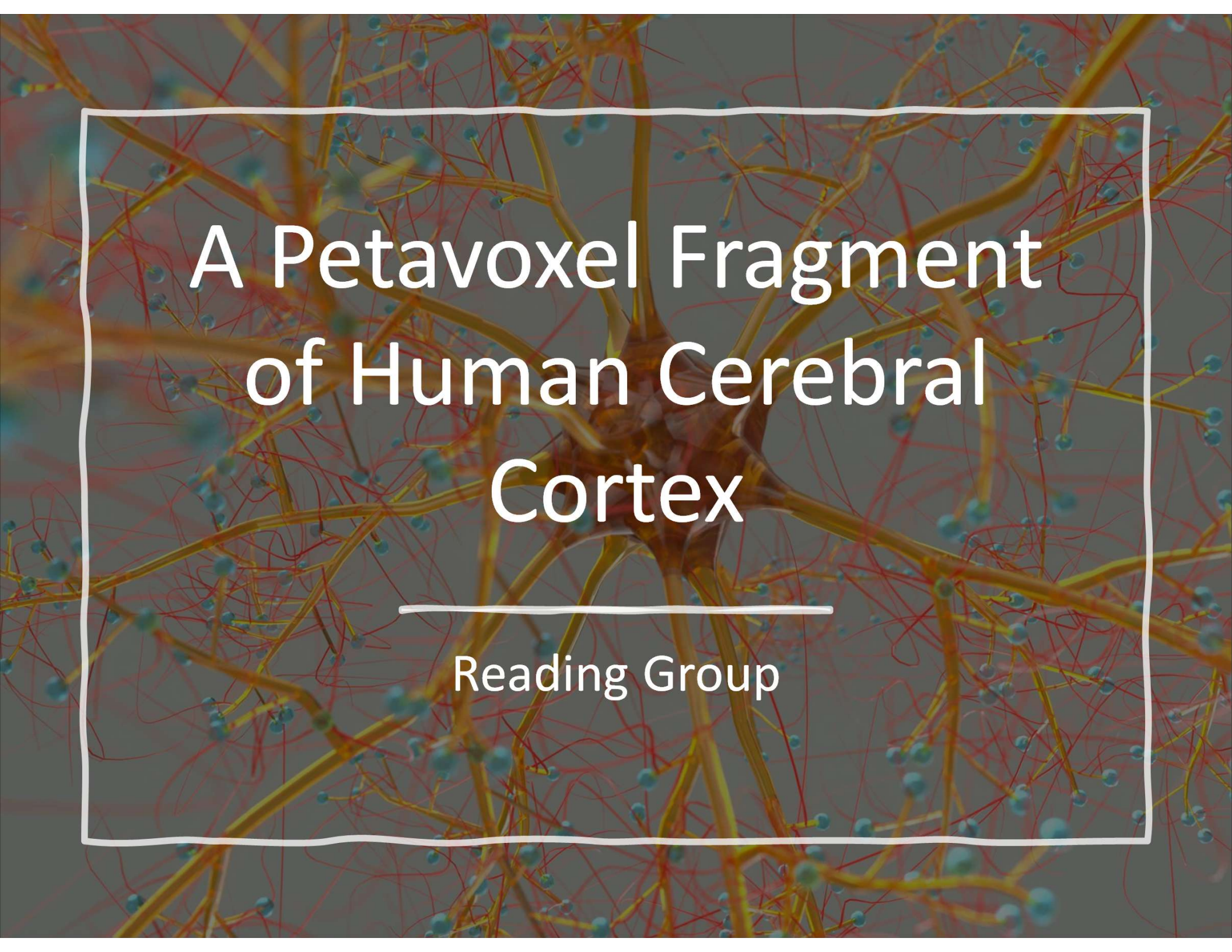


A Petavoxel Fragment of Human Cerebral Cortex

Reading Group

Overview

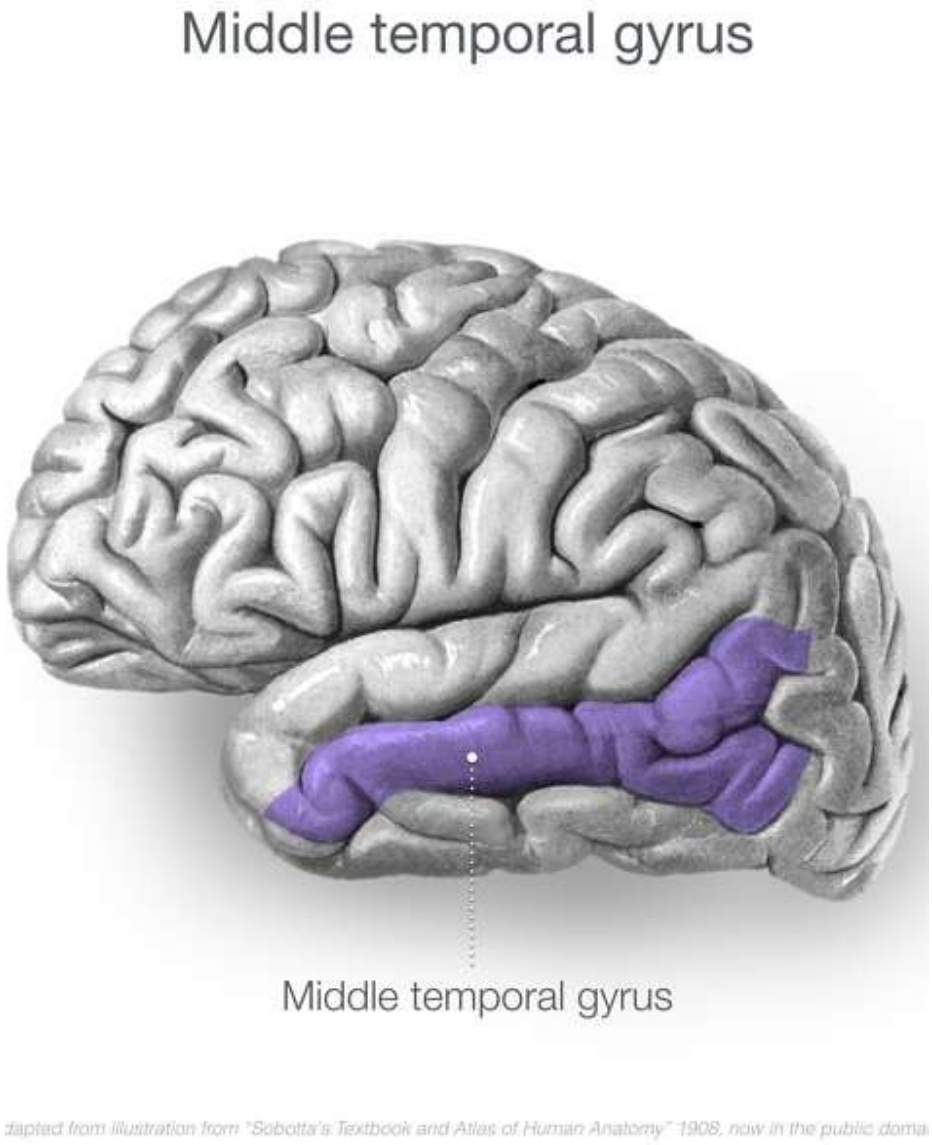
- This talk presents **H01**, a **~1 mm³** sample of human **left anterior middle temporal gyrus** imaged at **nanometer** resolution.
- The workflow converts tissue into a **conservative graph**: **tissue → images → aligned volume → segments → synapses → graph → claims**.
- Each step includes safeguards against **topological errors**, and the **dataset and tools are public** for replication and audits.
- Two focused **findings** are highlighted, each presented with **quantified uncertainty**.

Why This Research Matters

- Enables questions not answerable with LM/MRI: target-specific placement, convergence/divergence, **synapse multiplicity**.
- Community science: public viewers and logs enable replication and adversarial checks.
- Treat as a measurement substrate, not a finished atlas.
- Petavoxel interpretation requires ML with **human control points**.
- Judge models by **graph-level stability** and **error propagation**, not voxel IoU alone.
- Dataset stresses distribution shift relative to mouse EM.

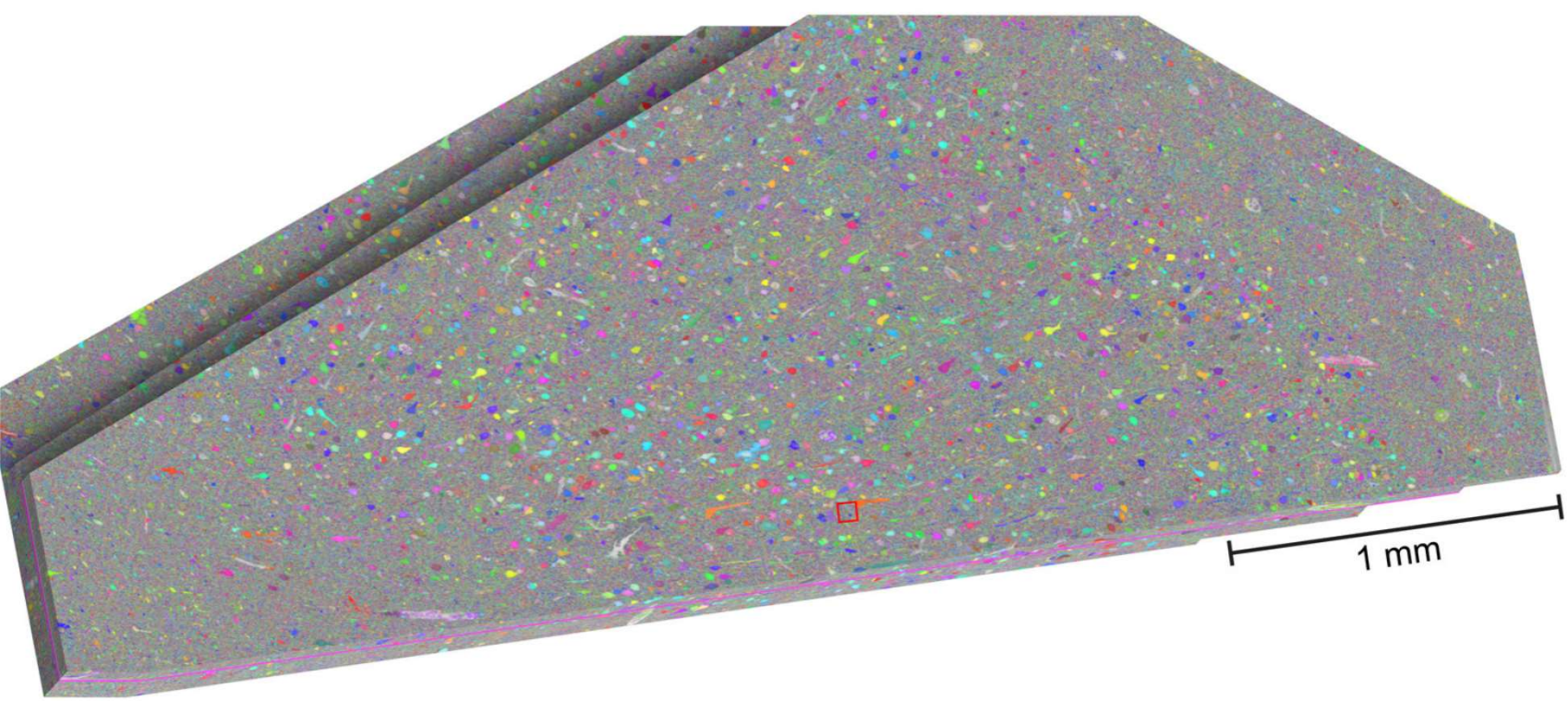
Where Is the Temporal Cortex?

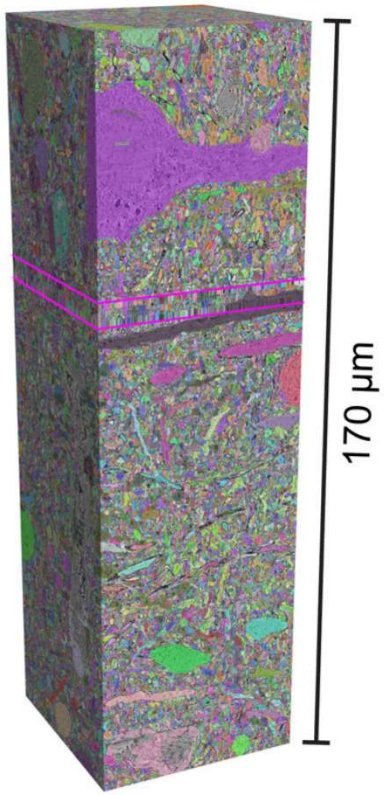
- The sample was from **left anterior middle temporal gyrus (MTG)** involved in **semantics** and **recognition**.
- Association cortex was expected to show **heterogeneous laminar inputs**, **local recurrence**, and **diverse inhibition**.
- Results should be read as **representative**, not a **complete census**, and broader generalizations remain **hypotheses**.



Why Epilepsy Surgery Tissue?

- Tissue access is **ethical**, and **light-microscopy histology** appears **normal**.
- Subtle **disease** or **medication** effects may persist, so quantitative **rates** are **hypothesis-generating**.
- This uncertainty **constrains** functional claims but does **not invalidate** the measurements.





Dataset Overview

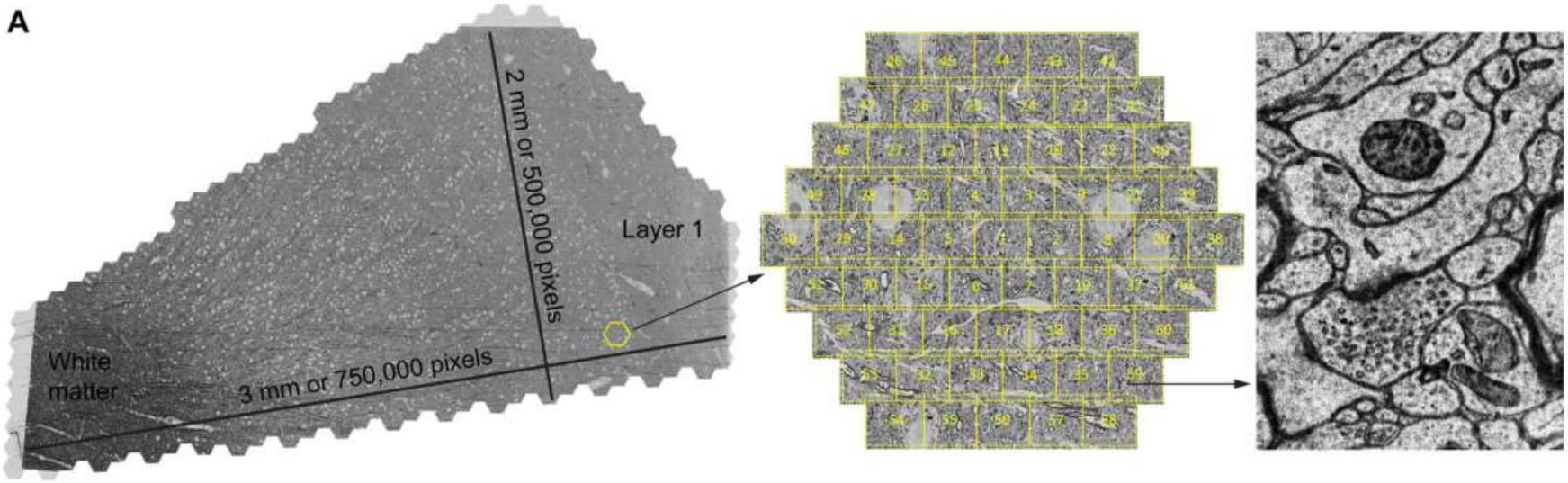
- First human cortical volume at this scale **with usable labels across layers**.
- **Stress-tested** seeding and **anisotropy**; exposed **thin-neurite omissions** and **implausible merges**.
- **Distribution shift** vs mouse EM was evident.

Methods: Sample Preparation (Tissue → Resin)

- The tissue came from **excess anterior middle temporal gyrus** removed during **drug-resistant epilepsy** surgery; **light-microscopy histology** judged the specimen **normal**.
- The sample underwent **rapid immersion fixation** (aldehyde-based) to preserve **ultrastructure** promptly after resection.
- Heavy-metal staining followed a **ROTO** sequence (**reduced osmium–thiocarbohydrazide–osmium**) with additional **heavy-metal contrasts** (for example, **uranyl** and **lead**, per Methods), enhancing membranes and synaptic densities.
- The block was **embedded in epoxy resin**, **trimmed**, and oriented **perpendicular to the pia**, aligning with the **fanned-out axes** of pyramidal **apical dendrites** and long-range **axons**.

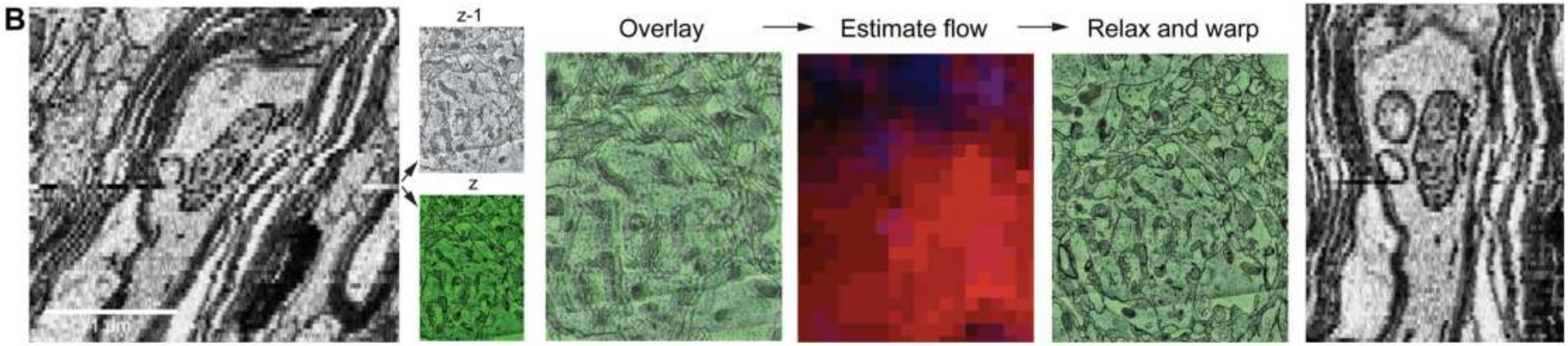
Methods — Sample & Imaging (ATUM–mSEM)

- The dataset reconstructed a **~1 mm³** slab of **left anterior MTG** at **nanometer** resolution.
- Sections were **~33–34 nm**; pixels were **~4 nm** across **5,019** sections.
- Imaging ran **326 days** and produced **~247M tiles (~1.8 PB raw)** → **~196M** stitched tiles.
- The aligned stack was **~1.4 PB**; sampling was **~5.55×4 nm**; volume was **1.05 mm³** (uncompressed).



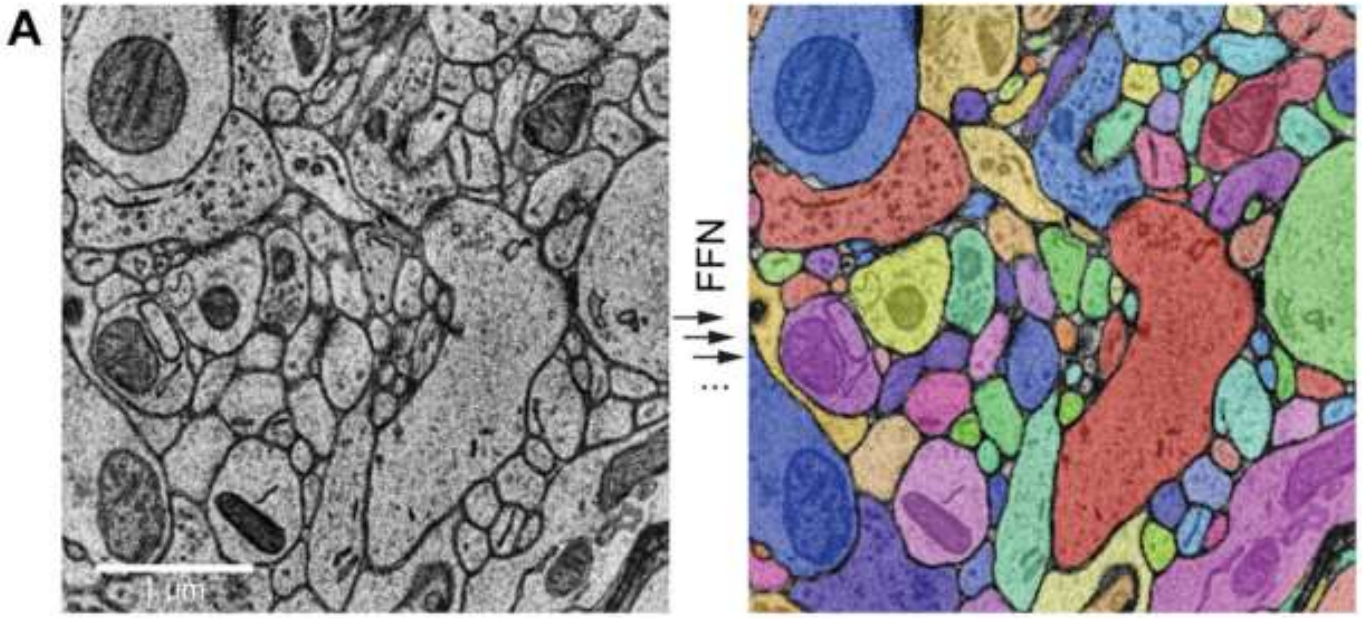
Methods — Alignment & Quality Control

- Elastic alignment corrected **phase joins, drift, tears, and compression**.
- **Residual misalignment** and **damaged sections** were **measured and quarantined**.
- Alignment statistics accompanied the volume and bounded **geometry error**.



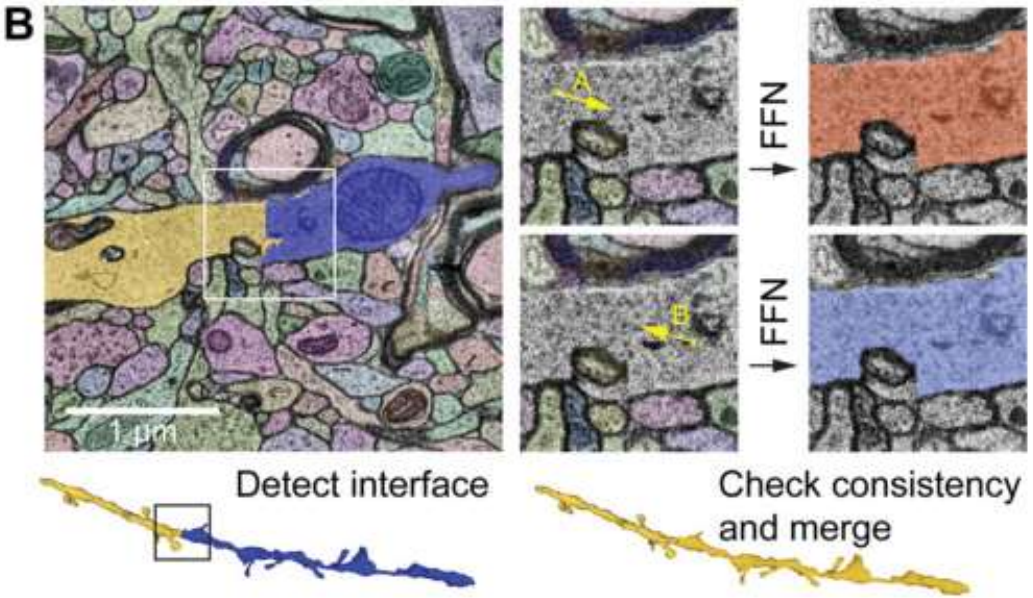
Methods — Segmentation (FFN) & Seeding/Anisotropy

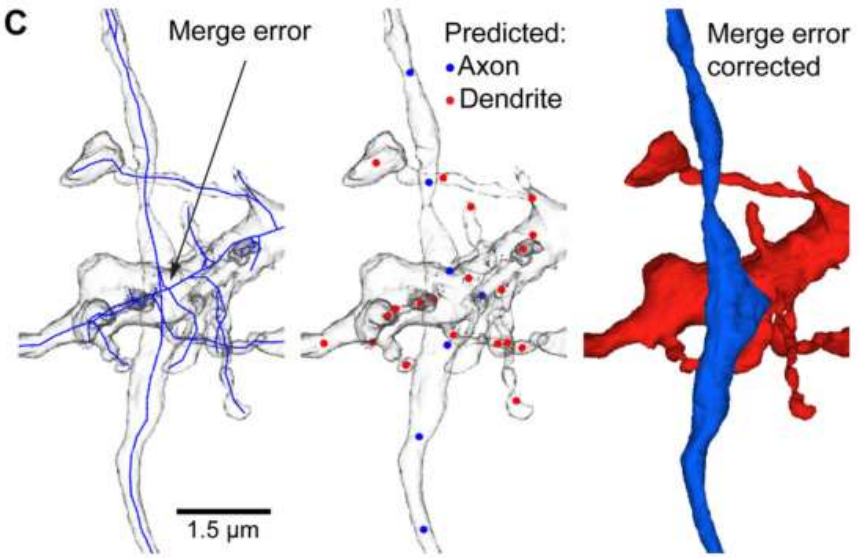
- Segmentation used **FFN** (seed-and-grow, local field of view).
- FFN handled **broken membranes**; **thin neurites** were sometimes missed.
- **Active reseeding** and **anisotropy augmentation** improved thin-structure recall.



Methods — Agglomeration & Operating Points (c2 vs c3)

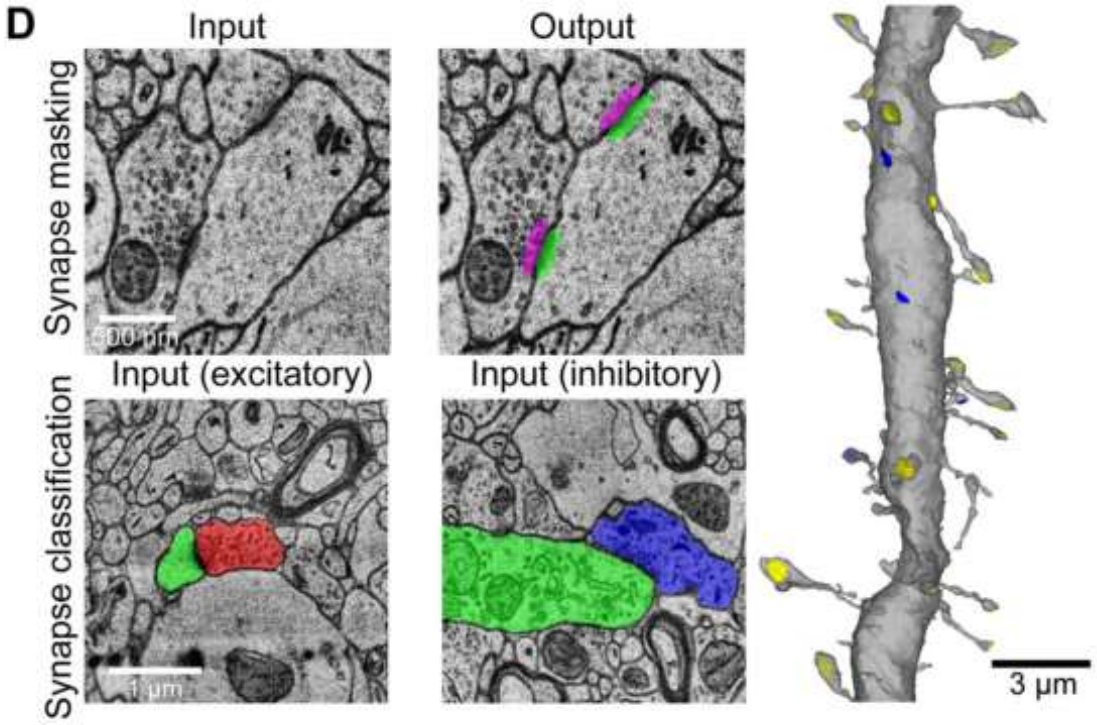
- Fragments merged only with **mutual re-segmentation** (conservative).
- In **104 neurons**: **merges 400→257** and **splits 238→504** (**c2→c3**).
- Detached spines changed **32.2%→33.7%**; analyses favored **c3** to avoid shortcuts.





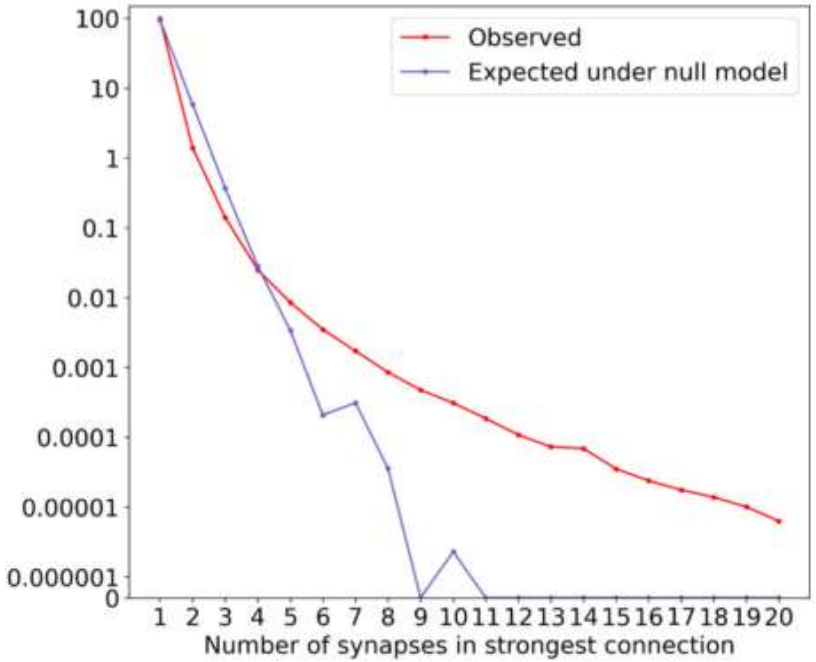
Methods — Synapse Detection & E/I Typing

- A **U-Net–style CNN** produced **voxel-wise synapse probability maps**.
- **Connected components** formed candidates; a **linking step** assigned **pre- and postsynaptic** partners.
- Thresholds kept **low FDR** (**~3.2% E**, **~2.7% I**) with **FN** at **~11% (E)** and **~35% (I)**.
- Outputs fed **E/I typing**, **proofreading triage**, and **error propagation** to metrics



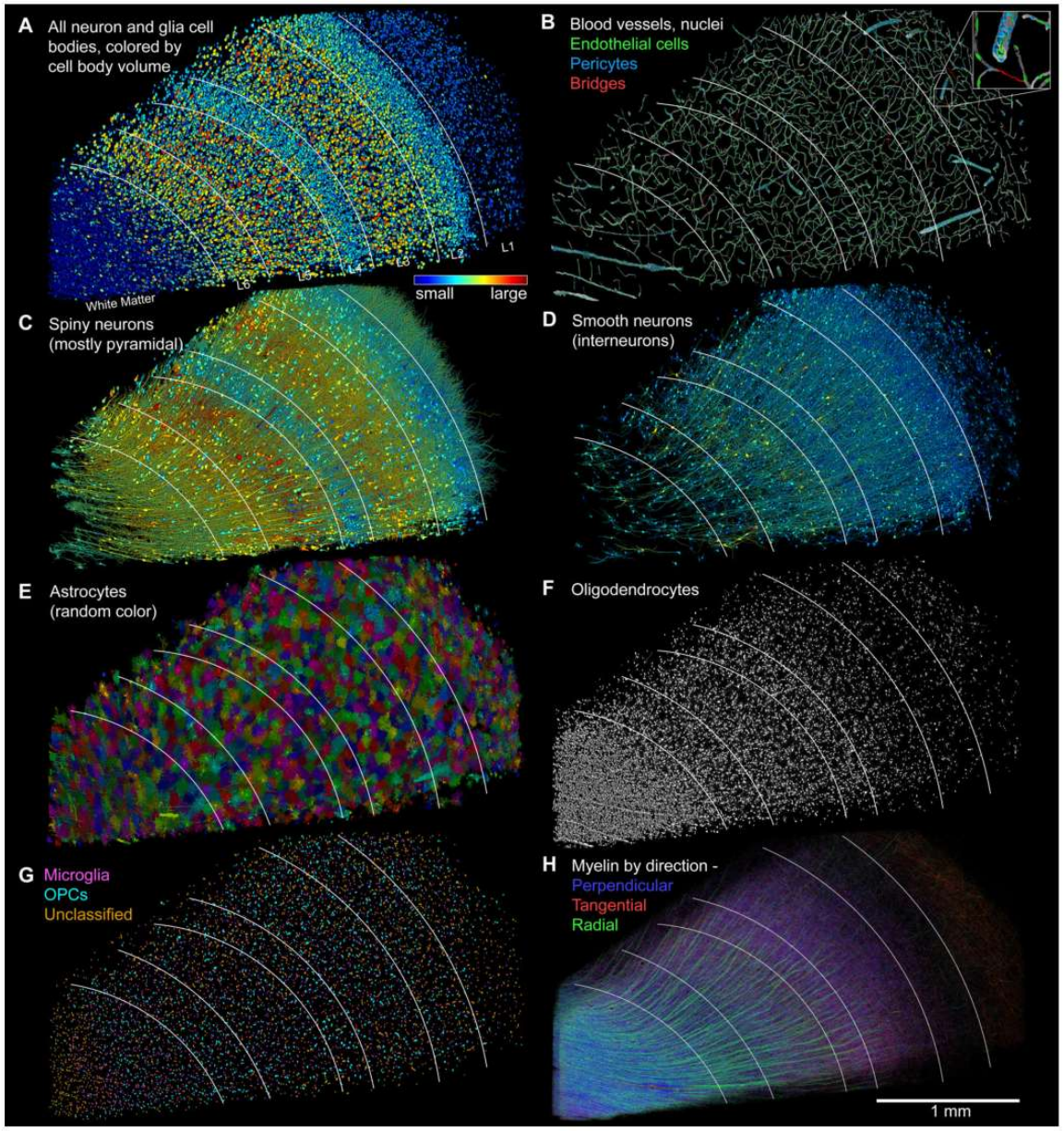
Methods — Graph, Border Controls, and Null

- The graph used **cells** as nodes; **edge weight = synapses per pair**.
- **Risky reattachments** were **avoided**; a **border jackknife** tested truncation.
- A **constrained shuffle** preserved **axon paths** and **bouton counts** but randomized targets.



Results — Cellular Composition

- Total labeled cells: **~57,000** (parenchymal **49,080** + vessel-associated **~8,136**).
- Neurons: **10,531 spiny**, **4,688 interneurons**; glia: **20,139 oligodendrocytes**, **5,474 astrocytes**, **6,702 microglia/OPCs**.
- Vasculature: **4,604 endothelial**, **3,549 perivascular**, and **74** bloodless bridges.
- Myelinated-axon orientation: **L1 tangential**, **cortex→WM radial**, **WM perpendicular**, **L3–L6 diagonal**.

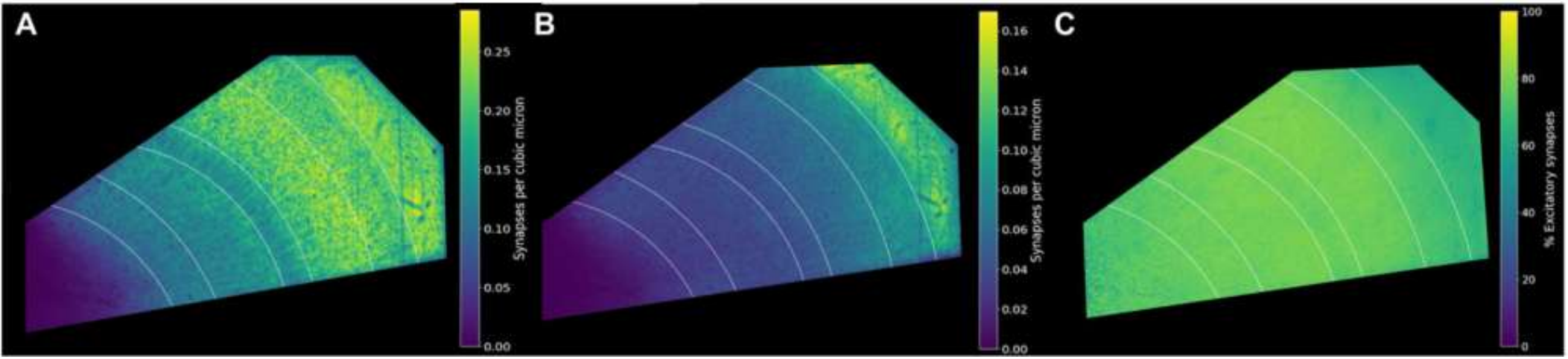


Results — Synapse Totals and E/I Balance

- The detector found **149,871,669 synapses** (**111,272,315 E**; **38,599,354 I**).
- After correction: **E ≈111.6M (61.2%)**; **I ≈70.7M (38.8%)**; **total ≈182.3M**.
- The rise mainly reflected **higher FN** for inhibitory synapses.

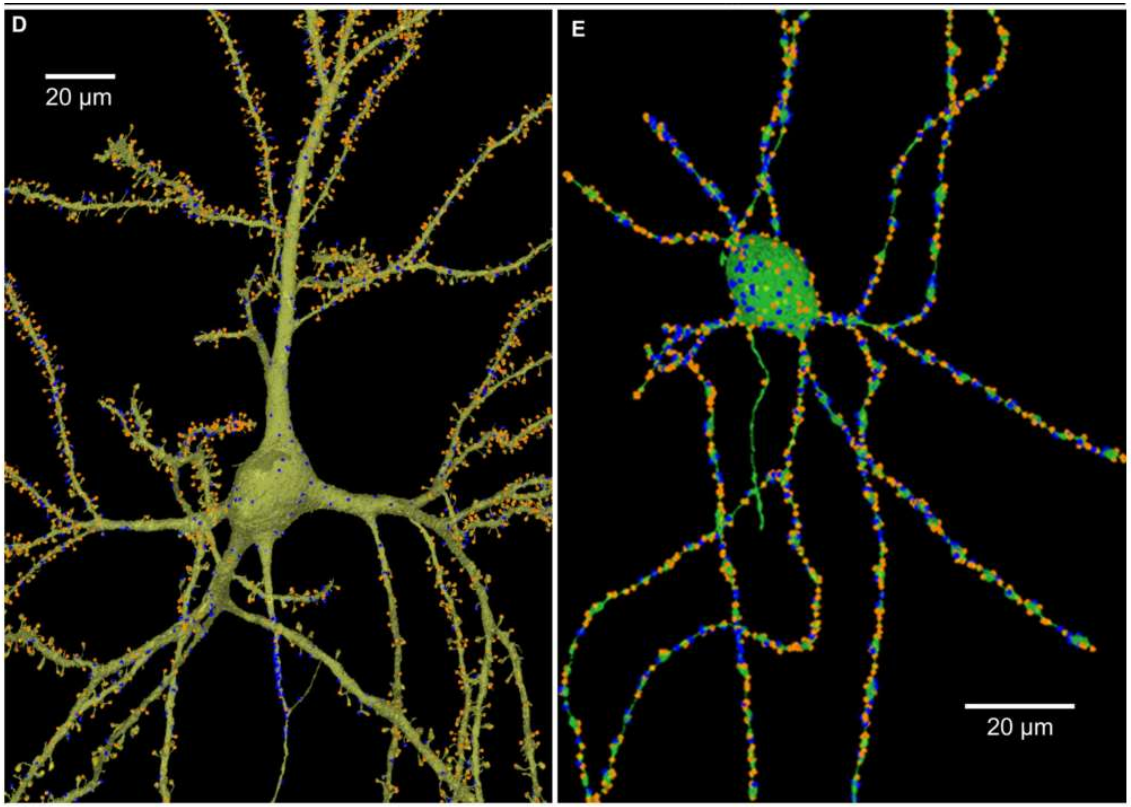
Results — Layer Profiles (E, I, and E-Fraction)

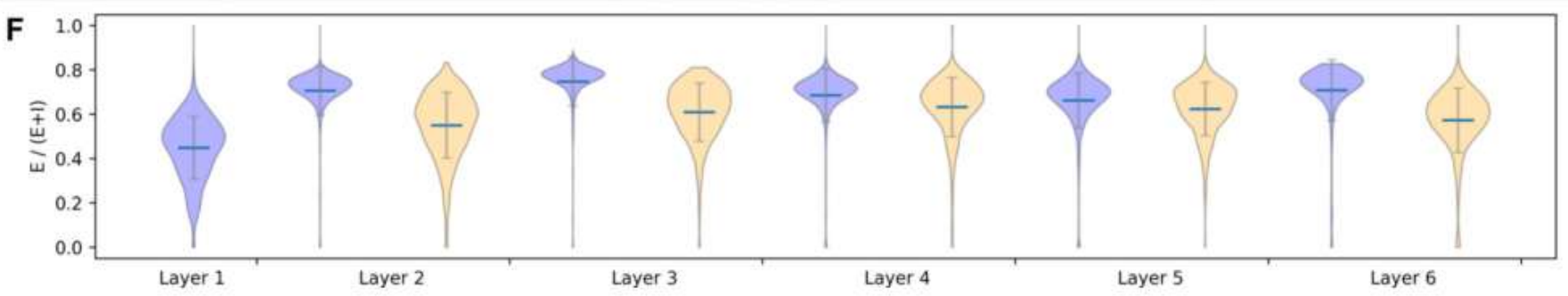
- **E density** was **highest in L1 and L3**.
- **I density** was **elevated in L1**.
- The **E/(E+I)** fraction was **higher in L1 and L3**.



Results — Postsynaptic Targets and Cell-Class Differences

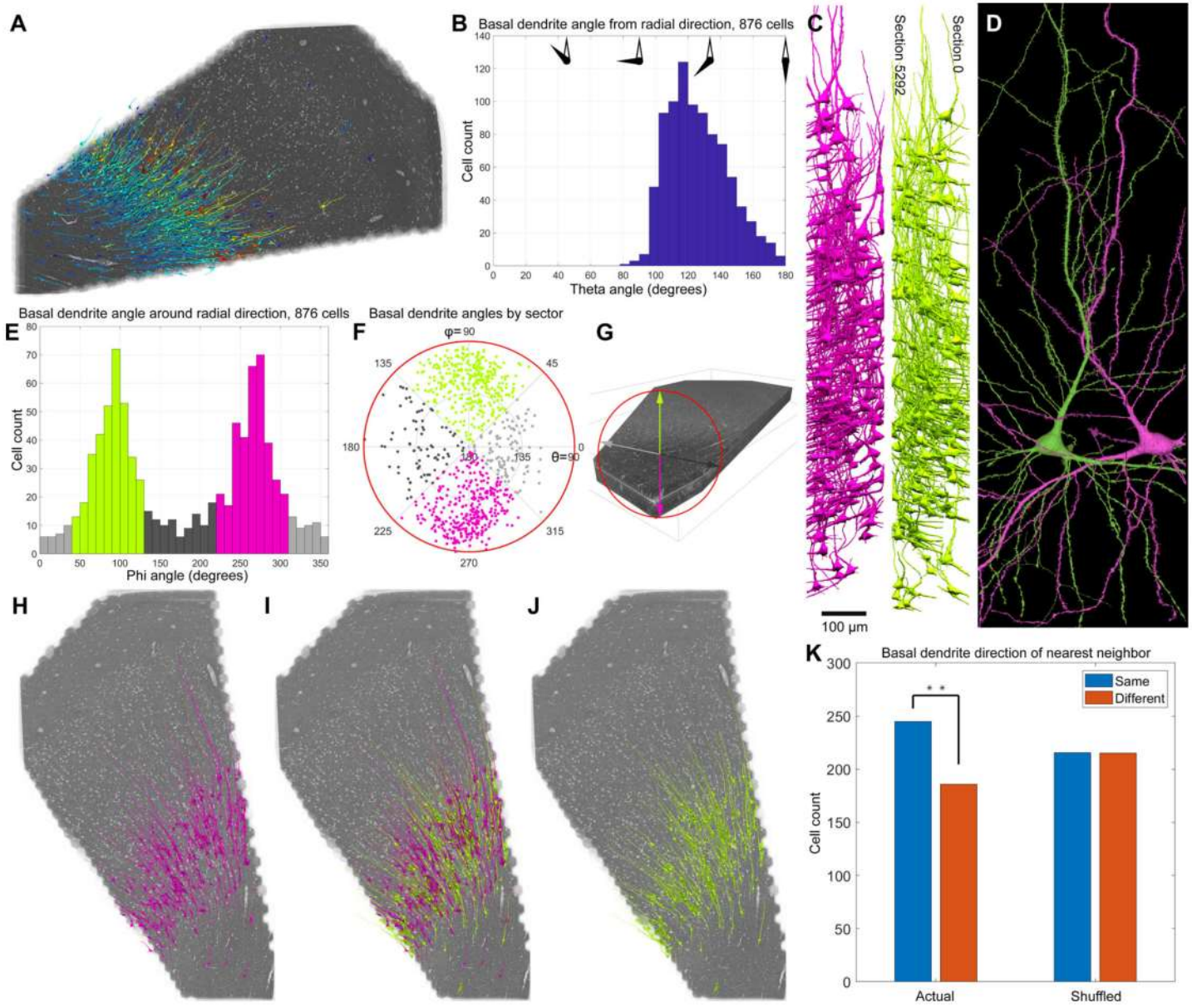
- Of **≈133.7M** with targets: **~99.4% dendrites**; **0.394% soma**; **0.197% AIS**.
- **Pyramidal neurons** received a **higher E fraction** than **interneurons** (**L2–L6**).





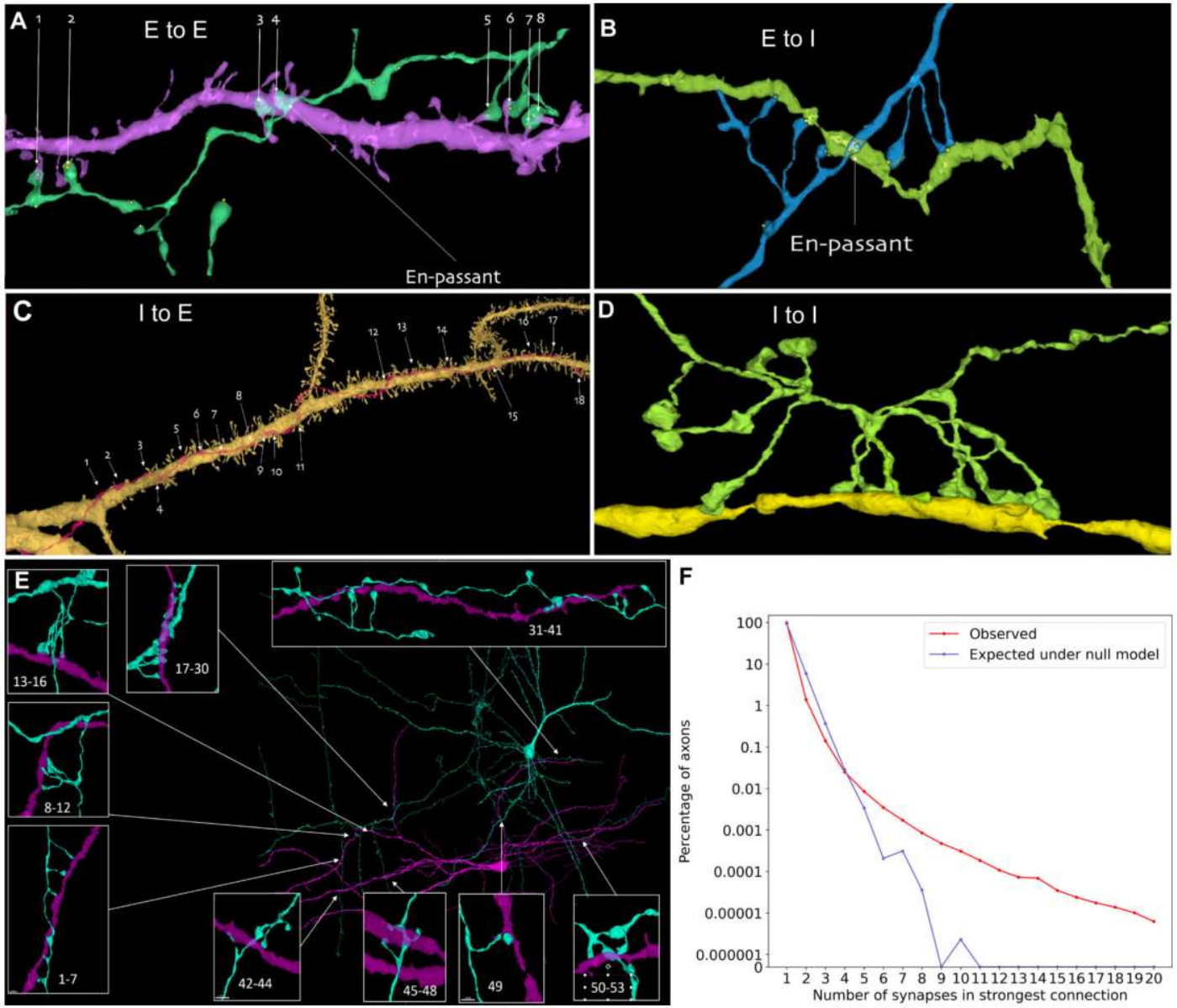
Finding 1 — “Compass” (Triangular) Neurons

- Deep-layer excitatory cells showed **bimodal basal-dendrite orientations**.
- Sample: **n=876** (central **n=431**); **like-with-like** clustering was **significant** (**Fisher p=0.005**).
- Mechanisms were proposed: **anisotropic input**, **guidance**, or **packing**.



Finding 2 — Rare Multi-Synaptic Pairs (Heavy Tail)

- Incidence by pair: **1: 96.49%**; **2: 2.99%**; **3: 0.35%**; **≥4: 0.092%** (examples to **~53**).
- The **observed** curve exceeded the **constrained null** for **N>3**.
- Among targets with **≥3,000 inputs**, **~39%** had a **≥7-synapse** partner; **c3** splits likely **underestimated** strength.



Limitations

- A **single subject** and **single area** limited **external validity**.
- **Truncation** from the **1 mm^3** field biased **paths** and **multiplicity**.
- **Residual merges/splits** and **E/I typing errors** persisted; **no physiology** anchored function.

Future Directions

- Test **global-context 3D transformers** against **FFN** with **topology-aware** metrics.
- Lower **I-synapse FN** and co-train **segmentation + typing**.
- Allocate proofreading for **maximum graph-uncertainty reduction**.

Summary, Conclusions & Validity

- **Human synapse-scale reconstruction** at **petavoxel** size was **feasible** with **risk-aware** choices that protect topology.
- **Public data and tools** enabled **replication** and **adversarial audits** by the community.
- Two **calibrated findings** were supported by **controls** and **sensitivity analyses** with **explicit scope** limits.
- Claims were **bounded** by **measured error rates** (FN/FDR), the **c2/c3** trade-off, and **truncation controls**.

Open Tools for ML and Neuroscience

- **Neuroglancer** provides **reproducible, shareable views** via stateful links.
- **CAVE** and **CREST** route **proofreading effort** and support **strong-pair** and **convergence** queries.
- **VAST** supports **manual edits and measurements** with **logged histories** for auditing.
- Viewer state is informative but does **not replace proofreading logs** as evidence.

https://h01-release.storage.googleapis.com/explore.html

Reference:

"A Petavoxel Fragment of Human Cerebral Cortex" (Science, Shapson-Coe et al., 2021)