First connectomics challenge in neuronal cultures Crowdsourcing network reconstruction

Javier G. Orlandi
Dept. ECM, University of Barcelona

Universitat de Barcelona

Connectome Workbench 2015

Outline

Neuronal cultures

Model system to study the interplay between activity and connectivity in neuronal systems

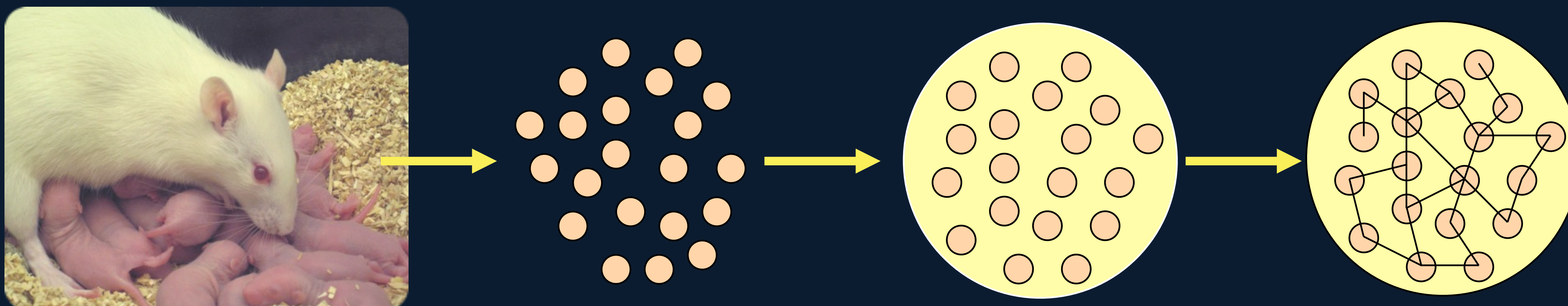
Mostly random network structure - dynamics leads to complex effective networks

You can easily manipulate the topology, for example by modifying the growth substrate (spatial embedding)

Develop and test network inference methods (causality) in neuroscience

First connectomics challenge in neuronal cultures

Neuronal cultures



cortex / hippocampus
17-19 days embryo or
0-1 days postnatal

neuron dissociation

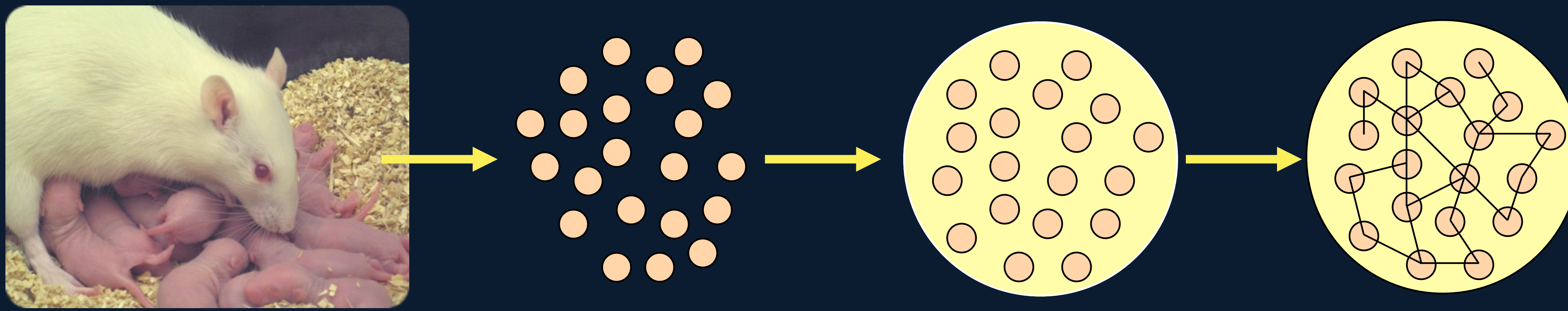
plating
(day 0)

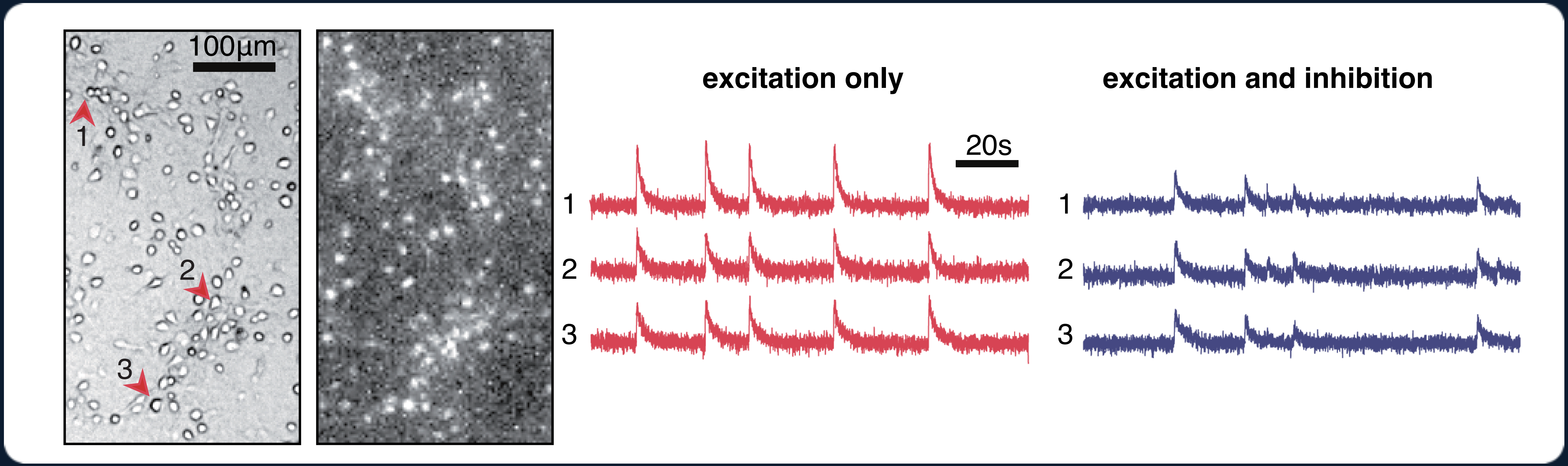
mature culture
(day 14)

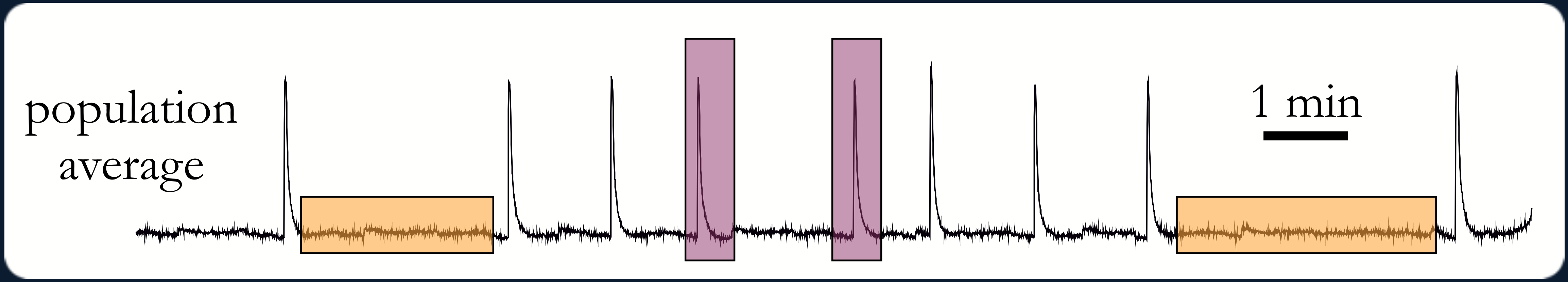
individual
firings

collective
firings

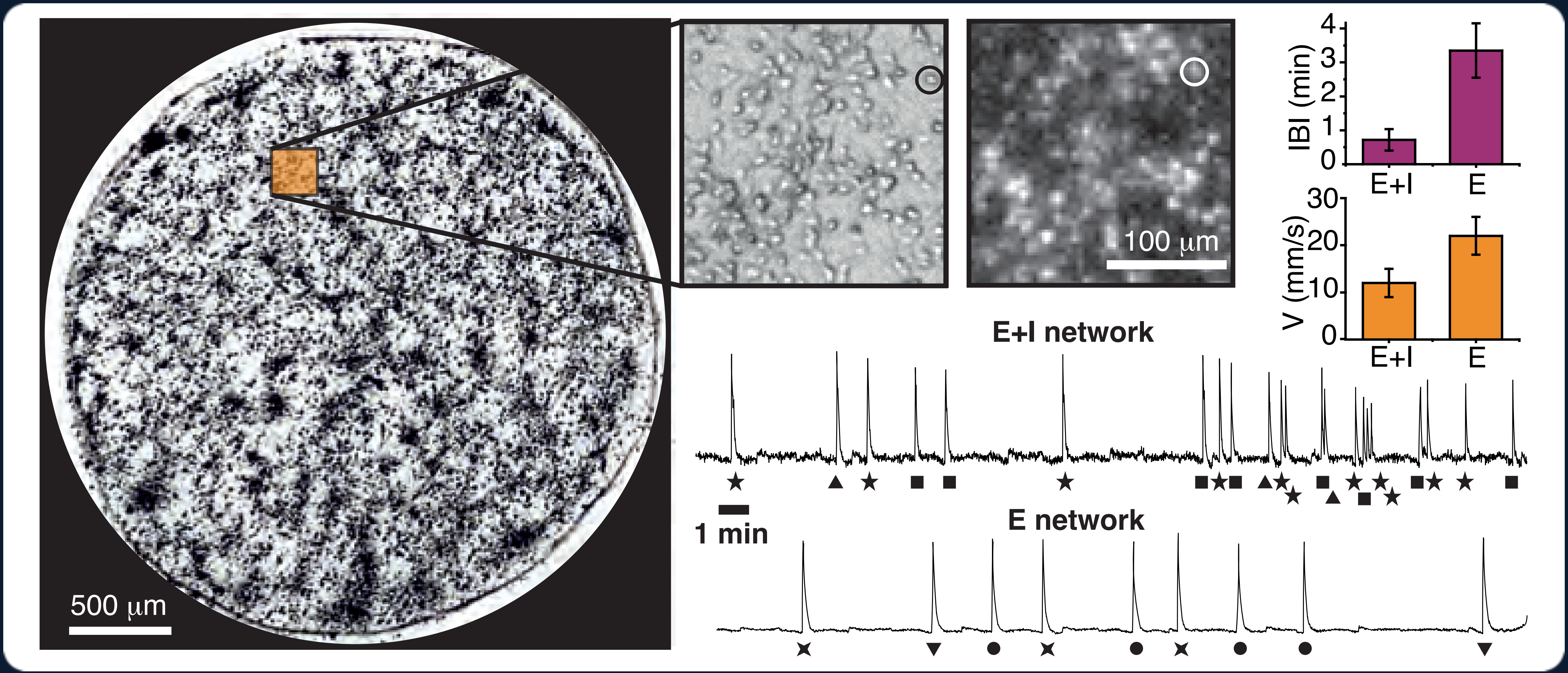
Neuronal cultures





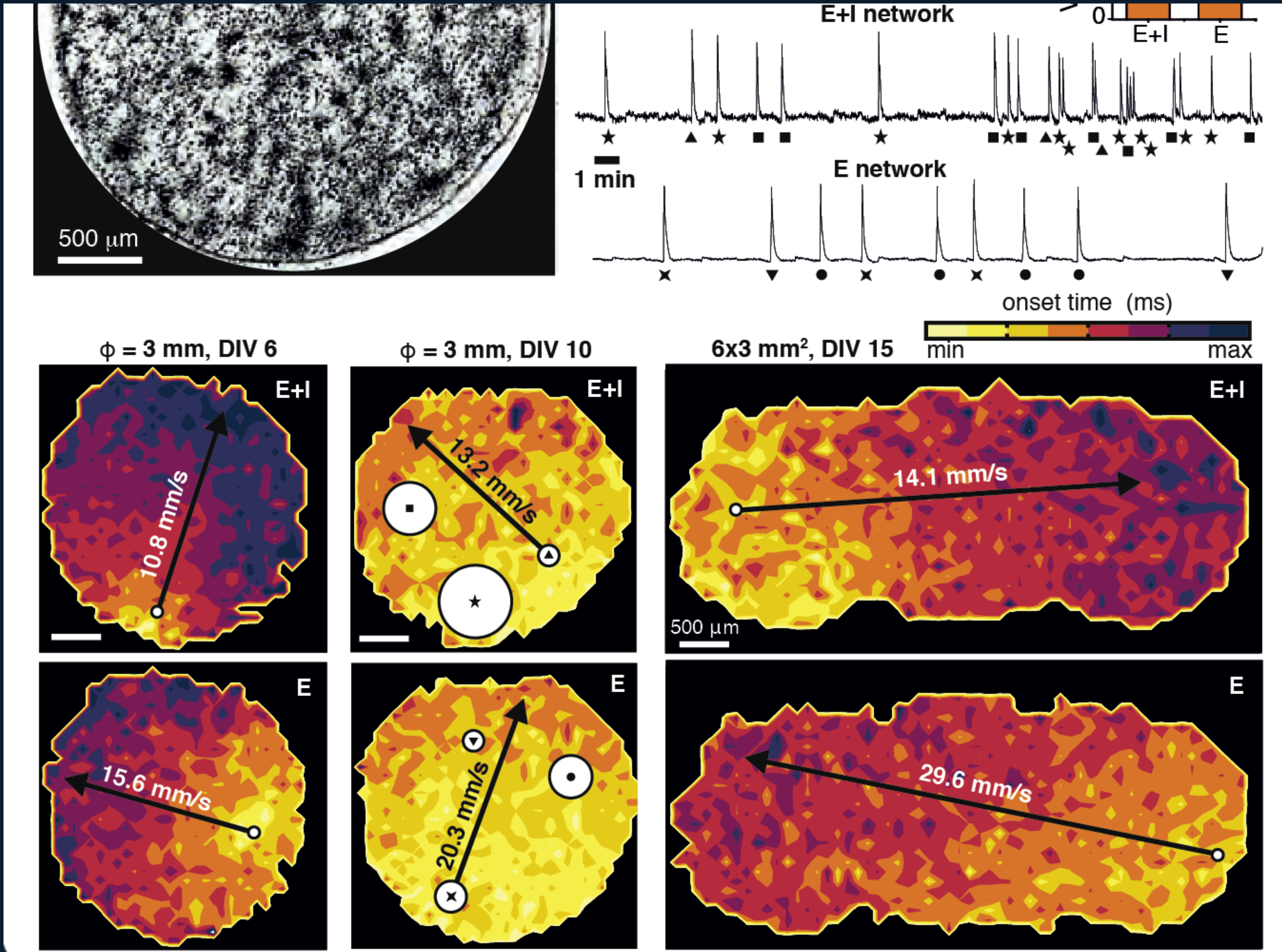


Dynamics in young cultures

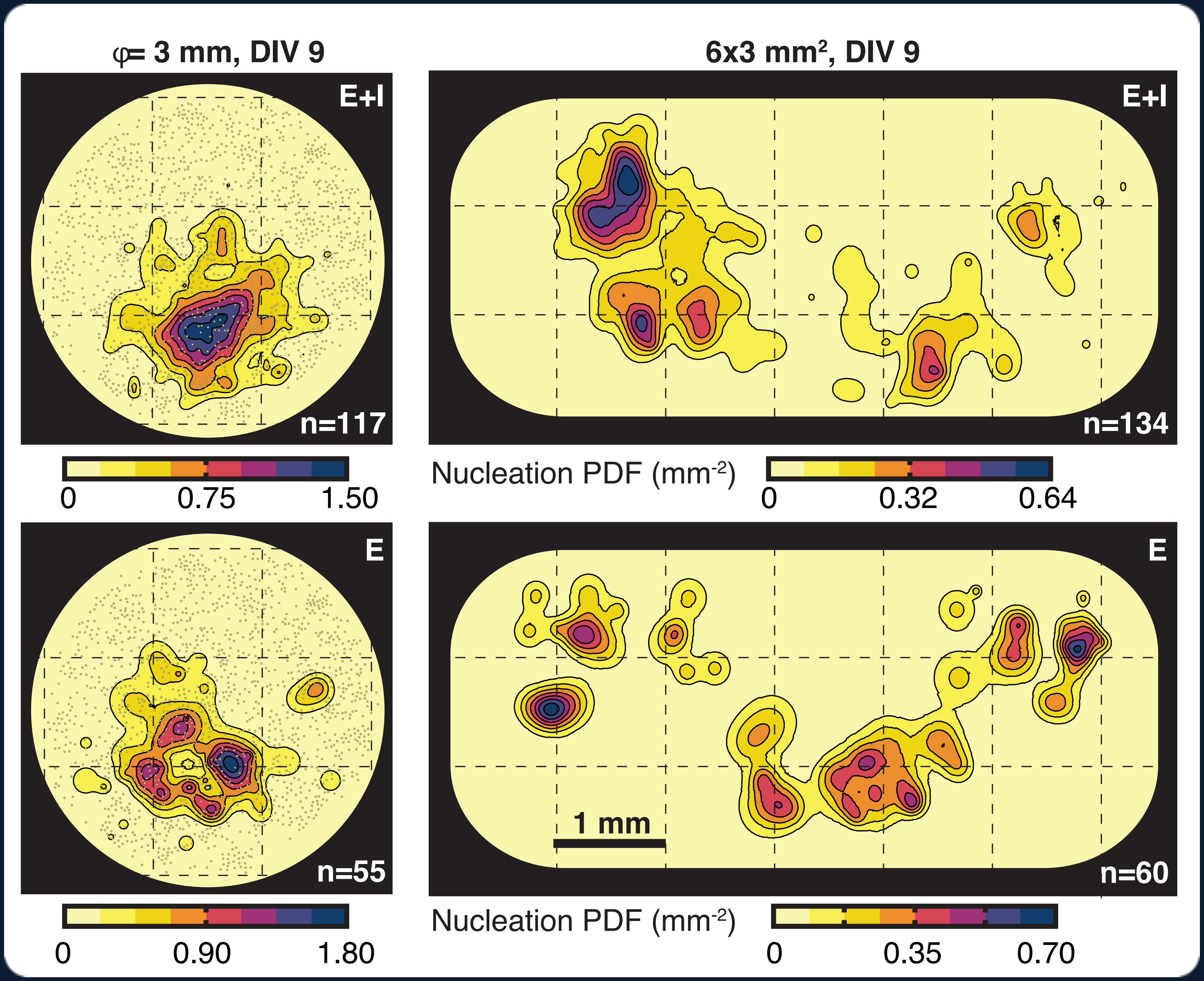


Each burst is in fact a propagating front of activity. And can only originate in specific locations of the culture (nucleation sites)

Dynamics in young cultures



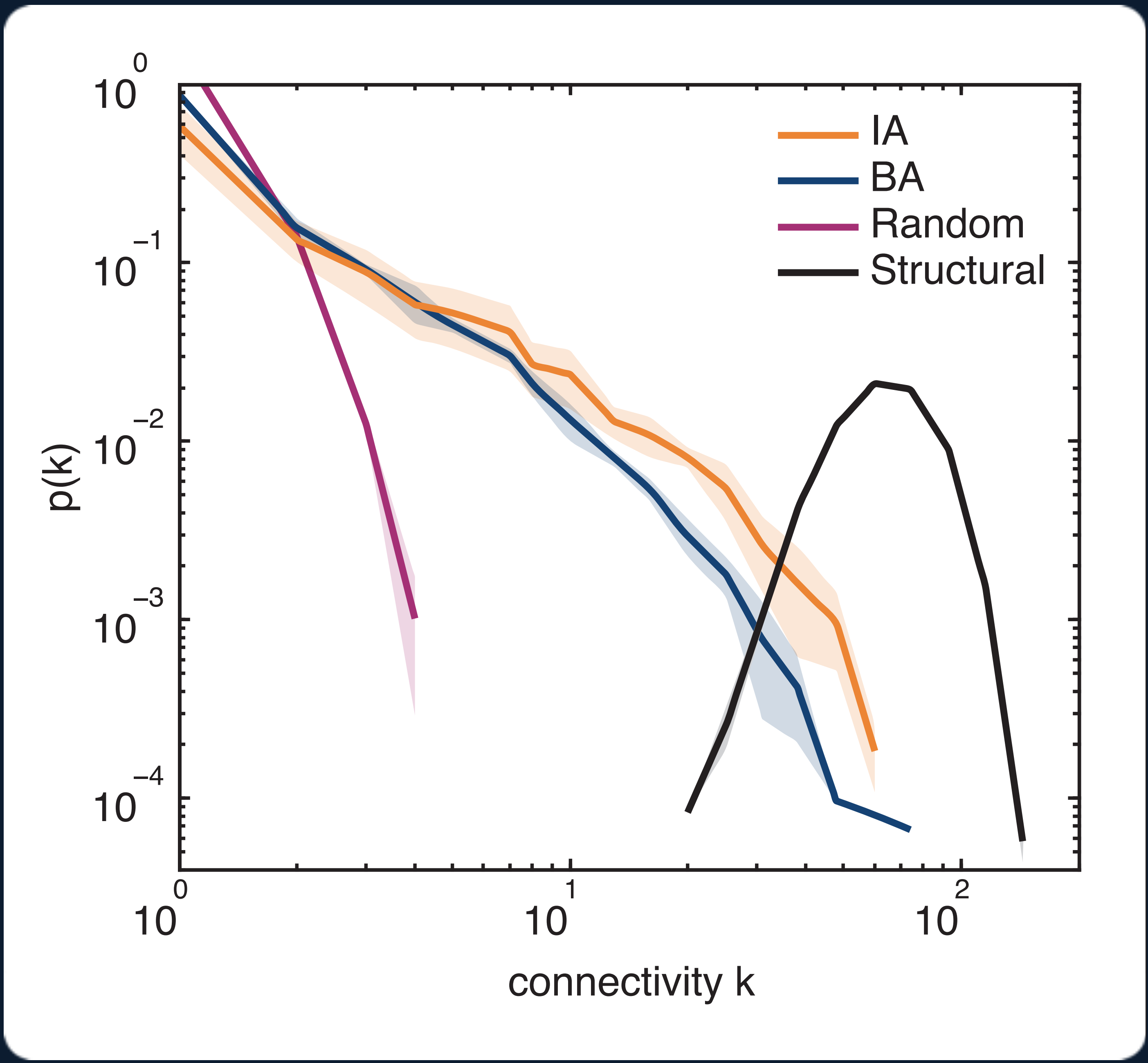
Distribution of nucleation sites



Bursts only originate in specific regions

Only a part of the network is capable of burst generation

Underlying network (through modeling)



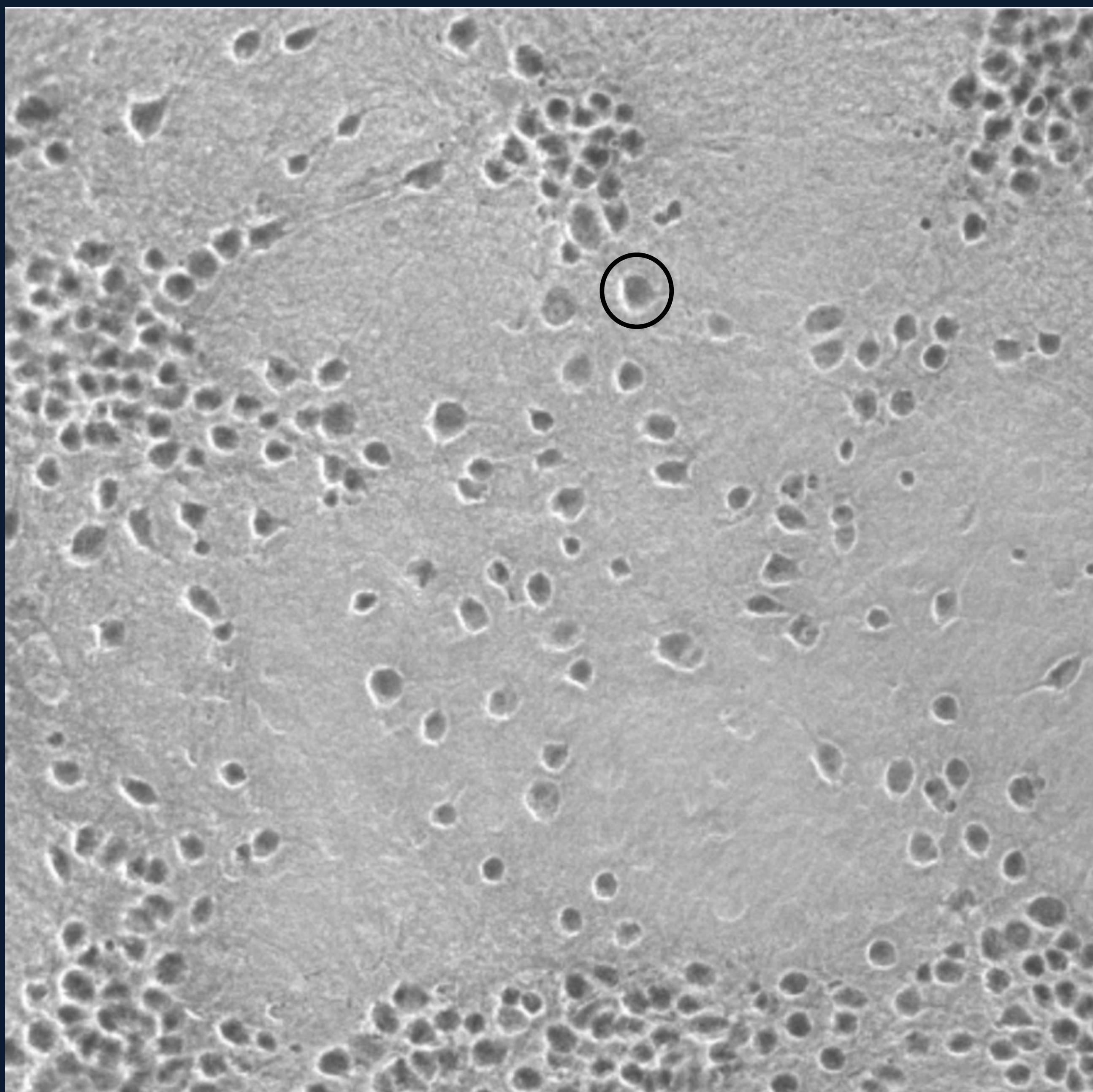
Orlandi et al, Nature Physics **2013**

Characterize the structure of the effective network responsible of burst initiation

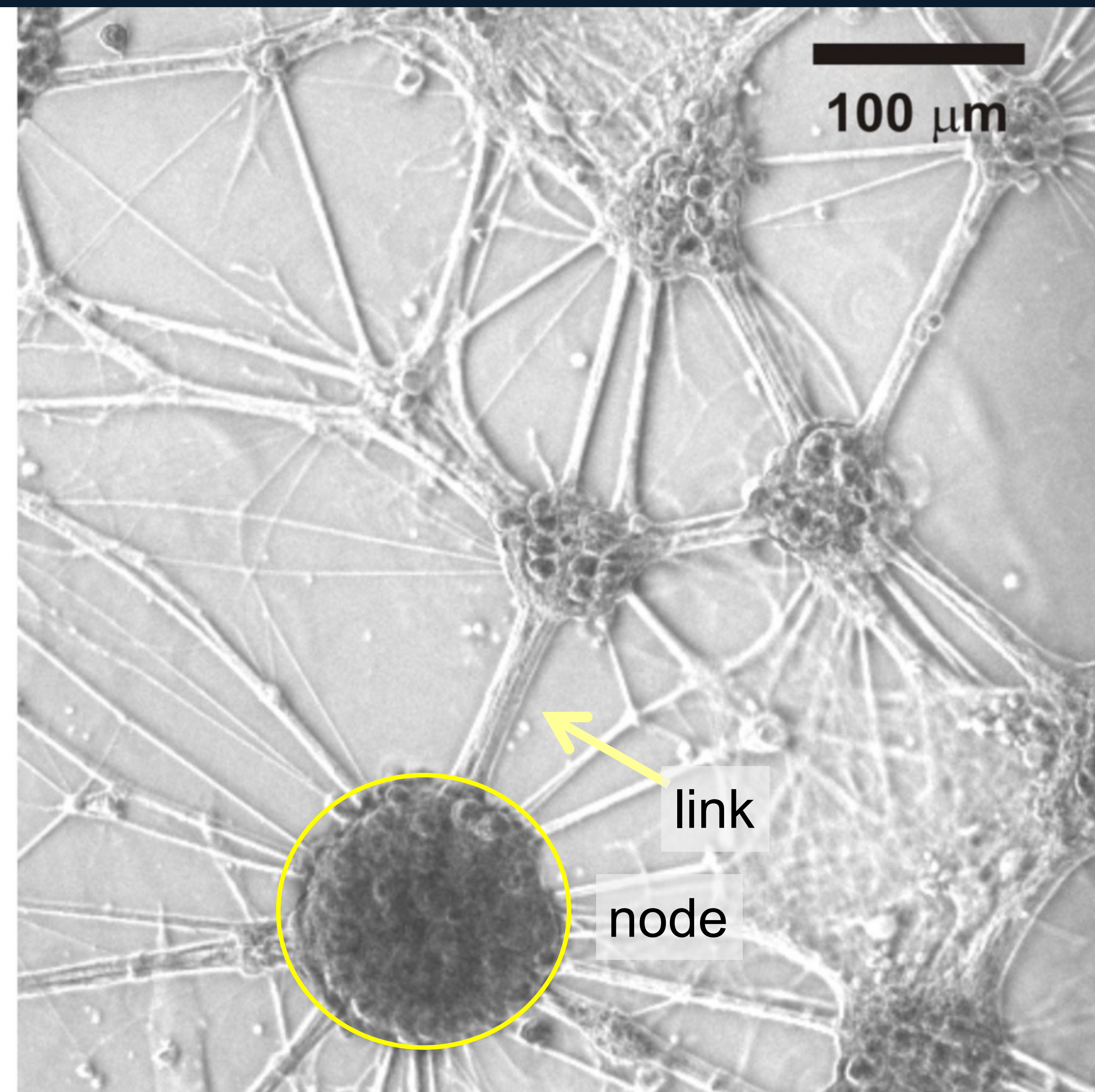
'trivial' structural connectivity gives rise to effective networks with broad-degree distributions

Clustered cultures

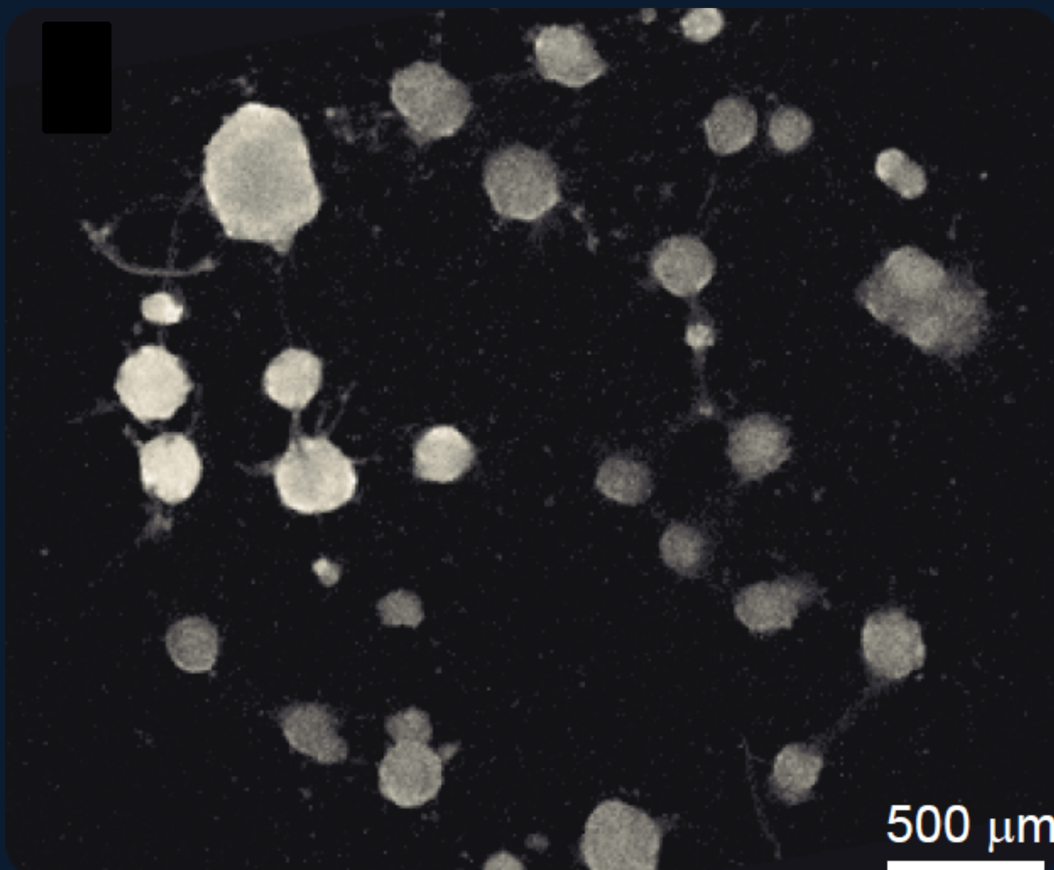
Homogeneous (coated substrate)

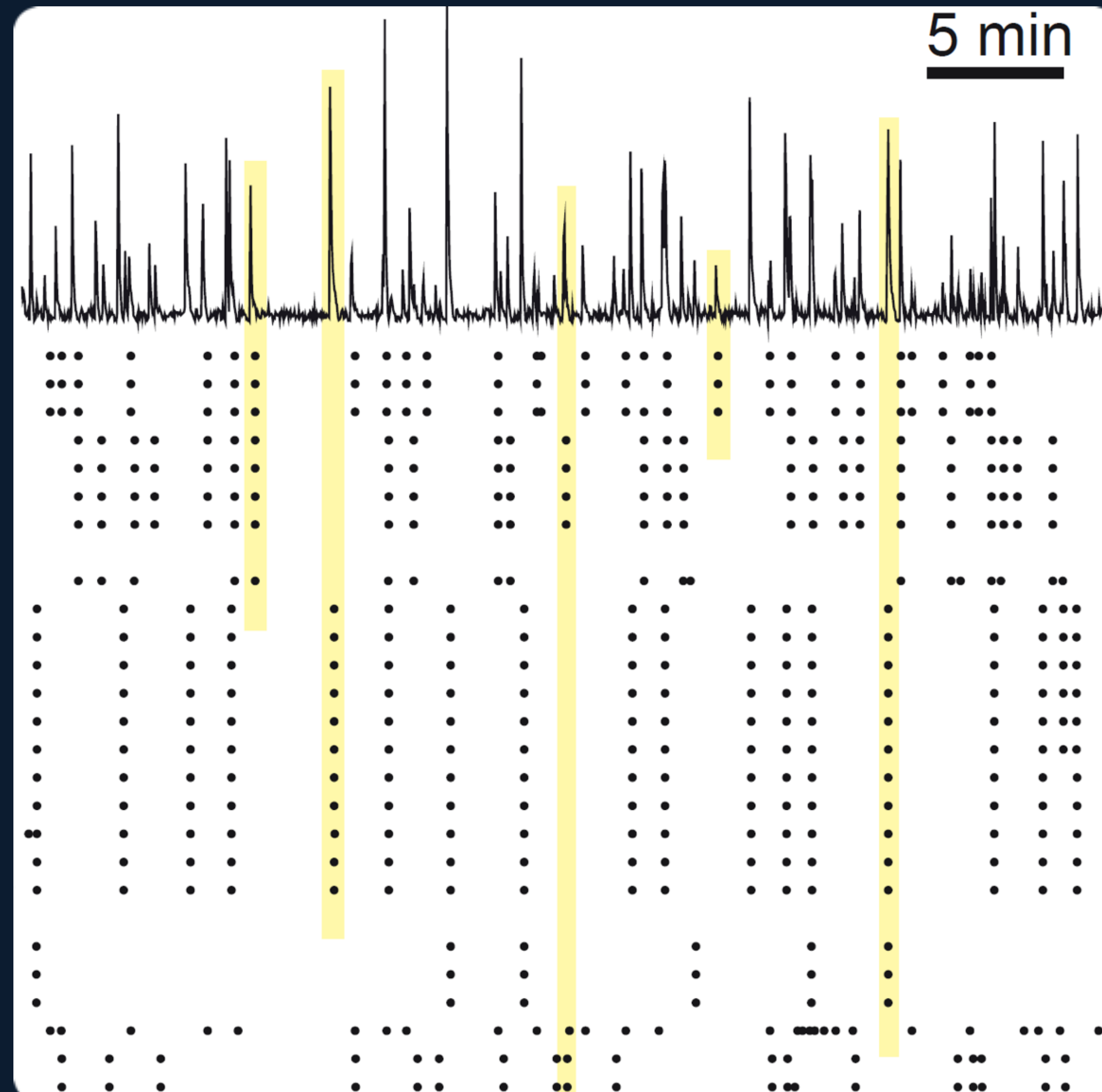


Clustered (uncoated)



Clustered cultures

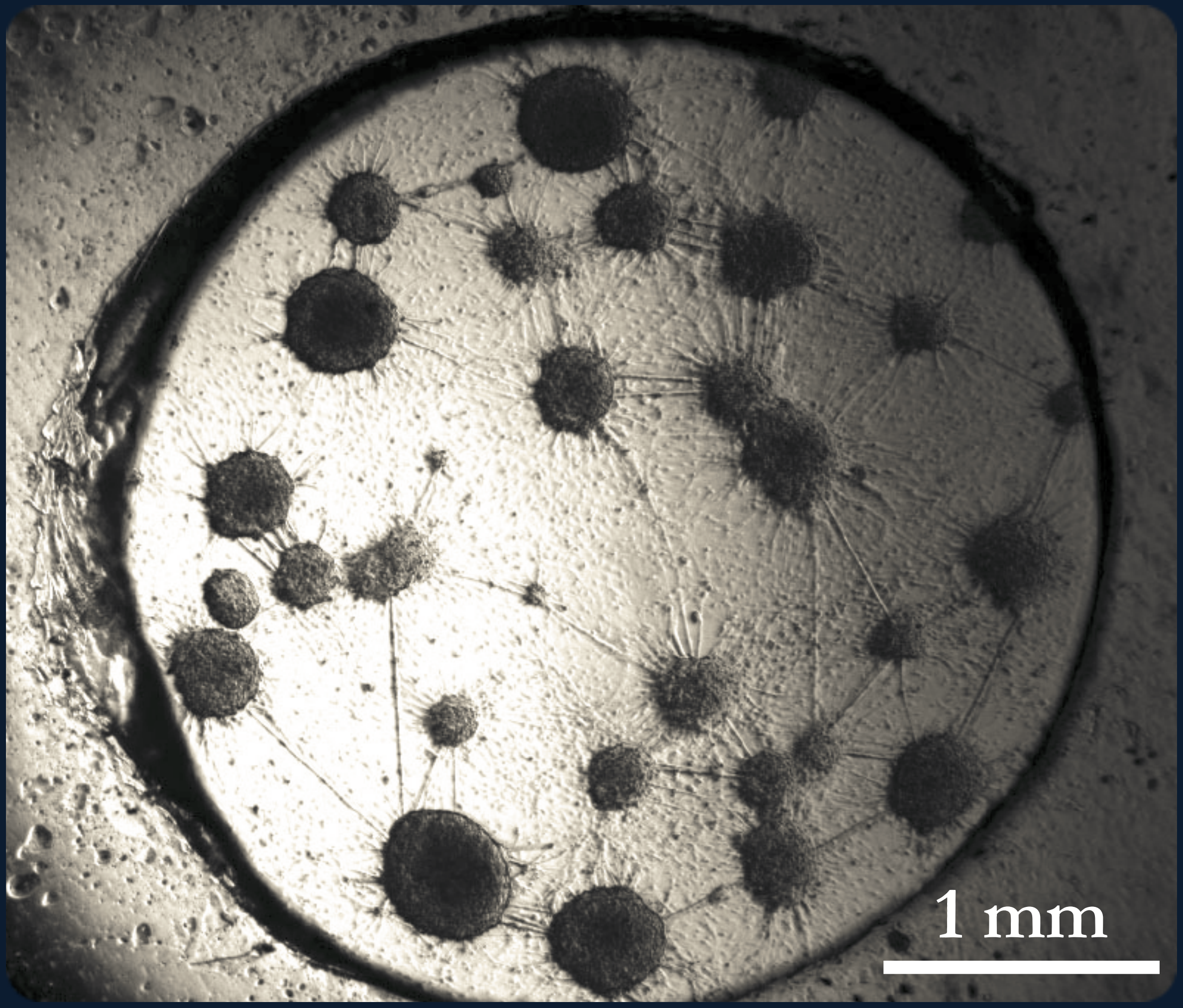




Each cluster acts as a single unit

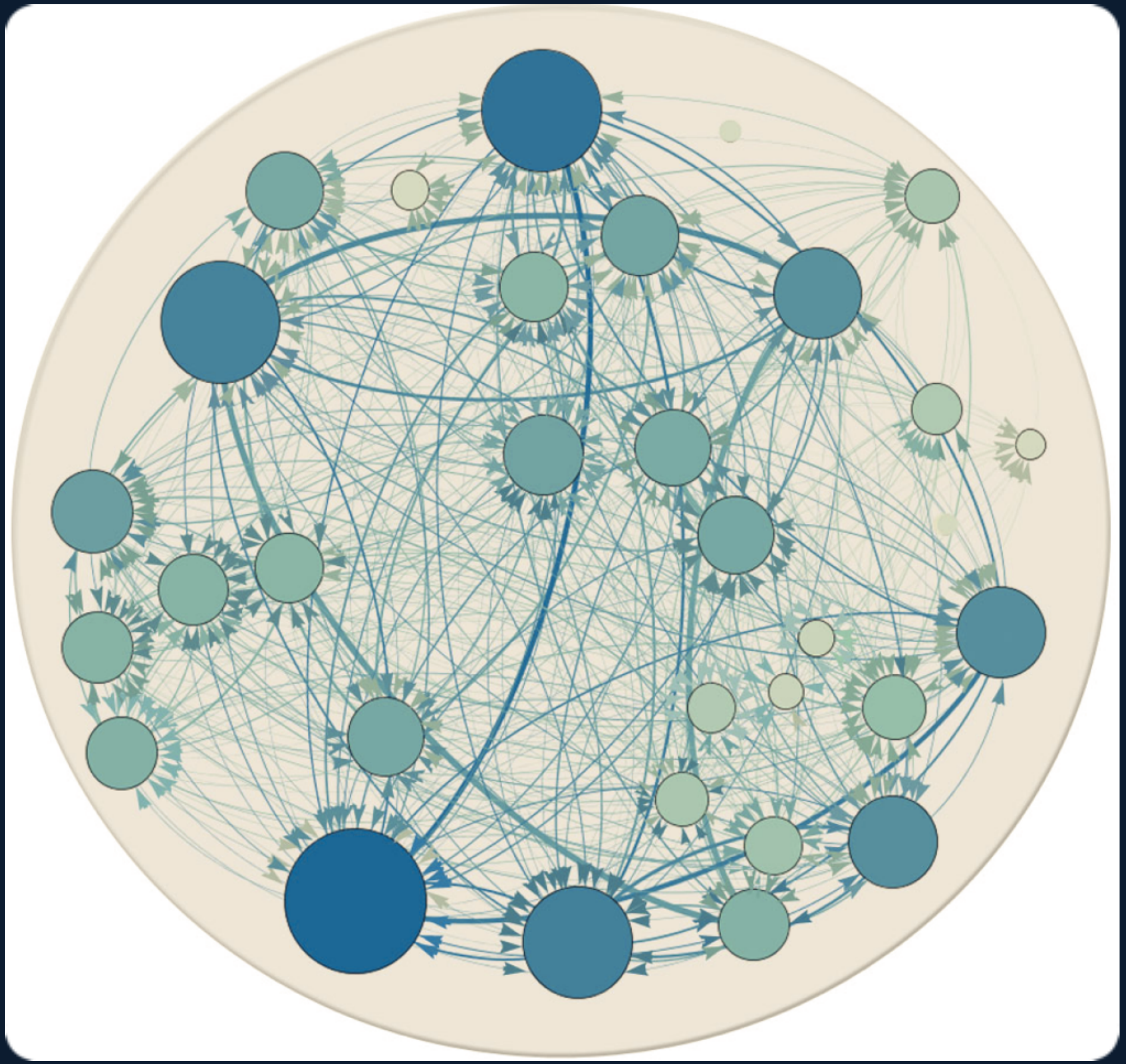
Modular firing patterns

Clustered cultures



'structural' network

effective network
(from causal inference)

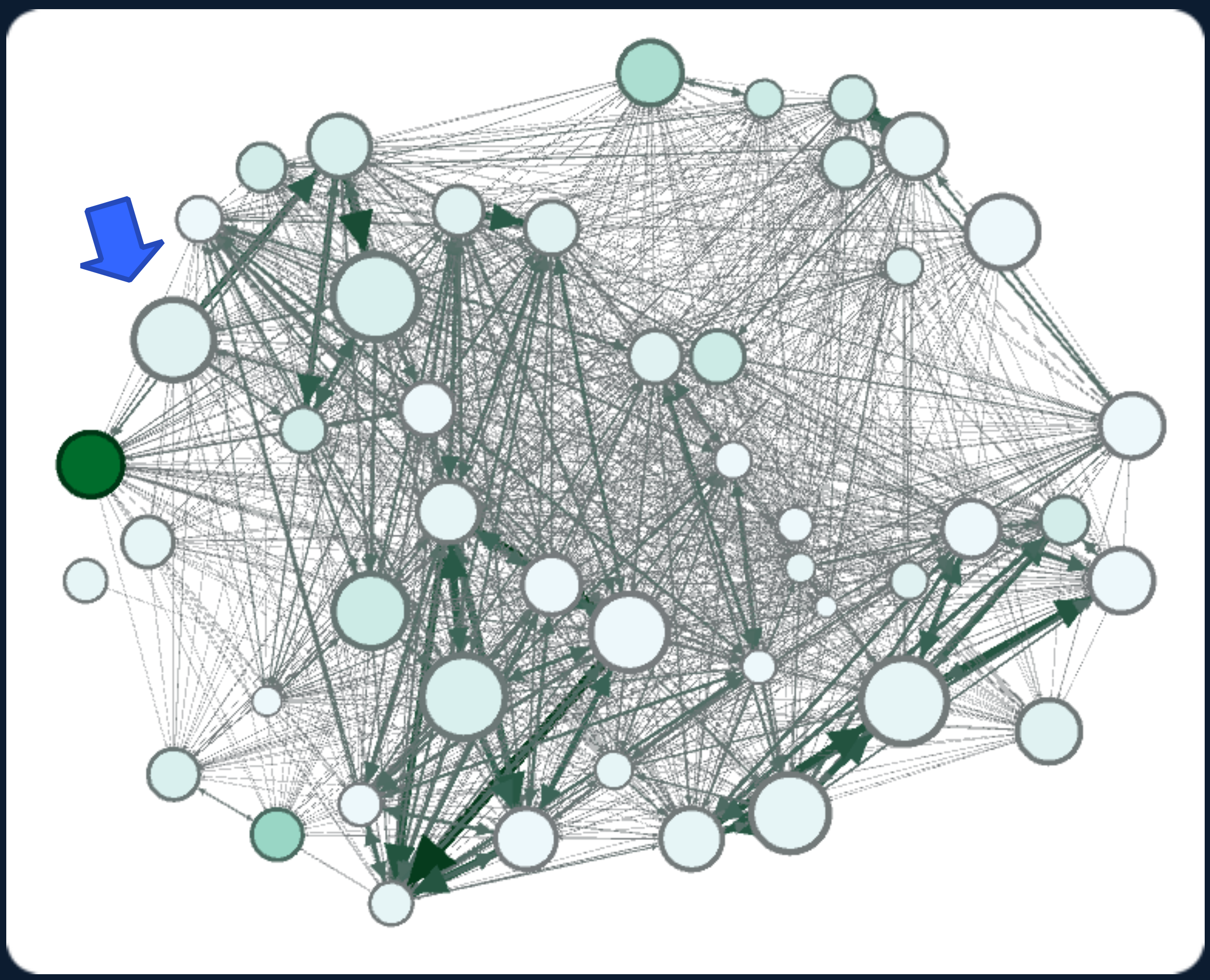


Assortative mixing
Rich club structure

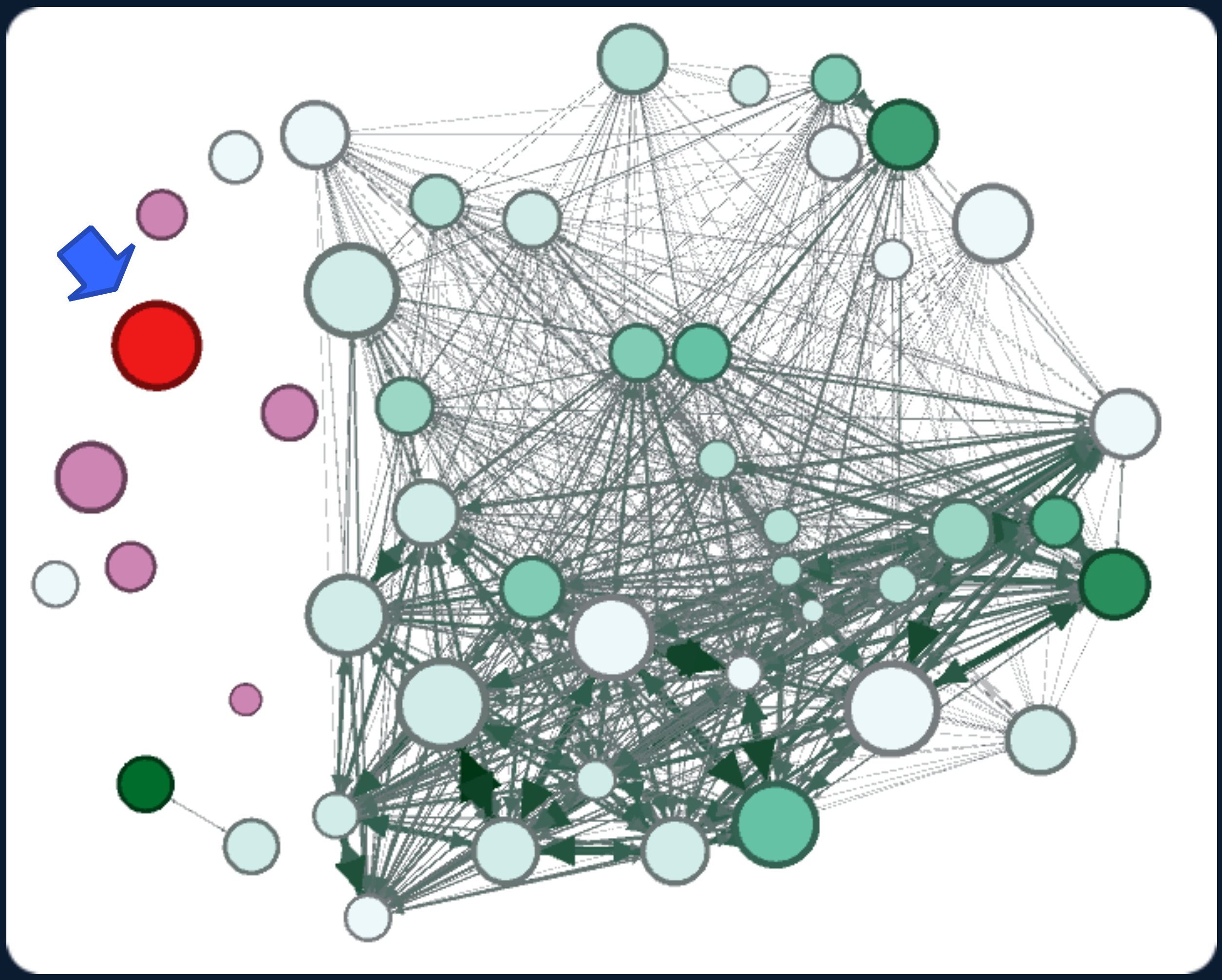
Teller et al, PLOS Comput Biol 2014

Clustered cultures

before



after



Evolution of the effective network after node removal (laser ablation)

Network inference

Indirect access to the network structure (from the dynamics)

Identify causal interactions in time series

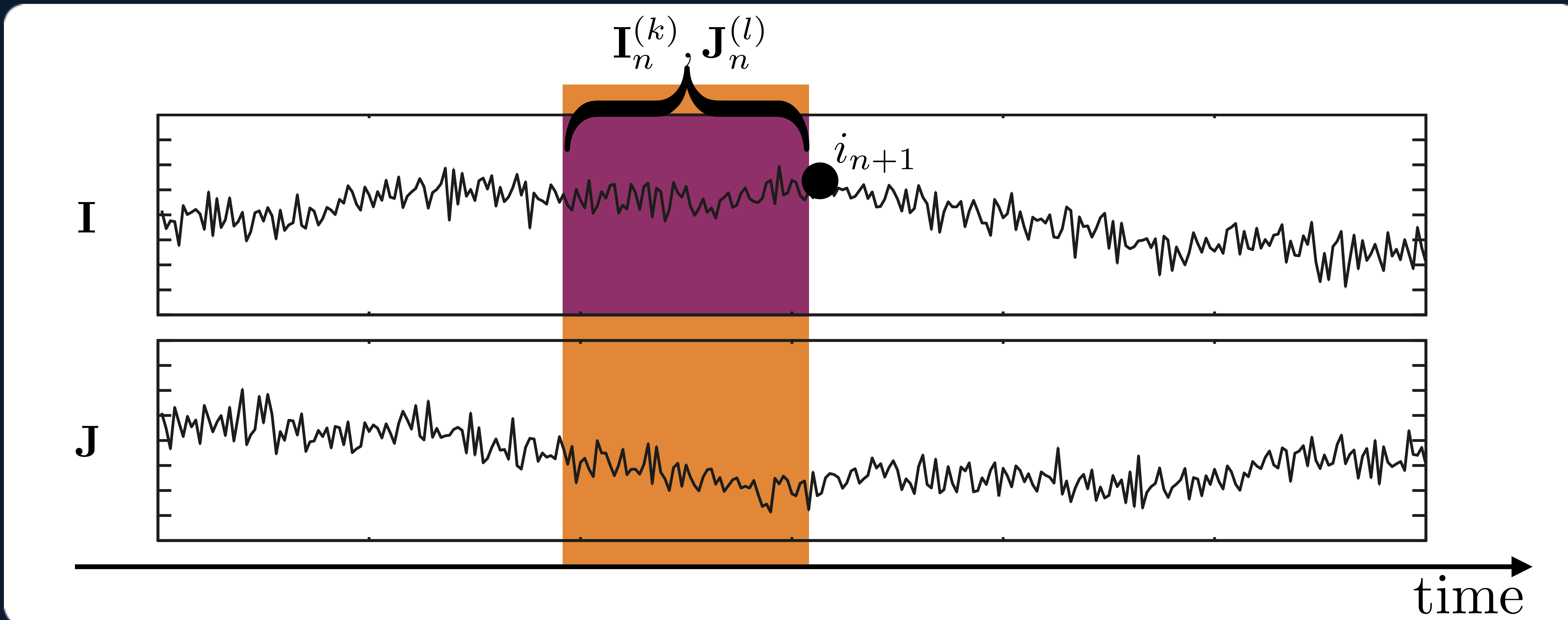
- Correlation
- Mutual information
- Granger causality
- Transfer entropy

Given sets of interdependent variables X and Y, it is said that "Y G-causes X" if, in an appropiate statistical sense, Y assists in predicting the future of X beyond the degree to which X already predicts its own future.

Barnett, PRL 2009

$$Y \longrightarrow X$$

Transfer Entropy



Schreiber, PRL 2000

$$T_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log \frac{p(i_{n+1}|i_n^{(k)}, j_n^{(l)})}{p(i_{n+1}|i_n^{(k)})}$$

Generalized Transfer Entropy

$$\mathrm{GTE}_{J\rightarrow I} = \sum P(i_n, i_{n-1}^{(k)}, j_{n-1+S}^{(k)} \,|\, g_n < \tilde{g}) \log \frac{P(i_n \,|\, i_{n-1}^{(k)}, j_{n-1+S}^{(k)}, g_n < \tilde{g})}{P(i_n \,|\, i_{n-1}^{(k)}, g_n < \tilde{g})}$$

Developed to deal with the singularities of culture dynamics

Fluorescence imaging

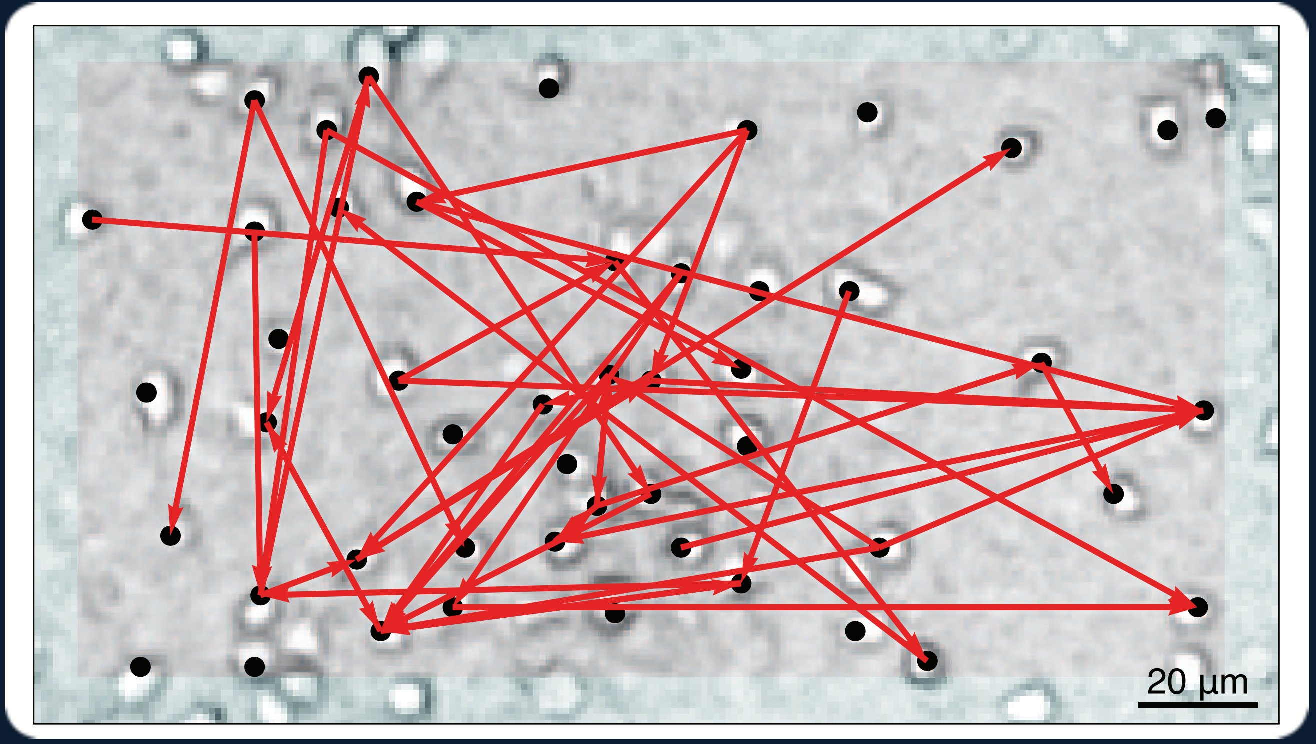
Bursting dynamics

Low acquisition speeds

Low signal to noise ratio

Light scattering

Purely excitatory networks

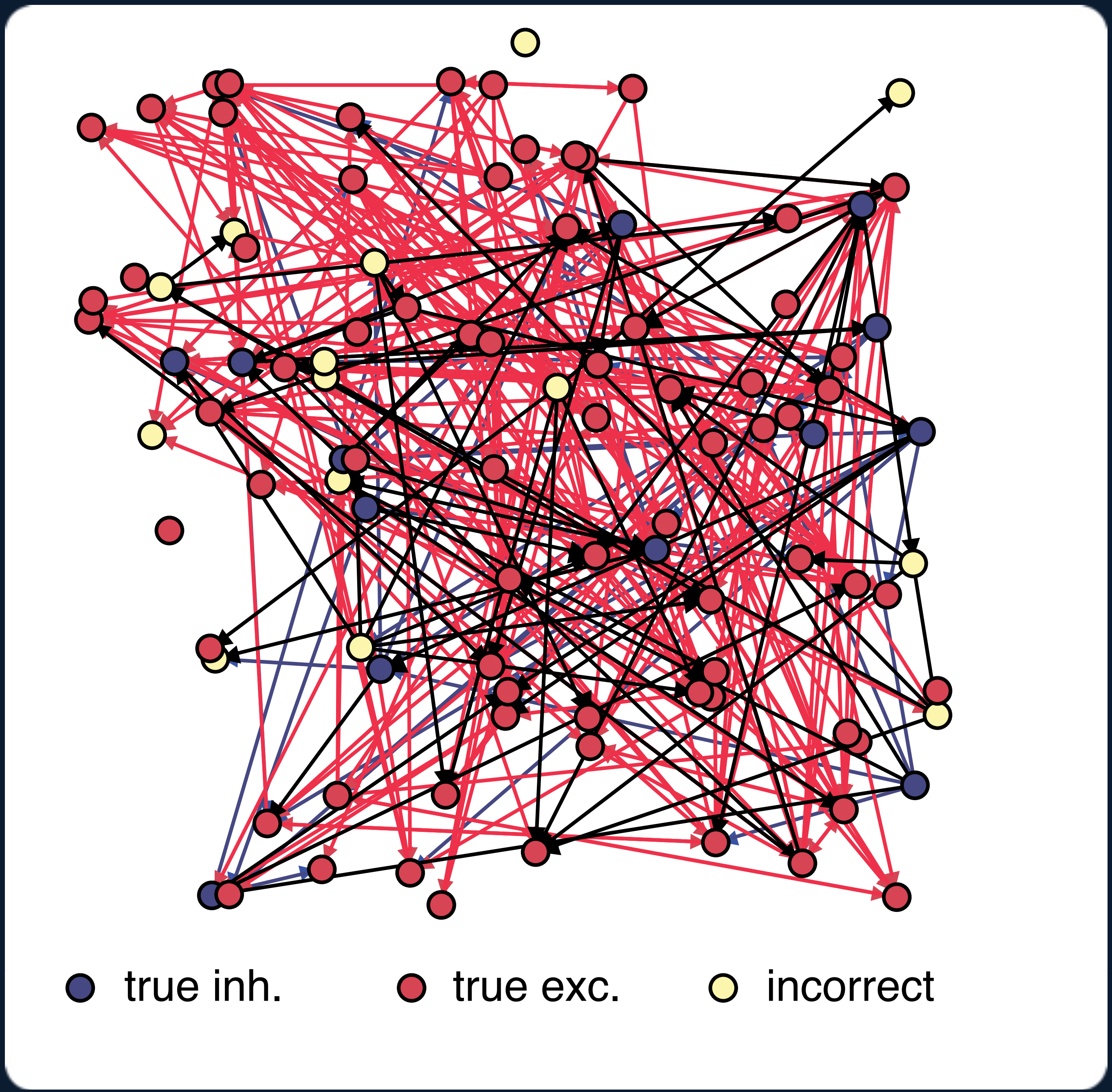


Stetter et al, PLOS Comput Biol **2012**

From the causal interactions, obtain the underlying network

Generalized Transfer Entropy

Extended to deal with inhibition



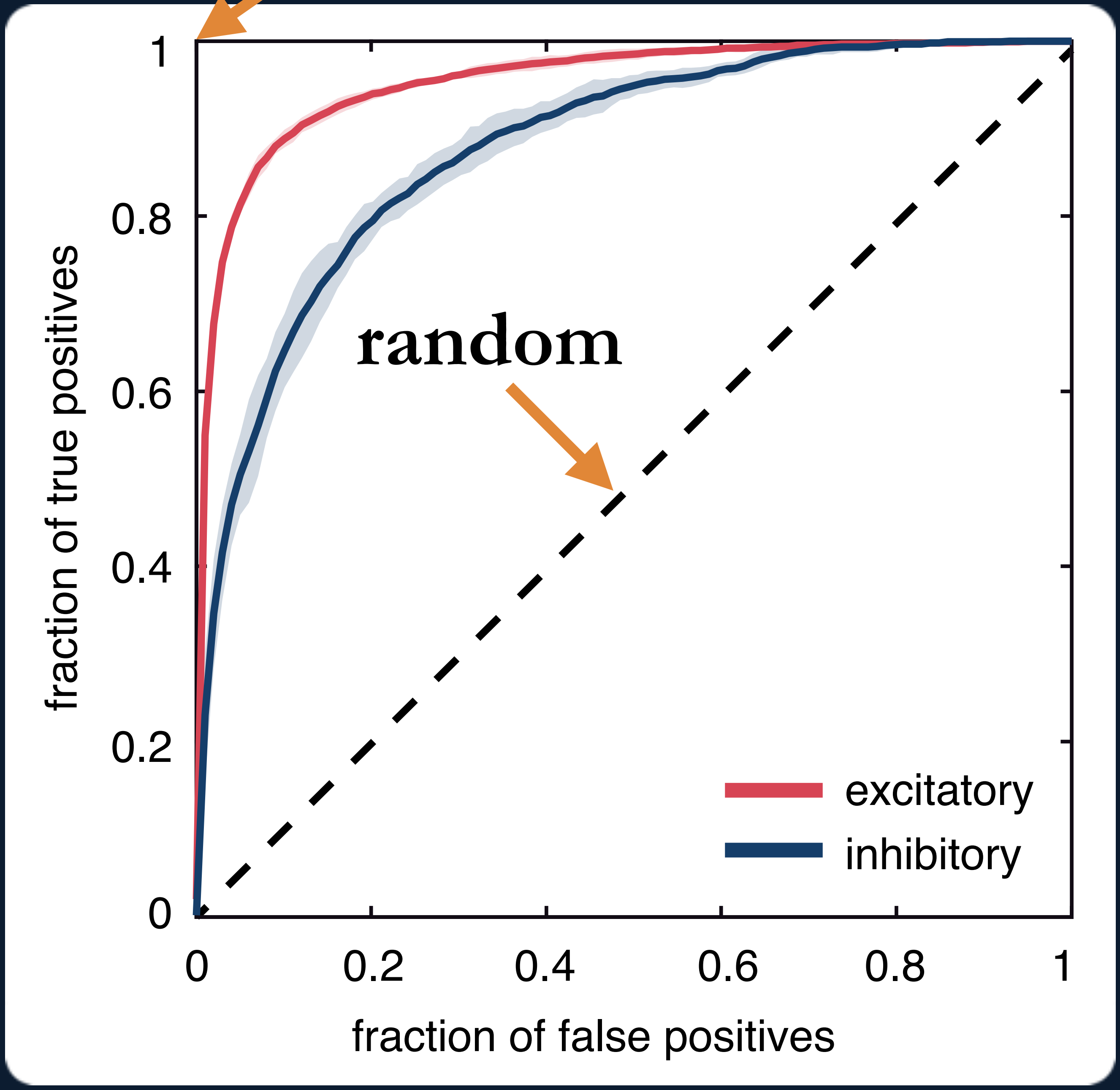
Also possible to infer neuronal type (excitatory/inhibitory)

Orlandi et al, PLOS ONE **2014**

Generalized Transfer Entropy

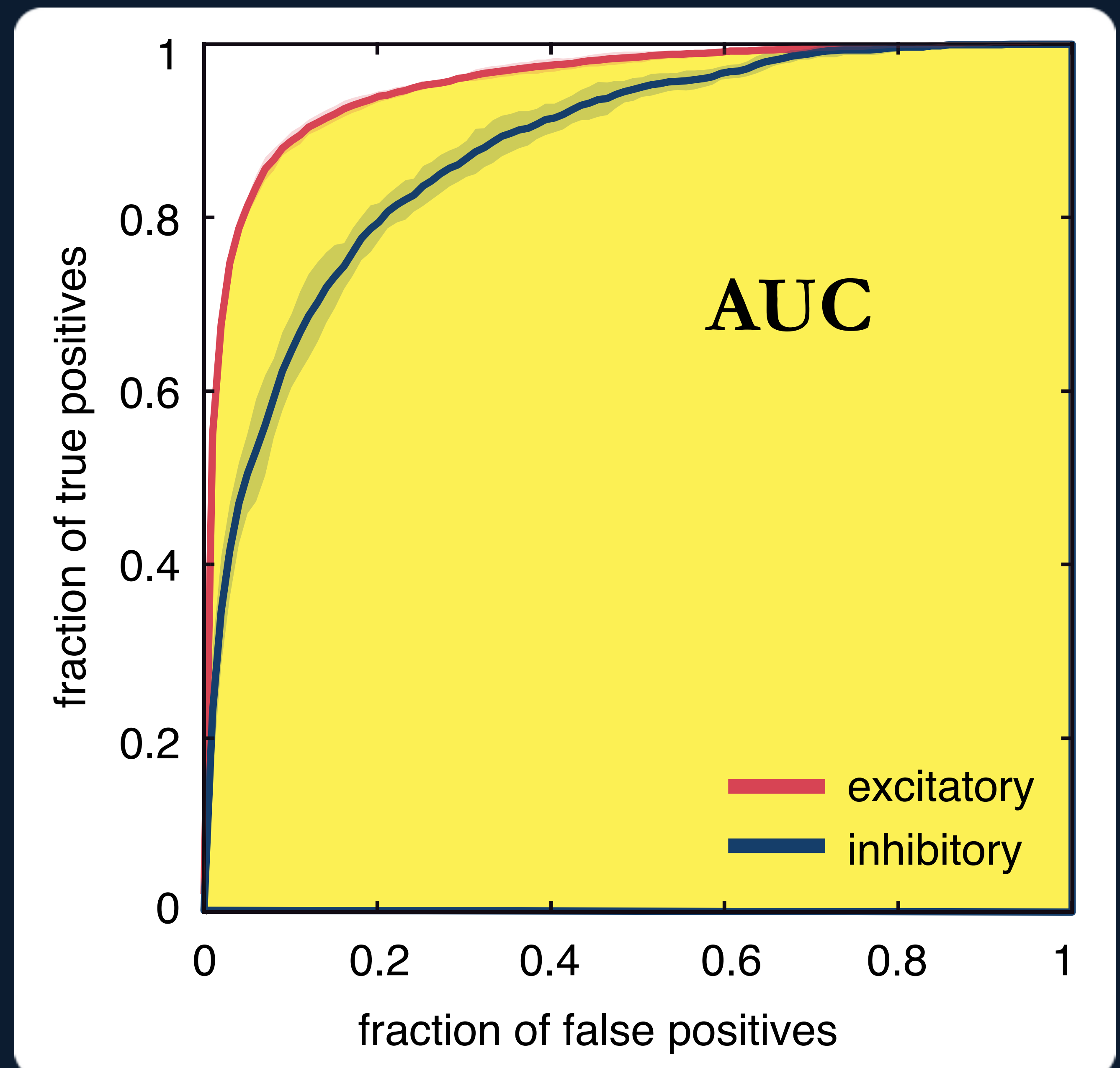
How good are the reconstructions?

perfect



Generalized Transfer Entropy

How good are the reconstructions?



AUC ~ 0.89

Still room for improvement!
A lot actually…

Also, do not trust any single method (for causal inference)

Network inference

Spend many years, (or many PhDs) improving the techniques

Or rather outsource it

Challenge

Challenges in Machine Learning (collaborative competitions)

Challenges to stimulate research across different fields

Launch the first (of a series) of online challenges in connectomics

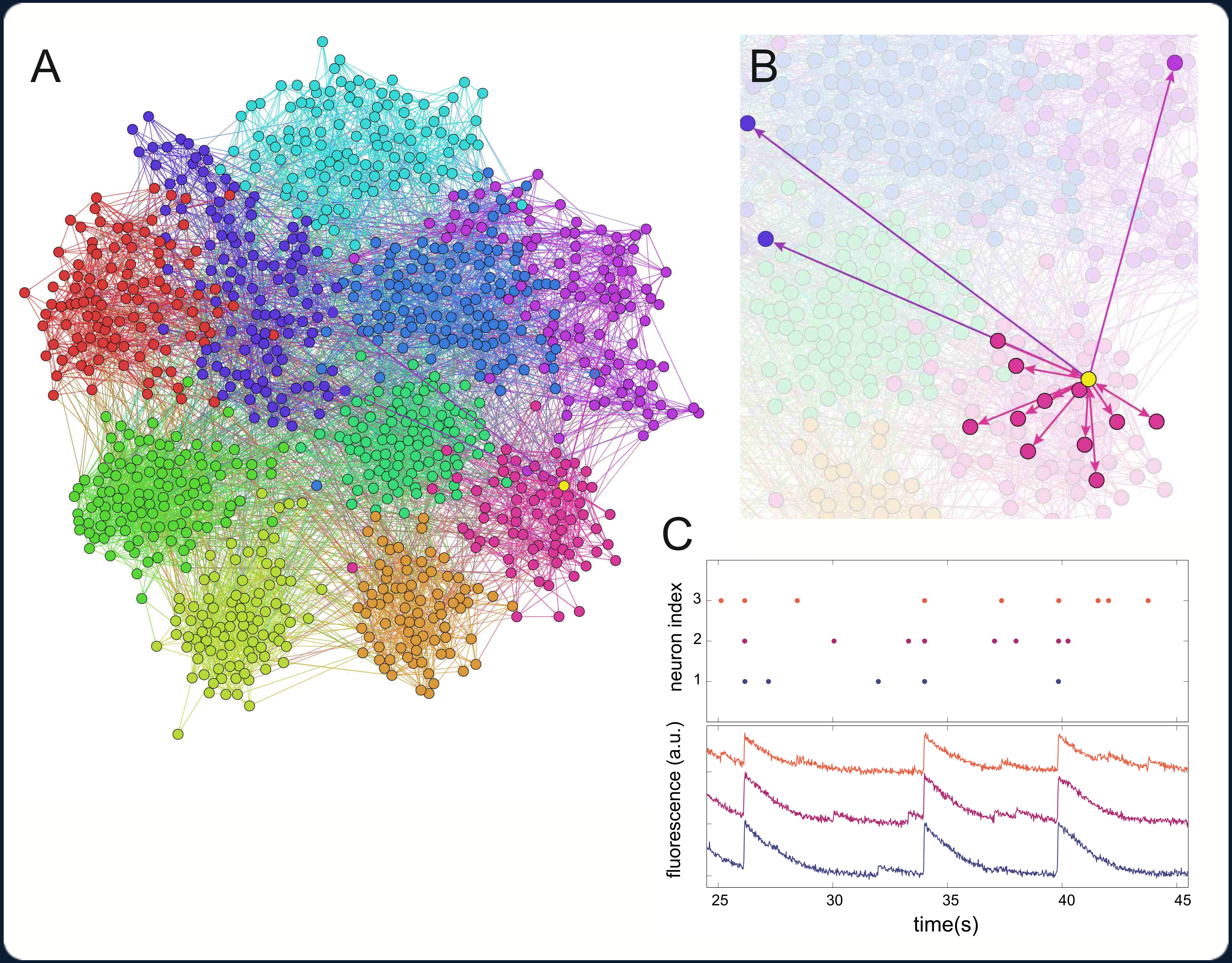
Infer the network structure of neuronal cultures based solely on simulated calcium fluorescence imaging data

Provided:
Fluorescence traces of 1000 neurons
Neuronal positions

10 sample datasets with known-truth topology with different conditions (signal to noise ratio, framerate, …)

Participants asked to predict the structural topology of two different networks with unknown topology - ranked based on AUC performance

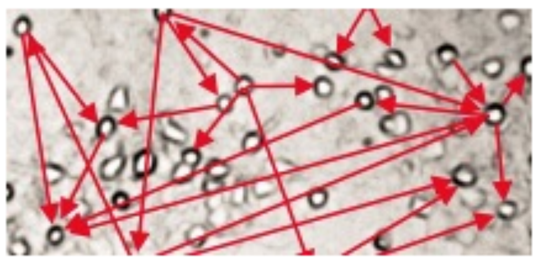
Challenge networks



Challenge overview

https://www.kaggle.com/c/connectomics/

kaggle | Customer Solutions | Competitions | Community | Javier G. Orlandi | Logout



CONNECTOMICS

Finished

Wednesday, February 5, 2014 | $3,000 • 144 teams | Monday, May 5, 2014

Dashboard

- Home
  - Data
  - Make a submission
- Information
  - Description
  - Evaluation
  - Rules
  - Prizes
  - Timeline
- Forum

Competition Details » Get the Data » Make a submission

Reconstruct the wiring between neurons from fluorescence imaging of neural activity

Understanding the brain structure and some of its disease alterations is key to research on the treatment of epilepsy, Alzheimer's disease, and other neuropathologies, as well as understanding the general function of the brain and its learning capabilities. The brain contains nearly 100 billion neurons with an average

Challenge overview

https://www.kaggle.com/c/connectomics/data

Deep CNN for Time Series Correlation Measurement
28 days ago

Fast Matlab code to get a score of .93985
30 days ago

144 teams

185 players

1459 entries

What's in the data?

Validation and test data include time series of activities of neurons "extracted" from simulated calcium fluorescence imaging data. The neurons are arranged on a flat surface simulating a neural culture and you get the coordinates of each neuron. The model takes into account light scattering effects. [Learn more...]

For "training" data, the network connectivity is also provided.

What am I predicting?

The network connectivity for the validation (valid) and test data.

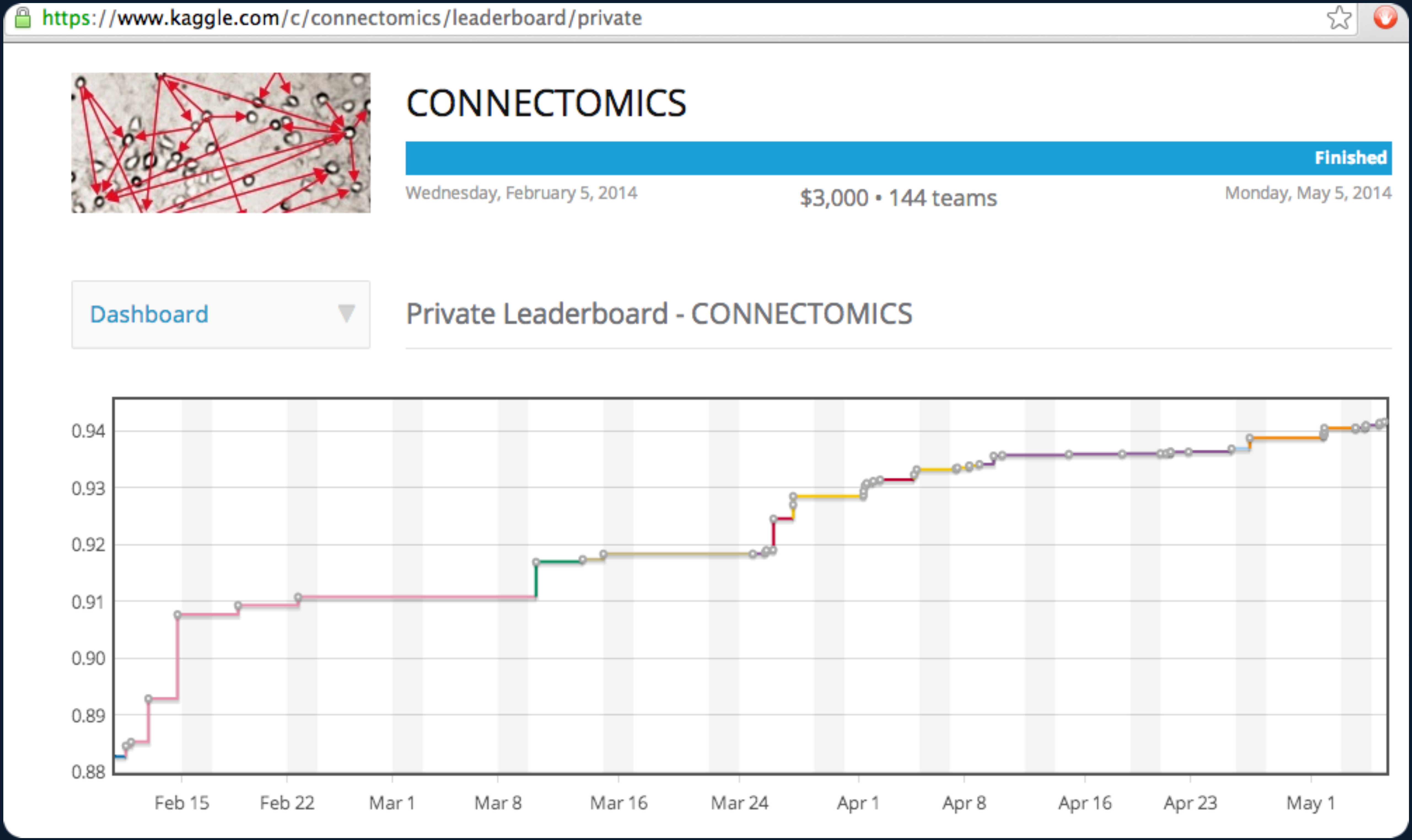
How do I get started?

We provide a sample submission corresponding to the Correlation benchmark, sample code and tutorial material. See also our frequently asked questions.

File types

- **"fluorescence" files (type F):** These are the time series of neural activities obtained from fluorescence signals. The neurons are in columns and the rows are time ordered samples. The signals are sampled at 20ms intervals.
- **"networkPosition" files (type P):** Each row represents a neuron. First column = X position; second column = Y position. The neurons span a $1mm^2$ square area.

Challenge evolution

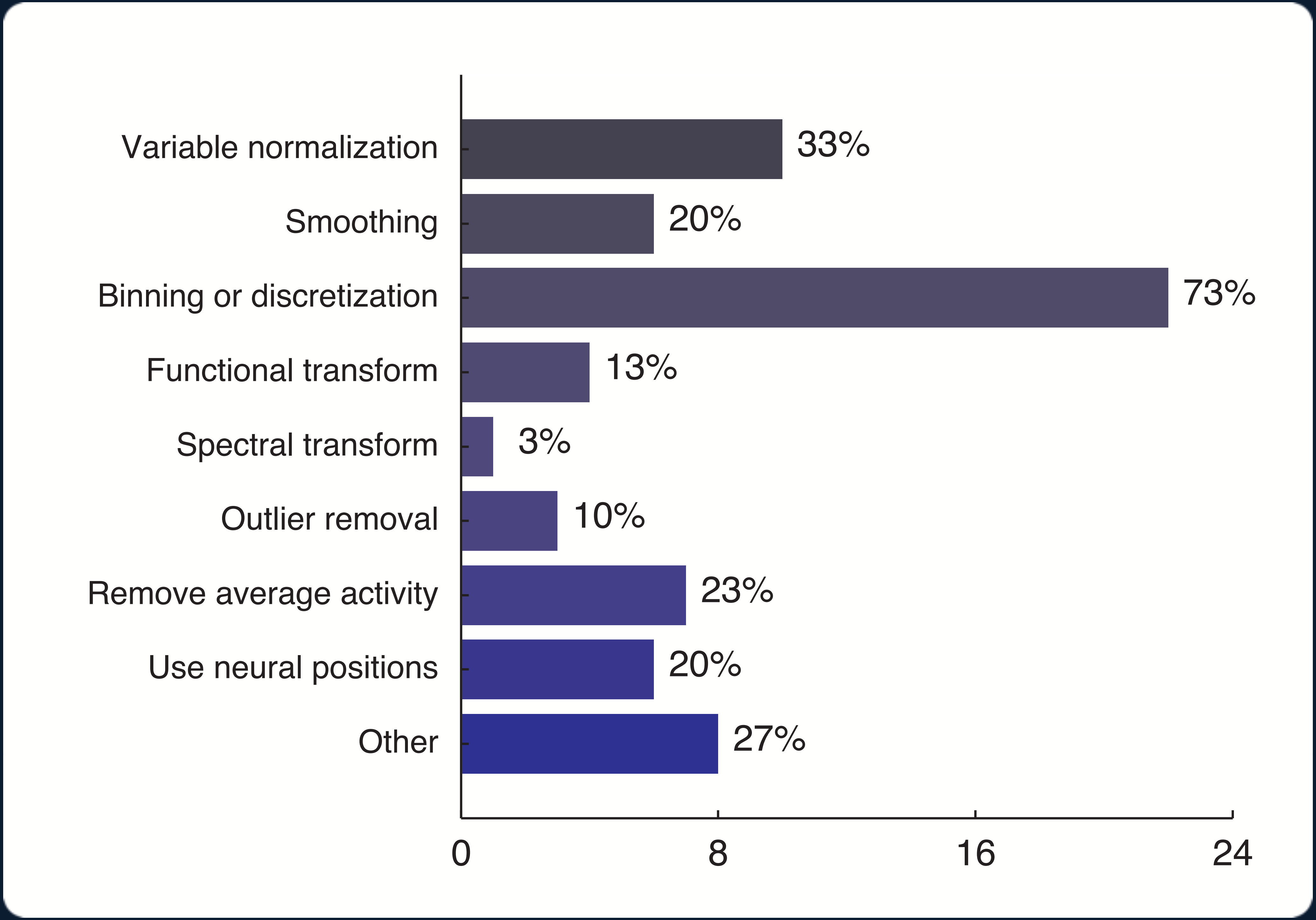


~20% performance increase over standard techniques (in 3 months)

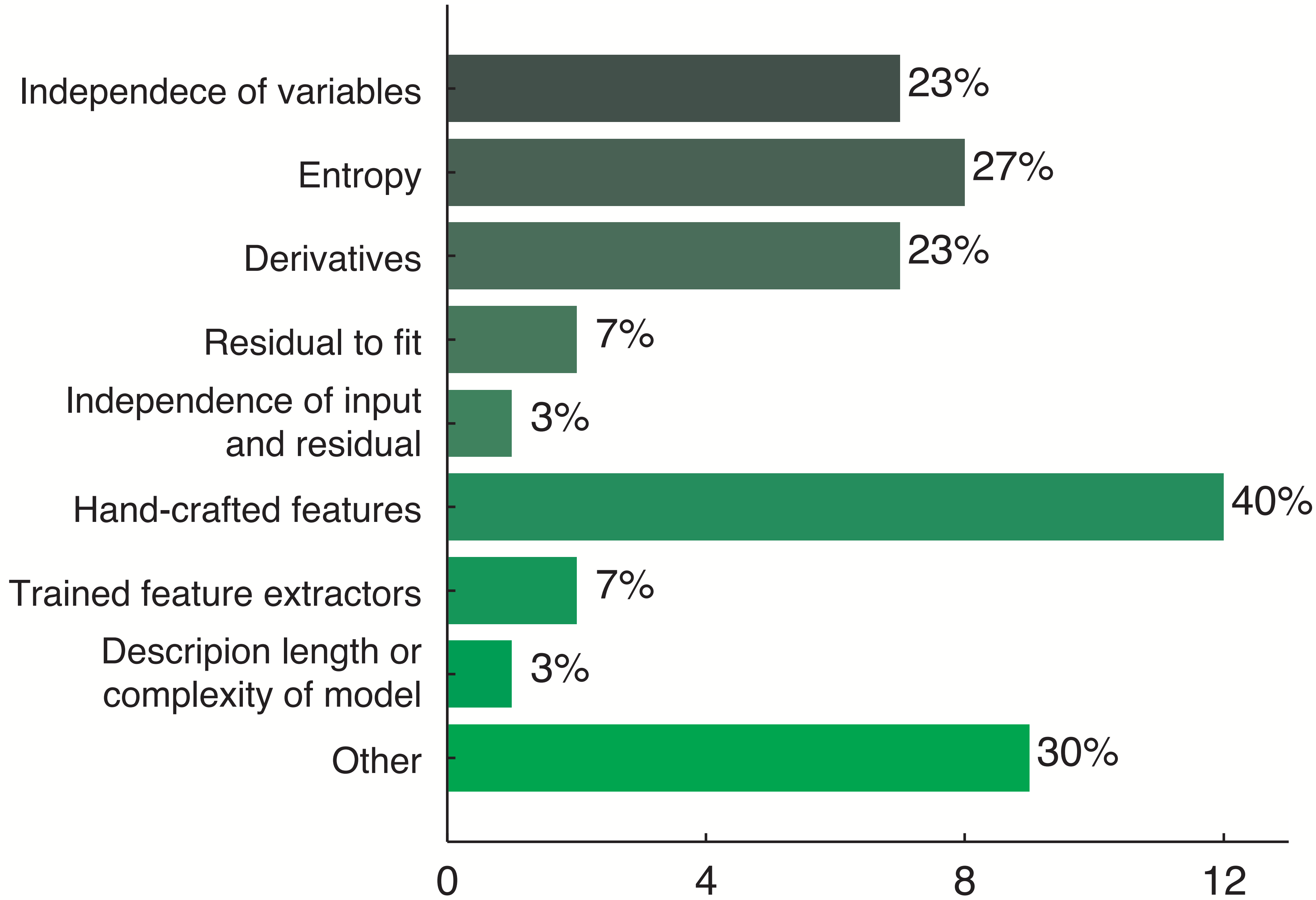
Challenge final scores

#	Δ1w	Team Name ‡ model uploaded * in the money	Score	Entries	Last Submission UTC (Best – Last Submission)
1	↑6	AAAGV ‡ *	0.94161	145	Mon, 05 May 2014 20:50:39
2	↑3	Matthias Ossadnik *	0.94102	71	Mon, 05 May 2014 19:27:34 (-0h)
3	↓2	Ildefons Magrans ‡ *	0.94063	21	Sun, 04 May 2014 13:47:26 (-0.3h)
4	↓2	Lukasz 8000	0.93956	101	Mon, 05 May 2014 15:54:48 (-18.1h)
5	↓2	Lejlot & Rafal	0.93826	97	Mon, 05 May 2014 15:39:39 (-20.9h)
6	—	Sium	0.93711	28	Mon, 05 May 2014 21:45:26 (-0.2h)
7	↓3	Alexander N & vopern	0.93666	7	Sat, 03 May 2014 23:46:07
8	—	gaucho81	0.93385	43	Sun, 04 May 2014 21:48:37 (-30.2h)
9	↑8	killertom	0.93011	13	Mon, 05 May 2014 12:50:08
10	new	dhanson	0.92885	6	Mon, 05 May 2014 21:02:20
11	↓2	DJMN	0.92609	20	Mon, 05 May 2014 19:44:13
12	↓2	Gideon & Alex	0.92420	48	Mon, 05 May 2014 16:16:57 (-19.7h)
13	↓1	Sandro	0.92039	18	Mon, 05 May 2014 23:15:48 (-8.7h)
14	↑5	Selfish Gene	0.92039	20	Sun, 04 May 2014 16:38:44
15	↑7	Nitai Dean	0.91945	3	Mon, 05 May 2014 12:43:11

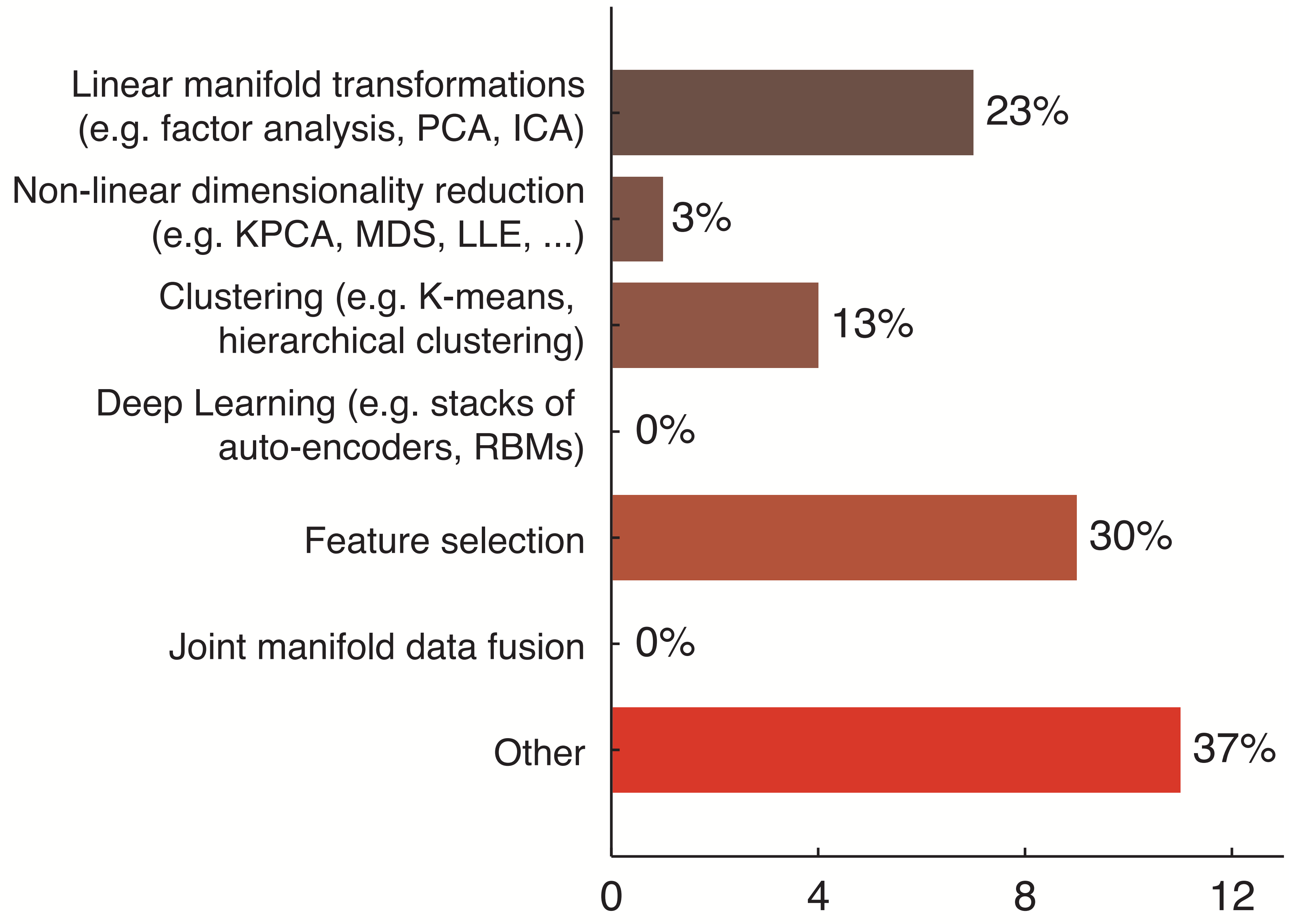
Preprocessing of fluorescence signals



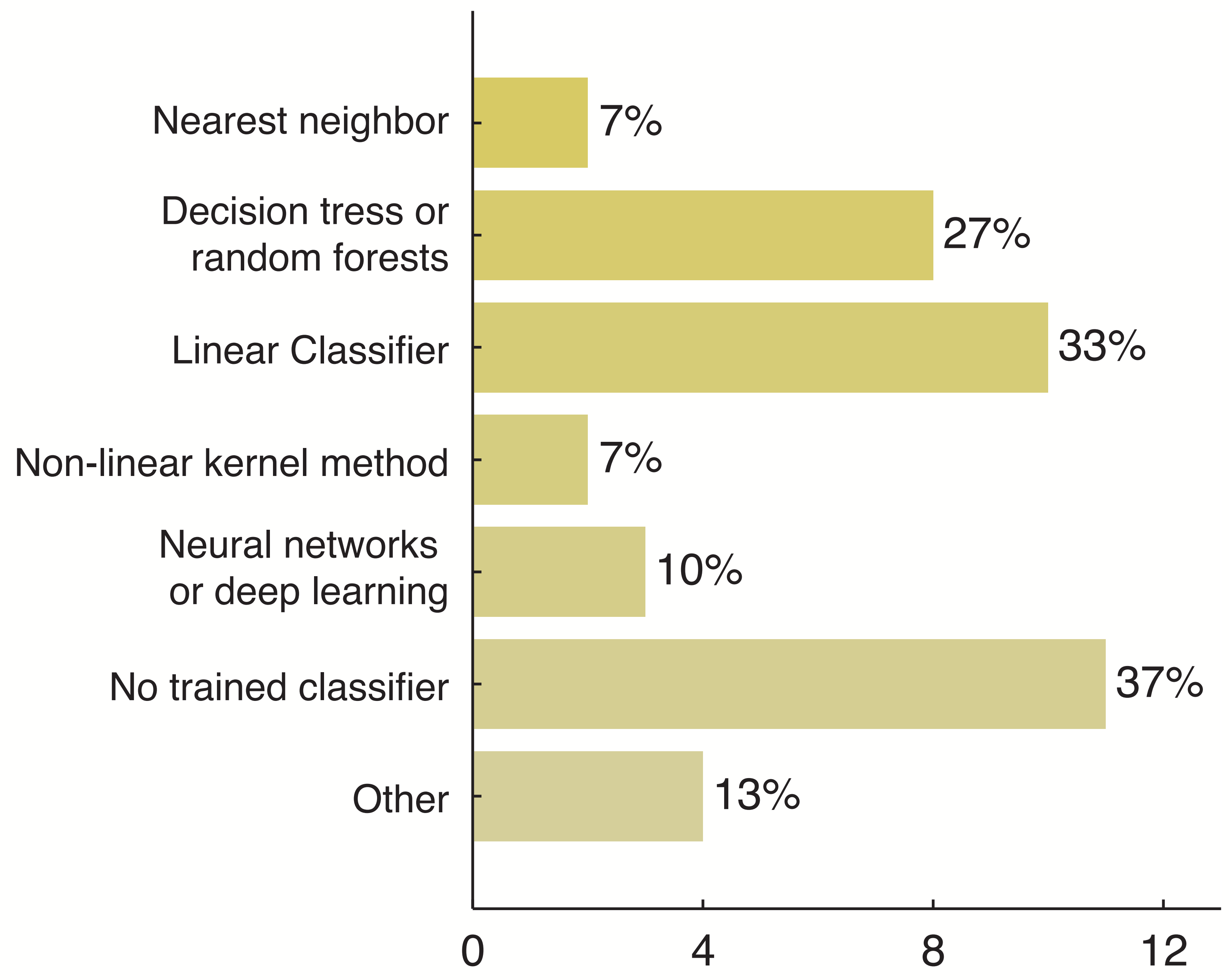
Feature extraction



Dimensionality reduction



Classification techniques



Post-challenge

Analysis of methods used by the participants:
Deep convolutional neural networks
State selection
Multivariate logistic regression of inferred spike trains
Inverse covariance matrix
Random forests and gradient boosting machines
Network deconvolution
…

Checking for robustness of the methods
Pool together the different approaches (wisdom of crowds)

Code from the winning teams is publicly available with open sources licenses (and all the data associated with the challenge) http://www.kaggle.com/c/connectomics

Conclusions

Neuronal cultures

Useful tool to test concepts and properties related to the connectome in a well controlled environment

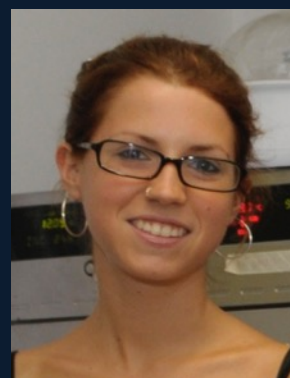
Challenges

A different approach to collaborative research

Acknowledgments

Universitat de Barcelona

Jordi Soriano



Sara Teller

Jaume Casademunt

Olav Stetter

Demian Battaglia

Theo Geisel

Max Planck Institute for Dynamics and Self-Organization

Isabelle Guyon, ChaLearn, California, USA

Alexander Statnikov, New York University, Langome Medical Center, NY, USA

Bisakha Ray, New York University, Langome Medical Center, NY, USA

Mehreen Saeed, Department of Computer Science, FAST, National University of Computer & Emerging Sciences, Lahore Campus, Pakistan