Supplementary Materials for

A Theory for the Emergence of Neocortical Network Architecture

Daniel Udvary[1], Philipp Harth[2], Jakob H. Macke[3], Hans-Christian Hege[2], Christiaan P.J. de Kock[4], Bert Sakmann[5], Marcel Oberlaender[1,*]

[1]Max Planck Group: In Silico Brain Sciences, Center of Advanced European Studies and Research (caesar), Ludwig-Erhard-Allee 2, 53175 Bonn, Germany. [2]Department of Visualization and Data Analysis, Zuse Institute Berlin, Takustraße 7, 14195 Berlin, Germany. [3]Machine Learning in Science, Tübingen University, Maria-von-Linden-Straße 6, 72076 Tübingen, Germany. [4]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU Amsterdam, De Boelelaan 1085, 1081 Amsterdam, the Netherlands. [5]Max Planck Institute of Neurobiology, Am Klopferspitz 18, 82152 Martinsried, Germany.

*** Editorial correspondence:** Max Planck Group: In Silico Brain Sciences, Center of Advanced European Studies and Research (caesar), Ludwig-Erhard-Allee 2, Bonn, 53175 Germany, marcel.oberlaender@caesar.de

Supplementary figures

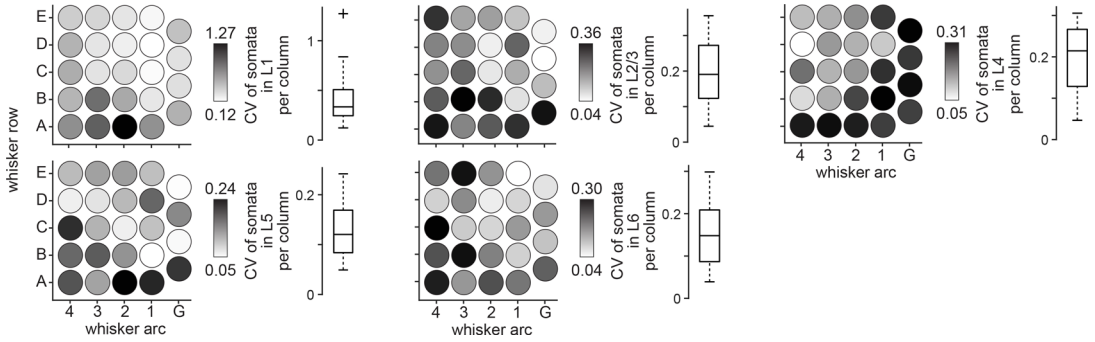

Figure S1: Inter-animal variability of cellular organization of the rat barrel cortex. CVs of somata for each barrel column and layer (N=4 rats).

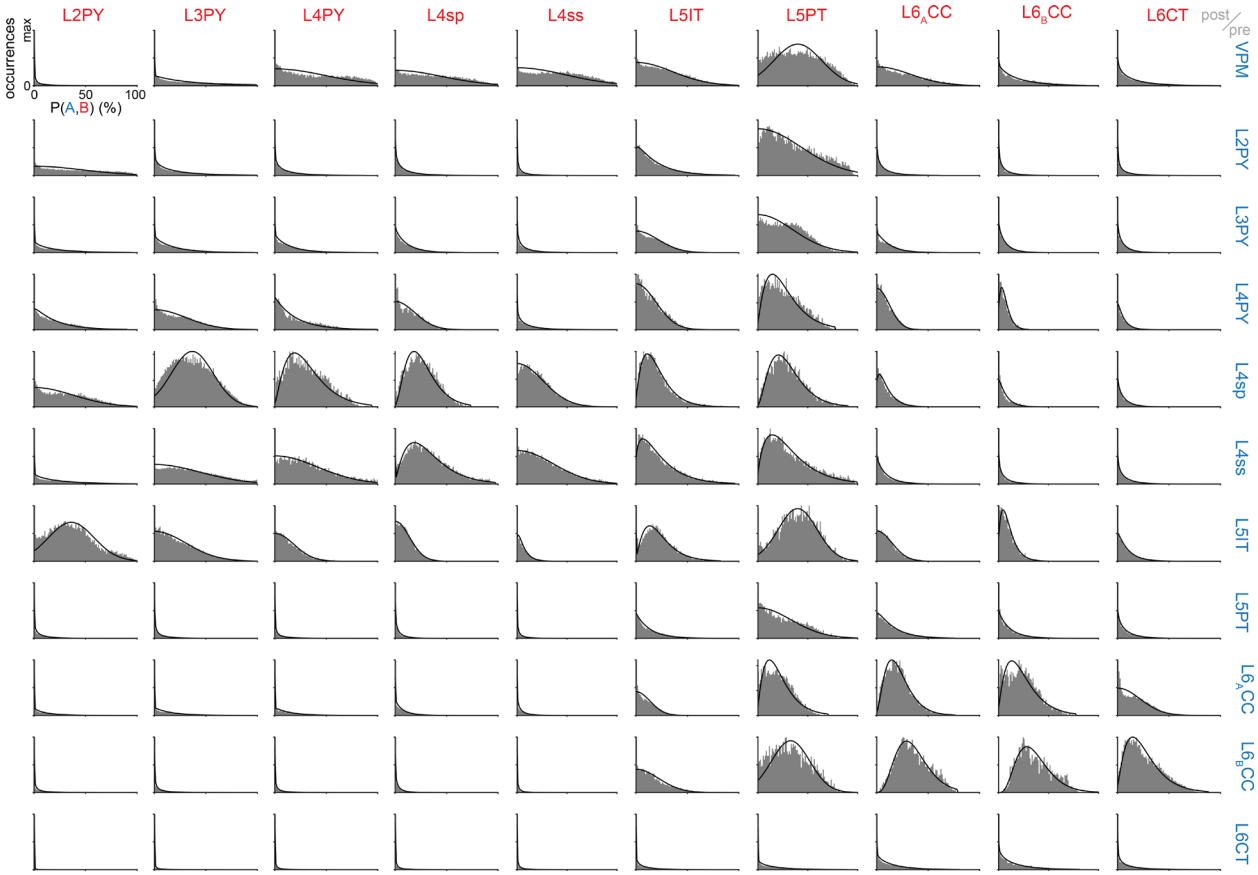

Figure S2: Distribution of connection probabilities for all cell type-dependent groupings within one barrel column. Histogram in gray, best fitting function as black line.

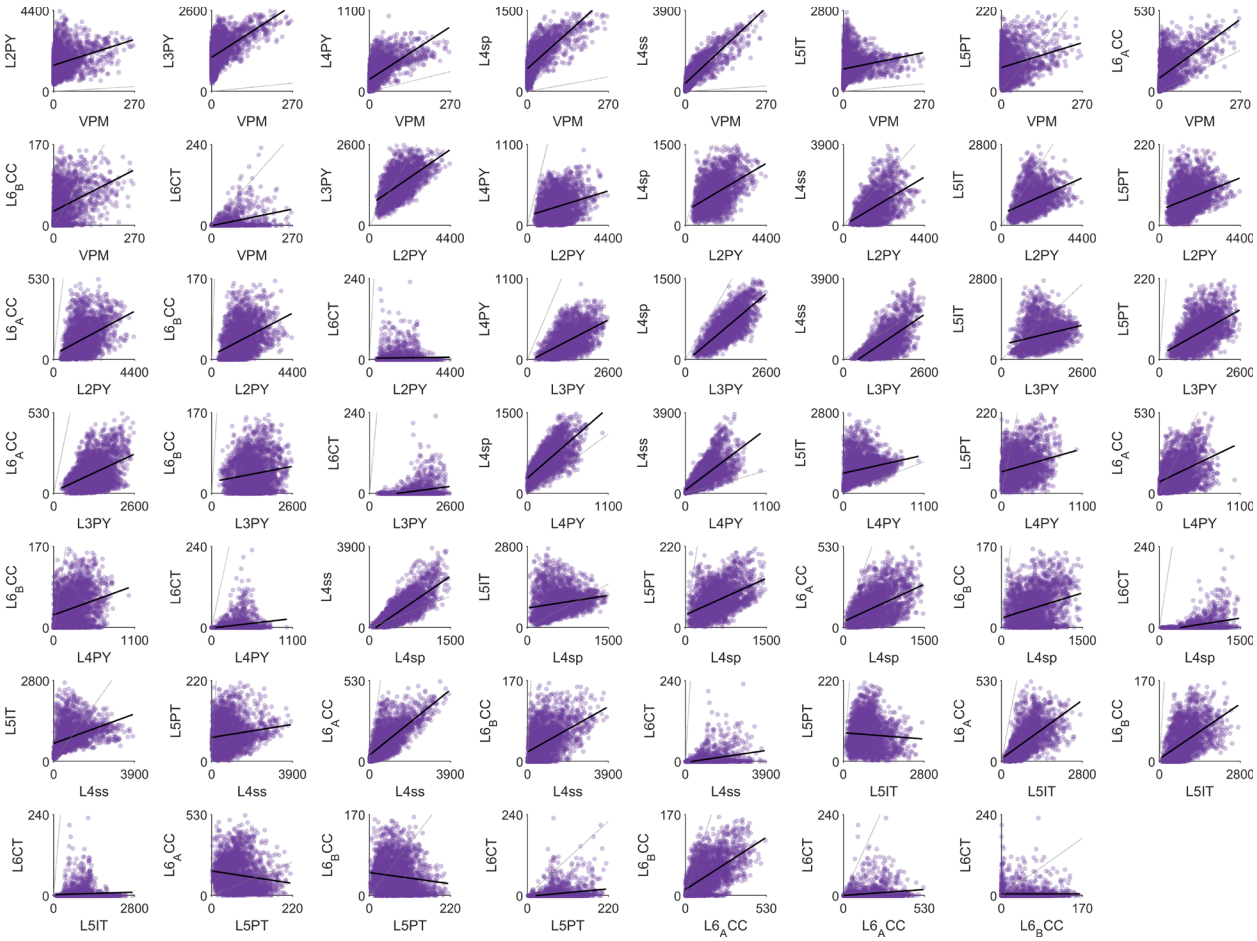

Figure S3: In-degree correlations between all cell type-dependent groupings A and B onto postsynaptic populations C. Y-axis represents in-degree(A), x-axis in-degree(B), postsynaptic population C=L2PY. All neurons sampled from C2 barrel column. Black line represents linear fit.

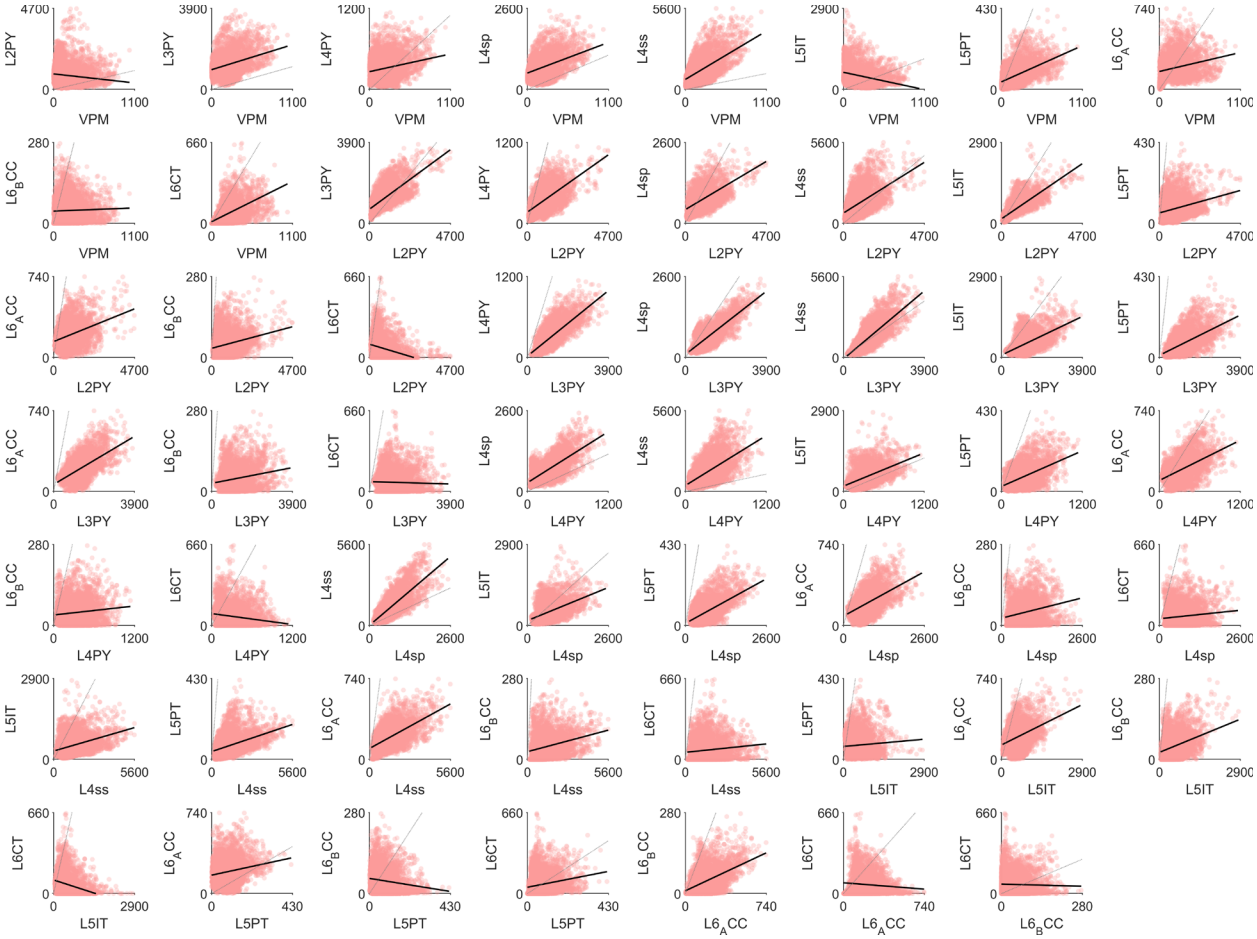

Figure S3 continued. Postsynaptic population C=L3PY.

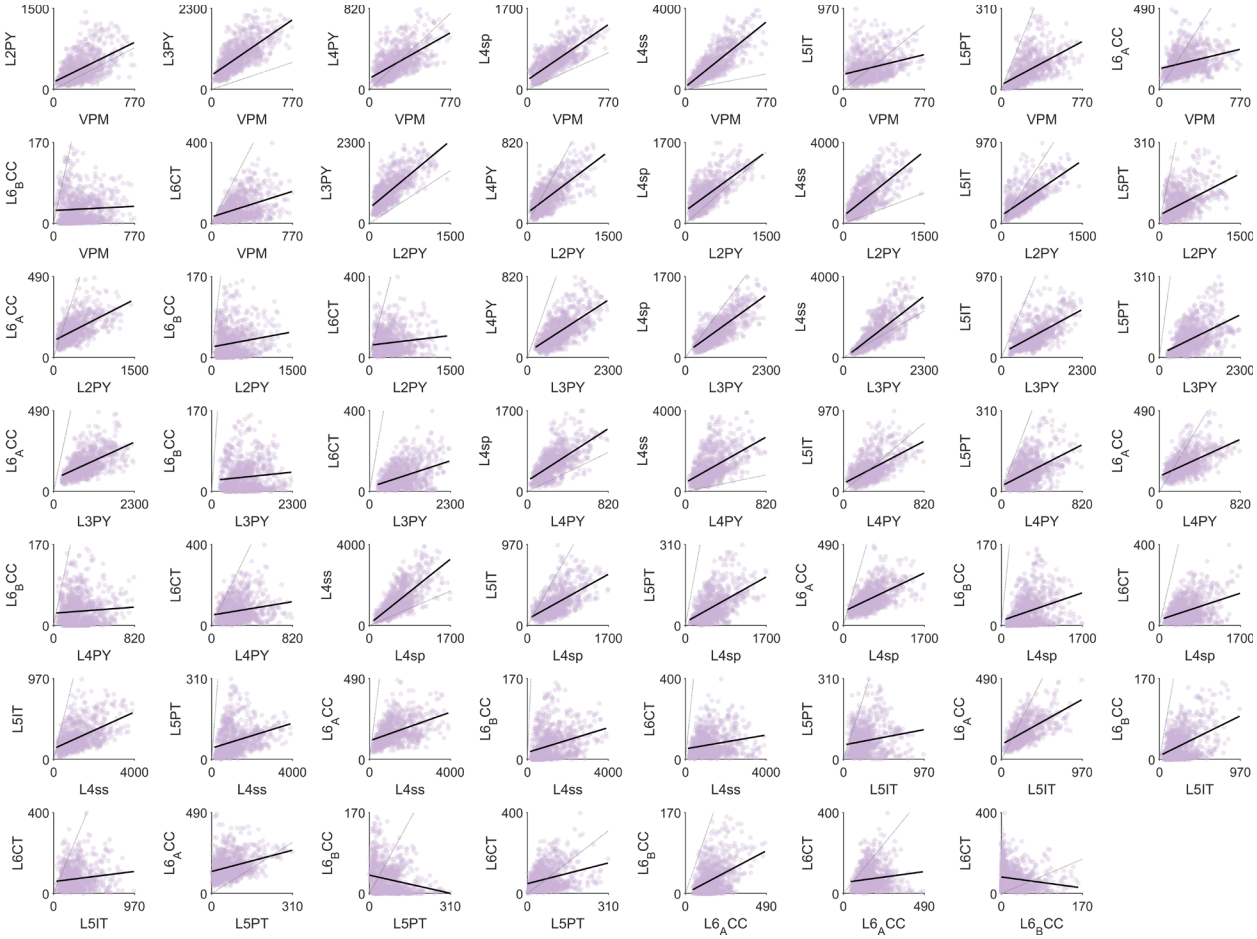

Figure S3 continued. Postsynaptic population C=L4PY.

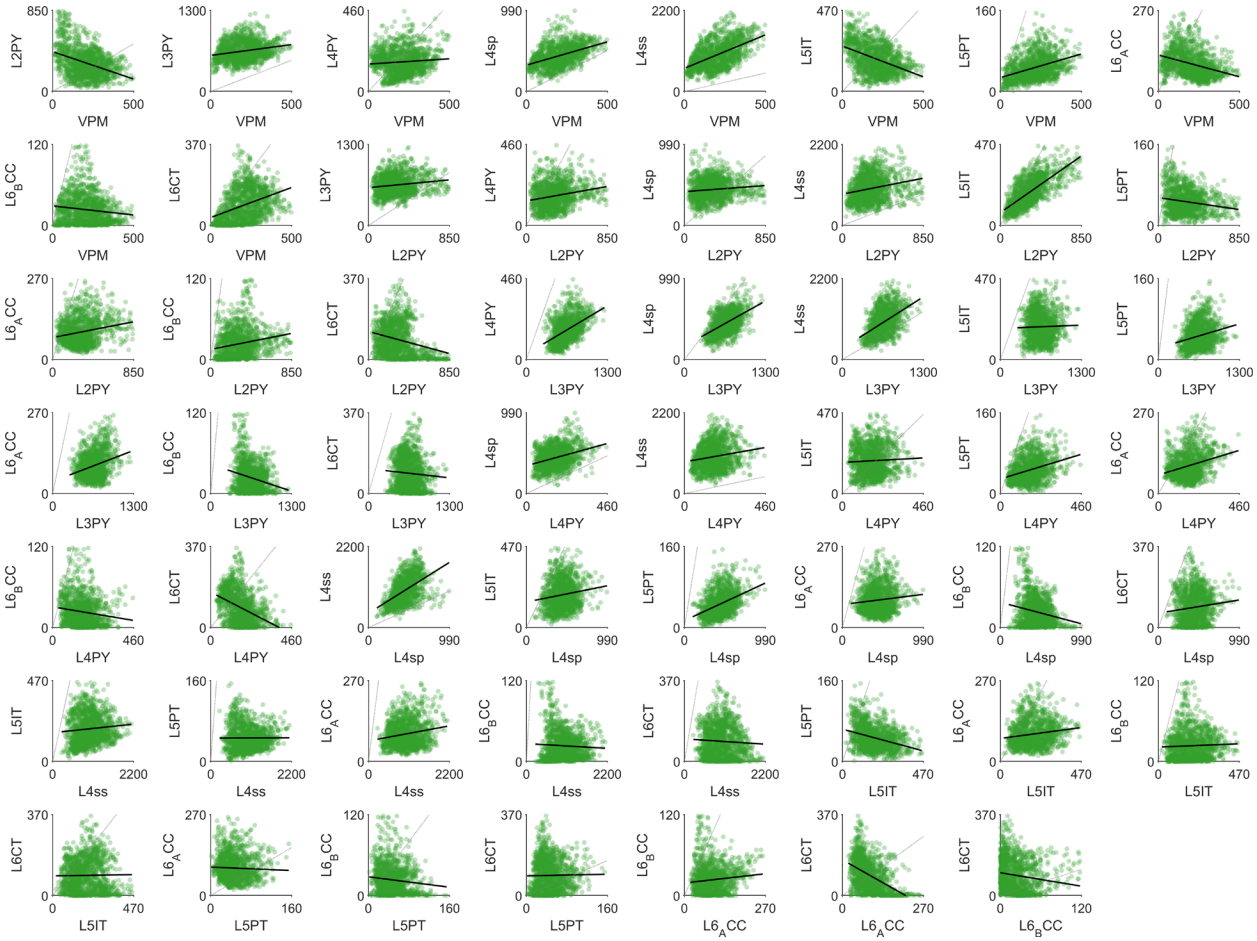

Figure S3 continued. Postsynaptic population C=L4sp.

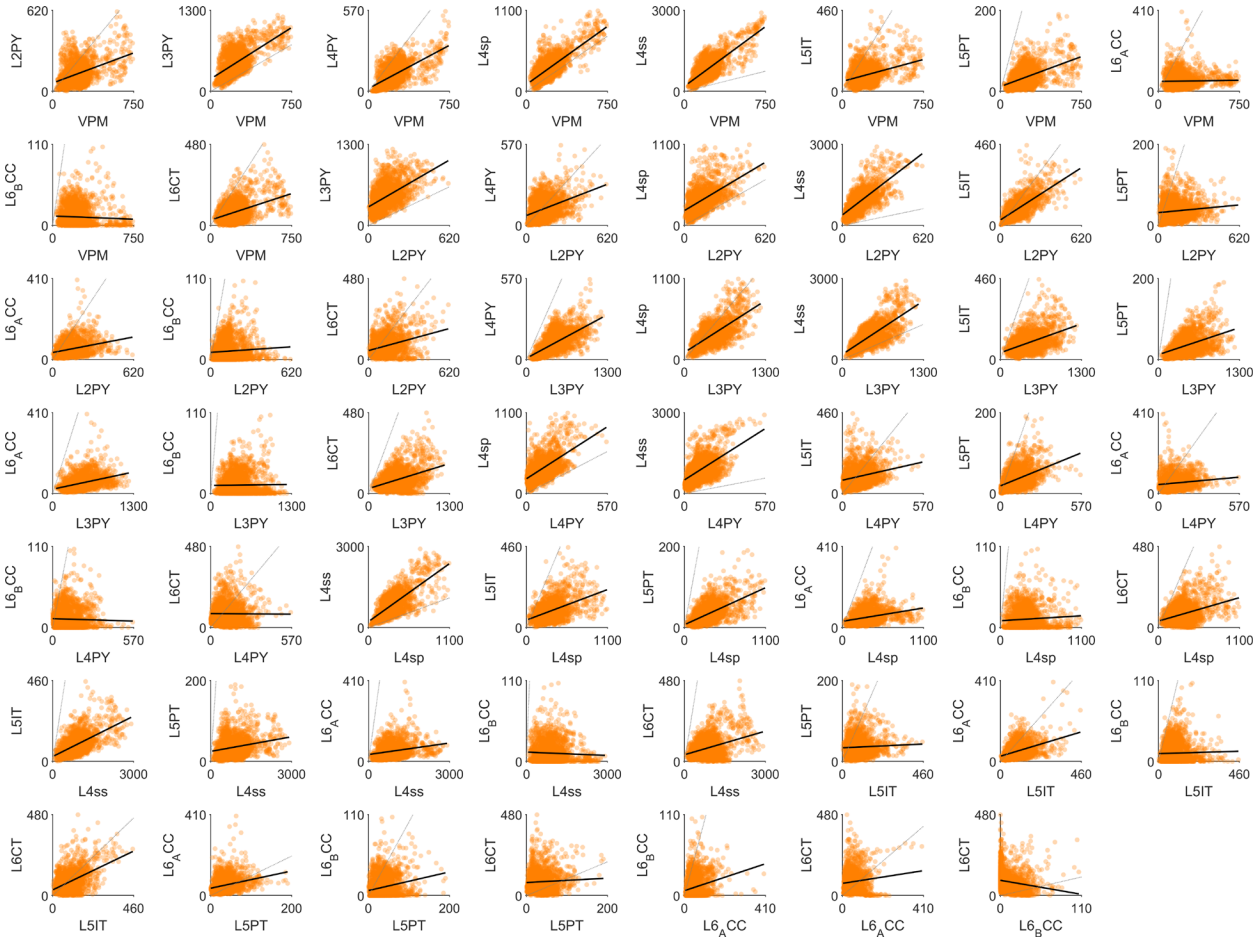

Figure S3 continued. Postsynaptic population C=L4ss.

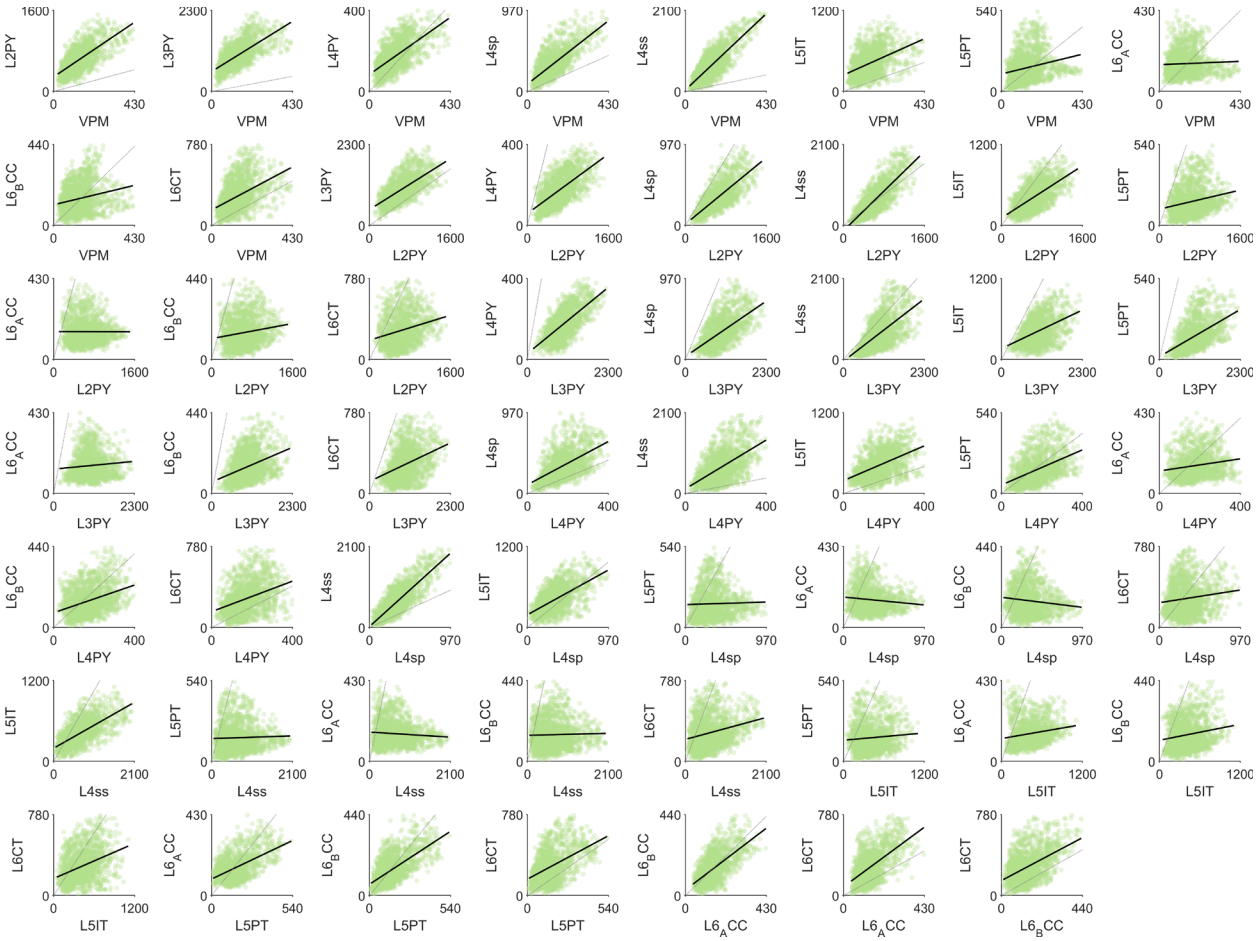

Figure S3 continued. Postsynaptic population C=L5IT.

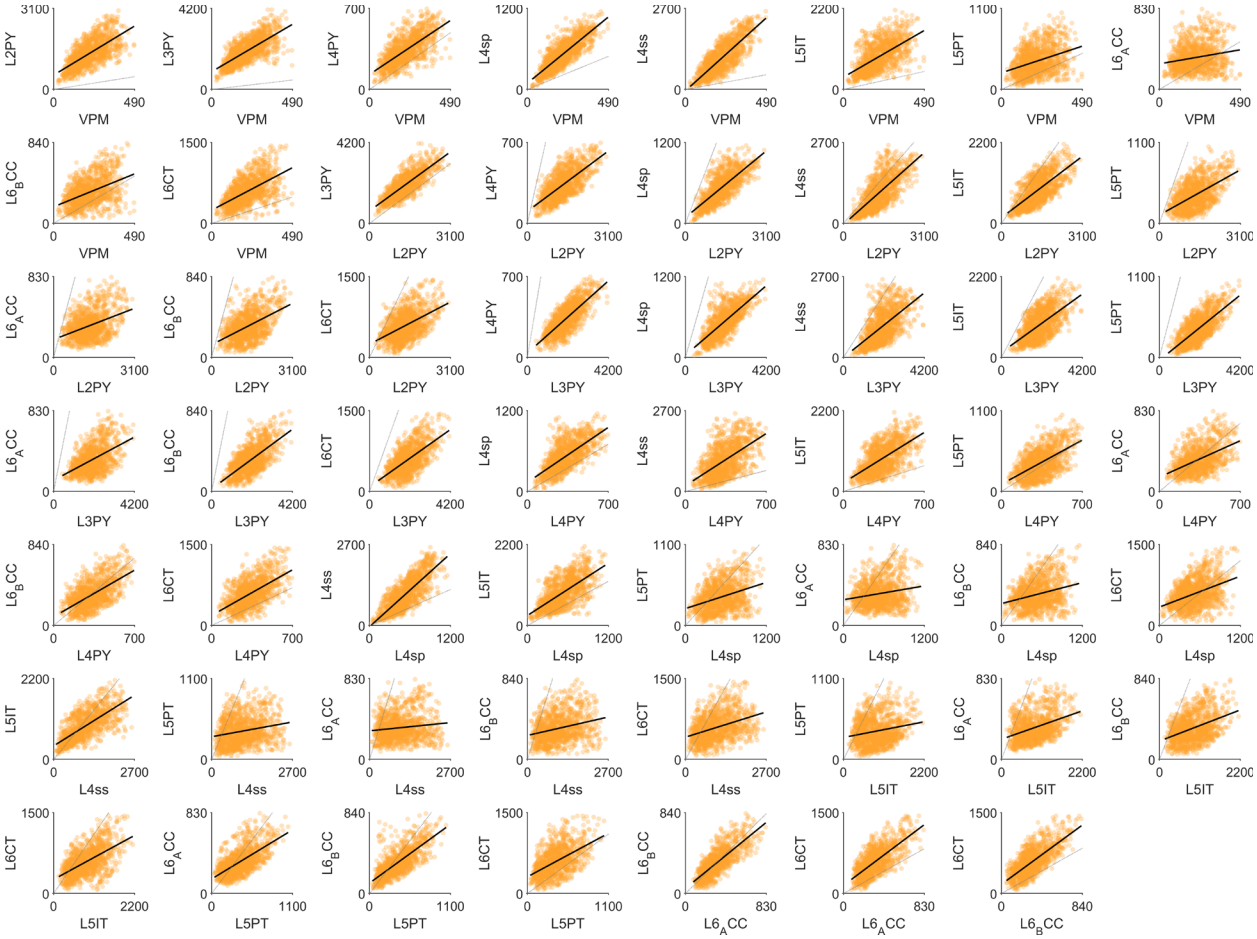

Figure S3 continued. Postsynaptic population C=L5PT.

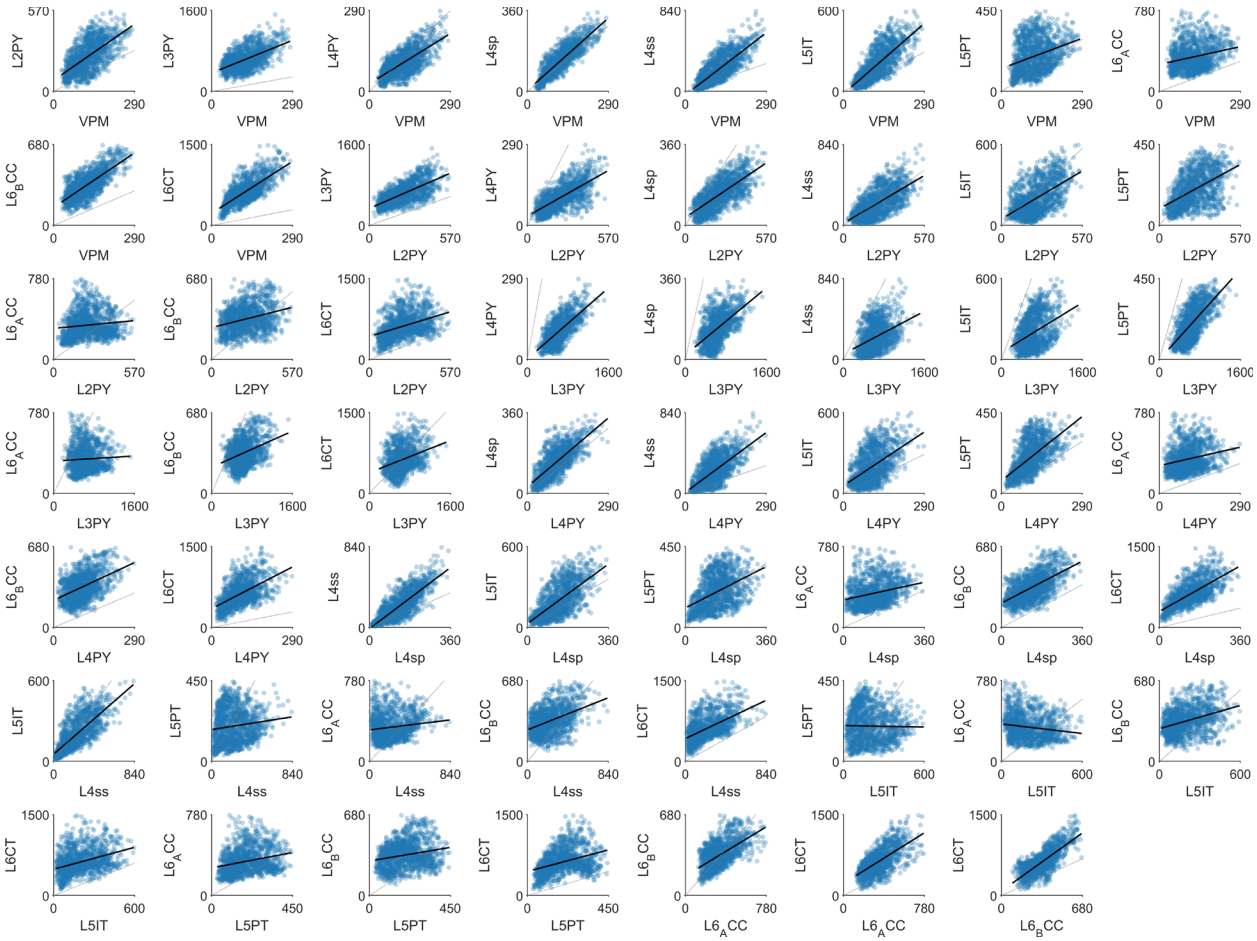

Figure S3 continued. Postsynaptic population C=L6$_A$CC.

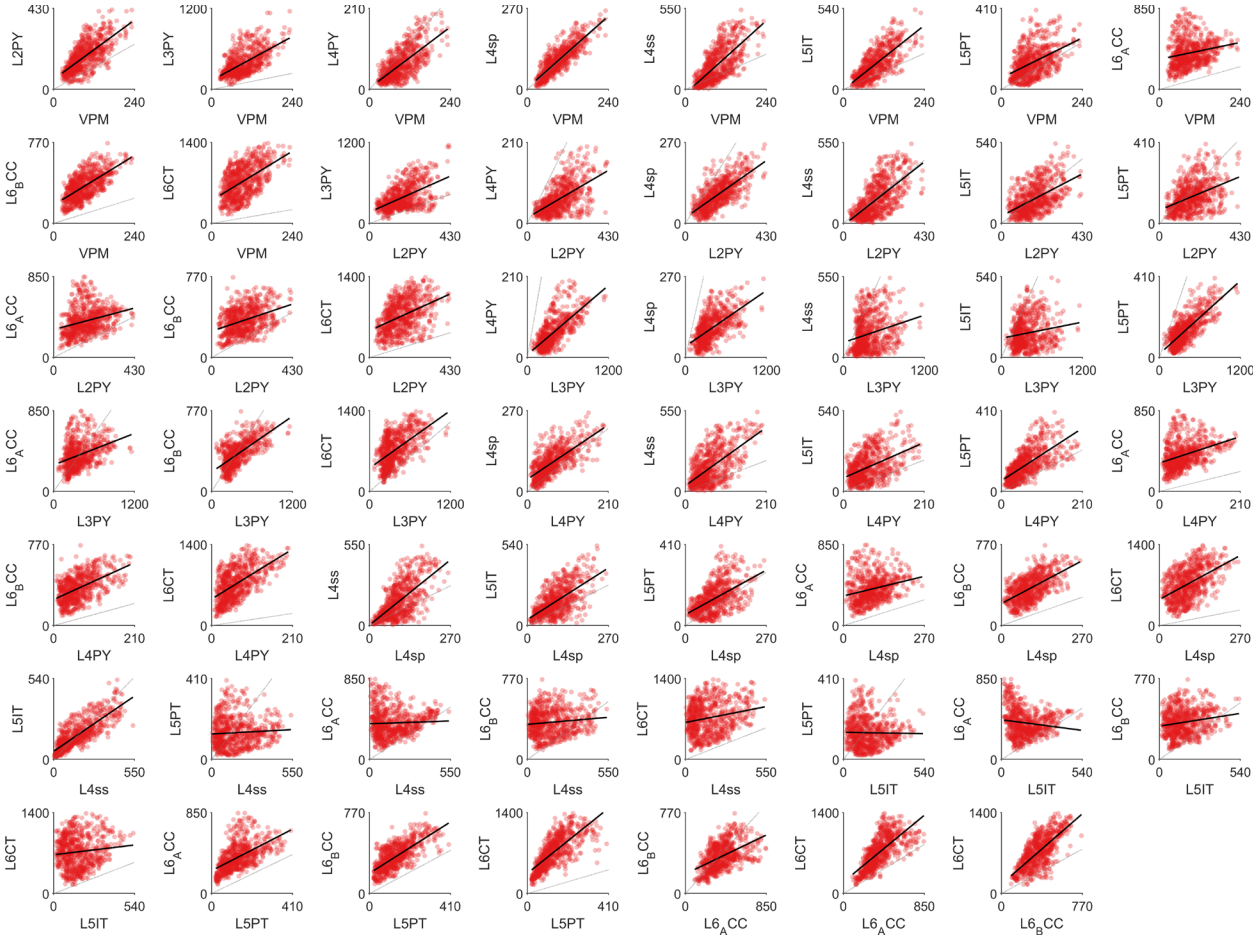

Figure S3 continued. Postsynaptic population C= $L6_BCC$.

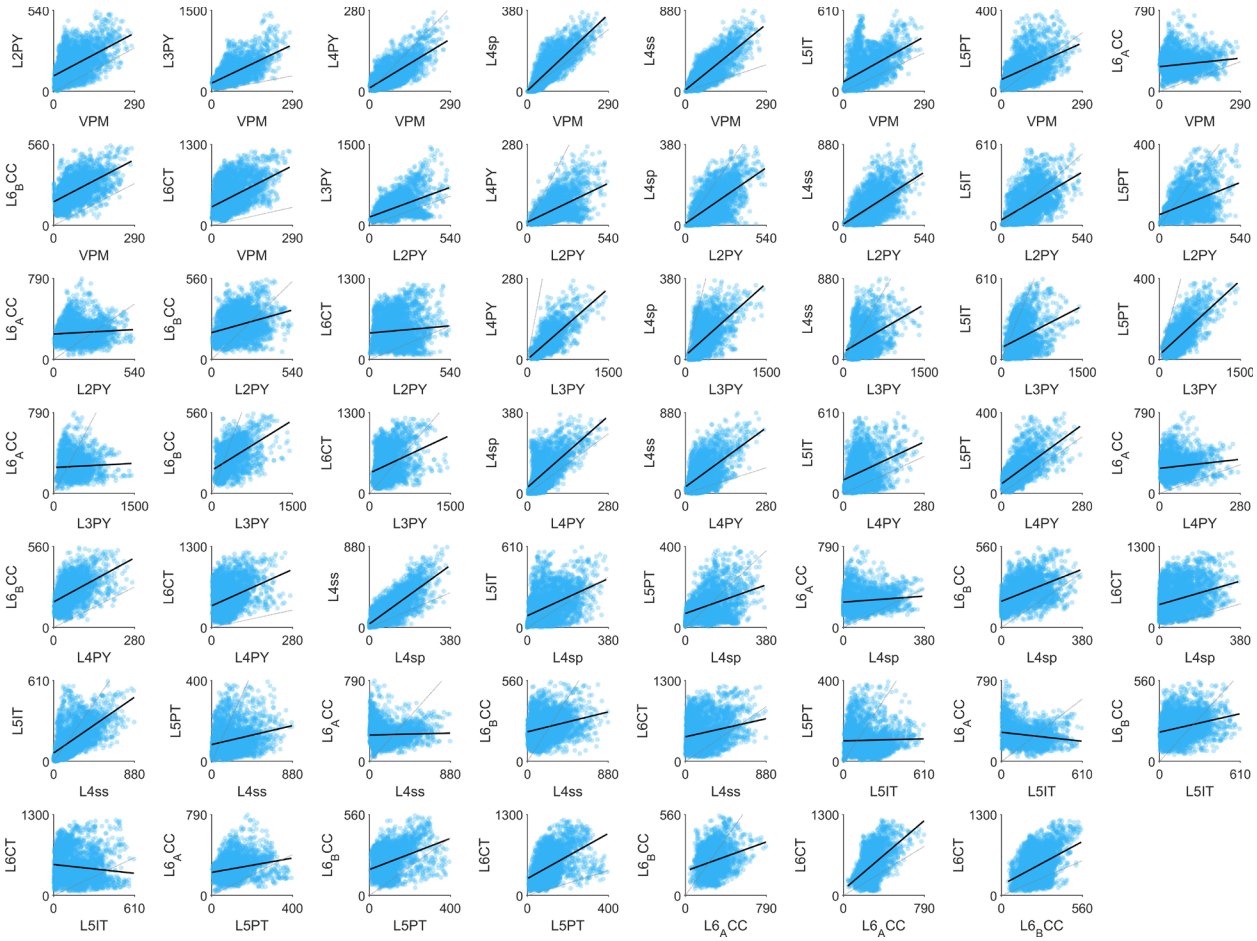

Figure S3 continued. Postsynaptic population C=L6CT.

Figure S4: Occurrences of triplet motifs deviate from those expected in random networks for all cell type groupings within one barrel column. Each neuron of the triplet sampled from different cell type.

Figure S4 continued. Neurons of the triplet sampled from one or two cell type(s).

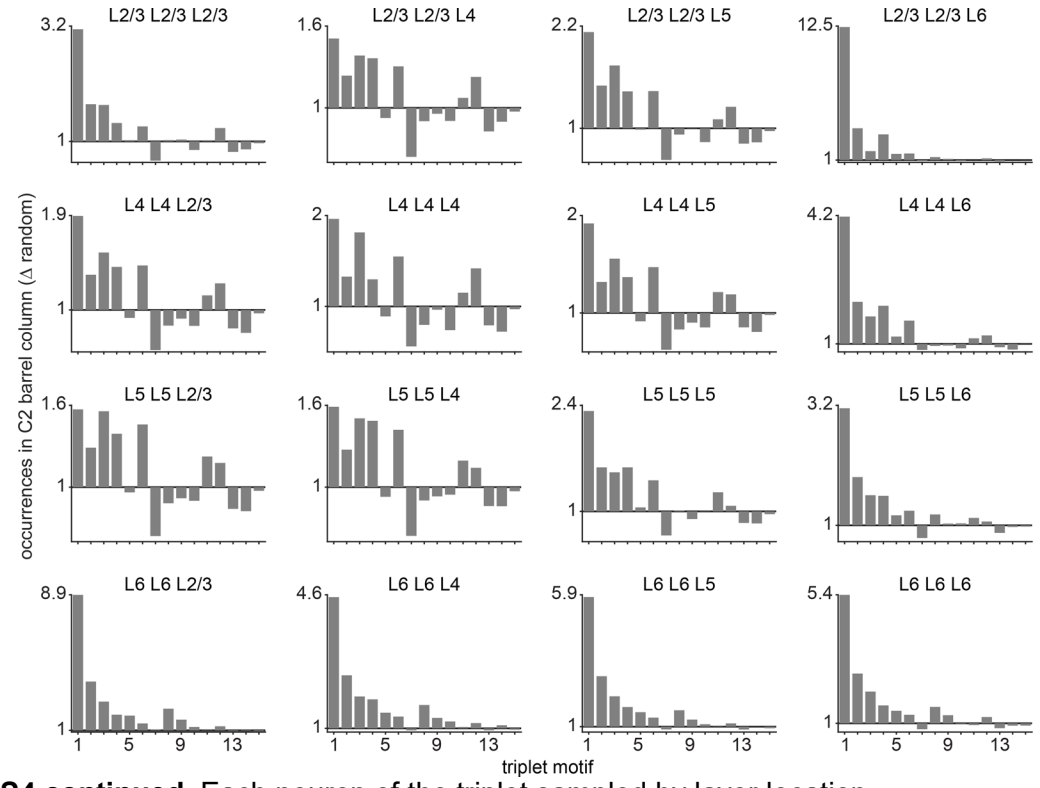

Figure S4 continued. Each neuron of the triplet sampled by layer location.

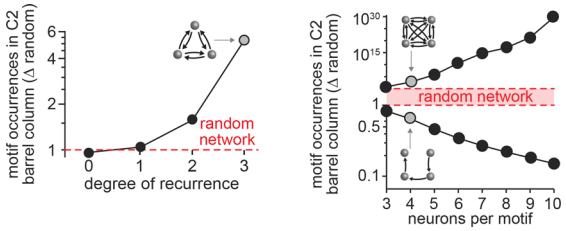

Figure S5: Predicted higher-order topologies deviate from those of random networks. Barrel cortex model predicts that the topological differences between barrel cortex and random networks increase depending on the number of bidirectional edges per motif (i.e., degree of recurrence), and the number of neurons (i.e., nodes) per motif, consistent with empirical observations (*42, 43*).

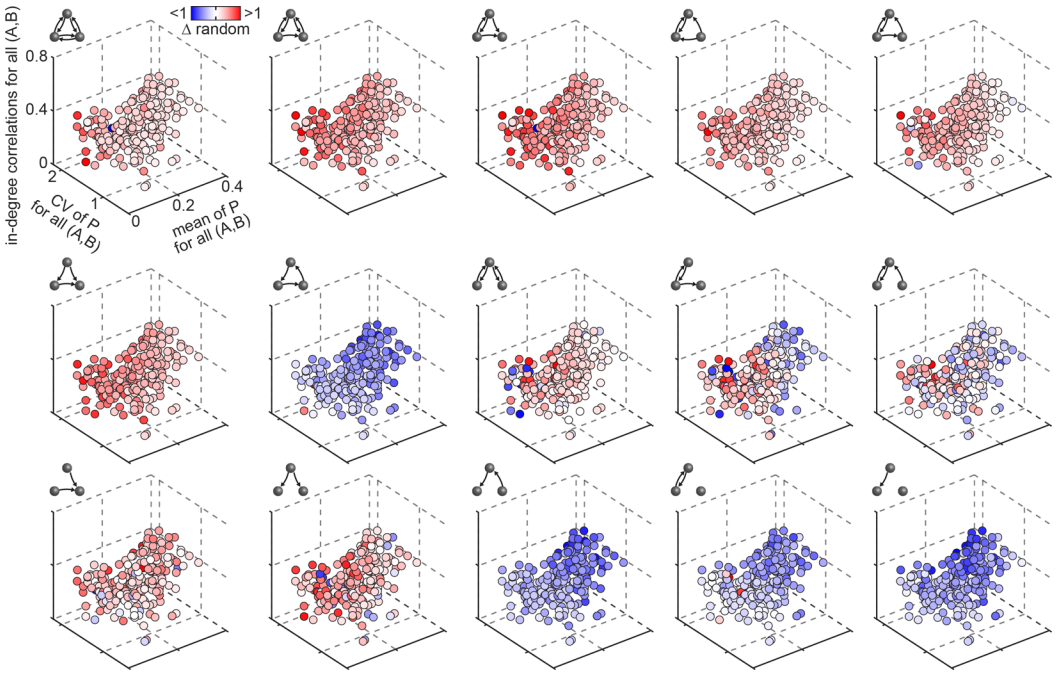

Figure S6: Relationship between connection probability distributions, their correlations, and deviations in the occurrences of triplet motifs. Deviations in the occurrences for all fifteen triplet motifs between the model and random networks as a function of the respective means and CVs of connection probability distributions, and in-degree correlations for all cell type-dependent groupings in the C2 barrel column.

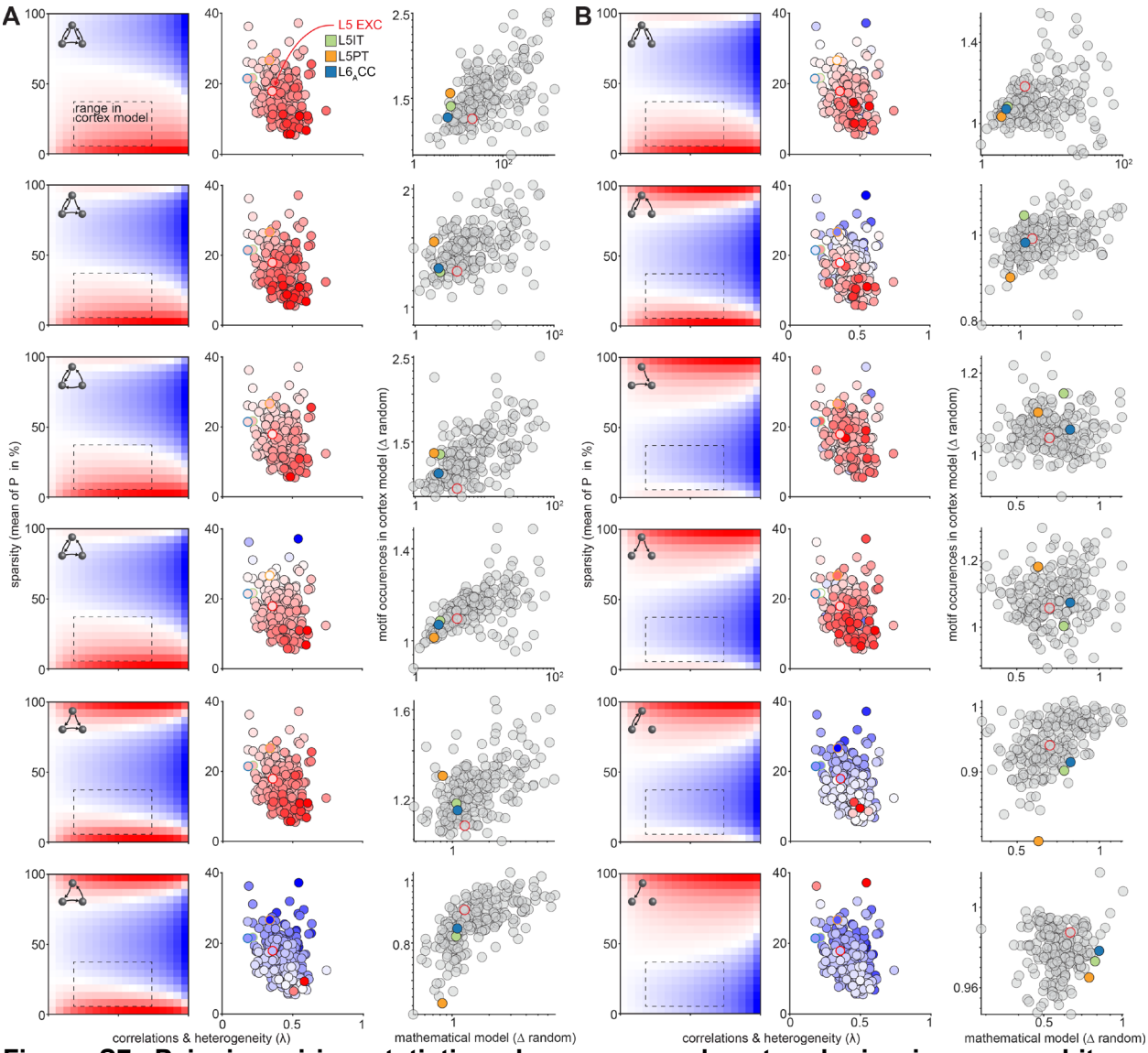

Figure S7: Pairwise wiring statistics shape non-random topologies in a non-arbitrary manner. (A) Left column: Mathematical model shows for triplet motifs (see **Fig. 7** for remaining triplet motifs) that the shapes and correlations of the underlying connection probability distributions define the specific non-random topology of networks. The dashed box reflects that range of sparsity and heterogeneity and correlations (i.e., λ) across all layer and/or cell type groupings in the barrel cortex model, as shown in the corresponding panels on the right. Center column: Respective distribution in barrel cortex model for all groupings. The colored circles represent the example groupings shown in **Fig. 6A**. Right column: Occurrences of motifs for all groupings in the barrel cortex model vs. the respective predictions by the mathematical model. **(B)** Same as in panel A but for remaining motifs.

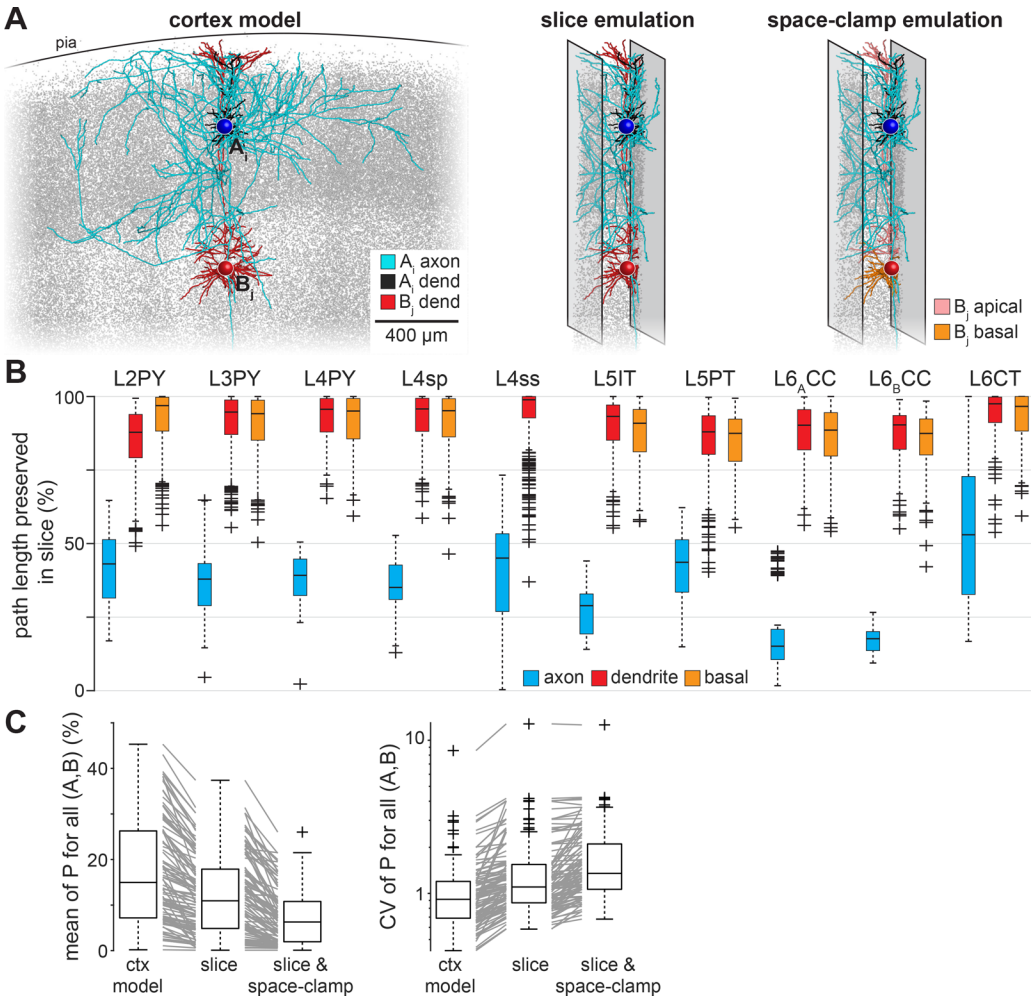

Figure S8: Emulating effects of slicing and space-clamping. (A) Illustration of slice and space-clamp emulation. Left: *In vivo* labeled dendrite (red, black) and axon (light blue) of two exemplary neurons (same as in **Fig. 5D)** in the barrel cortex model. Gray spheres represent somata. Center: Emulation of slicing procedure leads to truncation of neurites. Right: Emulation of space-clamping by restricting connections to those along proximal (i.e., basal) dendrites (orange). Connections to distal (i.e., apical) dendrites (light red) are ignored. **(B)** Percentages of path lengths preserved for axons (light blue), dendrites (red), and basal dendrites (orange) in multiple thalamocortical slices of 300 µm thickness. Neuron somata were sampled between 31 µm to 130 µm from slicing surface. **(C)** Impact of slicing and space-clamping onto the mean (left) and CV (right) of connection probabilities across all layer and/or cell type groupings (n=110).

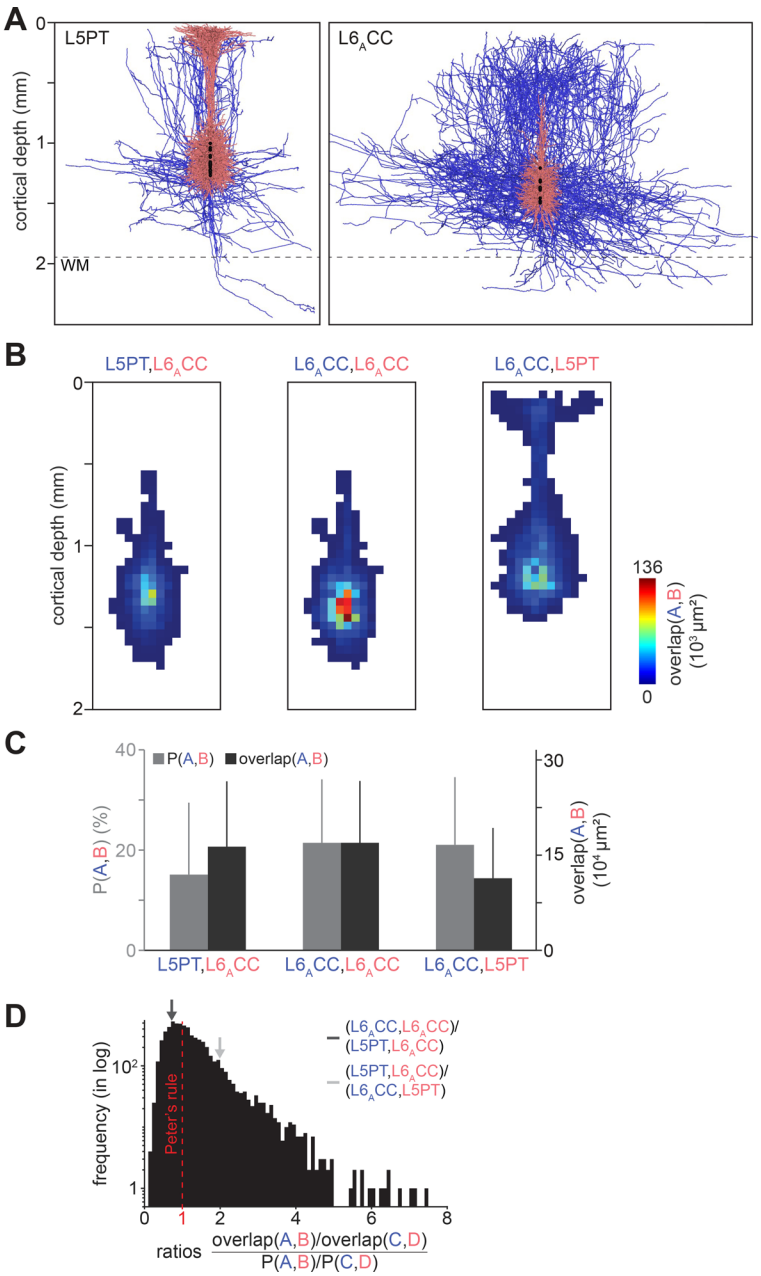

Figure S9: Predicted connection probabilities vs. axo-dendritic overlap. (A) Sample of L5PT and L6$_A$CC neuron morphologies horizontally aligned by their somata. Dendrite in red, axon in blue, somata as black spheres. **(B)** 2D density of axo-dendritic overlap (analogous to (*35*)) between three combinations of L5PT and L6$_A$CC neurons. **(C)** Predicted connection probabilities (gray) and axo-dendritic overlap (black) for the three combinations of panel B. **(D)** Histogram of ratios between axo-dendritic overlap and connection probability ratios for all cell type groupings. Ratio of one (dashed red line) indicates one-to-one relationship between differences in axo-dendritic overlap and connection probabilities.

Supplementary tables

Ref.	Grouping				Empirical			Model		Comparison	
	Presynaptic A		Postsynaptic B		*in vivo*	n	P(A,B) (%)	P(A,B) MEAN (%)	P(A,B) STD (%)	dev	prctl.
	layer	type	layer	type							
(*39*)	VPM		L4	n/s	yes	40	**43**	**44**	28	-0.04	50
(*73*)					yes	62	**37**			-0.23	44
(*39*)*					yes	14	**43**	**39**	28	0.13	56
(*73*)*					yes	21	**38**			-0.04	52
(*39*)				L4sp	yes	24	**42**	**45**	26	-0.11	45
(*39*)				L4ss	yes	11	**64**	**42**	27	0.82	76
(*74*)			L5	L5IT	yes	18	**17**	**28**	21	-0.55	36
(*74*)				L5PT	yes	9	**44**	**40**	24	0.19	57
(*74*)			L6	n/s	yes	11	**9**	**18**	19	-0.48	45
(*86*)	L2	n/s	L2	n/s	no	950	**9**	**16**	19	-0.34	51
(*83*)					yes	878	**7**	**31**	27	-0.89	26
(*86*)			L3	n/s	no	183	**5**	**15**	16	-0.58	39
(*86*)			L4	n/s	no	208	**1**	**9**	12	-0.64	38
(*86*)			L5A	n/s	no	211	**9**	**8**	10	0.12	68
(*86*)			L5B	n/s	no	108	**8**	**7**	9	0.18	71
(*86*)			L6	n/s	no	50	**0**	**1**	3	-0.39	69
(*69*)	L2/3	n/s	L2/3	n/s	no	95	**17**	**16**	18	0.05	63
(*77*)					no	-	**10**			-0.33	50
(*81*)					no	542	**5**			-0.61	38
(*82*)					no	112	**2**			-0.79	28
(*91*)					no	247	**26**			0.58	77
(*41*)					no	112	**20**			0.21	68
(*80*)					no	235	**19**			0.18	67
(*84*)					yes	774	**7**	**25**	23	-0.76	30
(*82*)			L5	L5IT	no	98	**4**	**7**	9	-0.34	51
(*82*)				L5PT	no	51	**4**	**9**	11	-0.51	41
(*86*)	L3	n/s	L2	n/s	no	182	**12**	**15**	18	-0.14	59
(*86*)			L3	n/s	no	513	**19**	**17**	18	0.10	64
(*86*)			L4	n/s	no	170	**2**	**11**	15	-0.61	40
(*91*)			L5	n/s	no	29	**55**	**8**	10	4.79	100
(*86*)			L5A	n/s	no	87	**6**	**9**	11	-0.28	53
(*86*)			L5B	n/s	no	164	**12**	**6**	9	0.69	82
(*86*)			L6	n/s	no	61	**0**	**1**	3	-0.39	68
(*86*)	L4	n/s	L2	n/s	no	208	**12**	**12**	14	0.03	64
(*78*)			L2/3	n/s	no	64	**15**	**15**	17	0.01	62
(*41*)					no	50	**20**			0.32	71
(*86*)			L3	n/s	no	172	**15**	**17**	18	-0.16	54
(*91*)					no	25	**28**			0.60	77
(*72*)			L4	n/s	no	89	**6**	**13**	16	-0.43	47
(*86*)					no	1046	**24**			0.72	80
(*79*)			L5A	n/s	no	-	**14**	**9**	10	0.50	77
(*86*)					no	276	**12**			0.27	71
(*86*)			L5B	n/s	no	136	**8**	**6**	8	0.24	71
(*86*)			L6	n/s	no	93	**3**	**1**	3	0.61	88

(*70*)		L4PY	L4	L4PY	no	528	**4**	**11**	12	-0.53	43
(*40*)		L4sp		L4sp	no	24	**21**	**19**	14	0.14	60
(*76*)		L4ss		L4ss	no	94	**26**	**21**	20	0.21	64
(*89*)					no	146	**36**			0.70	77
(*40*)					no	24	**21**			-0.02	57
(*91*)	L5	n/s	L3	n/s	no	29	**3**	**10**	15	-0.41	56
(*82*)			L5	n/s	no	150	**0**	**8**	10	-0.78	22
(*87*)					no	500	**10**			0.22	71
(*90*)					no	1450	**12**			0.42	76
(*91*)					no	163	**9**			0.14	69
(*82*)		L5IT	L2/3	n/s	no	98	**0**	**10**	13	-0.77	19
(*35*)			L5	L5IT	no	118	**5**	**11**	9	-0.62	34
(*82*)					no	66	**0**			-1.17	3
(*35*)				L5PT	no	86	**19**	**17**	13	0.10	59
(*82*)					no	36	**0**			-1.31	3
(*82*)		L5PT	L2/3	n/s	no	51	**0**	**2**	5	-0.45	55
(*35*)			L5	L5IT	no	86	**5**	**7**	8	-0.32	50
(*82*)					no	36	**0**			-0.88	13
(*35*)				L5PT	no	225	**7**	**16**	14	-0.65	34
(*82*)					no	12	**0**			-1.14	6
(*42*)					no	3446	**13**			-0.27	49
(*43*)					no	8050	**12**			-0.33	47
(*86*)	L5A	n/s	L2	n/s	no	209	**4**	**10**	14	-0.40	53
(*86*)			L3	n/s	no	89	**2**	**13**	17	-0.66	37
(*86*)			L4	n/s	no	275	**1**	**10**	14	-0.65	35
(*86*)			L5A	n/s	no	934	**19**	**9**	10	1.04	87
(*86*)			L5B	n/s	no	175	**8**	**8**	10	0.00	65
(*86*)			L6	n/s	no	158	**3**	**2**	4	0.34	82
(*86*)	L5B	n/s	L2	n/s	no	104	**1**	**3**	8	-0.23	77
(*86*)			L3	n/s	no	167	**2**	**3**	7	-0.14	75
(*86*)			L4	n/s	no	137	**1**	**2**	4	-0.29	68
(*86*)			L5A	n/s	no	174	**2**	**5**	8	-0.48	48
(*85*)			L5B	n/s	no	269	**9**	**11**	12	-0.11	58
(*86*)					no	555	**7**			-0.29	51
(*86*)			L6	n/s	no	100	**7**	**5**	9	0.19	75
(*86*)	L6	n/s	L2	n/s	no	50	**0**	**0**	2	-0.19	91
(*86*)			L3	n/s	no	64	**0**	**1**	4	-0.26	85
(*86*)			L4	n/s	no	94	**0**	**1**	4	-0.31	74
(*86*)			L5A	n/s	no	160	**1**	**4**	6	-0.48	53
(*86*)			L5B	n/s	no	100	**2**	**9**	11	-0.70	33
(*71*)			L6	n/s	no	102	**2**	**7**	11	-0.49	51
(*86*)					no	532	**3**			-0.41	54
(*88*)					no	27	**4**			-0.33	57
(*75*)	L6A	L6CC	L6A	L6CT	no	43	**9**	**7**	9	0.22	68
(*75*)	L6A	L6CT	L6A	L6CC	no	40	**0**	**10**	12	-0.89	14

Table S1. Statistical connectome vs. empirical data. *denotes connectivity measurements of neurons located between barrel columns (i.e., within the septum). n denotes the number of pairs tested. Deviation between empirically observed connection probability and the model's prediction: dev = (P(A,B) empirical – mean of P(A,B) model) / (STD of P(A,B) model). prctl denotes percentage of predicted connection probabilities that are lower or equal than P(A,B) empirical

(e.g., prctl = 0: none of the predicted connection probabilities are lower or equal than P(A,B) empirical; prctl = 100: all predicted connection probabilities are lower or equal than P(A,B) empirical).

Inter-somatic distance (μm)	P(A,B) empirical (%)	median (%)	P(A,B) model 25th percentile (%)	75th percentile (%)
18	22	21	11	33
53	17	19	9	30
88	15	16	7	27
123	13	11	5	23
158	10	8	2	19
193	10	6	1	14
228	7	2	0	11
263	4	2	0	10
298	7	5	0	13

Table S2. Comparison with inter-somatic distance-dependent connection probabilities in L5. Empirical data from (*42*).

Inter-somatic distance (μm)	P(A,B) empirical (%)	median (%)	P(A,B) model 25th percentile (%)	75th percentile (%)
20	21	19	6	36
60	17	17	5	32
100	13	13	3	26
140	14	9	1	21

Table S3. Comparison with inter-somatic distance-dependent connection probabilities in L2/3. Empirical data from (*69*).