Decoding Cortical Circuits: Synaptic Signatures and Disease Vulnerabilities of Layer 5 Pyramidal Neuron Types

Gabriele Marcassa[1,2], Dan Dascenco[1,2], Blanca Lorente-Echeverría[1,2], Jeroen Vanderlinden[1,2], Elke Leysen[1,2], Lucas Baltussen[1,2], Andrew J. M. Howden[3] and Joris de Wit[1,2,4,*]

[1] VIB Center for Brain & Disease Research, Herestraat 49, 3000 Leuven, Belgium

[2] KU Leuven, Department of Neurosciences, Leuven Brain Institute, Herestraat 49, 3000 Leuven, Belgium

[3] Cell Signalling and Immunology, University of Dundee, Dundee, UK

[4] Lead contact

*Correspondence: joris.dewit@kuleuven.be

Abstract

Cortical layer 5 (L5) intratelencephalic (IT) and pyramidal tract (PT) neurons are embedded in distinct information processing pathways. Despite their significance for cognitive function, it remains poorly understood how L5 neuron types are connected into circuits and how this might be perturbed in disease. Utilizing an optimized proximity biotinylation workflow, we characterize the excitatory postsynaptic proteomes of L5 IT and PT neurons in intact somatosensory circuits. We find distinct molecular synaptic signatures of L5 IT and PT neurons defined by proteins regulating synaptic transmission and organization, suggesting that these contribute to function and input connectivity of L5 neuron types. The synaptic signature of L5 IT, but not of PT, neurons displays a marked enrichment of autism risk genes, aligning with previous studies highlighting the vulnerability of IT neurons and suggesting that the excitatory postsynaptic compartment of L5 IT neurons is notably susceptible in autism. Our analysis sheds light on the proteins that govern circuit integration and function of L5 neuron types and underlie their susceptibility in disease. Moreover, our approach is broadly applicable to other neuron types and circuits.

Introduction

How different types of neurons become embedded in neural circuits, the basis of the brain's capacity for information processing, is a major question in neuroscience. Cortical L5 pyramidal neurons (PNs) play a crucial role in information processing. These neurons integrate feedforward sensory information converging on their basal dendritic compartment and feedback contextual information converging on their apical dendrites. Their axons project to cortical and subcortical regions. L5 PNs are thus at the heart of cortical input/output processing[1–6].

In the somatosensory cortex, two main populations of L5 PNs can broadly be distinguished: intratelencephalic (IT) and pyramidal tract (PT) L5 PNs, which differ in their morphology, projection targets and electrophysiological properties. IT neurons are thin-tufted, intracortically projecting and regular spiking PNs mostly located in superficial L5 (L5a), whereas PT neurons are thick-tufted, subcortically projecting and bursting PNs mostly found in deeper L5 (L5b). The two L5 PN types also differ in input connectivity, which is best characterized for excitatory input from the thalamus. L5 IT neurons primarily receive input from the higher-order thalamic posteromedial nucleus (POm), whereas PT neurons preferentially receive input from the primary thalamic ventral posteromedial nucleus (VPM)[1,7–11]. Functionally, L5 IT and PT neurons are part of different cortical information processing pathways. Perception of tactile stimuli depends on activation of apical dendrites of L5 PT, but not IT, neurons[12]. In the motor cortex, L5 IT neurons are involved in preparation of movement, whereas PT neurons are important for initiation and execution[13,14]. Thus, the two L5 PN types are embedded in different circuits.

Despite their importance for cortical information processing, the molecular mechanisms regulating circuit connectivity and functional properties of L5 IT and PT neurons remain poorly understood. Single-cell genomics has provided insight in transcriptome composition[15–17] and differential vulnerability to disease of L5 PN types. In autism spectrum disorder (ASD) for example, RNA sequencing studies of patient

samples identify cortico-cortical projecting (IT) PNs across cortical layers, including L5, among the most affected cell types[18–20]. However, a direct analysis of the molecular composition of L5 PN synapses in intact circuits is missing. Examining such cell type-specific molecular synaptic signatures will be instrumental to start comprehending how L5 PN types are connected into circuits and how perturbations in this process might contribute to diseases, particularly those characterized by a strong synaptic component, such as neurodevelopmental disorders[21].

Characterizing the molecular composition of synaptic connections with cell type specificity is challenging[22–24]. Bulk synaptosome preparations provide depth but lack cell type-specific resolution[25–29]. Cell type-specific expression of a fluorescent presynaptic reporter in combination with fluorescence-activated synaptosome sorting (FASS) allows molecular profiling of mainly presynaptic compartments from genetically identified neurons and is a scalable approach[30–32]. Profiling the proteome of postsynaptic compartments of specific neuron types on the other hand has been possible only in selected cases that generated a transgenic mouse or took advantage of specific structural and molecular features of those synapses[33–35]. Proximity biotinylation, which relies on the ability of promiscuous biotin ligases or peroxidases to covalently attach a biotin molecule to nearby proteins within a short radius[36], offers a more generalizable approach to characterize the molecular composition of postsynaptic compartments *in vivo*[37]. Proximity biotinylation has thus far been used to characterize postsynaptic proteome composition of whole hippocampus and cortex but has not yet been applied to single cortical populations with compartment-specificity in intact circuits in vivo.

Here, we optimize a cell type-specific proximity biotinylation-based workflow using the engineered biotin ligase TurboID[38] to map the protein composition of excitatory postsynaptic densities of L5 IT and PT neurons in the somatosensory cortex. We find that differential expression of ion channels, neurotransmitter receptors and cell-surface proteins (CSPs) characterizes the molecular synaptic signatures of L5 IT and PT neurons. The identified proteins regulate synaptic transmission and organization of synapses, suggesting that they contribute to the specification of input connectivity and functional properties of the two L5 PN types. To assess whether the molecular synaptic signatures of L5 IT and PT neurons are vulnerable to disease, we analyze the identified proteomes for disease association. We find a striking enrichment of ASD risk genes in the excitatory postsynaptic proteome of L5 IT neurons, but not of PT neurons. Our findings align with previous studies highlighting the vulnerability of IT neurons in individuals with autism[18–20]. Moreover, our results suggest that the excitatory postsynaptic compartment of L5 IT neurons is notably susceptible in the context of ASD. Our analysis sheds light on the proteins that govern circuit integration and function of L5 PN types and contribute to their vulnerability in disease, and can be broadly applied to create a protein-based, synaptic atlas of cortical circuits.

Results

Efficient, compartment-specific protein biotinylation in genetically identified neurons

To compare the molecular synaptic signatures of L5 IT and PT neurons in the somatosensory cortex, we first optimized a workflow to identify postsynaptic proteomes in a cell-type specific manner, with

optimal capture of biotinylated proteins from a small population of neurons in brain region. This optimized workflow was based on the high-efficiency biotin ligase turboID (Figure 1A), which has a reported ~10-fold increase in *in vitro* activity compared to previously published biotin ligases[38]. To restrict the activity of turboID to the excitatory postsynaptic density (PSD) and thus only label proteins in this compartment, we generated a fusion construct with the scaffold protein PSD95 (PSD.turboID) (Figure 1B). As a control, we generated a turboID construct without any localization sequence (cytosolic.turboID) (Figure 1B). To target turboID to L5 PNs in a cell-type and region-specific manner, we packaged Cre-dependent turboID expression constructs into adeno associated viral vectors (AAV) and injected them in the primary somatosensory (S1) cortex of postnatal day (P)28 mouse lines expressing Cre recombinase in L5 PNs. A week after stereotactic injection, we subcutaneously injected the turboID substrate biotin and collected brains for immunofluorescence or biochemical analysis (Figure 1C). Tissue sections stained for HA-tagged turboID and for biotinylated proteins using fluorescent streptavidin showed colocalizing signal restricted to S1 and concentrated in L5 (Figure 1D). A control sample in which only biotin but no virus was injected (no.turboID) showed that background biotinylation was below detection levels compared to other conditions (Figure 1D). High-resolution confocal imaging showed a diffuse pattern of biotinylated proteins throughout the neuronal cytoplasm in cytosolic.turboID samples, while in PSD.turboID samples streptavidin-labeled biotinylated proteins displayed a punctate pattern, which colocalized with the excitatory postsynaptic marker Homer1, confirming turboID activity at the excitatory postsynaptic compartment (Figure 1E).

To control for potential PSD95 overexpression artefacts, we analyzed spine density and morphology in turboID-expressing samples using the Supernova labelling system, which allows for sparse yet bright GFP expression in genetically identified neurons[39] (Figure S1A). Dendritic spine reconstruction and classification found no difference in spine density or morphology across conditions (Figure S1B and 1C). These results show that our approach achieves efficient region-, cell type- and compartment-specific protein biotinylation *in vivo* without disrupting spine homeostasis.

To optimize *in vivo* biotinylation, we expressed PSD.turboID in S1 L5 PNs and provided biotin for different durations before collecting tissue for biochemical analysis. Western blot quantification showed a maximum, 4-fold increase in biotinylation levels compared to control conditions at 3h following biotin injection (Figure 1F). We therefore selected 3h post-biotin injection for subsequent experiments to achieve maximum biotinylation while limiting any off-target labelling[40].

We then compared *in vivo* expression and activity of the different turboID constructs. HA immunoblot showed that the turboID constructs run at the expected size of 35kDa for cytosolic.turboID and ~135 kDa for PSD.turboID (Figure 1G). Immunoblotting for biotinylated proteins showed two strong bands in the no.turboID condition, likely corresponding to endogenously biotinylated proteins[41], while cytosolic and PSD.turboID produced strong biotinylation across all molecular weights (Figure 1G). Streptavidin pulldown from the same lysate detected postsynaptic proteins Gria1 and Homer1 in the PSD.turboID sample, but not in the other conditions (Figure 1H). Together, these results show that our proximity biotinylation strategy reliably targets the excitatory postsynaptic compartment of genetically identified L5 PNs *in vivo* and efficiently yields biotinylated proteins in a short timeframe.

Identification of the excitatory postsynaptic proteome of cortical L5 IT and PT neurons

We next applied the optimized proximity biotinylation workflow to analyze the postsynaptic proteome composition of two closely related L5 PN populations: IT and PT neurons, genetically identified by the Tlx3-Cre and Sim1-Cre mouse lines[11,12,42,43], respectively (Figure 2A-B). Imaging of tissue sections from Tlx3-Cre and Sim1-Cre mice injected with an AAV expressing Cre-dependent GFP confirmed the segregation of labeled cell bodies to superficial and deep L5, respectively (Figure 2B). To identify the postsynaptic proteome of the two L5 PN populations, we injected P28 Tlx3-Cre or Sim1-Cre mice with cytosolic.turboID or PSD.turboID AAVs. At P35, mice were injected with biotin for 3h before tissue dissection and collection of the somatosensory cortex (Figure 2C). Negative control samples (no.turboID) were prepared with mice injected only with biotin, to exclude endogenously biotinylated proteins and contaminants from the analysis (Figure 2D). Western blot analysis of tissue lysate from each replicate, consisting of tissue from 2 mice (around 1mg of total dissected tissue), confirmed expression and successful biotinylation in turboID samples (Figure S2A). Biotinylated proteins were purified from tissue lysates using streptavidin-coupled beads, eluted and digested to peptides by trypsin and identified by label-free liquid chromatography-tandem mass spectrometry (LC-MS/MS) run in data-independent acquisition (DIA) mode (Figure 2C) (see Methods). Protein identification and quantification performed with DIA-NN software[44] yielded more than 4500 different proteins among all samples, which showed consistent intensity distribution at the whole sample level (Figure S2B).

As expected, we detected high levels of endogenously biotinylated proteins in all samples (Figure S2C). On the other hand, known postsynaptic proteins showed selective enrichment in the PSD.turboID replicates and depletion in the no.turboID samples (Figure S2D). PCA analysis showed replicates clustering together by condition with little separation relative to the mouse line used (Figure 2E). We used an approach similar to a previously published ratiometric analysis[38] to create a list of proteins detected in the no.turboID condition and selected contaminants to be excluded from the analysis (Figure S2E) (see methods). To determine enrichment of postsynaptic proteins in the PSD.turboID samples, we first compared the cytosolic.turboID and PSD.turboID conditions for each mouse line (Figure 2F). Core constituents of the excitatory postsynapse, such as glutamate receptors and scaffolding proteins, were found to be enriched in both L5 IT and PT neurons (Figure 2F). We then combined the list of enriched proteins (FDR < 0.1, s0 > 0.1) from the Tlx3-Cre and Sim1-Cre lines for the cytosolic.turboID samples, and for the PSD.turboID samples. In this way, we detected a total of 894 proteins enriched in the cytosolic.turboID samples and 492 proteins enriched in the PSD.turboID samples. This analysis workflow is summarized in Figure S2F. Gene ontology (GO) analysis (cellular compartment) revealed enriched terms for the cytoskeleton and the nuclear compartment in the cytosolic.turboID samples, while PSD.turboID-enriched proteins showed strong excitatory postsynaptic localization (Figure 2G). Taken together, these results show how an optimized proximity biotinylation workflow combined with DIA-MS can reliably detect proteins from specific subcellular compartments of genetically identified neurons in intact circuits.

Molecular synaptic signatures of cortical L5 IT and PT neurons

We next compared the postsynaptic protein composition of L5 IT and PT neurons. We performed a differential expression analysis of the synaptic proteins of the two cell types to map the proteins that specify the molecular postsynaptic signatures of L5 IT and PT neurons. We found a total of 135 differentially expressed postsynaptic proteins between IT (93 proteins) and PT (42 proteins) neurons (Figure 3A).

We first explored the function of the differentially expressed postsynaptic proteins using several GO databases[45]. Differentially expressed proteins showed strong postsynaptic localization (Figure S3A). 'Regulation of postsynaptic neurotransmitter receptor activity' and 'synapse organization' were among the most represented GO terms in the Biological Process category (Figure 3B). The Molecular Function category showed enrichment mostly in terms related to PSD interactions (Figure S3B). To further characterize the differentially expressed proteins, we queried the synaptic protein database published in Sorokina, Mclean, Croning et al.[46] (from now on referred to as the 'synapse proteome' database). This database lists the synaptic proteins detected in proteomic studies performed using various approaches in the last 20 years and collects and integrates information with several external databases. Of the identified postsynaptic proteins in L5 IT and PT neurons, 84% were represented in this database (Figure S3C). We finally cross-referenced differentially expressed postsynaptic proteins with the curated synaptic protein database SYNGO[47]. While SYNGO provides the most detailed annotation tool for synaptic proteins, only 44% of identified postsynaptic proteins were annotated, due to its relatively limited coverage (Figure S3D). Nevertheless, in agreement with classic GO analysis, we found high enrichment in terms related to synaptic transmission and synapse organization (Figure S3E).

We next explored the proteins in the synaptic transmission and synapse organization modules, as they could be the basis for functional properties and connectivity of L5 IT and PT neurons. To do so, we started from the proteins annotated in each term of the Biological Process category (Figure 3B and Figure S3F) and integrated these results with PANTHER[48] and Uniprot[49] information to assign a function to the differentially expressed proteins (see Table S1). In the synaptic transmission module, neurotransmitter receptors and channels were the most enriched protein classes (Figure 3C). Among differentially expressed receptors, kainate receptors (Grik2, Grik3) and their auxiliary subunit Neto1 showed enrichment in L5 PT neurons (Figure 3C and Figure S3F). Kainate receptors have been observed to be enriched in deep cortical layers of the mouse brain using immunohistochemistry[50] and receptor autoradiography[51,52], but an enrichment in L5 PT neurons has not been reported to our knowledge. NMDA receptor subunits (Grin2b) were enriched in IT neurons, whereas potassium and calcium ion channels showed different subunits specifically enriched in one of the two L5 neuron types (Figure 3C and Figure S3F).

In the synapse organization module, CSPs were the most abundant protein class (Figure 3D). Strikingly, the extracellular leucine-rich repeat (LRR) protein family was strongly represented (Elfn2, Flrt2, Flrt3, Lrrtm3, Lrrtm4, Lrrc4c/Ngl-1, Ntrk3/Trkc, Lrfn2/Salm1, Slitrk2) among the differentially expressed CSPs in L5 IT and PT neurons (Figure 3D). LRR proteins control precise wiring and confer synaptic properties across species[53–55], but their role in cortical circuit formation and function remains largely unexplored.

Interestingly, Flrt2 and Flrt3 belong to the same LRR subfamily but showed differential enrichment in IT and PT neurons (Figure 3D).

Due to the limits of antibody validation of proteomic data[56,57], we decided to validate our results using an orthogonal approach measuring cell-type specific gene expression. While it has been reported that mRNA and protein rarely match at absolute levels, it has also been shown that mRNA can predict proteomic differences at a global level[58–60]. We selected 4 differentially expressed proteins from the synaptic transmission and from the synapse organization modules (Figure 3E-H). We injected AAVs expressing Cre-dependent GFP in Tlx3-Cre or Sim1-Cre mice and performed single-molecule Fluorescence In Situ Hybridization (smFISH) experiments for the selected genes (Figure 3E and G). Quantification of mRNA puncta showed significant agreement between cell type enrichment of synaptic proteins and their corresponding transcripts, with one exception (Figure 3F and H). *Grik3* in the synaptic transmission module did not show a significant difference between L5 IT and PT neurons but showed a bimodal distribution with cells expressing high levels of mRNA and cells expressing none (Figure S3G). In conclusion, these results show that differential expression of neurotransmitter receptors, ion channels and CSPs defines the molecular synaptic signatures of L5 IT and PT neurons.

Differential vulnerability of L5 IT and PT synaptic signatures in neurodevelopmental disorders

Neurodevelopmental disorders, such as ASD, are associated with altered synaptic function and development[21,61–63]. We therefore tested whether the molecular synaptic signatures of L5 IT and PT neurons might provide insight in disease vulnerability of these two L5 neuron types.

We first ran the differentially expressed synaptic proteins in L5 IT and PT neurons (Figure 3A) through the GeneTrek database[64], which provides association of genes with neurodevelopmental disorders (Figure 4A). Based on several disease association databases, genes are given a 'high confidence' or 'candidate' score. The molecular synaptic signatures of both IT and PT neurons showed a high proportion of genes in each of these two categories (Figure 4A). Next, we queried the synapse proteome database for disease association and graphed the results separately for IT and PT neurons, normalized to the total number of synaptic proteins identified in each neuron type (Figure 4B and Table S1). Overall, neurodevelopmental disorders showed a higher percentage of enriched genes compared to neurodegenerative diseases such as Parkinson's or Alzheimer's in L5 IT and PT neurons (Figure 4B and Table S1). The molecular synaptic signature of PT neurons showed a larger fraction of genes associated with schizophrenia and psychotic disorder, whereas that of IT neurons showed a larger fraction of genes associated with ASD and intellectual disability (Figure 4B and Table S1).

We next cross-referenced the molecular synaptic signatures of L5 IT and PT neurons with the SFARI gene database[65], which has a curated, up to date annotation of ASD genes scored by clinical and scientific evidence. This analysis revealed a marked enrichment of ASD-associated genes in the molecular synaptic signature of L5 IT neurons compared to that of PT neurons (Figure 4C). All differentially expressed synaptic proteins with the highest ASD risk score annotation (‘syndromic’ and ‘high confidence’) in the SFARI database, such as Frmpd4, Cdkl5, Shank3, Cnksr2 and Slitrk2, were enriched in L5 IT neurons (Figure 4C and D). These findings show that the excitatory postsynaptic

compartment of L5 IT neurons, but not of PT neurons, is notably vulnerable in ASD. Together, these observations highlight how cell type-specific molecular synaptic signatures contribute to selective neuronal vulnerability to neurodevelopmental disorders such as ASD.

Discussion

Here, we analyze the synaptic molecular signatures of two closely related L5 cortical neuron types, using an optimized proximity biotinylation workflow to characterize cell-type-specific excitatory postsynaptic proteomes in intact circuits *in vivo*. We find that differential expression of neurotransmitter receptors, ion channels and CSPs with a role in synaptic transmission and synapse organization defines the synaptic signatures of L5 IT and PT neurons. Moreover, we find that the L5 IT and PT neuron-specific synaptic signatures contribute to differential vulnerability in neurodevelopmental disorders such as schizophrenia and ASD.

To our surprise, L5 IT neurons, which have a less extensive dendritic arborization compared to PT neurons, displayed a higher number of differentially expressed proteins related to synaptic transmission and organization than PT neurons. PCA and sample level intensity analysis exclude differences in labelling efficiency between the two cell types, suggesting that the observed differences in protein numbers reflect the diverse protein sets available to IT and PT neurons to specify their synapses. These cell type-specific molecular synaptic signatures represent promising targets for future investigation into regulators of synaptic properties, physiology and connectivity of L5 IT and PT neurons. Especially striking in this context are the LRR family proteins, which dominate the CSPs in the L5 IT and PT neuron synapse organization modules. LRR proteins are important regulators of connectivity and synaptic properties[66–68]. Input-specific combinations of postsynaptic LRR proteins define synaptic structure and function in hippocampal PNs[69]. Analogous to hippocampal PNs[69–71], LRR proteins may differentially distribute across the dendritic tree of cortical PNs to regulate connectivity, structure and function of specific inputs. While our approach does not provide spatial information about synaptic protein distribution in L5 PN dendrites, this remains an important point that will need to be explored in further studies. The two L5 PN neuron types we examined here have excitatory input in common[11] but also show cell type-specific differences in input connectivity patterns. Primary sensory input from the VPM thalamic nucleus preferentially, but not exclusively, targets L5 PT neurons, whereas higher-order POm input is concentrated on L5 IT neurons[1,7–9,72]. It will be of interest to determine whether differentially expressed synaptic proteins in L5 IT and PT neurons contribute to differences in thalamic input connectivity.

We find that the molecular synaptic signatures of L5 IT and PT neurons display differential vulnerability to distinct neurodevelopmental disorders. L5 PT neurons showed a larger fraction of synaptic proteins associated with schizophrenia, whereas IT neurons showed a larger fraction of synaptic proteins associated with ASD. Cross-referencing with the SFARI gene database revealed a striking enrichment of syndromic and high confidence ASD risk genes in the molecular synaptic signatures of L5 IT, but not of PT, neurons. These findings are in line with genomic studies that find upper layer (L2/3) neurons[18,73] or deeper layer (L5/6) neurons[19] to be the most affected cell types in ASD. These studies generally

support the notion that cortico-cortical projecting (IT) neurons are vulnerable cell types and, based on cell type, suggest altered intracortical processing in ASD[20,74]. Our synaptic compartment-specific proteome data indicates that compared to PT neurons, excitatory input of L5 IT neurons is particularly vulnerable in ASD. Considering the differences in excitatory input connectivity between L5 IT and PT neurons, it is tempting to speculate that higher-order thalamic input on IT neurons is affected in ASD. Brain imaging studies have indeed observed altered thalamocortical functional connectivity in individuals with autism[75–78]. Whether this involves projections from higher-order thalamic nuclei to specific cortical neuron types is unknown.

In addition to formulating hypotheses on L5 IT and PT connectivity and function, the molecular synaptic signatures we identify here can also be used to explore specific hypotheses on protein function. An interesting synaptic protein from our analysis is Tbr1, which is enriched in the L5 IT synaptic signature and a high confidence ASD risk gene. Tbr1 is a transcription factor but shows a dendritic distribution in pyramidal neurons and localizes to synaptic fractions in biochemical analysis of adult brain[79]. Possibly, PSD-associated Tbr1 shuttles between synapse and nucleus, where it can regulate expression levels of Grin2b[80], which we also find to be enriched in the L5 IT synaptic signature. Another interesting synaptic protein we identified is Pelo in the synaptic signature of L5 IT neurons. Pelo rescues stalled ribosomes and triggers degradation of damaged mRNA in a process termed No-Go Decay[81,82] but its role at synapses is unknown. Taken together, our results imply that many of the proteins that define the molecular synaptic signatures of L5 IT and PT neurons also contribute to their differential vulnerability to neurodevelopmental disorders.

To obtain these results, we optimized every step of the proximity biotinylation workflow, resulting in the identification of thousands of proteins starting from single, genetically-identified L5 neuron populations in the somatosensory cortex. We found that a single subcutaneous biotin injection 3 hours before sample collection is sufficient to produce strong protein biotinylation in the brain. Previous *in vivo* studies have relied on multiple injections of biotin over several days[37,83–85]. As prolonged labeling with biotin increases the number of proteins identified at the expense of specificity[38], minimizing the biotinylation window is important for the analysis of compartment-specific protein composition. Short biotinylation windows also open new possibilities, for example to study how the composition of synaptic proteomes changes upon neuronal activity. For similar reasons, we limited the duration of expression of turboID constructs to 1 week (compared to multiple weeks or months in previous studies[37,83,85–87]), as low levels of protein biotinylation using endogenous biotin can occur while turboID is expressed.

It is important to note that the success of our approach does not only depend on the experimental steps, but also on the DIA-MS setup used and subsequent data analysis. DIA-MS allows for increased proteome data coverage[88], which is critical for these experiments, as synaptic proteins are often low-abundant proteins that may be missed in data-dependent DDA-MS experiments. The power of DIA-MS in this context has been reported in the proteomic profiling of *C. elegans* AFD neurons, of which only 2 neurons are present per animal[89]. The data analysis workflow we developed here creates a list of high confidence synaptic proteins in two steps. First, common contaminants identified in the no.turboID condition are removed from the detected proteins. Then, each PSD.turboID sample is compared to its

matched cytosolic.turboID control to select proteins enriched at the synapse. Our two-step analysis will reject some synaptic proteins such as the metabotropic glutamate receptor GRM1, which localizes to the periphery of the postsynaptic density[90,91], but results in a final list of high-confidence synaptic proteins of cell type-specific excitatory PSD composition.

A possible limitation of our approach is the use of PSD95 to target turboID to the postsynaptic compartment. We carefully titrated our viral vectors to control PSD95 and turboID expression levels and analyzed spine density and morphology *in vivo* to verify the absence of overexpression artefacts. This strategy is suitable for mature circuits but carries a risk of driving precocious postsynaptic maturation in developing neurons. CRISPR/Cas9-mediated knock-in approaches[86] might be one alternative. These approaches avoid the overexpression of exogenous proteins, however, the low percentage of successful recombination makes it challenging to target single neuronal populations in specific cortical regions[86,92]. A possible way to avoid modifying expression of synaptic proteins, while retaining short duration of turboID expression and high efficiency of labeling, might be the use of recently published PSD95 binders[93], which are reported not to alter PSD95 interactions.

In conclusion, our analysis of cell type- and compartment-specific protein composition sets a foundation for the systematical description of cortical circuit connectivity at the molecular synaptic level. As our approach enables the comparison of the synaptic proteome of single neuronal populations, and mouse driver lines for cortical neuronal cell types are increasingly available[43,94], systematic application of this workflow could create a protein-based, synaptic atlas of the cortex.

Acknowledgements

We thank Esther Klingler, Patrik Verstreken, Pierre Vanderhaeghen and Anthony Holtmaat for critical reading of the manuscript, and De Wit lab members for helpful discussion and comments. We are grateful to Joris Vandenbempt for mouse colony maintenance and genotyping. We thank Charles Gerfen (National Institute of Mental Health, USA) and Nathaniel Heintz (The Rockefeller University, USA) for generously providing the Tlx3_P56.cre and Sim1_KJ18.cre mice. We thank Aya Takeoka (VIB-Imec-KU Leuven NeuroElectronics Research Flanders) for sharing Tlx3-Cre and Rbp4-cre mice. We thank Kristofer Davie (VIB Center for Brain & Disease Research Single-Cell Bioinformatics Expertise Unit) and Danie Daaboul (De Wit lab) for bioinformatic support. G.M. is supported by Fonds Wetenschappelijk Onderzoek (FWO, Belgium) PhD fellowship 11F1219N and 11F1221N; D.D. is supported by FWO Postdoctoral fellowship 12W5218N, 12W5221N and FWO Research Grant 1513320N; B.L-E. is supported by FWO PhD fellowship 1120821N and 1120823N; L.B. is supported by FWO Postdoctoral fellowship 12ZK221N. J.d.W. is supported by FWO Project Grants G0C4518N, G0A8720N and G0A8320N; FWO EOS Grant G0H2818N; ERANET-NEURON TAO2PATHY G0I3118N and a Methusalem Grant of KU Leuven/Flemish Government.

Declaration of interests

J.d.W. is scientific co-founder and served as scientific advisory board member of Augustine Tx.

Methods

Animals. All animal experiments were conducted according to the KU Leuven ethical guidelines and approved by the KU Leuven Ethical Committee for Animal Experimentation (ECD P183/2017 and P162/2021). Mice were maintained in a specific pathogen-free facility under standard housing conditions with continuous access to food and water. Mice used in the study were 1–5-weeks old and were maintained on 14 h light, 10 h dark light cycle from 7 to 21 h. Wild-type (WT) C57BL/6J mice were obtained from JAX. Rpb4.cre and Tlx3.cre lines were obtained from GENSAT. Sim1.cre was obtained from Dr. Charles Gerfen (National Institute of Mental Health, USA). Genotypes were regularly checked by PCR analysis. Heterozygous cre-positive males were breed to Bl6 females. For euthanasia, animals were either anesthetized with isoflurane, and decapitated or injected with an irreversible dose of ketamine–xylazine.

Cell Lines. HEK293T-17 human embryonic kidney cells were obtained from American Type Culture Collection (ATCC) cat# CRL-11268. HEK293T-17 cells were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (FBS; Invitrogen) and penicillin/streptomycin (Invitrogen).

Plasmids TurboID sequence was a gift from Alice Ting (Addgene plasmid #107169). Mouse PSD95 sequence was a gift from Gary Bassell (Addgene plasmid #102949). TurboID was fused at the C-terminus of PSD95 and inserted into a cre-dependent AAV expression vector under the synapsin promoter by Gibson assembly (NEB). Plasmids will be deposited at plasmids.eu.

AAV purification High titer AAV production and purification was carried out as previously reported [95]. Briefly, 6 plates of 70% confluent HEK cells were transfected with 20ug of pDelta F6 plasmid, 10ug of RepCap 2/9 plasmid and 10ug AAV genome plasmid per plate using PEI transfection. 3 days later cells were collected by scraping and centrifugation. AAV particles in the supernatant were precipitated by PEG and added to cell lysates. Lysates were treated by Benzonase nuclease to remove cellular DNA. Lysates were loaded onto a iodixanol gradient (60%, 40%, 25%, 15%) and ultracentrifuged 1:40 hours at 50.000 RPM at 12C. The 40% fraction containing the purified AAV was carefully collected and desalted and concentrated on a 4ml Amicon column (Sigma). AAV purity was tested by silver staining using the ProteoSilver™ Silver Stain Kit (Sigma PROTSIL1-1KT). AAV titer was determined by qPCR to normalize titers between different viral vector batches using LightCycler 480 SYBR Green I Master (Roche 04707516001). Primer for qPCR targeting the synapsin promoter were as follows: FWD tgataggggatgcgcaatttgg REV gtgcaagtgggttttaggacca.

Stereotactic injections P28 mice were anesthetized with 5% isoflurane and Duratears was applied to the eyes to prevent them from drying out. Mice were placed in a mouse stereotact (KOPF). During the rest of the procedure 2.5% isoflurane was constantly administered. After shaving and disinfecting the mouse's head, local anesthesia was administered by a subcutaneous injection with 100 µl lidocaine (xylocain 1%). An incision was made on the skin to reveal the skull. A hole was drilled at coordinates for somatosensory cortex and 200nl of AAV mix injected using a Nanoject III (Drummond) though a beveled

capillary at 2 nl/s. After a 5 min recovery, the capillary was pulled out at ~0.1 mm/5 s. The incision was stitched with surgical glue (Millpledge Veterinary). After 6 h, their health was examined and mice were injected with 0.1 mg/kg buprenorphine. Two injections per hemisphere were performed for biochemistry experiments to cover a larger area of the somatosensory cortex and recover more material (x: 3.2, y: 0.3 and 1.5 from bregma, z: 0.8 from the surface of the brain). One injection per hemisphere was performed for every other experiment (x: 3.2, y: 0.9 from bregma, z: 0.8 from the surface of the brain). Cytoplasmic.turboID was injected at a ~1.6*10^12 GC/mL, PSD.turboID was injected at a ~3*10^11 GC/mL, core-dependent GFP was injected at ~10^12 GC/mL.

Biotinylated protein collection One week after AAV injection, mice were injected subcutaneously with 1ml of 2mg/mL biotin in 1X PBS, adjusting the pH to 7.4 to increase biotin solubility. 3 hours later mice were anaesthetized by isoflurane and brains collected in ice cold HBSS. 500um coronal brain sections were produced using an ice-cold brain matrix (AgnThos 69-2165-1), and somatosensory cortices dissected under a fluorescence stereomicroscope using cre-dependent GFP signal expressed by infected cells and morphological landmarks. Tissue pieces from each mouse were collected in a separate tube, flash-frozen in liquid nitrogen and stored at -80C until needed. For protein extraction and pulldown, we extensively optimized published protocols. Frozen tissue was slowly thawed on ice and tissue from 2 animals was combined in a single tube as a replicate. Tissue was homogenized in 500ul 1% SDS RIPA buffer (50mM Tris pH 8, 150mM NaCl, 1% SDS, 0.5% Sodium Deoxycholate, 1% Triton-X 100, 1X protease inhibitors) using a small glass Dounce homogenizer. Lysates were sonicated 3x30s with 30s rest on ice in between. Lysates were boiled at 95C for 5 minutes to dissociate PSD and transsynaptic complexes and diluted to 0.5% SDS with ice cold SDS-free RIPA buffer as reported in [96]. After 1 hour incubation at 4C with rotation, lysates were cleared by ultracentrifugation at 100.000G for 30 minutes at 4C and supernatant collected. BCA assay was used to quantify protein concentration. 10ug protein input was used for Western blot experiments in Figure 1F and Figure 1G. 200 ug protein input was used for pulldown for western blots in Figure 1H. Around 1000ug protein input was used for pulldown for MS. 5ul of streptavidin magnetic beads (Pierce) per 100ug of protein input was used for every condition. Beads were washed 3 times in 0.5% SDS RIPA buffer using a magnetic rack and incubated overnight at 4C with rotation with protein input in 1ml of 0.5% SDS RIPA buffer final volume. The day after beads were washed 3x with 0.5% SDS RIPA buffer, 1X with 1M KCl, 1x with 0.1M Na2CO3, 1x with 2M Urea in 50mM Tris-HCl ph 8. MS sample preparation steps are described below, for western blot analysis beads were further washed 3x with 0.5% SDS RIPA buffer and proteins eluted in 1X protein loading buffer supplemented with 10mM DTT and 2mM biotin 5 mins at 95C.

Western blot After boiling 5 minutes at 95C, samples were loaded on 4-20% or 7.5% polyacrylamide gels and run at 180V. Proteins were transferred to 0.2um nitrocellulose membrane using semi-dry transfer (Biorad) using the mixed-molecular weight program. Membranes were blocked in 5% milk in TBS-T buffer (25mM Tris-base pH 7.5, 300mM NaCl, 0.05% Tween-20) 1h at RT. Primary and secondary antibodies were diluted in 5% milk in TBS-T. Primary antibodies were incubated O/N at 4C, secondary antibodies were incubated 1h at RT. When blotting for biotinylated proteins, HRP-conjugated streptavidin was diluted 1:500 in 5% BSA to avoid binding to the biotin present in the milk. We obtain best results in streptavidin signal while blocking membranes in milk, however we noticed that some

powdered milk gave higher background than other. Blocking in 5% BSA is an alternative. After primary and secondary antibodies membranes were washed 5 times with TBS-T and once with 1X PBS before development at ImageQuant 800 (Cytiva) using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher 34577).

MS sample preparation Beads were washed 2 more times with 2M Urea in 50mM Tris-HCl pH 8. Proteins were eluted from the beads and digested to peptides using S-Trap kit (Protify) following manufacturer instructions. Beads were resuspended in 92 ul of 5% SDS 50 mM TEAB pH 8, boiled 5 minutes at 95C and let cool down. Eluates were reduced with 4 ul of a 120 mM TCEP buffer at 55 degrees for 15 mins. Eluates were alkylated with 4 ul of 500 mM MMTS in isopropanol at RT for 10 mins in the dark. Eluates were acidified with 10 ul of 27.5% phosphoric acid in dH2O. Using a magnetic rack to separate the beads, supernatants were collected in a different tube and diluted in 660ul of binding/wash buffer. Supernatant were loaded onto an S-trap column, centrifuged at 4000G 1 minute. Columns were washed 10 times with binding/wash buffer. Proteins were digested to peptides applying 1ug of trypsin diluted in 20ul of digestion buffer directly to the column and incubated overnight at 37C in a humidified oven. Peptides were sequentially eluted from the column using 40ul of elution buffer 1, 2 and 3 and dried using a speedVac before being shipped for LC-MS/MS.

Mass Spectrometry Peptides were analysed by data independent acquisition (DIA) mass spectrometry as described previously[97,98] . In summary, peptides were injected onto a nanoscale C18 reverse-phase chromatography system (UltiMate 3000 RSLC nano, Thermo Scientific) and electrosprayed into an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher). For liquid chromatography the following buffers were used: buffer A (0.1% formic acid in Milli-Q water (v/v)) and buffer B (80% acetonitrile and 0.1% formic acid in Milli-Q water (v/v). Samples were loaded at 10 µL/min onto a trap column (100 µm × 2 cm, PepMap nanoViper C18 column, 5 µm, 100 Å, Thermo Scientific) equilibrated in 0.1% trifluoroacetic acid (TFA). The trap column was washed for 3 min at the same flow rate with 0.1% TFA then switched in-line with a Thermo Scientific, resolving C18 column (75 µm × 50 cm, PepMap RSLC C18 column, 2 µm, 100 Å). Peptides were eluted from the column at a constant flow rate of 300 nl/min with a linear gradient from 3% buffer B to 6% buffer B in 5 min, then from 6% buffer B to 35% buffer B in 115 min, and finally to 80% buffer B within 7 min. The column was then washed with 80% buffer B for 4 min and re-equilibrated in 3% buffer B for 15 min. Two blanks were run between each sample to reduce carry-over. The column was kept at a constant temperature of 50oC.

The data was acquired using an easy spray source operated in positive mode with spray voltage at 2.445 kV, and the ion transfer tube temperature at 250oC. The MS was operated in DIA mode. A scan cycle comprised a full MS scan (m/z range from 350-1650), with RF lens at 40%, AGC target set to custom, normalized AGC target at 300%, maximum injection time mode set to custom, maximum injection time at 20 ms, microscan set to 1 and source fragmentation disabled. MS survey scan was followed by MS/MS DIA scan events using the following parameters: multiplex ions set to false, collision energy mode set to stepped, collision energy type set to normalized, HCD collision energies set to 25.5, 27 and 30%, orbitrap resolution 30000, first mass 200, RF lens 40%, AGC target set to custom,

normalized AGC target 3000%, microscan set to 1 and maximum injection time 55 ms. Data for both MS scan and MS/MS DIA scan events were acquired in profile mode.

MS data analysis RAW MS data were analysed by DIA-N-N version 1.8.1[44]. Uniprot *Mus musculus* UP000000589 reference proteome was used for spectral library preparation. Precursor m/z range was set at 350 to 1650. Neural network classifier was set at double-pass mode. Every other parameter was kept at default. The resulting file was loaded onto Perseus for analysis[99]. A list of contaminants was created from the no.turboID samples. First, only proteins identified in all no.turboID samples were kept and ranked by intensity. Then, known PSD proteins were used to set a threshold defining true contaminants (Figure S2E). All keratins were considered contaminants and included in this list. Proteins in this list were removed from all other samples. Samples were log2 transformed and normalized by 'substract median' method. Only proteins present at least 3 times in at least one condition were kept (e.g. in all 3 Tlx3-Cre PSD.turboID replicates) and missing values imputed with standard parameters. To enrich for synaptic proteins, we applied a two samples t-test with permutation-based FDR at 10% and s0 > 0.1 between the cytoplasmic.turboID and PSD.turboID samples for each cre line. We then merged PSD.turboID significant proteins from Tlx3-Cre and Sim1-Cre to create the 'synaptic protein list' used for gene ontology and further comparison. To identify differentially expressed proteins between L5a and L5b neurons a t test was performed on the synaptic proteins with using an FDR < 0.05 and s0 > 0.1 to calculate significance. Raw MS data as well as the analysis workflow will be deposited at the ProteomeXchange data repository upon publication.

Immunofluorescence P35 mice were anesthetized by intraperitoneal injection with a lethal dose of 1 µl/g xylazine (VMB Xyl-M 2%), 2 µl/g ketamine (Eurovet Nimatek, 100 mg/ml), and 3 µl/g 0.9% saline. Next, mice were transcardially perfused with 4% PFA, in PBS. Brains were dissected, postfixed in 4% PFA in PBS at 4 °C overnight, washed in PBS and sliced to 50um thick by vibratome. Sections were blocked in 10% NHS, 0.5M Glycine, 0.5% Triton-X 100, 0.02% gelatin in 1X PBS for 2h before incubation with antibodies in 5% NHS, 0.5% Triton-X 100, 0.02% gelatin in 1X PBS with extensive washing with 0.05% Triton-X 100 in 1X PBS between primary and secondary antibodies. Sections were counterstained with DAPI before being mounted on microscope slides with Mowiol-4-88 (Millipore). Imaging was performed on a Slide Scanner Axio Scan.Z1 (Zeiss) or LSM880 with Airyscan (Zeiss).

For spine density analysis, mice were perfused with 4% PFA (EM grade, Science Services), 2% Sucrose in 0.1M PB buffer pH 7.4 and post-fixed 2 h at RT. Brains were sliced at 120um to increase the portion of dendritic tree for each positive neuron.

Mice injected with turboID AAVs were injected with 2mg/ml biotin in PBS 3h before perfusion.

RNAscope in situ hybridization Mice injected with cre-dependent GFP AAV were transcardially perfused at P35 as described. 20um slices were produced on a Vibrating Microtome 7000 (Campden Instruments LTD), mounted on SuperFrost Ultra Plus adhesion slides (Thermo-Fisher) and dried at RT. Sections were stained using the RNAscope™ Multiplex Fluorescent Detection Kit v2 (ACDbio, 323110) following manufacturer protocols. Briefly, sections were baked at 40C for 30 minutes to improve adhesion, treated for 10 minutes with hydrogen peroxide solution and 12 minutes with protease IV

washing with 1X PBS in between. RNAscope probes were hybridized at 40C for 2h and subsequently HRP signal developed using OPAL dyes. To counterstain the GFP, sections were blocked for 1h at RT in 5% NHS, 0.3% TritonX-100, in PBS. GFP antibody was diluted in blocking solution 1:500 and added O/N at 4C. after 3 washes in 1X PBS, anti-chicken 488 secondary antibody was added 1:1000 for 1h at RT. After 3 more washes in 1X PBS, sections were counterstained with DAPI for 5 minutes, and coverslipped with Mowiol. Images of GFP positive neurons were acquired on a Zeiss LSM900 using a 20X objective. Same laser power and gain settings were used to image the same gene in the two different mouse lines. Single dots were counted manually.

Dendritic spine imaging and analysis Adult mice were injected in the barrel cortex with 200nl of AAV mix consisting of pAAV-TRE-DIO-FLPo (Addgene #118027), pAAV-TRE-fDIO-GFP-IRES-tTA (Addgene #118026) and either cytoplasmic.turboID, PSD.turboID or nothing. A week later, mice were injected with 1ml of 2mg/ml biotin in 1X PBS for 3 hours before perfusion with 4% PFA, 2% Sucrose in 0.1M PB buffer pH 7.4 and post-fixed 2 h at RT. Brains were sliced at 120um to increase the portion of dendritic tree for each positive neuron. Brain slices were stained for GFP (antibody) and biotin (antibody). Dendritic spines images were acquired on a Zeiss LSM880 with airyscan detector using a 63X objective and 2.0 zoom. Optimal settings suggested by the software were used for pixel size and optical section size. Laser intensity and gain were adjusted for each image to obtain the best signal to noise ratio for GFP channel. Laser intensity and gain were kept the same for all conditions for the biotin channel. For each brain, 3 dendrites were imaged. Dendritic spines were reconstructed using neurolucida software and classified by morphology.

Antibodies

rat anti HA (Merck, 11867423001, WB: 1:1000, IHC: 1:1000),

Alexa 647-conjugated streptavidin (Thermo Fisher Scientific, S32357, IHC: 1:500),

rabbit Homer1 (Synaptic Systems, 160 003, WB: 1:5000, IHC: 1:1000).

HRP conjugated streptavidin (Thermo Fisher Scientific, SA10001, WB: 1:5000),

mouse anti Gria1 (Millipore , MAB2263, WB: 1:500),

chicken anti-GFP (Aveslabs, GFP-1020, IHC: 1:500),

rabbit anti Tubulin 3 (Abcam, ab18207, WB: 1:2500),

rabbit anti VGLUT2 (Synaptic Systems, 138-403, IHC: 1:5000),

Alexa fluor conjugated secondary antibodies from Invitrogen (1:1000).

Statistical analysis and software versions

Gene ontology analysis was performed using the R package gprofiler2 (0.2.2) based on gProfiler[45]. All proteins identified by MS were used as a reference background. Data were plotted using the R package ggplot2 (3.4.2). Data from Sorokina, Mclean, Croning[46] were accessed with the R package RSQLite (2.3.1) or directly with MySQLlite software. SYNGO[47] database was imported to R from https://syngoportal.org/. Genetrek database was imported to R from https://genetrek.pasteur.fr/ . SFARI

database was imported to R from https://gene.sfari.org/ . MS data analysis software and statistic is described in the methods, statistical tests and significance levels are indicated in the relevant figures.

References

1. Petreanu, L., Mao, T., Sternson, S. M. & Svoboda, K. The subcellular organization of neocortical excitatory connections. *Nature* **457**, 1142–1145 (2009).
2. Harris, K. D. & Mrsic-Flogel, T. D. Cortical connectivity and sensory coding. *Nature* **503**, 51–58 (2013).
3. Feldmeyer, D. Excitatory neuronal connectivity in the barrel cortex. *Frontiers in Neuroanatomy* **6**, (2012).
4. Spruston, N. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience* **9**, 206–221 (2008).
5. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat Neurosci* **18**, 170–181 (2015).
6. Suzuki, M. & Larkum, M. E. General Anesthesia Decouples Cortical Pyramidal Neurons. *Cell* **180**, 666-676.e13 (2020).
7. Sermet, B. S. *et al.* Pathway-, layer- and cell-type-specific thalamic input to mouse barrel cortex. *eLife* **8**,.
8. Audette, N. J., Urban-Ciecko, J., Matsushita, M. & Barth, A. L. POm Thalamocortical Input Drives Layer-Specific Microcircuits in Somatosensory Cortex. *Cerebral Cortex* **28**, 1312–1328 (2018).
9. Bureau, I., von Saint Paul, F. & Svoboda, K. Interdigitated Paralemniscal and Lemniscal Pathways in the Mouse Barrel Cortex. *PLoS Biology* **4**, e382 (2006).
10. Yao, S. *et al.* A whole-brain monosynaptic input connectome to neuron classes in mouse visual cortex. *Nat Neurosci* **26**, 350–364 (2023).
11. Kim, E. J., Juavinett, A. L., Kyubwa, E. M., Jacobs, M. W. & Callaway, E. M. Three Types of Cortical Layer 5 Neurons That Differ in Brain-wide Connectivity and Function. *Neuron* **88**, 1253–1267 (2015).

12. Takahashi, N. *et al.* Active dendritic currents gate descending cortical outputs in perception. *Nature Neuroscience* **23**, 1277–1285 (2020).
13. Baker, A. *et al.* Specialized Subpopulations of Deep-Layer Pyramidal Neurons in the Neocortex: Bridging Cellular Properties to Functional Consequences. *J. Neurosci.* **38**, 5441–5455 (2018).
14. Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R. & Svoboda, K. A motor cortex circuit for motor planning and movement. *Nature* **519**, 51–56 (2015).
15. Favuzzi, E. *et al.* Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science* **363**, 413–417 (2019).
16. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
17. Di Bella, D. J. *et al.* Molecular Logic of Cellular Diversification in the Mouse Cerebral Cortex. *Nature* **595**, 554–559 (2021).
18. Velmeshev, D. *et al.* Single-cell genomics identifies cell type–specific molecular changes in autism. *Science* **364**, 685–689 (2019).
19. Willsey, A. J. *et al.* Coexpression Networks Implicate Human Midfetal Deep Cortical Projection Neurons in the Pathogenesis of Autism. *Cell* **155**, 997–1007 (2013).
20. Wamsley, B. *et al.* Molecular cascades and cell-type specific signatures in ASD revealed by single cell genomics. 2023.03.10.530869 Preprint at https://doi.org/10.1101/2023.03.10.530869 (2023).
21. Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat Rev Neurosci* **16**, 551–563 (2015).
22. Marcassa, G., Dascenco, D. & de Wit, J. Proteomics-based synapse characterization: From proteins to circuits. *Current Opinion in Neurobiology* **79**, 102690 (2023).
23. Yim, Y. Y. & Nestler, E. J. Cell-Type-Specific Neuroproteomics of Synapses. *Biomolecules* **13**, 998 (2023).

24. Hindley, N., Sanchez Avila, A. & Henstridge, C. Bringing synapses into focus: Recent advances in synaptic imaging and mass-spectrometry for studying synaptopathy. *Frontiers in Synaptic Neuroscience* **15**, (2023).
25. Fernández, E. *et al.* Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins. *Mol Syst Biol* **5**, 269 (2009).
26. Wang, L. *et al.* A cross-species proteomic map reveals neoteny of human synapse development. *Nature* 1–8 (2023) doi:10.1038/s41586-023-06542-2.
27. Roy, M. *et al.* Proteomic analysis of postsynaptic proteins in regions of the human neocortex. *Nat Neurosci* **21**, 130–138 (2018).
28. Butko, M. T. *et al.* In vivo quantitative proteomics of somatosensory cortical synapses shows which protein levels are modulated by sensory deprivation. *PNAS* **110**, E726–E735 (2013).
29. Roy, M. *et al.* Regional Diversity in the Postsynaptic Proteome of the Mouse Brain. *Proteomes* **6**, 31 (2018).
30. Van Oostrum, M. *et al.* The proteomic landscape of synaptic diversity across brain regions and cell types. *Cell* S0092867423010826 (2023) doi:10.1016/j.cell.2023.09.028.
31. Biesemann, C. *et al.* Proteomic screening of glutamatergic mouse brain synaptosomes isolated by fluorescence activated sorting. *The EMBO Journal* **33**, 157–170 (2014).
32. Paget-Blanc, V. *et al.* A synaptomic analysis reveals dopamine hub synapses in the mouse striatum. *Nat Commun* **13**, 3102 (2022).
33. Apóstolo, N. *et al.* Synapse type-specific proteomic dissection identifies IgSF8 as a hippocampal CA3 microcircuit organizer. *Nat Commun* **11**, 5171 (2020).
34. Selimi, F., Cristea, I. M., Heller, E., Chait, B. T. & Heintz, N. Proteomic Studies of a Single CNS Synapse Type: The Parallel Fiber/Purkinje Cell Synapse. *PLoS Biol* **7**, e1000083 (2009).
35. Zhu, F. *et al.* Cell-type-specific visualisation and biochemical isolation of endogenous synaptic proteins in mice. *European Journal of Neuroscience* **51**, 793–805 (2020).

36. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of Cell Biology* **196**, 801–810 (2012).
37. Uezu, A. *et al.* Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353**, 1123–1129 (2016).
38. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880–887 (2018).
39. Luo, W. *et al.* Supernova: A Versatile Vector System for Single-Cell Labeling and Gene Function Studies *in vivo*. *Scientific Reports* **6**, 35747 (2016).
40. May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* **9**, 1070 (2020).
41. Kirkeby, S., Moe, D., Bøg-Hansen, T. C. & van Noorden, C. J. F. Biotin carboxylases in mitochondria and the cytosol from skeletal and cardiac muscle as detected by avidin binding. *Histochemistry* **100**, 415–421 (1993).
42. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat Neurosci* **22**, 1182–1195 (2019).
43. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC Cre-Recombinase Driver Lines to Study the Functional Organization of Cerebral Cortical and Basal Ganglia Circuits. *Neuron* **80**, 1368–1383 (2013).
44. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, 41–44 (2020).
45. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* **47**, W191–W198 (2019).
46. Sorokina, O. *et al.* A unified resource and configurable model of the synapse proteome and its role in disease. *Sci Rep* **11**, 9967 (2021).

47. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).
48. Thomas, P. D. *et al.* PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science* **31**, 8–22 (2022).
49. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**, D523–D531 (2023).
50. Kohama, S. G. & Urbanski, H. F. Distribution of glutamate receptor subunits in the primate temporal cortex and hippocampus. *Brain Research* **769**, 44–56 (1997).
51. Cremer, C. M., Lübke, J. H. R., Palomero-Gallagher, N. & Zilles, K. Laminar distribution of neurotransmitter receptors in different reeler mouse brain regions. *Brain Struct Funct* **216**, 201–218 (2011).
52. Bailey, A. *et al.* Regional mapping of low-affinity kainate receptors in mouse brain using [3H](2S,4R)-4-methylglutamate autoradiography. *European Journal of Pharmacology* **431**, 305–310 (2001).
53. de Wit, J. *et al.* Unbiased Discovery of Glypican as a Receptor for LRRTM4 in Regulating Excitatory Synapse Development. *Neuron* **79**, 696–711 (2013).
54. Cao, Y. *et al.* Interplay between cell-adhesion molecules governs synaptic wiring of cone photoreceptors. *Proc Natl Acad Sci U S A* **117**, 23914–23924 (2020).
55. Choi, Y. *et al.* NGL-1/LRRC4C Deletion Moderately Suppresses Hippocampal Excitatory Synapse Development and Function in an Input-Independent Manner. *Frontiers in Molecular Neuroscience* **12**, (2019).
56. Mehta, D., Ahkami, A. H., Walley, J., Xu, S.-L. & Uhrig, R. G. The incongruity of validating quantitative proteomics using western blots. *Nat. Plants* **8**, 1320–1321 (2022).
57. Ayoubi, R. *et al.* Scaling of an antibody validation procedure enables quantification of antibody performance in major research applications. *eLife* **12**, (2023).
58. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

59. Carlyle, B. C. *et al.* A multiregional proteomic survey of the postnatal human brain. *Nat Neurosci* **20**, 1787–1795 (2017).
60. Li, J. *et al.* Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. *Cell* **180**, 373-386.e15 (2020).
61. Hall, J. & Bray, N. J. Schizophrenia Genomics: Convergence on Synaptic Development, Adult Synaptic Plasticity, or Both? *Biol Psychiatry* **91**, 709–717 (2022).
62. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
63. de Wit, J. & Ghosh, A. Control of neural circuit formation by leucine-rich repeat proteins. *Trends in Neurosciences* **37**, 539–550 (2014).
64. Leblond, C. S. *et al.* Operative list of genes associated with autism and neurodevelopmental disorders based on database review. *Molecular and Cellular Neuroscience* **113**, 103623 (2021).
65. Abrahams, B. S. *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism* **4**, 36 (2013).
66. de Wit, J., Hong, W., Luo, L. & Ghosh, A. Role of Leucine-Rich Repeat Proteins in the Development and Function of Neural Circuits. *Annual Review of Cell and Developmental Biology* **27**, 697–729 (2011).
67. Schroeder, A. & de Wit, J. Leucine-rich repeat-containing synaptic adhesion molecules as organizers of synaptic specificity and diversity. *Exp Mol Med* **50**, 10 (2018).
68. Ko, J. & Kim, E. Leucine-rich repeat proteins of synapses. *J Neurosci Res* **85**, 2824–2832 (2007).
69. Schroeder, A. *et al.* A Modular Organization of LRR Protein-Mediated Synaptic Adhesion Defines Synapse Identity. *Neuron* **99**, 329-344.e7 (2018).
70. DeNardo, L. A., de Wit, J., Otto-Hitt, S. & Ghosh, A. NGL-2 Regulates Input-Specific Synapse Development in CA1 Pyramidal Neurons. *Neuron* **76**, 762–775 (2012).

71. Nishimura-Akiyoshi, S., Niimi, K., Nakashiba, T. & Itohara, S. Axonal netrin-Gs transneuronally determine lamina-specific subdendritic segments. *Proceedings of the National Academy of Sciences* **104**, 14801–14806 (2007).
72. Young, H., Belbut, B., Baeta, M. & Petreanu, L. Laminar-specific cortico-cortical loops in mouse visual cortex. *eLife* **10**, e59551 (2021).
73. Parikshak, N. N. *et al.* Integrative Functional Genomic Analyses Implicate Specific Molecular Pathways and Circuits in Autism. *Cell* **155**, 1008–1021 (2013).
74. Gandal, M. J. *et al.* Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD. *Nature* **611**, 532–539 (2022).
75. Cerliani, L. *et al.* Increased Functional Connectivity Between Subcortical and Cortical Resting-State Networks in Autism Spectrum Disorder. *JAMA Psychiatry* **72**, 767–777 (2015).
76. Woodward, N. D., Giraldo-Chica, M., Rogers, B. & Cascio, C. J. Thalamocortical dysconnectivity in autism spectrum disorder: An analysis of the Autism Brain Imaging Data Exchange. *Biol Psychiatry Cogn Neurosci Neuroimaging* **2**, 76–84 (2017).
77. Iidaka, T., Kogata, T., Mano, Y. & Komeda, H. Thalamocortical Hyperconnectivity and Amygdala-Cortical Hypoconnectivity in Male Patients With Autism Spectrum Disorder. *Front Psychiatry* **10**, 252 (2019).
78. Ayub, R. *et al.* Thalamocortical connectivity is associated with autism symptoms in high-functioning adults with autism and typically developing adults. *Transl Psychiatry* **11**, 93 (2021).
79. Hong, C.-J. & Hsueh, Y.-P. Cytoplasmic distribution of T-box transcription factor Tbr-1 in adult rodent brain. *J Chem Neuroanat* **33**, 124–130 (2007).
80. Chuang, H.-C., Huang, T.-N. & Hsueh, Y.-P. Neuronal excitation upregulates Tbr1, a high-confidence risk gene of autism, mediating Grin2b expression in the adult brain. *Front Cell Neurosci* **8**, 280 (2014).
81. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006).

82. Guydosh, N. R. & Green, R. Dom34 Rescues Ribosomes in 3′ Untranslated Regions. *Cell* **156**, 950–962 (2014).

83. Takano, T. *et al.* Chemico-genetic discovery of astrocytic control of inhibition in vivo. *Nature* **588**, 296–302 (2020).

84. Sun, X. *et al.* Deep single-cell type proteome profiling of mouse brain by nonsurgical AAV-mediated proximity labeling. *Anal Chem* **94**, 5325–5334 (2022).

85. Rayaprolu, S. *et al.* Cell type-specific biotin labeling in vivo resolves regional neuronal and astrocyte proteomic differences in mouse brain. *Nat Commun* **13**, 2927 (2022).

86. Gao, Y. *et al. Proximity Analysis of Native Proteomes Reveals Interactomes Predictive of Phenotypic Modifiers of Autism and Related Neurodevelopmental Conditions*. http://biorxiv.org/lookup/doi/10.1101/2022.10.06.511211 (2022) doi:10.1101/2022.10.06.511211.

87. Sharma, K. *et al.* Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci* **18**, 1819–1831 (2015).

88. Li, K. W., Gonzalez-Lozano, M. A., Koopmans, F. & Smit, A. B. Recent Developments in Data Independent Acquisition (DIA) Mass Spectrometry: Application of Quantitative Analysis of the Brain Proteome. *Front Mol Neurosci* **13**, 564446 (2020).

89. Artan, M. *et al.* Interactome analysis of Caenorhabditis elegans synapses by TurboID-based proximity labeling. *Journal of Biological Chemistry* **297**, (2021).

90. Luján, R., Roberts, J. D. B., Shigemoto, R., Ohishi, H. & Somogyi, P. Differential plasma membrane distribution of metabotropic glutamate receptors mGluR1α, mGluR2 and mGluR5, relative to neurotransmitter release sites. *Journal of Chemical Neuroanatomy* **13**, 219–241 (1997).

91. Nusser, Z., Mulvihill, E., Streit, P. & Somogyi, P. Subsynaptic segregation of metabotropic and ionotropic glutamate receptors as revealed by immunogold localization. *Neuroscience* **61**, 421–427 (1994).

92. Willems, J. *et al.* ORANGE: A CRISPR/Cas9-based genome editing toolbox for epitope tagging of endogenous proteins in neurons. *PLOS Biology* **18**, e3000665 (2020).

93. Rimbault, C. *et al.* Engineering selective competitors for the discrimination of highly conserved protein-protein interaction modules. *Nat Commun* **10**, 4521 (2019).
94. Daigle, T. L. *et al.* A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* **174**, 465-480.e22 (2018).
95. Challis, R. C. *et al. Widespread and targeted gene expression by systemic AAV vectors: Production, purification, and administration*. http://biorxiv.org/lookup/doi/10.1101/246405 (2018) doi:10.1101/246405.
96. Loh, K. H. *et al.* Proteomic Analysis of Unbounded Cellular Compartments: Synaptic Clefts. *Cell* **166**, 1295-1307.e21 (2016).
97. Walgrave, H. *et al.* microRNA-132 regulates gene expression programs involved in microglial homeostasis. *iScience* **26**, 106829 (2023).
98. Molina-Gonzalez, I. *et al.* Astrocyte-oligodendrocyte interaction regulates central nervous system regeneration. *Nat Commun* **14**, 3372 (2023).
99. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731–740 (2016).

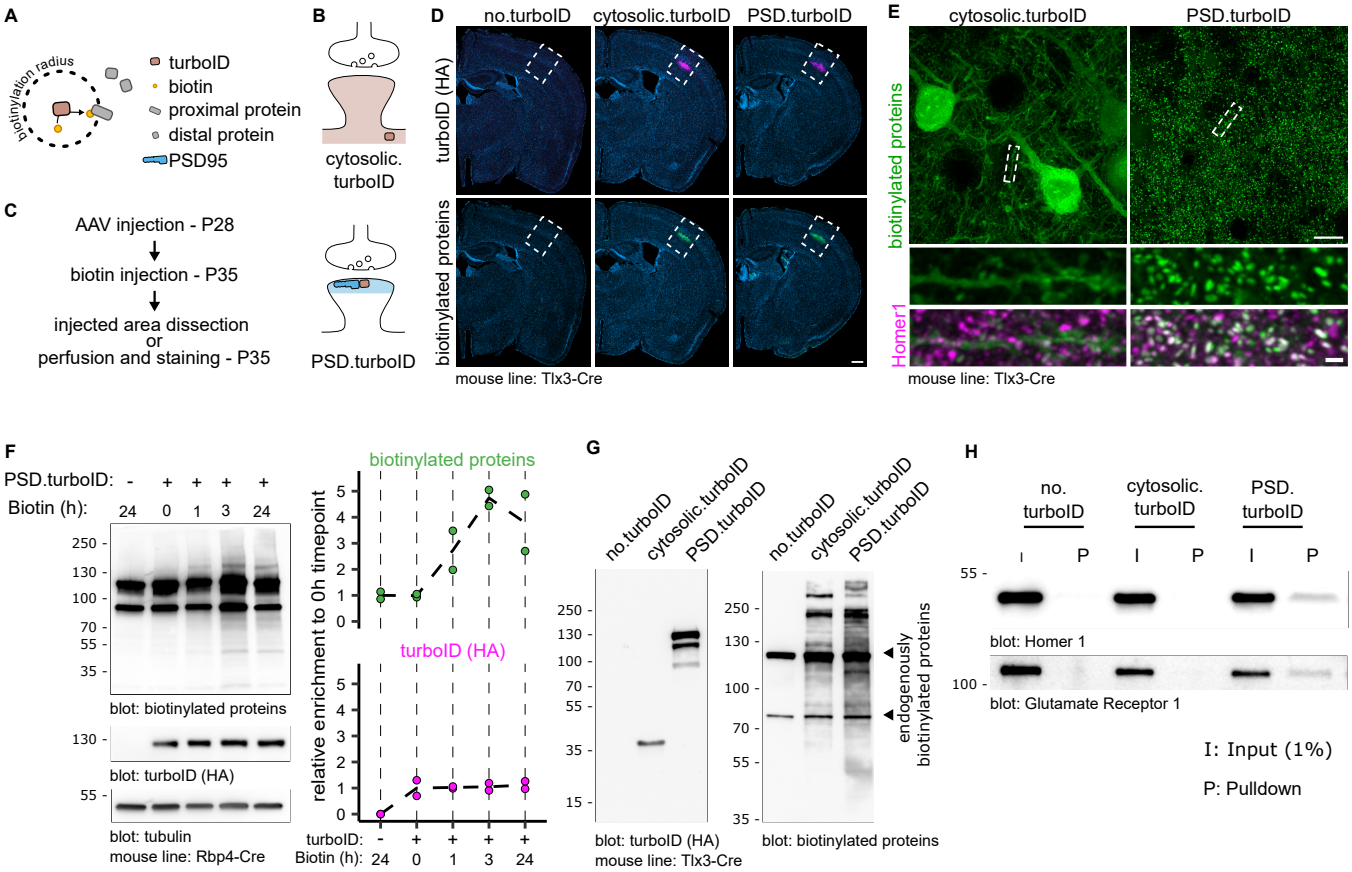
A
biotinylation radius
turboID
biotin
proximal protein
distal protein
PSD95
B
cytosolic. turboID
PSD.turboID
C
AAV injection - P28
biotin injection - P35
injected area dissection
or
perfusion and staining - P35
D
no.turboID
cytosolic.turboID
PSD.turboID
turboID (HA)
biotinylated proteins
mouse line: Tlx3-Cre
E
cytosolic.turboID
PSD.turboID
biotinylated proteins
Homer1
mouse line: Tlx3-Cre
F
PSD.turboID: - + + + +
Biotin (h): 24 0 1 3 24
blot: biotinylated proteins
blot: turboID (HA)
blot: tubulin
mouse line: Rbp4-Cre
biotinylated proteins
turboID (HA)
relative enrichment to 0h timepoint
turboID: - + + + +
Biotin (h): 24 0 1 3 24
G
no.turboID
cytosolic.turboID
PSD.turboID
blot: turboID (HA)
mouse line: Tlx3-Cre
endogenously biotinylated proteins
blot: biotinylated proteins
H
no. turboID
cytosolic. turboID
PSD. turboID
I P I P I P
blot: Homer 1
blot: Glutamate Receptor 1
I: Input (1%)
P: Pulldown

Figure 1. Fast and specific postsynaptic proximity biotinylation in genetically identified cortical layer 5 neurons.

A, TurboID covalently binds a biotin molecule to nearby proteins. **B**, A cytosolic.turboID construct labels cytosolic proteins. Fusion of turboID to PSD95 (PSD.turboID) directs its enzymatic activity to the postsynaptic compartment. **C,** Workflow for turboID construct testing. AAVs expressing different constructs are stereotactically injected in the somatosensory cortex of L5-specific Cre mice at P28. At P35, biotin is provided and mice are perfused for immunofluorescence analysis. For biochemical analysis of biotinylation, the transduced cortical area is dissected from fresh tissue. **D,** Whole-slice imaging of Tlx3-Cre mice injected with Cre-dependent turboID constructs labelled for HA-tagged turboID enzyme and biotinylated proteins. Scalebar is 200 μm. **E,** High resolution imaging shows biotinylated protein in cytosolic.turboID and PSD.turboID samples. Homer1 immunostaining labels postsynapses. Scalebar is 10 μm and 1 μm in the cropped region. **F,** Western blot quantification shows high levels of biotinylated proteins in dissected somatosensory cortices after 3 hours of biotin injection. Tubulin is used as loading control. (N = 2 mice per time point). **G,** Western blot for turboID confirms expression of constructs at the expected size. Both turboID constructs produce strong biotinylated protein signal across molecular weights compared to no.turboID sample which shows only bands for endogenously biotinylated proteins. **H,** After streptavidin pulldown, postsynaptic marker proteins Gria1 and Homer1 are only detected in PSD.turboID samples.

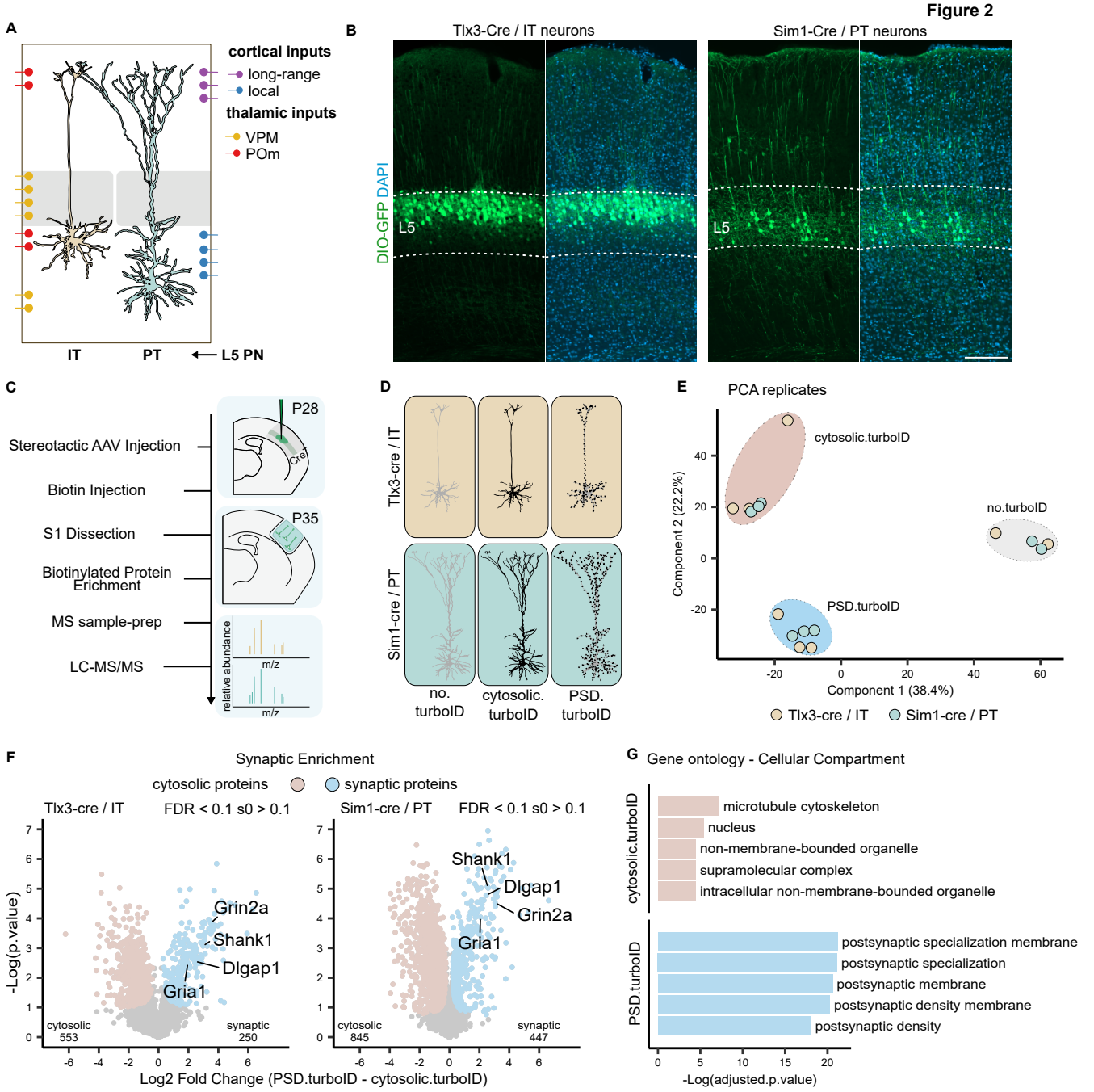
A
cortical inputs
long-range
local
thalamic inputs
VPM
POm
IT
PT
L5 PN
B
Tlx3-Cre / IT neurons
Sim1-Cre / PT neurons
DIO-GFP DAPI
L5
C
Stereotactic AAV Injection
Biotin Injection
S1 Dissection
Biotinylated Protein Erichment
MS sample-prep
LC-MS/MS
P28
Cre
P35
relative abundance
m/z
D
Tlx3-cre / IT
Sim1-cre / PT
no. turboID
cytosolic. turboID
PSD. turboID
E
PCA replicates
cytosolic.turboID
no.turboID
PSD.turboID
Component 2 (22.2%)
Component 1 (38.4%)
Tlx3-cre / IT
Sim1-cre / PT
F
Synaptic Enrichment
cytosolic proteins
synaptic proteins
Tlx3-cre / IT
FDR < 0.1 s0 > 0.1
Sim1-cre / PT
FDR < 0.1 s0 > 0.1
-Log(p.value)
Grin2a
Shank1
Dlgap1
Gria1
cytosolic 553
synaptic 250
cytosolic 845
synaptic 447
Log2 Fold Change (PSD.turboID - cytosolic.turboID)
G
Gene ontology - Cellular Compartment
cytosolic.turboID
microtubule cytoskeleton
nucleus
non-membrane-bounded organelle
supramolecular complex
intracellular non-membrane-bounded organelle
PSD.turboID
postsynaptic specialization membrane
postsynaptic specialization
postsynaptic membrane
postsynaptic density membrane
postsynaptic density
-Log(adjusted.p.value)

Figure 2. Identification of the postsynaptic proteome of layer 5 neurons.

A, Representation of L5 IT and PT neuron position and morphology in the somatosensory cortex with overview of input localization from other local and long-range excitatory neurons. **B,** AAV-mediated Cre-dependent GFP expression in Tlx3-Cre and Sim1-Cre mouse lines recapitulates soma position for L5 IT and PT neurons, respectively. Layer 5 borders were identified by density of DAPI-positive nuclei. Scalebar is 200 μm. **C,** Mice are injected with AAVs in the somatosensory cortex at P28. At P35, biotin is provided 3 hours before collection of somatosensory cortices. Biotinylated proteins are enriched by streptavidin pulldown and prepared for MS analysis. **D,** MS replicates were collected from Tlx3-Cre and Sim1-Cre mice injected with no.turboID, cytosolic.turboID or PSD.turboID. **E,** PCA analysis shows replicates clustering by condition but not by Cre line. **F,** Statistical analysis of cytosolic.turboID and PSD.turboID samples for each Cre line allows for synaptic protein enrichment. Core constituents of the excitatory postsynapse are annotated. Numbers of identified cytosolic and synaptic proteins for Tlx3-Cre and Sim1-Cre lines are indicated in bottom left and right corners of each graph. Unpaired t-test with permutation-based FDR correction at 10%, s0 > 0.1. **G,** Gene ontology analysis of the cytosolic and synaptic proteins from (f) showing cellular localization of identified proteins.

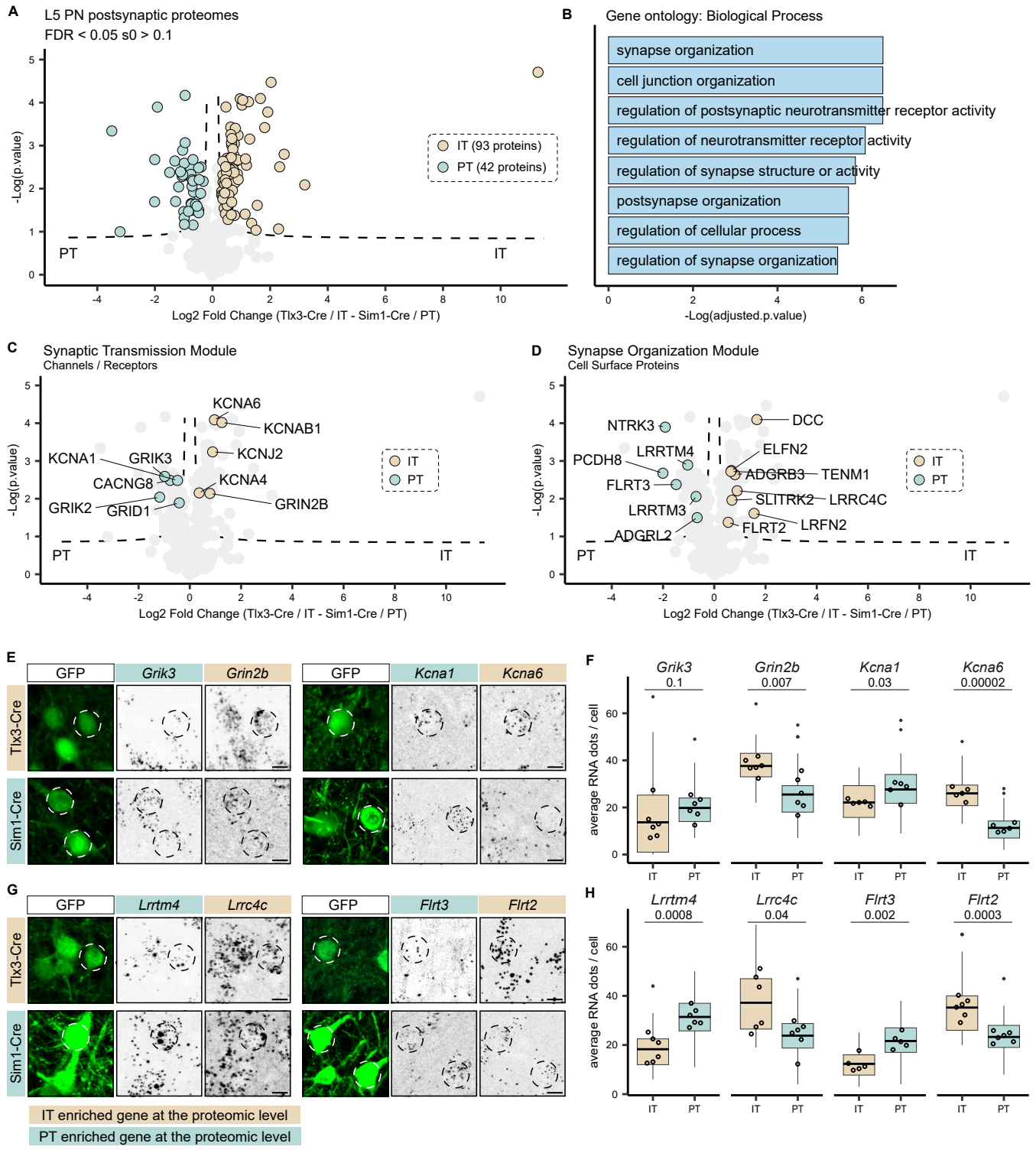
A
L5 PN postsynaptic proteomes
FDR < 0.05 s0 > 0.1
-Log(p.value)
Log2 Fold Change (Tlx3-Cre / IT - Sim1-Cre / PT)
IT (93 proteins)
PT (42 proteins)
PT
IT
B
Gene ontology: Biological Process
synapse organization
cell junction organization
regulation of postsynaptic neurotransmitter receptor activity
regulation of neurotransmitter receptor activity
regulation of synapse structure or activity
postsynapse organization
regulation of cellular process
regulation of synapse organization
-Log(adjusted.p.value)
C
Synaptic Transmission Module
Channels / Receptors
KCNA6
KCNAB1
KCNJ2
KCNA1
GRIK3
CACNG8
KCNA4
GRIK2
GRID1
GRIN2B
D
Synapse Organization Module
Cell Surface Proteins
NTRK3
DCC
LRRTM4
ELFN2
PCDH8
ADGRB3
TENM1
FLRT3
SLITRK2
LRRC4C
LRRTM3
FLRT2
LRFN2
ADGRL2
E
GFP
Grik3
Grin2b
Kcna1
Kcna6
Tlx3-Cre
Sim1-Cre
F
Grik3 0.1
Grin2b 0.007
Kcna1 0.03
Kcna6 0.00002
average RNA dots / cell
IT
PT
G
Lrrtm4
Lrrc4c
Flrt3
Flrt2
H
Lrrtm4 0.0008
Lrrc4c 0.04
Flrt3 0.002
Flrt2 0.0003
IT enriched gene at the proteomic level
PT enriched gene at the proteomic level

Figure 3. Molecular synaptic signatures of two populations of layer 5 neurons.

A, Relative protein enrichment in L5 IT and PT neurons. Significant proteins (unpaired t-test with permutation-based FDR correction at 5%, s0 > 0.1) are labelled according to Cre line (93 in Tlx3-Cre/IT and 42 in Sim1-Cre/PT neurons). **B,** Biological Process gene ontology analysis of differentially expressed proteins. **C,** Ion channels and neurotransmitter receptors are annotated to represent the 'synaptic transmission' module. **D,** CSPs are annotated to represent the 'synapse organization' module. **E, G,** smFISH analysis of candidate genes in L5 IT and PT neurons labelled by Cre-dependent GFP expression in Tlx3-Cre and Sim1-Cre mice, respectively. Cells that have been quantified are indicated by dotted circles. Scale bars 10 μm. **F, H,** Quantification of RNA expression (average number of RNA puncta per cell). N = 5 or 6 mice per Cre line, 8 cells per mouse are quantified. Boxplots (line across the box describes the mean) show distribution of single cells, empty dots display the average of each biological replicate. Statistic analysis is performed as an unpaired t-test, alpha = 0.05, p-value is reported under transcript name.

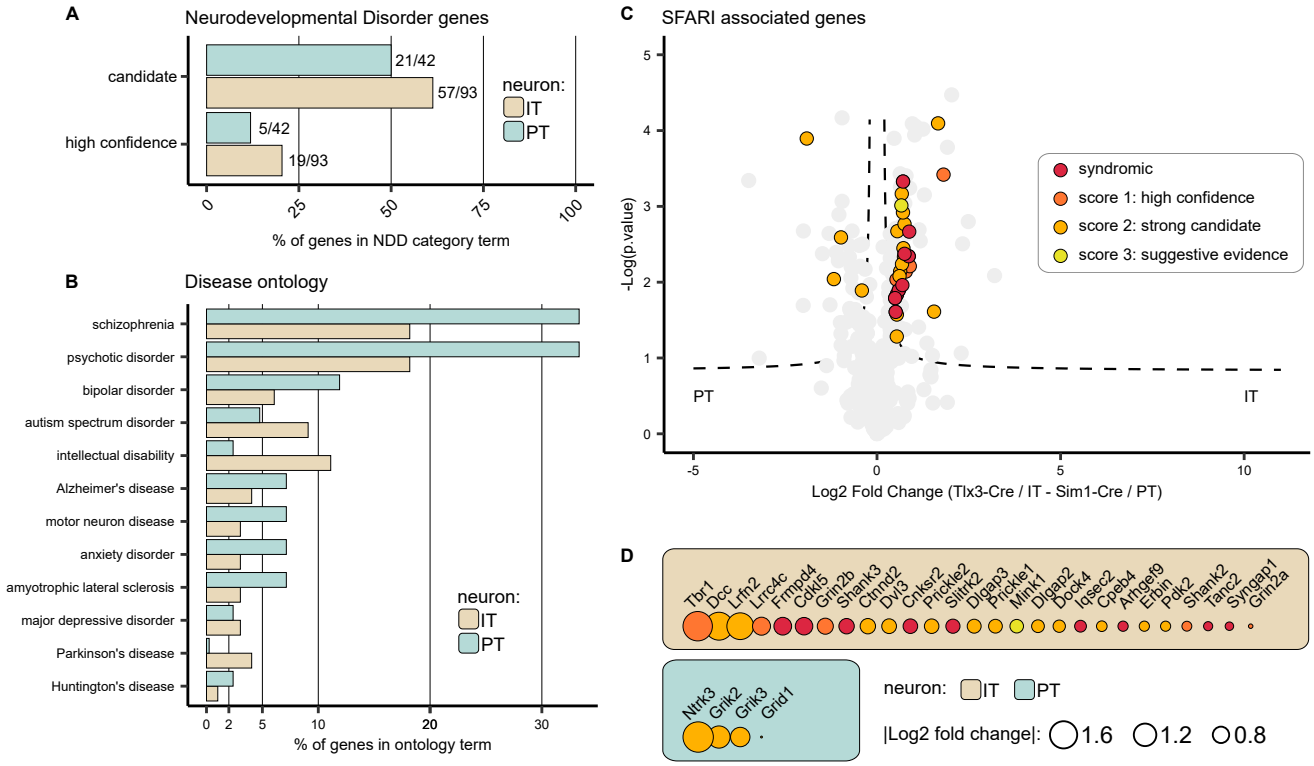
A
Neurodevelopmental Disorder genes
candidate
high confidence
21/42
57/93
5/42
19/93
neuron:
IT
PT
0
25
50
75
100
% of genes in NDD category term
B
Disease ontology
schizophrenia
psychotic disorder
bipolar disorder
autism spectrum disorder
intellectual disability
Alzheimer's disease
motor neuron disease
anxiety disorder
amyotrophic lateral sclerosis
major depressive disorder
Parkinson's disease
Huntington's disease
neuron:
IT
PT
0
2
5
10
20
30
% of genes in ontology term
C
SFARI associated genes
-Log(p.value)
syndromic
score 1: high confidence
score 2: strong candidate
score 3: suggestive evidence
PT
IT
-5
0
5
10
Log2 Fold Change (Tlx3-Cre / IT - Sim1-Cre / PT)
D
Tbr1
Dcc
Lrfn2
Lrrc4c
Frmpd4
Cdkl5
Grin2b
Shank3
Ctnnd2
Dvl3
Cnksr2
Prickle2
Sltrk2
Dlgap3
Prickle1
Mtnr1
Dlgap2
Dock4
Iqsec2
Cpeb4
Arhgef9
Erbin
Pdk2
Shank2
Tanc2
Syngap1
Grin2a
Ntrk3
Grik2
Grik3
Grid1
neuron:
IT
PT
|Log2 fold change|:
1.6
1.2
0.8

Figure 4. Differential disease vulnerability of layer 5 neuron molecular synaptic signatures.

A, Percentage of synaptic genes per cell type present in the 'high confidence' or 'candidate' neurodevelopmental disorder lists based on the Genetrek database published in Leblond et al., 2021[32]. **B,** Percentage of synaptic genes per cell type present in each disorder term based on the synapse proteome database[6]. **C,** Synaptic proteins in L5 IT and PT neurons with an ASD risk score annotation in the SFARI gene database[33] are indicated on the volcano plot showing protein distribution. **D,** ASD proteins identified in the SFARI database[33] listed per cell type and ranked by protein log fold change.

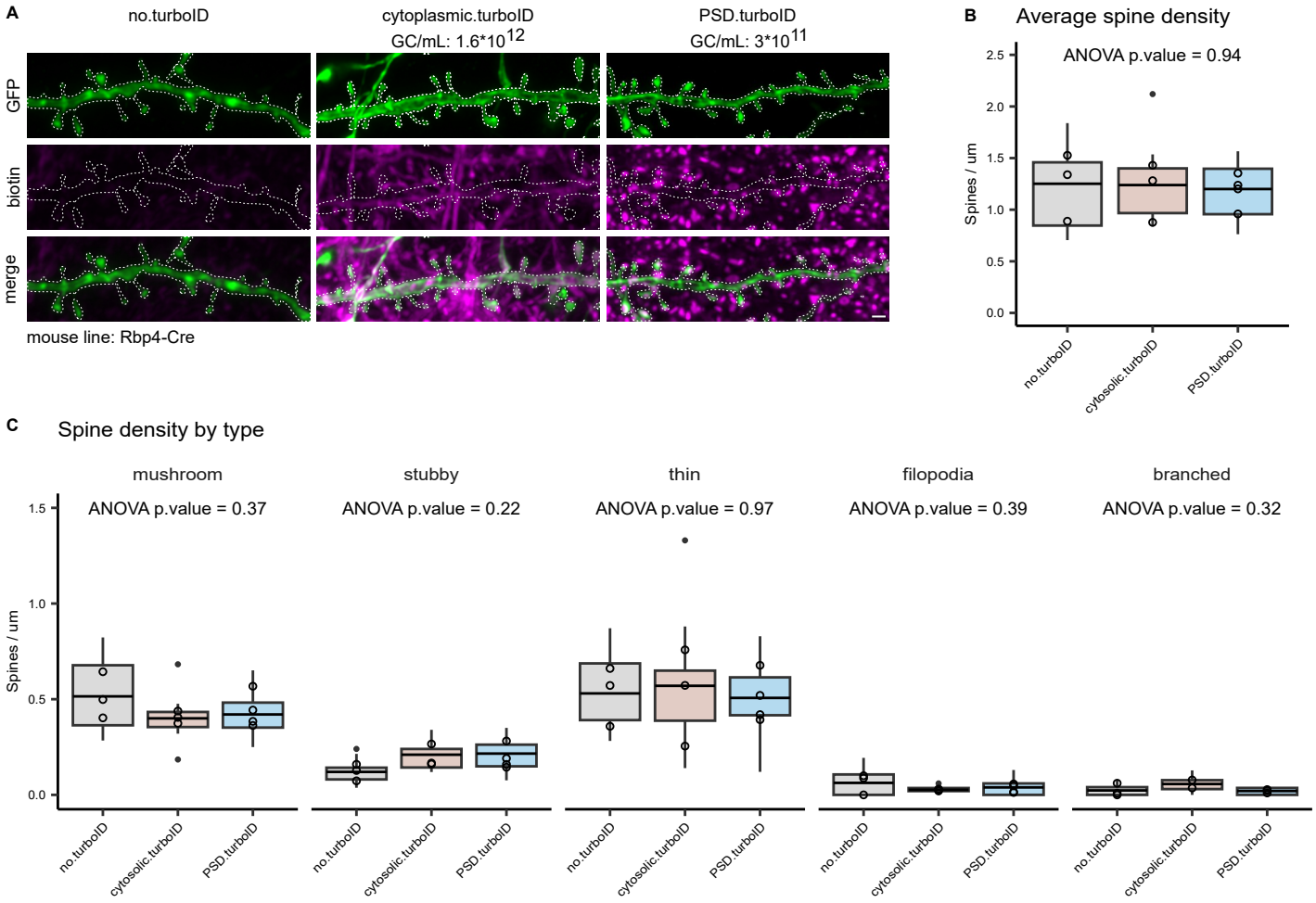
A
no.turboID
cytoplasmic.turboID
GC/mL: 1.6*10^12
PSD.turboID
GC/mL: 3*10^11
GFP
biotin
merge
mouse line: Rbp4-Cre
B
Average spine density
ANOVA p.value = 0.94
Spines / um
2.5
2.0
1.5
1.0
0.5
0.0
no.turboID
cytosolic.turboID
PSD.turboID
C
Spine density by type
mushroom
ANOVA p.value = 0.37
stubby
ANOVA p.value = 0.22
thin
ANOVA p.value = 0.97
filopodia
ANOVA p.value = 0.39
branched
ANOVA p.value = 0.32
Spines / um
1.5
1.0
0.5
0.0
no.turboID
cytosolic.turboID
PSD.turboID

Figure S1 related to Figure 1. TurboID expression has no effect on dendritic spine density or morphology.

A, Representative images of dendrites of L5 neurons expressing different turboID constructs. Viral titer is indicated as Genome Copies (GC) per mL. Scalebar is 1 μm. **B,** Quantification of spine density. **C,** Quantification of spine density by spine type. N = 3 mice per condition, 2 dendrites per mouse. Boxplots (line across the box describes the mean) show spine density distribution of individual dendrites, empty dots display the average spine density per each replicate. ANOVA statistical analysis is performed on the replicate means.

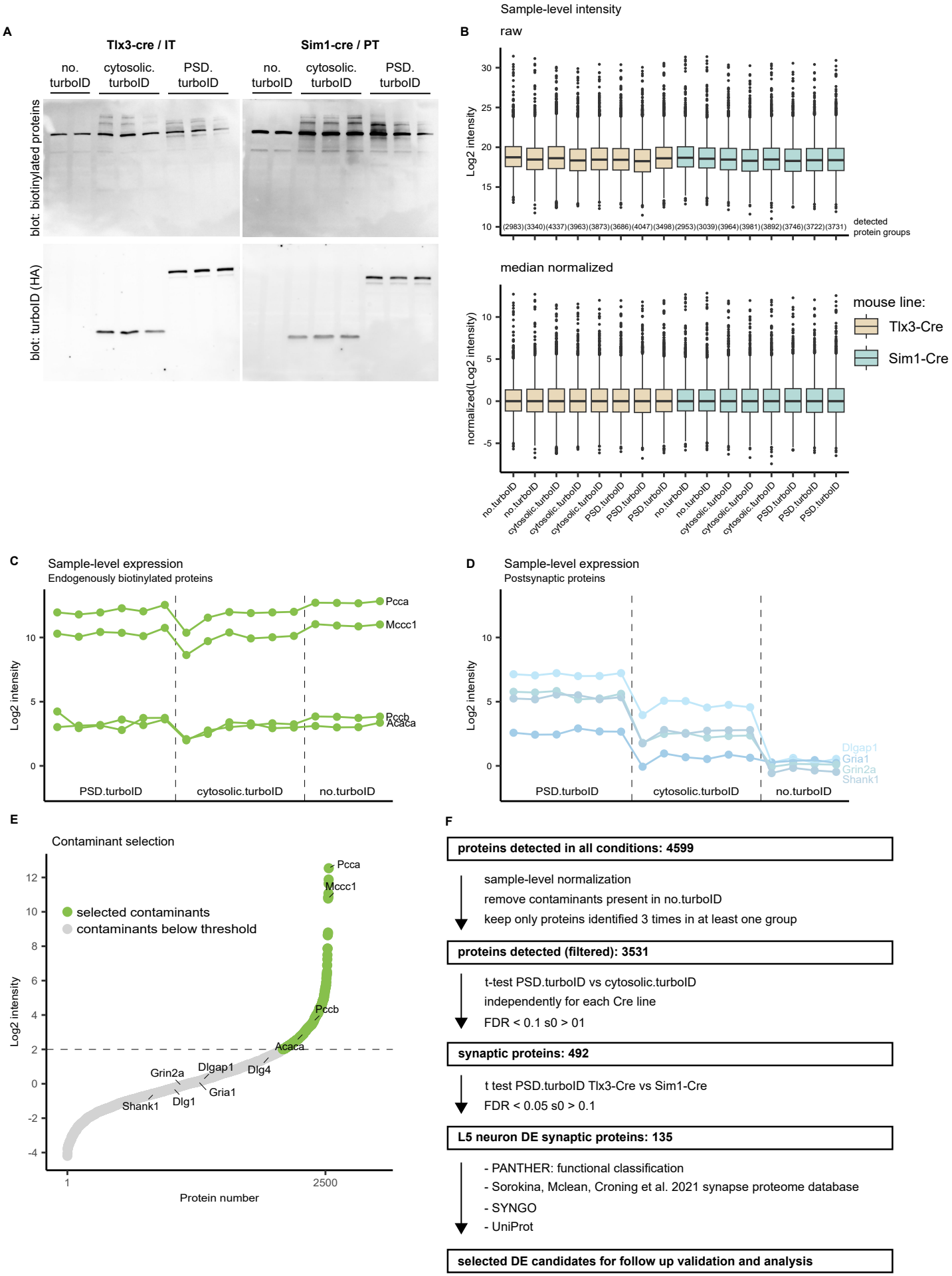
A
Tlx3-cre / IT
Sim1-cre / PT
no. turboID
cytosolic. turboID
PSD. turboID
no. turboID
cytosolic. turboID
PSD. turboID
blot: biotinylated proteins
blot: turboID (HA)
B
Sample-level intensity
raw
Log2 intensity
(2983)(3340)(4337)(3963)(3873)(3686)(4047)(3498)(2953)(3039)(3964)(3981)(3892)(3746)(3722)(3731)
detected protein groups
median normalized
normalized(Log2 intensity)
mouse line:
Tlx3-Cre
Sim1-Cre
no.turboID
cytosolic.turboID
PSD.turboID
C
Sample-level expression
Endogenously biotinylated proteins
Log2 intensity
Pcca
Mccc1
Pccb
Acaca
PSD.turboID
cytosolic.turboID
no.turboID
D
Sample-level expression
Postsynaptic proteins
Log2 intensity
Dlgap1
Gria1
Grin2a
Shank1
PSD.turboID
cytosolic.turboID
no.turboID
E
Contaminant selection
selected contaminants
contaminants below threshold
Log2 intensity
Pcca
Mccc1
Pccb
Acaca
Dlg4
Dlgap1
Grin2a
Gria1
Shank1
Dlg1
Protein number
1
2500
F
proteins detected in all conditions: 4599
sample-level normalization
remove contaminants present in no.turboID
keep only proteins identified 3 times in at least one group
proteins detected (filtered): 3531
t-test PSD.turboID vs cytosolic.turboID
independently for each Cre line
FDR < 0.1 s0 > 01
synaptic proteins: 492
t test PSD.turboID Tlx3-Cre vs Sim1-Cre
FDR < 0.05 s0 > 0.1
L5 neuron DE synaptic proteins: 135
- PANTHER: functional classification
- Sorokina, Mclean, Croning et al. 2021 synapse proteome database
- SYNGO
- UniProt
selected DE candidates for follow up validation and analysis

Figure S2 related to Figure 2. Quality control and analysis for turboID synaptic proteome workflow.

A, Western blots for biotinylated proteins and turboID-HA on replicates confirm successful biotinylation in every sample and correct turboID construct expression. **B,** Sample-level intensity values for each replicate before and after median normalization. Protein groups detected in each replicate are reported on the top graph. **C,** Intensity levels in each sample for 4 endogenously biotinylated proteins. **D,** Intensity levels in each sample for 4 known PSD proteins. **E,** Proteins detected in all 4 no.turboID replicates are ranked by intensity. Endogenously biotinylated proteins and known PSD proteins are labelled. Threshold for selecting contaminants is based on the expression of endogenously biotinylated proteins and known PSD components. **F,** Proteomic analysis workflow to enrich for synaptic proteins and select candidates from differentially expressed proteins.

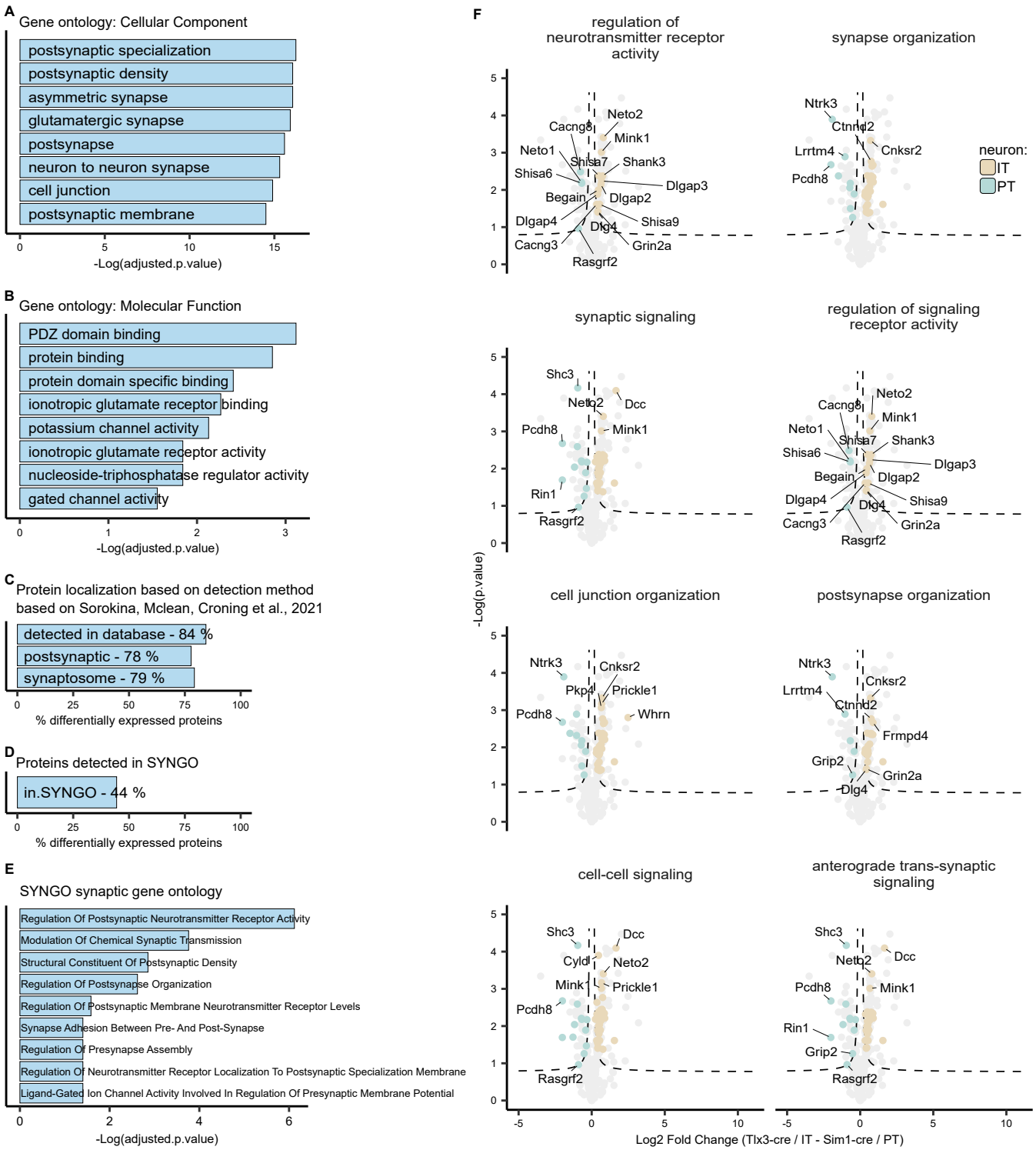
A
Gene ontology: Cellular Component
postsynaptic specialization
postsynaptic density
asymmetric synapse
glutamatergic synapse
postsynapse
neuron to neuron synapse
cell junction
postsynaptic membrane
-Log(adjusted.p.value)
B
Gene ontology: Molecular Function
PDZ domain binding
protein binding
protein domain specific binding
ionotropic glutamate receptor binding
potassium channel activity
ionotropic glutamate receptor activity
nucleoside-triphosphatase regulator activity
gated channel activity
-Log(adjusted.p.value)
C
Protein localization based on detection method
based on Sorokina, Mclean, Croning et al., 2021
detected in database - 84 %
postsynaptic - 78 %
synaptosome - 79 %
% differentially expressed proteins
D
Proteins detected in SYNGO
in.SYNGO - 44 %
% differentially expressed proteins
E
SYNGO synaptic gene ontology
Regulation Of Postsynaptic Neurotransmitter Receptor Activity
Modulation Of Chemical Synaptic Transmission
Structural Constituent Of Postsynaptic Density
Regulation Of Postsynapse Organization
Regulation Of Postsynaptic Membrane Neurotransmitter Receptor Levels
Synapse Adhesion Between Pre- And Post-Synapse
Regulation Of Presynapse Assembly
Regulation Of Neurotransmitter Receptor Localization To Postsynaptic Specialization Membrane
Ligand-Gated Ion Channel Activity Involved In Regulation Of Presynaptic Membrane Potential
-Log(adjusted.p.value)
F
regulation of neurotransmitter receptor activity
synapse organization
synaptic signaling
regulation of signaling receptor activity
cell junction organization
postsynapse organization
cell-cell signaling
anterograde trans-synaptic signaling
-Log(p.value)
neuron:
IT
PT
Log2 Fold Change (Tlx3-cre / IT - Sim1-cre / PT)
G
Grik3
Grin2b
Kcna1
Kcna6
Lrrtm4
Lrrc4c
Flrt3
Flrt2
Dots / cell
IT
PT

Figure S3 related to Figure 3. Gene ontology analysis of differentially expressed proteins in layer 5 neurons.

A, Gene ontology for Cellular Compartment and **B,** Molecular Function categories using differentially expressed proteins as input. **C,** Analysis of differentially expressed proteins using the synapse proteome database[63]. 85% were detected in the database. 79% and 82% were identified in postsynaptic fractionation studies and synaptosome fractionation studies, respectively. **D,** 46% of differentially expressed proteins were detected in the SYNGO database[46]. **E,** SYNGO gene ontologies. **F,** Annotation of proteins found in the Biological Process gene ontology terms (Figure 3b). Not all protein names are displayed due to space limitations, see Table 1 for complete annotation. **G,** Plots from Figure 3f and 3h showing distribution of the analyzed cells. Each color represents a different mouse, each dot represents a single cell.