JC

RK

09 - 12 - 2015

proteomics and the brain

Cell type– and brain region–resolved mouse brain proteome

Kirti Sharma[1], Sebastian Schmitt[2,3], Caroline G Bergner[2,3], Stefka Tyanova[1], Nirmal Kannaiyan[4], Natalia Manrique-Hoyos[2,3], Karina Kongi[5], Ludovico Cantuti[2,3], Uwe-Karsten Hanisch[6], Mari-Anne Philips[5], Moritz J Rossner[2,4], Matthias Mann[1] & Mikael Simons[2,3]

Brain transcriptome and connectome maps are being generated, but an equivalent effort on the proteome is currently lacking. We performed high-resolution mass spectrometry–based proteomics for in-depth analysis of the mouse brain and its major brain regions and cell types. Comparisons of the 12,934 identified proteins in oligodendrocytes, astrocytes, microglia and cortical neurons with deep sequencing data of the transcriptome indicated deep coverage of the proteome. Cell type–specific proteins defined as tenfold more abundant than average expression represented about a tenth of the proteome, with an overrepresentation of cell surface proteins. To demonstrate the utility of our resource, we focused on this class of proteins and identified Lsamp, an adhesion molecule of the IgLON family, as a negative regulator of myelination. Our findings provide a framework for a system-level understanding of cell-type diversity in the CNS and serves as a rich resource for analyses of brain development and function.

proteomic profiling
of human samples
for CEGS/psychENCODE

overview

BRAINS!

PROTEINS!

1. approach to getting max coverage of the proteome (comp. to RNA)
2. analysis of basic CNS cell-types
3. analysis of brain enriched proteins
4. analysis of brain-region enriched proteins
5. some specific in-vivo followup/validation

Figure 1 | transcriptome vs. proteome

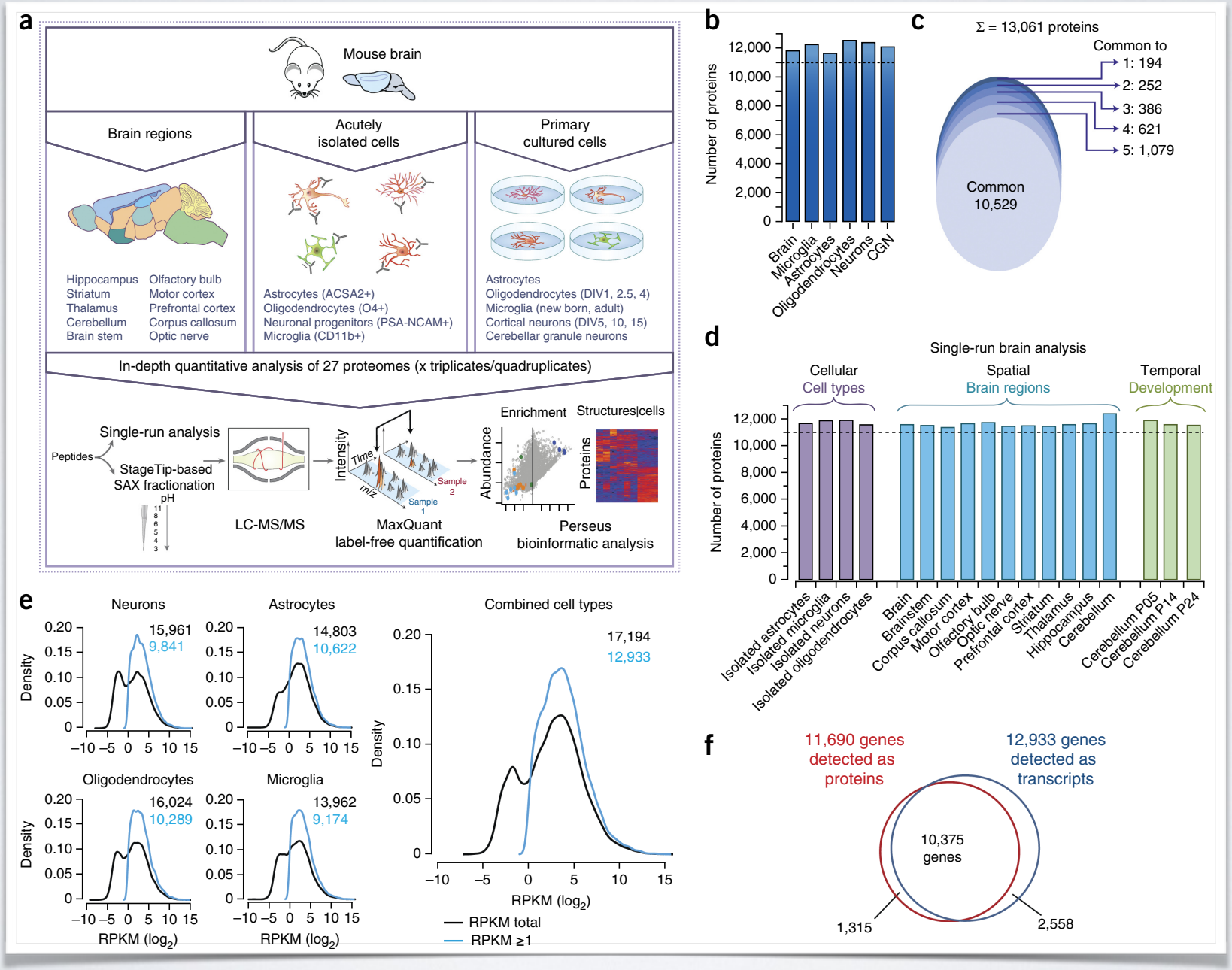
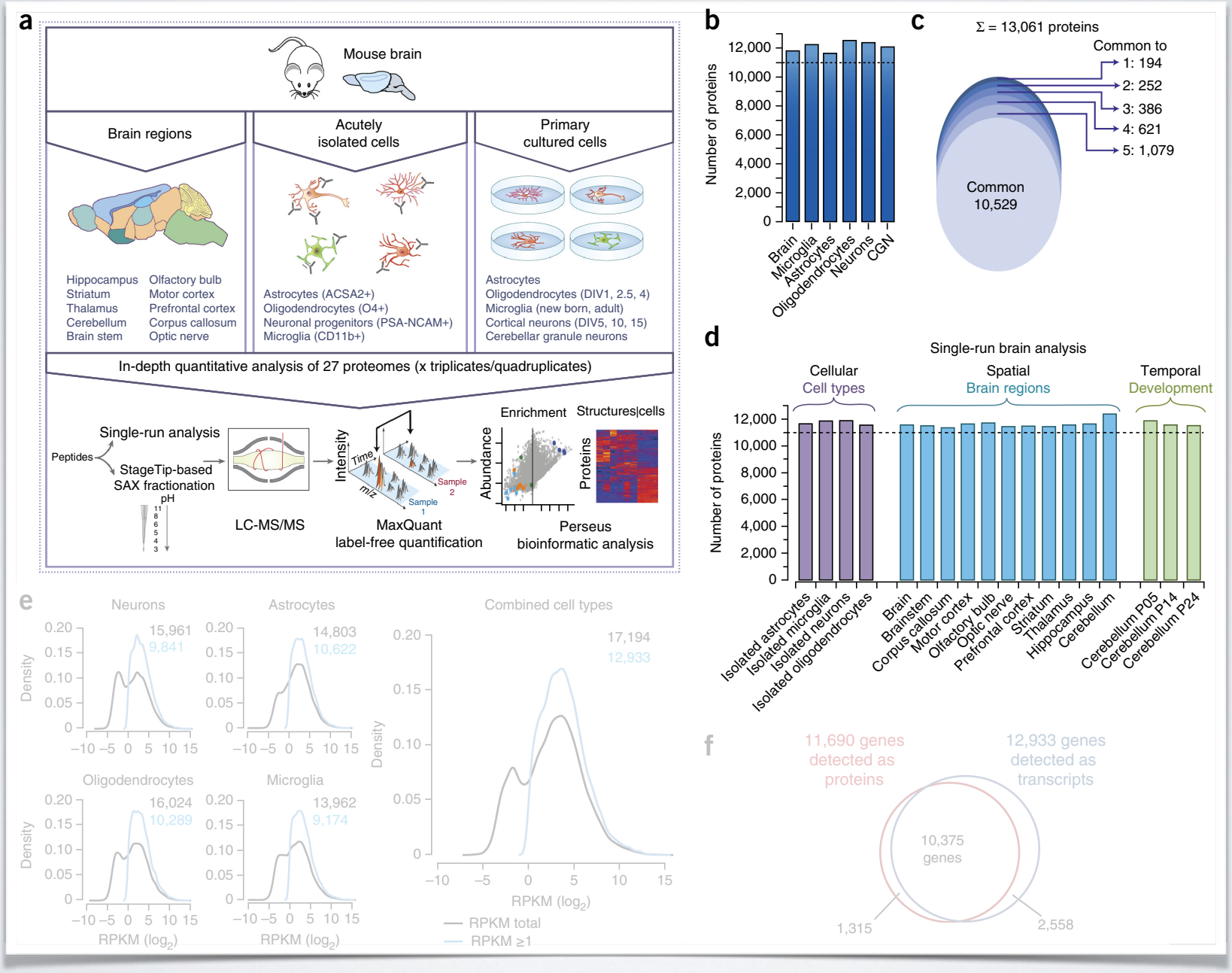


Figure 1 | transcriptome vs. proteome



experiment design rationale | Figure 1 (a-d)

- mapping spectra to the entire proteome/transcriptome can produce lots of false positives (this is what makes the Gencode team rather unhappy)

- to reduce the 'search space' they performed really-deep proteomics on a small number of samples:
  - fractionate peptides in before LC and run 6 injections for each sample
  - this is not quantitative, can only really assert presence/absence
  - basically collect as many high-quality spectra as possible
  - remove low-confidence peptides
  - create a library of high-quality peptides

- then for the quantitative runs of each sample (no fractionation) map the spectra to their custom library of peptides instead of the whole proteome

- this is what allows them to quantitate lots (>11,000) of proteins

Figure 1 | transcriptome vs. proteome

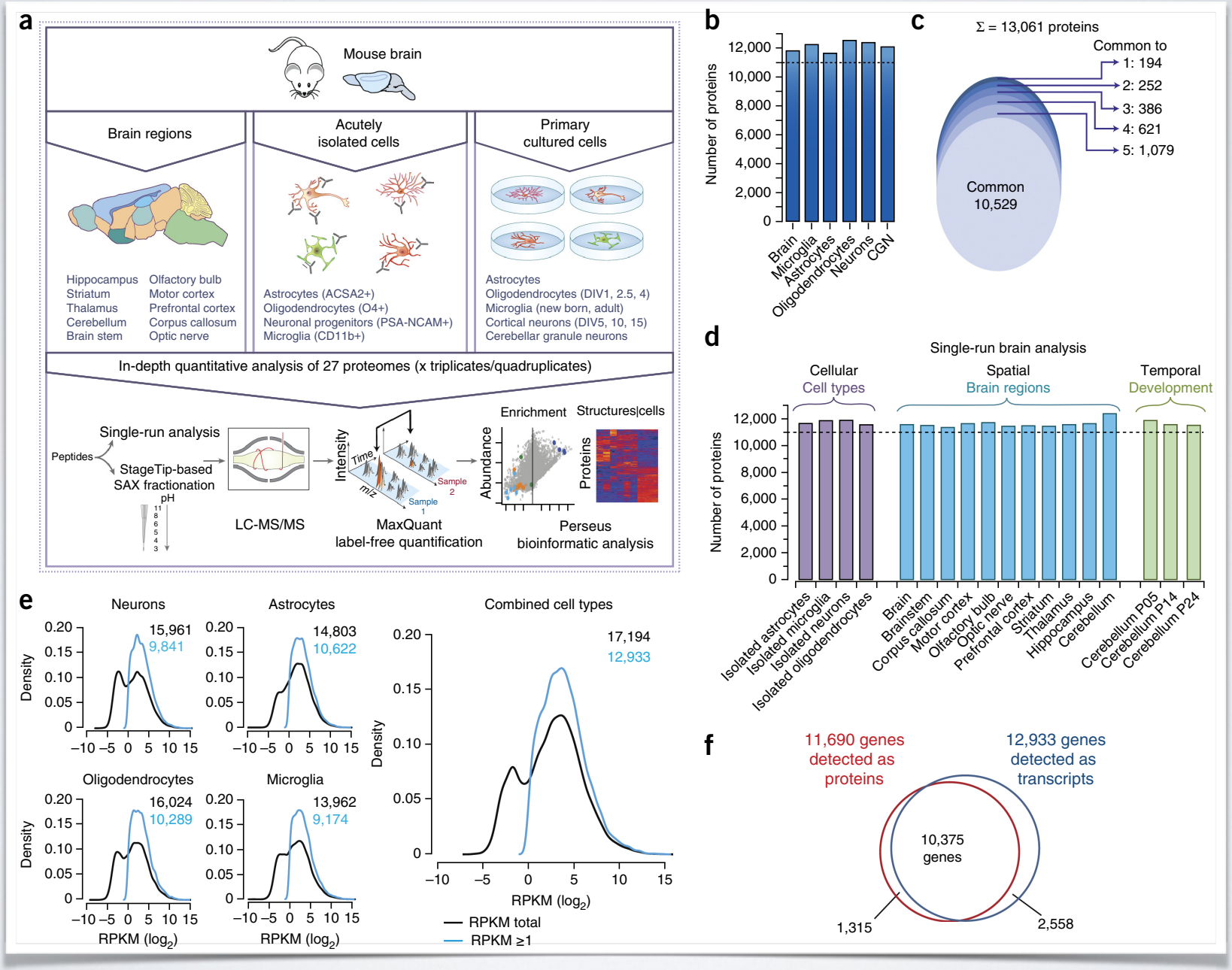


Figure 2 | cell-type specific proteomes

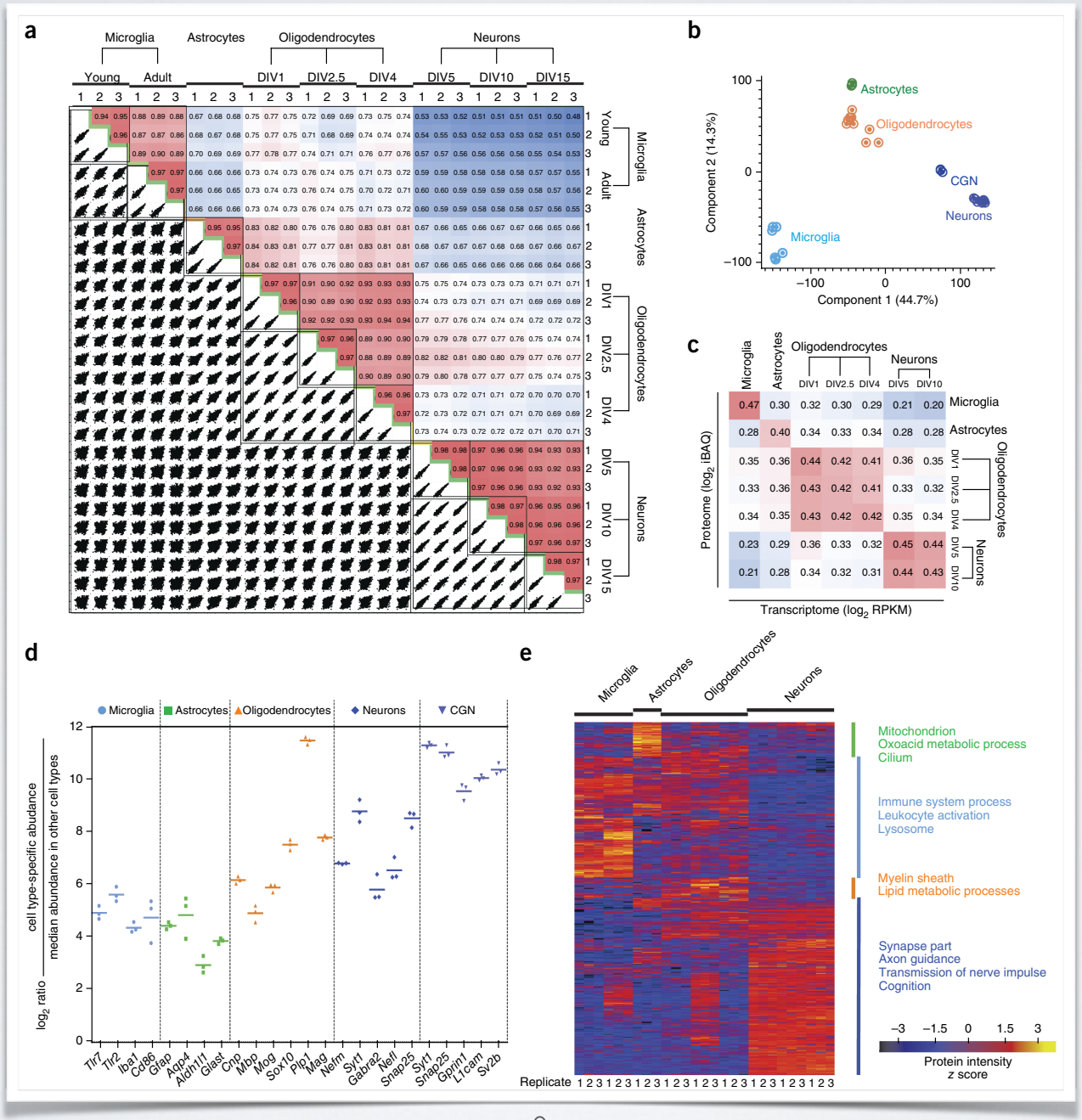


Figure 2 | cell-type specific proteomes

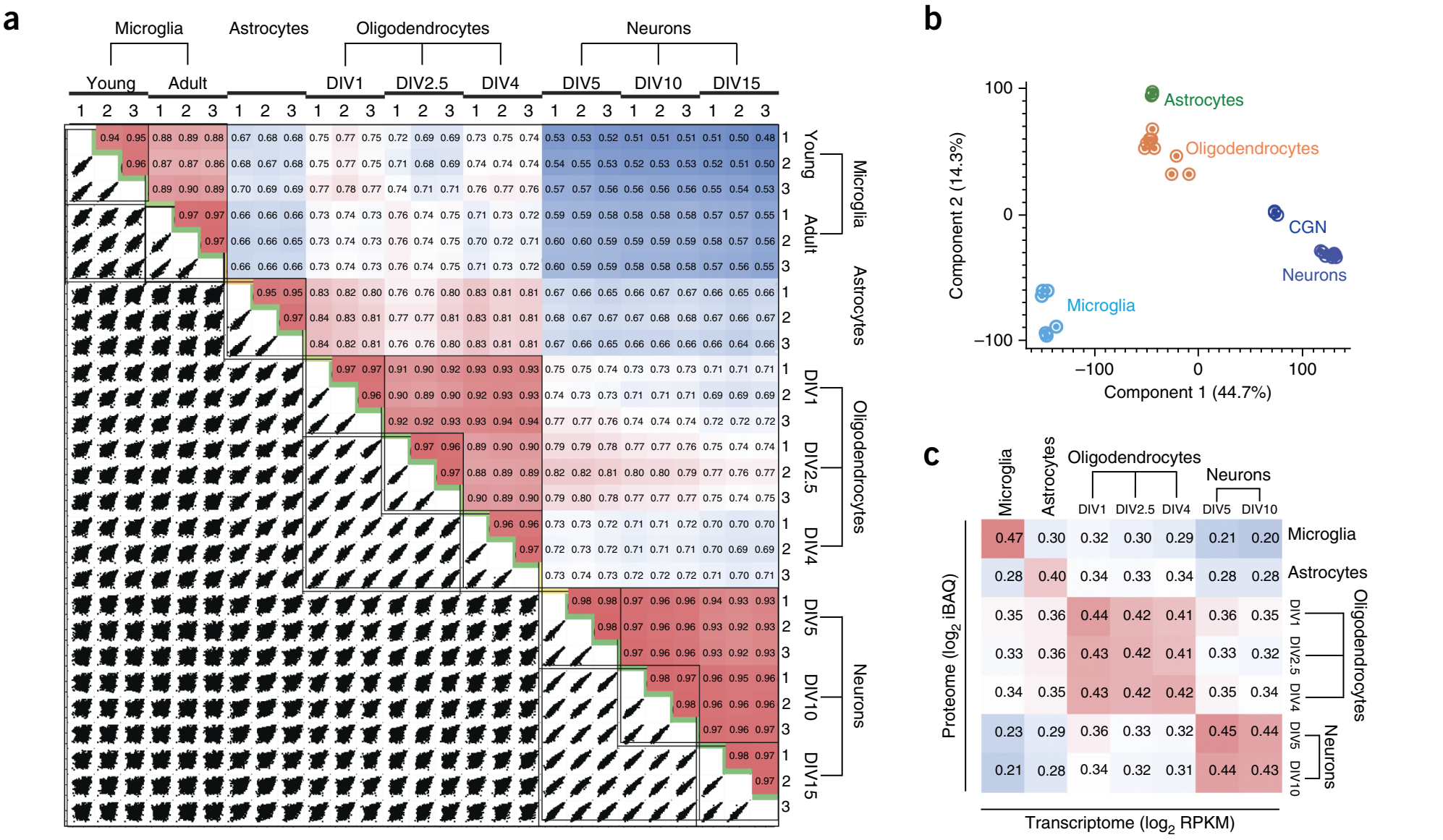


Figure 2 | cell-type specific proteomes

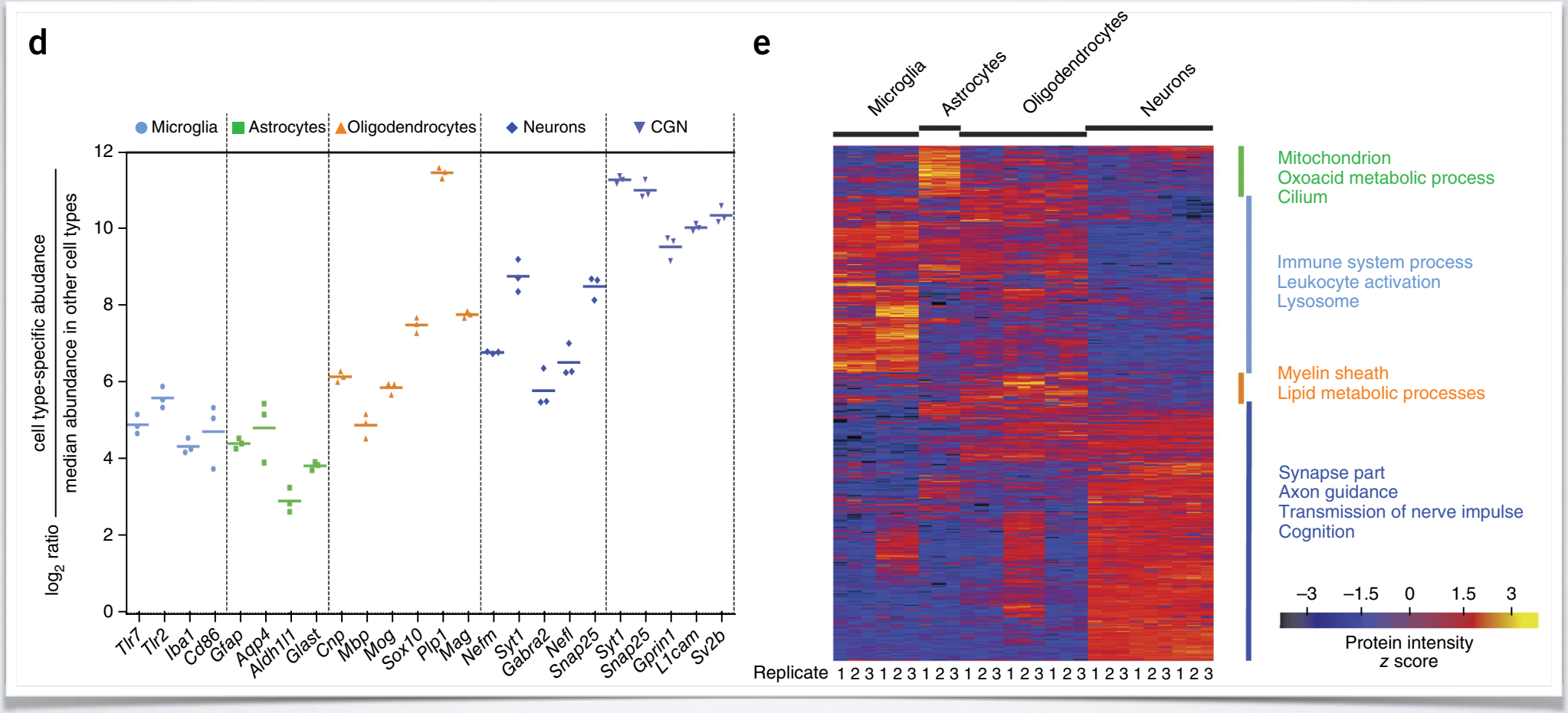


Figure 3 | cell-type abundances

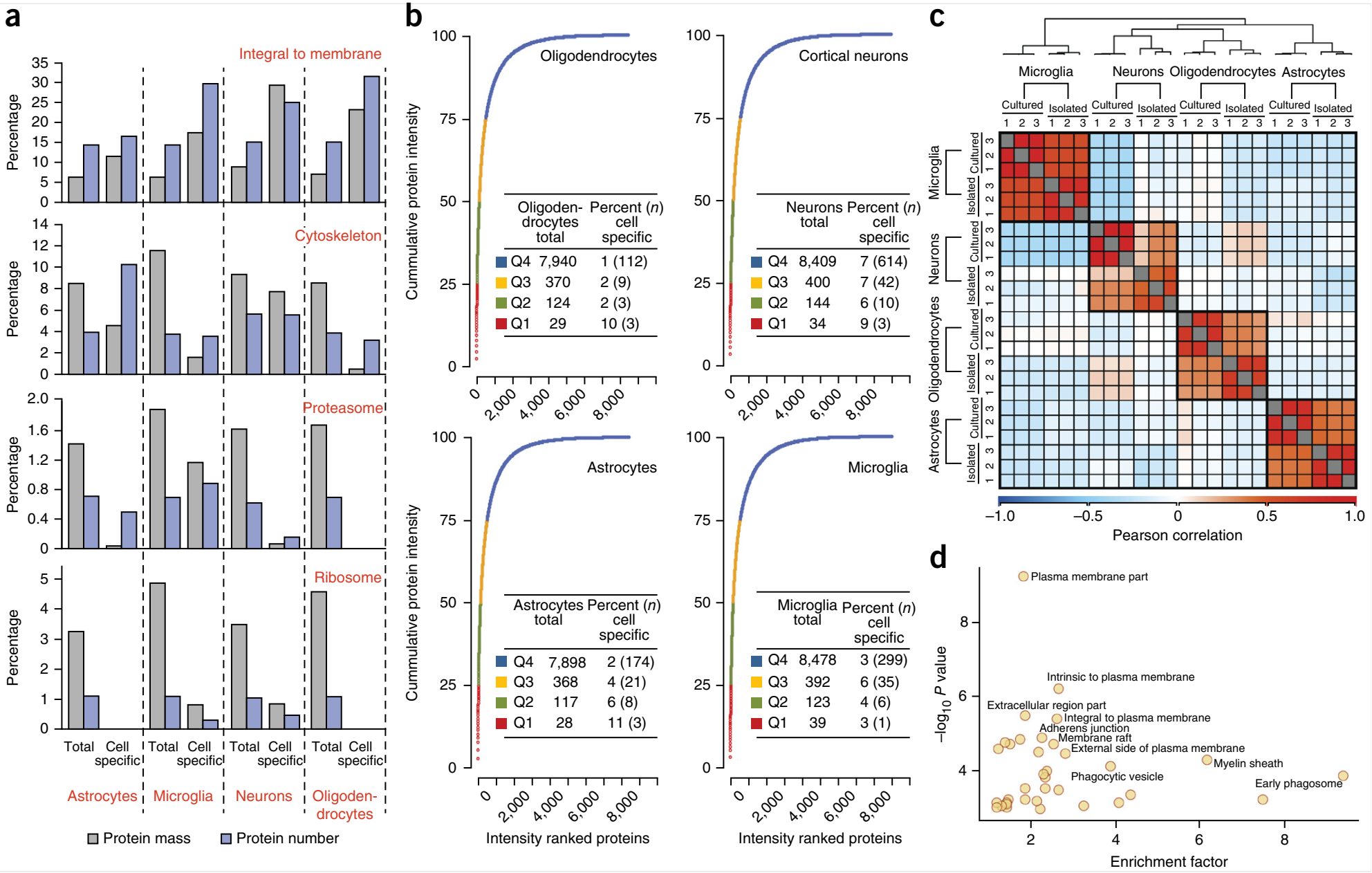


Figure 4 | brain enriched proteins

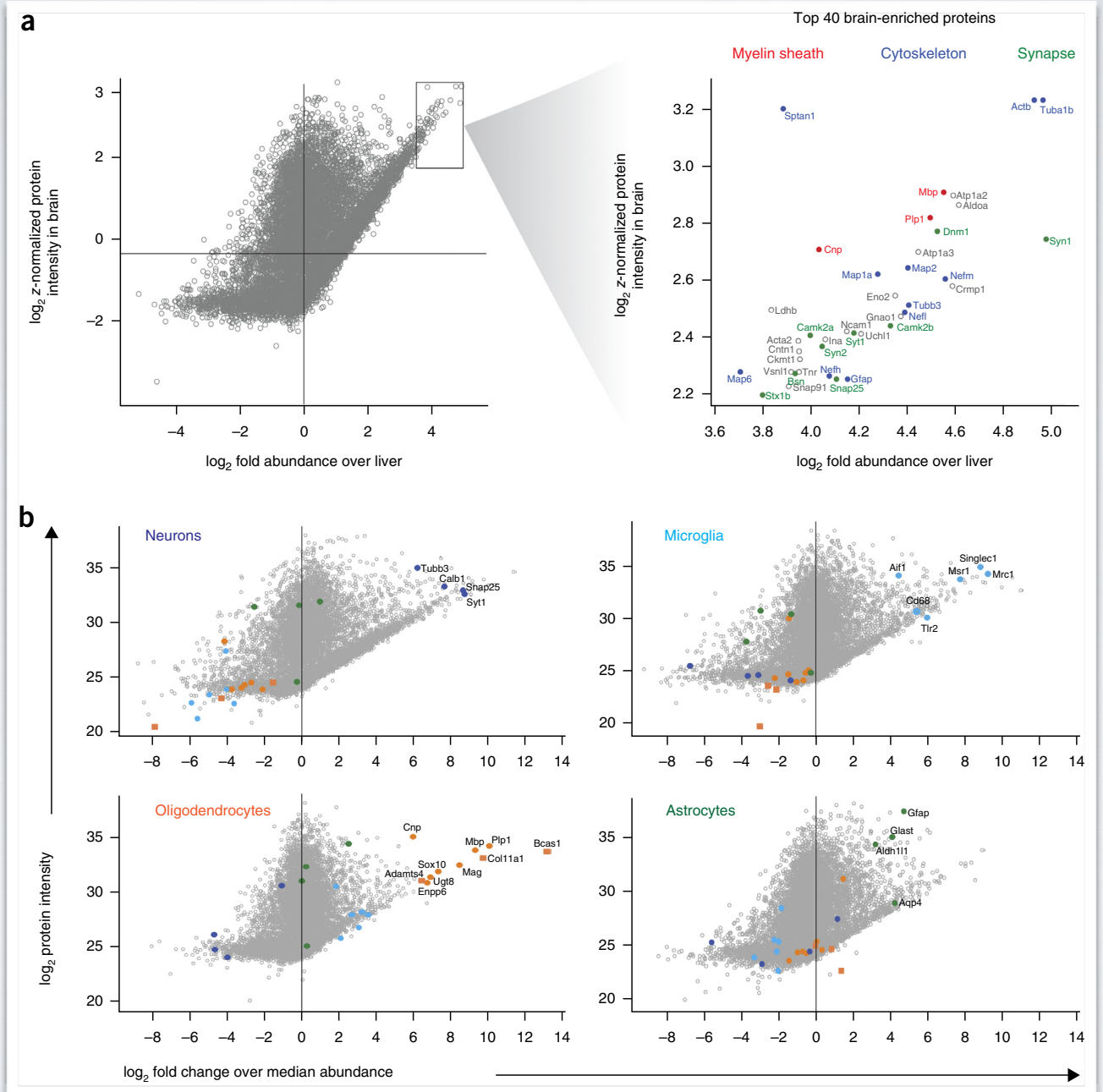


Figure 4 | brain enriched proteins

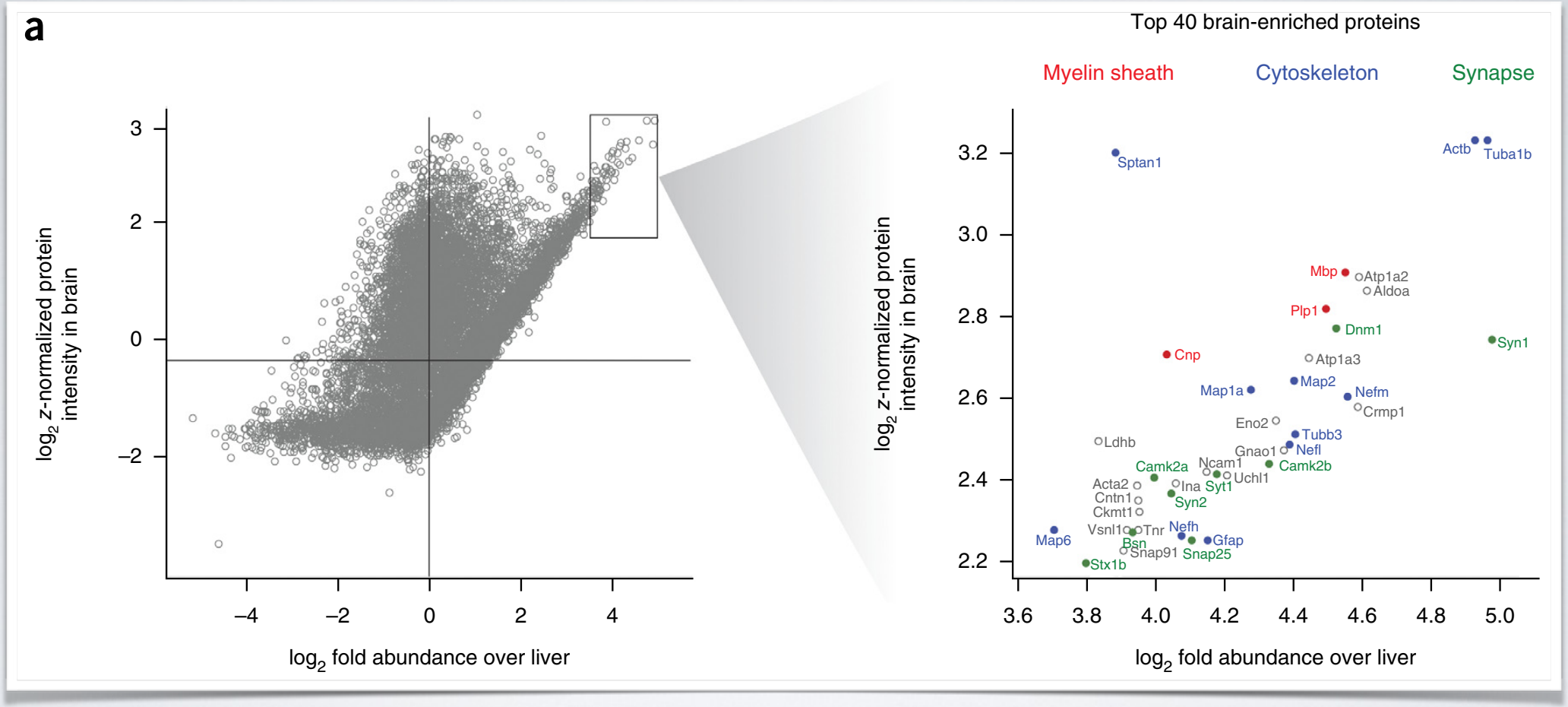


Figure 4 | brain enriched proteins

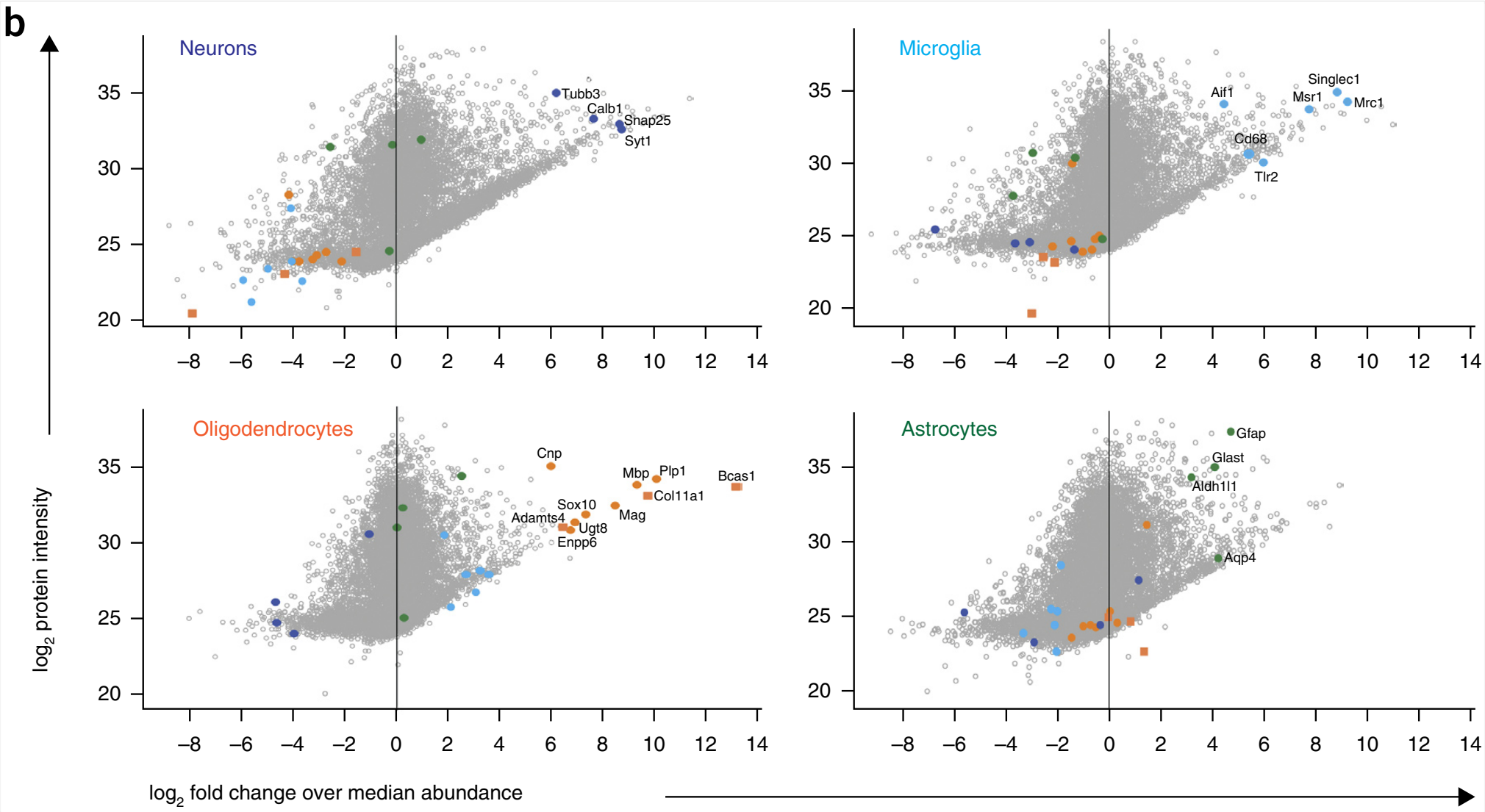


Figure 5 | brain region enriched proteins

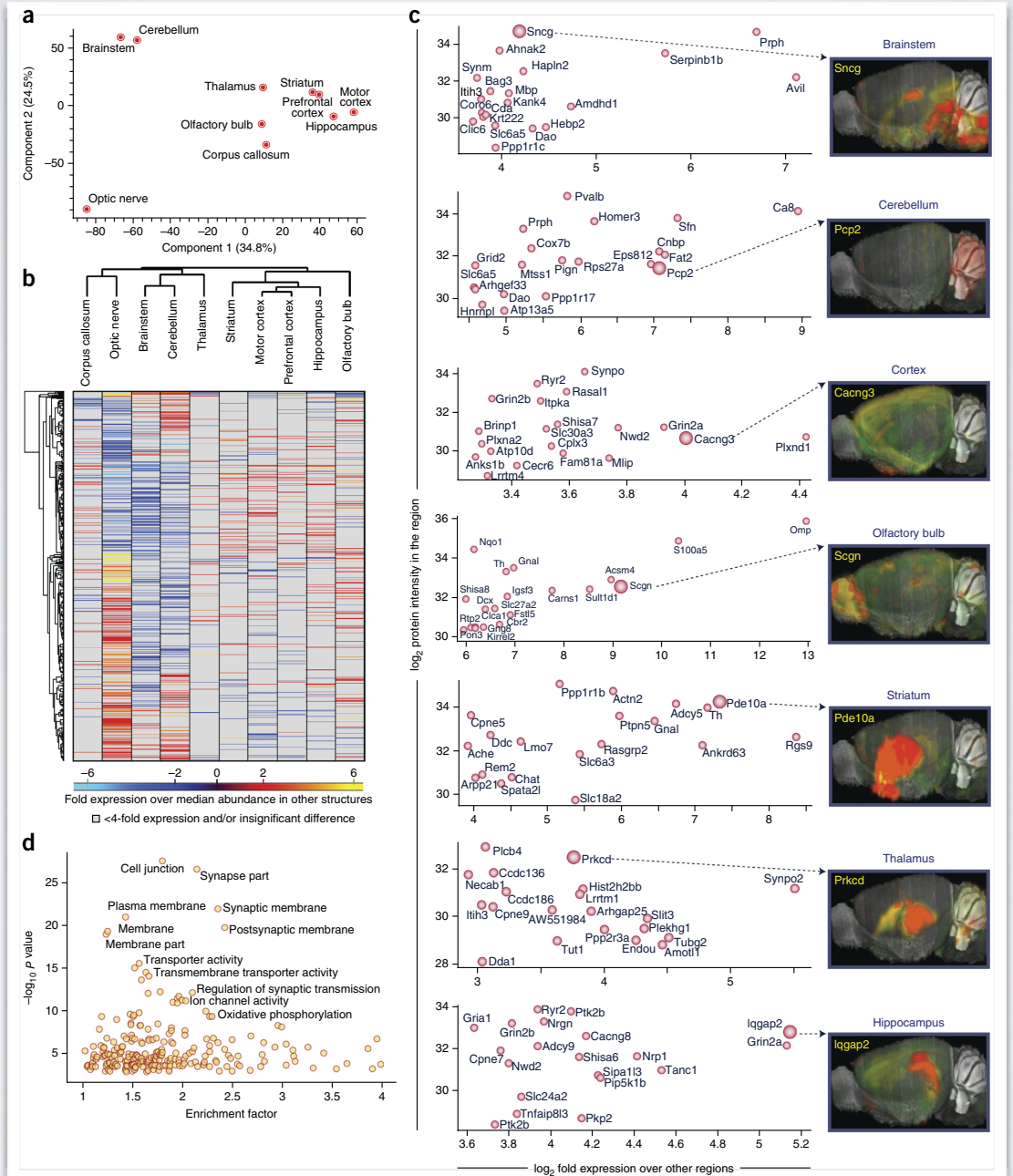


Figure 5 | brain region enriched proteins

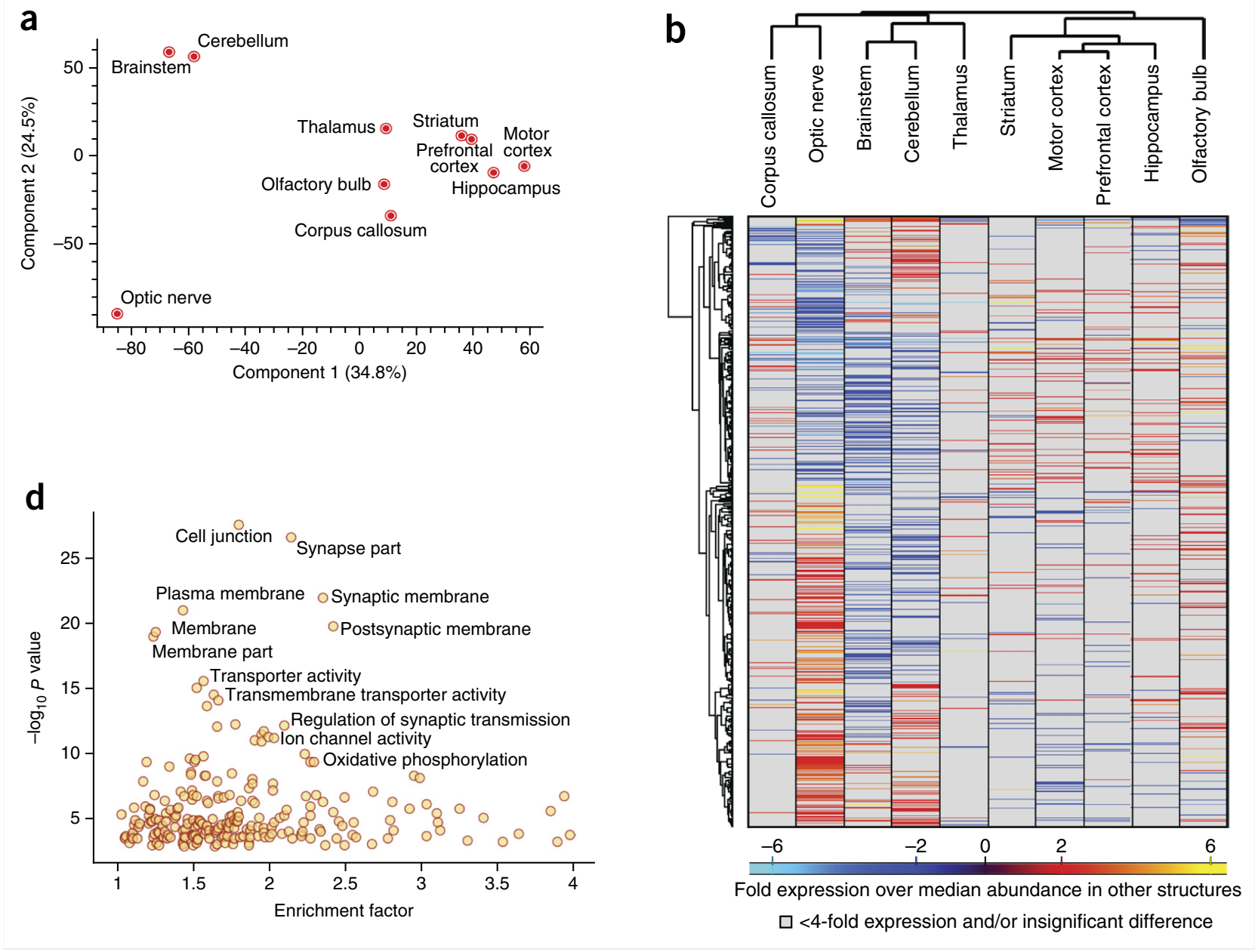


Figure 5 | brain region enriched proteins

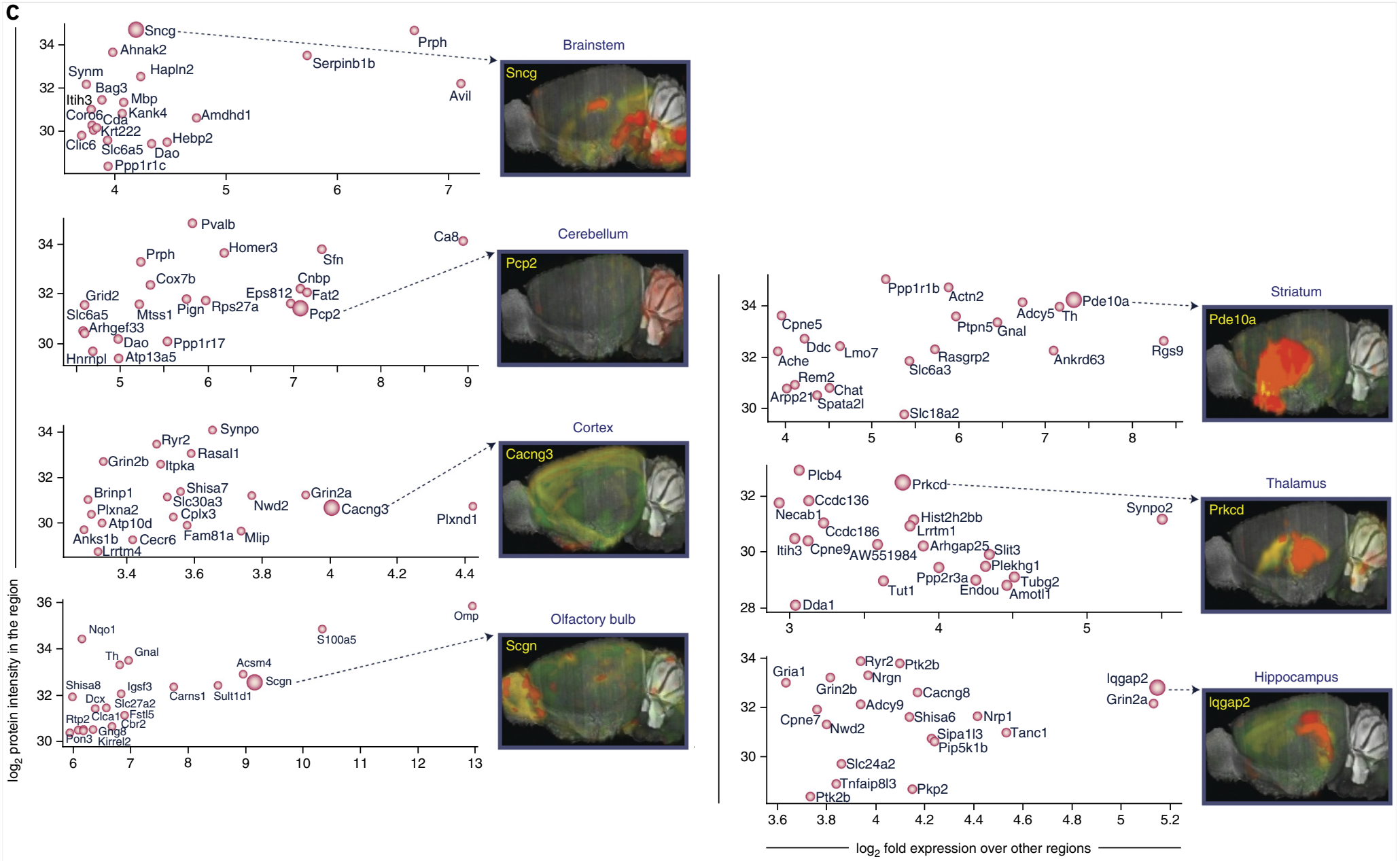


Figure 6 | cell-type pathway analysis

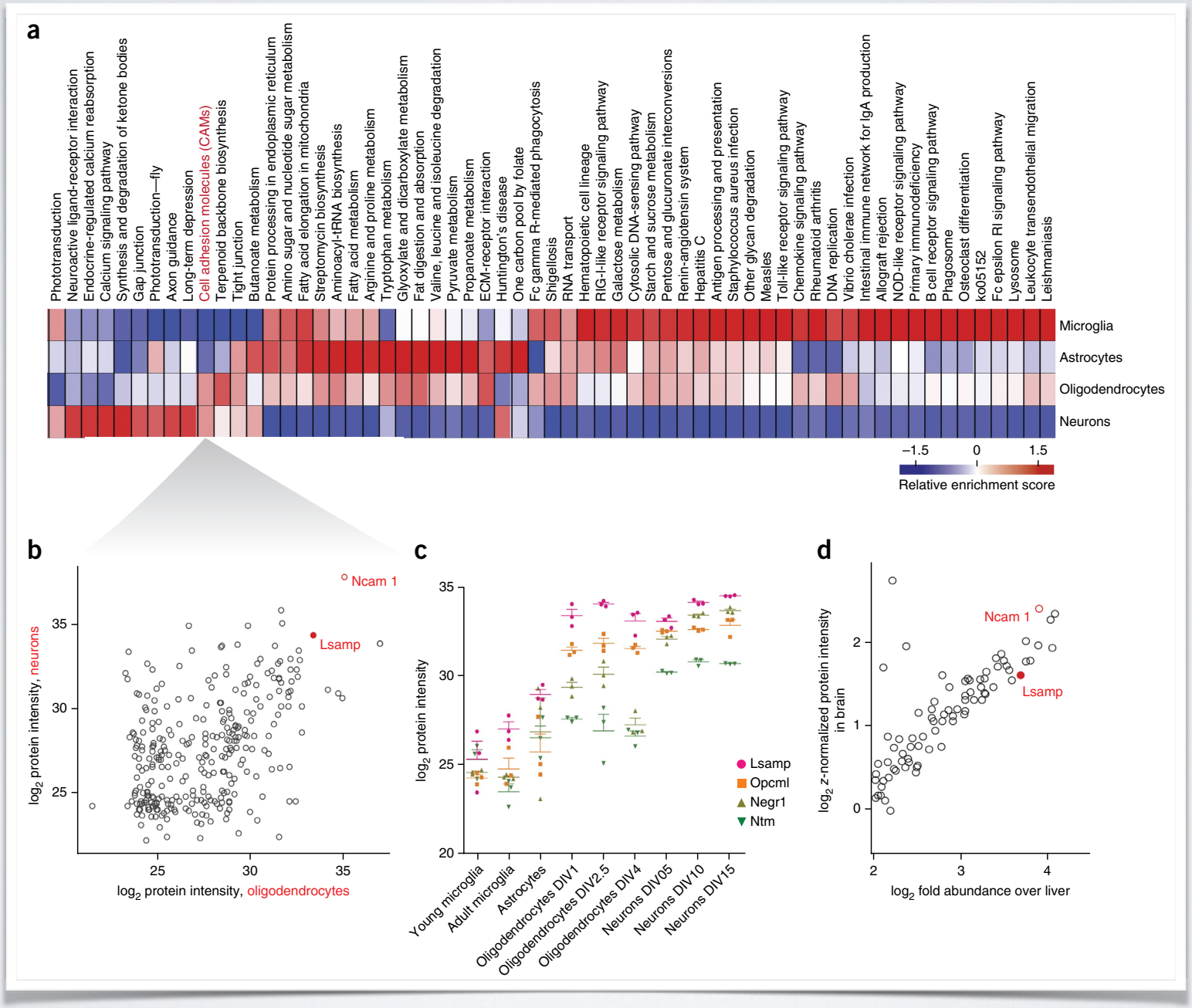


Figure 7 | specific gene followup

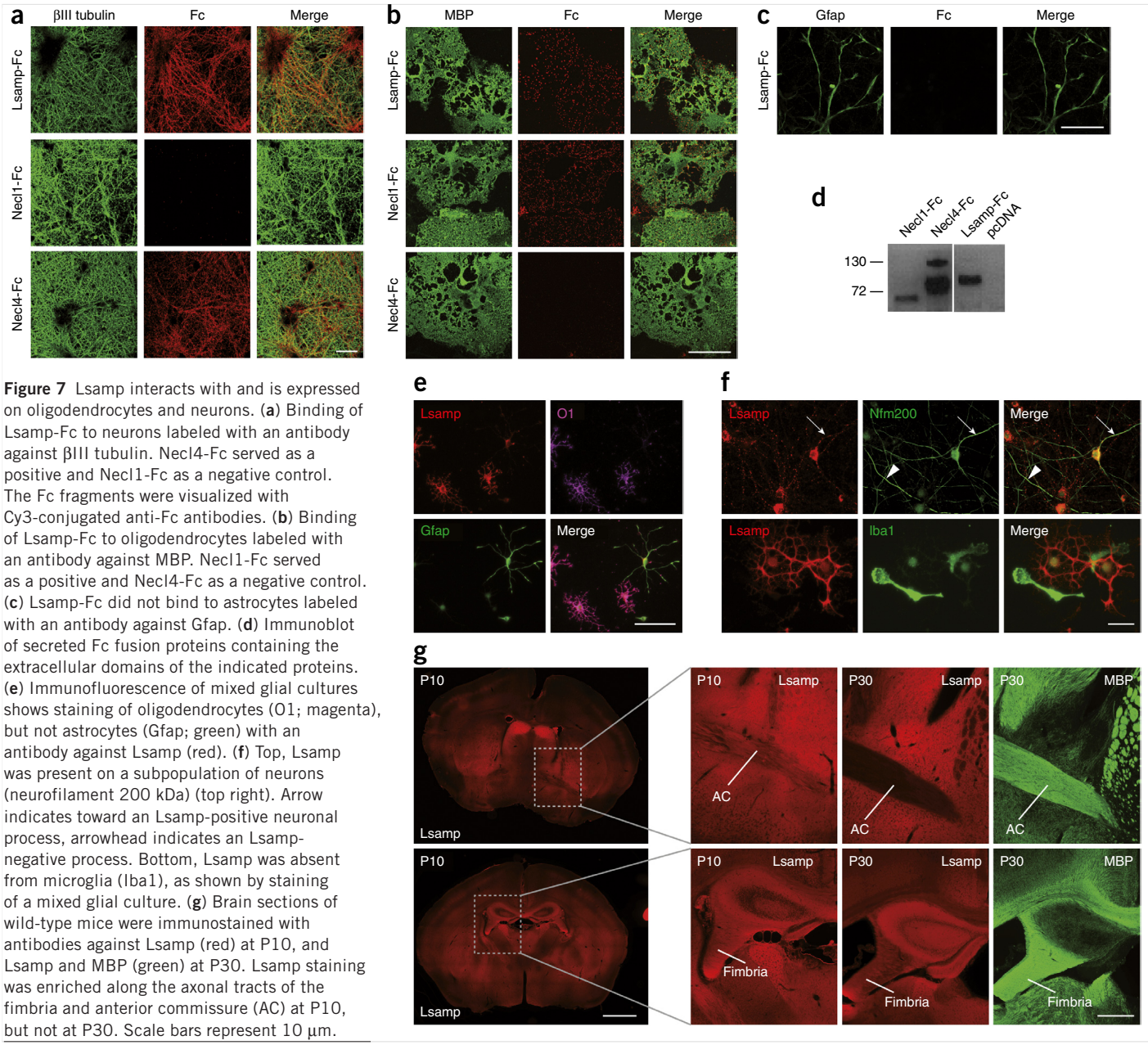


Figure 7 Lsamp interacts with and is expressed on oligodendrocytes and neurons. (**a**) Binding of Lsamp-Fc to neurons labeled with an antibody against βIII tubulin. Necl4-Fc served as a positive and Necl1-Fc as a negative control. The Fc fragments were visualized with Cy3-conjugated anti-Fc antibodies. (**b**) Binding of Lsamp-Fc to oligodendrocytes labeled with an antibody against MBP. Necl1-Fc served as a positive and Necl4-Fc as a negative control. (**c**) Lsamp-Fc did not bind to astrocytes labeled with an antibody against Gfap. (**d**) Immunoblot of secreted Fc fusion proteins containing the extracellular domains of the indicated proteins. (**e**) Immunofluorescence of mixed glial cultures shows staining of oligodendrocytes (O1; magenta), but not astrocytes (Gfap; green) with an antibody against Lsamp (red). (**f**) Top, Lsamp was present on a subpopulation of neurons (neurofilament 200 kDa) (top right). Arrow indicates toward an Lsamp-positive neuronal process, arrowhead indicates an Lsamp-negative process. Bottom, Lsamp was absent from microglia (Iba1), as shown by staining of a mixed glial culture. (**g**) Brain sections of wild-type mice were immunostained with antibodies against Lsamp (red) at P10, and Lsamp and MBP (green) at P30. Lsamp staining was enriched along the axonal tracts of the fimbria and anterior commissure (AC) at P10, but not at P30. Scale bars represent 10 μm.

Figure 8 | moar validation

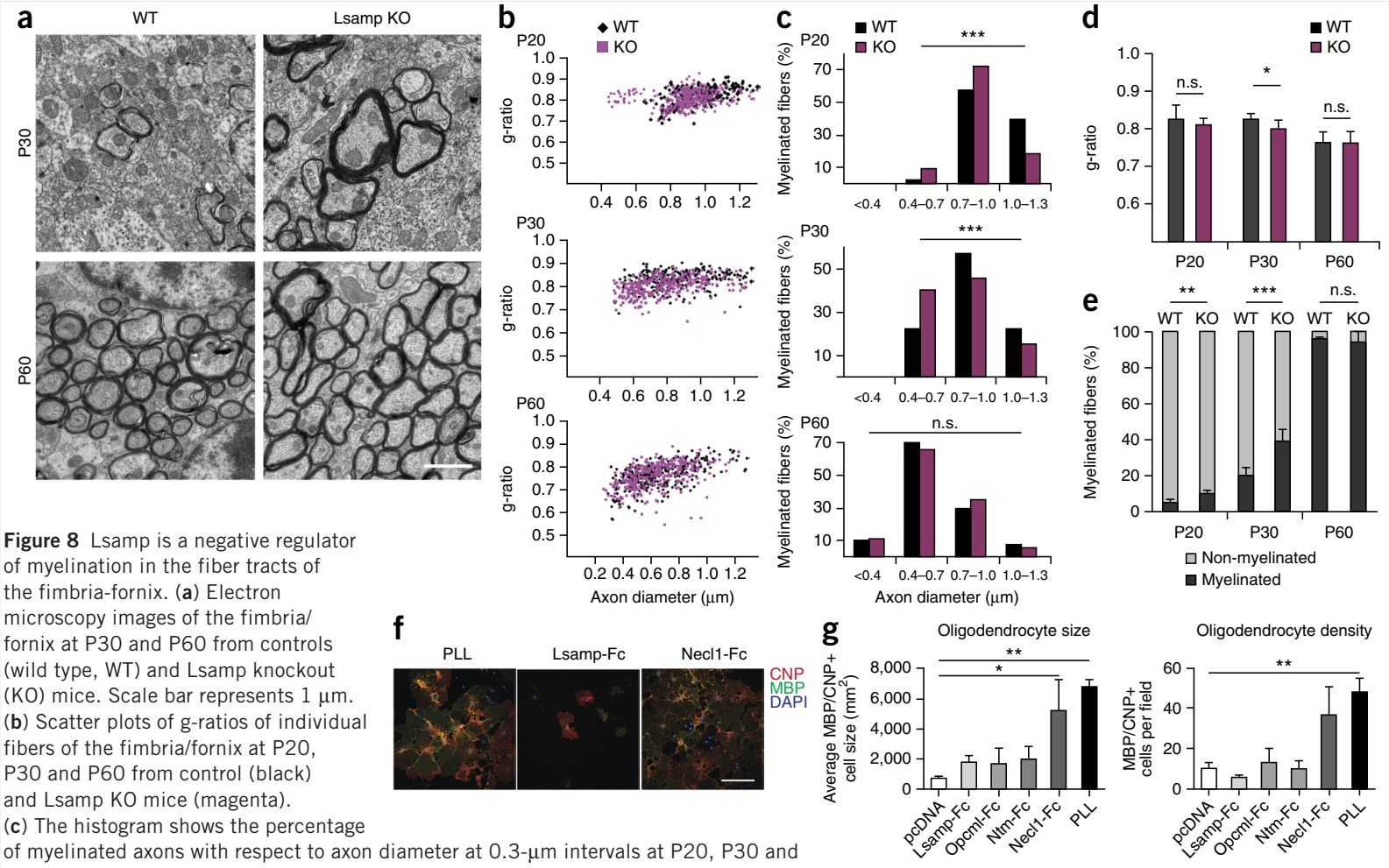


Figure 8 Lsamp is a negative regulator of myelination in the fiber tracts of the fimbria-fornix. (**a**) Electron microscopy images of the fimbria/fornix at P30 and P60 from controls (wild type, WT) and Lsamp knockout (KO) mice. Scale bar represents 1 μm. (**b**) Scatter plots of g-ratios of individual fibers of the fimbria/fornix at P20, P30 and P60 from control (black) and Lsamp KO mice (magenta). (**c**) The histogram shows the percentage of myelinated axons with respect to axon diameter at 0.3-μm intervals at P20, P30 and P60 for wild-type and Lsamp KO mice. There was a shift toward myelination of low-caliber axons in the mutant as compared with the control (chi-square test; from top to bottom: $P = 4.5 \times 10^{-10}$, $P = 2.8 \times 10^{-5}$, $P = 0.36$, $^{***}P < 0.0001$). More than ~250 axons for each genotype were counted (three animals per genotype) (**d**) Average g-ratio at P20, P30 and P60 for wild-type and Lsamp KO mice (Student's *t* test, $P = 0.0105$; $n = 3$ mice per genotype). Error bars represent s.d. (**e**) Percentage of myelinated and unmyelinated axons counted at P20, P30 and P60. More than 1,500 axons were counted for each time point ($n = 5$ mice for P20 and P30, $n = 4$ for P60) per genotype (bars show mean ± s.d.; Student's *t*-test; $^{**}P = 0.0055$, $^{***}P = 0006$). (**f**,**g**) Coverslips were coated with 10 μg ml^{-1} Fc-fusion proteins (IgLON family proteins and control), and oligodendrocyte precursor cells were plated and allowed to adhere and grow for 4 d. PLL and Necl1-Fc coating were used as positive controls. The purified supernatant of HEK 293T cells transfected with an empty vector (pcDNA) was used as negative control (bars show mean ± s.d.; ANOVA, $P < 0.05$, Dunnet *post hoc* test with pcDNA as control; $n = 3$ experiments; $^{*}P = 0.0087$, $^{**}P = 0.0025$). Scale bar represents 20 μm.

proteomics and the brain

Cell type– and brain region–resolved mouse brain proteome

Kirti Sharma[1], Sebastian Schmitt[2,3], Caroline G Bergner[2,3], Stefka Tyanova[1], Nirmal Kannaiyan[4], Natalia Manrique-Hoyos[2,3], Karina Kongi[5], Ludovico Cantuti[2,3], Uwe-Karsten Hanisch[6], Mari-Anne Philips[5], Moritz J Rossner[2,4], Matthias Mann[1] & Mikael Simons[2,3]

Brain transcriptome and connectome maps are being generated, but an equivalent effort on the proteome is currently lacking. We performed high-resolution mass spectrometry–based proteomics for in-depth analysis of the mouse brain and its major brain regions and cell types. Comparisons of the 12,934 identified proteins in oligodendrocytes, astrocytes, microglia and cortical neurons with deep sequencing data of the transcriptome indicated deep coverage of the proteome. Cell type–specific proteins defined as tenfold more abundant than average expression represented about a tenth of the proteome, with an overrepresentation of cell surface proteins. To demonstrate the utility of our resource, we focused on this class of proteins and identified Lsamp, an adhesion molecule of the IgLON family, as a negative regulator of myelination. Our findings provide a framework for a system-level understanding of cell-type diversity in the CNS and serves as a rich resource for analyses of brain development and function.

proteomic profiling
of human samples
for CEGS/psychENCODE

CEGS human adult | analysis strategy

- these are the adult human samples we have that match brainspan:
  - subjects: **HSB123 HSB126 HSB135 HSB136 HSB145**
  - regions: **PFC V1C HIP AMY STR MD CBC**

- perform highly fractionated proteomics for each of region (combine 5 individuals)

- perform single-shot quantitative proteomics for each sample (subject+region)

- compare benefit (sensitivity/specificity) of mapping single-shot peptides to:

  1 - entire annotated proteome (**standard approach**; high-FDR)
  2 - expressed **genes/transcripts** from RNA data (**our favourite**)
  3 - detected **genes/isoforms** from fractionated proteomics
  4 - detected **peptides** from fractionated proteomics (**Mann approach**)

CEGS human adult | basic protein detection stats

- **single-shot** MS/MS runs mapped to the **entire proteome** (approach 1):
  ~ **14,000 peptides (1% FDR)** from
  ~ **6,000 genes**

- the question is: to what degree the highly-fractionated reference proteome or the RNA-seq can increase the number of reliably detected peptides and genes

- **fractionated** MS/MS **'master' region proteomes** (for approaches 3 & 4):
  ~ **110,000 peptides (1% FDR)** from
  ~ **10,000 genes**

- getting pretty close to the **~12,000 'proteins'** (incl. isoforms) detected in **whole brain** from the Mann paper

CEGS human development | analysis strategy

- strategy depends on the outcome of the adult analysis
- we expect, in terms of maximising # detected proteins by minimising FDR:

  master region peptides/proteins > **RNA-seq** expressed > **whole proteome**

- how much better is the fractionated MS/MS compared to RNA-seq?
- fractionated proteomics is very expensive, RNA-seq is cheap (and data already exist!)
- if RNA-seq does a decent job, there are big implications for usefulness of **gencode/ GTEx** and the **human proteome map**

CEGS | timeline

year 1

- optimisation of protein extraction
- 34 adult brain samples profiled by single-shot MS/MS
- optimisation of fractionation

~$18,000

year 2

- run highly fractionated MS/MS for 7 adult human 'master' regions
- full integrated analysis of adult human samples
- select brains + regions for developmental analysis

~$19,000

year 3

- run developmental samples
- select brains + regions for NHP analysis

$TBD

fin