



# Complexity and graded regulation of neuronal cell-type-specific alternative splicing revealed by single-cell RNA sequencing

Huijuan Feng<sup>a,b,c</sup>, Daniel F. Moakley<sup>a,b,c</sup>, Shuonan Chen<sup>a,b,c</sup>, Melissa G. McKenzie<sup>a,b,c</sup> , Vilas Menon<sup>d</sup>, and Chaolin Zhang<sup>a,b,c,1</sup> 

<sup>a</sup>Department of Systems Biology, Columbia University, New York, NY 10032; <sup>b</sup>Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032; <sup>c</sup>Center for Motor Neuron Biology and Disease, Columbia University, New York, NY 10032; and <sup>d</sup>Department of Neurology, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY 10032

Edited by Donald C. Rio, University of California, Berkeley, CA, and accepted by Editorial Board Member Michael Rosbash February 3, 2021 (received for review June 23, 2020)

The enormous cellular diversity in the mammalian brain, which is highly prototypical and organized in a hierarchical manner, is dictated by cell-type-specific gene-regulatory programs at the molecular level. Although prevalent in the brain, the contribution of alternative splicing (AS) to the molecular diversity across neuronal cell types is just starting to emerge. Here, we systematically investigated AS regulation across over 100 transcriptomically defined neuronal types of the adult mouse cortex using deep single-cell RNA-sequencing data. We found distinct splicing programs between glutamatergic and GABAergic neurons and between subclasses within each neuronal class. These programs consist of overlapping sets of alternative exons showing differential splicing at multiple hierarchical levels. Using an integrative approach, our analysis suggests that RNA-binding proteins (RBPs) *Celf1/2*, *Mbnl2*, and *Khdrbs3* are preferentially expressed and more active in glutamatergic neurons, while *Elavl2* and *Qk* are preferentially expressed and more active in GABAergic neurons. Importantly, these and additional RBPs also contribute to differential splicing between neuronal subclasses at multiple hierarchical levels, and some RBPs contribute to splicing dynamics that do not conform to the hierarchical structure defined by the transcriptional profiles. Thus, our results suggest graded regulation of AS across neuronal cell types, which may provide a molecular mechanism to specify neuronal identity and function that are orthogonal to established classifications based on transcriptional regulation.

alternative splicing | neuronal cell type diversity | single-cell RNA-sequencing | RNA-binding proteins | graded regulation

The structural and functional complexity of the mammalian cortex is determined by the enormous diversity of neuronal cell types and their connections that form intricate neural circuitries. Traditionally, neuronal cell types are defined by their morphological or electrophysiological properties, connectivity, or a small set of molecular markers (1). Recent advances in single-cell RNA sequencing (scRNA-seq) have enabled the unbiased discovery and characterization of neuronal cell types based on global transcriptional profiles (2–6). These studies demonstrated unprecedented power in revealing over 100 previously known and novel putative cell types. It has been suggested that these cell types can be organized with a hierarchical taxonomy consisting of glutamatergic (excitatory) neurons and GABAergic (inhibitory) interneurons as major neuronal classes. Among them, glutamatergic neurons are composed of five major branches, including L6 corticothalamic (CT), L5/6 near-projecting (NP), L5 pyramidal tract (PT), intratelencephalic (IT), and L6b subclasses, while GABAergic neurons can be divided based on their embryonic origins in the medial and caudal ganglionic eminences (MGE and CGE), which can be further divided into *Pvalb* and *Sst* subclasses for MGE neurons and *Vip*, *Lamp5*, *Sncg*, and *Serpinf1* subclasses for CGE neurons (2).

However, most of these transcriptional cell types remain poorly characterized.

One limitation in previous analysis of transcriptomic neuronal cell types is that those studies almost exclusively focused on the steady-state transcript level, while mammalian genes undergo multiple steps of posttranscriptional regulation that are essential for diversifying the final protein products in time and space. Alternative splicing (AS) of precursor mRNA (pre-mRNA) is a molecular mechanism to produce multiple transcript and protein isoforms through different combinations of exons. Transcriptomic analysis based on RNA-seq suggests that AS is ubiquitous, occurring in >90% of multiexon genes in human and other mammalian species (7, 8). The use of AS in the nervous system is particularly extensive, and previous studies have revealed distinct splicing profiles of the brain compared to nonneuronal organs (7–9), neurons compared to nonneuronal cells in the cortex (10, 11), central nervous system neurons compared to peripheral sensory neurons (12), as well as at different developmental stages in the cortex (12). While the functional significance of the majority of these neuron-specific alternative exons has yet to be demonstrated, there is no lack of examples in which individual alternative exons

## Significance

Alternative splicing (AS) is extensively used in the mammalian brain, but its contribution to the molecular and cellular diversity across neuronal cell types remains poorly understood. Through systematic and integrative analysis of over 100 transcriptomically defined cortical neuronal types, we found neuronal subclass-specific splicing-regulatory programs consist of overlapping alternative exons showing differential splicing at multiple hierarchical levels. Evidence is provided that this graded AS regulation is controlled by unique combinations of RNA-binding proteins (RBPs). Importantly, these RBPs also contribute to splicing dynamics across neuronal cell types that do not conform to the hierarchical taxonomy established based on transcriptional profiles, suggesting that the graded AS regulation may provide a molecular mechanism orthogonal to transcriptional regulation in specifying neuronal identity and function.

Author contributions: H.F. and C.Z. designed research; H.F., D.F.M., S.C., and M.G.M. performed research; H.F., D.F.M., S.C., M.G.M., V.M., and C.Z. analyzed data; and H.F. and C.Z. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. D.C.R. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

<sup>1</sup>To whom correspondence may be addressed. Email: cz2294@columbia.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2013056118/-DCSupplemental>.

Published March 5, 2021.

play critical roles in nervous system development and function, such as neuronal migration, axon outgrowth, synaptic formation, and transmission (13). In addition, emerging evidence suggests global splicing differences between major excitatory and inhibitory neuronal cell types from recent bulk RNA-seq analysis of specific cell types or ribosome-associated transcriptomes purified from transgenic mice (14, 15). Therefore, elucidating cell-type-specific splicing regulation is a key step toward understanding the underpinnings of neuronal cell-type diversity and function.

Cell-type-specific AS is largely controlled by RNA-binding proteins (RBPs) that recognize specific regulatory sequences embedded in pre-mRNA transcripts. Technological advances have also made it possible to define the comprehensive RBP target networks by integrating global splicing profiles upon RBP depletion and genome-wide maps of protein-RNA interactions, as demonstrated by multiple studies, including our own (16–18). For example, RBPs specifically expressed or enriched in neurons, such as Nova, Rbfox, Ptp2, nElavl, Srrm4 (nSR100), and Mbnl2, have been demonstrated to regulate AS of numerous neuronal transcripts (reviewed in refs. 13 and 19). In addition, KH domain-containing proteins Khdrbs2 (Slm1) and Khdrbs3 (Slm2) were recently shown to be preferentially expressed in pyramidal cells over Pvalb interneurons in hippocampus to regulate splicing of a highly specific set of transcripts involved in excitatory synaptic formation (20, 21).

Despite this exciting progress, the extent to which AS contributes to the diversity of the wide range of neuronal cell types and the underlying regulatory mechanisms has not been systematically investigated. In the present study, we took advantage of deep scRNA-seq data of adult mouse neocortex with whole transcript coverage to identify differentially spliced (DS) alternative exons in major neuronal classes and subclasses at different hierarchical levels. By integrating de novo motif analysis, RBP expression profiles, their target networks, and position-dependent RNA-maps, our analysis revealed several major regulators of the neuronal cell-type-specific AS and “graded,” combinatorial regulatory programs that contribute to the complex landscape of splicing dynamics across cell types.

## Results

### Major Neuronal Cell Types Revealed by Single-Cell Splicing Profiles.

To study neuronal cell-type-specific AS on a global scale, we utilized the two sets of scRNA-seq data derived from adult mouse cortex with deep and full-length transcript coverage, generated by the Allen Institute for Brain Science (2, 3) (denoted as the Tasic 2016 and Tasic 2018 datasets, hereafter). The Tasic 2016 dataset includes 1,809 cells from the primary visual cortex (VISp), while the Tasic 2018 dataset is composed of 15,413 cells from VISp and 10,068 cells from anterior lateral motor cortex (ALM). In our analysis, we used the 1,424 and 21,154 core cells that were unambiguously assigned to different transcriptional cell types as defined in the two original studies (totaling 14 billion reads and 54 billion read pairs, respectively), and quantified the inclusion level (percent spliced in or PSI, denoted  $\Psi$ ) of over 16,000 cassette exons using the Quantas pipeline (10). Overall, a comparable number of exons have sufficient exon junction reads (i.e.,  $\geq 20$ ) for splicing quantification in glutamatergic and GABAergic neuronal populations in each dataset, while a lower number of exons are quantifiable in non-neuronal cells (median = 2,328, 2,114, and 813 for glutamatergic neurons, GABAergic neurons, and nonneuronal cells, respectively, in Tasic 2016 dataset, and 2,176, 2,218, and 1,128 in Tasic 2018 dataset) (SI Appendix, Fig. S1 and Dataset S1), which is consistent with lower RNA yield in nonneuronal cells. Given its much larger number of cells, we primarily focused on the Tasic 2018 dataset and integrated results from the Tasic 2016 dataset when appropriate.

We first asked whether single-cell splicing profiles can be used to infer neuronal cell types in the cortex. To this end, we performed *t*-distributed neighbor embedding (*t*-SNE) analysis and clustered cells using single-cell splicing profiles (Materials and Methods). For

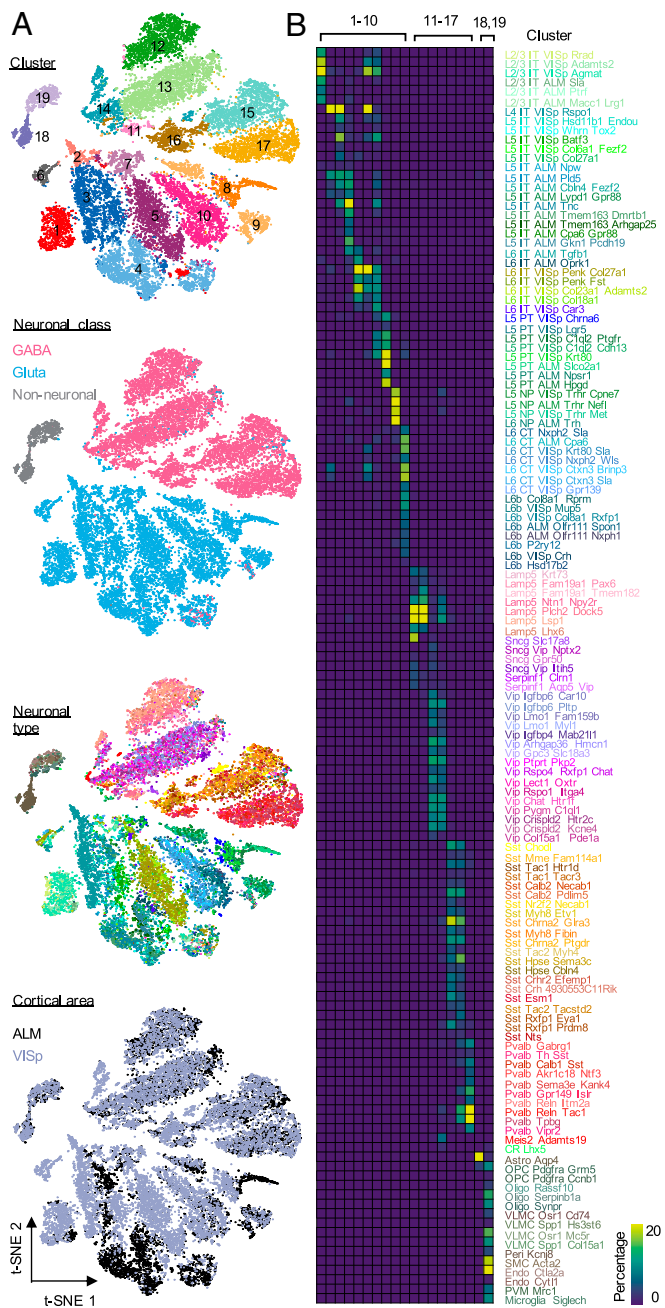
the Tasic 2018 dataset, we identified 19 major clusters (Fig. 1A). When compared to the 133 transcriptional cell types defined originally based on gene-expression profiles, which consist of 56 glutamatergic, 61 GABAergic, and 16 nonneuronal cell types, the clusters defined by splicing profiles clearly separated glutamatergic (clusters 1 to 10), GABAergic (clusters 11 to 17), and nonneuronal cells (clusters 18 to 19). Different subclasses within each of these major cell classes can also be distinguished, although they overlap to some extent in the *t*-SNE plots. GABAergic neuron clusters correspond to four major subclasses: Lamp5-Sncg-Serpinf1 (clusters 11 and 12), Vip (clusters 13 and 14), Sst (clusters 15 and 16), and Pvalb (cluster 17). In the glutamatergic neuronal population, IT neurons from different cortical layers, PT, NP, CT, and L6b neurons are separated clearly by clusters 1 to 7, 8, 9, and 10, respectively (Fig. 1B and SI Appendix, Fig. S2). The nonneuronal cells are clearly separated into two major clusters with one corresponding to astrocytes (cluster 18) and the other representing oligodendrocytes, endothelial, vascular and microglia cells (cluster 19). Similar observations were made from the Tasic 2016 dataset, albeit at lower resolution (SI Appendix, Fig. S3 and Dataset S1). When we colored the single cells by cortical regions, we observed several dense areas with ALM/VISp-derived cells in the glutamatergic neuron clusters, but no obvious regional aggregation of cells in the GABAergic neuron clusters (Fig. 1A), which is similar to the observation by Tasic et al. (2) based on gene-expression profiles. These data suggest that the global splicing profile at the single-cell level can clearly define the two primary neuronal classes, and to some extent, several major subclasses in the cortex.

### Differential Splicing between Glutamatergic and GABAergic Neurons.

Having demonstrated the neuronal cell-type specificity carried in the splicing profile, we next sought to identify alternative exons differentially regulated between different neuronal classes and subclasses. To optimize our analysis pipeline, we initially focused on the two primary neuronal classes, glutamatergic and GABAergic neurons, and detected 937 and 901 DS cassette exons in the Tasic 2016 and Tasic 2018 datasets, respectively (changes in PSI or  $|\Delta\Psi| \geq 0.1$ , false-discovery rate [FDR]  $\leq 0.05$ ; cells from ALM and VISp were aggregated together for this analysis) (Materials and Methods). Among them, 421 exons are common ( $P < 7.8e-305$ ; hypergeometric test) and all of them show splicing differences in the same direction in the two datasets (Fig. 2A); for the remaining exons that are significant in one dataset only,  $\sim 80\%$  of quantifiable exons show splicing differences in the same direction in the other dataset, although the difference is below the statistical cutoff. With more stringent criteria ( $|\Delta\Psi| \geq 0.2$ ), 342 and 269 cassette exons show differential splicing, totaling 469 DS exons from 368 genes used for additional characterization below (Fig. 2A and Dataset S2). Quantitatively, the  $\Delta\Psi$  values of the 469 unique DS exons are correlated well between the two datasets (Pearson's correlation  $R = 0.44$ ) (Fig. 2B and Dataset S2).

To assess potential regional differences, we also analyzed cells from ALM and VISp in the Tasic 2018 dataset separately. Splicing differences between glutamatergic and GABAergic neurons detected in the two regions are consistent in directions and also well correlated quantitatively ( $R = 0.38$ ) (SI Appendix, Fig. S4 and Dataset S2). Therefore, the differential splicing-regulatory program between glutamatergic and GABAergic neurons appears to be largely shared across the two cortical regions, justifying the aggregation of cells across regions in our analysis.

To further validate the reliability of DS exons detected in scRNA-seq data, we analyzed three bulk RNA-seq datasets of purified glutamatergic and GABAergic neurons (Dataset S1). A total of 320 DS exons were detected in one or more bulk RNA-seq dataset, with an average of 147 exons detected in individual datasets (Dataset S3). These exons have large overlap with the 469 DS exons we detected in scRNA-seq data with 160 exons in common ( $P < 1.4e-165$ , hypergeometric test) and splicing differences in the



**Fig. 1.** Clustering analysis of single-cell splicing profiles reveals major neuronal and glial cell classes. (A) Two-dimensional t-SNE plots of 19,853 core cells from the Tasic 2018 scRNA-seq dataset. Cells are colored by cluster based on the splicing profile, cortical area of the cell, neuronal class, and neuronal types defined based on gene-expression profile from the original paper. (B) Heatmap showing the overlap between the clusters defined based on splicing profiles and the original neuronal cell types defined based on expression profiles. The color key represents the percentage of cells in each cluster defined based on splicing profiles overlapped with the original transcriptional cell types.

same direction (Fig. 2C). We also examined DS exons previously validated in a previous study (14); our analysis detected significant differences for five of eight cassette exons annotated in our database, all in the correct direction.

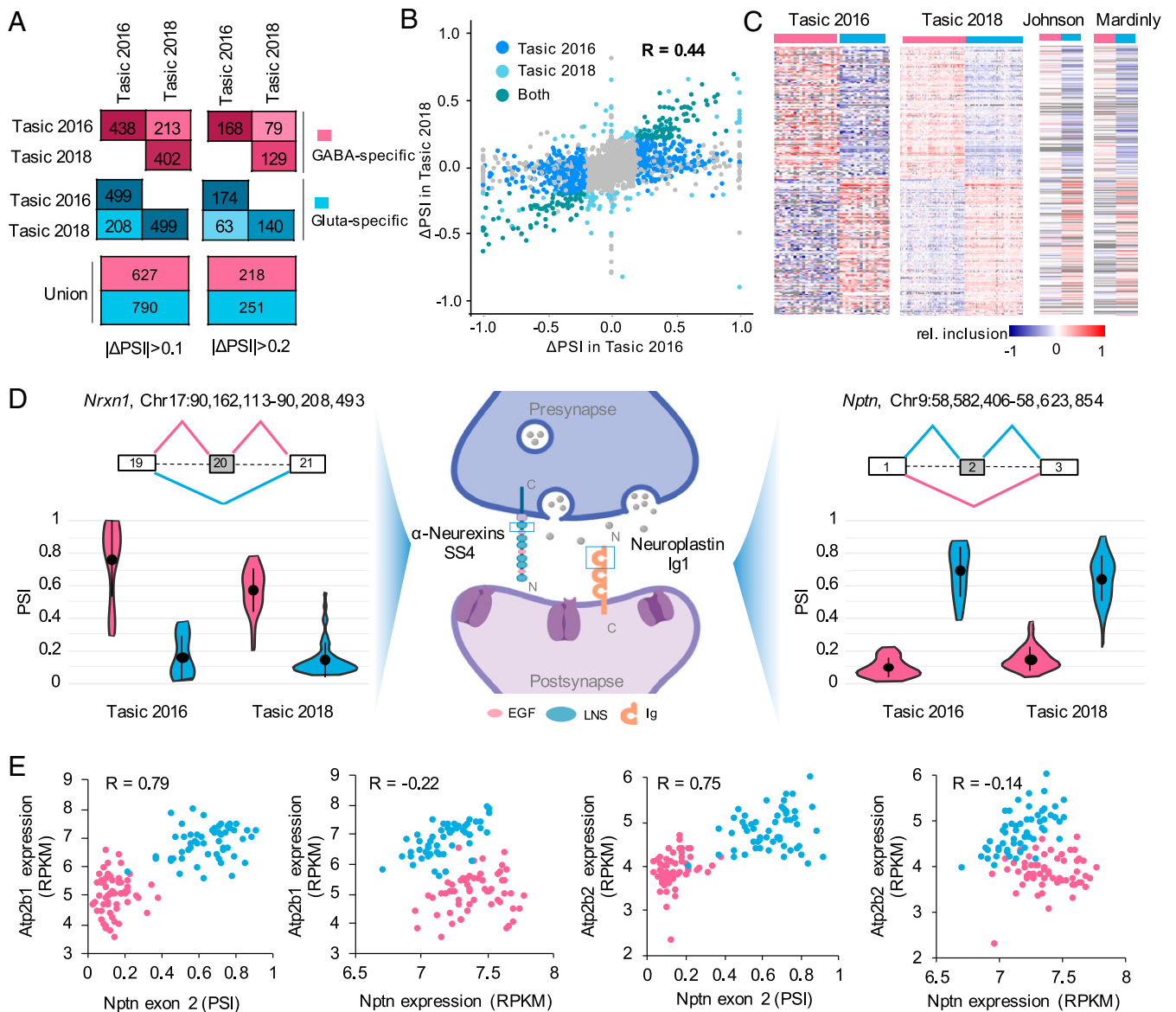
Compared to all cassette exons, the DS exons show characteristics of highly regulated, potentially functional alternative exons, with a higher percentage of in-frame coding exons (>75% vs. 54%) and

more conserved AS patterns in human and/or rat (>49% vs. 24%). In addition, a larger proportion of DS exons show neuron-specific splicing when comparing neuronal vs. nonneuronal cells in the cortex (>24% vs. 6%) and dynamic splicing switches at different developmental stages (~60% vs. 12%) (SI Appendix, Fig. S5A and Datasets S2 and S3). Gene ontology (GO) analysis using the 368 genes containing DS exons revealed significant enrichment of genes under three major GO categories, including synapse, cell projection, and ion channel activity (SI Appendix, Fig. S5B and C), in agreement with previous observations (14, 15). Therefore, our computational pipeline for scRNA-seq data is able to reliably detect differential splicing between distinct neuronal populations with comparable or improved sensitivity compared with bulk RNA-seq data.

As individual examples, our analysis detected GABAergic neuron-specific inclusion of the well-studied *Nrxn1* exon 20 (also known as alternatively spliced site 4, or SS4), which encodes a peptide overlapping with the sixth LNS domain of the Neuexin 1 protein (Fig. 2D, Left and Center). AS of this exon is known to be neuronal activity-dependent and may contribute to neuron-type-specific synaptic development by modulating its postsynaptic binding partners (22, 23). As a second, less characterized example, we identified *Nptn* exon 2 to be highly included in glutamatergic neurons but predominantly skipped in GABAergic neurons (Fig. 2D, Center and Right). *Nptn* encodes a cell-adhesion molecule that mediates formation and stabilization of excitatory synapses, activity-dependent long-term synaptic plasticity, as well as excitatory/inhibitory balance (24). Specific depletion of *Nptn* in glutamatergic neurons in mice results in altered neuronal activity and behavioral deficits by reducing the level of plasma membrane calcium ATPase (*Pmca*) genes *Atp2b1* and *Atp2b2*, and thus increasing intracellular  $Ca^{2+}$  concentration (25). Inclusion of exon 2 generates a longer isoform Np65 containing three Ig domains, which was demonstrated to be brain- and neuron-specific and localized between pre- and postsynaptic components of excitatory synapses (24). Interestingly, our analysis suggests that the inclusion of exon 2, but not the overall *Nptn* mRNA level, is strongly correlated with *Pmca* gene expression across neuronal cell types ( $R = 0.79$  vs.  $-0.22$  for *Atp2b1*;  $R = 0.75$  vs.  $-0.14$  for *Atp2b2*) (Fig. 2E). Therefore, Np65 might be the major contributor that regulates distinct *Pmca* expression between glutamatergic and GABAergic neurons.

### Neuronal Subclass-Specific Splicing Programs at Different Hierarchical Levels.

Compared to bulk RNA-seq data derived from a priori-defined neuronal cell types, scRNA-seq data with in-depth sampling of cells such as the Tasic 2018 dataset offers the unique opportunity to interrogate differential splicing between neuronal subclasses defined at different hierarchical levels. Our t-SNE analysis of splicing profiles showed aggregation of single cells reflecting subclasses of glutamatergic and GABAergic neuronal populations, although they overlap in the low-dimensional representation afforded by t-SNE (Fig. 1A). This overlap could be partly due to insufficient read coverage that resulted in the uncertainty in the single-cell splicing profiles. To reduce the uncertainty of splicing quantification, we performed an unbiased hierarchical clustering analysis of the splicing profiles of 55 glutamatergic and 60 GABAergic transcriptional cell types, respectively, by pooling the single cells assigned to each neuronal cell type (Fig. 3A). This analysis revealed clear separation of CGE- and MGE-originating GABAergic neurons, as well as *Vip*, *Sncg*, *Lamp5*, and *Serpinf1* subclasses in the CGE cluster and *Pvalb* and *Sst* subclasses in the MGE cluster. The glutamatergic neurons are mainly divided by projection subtypes, with IT neurons constituting the largest branch, followed by PT, NP, and L6b neuronal subclasses, while CT neurons are distributed across the branches (Fig. 3B). Within the IT branch, we observed higher similarity between L2/3 IT and L6 IT neurons as compared to L4/5 neurons (Fig. 3B). These observations confirmed that the molecular



**Fig. 2.** A robust set of DS cassette exons between glutamatergic and GABAergic neurons identified from scRNA-seq data. (A) Number of DS exons with higher inclusion in either GABAergic neurons (GABAergic-specific exons, pink) or glutamatergic neurons (glutamatergic-specific exons, blue) detected in Tasic 2016 and Tasic 2018 datasets. (B) Correlation of  $\Delta\Psi$  between GABAergic and glutamatergic neurons for all cassette exons estimated from Tasic 2016 and Tasic 2018 datasets. Dark blue points represent the DS exons ( $\Delta\Psi \geq 0.2$ ) only detected in Tasic 2016, light blue points represent the DS exons only detected in Tasic 2018, green points represent the overlap of DS exons detected in both datasets and gray points represent the other non-DS exons. The Pearson correlation is indicated. (C) Heatmaps showing exon inclusion level (with median subtraction) of DS exons in the two scRNA-seq datasets and two independent bulk RNA-seq datasets. Only DS exons with quantifiable exon inclusion in  $\geq 50\%$  of neuronal cell types in both Tasic 2016 and 2018 datasets were included for visualization. (D) Examples of DS exons in synaptic-related genes *Nrxn1* and *Nptn* showing GABAergic or glutamatergic-specific inclusion, respectively. The violin plots show the distribution of PSI in single cells of GABAergic and glutamatergic neuronal classes. Median and the interquartile range are indicated. (E) *Nptn* exon 2 inclusion but not gene expression is correlated with *Pmca* (encoded by *Atp2b1* and *Atp2b2*) expression. Glutamatergic and GABAergic neuronal cell types are colored in blue and pink, respectively. Pearson's correlation is indicated in each scatter plot.

diversity at the splicing level not only contributes to the distinction of glutamatergic vs. GABAergic neuronal classes, but also the delineation of subclasses in each population at multiple hierarchical levels.

Consistent with the unsupervised clustering analysis, we identified robust sets of DS exons by contrasting neuronal subclasses defined at different hierarchical levels and these exons also showed characteristics of highly regulated AS exons (SI Appendix, Fig. S6). For the GABAergic neuronal types, 230 exons are DS between MGE- and CGE-originating neurons, and between 75 and 314 exons show differential splicing when the four major

subclasses—*Pvalb*, *Sst*, *Vip*, and *Lamp5-Sncg-Serpinf1* neurons—were compared ( $\Delta\Psi \geq 0.1$ ,  $FDR \leq 0.05$ ) (Fig. 3B and Dataset S4). Similarly, between 93 and 384 DS exons were identified between major glutamatergic subclasses, including IT, PT, NP, CT, and L6b neurons, with the largest number observed from NP neurons (Fig. 3B and Dataset S4). In addition, IT neurons represent a heterogeneous group, and subsets of exons show distinct splicing patterns depending on the cortical layers from which they are derived.

To inform the potential function associated with neuronal class- or subclass-specific AS at different hierarchical levels, we examined

GO terms enriched in lists of genes containing DS exons for each neuronal subclass. We found that terms related to cell–cell adhesion, synapse, and ion channel/receptor activity are shared by multiple subclasses, suggesting that an enormous AS-mediated diversity is required for these genes related to fundamental neuronal properties (Fig. 3C and Dataset S5). In particular, we found that both ligand-gated and voltage-gated ion channel genes frequently contain specific isoforms enriched or depleted in multiple subclasses (Fig. 3D). On the other hand, our analysis also identified GO terms that are specifically enriched in certain subclasses, such as “excitatory extracellular ligand-gated ion channel activity” in L6b neurons and “calmodulin binding” and “metal ion transport” in Pvalb neurons (Fig. 3C and Dataset S5).

In accordance with the complex splicing regulation of ion channel gene families, comparison between lists of DS exons contrasting neuronal classes and subclasses at different hierarchical levels showed interesting overlaps (Fig. 4A). In particular, we noticed a significant overlap between exons DS in GABAergic vs. glutamatergic neurons and those DS in NP vs. other subclasses of glutamatergic neurons ( $P = 1e-151$ ; Fisher’s exact test), with similar splicing patterns in GABAergic and NP neurons. One major characteristic feature shared between GABAergic interneurons and NP neurons, but distinct from the other types of glutamatergic neurons, is the range of projection. To investigate whether a shared splicing program underlies this property, we took advantage of a subclass of L6b neurons, which also have local projections (2) and identified DS exons between local vs. long projecting L6b neurons. Comparison of the three lists of DS exons allowed us to identify 41 common exons with essentially all exons showing consistent splicing patterns with respect to projection types (Fig. 4B and C and Dataset S6). Genes containing these exons have functions associated with dendrite, synapse, and cell junctions, including multiple genes, such as *Tsc2*, *Arhgap44*, and *Macf1*, which were previously demonstrated to play a role in axon outgrowth and pathfinding (26–28). In addition, nine genes were shown to undergo dynamic AS switches during axonogenesis (29) (odds ratio = 11,  $P = 8.2e-7$ , Fisher’s exact test) (Fig. 4D). While further experimentation is required, these data suggest a splicing program correlated with and potentially contributing to near vs. long-range projection. Taken together, the hierarchical analysis suggests a complex landscape of AS contributing to the molecular diversity across neuronal cell types.

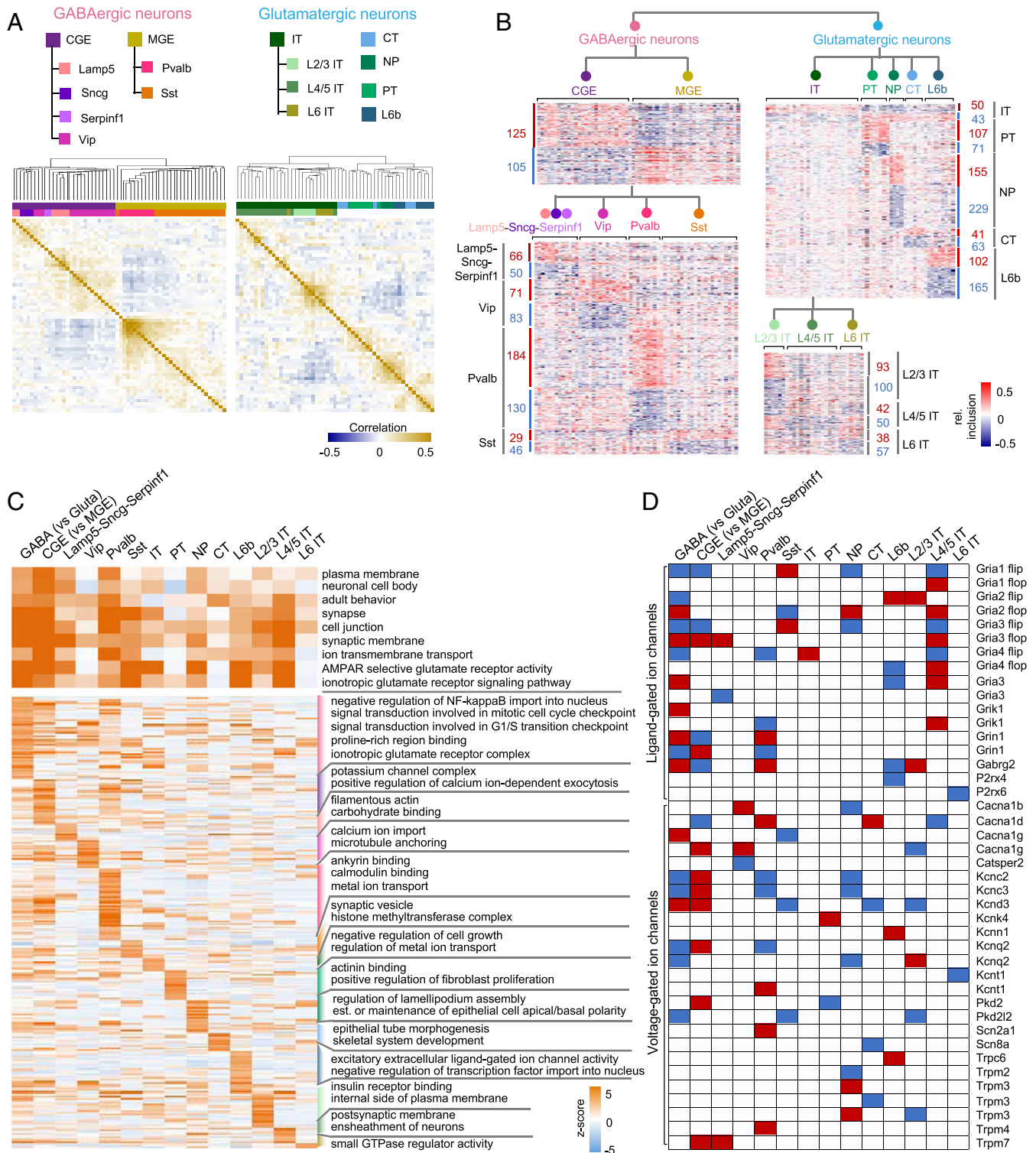
**RBP Regulating Differential Splicing between Glutamatergic and GABAergic Neurons.** We next asked how the distinct splicing programs across neuronal cell types are controlled. Unbiased clustering analysis of RBP expression clearly segregated glutamatergic and GABAergic neurons, implying their potential role in regulating differential splicing between neuronal classes (SI Appendix, Fig. S7 and Dataset S7). Differential expression analysis identified 17 RBPs (e.g., Celf2/5, Khdrbs2/3, Rbfox3, and Mbnl2) showing higher expression in glutamatergic neurons and 24 RBPs (e.g., Qk, Elavl2, Rbmx, and Upf3b) showing higher expression in GABAergic neurons ( $FDR \leq 0.05$ , fold-change  $\geq 1.5$ ) (SI Appendix, Fig. S8 and Dataset S7). In addition, we identified 78 RBPs that are differentially expressed in specific neuronal subclasses within glutamatergic or GABAergic neuronal classes at different hierarchical levels (SI Appendix, Fig. S8 and Dataset S7). Further analysis of some of these RBPs suggest that their expression pattern is highly reproducible in independent bulk RNA-seq data (SI Appendix, Fig. S9) and can also be validated by immunostaining at the protein level (SI Appendix, Fig. S10A and B).

To identify regulators of differential splicing between glutamatergic and GABAergic neurons, we began with an unbiased de novo motif analysis by searching for  $k$ -mers ( $k = 4, 5, 6$ ) enriched in the glutamatergic or GABAergic neuron-specific alternative exons and flanking intronic sequences (within 200 nucleotide region). This analysis identified 78 4-mers, 71 5-mers, and 56 6-mers

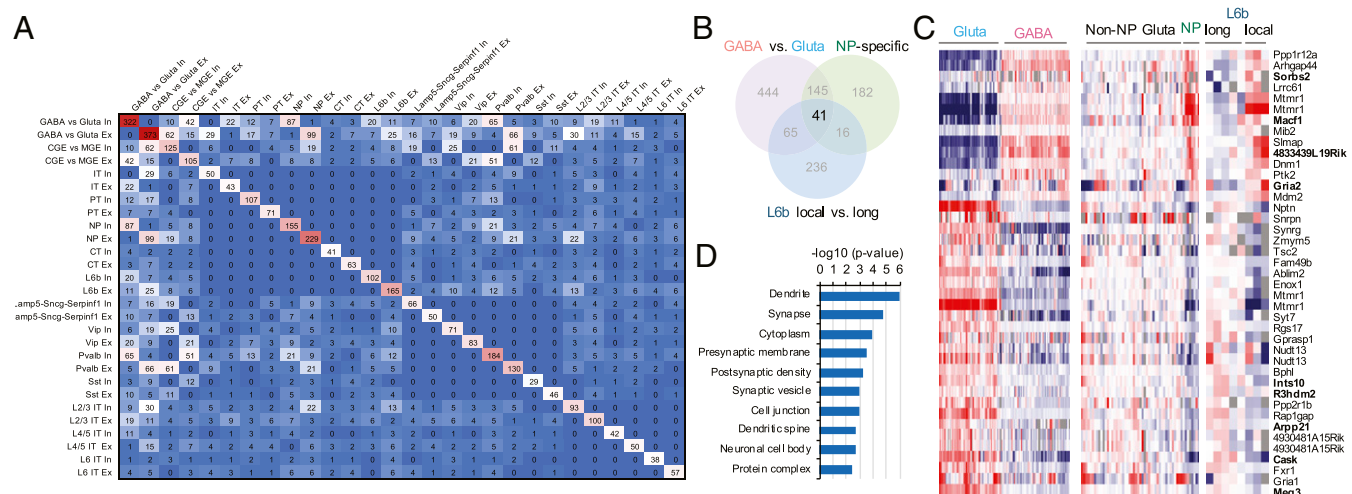
that are enriched in or around glutamatergic or GABAergic neuron-specific exons ( $FDR \leq 0.05$ , hypergeometric test) (Dataset S8). We clustered the significant  $k$ -mers to generate a  $k$ -mer graph for visualization based on their sequence similarity, as well as position-dependent enrichment patterns, since many RBPs activate or repress exon inclusion depending on where they bind (frequently activate exon inclusion when binding to the downstream and repress exon inclusion when binding to the alternative exon or upstream) (Materials and Methods). The enriched  $k$ -mers can be organized into several clusters resembling RBP binding sites known to regulate AS in neurons (Fig. 5). The largest cluster consists of UG- or GU-rich sequences that are recognized by the Celf family RBPs; these  $k$ -mers are enriched in the downstream intron for exons preferentially included in glutamatergic neurons and in the upstream intron for exons preferentially included in GABAergic neurons. The position-dependent enrichment pattern is most consistent with the higher expression of Celf2 in glutamatergic neurons (and to some extent, also Celf1, which has significant albeit more moderate differences), so that Celf1/2-dependent splicing activation by binding to the downstream intron results in preferential exon inclusion in glutamatergic neurons while Celf1/2-dependent splicing repression by binding to the upstream intron results in preferential exon skipping in glutamatergic neurons (i.e., preferential exon inclusion in GABAergic neurons). It is similarly true for  $k$ -mers containing YGCY tetramers ( $Y = C/U$ ), representing the consensus binding motif of Mbnl proteins, especially Mbnl2, which is also more highly expressed in glutamatergic neurons.

In contrast, a cluster of U- or AU-rich  $k$ -mers shows enrichment in the downstream intron of GABAergic neuron-specific exons and the upstream intron of glutamatergic-specific exons. This enrichment pattern is most consistent with Elavl2, which is more highly expressed in GABAergic neurons, so that its binding in the downstream intron results in splicing activation and preferential exon inclusion in GABAergic neurons, while its binding in the upstream intron results in splicing repression and preferential exon skipping in GABAergic neurons (i.e., preferential exon inclusion in glutamatergic neurons). The other two clusters contain UWAA ( $W = U/A$ ) and C-rich motifs, which are known as the binding consensus of Khdrbs and Pcbp families, respectively.

To further establish that differential RBP expression directly contributes to the distinct splicing patterns between glutamatergic and GABAergic neurons, we examined splicing targets of several neuronal RBP families as determined by RBP depletion followed by RNA-seq or, whenever possible, integrative modeling of multiple types of data, including RBP-dependent splicing and direct protein–RNA interactions, as we described previously (12, 16) (Dataset S9). Indeed, for RBPs with higher expression in GABAergic neurons (Elavl2 and Qk), exons with RBP-dependent inclusion are preferentially included in GABAergic neurons, while exons with RBP-dependent repression are preferentially included in glutamatergic neurons, consistent with the notion that higher activity of these RBPs in GABAergic neurons directly contributes to GABAergic neuron-specific splicing of their targets. Therefore, the directional splicing changes caused by RBP depletion would reduce or diminish the differential splicing of these exons between glutamatergic and GABAergic neurons, indicating a causative role of these RBPs. The opposite is true for Mbnl2 and Celf, with higher expression in glutamatergic neurons. Khdrbs3 is known to mostly repress exon inclusion, consistent with the preferential skipping we see of its target exons in glutamatergic neurons (Fig. 6A). In contrast, Rbfox, Ptpb, and Nova do not seem to have a systematic impact on the differential splicing between glutamatergic and GABAergic neurons, although Rbfox3 is more highly expressed in glutamatergic neurons. When we projected splicing profiles into the low-dimensional space using principal component analysis, we observed shifts of splicing profiles toward GABAergic neurons upon depletion of Mbnl1/2, Celf1, and Khdrbs3, and shifts toward glutamatergic neurons upon depletion of Qk (no RNA-seq data



**Fig. 3.** Neuronal subclass-specific splicing identified at different hierarchical levels. (A) Heatmap showing the pairwise Pearson's correlation coefficient of splicing profiles across neuronal cell types within the GABAergic (Left) or glutamatergic (Right) neuronal class. In each panel, exons quantifiable in  $\geq 50\%$  of cell types in the respective neuronal class with  $SD \geq 0.1$  were used for analysis. Cell types are ordered by hierarchical clustering with the dendrogram tree shown at the top. Cell types are colored by their developmental origin (MGE or CGE) and major subclasses (Pvalb, Sst, Vip, Lamp5, Sncg, and Serpinf1 for GABAergic neurons and IT, CT, PT, NP, and L6b for glutamatergic neurons; IT subclasses further divided by layer). (B) Heatmaps showing relative exon inclusion level (with median subtraction) of DS exons between neuronal subclasses detected at different hierarchical levels. Each row represents an exon and each column represents a cell type. The number of DS exons in each comparison is labeled. (C) GO terms associated with DS exons contrasting neuronal classes or subclasses at different hierarchical levels. Shared GO terms by multiple comparisons (i.e., significant in  $\geq 5$  comparisons) are shown at the top, while GO terms enriched in specific subclasses are shown at the bottom. GO term enrichment z-scores are shown in the heatmap with color scale indicated. (D) Neuronal class or subclass-specific exons in ligand- or voltage-gated ion channel genes. Red and blue indicate significant inclusion or skipping in the respective neuronal class or subclass, respectively.



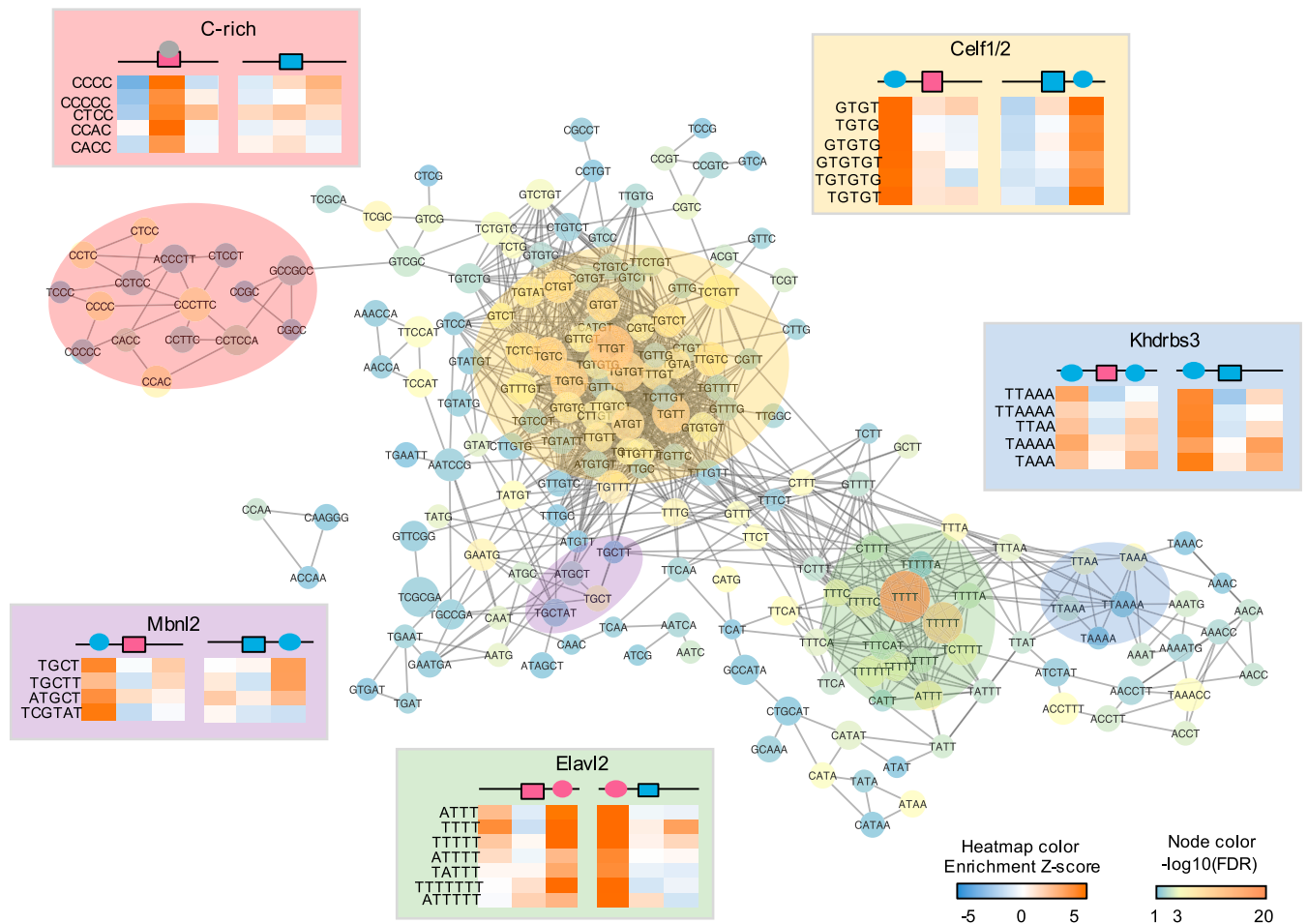
**Fig. 4.** Graded regulation of splicing across neuronal cell types. (A) Pairwise comparison showing overlap of DS exons detected at different hierarchical levels. Exons with neuronal class or subclass-specific inclusion and skipping are shown separately. (B) Overlap of DS exons identified in comparison of neuronal subclasses at different hierarchical levels potentially associated with range of neuronal projection. (C and D) Splicing profile (C) and associated GO terms (D) for 41 common DS exons identified in B. Exons in genes showing axonogenesis-dependent AS switches (29) are highlighted in bold in C.

with neuronal Elavl depletion), suggesting the global impact of these RBPs (Fig. 6B). Our analyses suggest that a combination of RBPs preferentially expressed in glutamatergic (Celf1/2, Mbnl2, and Khdrbs3) and GABAergic neurons (Elavl2 and Qk) together control the distinct splicing programs in the two neuronal classes.

It is particularly intriguing that about 20% of Mbnl-regulated exons show differential splicing between glutamatergic and GABAergic neurons (12% of DS exons of the two neuronal classes), presumably due to the higher expression and activity of Mbnl2 in glutamatergic neurons. The importance of Mbnl2 in regulating differential splicing between glutamatergic and GABAergic neurons has not been characterized, but is further supported by several lines of evidence in our analysis. The direct regulation of the DS exons by Mbnl2 is not only supported by position-dependent binding sites evident from motif analysis, but also from Mbnl2 binding footprints mapped by CLIP (Fig. 7A). In addition, 44% of DS exons between glutamatergic and GABAergic neurons show developmental splicing switches observed in the mouse cortex, and exons with late splicing switches between P4 and P15 are particularly enriched, consistent with Mbnl2 as a major regulator that promotes the mature splicing program in the cortex (12, 30) (Fig. 7B). Strikingly, among DS exons regulated by Mbnl2, the differential splicing pattern between glutamatergic and GABAergic neurons is very similar to that between adult and embryonic cortices, and can be explained by Mbnl2-dependent activation or repression (Fig. 7C). As a specific example, exon 8 of metabotropic glutamate receptor 5 gene (*Gm5*, also known as mGluR5), which encodes 32 amino acids in the C-terminal region, is more highly included in glutamatergic over GABAergic neurons (PSI = 0.81 vs. 0.41) (Fig. 7D–F) and developmentally regulated (Fig. 7G). The cell type and developmental stage-specific splicing is due, at least in part, to splicing activation by Mbnl2 by binding to the downstream intron, as depletion of Mbnl2 and its homolog Mbnl1 dramatically reduces exon inclusion toward the level observed in GABAergic neurons (PSI = 0.83 to 0.56) (Fig. 7H). Interestingly, inclusion or skipping of this exon generates protein isoforms with distinct roles in neurite outgrowth (31), and therefore, may contribute to morphological and functional differences between glutamatergic and GABAergic neurons. Together, these data support the notion that Mbnl2 plays a pivotal role in regulating the glutamatergic neuron-specific splicing program during brain development.

**Graded, Highly Combinatorial Splicing Regulation across Diverse Neuronal Cell Types.** We noted that ~46% (810 of 1,748) of neuronal class or subclass-specific DS exons show evidence of differential splicing at multiple hierarchical levels (*SI Appendix*, Fig. S11A and *Dataset S4*; see examples in Fig. 4D). In parallel, over half (49 of 86) of differentially expressed RBPs showed differences between neuronal subclasses defined at multiple hierarchical levels (*SI Appendix*, Figs. S11B and S12). We argue that the rich dynamics of RBP expression and downstream AS events across neuronal cell types can be used to infer the activity of an RBP in regulating differential splicing in specific neuronal subclasses. Target exons activated by an RBP should show positive correlation with RBP expression, while target exons repressed by the RBP should show negative correlation. Therefore, the impact of the RBP in regulating global differential splicing across all or subsets of neuronal cell types can be evaluated using a gene set enrichment analysis (GSEA) (32) that examines the nonrandom distribution of RBP targets among all exons ranked by their correlation with RBP expression.

We applied GSEA analysis to 16 RBPs (from six RBP families) whose splicing targets were defined as described above. Strikingly, GSEA analysis revealed significant contribution of nine RBPs, such as Mbnl2, Qk, and Nova1, and many of them across neuronal subclasses at multiple hierarchical levels (Fig. 8A and B). For example, Mbnl target exons show nonrandom distribution when ranked based on their correlation with Mbnl2 expression across all neuronal cell types, with activated exons at the top (positive correlation) and repressed exons at the bottom (negative correlation) (Fig. 8B). This is consistent with our analysis above, indicating Mbnl2 as a major regulator of differential splicing between glutamatergic and GABAergic neurons. Importantly, similar patterns were observed when GSEA analysis was applied to GABAergic and glutamatergic neuronal cell types separately, suggesting that Mbnl2 also has a global impact on splicing dynamics within the two major neuronal classes. Indeed, we found 51 Mbnl target exons with differential splicing between L2/3 IT neurons and L6b neurons, which have the highest and lowest Mbnl2 expression within glutamatergic subclasses, respectively (Fig. 8C). Among them is GABA receptor- $\gamma$ 2 subunit (*Gabrg2*) exon 9, whose inclusion is activated by Mbnl2 directly through a downstream intronic binding site (12). Inclusion of this exon is positively correlated with Mbnl2 expression ( $R = 0.76$ ), with the lowest inclusion observed in L6b as compared to the other glutamatergic subclasses (Fig. 8D). In contrast, *Spint 2* exon 4, which is repressed by Mbnl most likely through exonic and



**Fig. 5.** De novo motif analysis of DS exons with glutamatergic or GABAergic neuron-specific splicing. The graph shows significantly enriched  $k$ -mers ( $k = 4$  to 6) in either glutamatergic or GABAergic neuron-specific exons or their flanking intronic sequences.  $k$ -mers with similar sequences and position-dependent enrichment patterns are connected. Nodes are colored by the FDR of the most significant region. Five clusters of  $k$ -mers (shaded circles) are highlighted. For each cluster, the most likely cognate RBP (when available) is indicated and the position-dependent enrichment patterns of representative  $k$ -mers are also shown using a heatmap. Blue and pink colors represent higher RBP expression or exon inclusion in glutamatergic and GABAergic neurons, respectively.

upstream intronic binding sites, shows negative correlation with Mbn12 expression ( $R = -0.72$ ), with the highest inclusion in L6b neurons (Fig. 8E).

Intriguingly, in some cases GSEA analysis revealed the global impact of RBPs on differential splicing across neuronal cell types, but the pattern of dynamics cannot be readily explained by the hierarchical taxonomy. For example, Nova1 and Nova2 do not show clear differential expression between adult glutamatergic and GABAergic neurons or between subclasses in our analysis (except down-regulation of Nova1 in Lamp5-Sncg-Serpinf1), and yet GSEA analysis detected a significant impact of Nova1/2 on the differential splicing dynamics across all neuronal cell types, or within glutamatergic (Nova2) or GABAergic neurons (Nova1) (Fig. 8A and B). Similar observations were made for Rbfox1 and Ptpb2 (Fig. 8A). These results suggest a complex picture of graded, highly combinatorial, regulation of splicing across diverse neuronal subtypes (Fig. 8F).

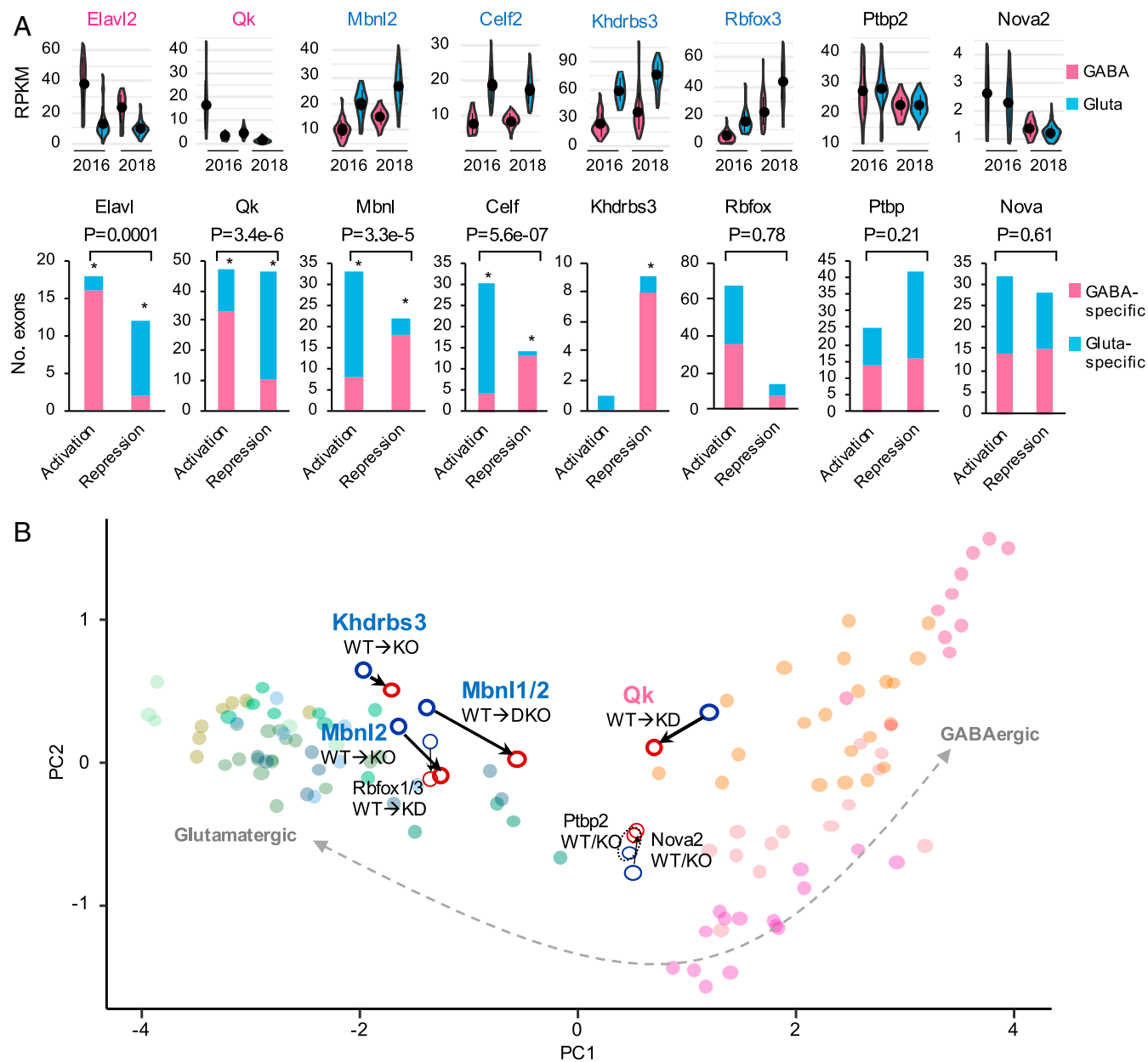
## Discussion

Here we report systematic analysis of neuronal cell-type-specific AS and the underlying regulatory mechanisms in the adult mammalian brain using deep scRNA-seq data (2, 3). Our analysis focused on cassette exons since they represent the most abundant type of AS events in mammals and inclusion of other types of AS events (such as alternative 5' or 3' splice sites) did not seem to

improve the resolution of clustering, probably because their quantification is more challenging using scRNA-seq data. Compared to recent work using bulk RNA-seq of purified neuronal populations or ribosome-engaged transcripts, this study offers two unique advantages. First, the unprecedented sequencing depth affords accurate quantification of AS and identification of DS exons with high sensitivity. Our analysis identified a larger number of DS exons from scRNA-seq data than from bulk RNA-seq data using stringent criteria (e.g., 469 vs. 320 cassette exons) (Datasets S2 and S3). Additional comparison with DS exons recently identified from bulk RiboTRAP-seq by another group suggests that the vast majority of DS exons we identified were not reported previously (14), and yet the identified DS exons in general show reproducible splicing patterns between the two scRNA-seq datasets and other independent bulk RNA-seq data, again confirming the reliability of our analysis (SI Appendix, Fig. S13).

Most importantly, the in-depth sampling of a large number of single cells without cell types defined a priori allows analysis of over 100 neuronal cell types that were defined from global transcription profiles rather than single reporter genes, including those missed in previous studies (e.g., the Lamp5-Sncg-Serpinf1 subclass of GABAergic interneurons) or treated as a homogeneous group (e.g., different subclasses of glutamatergic neurons). The hierarchical analysis of neuronal classes or subclasses afforded by scRNA-seq data provided several insights into the complexity



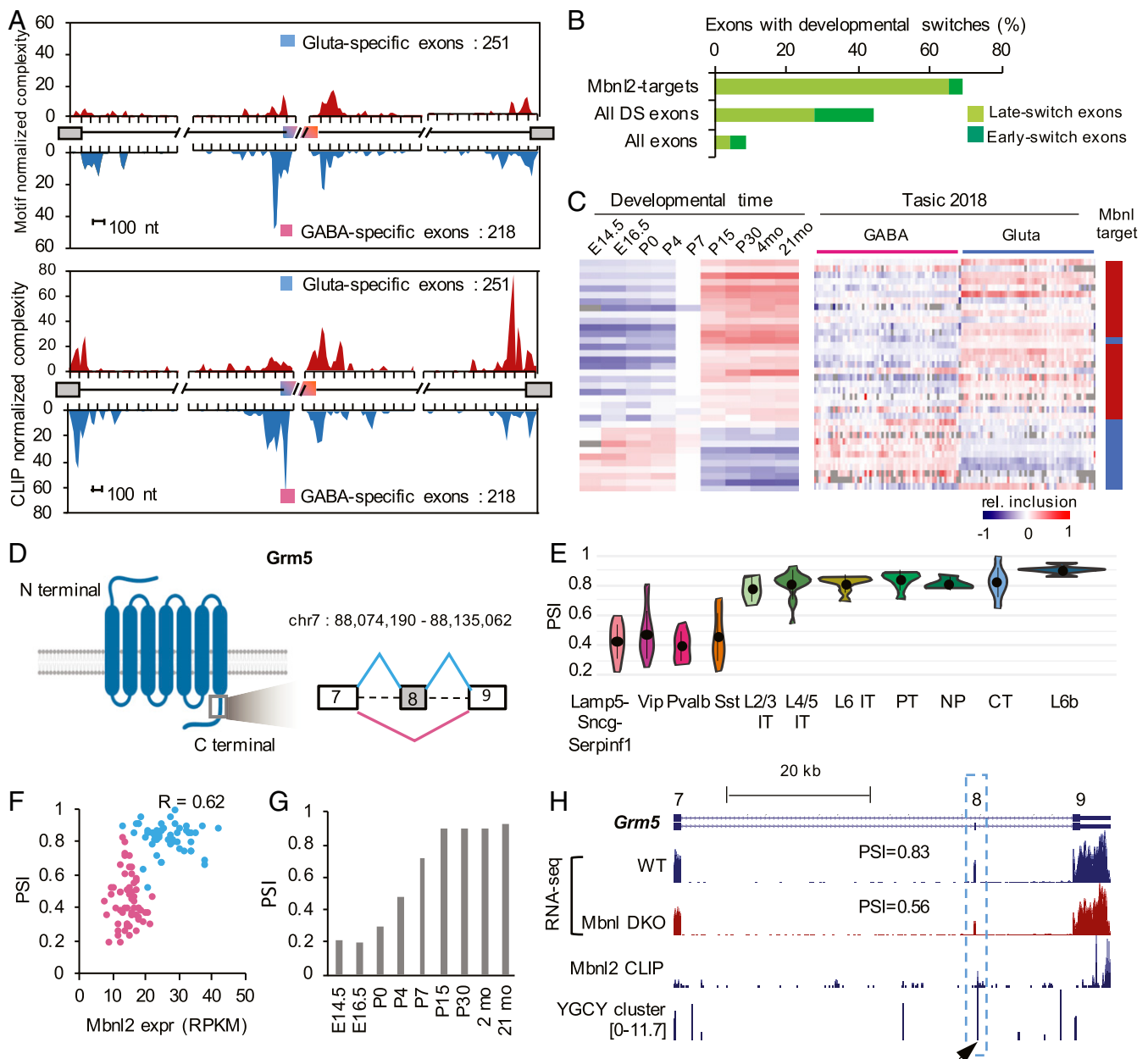


**Fig. 6.** RBPs regulating differential splicing between glutamatergic and GABAergic neurons. (A) Exons regulated by several RBP or RBP families showing glutamatergic or GABAergic neuron-specific exon inclusion. For each family, the representative RBP with differential expression between glutamatergic and GABAergic neurons estimated from Tasic 2016 and 2018 datasets is shown at the top. Note for Elavl2 and Qk with preferential expression in GABAergic neurons, exon activation or repression leads to a GABAergic neuron-like splicing profile, while for Mbnl2, Celf2 and Khdrbs3 with preferential expression in glutamatergic neurons, exon activation, or repression leads to a glutamatergic neuron-like splicing profile. An asterisk indicates significant bias toward glutamatergic or GABAergic neuron-specific inclusion for exons activated or repressed by the respective RBP, as compared to all DS exons between glutamatergic and GABAergic neurons ( $P < 0.05$ ). Exons activated by an RBP and those repressed by the RBP are also compared and the resulting  $P$  value is shown at the top. (B) A global impact of RBP depletion on the shift of splicing profile toward glutamatergic or GABAergic neurons. In the scatter plot, each dot is a sample, which is projected into the two-dimensional space defined by the first two principal components (PCs) of splicing profiles of glutamatergic or GABAergic neuronal cell types.

and regulation of the transcriptome at the splicing level. One characteristic feature of AS regulation that became clear from our analysis is the “graded” rather than dichotomic or on/off regulation more commonly seen at the transcription level. This is supported by differential splicing of many exons detected at multiple hierarchical levels (Fig. 3D and *SI Appendix*, Fig. S11A). In particular, a group of exons show coordinated differential splicing between GABAergic and glutamatergic neurons, between NP and other subclasses of glutamatergic neurons, and between L6b neuronal subclasses with local and long-range projection (Fig. 4). This

putative splicing program should be further assessed for its possible role in regulating neuronal morphology and types of synapses related to the range of neuronal projection for multiple neuronal subclasses.

Very little is currently known about neuronal cell-type-specific AS controlled by RBP splicing factors. It was recently shown that Nova2 is enriched in excitatory neurons and has distinct binding profiles in excitatory vs. inhibitory neurons in the developing cortex and cerebellum. Neuronal class-specific depletion of Nova2 affected splicing of a large number of exons, including a

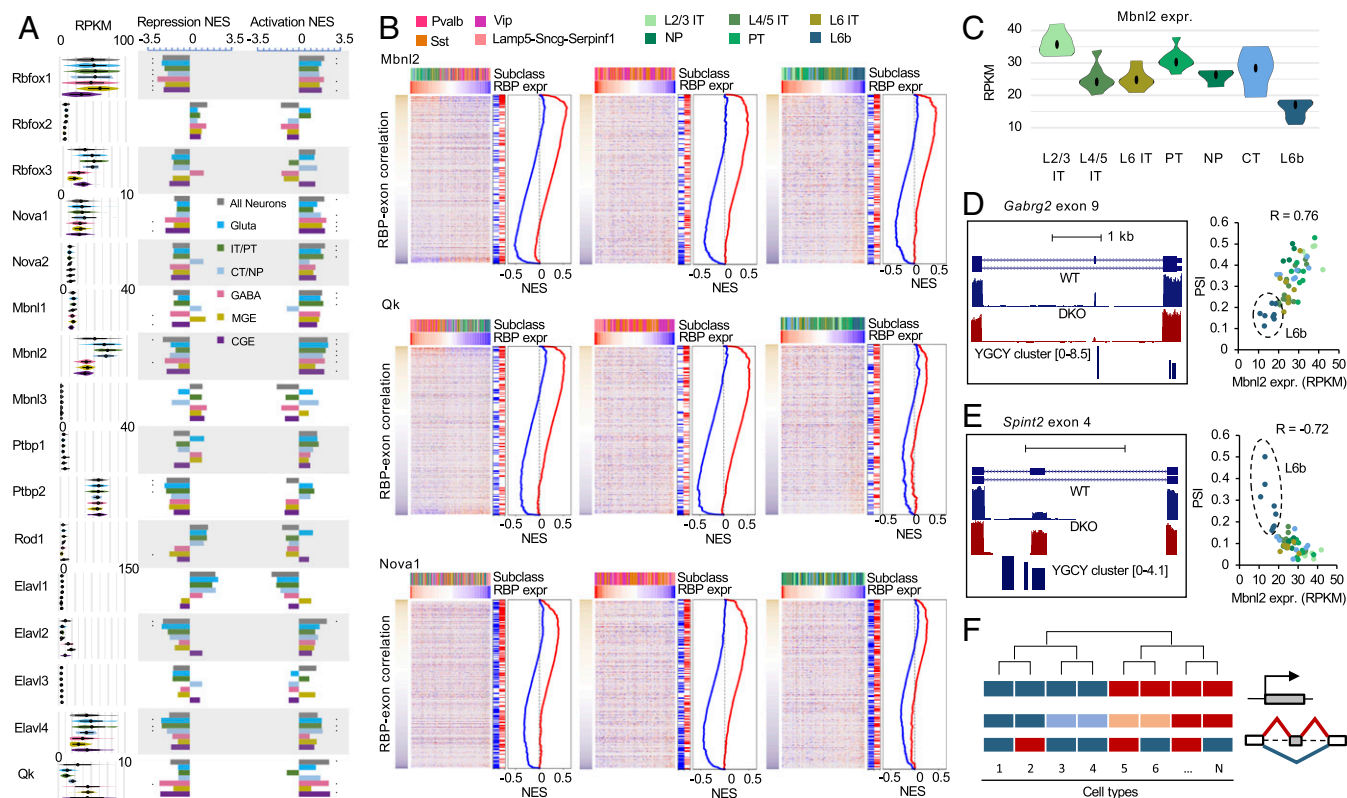


**Fig. 7.** Mbnl2 drives a glutamatergic neuron-specific splicing profile. (A) RNA maps showing enrichment of Mbnl binding sites (predicted YGCY clusters at the top and Mbnl2 CLIP tags at the bottom) in the upstream intron of cassette exons with GABAergic neuron-specific inclusion and in downstream intron of cassette exons with glutamatergic neuron-specific inclusion. (B) DS exons regulated by Mbnl are enriched in late-switch exons during cortical development as compared to all DS exons or all cassette exons. Exons showing early or late developmental splicing switches were obtained from Weyn-Vanhentenryck et al. (12). (C) Comparison of splicing profiles of DS exons regulated by Mbnl during cortical development (12) (Left) and across 117 neuronal cell types from Tasic 2018 (Right). Exons activated or repressed by Mbnl are indicated in red and blue, respectively, in the scale bar on the right. (D–H) Regulation of *Grm5* exon 8 by Mbnl2 is correlated with glutamatergic neuron-specific inclusion. (D) Schematic diagram of *Grm5* protein domain architecture and the region encoded by exon 8 with preferential inclusion in glutamatergic neurons. (E) Violin plots showing *Grm5* exon 8 inclusion in glutamatergic and GABAergic neuronal subclasses. Median and the interquartile range are indicated. (F) Correlation of exon inclusion and Mbnl2 expression across 117 neuronal cell types. Glutamatergic and GABAergic neuronal cell types are colored in blue and pink, respectively. Pearson's correlation coefficient is indicated. (G) *Grm5* exon 8 inclusion during cortical development. (H) Mbnl2 bind to the downstream intron of *Grm5* exon 8, as shown by Mbnl2 CLIP data and predicted Mbnl-binding YGCY clusters (arrowhead). Exon inclusion in the cortex is reduced upon depletion of Mbnl1/2 (Mbnl DKO).

subset with differential splicing between excitatory and inhibitory neurons (33). Another study demonstrated that *Rbfox1* is preferentially expressed in MGE- over CGE-derived cortical interneurons (34). Conditional depletion of *Rbfox1* in two subclasses of MGE-derived, PV, and SST interneurons resulted in distinct defects in synaptic connectivity and network excitability in mice. Hundreds of *Rbfox1*-dependent exons were reported in

PV and SST neurons with minimal overlaps, leading the authors to argue a unique role of *Rbfox1* in each interneuron subclass. Interestingly, the differential expression of these two RBPs between glutamatergic vs. GABAergic neurons is diminished in adult cortex from our analysis.

By integrating unbiased de novo motif discovery, differential RBP expression, RBP targets defined by RBP perturbation and



**Fig. 8.** GSEA analysis of graded splicing regulation across diverse neuronal subtypes. (A) RBPs with significant contribution to splicing regulation across neuronal cell types in various subclasses in are shown by the normalized enrichment score (NES) by GSEA. Exons activated and repressed by each RBP family are analyzed separately. \*FDR  $\leq 0.01$  and NES  $\geq 1.8$ . (B) GSEA plots of representative RBPs (Mbnl2, Qk, and Nova1) across all neuronal cell types (Left), across GABAergic cell types (Center), and across glutamatergic cell types (Right). In each plot, exon inclusion across cell types is shown in the heatmap, in which exons are ranked by their correlation with RBP expression and the neuron types are ordered by the corresponding RBP expression in decreasing order. The color scales of correlation and RBP expression are shown on the left and top of the heatmaps. The neuron subclasses are indicated at the top using different colors. The repressed and activated targets of the corresponding RBP is indicated on the right of the heatmap together with the NES plots. (C) Differential expression of Mbnl2 between L2/3 IT and L6 IT neurons from Tasic 2018 dataset. (D) *Gabrg2* exon 9 activation by Mbnl2 through downstream intronic binding sites results in differential splicing across glutamatergic neuronal types. Tracks show exon inclusion in WT and Mbnl1/2 DKO cortices and predicted Mbnl-binding YGCY clusters (Left). The correlation of exon inclusion and Mbnl2 expression across glutamatergic neuronal types is shown (Right). (E) *Spint2* exon 4 repression by Mbnl2 through exonic and upstream intronic binding sites results in differential splicing across glutamatergic neuronal types. See legends of D for more detail. (F) A schematic of the proposed model showing graded AS regulation that does not always conform to the hierarchical taxonomy of neuronal cell types defined based on transcriptional profiles.

protein–RNA interactions, position-dependent RNA-maps, and GSEA, our analysis highlighted several RBPs with global impact on the distinct splicing profiles of glutamatergic and GABAergic neurons. Mbnl2, Celf1/2, and Khdrbs3 have preferential expression and higher activity in glutamatergic than GABAergic neurons, and therefore splicing activation or repression by these RBPs results in a glutamatergic neuron-specific splicing profile. On the other hand, Elavl2 and Qk are more highly expressed and potent in GABAergic neurons, driving toward a GABAergic neuron-specific splicing profile. Together, these RBPs regulate about one-third of DS exons. Importantly, RBP perturbation integrated in our target network analysis indicates that the regulation is most likely causal and direct. Consistent with lack of differential expression, a limited, if any, impact of Rbfox and Nova RBP families was detected on the distinct splicing programs between adult glutamatergic and GABAergic neurons. The role of Khdrbs3 RBP family in regulating excitatory neuron-specific splicing was previously established through extensive studies of the alternative exon SS4 of neurexin genes (20, 21, 23), although their global impact is less clear. We found 10 of 13 Khdrbs3-dependent exons are DS between cortical glutamatergic and GABAergic neurons, suggesting a striking degree of cell-type-specificity of this splicing program. On the other hand, the involvement of the other

RBPs in regulating neuronal cell-type-specific AS has not been reported and is unexpected to some extent. For example, Mbnl2 regulates developmental splicing switches in the brain, which are disrupted in myotonic dystrophy, but the exact cell types in which it acts were unclear (12, 30). Qk was mostly studied in oligodendrocytes with an important role in regulating myelination (35) and, more recently, in embryonic neural stem cells affecting cellular differentiation (36). *Caenorhabditis elegans* homologs of Celf (UNC-75) and Elavl (EXC-7) were previously shown to regulate cholinergic- and GABAergic-specific splicing in worms (37), although their function has likely diverged in vertebrates due to the dramatic expansion of these RBP families. Our results suggest potential productive avenues to revealing the functional significance of neuronal cell-type-specific splicing by careful examination of neuronal phenotypes caused by depletion of these RBPs using in vitro and in vivo model systems.

Consistent with the proposed graded AS regulation, differential expression of RBPs are mostly quantitative, and frequently detected between neuronal classes or subclasses at multiple hierarchical levels (SI Appendix, Figs. S11B and S12), lending power to GSEA analysis to infer RBP activity by correlating RBP expression and exon inclusion across neuronal cell types without explicitly imposing the hierarchical structure. By applying GSEA analysis to

different subsets of neuronal cell types, we found that RBPs frequently have a global impact on the splicing dynamics at multiple hierarchical levels. The most surprising finding resulting from our GSEA analysis is probably the inferred activity of RBPs—such as *Rbfox1*, *Nova1/2*, and *Ptbp2*—even when they do not show clear differential expression pattern between established neuronal subclasses. Expression of these RBPs varies across diverse neuronal cell types composed of more homogeneous cell populations, and such variation clearly contributes to dynamics of AS across these cell types through splicing activation and repression, yet the dynamics do not always conform to the hierarchical taxonomy established using transcriptional profiles (Fig. 8 A and B). Although *Rbfox3* shows more distinct expression between glutamatergic and GABAergic neurons, *Rbfox1* explains the splicing dynamics across neuronal cell types better than *Rbfox3*. So what is the consequence of high vs. low expression of these RBPs on neuronal identity and function? We conjectured that the resulting differential splicing may contribute to neuronal properties that are distinct from those used to establish current classifications (e.g., long vs. short range of neuronal projection, rather than types of neural transmitters) (Fig. 4).

Taking these data together, we find that this study suggests a model in which transcriptomic diversification by AS in the adult cortex is not simply downstream of transcriptional regulation, but could provide an orthogonal mechanism to specify neuronal cell type identity and function through complex and graded regulation

(Fig. 8F). Further validation of this model would benefit from improved sensitivity of scRNA-seq as well as new analytical approaches, which will allow more unbiased analysis at higher resolution without relying on the established taxonomy.

## Materials and Methods

Raw reads from the Tasic scRNA-seq datasets (2, 3) were processed using the Quantas pipeline (<https://zhanglab.c2b2.columbia.edu/index.php/Quantas>), as we described previously (10). To identify exons DS between neuronal classes or subclasses, we pooled read counts from single cells assigned to each neuronal cell types and treated these cell types within a group as biological replicates. This allowed us to reduce sampling noise and technical variabilities at the single-cell level while accounting for the uniform representation and consistency of splicing across neuronal cell types within each group. This method resulted in more accurate identification of DS exons compared to other methods of pooling individual cells. We also found filtering based on fraction of quantifiable neuronal cell types is effective in minimizing false positives, as it imposed reproducible differences across cell types of the compared groups. A detailed explanation of the materials and methods can be found in [SI Appendix](#).

**Data Availability.** All study data are included in the article and/or supporting information.

**ACKNOWLEDGMENTS.** We thank members of the C.Z. laboratory for helpful discussion about the project. This study was supported by NIH Grants R01NS089676 and R01GM124486 (to C.Z.). High-performance computation was supported by NIH Grants S10OD012351 and S10OD021764.

1. H. Zeng, J. R. Sanes, Neuronal cell-type classification: Challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).
2. B. Tasic *et al.*, Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
3. B. Tasic *et al.*, Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
4. A. Zeisel *et al.*, Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
5. B. B. Lake *et al.*, Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
6. M. B. Johnson *et al.*, Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015).
7. E. T. Wang *et al.*, Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
8. Q. Pan, O. Shai, L. J. Lee, B. J. Frey, B. J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
9. J. C. Castle *et al.*, Expression of 24,426 human alternative splicing events and predicted *cis* regulation in 48 tissues and cell lines. *Nat. Genet.* **40**, 1416–1425 (2008).
10. Q. Yan *et al.*, Systematic discovery of regulated and conserved alternative exons in the mammalian brain reveals NMD modulating chromatin regulators. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3445–3450 (2015).
11. Y. Zhang *et al.*, An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
12. S. M. Weyn-Vanhenryck *et al.*, Precise temporal regulation of alternative splicing during neural development. *Nat. Commun.* **9**, 2189 (2018).
13. C. K. Vuong, D. L. Black, S. Zheng, The neurogenetics of alternative splicing. *Nat. Rev. Neurosci.* **17**, 265–281 (2016).
14. E. Furlanis, L. Traunmüller, G. Fucile, P. Scheiffele, Landscape of ribosome-engaged transcript isoforms reveals extensive neuronal-cell-class-specific alternative splicing programs. *Nat. Neurosci.* **22**, 1709–1717 (2019).
15. M. A. Huntley *et al.*, Genome-wide analysis of differential gene expression and splicing in excitatory neurons and interneuron subtypes. *J. Neurosci.* **40**, 958–973 (2020).
16. C. Zhang *et al.*, Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* **329**, 439–443 (2010).
17. S. M. Weyn-Vanhenryck *et al.*, HITS-CLIP and integrative modeling define the *Rbfox* splicing-regulatory network linked to brain development and autism. *Cell Rep.* **6**, 1139–1152 (2014).
18. M. Jacko *et al.*, *Rbfox* splicing factors promote neuronal maturation and axon initial segment assembly. *Neuron* **97**, 853–868.e6 (2018).
19. B. Raj, B. J. Blencowe, Alternative splicing in the mammalian nervous system: Recent insights into mechanisms and functional roles. *Neuron* **87**, 14–27 (2015).
20. T. Iijima, Y. Iijima, H. Witte, P. Scheiffele, Neuronal cell type-specific alternative splicing is regulated by the KH domain protein SLM1. *J. Cell Biol.* **204**, 331–342 (2014).
21. L. Traunmüller, A. M. Gomez, T. M. Nguyen, P. Scheiffele, Control of neuronal synapse specification by a highly dedicated alternative splicing program. *Science* **352**, 982–986 (2016).
22. T. Iijima *et al.*, SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. *Cell* **147**, 1601–1614 (2011).
23. T. M. Nguyen *et al.*, An alternative splicing switch shapes neurexin repertoires in principal neurons versus interneurons in the mouse hippocampus. *eLife* **5**, e22757 (2016).
24. P. W. Beesley, R. Herrera-Molina, K. H. Smalla, C. Seidenbecher, The neuroplastin adhesion molecules: Key regulators of neuronal plasticity and synaptic function. *J. Neurochem.* **131**, 268–283 (2014).
25. R. Herrera-Molina *et al.*, Neuroplastin deletion in glutamatergic neurons impairs selective brain functions and calcium regulation: Implication for cognitive deterioration. *Sci. Rep.* **7**, 7273 (2017).
26. Y. J. Choi *et al.*, Tuberous sclerosis complex proteins control axon formation. *Genes Dev.* **22**, 2485–2495 (2008).
27. M. Galic *et al.*, Dynamic recruitment of the curvature-sensitive protein ArhGAP44 to nanoscale membrane deformations limits exploratory filopodia initiation in neurons. *eLife* **3**, e03116 (2014).
28. N. Sanchez-Soriano *et al.*, Mouse ACF7 and drosophila short stop modulate filopodia formation and microtubule organisation during neuronal growth. *J. Cell Sci.* **122**, 2534–2542 (2009).
29. M. Zhang *et al.*, Axonogenesis is coordinated by neuron-specific alternative splicing programming and splicing regulator PTBP2. *Neuron* **101**, 690–706.e10 (2019).
30. K. Charizanis *et al.*, Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* **75**, 437–450 (2012).
31. S. Mion *et al.*, Bidirectional regulation of neurite elaboration by alternatively spliced metabotropic glutamate receptor 5 (mGluR5) isoforms. *Mol. Cell. Neurosci.* **17**, 957–972 (2001).
32. A. Subramanian *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
33. Y. Saito *et al.*, Differential NOVA2-mediated splicing in excitatory and inhibitory neurons regulates cortical development and cerebellar function. *Neuron* **101**, 707–720.e5 (2019).
34. B. Wamsley *et al.*, *Rbfox1* mediates cell-type-specific splicing in cortical interneurons. *Neuron* **100**, 846–859.e7 (2018).
35. L. Darbelli, S. Richard, Emerging functions of the Quaking RNA-binding proteins and link to human diseases. *Wiley Interdiscip. Rev. RNA* **7**, 399–412 (2016).
36. Y. Hayakawa-Yano *et al.*, An RNA-binding protein, Qki5, regulates embryonic neural stem cells through pre-mRNA processing in cell adhesion signaling. *Genes Dev.* **31**, 1910–1925 (2017).
37. A. D. Norris *et al.*, A pair of RNA-binding proteins controls networks of splicing events contributing to specialization of neural cell types. *Mol. Cell* **54**, 946–959 (2014).