

# Investigating the connectome of a larval *Drosophila* brain

**Benjamin D. Pedigo**

(he/him)

[NeuroData lab](#)

Johns Hopkins University - Biomedical Engineering

 [bpedigo@jhu.edu](mailto:bpedigo@jhu.edu)

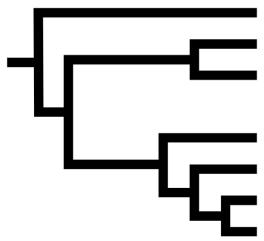
 [@bdpedigo](#) (Github)

 [@bpedigod](#) (Twitter)

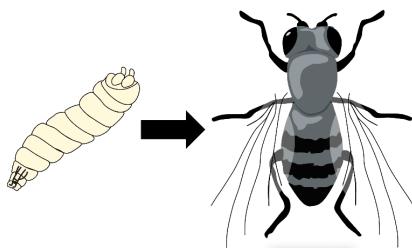
 [bpedigo.github.io](http://bpedigo.github.io)

These slides at:

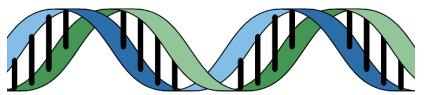
# Many goals of connectomics involve linking connectome to other properties



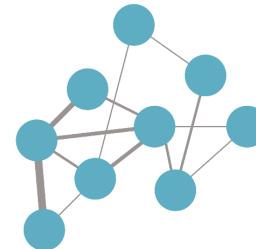
Evolution



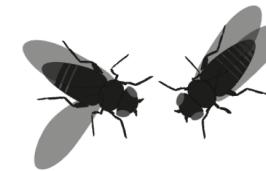
Development



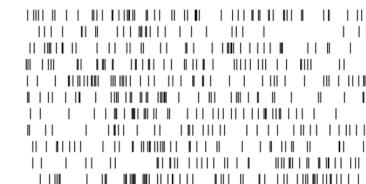
Genome



Connectome



Behavior



Activity

# What do we do with these datasets once we have them?

- Hypothesis *generation*
  - How could signals travel on this network, from sensory inputs to motor outputs?
  - What cells have similar patterns of connectivity?
- Hypothesis *testing*
  - Connectome (network) as an object that we want to "do inference" on

# Comparative connectomics as a potential solution?

- Map connectomes from related individuals/organisms which may differ in feature  $X$ :
  - Genome
  - Behavioral patterns/habits
  - Life experience
  - Developmental stage
- Compare connectomes
- Understand how  $X$  {affects, is affected by, is associated with} connectome structure

# Why is comparative connectomics hard?

Collecting the data is still a large effort...

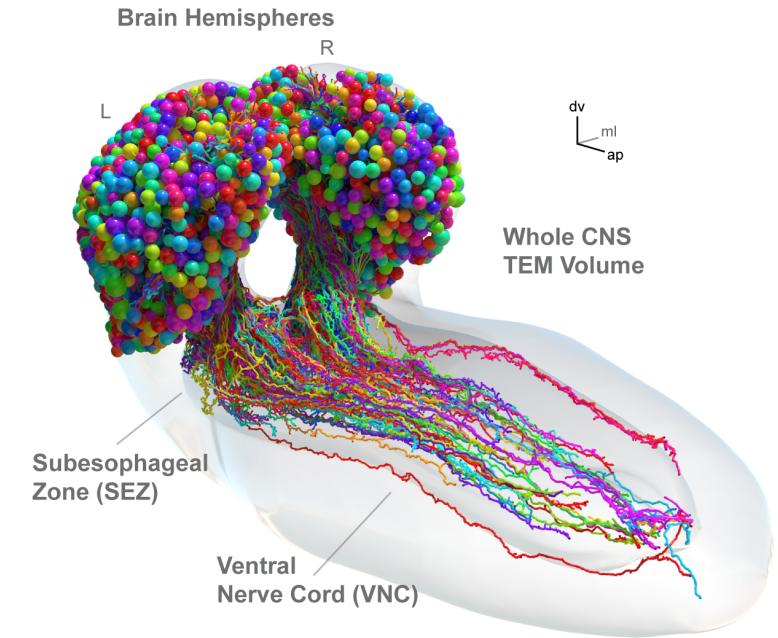
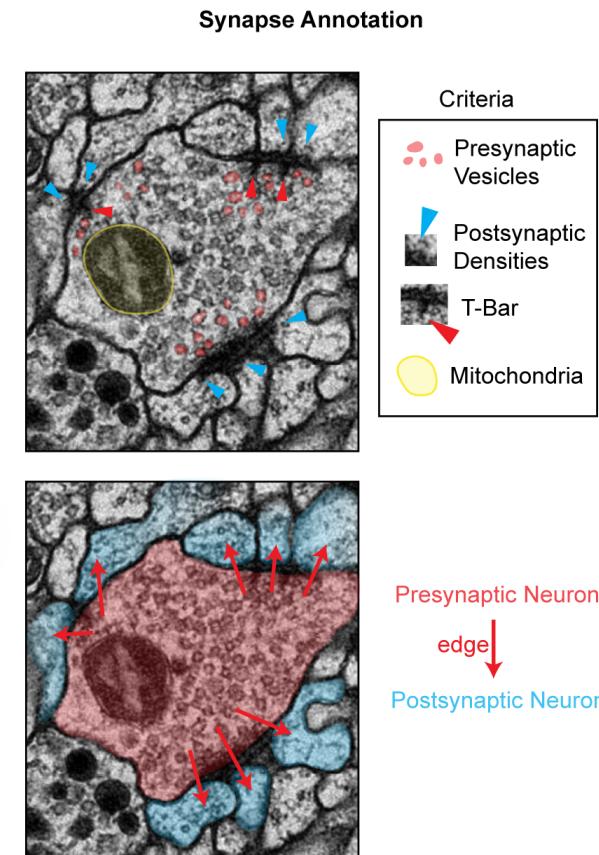
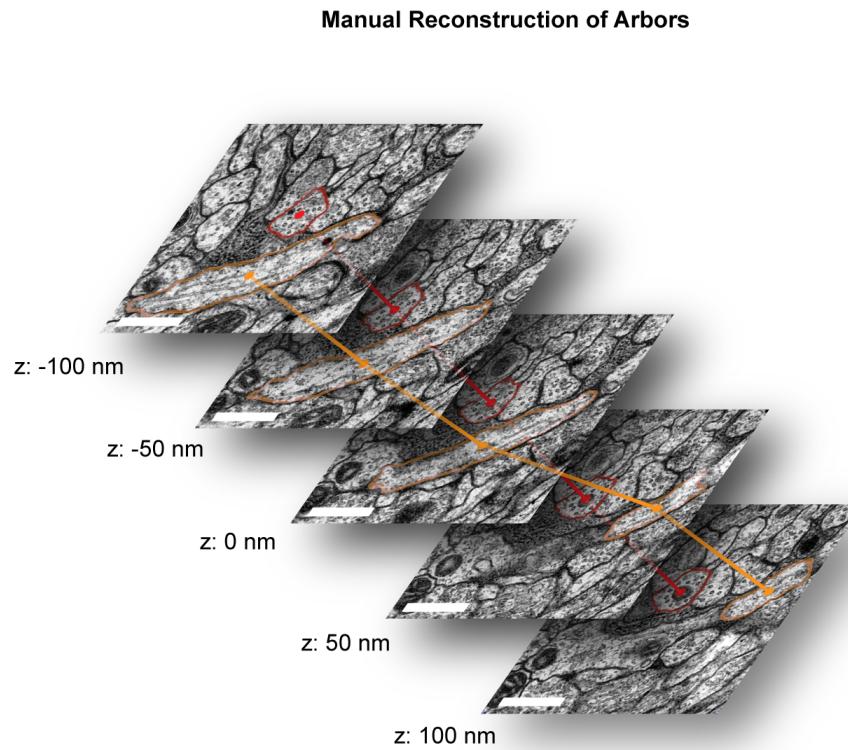
But how do we even compare connectomes once we have them?

- Data are networks
  - Data are networks with rich attributes
- Data will always have noise
  - "Experimental noise"
  - "Biological noise"
- Data are big (and getting bigger)

# Outline

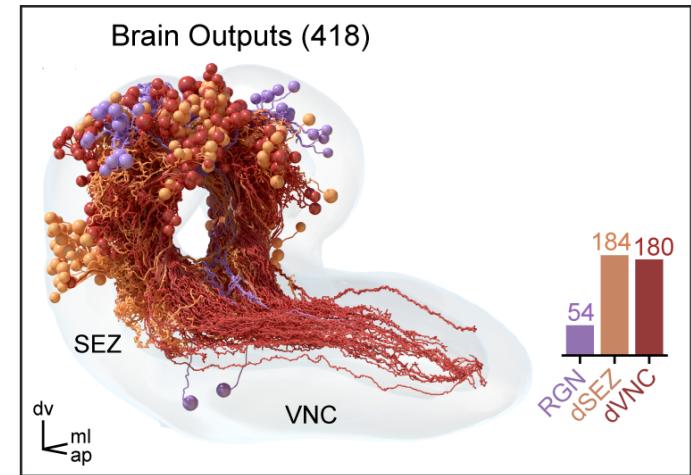
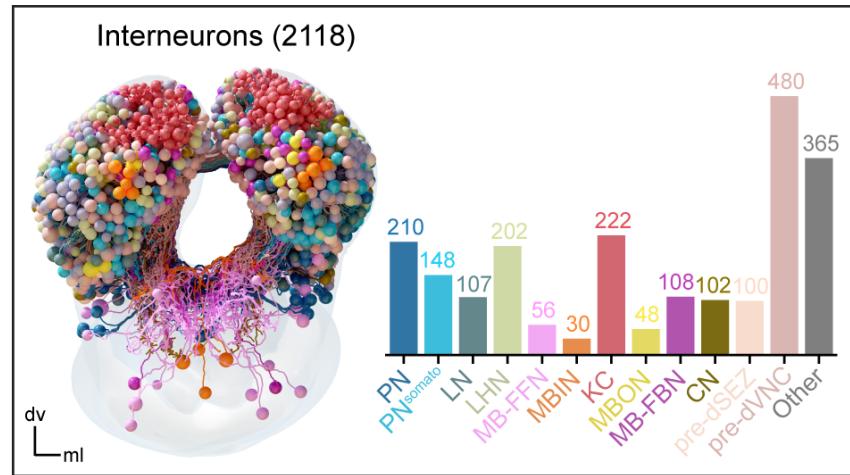
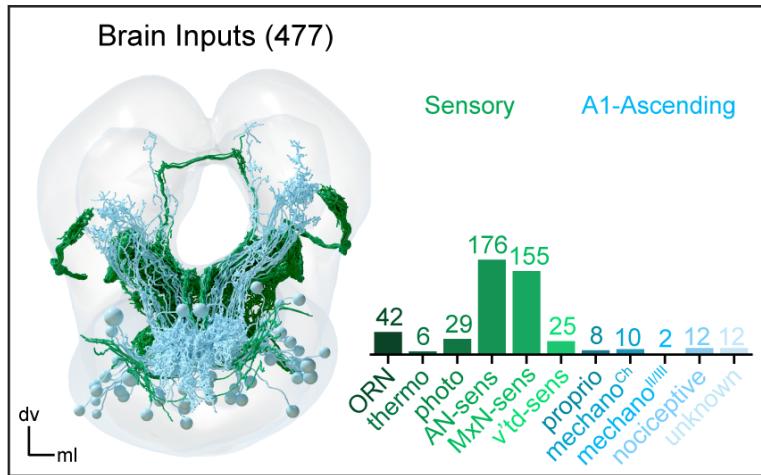
- **Larval connectome dataset**
  - Flow
  - Connectivity based cell types
- **Connectome comparison via network hypothesis testing**
- **Pairing neurons across connectomes via graph matching**
- **Ongoing extensions/applications**

# Larval *Drosophila* brain connectome

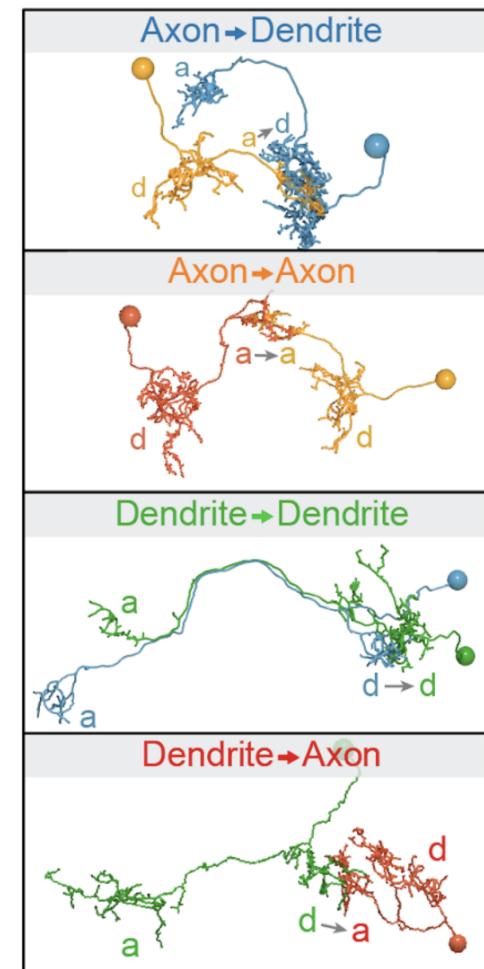
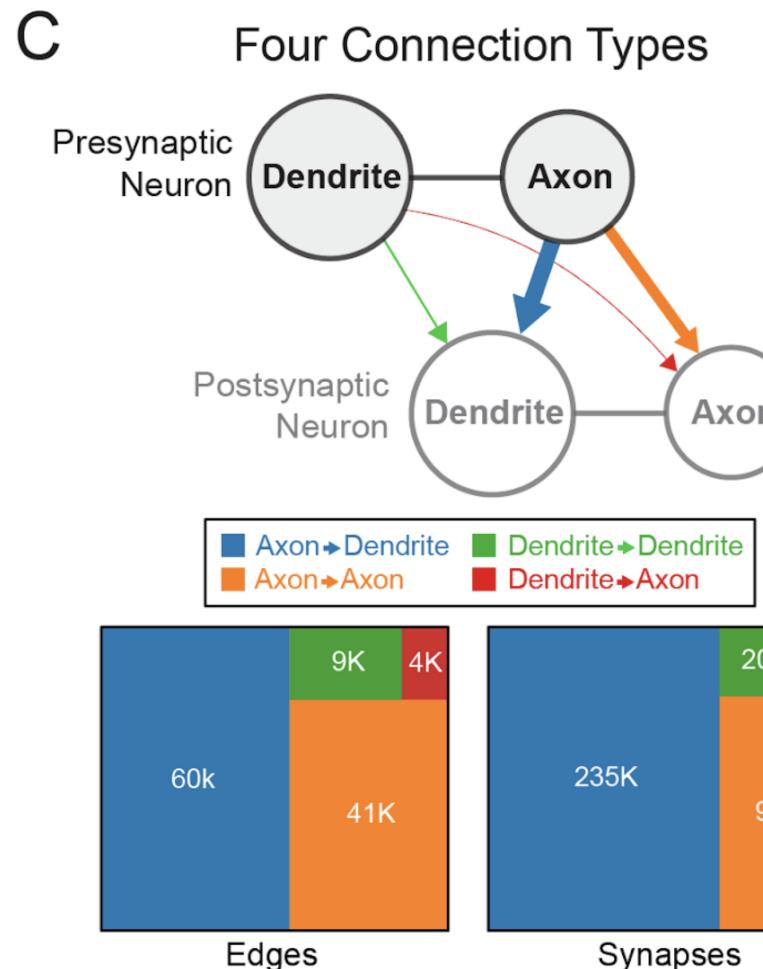
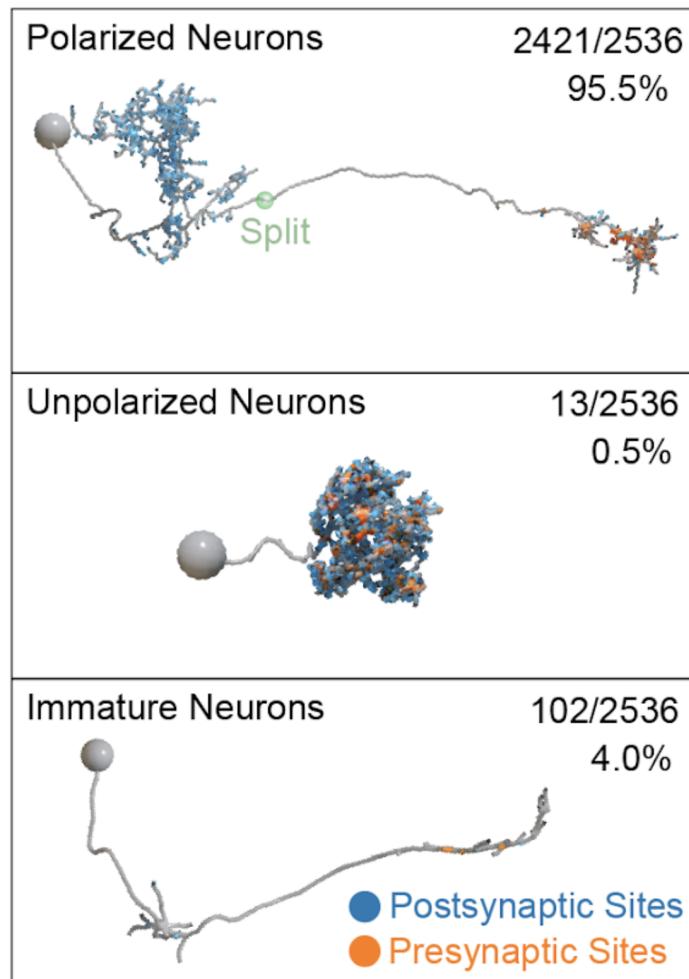


~3k neurons, ~550K synapses  
Both hemispheres

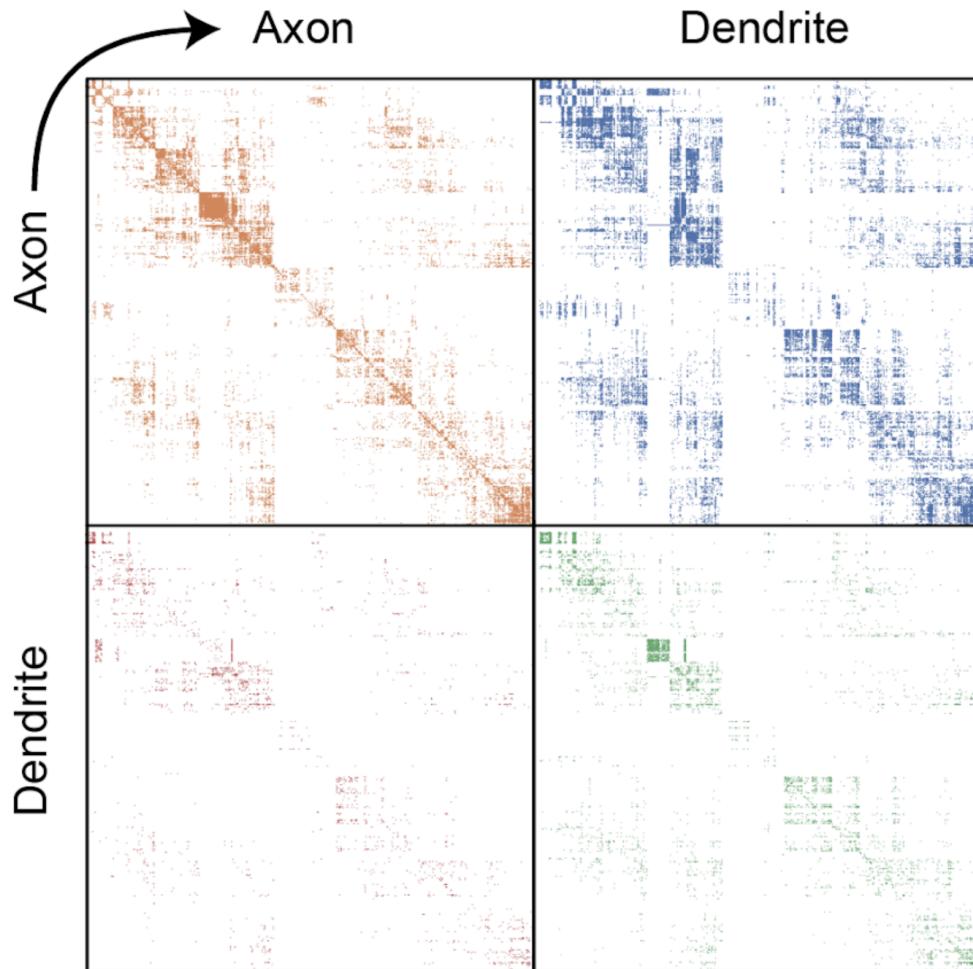
# High level (mostly anatomical) cell types



# Axons/dendrites can be split



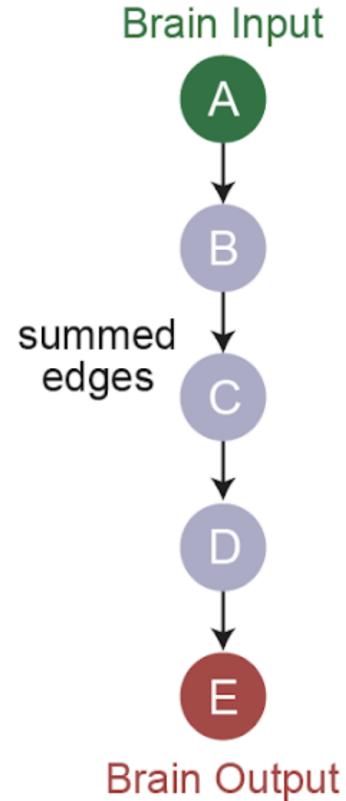
# This split induces 4 graphs (or layers)



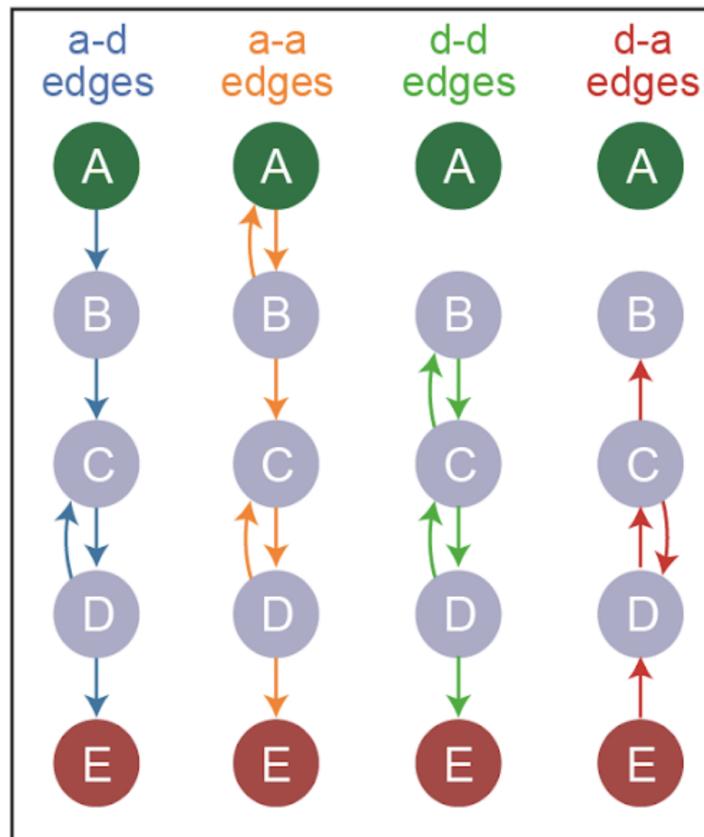
Edge	Nodes	Density	Max Degree
A→D	2880	0.77%	203
A→A	2887	0.49%	241
D→D	2204	0.19%	143
D→A	1907	0.10%	69

# What are these different "channels" doing?

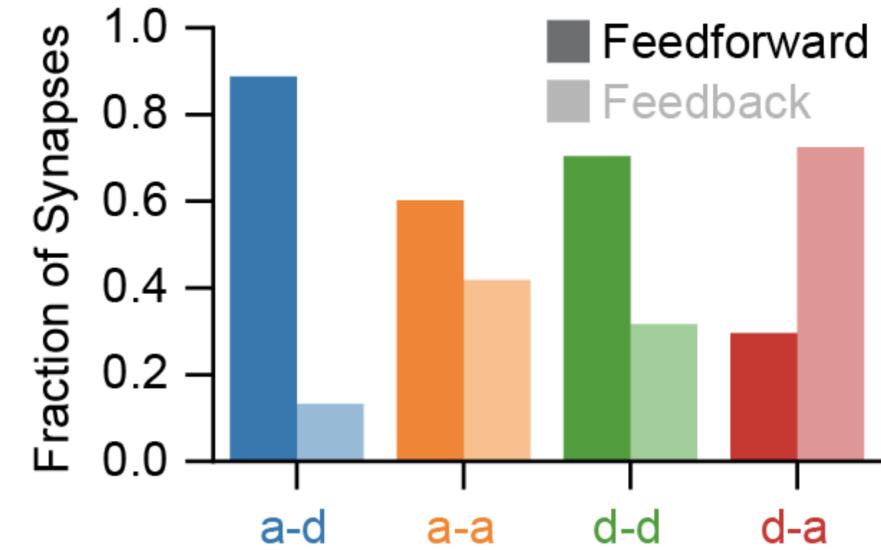
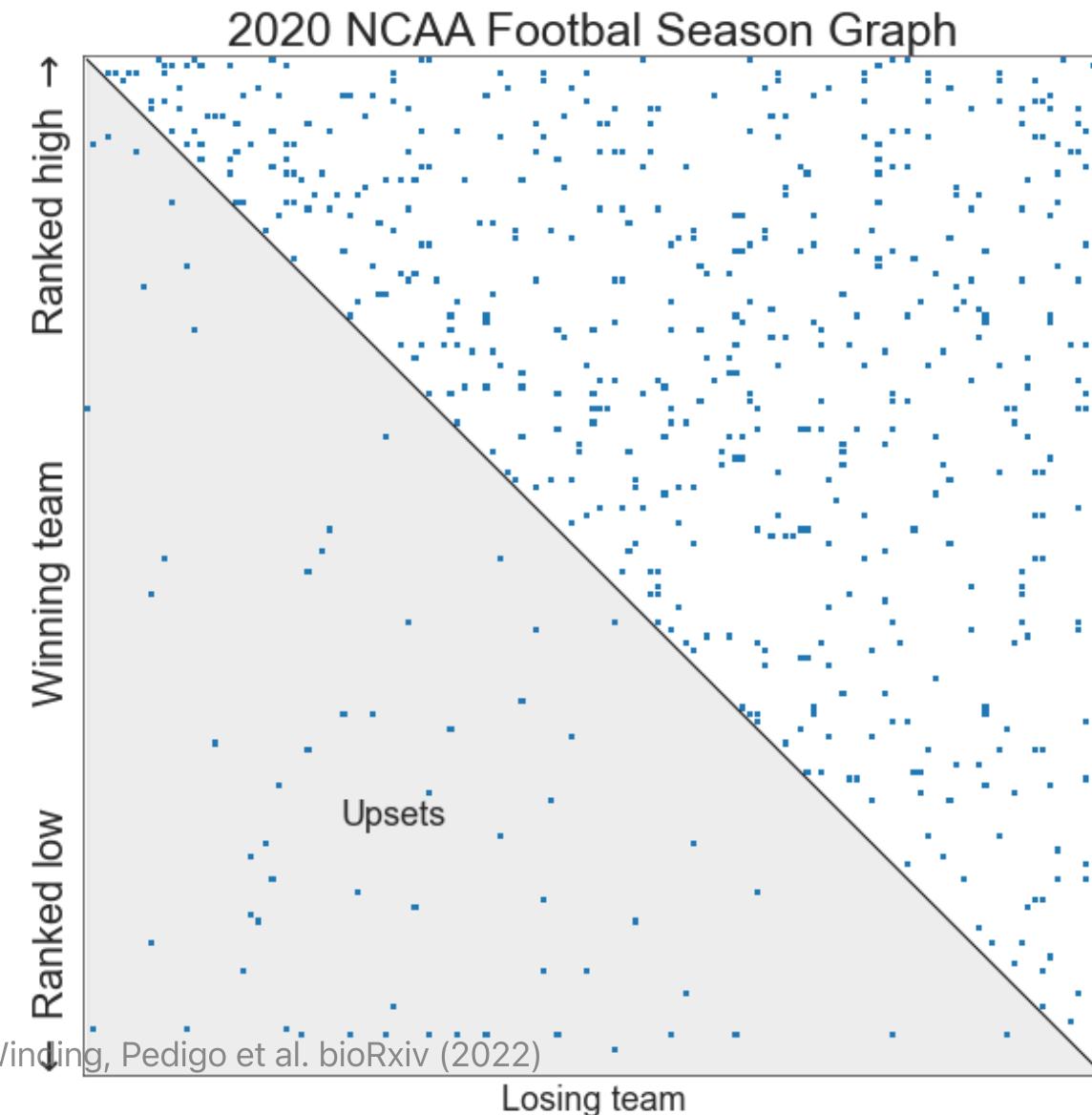
Sort based on  
summed connectivity



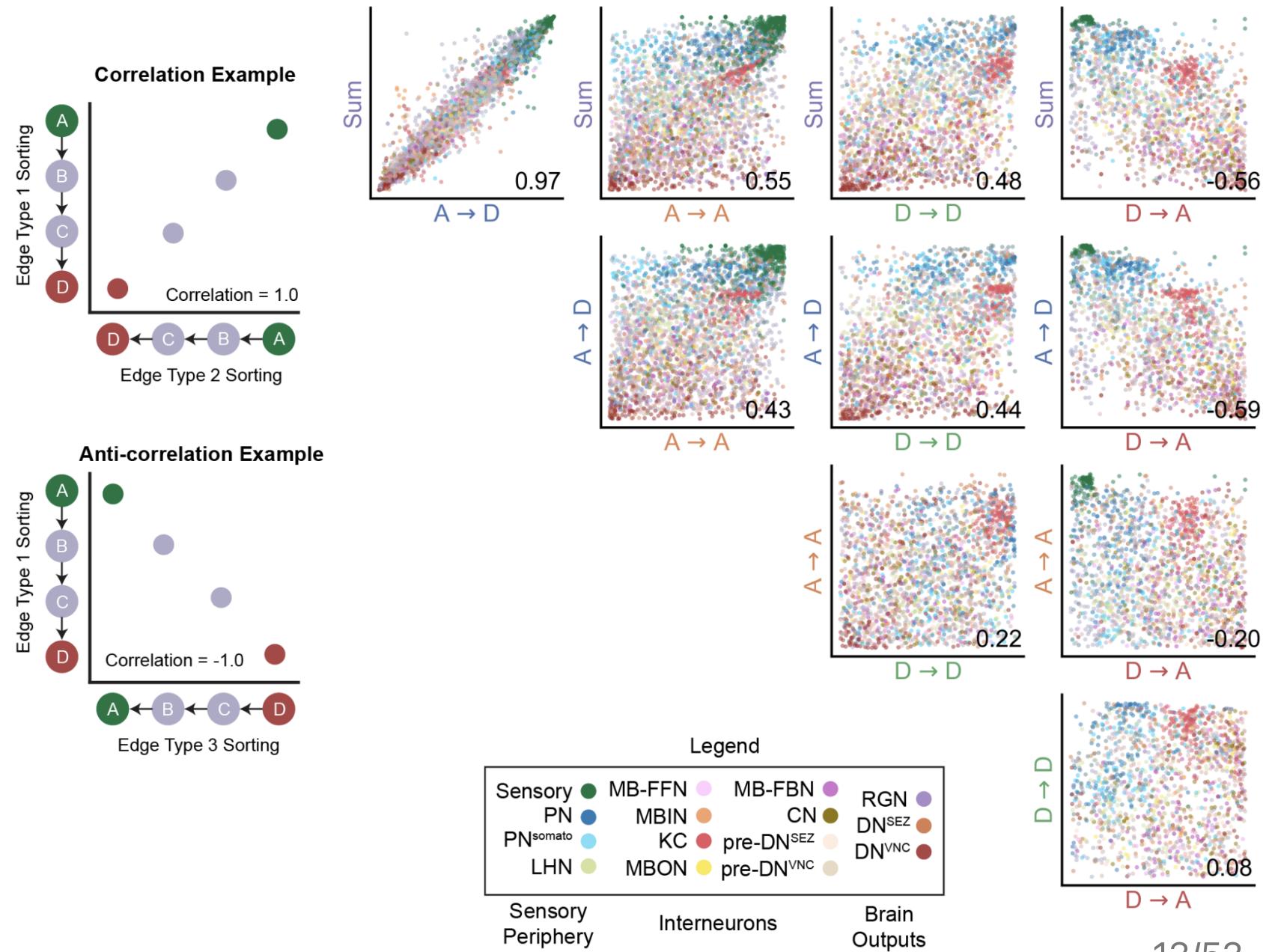
Analyze each edge type  
based on this sorting



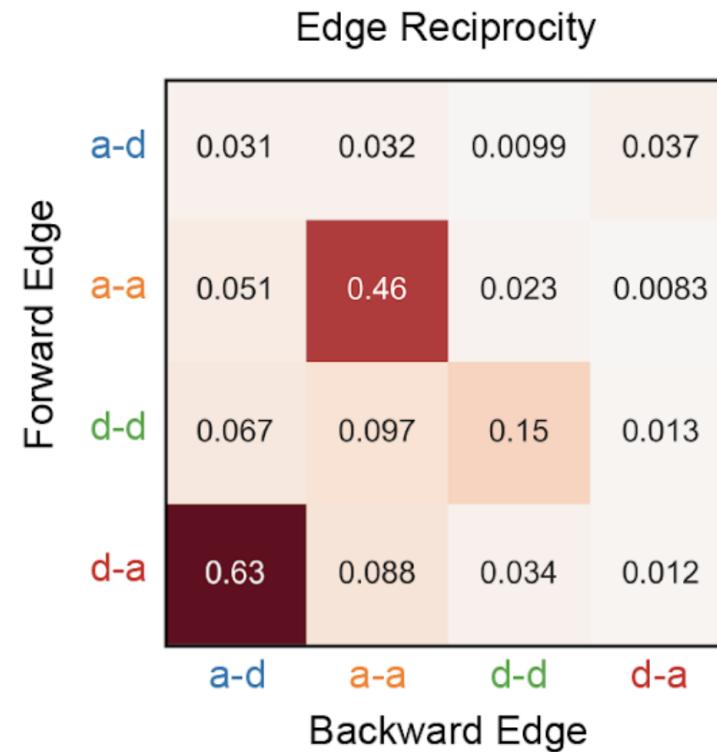
# Quantifying high-level "feedforward/feedback"



# Comparing independently sorted channels



# Edge reciprocity



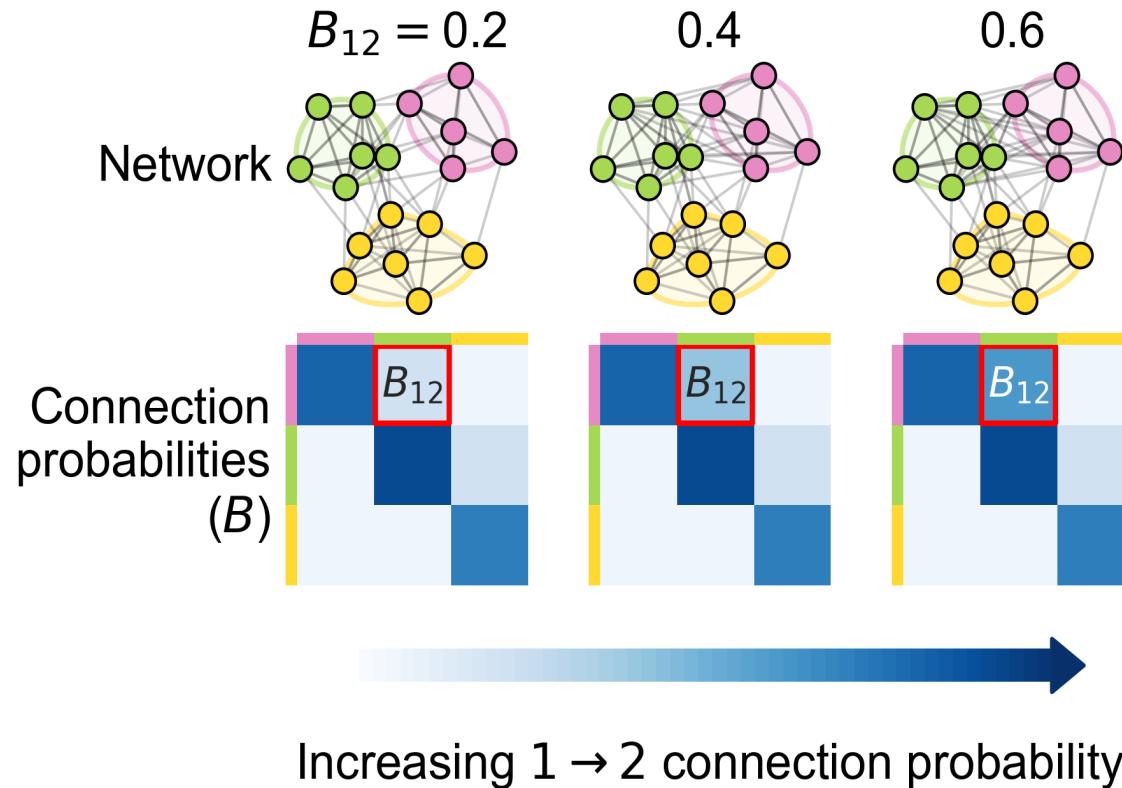
# Outline

- **Larval connectome dataset**
  - Flow
  - Connectivity based cell types
- **Connectome comparison via network hypothesis testing**
- **Pairing neurons across connectomes via graph matching**
- **Ongoing extensions/applications**

# Spectral embedding

- Spectral decomposition of the adjacency matrix (or Laplacian)
- Spectral decomp. + clustering how to be a consistent estimator of block model

# Stochastic block model



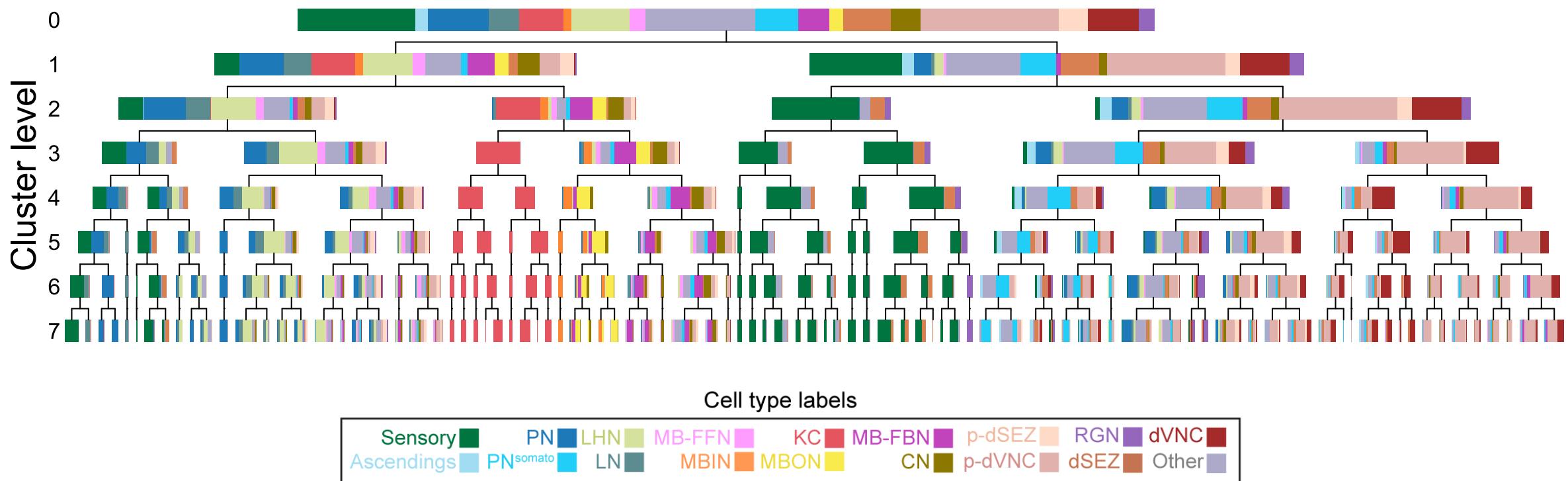
- Each node is assigned to a group
- $B$  is a matrix of connection probabilities between groups

# Diagram of procedure

- mention embedding
- cluster - when to stop

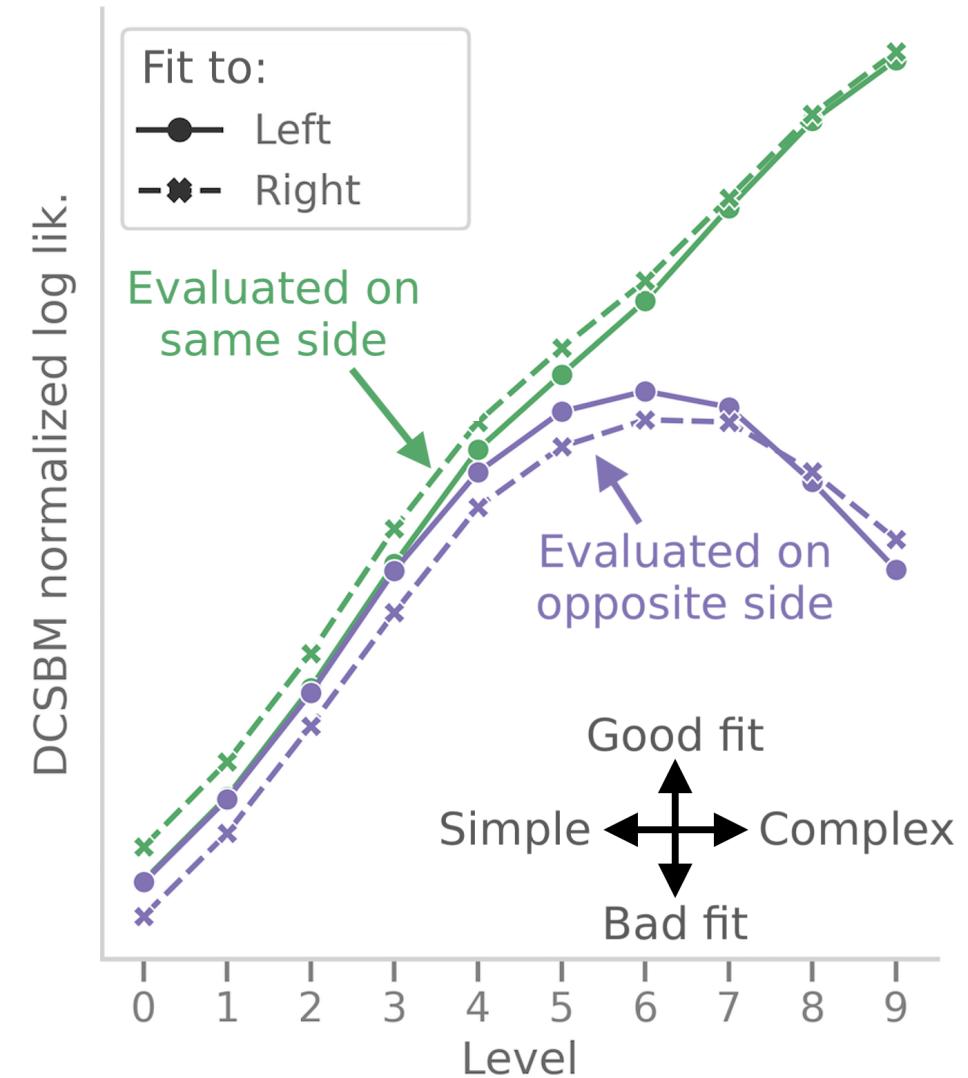
# Neurons clustered by connectivity using recursive spectral clustering

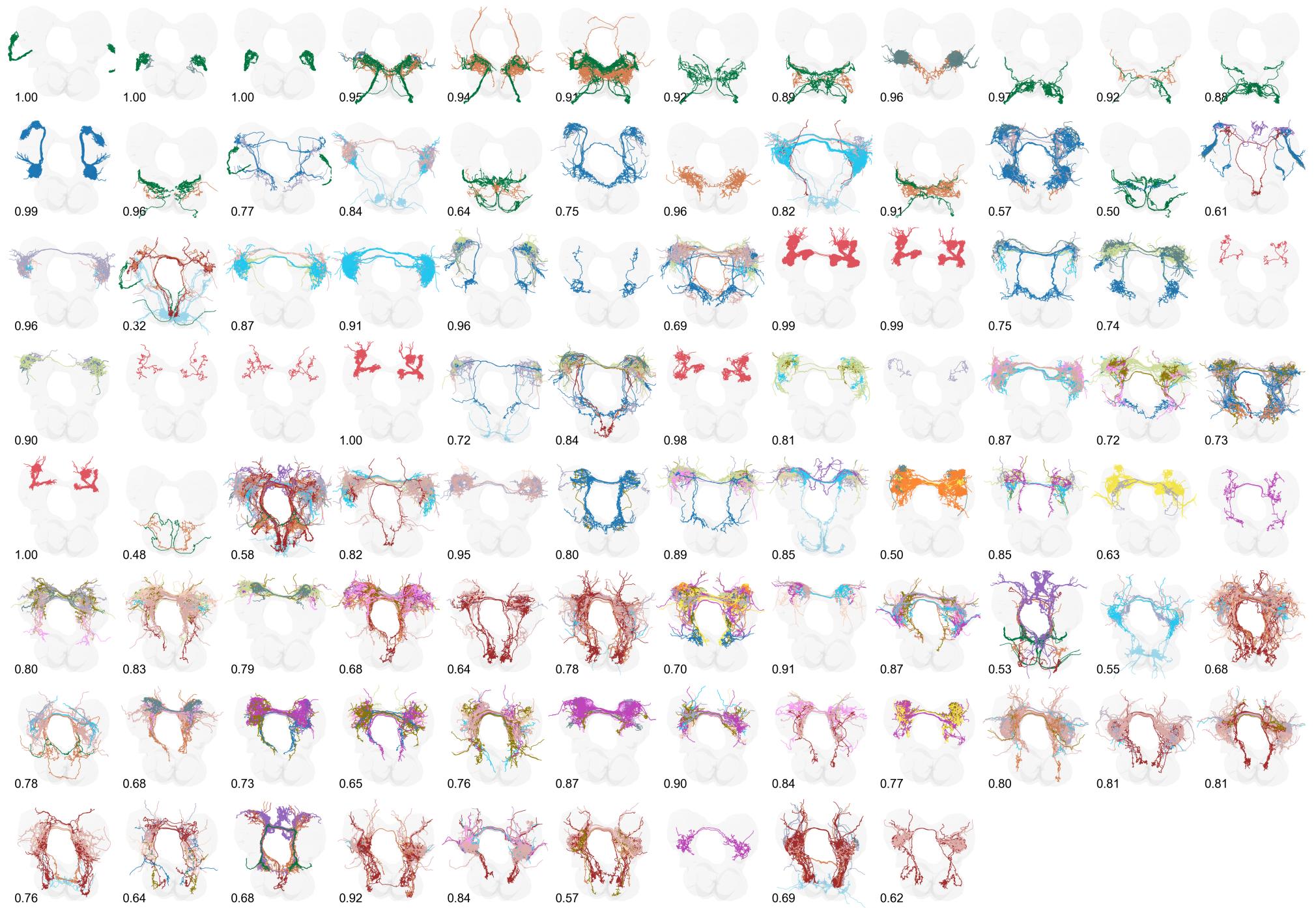
Where to stop splitting?



# Using models to evaluate cell type groupings

- Clustering nodes corresponds with inferring groups in a stochastic block model (DCSBM)
- How well do these models generalize to the other side of the brain (let alone the next maggot)?





# Bilateral symmetry

"This brain is bilaterally symmetric."

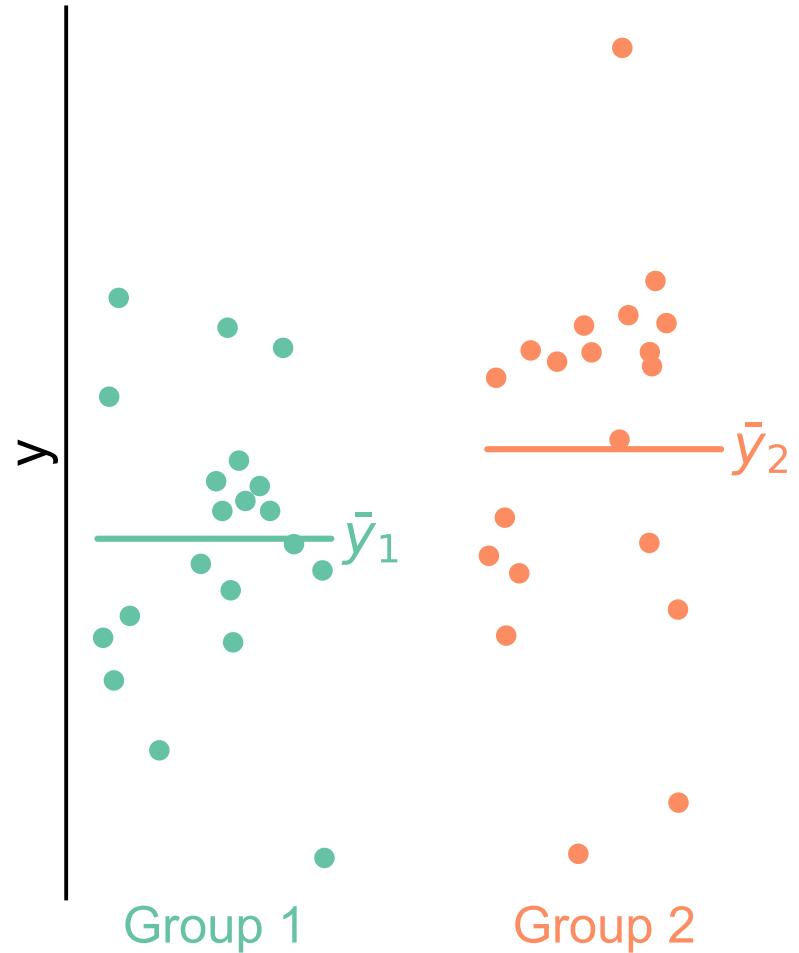
"What does that even mean? And how would we know if it wasn't?"

**Are the *left* and *right* sides of this connectome  
*different?***

# Outline

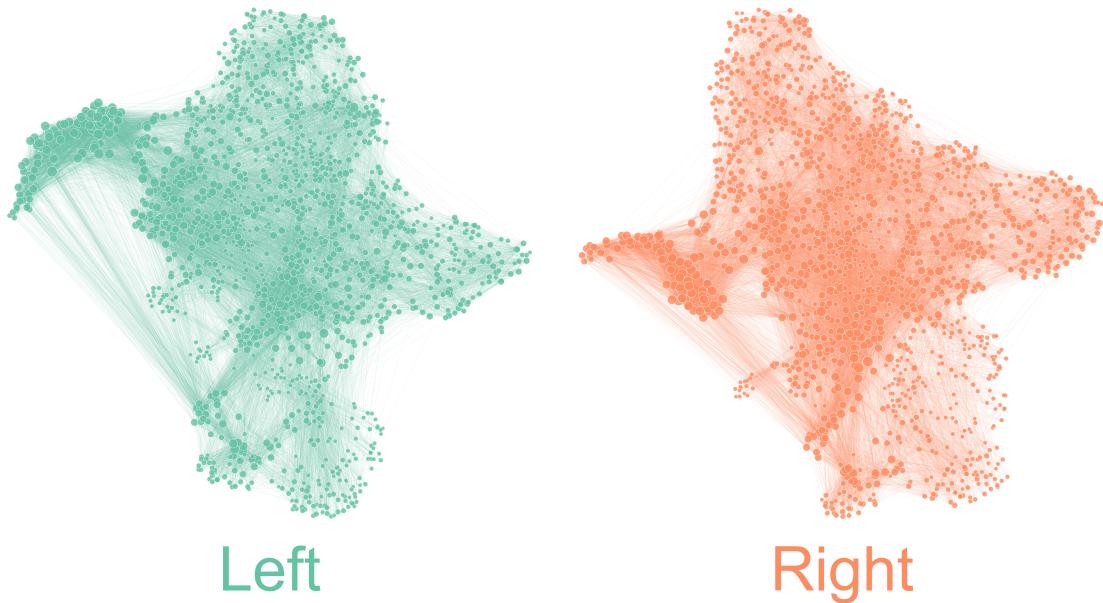
- Larval connectome dataset
- Connectome comparison via network hypothesis testing
- Pairing neurons across connectomes via graph matching
- Ongoing extensions/applications

# Are these populations different?



- Known as two-sample testing
- $Y^{(1)} \sim F^{(1)}$ ,  $Y^{(2)} \sim F^{(2)}$
- $H_0 : F^{(1)} = F^{(2)}$   
 $H_A : F^{(1)} \neq F^{(2)}$

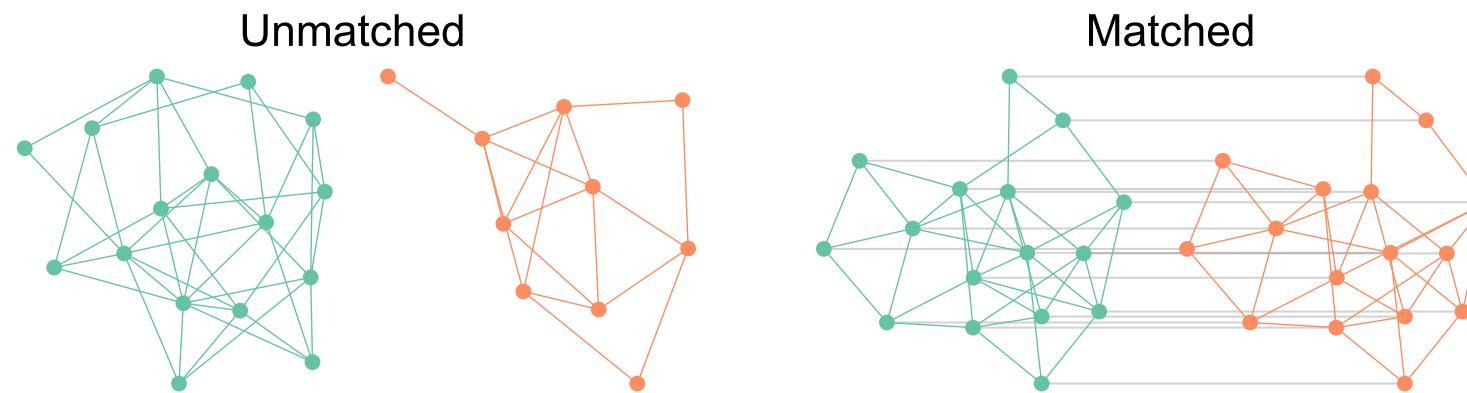
# Are these *networks* different?



- Want a two-network-sample test!
- $A^{(L)} \sim F^{(L)}, A^{(R)} \sim F^{(R)}$
- $H_0 : F^{(L)} = F^{(R)}$   
 $H_A : F^{(L)} \neq F^{(R)}$

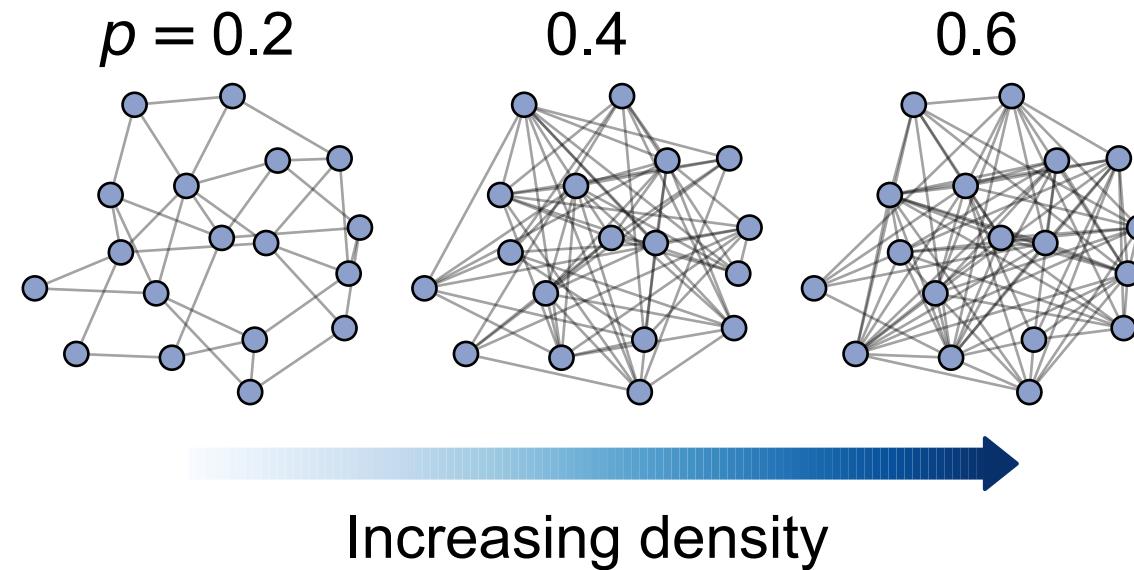
# Assumptions

- Know the direction of synapses, so network is *directed*
- For simplicity (for now), consider networks to be *unweighted*
- For simplicity (for now), consider the **left → left** and **right → right** (*ipsilateral*) connections
- Not going to assume any nodes are matched



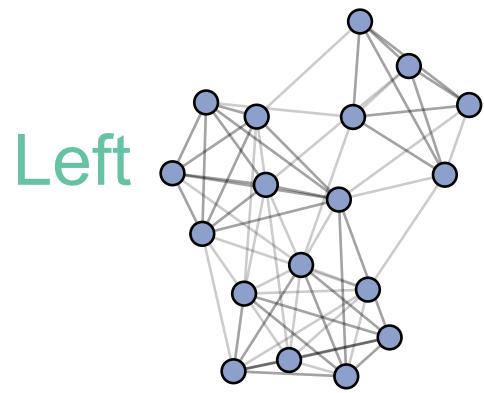
# Erdos-Renyi model

- All edges are independent
- All edges generated with the same probability,  $p$

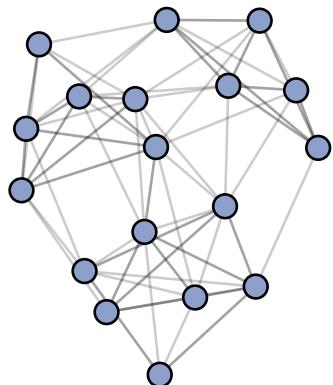


# Detect a difference in density

Compute global connection density



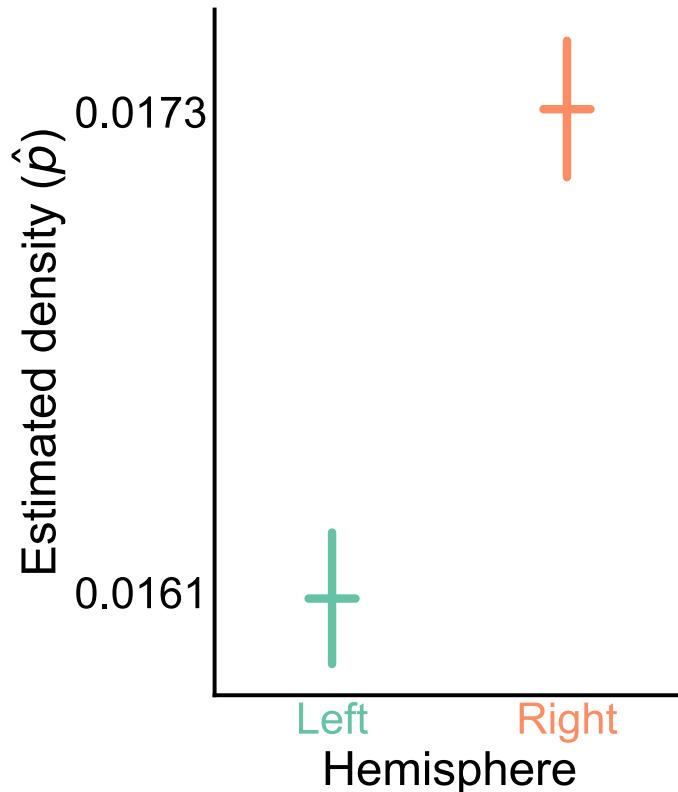
Right



Compare ER models

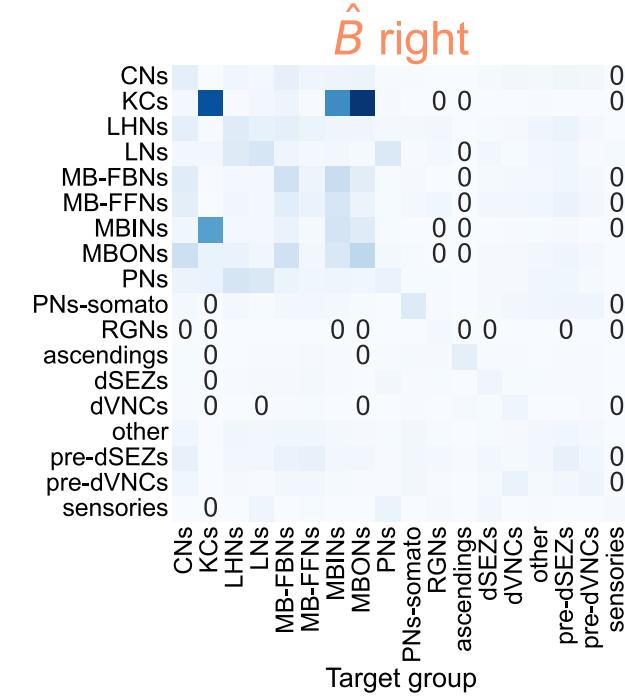
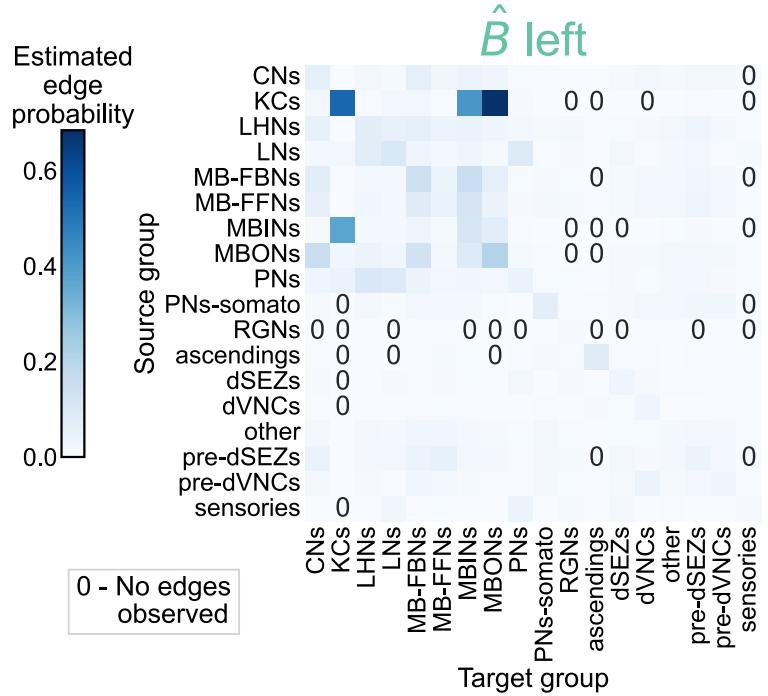
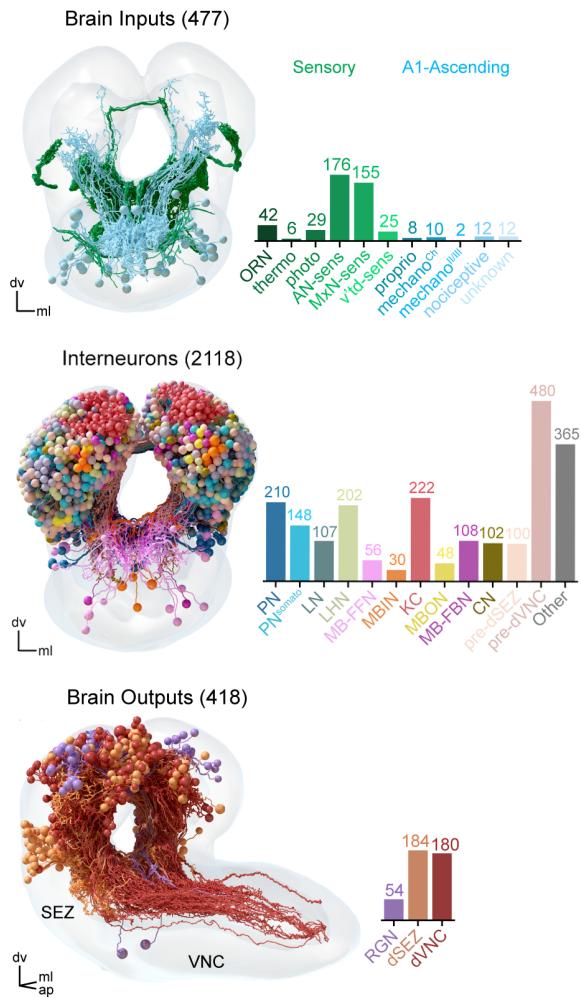
$$p = \frac{\text{\# edges}}{\text{\# potential edges}}$$

$$\begin{aligned} H_0: p^{(L)} &= p^{(R)} \\ H_A: p^{(L)} &\neq p^{(R)} \end{aligned}$$



p-value <  $10^{-22}$

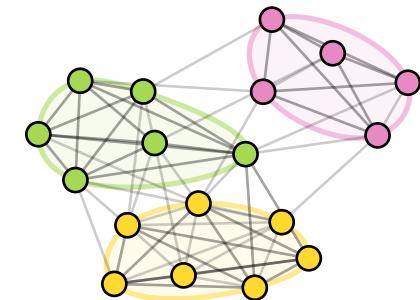
# Connection probabilities between groups



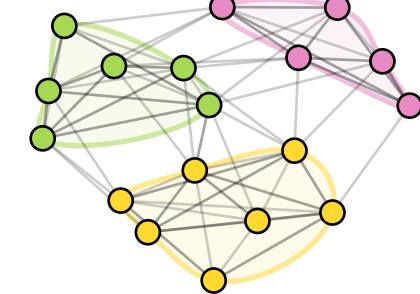
# Group connection test

Group neurons

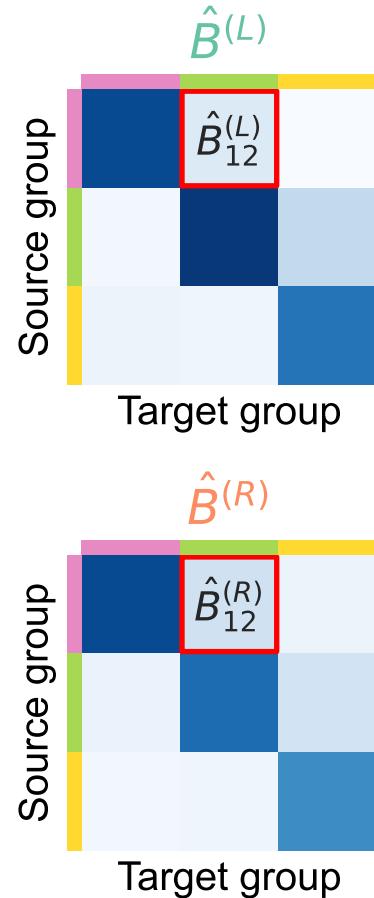
Left



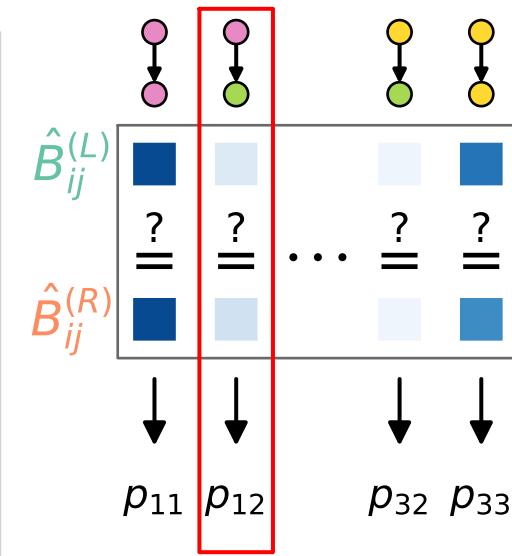
Right



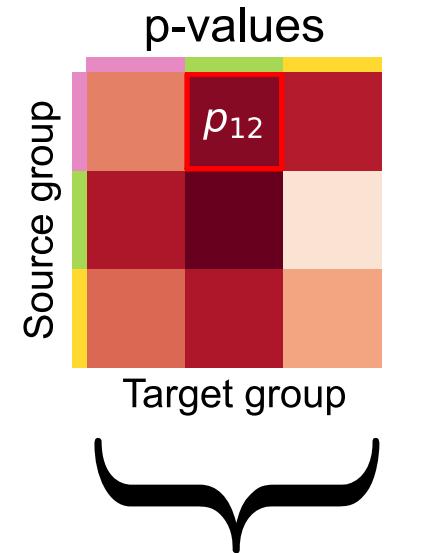
Estimate group connection probabilities



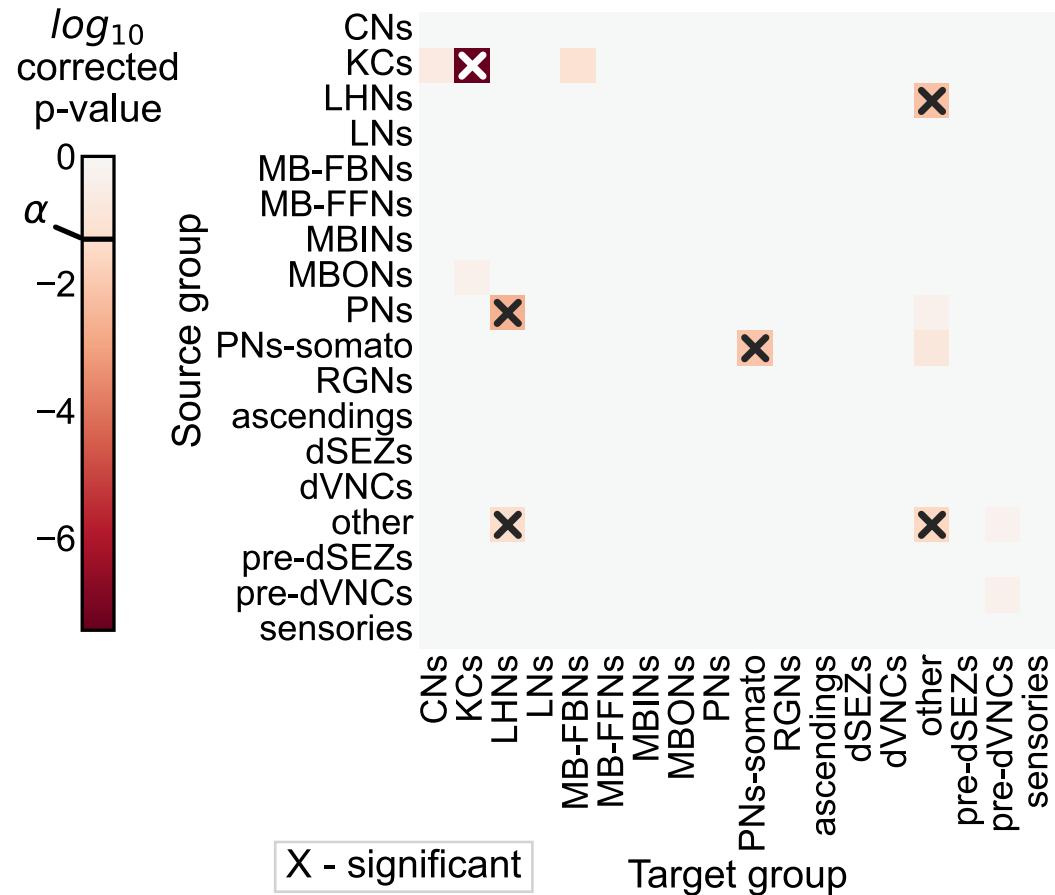
Compare probabilities, compute p-values



Combine p-values for overall test



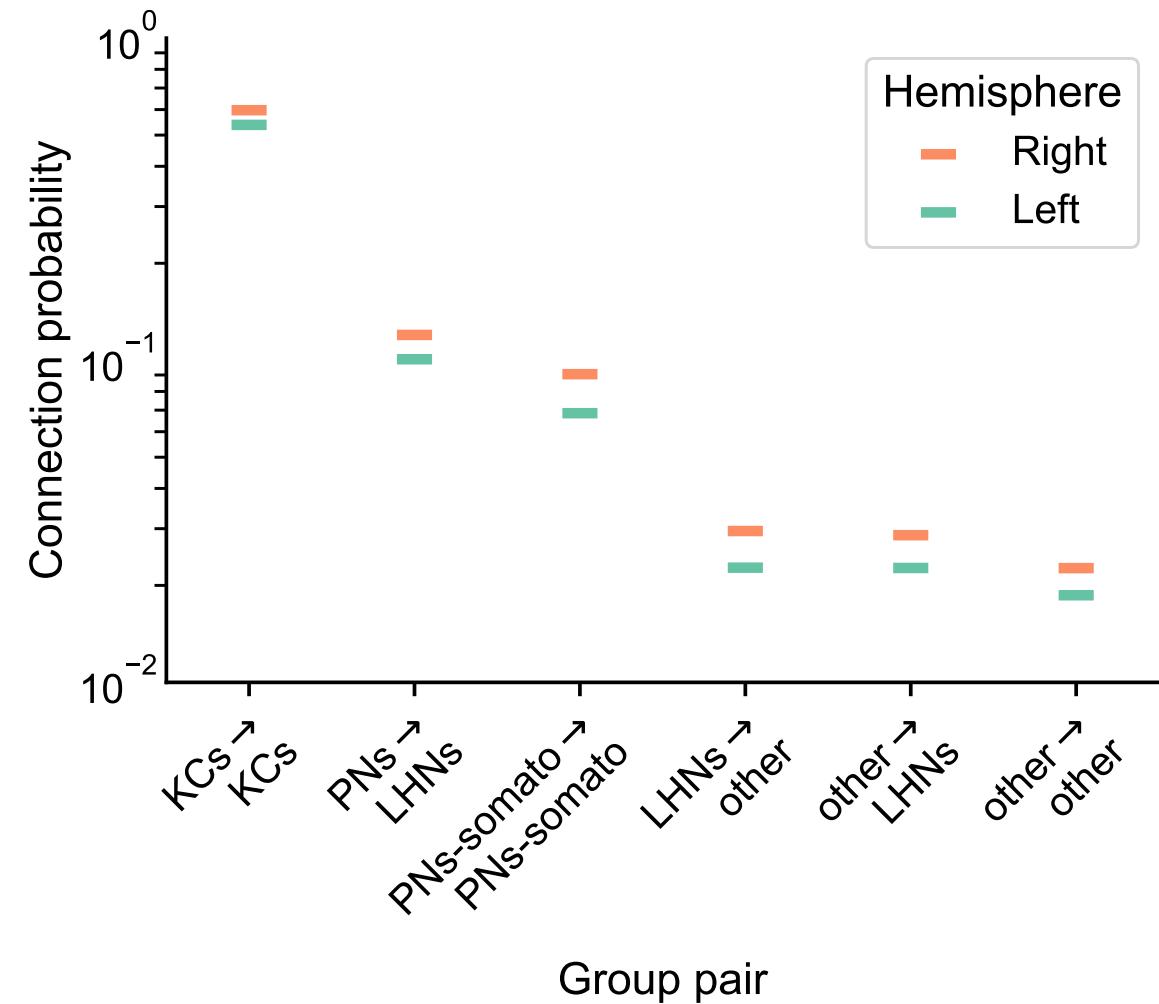
# Detect differences in group connection probabilities



- 5 group-to-group connections are significantly different (after multiple comparisons correction)
- Overall test (comparing all blocks):  
p-value  $< 10^{-7}$

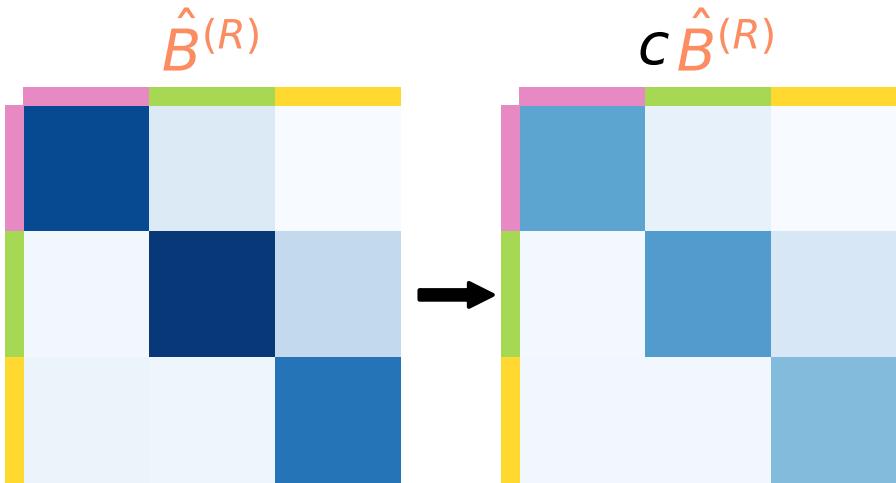
# Should we be surprised?

- Already saw that even the overall densities were different
- For all significant comparisons, probabilities on the right hemisphere were higher
- Maybe the right is just a "scaled up" version of the left?
  - $H_0 : B^{(L)} = cB^{(R)}$  where  $c$  is a density-adjusting constant,  $\frac{p^{(L)}}{p^{(R)}}$

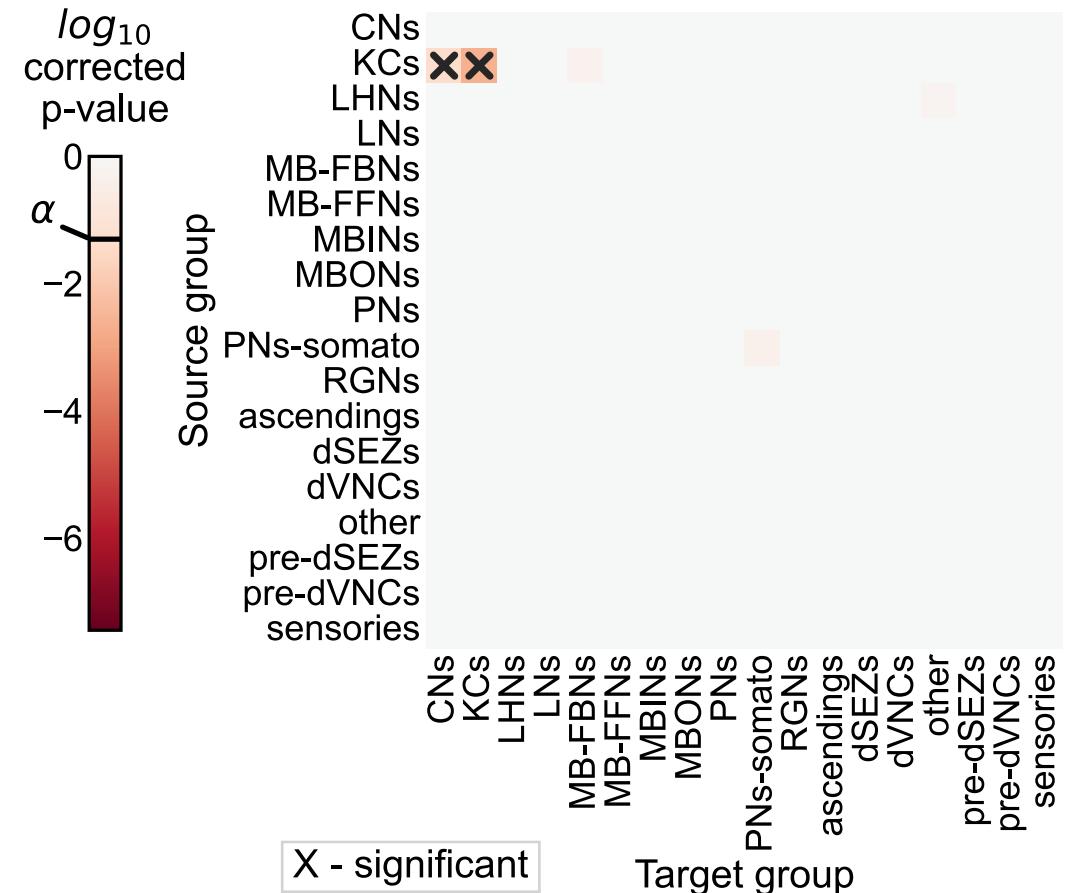


# After adjusting for density, differences are in KCs

Scale connection probabilities  
to match densities



$$\begin{aligned} H_0: B^{(L)} &= CB^{(R)} \\ H_A: B^{(L)} &\neq CB^{(R)} \end{aligned}$$



Overall p-value:  $< 10^{-2}$

# To sum up...

"This brain is bilaterally symmetric."

Depends on what you mean...

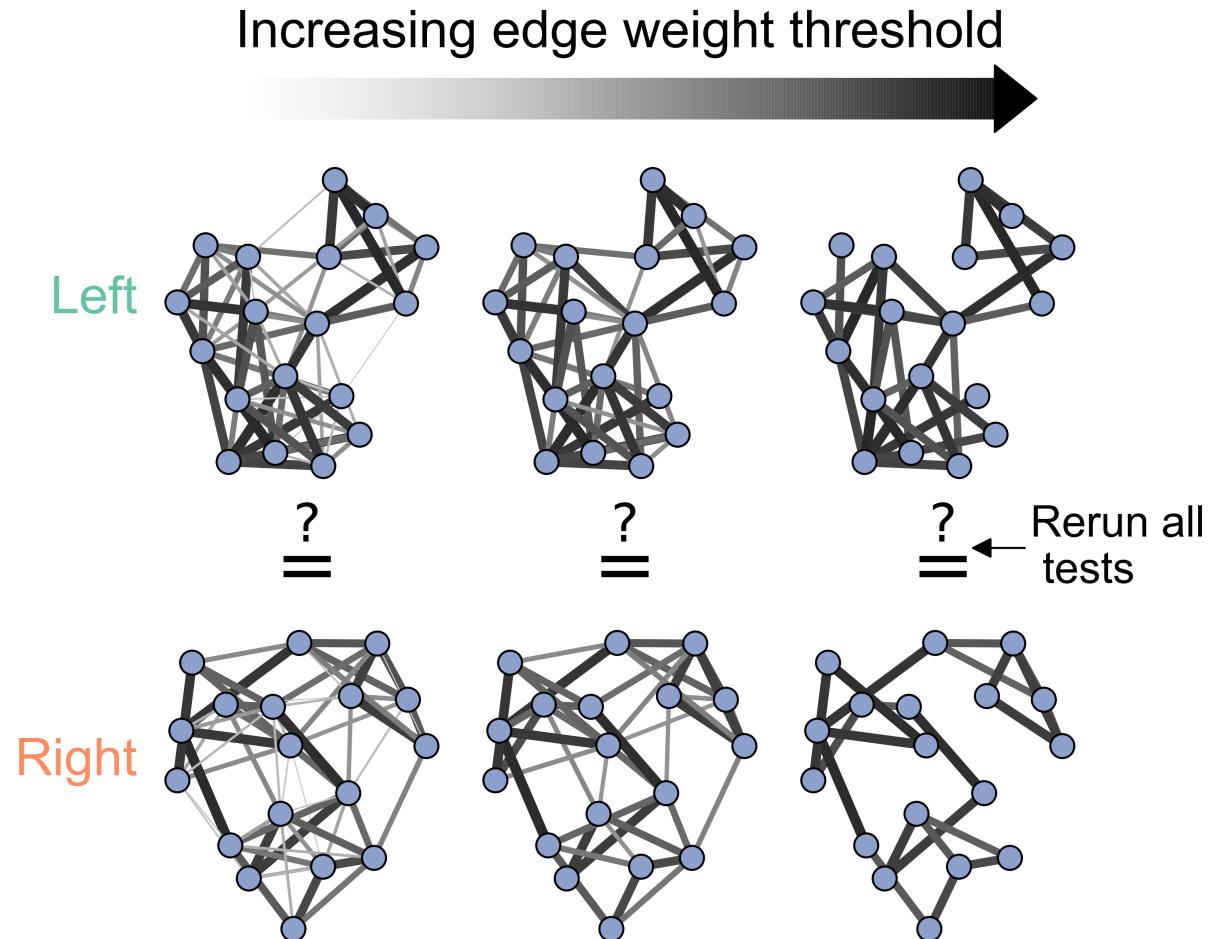
## With Kenyon cells

Model	$H_0$ (vs. $H_A \neq$ )	p-value
ER	$p^{(L)} = p^{(R)}$	$<10^{-23}$
SBM	$B^{(L)} = B^{(R)}$	$<10^{-7}$
daSBM	$B^{(L)} = cB^{(R)}$	$<10^{-2}$

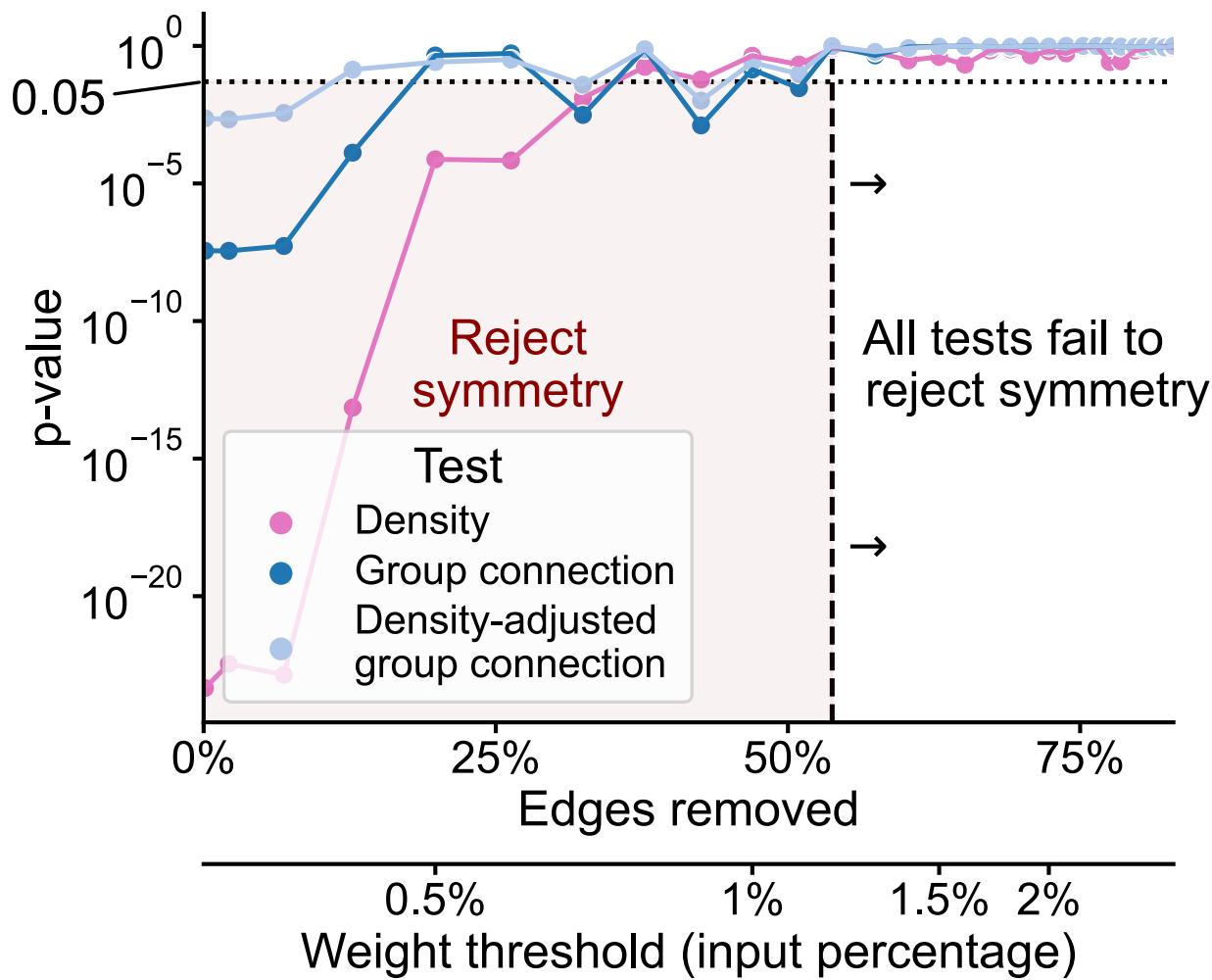
## Without Kenyon cells

Model	$H_0$ (vs. $H_A \neq$ )	p-value
ER	$p^{(L)} = p^{(R)}$	$<10^{-26}$
SBM	$B^{(L)} = B^{(R)}$	$<10^{-2}$
daSBM	$B^{(L)} = cB^{(R)}$	$\approx 0.51$

# Examining the effect of edge weights



# Highest edge weight networks show no asymmetry

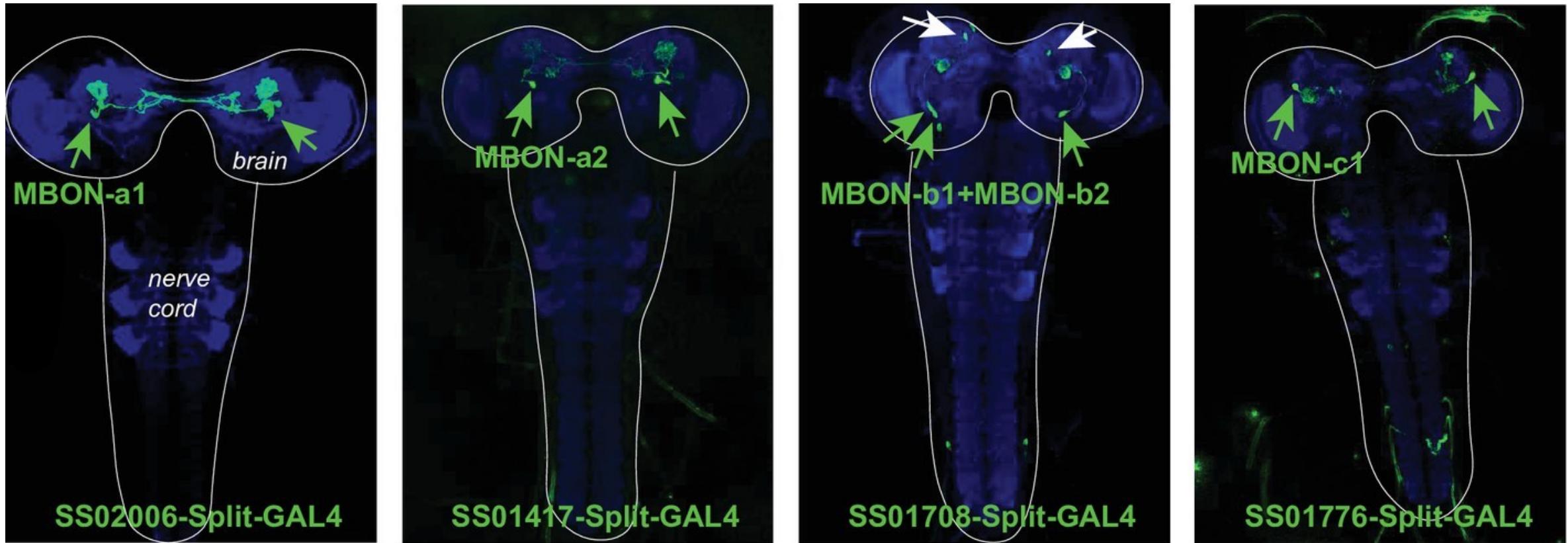


# Outline

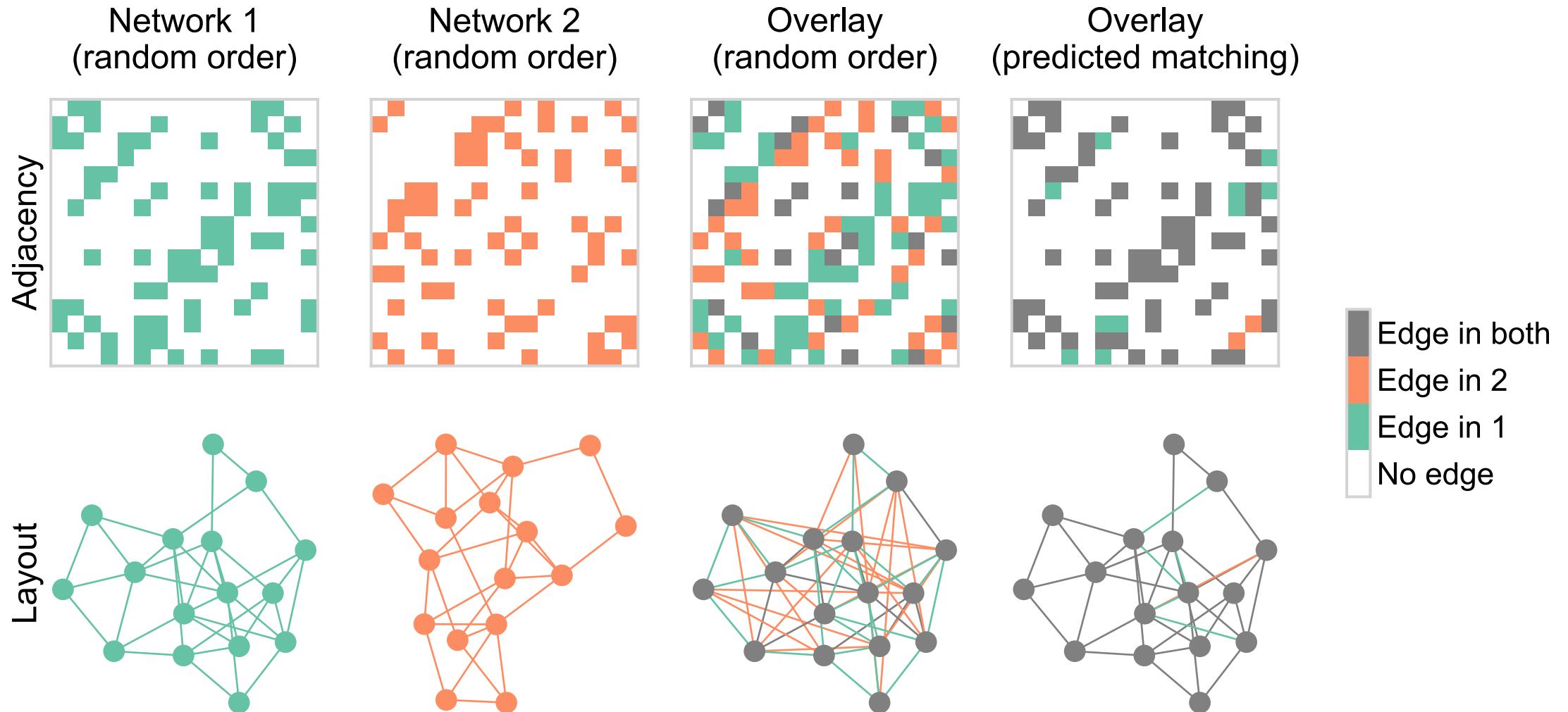
- Larval connectome dataset
- Connectome comparison via network hypothesis testing
- Pairing neurons across connectomes via graph matching
- Ongoing extensions/applications

# Bilaterally homologous neuron pairs

We believe a matching exists!



# What is graph matching?



# How do we measure network overlap?

$$\min_{P \in \mathcal{P}} \underbrace{\|A_1 - \underbrace{PA_2P^T}_{\text{distance between adj. mats.}}\|_F^2}_{\text{reordered } A_2}$$

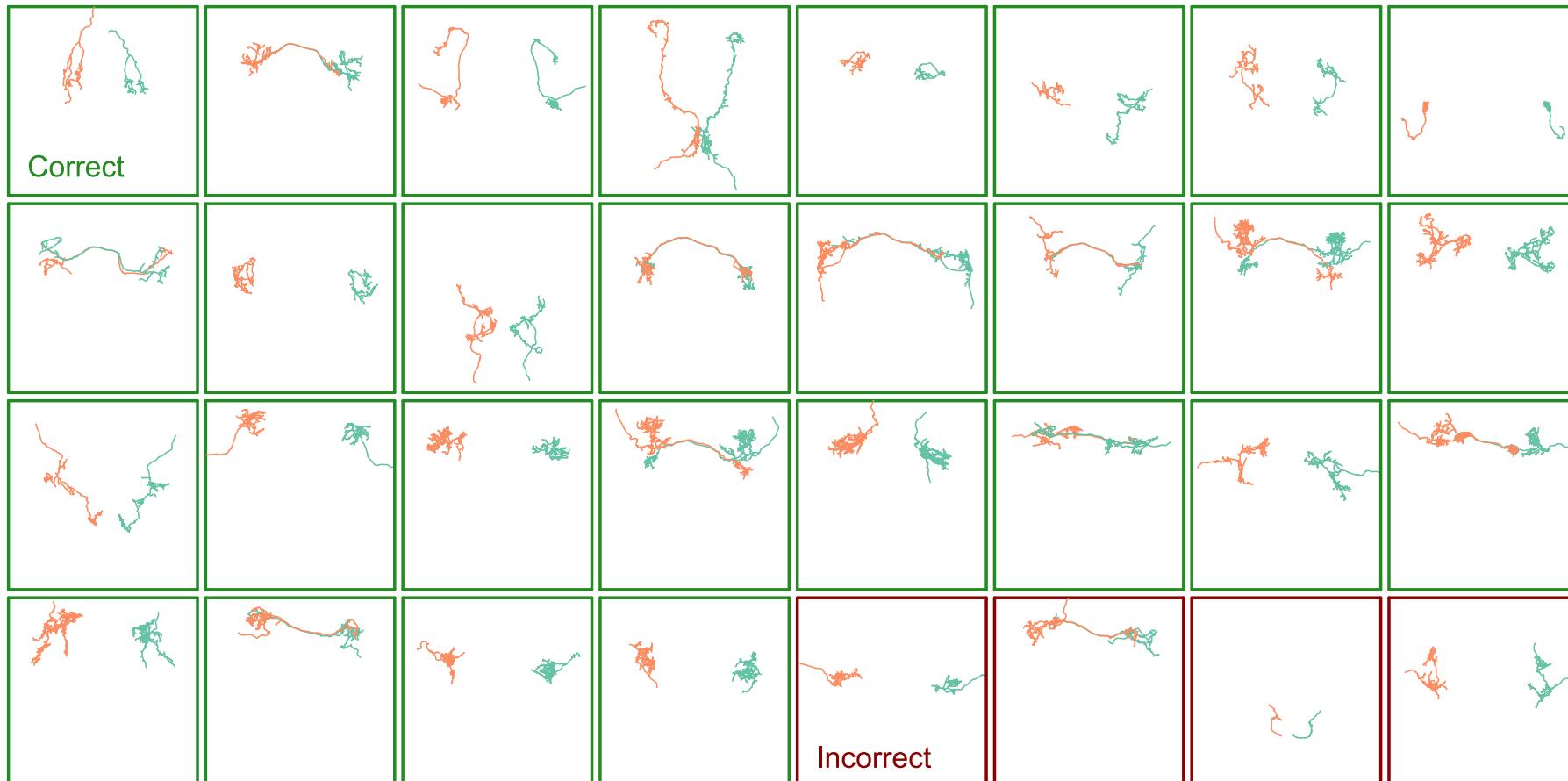
where  $\mathcal{P}$  is the set of permutation matrices

- Measures the number of edge disagreements for unweighted networks,
- Norm of edge disagreements for weighted networks

# How do we do graph matching?

- Relax the problem to a continuous space
  - Convex hull of permutation matrices
- Minimize a linear approximation of objective function (repeat)
- Project back to the closest permutation matrix

# Matching (by connectivity only) performs fairly well



With "vanilla" graph matching: ~80% correct (according to expert annotator)

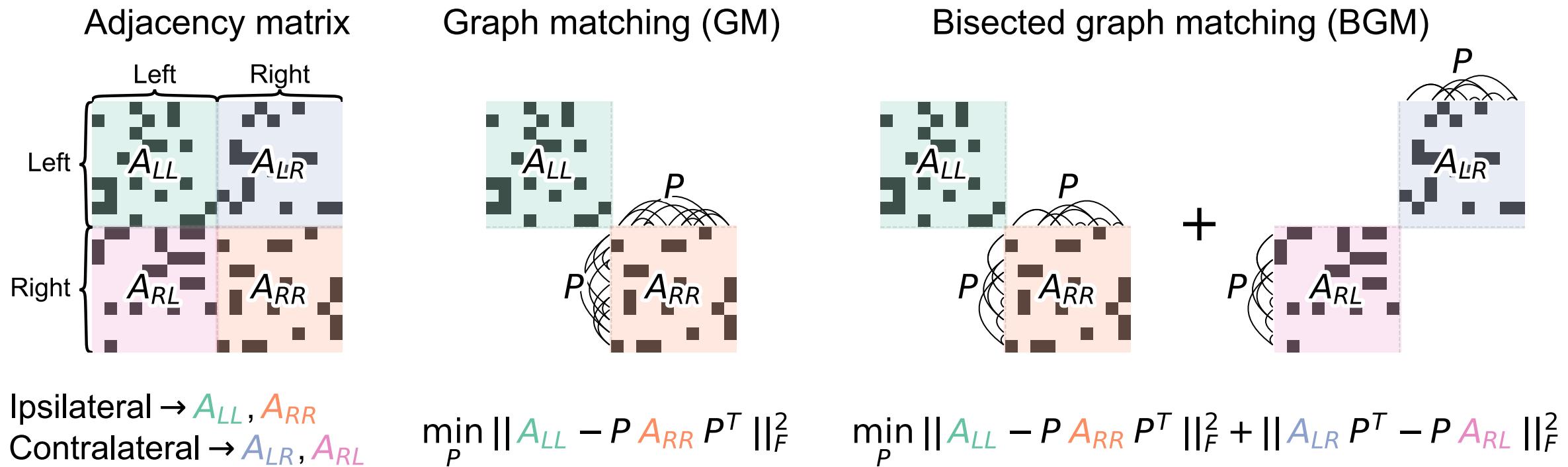
# Many ways to try to improve on this...

- Edge types allow for "multilayer" graph matching
- Partial knowledge of the matching (seeds)
- Morphology (e.g. NBLAST)

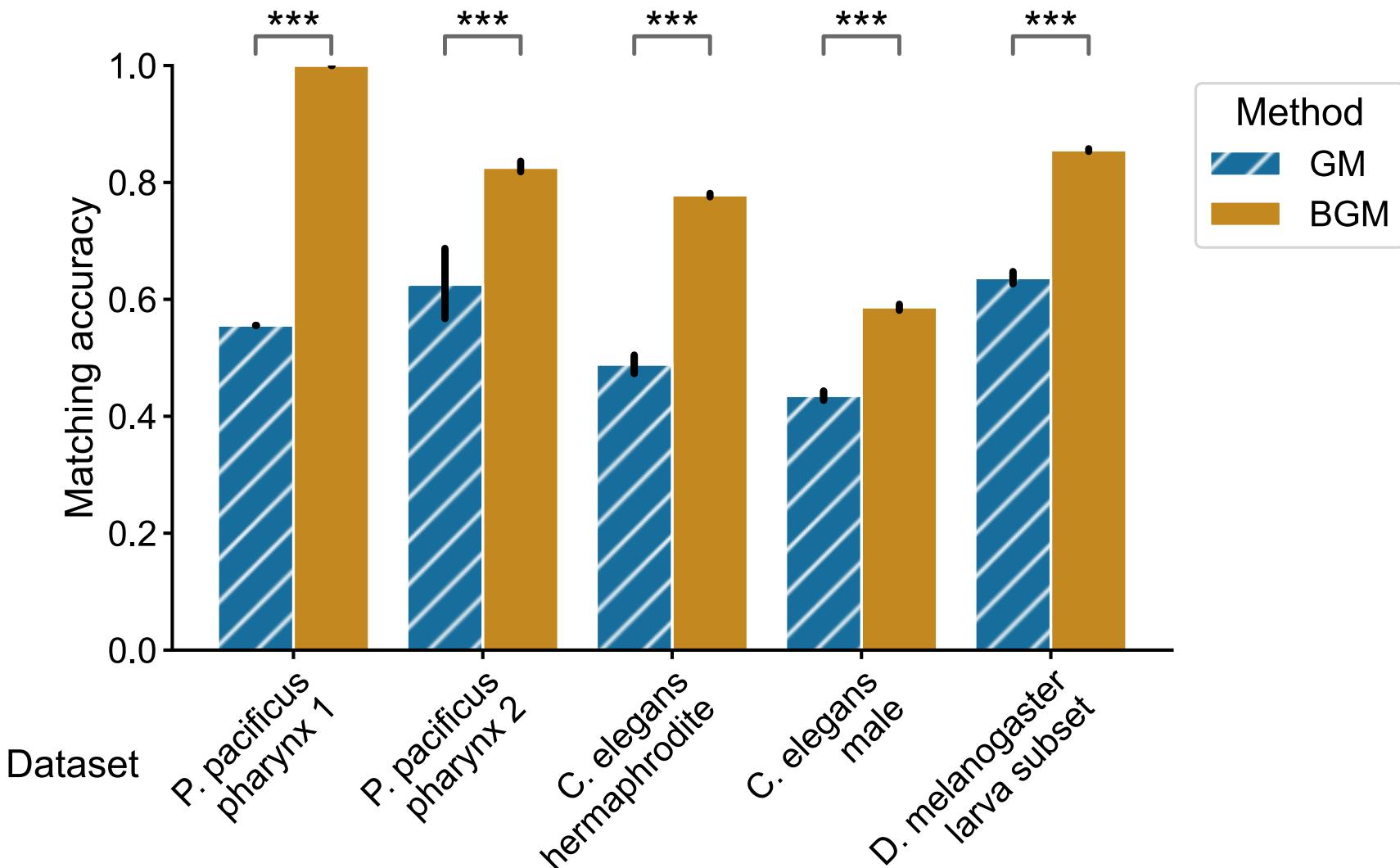
**Thus far, we've not used the contralateral connections**

**These are about 1/3 of the edges in the brain!**

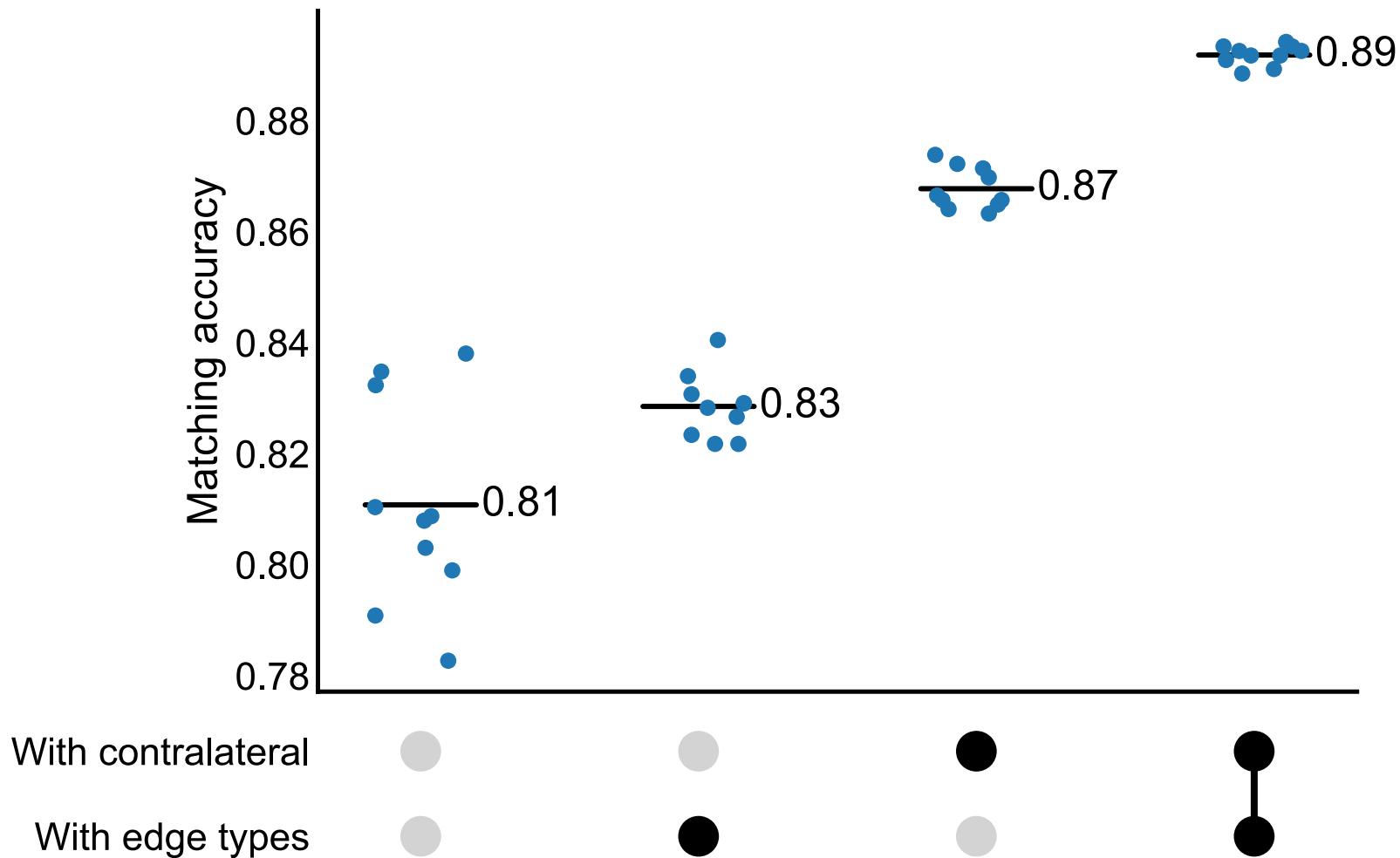
# From graph matching to bisected graph matching



# Contralateral connections are helpful!



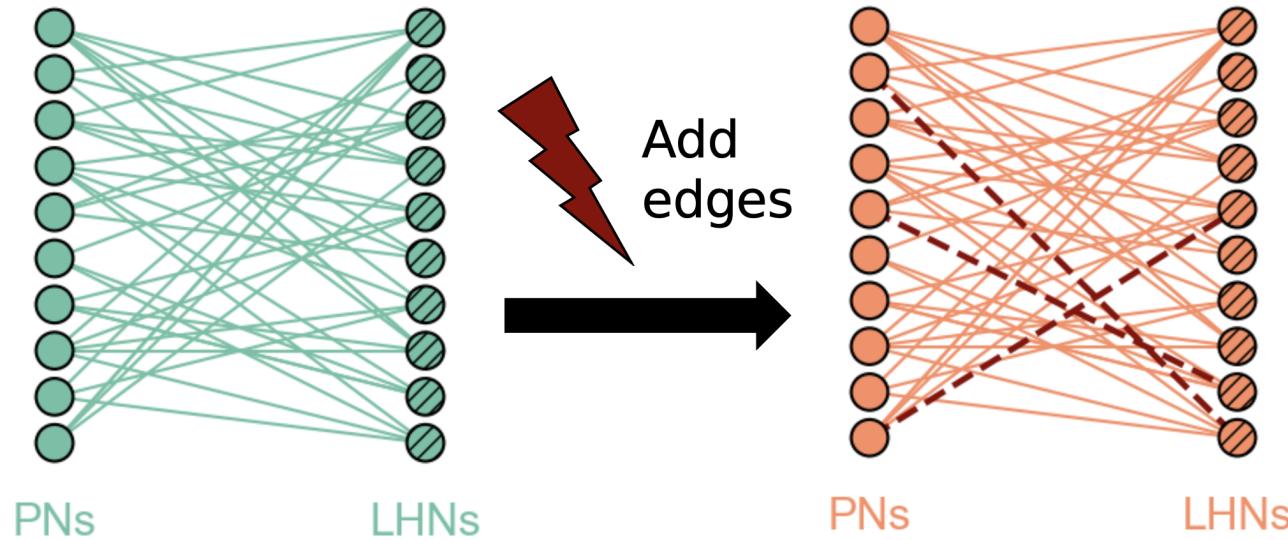
# Performance improvement on the full brain



# Outline

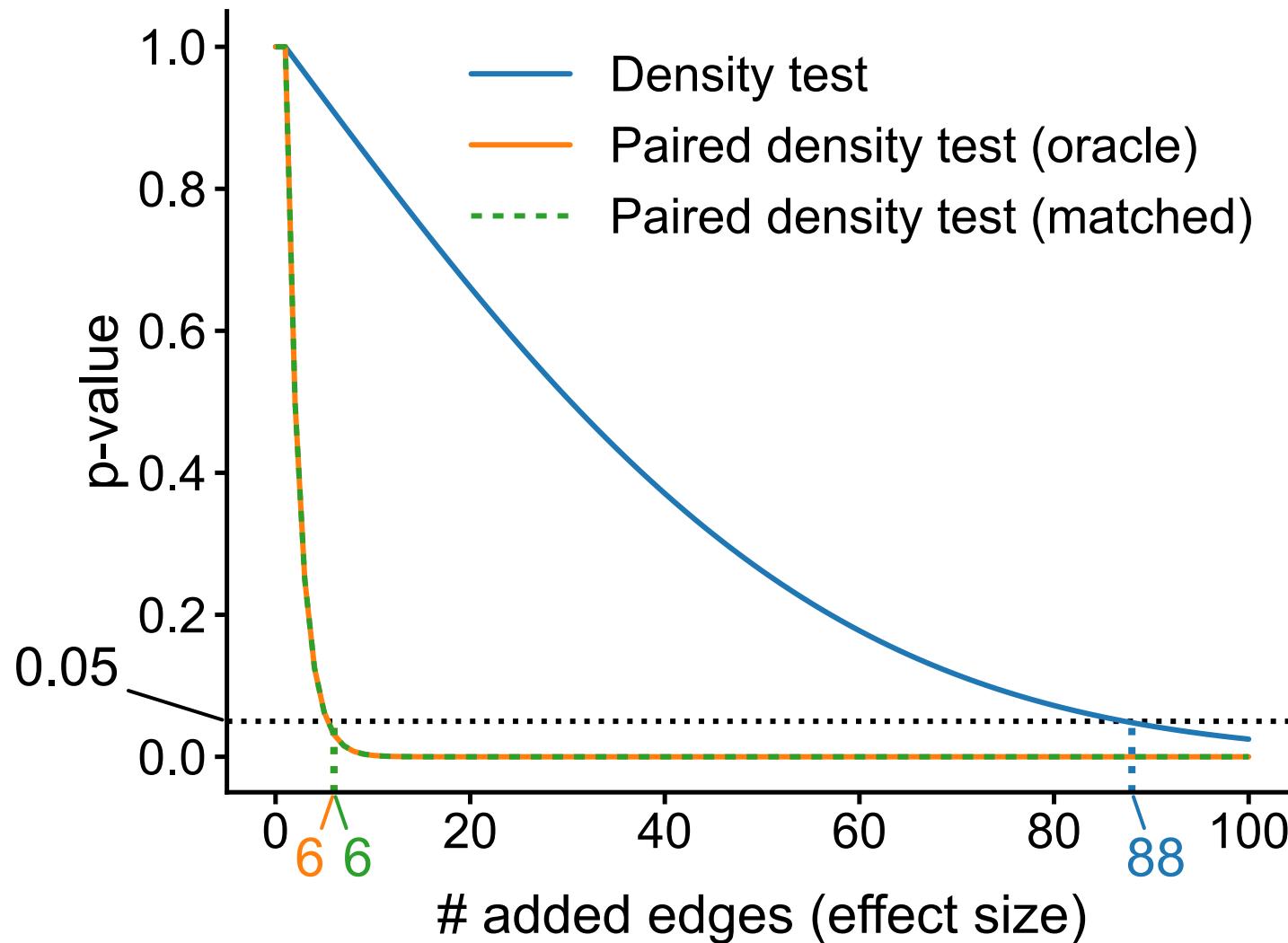
- Larval connectome dataset
- Connectome comparison via network hypothesis testing
- Pairing neurons across connectomes via graph matching
- Ongoing extensions/applications

# A hypothetical difference we want to detect

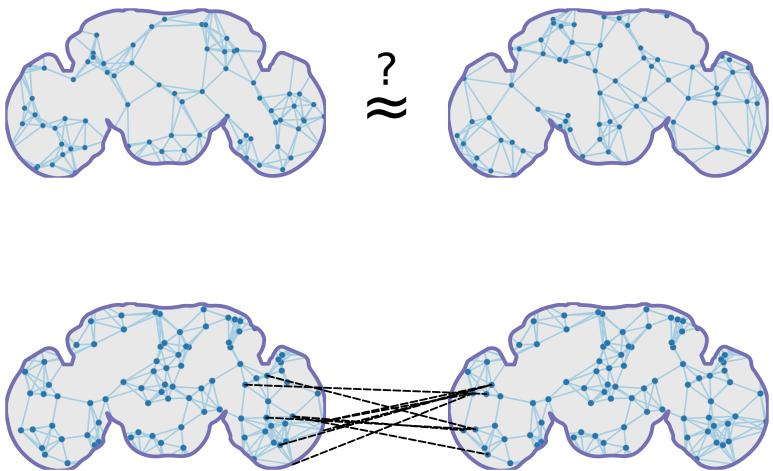


- Start from some subgraph in the connectome,  $A$
- Perturb a copy of it,  $B$  (add edges)
- Test for differences between  $A$  and  $B$

# Pairs facilitate more powerful tests



# Summary



- Model-based network comparison enables testing (and refining) hypotheses about connectomes
  - We proposed a few tests, but just the beginning!
- Graph matching can pair neurons across datasets
  - Helpful to adapt off-the-shelf algos. to use biological info (e.g contralaterals, edge types)

**Aim to apply these (and other) tools to make inferences from connectome comparisons!**

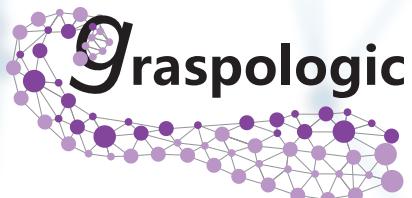
# References

Winding, M. & Pedigo, B.D. et al. The connectome of an insect brain. bioRxiv 2022.11.28.516756 (2022).

Pedigo, B. D. et al. Generative network modeling reveals quantitative definitions of bilateral symmetry exhibited by a whole insect brain connectome. bioRxiv 2022.11.28.518219 (2022).

Pedigo, B. D. et al. Bisected graph matching improves automated pairing of bilaterally homologous neurons from connectomes. Network Neuroscience (2022).

# Code



[github.com/microsoft/graspologic](https://github.com/microsoft/graspologic)

downloads 167k  
Stars 273

[github.com/neurodata/maggot\\_models](https://github.com/neurodata/maggot_models)  
[github.com/neurodata/bilateral-connectome](https://github.com/neurodata/bilateral-connectome)  
[github.com/neurodata/bgm](https://github.com/neurodata/bgm)

# Acknowledgements

## Team



Michael  
Winding



Mike Powell



Eric  
Bridgeford



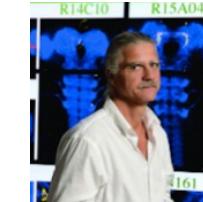
Ali  
Saad-Eldin



Marta Zlatic



Albert  
Cardona



Carey Priebe

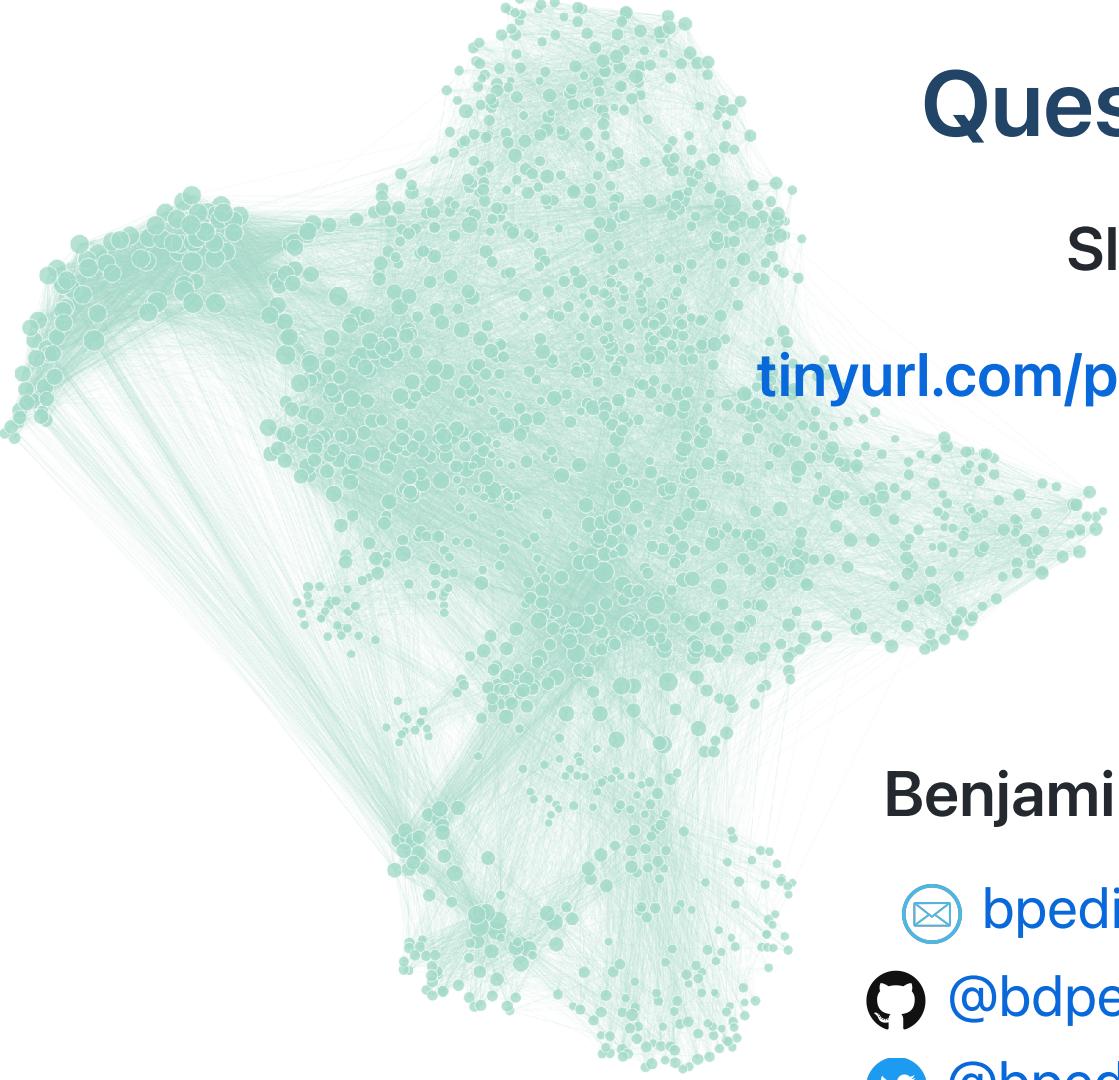


Joshua  
Vogelstein

Tracers who contributed to larva connectome, Heather Patsolic, Youngser Park, NeuroData lab, Microsoft Research  
Figures from Scidraw + Noun Project (Alexander Bates, Xuan Ma, Gil Costa, Vivek Kumar, Leslie Coonrod)

## Funding

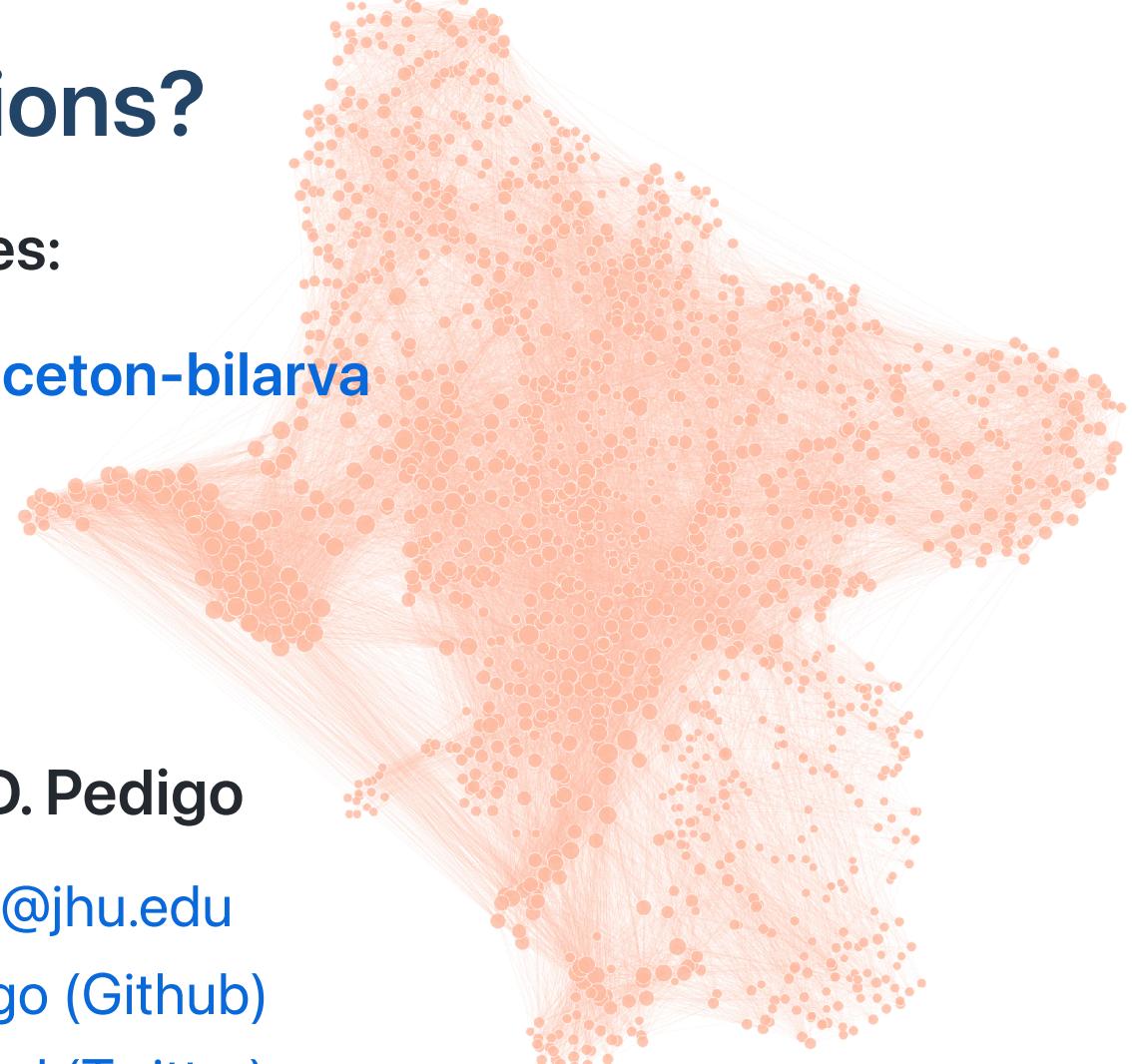
NSF Graduate Research Fellowship (B.D.P.), NSF CAREER Award (J.T.V.), NSF NeuroNex Award (J.T.V and C.E.P.),  
NIH BRAIN Initiative (J.T.V.)



# Questions?

Slides:

[tinyurl.com/princeton-bilarva](http://tinyurl.com/princeton-bilarva)



**Benjamin D. Pedigo**

-  [bpedigo@jhu.edu](mailto:bpedigo@jhu.edu)
-  [@bdpedigo \(Github\)](https://github.com/bdpedigo)
-  [@bpedigod \(Twitter\)](https://twitter.com/bpedigod)
-  [bdpedigo.github.io](https://bdpedigo.github.io)

Left

Right