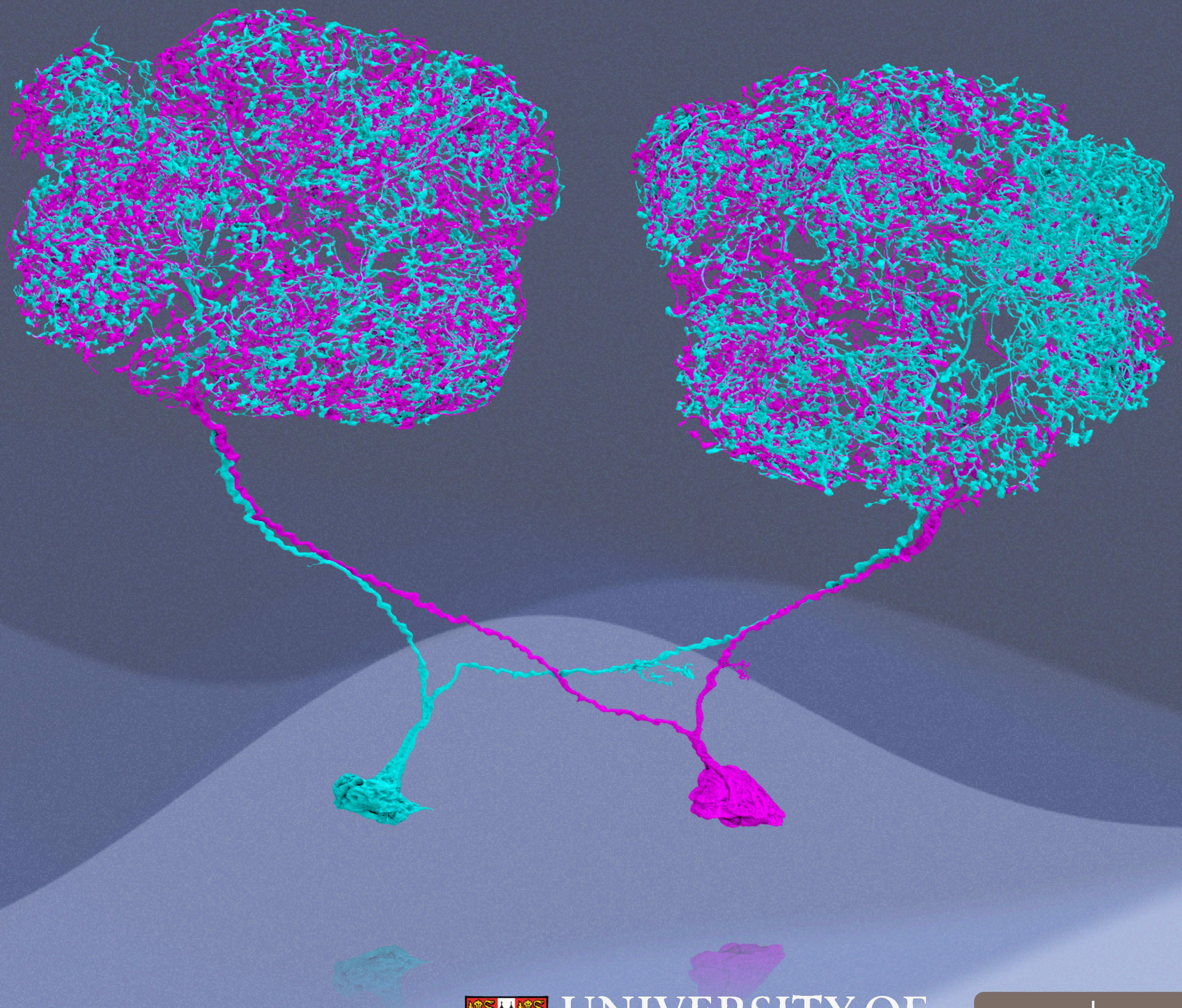
Insights from Comparative Connectomics in *Drosophila*



Philipp Schlegel, Jefferis Lab, University of Cambridge

IPAM, UCLA Workshop: Mathematical Approaches for Connectomics Analysis, 2024

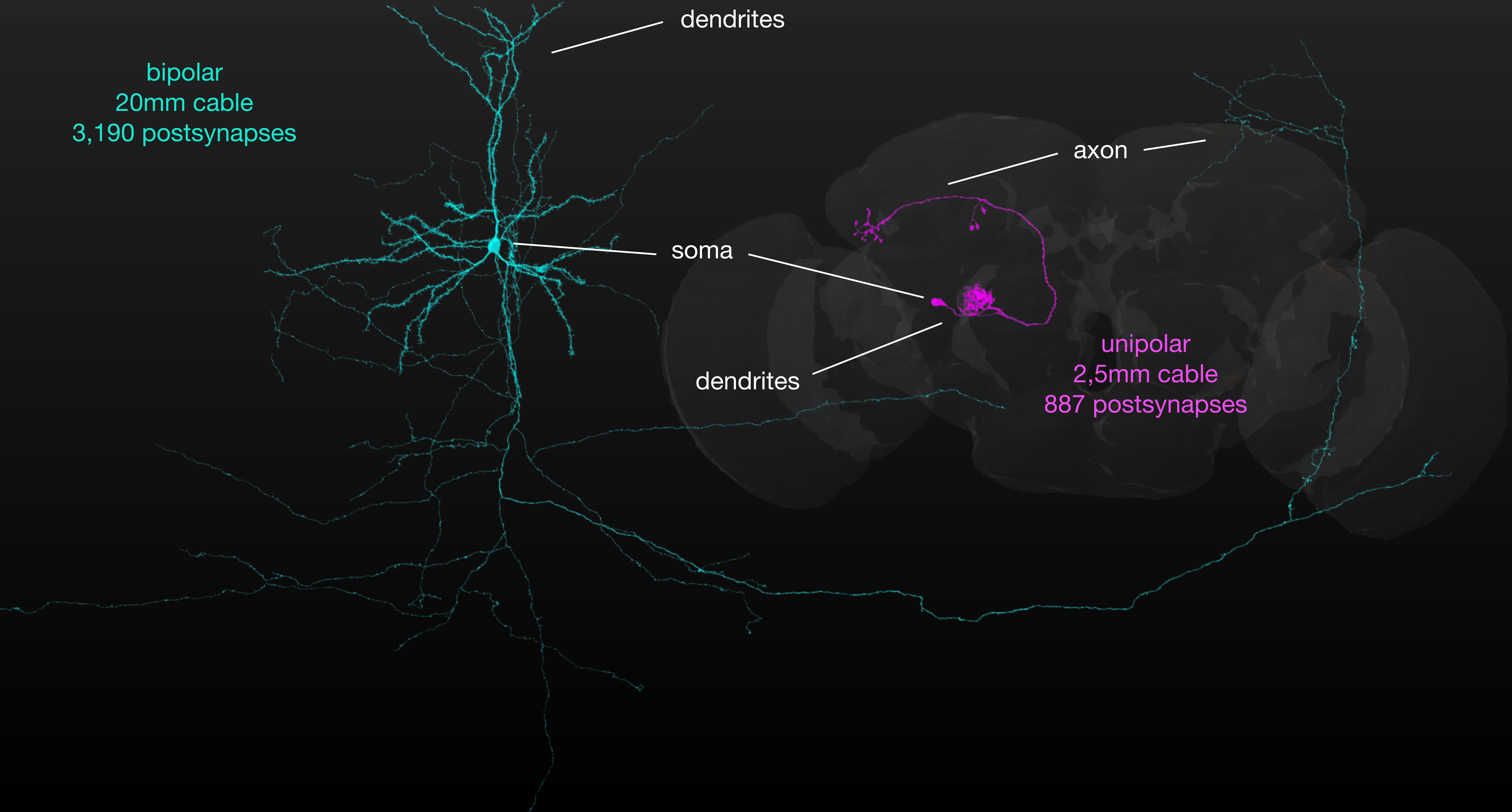
UNIVERSITY OF CAMBRIDGE

MRC Laboratory of Molecular Biology

Overview

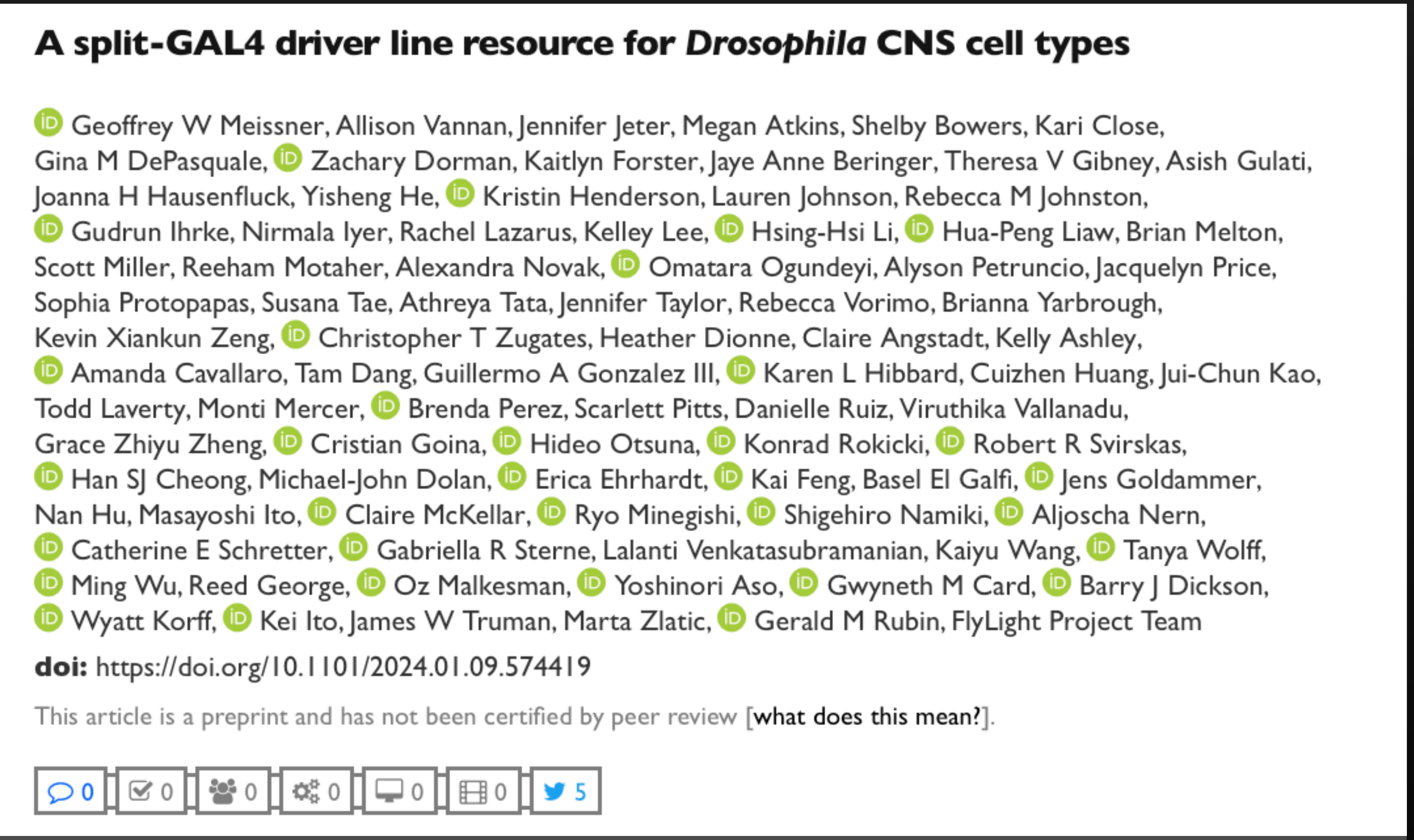
1. The new & shiny FlyWire connectome
2. Lessons learned from comparing two connectomes
3. What is a cell type and why does that matter?


central brain
cortex
neuropil
optic lobes
ventral nerve cord
eye
ommatidium
0.5 mm


bipolar
20mm cable
3,190 postsynapses
dendrites
soma
axon
dendrites
unipolar
2,5mm cable
887 postsynapses


navigation
(e.g. Jayaraman)
learning & memory
(e.g. Rubin, Waddell)
olfactory processing
& innate behaviour
(e.g. Jefferis, Wilson)
visual processing
(e.g. Card, Kim, Turaga, Reiser)
motor control &
motor sequences
(e.g. Card, Lee, Simpson, Tuthill, Dickinson)
• sexually dimorphic behaviour
• decision making
• sensory integration

3,063 genetic driver lines
targeting specific neuronal cell types

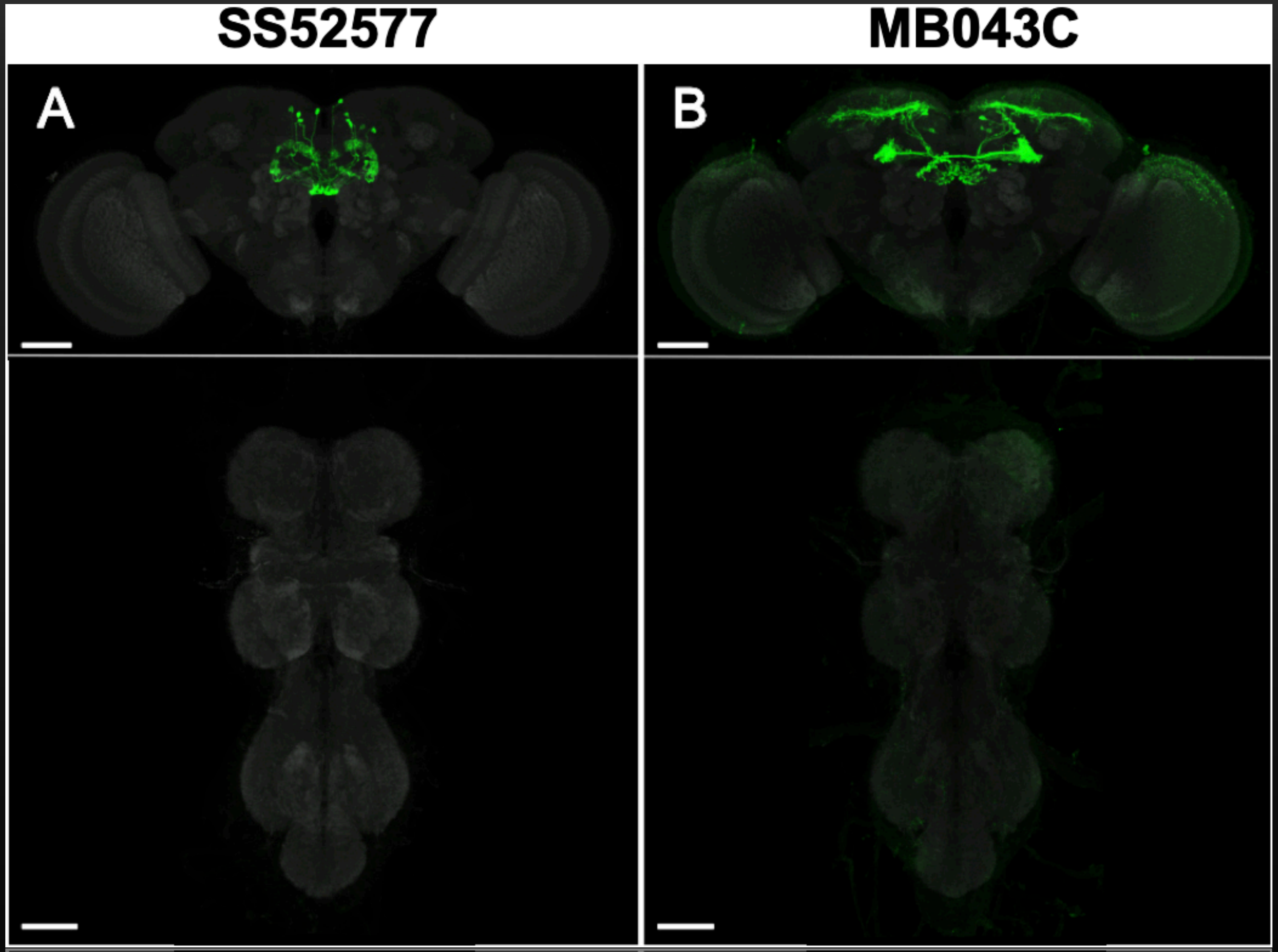


A split-GAL4 driver line resource for *Drosophila* CNS cell types

Geoffrey W Meissner, Allison Vannan, Jennifer Jeter, Megan Atkins, Shelby Bowers, Kari Close, Gina M DePasquale, Zachary Dorman, Kaitlyn Forster, Jaye Anne Beringer, Theresa V Gibney, Asish Gulati, Joanna H Hausenfluck, Yisheng He, Kristin Henderson, Lauren Johnson, Rebecca M Johnston, Gudrun Ihrke, Nirmala Iyer, Rachel Lazarus, Kelley Lee, Hsing-Hsi Li, Hua-Peng Liaw, Brian Melton, Scott Miller, Reeham Motaher, Alexandra Novak, Omatara Ogundeyi, Alyson Petruncio, Jacquelyn Price, Sophia Protopapas, Susana Tae, Athreya Tata, Jennifer Taylor, Rebecca Vorimo, Brianna Yarbrough, Kevin Xiankun Zeng, Christopher T Zugates, Heather Dionne, Claire Angstadt, Kelly Ashley, Amanda Cavallaro, Tam Dang, Guillermo A Gonzalez III, Karen L Hibbard, Cuizhen Huang, Jui-Chun Kao, Todd Laverty, Monti Mercer, Brenda Perez, Scarlett Pitts, Danielle Ruiz, Viruthika Vallanadu, Grace Zhiyu Zheng, Cristian Goina, Hideo Otsuna, Konrad Rokicki, Robert R Svirskas, Han SJ Cheong, Michael-John Dolan, Erica Ehrhardt, Kai Feng, Basel El Galfi, Jens Goldammer, Nan Hu, Masayoshi Ito, Claire McKellar, Ryo Minegishi, Shigehiro Namiki, Aljoscha Nern, Catherine E Schretter, Gabriella R Sterne, Lalanti Venkatasubramanian, Kaiyu Wang, Tanya Wolff, Ming Wu, Reed George, Oz Malkesman, Yoshinori Aso, Gwyneth M Card, Barry J Dickson, Wyatt Korff, Kei Ito, James W Truman, Marta Zlatic, Gerald M Rubin, FlyLight Project Team

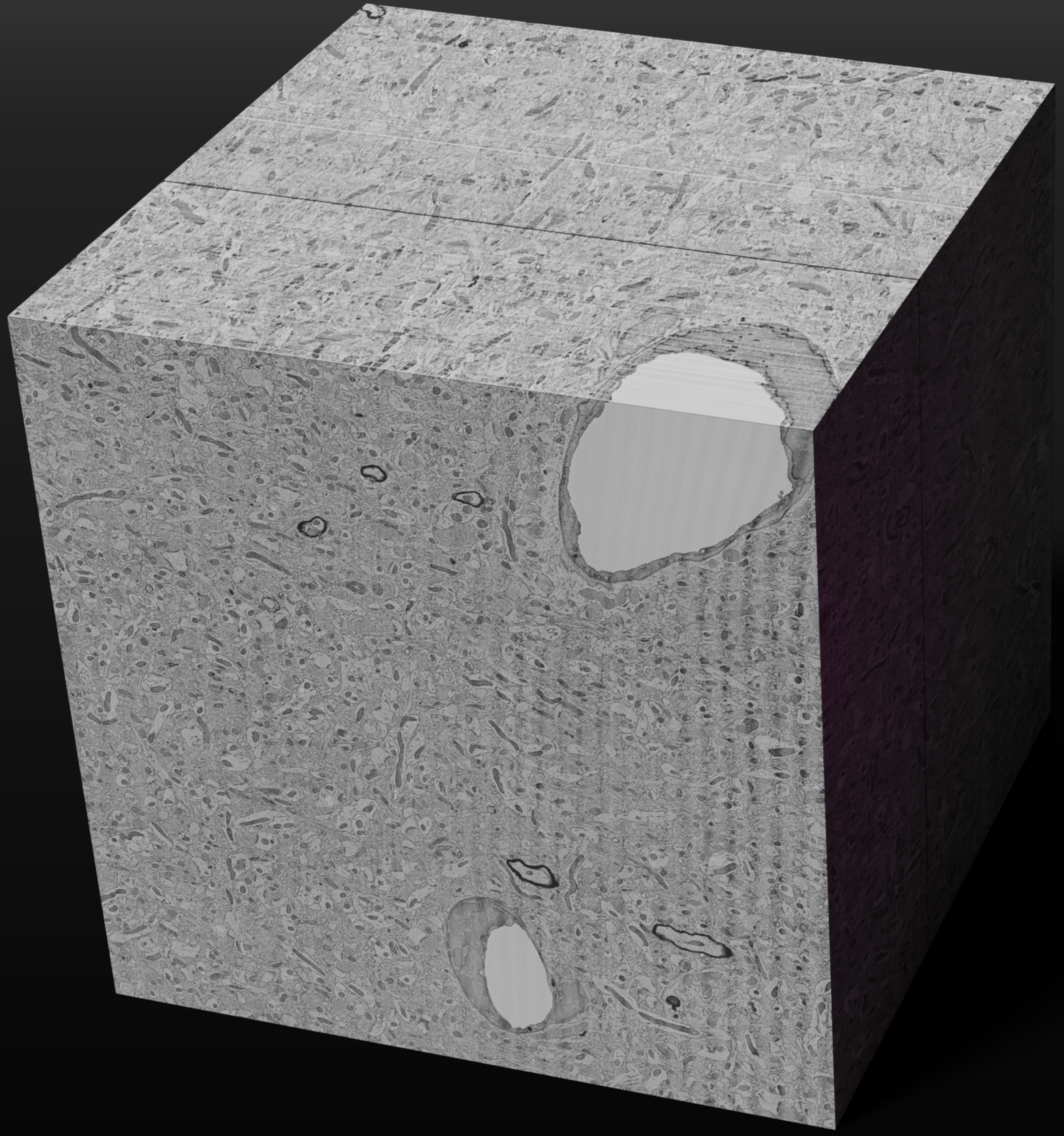
doi: https://doi.org/10.1101/2024.01.09.574419

This article is a preprint and has not been certified by peer review [what does this mean?].

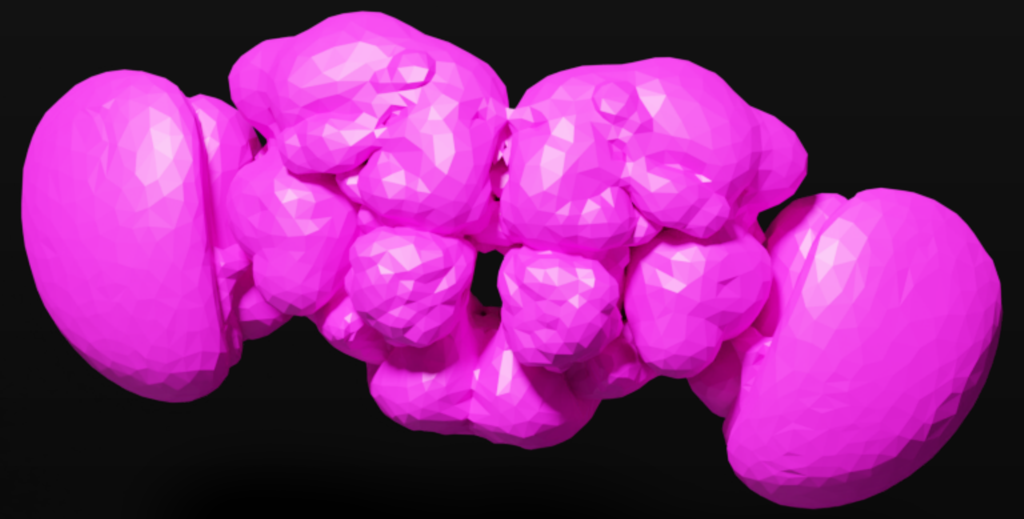


brain

ventral nerve cord

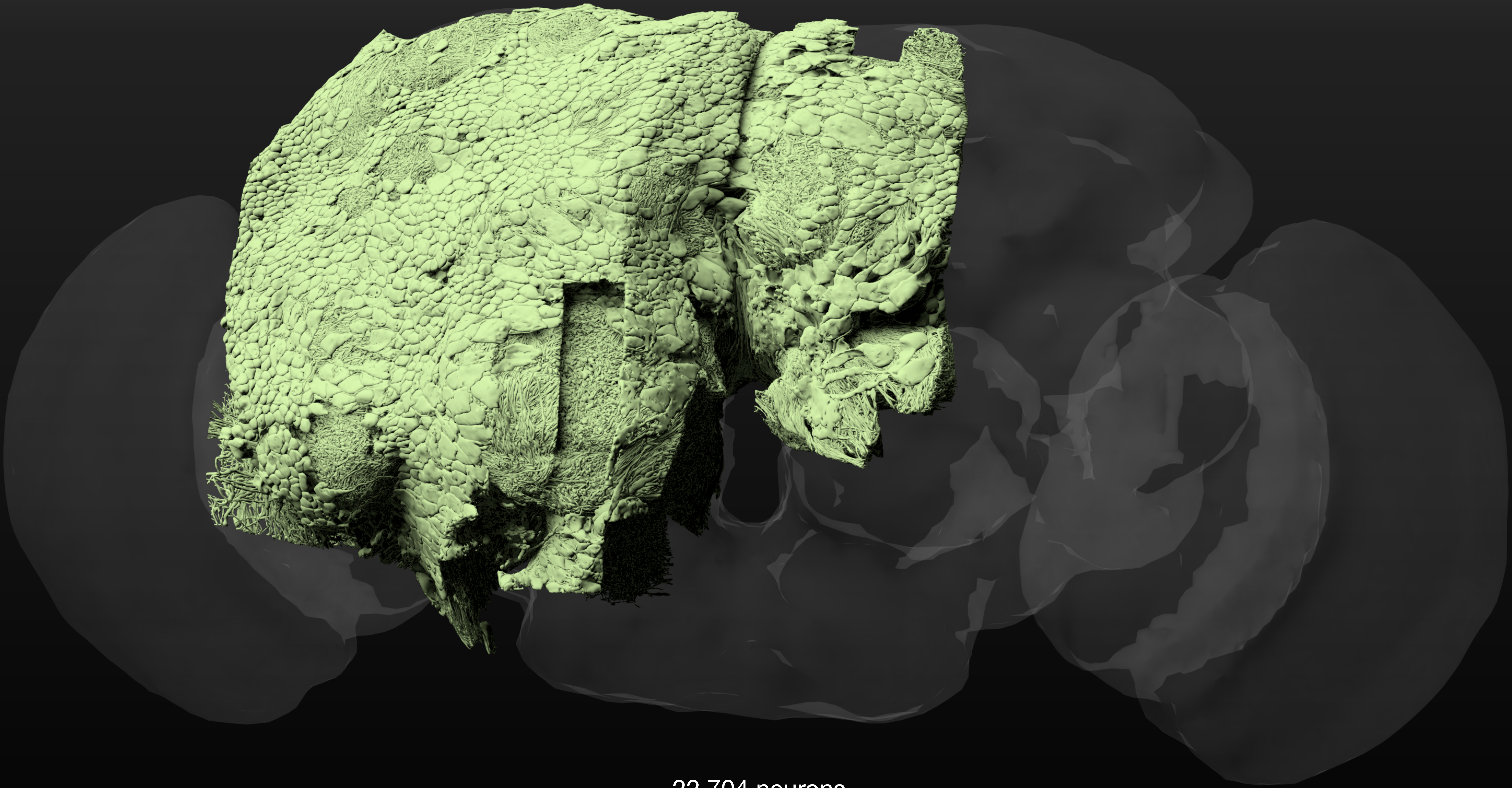


cortical mm³
*artistic rendering



fly brain

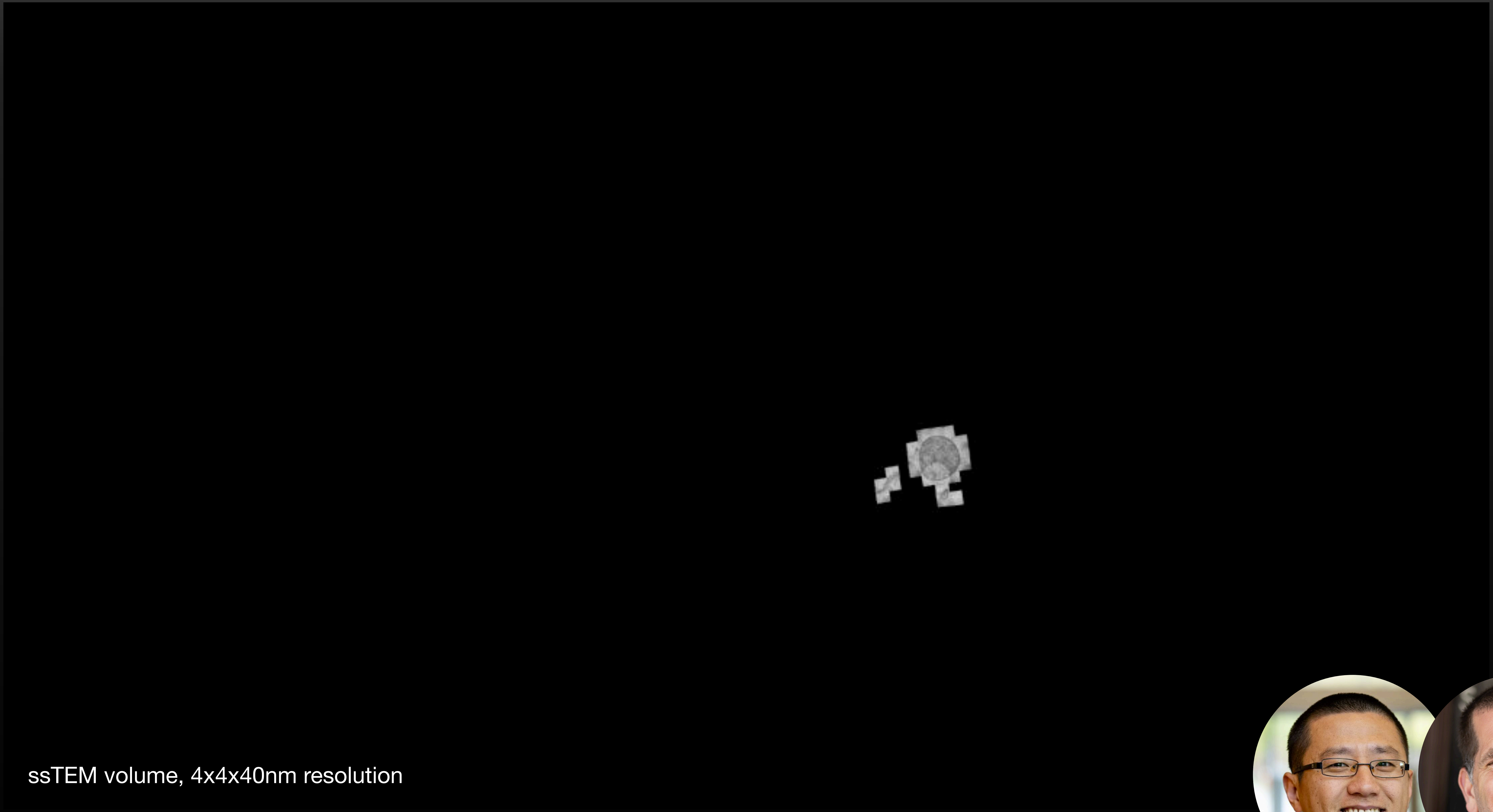
The Janelia "hemibrain" connectome



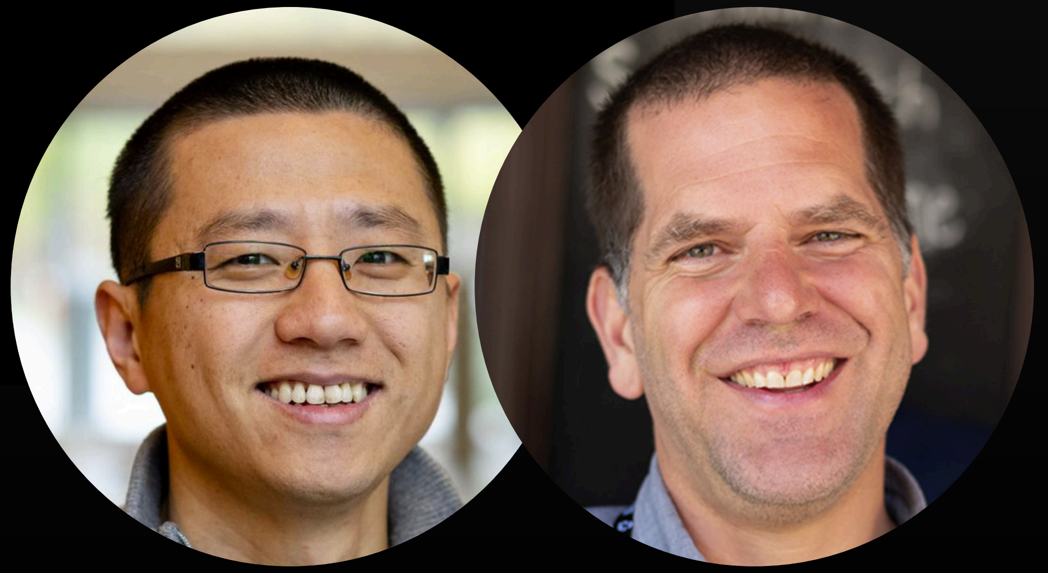
22,704 neurons
5,235 cell types

Scheffer *et al.*, eLife (2020)

FAFB, “Full Adult Fly Brain” image volume



ssTEM volume, 4x4x40nm resolution



Z. Zheng D. Bock

Zheng, Lauritzen *et al.*, Nature Methods (2018)

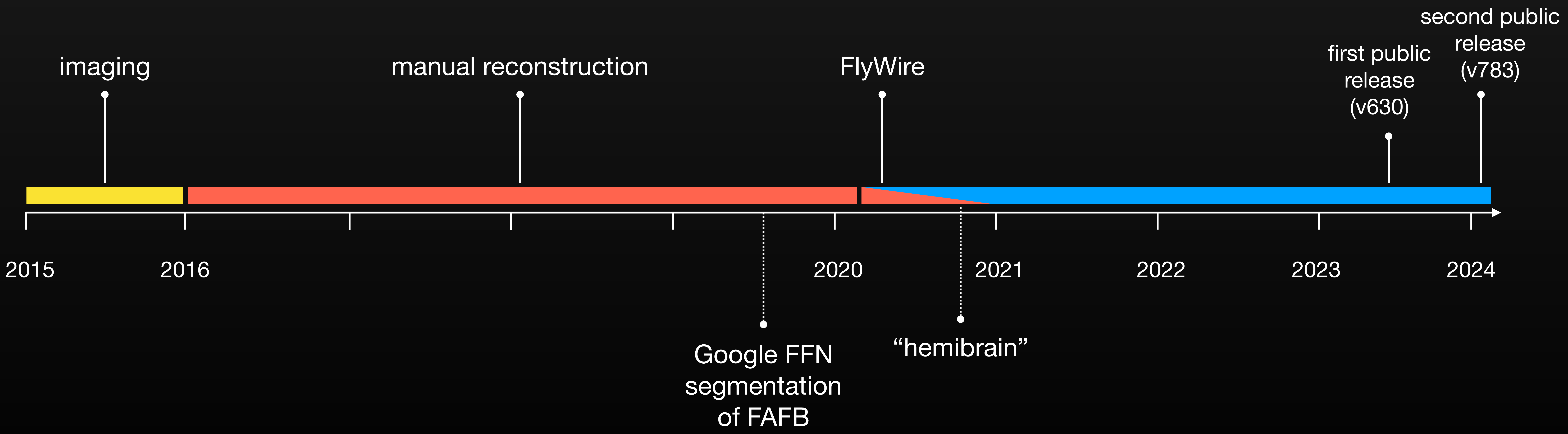
FlyWire

Collaboration between Princeton, Cambridge and others

Z. Zheng D. Bock

M. Murthy S. Seung S. Dorkenwald G. Jefferis

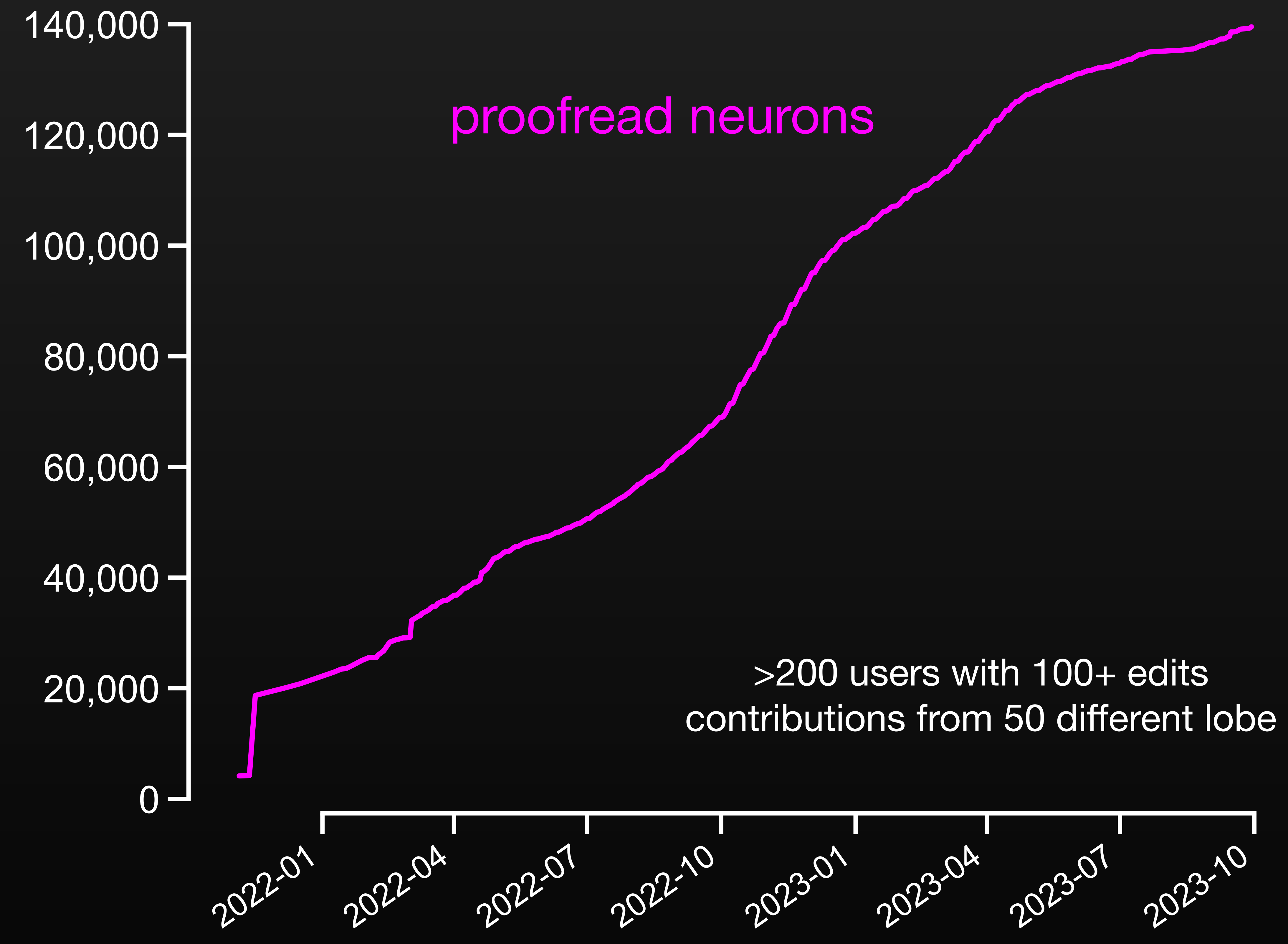
+ FlyWire community

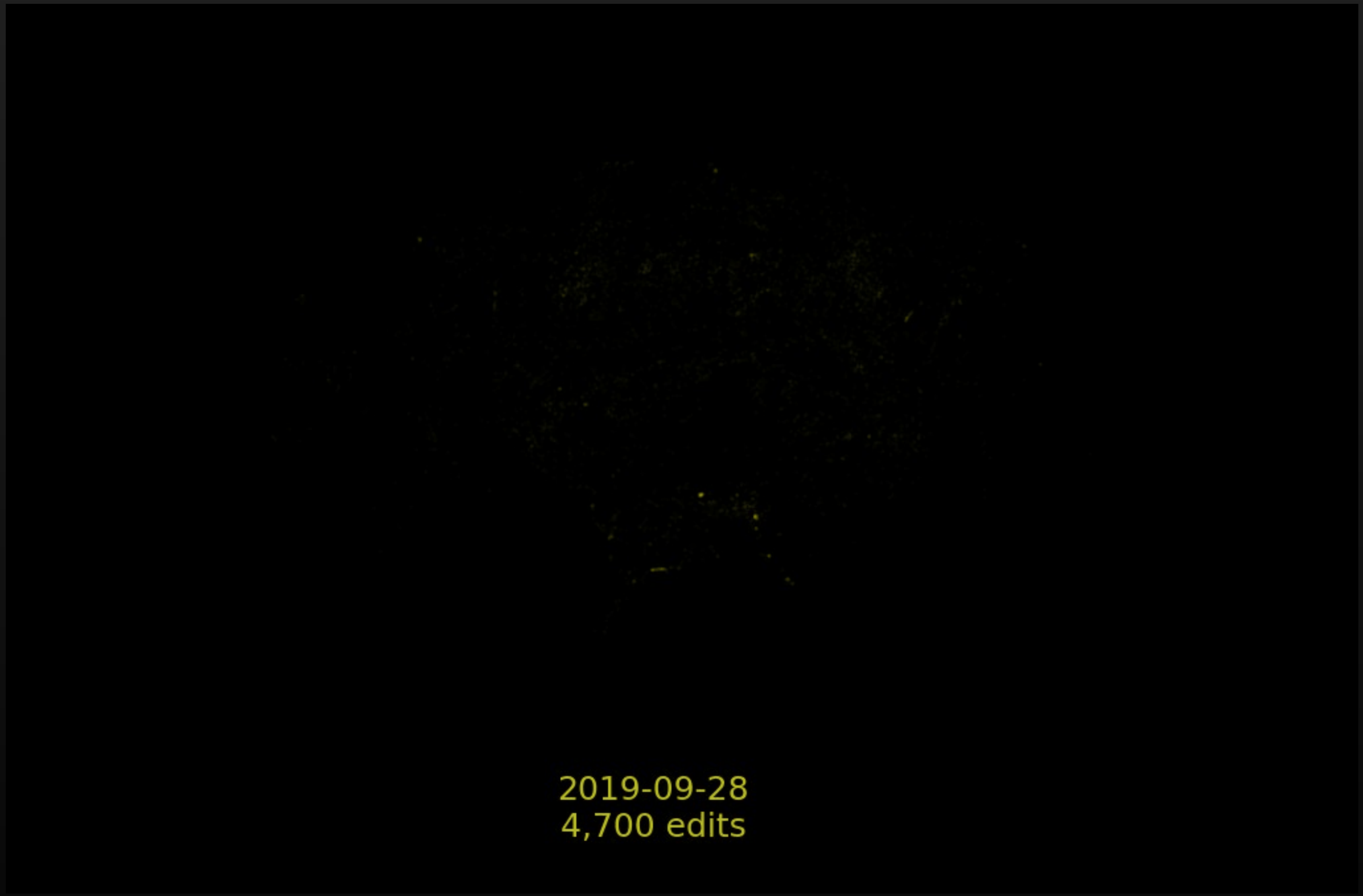


Dorkenwald et al., Nature Methods (2021)
Popovych et al., Nature Comm. (2024)

FlyWire

Collaboration between Princeton, Cambridge and others





30 person-years

FlyWire

Collaboration between Princeton, Cambridge and others

3.1M total edits

estimated 30 person-years
(hemibrain: 50 person-years)

0.02mm^3
(1/25,000 of a mouse brain)

139,000 neurons

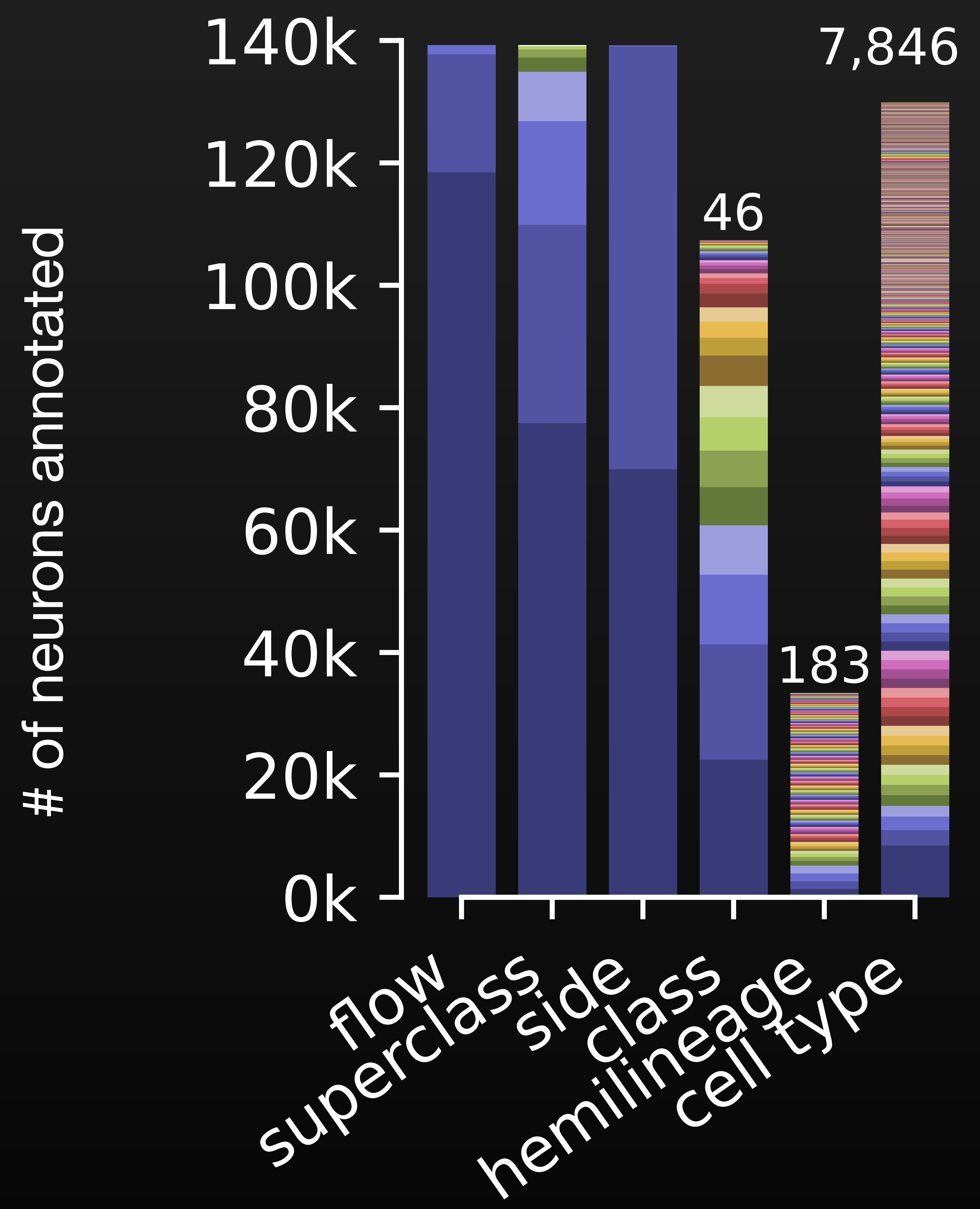
150m of cable

130,000,000 synapses

Schlegel et al. (2023)
Dorkenwald et al. (2023)
FlyWire.ai

FlyWire

Collaboration between Princeton, Cambridge and others



860,000 annotations in total
(hemibrain: ~64,000)

estimated 3 person years

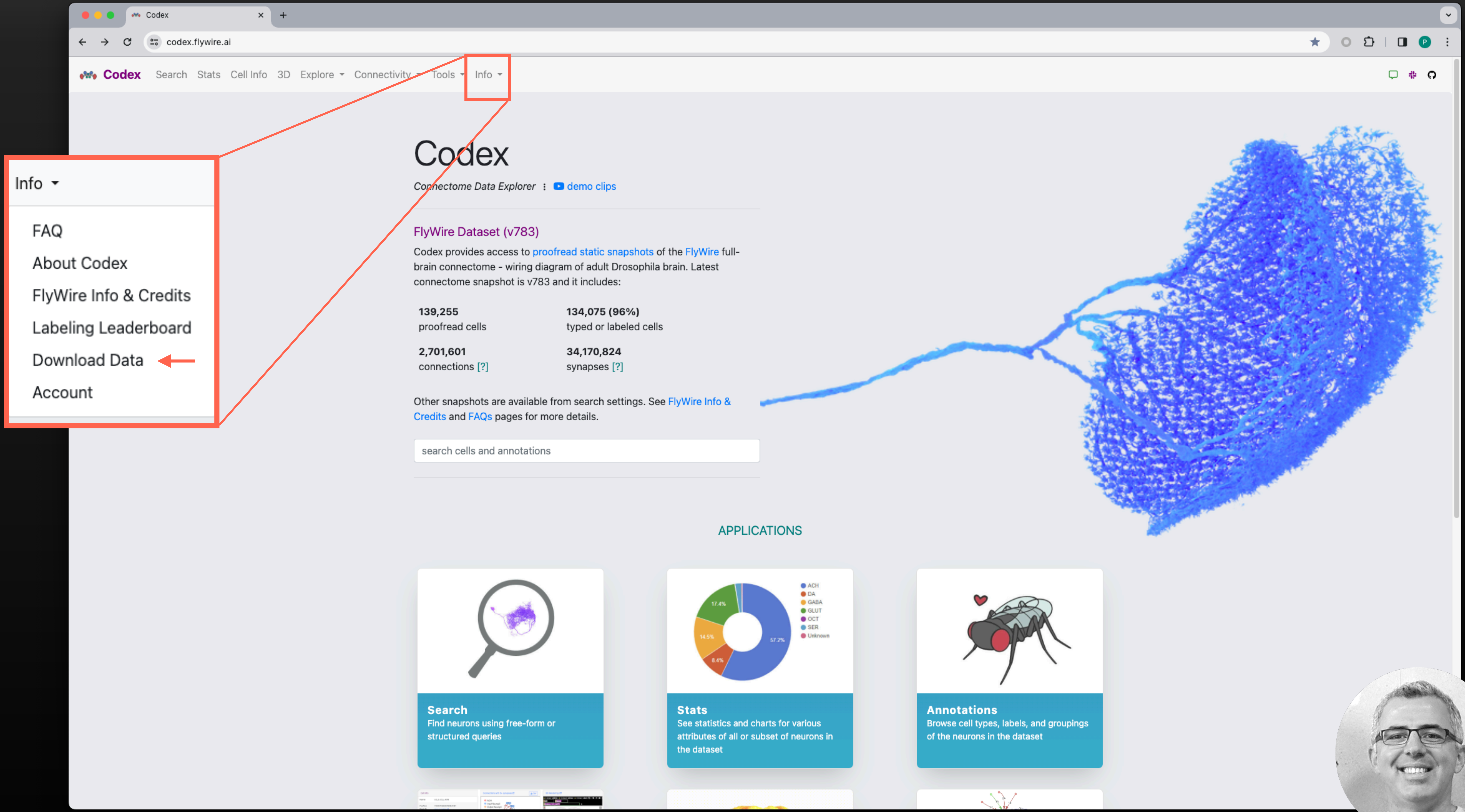
FlyWire

Collaboration between Princeton, Cambridge and others

Schlegel et al. (2023)
Dorkenwald et al. (2023)
FlyWire.ai

Where can I get the data?

codex.flywire.ai



Arie Matsliah

R

```
> library(fafbseg)

> dl4df = flytable_meta('DL4.*')
> dl4_p = flywire_partner_summary(dl4df)
```

Python

```
>>> from fafbseg import flywire
>>> NC = flywire.NeuronCriteria

>>> dl4_p = flywire.get_connectivity(
...             NC(type='DL4_adPN')
...             )
```

https://natverse.org/fafbseg/

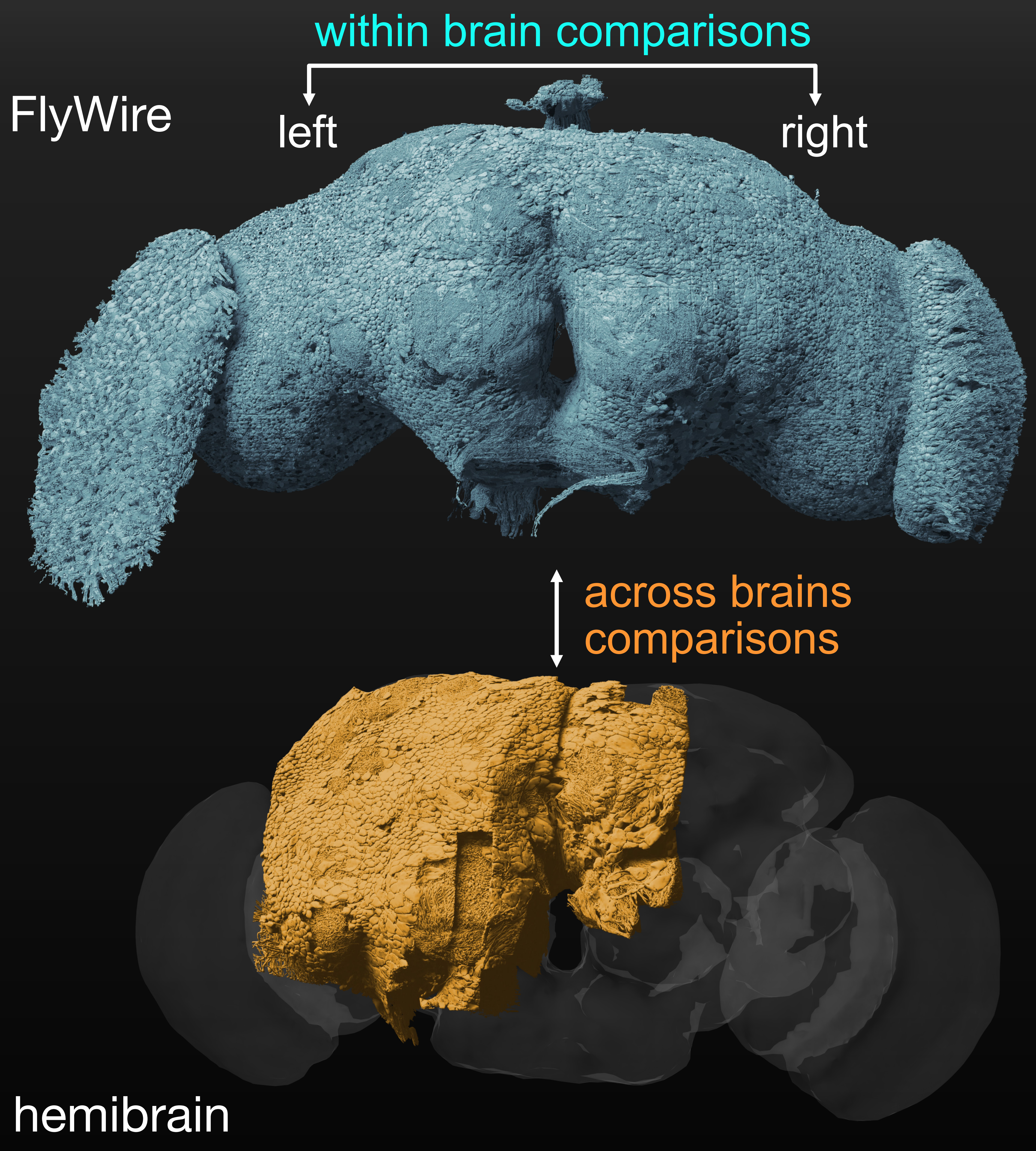
https://github.com/natverse/fafbseg/

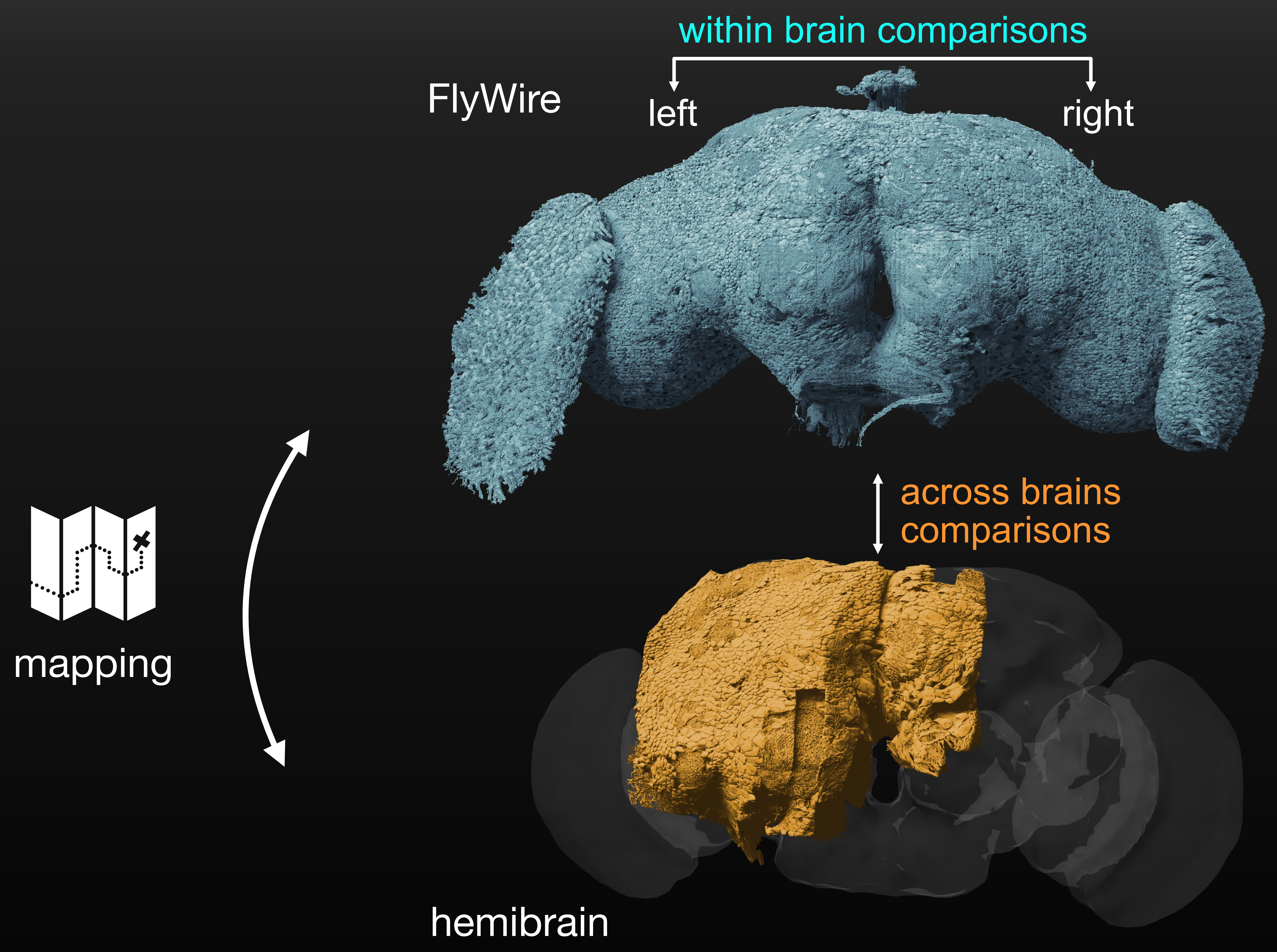
https://fafbseg-py.readthedocs.io/

https://github.com/navis-org/fafbseg-py

How stereotyped are brains?

Are we collecting snowflakes?

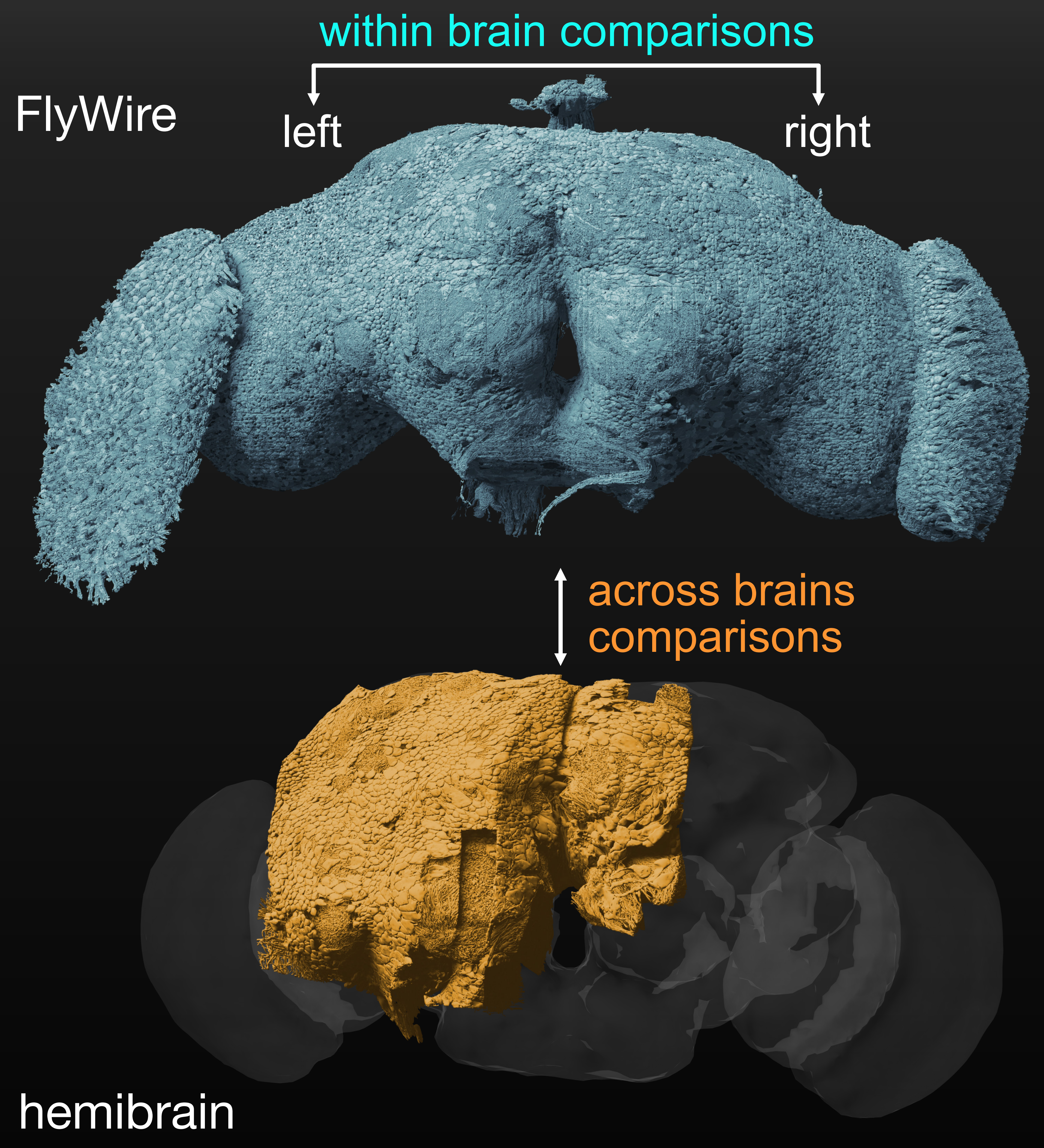


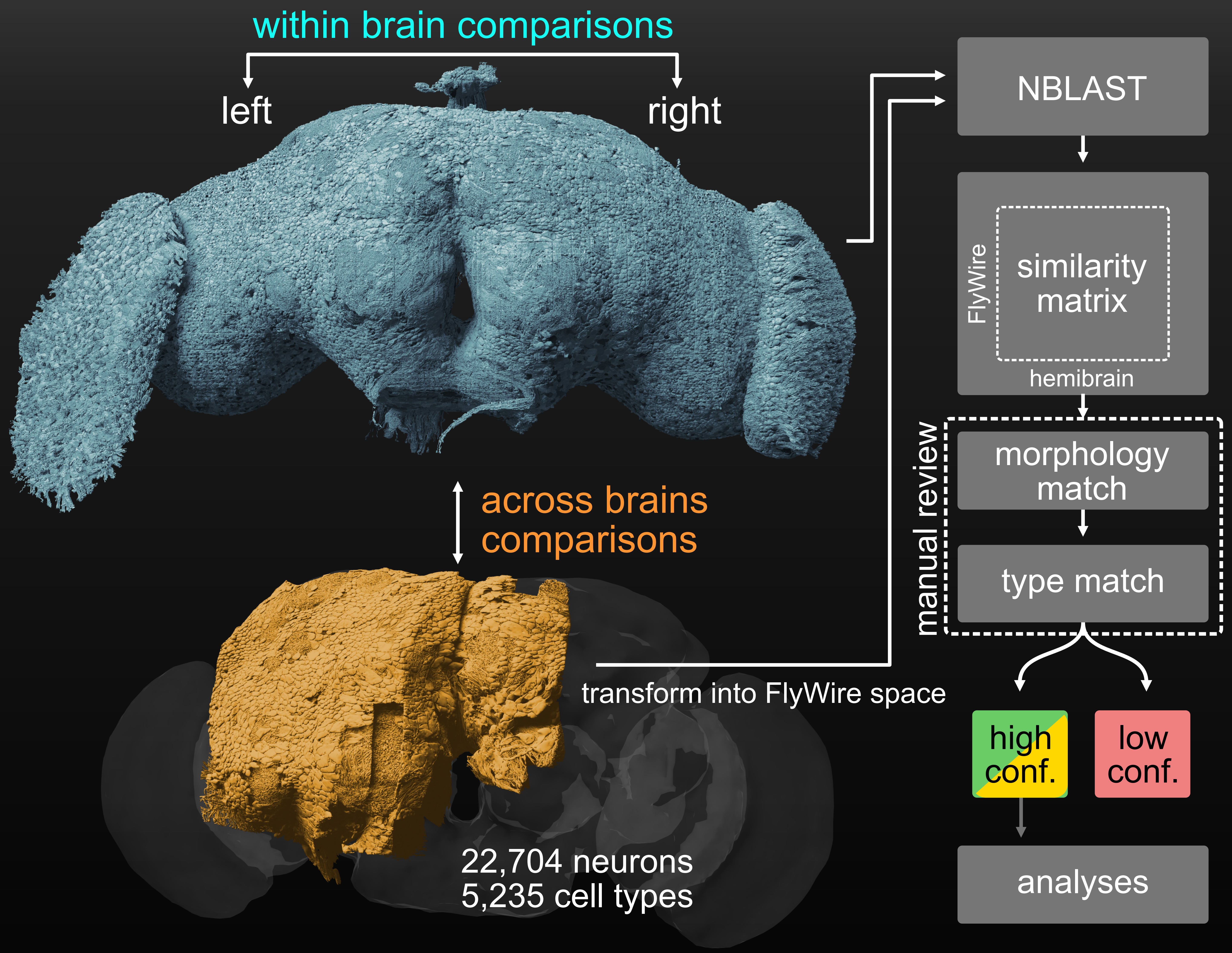

within brain comparisons
FlyWire
left
right
across brains
comparisons
mapping
hemibrain

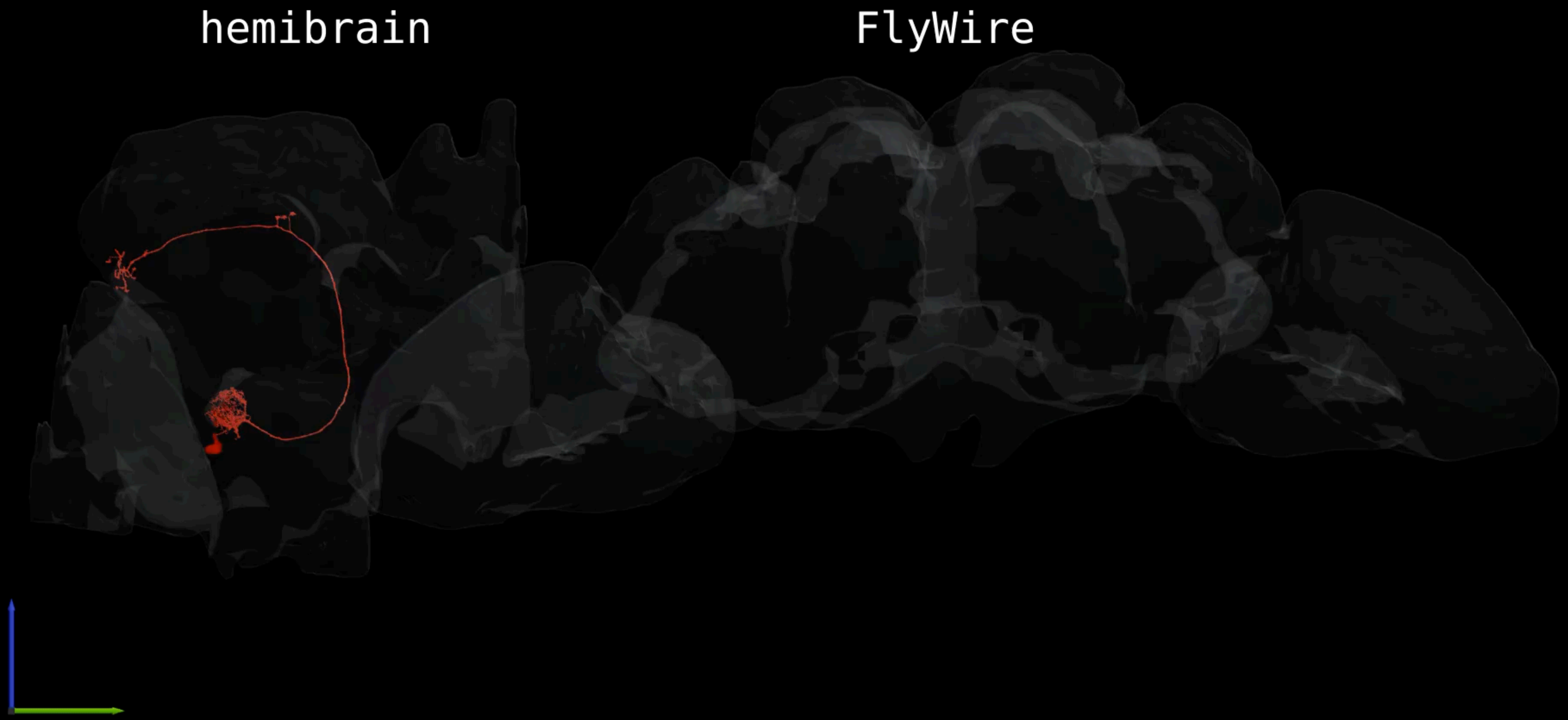
mapping

Morphology (+) Connectivity

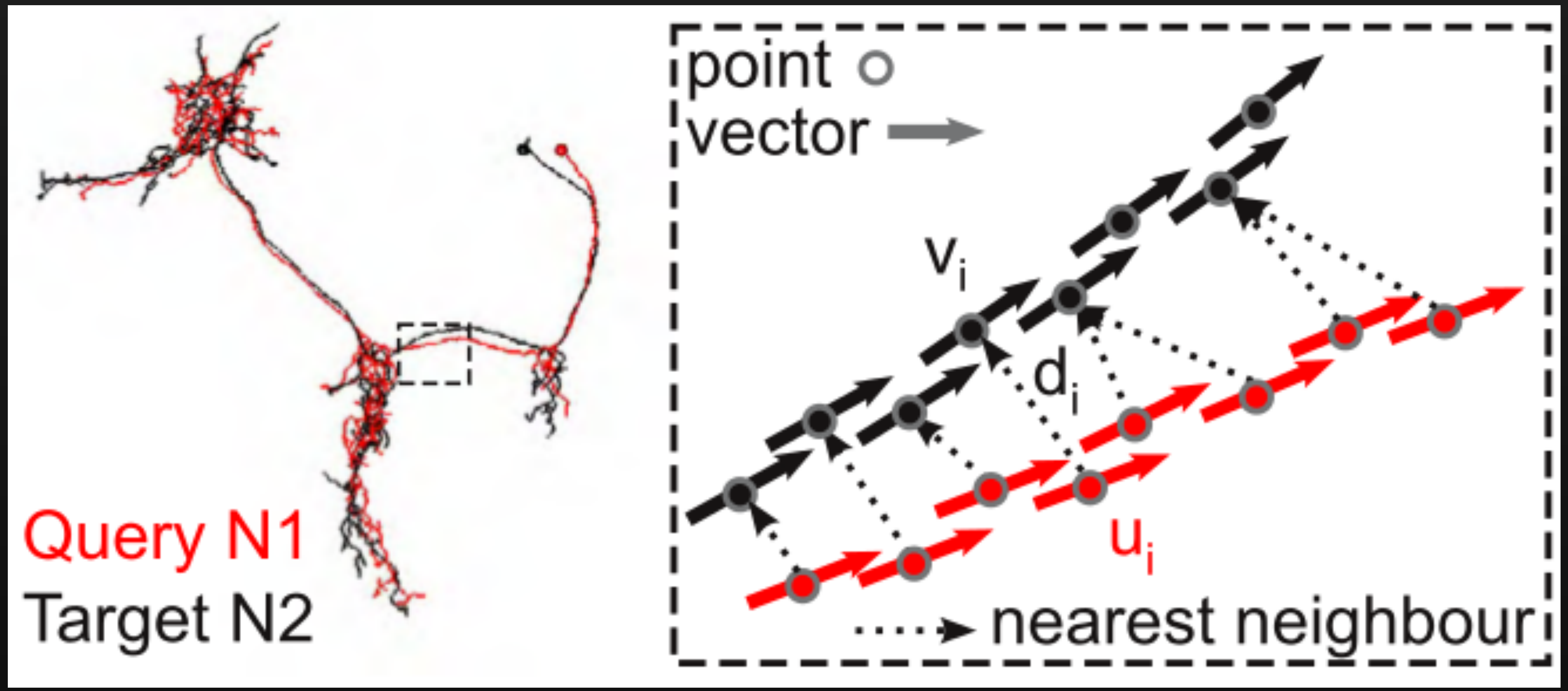
requires spatial transform

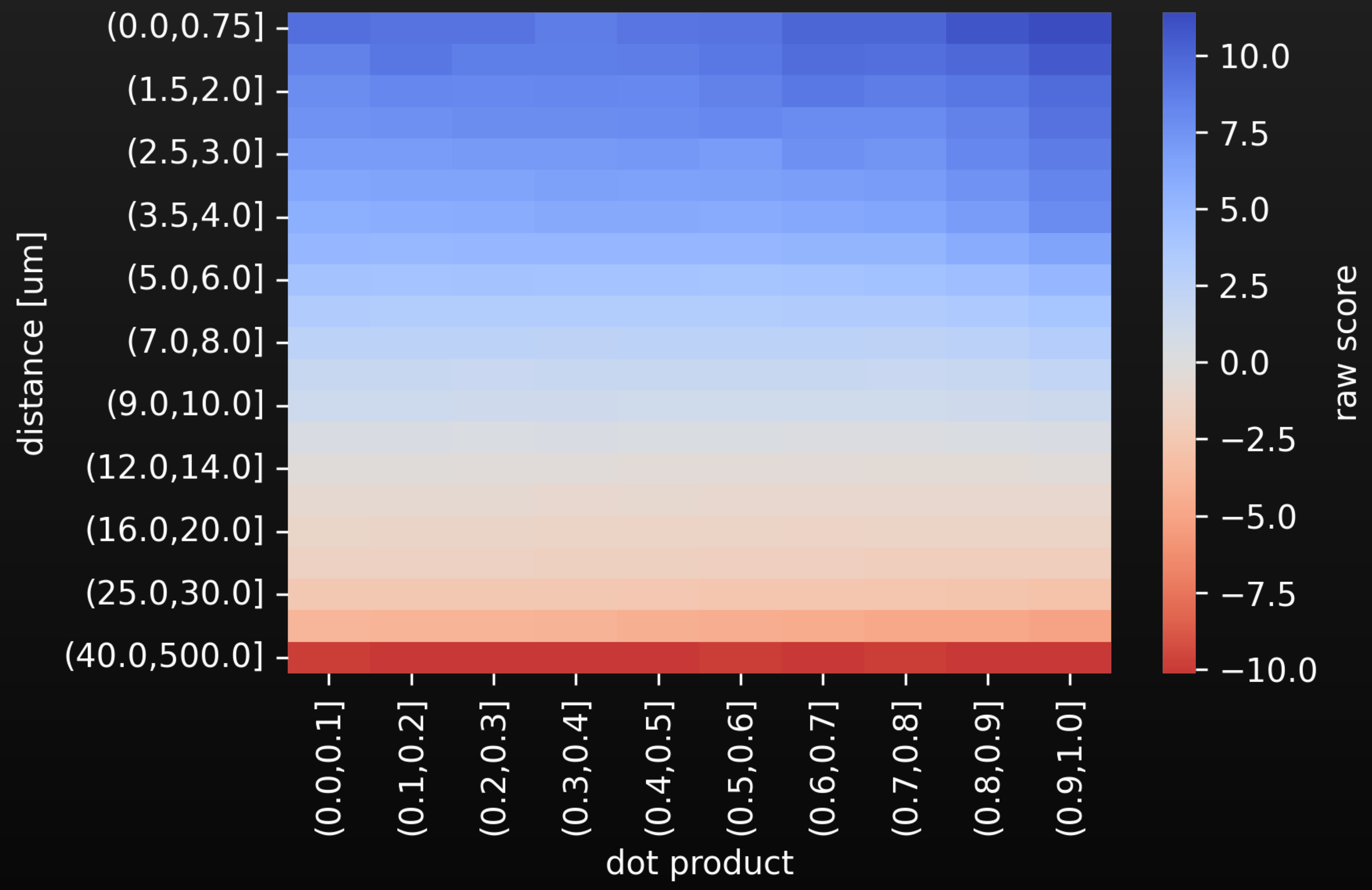



within brain comparisons
left
right
across brains
comparisons
22,704 neurons
5,235 cell types
transform into FlyWire space
NBLAST
FlyWire
similarity
matrix
hemibrain
manual review
morphology
match
type match
high
conf.
low
conf.
analyses

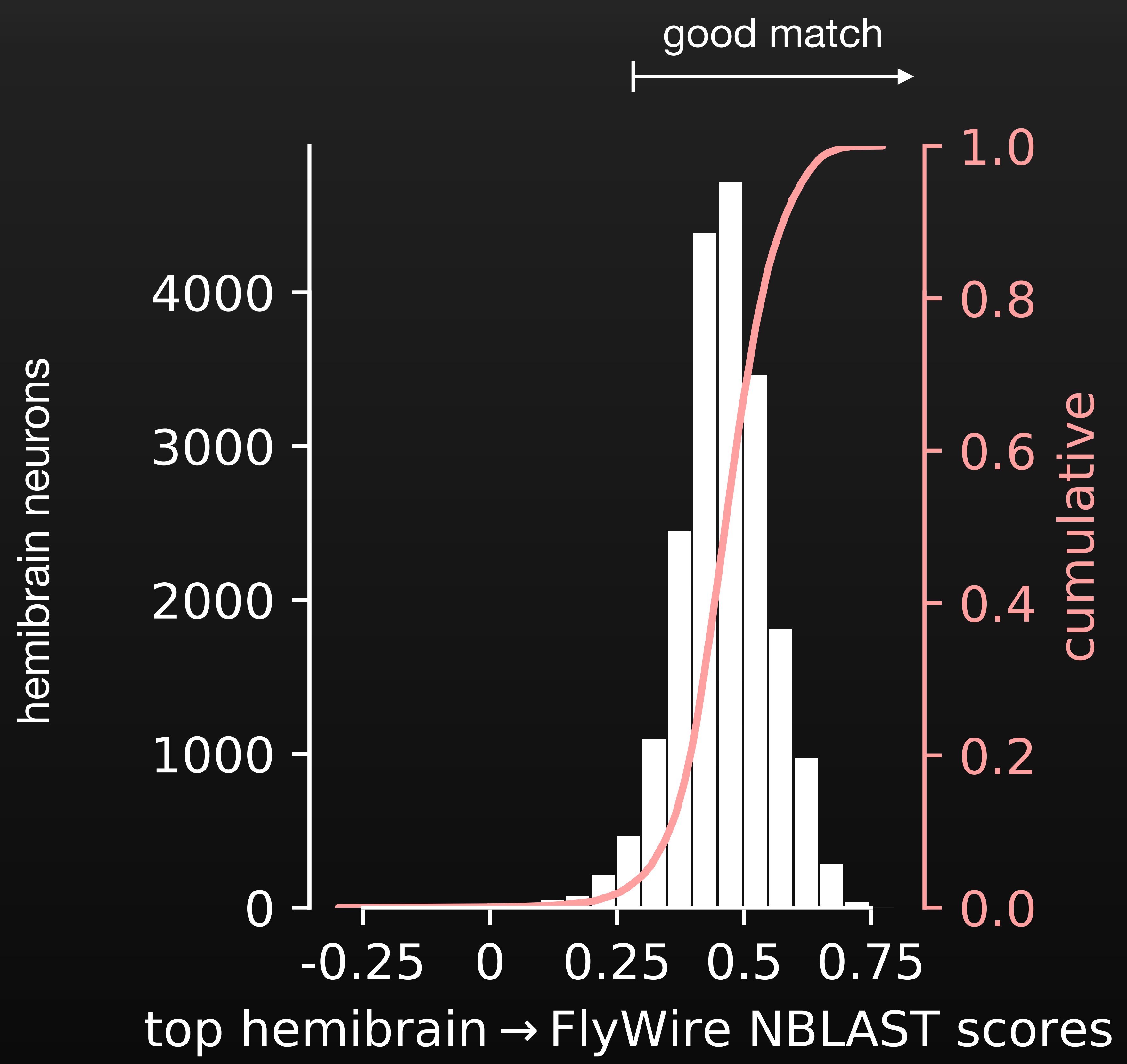

hemibrain
FlyWire

NBLAST for rapid comparison of neurons

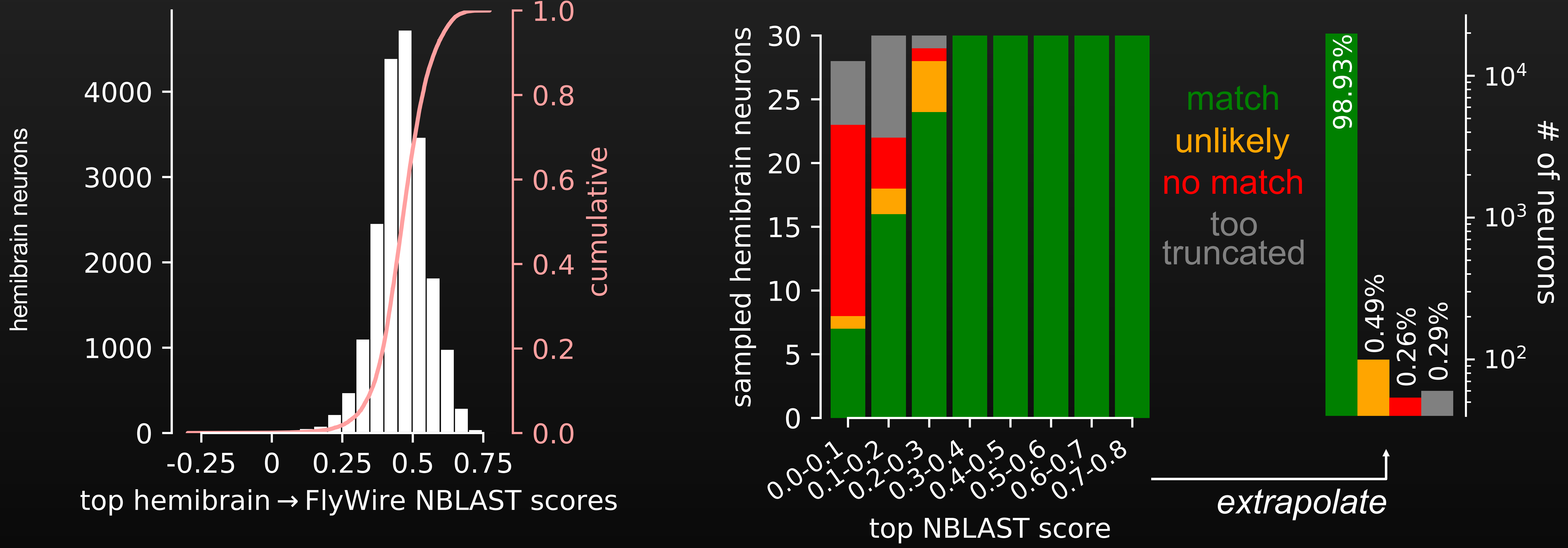




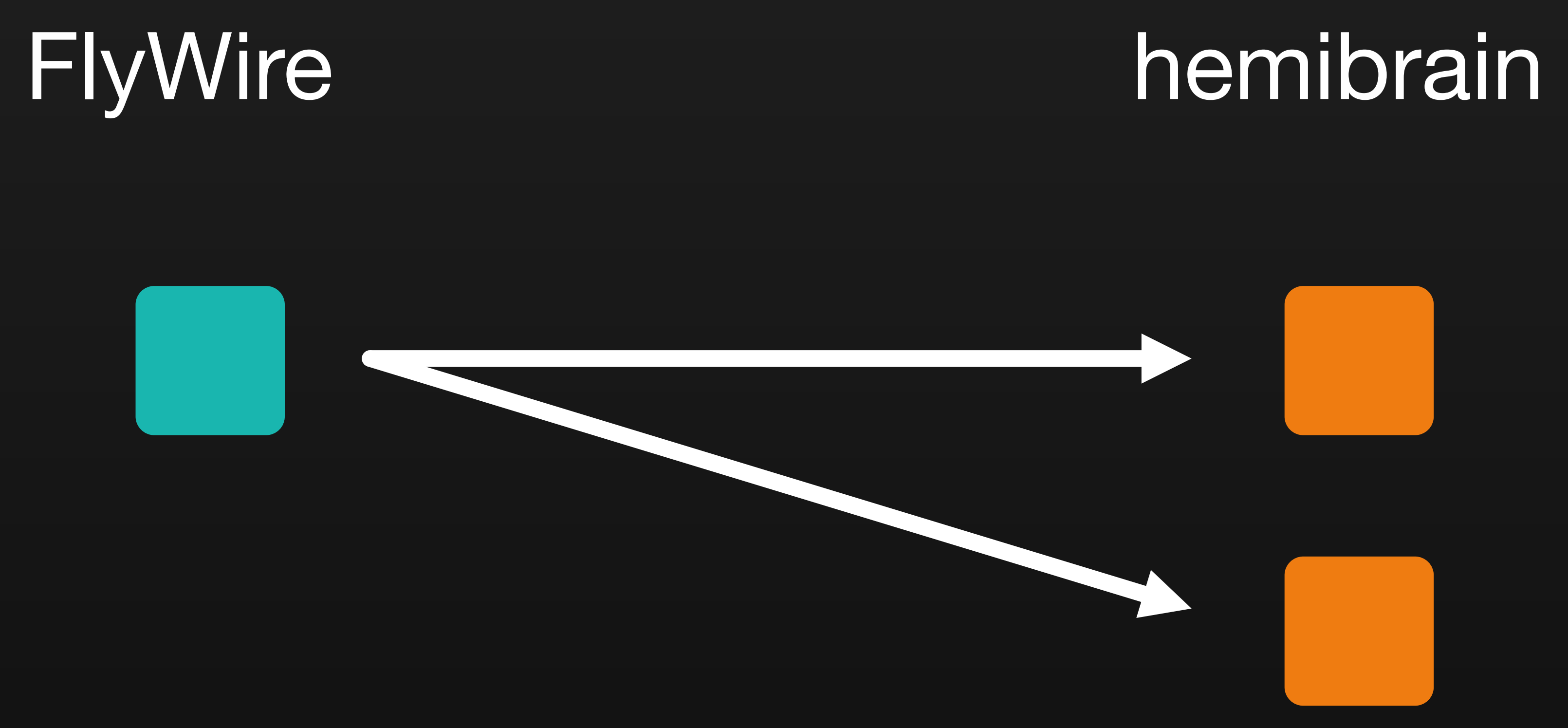
Costa *et al.*, *Neuron (2016)*


good match
hemibrain neurons
4000
3000
2000
1000
0
-0.25
0
0.25
0.5
0.75
1.0
0.8
0.6
0.4
0.2
0.0
cumulative
top hemibrain → FlyWire NBLAST scores

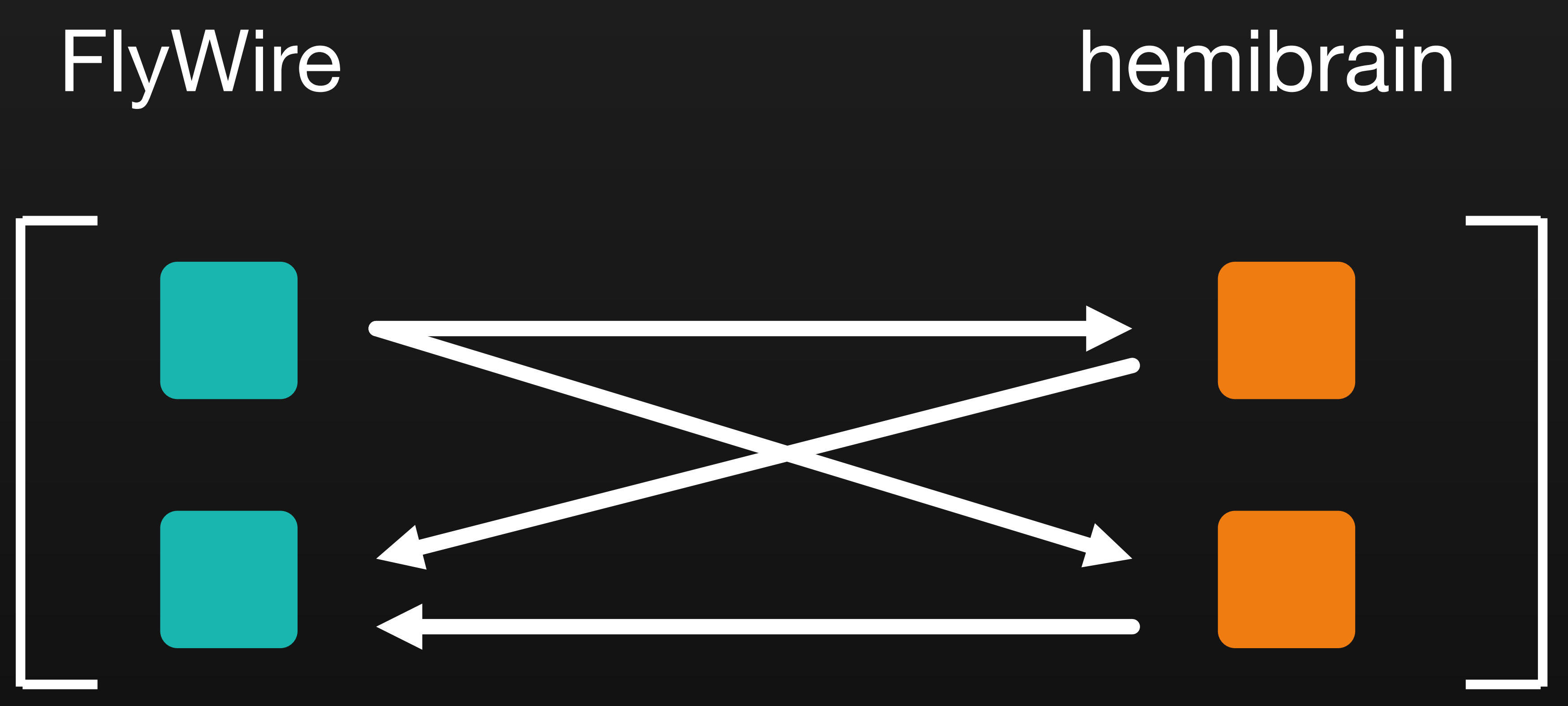
99% of neurons have a match in another brain



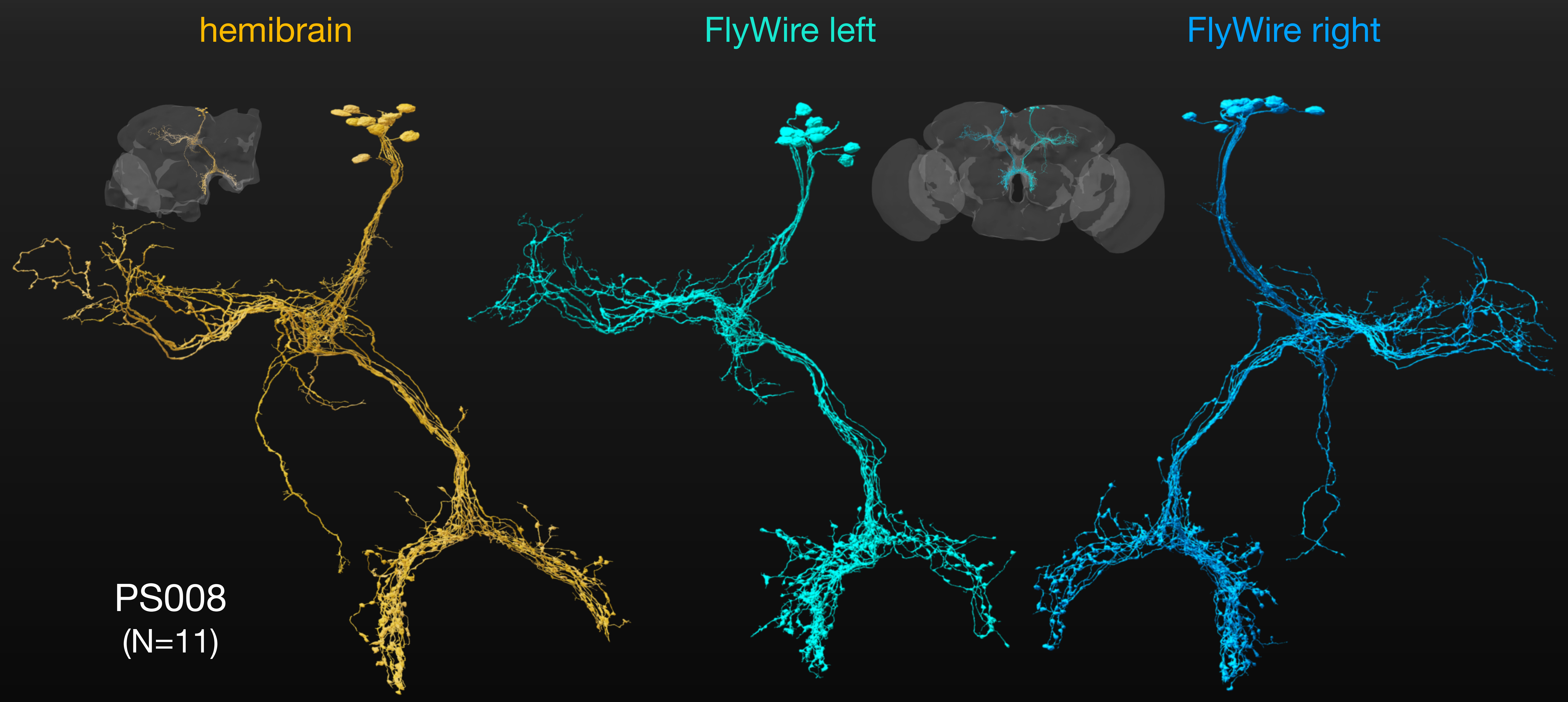

FlyWire
hemibrain

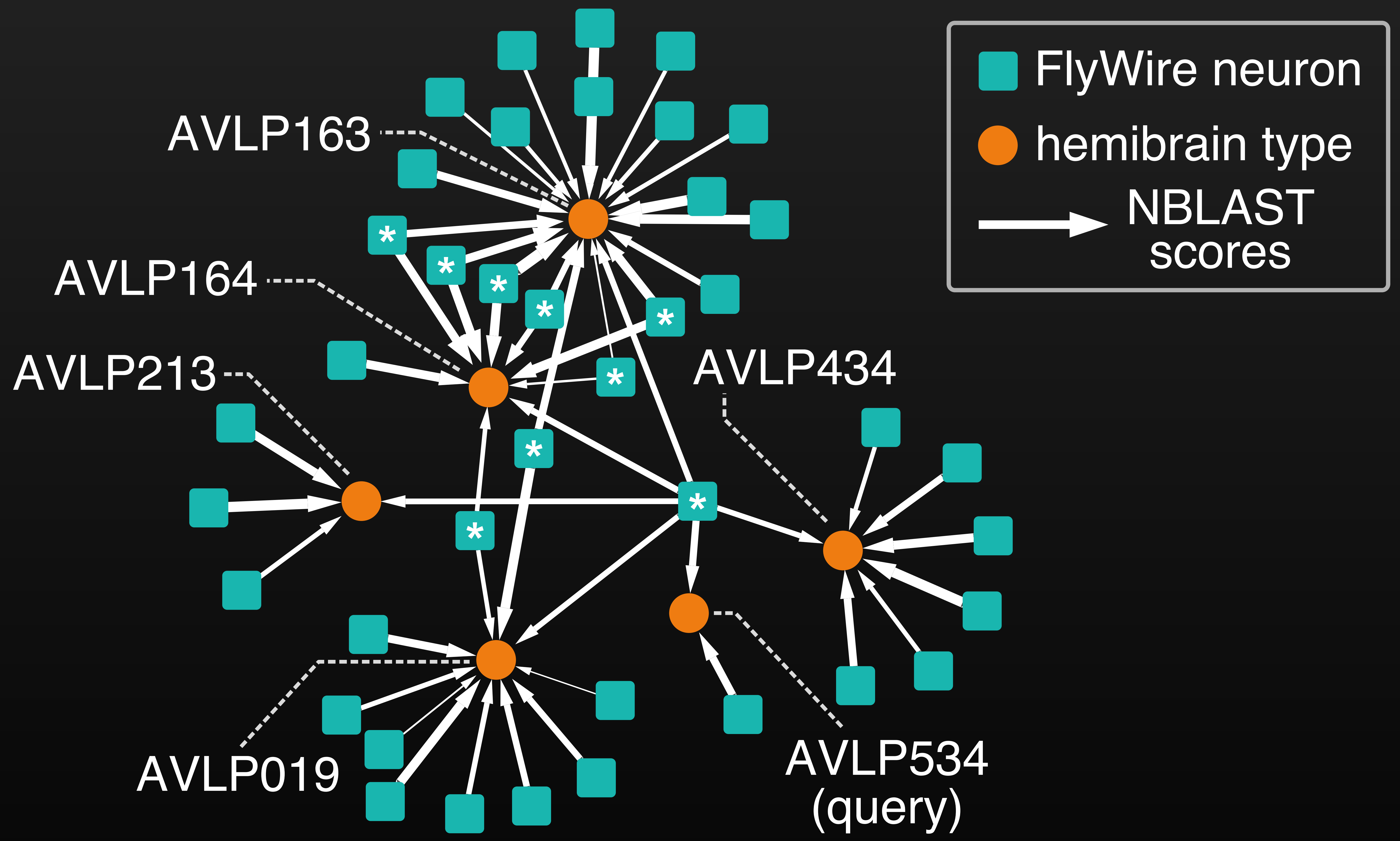

FlyWire
hemibrain


FlyWire
hemibrain



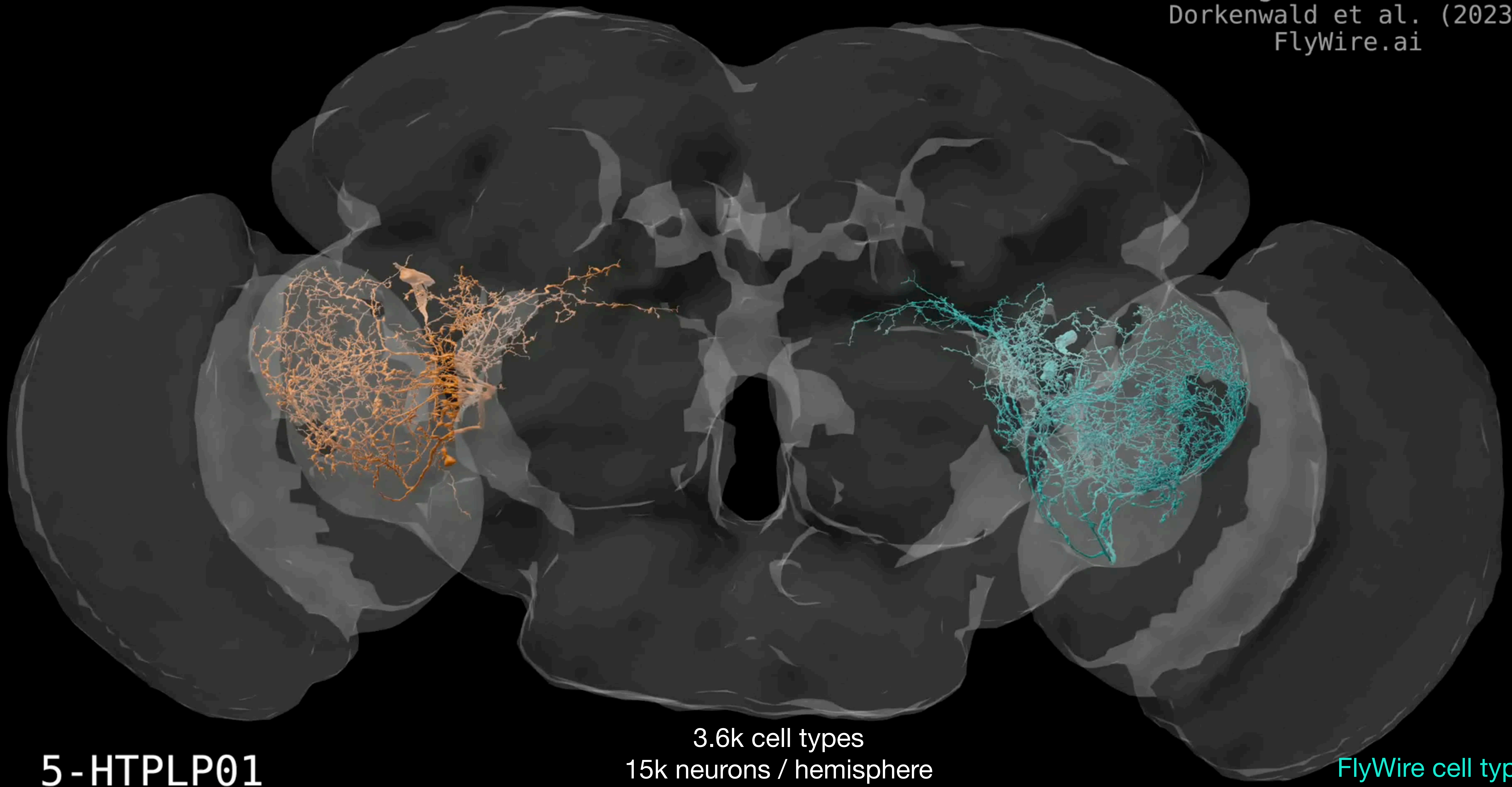
“cell type” is the smallest unit of conservation across brains/hemispheres

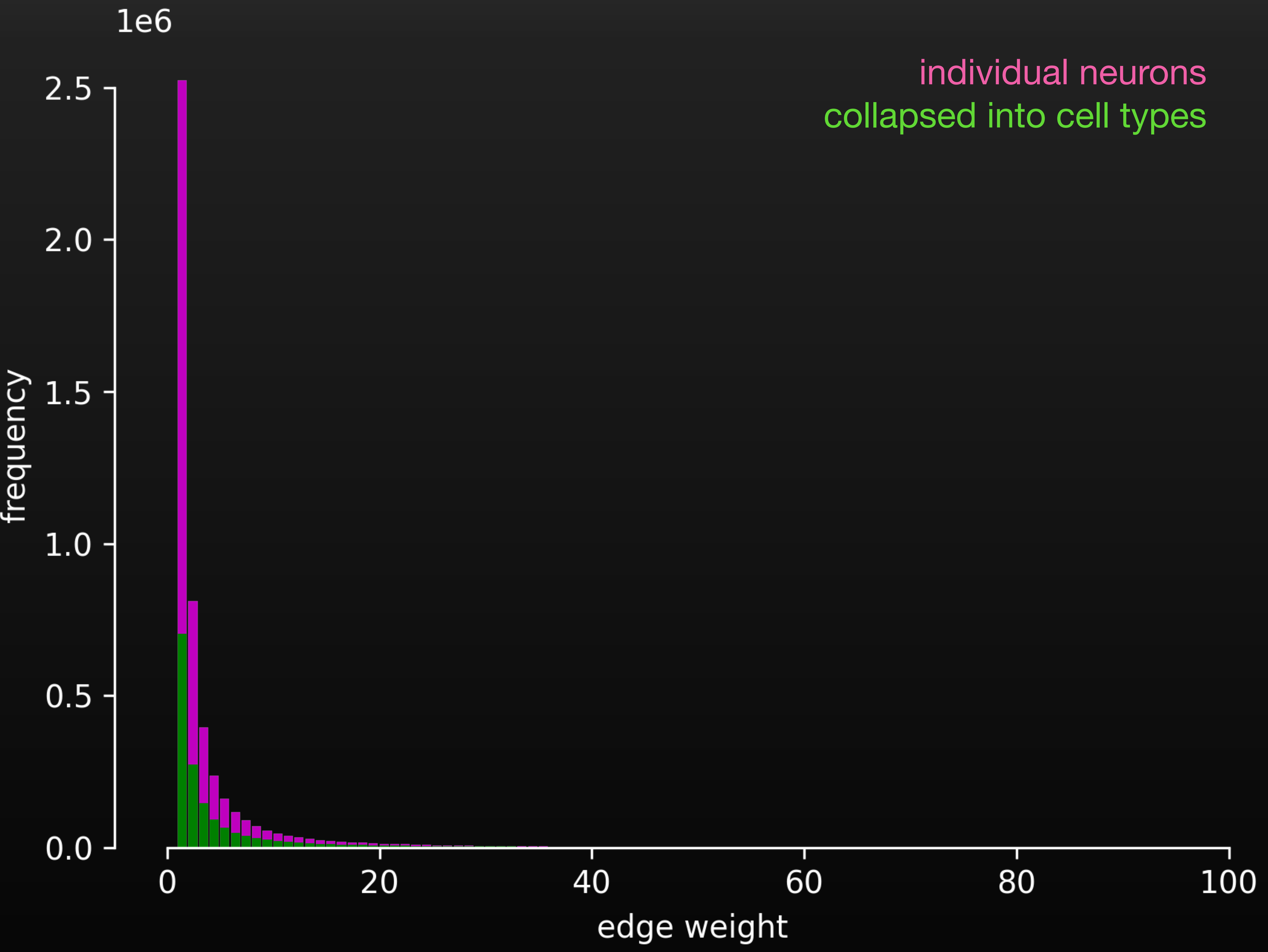

hemibrain
FlyWire left
FlyWire right
PS008
(N=11)

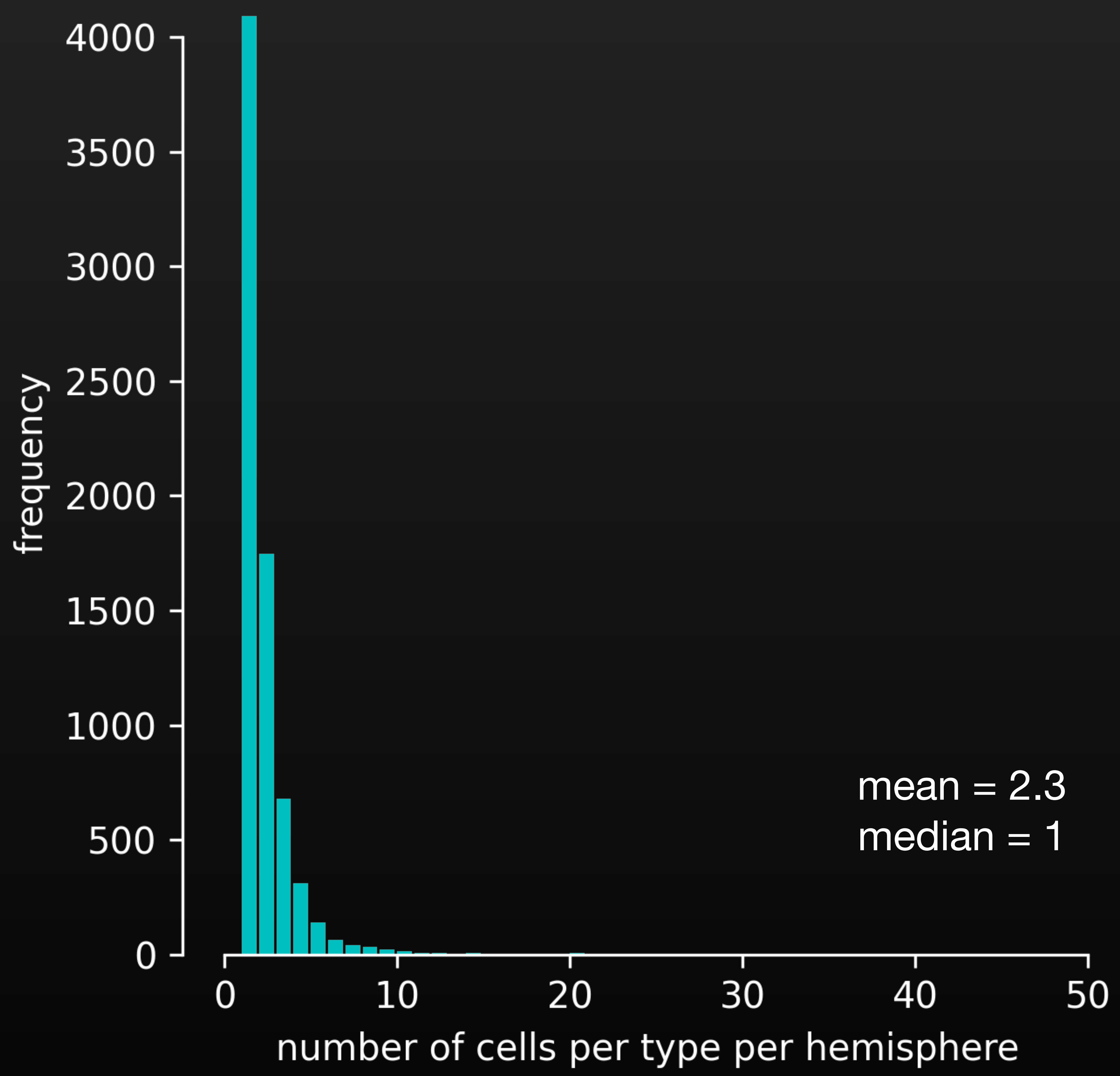


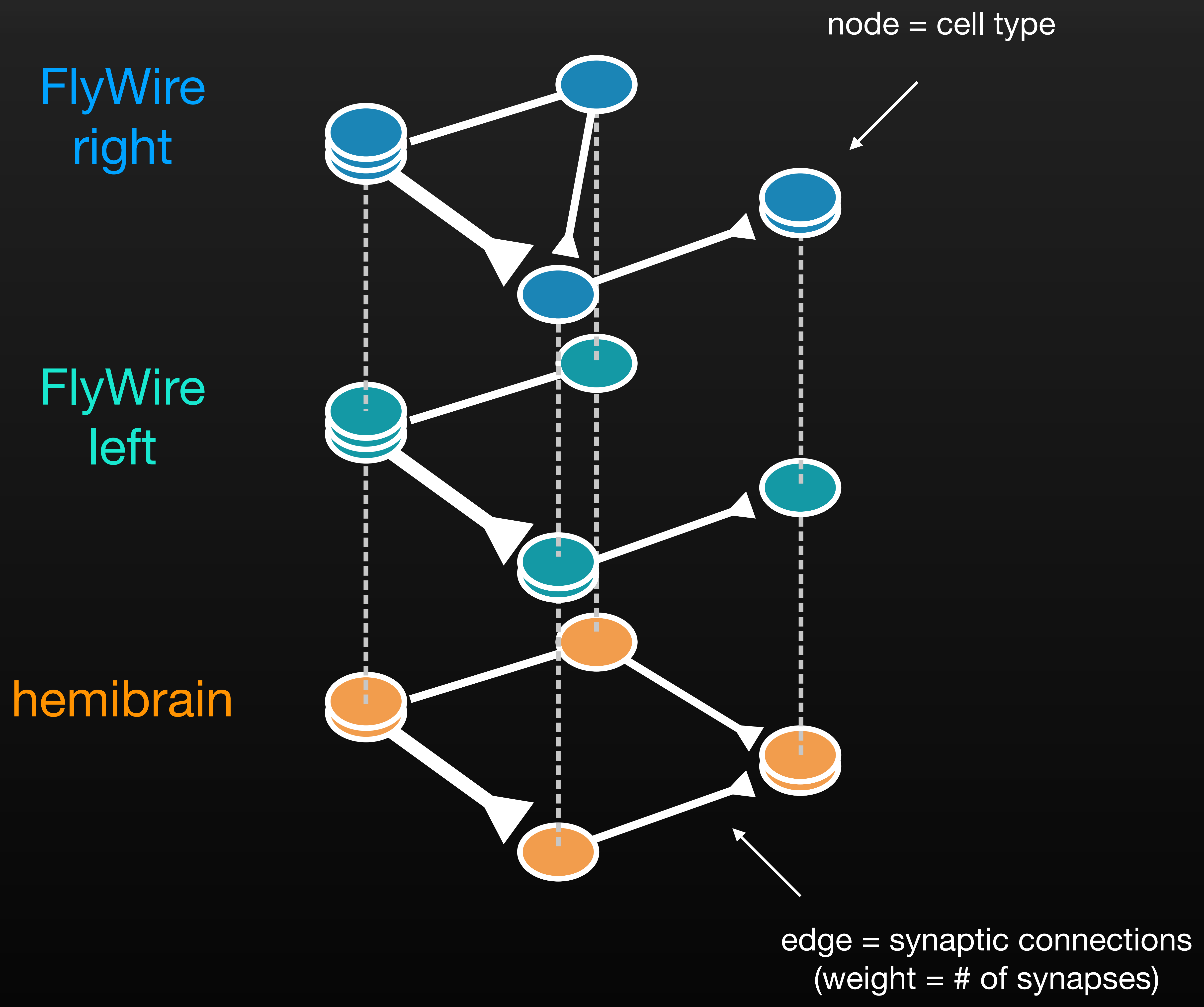
hemibrain cell types identified:

- 55% matched 1:1
- 12% found but modified
- 33% likely need revision

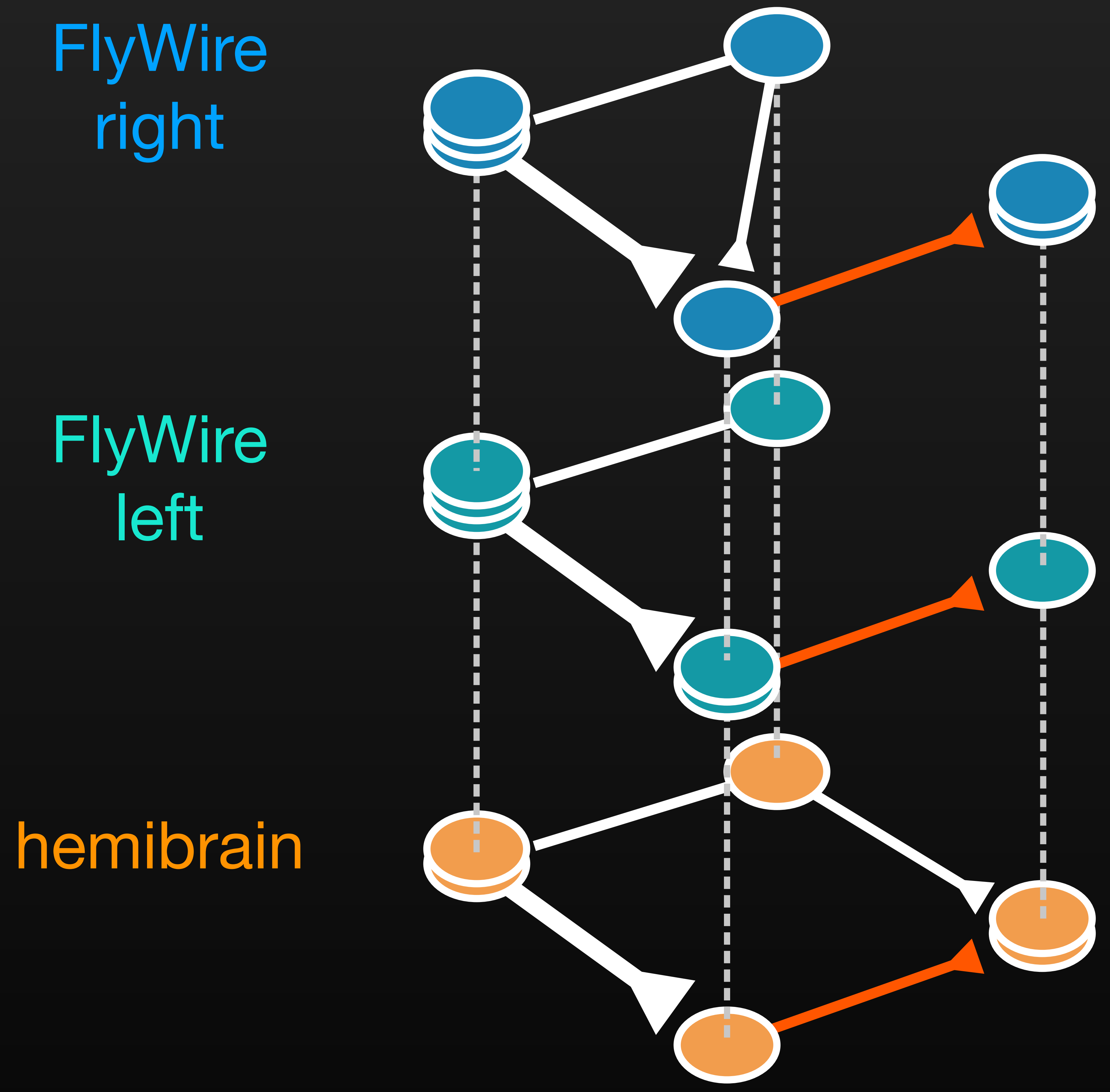

Schlegel et al. (2023)
Dorkenwald et al. (2023)
FlyWire.ai
5-HTPLP01
3.6k cell types
15k neurons / hemisphere
coverage ~50% of central brain graph
FlyWire cell type
hemibrain cell type


1e6
2.5
2.0
1.5
1.0
0.5
0.0
frequency
individual neurons
collapsed into cell types
0
20
40
60
80
100
edge weight


4000
3500
3000
2500
2000
1500
1000
500
0
frequency
mean = 2.3
median = 1
0
10
20
30
40
50
number of cells per type per hemisphere

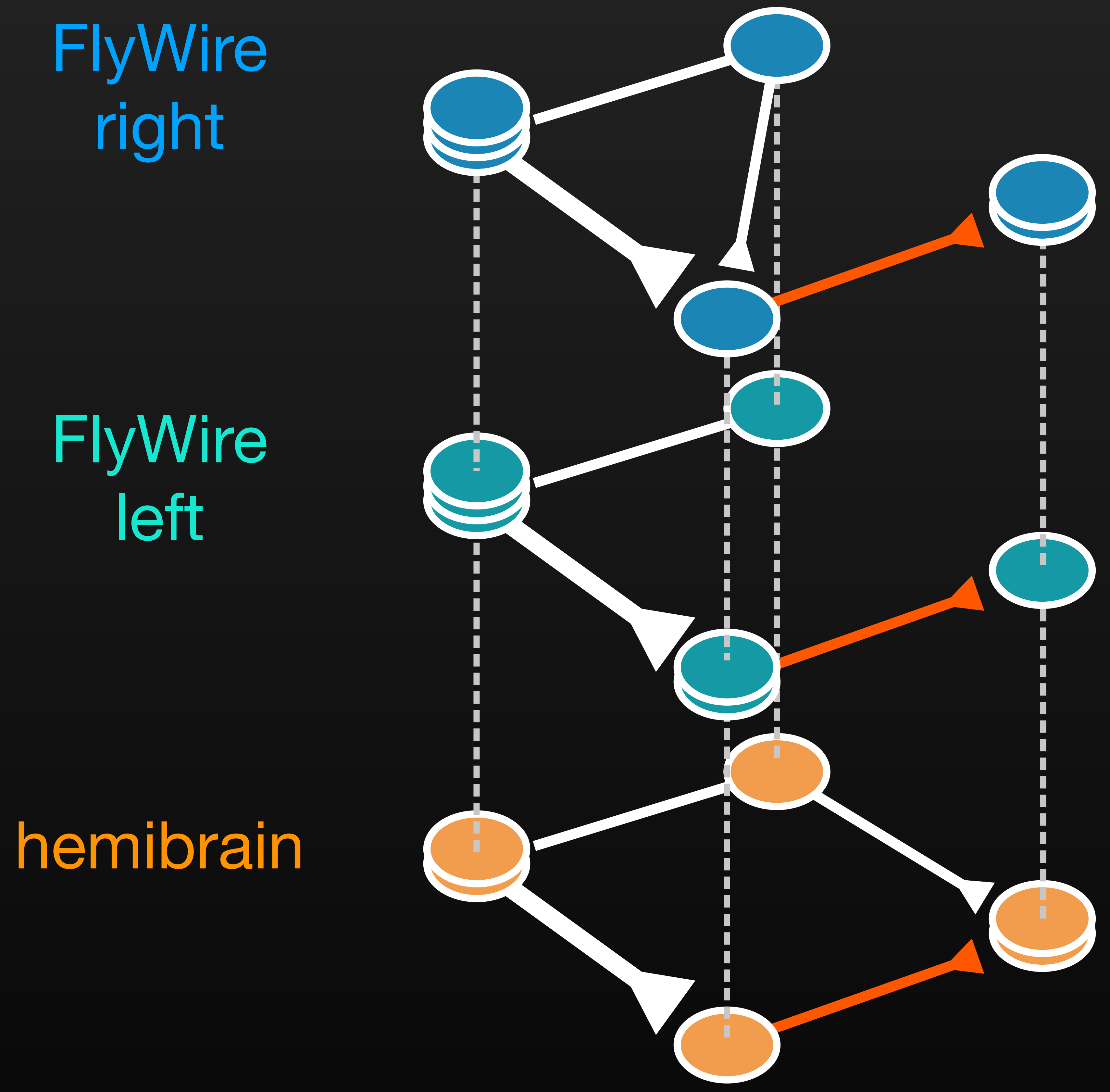


3,6k cell types
15k neurons / hemisphere

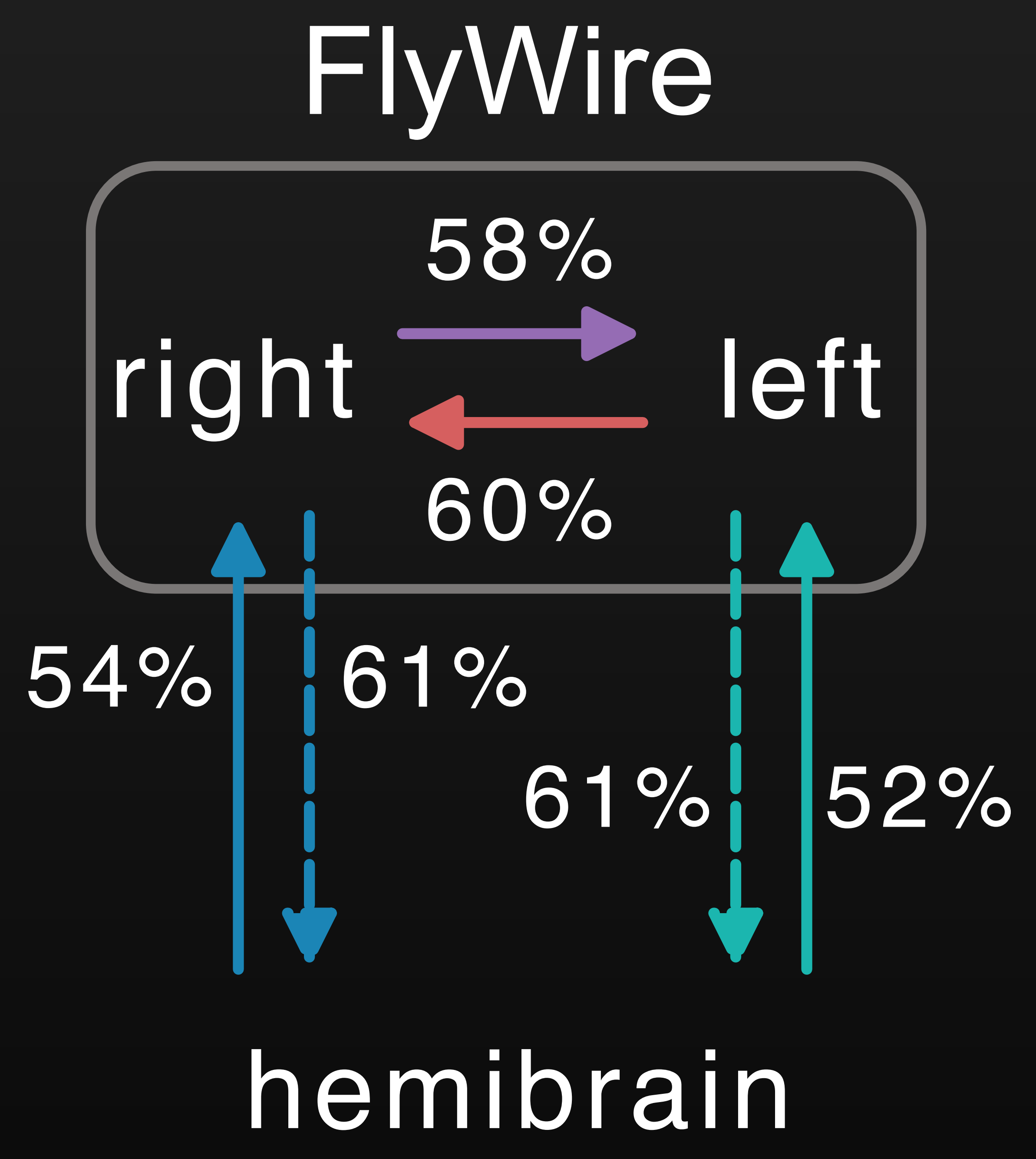


Does this edge exist across all hemispheres?

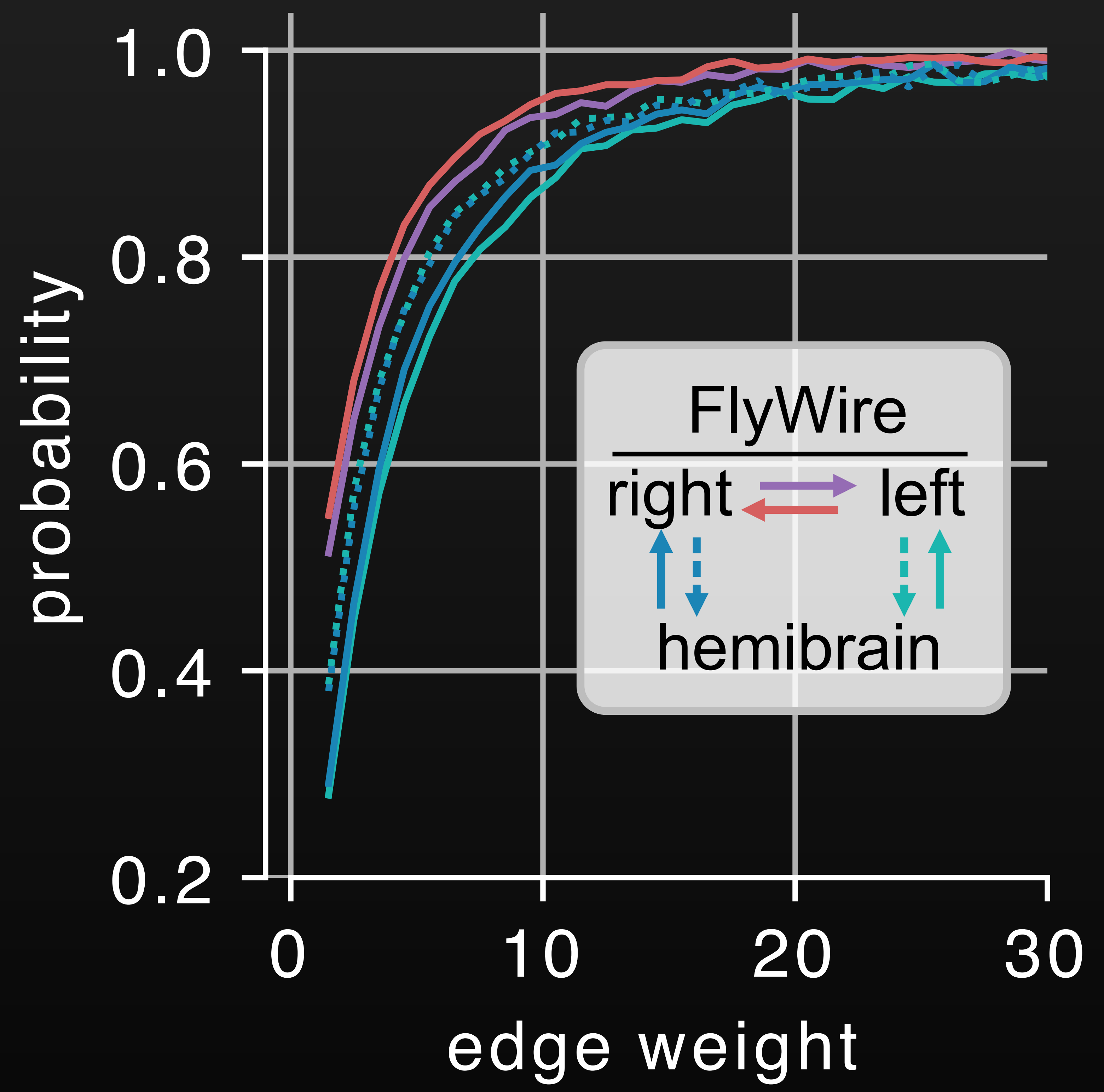
If so, how consistent is its weight?

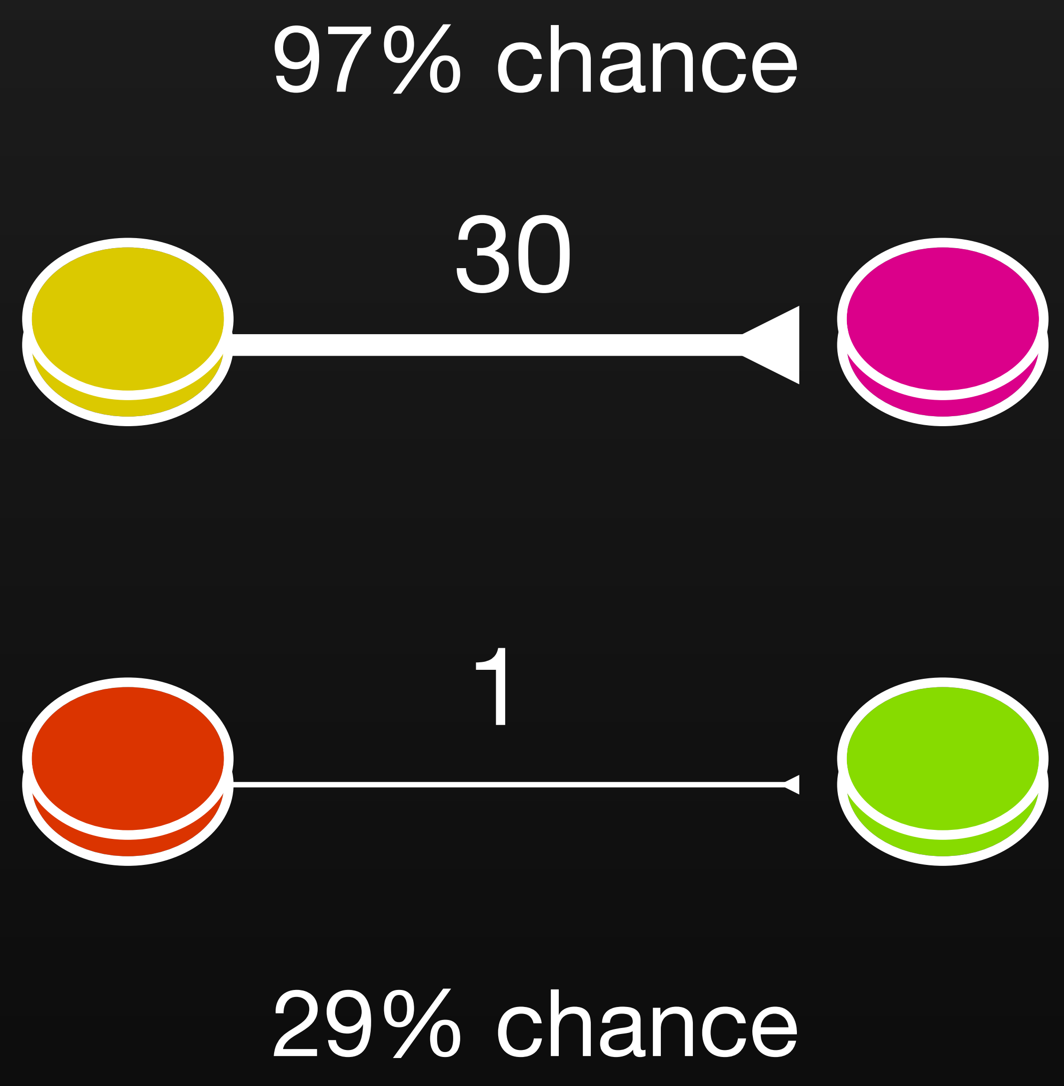


fraction of edges present across hemispheres

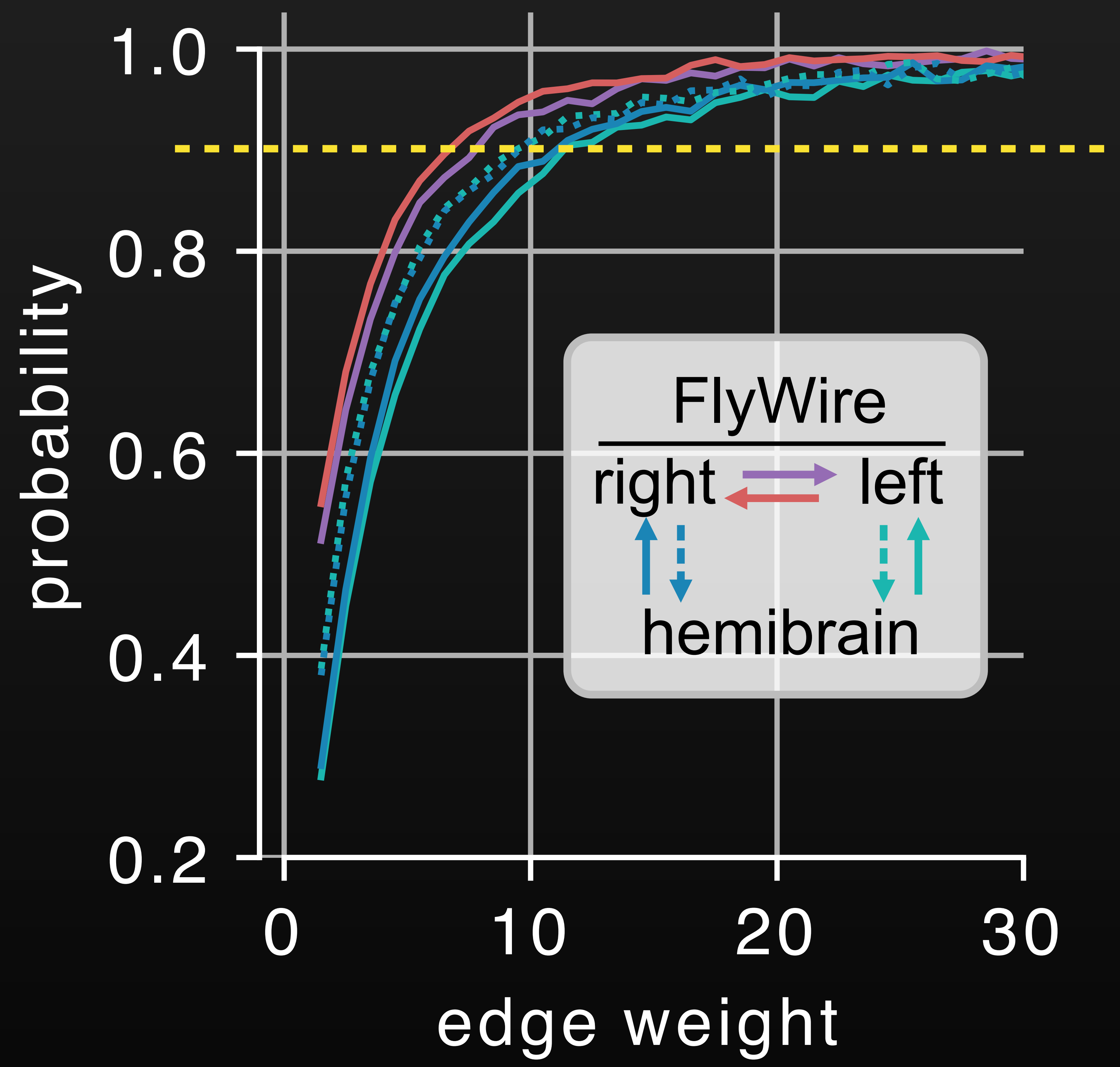


stronger edges are more likely to be present across hemispheres

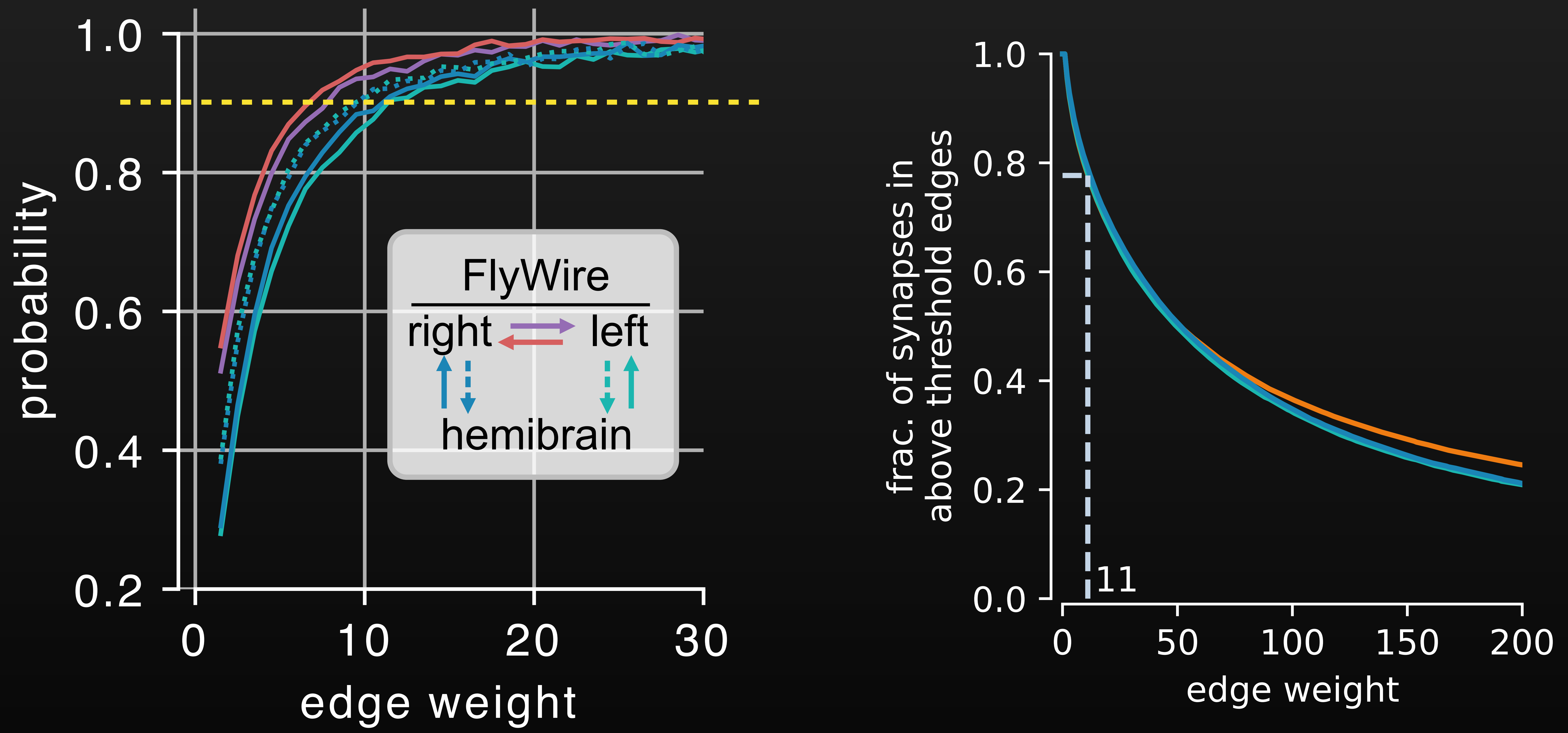


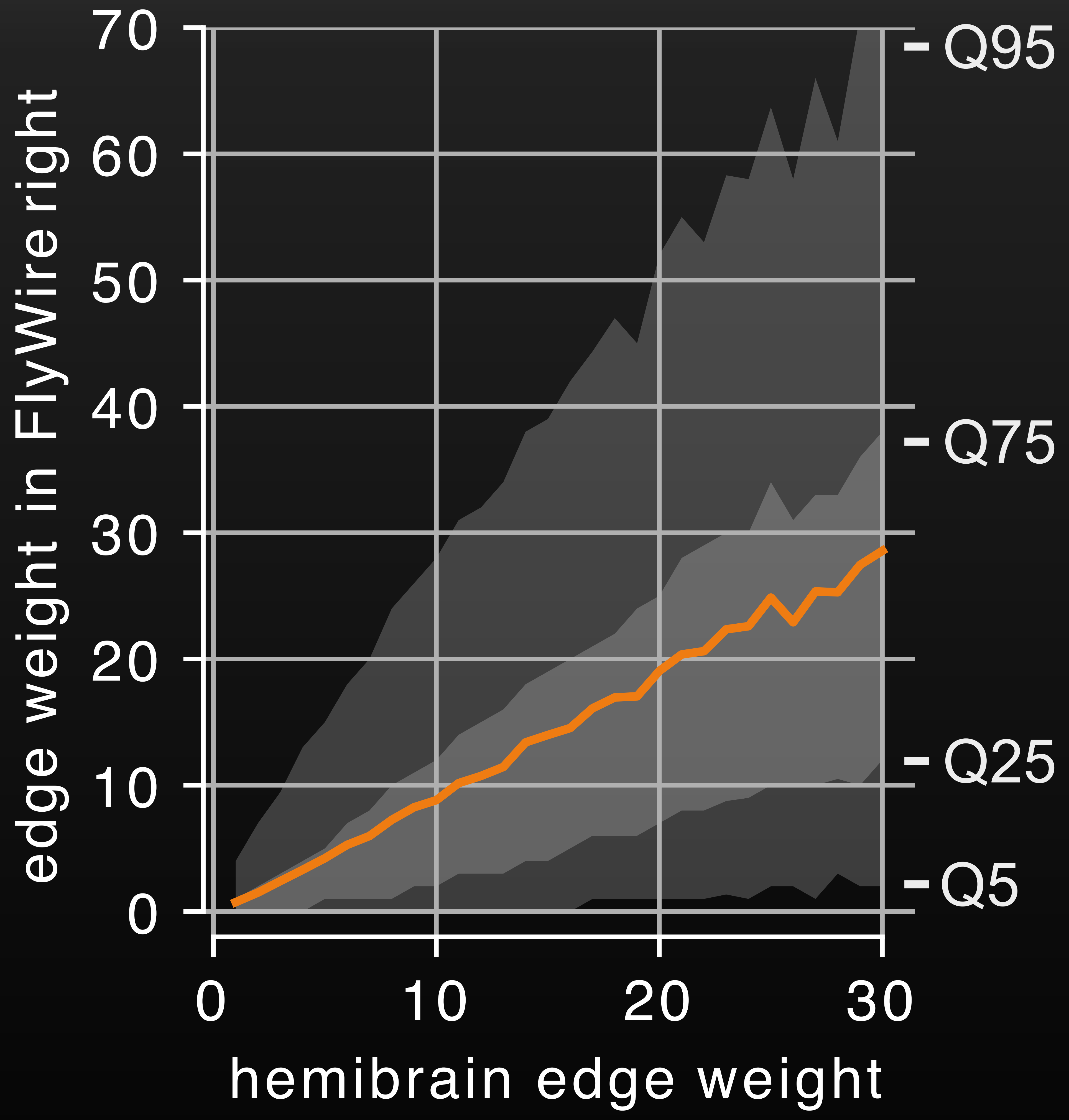


edges with >10 synapses
(or >=0.9% of target's input)
can be reproducibly (>90% chance) found across datasets

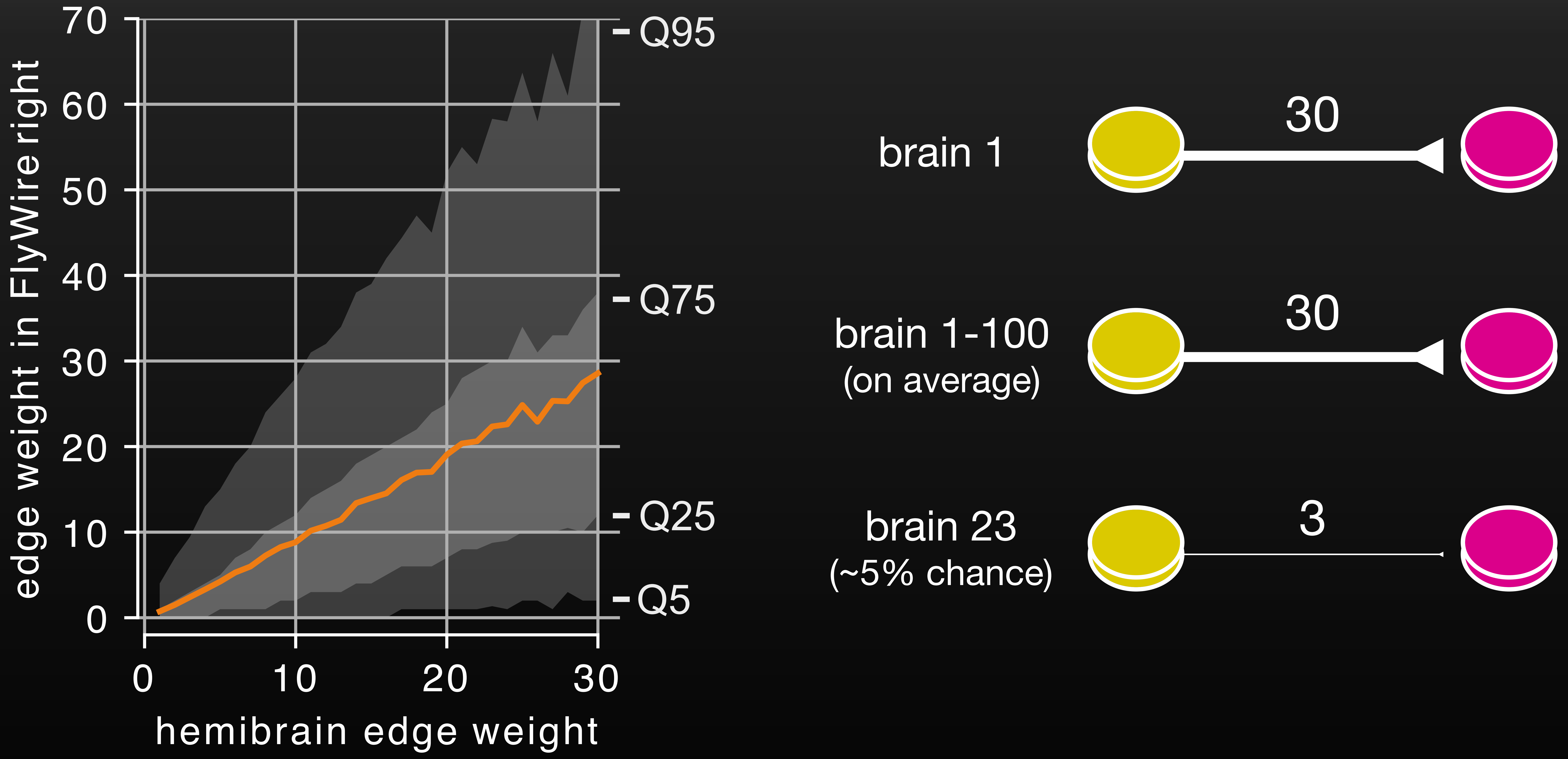


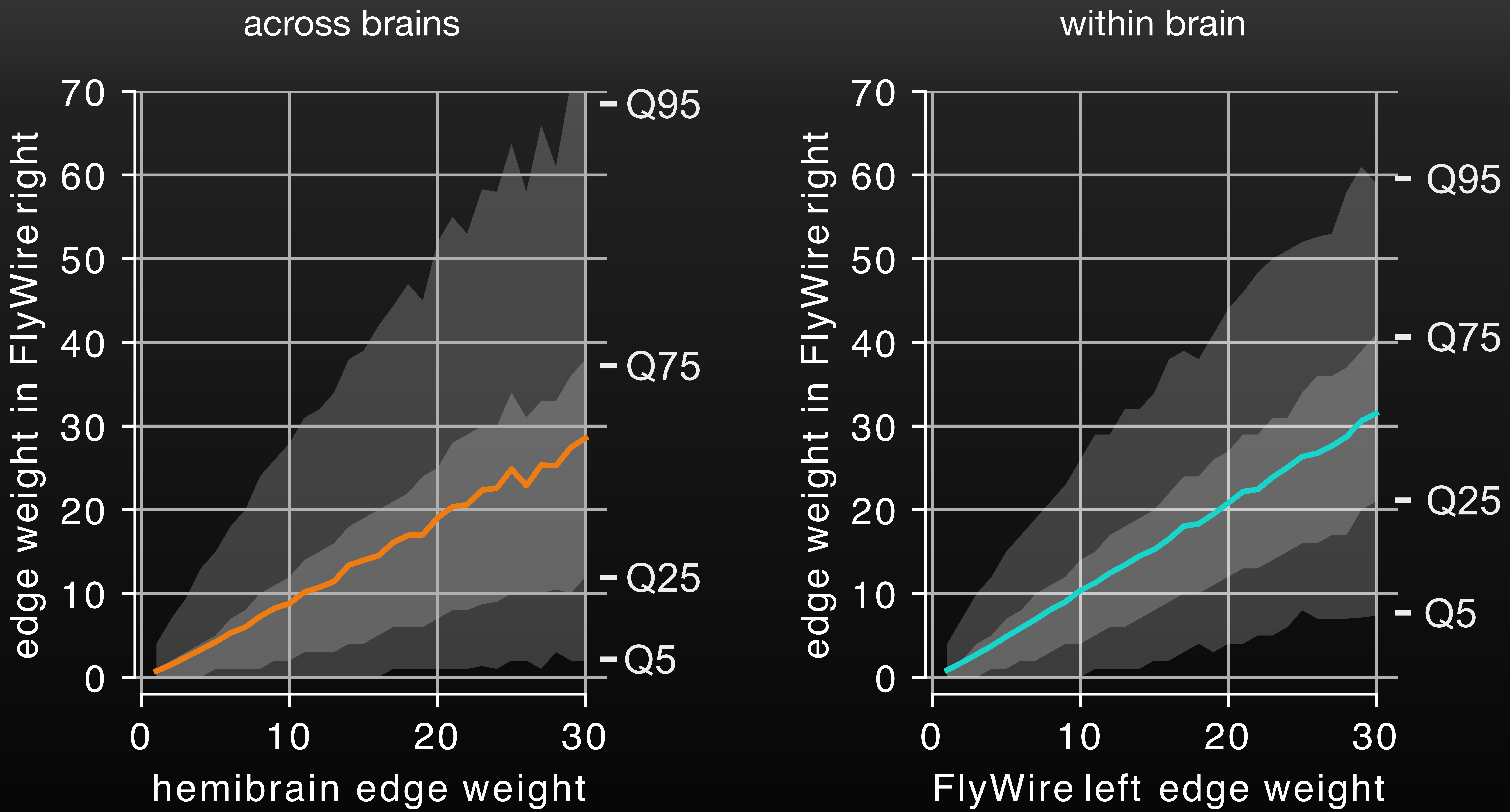
16% of edges are >10 but contain ~80% of all synapses
7% of edges are stronger than 1% but contain ~50% of all synapses





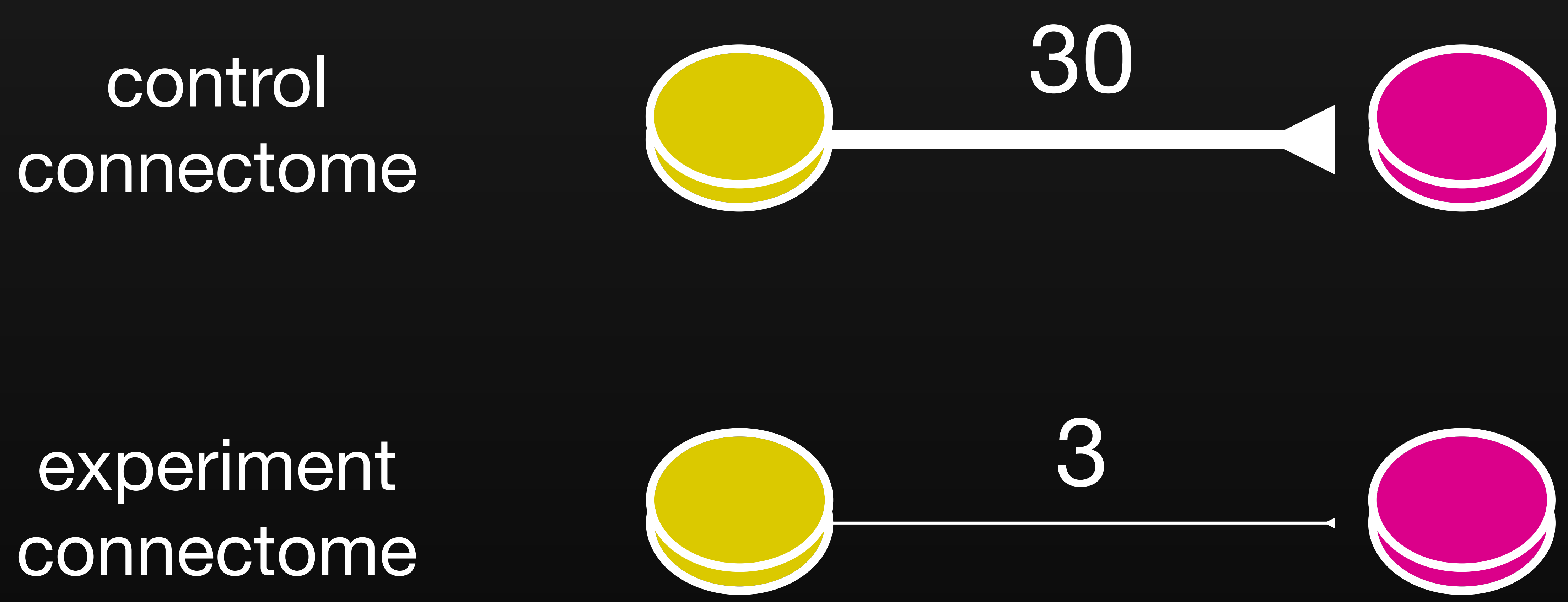
edge weight in FlyWire right
70
60
50
40
30
20
10
0
0
10
20
30
hemibrain edge weight
Q95
Q75
Q25
Q5


edge weight in FlyWire right
70
60
50
40
30
20
10
0
0
10
20
30
hemibrain edge weight
Q95
Q75
Q25
Q5
brain 1
30
brain 1-100
(on average)
30
brain 23
(~5% chance)
3

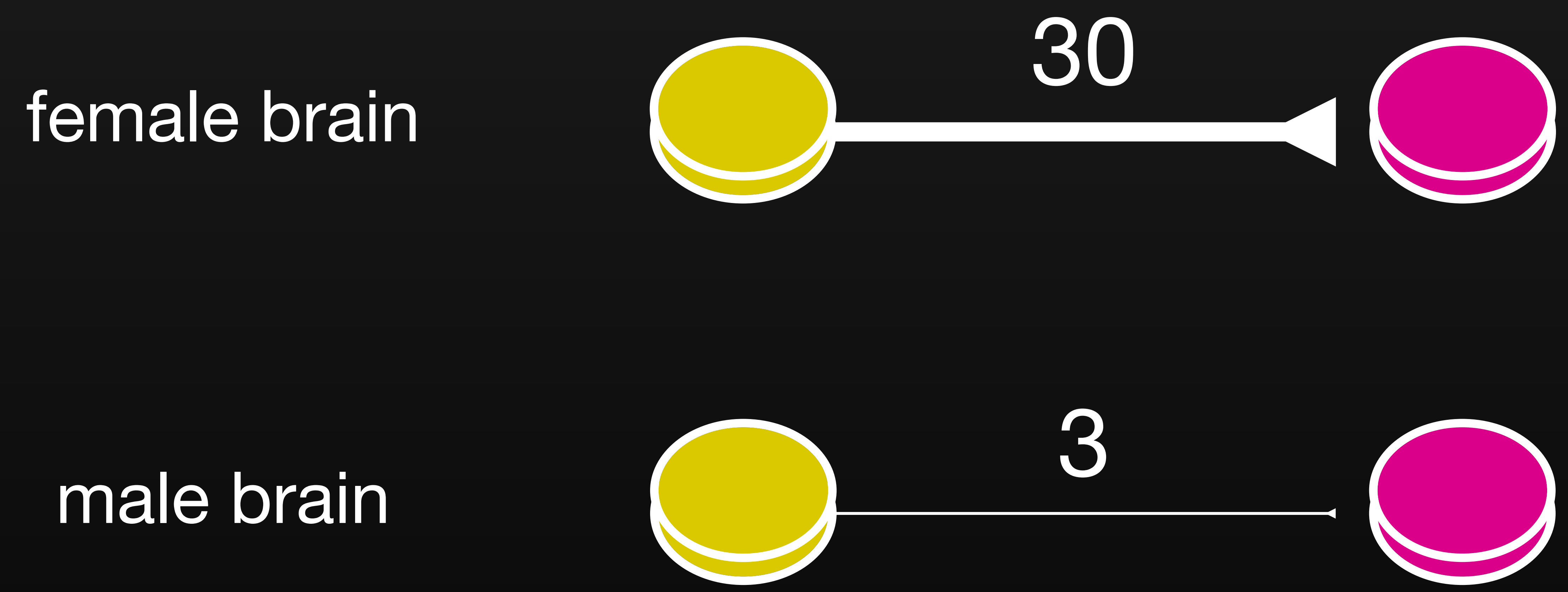


across brains
within brain
edge weight in FlyWire right
hemibrain edge weight
FlyWire left edge weight
Q95
Q75
Q25
Q5
0
10
20
30
40
50
60
70

Why does this matter?

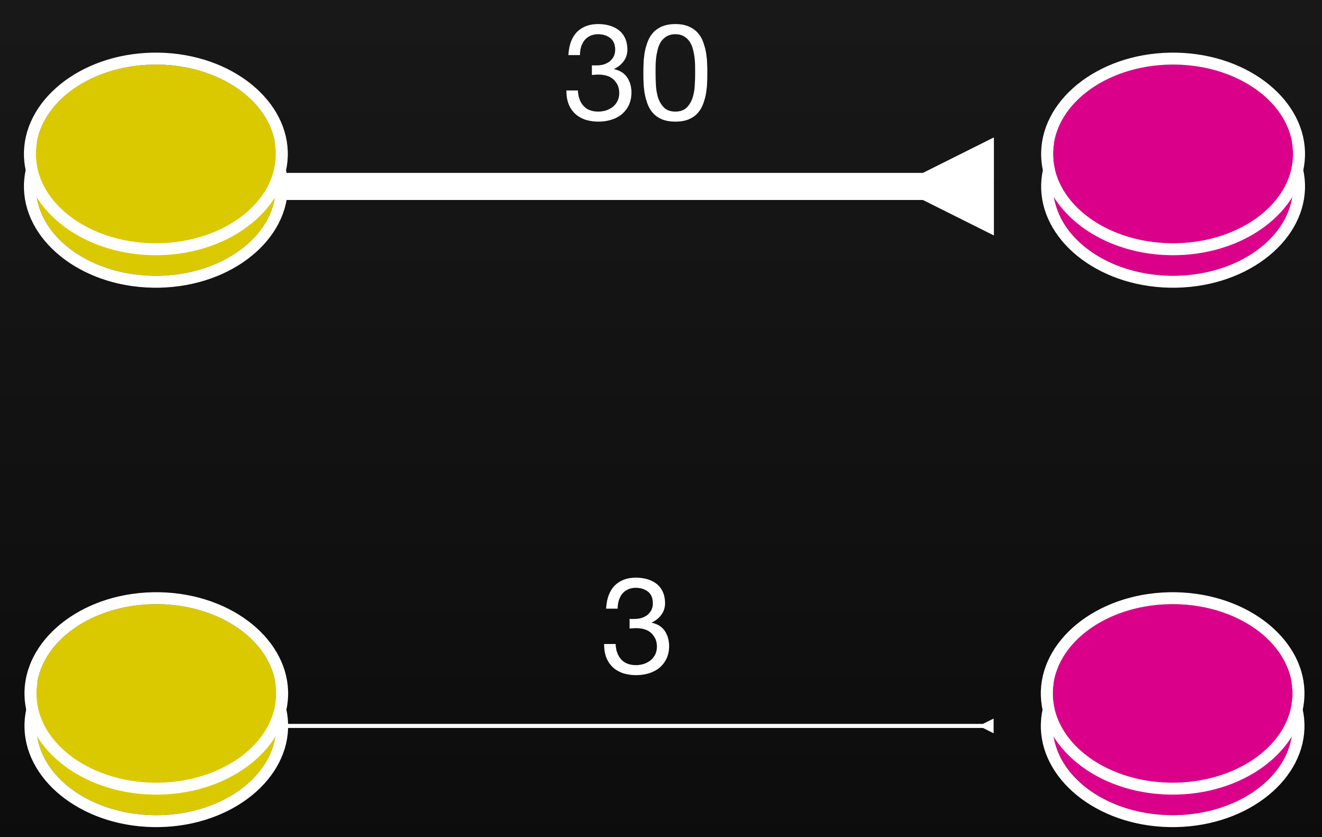


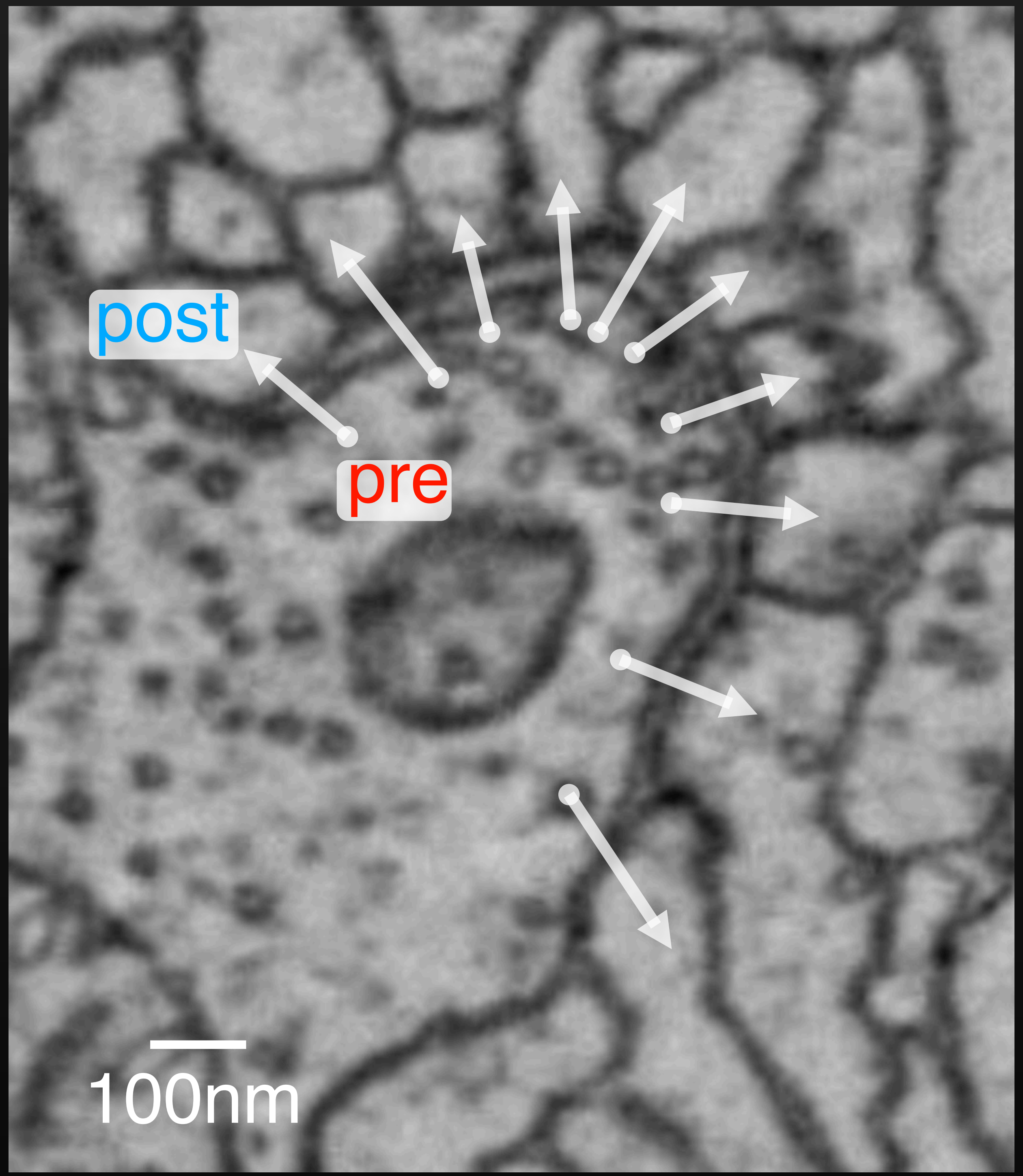
Why does this matter?



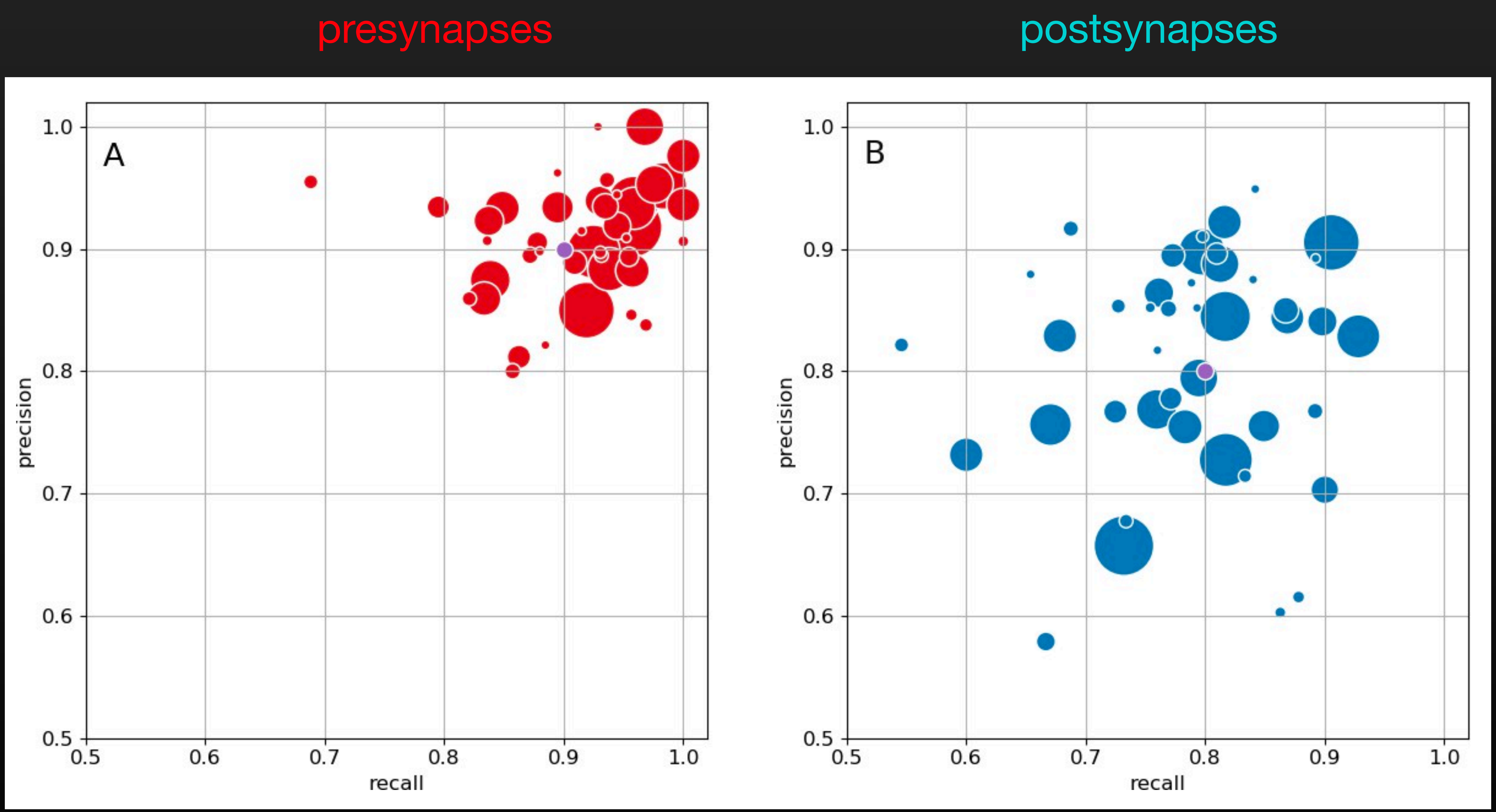
Reasons connections could be different between brains?

- neurons proofread to different degrees
- synapse detection
- developmental abnormality
- biological variability



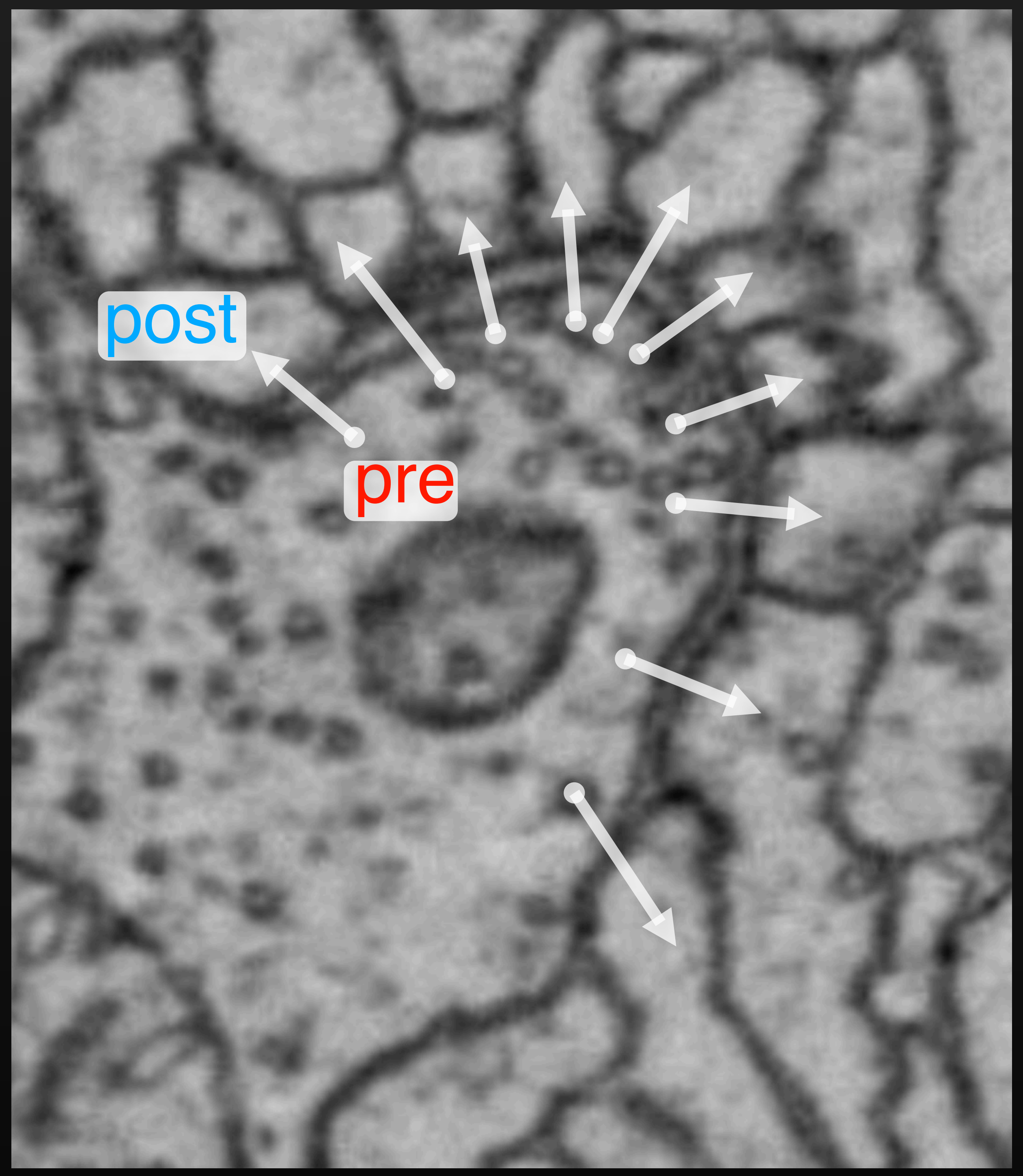


FAFB/FlyWire synapse

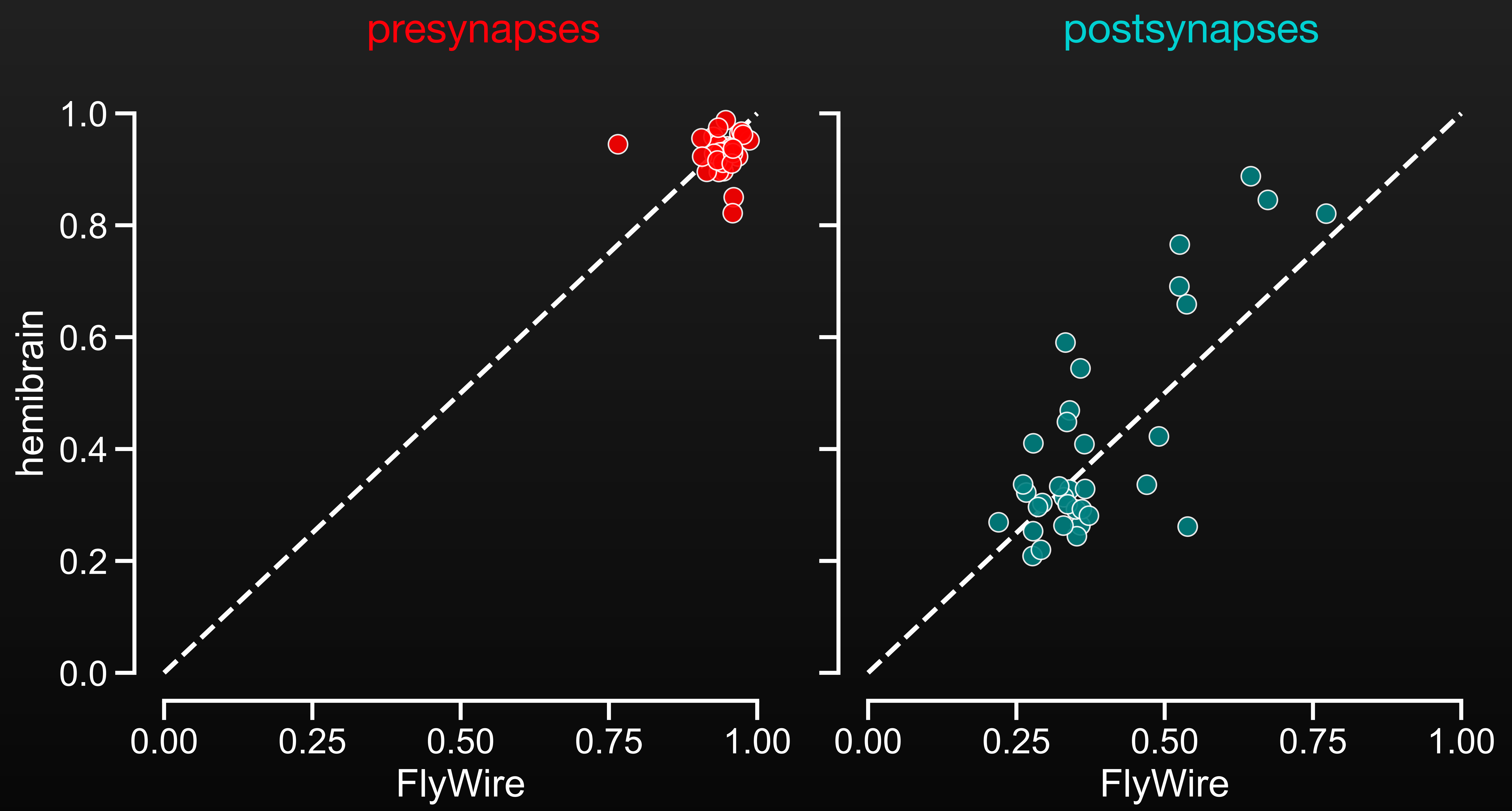


Scheffer *et al., eLife (2020)*

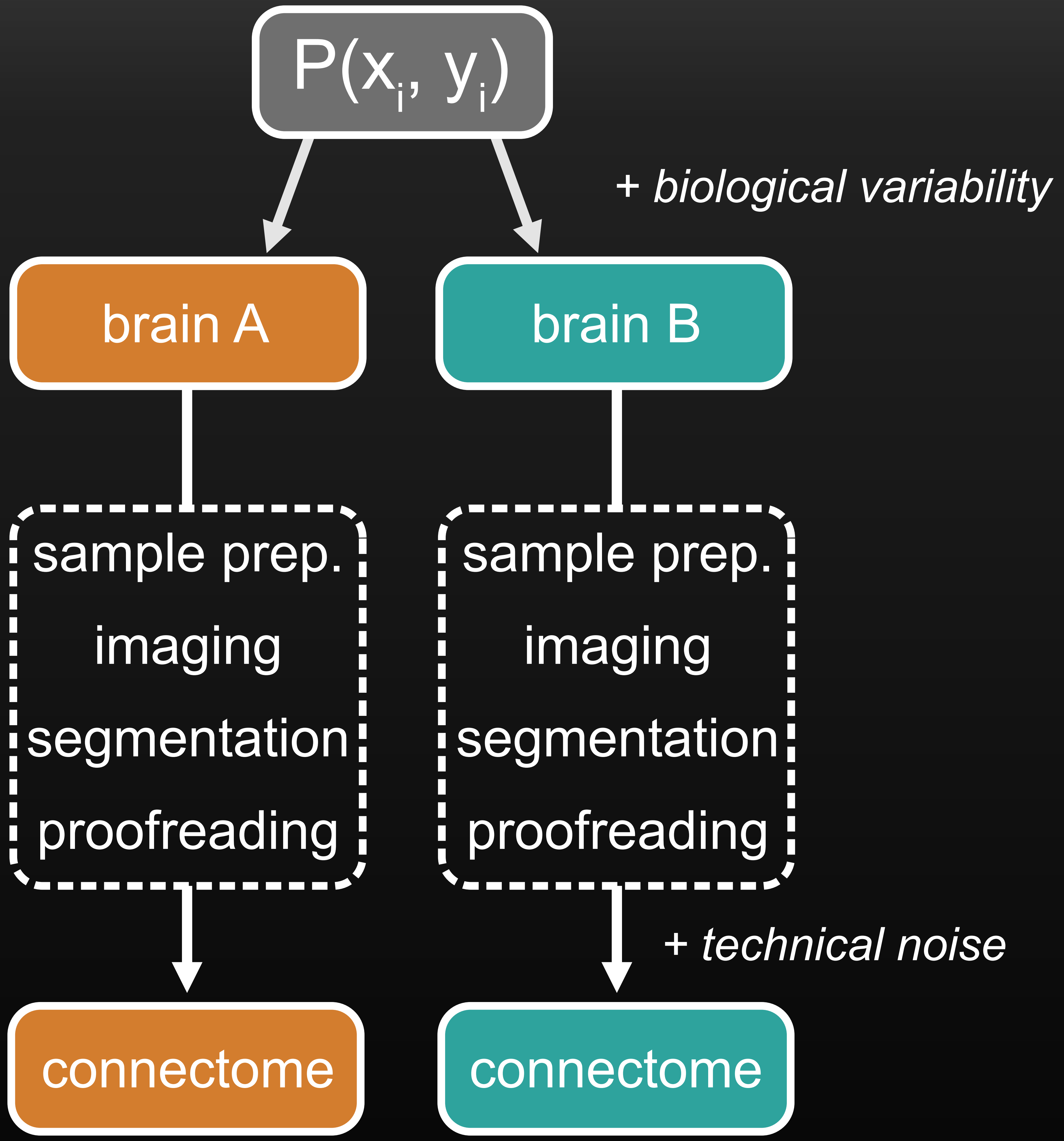
accuracy of synapse detection across neuropils in hemibrain



FAFB/FlyWire synapse

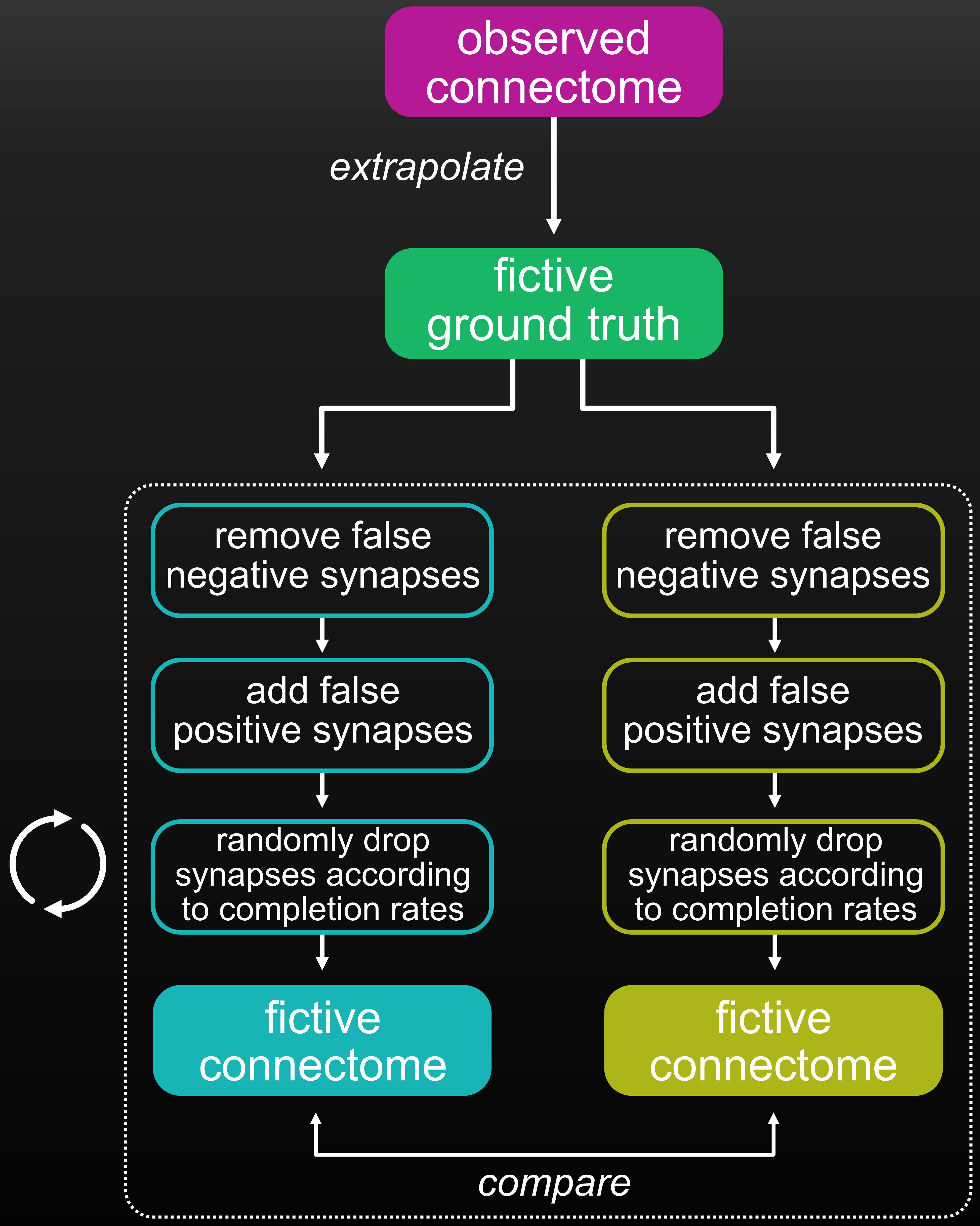


completion rates per brain area

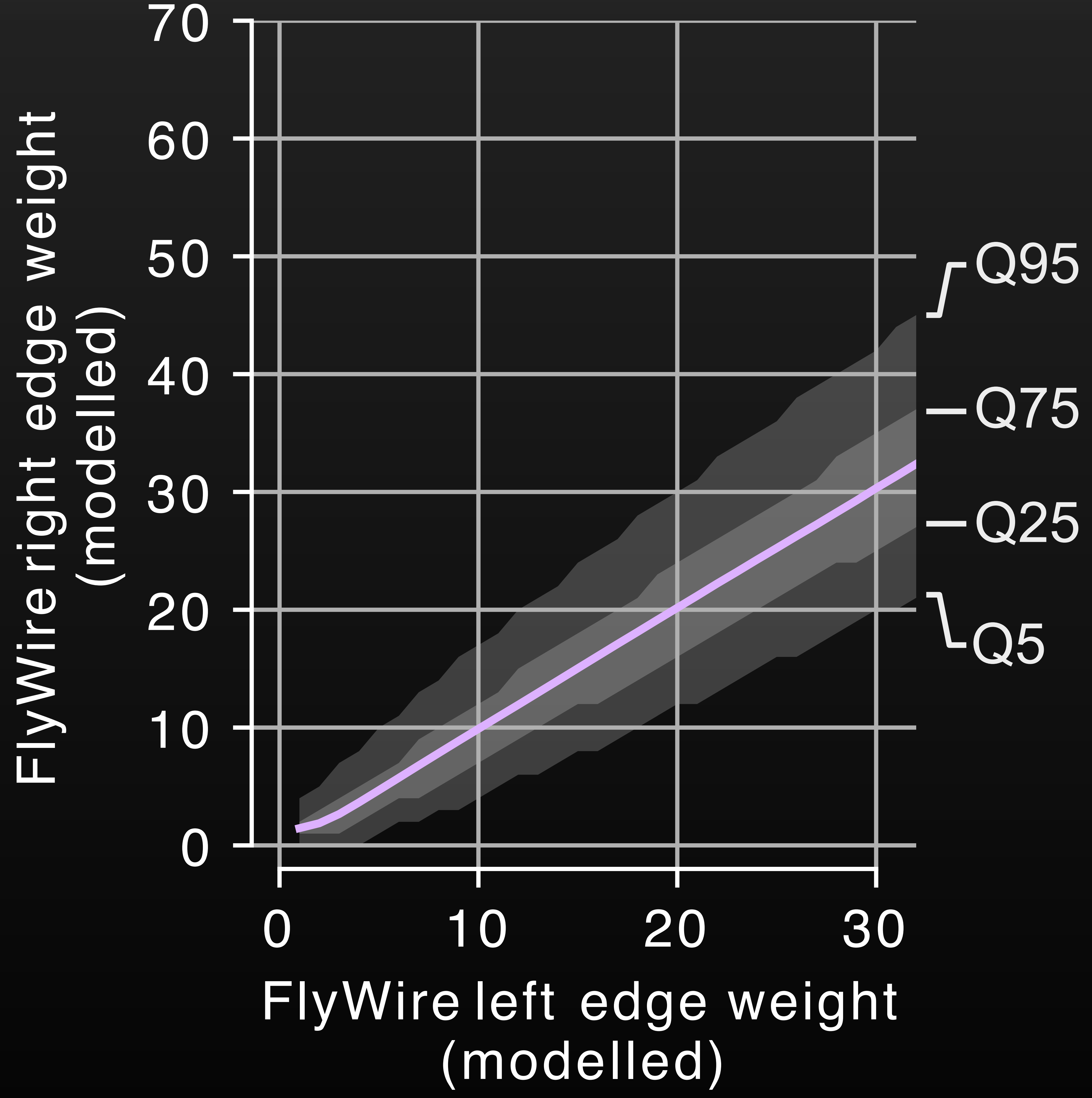


P(x_i, y_i)
+ biological variability
brain A
brain B
sample prep.
imaging
segmentation
proofreading
sample prep.
imaging
segmentation
proofreading
+ technical noise
connectome
connectome

modelling impact of technical noise

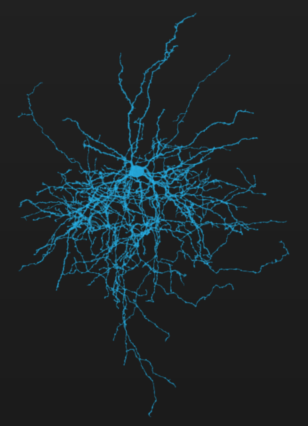


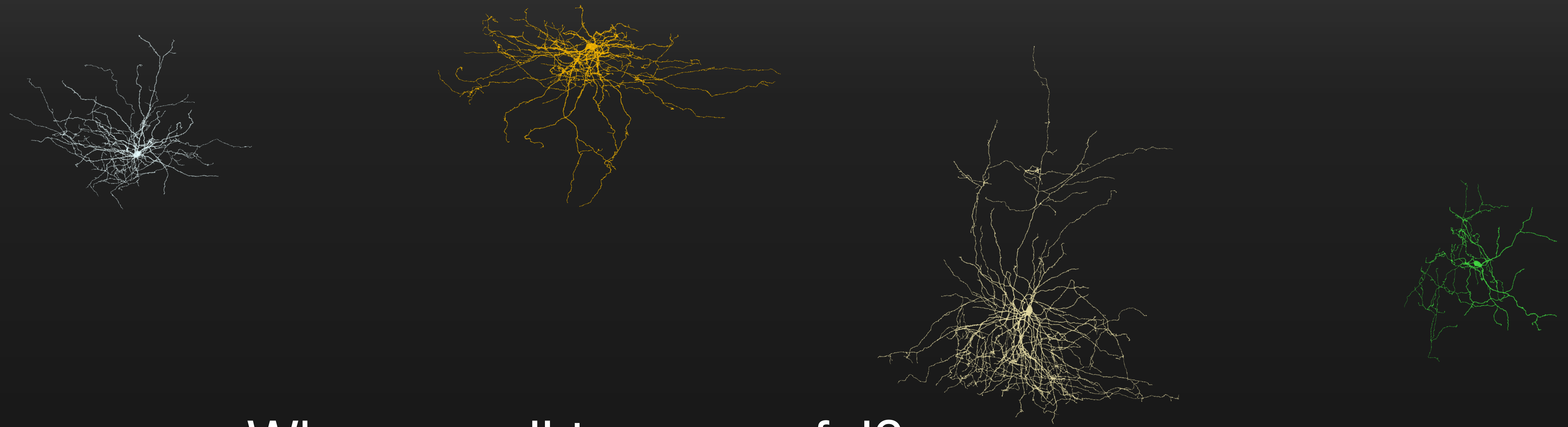
modelled technical noise



edge weight differences of +/- 30%
could be entirely due to technical noise

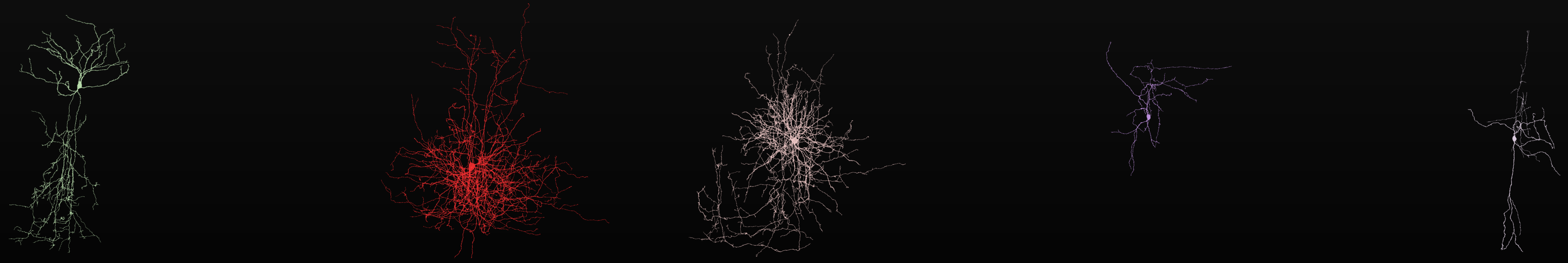
Cell types



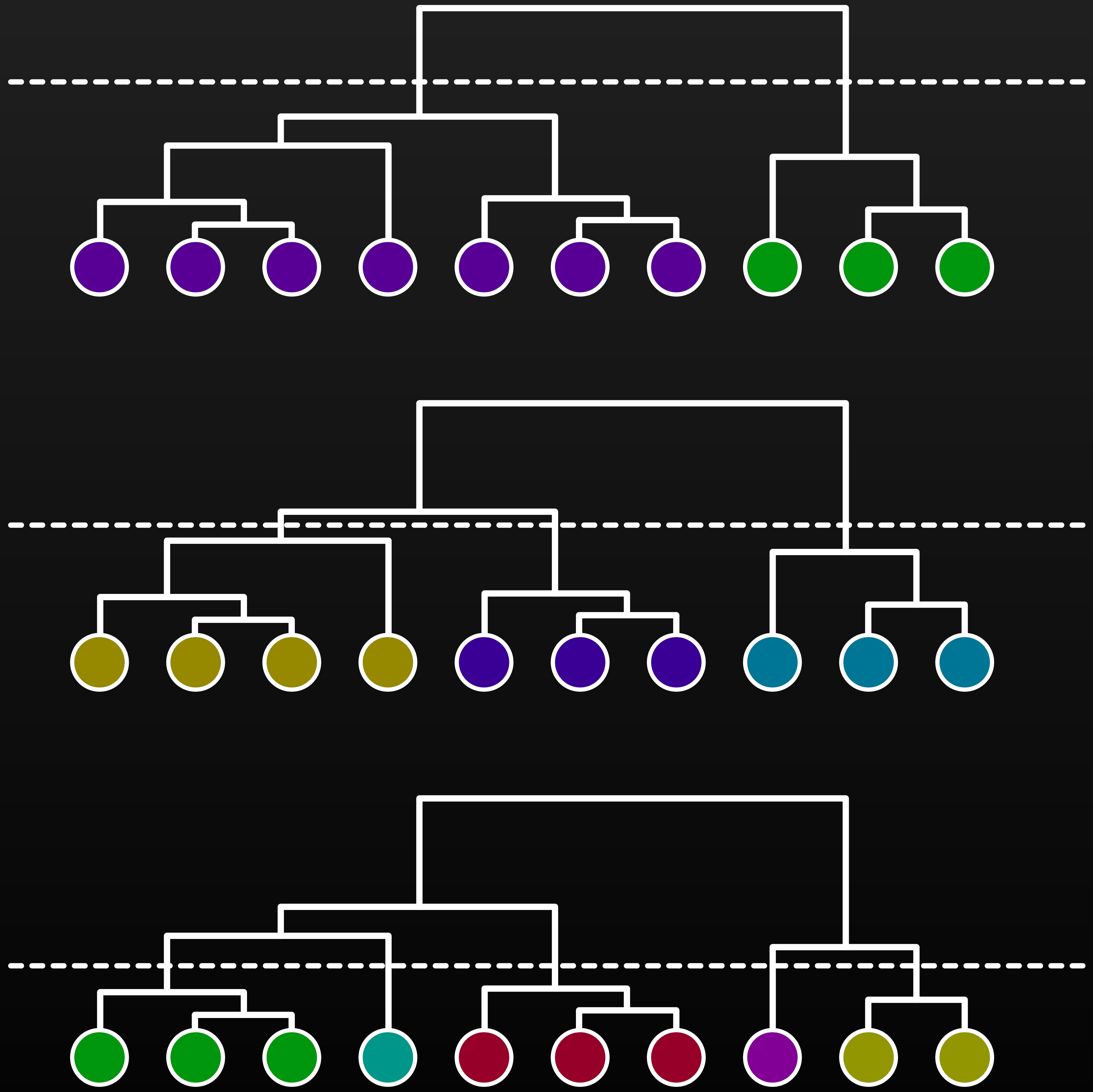


Why are cell types useful?

1. Easier to think about types than individual cells
2. Compress the data
3. Link neurons across datasets/modality (“unit of conservation”)



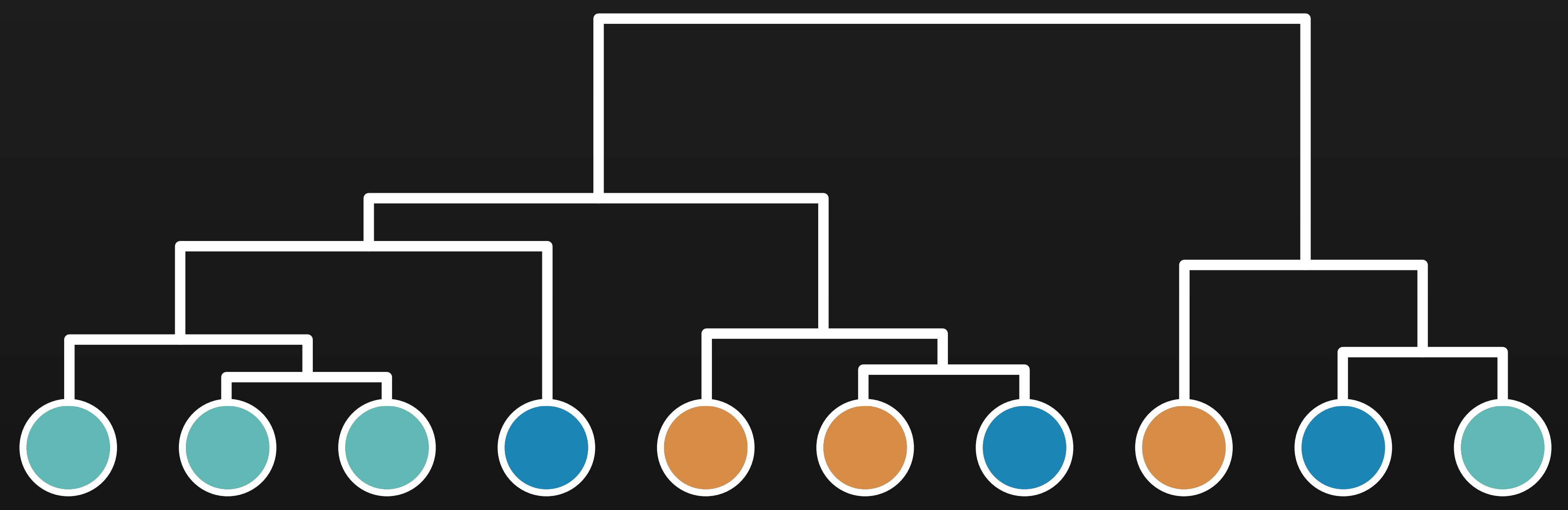
How to define cell types?



Cell types

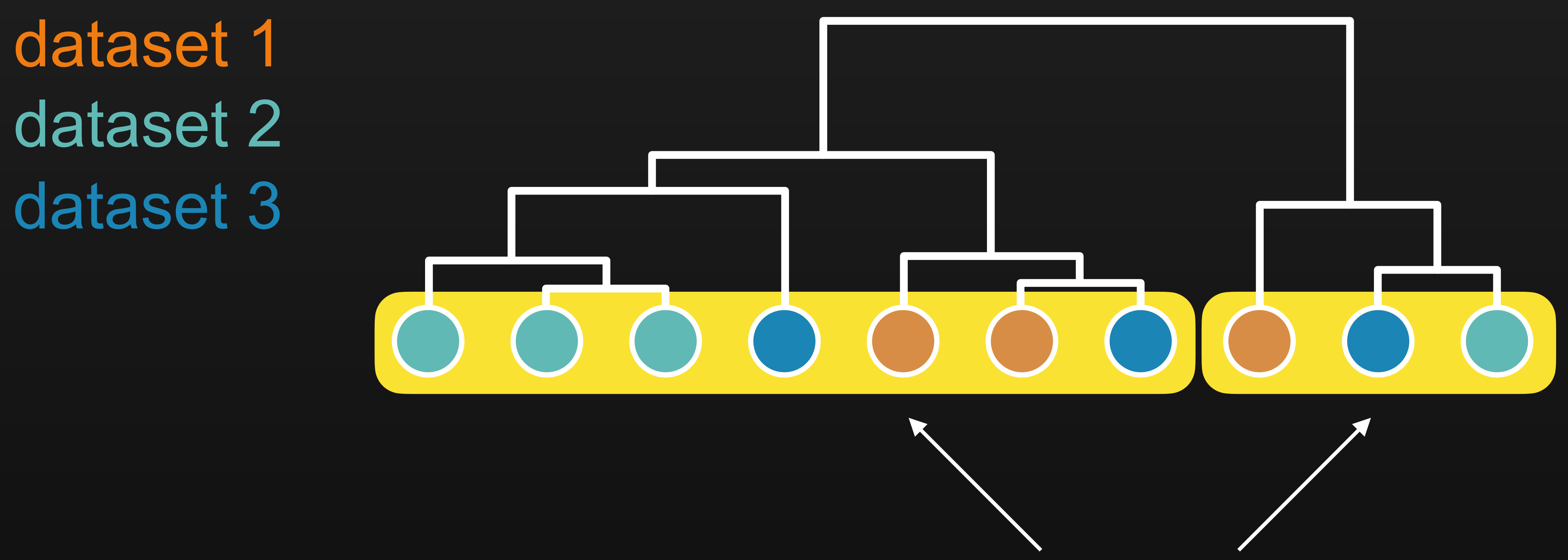
How to define cell types?

dataset 1
dataset 2
dataset 3



Cell types

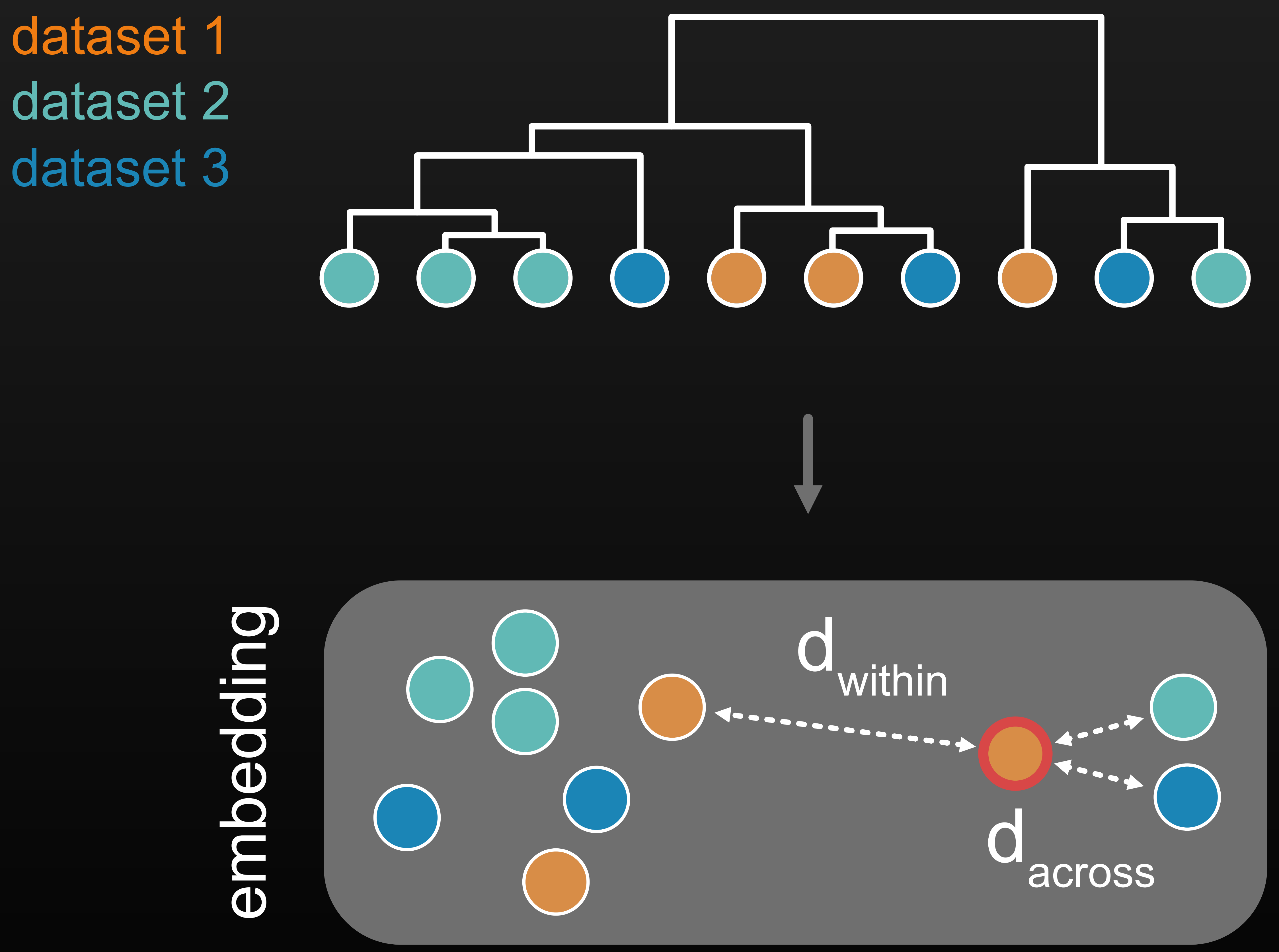
How to define cell types?



cell types = clusters with representatives from all datasets

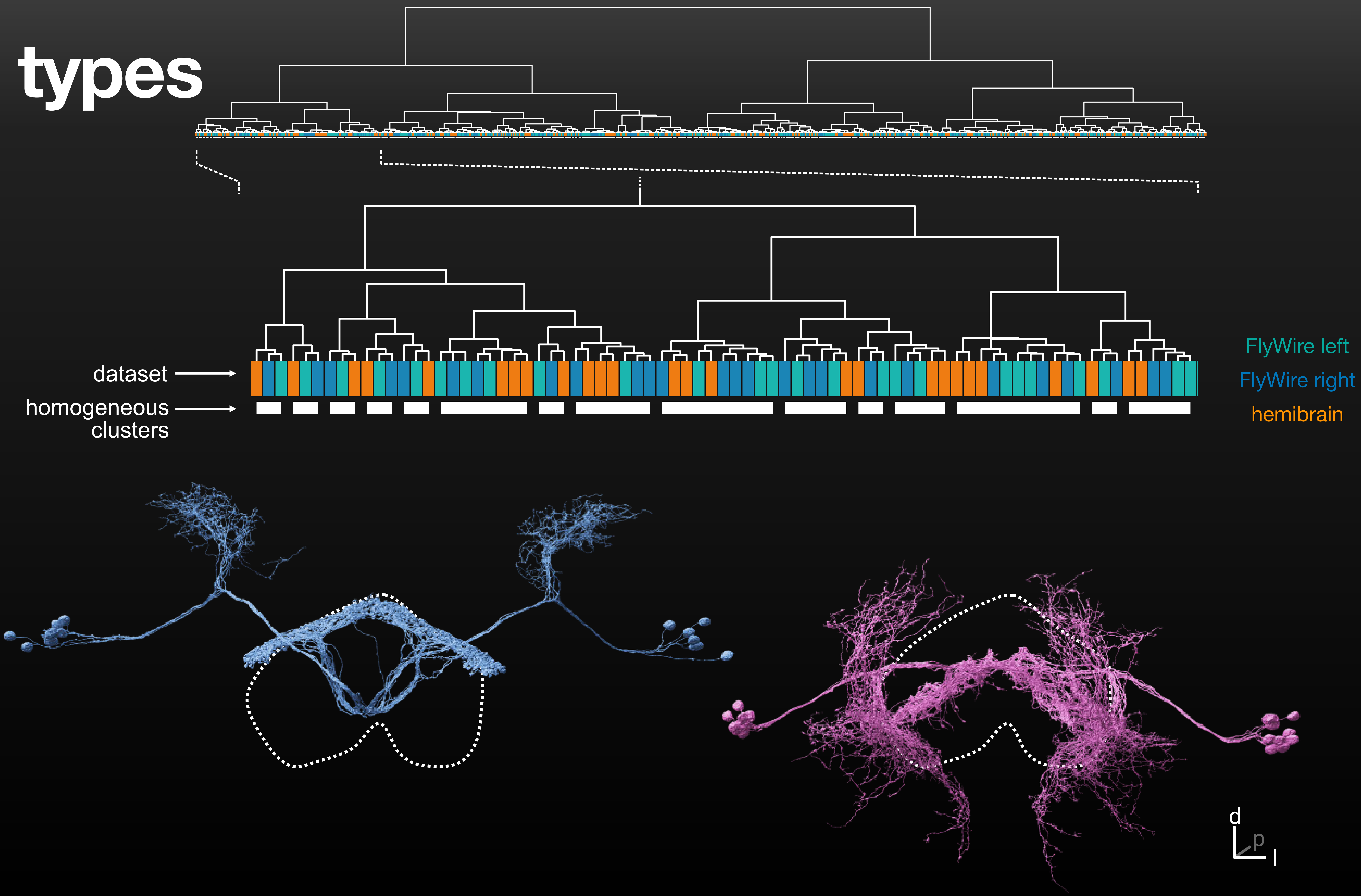
Cell types

How to define cell types?



"A cell type is a group of neurons that is more similar to a group of neurons in another brain than to any other neuron in the same brain"

Cell types



dataset
homogeneous clusters
FlyWire left
FlyWire right
hemibrain
d
p
l

Take homes

1. New adult fly brain connectome now publicly available (as well as complete nerve cord)
2. (Insect) brains aren't snowflakes but the observed variability includes both biological and technical components
3. Robust cell types have to be defined in a way that takes inter-individual varianility into account

Acknowledgments

Cambridge

Alex Bates
Alexandra Fragniere
Andrew Champion
Christopher Dunne
Eva Munelly
Greg Jefferis
Griffin Badalemente
Katharina Eichler
Ladann Kiassat
Laia Serratosa
Márcia Santos
Markus Pleijzier
Marina Gkantia
Marta Costa
Paul Brooks
Tomke Stuerner
Varun Sane
Yijie Yin

Princeton

Mala Murthy
Sebastian Seung

Amy Sterling
Arie Matsliah
Claire McKellar
Sven Dorkenwald
William Silversmith

Allen

Casey Schneider-Mizell
Forrest Collman

FlyWire community

U. Vermont

Davi Bock
Eric Perlman
Tom Kazimiers

UCLA

Volker Hartenstein

Columbia

Gwyneth Card

Janelia

Gerry Rubin

Oxford

Scott Waddell

UNIVERSITY OF CAMBRIDGE

UKRI MRC Laboratory of Molecular Biology

NSF

NIH

Supported by wellcome trust

BRAIN Initiative DBI-2014862
NSF NeuroNex 1RF1MH120679-01