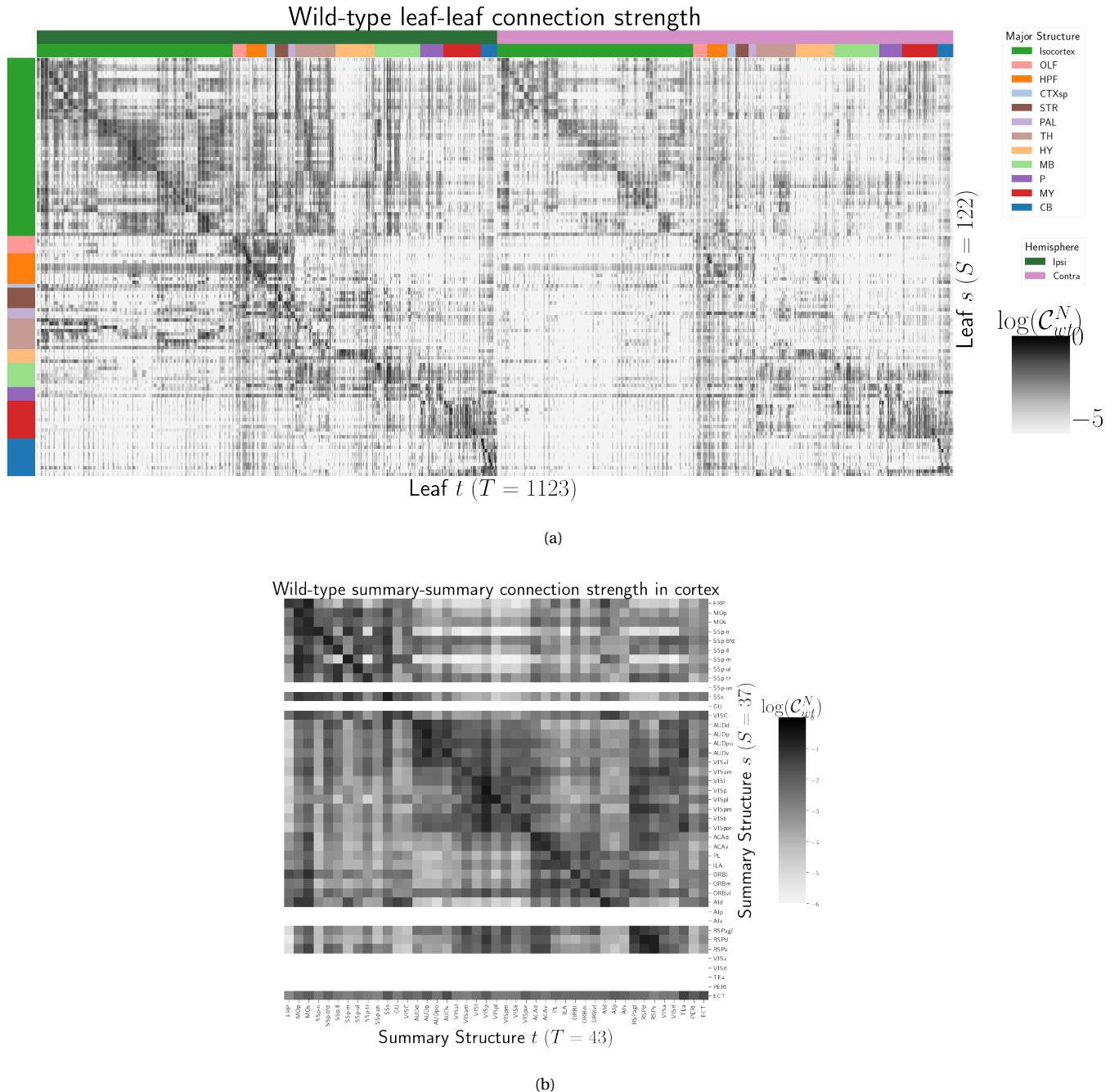


Figure 2: Wild-type connectivities. 2a Log wild-type leaf-to-leaf connectivity matrix $\log \mathcal{C}(s, t, v_{wt})$. Sources and target major brain division structure are shown. 2b Log wild-type intracortical connectivity matrix at the summary structure level. Summary structures without an injection centroid are left blank.

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

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

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

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

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

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

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

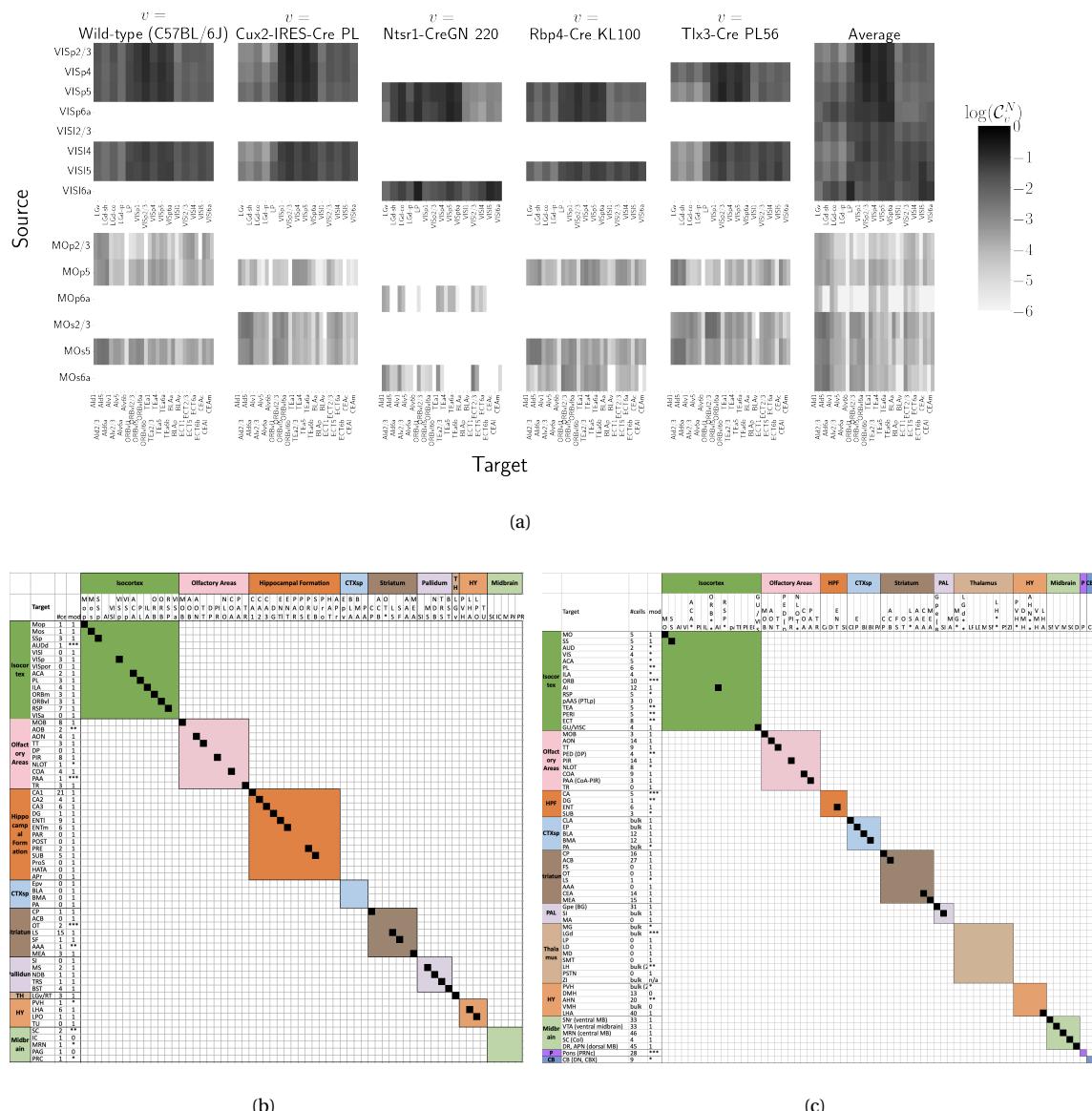


Figure 3: Cell-class specificity. 3a Selected cell-class and layer specific connectivities from two visual and two motor areas. Sources without an injection of a Cre-line are not estimated due to lack of data. 3b Targets reported for cholinergic cells in MS/NDB in Li et al. (2018) and above the 90th 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.

327 Connectivity Analyses

328 While the manual analysis in Section 3 is valuable for validation, scaling our interpretation of our
329 connectivity estimates motivated us to apply dimension reduction methods to our connectivity
330 estimates and understand whether the learned structure agrees with the expected biology. For
331 example, Supplemental Section 7 shows a collection of connectivity strengths generated using
332 Cre-specific models for wild-type, Cux2, Ntsr1, Rbp4, and Tlx3 cre-lines from VIS areas at leaf level in
333 the cortex to cortical and thalamic nuclei. Sorting source and target structure/cell-class combinations
334 hierarchical clustering shows, for example, that layer 6-specific Ntsr1 Cre-lines source regions cluster
335 together. This makes sense, since layer 6-specific Ntsr1 Cre-line distinctly projects to thalamic nuclei,
336 regardless of source summary structure, and in contrast with the tendency of other cell-classes to
337 project to nearby regions within the cortex.

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

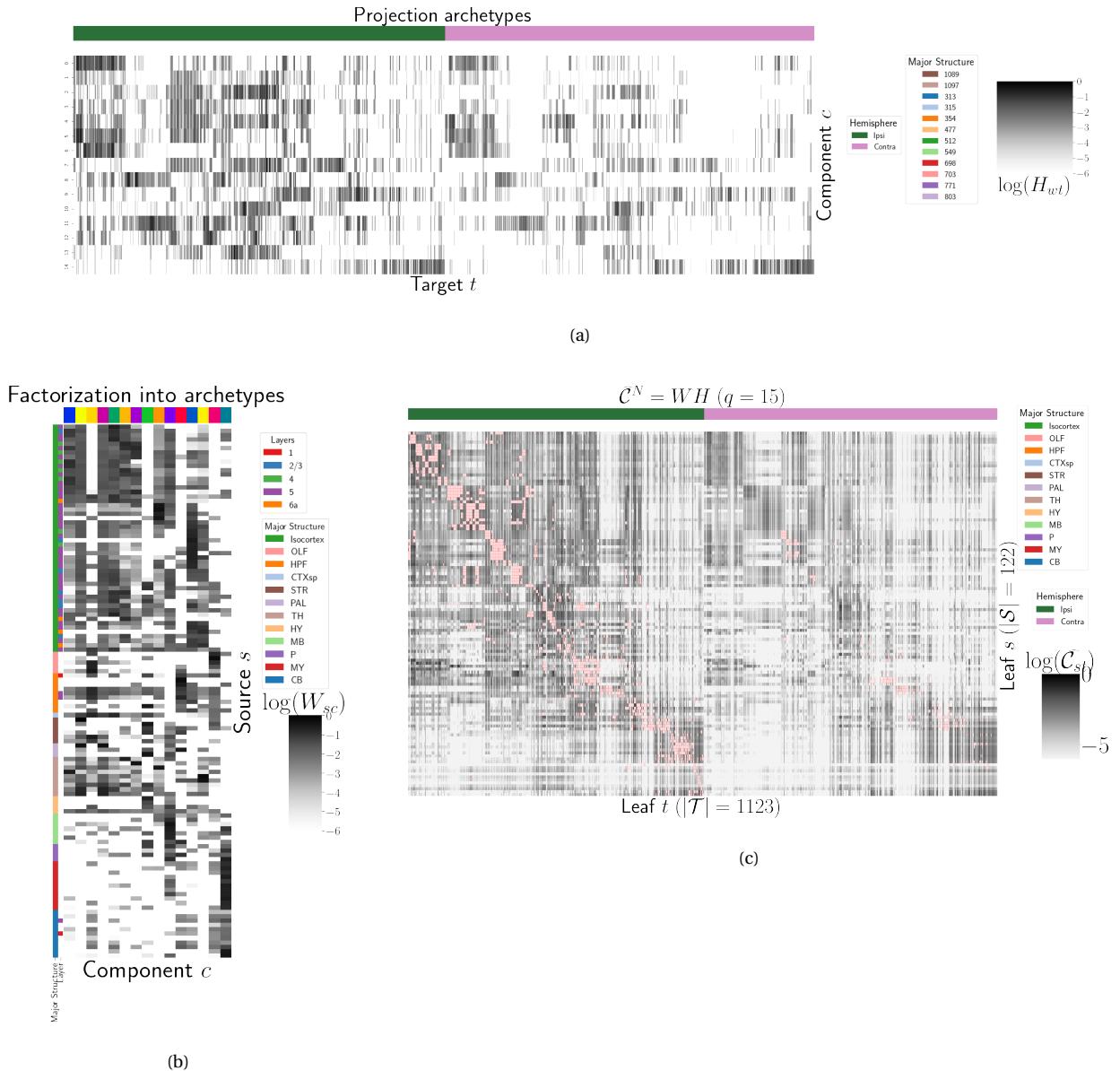


Figure 4: Non-negative matrix factorization results $\mathcal{C}_{wt}^N = WH$ for $q = 15$ components. 4a Latent space coordinates H of \mathcal{C} . Target major structure and hemisphere are plotted. 4b Loading matrix W . Source major structure and layer are plotted. 4c Reconstruction of the normalized 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

353 The model presented here is among the first cell-type specific whole brain projectome models for a
 354 mammalian species, and it opens the door for a large number of models linking brain structure to
 355 computational architectures. Overall, we find expected targets, based on our anatomical expertise
 356 and published reports, but underscore that the core utility of this bulk connectivity analysis is not
 357 only in validation of existing connection patterns, but also in identification of new ones. We note that
 358 although the concordance appeared stronger for the cholinergic cells than the serotonergic cells, any
 359 differences might still be explained by the lack of high quality “ground-truth” datasets to validate
 360 these Cre-line connectome models. It is important to note several limitations of the current analyses.
 361 Short-range connections can be affected by saturating signals near injection sites, as well as the
 362 segmentation algorithm capturing dendrites as well as axons. Furthermore, larger numbers of single
 363 cell reconstructions that saturate all possible projection types would be a better gold standard than
 364 the small number of cells reported here. Future iterations of connectome models may also take into
 365 account single cell axon projection data, or synthesize with retrograde tracing experiments.

366 The Nadaraya-Watson estimator using the cell-type space based on similarities of projections, and
 367 theoretical justification of the use of an intermediate shape-constrained estimator, provides an
 368 empirically useful new tool for categorical modeling. Ours is not the first cross-validation based
 369 model averaging method Gao, Zhang, Wang, and Zou (2016), but our use of shape-constrained
 370 estimator in target-encoded feature space is novel and fundamentally different from
 371 Nadaraya-Watson estimators that use an optimization method for selecting the weights (Saul &
 372 Roweis, 2003). The properties of this estimator and its relation to estimators fit using an optimization
 373 algorithm are therefore a possible future avenue of research (Groeneboom & Jongbloed, 2018; Salha &
 374 El Shekh Ahmed, n.d.). Since the impact of the virus depends on the particular injection region, a
 375 deep model such as Lotfollahi, Naghipourfar, Theis, and Alexander Wolf (2019) could be appropriate,
 376 provided enough data was available, but our sample size seems too low to utilize a fixed or mixed
 377 effect model generative model. In a sense both the non-negative least squares Oh et al. (2014) and NW
 378 models can be thought of as improvements over the structure-specific average, and so is also possible
 379 that a yet undeveloped residual-based data-driven blend of these models could provide improved
 380 performance. Finally, we note that a Wasserstein-based measure of injection similarity per structure

381 could naturally combine both the physical simplicity of the centroid model while also incorporating
382 the full distribution of the injection signal.

383 The factorization of the connectivity matrix could also be improved. Non-linear data
384 transformations or matrix decompositions, or tensor factorizations that account for correlations
385 between cell-types could better capture the true nature of latent neural connections (K. D. Harris et
386 al., 2016). Cre- or layer-specific signal recovery as performed here could be used to evaluate a range of
387 matrix decompositions. This could help for example to understand the influence of traveling fibers on
388 the observed connectivity (Llano & Sherman, 2008). From a statistical perspective, stability-based
389 method for establishing archetypal connectivities in NMF is similar to those applied to genomic data
390 Kotliar et al. (2019); Wu et al. (2016). Regardless of statistical approach, as in genomics, latent
391 low-dimensional organization in connectivity should inspire search for similarly parsimonious
392 biological correlates.

ACKNOWLEDGMENTS

³⁹³ We thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement,
³⁹⁴ and support.

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

5 SUPPLEMENTAL INFORMATION

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

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

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

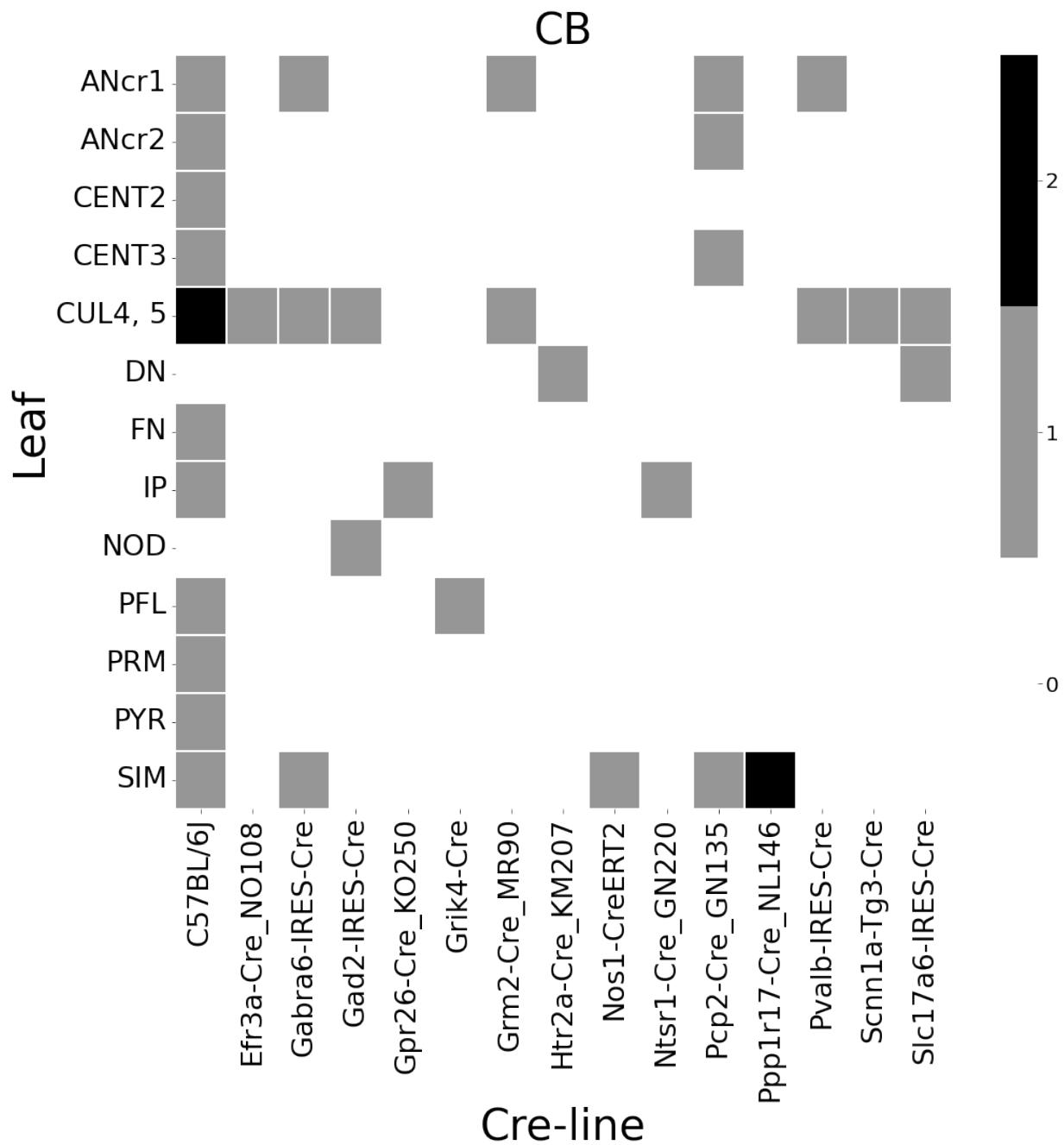


Figure 5: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

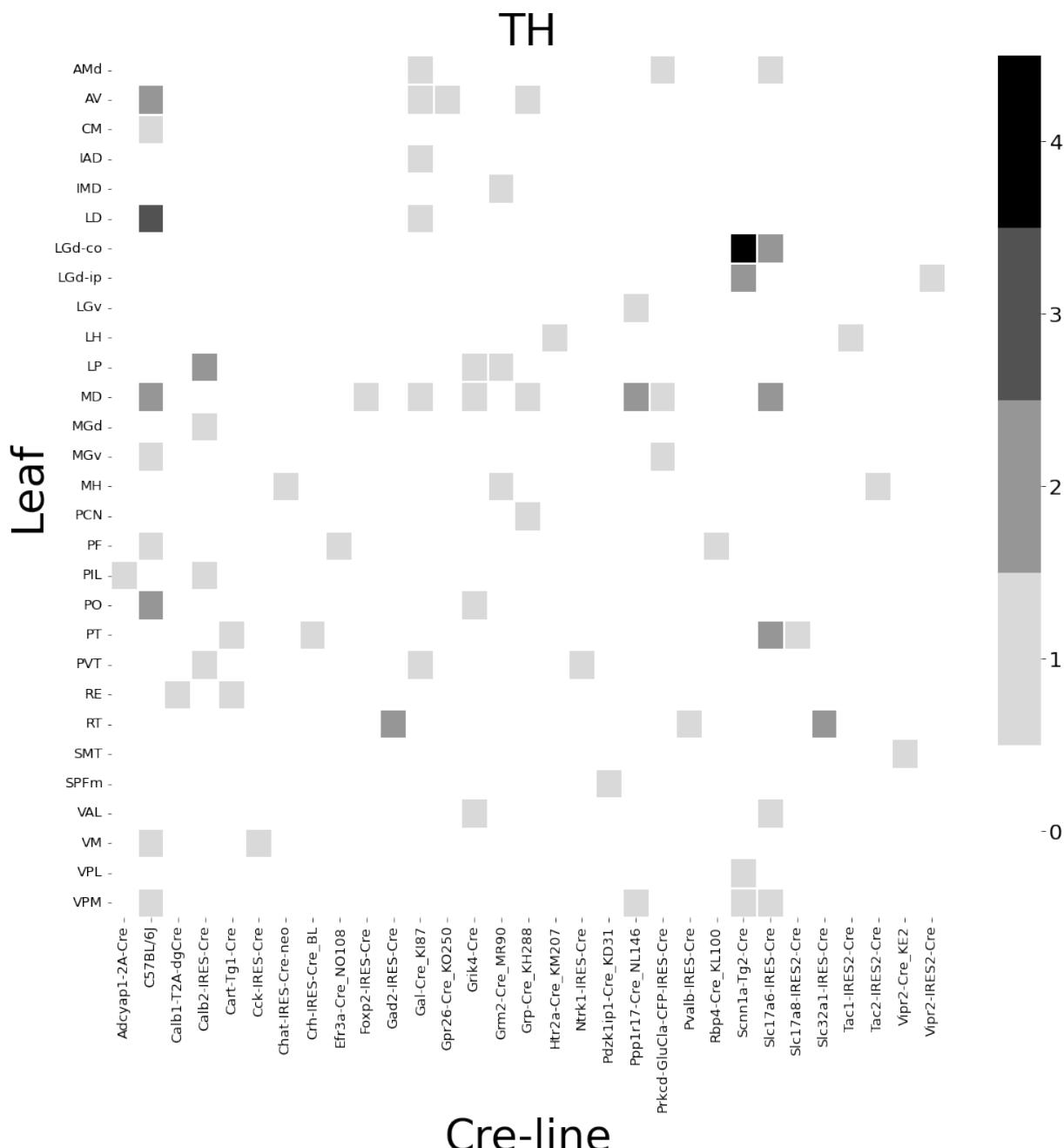


Figure 6: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

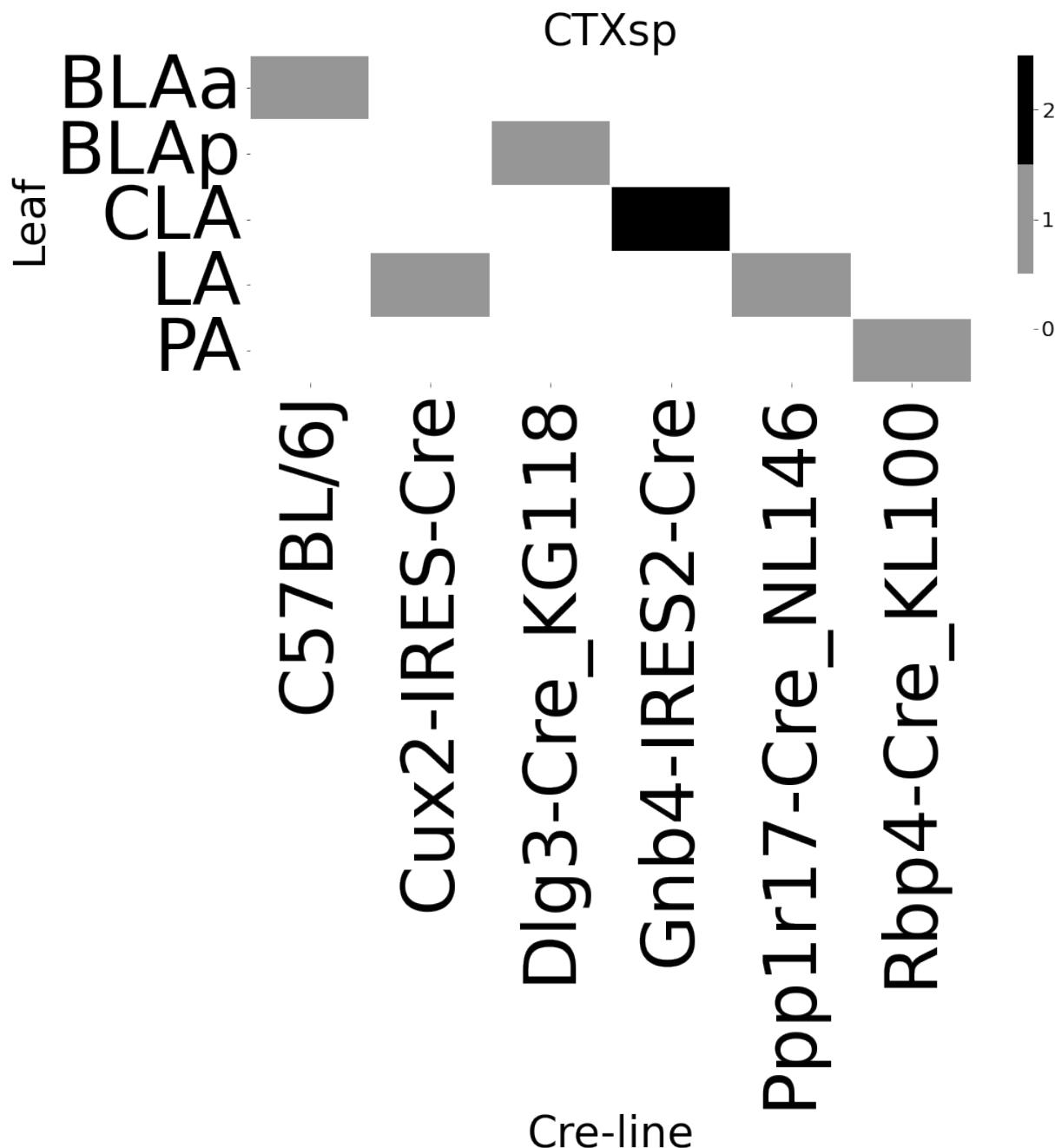


Figure 7: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

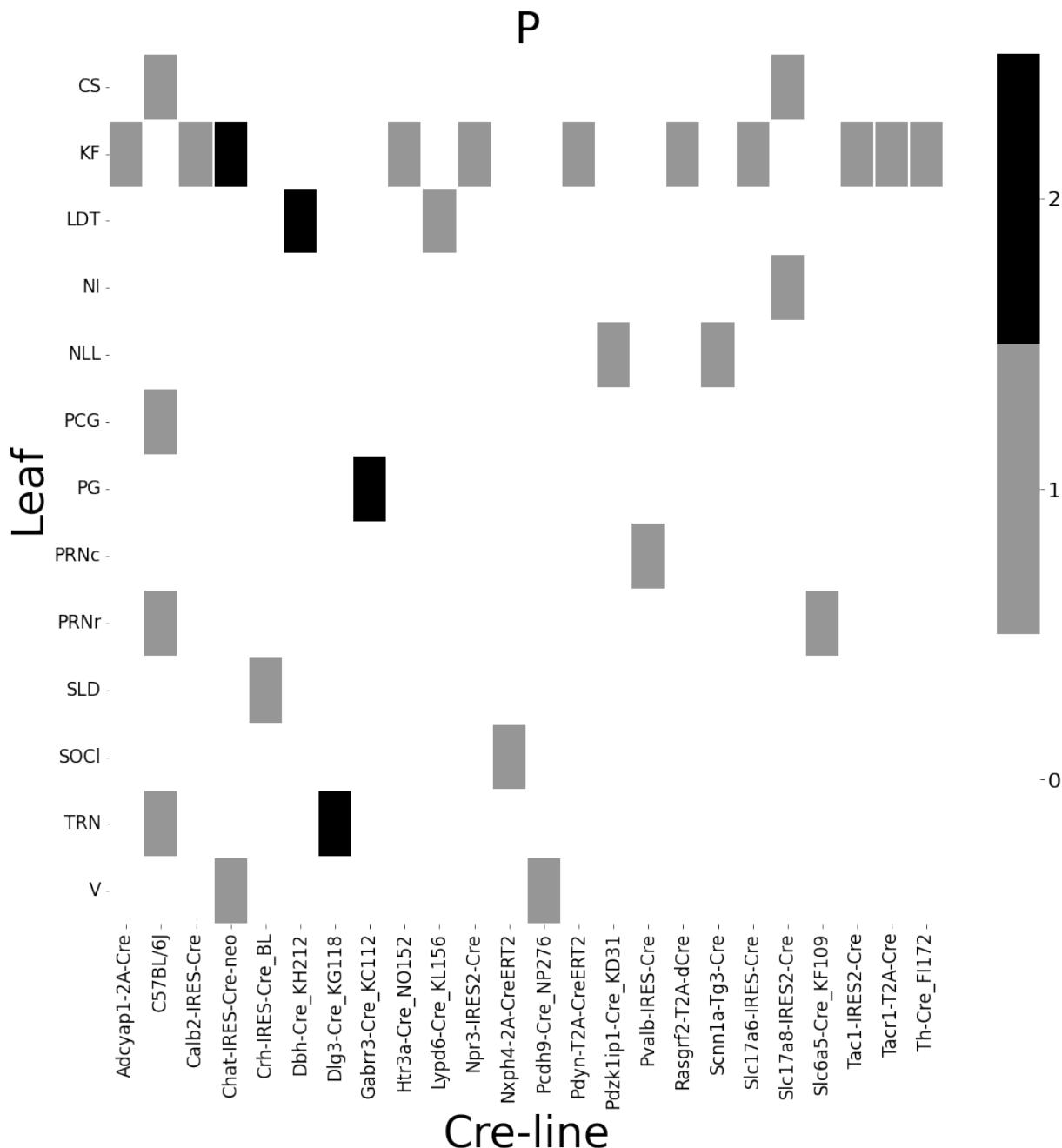


Figure 8: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

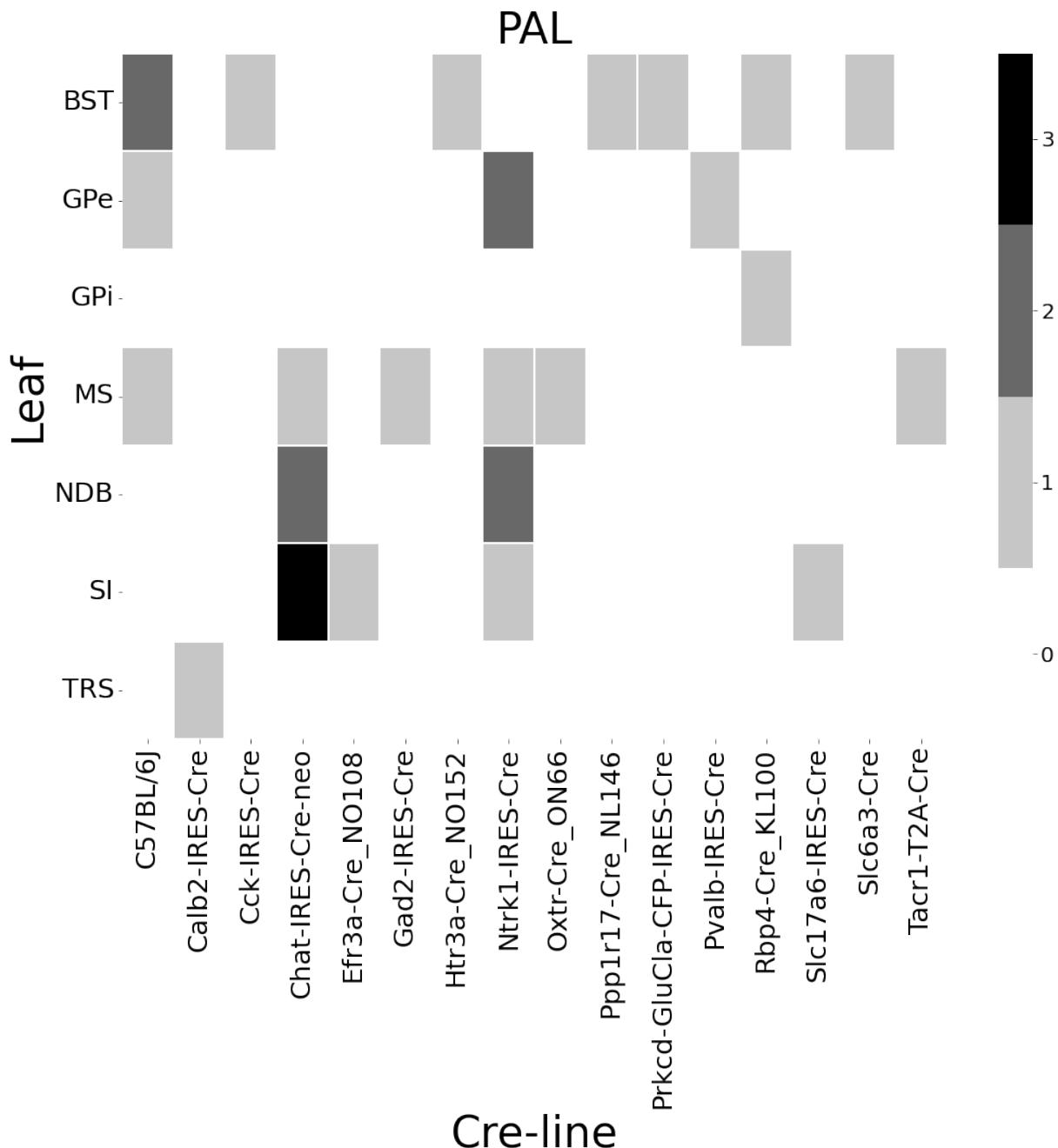


Figure 9: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

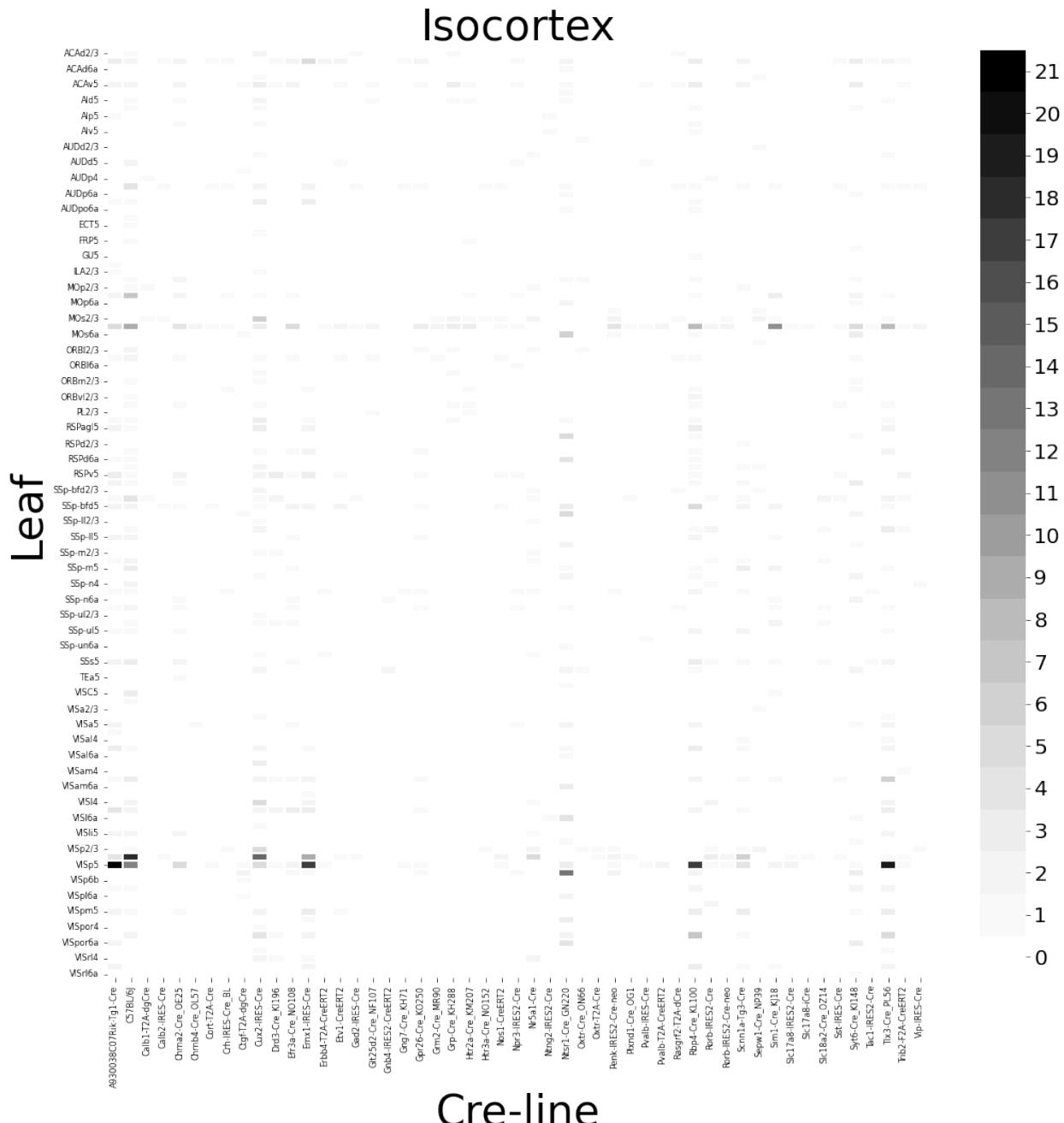


Figure 10: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

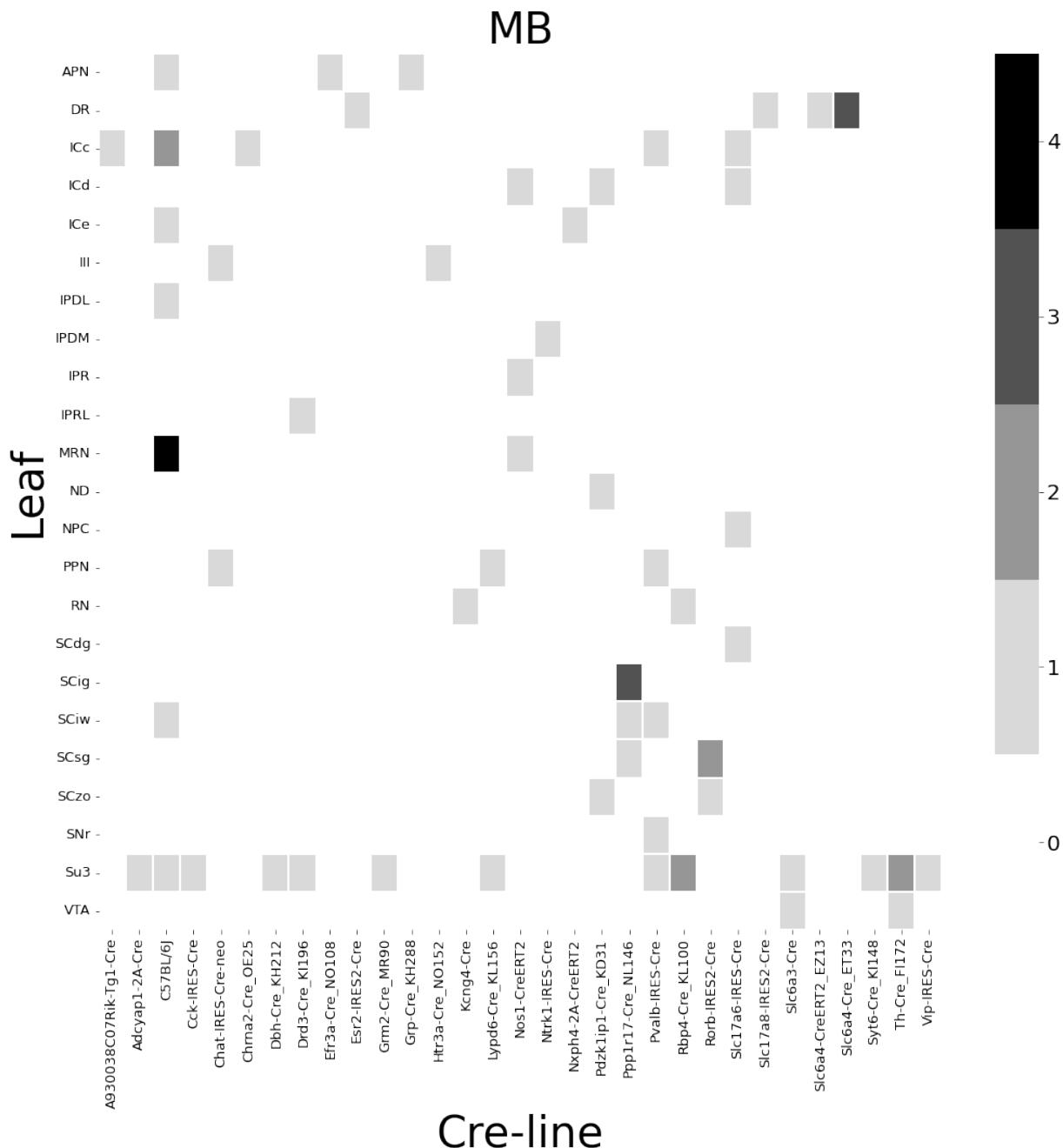


Figure 11: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

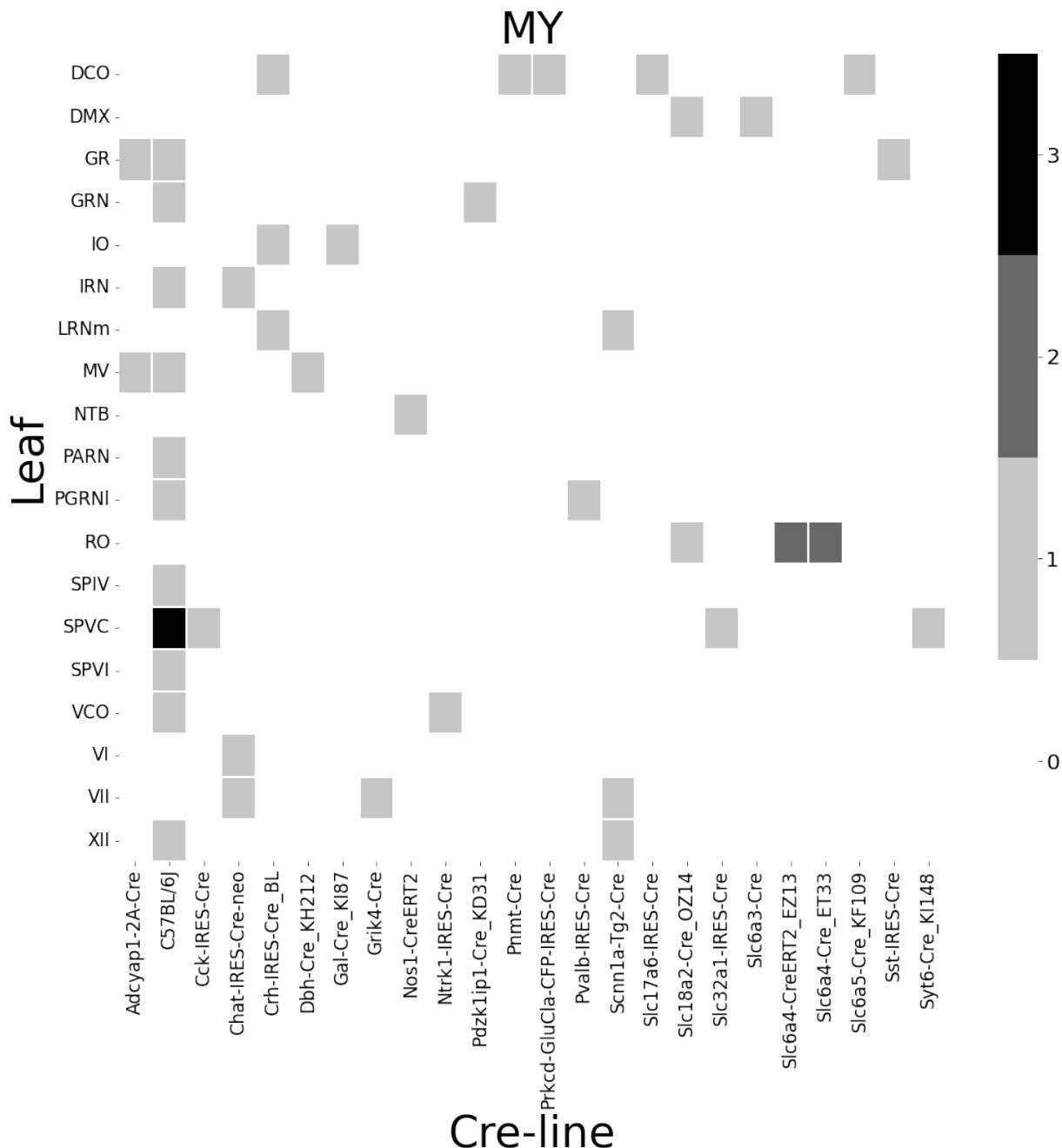


Figure 12: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

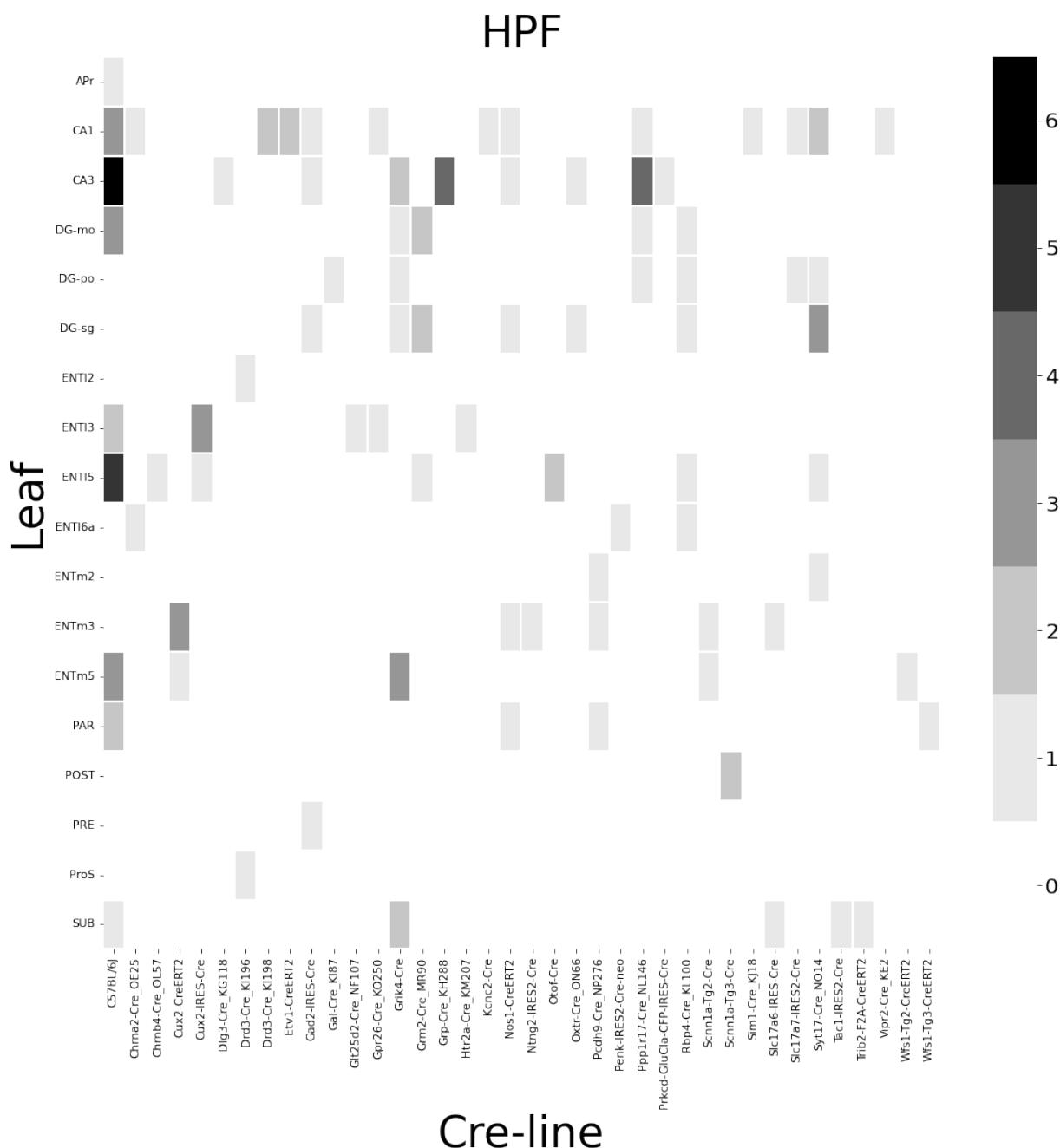


Figure 13: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

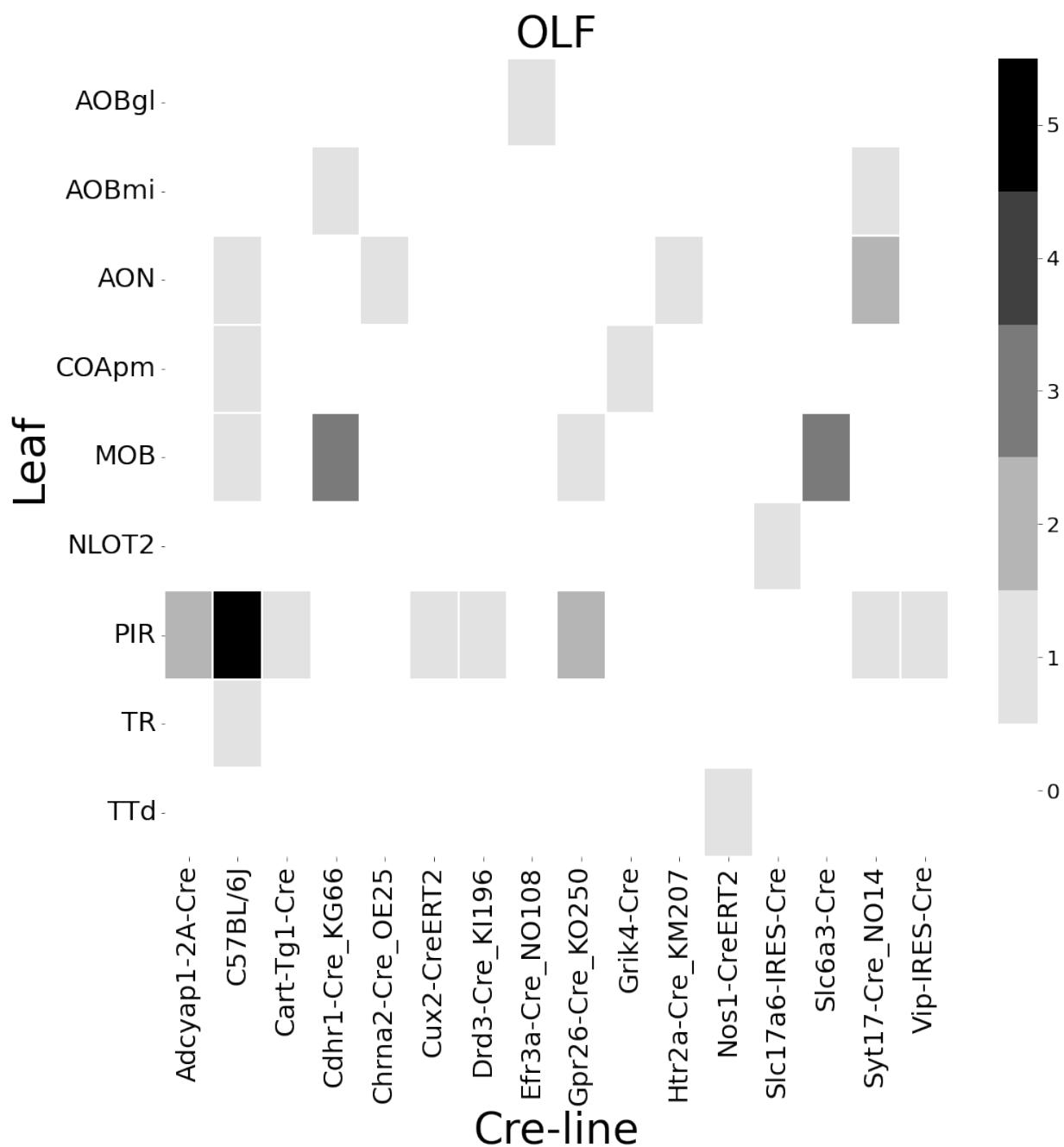


Figure 14: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

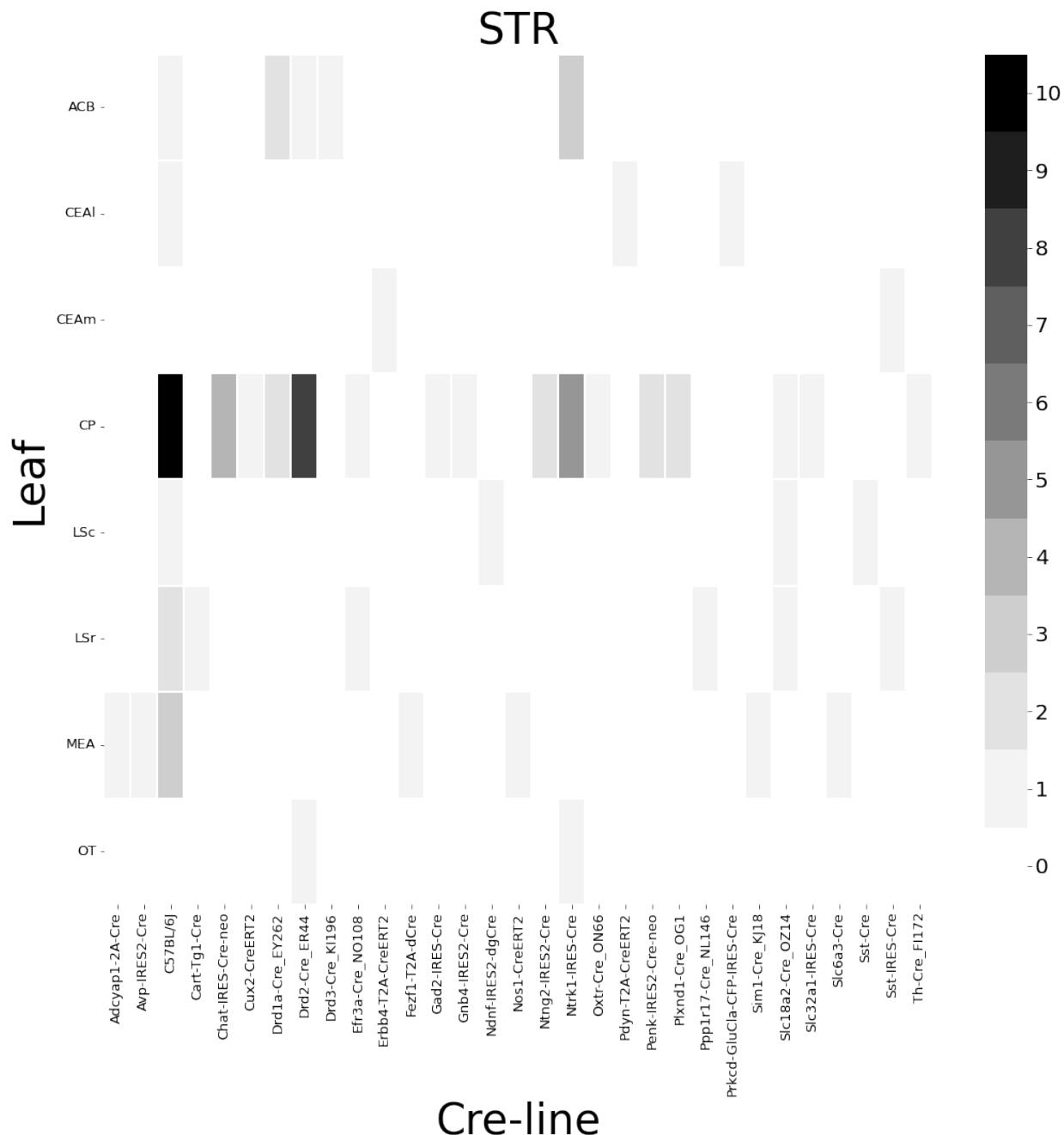


Figure 15: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

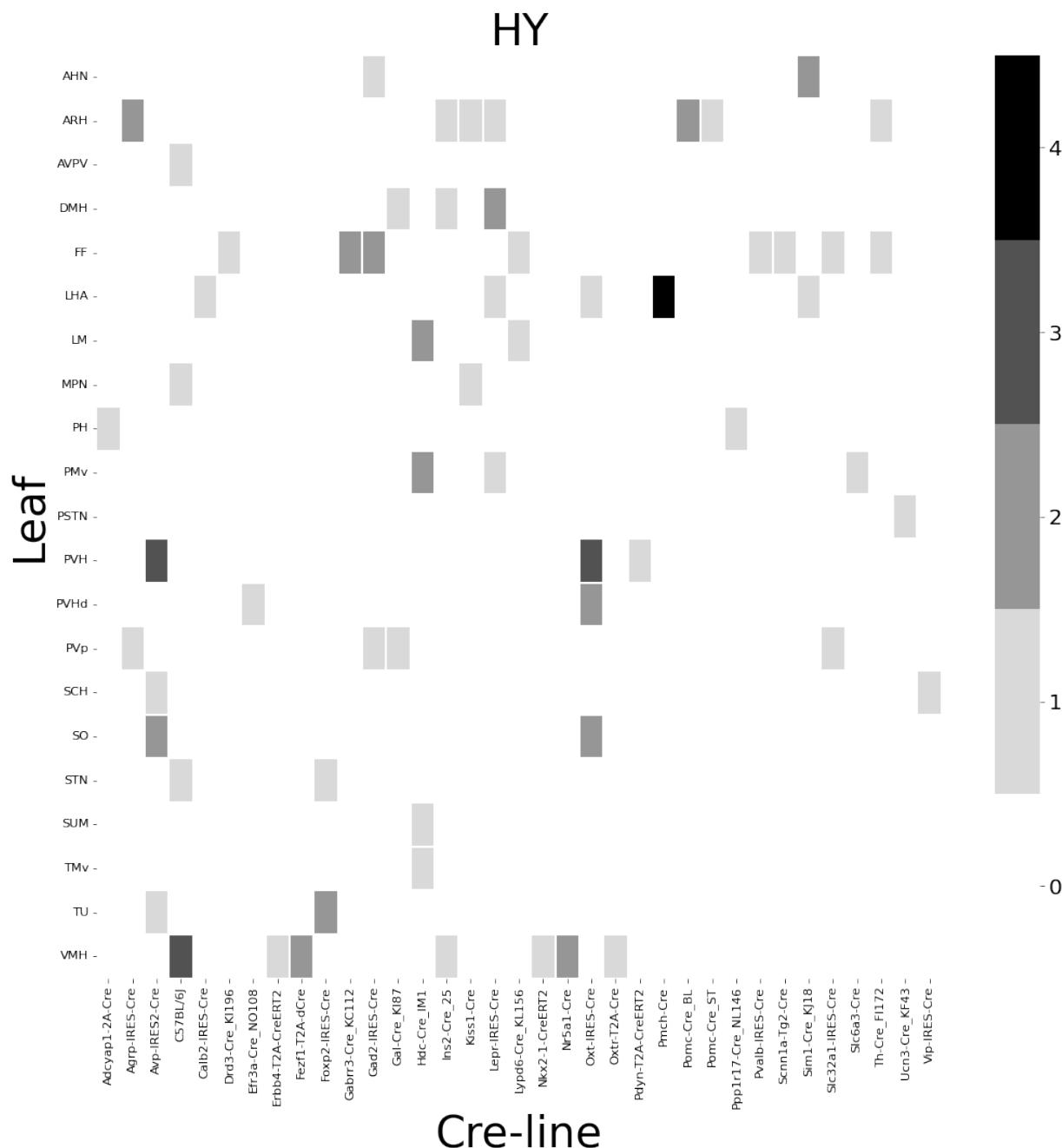


Figure 16: Frequencies of Cre-line and leaf-centroid combinations in our dataset.

406 ***Distances between structures***

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

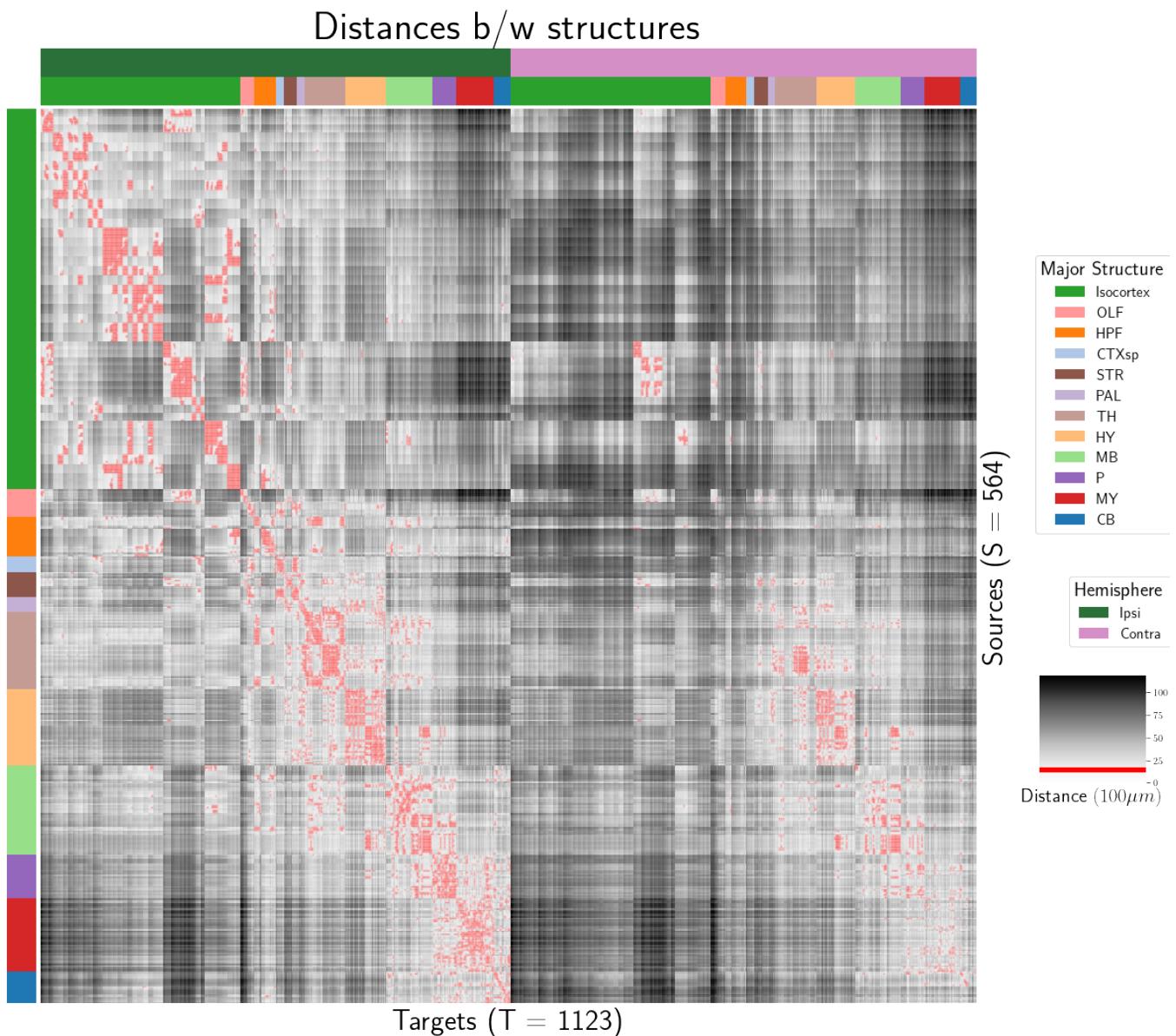


Figure 17: Distance between structures. Short-range connections are masked in red.

410 Model evaluation

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

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

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

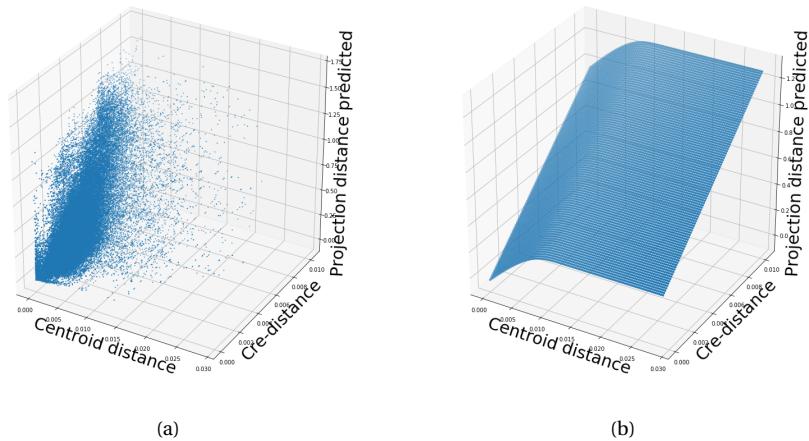


Figure 18: Fitting g . 18a Distribution of projection errors against centroid distance and cre-distance in Isocortex. 18b estimated \hat{g} using B-splines. Projection distance is $\|y_{\mathcal{T}}(i) - y_{\mathcal{T}}(i')\|_2$, Cre-distance is $\|\bar{y}_{\mathcal{T},s,v} - \bar{y}_{\mathcal{T},s,v}(i)\|_2$, and centroid-distance is $\|c(i') - c(i)\|_2$.

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 (Cai, 2001; Salha & El Shekh Ahmed, n.d.). While we do not attempt a
 508 detailed theoretical study of this estimator, we do establish the need for the shape constraint in our
 509 spline estimator. Though this fact is probably well known, we prove a (slightly stronger) version here
 510 for completeness.

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

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

- 515 ▪ The first property of a metric is that $f(x, x') = 0 \iff x = x'$. The left implication:
 516 $x = x' \implies f(x_1, x'_1, \dots, x_n, x'_n) = g(0, \dots, 0)$, since d are metrics. Then, since $g(0) = 0$, we have that
 517 $f(x, x') = 0$. The right implication: $f(x, x') = 0 \implies d = 0 \implies x = x'$ since d are metrics.
- 518 ▪ The second property of a metric is that $f(x, x') = f(x', x)$. This follows immediately from the
 519 symmetry of the d_i , i.e. $f(x, x') = f(x_1, x'_1, \dots, x_n, x'_n) = g(d_1(x_1, x'_1), \dots, d_n(x_n, x'_n)) =$
 520 $g(d_1(x'_1, x_1), \dots, d_n(x'_n, x_n)) = f(x'_1, x_1, \dots, x'_n, x_n) = f(x', x)$.
- 521 ▪ The third property of a metric is the triangle inequality: $f(x, x') \leq f(x, x^*) + f(x^*, x')$. To show this
 522 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
 523 non-decreasing and by the triangle inequality of d . Then, since g is concave,
 524 $g(d(x, x^*) + d(x^*, x')) \leq g(d(x, x^*)) + g(d(x^*, x')) = f(x, x^*) + f(x^*, x')$.

526 ***Setting a lower detection threshold***

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

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

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

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

537 ***Decomposing the connectivity matrix***

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

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

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

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

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

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

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

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

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

570 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

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

571 *Stabilizing NMF* The NMF program is non-convex, and, empirically, individual replicates will not
 572 converge to the same optima. One solution therefore is to run multiple replicates of the NMF
 573 algorithm and cluster the resulting vectors. This approach raises the questions of how many clusters
 574 to use, and how to deal with stochasticity in the clustering algorithm itself. We address this issue
 575 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

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

⁵⁷⁶ A review of this approach is found in von Luxburg (2010b). Intuitively, archetype vectors that cluster
⁵⁷⁷ together frequently over clustering replicates indicate the presence of a stable clustering. For d , we
⁵⁷⁸ utilize the adjusted Rand Index - a simple dissimilarity measure between clusterings. Note that we
⁵⁷⁹ expect to select slightly more than the q components suggested by cross-validation, since archetype
⁵⁸⁰ vectors which appear in one NMF replicate generally should appear in others. We then select the q
⁵⁸¹ clusters with the most archetype vectors - the most stable NMF results - and take the median of each
⁵⁸² cluster to create a sparse representative archetype Kotliar et al. (2019); Wu et al. (2016). We then find
⁵⁸³ the according H using Program 3. Experimental results for these cross-validation and stability
⁵⁸⁴ selection approaches are given in Supplemental Section 7.

7 SUPPLEMENTAL EXPERIMENTS

585 The supplemental experiments show results on lower limit of detection, performance of our estimator
586 for different regions and cell-classes, heirarchical clustering of connectivities, and stability and
587 component analysis of our NMF results.

588 *Setting detection threshold τ*

589 We give results on the false detection rate at different limits of detection. These conclusively show that
590 10^{-6} is the good threshold for our normalized data.

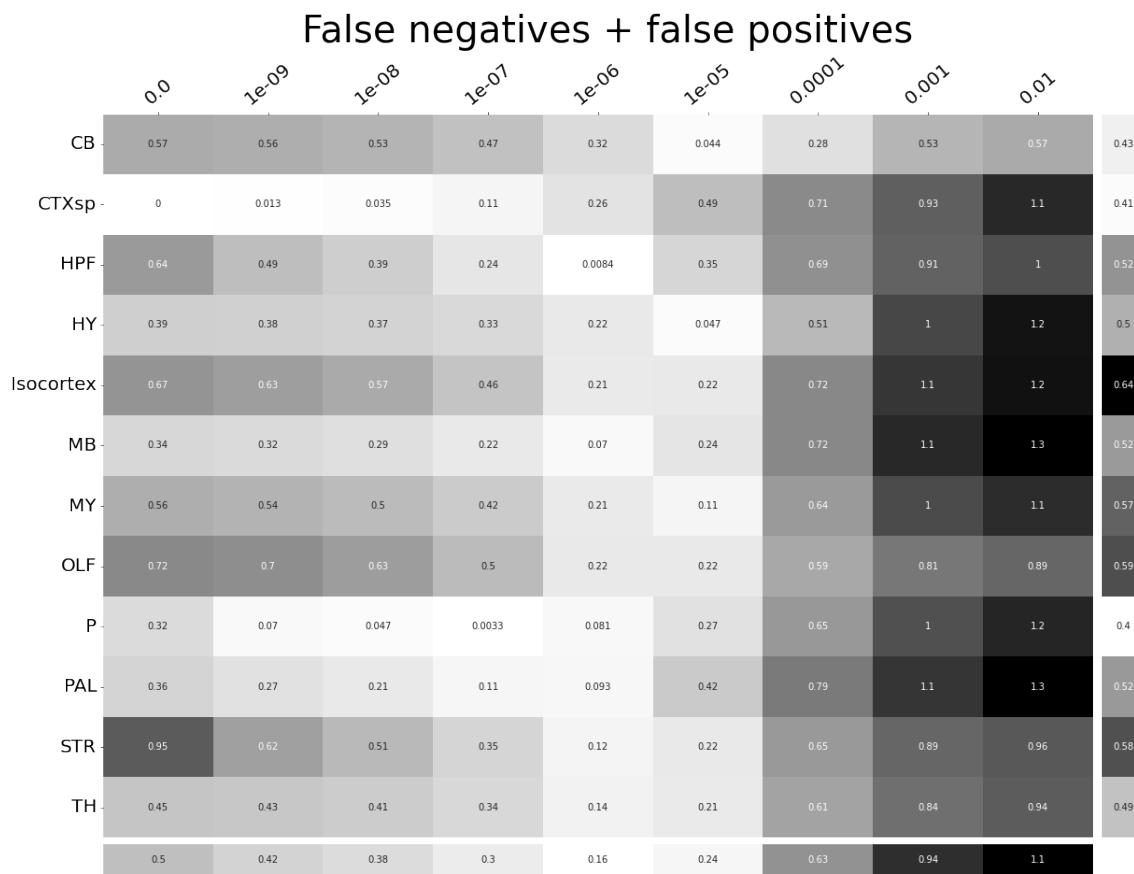


Figure 19: τ at different limits of detection in different major structures. 10^{-6} is the optimal detection threshold.

591 ***Loss subsets***

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



Figure 20: Weighted loss for Cre-leaf combinations in CB. Missing values are omitted. For example, this figure has one present and three missing values. Row and column averages are also plotted.

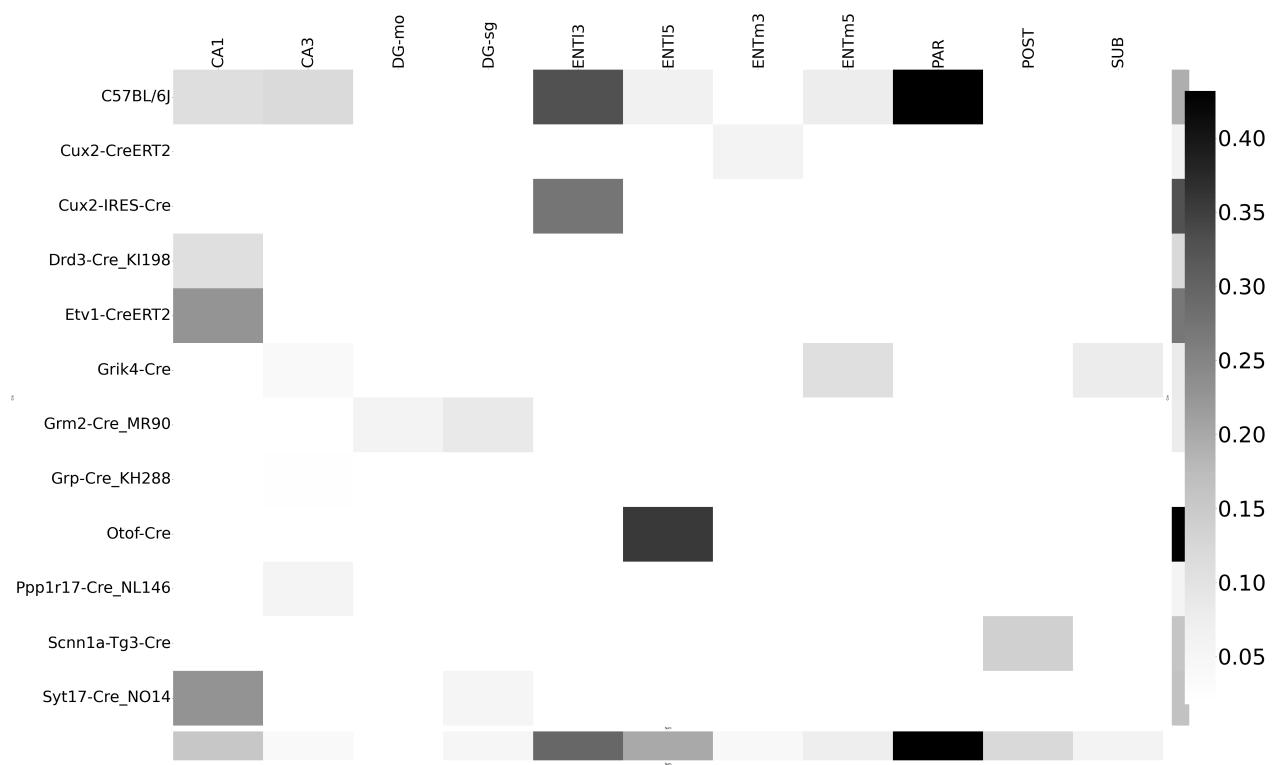


Figure 21: Weighted loss for Cre-leaf combinations in HPF. Missing values are omitted. Row and column averages are also plotted.

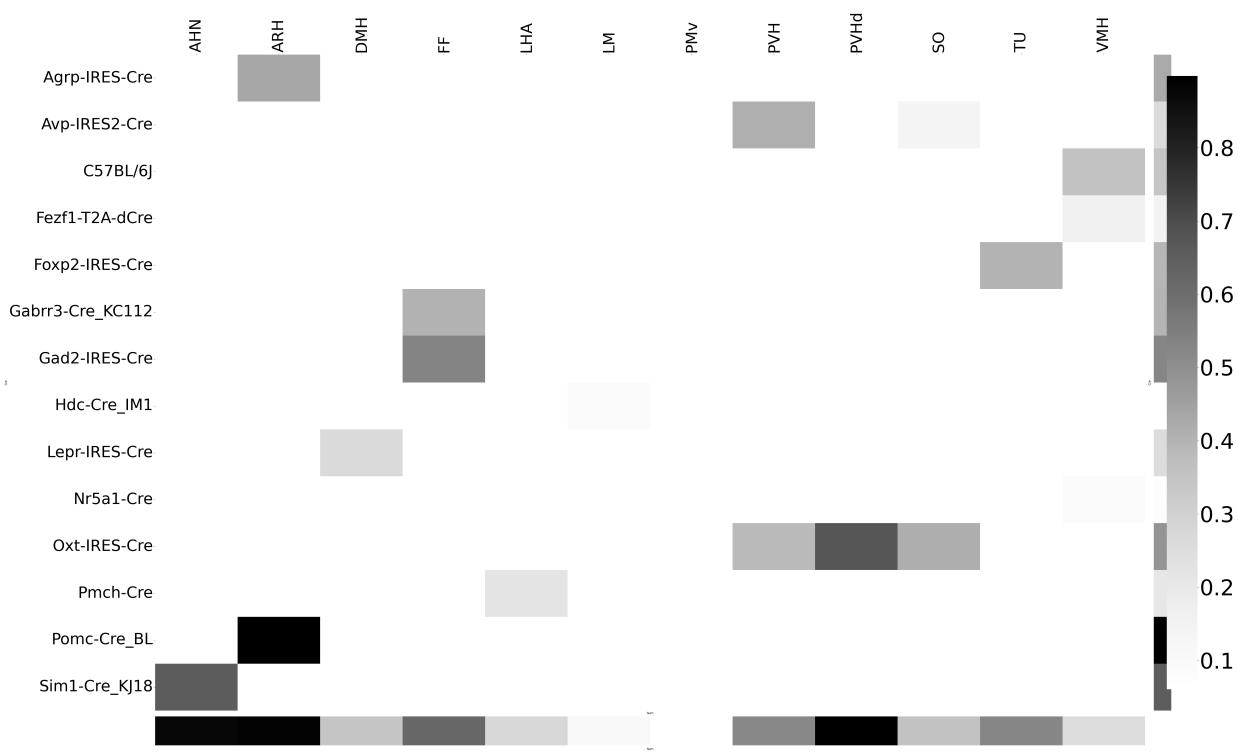


Figure 22: Weighted loss for Cre-leaf combinations in HY. Missing values are omitted. Row and column averages are also plotted.

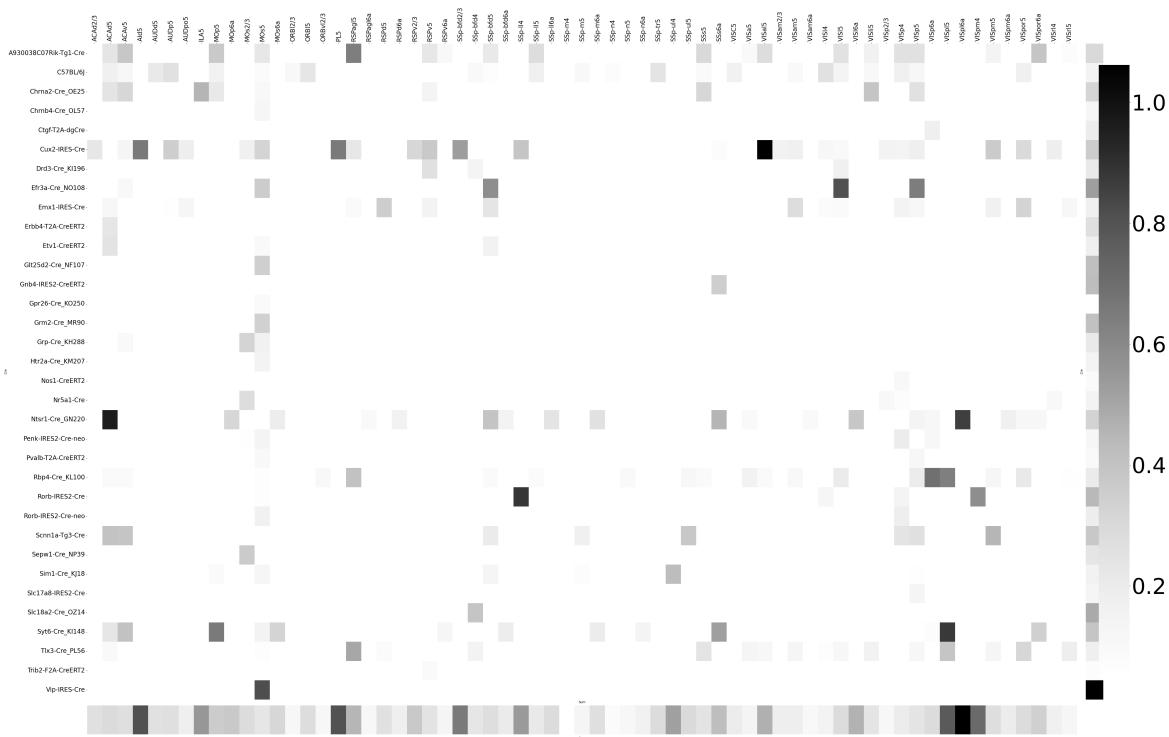


Figure 23: Weighted loss for Cre-leaf combinations in Isocortex. Missing values are omitted. Row and column averages are also plotted.

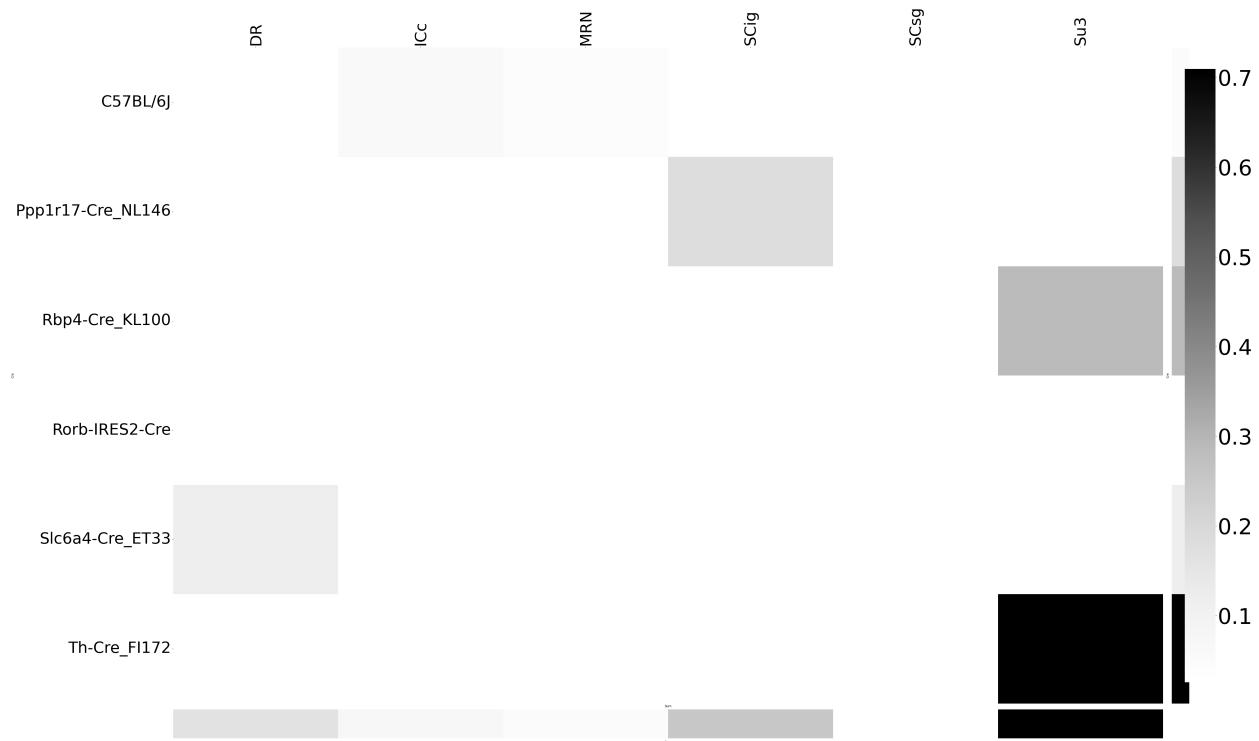


Figure 24: Weighted loss for Cre-leaf combinations in MB. Missing values are omitted. Row and column averages are also plotted.

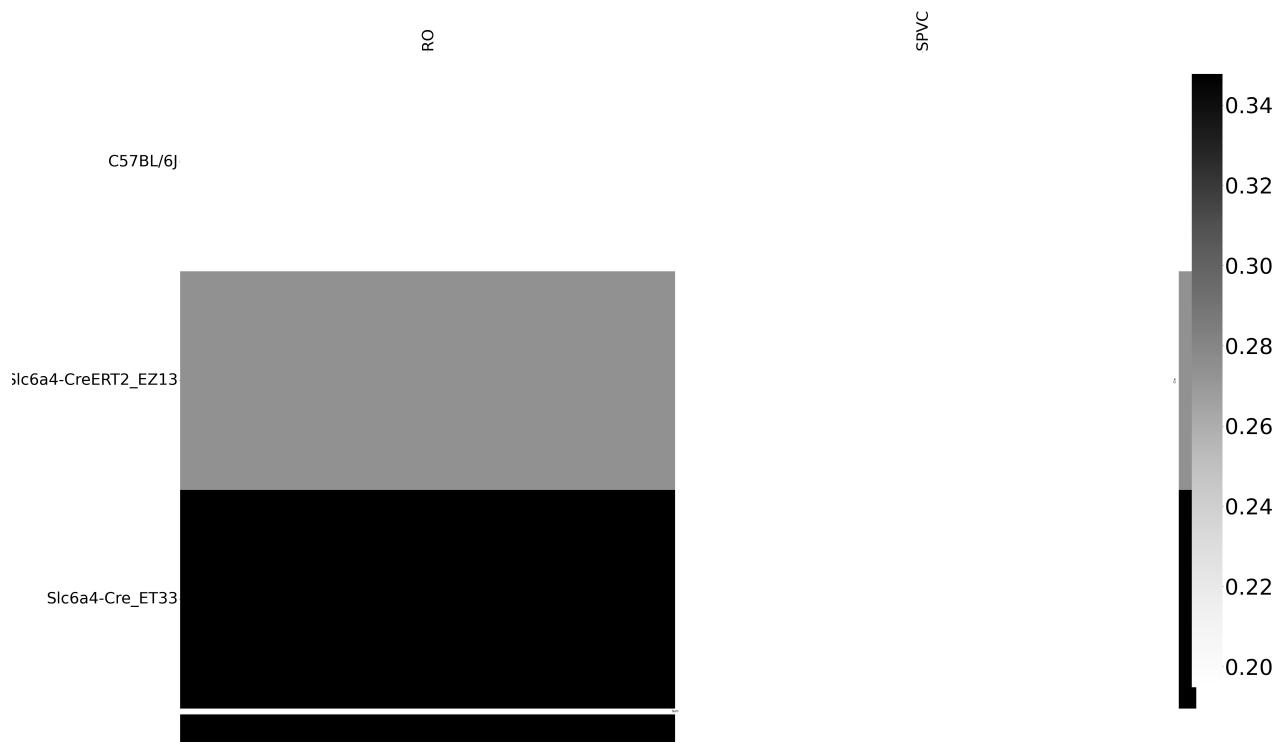


Figure 25: Weighted loss for Cre-leaf combinations in MY. Missing values are omitted. Row and column averages are also plotted.

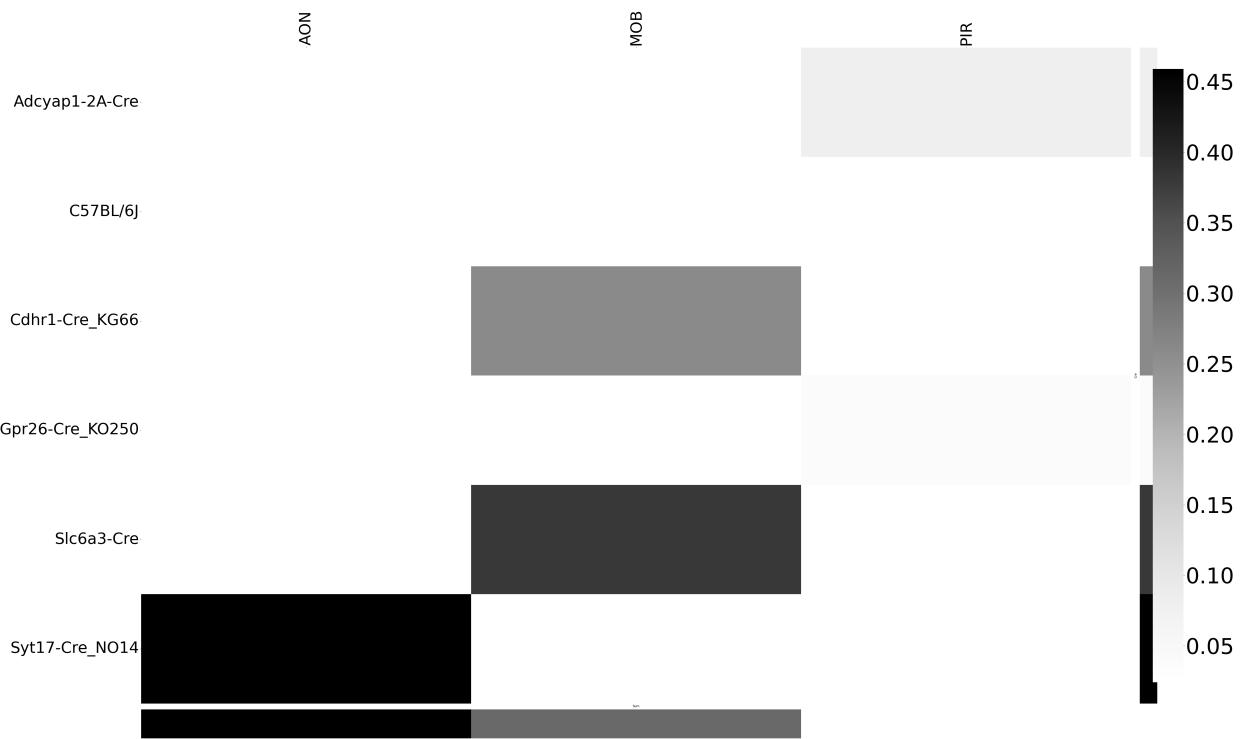


Figure 26: Weighted loss for Cre-leaf combinations in OLF. Missing values are omitted. Row and column averages are also plotted.

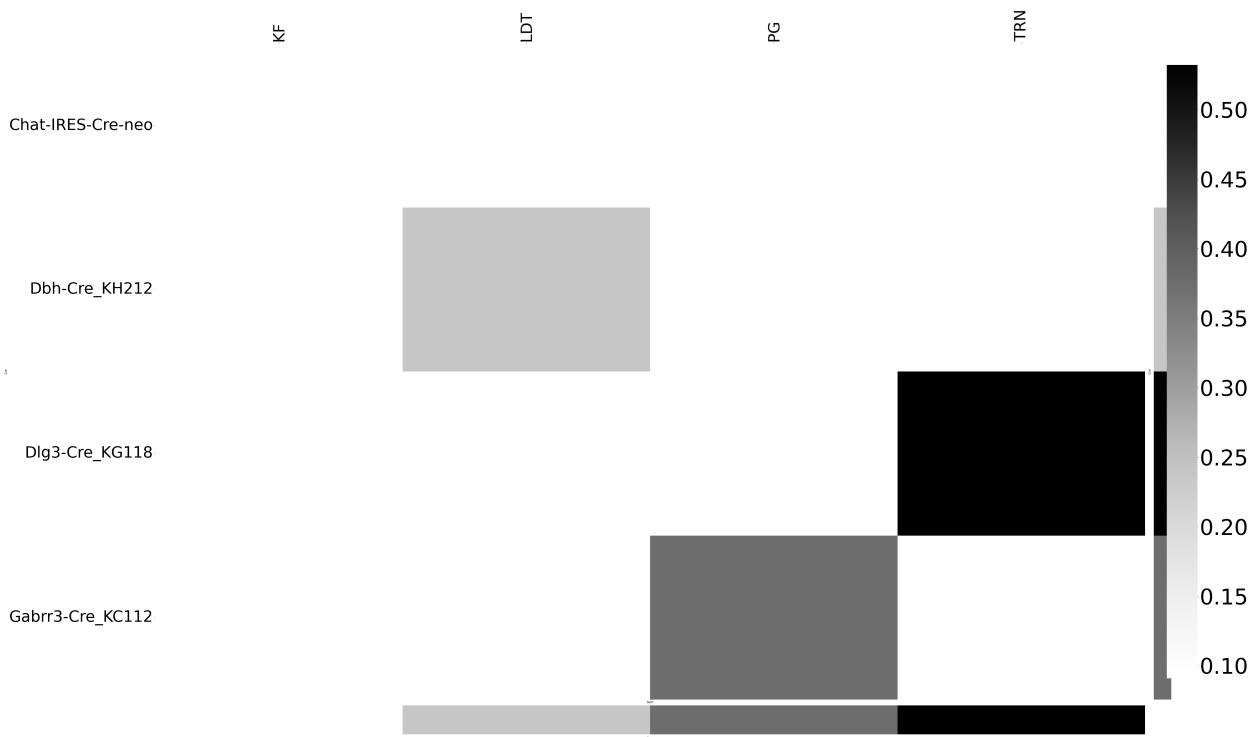


Figure 27: Weighted loss for Cre-leaf combinations in P. Missing values are omitted. Row and column averages are also plotted.

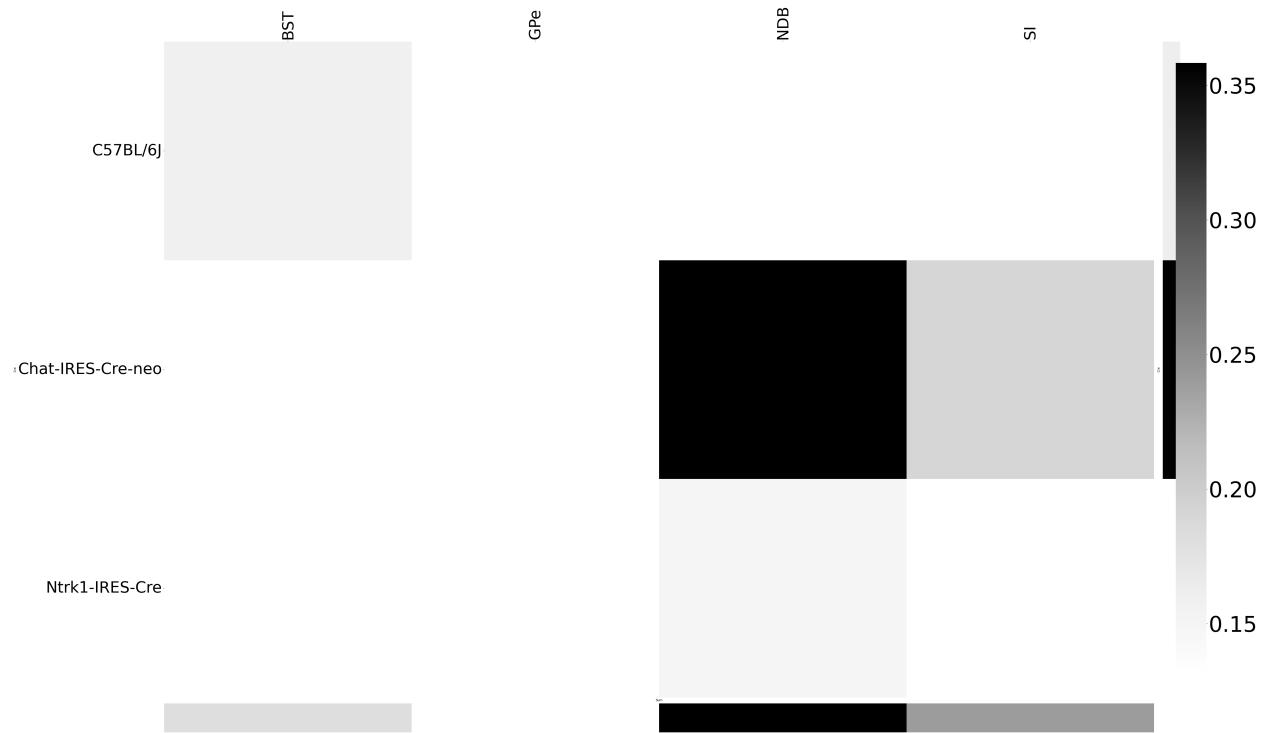


Figure 28: Weighted loss for Cre-leaf combinations in PAL. Missing values are omitted. Row and column averages are also plotted.

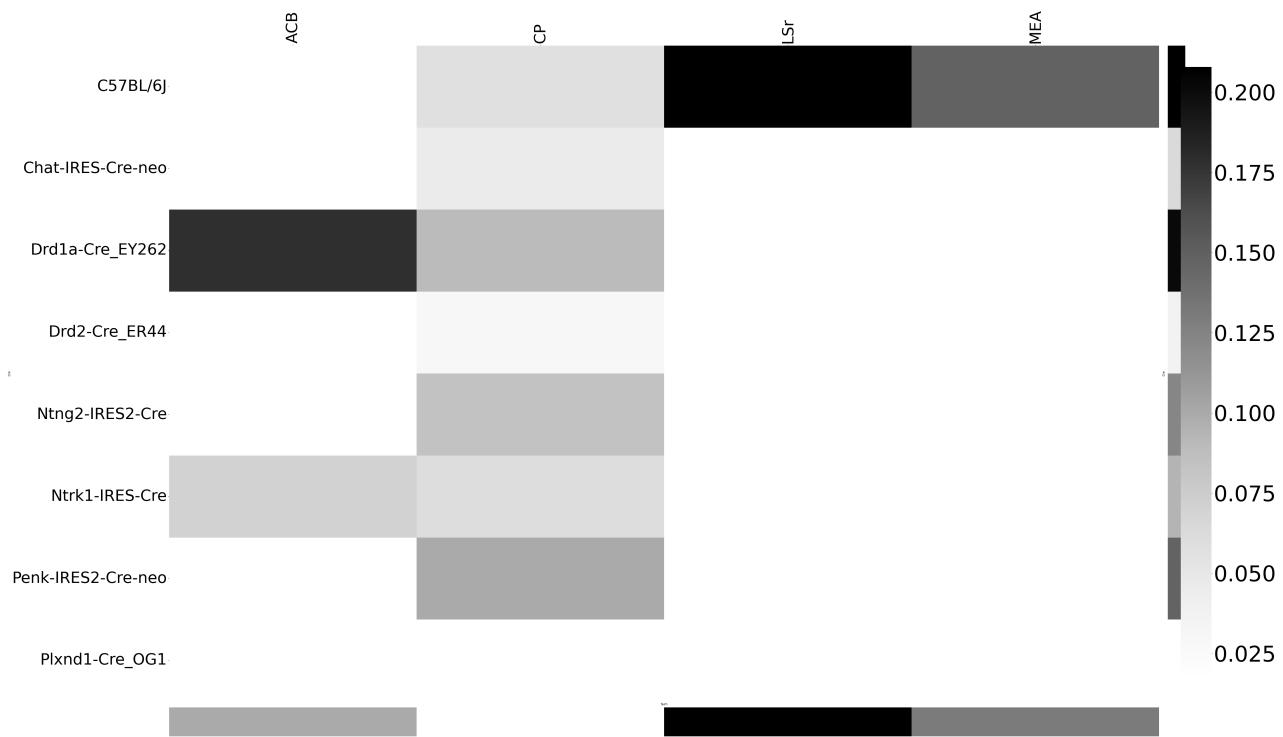


Figure 29: Weighted loss for Cre-leaf combinations in STR. Missing values are omitted. Row and column averages are also plotted.

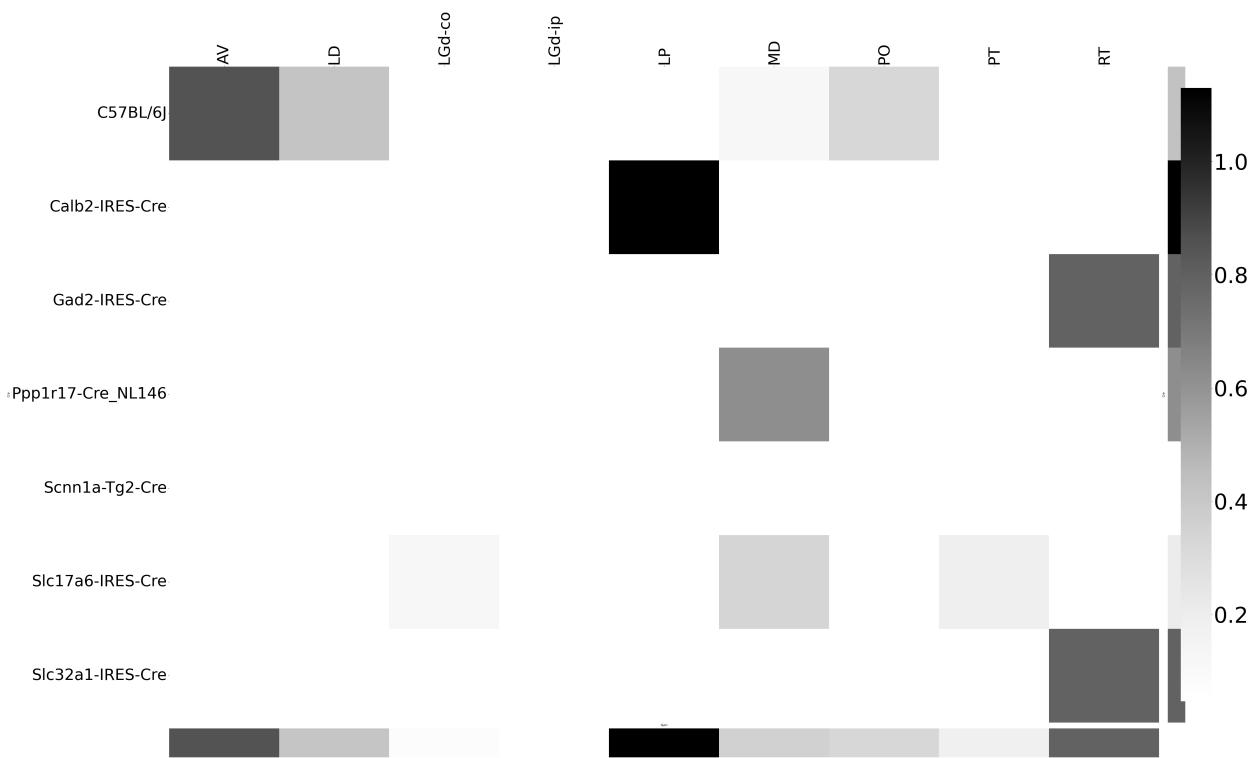


Figure 30: Weighted loss for Cre-leaf combinations in TH. Missing values are omitted. Row and column averages are also plotted.

595 ***Cell-type specificity***

596 We performed hierarchical clustering using the default method in Seaborn (Waskom, 2021) to
 597 investigate shared projection patterns across Cre-lines. That is, we used agglomerative clustering with
 598 Ward's criterion Hastie et al. (n.d.); Lalloué et al. (2013). This showed clustering of Ntsr1 projections to
 599 Thalamic nuceli.

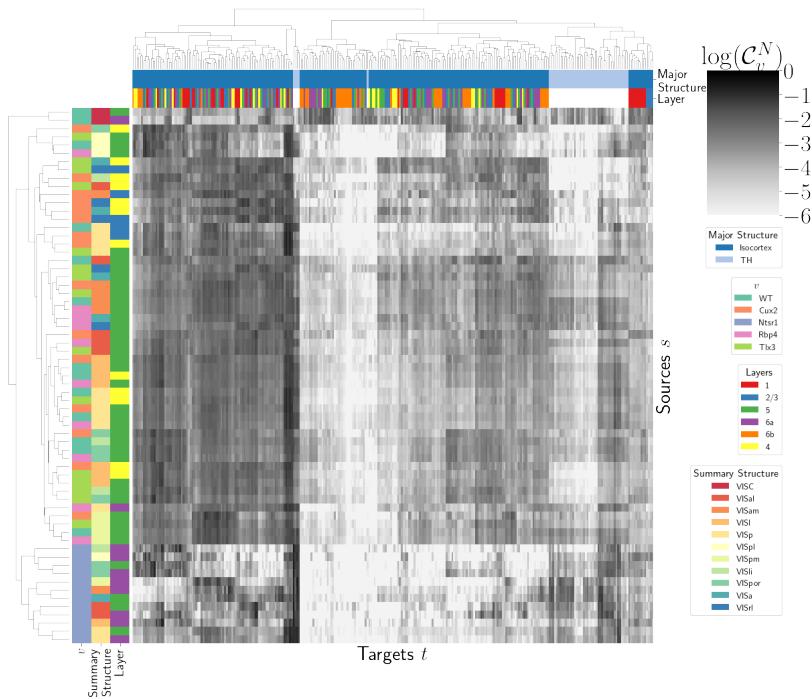


Figure 31: Hierarchical clustering of connectivity strengths from visual cortex cell-types to cortical and thalamic targets. Cre-line, summary structure, and layer are labelled on the sources. Major brain division and layer are labelled on the targets.

600 **Matrix Factorization**

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

604 *Cross-validation* We set $\alpha = 0.002$ and run Program 2 on \mathcal{C}_{wt} . We use a random mask with $p = .3$ to
 605 evaluate prediction accuracy of models trained on the unmasked data on the masked data. To
 606 account for stochasticity in the NMF algorithm, we run $R = 8$ replicates at each potential dimension q .
 607 The lowest mean test error was observed at $\hat{q} = 70$, indicating that even more components could be
 608 estimated. However, the low decrease in reconstruction error at higher values of q and need for
 609 brevity in our figures motivated us to choose $q = 15$ for the purposes of display.

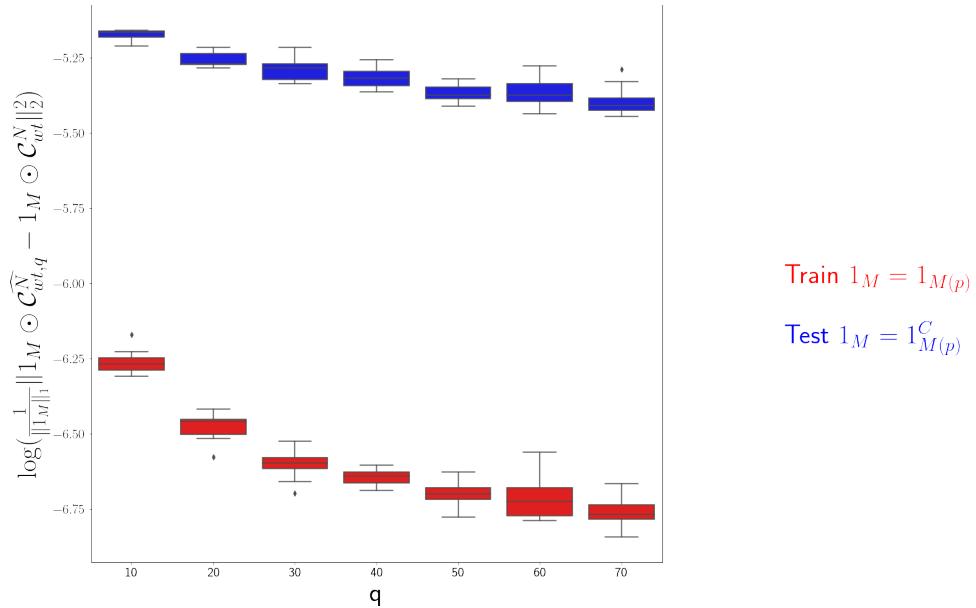


Figure 32: Train and test error using NMF decomposition.

610 *Stability* To address the instability of the NMF algorithm in identifying components, we k-means
 611 cluster components over $R = 10$ replicates with $k \in \{10, 15, 20, 25, 30\}$. Since the clustering is itself

612 unstable, we repeat the clustering 25 times and select the k with the largest Rand index, a standard
 613 method of clustering stability (Meila, 2007; Rand, 1971).

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

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

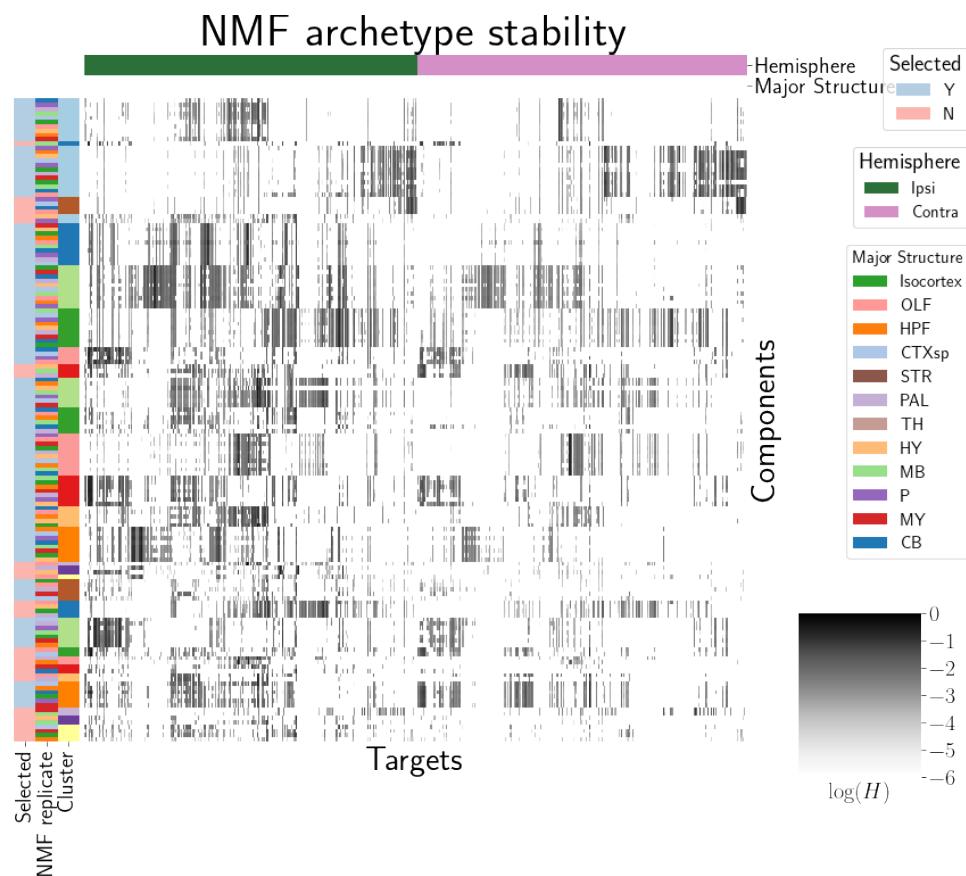


Figure 33: Stability of NMF results across replicates. Replicate and NMF component are shown on rows. Components that are in the top 15 are also indicated.

617 We plot the medians of these components in Figure 4a and in the main text. These are the
 618 connectivity archetypes. We then fit a non-negative least squares (the second step in the standard
 619 NMF optimization algorithm) to determine W (Lee & Seung, 2000)

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

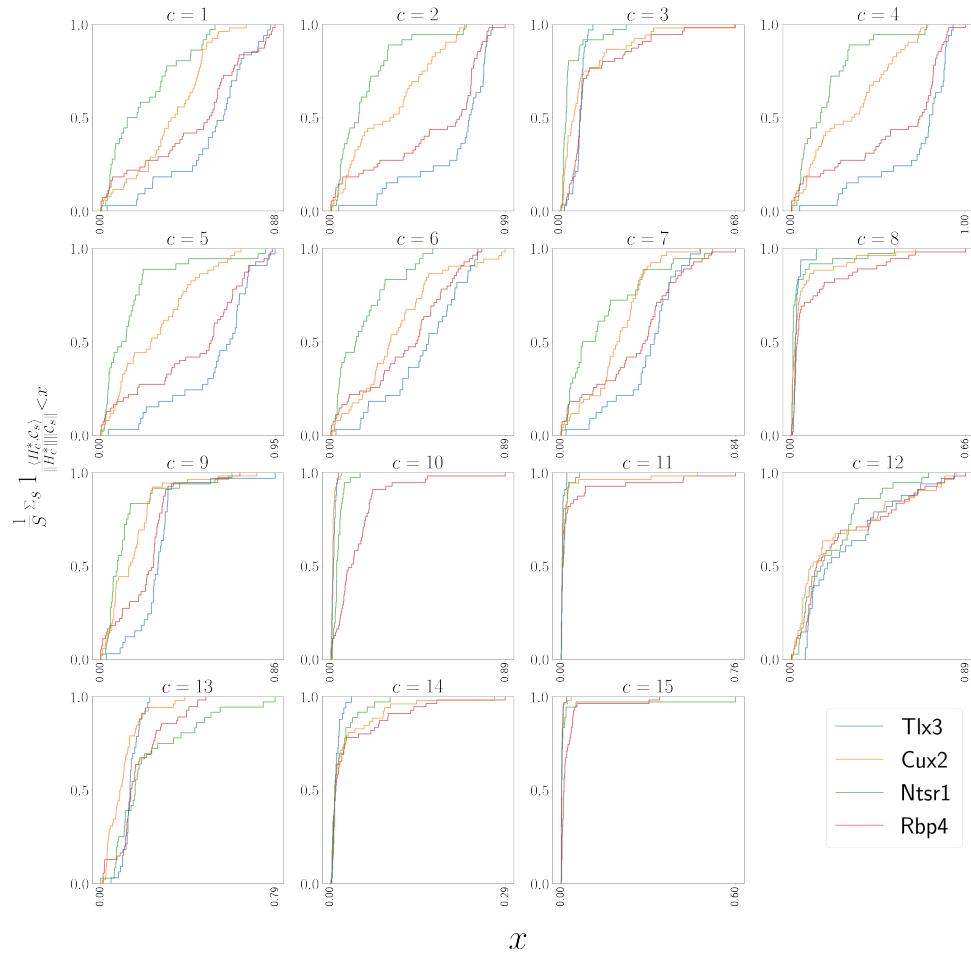


Figure 34: Empirical cumulative distributions of cosine similarities between source structures and connectivity components for four different Cre-lines.

624

625 **REFERENCES**

626

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

629 Cai, Z. (2001). Weighted Nadaraya-Watson regression estimation. *Stat. Probab. Lett.*, 51(3), 307–318.

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

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

634 Devarajan, K. (2008). Nonnegative matrix factorization: an analytical and interpretive tool in computational biology.
 635 *PLoS Comput. Biol.*, 4(7), e1000029.

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

638 Gämänuț, R., Kennedy, H., Toroczkai, Z., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., & Burkhalter, A. (2018). The
 639 mouse cortical connectome, characterized by an Ultra-Dense cortical graph, maintains specificity by distinct
 640 connectivity profiles. *Neuron*, 97(3), 698–715.e10.

641 Gao, Y., Zhang, X., Wang, S., & Zou, G. (2016). Model averaging based on leave-subject-out cross-validation. *J. Econom.*,
 642 192(1), 139–151.

643 Groeneboom, P., & Jongbloed, G. (2018). Some developments in the theory of shape constrained inference. *Stat. Sci.*,
 644 33(4), 473–492.

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

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

- 649 Harris, K. D., Mihalas, S., & Shea-Brown, E. (2016). Nonnegative spline regression of incomplete tracing data reveals high
650 resolution neural connectivity.
- 651 Hastie, T., Tibshirani, R., & Friedman, J. (n.d.). *The elements of statistical learning*. Springer New York.
- 652 Huang, K. W., Ochandarena, N. E., Philson, A. C., Hyun, M., Birnbaum, J. E., Cicconet, M., & Sabatini, B. L. (2019).
653 Molecular and anatomical organization of the dorsal raphe nucleus. *Elife*, 8.
- 654 Jackson, K. L., Dayton, R. D., Deverman, B. E., & Klein, R. L. (2016). Better targeting, better efficiency for Wide-Scale
655 neuronal transduction with the synapsin promoter and AAV-PHP.B. *Front. Mol. Neurosci.*, 9, 116.
- 656 Jeong, M., Kim, Y., Kim, J., Ferrante, D. D., Mitra, P. P., Osten, P., & Kim, D. (2016). Comparative three-dimensional
657 connectome map of motor cortical projections in the mouse brain. *Sci. Rep.*, 6, 20072.
- 658 Knox, J. E., Harris, K. D., Graddis, N., Whitesell, J. D., Zeng, H., Harris, J. A., ... Mihalas, S. (2019). High-resolution
659 data-driven model of the mouse connectome. *Netw Neurosci*, 3(1), 217–236.
- 660 Kotliar, D., Veres, A., Nagy, M. A., Tabrizi, S., Hodis, E., Melton, D. A., & Sabeti, P. C. (2019). Identifying gene expression
661 programs of cell-type identity and cellular activity with single-cell RNA-Seq. *Elife*, 8.
- 662 Kuan, L., Li, Y., Lau, C., Feng, D., Bernard, A., Sunkin, S. M., ... Ng, L. (2015). Neuroinformatics of the allen mouse brain
663 connectivity atlas. *Methods*, 73, 4–17.
- 664 Kügler, S., Kilic, E., & Bähr, M. (2003). Human synapsin 1 gene promoter confers highly neuron-specific long-term
665 transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther.*,
666 10(4), 337–347.
- 667 Lalloué, B., Monnez, J.-M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., & Deguen, S. (2013). A statistical
668 procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int. J. Equity Health*, 12, 21.
- 669 Lee, D., & Seung, H. S. (2000). Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.*, 13.
- 670 Li, X., Yu, B., Sun, Q., Zhang, Y., Ren, M., Zhang, X., ... Qiu, Z. (2018). Generation of a whole-brain atlas for the
671 cholinergic system and mesoscopic projectome analysis of basal forebrain cholinergic neurons. *Proc. Natl. Acad. Sci.
U. S. A.*, 115(2), 415–420.

- 673 Llano, D. A., & Sherman, S. M. (2008). Evidence for nonreciprocal organization of the mouse auditory
674 thalamocortical-corticothalamic projection systems. *J. Comp. Neurol.*, 507(2), 1209–1227.
- 675 Lotfollahi, M., Naghipourfar, M., Theis, F. J., & Alexander Wolf, F. (2019). Conditional out-of-sample generation for
676 unpaired data using trVAE.
- 677 Meila, M. (2007). Comparing clusterings—an information based distance. *J. Multivar. Anal.*, 98(5), 873–895.
- 678 Mohammadi, S., Ravindra, V., Gleich, D. F., & Grama, A. (2018). A geometric approach to characterize the functional
679 identity of single cells. *Nat. Commun.*, 9(1), 1516.
- 680 Muzerelle, A., Scotto-Lomassese, S., Bernard, J. F., Soiza-Reilly, M., & Gaspar, P. (2016). Conditional anterograde tracing
681 reveals distinct targeting of individual serotonin cell groups (B5-B9) to the forebrain and brainstem. *Brain Struct.
Funct.*, 221(1), 535–561.
- 683 Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., ... Zeng, H. (2014). A mesoscale connectome of the
684 mouse brain. *Nature*, 508(7495), 207–214.
- 685 Perry, P. O. (2009). Cross-Validation for unsupervised learning.
- 686 Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.*, 66(336), 846–850.
- 687 Ren, J., Friedmann, D., Xiong, J., Liu, C. D., Ferguson, B. R., Weerakkody, T., ... Luo, L. (2018). Anatomically defined and
688 functionally distinct dorsal raphe serotonin sub-systems. *Cell*, 175(2), 472–487.e20.
- 689 Ren, J., Isakova, A., Friedmann, D., Zeng, J., Grutzner, S. M., Pun, A., ... Luo, L. (2019). Single-cell transcriptomes and
690 whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *Elife*, 8.
- 691 Salha, R. B., & El Shekh Ahmed, H. I. (n.d.). Reweighted Nadaraya-Watson estimator of the regression mean.
692 *International Journal of Statistics and Probability*, 4(1), 2015.
- 693 Saul, L. K., & Roweis, S. T. (2003). Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *J.
Mach. Learn. Res.*, 4(Jun), 119–155.
- 695 Saunders, A., Johnson, C. A., & Sabatini, B. L. (2012). Novel recombinant adeno-associated viruses for cre activated and
696 inactivated transgene expression in neurons. *Front. Neural Circuits*, 6, 47.

- 697 Servén D., B. C. (n.d.). *pygam: Generalized additive models in python.*
- 698 von Luxburg, U. (2010a). Clustering stability: An overview.
- 699 von Luxburg, U. (2010b). Clustering stability: An overview.
- 700 Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., ... Ng, L. (2020). The allen mouse brain common coordinate
701 framework: A 3D reference atlas. *Cell*, 181(4), 936–953.e20.
- 702 Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. Retrieved
703 from <https://doi.org/10.21105/joss.03021> doi: 10.21105/joss.03021
- 704 Watson, C., Paxinos, G., & Puelles, L. (2012). The mouse nervous system..
- 705 Wu, S., Joseph, A., Hammonds, A. S., Celtniker, S. E., Yu, B., & Frise, E. (2016). Stability-driven nonnegative matrix
706 factorization to interpret spatial gene expression and build local gene networks. *Proc. Natl. Acad. Sci. U. S. A.*, 113(16),
707 4290–4295.
- 708 Zaborszky, L., Csordas, A., Mosca, K., Kim, J., Gielow, M. R., Vadász, C., & Nádasdy, Z. (2015). Neurons in the basal
709 forebrain project to the cortex in a complex topographic organization that reflects corticocortical connectivity
710 patterns: an experimental study based on retrograde tracing and 3D reconstruction. *Cereb. Cortex*, 25(1), 118–137.

8 TECHNICAL TERMS

- 711 **Technical Term** a key term that is mentioned in an NETN article and whose usage and definition may
712 not be familiar across the broad readership of the journal.
- 713 **Cre-line** The combination of Cre-recombinase expression in transgenic mouse and Cre-induced
714 promotion in the vector that induces labelling of projection
- 715 **Cell class** The projecting neurons targeted by a particular Cre-line
- 716 **Structural connectivities** connectivity between structures
- 717 **Voxel** A $100\mu m$ cube of brain
- 718 **Structural connection tensor** Connectivities between structures given a neuron class
- 719 **dictionary-learning** A family of algorithms for finding low-dimensional data representations.
- 720 **Shape constrained estimator** A statistical estimator that fits a function of a particular shape (e.g.
721 monotonic increasing, convex).
- 722 **Nadaraya-Watson** A simple local smoothing estimator.
- 723 **Connectivity archetypes** Projection patterns from which we can linearly reconstruct the
724 connectome
- 725 **Expected loss** Our new estimator that weights different features by their estimated predictive
726 power.