

¹ RESEARCH

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

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

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

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

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

ABSTRACT

⁹ The Allen Mouse Brain Connectivity Atlas consists of anterograde tracing experiments targeting
¹⁰ diverse structures and classes of projecting neurons. Beyond regional anterograde tracing done in
¹¹ C57BL/6 wild type mice, a large fraction of experiments are performed using transgenic Cre-lines.
¹² This allows access to cell-class specific whole brain connectivity information, with class defined by
¹³ the transgenic lines. However, even though the number of experiments is large, it does not come close
¹⁴ to covering all existing cell classes in every area where they exist. Here, we study how much we can fill
¹⁵ in these gaps and construct a voxel-based connectivity matrix given the observations that nearby
¹⁶ voxels have similar connections when they are in the same area, that connections can change
¹⁷ dramatically at area boundaries, and that particular cell classes can have similar connections, but that
¹⁸ this similarity is region-dependent.

¹⁹ This paper describes the conversion of the Cre-line tracer experiments into class-specific
²⁰ connectivity matrices representing the connection strengths between source and target structures.

21 We introduce and validate a novel statistical model for creation of connectivity matrices. We expand a
22 Nadaraya-Watson kernel learning method which we previously used to fill in spatial gaps, to also fill in
23 a gaps in cell class connectivity information. To do this, we construct a "cell-class space" and combine
24 smoothing in 3D space as well as in this abstract space to share information between similar neuron
25 classes. Using this method we construct a set of connectivity matrices using multiple levels of
26 resolution at which discontinuities in connectivity are assumed. We show that the connectivities
27 obtained from this model display expected cell-type and structure specific connectivities. Inspired by
28 how this complexity arises from a relatively parsimonious set of genetic information during
29 development, we also show that the wild type connectivity matrix can be factored using a sparse set of
30 factors, and analyze the informativeness of this underlying latent structure.

AUTHOR SUMMARY

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

1 INTRODUCTION

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

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

57 Several statistical models for the conversion of such experiment-specific signals into generalized
58 estimates of connectivity strength have been proposed (K. D. Harris, Mihalas, & Shea-Brown, 2016;
59 Knox et al., 2019; Oh et al., 2014; ?). Of these, Oh et al. (2014) and Knox et al. (2019) model **regionalized**
60 **connectivities**, which are voxel connectivities integrated by region. The value of these models is that
61 they provide some improvement over simply averaging the projection signals of injections in a given
62 region. However, these previous works only model connectivities observed in wild type mice in which
63 all neuron types were labeled, and so are poorly suited for extension to tracing experiments that
64 induce cell-type specific fluorescence (J. A. Harris et al., 2019). In particular, GFP expression is
65 induced by Cre-recombinase in cell-types specified by transgenic strain. Thus, this paper introduces a
66 **cell class**-specific statistical model to deal with the diverse set of **Cre-lines** described in J. A. Harris et
67 al. (2019), and expands this model to the entire mouse brain. In cortex a large number of transgenic
68 lines were chosen for their laminar specific expression.

69 Our model is a to-our-knowledge novel estimator that takes into account both the spatial position
70 of the labelled source, as well as the categorical cell class. Like the previously state-of-the-art model in
71 Knox et al. (2019), this model predicts regionalized connectivity as an average over positions within
72 the structure, with nearby experiments given more weight. However, our model weighs class-specific
73 behavior in a particular structure against spatial position, so a nearby experiment targeting a similar
74 cell class would be relatively up-weighted, while a nearby experiment targeting a dissimilar class
75 would be down-weighted. This model outperforms the model of Knox et al. (2019) based on their
76 ability to predict held-out experiments in leave-one-out cross-validation. We establish a lower-limit of
77 detection, and then use the trained model to estimate overall connectivity matrices for each assayed
78 cell class.

79 The resulting cell-type specific connectivity is a directed weighted multigraph which can be
80 represented as a tensor with missing values. We do not attempt an exhaustive analysis of this data, but
81 do manually verify several cell-type specific connectivity patterns found elsewhere in the literature,
82 and show that these cell-type specific signals are behaving in expected ways. Finally, we decompose
83 the wild type connectivity matrix into factors representing archetypal connectivity patterns using
84 non-negative matrix factorization. These components are themselves novel and of some independent
85 interest since they allow approximation of the regionalized connectivity using a small set of latent
86 components. Such components can provide a link to the genetic origin of the regionalized
87 connectivity.

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

2 METHODS

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

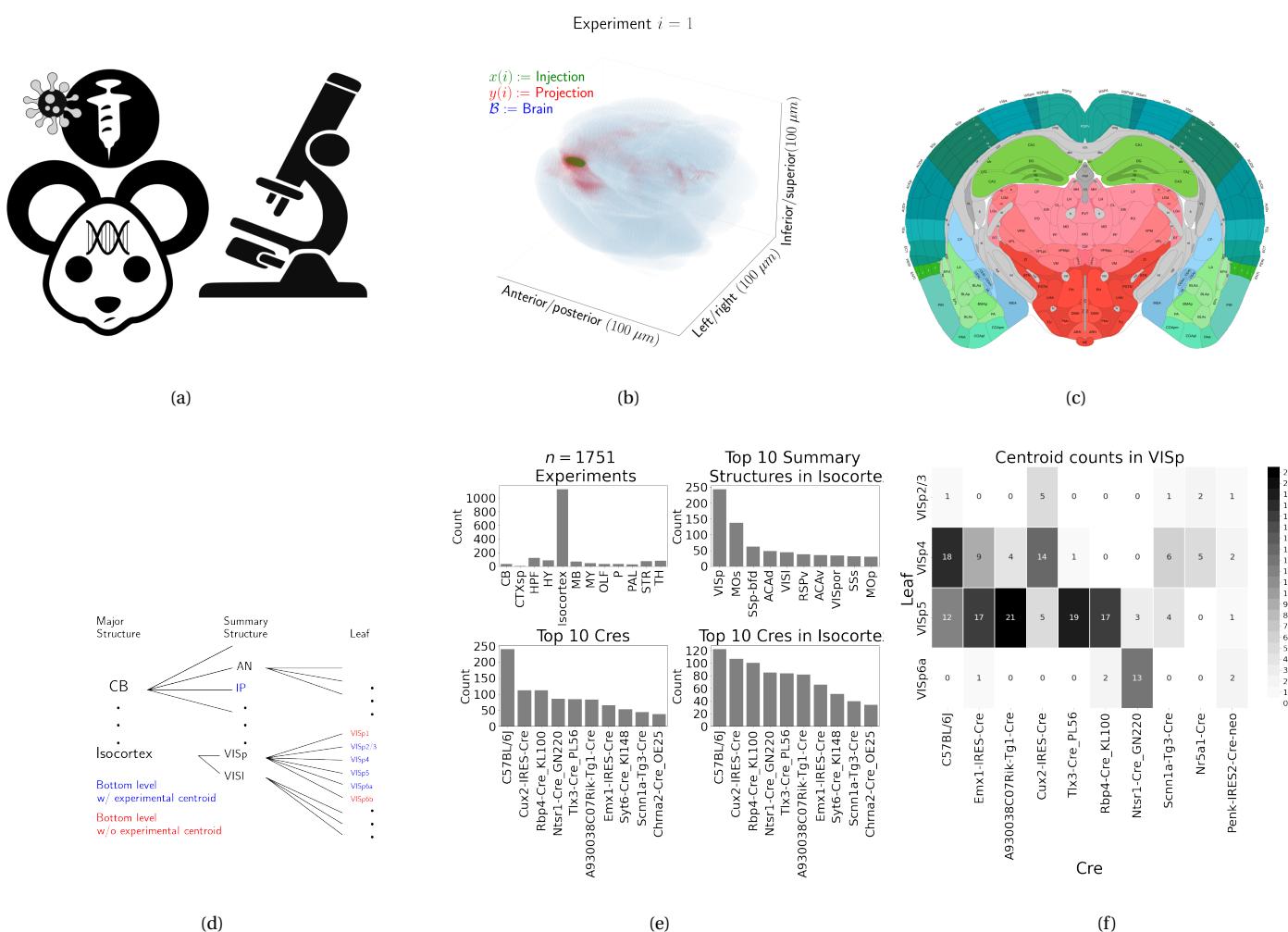


Figure 1: Experimental setting. 1a For each experiment, a Cre-dependent GFP-expressing transgene cassette is transduced by stereotaxic injection into a Cre-driver mouse, followed by serial two-photon tomography imaging. 1b An example of the segmentation of projection (targets) and injection (source)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

98 **Data**

99 Our dataset \mathcal{D} consists of $n = 1751$ publicly available murine brain viral tracing experiments from the
100 Allen Mouse Brain Connectivity Atlas. Figure 1a summarizes the multistage experimental process
101 used to generate this dataset. In each experiment, a mouse is injected with AAV encoding GFP into a
102 single location in the brain. For cell type-specific labeling, we used a double-inverted floxed (Cre
103 inducible) AAV in a Cre driver mouse, producing fluorescence that depends on Cre expression in
104 infected cells. Cre driver mice are engineered to express Cre under the control of a specific and single
105 gene promoter. This single gene promoter-mediated expression of Cre in some rare cases corresponds
106 to a very specific transcriptomic cell-type, but more frequently it corresponds to a combination of fine
107 level transcriptomic cell-types that represent broader cell classes. For example, we include
108 experiments from Cre driver lines that selectively label cell classes located in distinct cortical layers or
109 other nuclei across the whole brain. In wild type mice injected with non-Cre-dependent viruses,
110 fluorescence is induced in all infected cells in the region (not just those expressing Cre), and so the full
111 set of projection target areas are visualized from the given injected brain region. We therefore use the
112 term "cell class" here to refer to all neurons expressing Cre in a specific Cre driver mouse line.

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

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

137 ***Modeling Regionalized Connectivity***

As a continuous function, cell-class specific connectivity maps $f : \mathcal{V} \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}_{\geq 0}$, giving the directed connection of a particular cell class from one position in the brain to another. In contrast to Knox et al. (2019), which only uses wild type C57BL/6J mice, our dataset has experiments from $V = 114$ different Cre-lines $\mathcal{V} := \{v\}$. As in Knox et al. (2019), we ultimately estimate an integrated regionalized connectivity defined with respect to a set of $S = 564$ source regions $\mathcal{S} := \{s\}$ and $T = 1123$ target regions $\mathcal{T} := \{t\}$, of which $1123 - 564 = 559$ are contralateral. That is, we define

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

$$\text{normalized regionalized connectivity strength } \mathcal{C}^N : \mathcal{V} \times \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_{\geq 0} \text{ with } \mathcal{C}^N(v, s, t) = \frac{1}{|s|} \mathcal{C}(v, s, t),$$

$$\text{normalized regionalized projection density } \mathcal{C}^D : \mathcal{V} \times \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_{\geq 0} \text{ with } \mathcal{C}^D(v, s, t) = \frac{1}{|s||t|} \mathcal{C}(v, s, t).$$

138 Since the normalized strength and densities are computable from the strength via a fixed
 139 normalization, our main statistical goal is to estimate $\mathcal{C}(v, s, t)$ for all v, s and t . In other words, we
 140 want to estimate matrices $\mathcal{C}_v \in \mathbb{R}_{\geq 0}^{S \times T}$. We call this estimator $\widehat{\mathcal{C}}$.

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. Mathematically, we represent these considerations as

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

141 This makes explicit the data featurization f_* , statistical estimator \widehat{f} , and any potential subsequent
 142 transformation f^* such as summing over the source and target regions. Denoting \mathcal{D} as a function of
 143 v, s , and t reflects that different data may be used to estimate different connectivities. Table 1 reviews
 144 estimators used for this data-type used in previous work, as well as our two main extensions: the
 145 **Cre-NW** and **Expected Loss** (EL) models. Additional information on these estimators is given in
 146 Supplemental Section 6.

| Name | f^* | \hat{f} | f_* | $\mathcal{D}(v, s)$ |
|------------------------|---------------------------------|-----------------|--|------------------------|
| NNLS (Oh et al., 2014) | $\hat{f}(S)$ | 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) | $X = c, Y = y_{\mathcal{T}}$ | I_m / I_m |
| Cre-NW | $\sum_{l_s \in s} \hat{f}(l_s)$ | NW(X,Y) | $X = c, Y = y_{\mathcal{T}}$ | $(I_l \cap I_v) / I_m$ |
| Expected Loss (EL) | $\sum_{l_s \in s} \hat{f}(s)$ | EL(X, Y, v) | $X = c, Y = y_{\mathcal{T}}, v$ | 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 leaves 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 structural projection signal as a function of structural injection signal. It generates estimated connectivities for individual structures. The Nadaraya-Watson model (NW) is a local smoothing model that generates a prediction for each voxel within a structure. These predictions are averaged to create estimate the structure-specific connectivity.

147 Our contributions have several differences from the previous methods. In contrast to the
 148 non-negative least squares (Oh et al., 2014) and Nadaraya-Watson (Knox et al., 2019) estimators that
 149 take into account sources s and targets t , but not cell classes v , our new estimators specifically
 150 account for cell class. The Cre-NW estimator only uses experiments from a particular class to predict
 151 connectivity for that class, while the EL estimator shares information between classes within a
 152 structure. A detailed mathematical description of our new estimator is given in Supplemental Section
 153 6. This estimator takes into account two types of covariate information about each experiment: the
 154 centroid of the injection, and the Cre-line. Like the NW and Cre-NW estimator, the EL estimator
 155 generates predictions for each voxel in a structure, and then sums them together to get the overall
 156 connectivity. However, in contrast to these alternative approaches, when predicting the projection
 157 pattern of a certain cell-class at a particular location, the EL estimator weights the average behavior of
 158 the class in the structure containing the location in question against the locations of the various

159 nearby experiments. Thus, nearby experiments with similar Cre-lines can help generate the
160 prediction, even when there are few nearby experiments of the cell-class in question.

161 ***Model evaluation***

162 We select optimum functions from within and between our estimator classes using **leave-one-out**
 163 **cross validation**, in which the accuracy of the model is assessed by its ability to predict experiments
 164 excluded from the training data. Equation 1 includes a deterministic step f^* included without input
 165 by the data. The performance of $\widehat{\mathcal{C}}(v, s, t)$ is thus determined by performance of $\widehat{f}(f_*(\mathcal{D}(v, s)))$. Thus,
 166 we can evaluate prediction on $f_{\mathcal{T}} : \mathbb{R}^3 \rightarrow \mathbb{R}_{\geq 0}^T$ - the regionalized connection strength at a given location.

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

We use weighted $l2$ -loss to evaluate these predictions.

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

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

173 This is a somewhat different loss from Knox et al. (2019), both because of the normalization of
 174 projection, and because of the increased weighting of rarer combinations of s and v implicit in the
 175 $\frac{1}{|I_s \cap I_v|}$ term in the loss. As a final modeling step, we establish a lower limit of detection. The EL model
 176 also contains a separate cross-validation step. These approaches are covered in Supplemental Section

178 ***Connectivity analyses***

179 We examine latent structure underlying our estimated connectome using two types of unsupervised
 180 learning. Our use of hierarchical clustering is standard, and so we do not review it here. However, our
 181 application of non-negative matrix factorization (NMF) to decompose the estimated long-range
 182 connectivity into **connectivity archetypes** that linearly combine to reproduce the observed
 183 connectivity of some independent interest. Non-negative matrix factorization refers to a collection of
 184 **dictionary-learning** algorithms for decomposing a non-negatively-valued matrix such as \mathcal{C} into
 185 positively-valued matrices called, by convention, weights $W \in \mathbb{R}_{\geq 0}^{S \times q}$ and hidden units $H \in \mathbb{R}_{\geq 0}^{q \times T}$.
 186 Unlike PCA, NMF specifically accounts for the fact that data are all in the positive orthant. The matrix
 187 H is typically used to identify latent structures with interpretable biological meaning, and the choice
 188 of matrix factorization method reflects particular scientific subquestions and probabilistic
 189 interpretations.

190 Our algorithm solves the following optimization problem

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

191 For this decomposition we ignore connections between source and target regions less than $1500\mu m$
 192 apart. This is because short-range projections resulting from diffusion dominate the matrices $\hat{\mathcal{C}}$, and
 193 represent a less-interesting type of biological structure. We set $\lambda = 0.002$ to encourage sparser and
 194 therefore more interpretable components. We use unsupervised cross-validation to determine an
 195 optimum q , and show the top 15 stable components (?). Stability analysis accounts for the
 196 difficult-to-optimize NMF program by clustering the resultant H from multiple replicates. Since the
 197 NMF objective is difficult to optimize and sensitive to initialization, we follow up with a stability
 198 analysis. The medians of the component clusters appearing frequently across NMF replicates are
 199 selected as **connectivity archetypes**. Details of these approaches are given in Supplementary Sections
 200 6 and 7.

3 RESULTS

²⁰¹ We provide several types of results. First, we show that the novel expected-loss (EL) estimator
²⁰² performs best in our validation assays. Second, qualitative exploratory analysis confirms that the
²⁰³ Cre-specific connectivity matrices generated using this model are consistent with known biology.
²⁰⁴ Third, statistical decomposition of the wild-type connectivity matrix using unsupervised learning
²⁰⁵ shows how archetypal components can combine to produce observed signals.

²⁰⁶ ***Model evaluation***

²⁰⁷ Our EL model generally performs better than the other estimators that we consider. Table 5 contains
²⁰⁸ weighted losses from leave-one-out cross-validation of candidate models, such as the NW Major-WT
²⁰⁹ model from Knox et al. (2019). The EL model combines the good performance of class-specific
²¹⁰ models like NW Leaf-Cre in regions like Isocortex with the good performance of class-agnostic models
²¹¹ in regions like Thalamus. Additional information on model evaluation, including class and structure-
²¹² specific performance, is given in Appendix 5. In particular, Supplementary Table 4 contains the sizes
²¹³ of these evaluation sets in each major structure, and Supplementary Section 7 contains the structure-
²¹⁴ and class specific losses.

| | Mean Leaf-Cre | NW Major-Cre | NW Leaf-Cre | NW Leaf | NW Major-WT | NW Major | EL |
|---------------|----------------|----------------|----------------|--------------|-------------------|--------------|--------------|
| \hat{f} | Mean | NW | NW | NW | NW | 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 | 0.239 | 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 |

Table 2: Losses from leave-one-out cross-validation of candidate models. **Bold** numbers are best for their major structure.

215 ***Connectivities***

216 Our main result is the estimation of matrices $\hat{\mathcal{C}}_v \in \mathbb{R}_{\geq 0}^{S \times T}$ representing connections of source structures
 217 to target structures for particular cre-lines v . We confirm the detection of several well-established
 218 connectivities within our tensor, although it is our expectation that additional interesting biological
 219 processes are also manifest. The connectivity tensor and code to reproduce it are available at
 220 https://github.com/AllenInstitute/mouse_connectivity_models/tree/2020.

221 *Overall connectivity* Several expected biological projection patterns are evident in the wild-type
 222 connectivity matrix \mathcal{C}_{wt} from leaf sources to leaf targets shown in Figure 2a. Intraareal connectivities
 223 are clear, as are ipsilateral connections between cortex and thalamus. The clear intrastructural and
 224 intraareal connectivities mirror previous estimates in Oh et al. (2014) and Knox et al. (2019) and
 225 descriptive depictions of individual experiments in J. A. Harris et al. (2019).

226 Our estimated wild-type connectivities appear more variable than those in Knox et al. (2019), which
 227 used the NW Major-WT model whose accuracy is evaluated in Table 5. This is plausibly because of
 228 both the layer-specific targeting of the different cre-lines, and also the layer-specificity of the selected
 229 model. Although layer-specificity is a major advantage of including distinct cre-lines, for comparison,
 230 we also plot coarser projections between summary-structure sources and targets in the cortex in
 231 Figure 2b. These are averages over component layers weighted by layer size. Grossly congruent with
 232 the previous work, these results also exhibit a larger range of connectivities than those in Knox et al.
 233 (2019). Importantly, as shown in Table 5 this finer spatial resolution corresponds to the increased
 234 accuracy of our EL model over the NW Major-WT model.

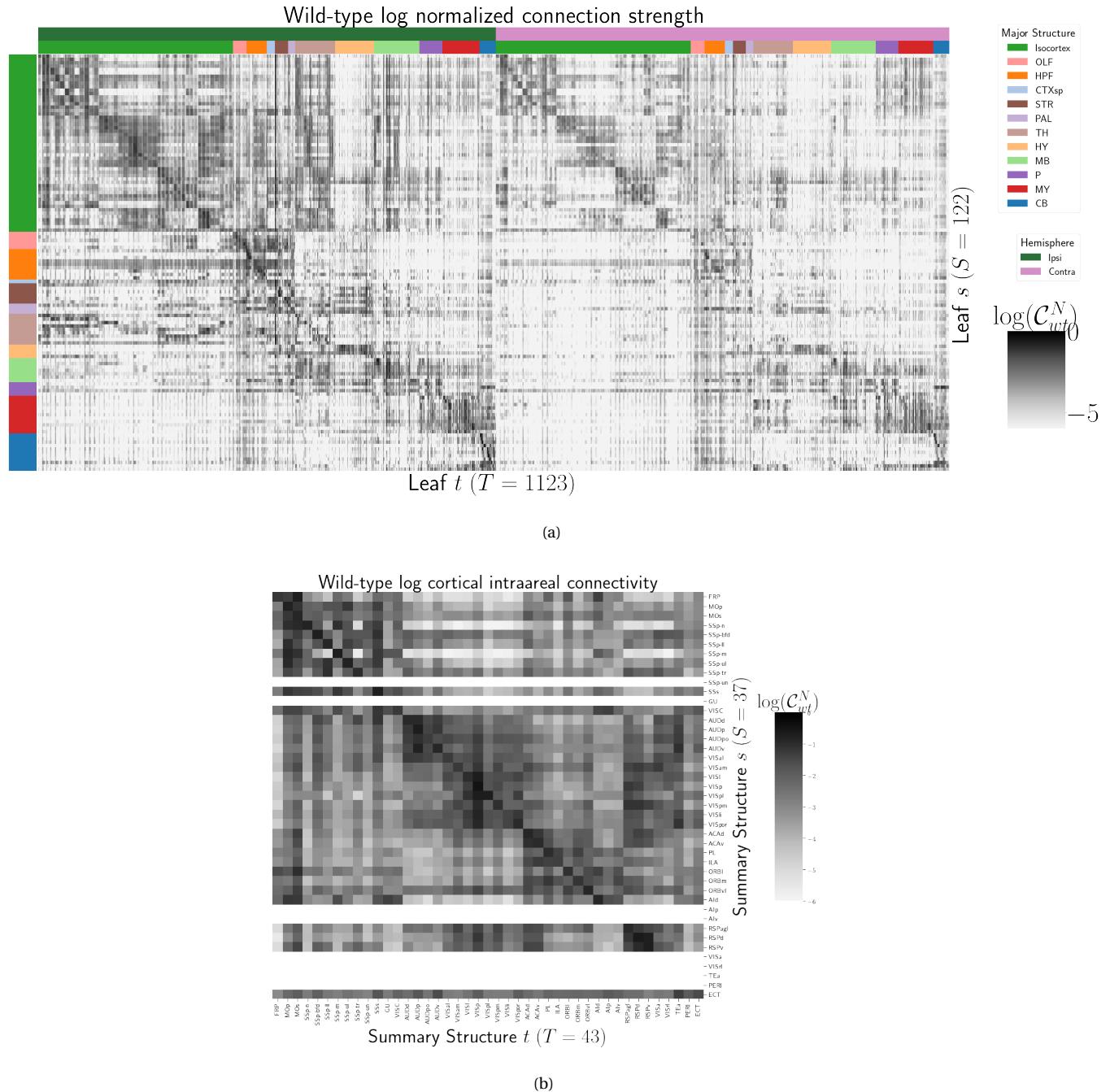


Figure 2: Wild-type connectivities. 2a Log wild-type connectivity matrix $\log \mathcal{C}(s, t, v_{wt})$. 2b Log wild-type intracortical connectivity matrix at the summary structure level.

235 *Class-specific connectivities* Source and cell-type combinations which project similarly indicate the
 236 network structure underpinning cognition. Although there is a rich anatomical literature using
 237 anterograde tracing data to describe projection patterns from subcortical sources to a small set of
 238 targets of interest, most accessible whole brain projection data are from the Allen Mouse Connectivity
 239 Atlas (MCA) project used here to generate the connectome models. Thus, to validate our results while
 240 avoiding a circular validation of the data used to generate the model weights, we confirm that these
 241 class-specific connectivities exhibit certain known behaviors. In particular, we focus on several cell
 242 types and source areas with extensive previous anatomical descriptions of projections using both bulk
 243 tracer methods with cell type specificity and single cell reconstructions: 1) thalamic-projecting
 244 neurons in the visual and motor cortical regions, 2) cholinergic neurons in the medial septum and
 245 nucleus of the diagonal band (MS/NDB); 3) cholinergic neurons in the caudoputamen, and 4)
 246 serotonergic neurons of the dorsal raphe nucleus (DR). We find that our inferred connections are in
 247 agreement with literature on these cell types.

248 DEPENDENCE OF THALAMIC CONNECTION ON CORTICAL LAYER. Visual cortical areas VISp and VISl and
 249 cortical motor areas MOp and MOs are ideal testbeds for our connectivities because there are
 250 well-established layer-specific projection patterns that can be labeled with the layer-specific cre-lines
 251 from the Allen datasets and others J. A. Harris et al. (2019); Jeong et al. (2016), and are also
 252 well-represented in our dataset. Figure 3a shows that in VISp, the Ntsr1-Cre line strongly targets the
 253 core part of the thalamic LGd nucleus while in VISl has a very strong projection to the LP nucleus. In
 254 VISp, the the Rbp-4Cre strongly targets LP as well. Recall that we display connectivity estimates for
 255 structures with at least one injection centroid in the structure. Thus, the position of non-zero rows in
 256 Figure 3a shows the localization of Rbp4-Cre and Ntsr1-Cre injection centroids to layers 5 and 6
 257 respectively. This is further examined in Supplemental Section ??). Thus, as a heuristic alternative
 258 model, to also synthesize information about leafs targeted by different cre-lines, we also display an
 259 average connectivity matrix over all cre-lines. This combined output is not evaluated in our testing,
 260 and is only a general stand-in for overall behavior, but provides a useful summary of results.

261 MS AND NDB PROJECTIONS IN THE CHAT-IRES-CRE-NEO MODEL. Cholinergic neurons in the MS and
 262 NDB are well-known to strongly innervate the hippocampus, olfactory bulb, piriform cortex,
 263 entorhinal cortex, and lateral hypothalamus (Watson, Paxinos, & Puelles, 2012; Zaborszky et al., 2015).

264 In the Allen MCA, cholinergic neurons were labeled by injections into Chat-IRES-Cre-neo mice. We
265 first checked the estimated connectome weights to targets in these major brain divisions from MS and
266 NDB. We observed that all these expected divisions were represented above the 90th percentile of
267 weights. Recently, a single cell whole brain mapping project using Chat-Cre mice fully reconstructed
268 n=50 cells, revealing these same major targets and also naming additional targets from MS/NDB (Li et
269 al., 2018). We compared our Chat-IRES-Cre connectome model data for MS and NDB with the targets
270 identified by Li et al. (2018). We identified 150 targets at the fine leaf structure level in the top 10% of
271 estimated weights. To directly compare our data across studies, we had to merge structures as needed
272 to get to the same ontology level and remove ipsilateral and contralateral information. After
273 formatting our data, we found 51 targets in the top 10%; Li et al. reported 47 targets across the 50 cells.
274 There was good consistency overall between the target sets; 35 targets were shared, 12 were unique to
275 the single cell dataset, and 16 unique to our model data. We checked whether targets missing from
276 our dataset were because of the threshold level. Indeed, lowering the threshold to the 75 th percentile
277 confirmed 6 more targets-in-common, and all but 2 targets from Li et al. (2018) were above the 50th
278 percentile weights in our model. Of note, the absence of a target in the single cell dataset that was
279 identified in our model data is most likely due to the sparse sampling of all possible projections from
280 only n=50 MS/NDB cells.

281 CP PROJECTIONS IN THE CHAT-IRES-CRE-NEO MODEL. Most cells in the caudoputamen (CP) are
282 GABAergic spiny projection neurons. These cells are also the only type that send projections outside
283 the CP. Cholinergic interneurons make up 1-2% of all CP cells and their axon terminals do not extend
284 beyond the CP borders. We confirmed that the model predictions for connection weights from CP
285 cholinergic cells were consistent with this known anatomy; the connection weight to CP was ~ 2-fold
286 higher than any other in the top 5%.

287 DR PROJECTIONS IN THE SLC6A4-CRE'ET33 MODEL. Serotonergic projections from cells in the
288 dorsal raphe (DR) are widely distributed and innervate many forebrain structures including isocortex
289 and amygdala. In the Allen MCA, serotonergic neurons were labeled using Slc6a4-Cre_ET33 and
290 Slc6a4-CreERT2_EZ13 mice. This small nucleus appears to contain a complex mix of molecularly
291 distinct serotonergic neuron subtypes with some hints of subtype-specific projection patterns (Huang
292 et al., 2019; Ren et al., 2018, 2019). We expect that the Cre lines we used here in the Allen MCA, which

293 utilize the serotonin transporter promoter (*Slc6a4-Cre* and *-CreERT2*), will lead to expression of tracer
 294 in all the serotonergic subtypes recently described in an unbiased way, but this assumption has not
 295 been tested directly. We compared our model data to a single cell reconstruction dataset consisting of
 296 n=50 serotonergic cells with somas in the DR that also had bulk tracer validation ([SK's](#)
 297 [comment:INSERT LINK TO FIGURE](#)). Ren et al. (2019) listed 55 targets across the single cell
 298 reconstructions. After processing our data to match the target structure ontology level across studies,
 299 we identified 37 targets from the DR with weights above the 90 th percentile; 27 of these targets
 300 matched those named by Ren et al. (2019). Overall there was good consistency between targets in
 301 olfactory areas, cortical subplate, CP, ACB and amygdala areas, as well in pallidum and midbrain.

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

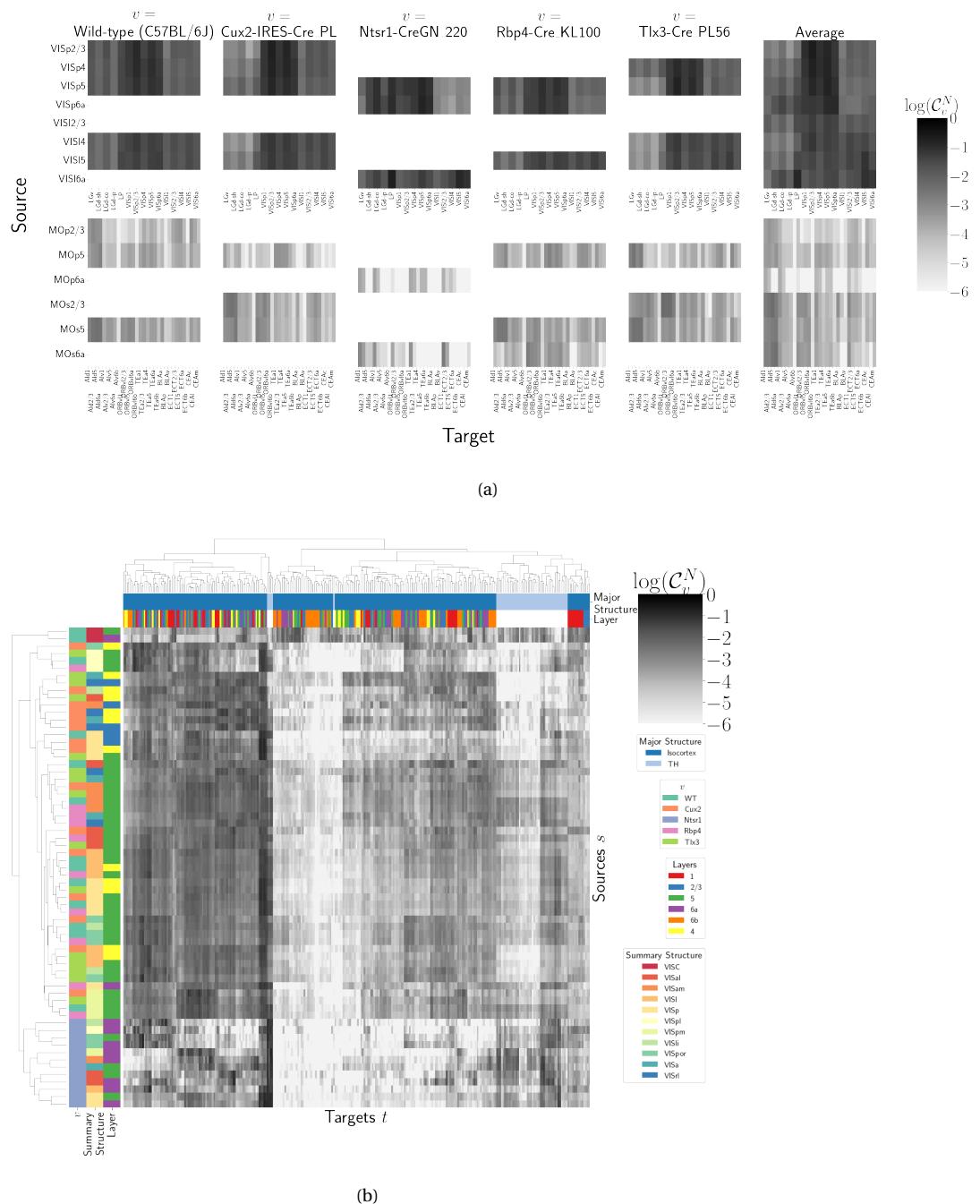


Figure 3: Cell-class specificity. 3a Selected cell-class and layer specific connectivities from two visual and two motor areas. Sources without an injection in the Cre driver line are not estimated due to lack of data for that Cre-line in that structure. 3b Hierarchical clustering of connectivity strengths 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.

318 ***Connectivity Analyses***319 **(SK's comment:move to connectivity analysis section)**

320 Cell-class, while often correlated with cortical layer, can be a stronger driver of connectivity than
 321 summary structure especially when looking at targets at major brain division level. Figure 3b shows a
 322 collection of connectivity strengths generated using cre-specific models for wild-type, Cux2, Ntsr1,
 323 Rbp4, and Tlx3 cre-lines from VIS areas at leaf level in the cortex to cortical and thalamic nuclei. We
 324 use hierarchical clustering to sort source structure/cell-class combinations by the similarity of their
 325 structural projections, and sort target structures by the structures from which they receive projections.
 326 Examining the former, we can see that the layer 6 Ntsr1 Cre-line distinctly projects to thalamic nuclei,
 327 regardless of source summary structure. This contrasts with the tendency of other cell-classes to
 328 project intracortically in a manner determined by the source structure. Similarly, layer 6 targets are
 329 not strongly projected to by any of the displayed Cre-lines. There are too many targeted summary
 330 structures to plot here, but we expect that the source profile of each target clusters by structure.

331 Each regionalized connectivity matrix is a high-dimensional realization of relatively few biological
 332 processes, and decomposition of neural signals to recover these processes is a fundamental goal in
 333 neuroscience. In this section, we apply non-negative matrix factorization to decompose the
 334 long-range wild-type connectivities into linear combinations of archetypal connectivities. This
 335 decomposes the remaining censored connectivity matrix into a linear model based off a relatively
 336 small number of distinct signals. This model is able to capture a large amount of the observed
 337 variability, and recovers structure-specific archetypal signals.

338 These signals are plotted in Figure 4, and technical details and intermediate results are given in
 339 Supplemental Sections 6 and 7, respectively. These details include a cross-validation based method
 340 for selecting the number of components, a masking method for focusing only on long range
 341 connections, and a stability method for ensuring that the decomposition is reliable across
 342 computational replicates. The plotted decomposition shows that these underlying connectivity
 343 archetypes correspond strongly to major brain division in both target and sources.

344 Inspection of the reconstructed distal normalized connection strength using the top 15
 345 components shows qualitatively shows that this relatively sparse decomposition is able to capture

³⁴⁶ much of the observed variability. Other connectivity patterns like cortical-cortical and
³⁴⁷ cortical-thalamic are also detected.

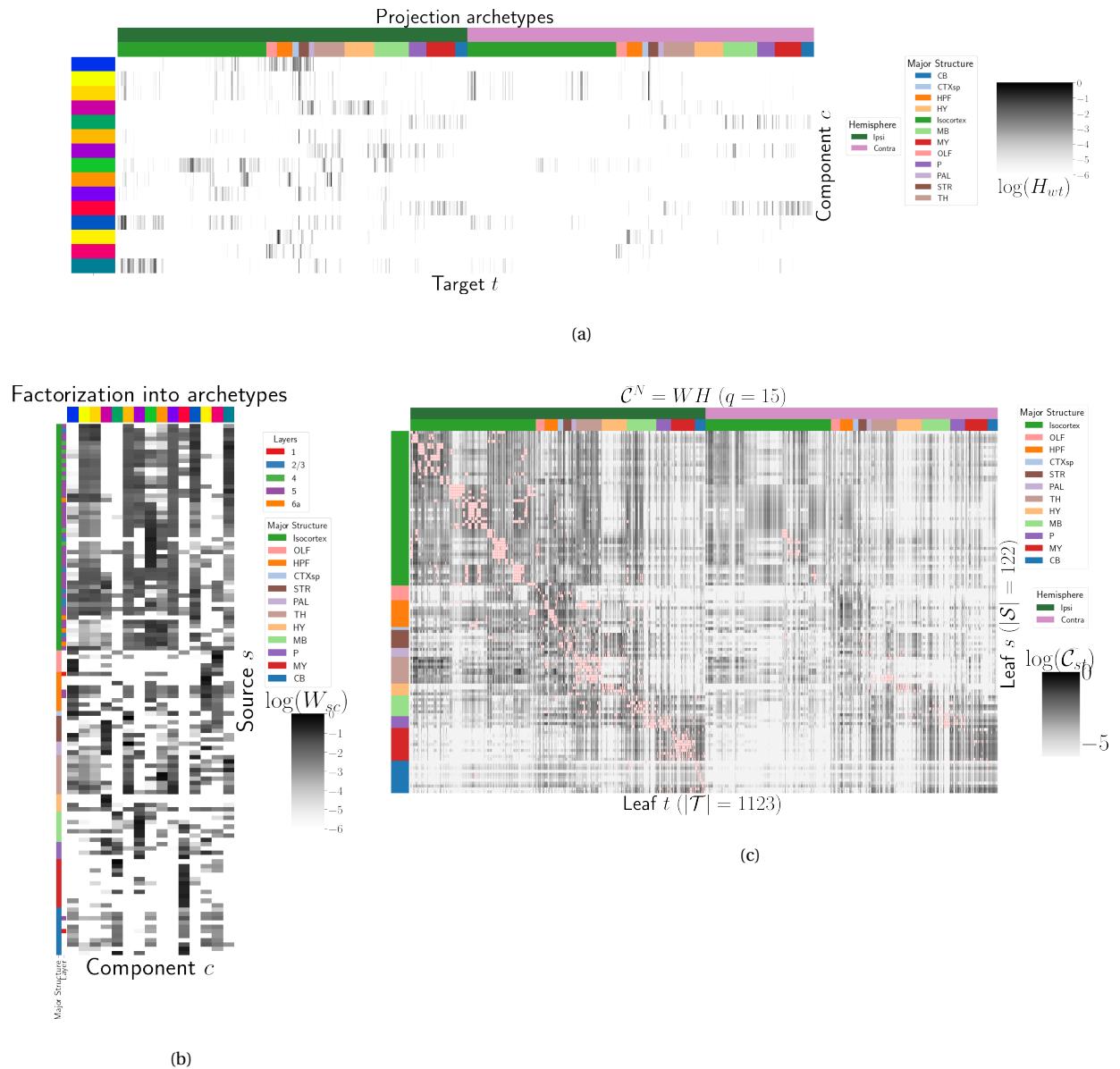


Figure 4: Non-negative matrix factorization results $\mathcal{C}_w^N = W H$ for $q = 15$ components. 4a Latent space coordinates H of \mathcal{C} . Target major structure and hemisphere are plotted. 4b Loading matrix W . Source major structure and layer are plotted. 4c Reconstruction of the normalized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

4 DISCUSSION

348 The model presented here is the first cell-type specific whole brain projectome model for a
 349 mammalian species, and it opens the door for a large number of models linking brain structure to
 350 computational architectures. Overall, we find expected targets, based on our anatomical expertise
 351 and published reports, but underscore that the core utility of this bulk connectivity analysis is not
 352 only in validation of existing connection patterns, but also in identification of new ones. We note that
 353 although the concordance appeared stronger for the cholinergic cells than the serotonergic cells, any
 354 differences might still be explained by the lack of high quality “ground-truth” datasets to validate
 355 these Cre connectome models. Larger numbers of single cell reconstructions, that saturate all
 356 possible projection types, would be a better gold standard than the small number of cells reported
 357 here (n=50 for each). ([SK's comment: add citation](#)) Perhaps future iterations of the connectome model
 358 may also take into account some single cell axon projection data.

359 The Nadaraya-Watson estimator presented here is novel. Beyond using a Nadaraya-Watson kernel
 360 regression defined in physical space, we define a cell-type space based on similarities of projections,
 361 and theoretically justify the use of an intermediate shape-constrained estimator. While methods like
 362 non-negative least squares can also account for covariates, the centroid method from Knox et al.
 363 (2019) was shown that the more precise notion of injection location than the non-negative least
 364 squares in Oh et al. (2014). Furthermore, our sample size seems too low to utilize a fixed or mixed
 365 effect, particularly since the impact of the virus depend on the particular injection region. In a sense
 366 both the NNLS and NW models can be thought of as improvements over the structure-specific
 367 average, and so is also possible that a yet undeveloped residual-based data-driven blend of these
 368 models could provide improved performance.

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

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

379 The factorization of the connectivity matrix could also be improved and better used. From a
380 statistical perspective, stability-based method for establishing archetypal connectivities in NMF is
381 similar to those applied to genomic data Kotliar et al. (2019); Wu et al. (2016). However, non-linear
382 data transformations or matrix decompositions, or tensor factorizations that account for correlations
383 between cell-types could better capture the true nature of archetypal neural connections. It would
384 also be of great interest to associate the archetypal signals detected from connectivity analysis with
385 undergirding gene expression patterns or functional information.

ACKNOWLEDGMENTS

³⁸⁶ We thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement,
³⁸⁷ and support.

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

5 SUPPLEMENTAL INFORMATION

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

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

395 This section describes the abundances of leaf and Cre-line combinations in our dataset. Users of the
396 connectivity matrices who are interested in a particular Cre-line or structure can see the quantity and
397 type of data used to compute and evaluate that connectivity.

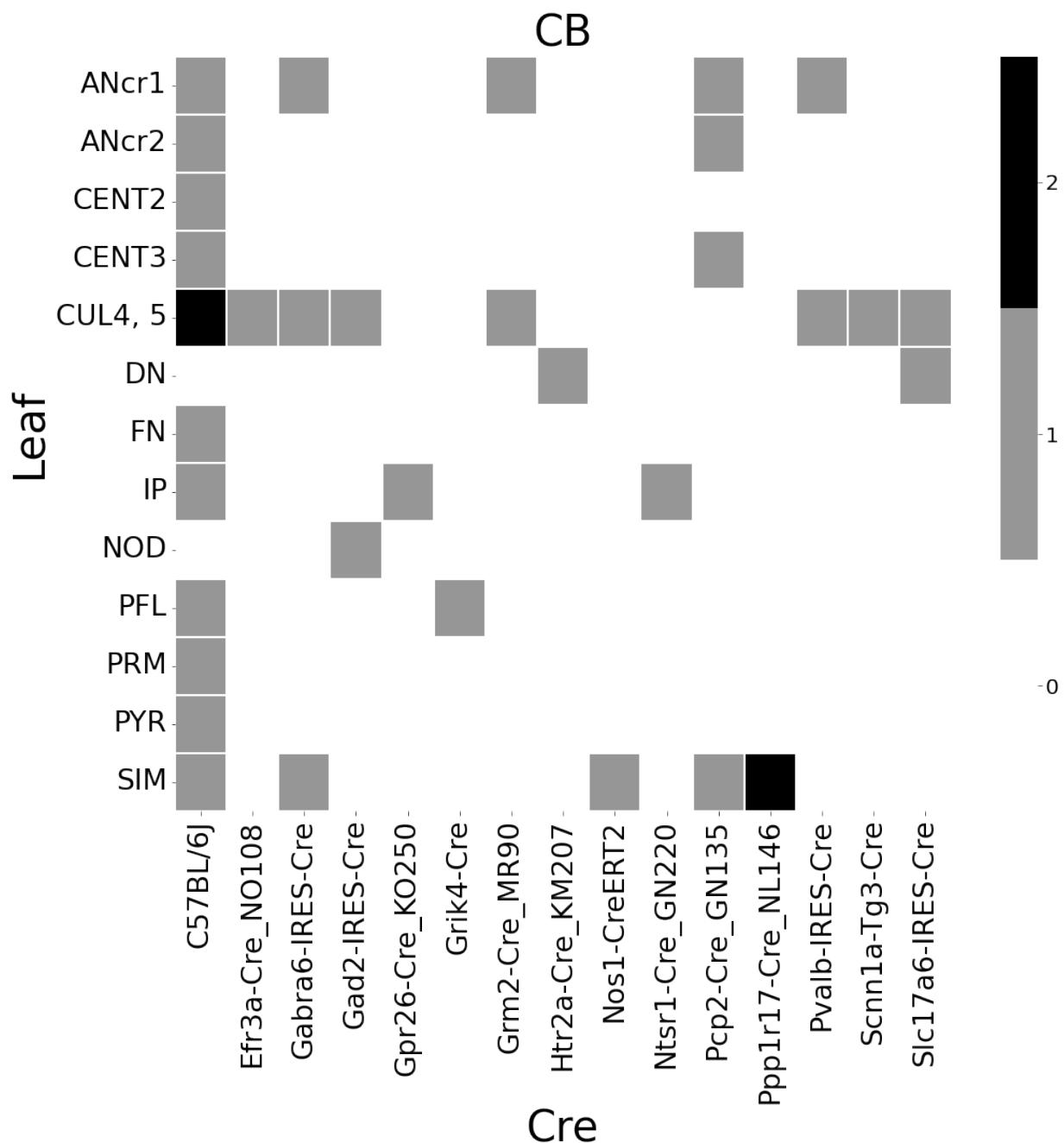


Figure 5: Abundances of Cre-line and leaf-centroid combinations.

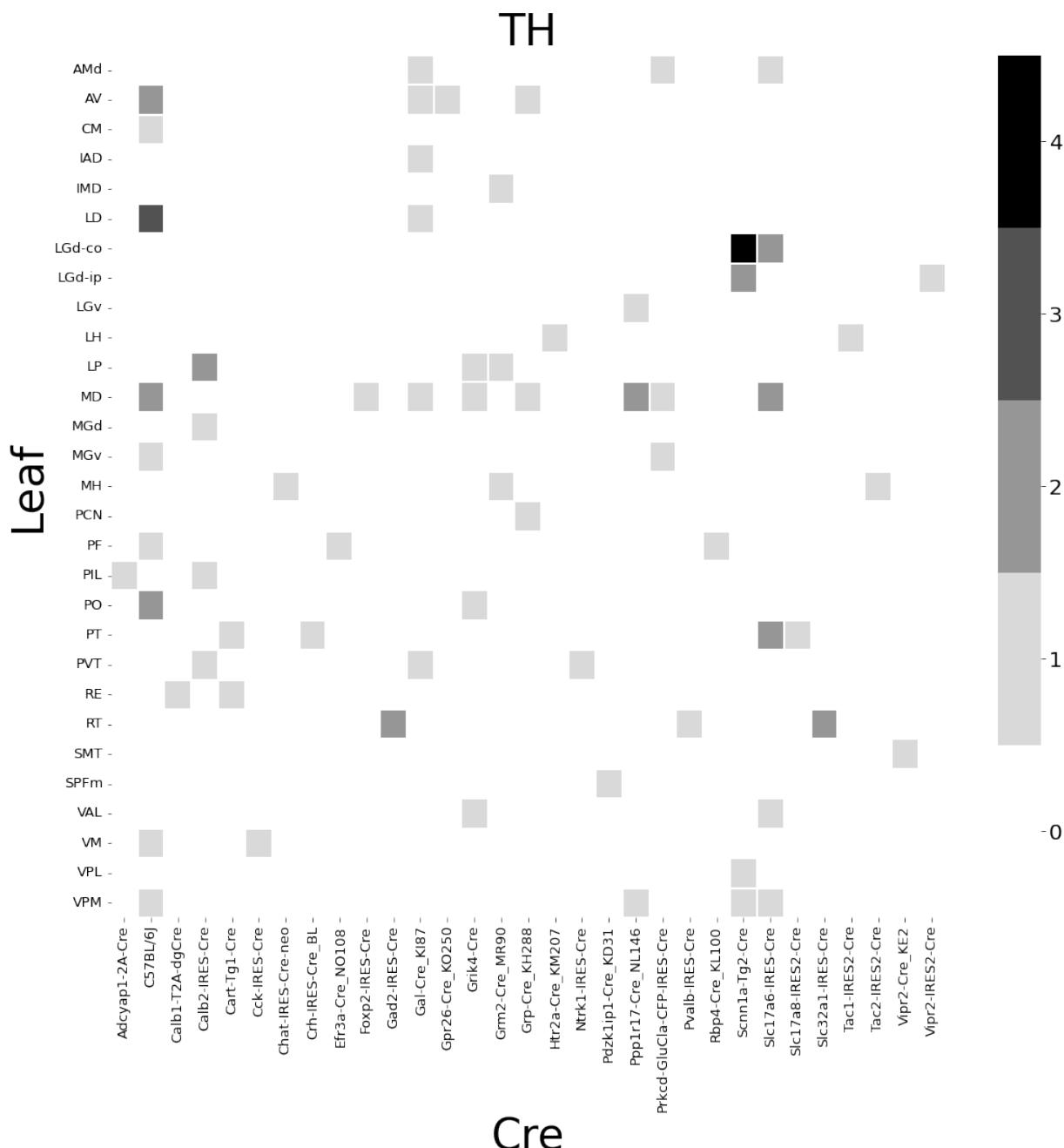


Figure 6: Abundances of Cre-line and leaf-centroid combinations.

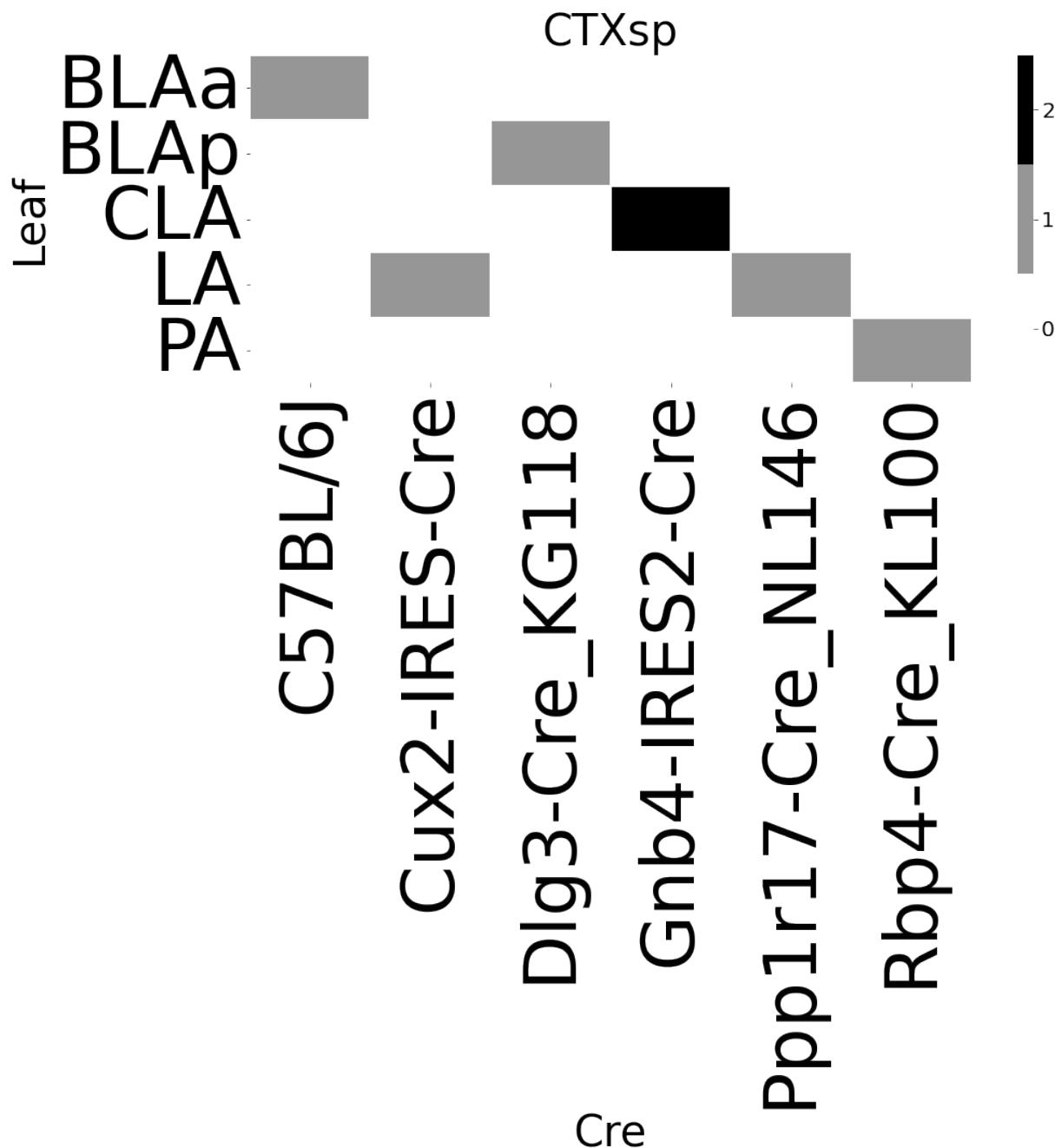


Figure 7: Abundances of Cre-line and leaf-centroid combinations.

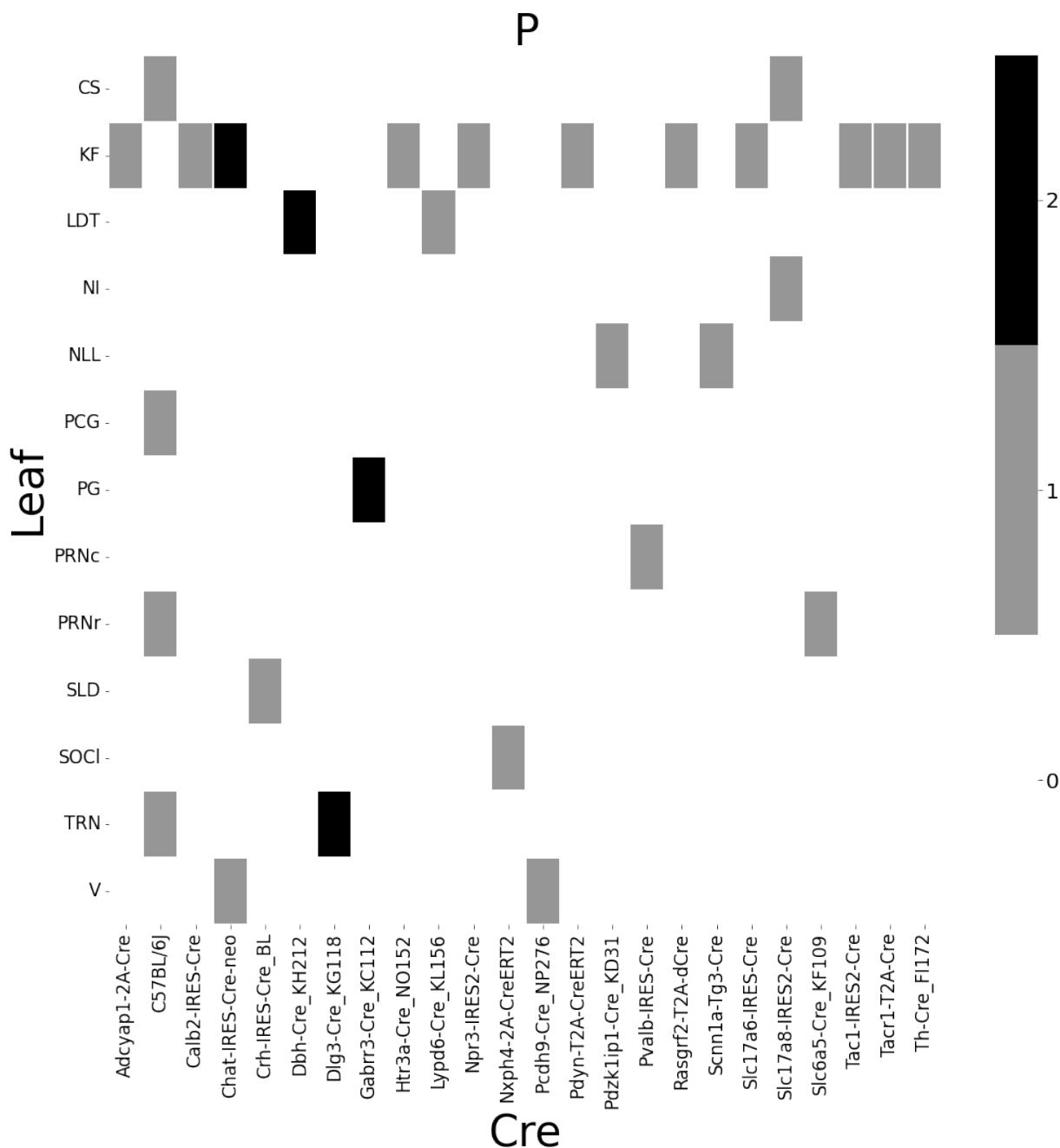


Figure 8: Abundances of cre-line and leaf-centroid combinations.

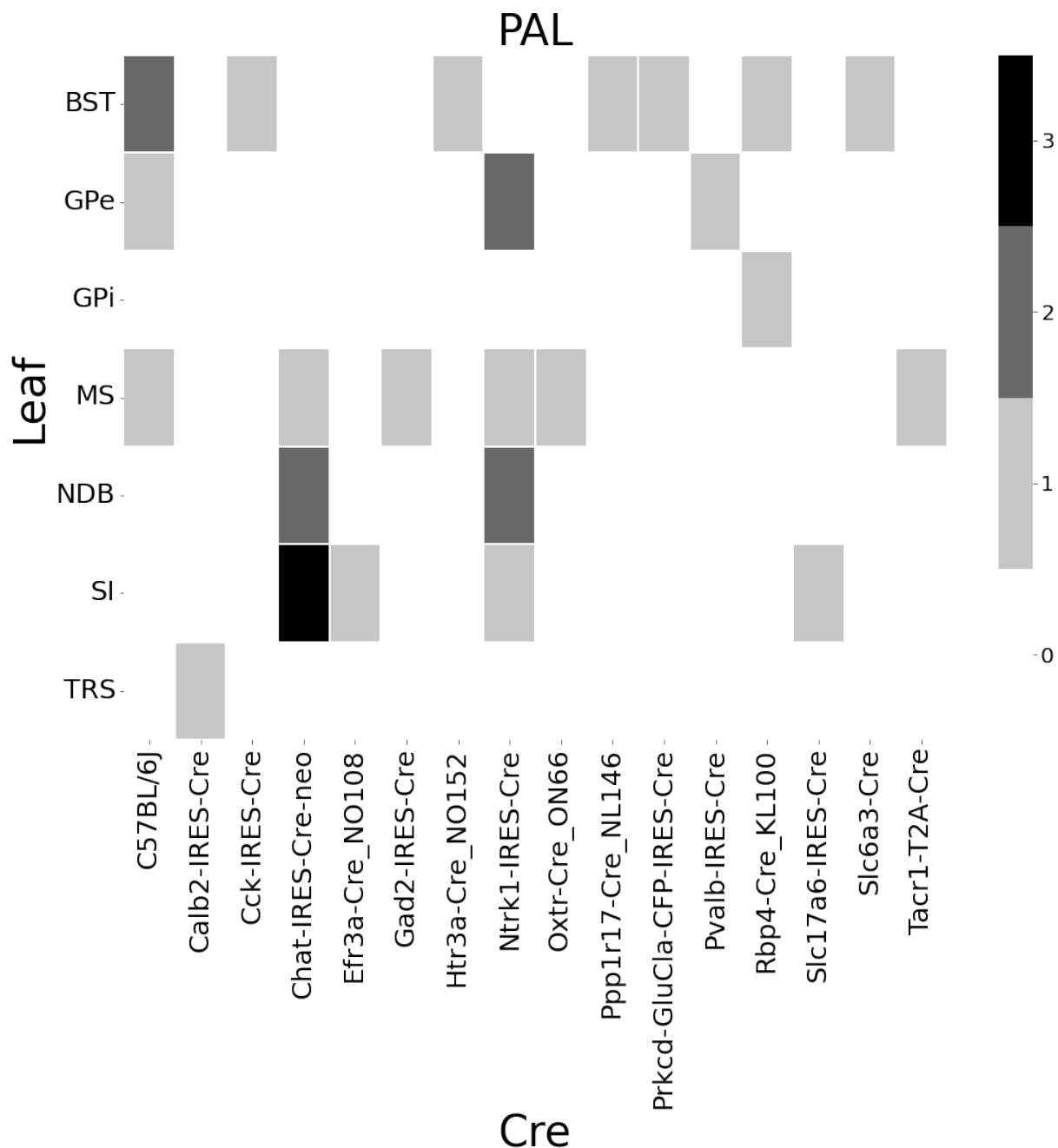


Figure 9: Abundances of Cre-line and leaf-centroid combinations.

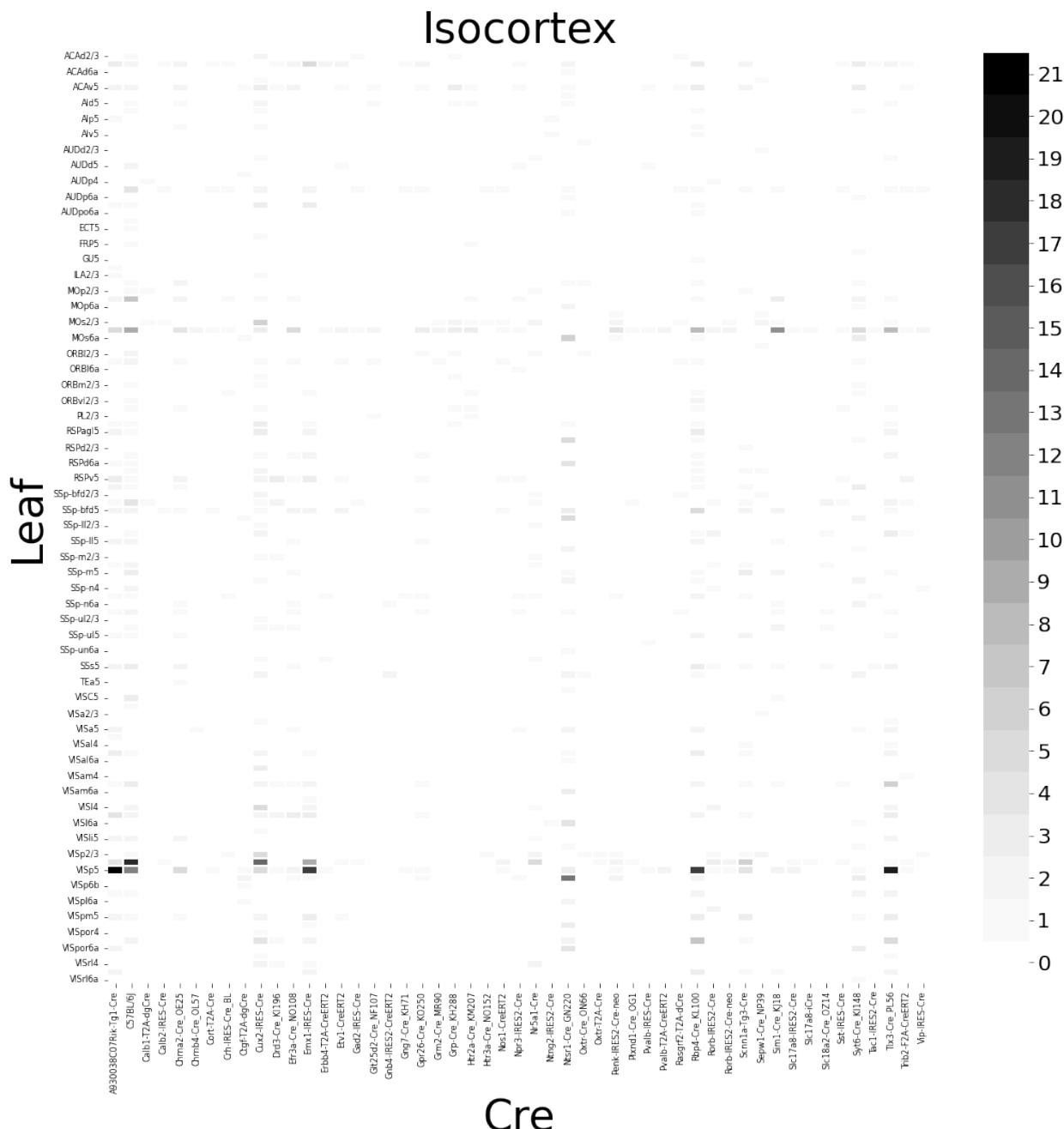


Figure 10: Abundances of Cre-line and leaf-centroid combinations.

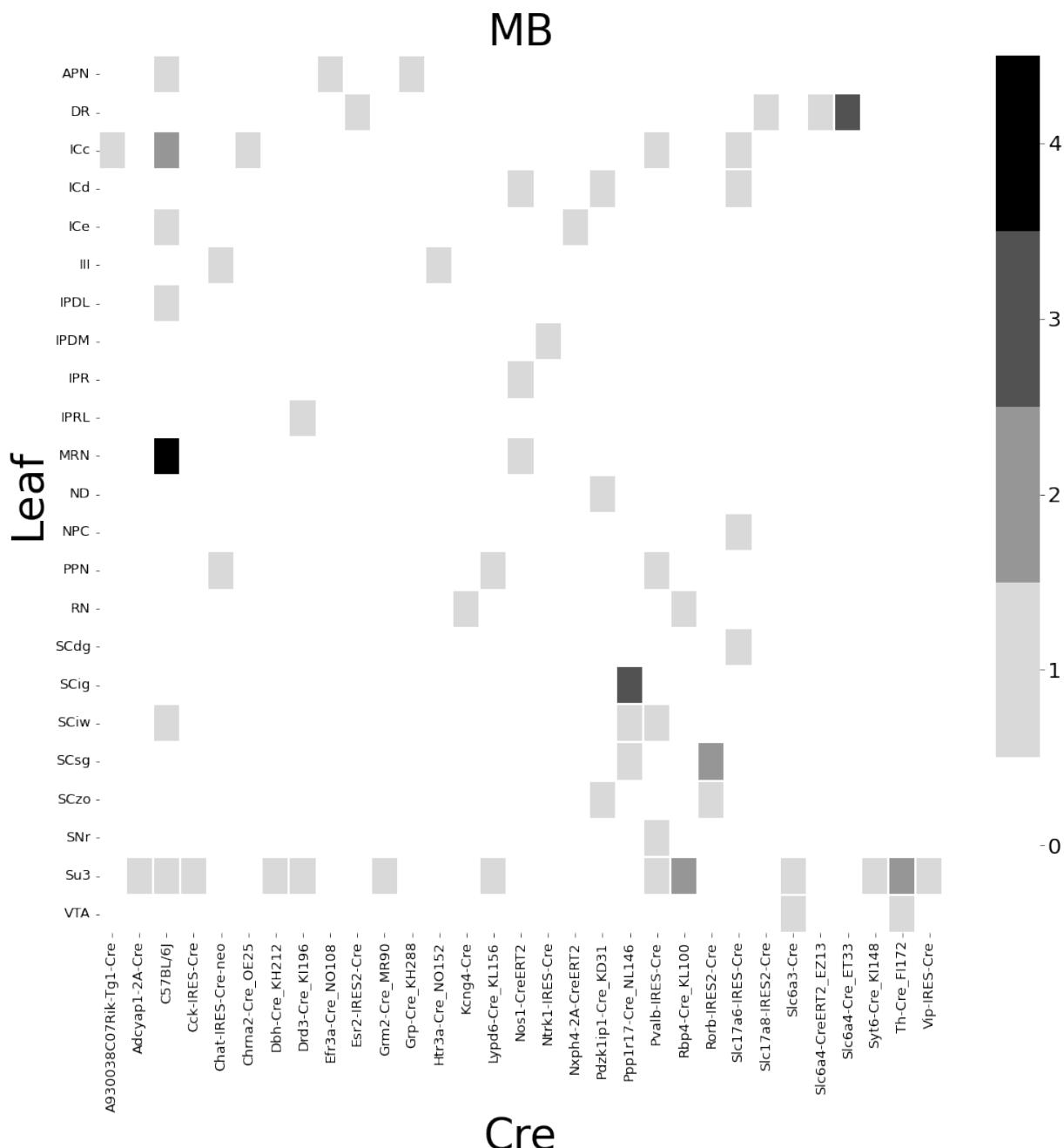


Figure 11: Abundances of Cre-line and leaf-centroid combinations.

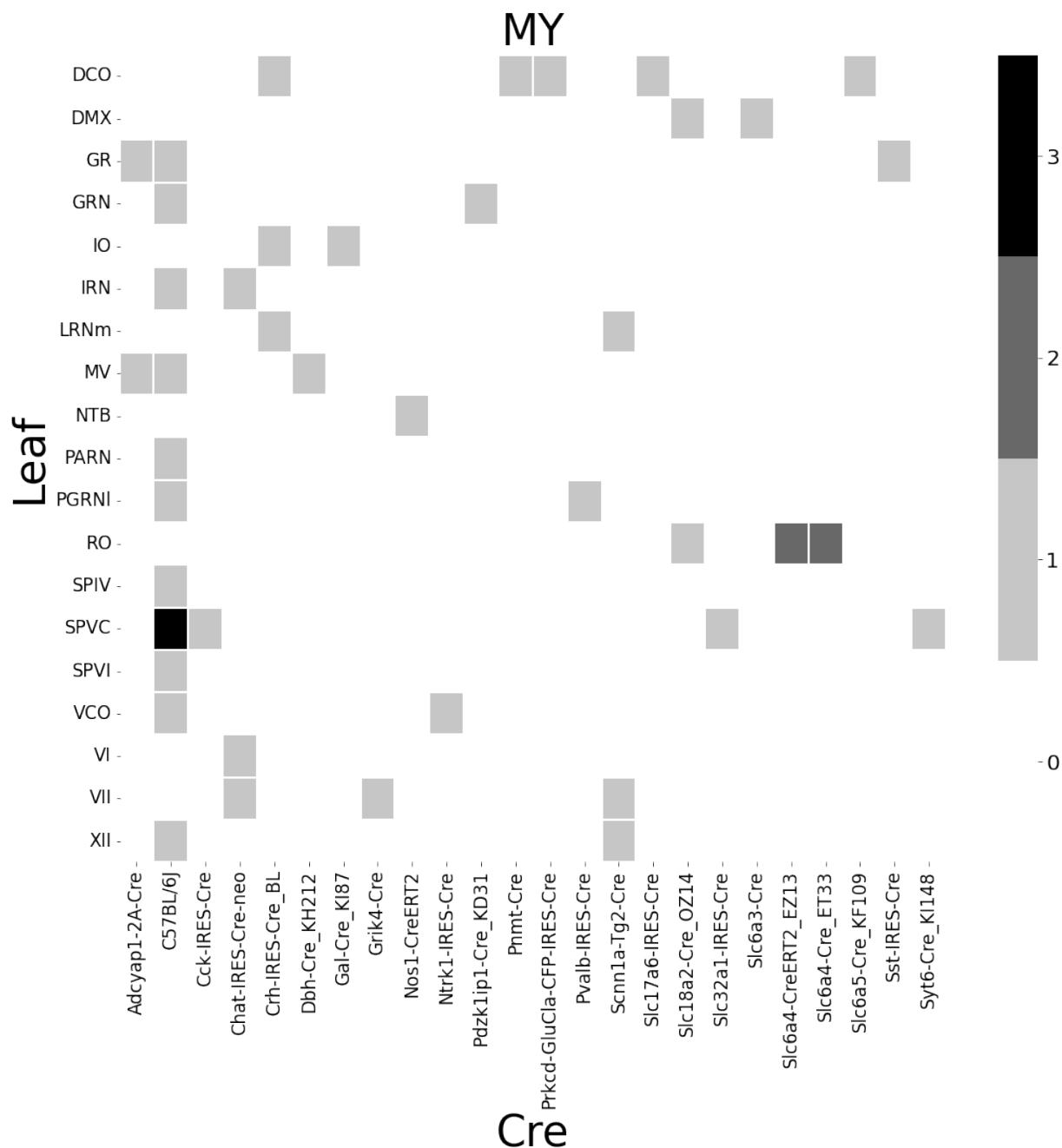


Figure 12: Abundances of Cre-line and leaf-centroid combinations.

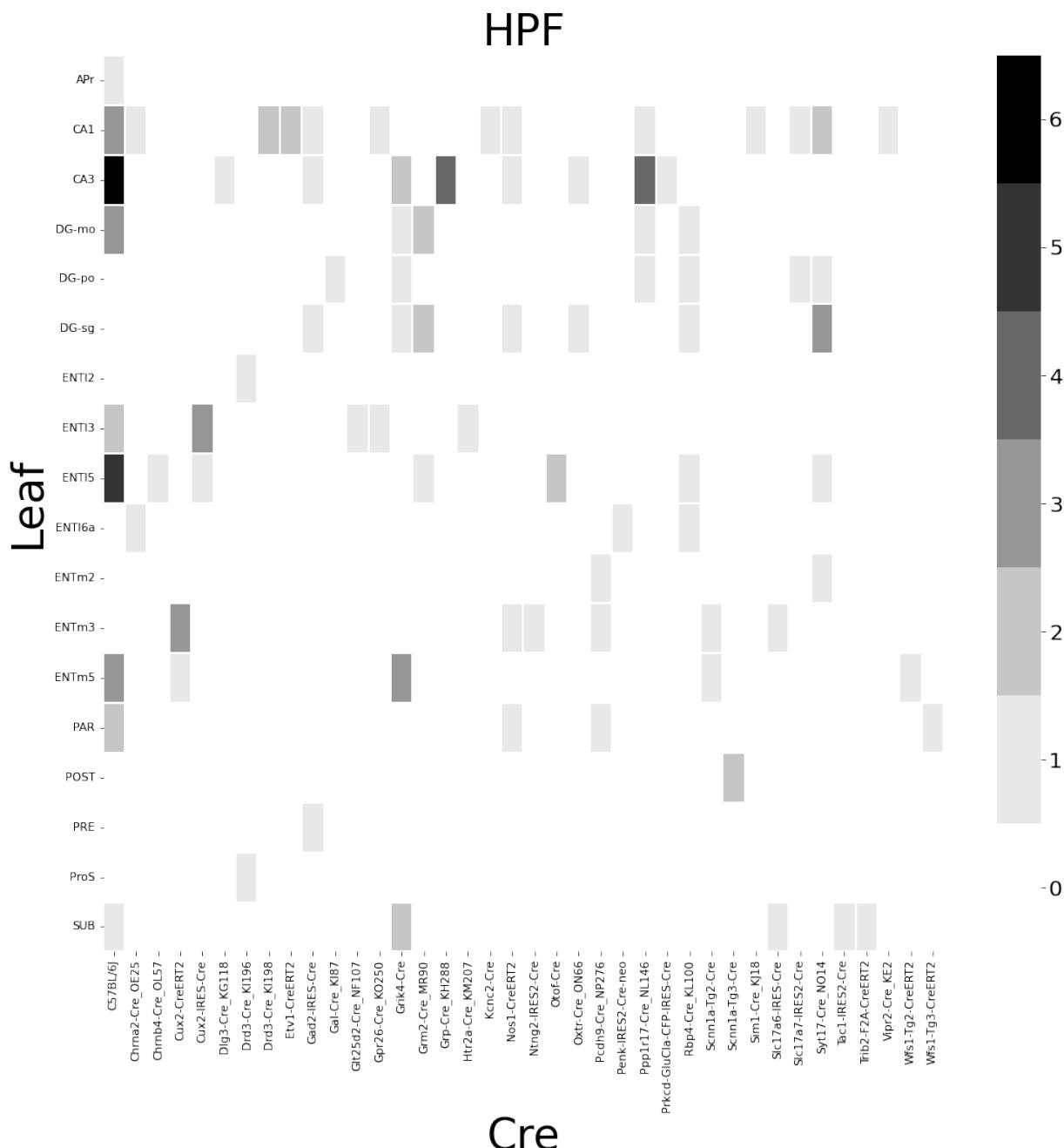


Figure 13: Abundances of Cre-line and leaf-centroid combinations.

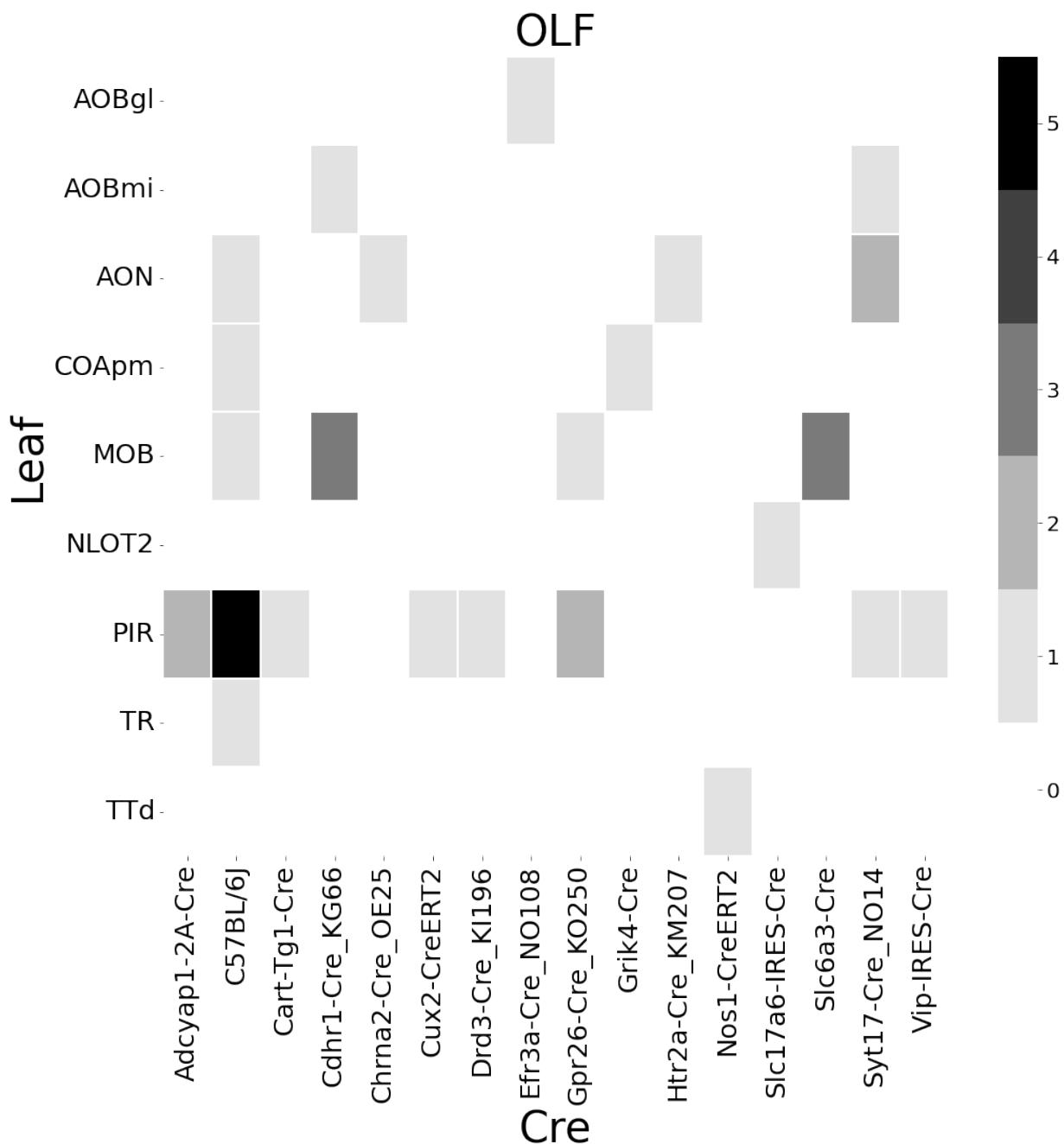


Figure 14: Abundances of Cre-line and leaf-centroid combinations.

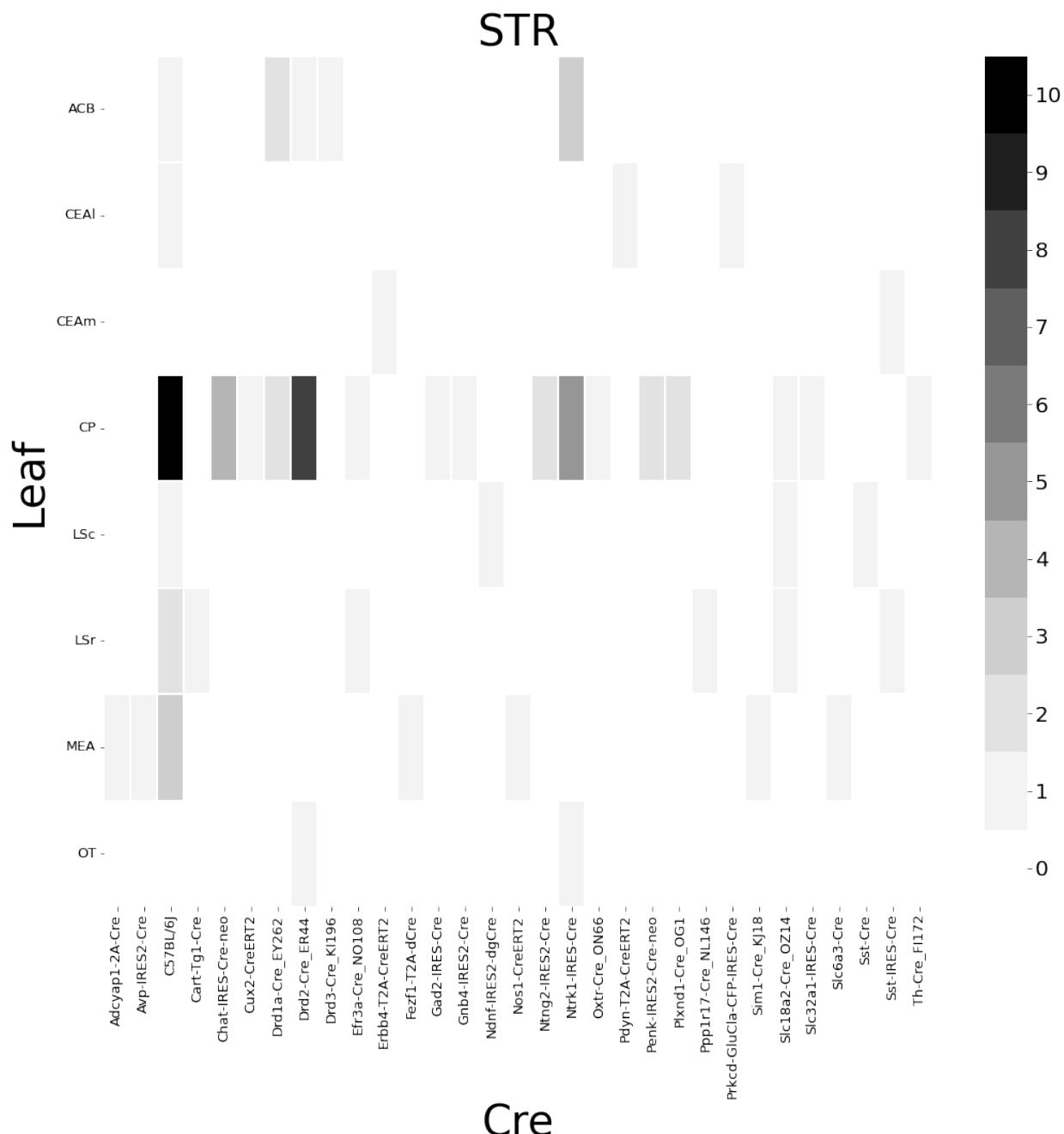


Figure 15: Abundances of Cre-line and leaf-centroid combinations.

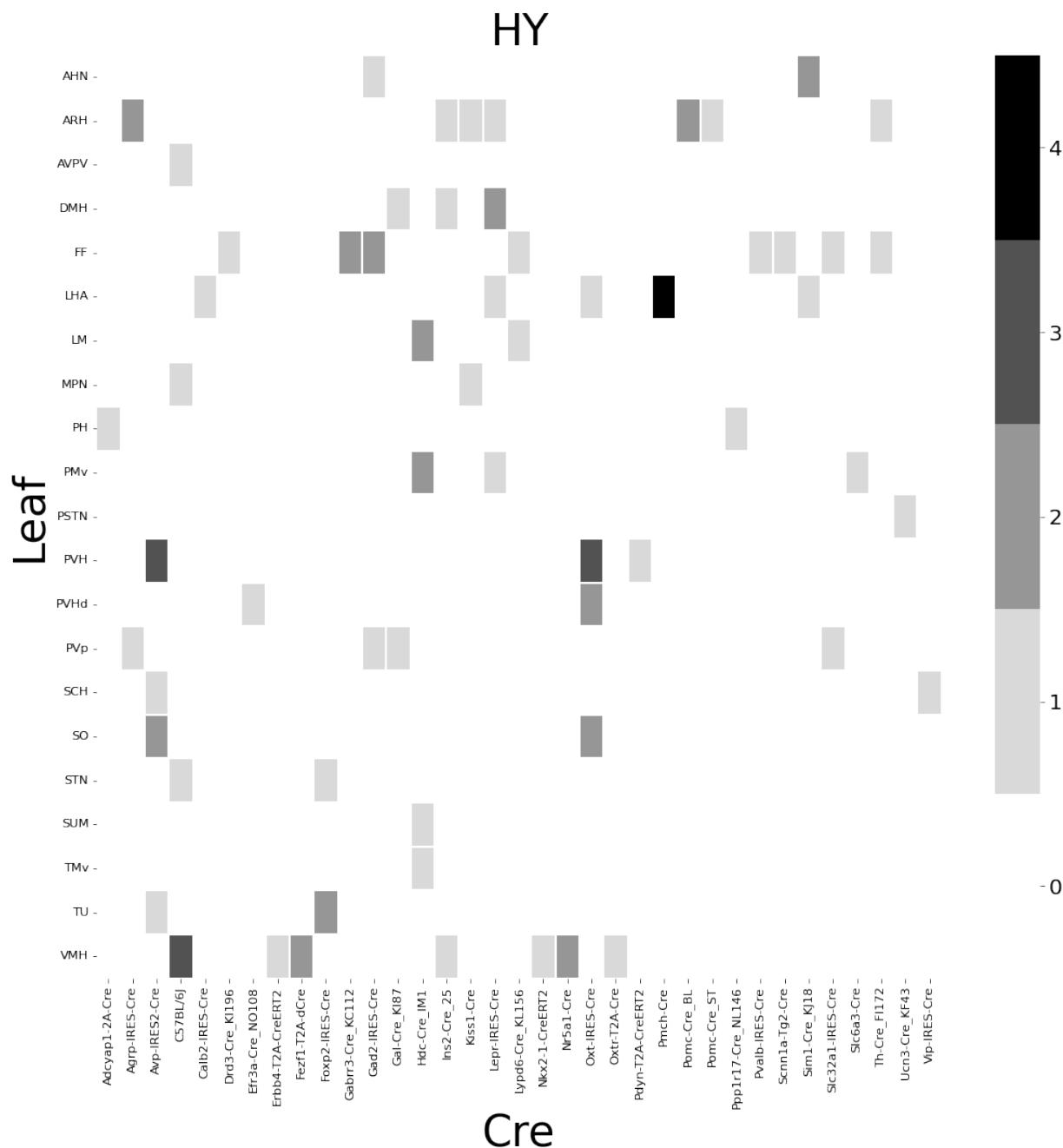


Figure 16: Abundances of Cre-line and leaf-centroid combinations.

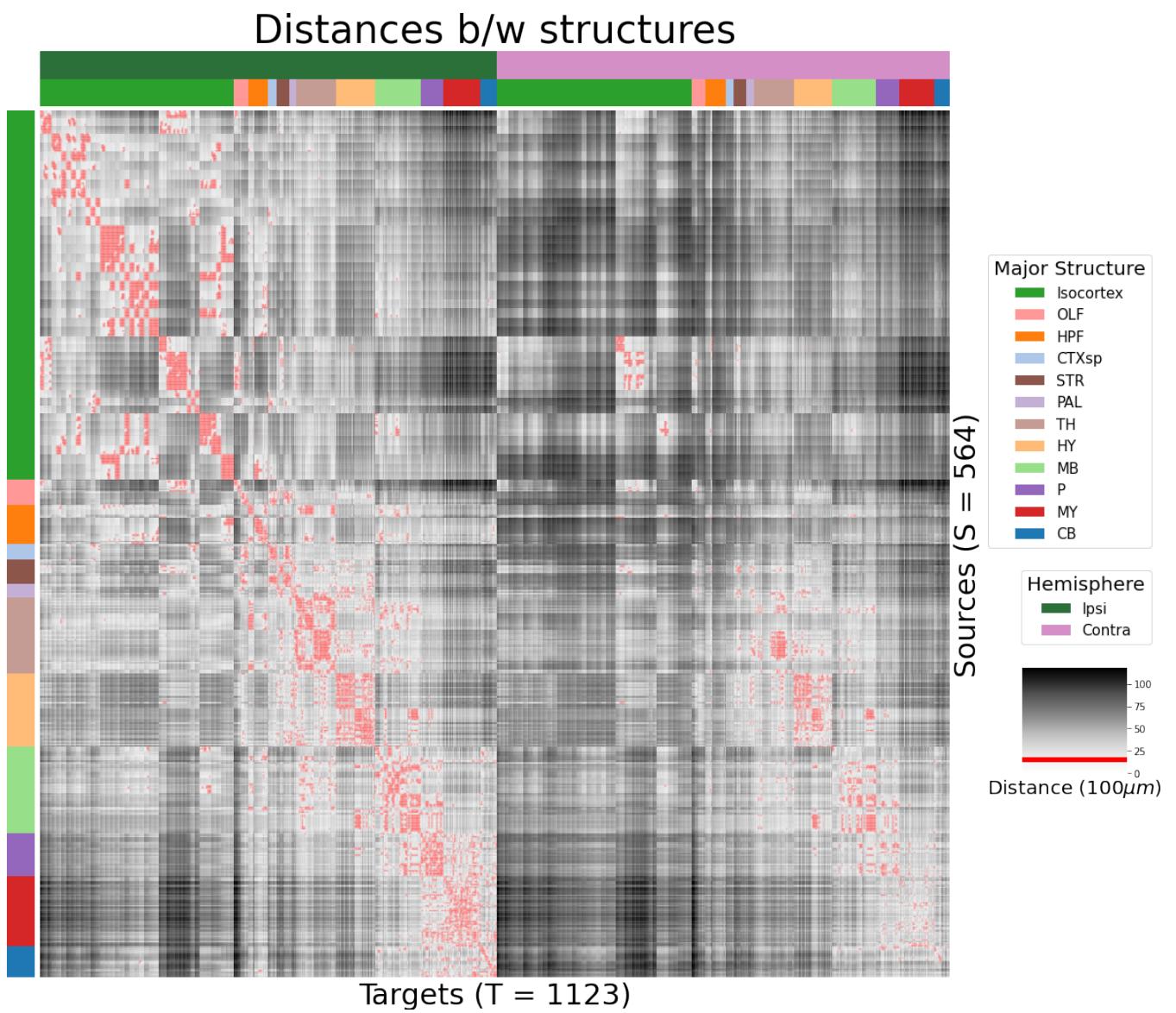
398 ***Distances between structures***

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

399 ***Model evaluation***

400 We give information on the quality of our models. This includes the sizes of our evaluation sets in
 401 leave-one-out cross-validation and additional losses in the injection-normalized case.

402 NUMBER OF EXPERIMENTS IN EVALUATION SETS In order to compare between methods, we therefore
 403 restrict to the smallest set of evaluation indices, which is to say, virus-leaf combinations that are
 404 present at least twice. This means that our evaluation set is smaller in size than our overall list of
 405 experiments.

| | Total | Cre-Leaf | Cre-leaf over threshold |
|-----------|-------|----------|-------------------------|
| Isocortex | 36 | 4 | 4 |
| OLF | 7 | 2 | 2 |
| HPF | 122 | 62 | 59 |
| CTXsp | 85 | 41 | 38 |
| STR | 1128 | 732 | 7 |
| PAL | 68 | 18 | 17 |
| TH | 46 | 7 | 7 |
| HY | 35 | 17 | 17 |
| MB | 33 | 8 | 8 |
| P | 30 | 11 | 11 |
| MY | 78 | 45 | 44 |
| CB | 83 | 29 | 29 |

Table 3: Number of experiments available to evaluate models in leave-one-out cross validation. Models that rely on a finer granularity of modeling have less data available to validate with.

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

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

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

⁴¹² PROJECTION-NORMALIZED LOSSES ON THRESHOLDED SET We also give results for the
⁴¹³ 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

⁴¹⁴ This section consists of additional information on preprocessing of the neural connectivity data,
⁴¹⁵ estimation of connectivity, and matrix factorization.

⁴¹⁶ **Data preprocessing**

⁴¹⁷ Several data preprocessing steps take place prior to evaluations of the connectivity matrices. These
⁴¹⁸ steps are described in Algorithm PREPROCESS. The arguments of this normalization process - injection
⁴¹⁹ signals $x(i)$, projection signals $y(i)$, injection fraction $F(i)$, and data quality mask $q(i)$ - were
⁴²⁰ downloaded using the Allen SDK. The injections and projection signals $\mathcal{B} \rightarrow [0, 1]$ were segmented
⁴²¹ manually in histological analysis. The projection signal gives the proportion of pixels within the voxel
⁴²² displaying fluorescence, and the injection signal gives the proportion of pixels within the
⁴²³ histologically-selected injection subset displaying fluorescence. The injection fraction $\mathcal{B} \rightarrow [0, 1]$ gives
⁴²⁴ the proportion of pixels within each voxel in the injection subset. Finally, the data quality mask
⁴²⁵ $\mathcal{B} \rightarrow \{0, 1\}$ gives the voxels that have valid data.

⁴²⁶ Our preprocessing makes use of the above ingredients, as well as several other essential steps. First,
⁴²⁷ we compute the weighted injection centroid

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

⁴²⁸ where $x(i)(l)$ is the injection density at location $l \in \mathbb{R}^3$. Given a regionalization \mathcal{R} from the Allen SDK,
⁴²⁹ we can also access regionalization map $R: \mathcal{B} \rightarrow \mathcal{R}$. This induces a functional of connectivities from
⁴³⁰ the space of maps $\{\mathcal{X} = x: \mathcal{B} \rightarrow [0, 1]$

$$\begin{aligned} 1_{\mathcal{R}}: \mathcal{X} &\rightarrow \mathcal{R} \times \mathbb{R}_{\geq 0} \\ x &\mapsto \sum_{l \in r} x(l) \text{ for } r \in \mathcal{R}. \end{aligned}$$

⁴³¹ We also can restrict a signal to a individual structure as

$$\begin{aligned} 1|_s: \mathcal{X} &\rightarrow \mathcal{X} \\ x(l) &= \begin{cases} x(l) & \text{if } l \in S \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

⁴³² Finally, given a vector or array $a \in \mathbb{R}^T$, we have the $l1$ normalization map

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

⁴³³ We define these objects as functions and functionals, but this is for notational convenience and

⁴³⁴ non-essential. A function $x(i) : \mathcal{B} \rightarrow [0, 1]$ is mathematically equivalent to the graph

⁴³⁵ $\mathcal{G}(x(i)) \in \mathcal{B} \times [0, 1]$. As an abuse of notation, we define $x \odot x' := z$ such that $z(l) = x(l)x'(l)$ for all $l \in \mathcal{B}$.

⁴³⁶ Also, denote $m(i)$ as the major structure containing experiment i . We then can write the

⁴³⁷ 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 $\tilde{y}_{\mathcal{T}}, c$

⁴³⁸ **Estimators**

⁴³⁹ As mentioned previously, we can consider our estimators as modeling a connectivity vector
⁴⁴⁰ $f_{\mathcal{T}}(\nu, s) \in \mathbb{R}_{\geq 0}^T$. Thus, for the remainder of this section, we will discuss only $f(\nu, s)$. We review the
⁴⁴¹ Nadaraya-Watson estimator from Knox et al. (2019), and describe its conversion into our cell-class
⁴⁴² specific Expected Loss estimator.

⁴⁴³ *Centroid-based Nadaraya-Watson* In the Nadaraya-Watson approach of Knox et al. (2019), the injection
⁴⁴⁴ is considered only through its centroid $c(i)$, and the projection is considered regionalized. That is,

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

⁴⁴⁵ Since the injection is considered only by its centroid, this model only generates predictions for
⁴⁴⁶ particular locations l , and the prediction for a structure s is given by integrating over locations within
⁴⁴⁷ the structure

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

⁴⁴⁸ Here, I is the training data, and \hat{f} is the Nadaraya-Watson estimator

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

⁴⁴⁹ where $\omega_{il} := \exp(-\gamma d(l, c(i))^2)$ and d is the Euclidean distance between centroid $c(i)$ and voxel with
⁴⁵⁰ position l .

⁴⁵¹ Several facets of the estimator are visible here. A smaller γ corresponds to a greater amount of
⁴⁵² smoothing, and the index set $I \subseteq \{1 : n\}$ generally depends on s and ν . Varying γ bridges between
⁴⁵³ 1-nearest neighbor prediction and averaging of all experiments in I . In Knox et al. (2019), I consisted
⁴⁵⁴ of experiments sharing the same brain division, i.e. $I = I_m$, while restricting of index set to only
⁴⁵⁵ include experiments with the same cell class gives the class-specific Cre-NW model. Despite this
⁴⁵⁶ restriction, we fit γ by leave-one-out cross-validation for each m rather than a smaller subset like s or
⁴⁵⁷ ν . That is,

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

⁴⁵⁸ *The Expected-Loss estimator* Besides location of the injection centroid, cell class also influences
⁴⁵⁹ projection. Thus, we introduce method for estimating the effect of Cre-distance, which we define as
⁴⁶⁰ the distance between the projections of the mean experiment of one (Cre,leaf) pair with another. This
⁴⁶¹ method assigns a predictive weight to each pair of training points that depends both on their
⁴⁶² centroid-distance and Cre-distance. This weight is determined by the expected prediction error of
⁴⁶³ each of the two feature types

⁴⁶⁴ We define Cre-line behavior as the average regionalized projection of a Cre-line in a given structure
⁴⁶⁵ (i.e. leaf). The vectorization of categorical information is known as **target encoding**

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

⁴⁶⁶ The Cre-distance is then defined a **Cre-distance** in a leaf to be the distance between the target-encoded
⁴⁶⁷ projections of two Cre-lines. The relative predictive accuracy of Cre-distance and centroid distance is
⁴⁶⁸ determined by fitting a surface of projection distance as a function of Cre-distance and centroid
⁴⁶⁹ distance.

⁴⁷⁰ as determined by cross-validation. When we use shape-constrained B-splines to estimate this
⁴⁷¹ weight, the weights then may be said to be used in a Nadaraya-Watson estimator.

⁴⁷² For this reason, we call this the Expected Loss Estimator. The resulting weights are then utilized in a
⁴⁷³ Nadaraya-Watson estimator in a final prediction step. increase the effective sample size of our
⁴⁷⁴ Nadaraya-Watson estimator,

⁴⁷⁵ In mathematical terms, our full feature set consists of the centroid coordinates and the
⁴⁷⁶ target-encoded means of the combinations of virus type and injection-centroid structure. That is,

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

⁴⁷⁷ f^* is defined as in (2). The expected loss estimator is then

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

⁴⁷⁸ where

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

⁴⁷⁹ and s is the structure containing l .

⁴⁸⁰ The key step therefore is finding a suitable function g with which to weight the positional and Cre
⁴⁸¹ information. Note that g must be a concave, non-decreasing function of its arguments with with
⁴⁸² $g(0, 0) = 0$. Then, g defines a metric on the product of the metric spaces defined by experiment
⁴⁸³ centroid and target-encoded cre-line, and \hat{f}_{EL} is a Nadaraya-Watson estimator. A derivation of this
⁴⁸⁴ fact is given later in this section.

⁴⁸⁵ therefore use shape-constrained B-splines to estimate g . In particular, B-splines is a method for
⁴⁸⁶ generating a predictive model

$$y =$$

⁴⁸⁷ In our setting, this gives us a surface the Cre-distance and centroid-distance to the
⁴⁸⁸ projection-distance response. Similarly to the Nadaraya-Watson model, we make the decision to fit a
⁴⁸⁹ g separately for each major brain division, and select \hat{g} as in 2. We compute our estimates using the
⁴⁹⁰ pyGAM Python package (Servén D., n.d.).

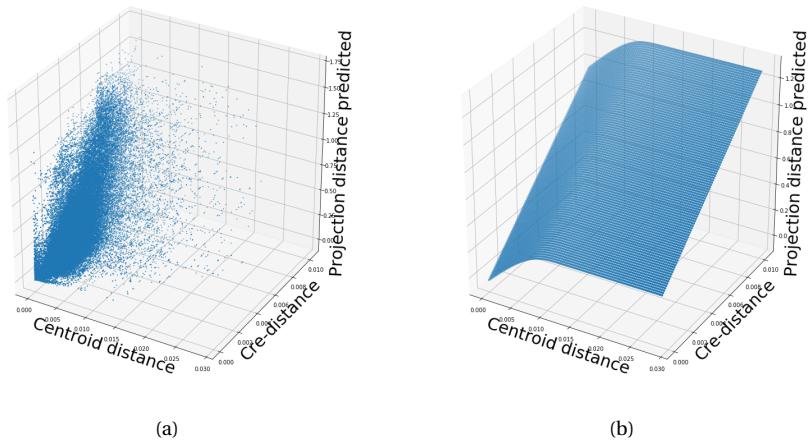


Figure 18: Fitting g . 18a Distribution of projection errors against centroid distance and cre-distance in Isocortex. 18b estimated \hat{g} using B-splines.

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

496 **Proposition 1.** *Given a collection of metric spaces X_1, \dots, X_n with metrics d_1, \dots, d_n (e.g. $d_{centroid}, d_{cre}$),
 497 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
 498 concave, non-decreasing and $g(d) = 0 \iff d = 0$.*

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

- 500 ▪ The first property of a metric is that $f(x, x') = 0 \iff x = x'$. The left implication:
 501 $x = x' \implies f(x_1, x'_1, \dots, x_n, x'_n) = g(0, \dots, 0)$, since d are metrics. Then, since $g(0) = 0$, we have that
 502 $f(x, x') = 0$. The right implication: $f(x, x') = 0 \implies d = 0 \implies x = x'$ since d are metrics.
- 503 ▪ The second property of a metric is that $f(x, x') = f(x', x)$. This follows immediately from the
 504 symmetry of the d_i , i.e. $f(x, x') = f(x_1, x'_1, \dots, x_n, x'_n) = g(d_1(x_1, x'_1), \dots, d_n(x_n, x'_n)) =$
 505 $g(d_1(x'_1, x_1), \dots, d_n(x'_n, x_n)) = f(x'_1, x_1, \dots, x'_n, x_n) = f(x', x)$.
- 506 ▪ The third property of a metric is the triangle inequality: $f(x, x') \leq f(x, x^*) + f(x^*, x')$. To show this
 507 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
 508 non-decreasing and by the triangle inequality of d . Then, since g is concave,
 509 $g(d(x, x^*) + d(x^*, x')) \leq g(d(x, x^*)) + g(d(x^*, x')) = f(x, x^*) + f(x^*, x')$.

511 ***Setting a lower detection threshold***

512 The lower detection threshold of our approach is a complicated consequence of our experimental and
 513 analytical protocols. For example, the Nadaraya-Watson estimator is likely to generate many small
 514 false positive connections, since the projection of even a single experiment within the source region
 515 to a target will cause a non-zero connectivity in the Nadaraya-Watson weighted average. On the other
 516 hand, the complexities of the experimental protocol itself and the image analysis and alignment can
 517 also cause spurious signals. Therefore, it is of interest to establish a lower-detection threshold below
 518 which we have very little power-to-predict, and set estimated connectivities below this threshold to
 519 zero. This should make our estimated connectivities more accurate, especially in the
 520 biologically-important sense of sparsity.

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

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

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

523 ***Decomposing the connectivity matrix***

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

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

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

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

538 The mask $\mathbf{1}_M$ specifies this objective for detecting patterns in long-range connections. We note the
 539 existence of NMF with alternative norms for certain marginal distributions, but leave utilization of
 540 this approach for future work (Brunet, Tamayo, Golub, & Mesirov, 2004).

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

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

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

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

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

A standard solution is to 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

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

557 *Stabilizing NMF* The NMF program is non-convex, and, empirically, individual replicates will not
 558 converge to the same optima. One solution therefore is to run multiple replicates of the NMF
 559 algorithm and cluster the resulting vectors. This approach raises the questions of how many clusters
 560 to use, and how to deal with stochasticity in the clustering algorithm itself. We address this issue
 561 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

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

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

7 SUPPLEMENTAL EXPERIMENTS

571 Setting detection threshold τ

572 We give results on the false detection rate at different limits of detection. These conclusively show that
 573 10^{-6} is the good threshold for our normalized data.

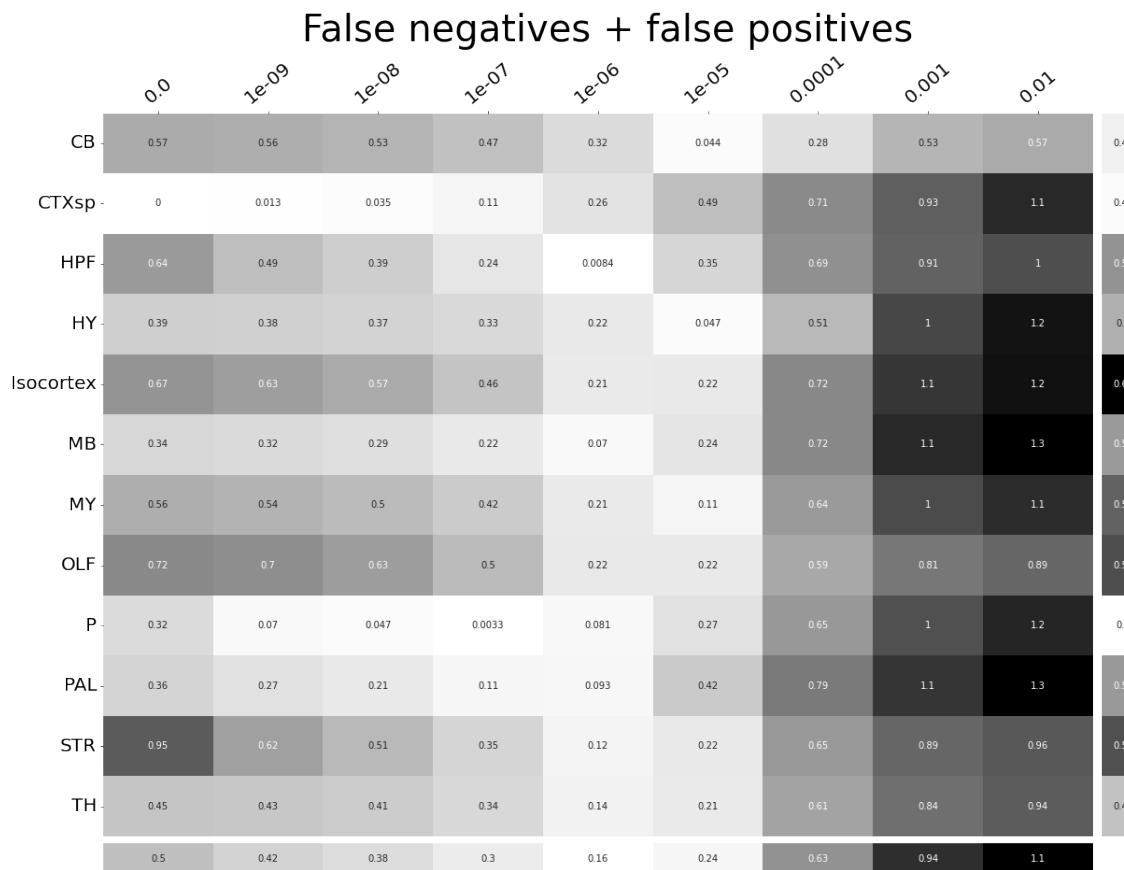


Figure 19: τ at different limits of detection in different major structures. 10^{-6} is the optimal detection threshold.

574 ***Loss subsets***

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

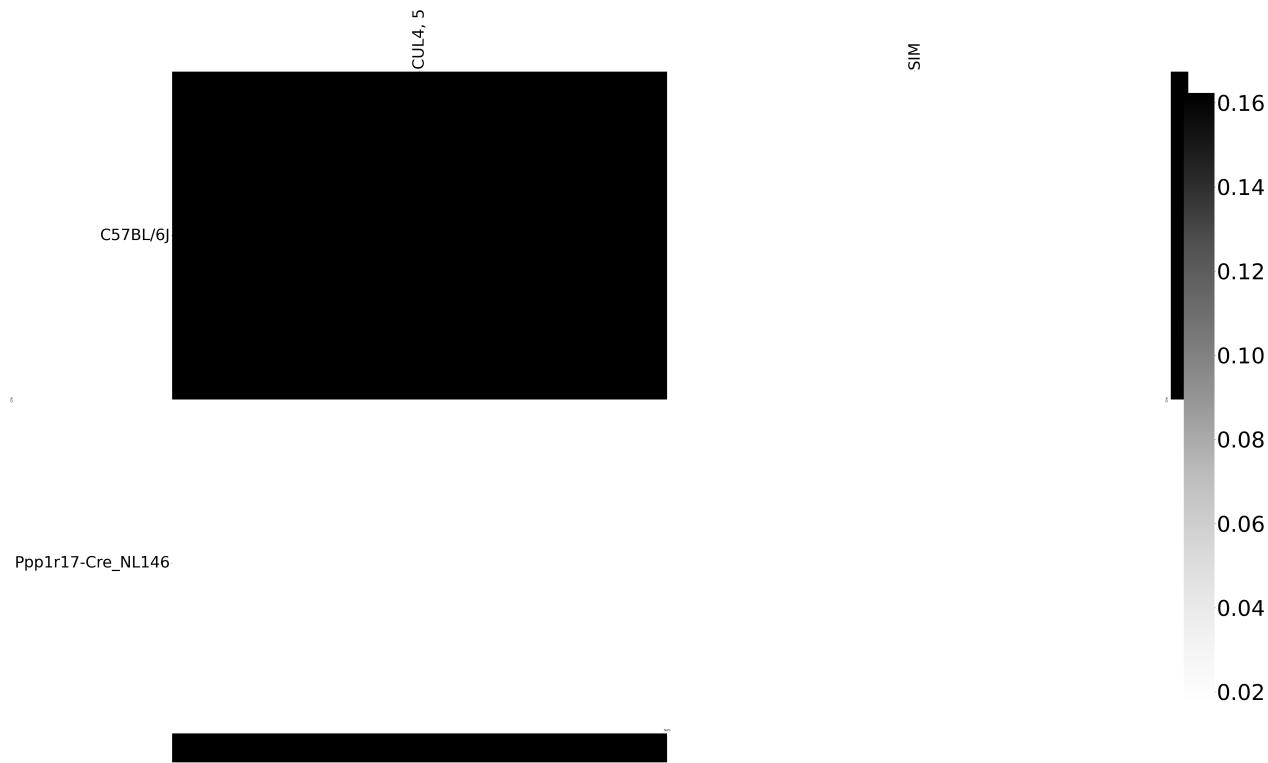


Figure 20: Weighted loss for cre-leaf combinations in CB. Missing values are omitted. Row and column averages are also plotted.

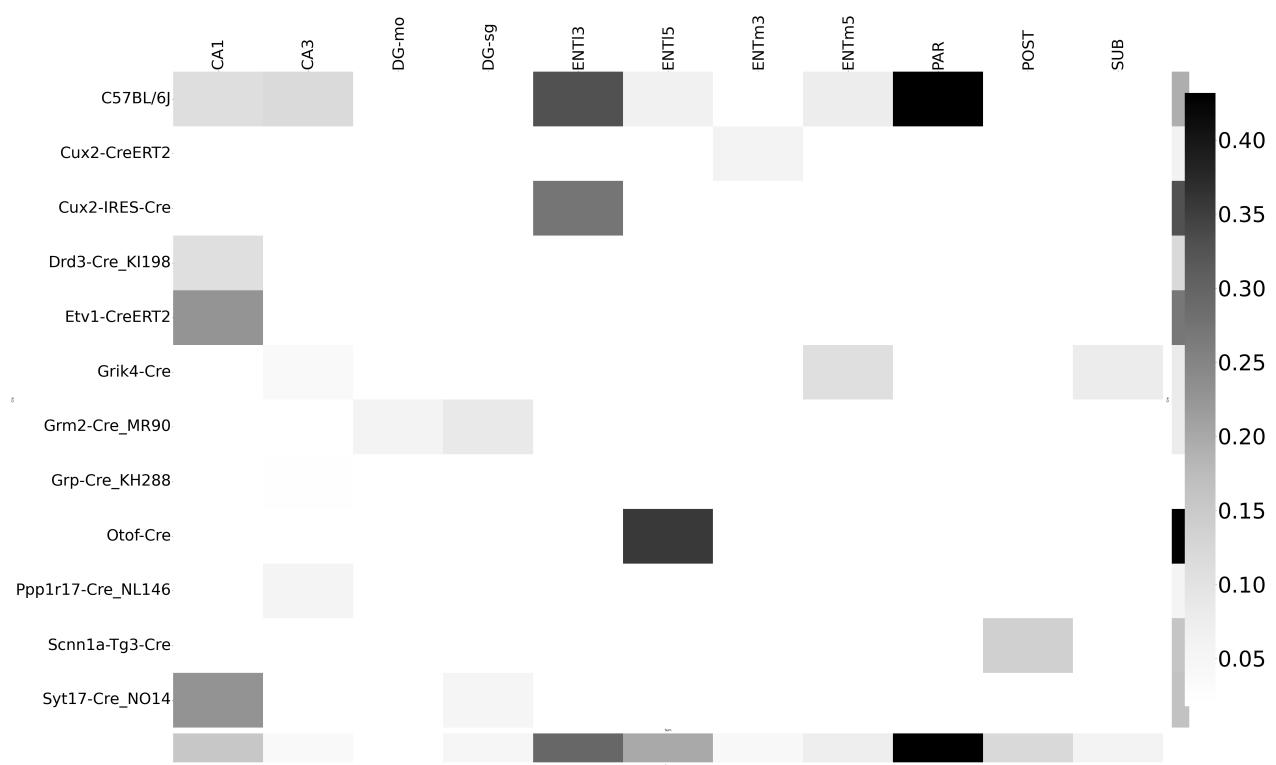


Figure 21: Weighted loss for cre-leaf combinations in HPF. Missing values are omitted. Row and column averages are also plotted.

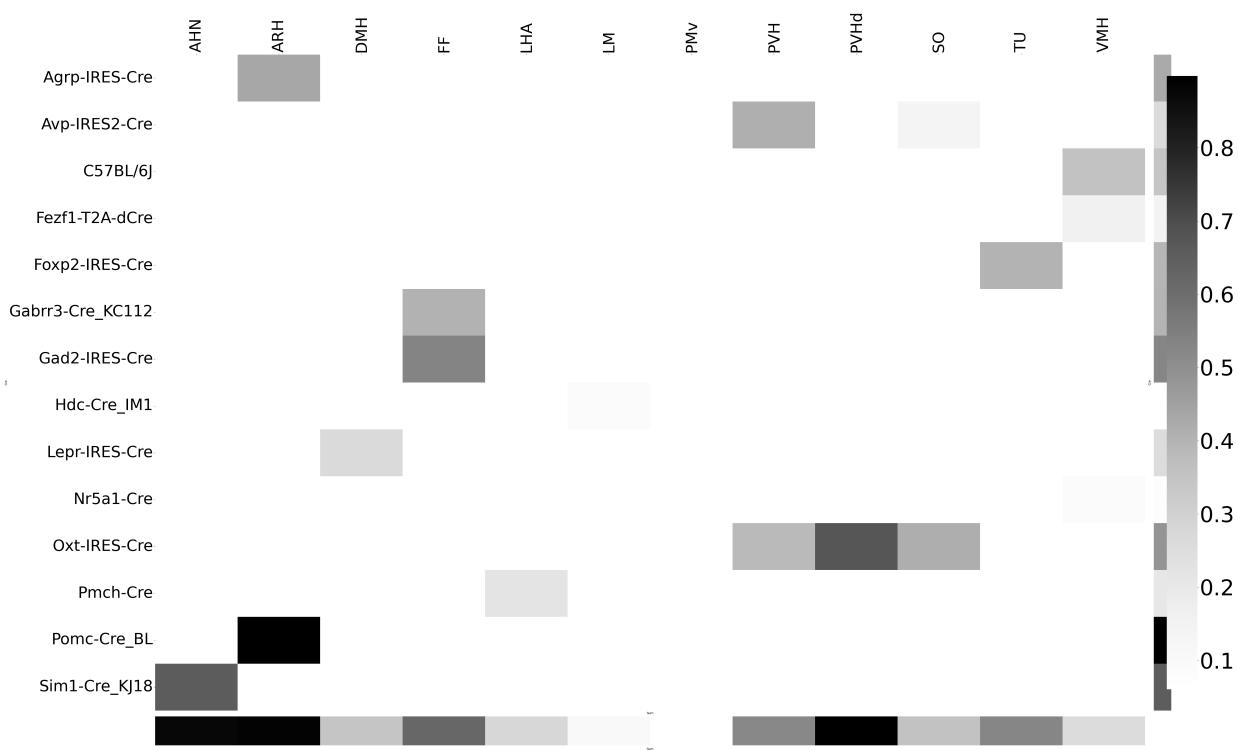


Figure 22: Weighted loss for cre-leaf combinations in HY. Missing values are omitted. Row and column averages are also plotted.

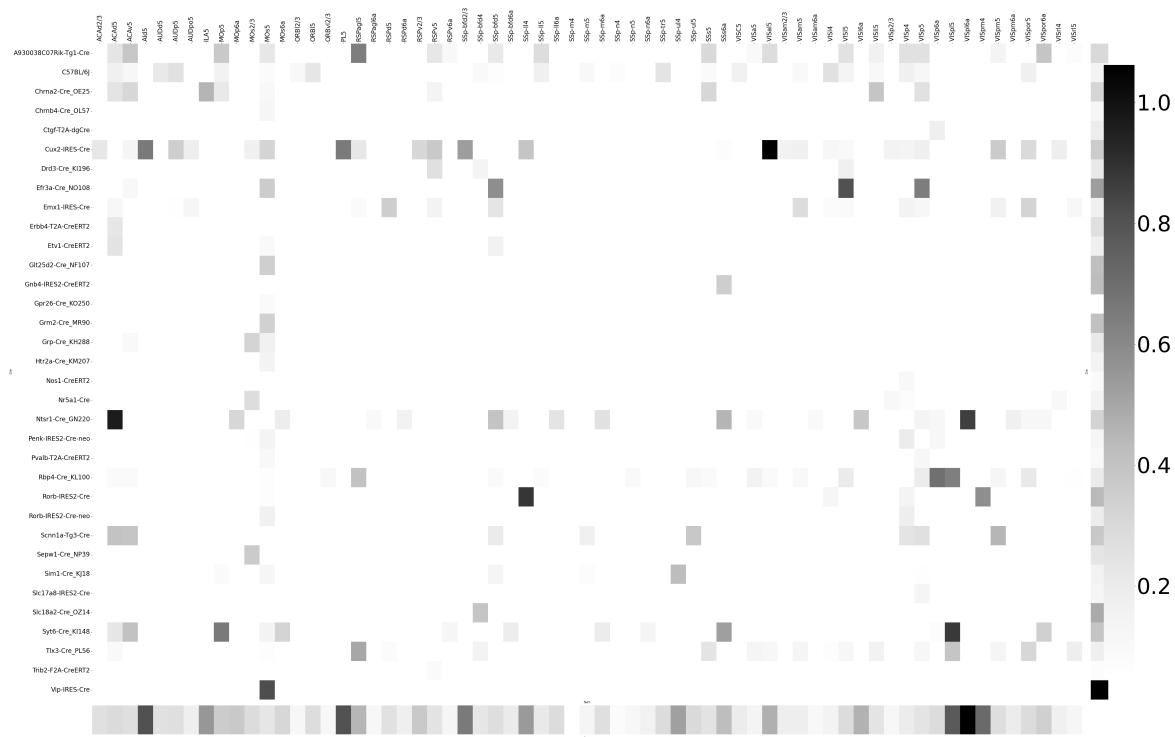


Figure 23: Weighted loss for cre-leaf combinations in Isocortex. Missing values are omitted. Row and column averages are also plotted.

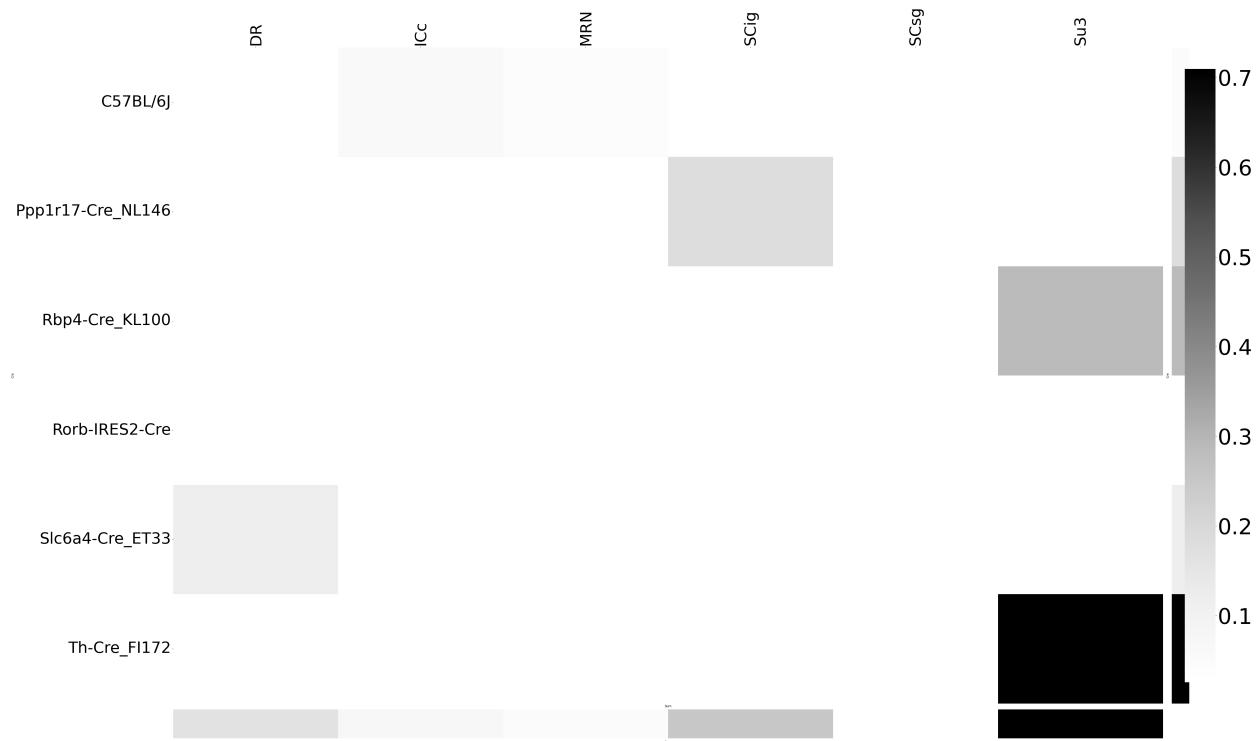


Figure 24: Weighted loss for cre-leaf combinations in MB. Missing values are omitted. Row and column averages are also plotted.

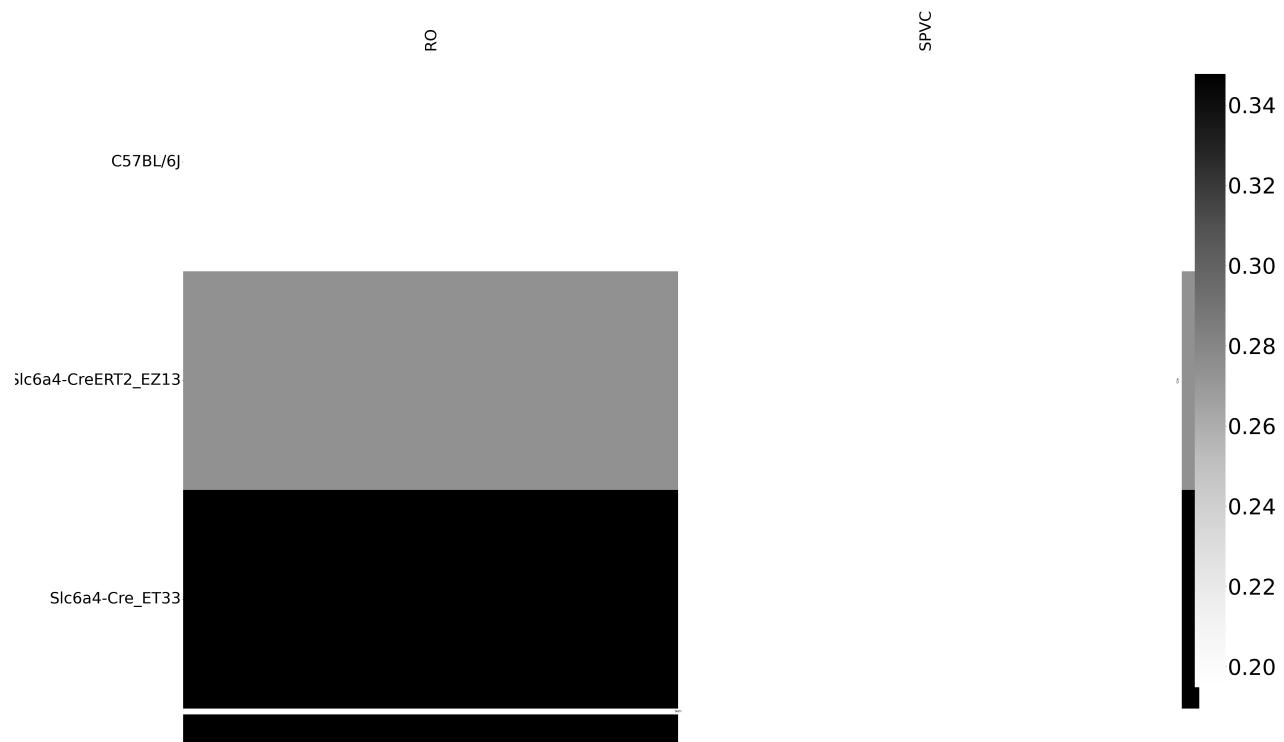


Figure 25: Weighted loss for cre-leaf combinations in MY. Missing values are omitted. Row and column averages are also plotted.

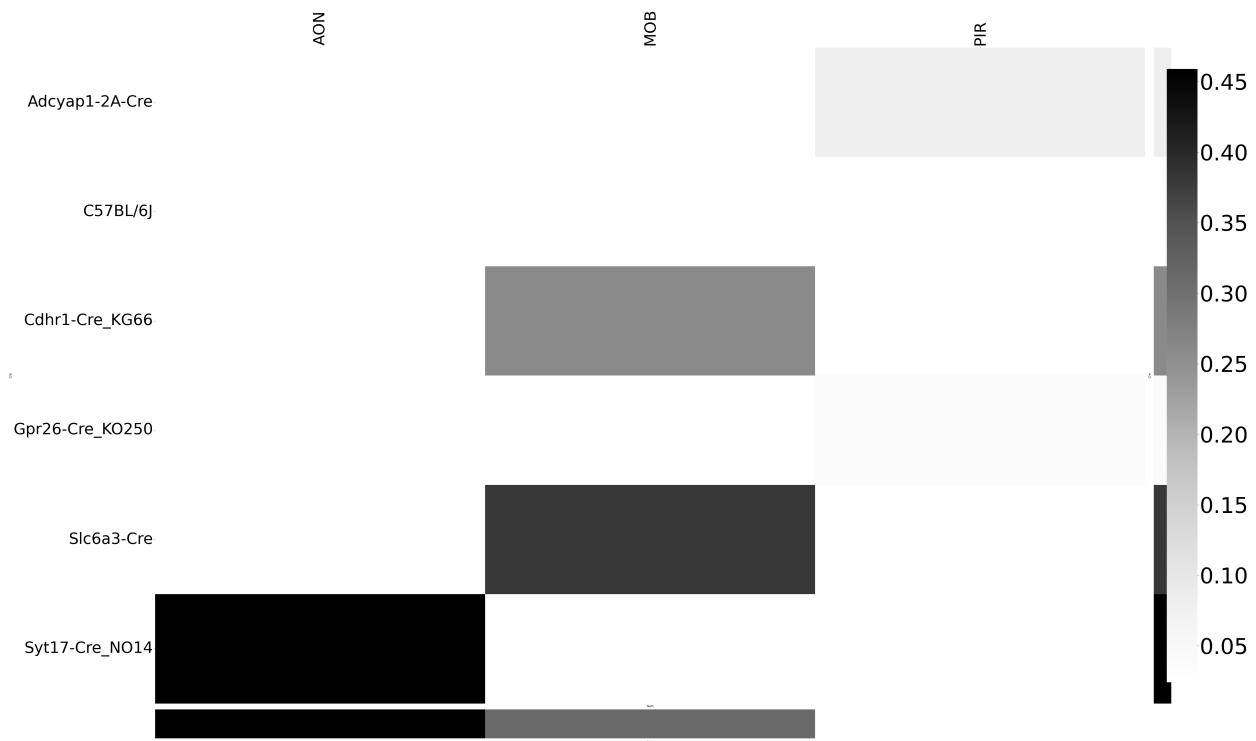


Figure 26: Weighted loss for cre-leaf combinations in OLF. Missing values are omitted. Row and column averages are also plotted.

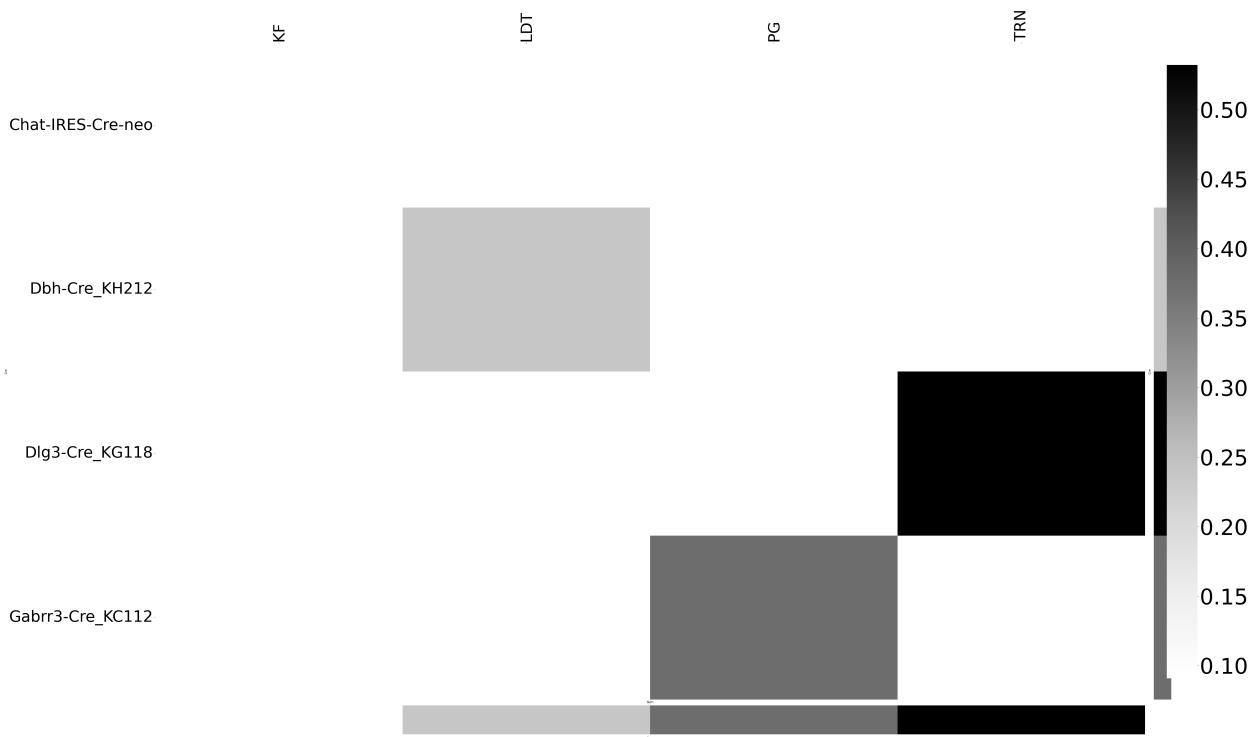


Figure 27: Weighted loss for cre-leaf combinations in P. Missing values are omitted. Row and column averages are also plotted.

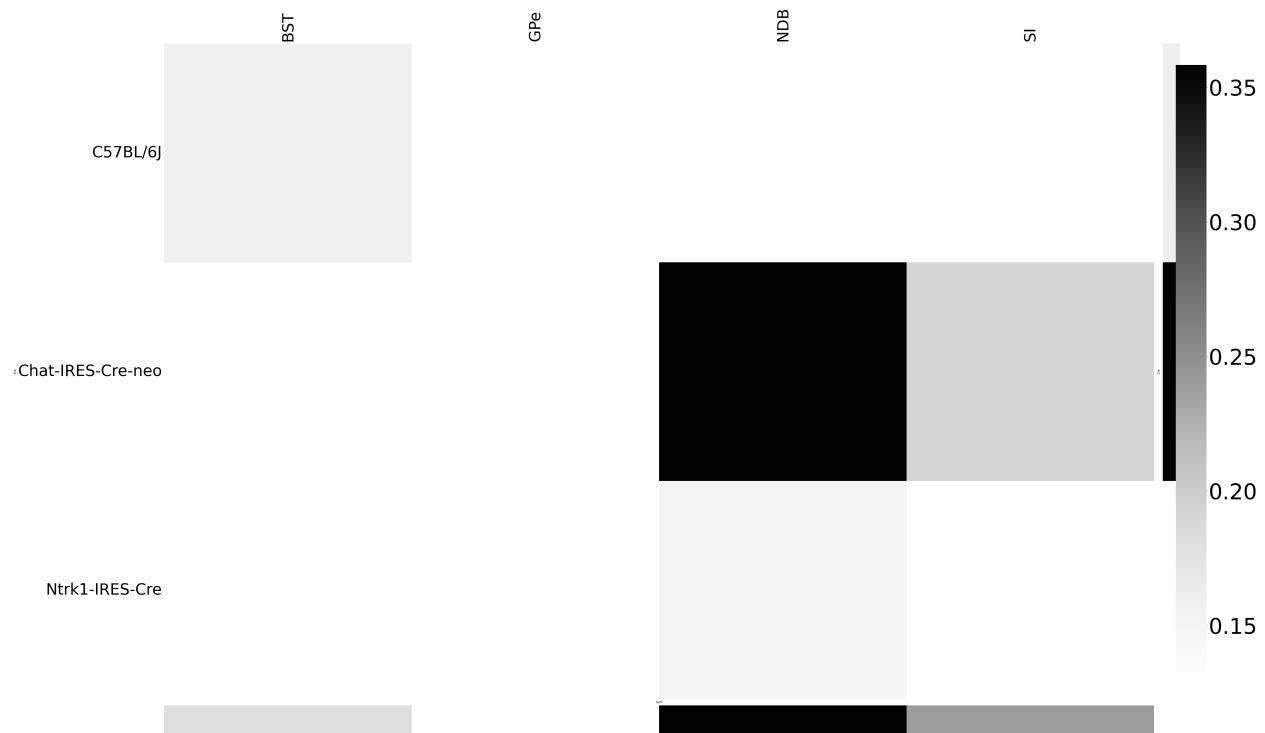


Figure 28: Weighted loss for cre-leaf combinations in PAL. Missing values are omitted. Row and column averages are also plotted.

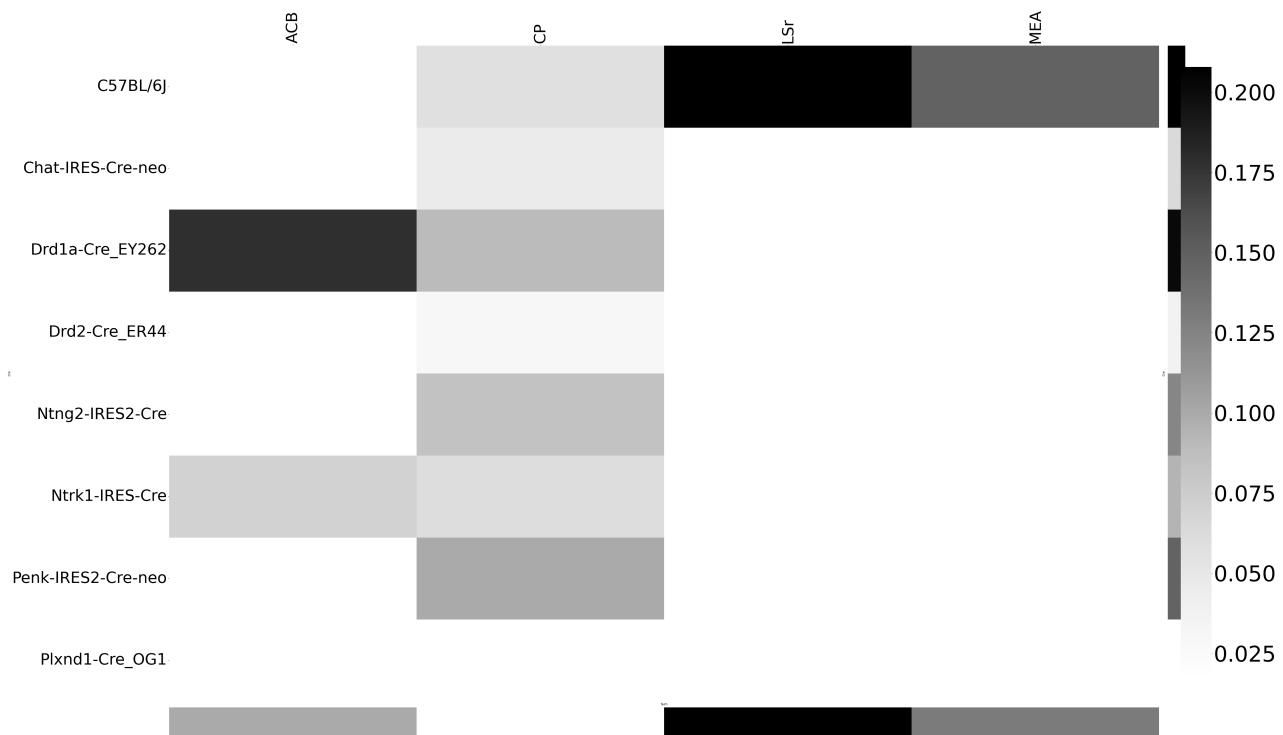


Figure 29: Weighted loss for cre-leaf combinations in STR. Missing values are omitted. Row and column averages are also plotted.

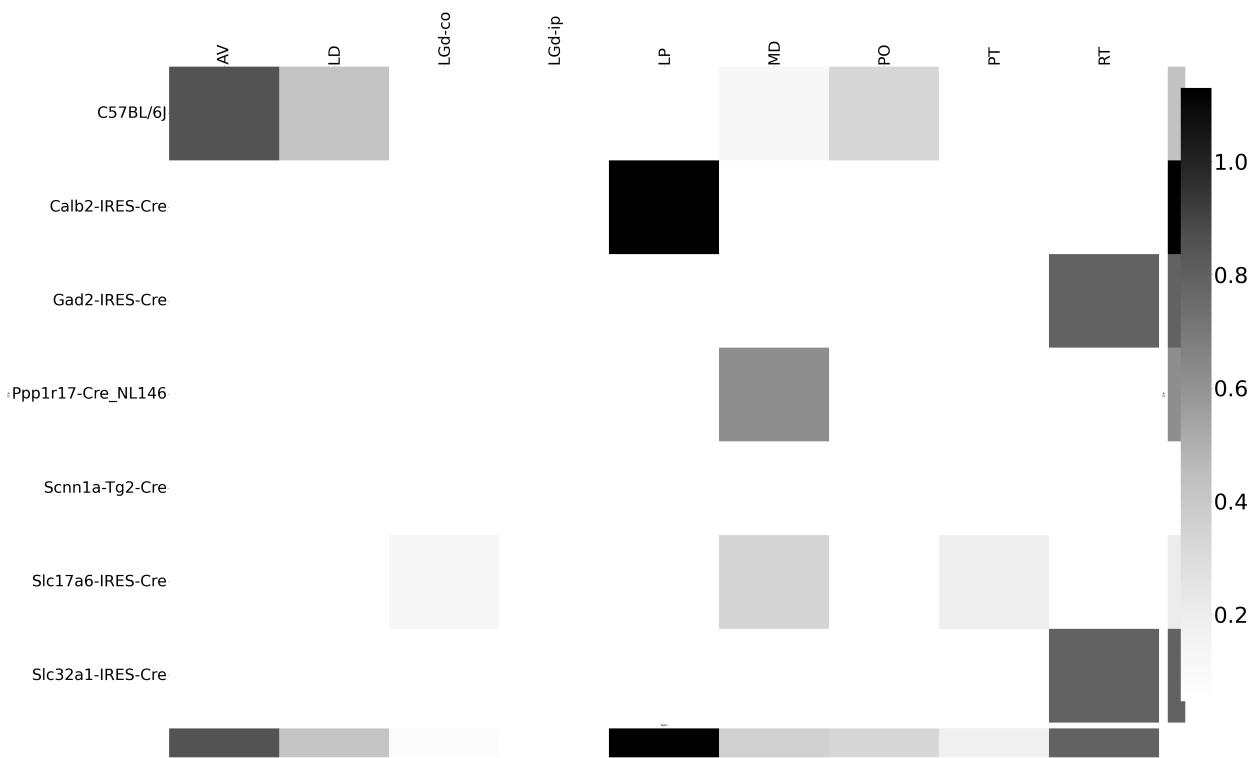


Figure 30: Weighted loss for cre-leaf combinations in TH. Missing values are omitted. Row and column averages are also plotted.

578 **Matrix Factorization**

579 We give additional results on the generation of the archetypal connectome patterns. These consist of
 580 cross-validation selection of q , the number of latent components, stability analysis, and visualization
 581 of the reconstructed wild-type connectivity.

582 *Cross-validation* We set $\alpha = 0.002$ and run Program 2 on \mathcal{C}_{wt} . We use a random mask with $p = .3$ to
 583 evaluate prediction accuracy of models trained on the unmasked data on the masked data. To
 584 account for stochasticity in the NMF algorithm, we run $R = 8$ replicates at each potential dimension q .
 585 This selects $\hat{q} = 60$.

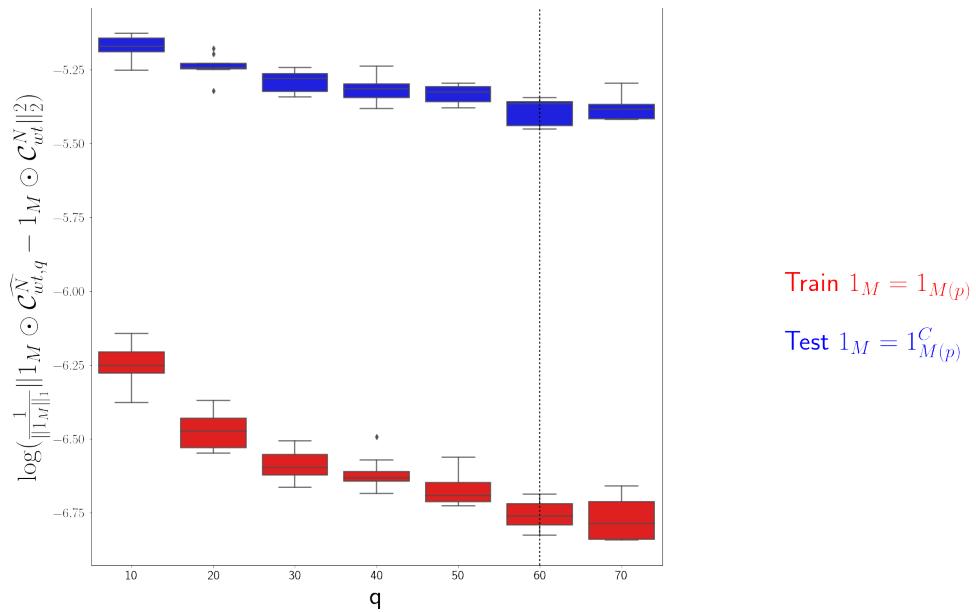


Figure 31: Train and test error using NMF decomposition.

586 *Stability* For the purposes of visualization and interpretability, we restrict to a $q = 15$ component
 587 model. To address the instability of the NMF algorithm in identifying components, we $k - means$
 588 cluster components over $R = 10$ replicates with $k \in \{10, 15, 20, 25, 30\}$. Since the clustering is itself
 589 unstable, we repeat the clustering 25 times and select the k with the largest Rand index.

| 590 q | 10.000000 | 20.000000 | 30.000000 | 40.000000 | 50.000000 | |
|-------|------------|-----------|-----------|-----------------|-----------|----------|
| | Rand index | 0.772544 | 0.844981 | 0.932957 | 0.929827 | 0.885862 |

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

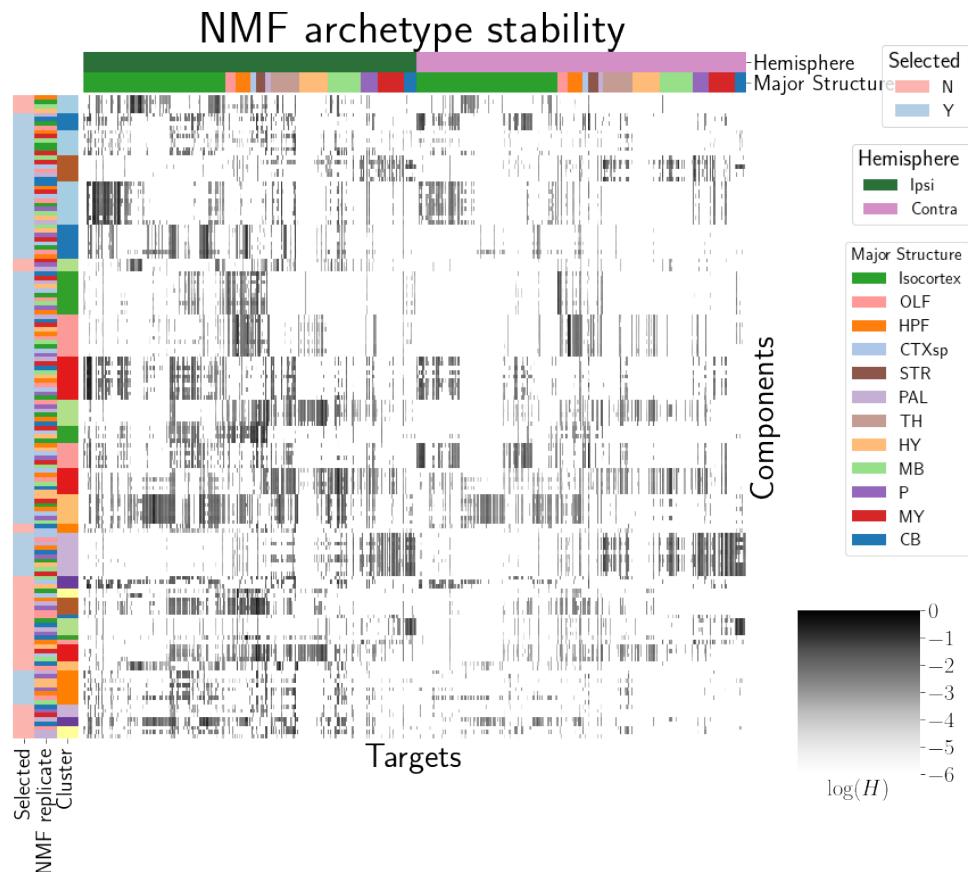


Figure 32: Stability of NMF results across replicates. Replicate and NMF component are shown on rows. Components that are in the top 15 are also indicated.

593 These are the components whose medians are plotted in Figure 4a.

8 COMPETING INTERESTS

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

596

REFERENCES

598

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

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

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

605 Gao, Y., Zhang, X., Wang, S., & Zou, G. (2016). Model averaging based on leave-subject-out cross-validation. *J. Econom.*,
606 192(1), 139–151.

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

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

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

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

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

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

619 Kotliar, D., Veres, A., Nagy, M. A., Tabrizi, S., Hodis, E., Melton, D. A., & Sabeti, P. C. (2019). Identifying gene expression
620 programs of cell-type identity and cellular activity with single-cell RNA-Seq. *Elife*, 8.

- 621 Li, X., Yu, B., Sun, Q., Zhang, Y., Ren, M., Zhang, X., ... Qiu, Z. (2018). Generation of a whole-brain atlas for the
 622 cholinergic system and mesoscopic projectome analysis of basal forebrain cholinergic neurons. *Proc. Natl. Acad. Sci.*
 623 *U. S. A.*, 115(2), 415–420.
- 624 Lotfollahi, M., Naghipourfar, M., Theis, F. J., & Alexander Wolf, F. (2019). Conditional out-of-sample generation for
 625 unpaired data using trVAE.
- 626 Muzerelle, A., Scotto-Lomassese, S., Bernard, J. F., Soiza-Reilly, M., & Gaspar, P. (2016). Conditional anterograde tracing
 627 reveals distinct targeting of individual serotonin cell groups (B5-B9) to the forebrain and brainstem. *Brain Struct.*
 628 *Funct.*, 221(1), 535–561.
- 629 Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., ... Zeng, H. (2014). A mesoscale connectome of the
 630 mouse brain. *Nature*, 508(7495), 207–214.
- 631 Ren, J., Friedmann, D., Xiong, J., Liu, C. D., Ferguson, B. R., Weerakkody, T., ... Luo, L. (2018). Anatomically defined and
 632 functionally distinct dorsal raphe serotonin sub-systems. *Cell*, 175(2), 472–487.e20.
- 633 Ren, J., Isakova, A., Friedmann, D., Zeng, J., Grutzner, S. M., Pun, A., ... Luo, L. (2019). Single-cell transcriptomes and
 634 whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *Elife*, 8.
- 635 Saul, L. K., & Roweis, S. T. (2003). Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *J.*
 636 *Mach. Learn. Res.*, 4(Jun), 119–155.
- 637 Servén D., B. C. (n.d.). *pygam: Generalized additive models in python*.
- 638 von Luxburg, U. (2010a). Clustering stability: An overview.
- 639 von Luxburg, U. (2010b). Clustering stability: An overview.
- 640 Watson, C., Paxinos, G., & Puelles, L. (2012). The mouse nervous system..
- 641 Wu, S., Joseph, A., Hammonds, A. S., Celiker, S. E., Yu, B., & Frise, E. (2016). Stability-driven nonnegative matrix
 642 factorization to interpret spatial gene expression and build local gene networks. *Proc. Natl. Acad. Sci. U. S. A.*, 113(16),
 643 4290–4295.
- 644 Zaborszky, L., Csordas, A., Mosca, K., Kim, J., Gielow, M. R., Vadasz, C., & Nadasdy, Z. (2015). Neurons in the basal
 645 forebrain project to the cortex in a complex topographic organization that reflects corticocortical connectivity

646 patterns: an experimental study based on retrograde tracing and 3D reconstruction. *Cereb. Cortex*, 25(1), 118–137.

9 TECHNICAL TERMS

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

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

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

652 **Structural connectivities** connectivity between structures

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

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

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

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

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

659 **Connectivity archetypes** Typical connectivity patterns

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