nature portfolio

Peer Review File

Synaptic Signatures and Disease Vulnerabilities of Layer 5 Pyramidal Neurons

Corresponding Author: Dr Joris De Wit

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Understanding the molecular diversity of synapses is a critical goal of synapse biology. Achieving this goal is more than a methodological exercise to define synapse types. It can provide insights into how different brain disorders impact distinct synapse types. This work by Marcassa and colleagues marks an important advance in achieving these goals. The authors optimized a Cre-dependent TurboID labeling approach using reporters based on PSD-95 and Homer 1. They utilized it to differentially mark postsynaptic proteomes of layer 5 IT and PT neurons. The differential expression of cell-surface proteins they find is intriguing and points to the concept that they modulate or even instruct synapse diversity. A control comparison of gene expression profiles of representative postsynaptic proteins differentially enriched in these neuron types overall matches the postsynaptic proteome data well, validating the approach. The mouse proteomic data also match human expression analyses of layer 5 IT and PT neurons. Last, and very interestingly, the postsynaptic proteomes of layer 5 IT and PT neurons are differentially enriched in proteins encoded by genes linked to increased risk for ASD and schizophrenia.

This study stands out first, for the careful optimization and validation of the approaches it applies to map synaptic proteomes and second, for providing a blueprint that guides investigators in applying TurboID to identify postsynaptic proteomes in neuron types that can be targeted using Cre mouse lines. The results are presented clearly, and the Discussion is well-written. Several points should be addressed, though, to strengthen this manuscript further.

1. How high is the variability between individual mice? It is not clear whether Figure 3e or Extended Data Figure 2 address this question. This is important because some experiments used only 2 mice per condition (line 221).

2. After virus titer titration to avoid spine density effects, how high was the overexpression level of the reporters compared to endogenous PSD-95 and Homer 1 proteins?

3. The idea that the PSD-95 and Homer reporter mark distinct sub-postsynaptic proteomes is not entirely convincing. The fact that their binding partners—which by definition are the closest proteins—are differently labeled is a good control for the restricted local labeling but should not be overinterpreted to mean that the reporters can be used to define sub-postsynaptic proteomes per se.

4. Can the authors discuss how trans-membrane proteins, which often have only short cytosolic tails, are efficiently biotinylated using the PSD-95 reporter?

5. Have the authors combined the PSD-95 TurboID reporter expression with synaptic fractionation to assess whether the selectivity of their current results for postsynaptic targets can be further improved? Or is the yield too low, with too many biotinylated proteins lost during a fractionation?

6. To what extent were presynaptic proteins identified? This is a quality control for the approach. Extended Data Figure 4 appears to show that this was rare, and this could be stated.

7. The authors could speculate whether differences in postsynaptic proteomes of layer 5 IT and PT neurons could instruct input-specific properties of presynaptic sites.

8. The approaches to analyzing the proteomic data are only briefly outlined. If the authors want their study to support the broader use of TurboID, a highly detailed methods section describing the filters and statistical methods will be necessary.

Thomas Biederer

Reviewer #2

(Remarks to the Author)
Marcassa et al.

In this study, Marcassa et al. optimize and use a proximity biotinylation approach to profile the protein composition of excitatory synapses in two genetically identified subtypes of cortical layer 5 (L5) pyramidal neurons, intratelencephalic (IT) and pyramidal tract (PT) ones. They test two different TurboID fusion proteins, one with the postsynaptic protein Homer1, the other with PSD95, for synaptic profiling and identify different subsynaptic complexes. Comparative analysis using PSD95-TurboID allows the authors to identify synaptic signatures specific to L5 IT versus L5 PT neurons. The specificity of enrichment is confirmed using smFISH analysis in the mouse brain and single cell transcriptomics data for human neurons. Using several databases, the authors show that the specific synaptic signature of L5 IT neurons correlates with their potential higher susceptibility in neurodevelopmental disorders such as autism spectrum disorders.
This is an excellent study, carefully designed, executed and interpreted. The limitations are acknowledged by the authors in the discussion. The combination of experimental approaches with database mining demonstrates the power of their comparative analysis and the relevance of this type of studies for our understanding of brain disorders. The strategy will likely be very broadly used in the future by many laboratories. I have only minor concerns.

1. Page 3, line 91: the authors state that they “demonstrate the high spatial resolution of their approach”. I think this statement is too strong, unless they provide experimental evidence that the proteins identified using the two different postsynaptic scaffolds are really not co-localized spatially. Similarly, page 6 line 205: talking about “subsynaptic resolution” seems a bit strong.
2. In figure 1, two different mouse lines are used, Rbp4-Cre and Tlx3-Cre. The rationale for using these two different lines is not explained in the text.
3. In figure 1d, the authors use anti-Homer1 co-immunolabeling with PSD95-TurboID biotinylated proteins and vice versa. Is that a way for the authors to demonstrate synaptic localization of the TurboID without relying on the tagged protein? What is the percentage of colocalization of the TurboID tagged proteins with the endogenous one? Is the localization exclusively synaptic? It seems so from the images in figure 1d but a quantification would strengthen the validity of their approach, given that the protocol then consists in affinity-purification from total protein lysates and not synaptic fractions.
4. In figure 1f, several bands at different sizes are visible for PSD95-TurboID. Can the authors comment on this?
5. In figure 1g, what about the enrichment of cytosolic proteins and inhibitory synapse proteins in these pull-downs?
6. In Figure 2f, it would be interesting to show the GO terms for the cytosolic.turboID specific proteins (like in figure 3) for comparison.
7. In Ext. Figure 2d, it is not completely clear to me if they define an intensity threshold that is then used to exclude certain proteins or whether they simply exclude the selected contaminants identified in this particular analysis ? In other words, is there a fold enrichment compared to the “no TurboID” that is used to exclude “background proteins” in the protein lists before they are used for comparative analysis? Or is the workflow the one presented in ext. figure 3f?
8. I am not sure I really understand the last paragraph of the results section, page 10. Does it mean that only a fourth of the synaptic proteins found to be enriched in mouse IT neurons are enriched in human IT neurons ? Is it possible to perform the same analysis with mouse single cell transcriptomics data (if available) to understand whether this result is due to human specific regulations? This would also nicely complement the results obtained by smFISH for differential expression between the two L5 neuron subtypes.
9. Page 3, line 100: I am not sure I understand this sentence, something might be missing.
10. Page 6, line 185: “Fig 2d, e” repeated twice
11. Page 7, line 234: it is Fig. 3g not h
12. Figure 4c: there seem to be a mistake in the color coding of the legend
13. In the material and methods section, the authors state they use 1-5 weeks old mice. Is the range right? If so, the authors should state in the manuscript when they use 1 week old mice?

Reviewer #3

(Remarks to the Author)
Marcassa et al. conducted detailed analysis of postsynaptic molecular constituents in Layer 5 intratelencephalic (IT) and pyramidal tract (PT) neurons of the somatosensory cortex. Using a cell type-specific proximity biotinylation-based method with TurboID, they mapped the protein composition of excitatory postsynaptic densities, enabling precise in vivo examination. The authors employed various bioinformatics tools and databases for protein identification, quantification, and functional annotation, integrating data from multiple sources to enhance their findings' robustness. Although IT and PT

neurons have been extensively studied for their electrophysiological properties and roles in behaviour and disease, their synaptic protein expression is less explored. Thus, their results will likely generate interest in both the methodology and the specific IT/PT findings and their disease implications. This study's clinical significance is highlighted by its focus on the varying susceptibility of L5 IT and PT neurons to neurodevelopmental disorders such as autism. The findings are clearly presented and thoroughly supported by figures and supplementary materials.

The methodological approach is innovative. A potential concern is whether ectopic expression of PSD95 and Homer could have unintended affects. However, it appears that the authors have exercised due diligence by titrating levels of the AAVs to ensure that the expression does not affect spine morphology. In the absence of in-frame genetic labeling, which has its own caveats, the authors have taken care to add to the rigor of the findings by assessing two different postsynaptic scaffold molecules. This considerably adds to the confidence in the findings. A minor suggestion is to assess some basic synaptic transmission properties to demonstrate that the ectopic expression does not affect function. There are some minor concerns about the specificity of the Cre mouse lines used, which the authors have taken good care to address, and so do not detract from my enthusiasm for the study.

The authors have made some important observations. First, their proteomics results agree with disease databases, and form the framework for future studies using similar methods. Second, this line of work will be important in assigning proteins to specific synapses. For example, the differential expression of LRR-containing synapse organizing proteins adds to the knowledge on synaptic specificity. Third, the study opens a new approach in assessing the local molecular environment of specific proteins. Overall, the study was technically challenging, addresses an important fundamental question, and conceptually advances the field.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors carefully and convincingly addressed all points this reviewer raised, including through extensive new experiments and added analyses. This further improved their excellent study.

Reviewer #2

(Remarks to the Author)
The authors have added new data and new analysis, modified the text and answered all my comments. Their manuscript is further strengthened and is suitable for publication.

Minor remark: I think two proteins might be missing in extended Figure 5c for mouse PT neurons (only 20 are represented, not 22 as stated in the figure).

Reviewer #3

(Remarks to the Author)
The authors have addressed all my concerns satisfactorily, and I have no further concerns.

Tabrez J. Siddiqui

Response to Reviewers "Synaptic signatures and disease vulnerabilities of layer 5 pyramidal neurons", Marcassa et al., manuscript NCOMMS-24-24182-T

We thank all reviewers and the editor for their time and effort and for their supportive feedback on our manuscript. We were very pleased with the reviewers' insightful and positive comments on our work. To address the reviewer comments, we have added experiments and performed additional data analysis, which have considerably strengthened the conclusions of our study. Below, we have outlined our point-by-point responses (in black) to the reviewers' comments (in blue).

Reviewer #1 (Remarks to the Author): Understanding the molecular diversity of synapses is a critical goal of synapse biology. Achieving this goal is more than a methodological exercise to define synapse types. It can provide insights into how different brain disorders impact distinct synapse types. This work by Marcassa and colleagues marks an important advance in achieving these goals. The authors optimized a Cre-dependent TurboID labeling approach using reporters based on PSD-95 and Homer 1. They utilized it to differentially mark postsynaptic proteomes of layer 5 IT and PT neurons. The differential expression of cell-surface proteins they find is intriguing and points to the concept that they modulate or even instruct synapse diversity. A control comparison of gene expression profiles of representative postsynaptic proteins differentially enriched in these neuron types overall matches the postsynaptic proteome data well, validating the approach. The mouse proteomic data also match human expression analyses of layer 5 IT and PT neurons. Last, and very interestingly, the postsynaptic proteomes of layer 5 IT and PT neurons are differentially enriched in proteins encoded by genes linked to increased risk for ASD and schizophrenia.

This study stands out first, for the careful optimization and validation of the approaches it applies to map synaptic proteomes and second, for providing a blueprint that guides investigators in applying TurboID to identify postsynaptic proteomes in neuron types that can be targeted using Cre mouse lines. The results are presented clearly, and the Discussion is well-written. Several points should be addressed, though, to strengthen this manuscript further.

1. How high is the variability between individual mice? It is not clear whether Figure 3e or Extended Data Figure 2 address this question. This is important because some experiments used only 2 mice per condition (line 221).

Response: We apologize if the initial description was not clear and have now clarified the text to better explain numbers of mice and replicates used per condition. In the first set of proteomic experiments (Figure 2), we pooled tissue from 4 mice to generate each replicate. We then collected 3 replicates for PSD95.turboID, 3 replicates for Homer1.turboID, and 2 replicates for cytosolic.turboID.

Based on our findings that 4 mice per replicate generated more than enough material for proteomic identification, we decided to pool tissue from 2 mice per replicate in subsequent proteomic experiments (Figures 3 and 4). We then collected 3 replicates for PSD95.turboID and 3 replicates for cytosolic.turboID per mouse line (Tlx3-cre and Sim1-cre), and 2 replicates for the no.virus condition per mouse line (Tlx3-cre and Sim1-cre).

Since each replicate in these experiments contains tissue from at least 2 mice, individual mouse variability is averaged per replicate. The variability between replicates per condition is limited, as illustrated in the PCA analysis in Figures 2c and 3e.

2. After virus titer titration to avoid spine density effects, how high was the overexpression level of the reporters compared to endogenous PSD-95 and Homer 1 proteins?

Response: This is an important consideration, and we thank the reviewer for raising this point. We initially used an empirical approach to avoid overexpression artefacts, diluting AAVs to avoid morphological effects on synapses but retain sufficient synaptic protein biotinylation. Direct assessment of the levels of overexpression using e.g. semi-quantitative western blot (WB) is not trivial because AAV-mediated expression will only target a subset of cells, and thus the actual levels will be underestimated. Nevertheless, to estimate this, we took advantage of the size difference between the endogenous protein and our virally expressed fusion protein (95 kDa vs. ~130 kDa for PSD95 and ~50kDa vs. ~90 kDa for Homer1). We injected mice with bait-TurboID fusion protein AAVs following the same workflow as in the other experiments in our study. After 1 week of expression and 3 hours of biotin injection, we micro-dissected tissue from layer 5 as closely as possible to the transduced region (visualized by fluorescence from co-injected AAV encoding a Cre-dependent GFP reporter) to maximize the fraction of cells expressing the bait-TurboID fusion. After extracting proteins, we compared band intensities in WB after blotting with an anti-PSD95 antibody (NeuroMab K28/43) or anti-Homer1 antibody (SySy #160003). **Figure R1** shows that the amount of overexpressed proteins is considerably lower than that of the endogenous protein. Even with the aforementioned caveats, this indicates that overexpression levels are comparatively low. We have added this data to Extended Data Figure 1g.

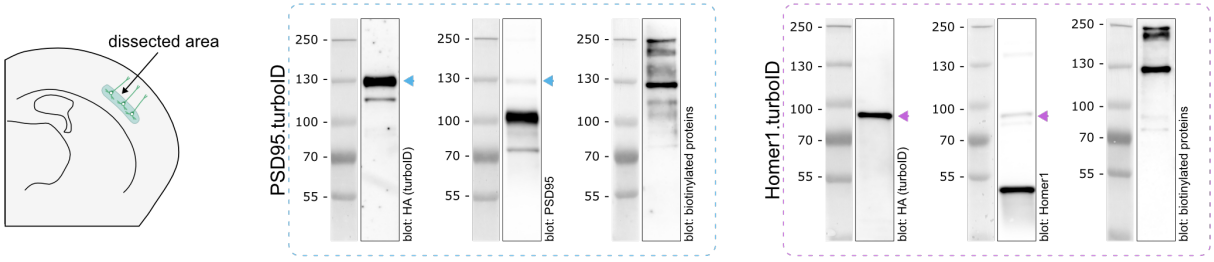

Figure R1. Overexpression levels of bait-TurboID fusion proteins. After AAV injection in the somatosensory cortex of Rbp4-Cre mice, we dissected transduced tissue as closely as possible to the GFP signal of a co-injected Cre-dependent GFP reporter (left panel). After extracting proteins and checking for comparable biotinylation levels as in other experiments (biotinylated protein blots), we compared the intensity of endogenous proteins and bait-TurboID fusion proteins, while checking the correct size of the latter with an anti-HA blot. Both in the case of PSD95- and Homer1-TurboID, the intensity of the overexpressed bait-TurboID fusion protein bands (colored arrows) is qualitatively lower than the intensity of the endogenous band.

3. The idea that the PSD-95 and Homer reporter mark distinct sub-postsynaptic proteomes is not entirely convincing. The fact that their binding partners—which by definition are the closest proteins—are differently labeled is a good control for the restricted local labeling but should not be overinterpreted to mean that the reporters can be used to define sub-postsynaptic proteomes per se.

Response: We agree and have modified the text accordingly. We now refer to differential spatial enrichment instead of sub-synaptic resolution for the two reporters.

4. Can the authors discuss how trans-membrane proteins, which often have only short cytosolic tails, are efficiently biotinylated using the PSD-95 reporter?

Response: TurboID biotinylates lysine residues of the proteins in its vicinity. A short cytosolic tail should be sufficient for biotinylation provided it contains an accessible lysine residue. We analyzed the amino acid composition of the cytosolic regions for the transmembrane proteins in Figures 4c, as annotated on Uniprot. We find that all these proteins have lysine residues in their cytoplasmic regions, ranging from 4 to 44 lysines, with an average of ~15 lysines/cytoplasmic region. Given the surfactant-rich composition of our pulldown buffer, we expect to be able to pulldown these proteins efficiently.

5. Have the authors combined the PSD-95 TurboID reporter expression with synaptic fractionation to assess whether the selectivity of their current results for postsynaptic targets can be further improved? Or is the yield too low, with too many biotinylated proteins lost during a fractionation?

Response: We have opted to avoid fractionation for two main reasons: First, our approach maximizes the yield of purified proteins from small populations of neurons, which would indeed be challenging with a fractionation approach due to loss of material. Second, by avoiding lengthy fractionation procedures, we acutely isolate biotinylated proteins from tissue, staying as close as possible to the situation in the intact circuit. In our approach, the synaptic selectivity of the identified proteins is obtained first by targeting TurboID to the postsynapse, and second by stringent filtering of the identified proteins during proteomic data analysis using the no.virus and cytosolic.turboID controls.

6. To what extent were presynaptic proteins identified? This is a quality control for the approach. Extended Data Figure 4 appears to show that this was rare, and this could be stated.

Response: Indeed, the gene ontology analysis in Extended Data Figure 4 indicates that identified proteins are postsynaptic. To better illustrate this, we have included a SynGO sunburst gene enrichment plot of the annotated SYNGO proteins for PSD95.turboID and Homer1.turboID (**Fig. R2a** and new Extended Data Figure 2h). Many proteins display GO:CC terms related to the postsynapse, resulting in higher Q-values for those terms. Additionally, none of the abundant synaptic vesicle and active zone proteins (i.e. proteins with molecules/bouton > 1000[1]) are present in our postsynaptic dataset, including "core" presynaptic proteins such as Bassoon, Munc13, SNAP25, VAMP2, VGluT1/2, Synapsin1 and Synaptophysin (Supplementary Table 1). Finally, we also analyzed the presence of the presynaptic marker Synaptophysin in our biotinylated protein pulldown samples by WB. We do not detect Synaptophysin in any condition (**Fig. R2b**). We have added this data to Figure 1g. Altogether, we can confidently conclude that the approach we use, together with the appropriate controls, specifically identifies postsynaptic proteins. We have modified the text to clarify this point.

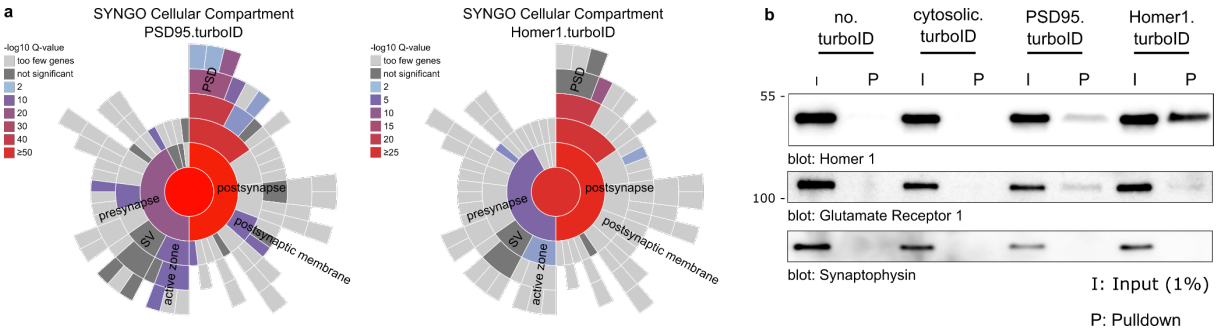

Figure R2. Postsynaptic specificity of PSD95- and Homer1-TurboID labeled proteins. a. SynGO Sunburst gene enrichment plot using the genes from the synaptic dataset in Figure 2d and 2e. **b.** Streptavidin pulldown from Tlx3-Cre mice injected with different turboID construct-expressing AAVs. Note the specific presence of post-synaptic proteins Homer1 and Gria1 in the PSD95-TurboID and Homer1-TurboID conditions, and the absence of Synaptophysin in any pulldown condition.

7. The authors could speculate whether differences in postsynaptic proteomes of layer 5 IT and PT neurons could instruct input-specific properties of presynaptic sites.

Response: IT and PT L5 pyramidal neurons share a large part of pre-synaptic input, such as input from primary motor cortex[2], so it is indeed tempting to speculate about an organizing role of the post-synaptic identity on the function of presynaptic inputs. However, it is currently unclear whether shared

inputs to L5 IT and PT neurons come from a homogenous population able to contact both postsynaptic partners, or from two distinct subtypes that would already display cell-intrinsic presynaptic differences. Conversely, it is also possible that slightly different presynaptic inputs shape the differences we see in postsynaptic proteomes, through e.g. transcriptional regulation along a synapse to nucleus axis. Retrograde tracing combined with precise genetic identification of presynaptic inputs would be a prerequisite for studying this question.

8. The approaches to analyzing the proteomic data are only briefly outlined. If the authors want their study to support the broader use of TurboID, a highly detailed methods section describing the filters and statistical methods will be necessary.

Response: This is indeed an important point. We have now extended the methods section, describing in detail the analysis steps taken and databases used. Similarly to Extended Data Fig. 3f, we now added an analysis workflow schematic in Extended Data Fig. 2i listing all the steps taken during the analysis of the Homer1.turboID and PSD95.turboID experiment.

Thomas Biederer

Reviewer #2 (Remarks to the Author): Marcassa et al. In this study, Marcassa et al. optimize and use a proximity biotinylation approach to profile the protein composition of excitatory synapses in two genetically identified subtypes of cortical layer 5 (L5) pyramidal neurons, intratelencephalic (IT) and pyramidal tract (PT) ones. They test two different TurboID fusion proteins, one with the postsynaptic protein Homer1, the other with PSD95, for synaptic profiling and identify different subsynaptic complexes. Comparative analysis using PSD95-TurboID allows the authors to identify synaptic signatures specific to L5 IT versus L5 PT neurons. The specificity of enrichment is confirmed using smFISH analysis in the mouse brain and single cell transcriptomics data for human neurons. Using several databases, the authors show that the specific synaptic signature of L5 IT neurons correlates with their potential higher susceptibility in neurodevelopmental disorders such as autism spectrum disorders.

This is an excellent study, carefully designed, executed and interpreted. The limitations are acknowledged by the authors in the discussion. The combination of experimental approaches with database mining demonstrates the power of their comparative analysis and the relevance of this type of studies for our understanding of brain disorders. The strategy will likely be very broadly used in the future by many laboratories. I have only minor concerns.

1. Page 3, line 91: the authors state that they “demonstrate the high spatial resolution of their approach”. I think this statement is too strong, unless they provide experimental evidence that the proteins identified using the two different postsynaptic scaffolds are really not co-localized spatially. Similarly, page 6 line 205: talking about “subsynaptic resolution” seems a bit strong.

Response: We agree and have modified the text accordingly. We now refer to differential spatial enrichment instead of sub-synaptic resolution for the two reporters.

2. In figure 1, two different mouse lines are used, Rbp4-Cre and Tlx3-Cre. The rationale for using these two different lines is not explained in the text.

Response: The use of the two different mouse Cre lines (both of which target layer 5 neurons) was based on our mouse availability at the time. The Rbp4-Cre line was only used for control experiments, such as determining the optimal time course for biotin injection (Figure 1e). We have now included a sentence in the main text (page 4) to clarify this.

3. In figure 1d, the authors use anti-Homer1 co-immunolabeling with PSD95-TurboID biotinylated proteins and vice versa. Is that a way for the authors to demonstrate synaptic localization of the TurboID without relying on the tagged protein? What is the percentage of colocalization of the TurboID tagged proteins with the endogenous one? Is the localization exclusively synaptic? It seems so from the images in figure 1d but a quantification would strengthen the validity of their approach, given that the protocol then consists in affinity-purification from total protein lysates and not synaptic fractions.

Response: Indeed, we have used anti-Homer1 stainings for the PSD95-TurboID condition (and vice versa) to show synaptic localization of biotinylated proteins because antibodies directed against the bait would otherwise recognize the virally expressed protein. To better quantify the localization of bait-TurboID proteins, we have performed IHC on brain sections from Tlx3-Cre mice injected with PSD95-TurboID and Homer1-TurboID AAVs (**Fig. R3a**). We used streptavidin staining as a proxy for the localization of bait-TurboID proteins,since viral expression levels are low (**Fig. R1**) and anti-HA stainings show only weak signal in our hands. We quantified the percentage of streptavidin-positive puncta that also co-localize with an endogenous postsynaptic marker and found that most of the streptavidin-positive objects are indeed postsynaptic (**Fig. R3b**), highlighting the high spatial specificity of TurboID-mediated proximity labeling. The colocalization of streptavidin-positive puncta with postsynaptic markers we observe is likely a slight underestimation, as we used single-plane confocal images for these quantifications. We have added this data to Extended Figure 1a-b and amended the main text accordingly (page 4).

Because of the high spatial specificity highlighted here, we opted to not use biochemical fractionation (such as synaptosome preps) because of the unavoidable loss of material in these procedures, thus keeping the number of animals required per replicate low (see also response to Reviewer #1 Point 5).

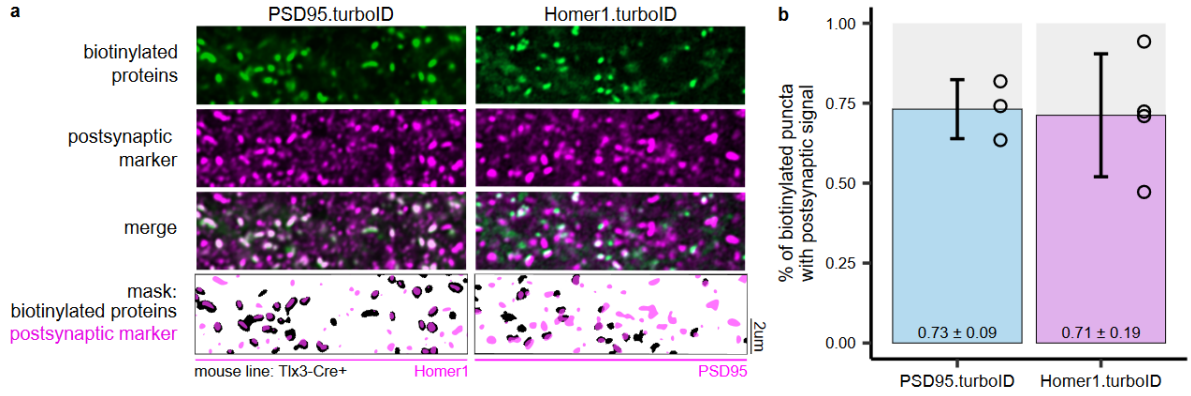

Figure R3. Co-localization of biotinylated proteins with postsynaptic markers. a, Example images of biotinylated proteins, postsynaptic markers and the masks used for the quantification. **b,** Quantification of biotinylated puncta that also have postsynaptic marker signal.

4. In figure 1f, several bands at different sizes are visible for PSD95-TurboID. Can the authors comment on this?

Response: PSD95 undergoes extensive post-translational modifications, including phosphorylation and palmitoylation. Several publications reported a similar 3-band pattern for endogenous PSD95 as

we found for our tagged PSD95 construct[3–6]. Antibodies producing a single band by western blot (WB) are likely targeting a specific post-translational modification on the protein.

To compare the pattern of PSD95.turboID to that of endogenous PSD95, we performed WBs using a KO-validated antibody against endogenous PSD95 (NeuroMab K28/43) on mice injected with PSD95.turboID, taking advantage of the size difference between the endogenous protein and our virally expressed fusion protein (95 kDa vs. ~130 kDa for PSD95). We injected mice with AAV-PSD95.turboID following the same workflow as in the other experiments in our study. After 1 week of expression and 3 hours of biotin injection, we micro-dissected tissue from layer 5 as closely as possible to the transduced region (visualized by fluorescence from co-injected AAV encoding a Cre-dependent GFP reporter) to maximize the fraction of cells expressing PSD95.turboID. After extracting proteins, we compared band patterns and intensities in WB after blotting with an anti-PSD95 antibody (NeuroMab K28/43). **Figure R4** (same as Figure R1, but copied here for convenience) shows that endogenous and virally expressed PSD95 show similar band patterns, and that the amount of overexpressed proteins is considerably lower than that of the endogenous protein. We have added this data to Extended Data Figure 1g. Please also see response to Reviewer #1 Point 2.

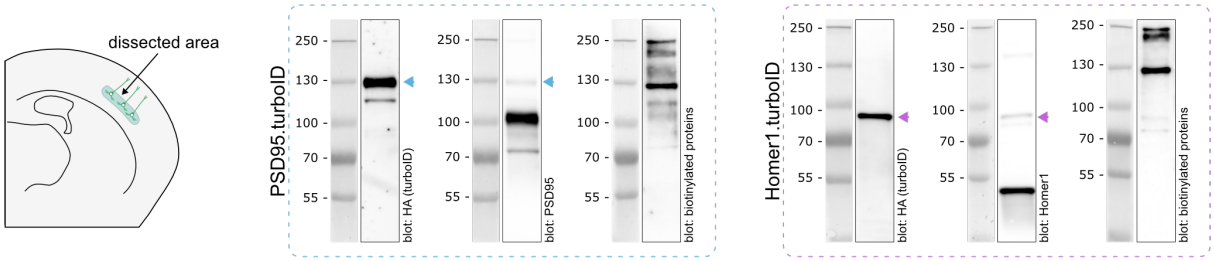

Figure R4. Overexpression levels of bait-TurboID fusion proteins. After AAV injection in the somatosensory cortex of Rbp4-Cre mice, we dissected transduced tissue as closely as possible to the GFP signal of a co-injected Cre-GFP reporter (left panel). After extracting proteins and checking for comparable biotinylation levels as in other experiments (biotinylated protein blots), we compared the intensity of endogenous proteins and bait-TurboID fusion proteins, while checking the correct size of the latter with an anti-HA blot. Both in the case of PSD95- and Homer1-TurboID, the intensity of the overexpressed bait-TurboID fusion protein bands (colored arrows) is qualitatively lower than the intensity of the endogenous band.

5. In figure 1g, what about the enrichment of cytosolic proteins and inhibitory synapse proteins in these pull-downs?

Response: We thank the reviewer for pointing this out. We have now analyzed presence of the presynaptic marker Synaptophysin in our biotinylated protein pulldown samples by WB. We do not detect Synaptophysin in any pulldown condition (**Fig. R5a,** same as Fig. R2b but copied here for convenience). We have added this data to Figure 1g.

In addition, we have analyzed the proteomics data we generated for the presence of inhibitory synapse proteins: some examples are highlighted in **Figure R5b**. Inhibitory synapse proteins (Slc32a1/Vesicular inhibitory amino acid transporter and gephyrin) show no enrichment in any condition, while cytoskeletal proteins (Map1b and Nefm) show stronger enrichment in the cytosolic.turboID condition compared to the synaptic conditions. We have added this data as a new plot in Extended Data Fig. 2d.

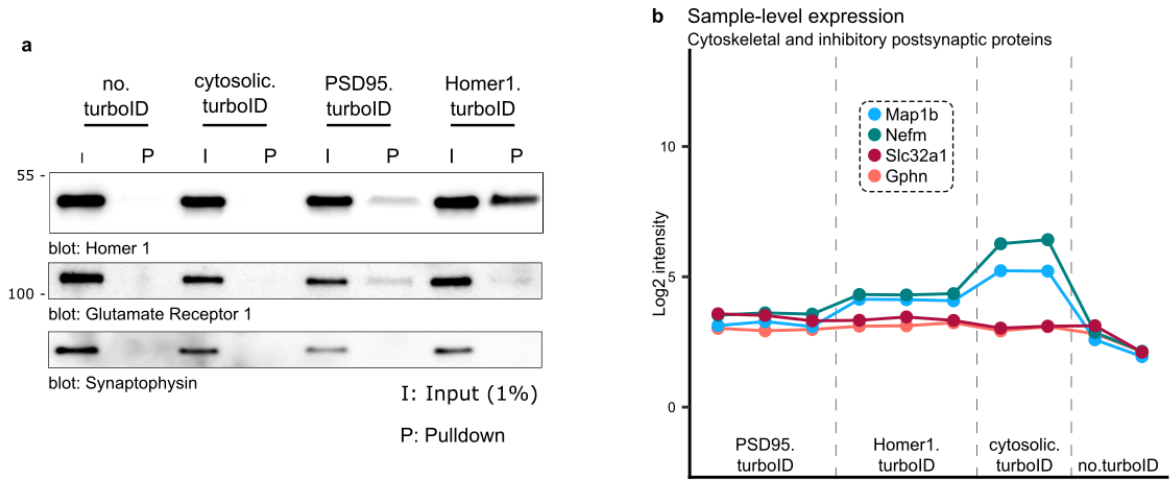

Figure R5. Excitatory postsynaptic specificity of PSD95- and Homer1-TurboID. a) Pre- and postsynaptic marker protein identification by western blot in different turboID conditions after streptavidin pulldown. The presynaptic marker Synaptophysin is not identified in any condition. b) Sample level expression of cytosolic and inhibitory synaptic proteins in the experiment described in Fig. 2. Note that no condition displays enrichment of the inhibitory presynaptic protein Slc32a1 (VGAT) or the inhibitory postsynaptic protein Gephyrin (Gphn). Synaptic proteomes labelled with either PSD95-TurboID or Homer1-TurboID are relatively depleted in the cytosolic proteins Map1b and Nefm compared to proteomes labelled by cytosolic-TurboID.

6. In Figure 2f, it would be interesting to show the GO terms for the cytosolic.turboID specific proteins (like in figure 3) for comparison.

Response: We have added those plots (**Fig. R6**) to Extended Data Fig. 2f as suggested by the reviewer.

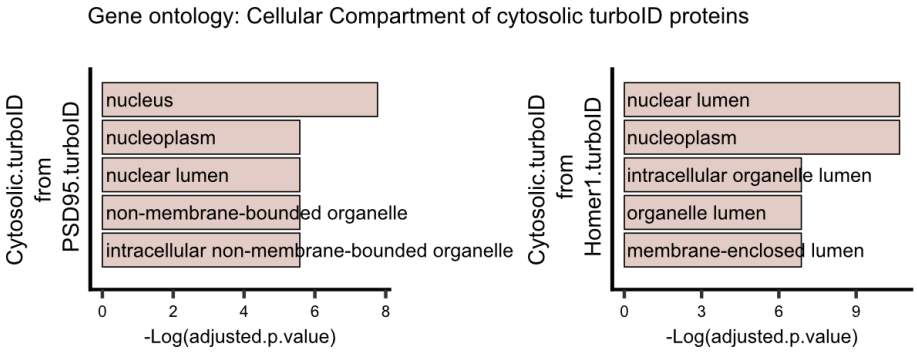

Figure R6. Top Cytosolic-TurboID associated GO terms. Top GO:CC terms in cytosolic.turboID enriched proteomes compared to PSD95.turboID (left) and Homer1.turboID (right). Data (refers to Figure 2f).

7. In Ext. Figure 2d, it is not completely clear to me if they define an intensity threshold that is then used to exclude certain proteins or whether they simply exclude the selected contaminants identified in this particular analysis ? In other words, is there a fold enrichment compared to the "no TurboID" that is used to exclude "background proteins" in the protein lists before they are used for comparative analysis? Or is the workflow the one presented in ext. figure 3f?

Response: We apologize if this was not clear. We followed the 'ratiometric approach', as described in studies from the Ting lab that developed TurboID[7,8]. Based on this approach, we first identify known contaminants and known synaptic proteins (labeled in Extended Data Figs. 2e and 3e) and draw a threshold between those two groups. This threshold is qualitative and varies between different

datasets, depending on the number of identified proteins. While it is only a qualitative threshold, we found that is a helpful step to remove high-abundance contaminants (for example endogenously biotinylated proteins and keratins) as well as non-specific binders that are not washed away during our protocol. We have now expanded the paragraph of data analysis in the Methods section to explain this step in detail.

8. I am not sure I really understand the last paragraph of the results section, page 10. Does it mean that only a fourth of the synaptic proteins found to be enriched in mouse IT neurons are enriched in human IT neurons ? Is it possible to perform the same analysis with mouse single cell transcriptomics data (if available) to understand whether this result is due to human specific regulations? This would also nicely complement the results obtained by smFISH for differential expression between the two L5 neuron subtypes.

Response: We thank the reviewer for this helpful suggestion. We have now performed the suggested analysis with mouse transcriptomics data, using the same publicly available dataset we previously used for human transcriptome data[9]. We find a better match between mouse transcriptome and mouse proteome data. We extracted and re-clustered mouse L5 IT and PT neuron transcriptomes from the Bakken et al. (2021) motor cortex single-nucleus RNA sequencing dataset and calculated differential gene expression between IT and PT neurons (**Fig. R7a**). We then calculated the overlap between differentially expressed synaptic proteins from our proteomic dataset and differentially expressed genes from the single-nucleus RNA sequencing dataset. Of 93 differentially expressed synaptic proteins in mouse IT neurons, we found 36 to be differentially expressed at the RNA level in mouse neurons; 94% (34/36) of which showed matching enrichment in mouse L5 IT neurons (**Fig. R7b, c**). Of 42 differentially expressed synaptic proteins in mouse PT neurons, 22 were differentially expressed at the RNA level in mouse neurons; 91% (20/22) of which showed matching enrichment in mouse L5 PT neurons (**Fig. R7b, c**). The better match between mouse proteome and mouse transcriptome data supports our smFISH experiments in Figure 4e-h. We have added this analysis as Extended Data Figure 5. In addition, we updated the panels in Figure 5f and Extended Data Figure 5b to better illustrate overlap between differentially expressed synaptic proteins and genes, and have better explained the way we compare protein and transcriptome datasets and calculate overlap in the text (page 10, 11).

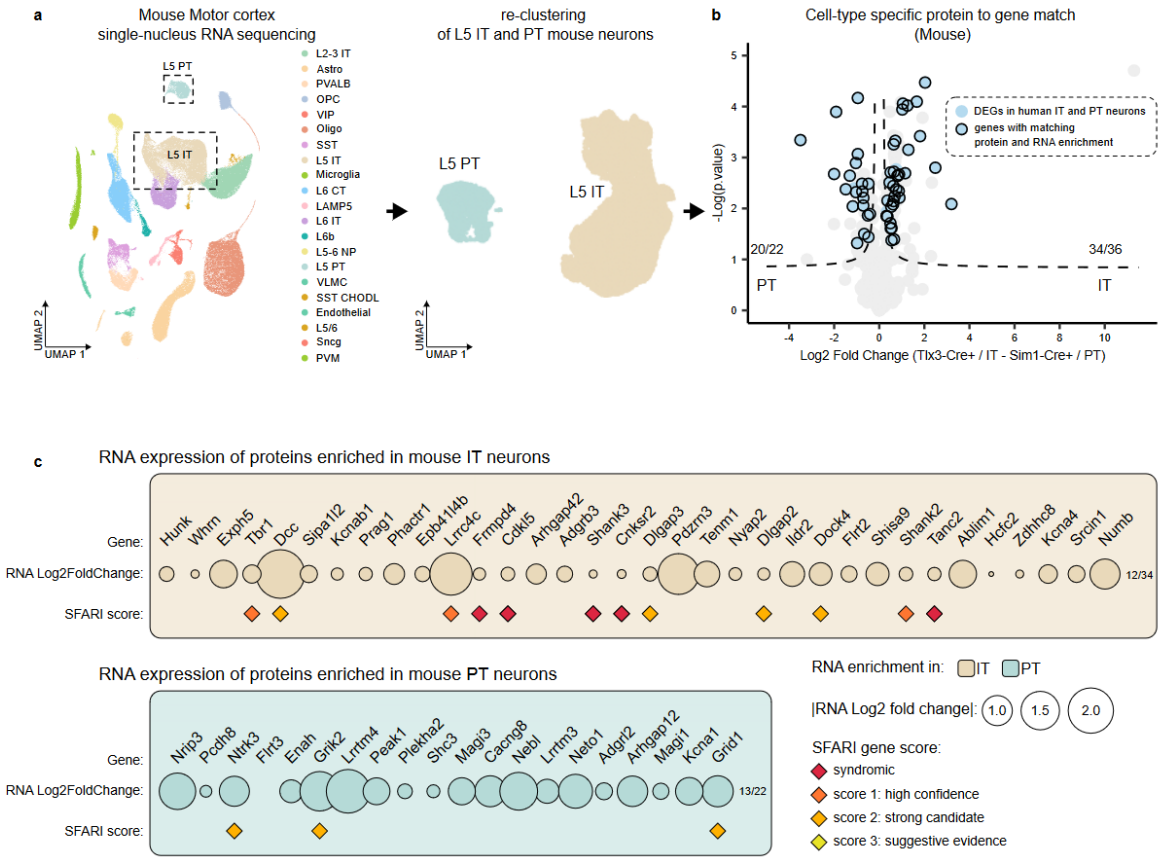

Figure R7. a, L5 IT and PT neurons were extracted from a published mouse motor cortex snRNA-Seq dataset[9] and re-clustered. **b,** Differentially expressed synaptic proteins in mouse were compared to differentially expressed genes in mouse motor cortex. Of the 93 differentially expressed synaptic proteins detected in mouse L5 IT neurons and 42 proteins in mouse PT neurons, 36 and 22 genes respectively showed differential expression also in mouse samples. **c,** Mouse proteins that show matching differential gene expression at the mouse transcript level are shown. SFARI scores are annotated.

9. Page 3, line 100: I am not sure I understand this sentence, something might be missing.

We have clarified this sentence.

10. Page 6, line 185: "Fig 2d, e" repeated twice.

Thank you, we have corrected this.

11. Page 7, line 234: it is Fig. 3g not h

We have corrected this.

12. Figure 4c: there seem to be a mistake in the color coding of the legend

Thank you for noticing; the legend was indeed incorrectly labeled. We have fixed this.

13. In the material and methods section, the authors state they use 1-5 weeks old mice. Is the range right? If so, the authors should state in the manuscript when they use 1 week old mice?

Thank you, we have corrected this. This was a typo: mice used in this study were 4-5 weeks old.

Reviewer #3 (Remarks to the Author): Marcassa et al. conducted detailed analysis of postsynaptic molecular constituents in Layer 5 intratelencephalic (IT) and pyramidal tract (PT) neurons of the somatosensory cortex. Using a cell type-specific proximity biotinylation-based method with TurboID, they mapped the protein composition of excitatory postsynaptic densities, enabling precise in vivo examination. The authors employed various bioinformatics tools and databases for protein identification, quantification, and functional annotation, integrating data from multiple sources to enhance their findings' robustness. Although IT and PT neurons have been extensively studied for their electrophysiological properties and roles in behaviour and disease, their synaptic protein expression is less explored. Thus, their results will likely generate interest in both the methodology and the specific IT/PT findings and their disease implications. This study's clinical significance is highlighted by its focus on the varying susceptibility of L5 IT and PT neurons to neurodevelopmental disorders such as autism. The findings are clearly presented and thoroughly supported by figures and supplementary materials.

The methodological approach is innovative. A potential concern is whether ectopic expression of PSD95 and Homer could have unintended affects. However, it appears that the authors have exercised due diligence by titrating levels of the AAVs to ensure that the expression does not affect spine morphology. In the absence of in-frame genetic labeling, which has its own caveats, the authors have taken care to add to the rigor of the findings by assessing two different postsynaptic scaffold molecules. This considerably adds to the confidence in the findings. A minor suggestion is to assess some basic synaptic transmission properties to demonstrate that the ectopic expression does not affect function. There are some minor concerns about the specificity of the Cre mouse lines used, which the authors have taken good care to address, and so do not detract from my enthusiasm for the study.

The authors have made some important observations. First, their proteomics results agree with disease databases, and form the framework for future studies using similar methods. Second, this line of work will be important in assigning proteins to specific synapses. For example, the differential expression of LRR-containing synapse organizing proteins adds to the knowledge on synaptic specificity. Third, the study opens a new approach in assessing the local molecular environment of specific proteins. Overall, the study was technically challenging, addresses an important fundamental question, and conceptually advances the field.

Response: We thank the reviewer for these positive and constructive remarks on our study. The reviewer rightfully points out that in-frame genetic labeling has its own caveats. To further address the point of ectopic expression of PSD95 and Homer1, we have analyzed the level of overexpression of our bait proteins, as also suggested by Reviewer #1, point 2. As shown in **Figure R1**, we find that levels of overexpressed proteins, estimated by western blot, are considerably lower than those of the respective endogenous protein, for PSD95-TurboID as well as Homer1-TurboID. In combination with the lack of effect on synapse morphology, we are confident that our approach is unlikely to drastically affect synapse function and composition.

References

1. Wilhelm, B. G. *et al.* Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* **344**, 1023–1028 (2014).
2. Petreanu, L., Mao, T., Sternson, S. M. & Svoboda, K. The subcellular organization of neocortical excitatory connections. *Nature* **457**, 1142–1145 (2009).
3. Kim, M. J. *et al.* Synaptic Accumulation of PSD-95 and Synaptic Function Regulated by Phosphorylation of Serine-295 of PSD-95. *Neuron* **56**, 488–502 (2007).
4. Nelson, C. D., Kim, M. J., Hsin, H., Chen, Y. & Sheng, M. Phosphorylation of threonine-19 of PSD-95 by GSK-3β is required for PSD-95 mobilization and long-term depression. *J. Neurosci. Off. J. Soc. Neurosci.* **33**, 12122–12135 (2013).
5. Chowdhury, D. & Hell, J. W. Ca2+/Calmodulin Binding to PSD-95 Downregulates Its Palmitoylation and AMPARs in Long-Term Depression. *Front. Synaptic Neurosci.* **11**, 433801 (2019).
6. Zheng, S. *et al.* PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat. Neurosci.* **15**, 381–388 (2012).
7. Cho, K. F. *et al.* Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* 1–29 (2020) doi:10.1038/s41596-020-0399-0.
8. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
9. Bakken, T. E. *et al.* Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111–119 (2021).

Point-by-point reply "Synaptic signatures and disease vulnerabilities of layer 5 pyramidal neurons", Marcassa et al., final revisions manuscript NCOMMS-24-24182-A

We thank the reviewers for their final input on our revised manuscript. We have revised the manuscript according to their remaining comment. All reviewer comments are copied below (in blue) and addressed in black.

Reviewer #1 (Remarks to the Author):
The authors carefully and convincingly addressed all points this reviewer raised, including through extensive new experiments and added analyses. This further improved their excellent study.
We thank the reviewer for these positive remarks.

Reviewer #2 (Remarks to the Author):
The authors have added new data and new analysis, modified the text and answered all my comments. Their manuscript is further strengthened and is suitable for publication.

Minor remark: I think two proteins might be missing in extended Figure 5c for mouse PT neurons (only 20 are represented, not 22 as stated in the figure).

We thank the reviewer for or these positive remarks and for noticing this typo. The graph represents the genes with matching expression at RNA and protein level in the same mouse neuron type (IT or PT). For PT neurons, these are 20 genes (see also extended Figure 5b). However, on the graph in Extended figure 5c we put the wrong number on the label which represents the number of SFARI genes out of the matching proteins (13/22 should be 3/20). This is now fixed in the latest version.

Reviewer #3 (Remarks to the Author):
The authors have addressed all my concerns satisfactorily, and I have no further concerns.
Tabrez J. Siddiqui
We thank the reviewer for these positive remarks.