

SUPPLEMENTARY MATERIAL

Homophilic wiring principles underpin neuronal network topology *in vitro*

Danyal Akarca^{1*}, Alexander W. E. Dunn^{2,3*}, Philipp J. Hornauer⁴, Silvia Ronchi⁴, Michele Fiscella⁴, Congwei Wang⁵, Marco Terrigno⁵, Ravi Jagasia⁵, Petra E. Vértes⁶, Susanna B. Mierau^{2,7,8}, Ole Paulsen², Stephen J. Eglén³, Andreas Hierlemann⁴, Duncan E. Astle^{1†}, Manuel Schröter^{4†}

1. MRC Cognition and Brain Sciences Unit, University of Cambridge, UK
2. Department of Physiology Development and Neuroscience, University of Cambridge, UK
3. Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK
4. Department of Biosystems Science and Engineering in Basel, ETH Zurich, Switzerland
5. NRD, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Basel, Switzerland
6. Department of Psychiatry, University of Cambridge, Cambridge, UK
7. Division of Cognitive and Behavioral Neurology, Brigham & Women's Hospital, Boston, USA
8. Harvard Medical School, Boston, USA

Corresponding author:

Dr Danyal Akarca

Email: danyal.akarca@mrc-cbu.cam.ac.uk

*Co-lead first authors

†Co-lead senior authors

CONTENTS

Supplementary Figures

Supplementary Figure 1. Tracking neuronal networks at cellular resolution on high-density microelectrode arrays.

Supplementary Figure 2. Global topological measures of sparse rodent cultures over time.

Supplementary Figure 3. Distribution of inter-neuronal Euclidean distances across rodent and human datasets.

Supplementary Figure 4. Generative network modeling results for the sparse rodent primary cortical cultures.

Supplementary Figure 5. Generative network model energy acquired from size- and density-matched random networks.

Supplementary Figure 6. Generative model fits recapitulate observed network statistics.

Supplementary Figure 7. Comparison of global network statistics across sparse and dense primary cortical rodent networks at DIV14.

Supplementary Figure 8. Generative network modeling results for the dense rodent primary cortical cultures.

Supplementary Figure 9. Homophilic generative mechanisms best account for local relationships in developing dense rodent neuronal cultures.

Supplementary Figure 10. Relationship between model energy and topological fingerprint dissimilarity.

Supplementary Figure 11. Cellular and network differences following chronic application of gabazine.

Supplementary Figure 12. Generative model comparisons between control and gabazine-treated sparse rodent PC cultures.

Supplementary Figure 13. Generative network modeling of human iPSC-derived neuronal cultures and cerebral organoids.

Supplementary Tables

Supplementary Table 1. Overview of all the datasets used in the study.

Supplementary Table 2. A list of all the value K_{ij} terms that were included in the generative modeling, as given in the wiring equation.

Supplementary Table 3. Statistical comparisons of rodent 50k neuronal culture energy comparisons across generative rules.

Supplementary Table 4. Statistical comparisons of rodent 100k neuronal culture energy comparisons across generative rules.

Supplementary Table 5. Statistical comparisons of rodent 50k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

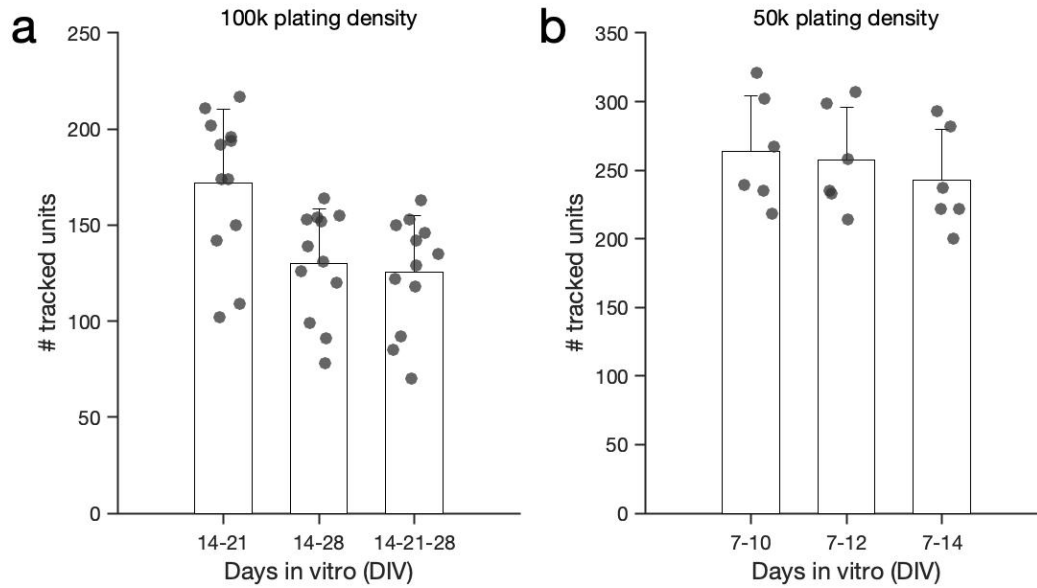
Supplementary Table 6. Statistical comparisons of rodent 100k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

Supplementary Table 7. Statistical comparisons of DIV28 human iPSC neuronal culture (glutamatergic neurons, motor neurons and dopaminergic neurons) energy comparisons across generative rules.

Supplementary Table 8. Statistical comparisons of human cerebral organoid energy comparisons across generative rules.

Supplementary Table 9 Overview of used antibodies.

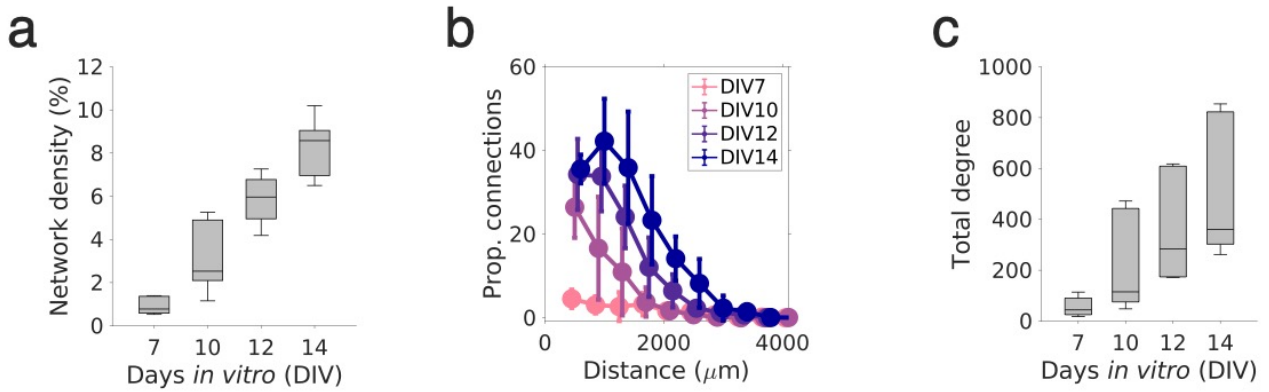
SUPPLEMENTARY FIGURES



Supplementary Figure 1.

Tracking neuronal networks at cellular resolution on high-density microelectrode arrays.

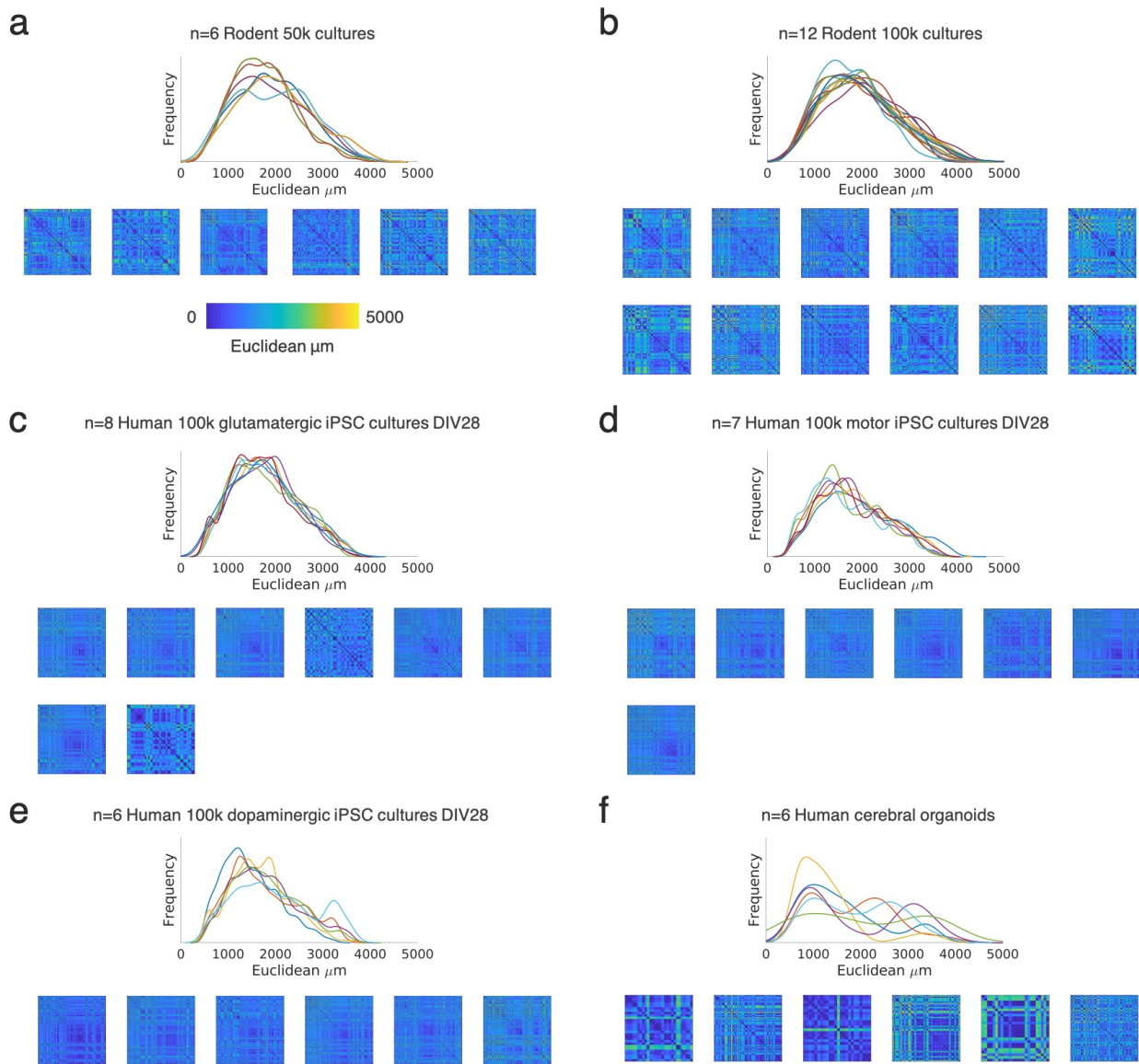
a Bar plot depicts the tracking result for the rodent data with 100k plating density. On average more than 100 neurons per HD-MEA could be tracked over three weeks in vitro. **b** The same plot as in panel a, for the primary cortical cultures plated at 50,000 cells per high-density microelectrode array (HD-MEA).



Supplementary Figure 2.

Global topological measures of sparse rodent cultures over time.

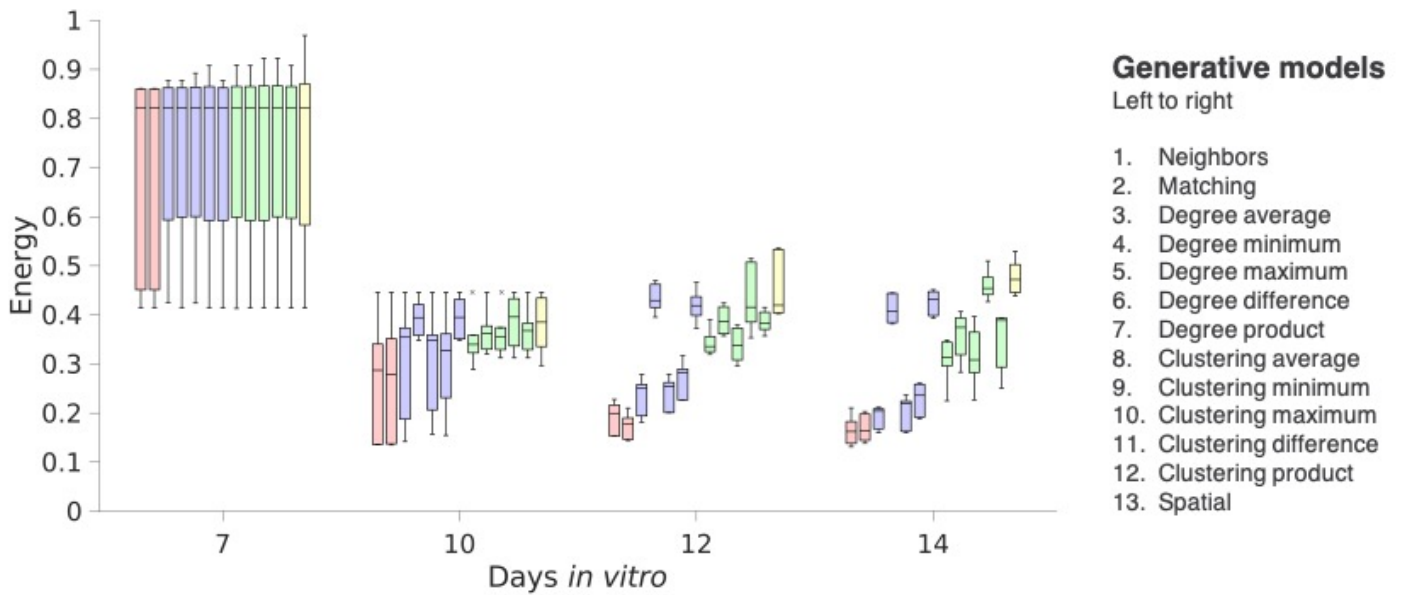
a Panel depicts the statistically inferred network density for the sparse (50,000 cell per well) primary cortical (PC) neuronal networks until days in vitro (DIV)14. **b** Proportion of extant connection by distance and grouped by DIV7 (pink), DIV10 (light purple), DIV12 (dark purple) and DIV14 (dark blue). **c** The total degree for sparse PC networks across development.



Supplementary Figure 3.

Distribution of inter-neuronal Euclidean distances across rodent and human datasets.

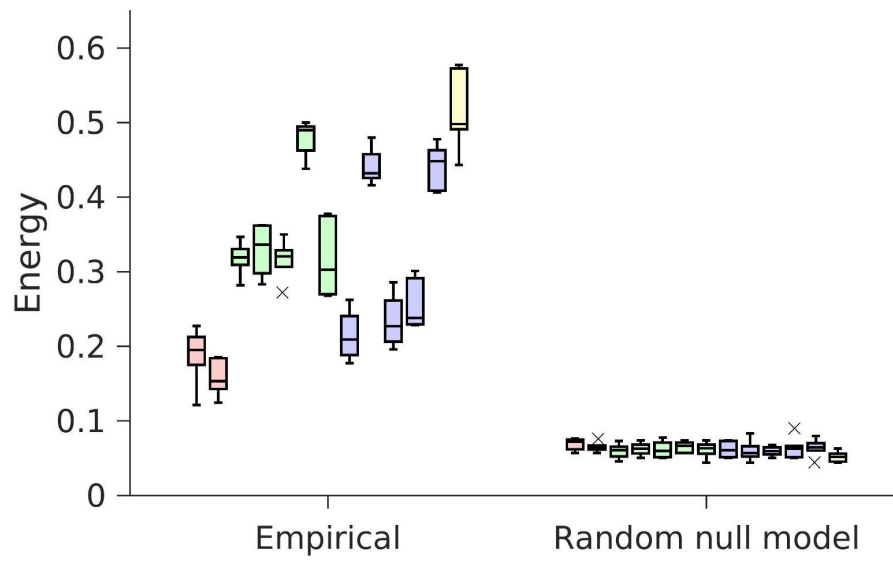
a Density plot of inter-neuronal Euclidean distances for each culture. Note, each matrix was used as the $D_{i,j}$ term for the generative model constructed for each time point of that culture. Panel **a** depicts Euclidean distances for the sparse PC rodent networks (50,000 neurons per well); the left panel shows the overall distributions; the right panel shows the individual distance matrices. **b** Euclidean distances for the dense PC rodent networks (100,000 neurons per well). Panels **c-e** show Euclidean distance distribution across the iPSC-derived human neuron lines at DIV28 (**c**, glutamatergic; **d**, motor and **e**, dopaminergic neurons). **f** Euclidean distances for the human cerebral organoid recordings.



Supplementary Figure 4.

Generative network modeling results for the sparse rodent primary cortical cultures.

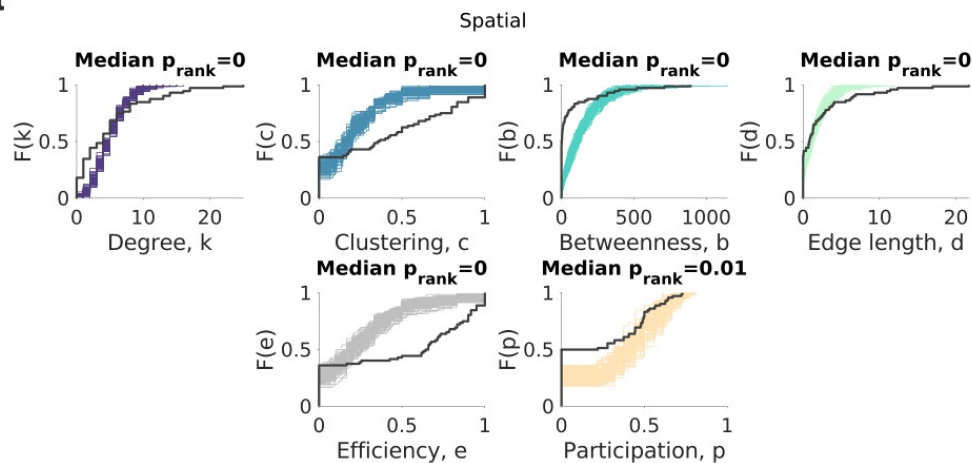
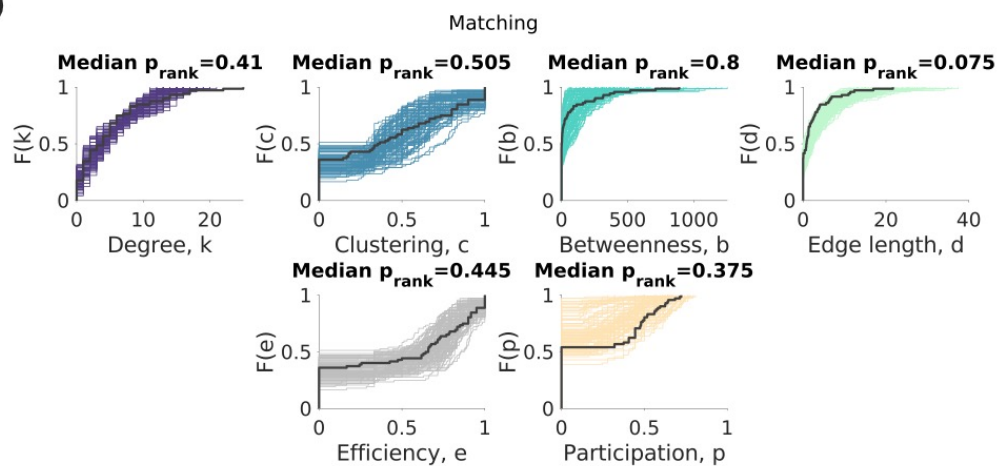
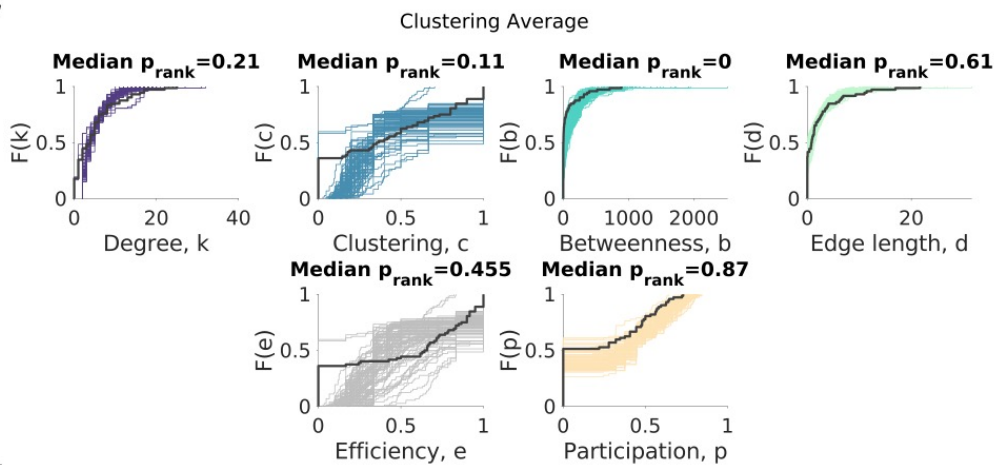
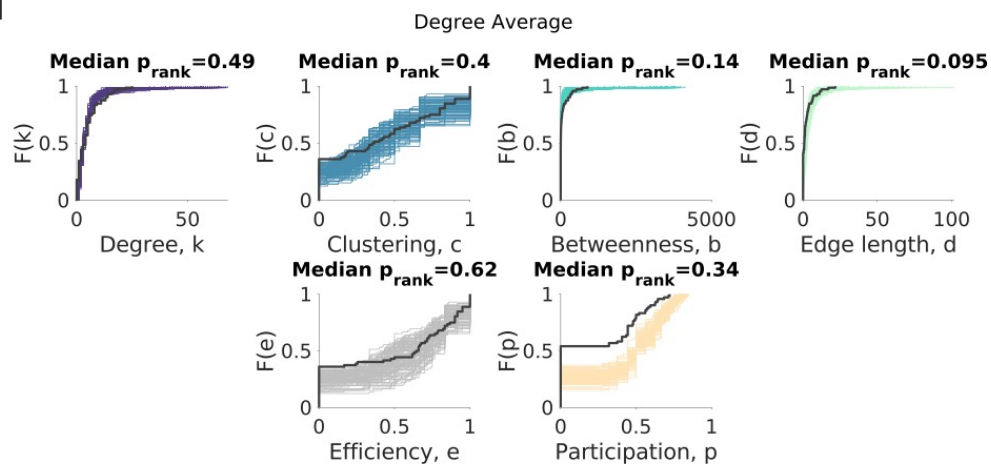
Generative model fits (energy), including all 13 wiring models, for the sparse primary rodent networks (50,000 cells per well; n=6 cultures) across development. Each boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of the top n=1 performing simulations are shown.



Supplementary Figure 5.

Generative network model energy acquired from size- and density-matched random networks.

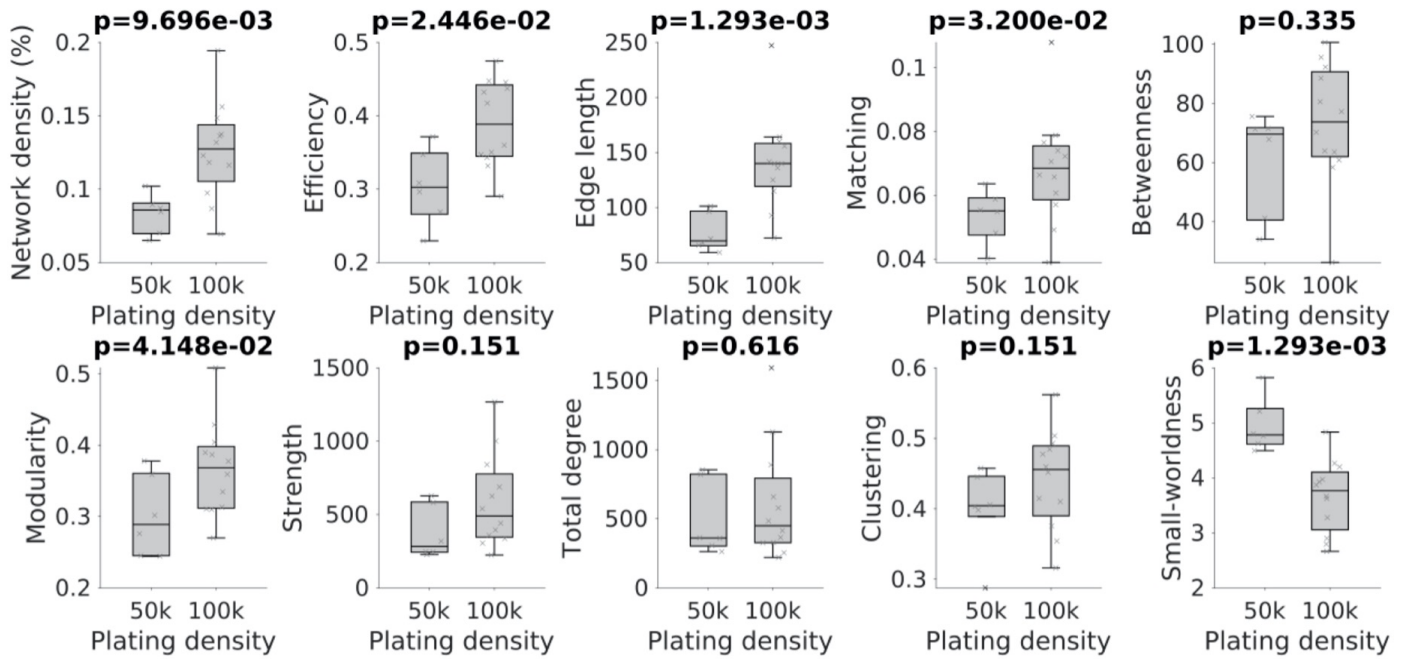
Each boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of top n=1 performing simulation is shown.

a**b****c****d**

Supplementary Figure 6.

Generative model fits recapitulate observed network statistics.

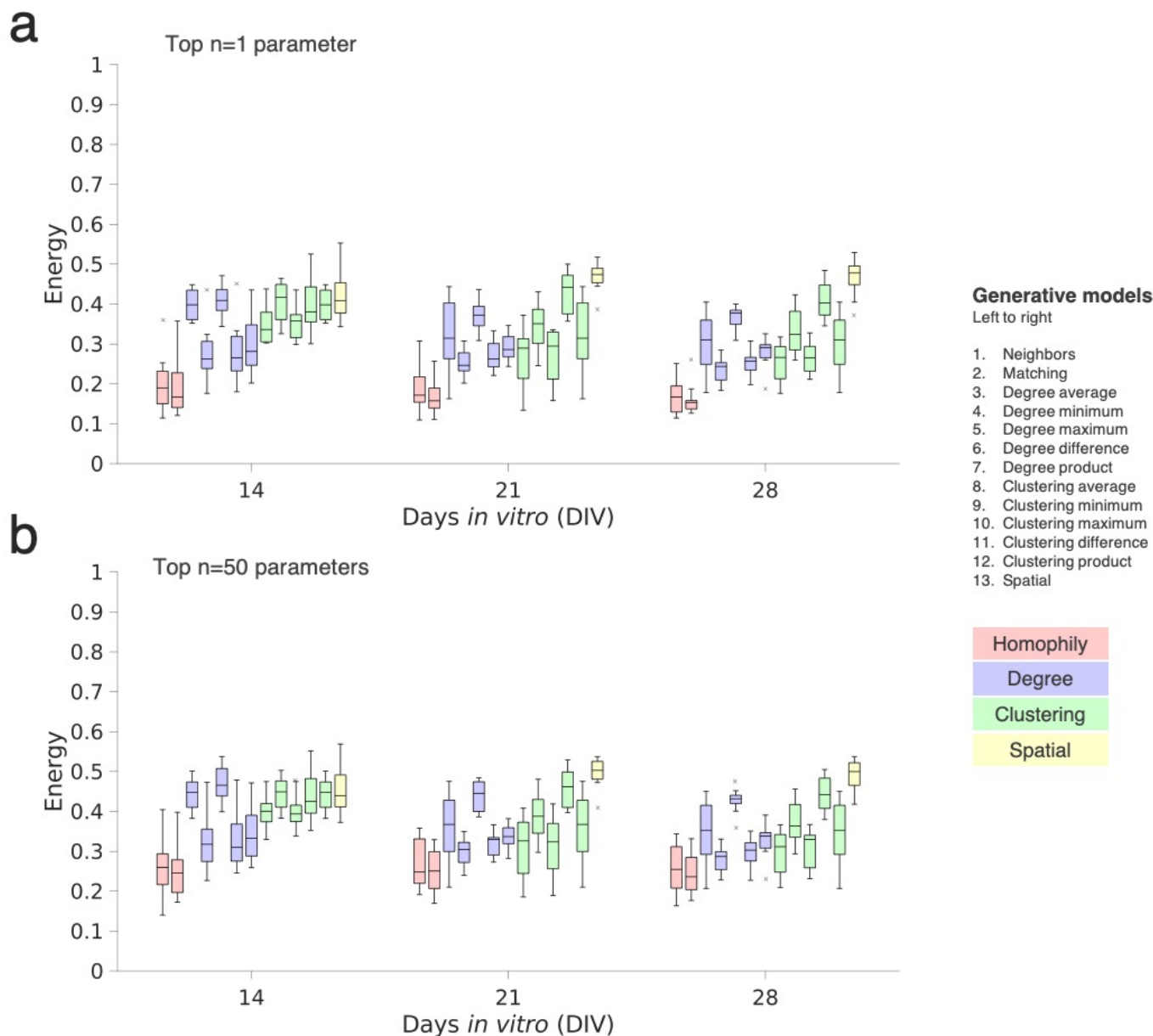
a Cumulative density functions (CDFs) of the top=99 simulations (of the total 20,000 simulations) for the best performing generative models in each class (**a**, spatial. **b**, matching. **c**, clustering average. **d**, degree average) across the four statistics included in the energy equation (top four panels) and two additional metrics - the local efficiency and participation coefficient - not included in the energy equation (bottom two panels). For visualization, we show only the solutions for a single sparse rodent PC culture. p values were computed using the Monte-Carlo bootstrapping procedure outlined in **Methods; Generative network models**.



Supplementary Figure 7.

Comparison of global network statistics across sparse and dense primary cortical rodent networks at DIV14.

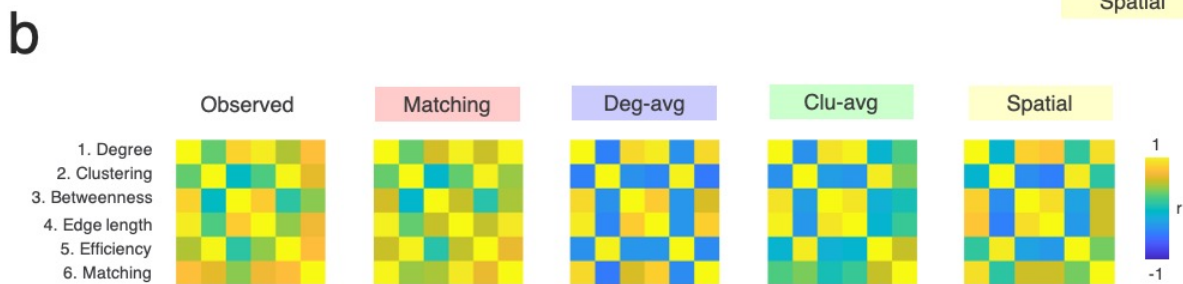
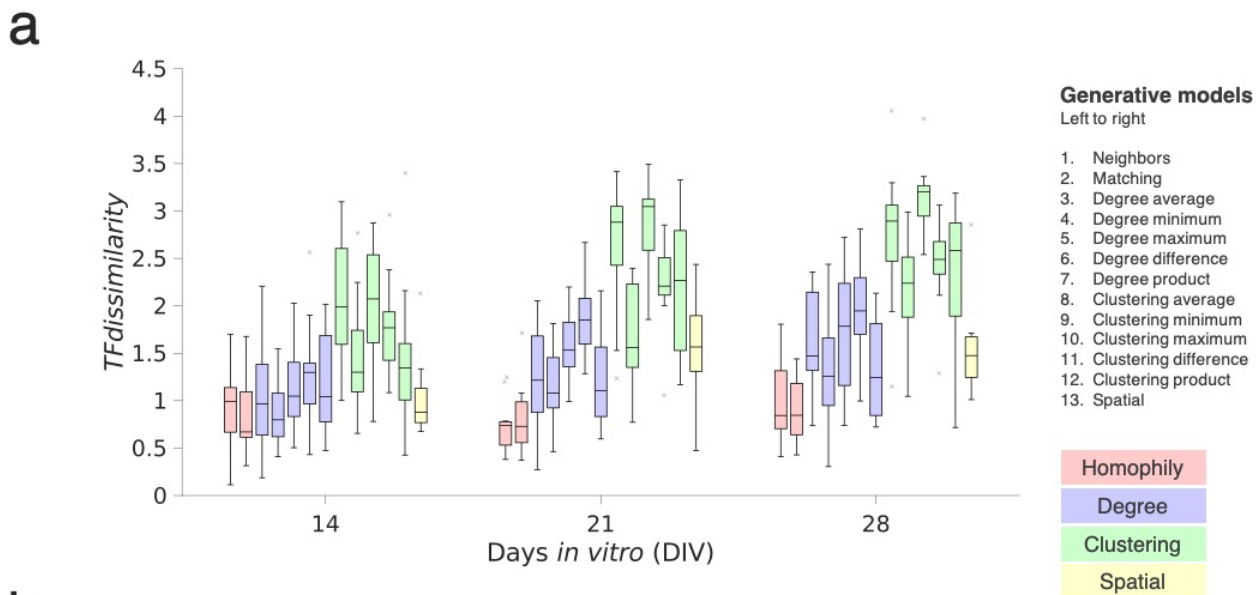
Comparisons were computed for network density, efficiency, edge length, matching, betweenness, modularity, strength, total degree, clustering and small worldness. For each comparison we provide the p-value computed from a Mann-Whitney U test. Sparse networks were plated at 50,000 cells per well, dense networks were plated at 100,000 cells per well.



Supplementary Figure 8.

Generative network modeling results for the dense rodent primary cortical cultures.

Generative model fits (energy), including all 13 wiring models, for the dense primary rodent networks (100,000 cells per well; n=12 cultures) across development. **a** Top performing n=1 parameter combination. **b** Top performing n=50 parameter combinations. Each boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of the top n=1 performing simulations are shown.

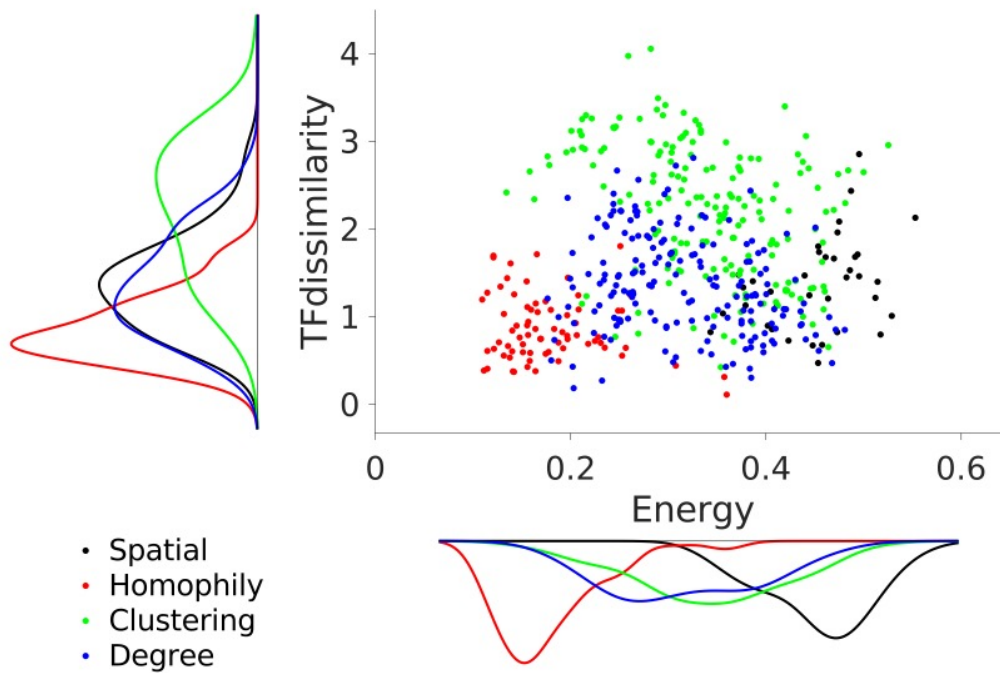


Supplementary Figure 9.

Homophilic generative mechanisms best account for local relationships in developing dense rodent neuronal cultures.

a Homophily generative models produce the lowest *TFdissimilarity* across all time-points (DIV14, 21 and 28), suggesting that it can reconstruct local connectivity patterns of *in vitro* neuronal networks. In total, there are $n=12$ data points (one per culture) shown in each of the 13 boxplots. Boxplots presents the median and IQR. Outliers are demarcated as small gray crosses, and are those which exceed 1.5 times the interquartile range away from the top or bottom of the box. **b** A visualization of the averaged topological organization matrix for the observed (left), matching (middle left), degree-average (middle), clustering-average (middle right), and spatial (right) models at DIV28.

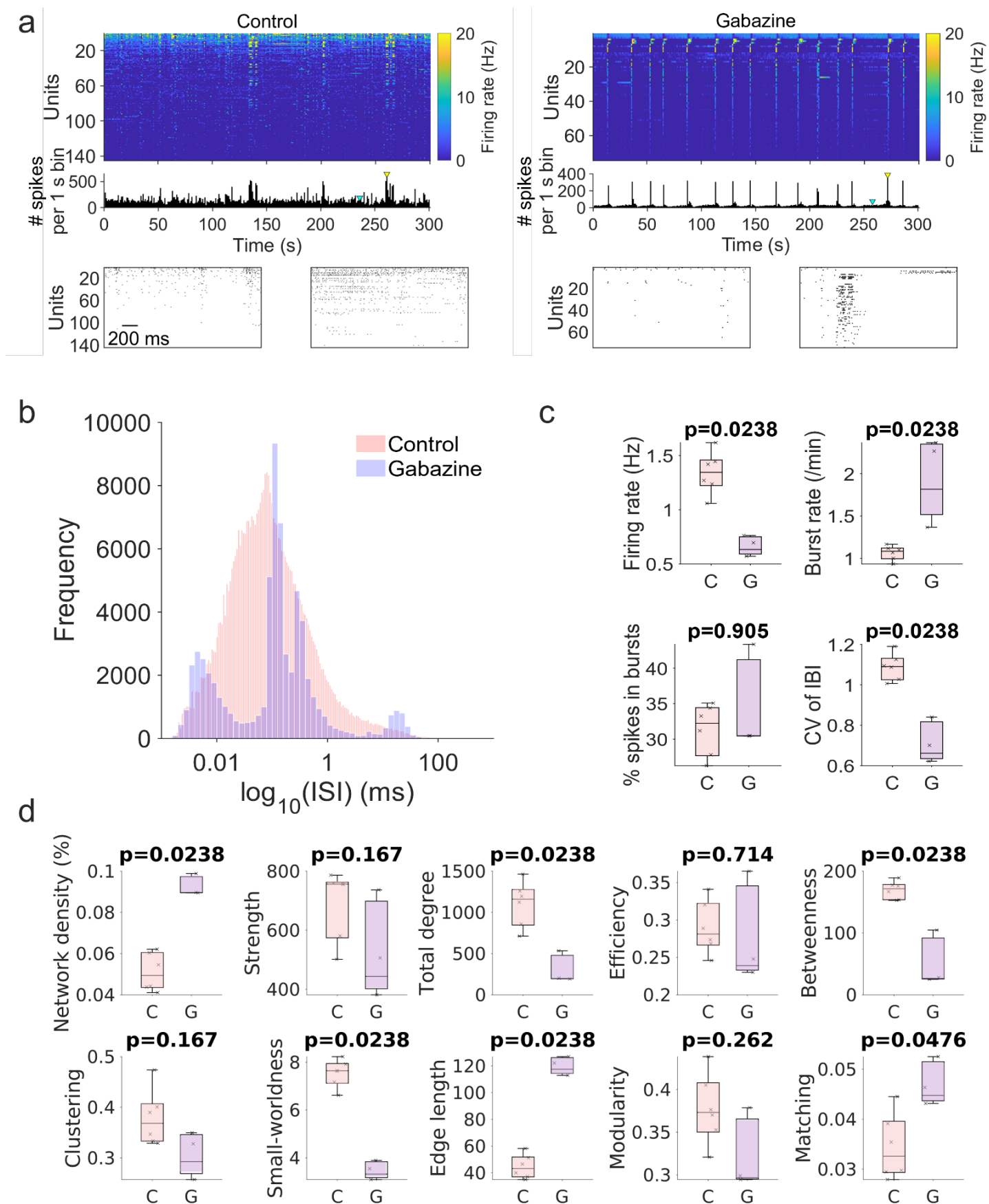
Rodent 100k cultures



Supplementary Figure 10.

Relationship between model energy and topological fingerprint dissimilarity in developing dense rodent neuronal cultures.

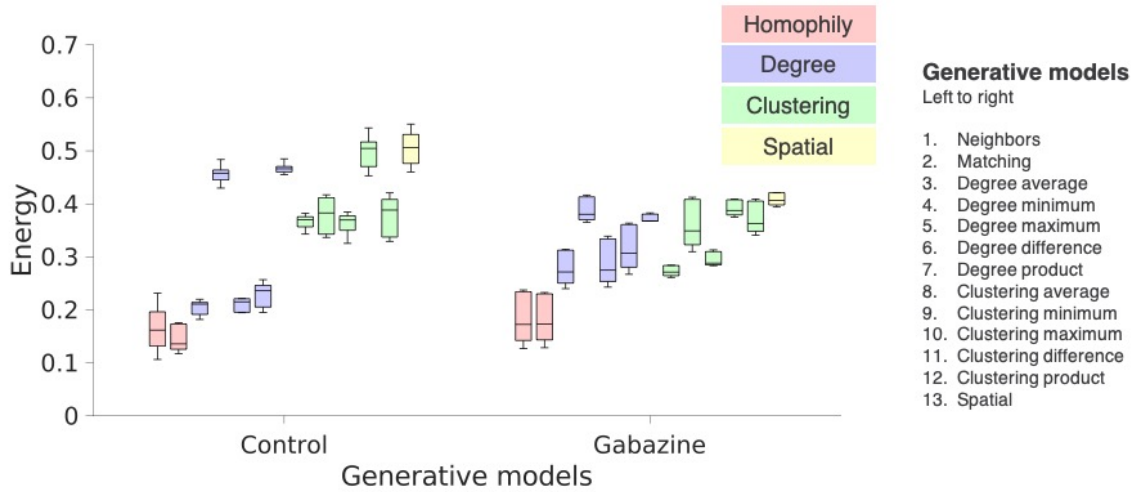
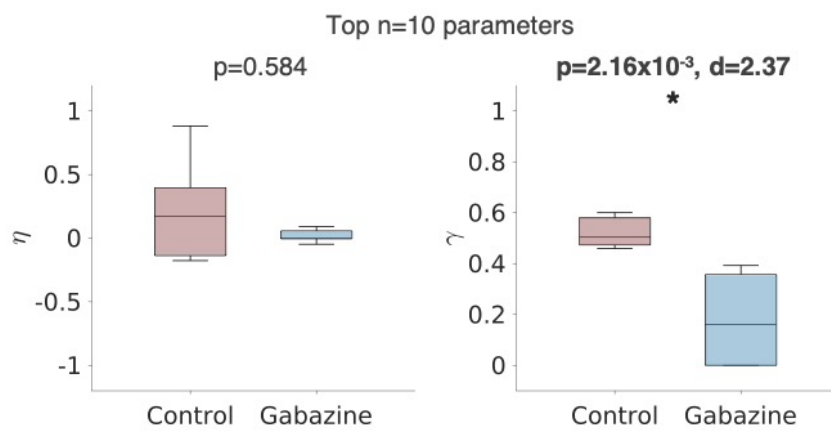
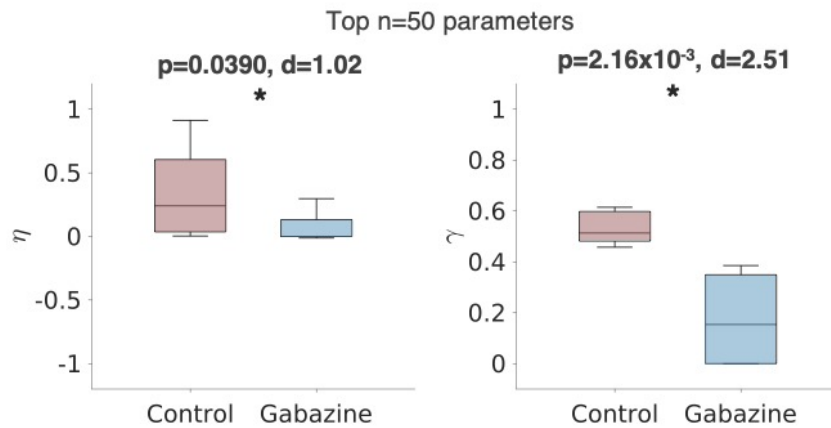
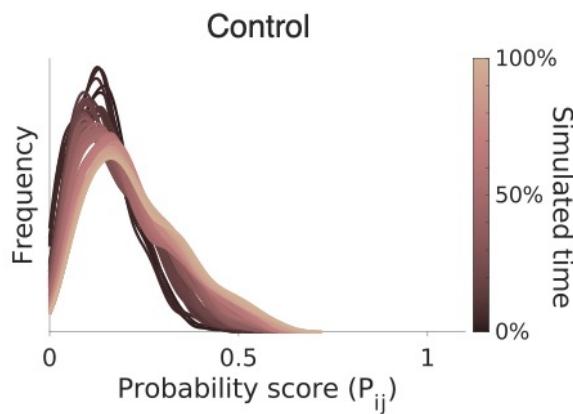
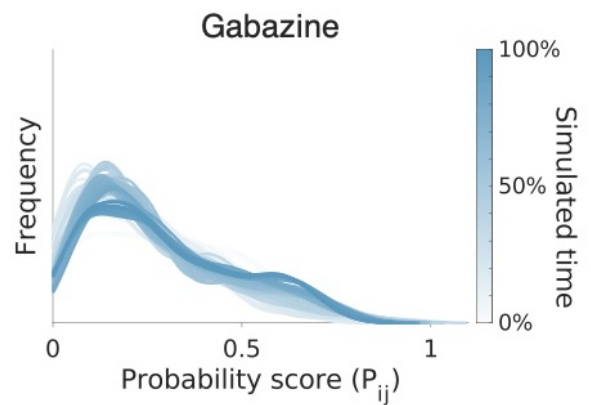
Using dense PC rodent network (100,000 cells per MEA), we show $n=468$ data points ($n=6$ cultures \times $n=3$ time-points \times $n=13$ generative model simulations) corresponding to the top $n=1$ performing simulation's energy and its topological organization dissimilarity performances. Distributions are plotted for each, showing homophily to achieve the best fits in both (bottom left). Note, that this is the same plot as shown in **Figure 5d**.



Supplementary Figure 11.

Cellular and network differences following chronic application of gabazine.

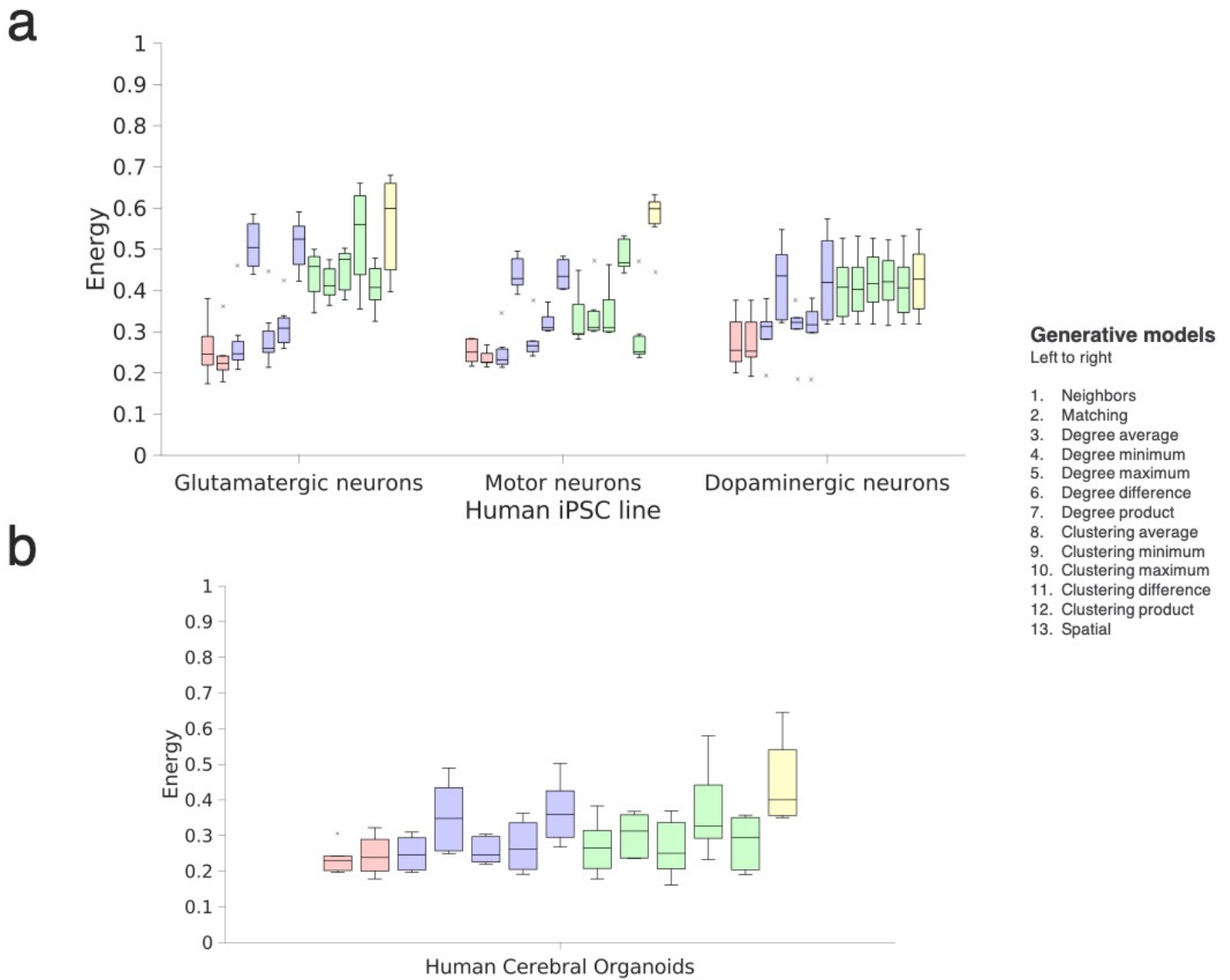
a Representative spike train raster plots from a control (left) and a gabazine-treated culture (right); the panel below shows the representative population activity vectors (activity is aggregated 1 s bins). Panels on the bottom: example close-ups of 1 s time windows taken from time bins comprising either median network activity (left, indicated by cyan triangle) or peak network activity (right, indicated by yellow triangle). **b** A logarithmic histogram of interspike intervals (ISIs) showing that gabazine relatively increases the proportion of very short ISIs and long ISIs, giving a tri-modal distribution, equivalent to periods of relative quiescence followed by periods of fast bursting. **c** Cultures treated with Gabazine showed decreased mean firing rate (Hz) of cellular activity compared to controls, but increased and more regular bursting activity (reduced coefficient of variation in interburst intervals). **d** A range of global topological metrics inferred from functional connectivity graphs of control (C) and Gabazine (G) treated PC cultures. For each comparison we provide the p-value computed from a Mann-Whitney U test.

a**b****c****d****e**

Supplementary Figure 12.

Generative model comparisons between control and gabazine-treated sparse rodent PC cultures.

a The average energy values of the top $n=1$ performing simulated networks for control (left) and gabazine-treated (right) sparse PC networks are shown for each evaluated network across 13 wiring rules used as the K_{ij} term for each recording. Boxplots represent the median and interquartile range (IQR). Outliers are demarcated as small black crosses, and are those which exceed 1.5x the IQR away from the top or bottom of the box. **b** Wiring parameters of control (left; colored in red) versus gabazine (right; colored in blue) derived from the top $n=10$ and **c** top $n=50$ performing matching simulation in terms of the wiring equation are shown. The probability distributions for **d** control and **e** gabazine networks were computed and scaled at 1% increments throughout the developmental course of the simulations. These distributions were averaged over time to form the mean probability distributions shown in **Figure 6d**.



Supplementary Figure 13.

Generative network modeling of human iPSC-derived neuronal cultures and cerebral organoids.

The energy of the top $n=1$ performing simulated networks are shown for each evaluated network across thirteen wiring rules used as the K_{ij} term for each recording. Boxplots represent the median and interquartile range (IQR). Outliers are demarcated as small black crosses, and are those which exceed 1.5x the IQR from the top or bottom of the box. **a** $n=18$ human iPSC-derived neuronal cultures ($n=6$ for each culture cell type), compared at DIV28. **b** The same analysis for the human ESC-derived cerebral organoids ($n=6$; recorded from 120-day old organoid slices).

SUPPLEMENTARY TABLES

Cell line	No. of units [total number]	Recording time point (s)	Plating density	Samples (<i>n</i>)
Primary rodent cortex (sparse)	152.8 ± 20.5 [917]	DIV 7, 10, 12, 14	50,000	6
Primary rodent cortex (dense)	139.9 ± 30.9 [1679]	DIV 14/15, 21, 28	100,000	12
Human motor neurons	210.7 ± 24.7 [1475]	DIV 28	100,000	7
Human glutamatergic neurons	191.6 ± 69.1 [1341]	DIV 28	100,000	8
Human dopaminergic neurons	211.2 ± 55.9 [1267]	DIV 28	100,000	6
Human cerebral organoids	47.3 ± 27.8 [284]	120 days	n/a	6 (slices), 3 organoids

Supplementary Table 1.

Overview of all the datasets used in the study.

Number of unit results are expressed as the mean ± SD of the respective cell line.

Name	$K_{i,j}$
Neighbors	$\sum_l A_{il}A_{jl}$
Matching	$\frac{ N_{i/j} \cap N_{j/i} }{ N_{i/j} \cup N_{j/i} }$
Clustering Average	$\frac{c_i}{2} + \frac{c_j}{2}$
Clustering Difference	$ c_i - c_j $
Clustering Maximum	$\max(c_i, c_j)$
Clustering Minimum	$\max(c_i, c_j)$
Clustering Product	$c_i c_j$
Degree Average	$\frac{k_i}{2} + \frac{k_j}{2}$
Degree Difference	$ k_i - k_j $
Degree Maximum	$\max(k_i, k_j)$
Degree Minimum	$\max(k_i, k_j)$
Degree Product	$k_i k_j$
Spatial	1

Supplementary Table 2.

List of all value $K_{i,j}$ terms that were included in the generative modeling, as given in the wiring equation.

A is the binary adjacency matrix, c is the local clustering coefficient, k is the node degree and $N_{i/j}$ represents the neighbors of node i , excluding node j . Note that the spatial model enforces $K_{i,j}=1$, which means that the value term has no effect on the generative process.

Rule A	Rule B	DIV 7		DIV10		DIV12		DIV14	
		ANOVA p=0.943	Cohen's d	ANOVA p=5.18e-3	Cohen's d	ANOVA p=1.74e-12	Cohen's d	ANOVA p=6.32e-11	Cohen's d
Homophily	Degree	n/a	n/a	0.0443	0.747	1.07e-06	1.2	0.000171	1.09
Homophily	Clustering	n/a	n/a	0.00415	0.928	3.80e-09	5.02	1.81e-08	3.07
Homophily	Spatial	n/a	n/a	0.0341	0.779	4.01e-09	4.56	4.52e-09	12.1
Degree	Clustering	n/a	n/a	0.688	0.248	0.0102	0.637	0.0144	0.526
Degree	Spatial	n/a	n/a	0.713	0.258	0.000351	2.32	2.59e-05	6.29
Clustering	Spatial	n/a	n/a	0.974	0.393	0.0871	1.03	0.0117	1.72

Supplementary Table 3.

Statistical comparisons of rodent 50k neuronal culture energy comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV14		DIV21		DIV28	
		ANOVA p=1.96e-21	Cohen's d	ANOVA p=3.94e-23	Cohen's d	ANOVA p=2.26e-26	Cohen's d
Homophily	Degree	3.77e-09	1.59	3.77e-09	1.53	3.77e-09	1.55
Homophily	Clustering	3.77e-09	2.7	3.77e-09	1.75	3.77e-09	1.72
Homophily	Spatial	3.77e-09	2.74	3.77e-09	6.12	3.77e-09	7.97
Degree	Clustering	0.00241	0.656	0.632	0.175	0.742	0.14
Degree	Spatial	0.0008	1	3.77e-09	3.64	3.77e-09	3.16
Clustering	Spatial	0.285	0.602	4.38e-09	2.16	3.77e-09	2.06

Supplementary Table 4.

Statistical comparisons of rodent 100k neuronal culture energy comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV7		DIV10		DIV12		DIV14	
		ANOVA p=0.644	Cohen's <i>d</i>	ANOVA p=0.0138	Cohen's <i>d</i>	ANOVA p=1.06e-08	Cohen's <i>d</i>	ANOVA p=2.93e-11	Cohen's <i>d</i>
Homophily	Degree	n/a	n/a	0.179	1.13	0.000681	1.25	0.000348	1.55
Homophily	Clustering	n/a	n/a	0.00931	0.879	7.44e-09	1.9	3.80e-09	2.33
Homophily	Spatial	n/a	n/a	0.846	0.616	0.0106	2.43	0.157	0.999
Degree	Clustering	n/a	n/a	0.459	0.453	0.000758	0.798	1.03e-05	1.02
Degree	Spatial	n/a	n/a	0.9	-0.642	0.963	0.365	0.815	-0.552
Clustering	Spatial	n/a	n/a	0.391	-0.675	0.26	-0.727	0.00127	-1.37

Supplementary Table 5.

Statistical comparisons of rodent 50k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV14		DIV21		DIV28	
		ANOVA p=8.05e-12	Cohen's <i>d</i>	ANOVA p=2.56e-23	Cohen's <i>d</i>	ANOVA p=1.23e-23	Cohen's <i>d</i>
Homophily	Degree	0.389	0.364	1.75e-05	1	4.56e-05	1.23
Homophily	Clustering	4.17e-09	1.33	3.77e-09	1.88	3.77e-09	1.93
Homophily	Spatial	0.906	0.234	0.000627	1.11	0.0201	0.855
Degree	Clustering	4.14e-09	0.96	3.77e-09	1.33	3.77e-09	1.27
Degree	Spatial	0.969	-0.193	0.873	0.227	0.998	-0.0489
Clustering	Spatial	0.000139	-1.26	5.69e-05	-0.986	2.95e-07	-1.56

Supplementary Table 6.

Statistical comparisons of rodent 100k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	Glutamatergic neurons		Motor neurons		Dopaminergic neurons	
		ANOVA p=9.67e-12	Cohen's d	ANOVA p=4.81e-12	Cohen's d	ANOVA p=1.35e-05	Cohen's d
Homophily	Degree	0.000116	0.881	0.000948	0.821	0.0219	0.642
Homophily	Clustering	6.15e-09	2.67	0.000202	1.6	1.59e-05	1.19
Homophily	Spatial	4.12e-09	2.32	3.77e-09	6.88	0.00173	1.14
Degree	Clustering	0.00832	0.42	0.941	0.0859	0.03	0.715
Degree	Spatial	3.59e-05	1.05	1.07e-08	2.13	0.204	0.854
Clustering	Spatial	0.0235	1.18	3.43e-08	1.89	0.985	0.51

Supplementary Table 7.

Statistical comparisons of DIV28 human iPSC neuronal culture (glutamatergic neurons, motor neurons and dopaminergic neurons) energy comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	ANOVA p=6.88e-05	Cohen's d
Homophily	Degree	0.192	0.986
Homophily	Clustering	0.159	0.858
Homophily	Spatial	2.17e-05	2.47
Degree	Clustering	0.999	0.0354
Degree	Spatial	0.000695	2.17
Clustering	Spatial	0.000897	1.85

Supplementary Table 8.

Statistical comparisons of human cerebral organoid energy comparisons across generative rules.

For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's *d* if the ANOVA was significant at $p < 0.05$. A positive Cohen's *d* reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Antibodies (human organoids)	Type	Dilution	Catalog number
Tau	Primary	1:500	#MN1000, ThermoFisher
NeuN	Primary	1:300	#M11954-3, Boster Bio, Pleasanton, CA, USA
GFAP	Primary	1:500	#NB300-141, Novus Biologicals, Englewood, CO, USA
goat anti-mouse IgG, Alexa Fluor Plus 488	Secondary	1:400	#A32723, ThermoFisher
goat anti-rabbit IgG, Alexa Fluor 568	Secondary	1:400	#A11036, ThermoFisher
goat anti-chicken IgY, Alexa Fluor Plus 647	Secondary	1:400	#A32933, ThermoFisher
Antibodies (human iPSC)	Type	Dilution	Catalog number
mouse anti-TH	Primary	1:500	#MAB318, Sigma-Aldrich
chicken anti-MAP2	Primary	1:1000	#CH22103, Neuromics (Edina, MN, USA)
rabbit anti-GFAP	Primary	1:500	#Z0334, Agilent (Santa Clara, CA, USA)
donkey anti-mouse 488	Secondary	1:250	#A-21202, ThermoFisher
goat anti-chicken 647	Secondary	1:500	#A-32933, ThermoFisher
donkey anti-rabbit 568	Secondary	1:250	#A-A10042, ThermoFisher
Antibodies (rodent PC)	Type	Dilution	Catalog number
mouse anti-Synaptophysin	Primary	1:100	#ab8049, Abcam
rabbit anti-NeuN	Primary	1:300	#ab177487, Abcam
chicken anti-beta III Tubulin	Primary	1:1000	#ab41489 Abcam
donkey anti-mouse IgG, Alexa Fluor 488	Secondary	1:500	#ab150105, Abcam
donkey anti-Chicken IgY (IgG)	Secondary	1:400	#703-165-155, Jackson ImmunoResearch, West Grove, USA
donkey anti-rabbit IgG, Alexa Fluor 405	Secondary	1:1000	#ab175651, Abcam

Supplementary Table 9. Primary and secondary antibodies.