Cortical Organoids Model Early Brain Development Disrupted by 16p11.2 Copy Number Variants in Autism

Jorge Urresti[1#], Pan Zhang[1#], Patricia Moran-Losada[1], Nam-Kyung Yu[2], Priscilla D. Negraes[3,4], Cleber A. Trujillo[3,4], Danny Antaki[1,3], Megha Amar[1], Kevin Chau[1], Akula Bala Pramod[1], Jolene Diedrich[2], Leon Tejwani[3,4], Sarah Romero[3,4], Jonathan Sebat[1,3,5], John R. Yates III[2], Alysson R. Muotri[3,4,6,7,*], Lilia M. Iakoucheva[1,*]

[1] Department of Psychiatry, University of California San Diego, La Jolla, CA, USA

[2] Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA, USA

[3] Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, CA USA

[4] Department of Pediatrics/Rady Children's Hospital San Diego, University of California, San Diego, La Jolla, CA, USA

[5] University of California San Diego, Beyster Center for Psychiatric Genomics, La Jolla, CA, USA

[6] University of California San Diego, Kavli Institute for Brain and Mind, La Jolla, CA, USA

[7] Center for Academic Research and Training in Anthropogeny (CARTA), La Jolla, CA, USA

[#] These authors contributed equally: Jorge Urresti, Pan Zhang

*Corresponding authors: Lilia M. Iakoucheva (lilyak@ucsd.edu) and Alysson R. Muotri (muotri@ucsd.edu).

Summary

Reciprocal deletion and duplication of 16p11.2 is the most common copy number variation (CNV) associated with Autism Spectrum Disorders, and has significant effect on brain size. We generated cortical organoids to investigate neurodevelopmental pathways dysregulated by dosage changes of 16p11.2 CNV. We show that organoids recapitulate patients' macrocephaly and microcephaly phenotypes. Deletions and duplications have "mirror" effects on cell proliferation, neuronal maturation and synapse number, consistent with "mirror" effects on brain development in humans. Excess neuron number along with depletion of neural progenitors in deletions, and "mirror" phenotypes in duplications, demonstrate dosage-dependent impact of 16p11.2 CNV on early neurogenesis. Transcriptomic and proteomic profiling revealed synaptic defects and neuron migration as key drivers of 16p11.2 functional effect. Treatment with the RhoA inhibitor Rhosin rescued neuron migration. We implicate upregulation of small GTPase RhoA as one of the pathways impacted by the 16p11.2 CNV. This study identifies pathways dysregulated by the 16p11.2 CNV during early neocortical development.

Introduction

Over the last decade it has been convincingly demonstrated deletions and duplications of large genomic regions, or *copy number variants* (CNVs), are associated with multiple neurodevelopmental disorders (Marshall et al., 2008; McCarthy et al., 2009; Pinto et al., 2010; Sebat et al., 2007). The deletions (DEL) of a genomic region spanning 29 genes on human chromosome 16, 16p11.2 CNV, has been identified as one of the strongest risk factors for Autism Spectrum Disorder (ASD) and Intellectual Disability (ID), whereas the duplications (DUP) of the same region were associated with ASD, ID, schizophrenia (SCZ) and Bipolar Disorder (BD) (Bijlsma et al., 2009; Malhotra and Sebat, 2012; Marshall et al., 2008; McCarthy et al., 2009; Weiss et al., 2008). Furthermore, DEL and DUP are associated with increased and decreased size of the brain, respectively (Martin-Brevet et al., 2018; Sonderby et al., 2018). This phenotype, however, has not been recapitulated in mouse models (Arbogast et al., 2016; Horev et al., 2011; Portmann et al., 2014). Some of the animal studies have reported mirror effects of 16p11.2 CNV on regional brain volumes along with changes in gene expression, brain cytoarchitecture, behavior and viability, but there is little direct concordance in the phenotypes between human and mouse models.

Significant progress has been made for implicating various biological mechanisms that may be impacted by the 16p11.2 CNV. RNA sequencing of cortex from 16p11.2 deletion and duplication mice identified altered expression of genes and networks that converged on general ASD-associated pathways including synaptic function, chromatin modification and transcriptional regulation (Blumenthal et al., 2014). Transcriptome profiling of lymphoblastoid cell lines of 16p11.2 CNV human carriers identified expression dysregulation of neuronal-related gene in deletion, but not in duplication (Luo et al., 2012). Dysregulation of ciliopathy genes (Migliavacca et al., 2015), ERK/MAPK signaling (Pucilowska et al., 2018; Pucilowska et al., 2015), and metabotropic glutamate receptor 5 (mGluR5)-dependent synaptic plasticity and protein synthesis (Tian et al., 2015) in mouse models were all shown to play a role. Despite the progress made with regard to understanding of the general mechanisms disrupted by the 16p11.2 CNV, the question of how 16p11.2 variants impact early human brain development remains unanswered.

Recent advances in stem cell technologies opened a window of opportunities for investigating brain disorders using *in vitro* systems (Adams et al., 2019). During the last decade, 3D cortical organoid models have been successfully developed (Lancaster et al., 2013). Characterization of these models demonstrated that they closely resemble fetal brain, forming

structures reminiscent of deeper cortical layers and sharing cell types and transcriptomic signatures (Amiri et al., 2018; Camp et al., 2015; Luo et al., 2016; Velasco et al., 2019). Previous studies have used 3D cortical organoids to model lissencephaly (Bershteyn et al., 2017; Iefremova et al., 2017), autosomal recessive primary microcephaly (Lancaster et al., 2013), Timothy syndrome (Birey et al., 2017) and non-syndromic autism (Mariani et al., 2015). Here, we extend this list of studies by 3D modeling of the most common autism subtype associated with deletions and duplications of the 16p11.2 CNV.

In this study, we generated induced pluripotent stem cells (iPSCs) and cortical organoids from 16p11.2 DEL and DUP patient fibroblasts and unrelated healthy control (CTRL) individuals, and investigated molecular and cellular processes that are disrupted by this genetic variant. We found that the diameter of deletion organoids is larger, and duplication organoids is smaller, recapitulating “mirror” effect on brain size in humans. Transcriptome and proteome profiling of organoids identified differentially expressed genes, proteins, and co-expression modules with synaptic and neuronal migration functions. These results were validated by a panel of assays, followed by phenotype rescue experiments (**Fig. 1A**). Cellular assays confirmed that 16p11.2 deletions and duplications exhibit defects in neuronal maturation, migration, morphology and synaptic abnormalities, implicating disrupted neurogenesis. These mechanisms have not been previously associated with 16p11.2-linked autism, likely due to lack of appropriate 3D models of human fetal brain. Although many different pathways could be disrupted by the 16p11.2 CNV, we identified upregulation of the active form of small GTPase RhoA as a likely contributor to decreased migration in both, deletions and duplications. Treatment with a small molecule inhibitor Rhosin rescued migration, but not synaptic phenotype. Our study makes significant contribution to understanding of neurobiological mechanisms that may be disrupted during early human neocortical development in the 16p11.2 CNV carriers.

Results

Cortical organoids maturation resembles stages of human brain development

To investigate how 16p11.2 CNV impacts early stages of human brain development, and what neurobiological mechanisms are dysregulated by this genetic variant, we generated cortical organoids from the 16p11.2 CNV carriers. We first obtained iPSCs by reprogramming patient- and control-derived fibroblasts using episomal transduction, and then differentiated iPSCs into cortical organoids as previously described (Trujillo et al., 2019). We selected six male 16p11.2 CNV carriers with extreme head size phenotypes (age-normalized head circumference Z score

range from 2.51 to 4.32 in DELs; and from -0.69 to -1.63 in DUPs). Details of patients' phenotypes are described in **Table S1**. Three gender-matched unrelated individuals that did not carry 16p11.2 CNV were used as controls. We performed rigorous quality control assessment of reprogrammed iPSCs clones by comparing them to parental fibroblasts using immunofluorescence (**Figs. S1-S2**), qRT-PCR and SNP array genotyping (**Figs. S3-S4**); and selected two clones per individual for organoids production. We performed bulk RNA sequencing (RNA-seq) of a total of 108 samples derived from iPSCs, 1 month-old (1M) and 3 month-old (3M) organoids (36 samples at each stage). We sequenced two clones per individual, and two replicates per clone for all three genotypes (DELs, DUPs and CTRLs) (**Fig. S5**). Sequencing quality control parameters are shown in **Table S2**.

To investigate whether developmental maturity and laminar organization of produced organoids resembles human brain, we compared transcriptional profiles of iPSCs and organoids with the atlas of the developing human brain (Kang et al., 2011) using CoNTExT (Stein et al., 2014). Transcriptional profiles of iPSCs from all individuals closely matched those of embryonic (4-8 PCW) and early fetal (8-10 PCW) human brain, independently validating successful conversion of fibroblasts into pluripotent state by reprogramming (**Fig. 1B**). Transcriptional profiles of 1 month-old organoids resembled those of early mid-fetal (13-16 PCW) through late mid-fetal (19-24 PCW) periods. Likewise, transcriptional profiles of 3 month-old organoids mostly recapitulated those of late mid-fetal (19-24 PCW) through neonatal-early infancy (birth to 6 months) developmental periods, and even further into late infancy (6-12 M).

Next, we examined the degree of overlap between *in vivo* cortical development of prenatal human brain and our *in vitro* differentiated organoids using TMAP (Stein et al., 2014). We compared organoids transcriptional profiles with those derived from laser capture microdissected cortical laminae of postmortem human fetal brain (15-21 PCW) (Miller et al., 2014). TMAP performs serialized differential expression analysis between any two *in vivo* developmental periods and any two *in vitro* differentiation time points, followed by quantification of overlap (Stein et al., 2014). Laminar matching by TMAP demonstrated transitions between proliferative layers (ventricular VZ, subventricular SZ and intermediate IZ zones) and post mitotic upper layers for both, 1M and 3M old organoids (**Fig. 1C**). However, in 3M organoids laminar transition into these upper layers manifests greater shift than in 1M organoids. For example, greater correspondence to upper layers (subplate SP, cortical plate inner CPi and outer CPo layers and marginal zone MZ) is visible at 3M compared to 1M, consistent with increased maturity at a later time point. We replicated this result using an additional independent dataset (Fietz et al., 2012) (**Fig. S6**)

Together, the results suggest that cortical organoids mature over time, closely recapitulating human brain neurodevelopment in terms of temporal transitions and laminar organization. Furthermore, organoids between 1M and 3M of differentiation most closely resemble human mid-fetal brain development, and represent suitable models for studying molecular basis of neurodevelopmental disorders, considering a proven role of this period in ASD and SCZ pathogenesis (Gulsuner et al., 2013; Lin et al., 2015; Willsey et al., 2013).

Patient-derived organoids recapitulate macrocephaly and microcephaly phenotypes

Since our patients with 16p11.2 DELs and DUPs were selected based on the extreme head circumference phenotypes, we investigated whether organoids recapitulate these phenotypes. We measured the diameter of 16p11.2 and control organoids at four time points, at day 6 (D6, neural induction), day 16 (D16, proliferation), 1 month (1M, early maturation) and 3 months (3M, late maturation). We observed significant size differences between CTRL and DEL or DUP organoids at almost all time points, with DEL organoids being larger and DUP organoids being smaller than CTRL (**Fig. 1D-E** and **Table S3**). This difference was most apparent at the late maturation stage, corresponding to 3M of differentiation. These results demonstrate that cortical organoids recapitulate 16p11.2 patients' brain size phenotypes.

Differential gene expression analysis points to dysregulation of neuronal migration and actin cytoskeleton by 16p11.2 CNV

To understand molecular pathways dysregulated by the 16p11.2 CNV, we performed differential gene expression analyses of 108 transcriptomes derived from iPSCs, 1M, and 3M organoids (**Fig. S5 and Methods**). An extensive quality control and normalization included sample outlier detection, principal component analyses, surrogate variable analysis, and covariates selection with multivariate adaptive regression splines (**Figs. S7-S8 and Methods**). For gene differential expression analyses, we implemented limma-voom model with 'duplicateCorrelation' component to account for duplicate samples (clones and replicas) from the same individuals, and to avoid pseudo-replication in the analyses.

We identified 1044, 118 and 430 differentially expressed genes (DEGs) in DELs *vs* DUPs at 10% FDR for iPSCs, 1M and 3M organoids, respectively (**Fig 2A-C and Table S4**). Genes from the 16p11.2 *locus* were most significantly dysregulated, confirming the expected *cis*-effect of CNV on gene expression. In addition, 16p11.2 CNV had a significant effect on the expression of many genes outside of the *locus*. Gene Ontology (GO) analyses of DEGs from iPSC revealed enrichment in cell migration and motility, actin filament-based movement and potassium ion

homeostasis; whereas DEGs from organoids were enriched in regulation of neuron migration, actin cytoskeleton-related functions, layer formation in cerebral cortex and neuron differentiation (**Fig. 2A-C** and **Table S5**). The enrichment of migration and motility-related functions were consistently observed across all three time points (iPSCs, 1M and 3M organoids), with neurogenesis and synaptic functions observed only in organoids, further reflecting their advanced maturation stages compared to iPSCs.

Gene co-expression analyses identifies migration and synaptic/neuronal modules

We performed weighted gene co-expression network analysis (WGCNA) (Zhang and Horvath, 2005) to identify modules of genes with correlated expression that are impacted by the 16p11.2 CNV. Overall, we identified 11, 63 and 41 modules in iPSC, 1M and 3M organoid, respectively (**Table S6**). When all modules were statistically tested for association with DEL and DUP genotypes, thirty five modules (7 in iPSC, 7 in 1M, and 21 in 3M organoids) were detected as positively or negatively associated with genotypes at 10% FDR (**Figs. S9-S11**). As expected, we detected a module comprising 16p11.2 CNV genes to be positively associated with DUPs and negatively associated with DELs (*10purple* in iPSCs, *11greenyellow* in 1M and *16lightcyan* in 3M organoids, respectively) (**Extended Data Fig. 12**). Other notable modules included cell migration and motility (*22darkgreen* in 1M and *32violet* in 3M organoids), synaptic signaling and neuron differentiation (*46brown4* in 1M and *25orange* in 3M organoids), chromatin organization (*11greenyellow* in iPSCs), cilium assembly (*19lightyellow* in 3M organoids), and mitochondrial respiration (*3brown* in 3M organoids) (**Table S7**). Whereas some modules were unique to iPSCs or organoids, modules with migration and synaptic/neuronal functions were shared between 1M and 3M organoids. Interestingly, these modules were associated with genotypes in the opposite directions – migration module had negative association with DELs (*22darkgreen*) and positive with DUPs (*32violet*), while synaptic/neuronal module had positive association with DELs (*46brown4)* and negative with DUPs (*25orange*)(**Extended Data Figs. 10-11**).

To further investigate the modules, we performed statistical enrichment against curated gene lists with previous evidence for involvement in autism. We observed one module in each dataset with similar enrichment signatures (*9magenta* M9 in iPSCs, *6red* M6 in 1M and *2blue* M2 in 3M organoids) (**Fig. 2d**). These modules were enriched in highly confident ASD risk genes, constrained and highly intolerant to mutations (pLI>0.99) genes, as well as CHD8 and FMRP target genes in all datasets. GO analyses revealed shared biological functions related to histone modification and chromatin organization, with many ASD risk genes found within these functions.

We also observed several modules enriched in presynaptic or postsynaptic genes. Both, DEG and WGCNA analyses, suggest that the processes dysregulated by the 16p11.2 CNV at the transcriptome level converge on migration, synaptic/neuronal and chromatin-related functions.

Proteomic profiling of organoids supports transcriptomic findings

To investigate how 16p11.2 CNV impacts organoids' proteome, we carried out Tandem Mass Tag mass spectrometry (TMT-MS) on 1M and 3M organoids from the same samples as those used for RNA-seq experiments (**Fig. S13**). We detected a total of 6,126 proteins in 1M and 5,481 proteins in 3M organoids, with 13 and 11 proteins from within 16p11.2 CNV, respectively. When DEL proteomes were compared to those of DUPs, we identified 305 and 970 differentially expressed proteins (DEPs) at 10% FDR in 1M and 3M organoids, respectively (**Fig. 3A-B** and **Table S8**). In proteomic data, the *cis*-effect of 16p11.2 CNV was weaker than the effect observed by RNA-seq, possibly due to lower dynamic range between RNA and protein detectability in transcriptomic and proteomic experiments (Zubarev, 2013). However, patterns of proteome-wide effect of the 16p11.2 CNV on proteins outside of the *locus* were similar to transcriptome-wide effect, with more DEPs observed in 3M organoids compared to 1M organoids. Biological functions of DEPs agree remarkably well with those of DEGs, with most notable pathways related to actin cytoskeleton, synaptic, neuronal functions and locomotion (**Fig. 3A-B** and **Table S9**). The DEPs shared by 1M and 3M organoids include synaptic (SYN1, STX1B, SYNJ1), cytoskeletal (MAPT, TUBB4A, TRIO) and cell adhesion (NCAM, CNTN1) proteins that are downregulated in DUPs and upregulated in DELs. Similar trend is observed for several high-confident autism-associated proteins detected in both datasets (ANK2, DPYSL2, STXBP1, DYNC1H1). This suggests that 16p11.2 CNV impacts proteins outside of the *locus,* with particular effect on cytoskeletal, synaptic, and autism-relevant proteins.

To further investigate the impact of 16p11.2 CNV on organoids' proteome, we performed weighted protein co-expression network analysis (WPCNA) using TMT-MS proteomic data. We identified 21 and 17 co-expression protein modules in 1M and 3M organoids, respectively (**Table S10**). Twelve modules (5 in 1M and 7 in 3M organoids) were associated with DEL or DUP genotypes at 10% FDR (**Figs. S14-S15**). In addition to a module that was comprised largely of 16p11.2 proteins (*16lighcyan* in 1M organoids), other dysregulated modules included those enriched in GO functions RNA splicing and chromatin organization (*5green* in 1M organoids), ribosome and translation (*2blue* in 3M organoids), cytoskeleton and microtubule (*7black* in 3M organoids), and mitochondrial respiration (*8pink* in 3M organoids) (**Table S11**). The most interesting module detected in 3M organoids by WPCNA was *1turquoise* (M1). It was enriched in

pre- and postsynaptic, constrained and FMRP target proteins (**Fig. S16**). It included proteins involved in neuron differentiation and neurogenesis, neuron projection development, synaptic signaling, cytoskeleton organization, actin filament processes, as well as migration and locomotion. All these functions were also identified by RNA-seq profiling.

Functional convergence of organoids transcriptome and proteome

To determine the extent of convergence between organoids transcriptomes and proteomes, we calculated correlation coefficient, and carried out module preservation analyses (**Fig. 3C-D**). Globally, we observed positive correlation between DEGs and DEPs in 1M organoids (Pearson r=0.25) (**Fig. 3C**) and in 3M organoids (Pearson r=0.1) (**Fig. S17**). Module preservation analyses demonstrated high degree of preservation at the RNA/protein level for most modules that were significantly associated with genotype (**Fig. S18**), and especially for 16p11.2, migration and neuronal/synaptic modules in both 1M (**Fig. 3D**) and 3M (**Fig. S19**) organoids. As expected, the module containing 16p11.2 genes had highest preservation scores within transcriptome modules, whereas lower preservation scores were observed for migration and synaptic modules (**Fig. 3D**). Overall, we observed remarkably high correlation between organoids transcriptomes and proteomes, given previously noted lack of conservation between the two in other studies (Liu et al., 2016; Wang et al., 2019).

To put our results into a context of previous 16p11.2 CNV- or autism organoids-focused studies, we performed enrichment analyses our DEGs and DEPs in other datasets. Specifically, our DEGs and DEPs were compared with transcriptomes of 16p11.2 patients' lymphoblast lines and cerebral cortex of 16p11.2 mice (Blumenthal et al., 2014), idiopathic ASD patient-derived organoids (Mariani et al., 2015) and CHD8 KO organoids, NPCs and neurons (Wang et al., 2017) (**Fig. 3E**). Overall, we observed greater overlap of our DEGs and DEPs with DEGs identified in idiopathic ASD organoid models, suggesting that our organoids share transcriptomic signatures with other ASD subtypes. With regards to transcriptomes from 16p11.2 patients' lymphoblastoid cell lines and 16p11.2 mouse cortex, our 3M organoids captured transcriptomic signatures of these models better than 1M organoids, potentially due to their more advanced stages of maturation. We observed lesser overlap of our DEPs with DEGs from human lymphoblastoid lines and adult mouse cortex, potentially reflecting general low correlation between protein and mRNA levels due to various biological and technical factors (Liu et al., 2016; Wang et al., 2019).

The dosage of 16p11.2 CNV impacts cell type composition of organoids

The findings from transcriptomic and proteomic analyses provide important insights into molecular processes that are disrupted by the 16p11.2 CNV in the context of brain development. Given complex cell type composition of human brain, there is likely cell type-specific effect manifested by this CNV. To better understand how 16p11.2 dosage may impact cell type composition of organoids, we performed cell type enrichment analyses of organoid transcriptomes using single cell RNA-seq (scRNA-seq) from developing human neocortex (Polioudakis et al., 2019).

We have previously demonstrated that at 1M, organoids primarily consist of progenitor cells, with smaller fractions of glutamatergic neurons, glial cells, and intermediate progenitors (IP) (Trujillo et al., 2019). Here, we used recent scRNA-seq data from fetal human neocortex (Polioudakis et al., 2019) to identify cell types significantly enriched in organoids at 1M and 3M. We observed significant enrichment of at least one relevant cell type in 12/63 co-expression modules for 1M organoids and in 12/41 co-expression modules for 3M organoids (**Fig. 4A**). The analyses of the significant modules by genotype revealed that DEL organoids were enriched in neuronal cell types and DUP in cycling progenitors at both, 1M and 3M time points (**Fig. 4B**). Interestingly, DUPs were also enriched in intermediate progenitors and radial glia at 1M (**Fig. 4C-E**). In support of cell type enrichment results, GO functions of most significant modules reflected processes typically associated with corresponding cell types (**Fig. 4B-E, right panels**). For example, 1M *45darkorange2* “Neuron” cell type module’s GO functions included “neurogenesis,”, “neuron development” and “neuron differentiation”. Likewise, the “CycProg” *1turquoise* module’s functions were mostly related to cell cycle, mitosis, chromosome organization and cytoskeleton. The GO functions of the “IP” *42lighcyan1* module included “pattern specification process”, “nervous system development”, and “generation of neurons”, along with differentiation- and proliferation-related functions. The “RG” *2blue* module’s functions captured cilium and microtubule-based processes. Overall, cell type enrichment results provided further insight into cell type composition of organoids and correlated with previous findings from ASD brain. Excess neuron number has been observed as a hallmark of brain overgrowth in ASD patients during first year of life (Courchesne et al., 2007; Stoner et al., 2014), supporting “Neuron” cell type enrichment and macrocephaly phenotype in DELs. Likewise, depletion of radial glia and cilial defects during neurogenesis have been previously associated with microcephaly in humans (Shohayeb et al., 2020), supporting “RG” depletion in DUP at 3M.

Altered neuronal maturation is observed in 16p11.2 organoids

Both transcriptome and proteome analyses pointed to dysregulation of synaptic and neuronal processes that may underlie some of the observed organoids' phenotypes. Transcriptomic module *46brown4* in 1M organoids was enriched in "Neuron" cell type (**Fig. 4A**), was significantly upregulated in DELs (**Fig. 5A**), and consisted of genes with neuronal and synaptic functions (**Fig. 5B**). The expression levels of genes from this module highly correlate with corresponding protein expression (Pearson correlation coefficient PCC=0.62), with 42.2% of genes within this module also detected by proteomics **(Fig. 5C)**. One of the high-confidence autism risk genes, *SCN2A* (Satterstrom et al., 2020), is one of highly connected hubs in this module (**Fig. 5D**).

To validate these findings and to better understand cellular basis of synaptic dysregulation in organoids, we quantified neural progenitors and neurons by immunohistochemistry in 1M organoid slices. We observed opposite cellular phenotypes in DEL and DUP organoids: DEL organoids had greater number of neurons (i.e. $NeuN^+$ cells) and lower number of neural progenitors ($Pax6^+$), and DUP organoids had the opposite phenotype (**Fig. 5E-F**). This suggests that DEL organoids mature faster than DUP organoids, and that progenitor proliferation dynamics is disrupted by the 16p11.2 CNV. Excess neuron number and larger size in DELs supports previously observed link between neuron number and early brain overgrowth in ASD patients (Courchesne et al., 2007; Stoner et al., 2014). The DEL organoids also had decreased proliferation rate, likely due to depletion of progenitor pool by 1M (**Fig. S20**). Accelerated proliferation of neural progenitors from iPSCs has been previously observed much earlier than 1M in other models (Marchetto et al., 2017), suggesting that by 1M we are capturing later or even terminal stages, at which progenitor pool in DELs have already been depleted. Cell cycle exit determined by the ratio of Edu^+ and $Ki67^-$ cells was not affected in 1M DEL or DUP organoids. Summary of this and all follow-up experiments by clones and replicates is shown in **Table S12**.

Neuronal morphology and synaptic defects are detected in 16p11.2 organoids

Previous study in 2D neuronal cultures derived from iPSCs of 16p11.2 patients observed differences in soma size and dendrite length DEL and DUP neurons (Deshpande et al., 2017). To replicate these findings in organoid models, we investigated neuron morphology by measuring soma size and neurite length in the dissociated 1M organoids (**Methods**). Since cortical organoids at 1M of differentiation produce functional cortical neurons (Trujillo et al., 2019), we stained dissociated cells with MAP2 neuronal marker and performed measurements eight days after plating. We observed increased soma size in DEL organoids compared to CTRL (one-way

ANOVA p=0.034) (**Fig. 5G-H**). The total neurite length was increased in DEL *vs.* CTRL (one-way ANOVA p=0.01), and in DEL *vs.* DUP (p=0.0007), with a trend for decreased neurite length in DUPs *vs* CTRL that did not reach statistical significance. These results suggest that one of many factors contributing to organoid size differences could be soma size and neurite length.

Changes in neuronal morphology together with altered neuronal maturation could impact synaptogenesis in organoids. We therefore analyzed synaptic puncta by co-staining 1M organoid slices with presynaptic marker Synapsin-I (SynI) and neuronal marker MAP2. We observed significant increase in the number of synaptic puncta in DEL organoids (one-way ANOVA p=0.0003) (**Fig. 5I-J**). This result correlates with the increased number of neurons, and with upregulation of neuronal/synaptic transcriptomic module observed in DELs.

Severe neuronal migration defects are observed in 16p11.2 organoids

Gene sets and modules involved in neuronal migration and locomotion were dysregulated across all transcriptomic (DEG and WGCNA) and proteomic (DEP and WPCNA) analyses. For example, transcriptomic *22darkgreen* module from 1M organoid was significantly downregulated in DELs, and annotated with locomotion, migration, and motility (**Fig. 6A-B**). Other GO functions within this module included Wnt signaling and regulation of Rho protein signal transduction, both related to cell migration and cytoskeletal functions. Similarly to synaptic module, gene and protein expression within this module were positively correlated (Pearson correlation coefficient PCC=0.46) **(Fig. 6C)**. One of the genes in this module, ARHGEF2, is a microtubule-regulated Rho guanine exchange factor, known to be involved in cell migration (**Fig. 6D**).

To investigate migration defects in the 16p11.2 organoids, we performed two orthogonal *in vitro* migration assays as previously described (Bershteyn et al., 2017; Boyden, 1962). First, we seeded organoids onto matrigel-coated plates, and quantified the number of migrated neurons and migration distance 72 hours after seeding. Within the first 24 hours after plating, protrusions of fibers from the organoid edges were observed. Then, neurons started to migrate along these fibers. While about 40% of neurons migrated to a distance of over 200 μm along the fibers in the CTRL organoids, only about 20% of neurons migrated to the same distance from the DEL or DUP organoids in this experiment (one-way ANOVA DEL/CTRL p=0.038; DUP/CTRL p=0.073) (**Fig. 6E-F**). Live imaging further verified that migration distance is shorter for both DEL and DUP organoids (**Fig. S21 and movie S1**). We verified by immunostaining that the fibers consist of both neurites and radial glia bundles, and that the migrating cells are neurons and not neural progenitors (**Fig. S22**). The orthogonal Boyden chamber assay further confirmed migration

defects in 16p11.2 organoids (**Fig. 6G-H**). These results suggest that neuron migration defects are observed in both, DEL and DUP organoids.

RhoA inhibition rescues migration defects in 16p11.2 organoids

Rho signaling is one of the pathways enriched in the migration module (**Fig. 6B**). We previously hypothesized that this pathway may be impacted by the 16p11.2 CNV (Lin et al., 2015). Recent follow-up studies in mice with deletion of individual 16p11.2 genes, KCTD13 (Escamilla et al., 2017) and TAOK2 (Richter et al., 2018) also observed changes in RhoA level.
RhoA is known to regulate neuronal migration, actin dynamics and neurite outgrowth during brain development (Azzarelli et al., 2014; Cappello et al., 2012; Govek et al., 2011). Since we observed defects in migration in DEL and DUP organoids, we tested for potential involvement of RhoA by Western Blot (**Fig. 7A-B and Fig. S23**). As expected, KCTD13 protein level was significantly decreased in DEL and significantly increased in DUP organoids compared to CTRL (ratio paired t-test CTRL *vs* DEL p=0.01, CTRL *vs* DUP p=0.02), mirroring 16p11.2 CNV dosage. Total RhoA levels had an inverse trend, as we have previously hypothesized (Lin et al., 2015). However, active GTP-bound RhoA (RhoA-GTP) was significantly upregulated in organoids of both genotypes (ratio paired t-test CTRL *vs* DEL p=0.02, CTRL *vs* DUP p=0.03), in agreement with decreased neuronal migration in DELs and DUPs (**Fig. 6E-H**).

Next, we used RhoA inhibitor Rhosin, shown to specifically block RhoA activation (Shang et al., 2012), to rescue neuron migration. After constitutive treatment of intact organoids with Rhosin starting from 6th day of differentiation, migration defects in both DEL and DUP were rescued to the levels indistinguishable from CTRL (one-way ANOVA CTRL *vs* DEL_Rh p=0.53; CTRL *vs* DUP_Rh p=0.36) (**Fig. 7C-E**). Boyden chamber experiments replicated migration phenotype rescue with Rhosin (**Fig. 7F-G**). This suggests that Rho signaling may be one of the pathways that is contributing to decreased migration phenotype observed in DEL and DUP organoids.

Discussion

Patient-derived and isogenic brain organoids are becoming popular models for investigating molecular mechanisms underlying neurodevelopmental disorders (Amin and Pasca, 2018; Setia and Muotri, 2019). Given lack of fetal brain tissues from ASD patients, there are numerous advantages in using brain organoids to model patient phenotypes. In the present study we model molecular and cellular mechanisms of ASD risk attributable to rare deletions and

duplications at 16p11.2 *locus* using cortical organoids derived from skin fibroblasts of ASD patients with macrocephaly or microcephaly, respectively.

Organoid models of 16p11.2 CNV exhibit defects in neuronal maturation, migration, morphology and synaptic abnormalities. These genetic effects are consistent with brain growth phenotypes observed in the patients carrying 16p11.2 DEL or DUP. Accelerated neuronal maturation in DEL and delayed in DUPs, along with impaired migration, are novel mechanisms that have not been previously implicated in 16p11.2-linked autism. Our study makes significant contribution to mechanistic understanding of cellular and molecular processes that may be disrupted during early neocortical development in the 16p11.2 CNV carriers.

In this study, we perform in-depth characterization of organoids' transcriptomes and proteomes in parallel, from the same samples, at different developmental time points. This strategy provides the leverage for comparing mirror phenotypes of DELs and DUPs at two levels of regulation, transcriptional and translational. We obtain highly consistent results at the level of genes and proteins, by two independent methods, RNAseq and quantitative proteomics.

At the molecular level, we observe perturbations of transcriptional programs associated with key processes involved in neurodevelopment. Dysregulation of genes with neuron differentiation, synaptic signaling and cortical layer formation functions are observed in organoids, but not in the iPSC lines. This suggests that disruption of neural processes, which may not be apparent at the very early embryonic stages in iPSCs, become more pronounced during organoids maturation. Transcriptional signatures of one month old organoids recapitulate those of the late mid-fetal human brain development, the most critical period for establishing network connectivity among 16p11.2 CNV genes (Lin et al., 2015). This period was also implicated in ASD by other studies (Parikshak et al., 2013; Willsey et al., 2013). Most importantly, dysregulated transcriptional modules associated with synaptic functions and neuronal migration identified in organoids were conserved at the proteomic level. Preservation of transcriptional signatures at the translational level further reinforces and validates our findings.

Our results are in agreement with other studies that include either organoids produced from idiopathic ASD individuals (Mariani et al., 2015), or transgenic mice with 16p11.2 CNV (Blumenthal et al., 2014). Significant overlap between differentially expressed genes and proteins from our study is observed with DEGs from Mariani (Mariani et al., 2015), suggesting shared signatures among different genetic subtypes of ASD. Importantly, synaptic gene co-expression module is also dysregulated in organoids from idiopathic ASD patients. The degree of overlap is

lower with DEGs from organoids engineered to knock-down CHD8, a top autism gene (Wang et al., 2017). It is likely that patient's overall genetic background may be an important contributor to the observed transcriptional signatures. With regard to the 16p11.2 mouse model, our 3M organoids capture greater signal from 16p11.2 adult mouse brain than 1M organoids, likely due to their more advanced stages of maturation with closer similarity to brain tissues.

As observed previously in other brain diseases, organoid models can recapitulate patient's microcephaly (Bershteyn et al., 2017; Cugola et al., 2016; Lancaster et al., 2013) and macrocephaly (Li et al., 2017; Wang et al., 2017) phenotypes. Here, we extend this finding by demonstrating that dosage changes of the same genetic variant leads to opposite organoids size. In addition, we replicate altered neuronal morphology observed in 2D studies (Deshpande et al., 2017), suggesting that it may be one of the phenotypes contributing to size differences between DELs and DUPs. Aberrant control of cell proliferation and excess neuron number has been previously hypothesized to cause early brain overgrowth in ASD patients (Courchesne et al., 2007; Vaccarino et al., 2009). Consistent with this hypothesis, we observed excess of neurons and depletion of neural progenitors in DEL organoids. The opposite phenotype was found in DUP organoids. We also found decreased cell proliferation in 1M old DEL organoids. We believe that accelerated proliferation in DELs at earlier time points causes premature depletion of neural progenitor pool by 1M. Interestingly, a previous study did not find differences in proliferation in 16p11.2-derived NPCs or iPSCs (Deshpande et al., 2017), whereas Mariani (Mariani et al., 2015) study shows accelerated proliferation in monolayer NPCs, but not in organoids, which is more consistent with our results. Another recent report implicated increased proliferation of iPSCs and NPCs in early brain overgrowth (Marchetto et al., 2017). Further investigation of proliferation rates at various developmental time points (iPSCs, NPCs, early maturation and late maturation in organoids) is needed, and could uncover time-dependent mechanisms of proliferation defects in ASD.

One of the important findings from our study is impaired cortical neuron migration in 16p11.2 organoids. We confirmed reduced migration in DEL and DUP organoids by two orthogonal methods, recordings from intact organoids, and experiments in dissociated organoids. Previously, neuronal migration defects have been observed in organoids derived from patients with lissencephaly (Bershteyn et al., 2017), periventricular heterotopia (Klaus et al., 2019), and in CHD8 deficient mice (Xu et al., 2018), but was not previously implicated in ASD caused by the 16p11.2 CNV. Our results suggest that aberrant neuronal migration might be present in the brains of subjects with 16p11.2 CNV. Notably, observations from post-mortem ASD brains show patches

of disorganized cortical neurons that may not be migrating properly during early brain development (Stoner et al., 2014).

Our observations about increased active RhoA levels and decreased neuron migration fit perfectly with earlier data that examined these relationships *in vivo* in developing mouse cortex (Cappello et al., 2012). First, this study revealed that RhoA-depleted neurons migrated faster and reached cortical plate sooner than GFP+ control neurons upon Cre electroporation into cerebral cortex of E14 mouse embryos. Second, the same study demonstrated that electroporation of spontaneously activated (''fast-cycling'') mutant of RhoA caused slower neuronal migration in the condition of activated RhoA. Thus, overactivation of RhoA stalls migration of neurons. These data were also consistent with previous work rescuing delayed neuronal migration by inactivation of RhoA or inhibition of ROCK, a direct target of RhoA (Hand et al., 2005; Kholmanskikh et al., 2003; Pacary et al., 2011). Here, we demonstrate that inhibition of RhoA with Rhosin also rescues delayed migration in organoids.

There are a myriad of biological pathways that could be dysregulated by the 16p11.2 CNV in ASD (Iakoucheva et al., 2019). Due to its polygenic nature, with 29 genes within the *locus* and hundreds of genes impacted outside of the *locus* (as demonstrated here), genetic and epistatic interactions among these genes are likely responsible for neuroanatomical and cellular phenotypes observed in the patients and animal models (Iyer et al., 2018; Qiu et al., 2019). The effect of this CNV due to combinatorial and synergistic effects of multiple genes (Grice et al., 2015), creates apparent challenges in implicating a specific pathway. Rather, dysregulation of multiple pathways could lead to observed cellular and molecular phenotypes. For example, a number of genes within 16p11.2 CNV (MAPK3, MVP and TAOK2), are involved in MAPK/ERK and phosphatidylinositol 3-kinase PI3K/AKT signaling pathways. These pathways, regulating cell cycle in proliferation of neural progenitors, were shown to be dysregulated in 16p11.2 deletion mouse model (Pucilowska et al., 2015), and are likely to also be impacted in our organoid model. Here, we identified modules with genes and proteins involved in Wnt/β-catenin signaling, suggesting that 16p11.2 CNV may also impact this pathway. Finally, Rho GTPase signaling that regulates neuronal migration, actin dynamics and neurite outgrowth during brain development is another likely candidate pathway for 16p11.2 CNV. Thus, pleiotropy and epistasis of 16p11.2 genes at the pathway level is a hallmark of its functional impact. Future studies using organoid models or fetal brain tissues from 16p11.2 CNV carriers are required to untangle the complexity of phenotype-pathway relationships in ASD.

Materials and Methods

Tissue collection

Skin fibroblasts of three patients with 16p11.2 deletions (DEL) and three patients with 16p11.2 duplications (DUP) were obtained from the Simons Searchlight https://www.sfari.org/resource/simons-searchlight/; formerly Simons Variation in Individuals Project or Simons VIP). Patients were selected based on fibroblasts availability, head circumference, ASD diagnosis, and were gender and age matched (see detailed information about the patients in **Table S1**). De-identified patients tissue samples are distributed to Simons Investigators following approved IRB protocol to Simons Foundation through Columbia University Medical Center (PIs Drs. Gerald Fischbach and Wendy Chung). Collection and use for research of fibroblasts from three de-identified control individuals (CTRL) was approved by UCSD IRB. Skin fibroblasts were maintained in DMEM F-12 (Life Technologies) containing 10% fetal bovine serum.

Generation and maintenance of iPSCs

To generate induced pluripotent stem cells (iPSCs), skin fibroblasts were infected with Sendai virus vectors containing coding sequences of human OCT4, SOX2, KLF4 and c-MYC (Cytotune reprogramming kit, Thermo Fisher). Four days post infection, fibroblasts were trypsinized to single cells, plated on the inactivated mouse embryonic fibroblast feeders, and cultured using human embryonic stem cell medium (Gibco). After 3–4 weeks, iPSC clones were manually picked and propagated clonally on feeders. After 8-10 passages, iPSCs were transferred to a feeder-free system and grown on matrigel-coated dishes (Corning) in mTeSR1 media (StemCell Technologies). The cells were passaged by manually picking colonies.

Quality Control of generated iPSC clones

The generated iPSC clones were examined for genomic integrity by microarray genotyping. Parental fibroblasts and 8 iPSC clones for each patient were genotyped using BeadChip Illumina microarray platform. Copy Number Variants (CNVs) were called using PennCNV (v1.0.3) (Wang et al., 2007) with default parameters. Deletions or duplications were stitched as previously described (Marshall et al., 2017; Sanders et al., 2015). Briefly, variants were joined if the gap between two CNVs of the same type was less than 50% of the number of

markers within the larger CNV. This rule was applied recursively until no more CNVs could be stitched. Only CNVs of over 100 kbp in size were retained for the subsequent analysis. In addition, if over 50% of the CNV overlapped with the regions that can confound CNV calling (such as assembly gaps, segmental duplications, centromeres, and telomeres), they were omitted from the analyses. We also removed CNVs if the number of markers supporting the call was less than 8 and/or if the PennCNV confidence score was less than 20. After applying these filters, we confirmed the presence of 16p11.2 DELs or DUPs in all fibroblast and iPSC clones. We then sought to remove those iPSC clones, for which CNV burden was significantly greater than that of parental fibroblasts. To compare iPSC clones and fibroblasts, we defined CNV burden as a total sum of base pairs that are copy number variable (excluding 16p11.2 CNV). For each patient, we defined the mean CNV burden as the CNV burden in fibroblasts, and standard deviation as the burden in all iPSC clones from the same patient. We then compared CNV burden between fibroblasts and iPSC clones for each patient, and discarded clones with the Z-scores > 1.5 SD. Most clones passed these filtering steps. Two iPSC clones with CNV burden closest to the parental fibroblasts of each patient were used for subsequent experiments.

Generation of cortical organoids

To generate cortical organoids from iPSCs, we used the protocol described in Trujillo et al. (Trujillo et al., 2019). Briefly, feeder-free iPSCs at passage 15 or later were fed daily with mTeSR1 for at least 7 days before differentiation. Colonies were dissociated using Accutase (Life Technologies) in PBS (1:1) for 10 minutes at 37°C and centrifuged for 3 minutes at 100x*g*. The cell pellet was resuspended in mTeSR1 supplemented with 10 μM SB431542 (SB, Tocris) and 1 μM Dorsomorphin (Dorso, R&D Systems). Approximately 5×10^6 cells were transferred to each well of a 6-well plate and kept in suspension under rotation (95 rpm) in the presence of 5 μM ROCK inhibitor (StemCell Technologies) for 24 hours to form free-floating spheres. Then, the media was replaced with mTeSR1 for additional 48 hours. After 72 hours, Media1 [Neurobasal (Life Technologies) supplemented with Glutamax, 2% Gem21 NeuroPlex (Gemini Bio-Products), 1% N2 NeuroPlex (Gemini Bio-Products), 1% MEM nonessential amino acids (NEAA, Life Technologies), 1% penicillin/streptomycin (PS; LifeTechnologies), 10 μM SB and 1 μM Dorso] was used for maintenance for 7 days, with media changes every other day. Subsequently, Media1 was replaced with Media2 [Neurobasal with Glutamax, 2% Gem21 NeuroPlex, 1% NEAA and 1% PS] supplemented with 20 ng/mL FGF2 (Life Technologies) for additional 7 days. Then, Media2 was supplemented with both 20 ng/mL FGF2 and 20 ng/mL EGF (Life Technologies) and spheres were cultured for additional 7 days with media changes every other day. Next, organoids were

transferred into Media3 [Media2 supplemented with 10 ng/mL BDNF, 10 ng/mL GDNF, 10 ng/mL NT-3 (all from Life Technologies), 200 μM L-ascorbic acid (Tocris) and 1 mM dibutyryl-cAMP (StemCell Technologies)] for another 7 days with media changes every other day. After 28 days, cortical organoids were maintained in Media2 for as long as needed, with media changes every 3-4 days. All organoids were generated, grown and used for all experiments in the same plate with one DEL, one DUP and one CTRL (called a “batch” thereafter) to reduce batch effect from genotypes.

Mycoplasma testing

All iPSC and organoid cultures were routinely tested for mycoplasma by PCR. Media supernatants (with no antibiotics) were collected, centrifuged, and resuspended in saline buffer. Ten microliters of each sample were used for Mycosplama testing using commercially available LookOut Mycoplasma PCR Detection Kit (Sigma Aldrich) following manufacturer’s instructions.

Organoid size analysis

The diameter of individual organoids was measured using ImageJ software. Size measurements for organoid batches (1 DEL, 1 DUP and 1 CTRL) followed normal distribution, as verified with Prism software (GraphPad). For size comparison, the “large” group was defined as a proportion of organoids with size higher than 1 Standard Deviation (SD) within the batch; the “small” group as a proportion of organoids with size lower than 1 SD within the batch, and the “medium” group comprised the remaining organoids. The proportions for each batch were averaged for final quantification.

Immunofluorescence staining

Cortical organoids were fixed in 4% paraformaldehyde (PFA) overnight. Next morning they were washed in PBS, transferred to a 30% sucrose solution and kept at 4°C. After the 3D structures sink, they were transferred into Tissue-Tek OCT medium (Sakura). Subsequently, 20 μm thick sections were obtained using a cryostat. For immunostaining of iPSC clones, cells were grown directly on Matrigel-coated coverslips.

Slides containing organoid slices were air-dried and then washed with PBS to remove excess OCT. Permeabilization and blocking was performed with 4% Fetal Bovine Serum (FBS, Gibco), 0.1% Triton X-100 (Sigma Aldrich) diluted in PBS for one hour at room temperature. The

slides were then incubated overnight at 4°C with primary antibodies diluted in solution containing 4% FBS and 0.1% Triton X-100. PBS was used to wash the primary antibodies and the slides were incubated with secondary antibodies in solution containing 4% FBS for 1h at room temperature. The following primary antibodies were used for immunostaining: NANOG (goat, 1:500, R&D Systems), Oct4 (mouse, 1:500, Abcam), Tra-1-60 (mouse, 1:500, Abcam), Lin28 (rabbit, 1:500, Abcam), PAX6 (mouse, 1:300; DSHB), NeuN (mouse, 1:500, Millipore), NeuN (rabbit, 1:500, Cell Signaling), MAP2 (chicken, 1:2000; Abcam), Ki67 (rabbit, 1:1000, Abcam), β-tubulin III (mouse, 1:500, Abcam), Synapsin I (rabbit, 1:500, Abcam), SOX2 (rabbit, 1:500, Abcam). Alexa Fluor Dyes (Abcam) were used at 1:1000 dilution as secondary antibodies. Nuclei were visualized with Hoechst 33258 (1:25000, Life Technologies). EdU was visualized using the Edu Staining Kit (Abcam) following manufacturer's instructions. Slides were mounted using ProLong Gold antifade reagent (Invitrogen) and analyzed under a fluorescence microscope (Leica SP8). Image analysis was performed with ImageJ software. All cells expressing a particular marker were counted on sections and normalized to the total number of cells.

Cell migration assay

For the *in vitro* migration assay, intact organoids were seeded in Matrigel-coated 24-well plates (3-4 organoids per well), using Media2. Organoids were allowed to attach to the bottom of the plate for 24h, then media was replaced with fresh Media2 carefully not to disrupt organoids attachment. After 72h, immunostaining was performed. Images were taken using EVOS FL Cell Imaging System. Cell counting and image analysis were performed with ImageJ software.

For live-imaging, intact organoids were seeded in Matrigel-coated p35 glass-bottom culture dishes (Greiner). After 24h, pictures were taken every 15min using a LEICA SP8 microscope. Videos were mounted using LEICA SP8 software.

Organoids dissociation

Cortical organoids were dissociated into single-cell suspension using Accumax (Sigma Aldrich) for 30min at 37°C with rotation (95 rpm). Then, organoids were disaggregated using a 1000µl pipette tip, incubated for another 10min at 37°C in suspension with rotation (95 rpm), and centrifuged for 3 minutes at 100x*g*. The cell pellet was resuspended in Media2 containing 5 µM of ROCK inhibitor, filtered through a 100µm mesh (Gibco) and centrifuged again for 3 minutes at 100x*g*. To further remove undissociated organoid tissue, the procedure was repeated but with

filtering through the 40µm mesh (Gibco). Cells from suspension were counted using a Bio-Rad TC10 Cell Counter.

Boyden chamber assay

Approximately 3 × 10^5 cells from the dissociated organoids were seeded on top of a Millicell Cell Culture 8µm Insert (Millipore) in 24-well plates. The bottom of the Cell Culture Insert was filled with 500µl of Media2 supplemented with 20 ng/mL of FGF2 and 10 ng/mL of BDNF as chemoattractants. Cells were left to freely migrate for 24h, then washed with PBS and fixed with 4% PFA for immunostaining.

After immunostaining, images were taken using EVOS FL Cell Imaging System, visualizing all cells on the Cell Culture Inserts. Then, cells on the top of the Cell Culture Insert were removed using a cell scrapper. After three washes with PBS, only cells on the bottom of the insert were visualized. Cell counting was performed with ImageJ software.

Pharmacological treatment of cortical organoids

For phenotype rescue experiments, organoids where grown in Rhosin-treated media. Rhosin (Tocris) was added to the media during differentiation stage starting from day 6 (Rhosin was first added to second Media1 to the final concentration of 1 µM). The same amount of Rhosin was added during all subsequent media changes. The organoids were grown to 1 month, at which cell migration assays were carried out. An equivalent amount of vehicle (0.1% Dimethylsulfoxide, DMSO) was added to grow untreated CTRL, DEL and DUP organoids.

Neuronal morphology analysis

Cortical organoids were dissociated and approximately 3 × 10^5 cells per well were seeded on a 24-well plate coated with poly-ornithine (Sigma Aldrich) and Laminin (Invitrogen). Media2 was changed after 24h to remove ROCK inhibitor, and second media change was performed after 3 days. Cells were fixed seven days after seeding for immunostaining. Images were taken with EVOS FL Cell Imaging System and analyzed with ImageJ software. For soma area calculation, the perimeter of the MAP2-positive cell body was manually outlined and measured. For total dendrite length, each dendrite or its branch was traced separately, and the dendrite length was calculated by adding individual lengths for every neuron.

Synaptic puncta quantification

Three channel z-stack images of organoid slices were taken using an oil-inverted 60x objective. Then, an average projection image of each stack was generated. At least six randomly selected image fields for each genotype from two different batches were used for quantification of number of synapses with Synapsin I staining. Only puncta overlapping MAP2-positive processes were scored. The number of Synapsin I puncta was quantified using a plug-in Puncta Analyzer from the Fiji analysis software platform (Ippolito and Eroglu, 2010).

Western Blot

Cortical organoids from a quarter to a half of a well were washed once with ice cold PBS (w/o Ca^{2+}/Mg^{2+}). Proteins were extracted using lysis buffer (20mM Tris, pH 7.4, 140mM NaCl, 10% glycerol, 2mM EDTA, 1mM EGTA, and 1% Triton X-100) supplemented with EDTA-free Complete protease inhibitor cocktail (Roche) and Phosphatase Inhibitor cocktail (Sigma Aldrich). The suspension was centrifuged at 16,000x*g* at 4°C for 30min, and supernatants were collected. Protein concentration was quantified by a modified Lowry assay (DC protein assay; Bio- Rad). Cell lysates were resolved by SDS-PAGE and transferred onto PVDF Immobilon-P membranes (Millipore). After blocking with 1X TBS, 0.1% Tween-20 containing 5% nonfat dry milk for 1h at room temperature, membranes were first probed with primary antibodies, and then after 1h of incubation with corresponding peroxidase-conjugated secondary antibody (Abcam). Membranes were developed using the EZ-ECL chemiluminescence detection kit (Denville Scientific). The following primary antibodies were used: anti-KCTD13 (1:500; Atlas Antibodies), anti-RhoA (1:1000; Cell Signaling), and anti-β-actin (1:5000; Sigma Aldrich) as a loading control. Quantification was performed by densitometry with ImageJ software.

RNA isolation for RNA-Seq and qPCR

Total RNA was extracted from undifferentiated iPSCs or cortical organoids at 1 month and 3 months of differentiation. Two clones from each patient were used for RNA isolation, for each time point analyzed (**Fig. S2**).

Total RNA was extracted using the QIAGEN RNAeasy isolation kit (QIAGEN) following manufacturer's instructions. RNA sequencing was performed using the same input amount of total RNA for each sample. RNA samples were ribodepleted using Ribo-Zero rRNA Removal Kit (Illumina) and library preparation was performed using the TrueSeq Stranded Total RNA kit for

Illumina Sequencing according to the manufacturer's instructions. Paired-end RNA sequencing (2x150bp) was performed on an Illumina HiSeq4000 to an average depth of 40M reads per sample.

For qPCR experiments, cDNA was synthesized, starting from 100ng of total RNA with the SuperScript III First-Strand Synthesis kit and random hexamers (Invitrogen). qPCR was performed using the CFX96 Touch™ Real-Time PCR Detection System (Bio Rad) using Power SYBR Green PCR Master Mix (Applied Biosystems). HPRT1 and β-actin were used as housekeeping genes for normalization. Fold change in expression was calculated using the ΔΔCt method.

RNA-sequencing Data Processing Pipeline

All 108 FASTQ files (36 iPSC, 36 one month organoids and 36 three months organoids paired-end fastq) (**Fig. S5**) were run through a unified RNA-Seq processing pipeline. Pipeline source code can be found on https://github.com/IakouchevaLab/16p11.2. All fastqs were trimmed for adapter sequence and low base call quality (Phred score < 30 at ends) using Cutadapt (v1.14). Trimmed reads were then aligned to the GRCH37.p13 (hg19) reference genome via STAR (2.5.3a) using comprehensive gene annotations from Gencode (v19) (**Fig. S7A**). Gene-level quantifications were calculated using RSEM (v1.3). Quality control metrics were calculated using RNA-SeQC (v1.1.8), featureCounts (v1.6.), PicardTools (v2.12), and Samtools (v1.3) (**Fig. S7B** and **Table S2**).

RNA-Seq Quality Control and Normalization

RNA-Seq Quality Control and Normalization Expected counts were compiled from gene-level RSEM quantifications and imported into R for downstream analyses. Expressed genes were defined as genes with TPM > 0.5 in at least 80% of samples from each genotype (CTRL, DEL or DUP). Only expressed genes were included into the analysis. A total of 15,788; 13,348, and 13,723 expressed genes from iPSC, 1M old organoids and 3M old organoids, respectively, were used in the downstream analysis using the GENCODE V19 annotation gtf file. Outliers were defined as samples with standardized sample network connectivity Z scores <-2 (Oldham et al., 2012), and were removed (**Fig. S7C**). Highly variable genes between clones from the same individual were filtered out using the Variance Partition (v3.5) R package (Hoffman and Schadt, 2016).

Covariate Selection

We compiled a set of 197 RNA-Seq quality control metrics from the outputs of cutadapt, STAR, RNA-SeQC, featureCounts and PicardTools (CollectAlignmentSummaryMetrics, CollectInsertSizeMetrics, CollectRnaSeqMetrics, CollectGcBiasMetrics, MarkDuplicates) for each group of samples (iPSCs, 1M old organoids and 3M old organoids) (**Table S2**, **Figs. S7A-B**). These measures were summarized by the top principal components, which explained the total variance of each group (**Fig. S8A**). Batch effects and possible hidden confounding factors were detected using the Surrogate Variable Analysis (SVA) (Leek and Storey, 2007). Multivariate Adaptive Regression Splines (MARS) implemented in the earth package in R was used to determine which covariates to include in the final differential expression model (**Fig. S8B**). The potential covariates included: genotype, run/batch, RIN, clone, seqPCs and SVs (**Fig. S8C**). These covariates were inputted into the earth model along with gene expression data (limma voom normalized, centered, and scaled). The model was run using linear predictors and otherwise default parameters. MARS selected SV1 as a covariate for iPSC, SV1 to SV5 as covariates for 1M old organoids, and SV1 to SV6 as covariates for 3M old organoids (**Fig. S8B**).

Differential Gene Expression

Differential Gene Expression (DGE) analyses was performed using limma-voom with 'duplicateCorrelation' function to account for duplicate samples (clones and replicas) from the same individuals, and to avoid pseudo-replication in the analyses. Covariates were included as fixed effects in the model. The biomaRt (Durinck et al., 2005; Durinck et al., 2009) package in R was used to extract gene names, gene biotypes and gene descriptions. Differential genes expression analyses was performed using all three datasets (CTRL, DEL and DUP) for all time points. The volcano plots for DEL *vs* CTRL and DUPs *vs* CTRL are shown in **Fig. S24**, and differentially expressed genes from these datasets are listed in **Table S4**.

Weighted gene co-expression network analysis on RNA-seq data

We used Weighted Gene Co-expression Network Analysis (WGCNA) (Zhang and Horvath, 2005) to define modules of co-expressed genes from RNA-seq data. All covariates except for genotype at the 16p11.2 locus were first regressed out from the expression datasets. The co-expression networks and modules were estimated using the blockwiseModules function with the following parameters: corType=bicorr; networkType=signed; pamRespectsDendro=F; mergeCutHeight=0.1. Some parameters were specific for each dataset. For iPSC data:

power=14; deepSplit=0; minModuleSize=100. For 1M old organoid data: power=16; deepSplit=2; minModuleSize=50. For 3M old organoid data: power=19; deepSplit=2; minModuleSize=70. Module eigengene/genotype associations were calculated using a linear mixed-effects model, using a random effect of individual, to account for multiple clones and replicas derived from the same subject. Significance p-values were FDR-corrected to account for multiple comparisons. Genes within each module were prioritized based on their module membership (kME), defined as correlation to the module eigengene. For selected modules, the top hub genes were shown. Module preservation was tested using the modulePreservation function from the WGCNA package in R.

Enrichment analysis of gene sets

Enrichment for Gene Ontology (GO; Biological Process and Molecular Function) was performed using gProfileR R package (Reimand et al., 2007). Background was restricted to the expressed set of genes by group (iPSC, 1M old organoids and 3M old organoids). An ordered query was used, ranking genes by FDR-corrected p-value for DGE analyses and by kME for WGCNA analyses.

Enrichment analyses were also performed using several established, hypothesis-driven gene sets including syndromic and highly ranked (1 and 2) genes from SFARI Gene database (https://gene.sfari.org/database/gene-scoring/); pre- and post-synaptic genes from SynaptomeDB(Pirooznia et al., 2012); genes with loss-of-function intolerance (pLI) > 0.99 as reported by the Exome Aggregation Consortium (Karczewski et al., 2017); highly constrained genes (Samocha et al., 2014); FMRP targets (Darnell et al., 2011) and CHD8 targets (Wilkinson et al., 2015). Statistical enrichment analyses were performed using permutation test. One thousand simulated lists with similar number of genes, gene length distribution and GC-content distribution as the target gene list were generated, and the overlaps between each of the simulated list and the hypothesis-driven gene sets were calculated to form the null distribution. Significance p-value was calculated by comparing the actual overlap between target list and hypothesis-driven gene sets to the null distribution. All results were FDR-corrected for multiple comparisons.

Cell type enrichment analysis

Cell-type enrichment analysis for each co-expression module was performed using the Expression Weighted Cell Type Enrichment (EWCE) package in R (Skene and Grant, 2016). Cell type-specific gene expression data was obtained from single cell sequencing (scRNA-seq) studies of human fetal neocortex (Polioudakis et al., 2019). The specificity metric of each gene for each cell type was computed as described (Skene and Grant, 2016). “Neuron” cell type includes a union of ExcNeu (excitatory neurons) and IntNeu (interneurons). Enrichment was evaluated using bootstrapping. Z-score was estimated by the distance of the mean expression of the target gene set from the mean expression of bootstrapping replicates. P-values were corrected for multiple comparisons using FDR.

CoNTExT analyses

Regional and temporal identify of organoids was assessed using CoNTExT(Stein et al., 2014) (https://context.semel.ucla.edu/).

Sample preparation, protein identification and quantification by TMT-Mass Spectrometry

TMT mass-spectrometry experiments were performed on the organoid samples from the same well as those used for RNA-seq, by splitting the content of each well into two approximately equal amounts (**Fig. S12**). Organoids were lysed in 100 mM TEAB with 1% SDS, protease inhibitor cocktails (Sigma) and PhosSTOP (Sigma) by 2-3 times of brief probe sonication and then centrifuged at 18,000x*g* for 15 min at 4°C. Supernatants were reduced (10 mM TCEP at 55°C for 20 min) and alkylated (50 mM chloroacetamide at room temperature in the dark for 20 min), and then MeOH/$CHCl_3$ precipitation was performed. Pellets were dissolved by adding 6M urea in 50 mM TEAB, and then LysC/Tryp (Promega) was added by 1:25 (w/w) ratio to the peptides. After 3-4 h incubation at 37°C, reaction mixture was diluted with 50 mM TEAB for urea to be less than 1 M. After the o/n digestion, peptide concentration was estimated by colorimetric peptide BCA assay (Thermo), and the peptides were labelled with TMT 10-plex reagents (Thermo) for one hour, followed by 15 min quenching with hydroxylamine according to the manufacturer’s protocol. Equal amount of reaction mixtures for each channel were pooled together and dried using SpeedVac.

Since the total number of samples exceeded the maximum number of TMT channels, samples were divided into multiple sets (one replicate per each set). To compare and normalize different sets of TMT-labeled samples, pooled peptides were labeled with 131N and 131C as

duplicates, and these samples were commonly included in all sets within each age (1M and 3M old organoids) set. A total of 100µg of peptides were fractionated using Pierce™ High pH reversed-phase peptide fractionation kit (Thermo) and then dried in SpeedVac. Dried peptides were dissolved with buffer A (5% acetonitrile, 0.1% formic acid), and half of each fraction was injected directly onto a 25cm, 100µm-ID column packed with BEH 1.7µm C18 resin (Waters). Samples were separated at a flow rate of 300 nL/min on nLC 1000 (Thermo). A gradient of 1–25% B (80% acetonitrile, 0.1% formic acid) over 200min, an increase to 50% B over 120 min, an increase to 90% B over another 30min and held at 90% B for a final 10min of washing was used for 360min total run time. Column was re-equilibrated with 20µL of buffer A prior to the injection of sample. Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer Orbitrap Fusion by application of 2.8kV voltage at the back of the column.

Fusion was operated in a data-dependent mode. Full MS1 scans were collected in the Orbitrap at 120k resolution. The cycle time was set to 3s, and within this 3s the most abundant ions per scan were selected for CID MS/MS in the ion trap. MS3 analysis with multi-notch isolation (SPS3) (McAlister et al., 2014) was utilized for detection of TMT reporter ions at 60k resolution. Monoisotopic precursor selection was enabled, and dynamic exclusion was used with exclusion duration of 10s. Tandem mass spectra were extracted from the raw files using RawConverter (He et al., 2015) with monoisotopic peak selection. The spectral files from all fractions were uploaded into one experiment on Integrated Proteomics Applications (IP2, Ver.6.0.5) pipeline. Proteins and peptides were searched using ProLuCID (Xu et al., 2015) and DTASelect 2.0 (Tabb et al., 2002) on IP2 against the UniProt *H. sapiens* protein database with reversed decoy sequences (UniProt_Human_reviewed_05-05-2016_reversed.fasta). Precursor mass tolerance was set to 50.0ppm, and the search space allowed all fully-tryptic and half-tryptic peptide candidates without limit to internal missed cleavage and with a fixed modification of 57.02146 on cysteine and 229.1629 on N-terminus and lysine. Peptide candidates were filtered using DTASelect parameters of -p 2 (proteins with at least one peptide are identified) -y 1 (partial tryptic end is allowed) --pfp 0.01 (protein FDR < 1%) -DM 5 (highest mass error 5 ppm) -U (unique peptide only). Quantification was performed by Census(Park et al., 2008) on IP2.

Differential protein expression

Proteomics data was first summarized to peptide level by adding up the intensities of constituting spectra. Quantitation results from different TMT runs were merged and normalized using the pooled samples channel which was present in all runs. For each peptide, multiple

measurements from the same subject were collapsed to one measurement by taking the median of all measurements. The data was then log2 transformed. Differential protein expression (DPE) was calculated by fitting a linear mixed-effects model for each protein, using the lme4 package in R(Bates et al., 2015). Genotype was included as fixed effect in the model. We included a random effect term for each peptide to account for the fact that different peptides from the same protein are not entirely independent. Significance p values were calculated using lmerTest package in R (Kuznetsova et al., 2017). Resulting P-values were FDR-corrected using the Benjamini-Hochberg method to control for multiple comparisons. The volcano plots for DEL *vs* CTRL and DUPs *vs* CTRL are shown in **Fig. S25**, and differentially expressed proteins from these datasets are listed in **Table S8**.

Weighted protein co-expression network analysis

Proteomics data was first summarized to protein level by adding up the intensities of constituting peptides. Quantitation results from different TMT runs were merged and normalized using the pooled samples channel which was present in all runs. The merged data was then log2 transformed. Outlier samples detection, highly variable proteins removal, surrogate variables calculation and covariates selection were subsequently performed using the same methods as described for RNA-seq data processing. All covariates except for genotype at the 16p11.2 locus were first regressed out from the expression datasets. Protein co-expression networks and modules were estimated using the blockwiseModules function with the following parameters: corType=bicorr; networkType=signed; pamRespectsDendro=F; mergeCutHeight=0.1. Some parameters were specific for each dataset. For 1M old organoid data: power=13; deepSplit=3; minModuleSize=40; and for 3M old organoid data: power=17; deepSplit=2; minModuleSize=10. Module eigengene/genotype associations were calculated as described for the RNA-seq WGCNA. Module preservation was tested using the modulePreservation function from the WGCNA package in R.

Quantification and statistical analyses

The statistical analyses for above experiments were performed using Prism software (GraphPad). Statistical tests used and exact values of n are described in Figure legends. Significance was defined as $p < 0.05$(*), $p < 0.01$(**), or $p < 0.001$(***). Blinded measurements were performed for any comparison between control and 16p11.2 genotypes. The samples used for each type of experiments are shown in **Table S12**.

Data availability

Source RNA-seq data is available at GEO repository accession number GSE142174 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE142174).

Source proteomics data is available from the public repository MassIVE (Mass Spectrometry Interactive Virtual Environment), a part of the ProteomeXchange consortium, with the identifier MSV000084727 (and PXD016855 for ProteomeXchange) through the following link (https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?accession=MSV000084727).

Materials availability

New induced pluripotent stem cells (iPSCs) lines generated in this study will be deposited after publication to Rutgers repository by the Simons Foundation for Autism Research that funded this study.

Acknowledgments

This work was supported by a grant to LMI and ARM from the Simons Foundation for Autism Research (#345469), and by grants from the National Institute of Mental Health to LMI and ARM (MH109885, MH108528), to LMI (MH105524, MH104766), to JRY and ARM (MH100175), and to JS (MH119746). We thank Gabriel Hoffman, Karen Messer, Minya Pu, and Ruifeng Chen for suggestions regarding data analyses. We thank Lucas Bazier, Nicholas Chew and Alexander Sun for help with image analysis; and Jiaye Chen for help with uploading transcriptomic data to GEO database. RNA-seq data was generated at the UC San Diego IGM Genomics Center, University of California San Diego (grant P30CA023100). The images were acquired at the UCSD School of Medicine Microscopy Shared Facility (grant NS047101).

Author Contributions

L.M.I. and A.R.M. conceived the study. J.U., P.Z., P.M-L., N-K.Y., P.D.N., C.A.T., D.A., J.S, J.R.Y. III, A.R.M. and L.M.I. designed the experiments and analyses. J.U., N-K.Y., P.D.N., C.T., M.A., J.D., L.T., and S.R. performed experiments and analyses. P.Z., P.M-L., D.A., K.C., and A.B.P. performed computational data processing and analyses. J.U. and L.M.I. wrote the manuscript, with input from all co-authors. Supervision was performed by J.S., J.R.Y. III, A.R.M, and L.M.I.

Corresponding authors

Correspondence to Lilia M. Iakoucheva (lilyak@ucsd.edu) and Alysson R. Muotri (muotri@ucsd.edu).

Ethics declarations

Dr. Muotri is a co-founder and has equity interest in TISMOO, a company dedicated to genetic analysis and human brain organogenesis, focusing on therapeutic applications customized for autism spectrum disorders and other neurological disorders origin genetics. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

Figure titles and legends

Figure 1. Generation and characterization of cortical organoids from 16p11.2 patient and control iPSCs

(A) Overview of Study Design and Analyses.

(B) Comparison of iPSCs and organoids transcriptomes with the human developing brain transcriptome using CoNTExT (Stein et al., 2014). The samples from 3 individuals of each genotype (CTRL, DEL, DUP), with 2 clones per individual and 2 replicates per clone (n=36 datasets) are shown for each group (iPSC, 1M and 3M organoids). PCW – post conception weeks, M – months, Y-years.

(C) Predicted laminar transitions for 1M and 3M organoids using TMAP (Stein et al., 2014) and the transcriptome of laser capture microdissected cortical laminae from postmortem human fetal brain (15–21 PCW) (Miller et al., 2014). Rank-rank hypergeometric overlap (RRHO) maps for CTRL organoids (n=12 datasets) from 3 patients, 2 clones per patient, 2 replicates per clone are shown, with CTRL iPSCs (n=12 datasets) used as a second time point. Each pixel represents the overlap between fetal brain and organoids transcriptome, color-coded according to the $-\log_{10}$ p-value of a hypergeometric test. On each map, the extent of shared upregulated genes is displayed in the bottom left corner, whereas shared downregulated genes are displayed in the top right corners.

(D) Representative images of cortical organoids for each genotype (CTRL, DEL, DUP) at different time points of differentiation: Induction (6 days of differentiation), Proliferation (16 days of differentiation), Early maturation (1M of differentiation) and Late maturation (3M of differentiation).

(E) Analysis of size differences between cortical organoids of each genotype (CTRL, DEL and DUP) at different time points of differentiation. Organoids diameter (n>100 for each genotype) was measured using ImageJ, and average diameter for each genotype was computed. P-values were calculated using one-way ANOVA. Stars above DEL and DUP represent comparison to CTRLs. ****p<0.0001, ***p<0.001, **p<0.01, ns – p>0.05.

Figure 2. Differential gene expression and gene co-expression analyses of iPSCs, 1M and 3M cortical organoids.

(A-C) Volcano plots of differentially expressed genes between DEL and DUP iPSCs, 1M and 3M organoids. Genes within 16p11.2 CNV are colored in pink. Genes colored in orange are upregulated in DUP and downregulated in DEL; genes colored in blue are downregulated in DUP and upregulated in DEL, in DUP *vs* DEL comparison. Gene Ontology enrichment analyses are shown as bar plots on the right. The contribution of up- or down-regulated genes to specific GO terms are shown in blue and red, respectively.

(D) Hierarchical clustering of gene co-expression modules by module eigengene. Module-genotype associations (* FDR<0.1) are shown below each module. Module enrichment analyses against literature-curated gene lists with previous evidence for involvement in autism are shown at the bottom (* FDR<0.05). The lists include syndromic and highly ranked (1 and 2) genes from SFARI Gene database (https://gene.sfari.org/database/gene-scoring/); pre- and post-synaptic genes from SynaptomeDB (Pirooznia et al., 2012); genes with probability of loss-of-function intolerance (pLI)>0.99 as reported by the Exome Aggregation Consortium (Karczewski et al., 2017); constrained genes (Samocha et al., 2014); FMRP target genes (Darnell et al., 2011), and CHD8 target genes (Wilkinson et al., 2015).

Figure 3. Proteomic analyses of 1M and 3M cortical organoids and correlation of protein expression with transcriptome.

(A-B) Volcano plots of differentially expressed proteins between DEL and DUP 1M and 3M organoids derived from LC-MS/MS with TMT labeling proteomic profiling. Proteins within 16p11.2 CNV are colored in pink. Proteins colored in orange are upregulated in DUP and downregulated in DEL; proteins colored in blue are downregulated in DEL and upregulated in DUP. Gene

Ontology enrichment analyses are shown as bar plots on the right. The contribution of up- or down-regulated proteins to specific GO terms are shown in blue and red, respectively.

(C) Correlation of differentially expressed genes and proteins in 1M organoids. Genes/proteins within 16p11.2 CNV locus are colored in pink.

(D) Module preservation analyses for migration, neuronal/synaptic and 16p11.2 modules detected in transcriptomic profiling of 1M organoids as compared to other datasets (iPSC and 3M transcriptomic, and 1M and 3M proteomic datasets). The neuronal/synaptic module *vs* 1M Protein is located behind 16p11.2 module and is not visible.

(E) Comparison of differentially expressed genes and proteins from this study with differentially expressed genes from other relevant studies. Idiopathic_ASD_TD11 and Idiopathic_ASD_TD31 are differentially expressed genes from organoids derived from individuals with idiopathic ASD at 11th and 31st day of differentiation from Mariani study (Mariani et al., 2015); 16p11.2_Lymph_Human and 16p11.2_Cortex_Mouse are differentially expressed genes from human lymphoblastoid cell lines of ASD patients with 16p11.2 CNV and 16p11.2 deletion mouse cortex, respectively, from Blumenthal study (Blumenthal et al., 2014); CHD8_NPC, CHD8_NEURON and CHD8_ORGANOID are differentially expressed genes from isogenic CRISPR/Cas9 generated heterozygous CHD8 neural progenitors, monolayer neurons and organoids from Wang study (Wang et al., 2017). Number of overlapped genes and odds ratio (in parenthesis) are indicated inside each cell, and provided only for FDR<0.05 and OR>1.

Figure 4. Cell type enrichment analyses of 1M and 3M organoid transcriptomes.

(A) Cell type enrichment analyses of RNA-seq co-expression modules from 1M and 3M old organoids using cell types from scRNA-seq of fetal neocortex (Polioudakis et al., 2019). Modules significantly enriched in at least one cell type are shown. “Neuron” category includes combination of excitatory and interneurons; IP – intermediate progenitors; RG – radial glia; CycProg – cycling progenitors. Enrichment was evaluated using bootstrapping. Z-score was estimated by the distance of the mean expression of the target gene set from the mean expression of bootstrapping replicates and is shown inside each cell. P-values were corrected for multiple comparisons using FDR.

(B-E) PC1 of enriched organoid modules plotted by genotype at two time points. PC1 at 1M and 3M was computed for a union of genes from all modules significantly enriched in a specific cell type. All comparisons between DELs and DUPs are significant using t-test statistics. ***p<0.001, **p<0.01. GO terms for one representative most enriched module from 1M dataset were obtained using g:Profiler and are shown on the right side.

Figure 5. Altered neuronal maturation, morphology and synaptic defects in 16p11.2 cortical organoids.

(A-B) Module eigengene and GO functional annotations for *46brown4* transcriptomic module in 1M organoids. Dots (n=12) correspond to each replicate derived from 3 patients (CTRL, DEL or DUP), 2 clones per patient and 2 replicates per clone. Two replicates were removed from CTRL before the analyses during outlier detection procedure (see **Methods**).

(C) Heat plot representing gene (RNA) and protein expression from *46brown4* module. Pearson correlation coefficient between RNA and protein expression and the proportion of genes whose protein products were also detected by proteomics are shown below the plot.

(D) Twenty top hub genes from *46brown4* module. Edges represent co-expression.

(E-F) Representative images of 1M organoid slices (CTRL, DEL and DUP) immunostained with DAPI, PAX6 and NeuN. Quantification of the percentage of positive cells for each marker is shown. Symbols represent organoids from the same differentiation batch, where batch is defined as CTRL, DEL, DUP from one patient, one clone and one replica. Data is presented as mean ± SEM (n=2 patients per genotype, at least 4 organoids per patient). Significance is calculated using one-way ANOVA; ***p<0.001, **p<0.01, *p≤0.05. Significance above bars represents comparison against CTRL.

(G-H) Representative images of neurons from dissociated 1M organoids immunostained with DAPI (blue) and MAP2 (green). Quantification of total neurite length and soma size is shown. Symbols represent neurons derived from organoids from the same differentiation batch. Data is presented as mean ± SD (n=2 patients per genotype, at least 15 neurons per patient). Significance is calculated using one-way ANOVA; *p≤0.05. Significance above bars represents comparison against CTRL.

(I-J) Representative images of 1M organoid slices immunostained with DAPI, MAP2 and SynI. Quantification of the total Synapsin I to estimate synaptic puncta is shown. Symbols represent organoids from the same differentiation batch. Data is presented as mean ± SEM (*n*=2 patients per genotype, at least 3 organoids per patient). Significance is calculated using one-way ANOVA; ***$p<0.001$. Significance above bars represents comparison against CTRL.

Figure 6. Neuronal migration defects in 16p11.2 cortical organoids.

(A-B) Module eigengene and GO functional annotations for *22darkgreen* transcriptomic module in 1M organoids. Dots (*n*=12) correspond to each replicate derived from 3 patients (CTRL, DEL or DUP), 2 clones per patient and 2 replicates per clone. Two replicates were removed from CTRL before the analyses during outlier detection procedure (see **Methods**).

(C) Heat plot representing gene (RNA) and protein expression from *22darkgreen* module. Pearson correlation coefficient between RNA and protein expression and the proportion of genes whose protein products were also detected by proteomics are shown below the plot.

(D) Twenty top hub genes from *22darkgreen* module. Edges represent co-expression.

(E-F) Representative images of 1M cortical organoids 72hrs after attachment to Matrigel-coated plates, immunostained with DAPI and β-tubulin III. Graph represents percentages of migrating cells to a distance with 200-μm bins of displacement (n=2 patients per genotype, at least 5 organoids per patient).

(G-H) Representative images of cells migrating from 1M dissociated organoids in Boyden chamber experiment. Immunostaining was performed with DAPI. Quantification of the percentages of migrated cells is shown. Symbols represent cells derived from organoids from the same differentiation batch. Data is presented as mean ± SD (*n*=2 patients per genotype, at least 5 images per patient). Significance is calculated using one-way ANOVA; ***$p<0.001$, **$p<0.01$. Significance above bars represents comparison against CTRL.

Figure 7. Rhosin treatment rescues neuronal migration deficits in 16p11.2 cortical organoids.

(A-B) Representative images of Western Blot analysis of 1M organoids for KCTD13, actin as loading control, total RhoA, and active RhoA (RhoA-GTP). Densitometry analysis of Western Blot is shown. Data is represented as mean ± SEM (*n*=7 differentiation batches with at least one batch for each patient or control for KCTD13, n=6 batches with at least one batch for each patient or control for total RhoA, n=4 batches from two patients or controls for active RhoA). Significance is calculated using ratio paired t-test; *$p \leq 0.05$. Significance above bars represents comparison against CTRL.

(C) Schematic representation of constitutive Rhosin treatment of organoids during differentiation.

(D-E) Representative images of 1M vehicle- or Rhosin-treated cortical organoids 72hrs after attachment to Matrigel-coated plates, immunostained with DAPI and β-tubulin III. Graph represents percentages of migrating cells to a distance with 200-µm bins of displacement (n=2 patients per genotype, at least 5 organoids per patient).

(F-G) Representative images of cells migrating from vehicle- or Rhosin-treated 1M dissociated organoids in Boyden chamber experiment. Immunostaining was performed with DAPI. Quantification of the percentages of migrated cells is shown. Symbols represent cells derived from organoids from the same differentiation batch. Data is presented as mean ± SD (*n*=2 patients per genotype, at least 5 images per patient). Significance is calculated using one-way ANOVA; **$p<0.01$, *$p<0.05$. Significance above bars represents comparison against CTRL.

Supplementary Figures

Figure S1. Immunohistochemical validation of iPSC's pluripotency. Representative image of patient-derived iPSCs immunostained with DAPI, NANOG and Lin28 (left), or DAPI, Tra-1-60 and Oct4 (right).

Figure S2. Transcriptional validation of iPSCs pluripotency. Six pluripotency markers were quantified in patient-derived fibroblasts and corresponding iPSCs by qRT-PCR. Graph shows the average of three biological replicates for each cell line. 16pA, 16pB and 16pC are lines derived from the 16p11.2 DEL patients; and 16pX, 16Y and 16pZ are lines derived from 16p11.2 DUP patients.

Figure S3. CNV burden analysis of patient-derived iPSC clones. Microarray genotyping of fibroblasts and iPSCs and CNV burden analyses using PennCNV is shown. The presence of 16p11.2 CNV in DELs and DUPs was confirmed in all fibroblast and iPSC clones, and 16p11.2

CNV was removed in subsequent burden analyses. Patients' mean CNV burden is defined as the CNV burden in fibroblasts, and standard deviation as the burden in all iPSC clones from the same patient. CNV burden between fibroblasts and iPSC clones for each patient was compared. Only iPSC clones with a CNV burden score <1 SD *vs* patients fibroblasts are shown. Fibroblasts CNV burden score is shown for reference. Arrows point to the iPSC clones selected for cortical organoids production.

Figure S4. Time course of the cell type markers. Transcriptional analysis of three cell type markers (NANOG, PAX6 and MAP2) in 16p11.2 patient-derived iPSCs, 1M and 3M cortical organoids by qRT-PCR. Graph shows the average of all patient cell lines (*n*=6), with three biological replicates per patient.

Figure S5. RNA-seq experimental design and data analysis workflow. A total of 108 transcriptomes have been sequenced in this study. A rigorous quality control including principal component analyses, sample connectivity analyses, surrogate variable analyses and multivariate adaptive regression spline (MARS) for covariate selection has been performed. The Limma-voom model with duplicate correlation function has been applied to account for clone replicates derived from the same patient in order to avoid pseudo-replication in the differential gene expression analyses.

Figure S6. Predicted laminar transitions for 1M and 3M organoids compared to fetal neocortex. The Figure was produced using TMAP (Stein et al., 2014) and the transcriptome of germinal zones of six 13–16 PCW human fetal neocortices (Fietz et al., 2012). Rank-rank hypergeometric overlap (RRHO) maps for CTRL organoids (n=12 datasets) from 3 patients, 2 clones per patient, 2 replicates per clone are shown, with CTRL iPSCs (n=12 datasets) used as a second time point. Each pixel represents the overlap between fetal brain and organoids transcriptome, color-coded according to the $-\log_{10}$ p-value of a hypergeometric test. On each map, the extent of shared upregulated genes is displayed in the bottom left corner, whereas shared downregulated genes are displayed in the top right corners. Ventricular zone (VZ), Inner Subventricular zone (ISVZ), Outer Subventricular zone (OSVZ) and Cortical Plate (CP) are shown.

Figure S7. Initial quality control metrics for RNA-seq data. (A) Sequencing metrics from STAR (2.5.3a) for each group of samples (iPSCs, 1M and 3M old organoids). **(B)** Sequencing metrics

from PicardTools (v2.12) for each group of samples (iPSC, 1M and 3M organoids). **(C)** Sample outlier removal performed with WGCNA package in R for each group of samples (iPSCs, 1M and 3M old organoids) based on Z-scores of standardized network connectivity. Outliers were defined as samples with Z scores of <(-2).

Figure S8. Quality control metrics for RNA-seq data. (A) Correlation plots among the top seqPCs (PCs that summarize the RNA-Seq QC metrics), surrogate variables selected by MARS, and other potential covariates (Run, RIN, Individual, Lab, Clone, Replica, Genotype, Z score) for each group of samples (iPSCs, 1M and 3M old organoids) generated by corrplot package in R. The spearman correlation coefficients values correspond to the areas of the circles. The legend shows the spearman correlation coefficient rho values. **(B)** Covariates for each group of samples (iPSCs, 1M and 3M old organoids) selected by MARS (implemented in earth package in R). **(C)** First two principal components (PCs) of gene expression values, calculated using "prcomp" function in R, are shown before (left panel) and after (right panel) covariate correction for the iPSC, 1M and 3M organoids. Colors represent saples (iPSCs, 1M and 3M organoids), and shapes represent genotypes: CTRL (circles), DEL (triangles) and DUP (squares).

Figure S9. Gene co-expression modules for iPSCs. Upper panel: WGCNA cluster dendrogram for iPSCs. Bottom panel: Module eigengene-genotype association. Comparisons are made between each genotype (DEL or DUP) and control (CTRL) for each module. Rows are genotypes (relative to CTRL) and columns are modules. Number in each tile is a beta value from linear mixed effect model, and color of each tile indicates statistical significance (see **Methods**). A total of 7 modules were significantly associated with DEL or DUP genotypes in iPSCs.

Figure S10. Gene co-expression modules for 1M old organoids. Upper panel: WGCNA cluster dendrogram for 1M old organoids. Bottom panel: Module eigengene-genotype association. Comparisons are made between each genotype (DEL or DUP) and control (CTRL) for each module. Rows are genotypes (relative to CTRL) and columns are modules. Number in each tile is a beta value from linear mixed effect model, and color of each tile indicates statistical significance (see **Methods**). A total of 7 modules were significantly associated with DEL or DUP genotypes in 1M old organoids.

Figure S11. Gene co-expression modules for 3M old organoids. Upper panel: WGCNA cluster dendrogram for 3M old organoids. Bottom panel: Module eigengene-genotype

association. Comparisons are made between each genotype (DEL or DUP) and control (CTRL) for each module. Rows are genotypes (relative to CTRL) and columns are modules. Number in each tile is a beta value from linear mixed effect model, and color of each tile indicates statistical significance (see **Methods**). A total of 21 modules were significantly associated with DEL or DUP genotypes in 3M old organoids.

Figure S12. 16p11.2 gene modules in iPSCs, 1M and 3M old organoids. The significantly associated with genotype 16p11.2 module was detected in all datasets (*10purple* in iPSCs, *11greenyellow* in 1M organoids, and *16lightcyan* in 3M old organoids). Left column of each panel: module-trait association. Middle column of each panel: module eigengene expression for DEL, DUP and CTRL datasets. Dots correspond to each replicate derived from 3 patients (CTRL, DEL or DUP), 2 clones per patient and 2 replicates per clone. Some replicates were removed before the analyses during outlier detection procedure (see **Mehtods**). Right column of each panel: top 20 hub genes (based on kME) from each module are shown.

Figure S13. Proteomics experimental design and data analysis workflow. A total of 72 proteomes have been processed in this study by LC-MS/MS with TMT 11-plex labeling. Protein Quantification was carried out by Census. A rigorous quality control including principal component analyses, sample connectivity analyses, surrogate variable analyses and multivariate adaptive regression spline (MARS) for covariate selection has been performed. Linear mixed effect model (LME) was implemented for differential protein expression analyses.

Figure S14. Protein co-expression modules for 1M old organoids. Upper panel: WPCNA cluster dendrogram for 1M old organoids. Bottom panel: Module eigengene-genotype association. Comparisons are made between each genotype (DEL or DUP) and control (CTRL) for each module. Rows are genotypes (relative to CTRL) and columns are modules. Number in each tile is a beta value from linear mixed effect model, and color of each tile indicates statistical significance (see **Methods**). A total of 5 modules were significantly associated with DEL or DUP genotypes in 1M old organoids.

Figure S15. Protein co-expression modules for 3M old organoids. Upper panel: WPCNA cluster dendrogram for 3M old organoids. Bottom panel: Module eigengene-genotype association. Comparisons are made between each genotype (DEL or DUP) and control (CTRL) for each module. Rows are genotypes (relative to CTRL) and columns are modules. Number in

each tile is a beta value from linear mixed effect model, and color of each tile indicates statistical significance (see **Methods**). A total of 7 modules were significantly associated with DEL or DUP genotypes in 3M old organoids.

Figure S16. Protein co-expression modules enrichment analyses. Hierarchical clustering of protein co-expression modules by module eigengene is shown. Module-genotype associations (* FDR<0.1) are shown below each module. Module enrichment analyses against literature-curated gene lists with previous evidence for involvement in autism are shown at the bottom (* FDR<0.05). The lists include syndromic and highly ranked (1 and 2) genes from SFARI Gene database (https://gene.sfari.org/database/gene-scoring/); pre- and post-synaptic genes from SynaptomeDB (Pirooznia et al., 2012); genes with probability of loss-of-function intolerance (pLI)>0.99 as reported by the Exome Aggregation Consortium (Karczewski et al., 2017); constrained genes (Samocha et al., 2014); FMRP target genes (Darnell et al., 2011), and CHD8 target genes (Wilkinson et al., 2015).

Figure S17. Correlation of differentially expressed genes and proteins in 3M organoids. Genes/proteins within 16p11.2 CNV locus are colored in pink.

Figure S18. Module preservation analyses between gene co-expression and protein co-expression modules. Module preservation scores of gene co-expression modules (vs. corresponding protein co-expression modules) for 1M old organoids (upper panel) and 3M old organoids (bottom panel) are shown. Gene modules significantly associated with genotypes (DEL and DUP) are marked. Z-scores above 2 are considered to be conserved, and above 10 are highly conserved.

Figure S19. Module preservation analyses for migration, neuronal/synaptic and 16p11.2 modules from 3M old organoids. Module preservation scores of 16p11.2 (*16lightcyan*), neuronal/synaptic (*25orange*) and migration (*32violet*) gene co-expression modules (vs. iPSC and 1M transcriptomic, and 1M and 3M proteomic modules) for 3M old organoids are shown. Z-scores above 2 are considered to be conserved.

Figure S20. Proliferation rate in 1M old organoids. Upper panel: **r**epresentative images of 1M organoid slices immunostained with DAPI, Ki67 and Edu. Bottom panel: quantification of the percentage of positive cells for each marker in each genotype and cell cycle exit ratio. Symbols

represent organoids from the same differentiation batch, where batch is defined as CTRL, DEL, DUP from one patient, one clone and one replica. Data is presented as mean ± SEM (*n*=2 patients per genotype, at least 4 organoids per patient). Significance is calculated using one-way ANOVA; ***$p<0.001$, **$p<0.01$, *$p<0.05$. Significance above bars represents comparison against CTRL.

Figure S21. Impaired migration in DEL and DUP organoids. Upper panel: **r**epresentative images of neurons migrating out of Matrigel-attached organoids at the indicated time points after the start of time-lapse. Arrows mark individual neurons. Bottom panel: quantification of total distance traveled by individual neurons (mean±SEM; one-way ANOVA, ***p<0.001, *p<0.05; n=4 neurons per organoid, 3-4 organoids per genotype). Tracing of cell movement of individual representative neurons for each genotype is shown on the right. Each dot represents location of the neuron after 1h time period.

Figure S22. Immunostaining of 1M old organoids with neuronal, developmental and intermediate filament markers. Representative images of 1M old organoid attached in Matrigel, stained with DAPI, Sox2 and NeuN (left), DAPI and Nestin (right).

Figure S23. Western Blots images of total and active RhoA. Western Blot images of 1M organoids for KCTD13, actin as loading control, total RhoA, and active RhoA (RhoA-GTP). Organoids were grown in the batches of CTRL, DEL and DUP for each experiment. Six batches were grown for total RhoA and four batches were grown for active RhoA experiments.

Figure S24. Differential gene expression (DEL *vs* CTRL and DUP *vs* CTRL) for iPSCs, 1M and 3M cortical organoids. Volcano plots of differentially expressed genes in DEL *vs* CTRL (left column) and DUP *vs* CTRL (right column) for iPSCs, 1M and 3M old organoids. Genes within 16p11.2 CNV are colored in pink. Genes colored in orange are upregulated in DEL or DUP compared to CTRL; genes colored in blue are downregulated in DEL or DUP compared to CTRL.

Figure S25. Differential protein expression (DEL *vs* CTRL and DUP *vs* CTRL) for 1M and 3M cortical organoids. Volcano plots of differentially expressed proteins in DEL *vs* CTRL and DUP *vs* CTRL for 1M and 3M old organoids. Proteins within 16p11.2 CNV are colored in pink. Genes colored in orange are upregulated in DEL or DUP compared to CTRL; genes colored in blue are downregulated in DEL or DUP compared to CTRL. Gene Ontology enrichment analyses are shown as bar plots on the right. The contribution of up- or down-regulated proteins to specific GO terms are shown in blue and red, respectively.

Supplementary information

Supplementary Tables titles and legends

Supplementary Table 1. 16p11.2 patient-derived fibroblast selection and patients' clinical information.

Supplementary Table 2. RNA-seq parameters and quality control metrics from Cutadapt, STAR, Picard, and RNA-SeQ for iPSCs, 1M and 3M old organoids representing 108 transcriptomes.

Supplementary Table 3. Cortical organoids size analyses.

Supplementary Table 4. Differentially Expressed Genes (DEGs) in iPSCs, 1M and 3M cortical organoids.

Supplementary Table 5. Gene Ontology enrichment analysis of differentially expressed genes.

Supplementary Table 6. Module membership from gene co-expression (WGCNA) analysis in iPSCs, 1M and 3M cortical organoids.

Supplementary Table 7. Gene Ontology enrichment analyses of significantly genotype-associated gene co-expression modules from WGCNA.

Supplementary Table 8. Differentially Expressed Proteins (DEPs) in 1M and 3M cortical organoids detected by LC-MS/MS with TMT labeling.

Supplementary Table 9. Gene Ontology enrichment analysis of differentially expressed proteins.

Supplementary Table 10. Module membership from protein co-expression (WPCNA) analysis in 1M and 3M cortical organoids.

Supplementary Table 11. Gene Ontology enrichment analyses of significantly genotype-associated protein co-expression modules from WPCNA.

Supplementary Table 12. Summary of experiments by patients, clones and replicates.

References

Adams, J.W., Cugola, F.R., and Muotri, A.R. (2019). Brain Organoids as Tools for Modeling Human Neurodevelopmental Disorders. Physiology (Bethesda) *34*, 365-375.

Amin, N.D., and Pasca, S.P. (2018). Building Models of Brain Disorders with Three-Dimensional Organoids. Neuron *100*, 389-405.

Amiri, A., Coppola, G., Scuderi, S., Wu, F., Roychowdhury, T., Liu, F., Pochareddy, S., Shin, Y., Safi, A., Song, L.*, et al.* (2018). Transcriptome and epigenome landscape of human cortical development modeled in organoids. Science *362*.

Arbogast, T., Ouagazzal, A.M., Chevalier, C., Kopanitsa, M., Afinowi, N., Migliavacca, E., Cowling, B.S., Birling, M.C., Champy, M.F., Reymond, A.*, et al.* (2016). Reciprocal Effects on Neurocognitive and Metabolic Phenotypes in Mouse Models of 16p11.2 Deletion and Duplication Syndromes. PLoS Genet *12*, e1005709.

Azzarelli, R., Kerloch, T., and Pacary, E. (2014). Regulation of cerebral cortex development by Rho GTPases: insights from in vivo studies. Front Cell Neurosci *8*, 445.

Bates, D., Machler, M., Bolker, B.M., and Walker, S.C. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software *67*, 1-48.

Bershteyn, M., Nowakowski, T.J., Pollen, A.A., Di Lullo, E., Nene, A., Wynshaw-Boris, A., and Kriegstein, A.R. (2017). Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. Cell Stem Cell *20*, 435-449 e434.

Bijlsma, E.K., Gijsbers, A.C., Schuurs-Hoeijmakers, J.H., van Haeringen, A., Fransen van de Putte, D.E., Anderlid, B.M., Lundin, J., Lapunzina, P., Perez Jurado, L.A., Delle Chiaie, B.*, et al.* (2009). Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. Eur J Med Genet *52*, 77-87.

Birey, F., Andersen, J., Makinson, C.D., Islam, S., Wei, W., Huber, N., Fan, H.C., Metzler, K.R.C., Panagiotakos, G., Thom, N.*, et al.* (2017). Assembly of functionally integrated human forebrain spheroids. Nature *545*, 54-59.

Blumenthal, I., Ragavendran, A., Erdin, S., Klei, L., Sugathan, A., Guide, J.R., Manavalan, P., Zhou, J.Q., Wheeler, V.C., Levin, J.Z.*, et al.* (2014). Transcriptional consequences of 16p11.2 deletion and duplication in mouse cortex and multiplex autism families. Am J Hum Genet *94*, 870-883.

Boyden, S. (1962). The chemotactic effect of mixtures of antibody and antigen on polymorphonuclear leucocytes. J Exp Med *115*, 453-466.

Camp, J.G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Brauninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M.*, et al.* (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc Natl Acad Sci U S A *112*, 15672-15677.

Cappello, S., Bohringer, C.R., Bergami, M., Conzelmann, K.K., Ghanem, A., Tomassy, G.S., Arlotta, P., Mainardi, M., Allegra, M., Caleo, M.*, et al.* (2012). A radial glia-specific role of RhoA in double cortex formation. Neuron *73*, 911-924.

Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P., and Morgan, J. (2007). Mapping early brain development in autism. Neuron *56*, 399-413.

Cugola, F.R., Fernandes, I.R., Russo, F.B., Freitas, B.C., Dias, J.L., Guimaraes, K.P., Benazzato, C., Almeida, N., Pignatari, G.C., Romero, S.*, et al.* (2016). The Brazilian Zika virus strain causes birth defects in experimental models. Nature *534*, 267-271.

Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W.*, et al.* (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell *146*, 247-261.

Deshpande, A., Yadav, S., Dao, D.Q., Wu, Z.Y., Hokanson, K.C., Cahill, M.K., Wiita, A.P., Jan, Y.N., Ullian, E.M., and Weiss, L.A. (2017). Cellular Phenotypes in Human iPSC-Derived Neurons from a Genetic Model of Autism Spectrum Disorder. Cell Rep *21*, 2678-2687.
Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439-3440.
Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc *4*, 1184-1191.
Escamilla, C.O., Filonova, I., Walker, A.K., Xuan, Z.X., Holehonnur, R., Espinosa, F., Liu, S., Thyme, S.B., Lopez-Garcia, I.A., Mendoza, D.B.*, et al.* (2017). Kctd13 deletion reduces synaptic transmission via increased RhoA. Nature *551*, 227-231.
Fietz, S.A., Lachmann, R., Brandl, H., Kircher, M., Samusik, N., Schroder, R., Lakshmanaperumal, N., Henry, I., Vogt, J., Riehn, A.*, et al.* (2012). Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. Proc Natl Acad Sci U S A *109*, 11836-11841.
Govek, E.E., Hatten, M.E., and Van Aelst, L. (2011). The role of Rho GTPase proteins in CNS neuronal migration. Dev Neurobiol *71*, 528-553.
Grice, S.J., Liu, J.L., and Webber, C. (2015). Synergistic interactions between Drosophila orthologues of genes spanned by de novo human CNVs support multiple-hit models of autism. PLoS Genet *11*, e1004998.
Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Nimgaonkar, V.L., Go, R.C.*, et al.* (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell *154*, 518-529.
Hand, R., Bortone, D., Mattar, P., Nguyen, L., Heng, J.I., Guerrier, S., Boutt, E., Peters, E., Barnes, A.P., Parras, C.*, et al.* (2005). Phosphorylation of Neurogenin2 specifies the migration properties and the dendritic morphology of pyramidal neurons in the neocortex. Neuron *48*, 45-62.
He, L., Diedrich, J., Chu, Y.Y., and Yates, J.R., 3rd (2015). Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. Anal Chem *87*, 11361-11367.
Hoffman, G.E., and Schadt, E.E. (2016). variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics *17*, 483.
Horev, G., Ellegood, J., Lerch, J.P., Son, Y.E., Muthuswamy, L., Vogel, H., Krieger, A.M., Buja, A., Henkelman, R.M., Wigler, M.*, et al.* (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. Proc Natl Acad Sci U S A *108*, 17076-17081.
Iakoucheva, L.M., Muotri, A.R., and Sebat, J. (2019). Getting to the Cores of Autism. Cell *178*, 1287-1298.
Iefremova, V., Manikakis, G., Krefft, O., Jabali, A., Weynans, K., Wilkens, R., Marsoner, F., Brandl, B., Muller, F.J., Koch, P.*, et al.* (2017). An Organoid-Based Model of Cortical Development Identifies Non-Cell-Autonomous Defects in Wnt Signaling Contributing to Miller-Dieker Syndrome. Cell Rep *19*, 50-59.
Ippolito, D.M., and Eroglu, C. (2010). Quantifying synapses: an immunocytochemistry-based assay to quantify synapse number. J Vis Exp.
Iyer, J., Singh, M.D., Jensen, M., Patel, P., Pizzo, L., Huber, E., Koerselman, H., Weiner, A.T., Lepanto, P., Vadodaria, K.*, et al.* (2018). Pervasive genetic interactions modulate neurodevelopmental defects of the autism-associated 16p11.2 deletion in Drosophila melanogaster. Nat Commun *9*, 2548.
Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G.*, et al.* (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483-489.

Karczewski, K.J., Weisburd, B., Thomas, B., Solomonson, M., Ruderfer, D.M., Kavanagh, D., Hamamsy, T., Lek, M., Samocha, K.E., Cummings, B.B.*, et al.* (2017). The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res *45*, D840-D845.
Kholmanskikh, S.S., Dobrin, J.S., Wynshaw-Boris, A., Letourneau, P.C., and Ross, M.E. (2003). Disregulated RhoGTPases and actin cytoskeleton contribute to the migration defect in Lis1-deficient neurons. J Neurosci *23*, 8673-8681.
Klaus, J., Kanton, S., Kyrousi, C., Ayo-Martin, A.C., Di Giaimo, R., Riesenberg, S., O'Neill, A.C., Camp, J.G., Tocco, C., Santel, M.*, et al.* (2019). Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia. Nat Med *25*, 561-568.
Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software *82*, 1-26.
Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373-379.
Leek, J.T., and Storey, J.D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet *3*, 1724-1735.
Li, Y., Muffat, J., Omer, A., Bosch, I., Lancaster, M.A., Sur, M., Gehrke, L., Knoblich, J.A., and Jaenisch, R. (2017). Induction of Expansion and Folding in Human Cerebral Organoids. Cell Stem Cell *20*, 385-396 e383.
Lin, G.N., Corominas, R., Lemmens, I., Yang, X., Tavernier, J., Hill, D.E., Vidal, M., Sebat, J., and Iakoucheva, L.M. (2015). Spatiotemporal 16p11.2 Protein Network Implicates Cortical Late Mid-Fetal Brain Development and KCTD13-Cul3-RhoA Pathway in Psychiatric Diseases. Neuron *85*, 742-754.
Liu, Y., Beyer, A., and Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell *165*, 535-550.
Luo, C., Lancaster, M.A., Castanon, R., Nery, J.R., Knoblich, J.A., and Ecker, J.R. (2016). Cerebral Organoids Recapitulate Epigenomic Signatures of the Human Fetal Brain. Cell Rep *17*, 3369-3384.
Luo, R., Sanders, S.J., Tian, Y., Voineagu, I., Huang, N., Chu, S.H., Klei, L., Cai, C., Ou, J., Lowe, J.K.*, et al.* (2012). Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. Am J Hum Genet *91*, 38-55.
Malhotra, D., and Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell *148*, 1223-1241.
Marchetto, M.C., Belinson, H., Tian, Y., Freitas, B.C., Fu, C., Vadodaria, K., Beltrao-Braga, P., Trujillo, C.A., Mendes, A.P.D., Padmanabhan, K.*, et al.* (2017). Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. Mol Psychiatry *22*, 820-835.
Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., Amenduni, M., Szekely, A., Palejev, D., Wilson, M.*, et al.* (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. Cell *162*, 375-390.
Marshall, C.R., Howrigan, D.P., Merico, D., Thiruvahindrapuram, B., Wu, W., Greer, D.S., Antaki, D., Shetty, A., Holmans, P.A., Pinto, D.*, et al.* (2017). Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet *49*, 27-35.
Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y.*, et al.* (2008). Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet *82*, 477-488.
Martin-Brevet, S., Rodriguez-Herreros, B., Nielsen, J.A., Moreau, C., Modenato, C., Maillard, A.M., Pain, A., Richetin, S., Jonch, A.E., Qureshi, A.Y.*, et al.* (2018). Quantifying the Effects of 16p11.2 Copy Number Variants on Brain Structure: A Multisite Genetic-First Study. Biol Psychiatry *84*, 253-264.

McAlister, G.C., Nusinow, D.P., Jedrychowski, M.P., Wuhr, M., Huttlin, E.L., Erickson, B.K., Rad, R., Haas, W., and Gygi, S.P. (2014). MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal Chem *86*, 7150-7158.
McCarthy, S.E., Makarov, V., Kirov, G., Addington, A.M., McClellan, J., Yoon, S., Perkins, D.O., Dickel, D.E., Kusenda, M., Krastoshevsky, O.*, et al.* (2009). Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet *41*, 1223-1227.
Migliavacca, E., Golzio, C., Mannik, K., Blumenthal, I., Oh, E.C., Harewood, L., Kosmicki, J.A., Loviglio, M.N., Giannuzzi, G., Hippolyte, L.*, et al.* (2015). A Potential Contributory Role for Ciliary Dysfunction in the 16p11.2 600 kb BP4-BP5 Pathology. Am J Hum Genet *96*, 784-796.
Miller, J.A., Ding, S.L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K.*, et al.* (2014). Transcriptional landscape of the prenatal human brain. Nature *508*, 199-206.
Oldham, M.C., Langfelder, P., and Horvath, S. (2012). Network methods for describing sample relationships in genomic datasets: application to Huntington's disease. BMC Syst Biol *6*, 63.
Pacary, E., Heng, J., Azzarelli, R., Riou, P., Castro, D., Lebel-Potter, M., Parras, C., Bell, D.M., Ridley, A.J., Parsons, M.*, et al.* (2011). Proneural transcription factors regulate different steps of cortical neuron migration through Rnd-mediated inhibition of RhoA signaling. Neuron *69*, 1069-1084.
Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell *155*, 1008-1021.
Park, S.K., Venable, J.D., Xu, T., and Yates, J.R., 3rd (2008). A quantitative analysis software tool for mass spectrometry-based proteomics. Nat Methods *5*, 319-322.
Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S.*, et al.* (2010). Functional impact of global rare copy number variation in autism spectrum disorders. Nature *466*, 368-372.
Pirooznia, M., Wang, T., Avramopoulos, D., Valle, D., Thomas, G., Huganir, R.L., Goes, F.S., Potash, J.B., and Zandi, P.P. (2012). SynaptomeDB: an ontology-based knowledgebase for synaptic genes. Bioinformatics *28*, 897-899.
Polioudakis, D., de la Torre-Ubieta, L., Langerman, J., Elkins, A.G., Shi, X., Stein, J.L., Vuong, C.K., Nichterwitz, S., Gevorgian, M., Opland, C.K.*, et al.* (2019). A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. Neuron *103*, 785-801 e788.
Portmann, T., Yang, M., Mao, R., Panagiotakos, G., Ellegood, J., Dolen, G., Bader, P.L., Grueter, B.A., Goold, C., Fisher, E.*, et al.* (2014). Behavioral abnormalities and circuit defects in the Basal Ganglia of a mouse model of 16p11.2 deletion syndrome. Cell Rep *7*, 1077-1092.
Pucilowska, J., Vithayathil, J., Pagani, M., Kelly, C., Karlo, J.C., Robol, C., Morella, I., Gozzi, A., Brambilla, R., and Landreth, G.E. (2018). Pharmacological Inhibition of ERK Signaling Rescues Pathophysiology and Behavioral Phenotype Associated with 16p11.2 Chromosomal Deletion in Mice. J Neurosci *38*, 6640-6652.
Pucilowska, J., Vithayathil, J., Tavares, E.J., Kelly, C., Karlo, J.C., and Landreth, G.E. (2015). The 16p11.2 deletion mouse model of autism exhibits altered cortical progenitor proliferation and brain cytoarchitecture linked to the ERK MAPK pathway. J Neurosci *35*, 3190-3200.
Qiu, Y., Arbogast, T., Lorenzo, S.M., Li, H., Tang, S.C., Richardson, E., Hong, O., Cho, S., Shanta, O., Pang, T.*, et al.* (2019). Oligogenic Effects of 16p11.2 Copy-Number Variation on Craniofacial Development. Cell Rep *28*, 3320-3328 e3324.
Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res *35*, W193-200.
Richter, M., Murtaza, N., Scharrenberg, R., White, S.H., Johanns, O., Walker, S., Yuen, R.K.C., Schwanke, B., Bedurftig, B., Henis, M.*, et al.* (2018). Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. Mol Psychiatry.

Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnstrom, K., Mallick, S., Kirby, A.*, et al.* (2014). A framework for the interpretation of de novo mutation in human disease. Nat Genet *46*, 944-950.
Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S.*, et al.* (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron *87*, 1215-1233.
Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L.*, et al.* (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell.
Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J.*, et al.* (2007). Strong association of de novo copy number mutations with autism. Science *316*, 445-449.
Setia, H., and Muotri, A.R. (2019). Brain organoids as a model system for human neurodevelopment and disease. Semin Cell Dev Biol.
Shang, X., Marchioni, F., Sipes, N., Evelyn, C.R., Jerabek-Willemsen, M., Duhr, S., Seibel, W., Wortman, M., and Zheng, Y. (2012). Rational design of small molecule inhibitors targeting RhoA subfamily Rho GTPases. Chem Biol *19*, 699-710.
Shohayeb, B., Ho, U., Yeap, Y.Y., Parton, R.G., Millard, S.S., Xu, Z., Piper, M., and Ng, D.C.H. (2020). The association of microcephaly protein WDR62 with CPAP/IFT88 is required for cilia formation and neocortical development. Hum Mol Genet *29*, 248-263.
Skene, N.G., and Grant, S.G. (2016). Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. Front Neurosci *10*, 16.
Sonderby, I.E., Gustafsson, O., Doan, N.T., Hibar, D.P., Martin-Brevet, S., Abdellaoui, A., Ames, D., Amunts, K., Andersson, M., Armstrong, N.J.*, et al.* (2018). Dose response of the 16p11.2 distal copy number variant on intracranial volume and basal ganglia. Mol Psychiatry.
Stein, J.L., de la Torre-Ubieta, L., Tian, Y., Parikshak, N.N., Hernandez, I.A., Marchetto, M.C., Baker, D.K., Lu, D., Hinman, C.R., Lowe, J.K.*, et al.* (2014). A quantitative framework to evaluate modeling of cortical development by neural stem cells. Neuron *83*, 69-86.
Stoner, R., Chow, M.L., Boyle, M.P., Sunkin, S.M., Mouton, P.R., Roy, S., Wynshaw-Boris, A., Colamarino, S.A., Lein, E.S., and Courchesne, E. (2014). Patches of disorganization in the neocortex of children with autism. N Engl J Med *370*, 1209-1219.
Tabb, D.L., McDonald, W.H., and Yates, J.R., 3rd (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res *1*, 21-26.
Tian, D., Stoppel, L.J., Heynen, A.J., Lindemann, L., Jaeschke, G., Mills, A.A., and Bear, M.F. (2015). Contribution of mGluR5 to pathophysiology in a mouse model of human chromosome 16p11.2 microdeletion. Nat Neurosci *18*, 182-184.
Trujillo, C.A., Gao, R., Negraes, P.D., Gu, J., Buchanan, J., Preissl, S., Wang, A., Wu, W., Haddad, G.G., Chaim, I.A.*, et al.* (2019). Complex Oscillatory Waves Emerging from Cortical Organoids Model Early Human Brain Network Development. Cell Stem Cell.
Vaccarino, F.M., Grigorenko, E.L., Smith, K.M., and Stevens, H.E. (2009). Regulation of cerebral cortical size and neuron number by fibroblast growth factors: implications for autism. J Autism Dev Disord *39*, 511-520.
Velasco, S., Kedaigle, A.J., Simmons, S.K., Nash, A., Rocha, M., Quadrato, G., Paulsen, B., Nguyen, L., Adiconis, X., Regev, A.*, et al.* (2019). Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. Nature *570*, 523-527.
Wang, D., Eraslan, B., Wieland, T., Hallstrom, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.H., Meng, C.*, et al.* (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol Syst Biol *15*, e8503.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res *17*, 1665-1674.
Wang, P., Mokhtari, R., Pedrosa, E., Kirschenbaum, M., Bayrak, C., Zheng, D., and Lachman, H.M. (2017). CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in cerebral organoids derived from iPS cells. Mol Autism *8*, 11.
Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T.*, et al.* (2008). Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med *358*, 667-675.
Wilkinson, B., Grepo, N., Thompson, B.L., Kim, J., Wang, K., Evgrafov, O.V., Lu, W., Knowles, J.A., and Campbell, D.B. (2015). The autism-associated gene chromodomain helicase DNA-binding protein 8 (CHD8) regulates noncoding RNAs and autism-related genes. Transl Psychiatry *5*, e568.
Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A.*, et al.* (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell *155*, 997-1007.
Xu, Q., Liu, Y.Y., Wang, X., Tan, G.H., Li, H.P., Hulbert, S.W., Li, C.Y., Hu, C.C., Xiong, Z.Q., Xu, X.*, et al.* (2018). Autism-associated CHD8 deficiency impairs axon development and migration of cortical neurons. Mol Autism *9*, 65.
Xu, T., Park, S.K., Venable, J.D., Wohlschlegel, J.A., Diedrich, J.K., Cociorva, D., Lu, B., Liao, L., Hewel, J., Han, X.*, et al.* (2015). ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. J Proteomics *129*, 16-24.
Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol *4*, Article17.
Zubarev, R.A. (2013). The challenge of the proteome dynamic range and its implications for in-depth proteomics. Proteomics *13*, 723-726.

Figure 1

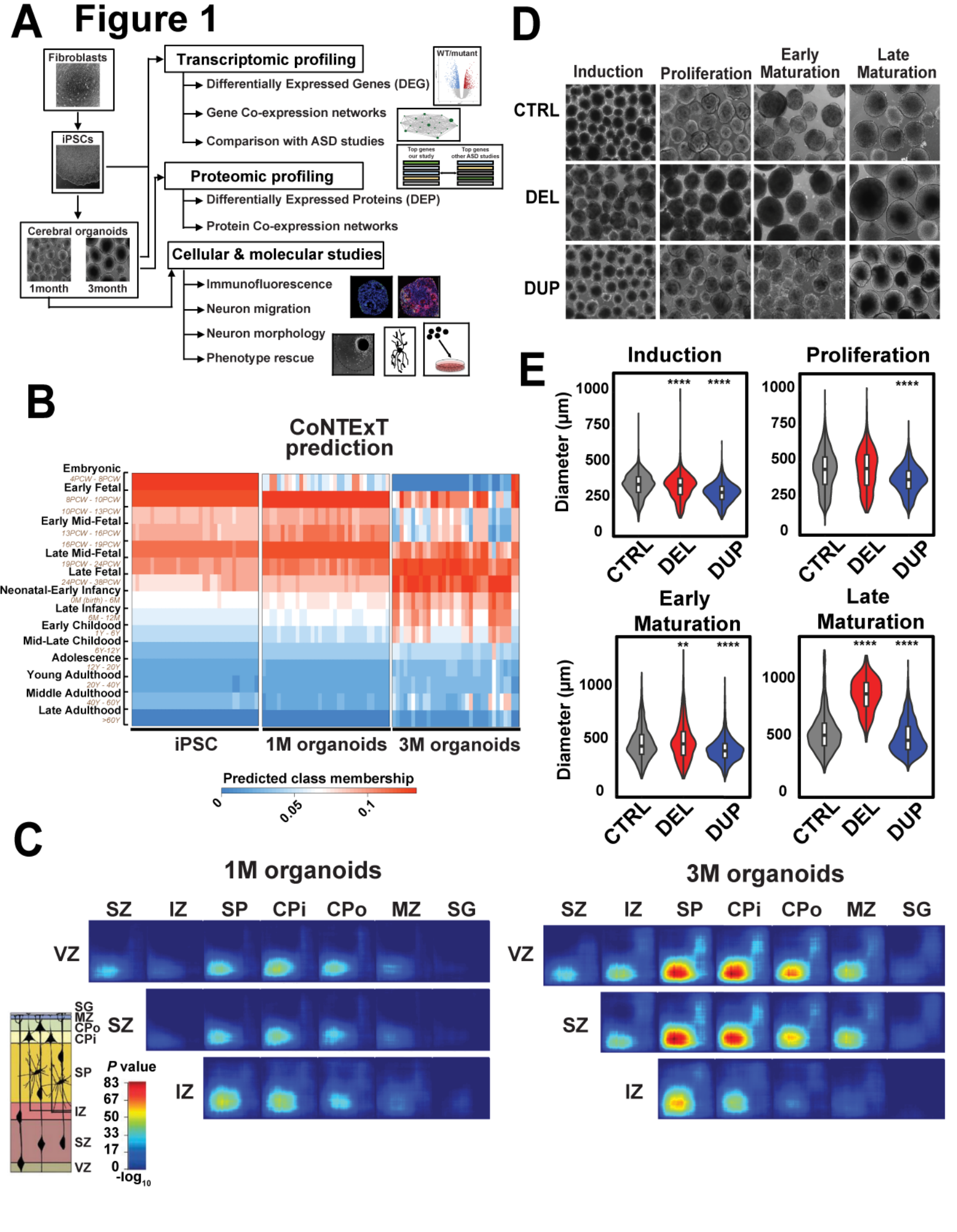

A
Fibroblasts
iPSCs
Cerebral organoids
1month
3month
Transcriptomic profiling
Differentially Expressed Genes (DEG)
WT/mutant
Gene Co-expression networks
Comparison with ASD studies
Top genes our study
Top genes other ASD studies
Proteomic profiling
Differentially Expressed Proteins (DEP)
Protein Co-expression networks
Cellular & molecular studies
Immunofluorescence
Neuron migration
Neuron morphology
Phenotype rescue
B
CoNTExT prediction
Embryonic
4PCW - 8PCW
Early Fetal
8PCW - 10PCW
10PCW - 13PCW
Early Mid-Fetal
13PCW - 16PCW
16PCW - 19PCW
Late Mid-Fetal
19PCW - 24PCW
Late Fetal
24PCW - 38PCW
Neonatal-Early Infancy
0M (birth) - 6M
Late Infancy
6M - 12M
Early Childood
1Y - 6Y
Mid-Late Childood
6Y-12Y
Adolescence
12Y - 20Y
Young Adulthood
20Y - 40Y
Middle Adulthood
40Y - 60Y
Late Adulthood
>60Y
iPSC
1M organoids
3M organoids
Predicted class membership
0
0.05
0.1
C
1M organoids
3M organoids
SZ
IZ
SP
CPi
CPo
MZ
SG
VZ
SZ
IZ
SG
MZ
CPo
CPi
SP
IZ
SZ
VZ
P value
83
67
50
33
17
0
-log10
D
Induction
Proliferation
Early Maturation
Late Maturation
CTRL
DEL
DUP
E
Induction
Proliferation
Early Maturation
Late Maturation
Diameter (µm)
1000
750
500
250
0
****
**
CTRL
DEL
DUP

Figure 2

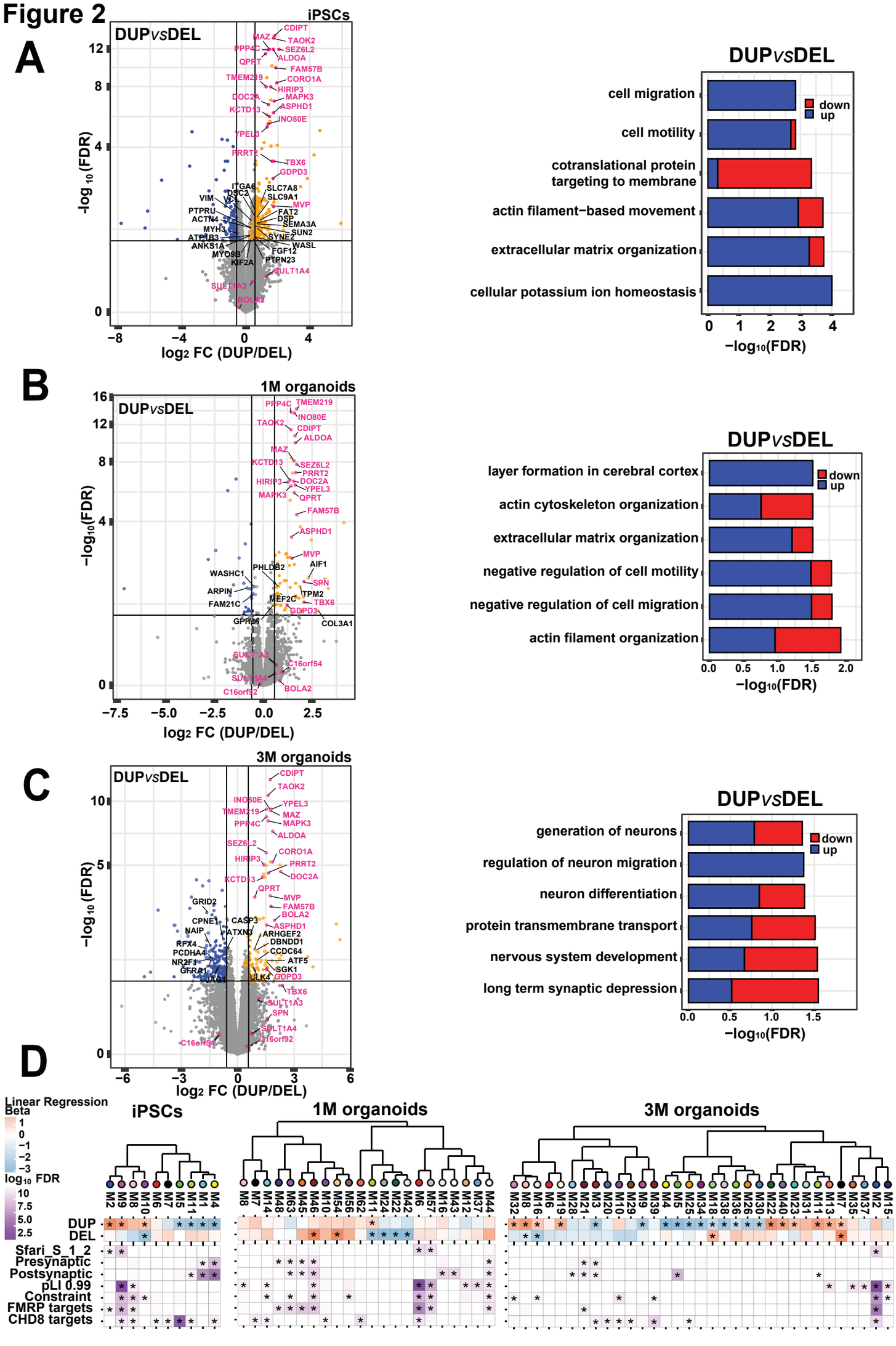

A
iPSCs
DUPvsDEL
-log10(FDR)
log2 FC (DUP/DEL)
CDIPT
TAOK2
MAZ
PPP4C
SEZ6L2
ALDOA
QPRT
FAM57B
TMEM219
CORO1A
HIRIP3
DOC2A
MAPK3
KCTD13
ASPHD1
INO80E
YPEL3
PRRT2
TBX6
GDPD3
ITGA6
SLC7A8
DSC2
VCL
SLC9A1
VIM
MVP
PTPRU
FAT2
ACTN4
DSP
SEMA3A
MYH3
SUN2
ATP1B3
SYNE2
ANKS1A
WASL
MYO9B
FGF12
KIF2A
PTPN23
SULT1A4
SULT1A3
BOLA2
DUPvsDEL
cell migration
cell motility
cotranslational protein targeting to membrane
actin filament−based movement
extracellular matrix organization
cellular potassium ion homeostasis
down
up
−log10(FDR)
B
1M organoids
DUPvsDEL
−log10(FDR)
log2 FC (DUP/DEL)
PPP4C
TMEM219
INO80E
TAOK2
CDIPT
ALDOA
MAZ
KCTD13
SEZ6L2
PRRT2
HIRIP3
DOC2A
YPEL3
MAPK3
QPRT
FAM57B
ASPHD1
MVP
AIF1
PHLDB2
WASHC1
SPN
ARPIN
TPM2
FAM21C
MEF2C
TBX6
GDPD3
GPR88
COL3A1
SULT1A3
C16orf54
SULT1A4
BOLA2
C16orf92
DUPvsDEL
layer formation in cerebral cortex
actin cytoskeleton organization
extracellular matrix organization
negative regulation of cell motility
negative regulation of cell migration
actin filament organization
down
up
−log10(FDR)
C
3M organoids
DUPvsDEL
−log10(FDR)
log2 FC (DUP/DEL)
CDIPT
TAOK2
INO80E
YPEL3
TMEM219
MAZ
PPP4C
MAPK3
ALDOA
SEZ6L2
CORO1A
HIRIP3
PRRT2
KCTD13
DOC2A
QPRT
MVP
GRID2
FAM57B
BOLA2
CPNE1
CASP3
ASPHD1
NAIP
ATXN3
ARHGEF2
RFX4
DBNDD1
PCDHA4
CCDC64
NR2F1
ATF5
GFRA1
SGK1
ULK4
GDPD3
TBX6
SULT1A3
SPN
SULT1A4
C16orf92
DUPvsDEL
generation of neurons
regulation of neuron migration
neuron differentiation
protein transmembrane transport
nervous system development
long term synaptic depression
down
up
−log10(FDR)
D
Linear Regression Beta
log10 FDR
iPSCs
1M organoids
3M organoids
DUP
DEL
Sfari_S_1_2
Presynaptic
Postsynaptic
pLI 0.99
Constraint
FMRP targets
CHD8 targets

Figure 3

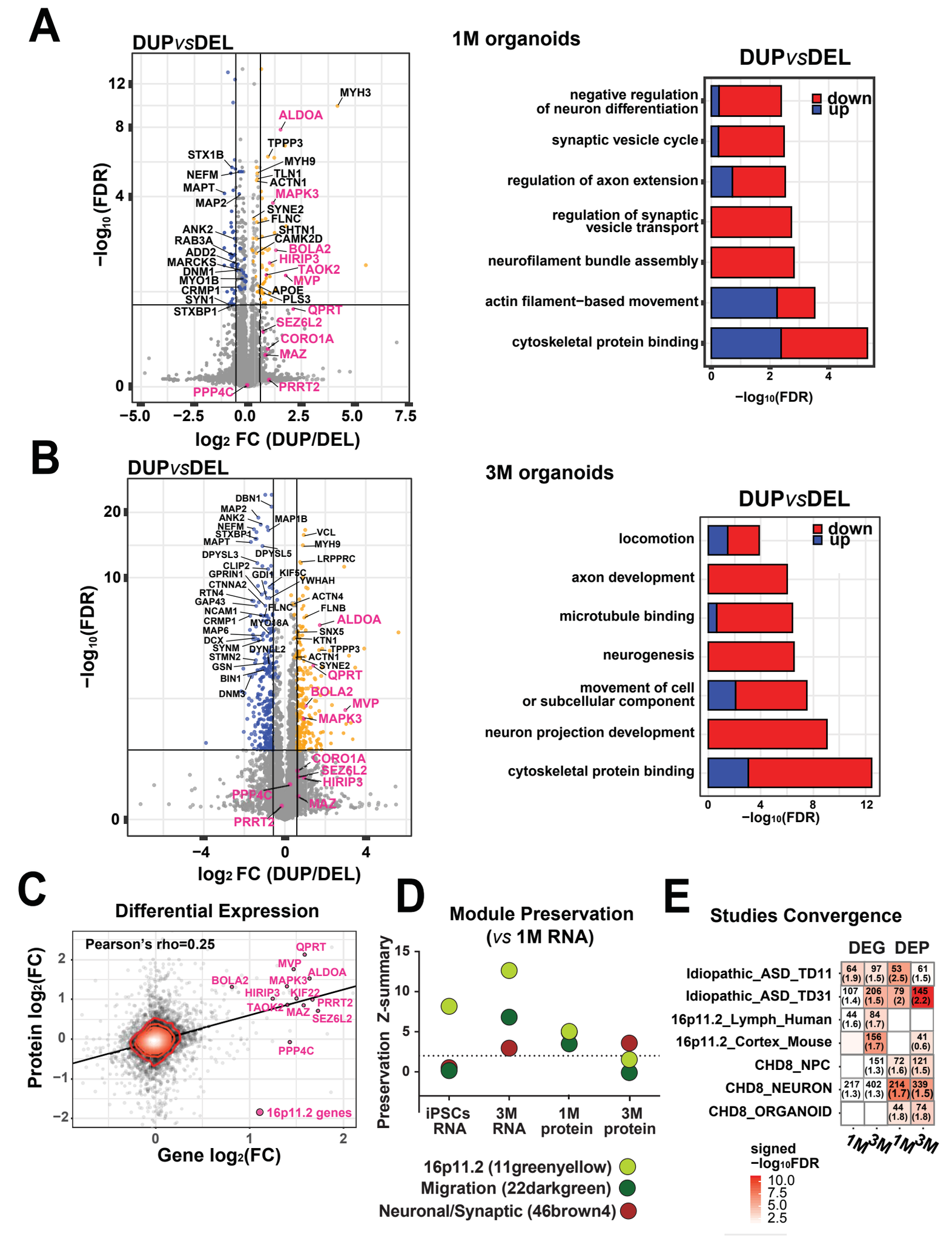

A
DUPvsDEL
−log10 (FDR)
log2 FC (DUP/DEL)
MYH3
ALDOA
TPPP3
MYH9
TLN1
ACTN1
MAPK3
SYNE2
FLNC
SHTN1
CAMK2D
BOLA2
HIRIP3
TAOK2
MVP
APOE
PLS3
QPRT
SEZ6L2
CORO1A
MAZ
PRRT2
PPP4C
STX1B
NEFM
MAPT
MAP2
ANK2
RAB3A
ADD2
MARCKS
DNM1
MYO1B
CRMP1
SYN1
STXBP1
1M organoids
DUPvsDEL
down
up
negative regulation of neuron differentiation
synaptic vesicle cycle
regulation of axon extension
regulation of synaptic vesicle transport
neurofilament bundle assembly
actin filament−based movement
cytoskeletal protein binding
−log10(FDR)
B
DUPvsDEL
−log10(FDR)
log2 FC (DUP/DEL)
DBN1
MAP2
ANK2
NEFM
STXBP1
MAPT
MAP1B
VCL
MYH9
DPYSL3
DPYSL5
LRPPRC
CLIP2
GPRIN1
GDI1
KIF5C
CTNNA2
YWHAH
RTN4
GAP43
ACTN4
NCAM1
FLNC
FLNB
CRMP1
MYO18A
ALDOA
MAP6
SNX5
DCX
KTN1
SYNM
DYNLL2
TPPP3
STMN2
ACTN1
GSN
SYNE2
QPRT
BIN1
DNM3
BOLA2
MVP
MAPK3
CORO1A
SEZ6L2
HIRIP3
PPP4C
MAZ
PRRT2
3M organoids
DUPvsDEL
down
up
locomotion
axon development
microtubule binding
neurogenesis
movement of cell or subcellular component
neuron projection development
cytoskeletal protein binding
−log10(FDR)
C
Differential Expression
Pearson's rho=0.25
Protein log2(FC)
Gene log2(FC)
QPRT
MVP
ALDOA
BOLA2
MAPK3
HIRIP3
KIF22
PRRT2
TAOK2
MAZ
SEZ6L2
PPP4C
16p11.2 genes
D
Module Preservation
(vs 1M RNA)
Preservation Z-summary
iPSCs RNA
3M RNA
1M protein
3M protein
16p11.2 (11greenyellow)
Migration (22darkgreen)
Neuronal/Synaptic (46brown4)
E
Studies Convergence
DEG DEP
Idiopathic_ASD_TD11 64 (1.9) 97 (1.5) 53 (2.5) 61 (1.5)
Idiopathic_ASD_TD31 107 (1.4) 206 (1.5) 79 (2) 145 (2.2)
16p11.2_Lymph_Human 44 (1.6) 84 (1.7)
16p11.2_Cortex_Mouse 156 (1.7) 41 (0.6)
CHD8_NPC 151 (1.3) 72 (1.6) 121 (1.5)
CHD8_NEURON 217 (1.3) 402 (1.3) 214 (1.7) 339 (1.5)
CHD8_ORGANOID 44 (1.8) 74 (1.8)
1M 3M 1M 3M
signed −log10FDR
10.0
7.5
5.0
2.5

Figure 4

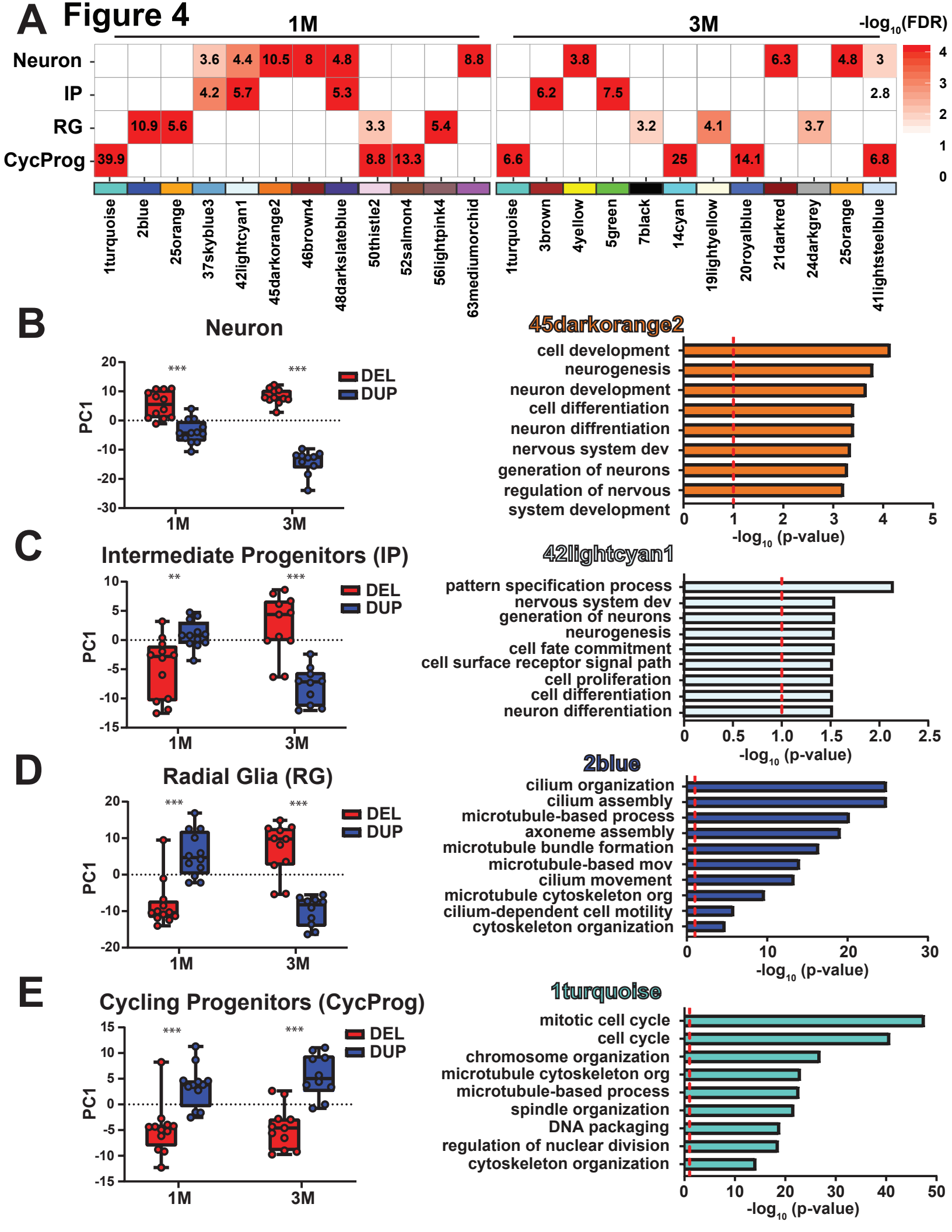

A
1M
3M
-log10(FDR)
Neuron
IP
RG
CycProg
1turquoise
2blue
25orange
37skyblue3
42lightcyan1
45darkorange2
46brown4
48darkslateblue
50thistle2
52salmon4
56lightpink4
63mediumorchid
3brown
4yellow
5green
7black
14cyan
19lightyellow
20royalblue
21darkred
24darkgrey
41lightsteelblue
B
Neuron
PC1
DEL
DUP
1M
3M
45darkorange2
cell development
neurogenesis
neuron development
cell differentiation
neuron diffrentiation
nervous system dev
generation of neurons
regulation of nervous system development
-log10 (p-value)
C
Intermediate Progenitors (IP)
42lightcyan1
pattern specification process
nervous system dev
generation of neurons
neurogenesis
cell fate commitment
cell surface receptor signal path
cell proliferation
cell differentiation
neuron differentiation
D
Radial Glia (RG)
2blue
cilium organization
cilium assembly
microtubule-based process
axoneme assembly
microtubule bundle formation
microtubule-based mov
cilium movement
microtubule cytoskeleton org
cilium-dependent cell motility
cytoskeleton organization
E
Cycling Progenitors (CycProg)
1turquoise
mitotic cell cycle
cell cycle
chromosome organization
microtubule cytoskeleton org
microtubule-based process
spindle organization
DNA packaging
regulation of nuclear division
cytoskeleton organization

Figure 5

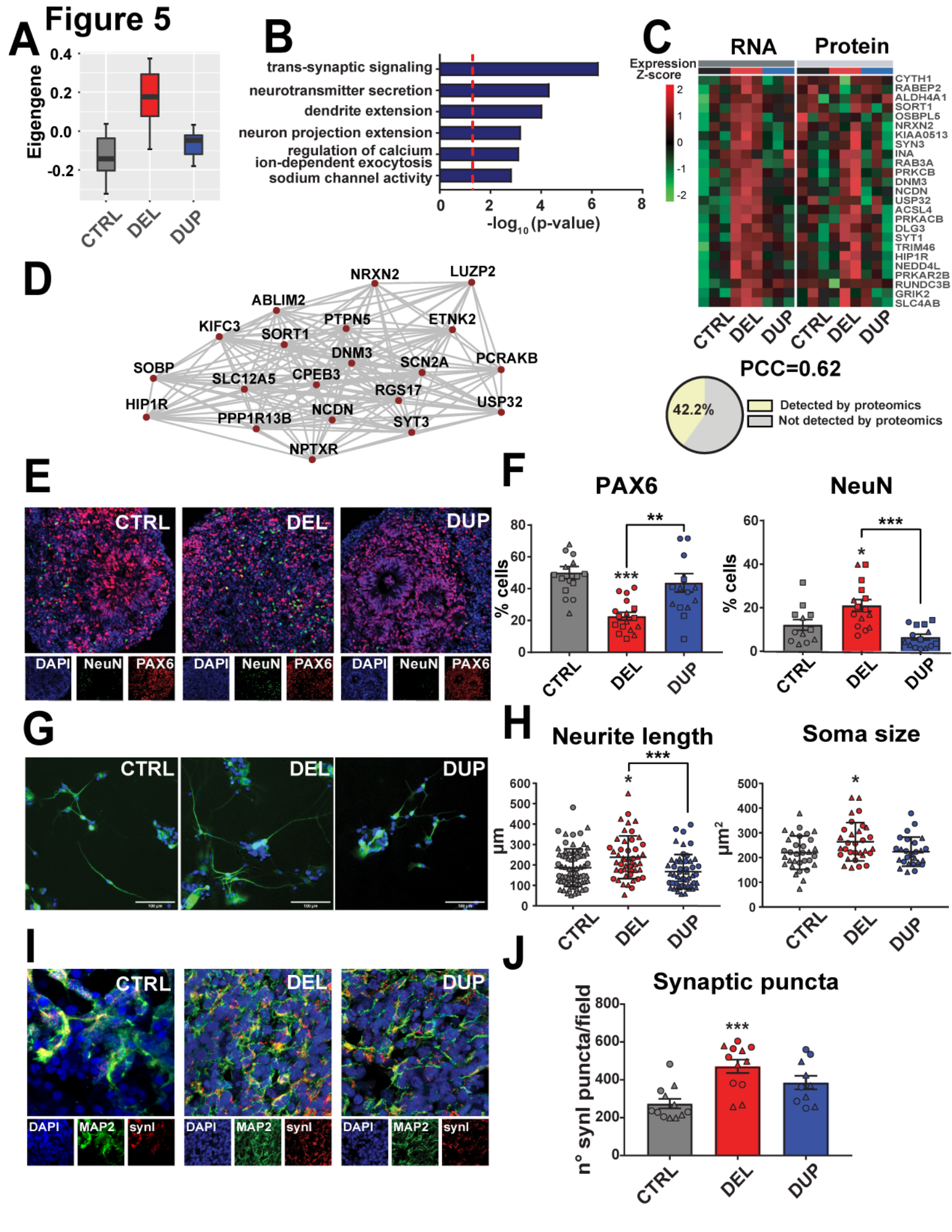

A
Eigengene
0.4
0.2
0.0
-0.2
CTRL
DEL
DUP
B
trans-synaptic signaling
neurotransmitter secretion
dendrite extension
neuron projection extension
regulation of calcium ion-dependent exocytosis
sodium channel activity
0
2
4
6
8
-log10(p-value)
C
RNA
Protein
Expression Z-score
2
1
0
-1
-2
CYTH1
RABEP2
ALDH4A1
SORT1
OSBPL5
NRXN2
KIAA0513
SYN3
INA
RAB3A
PRKCB
DNM3
NCDN
USP32
ACSL4
PRKACB
DLG3
SYT1
TRIM46
HIP1R
NEDD4L
PRKAR2B
RUNDC3B
GRIK2
SLC4AB
CTRL
DEL
DUP
CTRL
DEL
DUP
PCC=0.62
42.2%
Detected by proteomics
Not detected by proteomics
D
NRXN2
LUZP2
ABLIM2
KIFC3
PTPN5
ETNK2
SORT1
DNM3
SCN2A
PCRAKB
SOBP
SLC12A5
CPEB3
RGS17
HIP1R
USP32
NCDN
PPP1R13B
SYT3
NPTXR
E
CTRL
DEL
DUP
DAPI
NeuN
PAX6
DAPI
NeuN
PAX6
DAPI
NeuN
PAX6
F
PAX6
NeuN
% cells
80
60
40
20
0
**
***
CTRL
DEL
DUP
% cells
60
40
20
0
***
*
CTRL
DEL
DUP
G
CTRL
DEL
DUP
H
Neurite length
Soma size
µm
600
500
400
300
200
100
0
***
*
CTRL
DEL
DUP
µm²
500
400
300
200
100
0
*
CTRL
DEL
DUP
I
CTRL
DEL
DUP
DAPI
MAP2
synI
DAPI
MAP2
synI
DAPI
MAP2
synI
J
Synaptic puncta
n° synI puncta/field
800
600
400
200
0
***
CTRL
DEL
DUP

Figure 6

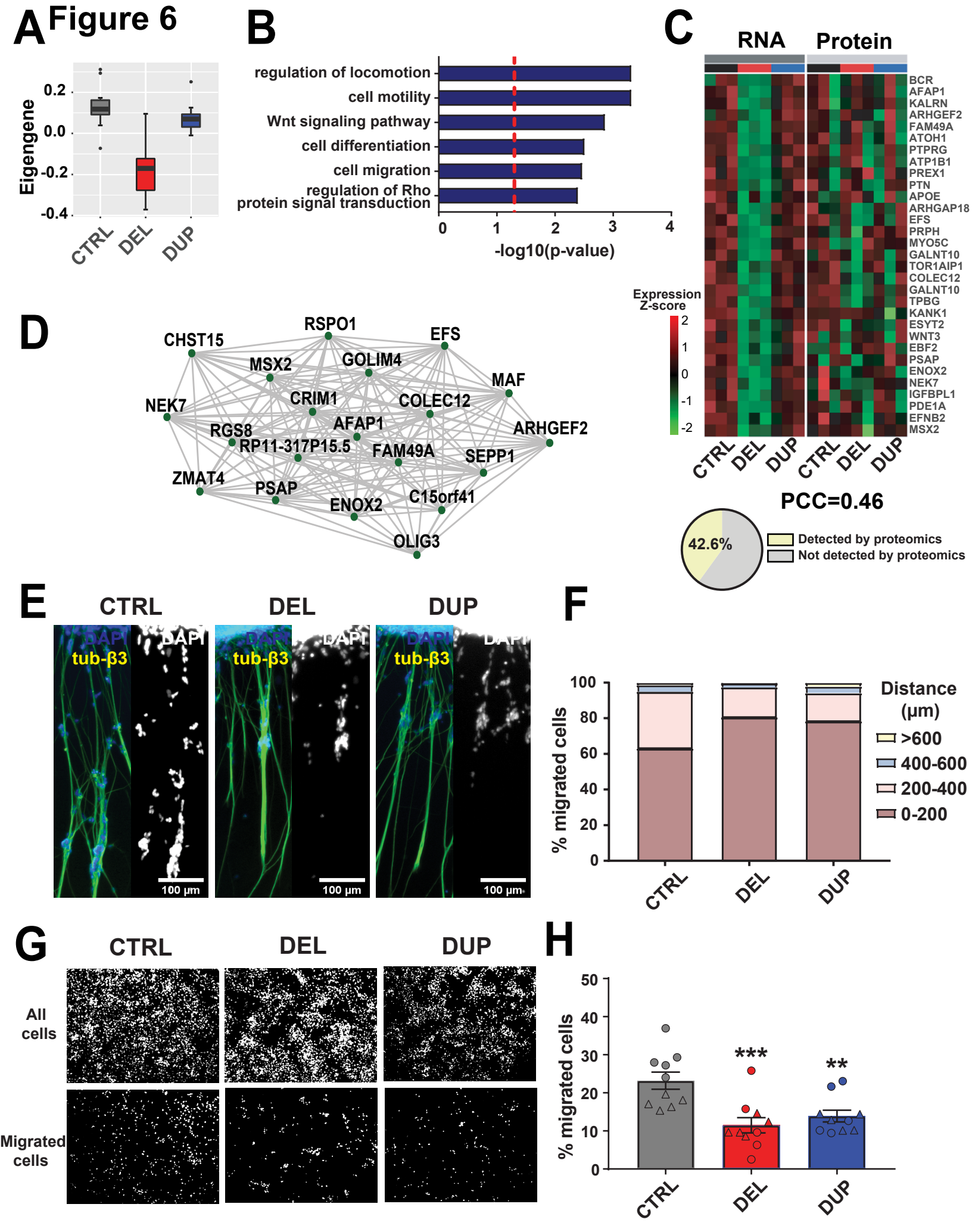
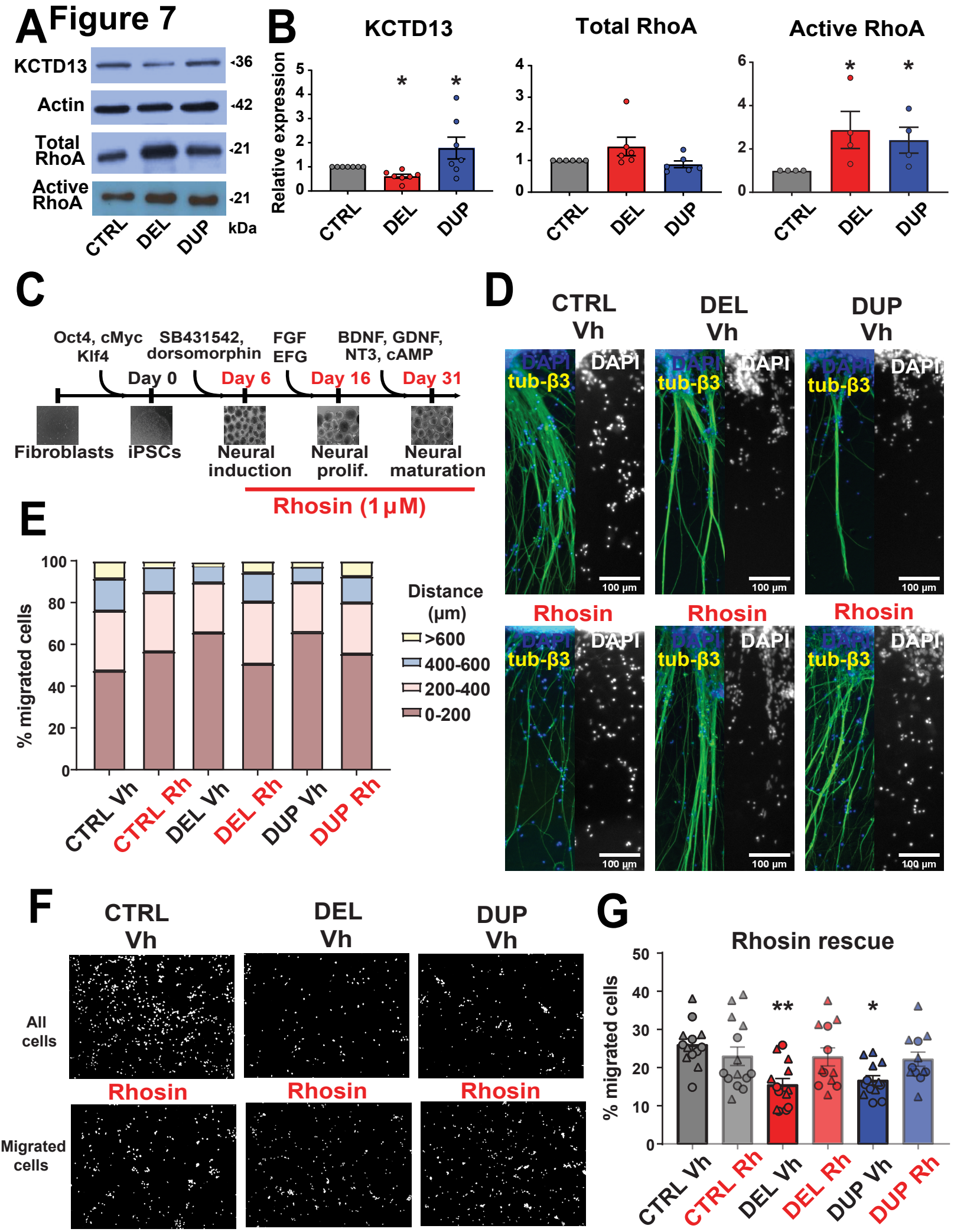

A
Eigengene
0.2
0.0
-0.2
-0.4
CTRL
DEL
DUP
B
regulation of locomotion
cell motility
Wnt signaling pathway
cell differentiation
cell migration
regulation of Rho protein signal transduction
0
1
2
3
4
-log10(p-value)
C
RNA
Protein
BCR
AFAP1
KALRN
ARHGEF2
FAM49A
ATOH1
PTPRG
ATP1B1
PREX1
PTN
APOE
ARHGAP18
EFS
PRPH
MYO5C
GALNT10
TOR1AIP1
COLEC12
GALNT10
TPBG
KANK1
ESYT2
WNT3
EBF2
PSAP
ENOX2
NEK7
IGFBPL1
PDE1A
EFNB2
MSX2
Expression Z-score
2
1
0
-1
-2
CTRL
DEL
DUP
CTRL
DEL
DUP
PCC=0.46
42.6%
Detected by proteomics
Not detected by proteomics
D
CHST15
RSPO1
EFS
MSX2
GOLIM4
MAF
NEK7
CRIM1
COLEC12
RGS8
AFAP1
ARHGEF2
RP11-317P15.5
FAM49A
SEPP1
ZMAT4
PSAP
C15orf41
ENOX2
OLIG3
E
CTRL
DEL
DUP
DAPI
tub-β3
DAPI
100 µm
F
% migrated cells
100
80
60
40
20
0
CTRL
DEL
DUP
Distance (µm)
>600
400-600
200-400
0-200
G
CTRL
DEL
DUP
All cells
Migrated cells
H
% migrated cells
50
40
30
20
10
0
***
**
CTRL
DEL
DUP

Figure 7


A
KCTD13
Actin
Total RhoA
Active RhoA
36
42
21
21
CTRL
DEL
DUP
kDa
B
KCTD13
Total RhoA
Active RhoA
Relative expression
CTRL
DEL
DUP
C
Oct4, cMyc
Klf4
SB431542,
dorsomorphin
FGF
EFG
BDNF, GDNF,
NT3, cAMP
Day 0
Day 6
Day 16
Day 31
Fibroblasts
iPSCs
Neural induction
Neural prolif.
Neural maturation
Rhosin (1 µM)
D
CTRL Vh
DEL Vh
DUP Vh
Rhosin
DAPI
tub-β3
100 µm
E
% migrated cells
Distance (µm)
>600
400-600
200-400
0-200
CTRL Vh
CTRL Rh
DEL Vh
DEL Rh
DUP Vh
DUP Rh
F
CTRL Vh
DEL Vh
DUP Vh
Rhosin
All cells
Migrated cells
G
Rhosin rescue
% migrated cells
CTRL Vh
CTRL Rh
DEL Vh
DEL Rh
DUP Vh
DUP Rh