Cell-type-resolved *NRXN1* isoforms across human brain tissues and hiPSC organoids

Lei Cao[1*], Yu Fan[1*], Sadaf Ghorbani[2*], Jessica Mariani[2], Yanchun Zhang[1], Michael B. Fernando[3], Jaroslav Bendl[1,4], John Fullard[1,4], Susana Isabel Ramos[5], Edward A. Mead[1], Gintaras Deikus[1], Kristin G Beaumont[1], Robert Sebra[1], Nadejda Tsankova[5], Panos Roussos[1,4,6], Kristen J. Brennand[2#] and Gang Fang[1#]

[1] *Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029*
[2] *Department of Psychiatry, Yale School of Medicine, New Haven, CT, 06520*
[3] *Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, 10029*
[4]*Center for Disease Neurogenomics; & Friedman Brain Institute; & Department of Psychiatry, New York, NY, USA*
[5]*Department of Pathology, Molecular, and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA*
[6]*Center for Precision Medicine and Translational Therapeutics, Mental Illness Research Education and Clinical Center (MIRECC), James J. Peters VA Medical Center, New York, 10468, USA*

[*]*These authors contributed equally.*

[#]*Correspondence: kristen.brennand@yale.edu, gang.fang@mssm.edu*

ABSTRACT

NRXN1 undergoes extensive alternative splicing that generates a highly diverse repertoire of isoforms, diversifying protein–protein interactions, shaping synaptic specialization, and contributing to neuropsychiatric disease when disrupted. However, the splicing landscape of *NRXN1* across distinct human brain cell types remains poorly defined. Because *NRXN1* expression level is relatively low in adult brains and human induced pluripotent stem cell (hiPSC) derived neurons, single-cell Iso-seq RNA sequencing often provides insufficient coverage to capture its full isoform diversity, particularly across heterogeneous cell populations. To address this gap, we developed an integrative sequencing strategy that combines single-cell transcriptomics, targeted enrichment of *NRXN1* transcripts, and long-read sequencing. This approach reveals a comprehensive catalog of cell-type resolved *NRXN1* isoforms across adult and fetal human postmortem brains as well as hiPSC-derived cortical organoids. From the adult prefrontal cortex (PFC) region, our analyses reveal distinct splicing programs across interneuron

subtypes, pyramidal neurons, and glial lineages. Comparisons between prenatal and adult brains indicate that *NRXN1* isoform profiles are established during early development and remain stable throughout neuronal maturation. In hiPSC organoid models, splicing profiles partially mirror both developmental and mature brain patterns, with a subset of splice sites exhibiting cell type- and stage-specific divergence. In the cerebellum of an autism case and in organoids derived from schizophrenia patients, carrying non-recurrent heterozygous *NRXN1* deletions, we identify disrupted isoform expression and mutant *NRXN1α* isoforms and characterize their distribution across diverse cell types. These findings help understand the complex alternative splicing of *NRXN1* in adult human brain and hiPSC-derived models and establish a framework for decoding cell type-specific isoform diversity for important genes with relatively low gene expression levels.

INTRODUCTION

Alternative splicing (AS) is a key post-transcriptional process in the human brain that generates diverse mRNA isoforms, contributing to the brain's functional complexity[1–4]. Presynaptic cell-adhesion molecules known as neurexins (*NRXN1-3*) stand out as canonical examples of AS-driven synaptic diversity, playing central roles in synapse formation, specification, and plasticity[5–9]. The *NRXN1* gene generates an exceptional number of isoforms through six major alternative splice sites (SS1-6)[10–12], which are essential for fine-tuning synaptic properties and diversification. For instance, alternative splicing at splice site 4 (SS4) is known to reshape ligand binding[13–16], while differences at SS3 in hippocampal GABAergic interneurons, specifically parvalbumin-positive (*PV*+) and cholecystokinin-positive (*CCK*+) subtypes, have been linked to distinct presynaptic calcium channel usage and modulatory receptors[17,18]. Aberrant alternative splicing due to *NRXN1*$^{+/-}$ deletions disrupt the excitatory–inhibitory balance by weakening excitatory transmission in glutamatergic neurons and enhancing inhibitory transmission in GABAergic neurons[19], underscoring how precise isoform regulation is critical for maintaining synaptic function.

A complete, cell-type-resolved isoform catalog of *NRXN1* remained elusive, particularly due to the technological limitations for detection of the immense number of neurexin isoforms and their extensive cellular diversity in the human brain. Methods such as patch pipette-based qPCR[20] and next generation sequencing[20] can examine various insertions at alternatively spliced segments but fail to assess the critical combinatorial nature of splicing events on a single mRNA molecule

66 and lack the throughput to capture the *NRXN1* splicing patterns in large populations of diverse
67 cells. Long-read platforms, coupled with single-cell RNA barcoding, represent a major
68 development to study splicing of complex genes such as *NRXN1*. However, due to the low
69 expression level of *NRXN1*, these transcriptomic approaches provide insufficient depth to
70 comprehensively assess the large repertoire of *NRXN1* isoforms.

71 To overcome these technical barriers, we developed an integrative sequencing strategy that
72 combines (1) probe-based targeted capture[21,22] with (2) single-cell/single-nucleus RNA isoform
73 sequencing for in-depth *NRXN1* isoform profiling, complemented with (3) an optimized RACE-seq
74 [23,24] (Rapid Amplification of cDNA Ends) based enrichment module to focus on specific mutant
75 isoforms of interest. As the core idea of this integrative strategy, single-cell transcriptomics is used
76 to provide cell-type deconvolution, while targeted capture is used to enrich *NRXN1* transcripts for
77 in-depth analysis; long read sequencing is then employed to achieve the crucial full-length isoform
78 resolution. The use of both probe-based targeting and RACE-seq (PCR-based) enrichment
79 complement each other in recovering the diversity of *NRXN1* isoforms (the strength of probes)
80 and detecting rare or low-abundance mutant isoforms (the strength of PCR). Using this integrative
81 strategy, we mapped comprehensive catalog of cell-type-specific *NRXN1* isoforms across
82 diverse neuronal and glial populations. From the PFC region of the human adult brain, we
83 revealed distinct splicing profiles across GABAergic interneuron, pyramidal neurons and glial
84 lineages. Interneurons derived from the same developmental origin exhibited similar AS patterns,
85 whereas those derived from different origins showed divergent profiles. Analyses of the AS
86 pattern in fetal brain suggest that *NRXN1* splicing profiles are defined in early development and
87 remain largely consistent during neural maturation. Using hiPSC cortical organoids derived from
88 patients with *NRXN1*$^{+/-}$ deletions and controls, we observed that glial progenitor cells and upper-
89 layer excitatory neurons (Ex-L2/3) are enriched for the mutant isoforms, and we found that the
90 organoids partially recapitulate the *NRXN1* splicing patterns of both developmental and adult
91 brains. In the postmortem brain tissue from an autistic patient case carrying a heterozygous
92 *NRXN1* deletion, we found the mutant isoforms transcribed from deletion allele were enriched in
93 molecular layer interneurons (MLI1/2) and astroglia cells, and decreased the *CBLN1*, a *NRXN1*
94 ligand, expression in granule neurons. These findings provide critical insights into the complex
95 alternative splicing of *NRXN1* and establish a framework for decoding cell type-specific isoform
96 diversity for important genes with low gene expression levels.

97

98 # RESULTS

99 **An integrated strategy reveals cell-type catalog of *NRXN1* isoforms across human brain**
100 **and hiPSC cortical o**

101 **rganoids**

102 Single-cell long-read transcriptomics provides a unique ability to resolve full-length *NRXN1*
103 isoforms across cell types, but whole transcriptomic profiling does not provide sufficient depth of
104 *NRXN1* due to its very low expression level. From the published human postmortem brain single-
105 nuclei Iso-seq data[25], *NRXN1* reads constituted only 0.09-0.2% of the total reads across different
106 cell types, underscoring the need for a targeted enrichment strategy (**Fig. 1a, upper panel**). To
107 maximize isoform recovery while preserving cell-type resolution, we used a targeted sequencing
108 strategy with two complementary components (**Fig. 1b**). To generate a comprehensive isoform
109 catalog, we designed capture probes tiled across all *NRXN1* exons at 1× density, which achieved
110 a ~36-96-fold enrichment of *NRXN1* reads over the non-targeted sequencing method (**Fig. 1a,**
111 **lower panel**). To recover cell type-specificity for isoforms of interest, particularly the mutant
112 isoforms transcribed from the allele carrying exonic deletions, we adapted the RACE-seq (Rapid
113 Amplification of cDNA Ends) method from short-read and optimized for long-read sequencing to
114 achieve a more sensitive detection of low-abundance isoforms. To link isoforms to specific cell
115 types, we matched the UMIs and cell barcodes on the *NRXN1* reads (from *NRXN1*-targeted long-
116 read sequencing data) with the UMIs from the cell cluster information by short-read single-cell
117 RNA-seq data (**Methods**). This design leveraged the sequencing depth and accuracy of short
118 reads together with full-length isoform resolution by long reads, enabling precise assignment of
119 complex *NRXN1* isoforms to distinct cell types.

120 We applied this workflow to single-cell or single-nucleus cDNA from multiple sources, including
121 the PFC of adult brains, gestational cortical samples and cerebellar tissue from an $NRXN1^{+/-}$
122 autism patient. In addition, cortical organoids derived from schizophrenia patients harboring
123 unique $NRXN1^{+/-}$ deletions were profiled to assess the impact of these variants on splicing
124 regulation. After isoform identification and quality filtering, using long-read Capture-seq, we
125 identified 50 distinct *NRXN1* isoforms across all datasets, including 27 previously annotated (five
126 α, five β, five γ, and twelve known others) and 23 novel isoforms not represented in Ensembl[26].
127 Expression of these novel isoforms was validated across our in-house datasets and confirmed
128 using independent long-read transcriptomes from postmortem human brains reported by
129 others[25,27] (**Fig. 1c, Supplementary Fig. 1**). Among all the *NRXN1* isoforms, **alpha** isoforms

130 constitute the largest and most structurally complex class. We applied long-read RACE-seq, a
131 more sensitive method to single-cell or single-nucleus cDNA of all samples and identified 103
132 **alpha** isoforms including 57 that has been reported previously. Together, our targeted sequencing
133 strategy will provide a comprehensive underrating of cell-type-specific *NRXN1* alpha isoforms
134 (**Supplementary Fig. 1**).

135

136 ***NRXN1* exhibits distinct splicing patterns across cell types in adult human brains**

137 To understand the *NRXN1* alternative splicing patterns across mature neurons in the human
138 brain, we applied the probe-based long-read Capture-seq on the single-nucleus cDNA from the
139 PFC region of 61 adult donors[28] (**Fig. 2a, b; Supplementary Fig. 2**). We generated 2.43M
140 *NRXN1* reads, among which 1.13M had 10x single cell barcodes, and 1.03M reads sharing the
141 barcodes with short-read data. By integrating the barcode information from short- and long-read
142 data, a total of 125,199 cells were identified across all donors. From the long-read Capture-seq
143 of the single nuclei cDNA pools, we identified 46 isoforms with confident cell barcodes and UMIs
144 matched to short-read data, including 6 alpha isoforms not annotated but reported previously, and
145 14 isoforms not annotated in Ensembl[26] but later validated by our in-house brain data set (**Fig.
146 1c, Supplementary Fig. 1**).

147 The brain comprises a great diversity of cell types, and studying neurons with shared
148 developmental markers offers a means to uncover how *NRXN1* alternative splicing contributes to
149 this cellular and functional diversity. Based on developmental markers[28], neurons were
150 categorized as Neurod2-expressing pyramidal cells[29] (PC), *Lhx6*-expressing interneurons[30]
151 (*Lhx6*$^+$-INs originated from medial ganglionic eminence, MGE), *Prox1*-expressing interneurons[31]
152 (*Prox1*$^+$-INs originated from caudal ganglionic eminence, CGE), and *Lamp5*$^+$ *Lhx6*$^+$ GABAergic
153 neurons (*Prox1*$^+$ labeled as *Lamp5*). *Lhx6*$^+$-INs can be further subclassified into *Pvalb*$^+$ and *Sst*$^+$
154 INs, while *Prox1*$^+$-INs are divided into *Lamp5*$^+$, *Sncg*$^+$, and *Vip*$^+$ INs. GABAergic INs are a highly
155 heterogeneous collection of cell types that are characterized by their unique spatial and temporal
156 capabilities to influence neuronal circuits[32–34]. We observed that GABAergic interneurons showed
157 the greatest *NRXN1* splicing diversity (**Fig. 2c, top**). Splicing patterns were more variable at SS1,
158 SS3, and SS4, whereas SS2 and SS6 were invariably excluded and SS5 was consistently
159 included. Within the *Prox1*$^+$-INs, *the Lamp5*$^+$*, Sncg*$^+$ and *Vip*$^+$ subtypes share similar patterns
160 across all 6 alternative splicing sites, while *Pvalb*$^+$ and *Sst*$^+$, subclasses of *Lhx6*$^+$-INs, differed at

161 SS1. We also identified a *Lamp5*$^{+}$ *Lhx6*$^{+}$ subclass that shared splicing patterns with the *Prox1*$^{+}$-
162 INs and *Sst*$^{+}$ of *Lhx6*$^{+}$- INs in (**Fig. 2c**). In PC neurons (**Fig. 2c, middle**), we found: (1) across
163 cortical layers, exons at SS1, SS3 and SS5 were dominantly spliced in ("in"; upward bars), while
164 exons at ASS6 and SS4 are dominantly spliced out. (2) layer-specific differences at SS2, where
165 exon 7 was often spliced out in PCs of layers L2/3-IT, L5-IT, and L6-IT-Car3, but spliced into most
166 PCs of layers L4-IT, L5/6-NP, L6B, L6-CT, and L6-IT. Among glial cells (astrocytes,
167 oligodendrocyte progenitor cells (OPCs), oligodendrocytes), exon 17 at SS6 was consistently
168 spliced out, whereas exon 23 at SS5 was spliced in. Oligodendrocytes showed spliced-in for
169 exon1 and exon5 at SS1, which were dominantly spliced-out in astrocytes and OPCs. At ASS4,
170 exon 21 was included in astrocytes but not in OPCs or oligodendrocytes, consistent with prior
171 studies showing that astrocytic neurexins splice SS4 at high ratios compared to excitatory
172 neuronal neurexins[35,36]. Building on the striking patterns of *NRXN1* alternative splicing observed
173 in the human brain, we conducted principal component analysis (PCA) of splicing levels across
174 cell types, which separately clustered *Prox1*$^{+}$-INs, *Lhx6*$^{+}$-INs and PCs classes (**Fig. 2d**). Overall,
175 these findings suggest that *NRXN1* splicing profiles exhibit substantial differences across various
176 neuronal populations and glial lineages. Within GABAergic neurons, interneurons derived from
177 the CGE (*Prox1*$^{+}$-INs subtypes) are closely clustered by the splicing patterns, yet distinct from INs
178 originating from the MGE (*Lhx6*$^{+}$-INs subtypes).

179

180 ***NRXN1* splicing profiles remain consistent across cells sharing a neurogenic origin during**
181 **neural maturation**

182 To determine whether the *NRXN1* splicing patterns of adult brain are defined as early as in the
183 human developing neocortex, we generated a snRNA-seq dataset from two unfixed snap-frozen
184 neocortex of postmortem samples obtained from the second and third trimesters of gestation (24
185 and 33 gestational weeks, gw). After quality control, we obtained the transcriptomic data from
186 52,426 nuclei, with a median of 2,910 genes, 6,156 UMIs from each nucleus. We identified 10
187 primary cell types from the single cell data: five neuronal clusters, including excitatory neuronal
188 subpopulations expressing layer-specific cortical marker genes (EX-L2/3, EX-L4/5/6); inhibitory
189 neuronal subpopulations (IN-CGE, IN-MGE and IN-Reelin); and glial cell populations, radial glia
190 and astrocyte (RG/AC) clusters, astrocytes, OPCs, transient glial intermediate progenitor cell
191 (gIPC) population, microglia and *FLT1*+ blood vessel cells (BVCs)[37] (**Fig. 3a**). We further validated
192 our cell annotations by comparing top differentially expressed genes (DEGs) per cell cluster to

193 published canonical lineage and proliferation markers[37,38] (**Fig. 3c**). Next, we analyzed how cell
194 type proportion varied between prenatal neocortex and adult PFC. The 33 gw neocortex sample
195 had more glial, INs-MGE and INs-CGE, whereas the 23 gw sample had more INs-Reelin and
196 excitatory neurons EX-L4/5/6 (**Fig. 3b**). Neuronal populations constituted most cells in the fetal
197 neocortex, whereas glial transient intermediate progenitor cells were exclusively present in fetal
198 neocortex and subsequently disappeared or transformed along with the cortical development in
199 adult PFC, which is consistent with the stages of neurogenesis in human cerebral cortex[39,40]
200 (**Supplementary Fig.3**).

201 To determine *NRXN1* splicing patterns by cell type in the fetal neocortex, we applied the probe-
202 based long-read Capture-seq on the single nuclei cDNA. Based on regional marker genes, exon
203 23 at SS5 was predominantly spliced in across all cell types, whereas exons at other splice sites
204 showed variable inclusion. In most (>60%) PCs across all layers, exons at SS1 and SS3 were
205 consistently spliced in, exon 21 at SS4 was spliced out, and exon 17 at SS6 showed variable
206 usage. In contrast, exon 7 at SS2 was differentially spliced between upper-layer (L2/3) and
207 deeper-layer (L4/5/6) PCs. Similarly, in GABAergic interneurons (INs), exons at SS1 and SS3
208 were largely spliced in, while exons at SS2, SS4, and SS6 were spliced out. Notably, the
209 proportion of IN-*Lhx6*$^{+}$ cells was higher than that of IN-*Prox1*$^{+}$ cells at SS3 (spliced in) and SS4
210 (spliced out). The glial cells were variably spliced at SS1, SS2, SS6 and SS4, but consistently
211 spliced-in at SS3 and SS5 (**Fig. 3d**). We next compared *NRXN1* splicing in matched neuronal
212 populations between prenatal neocortex and adult PFC. Exon inclusion level across 6 splicing
213 sites were highly correlated between prenatal neocortex and adult PFC for Lhx6^{+}-INs (*Pearson's*
214 $r = 0.82$, $p = 0.01$), *Prox1*$^{+}$-INs (*Pearson's* $r = 0.87$, $p = 0.005$), PCs (*Pearson's* $r = 0.87$, $p = 0.005$)
215 and glial populations (OPC, *Pearson's* $r = 0.9$, $p = 0.003$; RG/AC, *Pearson's* $r = 0.82$, $p = 0.01$)
216 (**Fig. 3e, Supplementary Fig. 3**). These findings suggest that *NRXN1* splicing profiles are defined
217 in early development and remain largely consistent during neural maturation.

218

219 **_NRXN1_$^{+/-}$ hCOs splicing profiles recapitulate developmental and mature features and**
220 **identify cell types expressing the mutant isoforms.**

221 From hiPSC-derived 2D neurons with a 3′-*NRXN1* deletion, we previously uncovered dozens of
222 mutant isoforms exclusively from the deletion allele (mainly affects alpha and beta isoforms)[12].
223 Later, we demonstrated that this exonic deletions at the 3' site disrupted normal *NRXN1* splicing
224 and isoform expression, leading to distinct phenotypes in iGLUT and iGABA neurons, primarily

225 reflected in altered synaptic frequency[19,41]. Building on these observations, we sought to explore
226 how the deletions affect the splicing events across cell types in a more complex
227 neurodevelopmental system. We generated human cortical organoids (hCOs) from the control
228 line and schizophrenia probands with the 3'deletion (3'-del) in parallel. hCOs were guided toward
229 the pallium of dorsal forebrain[42] (**Methods, Supplementary Fig. 4**). To profile cell types in the
230 hCOs, we performed single cell RNA-seq on hCOs from two 3'-del cases (line 581 and line 641)
231 and one control line (line 690) at day 150 with the 10x NextGEM 5pv2 kit, when neural activity is
232 well characterized[43–45]. After quality control data processing steps, we merged high-quality
233 scRNA-seq data from 3 cortical organoid samples into a core dataset of 27,325 cells. Following
234 unsupervised clustering, we identified and annotated major cell types using an extensive curated
235 list of known marker genes[19,42], which consisted of transit-amplifying cell (TAC), neuronal
236 intermediate progenitor cells (nIPCs) giving rise to distinct subpopulations of glutamatergic
237 excitatory neurons (Ex-L2/3, Ex-L4/5/6) and GABAergic inhibitory neurons (INs, IN-SST,IN-
238 Reelin); gIPCs leading to the subpopulations of radial glia and astrocyte clusters (RG/AC); as well
239 as BVCs (**Fig. 4a**). We further validated our cell annotations by comparing gene expression
240 signatures of our scRNAseq cell clusters to published brain organoid datasets as reference (**Fig.
241 4b**). We then compared cortical organoids derived from control iPSC line with the prenatal
242 neocortex in terms of cell-type composition. We found gIPCs (OR = 2.06, *Fisher's test*) and Ex-
243 L2/3 (OR = 1.67, *Fisher's test*) were significantly enriched in hCOs from the 3'del lines than in
244 control lines, while cell number of IN-Reelin were significantly low in 3'-del lines (OR = 0.03,
245 *Fisher's test*) (**Fig. 4c**).

246 To more comprehensively understand the greater splicing patterns of *NRXN1* across the cortical
247 organoid cell types, we applied the targeted Capture-seq on the full-length single-cell cDNA. From
248 the control lines and those with the 3' deletion, we generated 317,521 *NRXN1* reads, of which
249 144,450 with 10x single cell barcodes and 120,873 reads sharing the barcodes with short-read
250 data. Integrating the barcode information across short- and long-read data, we identified 23,187
251 cells from all cortical organoid samples (**Supplementary Fig. 4**). Using these targeted single cell
252 data, we first analyzed the *NRXN1* AS patterns in excitatory neurons, inhibitory neurons and non-
253 neuronal cells for both 3'del lines and control lines at the advanced maturation stage. Within the
254 glutamatergic excitatory neurons (the major population in hCOs), *NRXN1* splicing patterns were
255 highly similar between control lines and 3'del lines, except for differences in exon inclusion level
256 at SS1 (exon 3) and SS3 (exon 12) (**Fig. 4d,e**). The ~136-kb deletion impacting SS4 and SS5
257 (exons 21-23)[12,19] slightly reduced exon inclusion at SS4 in the 3'del lines (**Fig. 4e**).

258 As the deletion in the 3'-region of *NRXN1* generates mutant alpha and mutant beta isoforms, we
259 sought to know which cell types are enriched for these mutant isoforms. As some of the mutant
260 isoforms are low in abundance and the Capture-seq works better with abundant isoforms,–we
261 employed long-read 5'RACE-seq to resolve alpha isoforms with single cell barcode information
262 from the 5'GEX single cell cDNA (Methods). We identified 21 α-isoforms across all cortical
263 organoids, of which 18 were not detected by long-read Capture-seq, including six mutant isoforms
264 from the two 3′-del *NRXN1*$^{+/-}$ organoids (**Fig. 4f**). In addition, we identified seven β-isoforms,
265 including one mutant β-isoform (**Supplementary Fig. 4**). Compared with the control lines, the
266 mutant α-isoforms were expressed widely across all cell types in both lines with 3′-*NRXN1*$^{+/-}$
267 (except for *DLX1/2*-INs in 641). In contrast, the mutant β-isoform was detected in fewer cell
268 clusters, showing consistent expression in excitatory neurons in both 3′-*NRXN1*$^{+/-}$ lines, and
269 exclusively in TACs in line 581 and RG/AC cells in line 641 (**Fig. 4g**). The composition of *NRXN1*
270 isoform by cell type demonstrated that mutant isoforms from the allele bearing 3'-deletion are
271 expressed in all cell types, contributing to altered synaptic activity in excitatory neurons and
272 *DLX1/2*-INs.

273 With *NRXN1* splicing profiles available across multiple developmental stages, we next asked
274 which stage the control hCOs most closely resemble in terms of *NRXN1* splicing– the fetal
275 neocortex or the adult PFC. We therefore compared the exon inclusion level across splicing sites
276 (SS1-6) for *NRXN1* in three major cell classes that appear in all samples, *i.e.,* GABAergic
277 inhibitory neurons, excitatory neurons, and RG/AC (**Fig. 4h**). In GABAergic neurons, exon
278 inclusion levels were highly correlated between organoids and the fetal neocortex (*Pearson's r =*
279 *0.97, p = 3.82E-04*) and between cortical organoids and the adult PFC (*Pearson's r = 0.88, p =*
280 *0.008*), although exons at SS2, SS3, and SS5 displayed differential splicing. Similarly, significant
281 agreement was observed between cortical organoids and the fetal neocortex in glutamatergic
282 excitatory (Glu-EX; *Pearson's r = 0.73, p = 0.04*) and radial glia/astrocyte (RG/AC; *Pearson's r =*
283 *0.78, p = 0.04*) populations, and between organoids and the adult PFC in Glu-EX (*Pearson's r =*
284 *0.88, p = 0.004*) and RG/AC (*Pearson's r = 0.92, p = 0.003*) populations (**Supplementary Fig. 4**).
285 Together, these findings indicate that cortical organoid splicing profiles partially recapitulate both
286 developmental and mature splicing patterns, with a subset of splice sites showing cell type- and
287 stage-specific divergence (**Fig.4h, Supplementary Fig. 4**).

288

289 ***Exonic* deletion impacts *NRXN1* splicing in cerebellum from an ASD brain**

290 *NRXN1* deletions are also strong risk factors for autism spectrum disorder (ASD)[46,47]. The
291 cerebellum shows consistent structural and functional alterations in ASD, and *NRXN1α* isoform
292 complexity is significantly different in the cerebellum compared to cortex. To uncover how *NRXN1*
293 isoform diversity contributes to ASD, we applied our probe-based long-read Capture-seq to the
294 postmortem cerebellum of an *NRXN1*$^{+/-}$ ASD patient with an exon 9-13 deletion (**Supplementary**
295 **Fig. 5**). We performed single-nucleus RNA-seq on this sample and obtained 9,689 high-quality
296 nuclei. Unsupervised clustering and annotation identified eight cerebellar cell classes: excitatory
297 granule cells; inhibitory neurons (Purkinje cells; two types of molecular layer interneurons MLI1,
298 MLI2; and Purkinje-layer interneurons); astrocytes; Bergmann glia; oligodendrocytes; and OPCs
299 (**Fig. 5a**). These annotations matched expected marker gene expression[48] (**Fig. 5b**).

300 When applying the probe-based long-read Capture-seq to the single nuclei cDNA of cerebellum,
301 we identified 9 *NRXN1* isoforms with cell barcodes matched to short-read data, including two
302 mutant isoforms associated with the exon 9-13 deletion (**Fig. 5c)**. Among all the *NRXN1* reads
303 with cell barcodes matched to short-read data, the mutant isoform transcript 3312 and transcript
304 4131 accounted for ~35% of reads (**Fig. 5d**). When assigning isoforms to cell types, we found the
305 two mutant isoforms were significantly enriched in MLI1 (*Fisher's Exact test*, *p* =5.72E-18), MLI2
306 (Fisher's Extact *test*, *p* =0.04), and astroglia cells (*Fisher's Extact test*, *p* =51.49E-13) compared
307 to wild-type *NRXN1* isoforms (**Fig. 5e**). Next, we examined *NRXN1* splicing at six alternative
308 splicing sites in these cell types. We found that: (1) over 70% of granule cells are spliced-in at
309 SS1, SS3, SS4 and SS5, while more than half spliced-in at SS2; granule cells predominantly
310 spliced out at SS6; (2) The two MLI subtypes showed different splicing trends at SS1, SS2 and
311 SS5; (3) within the glia lineage, OPC and oligodendrocytes are uniformly spliced-in at SS5 but
312 spliced-out at ASS6. As the exonic deletion expanded the SS3 (exon12) and produced mutant
313 alpha isoforms, we asked whether mutant isoform expression correlated with synaptic ligands.
314 Among granule neurons, cells expressing mutant-type *NRXN1* isoforms showed significantly
315 decreased expression of cerebellin genes than cells expressing wildtype isoforms. Because
316 granule cells regulate Purkinje cells via the *CBLN1–NRXN1* complex that organizes and
317 maintains Purkinje cell synapses[49], our results link *NRXN1* deletions to abnormal Purkinje cell
318 signaling and highlight *NRXN1*'s functional role in ASD.

319

320 **DISCUSSION**

321 We applied an integrative isoform sequencing strategy to generate a cell-type-resolved catalog
322 of *NRXN1* isoform diversity across human cortical organoids, prenatal cortex, adult PFC, and
323 cerebellum samples. *NRXN1* demonstrated extensive cell type- and cell subtype specificity. For
324 example, in hCO organoids and adult PFC, *NRXN1* splicing displayed apparent cell-type
325 specificity, while the cerebellum of an ASD case exhibited distinct deletion-associated isoforms
326 enriched in molecular-layer interneurons and astroglia. Together, these results establish a
327 framework for decoding *NRXN1* alternative splicing.

328 Our integrative pipeline, which combines probe-based capture, single-cell barcoding, and long-
329 read sequencing, overcomes major obstacles in detecting low-abundance, full-length *NRXN1*
330 transcripts. Targeted enrichment increased on-target recovery by ~50-fold on average compared
331 with untargeted long-read methods, while preserving single-cell resolution. The workflow links full-
332 length isoforms to their cell identity, enabling quantitative mapping of *NRXN1* splicing patterns
333 across neuronal and glial populations, linking the genetic variations with functional phenotypes.
334 RACE-seq and Capture-seq represent complementary strategies for characterizing transcript
335 isoforms. Capture-seq employs probe-based hybridization to enrich targeted transcripts from
336 barcoded cDNA, enabling detection of novel isoforms (alternative transcription start sites/
337 transcription termination sites) for many gene targets in one assay across cell types. This method
338 preserves relative transcript abundance and is compatible with single-cell barcoding[50], making it
339 well-suited for analyzing splicing events at defined sites. However, Capture-seq preferentially
340 detects abundant isoforms and shows reduced sensitivity for low-abundance or long α-isoforms.
341 In contrast, long-read RACE-seq selectively amplifies transcripts from defined 5′ or 3′ ends,
342 allowing accurate reconstruction of specific full-length isoforms, including those spanning deletion
343 junctions in mutant alleles. Although long-read RACE-seq is less quantitative and lower
344 throughput, it offers greater sensitivity and completeness for targeted transcripts, while being
345 more susceptible to RNA degradation. Together, these approaches provide a comprehensive
346 view of transcript diversity: long-read Capture-seq defining global splicing architecture and long-
347 read RACE-seq validating and extending isoform-level resolution in selected contexts, though
348 both are limited by RNA integrity and read length, particularly in postmortem brain. This approach
349 is applicable to other synaptic genes with extensive splicing, bridging the gap between single-cell
350 transcriptomics and isoform-level resolution.

351 As the dominant population in the cortex, glutamatergic excitatory neurons demonstrated
352 remarkable concordance of *NRXN1* splicing patterns between prenatal neocortex and adult PFC.
353 Likewise, glutamatergic neurons from iPSC-derived organoids and prenatal cortex shared similar

354 *NRXN1* splicing, indicating that patient-specific *NRXN1*$^{+/-}$ organoids captured key developmental
355 splicing events. GABAergic interneurons in the PFC show extensive molecular and functional
356 diversity, reflected in their distinct *NRXN1* splicing profiles. Although *Pvalb*-expressing and *Sst*-
357 expressing interneurons both originate from the MGE (*Lhx6*+), they exhibited subtype-specific
358 exon usage patterns, consistent with their previously observed phenotypic differences[51,52].
359 Notably, *NRXN1* splicing programs appear to be established early in development and remain
360 largely stable, as evidenced by the strong concordance of exon inclusion levels in *Prox1*$^{+}$
361 interneurons between the prenatal neocortex and adult PFC[20].

362 Extending this analysis to an *NRXN1*$^{+/-}$ cerebellum from an ASD patient revealed unique mutant
363 isoforms (exon 9-13 deletion) that were highly expressed in molecular-layer interneurons and
364 astroglia, whereas granule cells retained WT isoforms. Cells expressing mutant *NRXN1* showed
365 decreased cerebellin ligand expression, implicating the mutant isoforms in disrupted synaptic
366 adhesion. These findings highlight that *NRXN1* deletions can differentially impact cortical and
367 cerebellar circuits, aligning with the cerebellar dysfunction frequently reported in ASD[53–55]. In the
368 cerebellum, granule cells influence Purkinje cells through the *Cbln1-NRXN1* complex, which
369 organizes and maintains the parallel fiber-Purkinje cell synapse. This trans-synaptic pathway is
370 essential for forming functional excitatory synapses and inducing long-term depression, a key
371 form of synaptic plasticity. Our findings suggest that disrupted Purkinje cell signaling may be
372 linked to *NRXN1* deletions. These findings reveal that *NRXN1* deletions disrupt cerebellar and
373 cortical circuits through precise, cell-type–specific mechanisms, paralleling the dysfunction
374 characteristic of ASD and schizophrenia. Given that *NRXN1* isoforms dictate partner specificity
375 at synapses, the isoform catalog provides a foundation for isoform-targeted interventions, such
376 as antisense oligonucleotides, to selectively restore normal *NRXN1* splicing and rescue circuit
377 function.

378

379 # FIGURE CAPTION

380 **Figure1. A targeting strategy reveals cell-type catalog of *NRXN1* isoforms.**

381 (a) *Upper panel* Bar graph showing the percentage of *NRXN1* reads detected by Capture-seq
382 (orange) and non-targeted iso-seq (blue) across major brain cell classes. GABA, GABAergic
383 neurons; Glu, glutamatergic neurons; Astro, astrocytes; Micro, microglia; Oligo,
384 oligodendrocytes; OPC, oligodendrocyte progenitor cells. *Lower panel:* Fold change of *NRXN1*-
385 mapped reads obtained by Capture-seq relative to the non-targeted single-cell Iso-seq.

(b) Schematic illustrating the comprehensive characterization of cell-type-specific *NRXN1* isoforms using the targeting strategy. Full-length single-cell/nuclei cDNA is divided into two parts: one processed for a short-read library using Tn5 transposase and sequenced to identify cell types; the other subjected to probe-based long-read Capture-seq and RACE-seq to detect *NRXN1* isoforms with matched barcodes and UMIs (see Methods).
(c) Schematic representation of *NRXN1* isoforms identified across samples by single-cell long-read Capture-seq. Each row represents a unique isoform, color-coded as α, α-mutant, β, γ, known others (other), and novel isoforms (other-novel). The heatmap displays the expression of each isoform in adult prefrontal cortex, fetal neocortex, and human cortical organoids.

Figure 2. *NRXN1* expression repertoires exhibit distinct splicing patterns across cell types in human brains.

(a) Workflow overview of long-read Capture-seq applied to single-nucleus cDNA from adult PFC.
(b) UMAP visualization of putative cell types in the adult PFC. Colored clusters represent neuronal and glial subtypes.
(c) Bar plots showing normalized *NRXN1* splicing events at SS1–SS6 in GABAergic interneurons (top), glutamatergic excitatory neurons (middle), and glial lineages (bottom). Each bar represents a cell class. Within GABAergic interneurons, *Prox1*$^{+}$-INs are colored in green, *Lhx6*$^{+}$-INs in red, and *Lamp5*$^{+}$ *Lhx6*$^{+}$-INs in orange. Exon inclusion ("splice-in") is shown upward and exclusion ("splice-out") downward. Data represent mean ± SEM.
(d) PCA plot based on *NRXN1* splicing patterns across splice sites SS1–SS6, showing that *Prox1*$^{+}$-INs, *Lhx6*$^{+}$-INs, and PCs form distinct, separately clustered classes. Colors correspond to cell types as in (c): *Prox1*$^{+}$- INs (green), *Lhx6*$^{+}$-INs (red), *Lamp5*$^{+}$ *Lhx6*$^{+}$-GABAergic neurons (orange), glutamatergic neurons (blue), and glial lineages (gray).

Figure 3. *NRXN1* splicing profiles remain consistent across cells sharing a neurogenic origin during neural maturation.

(a)UMAP visualization of putative cell types in the neocortex from two fetal brains. *Left:* merged UMAP of combined datasets; *Right:* separate UMAPs showing neocortex samples from 23 gw and 33 gw. Colored clusters represent neuronal and glial subtypes in the fetal neocortex. Abbreviations: gIPC, glutamatergic intermediate progenitor cells; RG/AC, radial glia/astrocyte lineage; Ex, excitatory neurons; IN, inhibitory neurons; BVC, blood vessel cells; MG, microglia.
(b) Bar graph showing the proportions of each cell class in 23 gw and 33 gw neocortex samples. Colored blocks indicate neuronal and glial subtypes.
(c) Heatmap of marker genes expression used for cell-type annotation. Each row represents a marker gene and each column a cell class; darker red indicates higher expression. Colored blocks indicate neuronal and glial subtypes in fetal neocortex.
(d) Bar plots showing normalized *NRXN1* splicing events at SS1–SS6 in *Prox1*$^{+}$-INs (green), *Lhx6*$^{+}$-INs (red), glutamatergic excitatory neurons (blue), and glial lineages (gray). Exon inclusion ("splice-in") is shown upward and exclusion ("splice-out") downward. Data represent

427 mean ± SEM.
428 (e) Pearson's correlation of *NRXN1* exon inclusion levels across splicing sites (SS1–SS6)
429 between adult PFC and prenatal neocortex in *Lhx6*$^{+}$-INs (left), *Prox1*$^{+}$-INs(middle), and PC
430 neurons (right).
431

432 **Figure 4. *NRXN1*$^{+/-}$ human cortical organoids (hCOs) splicing profiles recapitulate**
433 **developmental and mature features and identify cell types expressing the mutant**
434 **isoforms.**

435 (a) *Left:* UMAP of putative cell types in hCOs, with colored clusters representing neuronal and
436 glial subtypes. *Right:* Proportions of each cell class in organoids derived from control line 690 or
437 3′del lines (581 and 641). Abbreviations: gIPC, glutamatergic intermediate progenitor cells;
438 nIPC, neuronal intermediate progenitor cells; TAC, transient amplifying cells; RG/AC, radial
439 glia/astrocyte lineage; Ex, excitatory neurons; IN, inhibitory neurons; BVC, blood vessel cells.
440 (b) Heatmap of marker gene expression used for cell-type annotation. Each row represents a
441 marker gene and each column a cell type; darker red indicates higher expression.
442 (c) Bar graph showing enrichment of specific cell types in 3′del organoids compared with control
443 lines.
444 (d) Bar plots showing normalized *NRXN1* splicing events at SS1–SS6 in all cell classes from
445 control hCOs (upper) and 3′del hCOs (lower). Exon inclusion is shown upward ("splice-in"),
446 exclusion downward ("splice-out"). Data represent mean ± SEM.
447 (e) Pearson's correlation of *NRXN1* exon inclusion levels across splicing sites (SS1–SS6)
448 between 3′del and control hCOs.
449 (f) *Left:* Schematic of *NRXN1α* isoform structures identified by long-read RACE-seq. Each row
450 represents a unique *NRXN1α* isoform, with blue exons indicating inclusion and blanks indicating
451 exclusion. *Right:* Presence or absence of each *NRXN1α* isoform across 3′del and control hCOs
452 (green = present; blank = absent).
453 (g) Bar graph showing the composition of *NRXN1* isoforms identified by long-read RACE-seq
454 within each cell type across 3′del and control hCOs.
455 (h) Bar graph showing exon inclusion levels at SS1–SS6 in GABAergic INs (top), glutamatergic
456 excitatory neurons (middle), and RG/AC (bottom) across hCOs (blue), fetal neocortex (orange),
457 and adult PFC (green). Statistical comparisons were performed using one-way ANOVA (*, $P <$
458 0.05; **, $P < 0.01$; ***, $P < 0.001$).Data represent mean ± SEM.
459

460 **Figure 5. Patient-specific exonic deletion impacts *NRXN1* splicing patterns across cell**
461 **types from an ASD cerebellum.**

462 (a) *Left:* UMAP of putative cell types in the cerebellum. *Right:* Proportions of each cell class in
463 the cerebellum.
464 (b) Dot plot showing expression of marker genes used for cell-type annotation. Each row
465 represents a marker gene; cell types are shown along the x-axis.
466 (c) Schematic of *NRXN1* isoform structures identified by long-read capture-seq, with each row
467 representing a unique *NRXN1* isoform. The 3′-*NRXN1*$^{+/-}$ specific mutant isoforms are highlighted
468 in red.

469 (d) Bar graph showing read counts of *NRXN1* isoforms with single-cell barcodes and UMIs by
470 long-read Capture-seq. The red box highlights 3′-*NRXN1*$^{+/-}$ specific mutant isoforms.
471 (e)Bar graphs show the proportion of cells in each cell class expressing any NRXN1 isoforms,
472 wild-type isoforms(*NRXN1*-WT) and mutant isoforms (*NRXN1*-MT).
473 (f) Bar plots showing normalized *NRXN1* splicing events at SS1–SS6 across all cell classes
474 from *NRXN1*$^{+/-}$ cerebellum. Exon inclusion is shown upward ("splice-in") and exclusion
475 downward ("splice-out"). Data represent mean ± SEM.
476 (g) Violin plots showing average expression of *NRXN1*-binding ligands in granule cells
477 expressing wild-type (WT-expressing) versus mutant *NRXN1*(MT-expressing).

478

479 # METHODS

480 **Data Availability**

481 All sequencing data generated for this study will be uploaded to the Sequence Read Archive
482 and made publicly available.

483 **Acquisition and processing of human post-mortem samples**

484 **The MSSM cohort single nuclei cDNA.** The sequencing data and full length single-nucleus
485 cDNA libraries of 61 individuals from the Mount Sinai School of Medicine cohort were generated
486 from the previous publication[28]. The leftover barcoded full length single-nucleus cDNA was used
487 as input for the current study. To obtain 1-1.5 µg of cDNA for the hybridization step, 5-10 ng of
488 the surplus barcoded 10x cDNA was amplified in two 50 µl reactions, which consists of 25 µL
489 2X PrimeSTAR GXL premix (Takara, Cat# R051A), 1 µL partial read 1 primer (10 µM, 5'-
490 CTACACGACGCTCTTCCGATCT-3'), 1 µL partial TSO primer (10 µM, 5'-
491 AAGCAGTGGTATCAACGCAGAG-3'). The amplification program is set as follows: 98 °C for 30
492 seconds, 8 cycles of 98 °C for 10 seconds, 65 °C for 15 seconds and 68 °C for 8 minutes, 1
493 cycle of 68 °C for 10 minutes. The concentration of PCR products was quantified by Qubit, and
494 additional cycles were added if the yield was lower than 500 ng per reaction. The PCR products
495 were cleaned up with 0.6× AMPure XP beads.

496 ***NRXN1*$^{+/-}$ cerebellum of human post-mortem brain.**

497 Postmortem human cerebellum of the ASD case with a heterozygous *NRXN1*$^{+/-}$ mutation were
498 collected and consented through the Department of Pathology at Western Michigan University
499 Homer Stryker MD School of Medicine under WCG IRB protocol #20111080 and the other
500 samples were consented through the National Institute of Mental Health Intramural Research
501 Program under NIH protocol #90-M-0142, and was acquired by LIBD via material transfer
502 agreement.

503 **Fetal brain.** Two human gestation cortical tissue samples(24 gw and 33 gw) were obtained
504 from previous study [37] in accordance with the policies and regulations at the Icahn School of
505 Medicine at Mount Sinai and its Institutional Review Board.

506 **Single cell/nucleus full-length barcode cDNA**. 60mg of the frozen post-mortem tissues was
507 submitted to Yale Center for Genome Analysis for single nuclei isolation. Briefly, single nuclei
508 were automatically extracted from the frozen tissue using the Singulator system, which includes
509 washing nuclei suspension in buffer supplemented with sucrose, and nuclei FACS sorting. The
510 isolated nuclei were then processed using the 10x Genomics 5'GEX V2 to create single nuclei
511 cDNA and matched short-read RNA-seq libraries per manufacturer's instructions.

512

513 **Patient-derived organoids**

514 Human induced pluripotent stem cell (hiPSC) lines used were validated using methods as
515 previously described. Genome-wide SNP genotyping was performed using Illumina genome-wide
516 GSAMD-24v2-0 SNP microarray at the Children's Hospital of Philadelphia. Cultures were tested
517 for and maintained Myco-plasma free. A total of 3 control hiPS cell lines derived from fibroblasts
518 collected from 3 healthy subjects were used for experiments. Approval for this study was obtained
519 from the Stanford IRB panel, and informed consent was obtained from all subjects.

520 **Stem cell maintenance**. Human iPSCs were obtained as previously described in Flaherty et
521 al., 2019. Under sterile conditions, 6-well plates (Corning 3516) were coated with 60 ug/mL
522 Geltrex matrix (Thermo Fisher A1413302) diluted in Dulbecco's Modified Eagle Medium
523 (DMEM, Gibco 11965-092) per well and incubating them for at least 30 min at 37 °C. All hiPSC
524 lines were seeded on coated plates and maintained in StemFlex (StemFlexTM Basal Medium
525 (Gibco A3349301) supplemented with StemFlexTM Supplement (Gibco A3349201)) at 37 °C.
526 After 4 days, hiPSC lines were passaged by removing media washing cells with PBS. Then cells
527 were incubated in 0.5 µM EDTA in PBS (Invitrogen AM9261) for 4 min at 37 °C. After removing
528 EDTA, colonies were broken into smaller pieces by triturating using a p1000 pipette and
529 StemFlex. It's worth noting that very small colonies have a very low survival chance and should
530 be supplemented with CHR. After dissociation, colonies were distributed in new coated plates
531 containing StemFlex in a ratio of 1:6 (250K cell per well of a 6 well plate). A complete media
532 change was performed the following day. Complete media changes were performed as needed
533 or every other day.

534 **Cortical organoid generation**. Cultures of hiPSC were dissociated at 70% confluency and
535 checked for signs of differentiation. Mature undifferentiated cells were washed with PBS,
536 followed by incubation with Accutase (Thermo Fisher Scientific) and chroman (CHR) for 10-20
537 min to achieve complete detachment. Cells were gently triturated to achieve a single-cell
538 suspension. The resulting cell suspension was transferred into DMEM (Thermo Fisher
539 Scientific) and centrifuged at 500 × g for 4 min. Cells were resuspended in StemFlex™ medium
540 (Thermo Fisher Scientific) supplemented with CHR and diluted at a density of 1.5 × 10^6
541 cells/mL. AggreWell™ plates (STEMCELL Technologies) were prepared with anti-adherent
542 solution according to the manufacturer's instructions. A volume of 2 mL of the hiPSC
543 suspension was seeded per well to promote spheroid formation overnight (defined as day 0).
544 Neural induction was subsequently performed following the human forebrain spheroid protocol
545 described by Sloan et al. (2018). In brief, medium was replaced daily until day 15, every other
546 day from days 15–42, and every 4 days from day 43 onward. On day 1, embryoid bodies were
547 dislodged and transferred to ultra-low adhesion 10 cm dishes containing Neuronal Induction
548 Media (StemFlex™ medium without StemFlex™ supplement), SMAD inhibitors SB431542 (100
549 nM; Tocris), LDN193189 (100 nM; Stemgent), XAV (100nM; Stemcell technology). Plates were
550 placed on an orbital shaker at 53 rpm and maintained at the same speed until day 6. From days
551 6–25, Neuronal Differentiation Media (Neurobasal-A (Thermo Fisher Scientific), 1x Antibiotic-

552 Antimycotic, 1x GlutaMAX™, and 1x B27 supplement minus vitamin A (Thermo Fisher
553 Scientific) was used supplemented with 100 nM Fibroblast Growth Factor (FGF) and 100 nM
554 Epidermal Growth Factor (EGF). Between days 25–42, Neuronal Differentiation Media was
555 supplemented with brain-derived neurotrophic factor (BDNF; 100 nM; PeproTech) and
556 neurotrophin-3 (NT3; 100 nM; PeproTech). From day 42 onwards, organoids were maintained in
557 Neuronal Differentiation Media.

558 **Cryoprotection and immunocytochemistry**. Organoids were washed with PBS and fixed in 4%
559 paraformaldehyde (PFA)/phosphate buffered saline (PBS) overnight at 4°C. Organoids were then
560 washed in PBS and transferred to 30% sucrose/PBS for 48 hours. Organoids were embedded in
561 gelatin/sucrose embedding solution by dissolving 7.5 g gelatin in 100 ml 10% sucrose solution at
562 37°C to obtain a 7.5% gelatin/10%sucrose embedding solution. For immunofluorescence
563 staining, 14-16 µm-thick sections were washed with PBS and blocked in 10% donkey serum, and
564 0.1% Triton-X in PBS for 1 h at room temperature followed by primary (overnight, 4 °C) and
565 secondary antibody (1–2 h, Jackson ImmunoResearch or ThermoFisher Scientific) incubation.
566 Slides were then mounted (VECTASHIELD, Vector Labs) and imaged on a Zeiss microscope with
567 an apotome module quipped with ZEN 3.3 (ZEN pro) software. Antibody list: BCL11B/CTIP2 (rat,
568 1:500, Abcam), FOXG1 (rabbit, 1:200, Abcam), GAD1/GAD67 (mouse, 1:1000, Millipore), HuC/D
569 (mouse, 1:500, Invitrogen), KI67 (rabbit, 1:500, Vector Labs), PAX6 (mouse, 1:200, BD
570 Bioscience), RELN (mouse, 1:100, MBL), SOX1 (goat, 1:20, R&D Systems), TBR1 (rabbit,
571 1:1000, Abcam).

572 The sections were then incubated overnight at 4°C with primary antibodies diluted in 1% BSA,
573 0.1% Triton-X, sodium azide. Cryosections were washed three times and were incubated with
574 secondary antibodies in 10% donkey serum, 0.1% Tri- ton-X. Aquamount (Polysciences, 18606)
575 was used to mount cover glass slides. Imaging was performed on a Leica TCS SP8 confocal
576 microscope or Zeiss upright fluorescent microscope. Images were processed and analyzed
577 using Fiji (NIH) and IMARIS (Oxford Instruments). To quantify percentage of oligodendrocyte
578 precursor cells, images of Hoechst and OLIG2+ nuclei were taken. Hoechst images were
579 converted into 8-bit images, and we performed automatic image threshold (Default, Dark
580 Background) into a binary image, converted to Mask, performed Watershed (FiJi, NIH). OLIG2
581 nuclei images were manually thresholded due to inconsistency of background signals by a
582 researcher blinded to molecular condition. Nucleus from Hoechst and OLIG2 images were
583 quantified using Analyze Particles with pixel size 30-1000, and the number of OLIG2 nuclei
584 counts were di- vided by Hoechst nuclei counts to obtain a percentage. Ventricular-like regions
585 were quantified by examining cross sections of organoids stained with Hoechst and counting
586 clearly identifiable ventricle-like structures.

587 **Single cell dissociation and RNA-seq library preparation.** Organoids with optimal morphology
588 from each line were selected in days 25, 60, and 120. Five organoids from each line were pooled
589 at each time point and transferred to a 5 cm plate dish and washed with PBS. Organoids were
590 chopped into small pieces using a sharp sterile blade. 300 ul of Accumacs was added on top of
591 the pieces without scattering them. Pieces were collected by a cut tip and were transferred into a
592 1.5 ml tube. The tubes were incubated at 37°C and 800 g rotation on a thermal shaker for 20 min.
593 After 20 min, cell suspension was very gently pipetted up to 5 times with a p 1000 tip, to facilitate
594 the cell dissociation. 500 ul fresh Accumacs was added into the suspension and incubated again
595 for 20 min. Cell suspension was pipetted up to 5 times more and passed through a filter in a 15
596 ml sterile tube containing 1 ml of HEPES buffer. The filter was washed with 1 ml of HEPES buffer.
597 The single cell suspension in a 15 ml tube was centrifuged at 300 g for 4 min to collect cells in a
598 pellet. After removing the supernatant, the pellet was resuspended in 200 ul HEPES buffer and
599 cells were counted using a hemocytometer. The suspension was diluted to reach 1200 cell per

600 ml and was submitted to Yale Genomics Core facility for cDNA library prep using 10x 5' NEXT
601 GEM V2 according to the manufacturer's instructions.

602 **NGS sequencing of single cell/ nucleus cDNA**

603 The prepared 10X Genomics single cell/nucleus libraries were submitted to Yale Center for
604 Genome Analysis and sequenced on an Illumina NovaSeq sequencer with 150-bp paired-end
605 sequencing on an S4 flow cell. For each sample, ~10,000 cells were sequenced at a minimum
606 depth of 50,000 reads per cell.

607 **Probe-panel design**

608 IDT Lockdown probes (Integrated DNA Technologies) are 120-nt long oligos that are
609 biotinylated at their 5' ends and were designed using the xGen Panel Design Tool
610 (https://www.idtdna.com/pages/tools/xgen-hyb-panel-design-tool) to ensure robust coverage of
611 the targeted regions. Briefly, a customized *fasta* file containing all Ensembl[26]-annotated exons
612 (across all isoforms) of target genes, including UTRs was uploaded as input file. The probe
613 sequence list was generated at the 1x probe tilling density using "Homo sapiens (Human) NCBI
614 GRCh37.p13 (hg19)" as the reference (Supplementary table 1).

615 The probe panel was further synthesized employing TEQUILA-seq[21]. First, the probe panel pool
616 of 150-nt long oligos, each comprising a 3' end 30-nt universal primer binding sequence (5'-
617 CGAAGAGCCCTATAGTGAGTCGTATTAGAA-3') followed by a 120-nucleotide target-specific
618 probe, was synthesized with Twist Bioscience (Supplementary table 1). Then the oligo pools
619 were amplified and biotin-labeled using nickase-induced linear strand displacement
620 amplification as described previously. Finally, synthesized probes were purified with 1.8×
621 volumes of AMPure XP beads and quantified by NanoDrop 2000 Spectrophotometer.

622 **Capture-seq on single cell/ nucleus cDNA**

623 All hybridization and capture steps were performed following the IDT "Hybridization capture of
624 DNA libraries using xGen Lockdown probes and reagents" protocol. Briefly, 1-1.5 µg of
625 amplified 10x cDNA together with 1 µL of each blocking oligos (1 mM, Supplementary Table1)
626 were dried using a speed vac (Thermo, Cat# SPD140P1). Next, the dried sample were re-
627 hydrated with 8.5 µL of 2× Hybridization Buffer, 2.7 µl Hybridization Buffer Enhancer (xGen IDT
628 1080577), and 3.8 µL nuclease-free water, followed by denaturation at 95 °C for 10 min. The
629 denatured mixture was then incubated with 200 ng of TEQUILA probes at 65 °C for 16 h. Next,
630 the capture step was performed by adding 100 µL of washed M-270 streptavidin beads
631 (Invitrogen, #65306) to the mixture, followed by incubation at 65 °C for 45 minutes with vortexing
632 every 10 minutes to ensure thorough mixing. After capture, the mixture was then immediately
633 washed with high-temperature and room temperature buffers and resuspended in in 40 µL
634 nuclear-free water.

635 The streptavidin bead-captured cDNA was amplified using PrimeSTAR GXL premix (Takara,
636 Cat# R051A) and 10x cDNA amplification primers by incubating at 95 °C for 3 min, followed by
637 12 cycles of (98 °C for 20 s, 65 °C for 15 s, 68 °C for 8 min), with a final extension at 68 °C for
638 8 min. Post-capture PCR products were purified using 0.7× volumes of AMPure XP beads, and
639 quantified by the Qubit dsDNA High Sensitivity assay and size checked by Agilent TapeStation.

640 **Gene-targeted single cell/ nucleus RACE-PCR**

641 To sequence the *NRXN1α* isoforms and 10x single cell barcode from one long-read, the
642 modified RACE-PCR was employed to amplify the target from the 10x Genomics 3'GEX or
643 5'GEX cDNA. 5'RACE was used for cDNA generated by 10x 5'GEX kit, while 3'RACE was used
644 for cDNA generated by 10x 3'GEX kit.

645 **PCR primer design**. As the expression level of *NRXN1α* is relatively low in adult brains, nested
646 PCR would be needed to achieve enough molecules. The gene-specific primers (GSPs) should
647 meet the criteria: 23-28 nt long to ensure specific annealing, with 50%-70% GC, a T_m 65°C
648 (best to have a T_m >70°C) and be specific to the target gene (Supplementary Table 2). Universal
649 primers: The universal primer for first -round PCR should be the partial read1 sequence from
650 10x cDNA. To increase sensitivity for the second-round PCR, a sequence of
651 "AATGATACGGCGACCACCGAGATCT" was added to the 5' of the read1 (Supplementary
652 Table 2), forming the second-round universal PCR primer (UP2) (Supplementary Table 2).

653 **Protocol**. The input of 1-50 ng of 10x 5'GEX or 3'GEX cDNA was used for the RACE-PCR. The
654 first round of RACE-PCR is carried out with 1 µL partial Read1 primer (10 µM) and 4 µL *NRXN1*
655 -specific primer 1 (10 µM) in a 50 µl-reaction, which consists of additional 25 µL 2X PrimeSTAR
656 GXL premix (Takara, Cat# R051A), 1- 50 ng of full-length single cell cDNA by 10x 5'GEX or
657 3'GEX kit. Choose the PCR program based on the GSP T_m . Program 1: when GSP T_m >70°C,
658 run 5 cycles of 94°C for 30s, 72°C for 8min; 5 cycles of 94°C for 30s, 70°C for 30s and 72°C for
659 8min; 15-20 cycles of 94°C for 30s, 68°C for 30s and 72°C for 8min. Program 2: when T_m = 60-
660 70°C, run 15-20 cycles of 94°C for 30s, 68°C for 30s and 72°C for 8min. The concentration of
661 PCR products was quantified by Qubit, and additional cycles can be added if the concentration
662 of PCR product is lower than 10 ng/ µL. After the PCR, 3 µL of products were analyzed with 1%
663 agarose/ GelRed gel (biotium, cat# 41002), while the rest products were purified using 0.7×
664 volumes of AMPure XP beads and eluted with 30 µL nuclease-free water. If the primary PCR
665 fails to show the distinct band(s) of interest or produces a smear, a second- round of RCE-PCR
666 should be performed.

667 The second-round of RACE-PCR is performed with 2.5 µL of UP2 (10 µM) and 2.5 µL *NRXN1* -
668 specific primer 2 (10 µM) in a 50 µl-reaction together with 25 µL 2X PrimeSTAR GXL premix
669 and 10 µL of the purified first-round RACE-PCR products. The PCR program follows the
670 program 2 and runs for 10-15 cycles based the final concentration of PCR products. After the
671 PCR, 3 µL of products were analyzed with 1% agarose/ GelRed gel for size check, which is
672 followed by the purification using 0.7× volumes of AMPure XP beads.

673 **Library preparation for Oxford Nanopore sequencing**

674 Oxford Nanopore SQK-NBD114.24 was used for library preparation. Amplified cDNA from post-
675 capture of RACE-PCR, approximately 200 ng of PCR products from each sample were used for
676 library preparation according to the ONT SQK-NBD114.24 protocol. Briefly, cDNA products
677 were end-repaired and dA-tailed with NEBNext Ultra II End Repair/dA-Tailing Module by
678 incubating at 20 °C for 30 min and 65 °C for 30 min. The cDNA was then purified with 1× volume
679 of AMPure XP beads and eluted in 10 µl of nuclease-free water. Native barcode ligation was
680 then performed using 10 µl of end-prepped DNA, 2.5 µl of native barcodes, and 12.5 µl of
681 Blunt/TA Ligase Master Mix (NEB, #M0367) by incubating at room temperature for 30 min. After
682 the ligation step, 2.5 µl of EDTA was added to each tube to inactivate the ligase. Barcoded
683 cDNA samples were then pooled and purified using 0.4× volume of AMPure XP beads and
684 eluted in 31 µl of nuclease-free water. Next, ligation of Native Adapter was performed with 30 µl
685 of pooled barcoded samples, 5 µl of native adapter, 10 µl of 5× NEBNext Quick ligation reaction
686 buffer, and 5 µl of NEBNext Quick T4 DNA ligase at room temperature for 30 min, followed by

687 the bead's purification using 0.45× volume of AMPure XP beads and twice wash with short
688 fragment buffer. The libraries were quantified with Qubit, and 50 fmol of final library was
689 sequenced with a R10.4.1 MinION or PromethION flow cell for 72 hours.

690 **Single-cell RNA-seq processing and analysis**

691 Raw 10x Genomics single-cell RNA-seq data were processed using Cell Ranger (v7.2.0) with the
692 GRCh38 reference genome to generate gene-cell UMI count matrices. Downstream analyses
693 were performed in Seurat (v5.0.3)[56] under R (v4.4.0). Cells with fewer than 200 detected genes,
694 more than 6,000 detected genes, or mitochondrial transcript fractions exceeding 20% were
695 excluded. Potential doublets were identified and removed using DoubletFinder[57].

696 Post-quality control, gene expression values were normalized by scaling each cell's total
697 expression to 10,000 counts and log-transforming the results. Data were then processed using
698 SCTransform[56], which performs regularized negative binomial regression to normalize and correct
699 for sequencing depth and mitochondrial content. Highly variable genes were selected for PCA,
700 and dimensionality reduction was performed using UMAP.

701 Clustering was conducted using a shared nearest neighbor (SNN) graph and the Louvain
702 algorithm to identify transcriptionally distinct populations. Cell type annotation was guided by
703 canonical marker gene expression and cross-referenced with public human brain (or tissue-
704 specific) single-cell cataloges. Differential expression analyses between groups or clusters were
705 performed using the Wilcoxon rank-sum test with Benjamini-Hochberg correction for multiple
706 testing (FDR < 0.05).

707 Following dimensionality reduction and clustering, known marker genes were used to guide the
708 assignment of biological cell identities. Each cluster was annotated based on the expression of
709 established cell-type-specific markers reported in previous studies.

710

711 **Long-read isoform sequencing analysis**

712 Full-length transcript isoforms were characterized using long-read RNA sequencing (Oxford
713 Nanopore or PacBio). Raw reads were quality filtered, adapter-trimmed, and aligned to GRCh38
714 using minimap2 (v2.28) with spliced alignment parameters.

715 Isoform reconstruction and quantification were performed using IsoQuant (v3.2) with GENCODE
716 v43 as reference annotation. IsoQuant[58] clusters reads into high-confidence transcript models,
717 classifying each as known, novel in catalog (NIC), or novel not in catalog (NNC) based on splice
718 junction composition. Only isoforms supported by at least three full-length reads were retained.

719 **Integration of long-read isoform data with short-read single-cell transcriptomes**

720 Integration between long-read and short-read data was performed at both the gene and isoform
721 levels.To associate long-read reads with individual single cells, we applied BLAZE (v1.4.0)[59],
722 which accurately detects 10x cellular barcodes and unique molecular identifiers (UMIs) directly
723 from Nanopore cDNA reads. BLAZE searches for barcode sequences within read adapters,
724 compares them to the 10x whitelist, and assigns valid barcodes (allowing up to one mismatch)
725 to each read. This enabled recovery of cell identities from long-read data and construction of a
726 cell × isoform count matrix.

727 ***NRXN1* Isoform Quantification using *miniQuant***

728 For the BA4 samples, *miniQuant*[60] was applied in hybrid mode, integrating both short-read and
729 long-read RNA-seq generated in-house to improve isoform-level quantification. For the FC and
730 hippocampus datasets[25,27], where only long-read RNA-seq data were available, *miniQuant* was
731 applied in long-read-only mode to estimate isoform abundances.

732 In *miniQuant*, long-read cDNA sequencing provides high-confidence isoform structures and
733 reduces ambiguity in assigning reads to specific isoforms, whereas short-read sequencing
734 improves quantification accuracy for low-abundance transcripts by reducing sampling error.
735 *miniQuant*'s gene-specific weighting model adaptively balances these complementary data
736 sources based on isoform structural complexity (K-value), transcript abundance, and
737 sequencing depth, producing optimized isoform abundance estimates. Isoform expression
738 levels were reported as transcripts per million (TPM) to allow direct comparison across samples.

739 To ensure quantification reliability, we filtered out low-confidence isoforms following the criteria
740 described in the original *miniQuant* framework, requiring both sufficient long-read support and a
741 minimum expression threshold. Only isoforms with TPM ≥ 1 in at least one sample were
742 retained for downstream analyses.

743 **Data visualization and statistical analysis**

744 All visualizations, including UMAPs, dot plots, violin plots, and bar plots, were generated in *R* or
745 *Python* using the *ggplot2*, *matplotlib*, and *seaborn* libraries. For cell type enrichment analyses,
746 statistical significance was evaluated using a two-tailed Fisher's exact test, with enrichment
747 ratios calculated as the observed versus expected frequencies. For continuous variables, such
748 as comparisons of AS event inclusion levels or isoform expression differences, we applied
749 *Welch's t-test*, which does not assume equal variances between groups.

750 **ACKNOWLEDGEMENTS**

751 We acknowledge Daniel Weinberger and the Lieber Institute for Brain Development at Johns
752 Hopkins School of Medicine, for sharing post-mortem brain tissues. This work is supported by
753 the National Institute of Mental Health grants RO1 MH125579 (GF and KJB) and RO1
754 MH121074 (KJB and GF), and U01MH116442 (PR) and R01MH125246 (PR). *Research*
755 *reported in this publication was supported by the National Institute of General Medical Sciences*
756 *of the National Institutes of Health under Award Number 1S10OD030363-01A1.* This work was
757 also supported in part through the computational resources and staff expertise provided by the
758 Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

759

760 **REFERENCES**

761 1. Fagnani, M. *et al*. Functional coordination of alternative splicing in the mammalian
762 central nervous system. *Genome Biol* **8**, R108 (2007).

763 2. Raj, B. & Blencowe, B. J. Alternative splicing in the mammalian nervous system:
764 recent insights into mechanisms and functional roles. *Neuron* **87**, 14–27 (2015).

765 3. Furlanis, E., Traunmüller, L., Fucile, G. & Scheiffele, P. Landscape of ribosome-
766 engaged transcript isoforms reveals extensive neuronal-cell-class-specific
767 alternative splicing programs. *Nat Neurosci* **22**, 1709–1717 (2019).

768 4. Lipscombe, D. & Soto, E. J. L. Alternative splicing of neuronal genes: new
769 mechanisms and new therapies. *Curr Opin Neurobiol* **57**, 26–31 (2019).

770 5. Gomez, A. M., Traunmüller, L. & Scheiffele, P. Neurexins: molecular codes for
771 shaping neuronal synapses. *Nature Reviews Neuroscience* vol. 22 137–151 Preprint
772 at https://doi.org/10.1038/s41583-020-00415-7 (2021).

773 6. De Wit, J. & Ghosh, A. Specification of synaptic connectivity by cell surface
774 interactions. *Nat Rev Neurosci* **17**, 4 (2016).

775 7. Südhof, T. C. Synaptic neurexin complexes: a molecular code for the logic of neural
776 circuits. *Cell* **171**, 745–769 (2017).

777 8. Harkin, L. F. *et al*. Neurexins 1-3 each have a distinct pattern of expression in the
778 early developing human cerebral cortex. *Cerebral Cortex* **27**, 216–232 (2017).

779 9. The role of neurexins and neuroligins in the formation, maturation, and function of
780 vertebrate synapses.

781 10. Schreiner, D. *et al*. Targeted Combinatorial Alternative Splicing Generates Brain
782 Region-Specific Repertoires of Neurexins. *Neuron* **84**, 386–398 (2014).

783 11. Treutlein, B., Gokce, O., Quake, S. R. & Südhof, T. C. Cartography of neurexin
784 alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc*
785 *Natl Acad Sci U S A* **111**, (2014).

786 12. Flaherty, E. *et al*. Neuronal impact of patient-specific aberrant NRXN1α splicing. *Nat*
787 *Genet* **51**, 1679–1690 (2019).

788 13. Uemura, T. *et al*. Trans-synaptic interaction of GluRδ2 and Neurexin through Cbln1
789 mediates synapse formation in the cerebellum. *Cell* **141**, 1068–1079 (2010).

790 14. Siddiqui, T. J. *et al.* An LRRTM4-HSPG complex mediates excitatory synapse
791 development on dentate gyrus granule cells. *Neuron* **79**, 680–695 (2013).

792 15. Boucard, A. A., Chubykin, A. A., Comoletti, D., Taylor, P. & Südhof, T. C. A splice
793 code for trans-synaptic cell adhesion mediated by binding of neuroligin 1 to α-and β-
794 neurexins. *Neuron* **48**, 229–236 (2005).

795 16. Ko, J., Fuccillo, M. V, Malenka, R. C. & Südhof, T. C. LRRTM2 functions as a neurexin
796 ligand in promoting excitatory synapse formation. *Neuron* **64**, 791–798 (2009).

797 17. Fuccillo, M. V *et al.* Single-cell mRNA profiling reveals cell-type-specific expression
798 of neurexin isoforms. *Neuron* **87**, 326–340 (2015).

799 18. Freund, T. F. & Katona, I. Perisomatic inhibition. *Neuron* **56**, 33–42 (2007).

800 19. Fernando, M. B. *et al.* Phenotypic complexities of rare heterozygous neurexin-1
801 deletions. *Nature* 1–11 (2025).

802 20. Lukacsovich, D. *et al.* Single-cell RNA-seq reveals developmental origins and
803 ontogenetic stability of neurexin alternative splicing profiles. *Cell Rep* **27**, 3752–3759
804 (2019).

805 21. Wang, F. *et al.* TEQUILA-seq: a versatile and low-cost method for targeted long-read
806 RNA sequencing. *Nat Commun* **14**, (2023).

807 22. Ray, T. A. *et al.* Comprehensive identification of mRNA isoforms reveals the diversity
808 of neural cell-surface molecules with roles in retinal development and disease. *Nat*
809 *Commun* **11**, (2020).

810 23. Ohara, O., Dorit, R. L. & Gilbert, W. One-sided polymerase chain reaction: the
811 amplification of cDNA. *Proceedings of the National Academy of Sciences* **86**, 5673–
812 5677 (1989).

813 24. Frohman, M. A., Dush, M. K. & Martin, G. R. Rapid production of full-length cDNAs
814 from rare transcripts: amplification using a single gene-specific oligonucleotide
815 primer. *Proceedings of the National Academy of Sciences* **85**, 8998–9002 (1988).

816 25. Hardwick, S. A. *et al.* Single-nuclei isoform RNA sequencing unlocks barcoded exon
817 connectivity in frozen brain tissue. *Nat Biotechnol* **40**, 1082–1092 (2022).

818 26. Dyer, S. C. *et al.* Ensembl 2025. *Nucleic Acids Res* **53**, D948–D957 (2025).

819 27. Joglekar, A. *et al.* Single-cell long-read sequencing-based mapping reveals
820 specialized splicing patterns in developing and adult mouse and human brain. *Nat*
821 *Neurosci* **27**, 1051–1063 (2024).

822 28. Ruzicka, W. B. *et al.* Single-cell multi-cohort dissection of the schizophrenia
823 transcriptome. *Science (1979)* **384**, eadg5136 (2024).

824 29. Telley, L. *et al.* Sequential transcriptional waves direct the differentiation of newborn
825 neurons in the mouse neocortex. *Science (1979)* **351**, 1443–1446 (2016).

826 30. Liodis, P. *et al.* Lhx6 activity is required for the normal migration and specification of
827 cortical interneuron subtypes. *Journal of Neuroscience* **27**, 3078–3089 (2007).

828 31. Miyoshi, G. *et al.* Prox1 regulates the subtype-specific development of caudal
829 ganglionic eminence-derived GABAergic cortical interneurons. *Journal of*
830 *Neuroscience* **35**, 12869–12889 (2015).

831 32. Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from
832 cellular properties to circuits. *Neuron* **91**, 260–292 (2016).

833 33. Cardin, J. A. Inhibitory interneurons regulate temporal precision and correlations in
834 cortical circuits. *Trends Neurosci* **41**, 689–700 (2018).

835 34. Huang, Z. J. & Paul, A. The diversity of GABAergic neurons and neural
836 communication elements. *Nat Rev Neurosci* **20**, 563–572 (2019).

837 35. Liakath-Ali, K. & Südhof, T. C. The Perils of Navigating Activity-Dependent Alternative
838 Splicing of Neurexins. Front. Mol. Neurosci. 14, 659681. Preprint at (2021).

839 36. Baxter, P. S., Dando, O. & Hardingham, G. E. Differential splicing choices made by
840 neurons and astrocytes and their importance when investigating signal-dependent
841 alternative splicing in neural cells. *Front Mol Neurosci* **16**, 1214439 (2023).

842 37. Ramos, S. I. *et al.* An atlas of late prenatal human neurodevelopment resolved by
843 single-nucleus transcriptomics. *Nat Commun* **13**, 7671 (2022).

844 38. Wang, L. *et al.* Molecular and cellular dynamics of the developing human neocortex.
845 *Nature* 1–10 (2025).

846 39. Zhu, K. *et al.* Multi-omic profiling of the developing human cerebral cortex at the
847 single-cell level. *Sci Adv* **9**, eadg3754 (2023).

848 40. Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-
849 cell atlas of early human brain development highlights heterogeneity of human
850 neuroepithelial cells and early radial glia. *Nat Neurosci* **24**, 584–594 (2021).

851 41. Sebastian, R. *et al.* Schizophrenia-associated NRXN1 deletions induce
852 developmental-timing-and cell-type-specific vulnerabilities in human brain
853 organoids. *Nat Commun* **14**, 3770 (2023).

854 42. Birey, F. *et al*. Assembly of functionally integrated human forebrain spheroids.
855 *Nature* **545**, 54–59 (2017).

856 43. Trujillo, C. A. *et al*. Complex oscillatory waves emerging from cortical organoids
857 model early human brain network development. *Cell Stem Cell* **25**, 558–569 (2019).

858 44. Fair, S. R. *et al*. Electrophysiological maturation of cerebral organoids correlates
859 with dynamic morphological and cellular development. *Stem Cell Reports* **15**, 855–
860 868 (2020).

861 45. Park, Y. *et al*. Modulation of neuronal activity in cortical organoids with bioelectronic
862 delivery of ions and neurotransmitters. *Cell Reports Methods* **4**, (2024).

863 46. Matsunami, N. *et al*. Identification of rare recurrent copy number variants in high-risk
864 autism families and their prevalence in a large ASD population. *PLoS One* **8**, e52239
865 (2013).

866 47. Kim, H.-G. *et al*. Disruption of neurexin 1 associated with autism spectrum disorder.
867 *The American Journal of Human Genetics* **82**, 199–207 (2008).

868 48. Kozareva, V. *et al*. A transcriptomic atlas of mouse cerebellar cortex
869 comprehensively defines cell types. *Nature* **598**, 214–219 (2021).

870 49. Uemura, T. *et al*. Trans-synaptic interaction of GluRδ2 and Neurexin through Cbln1
871 mediates synapse formation in the cerebellum. *Cell* **141**, 1068–1079 (2010).

872 50. Peng, H. *et al*. Single-cell Rapid Capture Hybridization sequencing reliably detects
873 isoform usage and coding mutations in targeted genes. *Genome Res* **35**, 942–955
874 (2025).

875 51. Lim, L., Mi, D., Llorca, A. & Marín, O. Development and functional diversification of
876 cortical interneurons. *Neuron* **100**, 294–313 (2018).

877 52. Nguyen, T.-M. *et al*. An alternative splicing switch shapes neurexin repertoires in
878 principal neurons versus interneurons in the mouse hippocampus. *Elife* **5**, e22757
879 (2016).

880 53. Biswas, M. S., Roy, S. K., Hasan, R. & Pk, M. M. U. The crucial role of the cerebellum
881 in autism spectrum disorder: neuroimaging, neurobiological, and anatomical
882 insights. *Health Sci Rep* **7**, e2233 (2024).

883 54. van der Heijden, M. E., Gill, J. S. & Sillitoe, R. V. Abnormal cerebellar development in
884 autism spectrum disorders. *Dev Neurosci* **43**, 181–190 (2021).

885 55. D'Mello, A. M. & Stoodley, C. J. Cerebro-cerebellar circuits in autism spectrum
886 disorder. *Front Neurosci* **9**, 408 (2015).

887 56. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell
888 analysis. *Nat Biotechnol* **42**, 293–304 (2024).

889 57. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: doublet detection in
890 single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst* **8**, 329–
891 337 (2019).

892 58. Prjibelski, A. D. *et al.* Accurate isoform discovery with IsoQuant using long reads. *Nat*
893 *Biotechnol* **41**, 915–918 (2023).

894 59. You, Y. *et al.* Identification of cell barcodes from long-read single-cell RNA-seq with
895 BLAZE. *Genome Biol* **24**, 66 (2023).

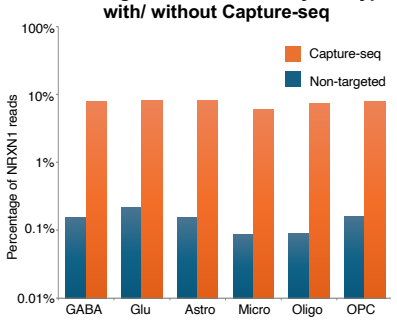
896 60. Li, H. *et al.* Improving gene isoform quantification with miniQuant. *Nat Biotechnol* 1–
897 13 (2025).

898

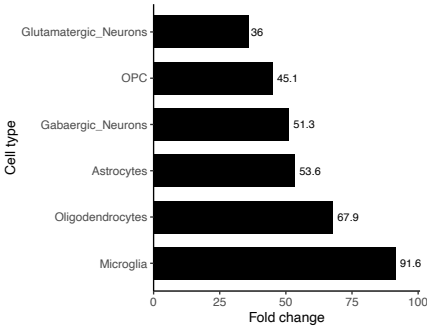
An Integrated Targeting Strategy Reveals Cell-Type Atlas of *NRXN1* Isoforms Across Human Brain Development and Genetic Backgrounds

a

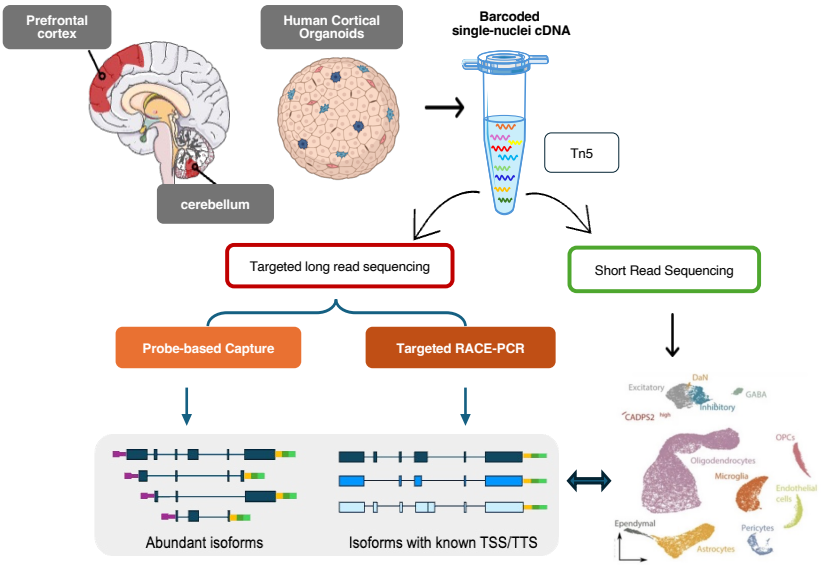
Percentage of *NRXN1* reads by cell type with/ without Capture-seq



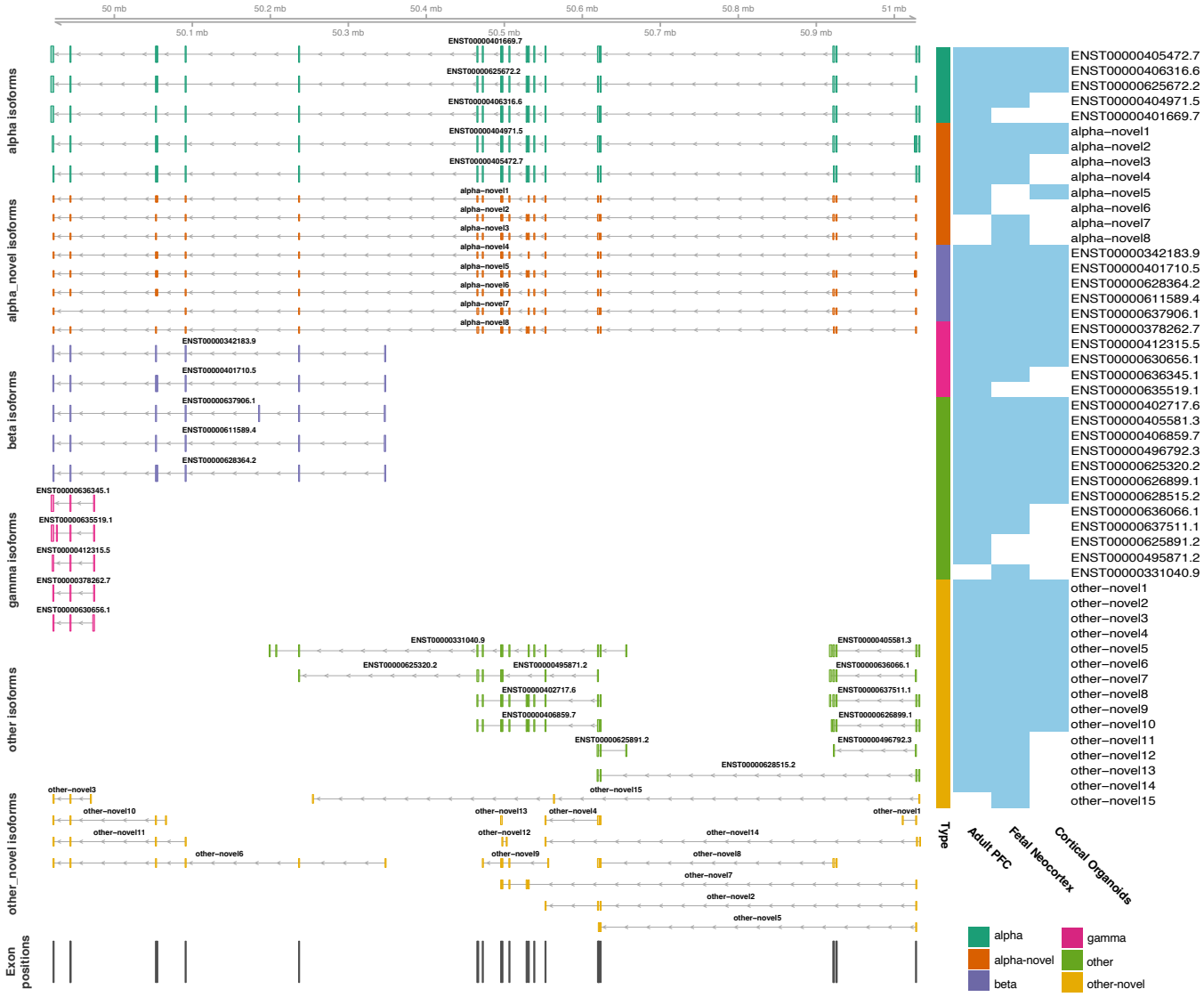
Fold enrichment of *NRXN1* reads by cell type



b **Workflow for comprehensive study of cell-type specific *NRXN1* isoforms**

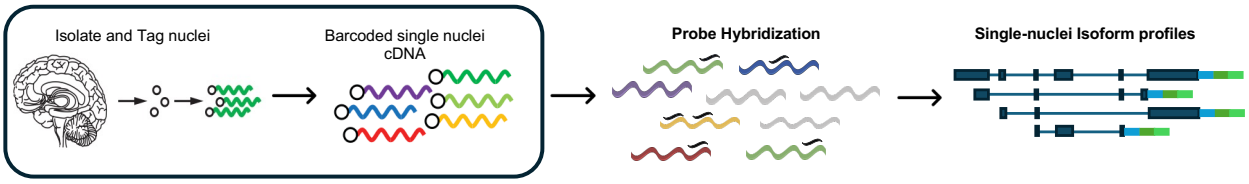


c **A schematic of the *NRXN1* isoforms identified across samples by single-cell long-read Capture-seq**

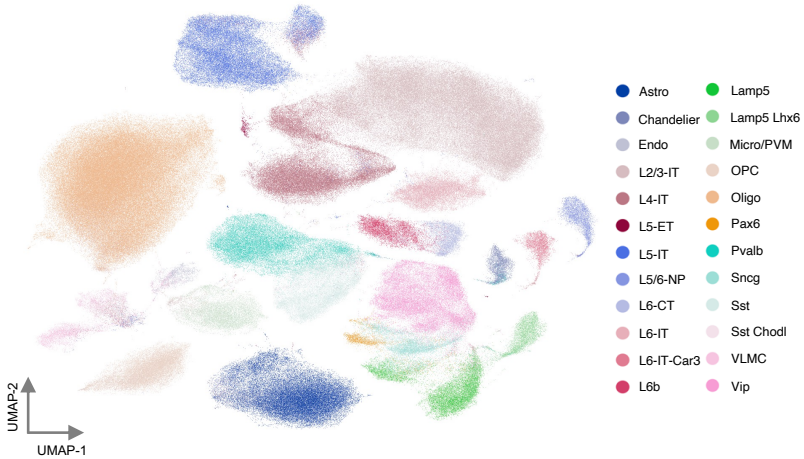


NRXN1 expression repertoires exhibit distinct splicing patterns across cell types in human brains

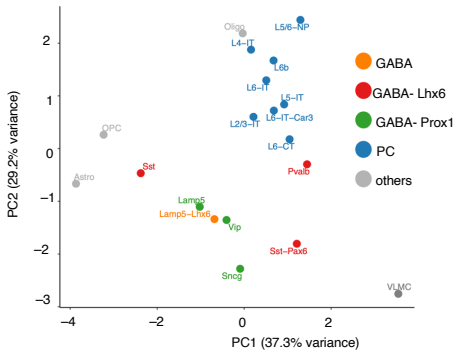
a Resolve cell-type specific *NRXN1* isoforms by probe-based long-read Capture-seq



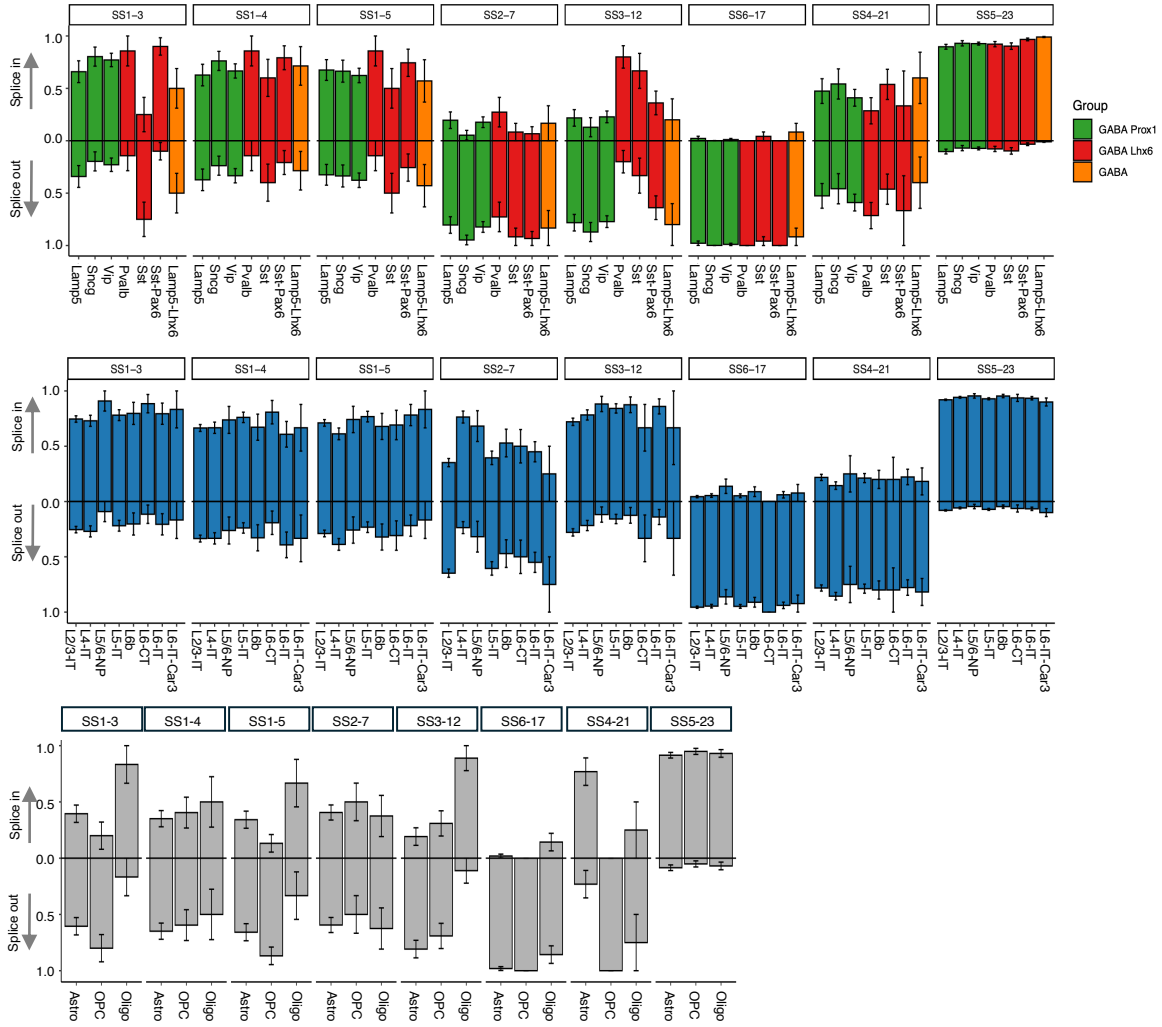
b UMAP visualization of cell types in the PFC of adult brains



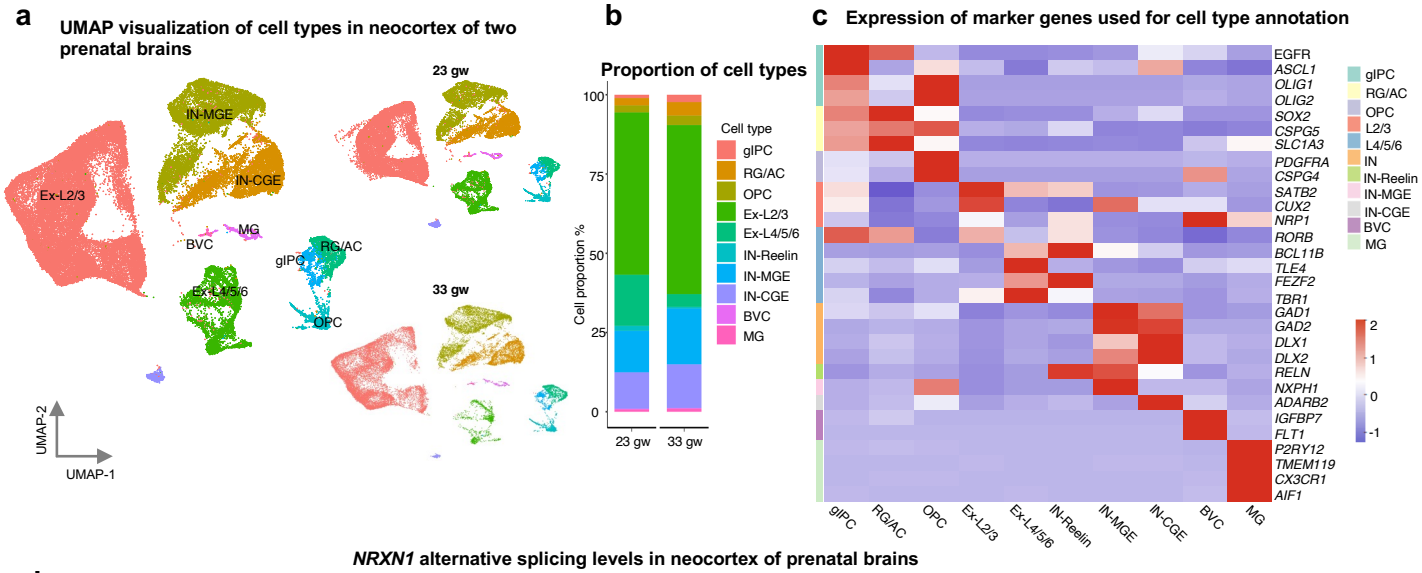
d PCA plot of *NRXN1* AS events across cell types

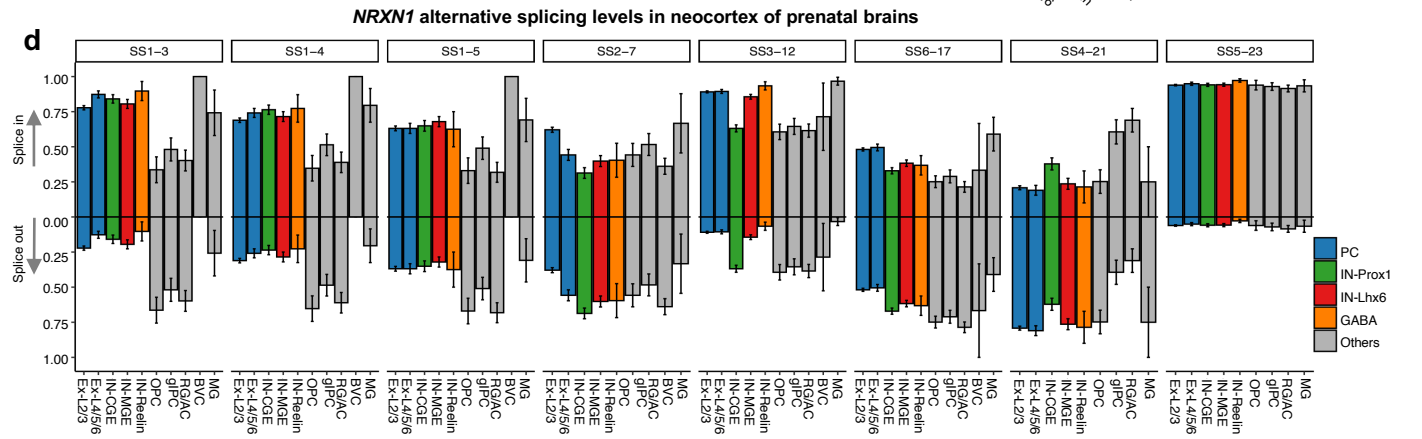


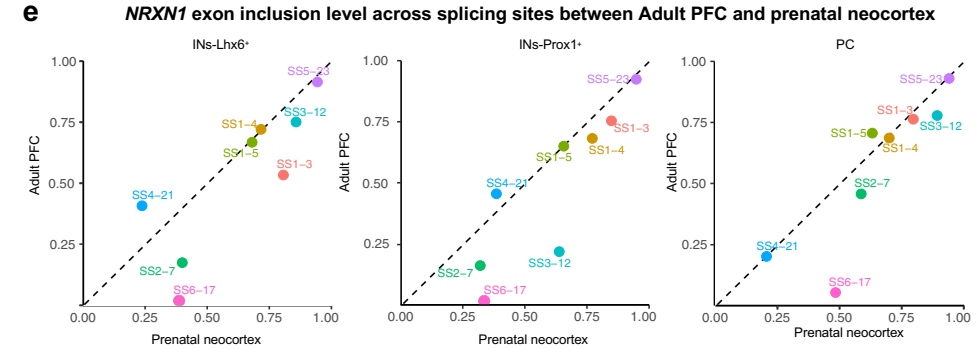
c Cell-type-specific *NRXN1* splicing patterns of SS1-6 revealed by long-read Capture-seq



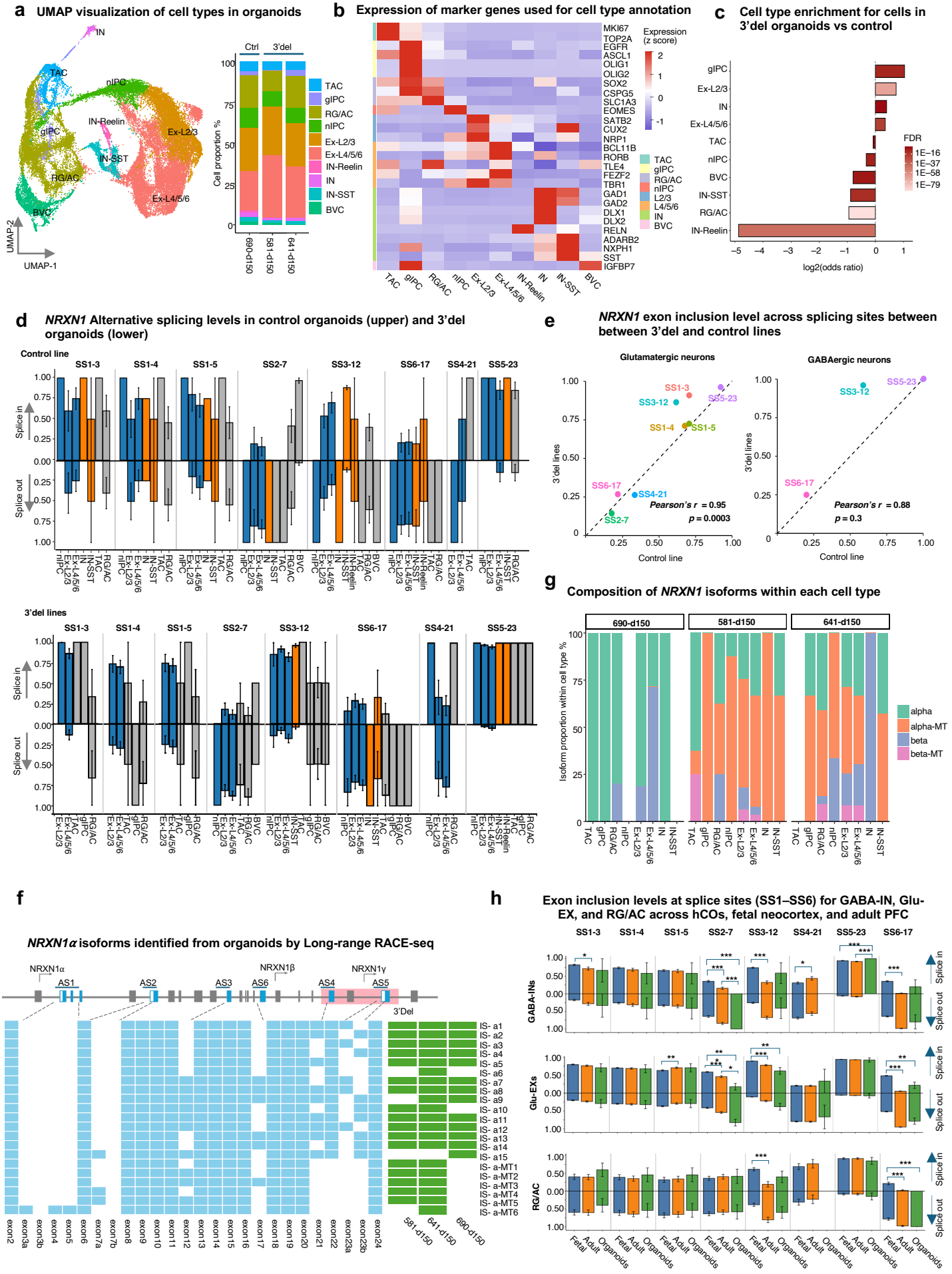
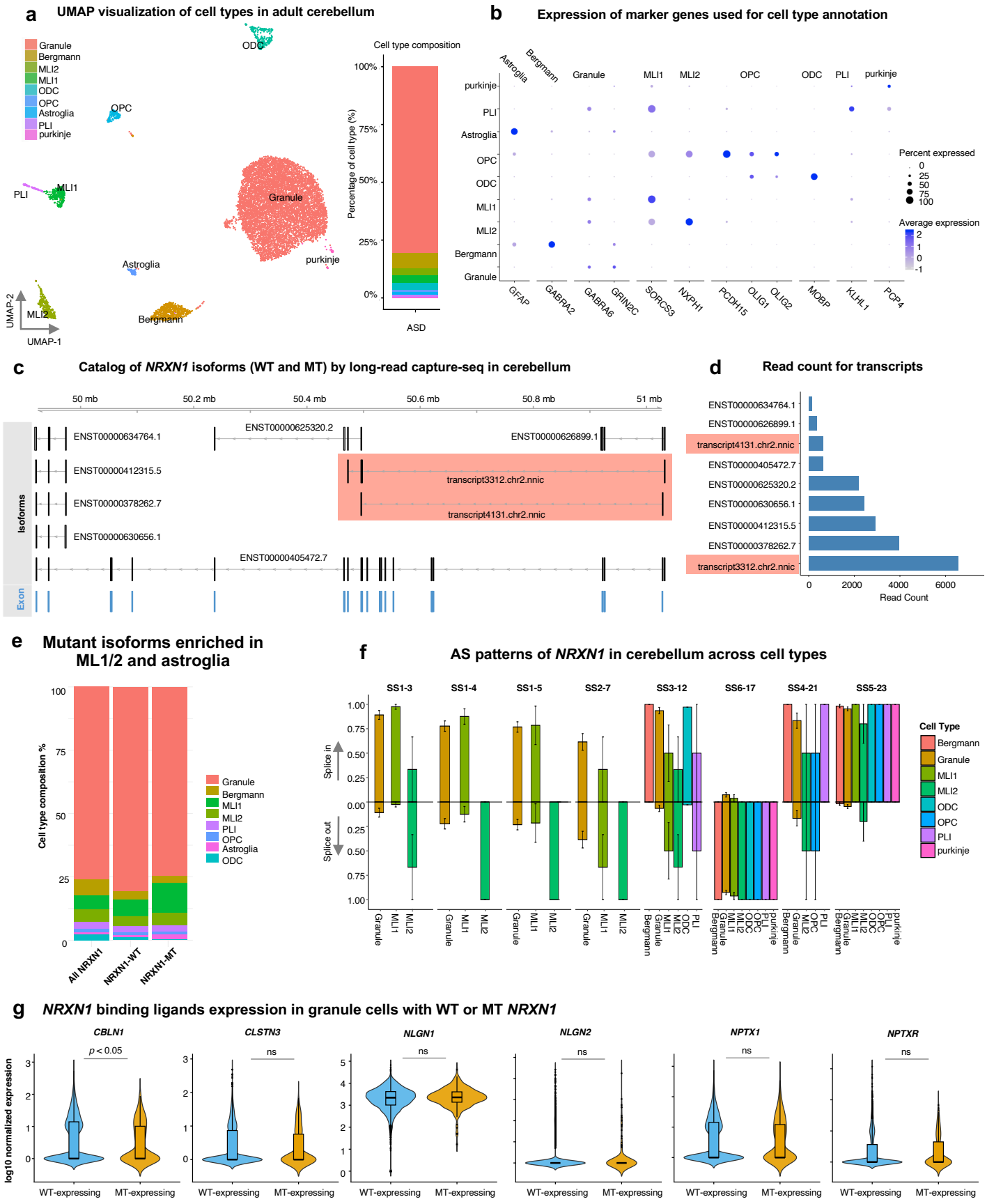
NRXN1 splicing profiles remain consistent across cells sharing a neurogenic origin during neural maturation







NRXN1$^{+/-}$ hCOs splicing profiles recapitulate developmental and mature features and identify cell types expressing the mutant isoforms



Unique exonic deletion impacts *NRXN1* splicing patterns across cell types in cerebellum of an ASD case