Dysregulated GABA/glutamate neuron differentiation in autism spectrum disorders with macrocephaly

Jessica Mariani[1,2*], Gianfilippo Coppola[1,2*], Ping Zhang[3], Alexej Abyzov[1,4,&], Lauren Provini[1,2], Livia Tomasini[1,2], Mariangela Amenduni[1,2], Anna Szekely[1,5], Dean Palejev[1,2¶], Michael Wilson[1,2], Mark Gerstein[1,4,6,7], Elena Grigorenko[1,2], Katarzyna Chawarska[1,2], Kevin Pelphrey[1,2], James Howe[3], Flora M. Vaccarino[1,2,8]

[1]Program in Neurodevelopment and Regeneration, Yale University, New Haven, Connecticut 06520, USA.

[2]Child Study Center, Yale University, New Haven, Connecticut 06520, USA.

[3]Department of Pharmacology, Yale University, New Haven, Connecticut 06520, USA.

[4]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, USA

[5]Department of Genetics, Yale University, New Haven, Connecticut 06520, USA.

[6]Program in Computation Biology and Bioinformatics, Yale University, New Haven, Connecticut 06520, USA.

[7]Department of Computer Science, Yale University, New Haven, Connecticut 06520, USA.

[8]Department of Neurobiology, Yale University, New Haven, Connecticut 06520, USA.

[*] These authors contributed equally to this work

[&] Present address: Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905

[¶]Present address: Bulgarian Academy of Sciences, Institute of Mathematics and Informatics, Sofia 1113, Bulgaria

Summary

Autism spectrum disorder (ASD) is a disorder of brain development believed, in most cases, to be of genetic origin. Currently, 80% of ASD cases are idiopathic, meaning that there is no knowledge of causative risk factors. It is hypothesized that ASD can be caused by rare penetrant mutations, but common genetic variants and environmental factors are known to contribute risk to the disorder. The inherent difficulty in reenacting altered trajectories of brain development and in obtaining samples of human brain resulted in the lack of a developmental model for ASD and has precluded understanding of the underlying pathophysiology. Here we use organoids differentiated from induced pluripotent stem cells (iPSCs) in patients with ASD and macrocephaly to investigate neurodevelopmental alterations that cause this form of ASD. By using transcriptome analyses, we identified modules of co-expressed genes that implicate cell proliferation, neuronal differentiation, and synaptic assembly as being significantly upregulated in ASD. On the cellular level, we demonstrate an accelerated cell cycle, synaptic overgrowth and an unexpected overproduction of GABAergic inhibitory neurons in organoids derived from individuals with ASD compared to non-ASD first-degree family members. Using RNA interference, we demonstrate that an increased expression of the telencephalic transcription factor FOXG1 is responsible for the overproduction of GABAergic neurons in ASD-derived neural cells. Finally, we show that altered expression of gene network modules and FOXG1 are positively correlated with symptom severity and increased head circumference in this patient cohort. Thus, for the first time, we identified molecular signatures and cellular alterations of idiopathic ASD during neurodevelopment and indicate potential drug targets. Our study also validates iPSC-derived organoid cultures as a clinically relevant model of early human telencephalic development.

Text:

About 80% of the cases of ASD has no clear etiology and no pathogenetic model. A large number of rare mutations have been identified in the context of syndromic and non-syndromic ASD and have been modeled in various organisms. However, these mutations are extremely heterogeneous and it is difficult to take their phenotype at face value, as in no instance they have been shown to be sufficient to cause ASD and they interact with other inherited and non-inherited risk factors. For example, some mutations involve synapse-associated molecules[1,2] and have led to the widespread notion that alterations in the assembly of synaptic connections are key in the pathophysiology of ASD. Others have formulated the hypothesis that there is an excitatory/inhibitory neuron imbalance in the disorder[3,4]. Yet, the heterogeneity of phenotypes found when modeling these mutations in animals and the inherent difficulty in creating behavioral phenotypes of ASD in rodents have complicated the construction of credible animal models of ASD. It is possible that the heterogeneity of rare mutations found in ASD, as currently conceptualized, denies a unified understanding of the pathophysiology of the disorder. However, emerging evidence suggests that current genomic data, when considered in the framework of gene network analyses, point to a common pathophysiological substrate in ASD rooted in the embryonic development of the cerebral cortex[5,6]. Here we have taken a different approach by directly analyzing early cortical development in probands with idiopathic ASD and increased head/brain size (macrocephaly), one of the most consistently replicated ASD phenotypes[7] and one that confers poorer clinical

outcomes[8,9]. Using induced pluripotent stem cells (iPSCs) obtained from affected families, we have produced cortical organoids that recapitulate transcriptional programs present in first trimester human cortical development. Transcriptome and cellular phenotype analyses in this model identified unexpected differences in cell cycle control and synaptic overgrowth together with a profound imbalance in GABA/glutamate neuronal differentiation in patients as compared to their unaffected family members.

Normal organization, phenotype and excitability of cells within iPSC-derived organoids of patients and controls

To analyze neurodevelopmental aspects of idiopathic ASD with macrocephaly, we generated iPSC lines from members of five families that each included an ASD proband with increased head circumference (HC) and one to three unaffected, first-degree family members (see **Extended Data Fig. 1** for family structure and **Supplementary Table 1** data collection and subject details). The characterization of the iPSC lines is shown in the **Supplemental Results** section and in **Extended Data Fig. 2 and 3**. Whole genome sequencing was obtained on all fibroblasts and iPSCs for members of all five families in this study. We did not find any *de novo* CNVs and/or any rare SNVs in probands that caused a known deleterious loss of function in the protein coding sequence of a gene previously involved in rare cases of syndromic or non-syndromic ASD, although probands 07-03 carries a deletion involving exon 2 in the PTEN gene inherited from his asymptomatic father and proband 03-03 carries a small intronic deletion in CNTNAP2 gene common in the general population (see **Supplemental Results** section and **Extended Data Fig. 4**). For four families, we differentiated two to three iPSC lines per person into telencephalic organoids using a modification of our free-floating tridimensional (3D) culture method[10] (see **Supplemental Results** and schematic outline in **Extended Data Fig. 1b**). Similarly to our previously published preparation[10,11], we obtained organoids characterized by polarized, autonomously organized layers of radial glia, intermediate progenitors and neurons (see **Supplemental Results** section) (**Fig. 2a,b; Extended Data Fig. 5a,b**). RNAseq analyses of the organoid's transcriptome in comparisons with the BrainSpan dataset[12] indicated that this preparation best reflected the transcriptome of the human cerebral cortex during the first trimester of gestation (**Supplemental Results** and **Extended Data Fig. 6**).

Electrophysiological characterization of the organoids by whole-cell patch-clamp recordings (see **Supplemental Results**) demonstrated expression of voltage-activated sodium and potassium currents which supported action potential firing (**Extended Data Fig. 7**) as well as spontaneous synaptic currents (**Extended Data Fig. 8).** In total, the electrophysiological data support the conclusion that the cells studied here display the phenotype of central neurons with no detectable difference between patients and controls.

Transcriptome analysis in organoids from ASD individuals

Next, we compared the transcriptomes of four probands to those of the unaffected family members (two to three iPSC clones per person) at two time points, TD11 and TD31. Differential gene expression (DGE) between the probands and the respective fathers, used as sex-matched normal controls, identified 1062 differentially expressed genes (DEGs) at TD11 and 2203 at TD31 (see **Supplementary Table 2**), hinting at a possibly divergent developmental trajectory between controls and probands. Validation by qPCR

of a subset of the DEGs identified by RNAseq revealed a 0.98 correlation coefficient between log2 fold changes from the two techniques and 100% concordance in direction of change (see **Supplementary Table 3**).

We then asked whether, at a system level, the DEGs underline coherent alterations in groups of co-expressed genes. Using weighted gene co-expression network analysis, WGCNA[13], we identified 24 modules of co-expressed genes across probands and controls at TD11 and TD31, all of which survived permutation analysis (**Supplementary Table 4)**. We then estimated the modules' eigengenes (i.e. the first principal component of the module's genes expression profiles) and assessed their changes over time and across diagnosis (see **Fig. 1a**). In the ASD probands, we found that the *yellow* and *green* modules (annotated by "vascular development" and "lipid metabolism" gene ontology (GO) categories, respectively) were enriched in downregulated DEGs, and their eigengenes were consistently downregulated; the *blue* and *magenta* modules (annotated by "neuronal differentiation" and "regulation of transcription" GO terms, respectively) were enriched in upregulated genes at both TD11 and TD31, and their eigengenes were consistently upregulated (in keeping with their developmental annotation). The *brown* and *tan* modules (annotated by "synaptic transmission" and "gated channel activity" GO terms, respectively) were enriched in upregulated genes only at TD31, and their eigengenes were upregulated more at TD31 than at TD11 (in line with their main synaptic functional annotation) (**Table 1; Supplementary Table 5**). Hierarchical clustering of module eigengenes also showed a tighter correlation among the *green* and *yellow*, the *blue* and *magenta*, and the *brown* and *tan* modules (**Fig. 1b**), suggesting similar functions and/or tighter regulatory interactions.

We decided to focus on the upregulated *magenta*, *blue*, *brown* and *tan* modules in light of their "neuronal" annotation and further investigated their functional pathways annotations (**Fig. 1a**). The *magenta* module was significantly enriched in transcription- and cell cycle-related canonical pathways (**Table 1; Supplementary Table 6; Supplementary Table 7).** Indeed, among the top upregulated genes in probands in this module are a number of transcription factors (TFs) crucial for acquisition of neural cell fates and precursor cell proliferation in the telencephalon, including DLX6-AS1, FOXG1, EOMES, POU3F3/BRN1, SOX3, SOX5, GSX2, ETV1, DLX1, DLX6, E2F2, and SYNE2 (**Supplementary Table 2; Supplementary Table 6**). Many of these TFs are also hub genes (**Fig. 1c**; **Supplementary Table 8**). Overall, genes in the *magenta* module function in the transcriptional regulation of cell fate and cell proliferation in the forebrain.

For the *blue* module, in keeping with its strong correlation with the *magenta* module, canonical pathway annotation was enriched in axon guidance and generic transcription pathways (**Supplementary Table 6; Supplementary Table 7).** The axon guidance genes, all upregulated in probands, included members of the neural cell adhesion family (NCAM1, NRCAM, L1CAM, NFASC), semaphorin (PLXNC1, PLXNB3), netrin (DCC, UNC5A), Rho GTPases (PAK3, PAK7), CDK5R1/P35, and DCX. Many of these molecules were among the top 100 hub genes, i.e., CDK5R1, DPYSL5, DCX, DPF1, APC2 (**Fig. 1c**; **Supplementary Table 8**). Hence, genes in the *blue* module are involved in cytoskeletal regulation of various cellular functions, including neurite outgrowth, axon guidance, cell proliferation, migration, and survival.

For the *brown* and the *tan* modules, the GO and top pathways annotation were mostly related to synaptic functions, ion channels, and ligand-receptor interactions (**Table 1, Supplementary Table 6, Supplementary Table 7**), which is in harmony with their more significant eigengene upregulation in probands at the later time point (**Fig. 1a)**. The *brown* module displayed the strongest enrichment with respect to the SFARI autism

gene dataset, a collection of genetic information that includes data from linkage and association studies, cytogenetic abnormalities, and specific mutations associated with ASD (https://gene.sfari.org) (SFARI category S to 4: syndromic, high confidence, strong confidence, suggestive evidence and minimal evidence)[14] (P-value = 6.51E-5). Among the top 100 hub genes of the *brown* module, 64 were upregulated at TD31, including molecules involved in synaptic assembly, ion transport, and post-synaptic signaling (i.e., NRXN1, NRXN2, SLITRK1, CAMK2B, CAMK1D, NRSN1, SYT13, GRIN1, SCN2A) (**Supplementary Table 8)**. Among the hub genes, NRXN1, SCN2A, and TSPAN7 were significantly upregulated and overlapped with genes in the SFARI autism database (**Fig. 1c).** The *tan* module was characterized by an upregulation at TD31 of transcripts for several potassium channels and for key components of GABAergic neurons, including the GABA synthetic enzyme GAD1 and three GABA receptor subunits, most of which were in the *tan* module's 100 hub genes (**Supplementary Table 8)**.

Altered neurodevelopmental processes in ASD with macrocephaly

In summary, the signatures that emerged from our transcriptome analyses as significantly perturbed in probands were transcriptional regulation of cell proliferation/cell fate, neuronal differentiation/process outgrowth, and synaptic transmission. To functionally validate these signatures, we performed morphometric cellular analyses and immunostaining for cell fate markers. We first studied the dynamics of the cell cycle in undifferentiated iPSCs and neuronal progenitors (TD11) by BrdU incorporation (**Extended Data Fig. 9a,b**). These experiments revealed a significant decrease in cell cycle length in ASD-derived iPSCs. We saw a similarly strong trend in early neuronal progenitors cultured as monolayers (**Supplementary Fig. 9a,b**). However, when we estimated the proportion of proliferating cells in more mature organoids (TD31), there was no significant difference in proliferation between ASD- and control-derived organoids (**Extended Data Fig. 9c,d**). Taken together, these results suggest that a decrease in cell cycle length could be an early event that is present at the iPSC undifferentiated stage and during the early stages of neuronal differentiation.

We then assessed neuronal maturation and synaptic formation in organoids at TD31. Quantification of microtubule-associated protein 2 (MAP2) showed a significant increase in its density in ASD-derived neurons (**Extended Data Fig. 9e,f**). Moreover, synapse number quantification revealed a significant increase in Synapsin I-immunoreactive $(SynI)^+$ puncta in ASD-derived neurons, suggesting increased neuronal maturation and synaptic overgrowth (**Extended Data Fig. 9e,g**). This is in agreement with the upregulated expression of the *blue* and *magenta* modules (**Fig. 1a**), whose signature suggested accelerated or increased neuronal differentiation in probands (**Table 1**).

Next, we directly tested whether there was any bias for differentiation into specific neuronal subtypes. In this analysis, we used as markers TFs which control fate choice, cell proliferation, and differentiation during normal telencephalic development, many of which were members of the *magenta* and *blue* modules (**Fig. 1c; Supplementary Table 6**). We found that the proportions of cortical excitatory neuron precursors of the subventricular zone expressing EOMES/TBR2, of layer 6 neurons expressing TBR1, and of early-born layer 5 CTIP2 neurons (also known as BCL11B) were not significantly different in ASD and control organoids at TD31 (**Fig. 2a-h**), although some of these cortical excitatory neuron markers (such as TBR1, TBR2, CTIP2 and SOX5) were upregulated in proband-derived organoids at the mRNA level.

We next investigated determinants of GABAergic inhibitory neuronal fates using DLX1-2 (two TFs which are among the earliest determinants of GABAergic fate in telencephalic

precursors cells and upregulated members of the *magenta* module) and GAD1/GAD67 (the GABA synthetic enzyme, an upregulated member of the *tan* module). The expression of these GABAergic markers was increased significantly in organoids derived from ASD individuals compared to those from unaffected family members. This increase was strong at TD31 and was already detectable at TD11 (**Fig. 2i-l**). Interestingly, GABA precursors arose in a segregated fashion in a restricted area of the organoids, devoid of TBR1+ excitatory neuron precursors (**Extended Data Fig. 10**), and then migrated widely throughout the aggregate. This suggests that these organoids contain autonomously regulated systems that mimic dorso-ventral mammalian telencephalic development[10]. Immunostaining for ASCL1/MASH1 and NKX2.1 (two TFs expressed by GABAergic progenitor cells) and the neurotransmitter GABA also showed an increase in their expression in ASD-derived organoids (**Extended Data Fig. 11**). Along with the upregulation of GSX2, ASCL1, DLX1, DLX2, DLX5, DLX6 and DLX6-AS1 (which is the top upregulated gene at TD11) (**Supplementary Table 2**) these cellular analyses strongly suggest an overproduction of progenitors and neurons of the GABAergic lineage as well as an altered balance between excitation and inhibition in organoids derived from probands (**Fig. 2m**).

We also found evidence for increased expression of GABAergic phenotypes electrophysiologically. Although the size of sodium currents in control- and proband-derived neurons was similar, there was substantial cell-to-cell variation in the voltage dependence of activation and inactivation, suggesting that different neurons express different proportions of brain sodium channel isoforms ($Na_v1.1$, $Na_v1.2$, $Na_v1.3$, $Na_v1.6$). **Figure 3** shows results for steady-state inactivation. Neurons from probands and their familial controls displayed substantial variation in the pre-test potential (V_{pre}) at which the peak sodium current was reduced to half its maximal amplitude ($E_{h1/2}$), and adequate fits to some of the data required two Hodgkin-Huxley components (**Fig. 3d,e**). Interestingly, sodium currents in proband-derived neurons tended to inactivate at more hyperpolarized membrane potentials than the corresponding currents in control neurons (**Fig. 3c-e;** n=7 ASD and 10 control neurons). The increased proportion of proband-derived neurons that gave $E_{h1/2}$ values in the range -72 to -65 mV (**Fig. 3f**) is consistent with increased expression of the $Na_v1.1$ isoform in these cells. This isoform is preferentially expressed in GABAergic interneurons[15-18] and the increased proportion of cells with this sodium channel phenotype is consistent with a greater proportion of GABAergic neurons in cortical organoids from probands with ASD.

Role of FOXG1 overexpression in causing deregulated cell differentiation in ASD organoids

Our DGE results show that DLX6-AS1, TMEM132C, FOXG1, C14orf23 and KLHDC8A are consistently among the top 10 upregulated genes at both TD11 and TD31 (**Supplementary Table 2**). Among these genes, FOXG1, which is one of the top 100 hubs in the *magenta* module, with an 8.5- and 13-fold increase in expression at TD11 and TD31, respectively (**Supplementary Table 2**), is a transcription factor important in the development of the telencephalon[19-21]; notably, loss of function mutations in FOXG1 have been found in patients with atypical Rett Syndrome[22-24] and confer a small brain size[25]. This is interesting, as the probands under investigation presented with large brain size, suggesting that FOXG1 may be, at least in part, involved in the modulation of the brain size phenotype, and possibly in the social disability component of the phenotype.

We therefore tested the hypothesis that abnormally high levels of FOXG1 and its downstream genes could be responsible for the phenotypic abnormalities identified in neuronal cells of macrocephalic ASD patients. To this end, using lentiviruses carrying short hairpin RNAs (shRNAs) targeting FOXG1, we tested whether an attenuation of the FOXG1 expression level in patients' neural cells was able to revert some of the neurobiological alterations.

As proof of principle, we generated four stable iPSCs lines (three of which stably expressed different shRNAs specifically targeting FOXG1 and one expressed a non-targeting shRNA control) from a proband-derived iPSC line (07-P#9). To confirm stable downregulation of FOXG1 expression, we performed qPCR analyses at TD11 (**Fig. 4a**). Introduction of two FOXG1-targeting shRNAs (shRNA-2 and 3) was able to down-regulate FOXG1 expression to a level comparable to that of the unaffected family member (**Fig. 4a,** compare bar 6 and 7 with bar 2). Immunostaining for FOXG1 confirmed that shRNA-2 and 3 were able to down-regulate its expression also at the protein level (**Fig. 4b-f**). We next analyzed the expression of GABAergic markers after FOXG1-RNA interference (RNAi) at the transcript and protein level. At TD11, organoids derived from the iPSC lines stably expressing FOXG1 shRNA-2 and shRNA-3 (07-P#9 shRNA-2 and 3) showed downregulation of DLX1, DLX2, and GAD1 transcripts as compared to the same iPSC line expressing the shRNA control (07-P#9 shRNA-C) (**Fig. 4g-i**). Immunostaining and stereological quantification of DLX1-2 and GAD1 positive cells showed that FOXG1 RNAi restored the normal level of GABAergic neuronal differentiation in proband-derived organoids at both TD11 and TD31 (**Fig. 4k-m**). These results suggest a causal role for FOXG1 and its downstream genes in the overproduction of neurons of the GABAergic lineage found in ASD-derived organoids. FOXG1 RNAi had no or minor effects on the transcript/protein expression levels of dorsal forebrain markers (such as PAX6) (**Fig. 4b-f**, and **j, l, m**), or on TFs directing cortical excitatory neuron differentiation (such as TBR1) (**Fig. 4m**).

To investigate the mechanism by which FOXG1 could affect the overproduction of GABAergic neurons, we compared cell proliferation in ASD- and control-derived organoids by BrdU incorporation with or without FOXG1 RNAi. Quantification of double-labeled $BrdU^+/Ki67^+$ cells at TD11 revealed no general changes in proliferation between proband- and control-derived organoids (**Extended Data Fig. 12a, b**). However, there was a significant increase in the number of $DLX2^+$ cells that incorporated BrdU in proband-derived organoids and an increased proportion of $BrdU^+$ cells that colocalized DLX1/2. Both effects were precluded by FOXG1-knockdown (**Fig. 5a-c**). Furthermore, at TD31, proband-derived organoids showed a greatly increased proportion of DLX^+ and $GAD1^+$ cells that had incorporated BrdU at TD11 (**Fig. 5d-h**), an effect that was also greatly attenuated by FOXG1 RNAi, which lowered the proportion of $BrdU^+/DLX^+$ and $BrdU^+/GAD1^+$ cells in ASD-derived organoids to levels comparable to those of the unaffected control (**Fig. 5a-h**). Moreover, $GAD1^+$ cells at TD31 were not aberrantly entering the cell cycle, suggesting that GABAergic neurons were terminally differentiated in the organoids (**Extended Data Fig. 12c, d**). Taken together, these data suggest that the early increase in proliferation of GABAergic neuronal progenitor cells in proband-derived organoids gave rise to an increased proportion of mature GABAergic interneurons, and that FOXG1 RNAi restored both these early and late effects to levels comparable to those found in unaffected individuals. Similar experiments revealed smaller or non-significant changes in the proliferation of $PAX6^+$ and $TBR1^+$ precursor cells after FOXG1 RNAi (**Extended Data Fig. 12**), suggesting that upregulated FOXG1 expression in ASD neural cells was driving an early proliferative effect in neuronal precursor cells of the GABAergic lineage.

Correlation with clinical phenotype

Finally, we sought to understand whether dysregulated gene expression in the *blue*, *magenta*, and *brown* modules, and in particular FOXG1, are incidental findings or might be important components of the ASD clinical phenotype in the five families in this study. First, we found that head circumference (HC) is highly positively correlated with the ADOS autism symptom severity score (r=0.84, nom p-value=0.07). The low number of individuals (n=5) rendered the p-value marginally significant, but the magnitude of the effect is very large and the results are consistent with recent reports suggesting that increased HC is correlated with lower IQ scores and with increased symptom severity in patients with ASD[8,9,26].
Next, we correlated the patients' HC and their autism symptom severity with gene expression indices. Despite the low number of samples, we saw unusually high correlation coefficients and low nominal p-values. We found a consistent positive correlation of HC and autism symptom severity with the upregulated modules as well as a negative correlation with the downregulated modules (**Fig.1d**). Specifically, the HC z-scores displayed a strong positive correlation with the *magenta* module's eigengene and level of FOXG1 gene expression at both time points (**Fig.1d**), implicating dysregulated gene expression in the *magenta* module, and FOXG1 in particular, as potential biomarkers of ASD with macrocephaly in our patient cohort. The autism severity score displayed a high positive correlation with the upregulated *blue, brown*, and *tan* modules, particularly at TD11. Together, the observed patterns of correlations suggest that the upregulation in the *magenta*, *blue*, *brown* and *tan* gene network is a maladaptive trait in these families and that it may represent an important pathophysiological antecedent of symptoms.

Discussion

We identified common neurobiological pathways for idiopathic cases of ASD with macrocephaly, using four families as bases for a genome-wide transcriptome analysis in iPSC-derived cortical organoid cultures that recapitulate human first trimester telencephalic development. Despite the likely heterogeneity in genotypes underlying ASD, we have identified perturbations in coherent programs of gene expression and associated features of altered neurodevelopment, i.e., an upregulation of cell proliferation, unbalanced inhibitory neuron differentiation and exuberant synaptic development.

Biologically, the iPSC lines derived from ASD patients in this cohort show decreased cell cycle length, suggesting a generalized increase in proliferative potential, which may be an underpinning of macrocephaly in these individuals. This increased proliferation is also shown at early stages, but not at later stages, of cortical neural development *in vitro*, suggesting powerful compensatory events that eventually restrict excessive production of neurons in the cortical organoids and presumably also *in vivo*. These findings are in line with earlier hypotheses stating that abnormal control of cell proliferation and overproduction of neurons might explain the accelerated brain growth in ASD[27-29].

Driven by the transcriptome changes, we found that there is biased production of neuronal subtypes from cortical neuron precursors of ASD probands, as compared with their fathers. Cortical organoids of ASD probands show, at all time points analyzed, exuberant GABAergic differentiation and no change in glutamate neuron types, which together cause an imbalance in glutamate/GABA neuron ratio. We also found that the overproduction of GABAergic cells is attributable, at least in part, to an early increase in FOXG1 gene expression, which drives an increased proliferation and number of GABA precursor cells expressing GABAergic TFs of the ventral telencephalon, including the DLX homeobox genes. FOXG1 inactivation in mouse causes premature lengthening of telencephalic progenitor cell cycles and a failure to specify ventral (GABAergic) telencephalic precursors, leading to a severely hypoplastic telencephalon[20,30]. Hence, our data in humans are consistent with the known roles of FOXG1 in telencephalic growth as well as in the determination of GABA neuron fate.

Our genomic data do not support the presence of any previously known deleterious mutation in an ASD candidate gene. Although we found no structural or sequence DNA variation in the FOXG1 locus that could explain the increase in FOXG1 gene expression (i.e., a duplication involving the FOXG1 locus or a mutation in a TF binding site in the proximal promoter region), we cannot exclude that uncharacterized SNVs in distal regions could affect FOXG1 expression. Future research may uncover whether the gene expression dysregulation that we identify in our patients with ASD and macrocephaly is due to novel mutations in genes that collectively affect the regulatory network of FOXG1.

In addition to GABAergic neuron overproduction, we also find evidence of exuberant cellular overgrowth of neurites and synapses at the transcriptome and cellular levels; this is consistent with increased spine densities found in the postmortem cerebral cortex of ASD subjects[31,32] and in the *fmr1* KO mouse, a model for Fragile X syndrome, one of the most common forms of inherited intellectual disability and ASD[33]. Consistently with the cellular phenotype, mRNA for the synaptic adhesion molecules NLGN1, NRXN1, 2, and 3 were all overexpressed in patient-derived organoids. In contrast to the present results, rare loss-of-function mutations in synaptic adhesion molecules (SHANK, NLGN, NRXN) suggest a deficiency in synaptic connections in some individuals with ASD, and deficient GABAergic synapses in particular. However, a presumed gain-of-function mutation in NLGN-3 (R451C), which is also found in ASD, confers an increase in GABAergic synaptic signaling[34-36]. Furthermore, both increases and decreases in the GABA synthesizing enzyme GAD have been reported in the *fmr1* KO mouse[37]. These apparently discrepant results may be reconciled by the suggestion that the reciprocal balance, rather than absolute numbers, of glutamate and GABA neurons is important for function.

It is also possible that overproliferation, imbalanced increase in GABA neuron fate, and synaptic overgrowth are features of the aberrant trajectory of cortical development found specifically in children with ASD and macrocephaly. These children represent a sizable portion of the “idiopathic” population, and their macroscopic overgrowth generally has been found to confer a poor outcome. Indeed, in our five-patient cohort we detected not only a correlation between degree of macrocephaly and symptomatology, but also a correlation between dysregulated modules of gene expression, and FOXG1 gene expression in particular, with symptomatology. Although the small number of subjects precluded a definitive conclusion, these correlations suggest that the transcriptome differences we identified are not incidental findings but represent a core feature of the

disorder with prognostic implications. Reinforcing the likely maladaptive function of altered levels of FOXG1, deletions and missense mutations in this gene have been associated with an atypical Rett Syndrome and small brain size[22,23,38]. This is interesting, as it suggests that deviations in FOXG1 levels during brain development, in excess and defects, cause opposite modulation in brain growth but a similarly disabling outcome, characterized by intellectual disability and ASD-like symptoms.
The results presented here demonstrate the first shared mechanism for idiopathic ASD. This work implies that a number of common and rare causative genetic risk factors can converge upon common mechanisms of pathogenesis. They also illustrate that directly studying neurodevelopmental processes in patients with neuropsychiatric disorders that have heterogeneous etiologies can open novel inroads into diagnosis and therapy[39].

References

1 Sudhof, T. C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903-911 (2008).
2 Zoghbi, H. Y. & Bear, M. F. Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harbor perspectives in biology* **4** (2012).
3 Rubenstein, J. L. Three hypotheses for developmental defects that may underlie some forms of autism spectrum disorder. *Current opinion in neurology* **23**, 118-123 (2010).
4 Casanova, M. F., Buxhoeveden, D. & Gomez, J. Disruption in the inhibitory architecture of the cell minicolumn: implications for autisim. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* **9**, 496-507 (2003).
5 Willsey, A. J. *et al.* Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997-1007 (2013).
6 Parikshak, N. N. *et al.* Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008-1021 (2013).
7 Courchesne, E. *et al.* Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology* **57**, 245-254 (2001).
8 Chaste, P. *et al.* Adjusting head circumference for covariates in autism: clinical correlates of a highly heritable continuous trait. *Biol Psychiatry* **74**, 576-584 (2013).
9 Chawarska, K. *et al.* Early generalized overgrowth in boys with autism. *Arch Gen Psychiatry* **68**, 1021-1031 (2011).
10 Mariani, J. *et al.* Modeling human cortical development in vitro using induced pluripotent stem cells. *Proc Natl Acad Sci U S A* **109**, 12770-12775 (2012).
11 Abyzov, A. *et al.* Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* **492**, 438-442 (2012).
12 Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483-489 (2011).
13 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559 (2008).
14 Abrahams, B. S. *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular autism* **4**, 36 (2013).

15 Yu, F. H. *et al.* Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. *Nat Neurosci* **9**, 1142-1149 (2006).
16 Ogiwara, I. *et al.* Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation. *J Neurosci* **27**, 5903-5914 (2007).
17 Cheah, C. S. *et al.* Specific deletion of NaV1.1 sodium channels in inhibitory interneurons causes seizures and premature death in a mouse model of Dravet syndrome. *Proc Natl Acad Sci U S A* **109**, 14646-14651 (2012).
18 Han, S. *et al.* Autistic-like behaviour in Scn1a+/- mice and rescue by enhanced GABA-mediated neurotransmission. *Nature* **489**, 385-390 (2012).
19 Xuan, S. *et al.* Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. *Neuron* **14**, 1141-1152 (1995).
20 Martynoga, B., Morrison, H., Price, D. J. & Mason, J. O. Foxg1 is required for specification of ventral telencephalon and region-specific regulation of dorsal telencephalic precursor proliferation and apoptosis. *Dev Biol* **283**, 113-127 (2005).
21 Hanashima, C., Li, S. C., Shen, L., Lai, E. & Fishell, G. Foxg1 suppresses early cortical cell fate. *Science* **303**, 56-59 (2004).
22 Ariani, F. *et al.* FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* **83**, 89-93 (2008).
23 Bahi-Buisson, N. *et al.* Revisiting the phenotype associated with FOXG1 mutations: two novel cases of congenital Rett variant. *Neurogenetics* **11**, 241-249 (2010).
24 Shoichet, S. A. *et al.* Haploinsufficiency of novel FOXG1B variants in a patient with severe mental retardation, brain malformations and microcephaly. *Human genetics* **117**, 536-544 (2005).
25 Kortum, F. *et al.* The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *J Med Genet* **48**, 396-406 (2011).
26 Campbell, D., Chang, J. & Chawarska, K. Early generalized overgrowth in Autism Spectrum Disorder: Prevalence rates, gender effects, and clinical outcomes. *in press*.
27 Courchesne, E. *et al.* Mapping early brain development in autism. *Neuron* **56**, 399-413 (2007).
28 Vaccarino, F. M., Grigorenko, E. L., Smith, K. M. & Stevens, H. E. Regulation of cerebral cortical size and neuron number by fibroblast growth factors: implications for autism. *J Autism Dev Disord* **39**, 511-520 (2009).
29 Courchesne, E. *et al.* Neuron number and size in prefrontal cortex of children with autism. *JAMA* **306**, 2001-2010 (2011).
30 Fasano, C. A. *et al.* Bmi-1 cooperates with Foxg1 to maintain neural stem cell self-renewal in the forebrain. *Genes Dev* **23**, 561-574 (2009).
31 Hutsler, J. J. & Zhang, H. Increased dendritic spine densities on cortical projection neurons in autism spectrum disorders. *Brain Res* **1309**, 83-94 (2010).
32 Tang, G. *et al.* Loss of mTOR-Dependent Macroautophagy Causes Autistic-like Synaptic Pruning Deficits. *Neuron* (2014).
33 Dolen, G. *et al.* Correction of fragile X syndrome in mice. *Neuron* **56**, 955-962 (2007).
34 Tabuchi, K. *et al.* A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* **318**, 71-76 (2007).

35 Etherton, M. *et al.* Autism-linked neuroligin-3 R451C mutation differentially alters hippocampal and cortical synaptic function. *Proc Natl Acad Sci U S A* **108**, 13764-13769 (2011).
36 Pizzarelli, R. & Cherubini, E. Developmental regulation of GABAergic signalling in the hippocampus of neuroligin 3 R451C knock-in mice: an animal model of Autism. *Frontiers in cellular neuroscience* **7**, 85 (2013).
37 Paluszkiewicz, S. M., Martin, B. S. & Huntsman, M. M. Fragile X syndrome: the GABAergic system and circuit dysfunction. *Developmental neuroscience* **33**, 349-364 (2011).
38 Mencarelli, M. A. *et al.* Novel FOXG1 mutations associated with the congenital variant of Rett syndrome. *J Med Genet* **47**, 49-53 (2010).
39 Volkmar, F. R., State, M. & Klin, A. Autism and autism spectrum disorders: diagnostic issues for the coming decade. *Journal of child psychology and psychiatry, and allied disciplines* **50**, 108-115 (2009).
40 Catterall, W. A., Goldin, A. L. & Waxman, S. G. International Union of Pharmacology. XLVII. Nomenclature and structure-function relationships of voltage-gated sodium channels. *Pharmacological reviews* **57**, 397-409 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements
We acknowledge support from the NIH and from the Harris Professorship fund. We also acknowledge the support team of the Yale University Biomedical High Performance Computing Center (in particular, Robert Bjornson and Nicholas Carriero). We thank Anita Huttner for culturing fibrobasts, Ying Zhang for preparing the sequencing libraries, and Nathaniel Calixto, Anahita Amiri for help in iPSC maintenance and characterization. We thank Elizabeth Jonas for facilities and help with the preliminary patch clamp experiments on dissociated cultures. We acknowledge the following grant support: NIMH MH089176, MH087879 (FMV), U54 MH066494 (Project 2 KC), and the State of Connecticut (FMV). This work was supported in part by grants from the Simons Foundation (SFARI #137055, FMV and SFARI #206929 R10981, KAP). We are grateful to all of the families at the participating Simons Simplex Collection (SSC) sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, E. Wijsman). We appreciate obtaining access to phenotypic data on SFARI Base. Approved researchers can obtain SSC population data by applying at https://base.sfari.org. We acknowledge the Yale Center for Clinical Investigation for clinical support in obtaining the biopsy specimens. We thank Dr. John Overton, the Yale Center for Genome Analysis and the Stanford Genomic Facility for advice in carrying out DNA and RNA sequencing.

Author contribution
The authors contributed this study at different levels, as described in the following. Study conception and design: F.M.V., J.M., G.C., J.H. Family selection and characterization:

K.C., K.P. Skin biopsy: A.S. iPSC generation and characterization: J.M., L.T., L.P., M.A. Neuronal differentiation: J.M., L.T, L.P. Electrophysiology: P.Z., J.H. RNAi: J.M. Sequencing library preparation: Y.Z. Processing and analysis of RNAseq data: G.C., D.P. Processing and analysis of DNAseq data, CNV and SNV discovery: A.A., M.W. qPCR validation: J.M. Human subjects: K.C. and K.P. Coordination of analyses: F.M.V. Display item preparation: J.M., G.C., F.M.V., L.T., J.H. Writing manuscript: J.M., G.C., F.M.V., J.H. All authors participated in discussion of results and manuscript editing.

Author Information

Competing Interest: The Authors declare no competing interest

Correspondence and requests for materials should be addressed to:
flora.vaccarino@yale.edu

Figure Legends

Figure 1. WGCNA network in neuronal cells from ASD patients and unaffected family controls and correlation with clinical phenotypes. **a**, Relationship between the modules' eigengenes with diagnosis and time in culture. TD11: 11 days of terminal differentiation; TD31: 31 days of terminal differentiation; F: fathers; P: probands. **b**, Module-to-module relationship by hierarchical clustering. **c**, Top 200 hub genes networks for three representative modules. Circles: genes; diamonds: genes overlapping with genes in the SFARI database classified as associated to ASD; red: genes overexpressed at either TD11 or TD31; grey: no changes in gene expression; larger font: genes differentially expressed at both TD11 and TD31. **d,** Pearson's correlation between modules' eigengenes with (log transformed) head circumference Z-score and ADOS severity score at TD11 and TD31. Also represented is the correlation between FOXG1 expression levels and the clinical scores.

Figure 2. ASD organoids show imbalance between glutamatergic and GABAergic neuron fate. Representative images of control-derived and ASD proband-derived organoids: **a-h**, Immunostaining and respective stereological quantification of SOX1/2 and PAX6 proliferating radial glia progenitors, cortical excitatory $TBR2^+$ intermediate progenitors, and more mature $TBR1^+$ and $CTIP2^+$ excitatory neurons at TD31. **i-l**, Immunostaining and quantification of GABAergic inhibitory progenitor cells ($DLX1\text{-}2^+$) and mature GABAergic interneurons ($GAD1^+$) at TD11 and TD31. **m**, Bar graph showing percentages of inhibitory ($GAD1^+$) and excitatory ($TBR1^+$, $CITIP2^+$) in ASD-derived and control-derived organoids at TD31. Total cells are estimated by counting $DAPI^+$ nuclei. C=Controls, P=Probands. ** $P < 0.01$, *** $P < 0.001$, t test analysis. Scale bars, 10 μm (**a**, **c**, **e**, **g**, **k**), 20 μm (**i**, **j**).

Figure 3. Sodium currents in iPSC-derived neurons from ASD patients and controls show different voltage dependence of steady-state inactivation. a, **b**, Inward sodium currents evoked by depolarizing test jumps to -20 mV from pre-test

potentials of -90 mV to -25 mV (**a**) or -45 mV (**b**) in steps of 5 mV. **c**, Peak inward current amplitudes as a function of pre-test potential from the data in **a** and **b**. The results were normalized to Imax values obtained for each neuron from fits to the raw data (as in **d** and **e**). **d**,**e**, Steady-state inactivation data obtained from 10 control neurons (**d**) and seven patient neurons (**e**). Data for individual cells are shown with different symbols and colors (normalized data in **c** are shown here with the same symbols: control, open black square; patient, filled black circle). For some neurons, the results were better fitted as the sum of two components (arrowheads point to $E_{h1/2}$ values for each component). **f**, Bar graph depicting the percentage of control and patient neurons (n = 10 and 7) with $E_{h1/2}$ values that fell within the indicated ranges. When two components were present, the fractional amplitudes of each were used in the calculation of mean percentages for each group. Cells that gave $E_{h1/2}$ values ≤ -65 mV gave half-activation voltages that ranged from -39.1 to -55.8 mV, a phenotype most consistent with the $Na_v1.1$ brain sodium channel isoform[40].

Figure 4. FOXG1 knockdown in ASD-derived organoids is able to restore the balance between GABA/glutamate neuron fate. Two out of three FOXG1-targeting shRNAs (shRNA-2, shRNA-3) downregulate FOXG1 expression in proband (line 07-P#9) to levels comparable to those present in the unaffected father (lines 07-F#1, 07-F#2). **a,** Relative expression levels of FOXG1 by qPCR among non virally transduced undifferentiated iPSCs from proband #9 (i07-P#9), differentiated organoids at TD11 from proband, his father, and proband's organoids harboring a non-targeting shRNA-control (shRNA-C) or three different shRNAs targeting FOXG1 (shRNA-1, shRNA-2, shRNA-3). **b-f**, Double immunostaining for FOXG1 and PAX6 in TD11 proband's organoids either non-virally transduced (**b**), or virally transduced with shRNA-C (**c**), shRNA-1 (**d**), shRNA-2 (**e**), shRNA-3 (**f**). **g-j,** qPCR for DLX1 (**g**), DLX2 (**h**), GAD1 (**i**), and PAX6 (**j**) in TD11 organoids from shRNA-C and shRNA-1/2/3. **k,** DLX1-2 and GAD1 double immunostaining in organoids derived from the father or the proband transduced with shRNA-C or shRNA-3 at TD11 and TD 30. **l, m,** Stereological quantification of immunocytochemical (ICC) staining for GABAergic (DLX1-2, GAD1) and glutamatergic markers (PAX6, TBR1) at TD11 (**l**) and TD 31 (**m**). (Sample names: 07= family name from which iPSCs were derived; F=Father; P=Proband; #=iPS clone number). Error bars are s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, t test analysis. Scale bars, 10 μm.

Figure 5. Increased proportion of proliferating GABAergic neuronal progenitors and mature GABAergic interneurons at TD31 in ASD-derived organoids. Representative images (**a, d**) and stereological quantification of $BrdU^+/DLX1\text{-}2^+$ proliferating cells (**b, c, e, f**) and $BrdU^+/GAD1^+$ neurons (**g, h**) in TD11 and TD31 organoids derived from 07-F#2 (father), 07-P#9 shRNA-C, and 07-P#9 shRNA-3 (proband's iPSC lines transduced with shRNA-C or with shRNA-3). The selectively increased proportion of DLX $1\text{-}2^+/BrdU^+$ double positive cells in patient-derived organoids at both TD11 (**b, c**) and TD31 (**e, f**), was restored to a physiological level after FOXG1-knockdown at both time points. The increased proportion of proliferating DLX1-2 GABAergic progenitors (**a** and **d**, upper panel) resulted in an overproduction of more mature GABAergic GAD1+/BrdU+ double positive interneurons (**d**, bottom panel), overproduction that was restored to physiological levels after FOXG1-knockdown (**g, h**). Error bars are s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, t test analysis. Scale bars, 10 μm.

Detailed Methods

Participants. Participants consisted of five probands diagnosed with Autistic Disorder or Pervasive Developmental Disorder – Not Otherwise Specified[41], referred in text jointly as Autism Spectrum Disorder (ASD) (See Supplementary Table 1 for full subjects description). The probands were recruited from a larger pool of participants evaluated through several research projects at the Yale Child Study Center Autism Program and Center for Translational Developmental Neuroscience. The inclusion criteria were male gender and head circumference (HC) above the 95th percentile compared to the CDC population norms[42] (n=3) or Roche HC norms (n=2)[43] at the time of entry into the feeder studies. HC was measured again at the time of the proband biopsy, and ranged from 92 to >95 percentile[43] (Supplementary Table 1). Informed consent was obtained from each participant enrolled in the study according to the regulations of the Institutional Review Board and Yale Center for Clinical Investigation at Yale University).

The autism symptom severity was assessed using the Autism Diagnostic Observation Schedule (ADOS)[44] and Autism Diagnostic Interview (ADI-R)[45]. Verbal and nonverbal functioning was assessed with the Mullen Scales of Early Learning (MSEL)[46] or Differential Ability Scales- Second Edition (DAS-II), and adaptive skills were assessed using the Vineland Adaptive Behaviors Scale – Second Edition (VABS-II)[47]. Diagnosis of ASD was assigned by a team of expert clinicians based on a review of medical and developmental history and comprehensive psychological and psychiatric assessment including the ADOS and ADI-R. The probands displayed a wide range of autism severity scores (ADOS CSS: 3 to 10) as well as verbal and nonverbal IQ scores (Standard Score range 25 – 107) (see Supplementary Table 1).[44]
The control group consisted of mothers (n=3), fathers (n=4), and unaffected siblings (n = 1). The sibling's classification as "unaffected" was confirmed through a comprehensive diagnostic evaluation. The diagnostic evaluation *ruled out*: 1) Diagnosed, referred, or suspected ASD, schizophrenia, or other developmental or psychiatric disorder; 2) An individual education plan for special education services, including speech/language therapy, occupational therapy, and social skills therapy; 3) Low or moderately low score on any domain of the Vineland Adaptive Behavior Scale-II; 4) Total T score >76 (severe range) on the SRS; 5) Clinical impression suggesting ASD, other developmental delay/disorder, or psychiatric disorder.

Neuronal differentiation of iPSCs. hiPSCs were generated by retroviral infection and characterized as previously described[11]. Several iPSC lines derived from subject 1120-01 were generated by a viral-free episomal reprogramming method[48].

For neuronal differentiation, undifferentiated iPSC colonies were pretreated for one hour with 50 μM Y27631 compound (Calbiochem), dissociated with Accutase (Millipore, 1:2 dilution in PBS 1X). A total of 3.2 million single cells were cultured for 2 days on Aggrewell™ 800 plates (STEMCELL Technologies) in DMEM/F12-GLUTAMAX-type medium containing 4% B27 supplement without vitamin A (Invitrogen), 1% N2 supplement (Invitrogen), and 55 μM 2-mercaptoethanol (2-ME). To induce forebrain fate, the medium was supplemented with 5μM Y-27632 and 200ng/ml recombinant mouse Noggin (R&D Systems, 1967-NG-025). After 2 days, embryoid bodies (EBs) were collected and plated onto 10-cm bacterial Petri dishes in 10 ml neuronal medium as above with the addition of B27 supplemented with vitamin A. The medium was supplemented with Y-27632 and Noggin as above. After two days (day 4), free-floating

EBs were collected and plated in neuronal medium, supplemented with only Noggin, onto 10 cm tissue culture dishes coated with growth factor-reduced Matrigel (BD Bioscience, diluted 1:30 with DMEM:F12 medium) to allow neural rosette formation. The next day (day 5), the neuronal medium was changed and supplemented with 20 ng/ml FGF2, 200 ng/ml Noggin, and 200 ng/ml rhDkk1 (R&D Systems, 5439-DK). After two or three days on Matrigel-coated dishes, the neural rosettes were manually dissected and replated as free-floating aggregates in 10-cm bacterial Petri dishes in 10 ml neuronal medium supplemented with FGF2 (10ng/ml) and EGF (10ng/ml). After five days in suspension, free-floating aggregates were plated onto ultra low-attachment 96-well plates (Corning) as single aggregates to avoid the formation of aggregate clumps in neuronal medium supplemented with FGF2 and EGF. Terminal differentiation was started the next day, using a NEUROBASAL-type medium supplemented with 1% N2, 2% B27 (with vitamin A), 15 mM HEPES, 1:100 Glutamax, 1:100 nonessential amino acids (NEAA) and 55 μM 2-ME. This medium was supplemented with 200 nM ascorbic acid, 10 ng/ml BDNF (R&D), 10 ng/ml GDNF (R&D) and 1 mM dibutyryl-cAMP (Sigma). Free-floating aggregates were analyzed after 11 or 31 days form the beginning of terminal differentiation. Half of the medium was changed twice a week.
To asses the cell cycle length at TD11 shown in **Extended Data Fig. 8b**, we used adherent monolayer cultures, where dissociated single cells (1-2 x 10^5 cells per cm^2) were plated onto poly-L-ornithine (50μg/ml) and laminin (5μg/ml) coated 8-well permanox chamber slides (Thermo Scientific) at the beginning of terminal differentiation.

shRNA knockdown experiments. Lentiviruses carrying three different shRNAs specifically targeting human FOXG1 (shRNA-1: TRC0000013949; shRNA-2: TRCN0000013952; shRNA-3: TRCN0000273934) and a non-targeting control-shRNA lentiviral vector (shRNA-C: SHC016V) were selected from the SIGMA MISSION^R shRNA Library and purchased as lentiviral transducing particles at a defined transducing units concentration (TU/ml). Two different multiplicities of infections (MOI=2 and MOI=5) were used for infection and puromycin selection of undifferentiated iPSCs. The optimal concentration of puromycin for the selection of transduced iPSC colonies was established by puromycin titration on the specific iPSC line used for all the experiments. Undifferentiated colonies of iPSCs were pretreated for one hour with 50 μM Y27631 compound, dissociated with Accutase (1:2 dilution in PBS 1X), and 1x10^5 cells were plated in each well of a 12-well plate (Corning) coated with Matrigel in iPSC medium (DMEM/F12 containing 1% N2 supplement, 2% B27 supplement, 2mM L-glutamine, 0.1mM NEAA, 1% form a 10.000units/ml penicillin and 10.000 μg/ml streptomycin stock solution, 0.5 mg/ml BSA Fraction V, 0.12 mM monothioglycerol) supplemented with 100 μM Y27631 compound and 80 ng/ml of FGF2. The next day, cells were infected with each shRNA-containing lentivirus at the MOI of 2 and 5, and the medium was changed after 24 h. 72 h after the infection, puromycin selection was started with 0.5 μg/ml of puromycin. Resistant iPSC colonies were expanded and maintained with 0.5 μg/ml of puromycin. Stable lines were used for neuronal differentiation experiments after two weeks of selection, with puromycin kept in all the neuronal media at 0.3 μg/ml.
Stable downregulation of FOXG1 was assessed by qPCR at TD11, and those cultures were selected for further neuronal differentiation experiments.

RNA isolation, qPCR, and RNA sequencing experiments. Total RNA was prepared from undifferentiated iPSCs and cortical organoids at TD11 and TD 30 *in vitro*. Total RNA was extracted using the PicoPure RNA isolation kit (Applied Bioscience). RNA sequencing was performed using the same input amount of total RNA with the

SMARTer® Ultra Low RNA kit for Illumina® Sequencing according to the manufacturer's instructions. For qPCR validations, cDNA was synthesized, starting from 100 ng of total RNA with the SuperScript III First-Strand Synthesis kit and random hexamers (Invitrogen). qPCR was performed using the StepOnePlus Real-Time PCR System (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems), and the data were normalized for GDPH expression. All primers used are listed in **Supplementary Table 9**. Paired-end RNA sequencing (2x75) was performed on an IIllumina Hi Seq2000 to an average depth of 30M reads per sample.

Immunostaining and data analysis. At TD11 and TD31 *in vitro*, individual aggregates were fixed in 4% PFA for two to three hours, incubated in 25% sucrose overnight, embedded in OCT, and cryosected at 12 μm. Cryosections were blocked in 10% normal donkey serum/0.1% Triton X-100 in PBS for 1 h at room temperature. Primary antibodies were diluted in 5% normal donkey serum/0.05% Triton X-100 in PBS and incubated overnight at 4°C. Samples were then washed in PBS/0.1% Triton X-100 three times for 5 min and incubated with secondary antibodies (1:500, Alexa or Jackson Laboratories) in 5% normal donkey serum/0.05% Triton X-100 in PBS 1 h at room temperature. For BrdU staining, sections were first incubated with 2M HCl at 37 °C for 30 min, followed by three washes in PBS before blocking. Primary antibodies were diluted as follows: NESTIN (mouse, Millipore, MAB5326, 1:200), PAX6 (mouse, BD Bioscience, 561462, 1:200), TUBB3 (mouse, Promega, G712A, 1:1000), BLBP (rabbit, Millipore, ABN14, 1:500), N-cadherin (mouse, BD Bioscience, 610920, 1:500), TBR1 (rabbit, Abcam, ab31940, 1:1000), TBR2 (rabbit, Abcam, ab23345, 1:1000), SOX2 (rabbit, Millipore, AB5603, 1:1000), SOX1 (goat, R&D System, AF3369, 1:20), GAD1/GAD67 (mouse, Millipore, MAB5406, 1:1000), CTIP2/BCL11B (rat, Abcam, ab18465, 1:500), SATB2 (mouse, Abcam, ab51502, 1:200), DLX1-2 (rabbit, gift of Yuri Morozov, Yale University, New Haven, CT, 1:1000), MASH1/ASCL1 (mouse, gift of Francois Guillemot, Medical Research Council National Institute for Medical Research, Mill Hill, London, UK, 1:200), NKX 2.1 (mouse, Thermo Scientific, MA5-16406, 1:200), BRN2 (goat, Santa Cruz, sc-6029, 1:100), Ki67 (rabbit, Vector, VP-RM04, 1:500), pH3 (rat, Sigma, H9908, 1:1000), BrdU (rat, OBT0030, AbD Serotec,1:200), GFP (chicken, Abcam, ab13970), FOXG1 (rabbit, NCFAB, StemCulture Incorporated, 1:500), MAP2 (mouse, Millipore, MAB3418, 1:1000), Synapsin I (rabbit, Millipore, AB1543, 1:500), GABA (rabbit, Sigma, A2052, 1:2000), DCX (rabbit, Abcam, ab18723, 1:1000), NeuN (mouse, Millipore, MAB377, 1:500).

All images were acquired using an ApoTome-equipped Axiovert 200M with Axiovision 4.5 software. Quantification of immunostained cells was performed by stereological analysis using a Carl Zeiss Axioskope 2 Mot Plus, connected to a computer running Stereoinvestigator Software (MicroBrigh-Field). Nuclear profiles of stained cells were counted using the optical fractionator probe using a 100x oil-immersion objective. Sample grids of 180 μm x 170 μm were used to obtain a coefficient of error ≤ 0.5. Tridimensional sampling boxes (50 μm x 50 μm x 10 μm) with three out of six exclusion borders were automatically placed by Stereoinvestigator at each grid intersection point. The relative amount of positive cells was calculated as a percentage of total $DAPI^+$ cells.

Quantification of the average MAP2 density was performed using four sections per organoid (and two organoids per sample), and images were acquired to cover the entire extent of the organoid. MAP2 density was estimated using Fiji software. MAP2 density (for each field) was calculated using the formula: Integrated Density-(Area of selected field*Mean fluorescence of background readings) and plotted as average density of the different fields. Quantification of Synapsin I (number of synapses) was performed using

at least 20 random images of each organoid acquired using a 100x oil-immersion objective, and the number of Synapsin I puncta was estimated using a plug in Puncta Analyzer under the Fiji analysis software platform[49].

Electrophysiology. Whole-cell patch-clamp recordings were made from single iPSC-derived neurons at room temperature with an EPC-9 amplifier and PatchMaster software (HEKA). In preliminary experiments, the cells were dissociated after rosette formation/dissection, plated onto glass coverslips, and maintained in vitro for 30-60 days. Most of the results were obtained directly from the 3D organoid preparation used for the other types of experiments and recordings were made from single cells that appeared to have migrated out from the organoids that had been plated onto 8-well permanox chamber slides, coated with poly-L-ornithine (50µg/ml) and laminin (5µg/ml). The slides were placed in a chamber on the stage of an upright Zeiss Axioskope microscope and viewed with Nomarski optics at 640x magnification. The external solution was (in mM): 150 NaCl, 3 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 5 glucose, and 10 HEPES (pH 7.4). Patch pipettes (open tip resistance 3–8 MΩ) were filled with a solution containing (in mM): 135 CsF, 33 CsOH, 2 $MgCl_2$, 1 $CaCl_2$, 11 Na_2EGTA, and 10 HEPES (pH 7.4). In voltage-clamp mode, currents were acquired with protocols specified in PatchMaster, low-pass filtered (4-pole Bessel) at 3 kHz, and sampled at 50 kHz. The digitized records were written directly to a Mac mini computer (Apple). All currents were leak subtracted before analysis. Series resistance and cell capacitance (12.2 $\pm$ 1.4 MΩ, 20.5 $\pm$ 1.6 pF, n=41) were estimated from cancellation of the capacitive transients evoked by a small voltage step (10 mV, 10 ms). Series resistance compensation was not used. Spontaneous synaptic currents were recorded at a holding potential of -70 mV and were filtered and sampled as above. To test for action potential firing, depolarizing current steps (1 s) were made in current-clamp mode from a membrane potential of -70 mV. For analysis, the records were exported to Igor (Wavemetrics). Measurements of current amplitudes were made manually. Examination of histograms of these amplitudes for both inward and outward currents from individual cells indicated that the current amplitudes (and current densities, pA/pF) were log-normally distributed and the amplitudes were converted to log10 values before statistical comparisons were made. Conductance-voltage data from I-V protocols and data from h-infinity protocols were fitted with Boltzmann functions to estimate $E_{a1/2}$ and $E_{h1/2}$ values (the membrane potentials at which activation and steady-state inactivation were half-maximal). The number of components required to fit the results adequately was determined from comparison of Chi-squared values (>10% improvement was considered significant). Decay time constants for synaptic currents were estimated from single exponential fits after aligning the initial rise of the currents and averaging 20 to 30 events. Means from multiple groups were compared with one-way ANOVA, and P-values $\leq$ 0.05 were considered statistically significant. Values are reported as mean $\pm$ s.e.m.

RNA-seq data analysis. We used Tophat[50] to map the reads to the human genome (hg19) and the GencodeV7[51] transcriptome annotation. The mapped reads in BAM file format were converted to SAM format, using SAMtools[52], and sorted. Gene expression in counts was estimated using the BEDTools[53] function coverageBed, after converting the data from SAM to BED format. Gene expression in RPKM was estimated using the RSEQTools[54] function MRFquantifier, after converting the data form SAM to MRF format. This resulted in estimated expression levels for 36664 genes.

Differential expression analysis. Counts data were first filtered requiring that at least 50% plus one of the samples have a level of expression of 1 count per million or more. About 20000 genes survived the filter (19806 at TD11 and 20194 at TD31) and were considered for analysis. Differential expression analysis was performed by edgeR[55] using the trended dispersion option to estimate the biological variance. An FDR cut-off of 0.05 was used for all the tests. We accounted for a matched pair design, proband vs. father within the same family, by using the generalized linear model capability of edgeR, with blocking applied to family.

Weighted Gene Co-expression Network Analysis. We used WGCNA[13] for co-expression network analysis using gene expression estimates in RPKM and all the available samples for probands and fathers at both TD11 and TD31, for a total of 29 samples. All the clones were then averaged, resulting in a dataset of 16 samples. Expression data were first filtered, requiring that at least 50% plus one of the samples have an equivalent level of expression of 1 count per million or more, resulting in a set of 20678 genes. This was necessary to produce a set of genes consistent with the one used in differential expression. The resulting gene expression levels were then normalized using log2(RPKM+1). We then estimated the co-expression network and modules using the function blockwiseModule with the following parameters: maxBlockSize=22000; corType=bicorr; power=12; networkType=signed; deepSplit=2; minModuleSize=40; minCoreKME=0.8; minKMEtoStay=0.5. The analysis produced a network of 23 modules, corresponding to about 17298 genes, plus the grey module of unassigned genes. We verified that the modules identified were not mere artifacts of the clustering procedure by permutation analysis. We assumed that the mean topological overlap of a network module has to be greater than the mean topological overlap a random set of genes in order for the module genes to be considered as co-expressed beyond chance. Therefore, for each module we estimated the average topological overlap[13], then randomly selected a number of genes equal to the number of module members and estimated the corresponding mean topological overlap. This operation was repeated 100 times. The p-value for the analysis was estimated by dividing the number of times the mean topological overlap of the random set of genes was greater than the mean topological overlap of the network module by the number of permutations (N=100). Finally, the p-values for all the modules were FDR corrected for multiple comparisons, and a cut-off of 0.05 was considered for significance.

Functional enrichment analysis. DAVID[56] and MSigDB v4.0[57] were used to test differentially expressed genes and WGCNA modules for overrepresentation in Gene Ontologies and Canonical Pathways.

BrainSpan Classification Analysis

Correlation between transcriptomes (by RNA-Seq) of neuronal precursor cells and neurons differentiated from iPSC, comprising a total of 45 samples at day 0, 11 and 31 of differentiation (and corresponding to a total age of 11, 22, 42 days in vitro if counted from the iPSC stage) within the BRAINSPAN database (n=524 brain samples across all ages) was used to classify iPSC-derived neurons as corresponding to a particular brain region and developmental age. We matched the gene IDs within our set with the ones for which expression levels were available within then BrainSpan data set. We filtered out any gene with a standard deviation less than 1, across the entire BrainSan data set. We computed the Spearman correlation matrix between the log transformed (i.e. RPKMlog = log2 (RPKM + 1)) expression levels of our samples and each of the BrainSpan sampes. We derived the 95% confidence interval for each correlation

coefficient by Fisher transform, identified the maximum correlation coefficient for each of our samples (with each BrainSpan samples) and any other correlation coefficient with 95% confidence interval overlapping with the maximum correlation coefficient 95% confidence interval. This approach implies that each of our samples could potentially overlap with more than one sample within the BrainSpan data set. Given the uneven representation of the regional classes within BrainSpan, we required each regional class to be represented by at least 25 samples, independent of its developmental age, resulting in a final BrainSpan subset with region class relatively evenly represented by a range of 26 to 36 samples. Class equalization was not possible for age classification, since our previous findings[10] showed that our samples likely represent the 8-9pcw developmental period and such classes are amongst the least represented. Neverthless, the analysis results show an overrepresentation of such developmental period, in spite of the bias, suggesting that unevenness is not affecting our age classification.

Genomic analyses

Using BWA 0.5.9-r16[58] aligner with options '-t 4 -q 15' we have aligned genomic sequence reads to the human reference genome GRCh37. Aligned reads were paired, mapped and sorted by BWA invoked with the following options '-a 1000 -n 1 -N 1'. As a result, for each sequenced sample we obtained a file with mapped reads in BAM format. To predict variants for each individual in the study, we combined sequencing data for the corresponding fibroblast sample and their iPSC samples. CNVs were predicted with the CNVnator[59] with 400 bp bins, while SNVs were predicted with GATK v2.6[60]. The following analyses were applied to call *de novo* variants.

***De novo* CNVs discovery and validation**

We discovered CNVs in children and genotyped them in parents. In a normal cell, CN should be a whole number (e.g., 0, 1, 2, etc.), however, estimate of copy number is a real number that is typically close to the one of the whole numbers. For diploid chromosomes, we call *de novo* deletion/duplication candidate if i) its estimated copy-number in proband was lower/higher than 1.5/2.5; ii) its estimated copy-number was higher/low than 1.5/2.5 in parents and iii) copy-number estimates in each parent compare to child are different by 0.5. For haploid chromosomes, the first two cutoffs were lowered by 1. We predicted from 1 to 12 *de novo* CNVs candidates per proband. The candidates were validated with quantitative PCR, by comparing copy-number of the region of interest in proband relative to each parent. Only one candidate *de novo* CNV was validated.

Rare and *de novo* SNVs discovery and validation

We discovered SNVs in all family members. Rare SNPs were defined as those not found in the 38M SNPs catalogued by the 1000 Genomes Project [61] in 1,092 healthy individuals. In probands, we selected *de novo* SNVs candidates as those not found in parents. For the candidates, we then looked for an alternative (i.e., non-reference) allele support in parents by examining aligned reads spanning the candidate sites. This was done with 'samtools mpileup' command. We then called *de novo* SNVs as those sites having i) at least 70% of average coverage in proband; ii) allele fraction between 35% and 65% in proband; iii) no more than one read supporting a third allele; iv) no support for the alternative allele in either parents. We first tested it on an external data set with known *de novo* SNVs[62]. We estimated that our pipeline has ~73% sensitivity for *de novo* discovery, while generating additional 20% of *de novo* calls that could be false positives. We predicted from 146 to 287 *de novo* SNVs per proband. Due to us combining sequencing data for samples from fibroblast and iPSC lines, some of the predicted *de*

novo SNVs could have been the mutations acquired during iPSC culturing or somatic mosaic variants revealed through clonal iPSC lines[11]. PCR validation of 20 random putative *de novo* SNVs revealed that 20% of calls were true *de novo* SNVs, 15% were variants observed only in iPSC lines, 30% could have been true *de novo* or somatic mosaic variants in fibroblasts, and 15% were false positives. For another 20% of sites we couldn't design suitable primers or PCR reaction yielded no clear bands.

Clinical scores correlation analysis.
The head circumference measurements were standardized using $HC\_Z = (HC\text{-}HC_{mean})/HC_{std}$, where HC_{mean} and HC_{std} are the mean and standard deviation (linearly interpolated for age) from the Roche Table of Smoothed Head Circumference Values of Boys[43]. All of the scores were then log transformed, and their Pearson correlation with the WGCNA modules' eigenvalues, as well as FOXG1 expression levels, were estimated using the R function cor.test.

Additional References

41 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. 4th edn, (1994).
42 Kuczmarski, R. J. *et al.* CDC growth charts: United States. *Advance data*, 1-27 (2000).
43 Roche, A. F., Mukherjee, D., Guo, S. M. & Moore, W. M. Head circumference reference data: birth to 18 years. *Pediatrics* **79**, 706-712 (1987).
44 Lord, C., Rutter, M., DiLavore, P. C. & Risi, S. *Autism Diagnostic Observation Schedule-Generic*. (Western Psychological Services, 2000).
45 Rutter, M., Le Couter, A. & Lord, C. *ADI-R: Autism Diagnostic Interview-Revised*. (Western Psychological Services, 2003).
46 Mullen, E. *Mullen Scales of Early Learning. AGS Edition*. (American Guidance Serivce, Inc., 1995).
47 Sparrow, S. S., Balla, D. A. & Cicchetti, D. V. *Vineland Adaptive Behavior Scales, Second Edition (Vineland-II)*. (American Guidance Service, 2005).
48 Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nat Methods* **8**, 409-412 (2011).
49 Ippolito, D. M. & Eroglu, C. Quantifying synapses: an immunocytochemistry-based assay to quantify synapse number. *J Vis Exp* (2010).
50 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
51 Harrow, J. *et al.* GENCODE: producing a reference annotation for ENCODE. *Genome Biol* **7 Suppl 1**, S4 1-9 (2006).
52 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
53 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
54 Habegger, L. *et al.* RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. *Bioinformatics* **27**, 281-283 (2011).
55 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

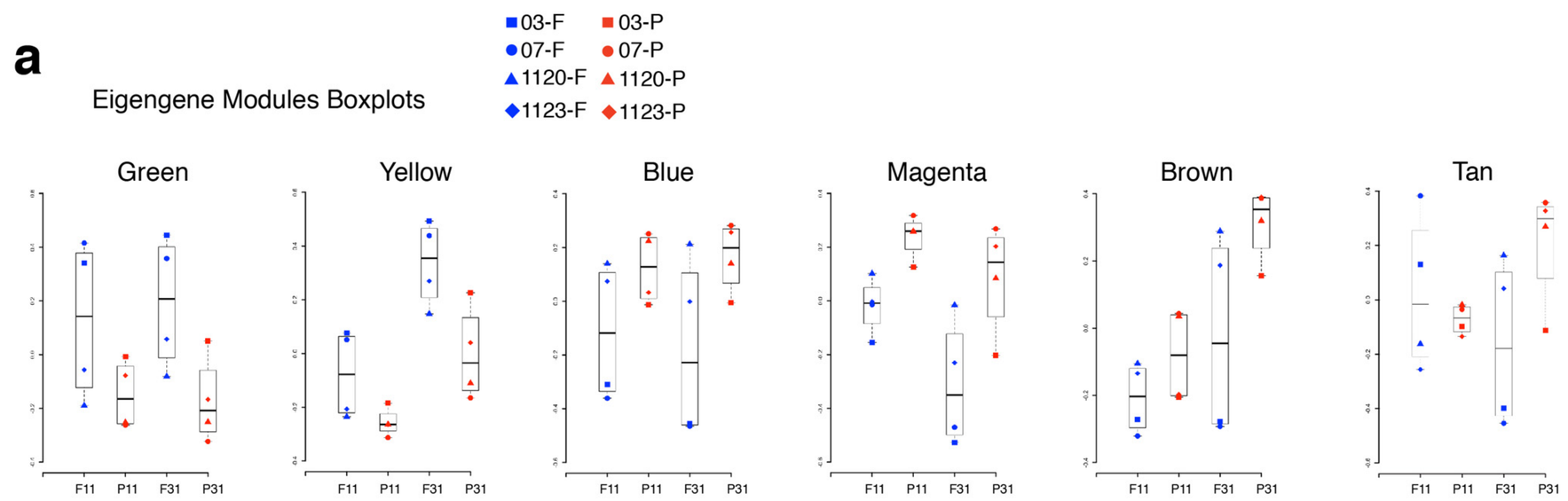
56 Dennis, G., Jr. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**, P3 (2003).
57 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
58 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
59 Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: An approach to discover, genotype and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* (2011).
60 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).
61 Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
62 Michaelson, J. J. *et al.* Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* **151**, 1431-1442 (2012).

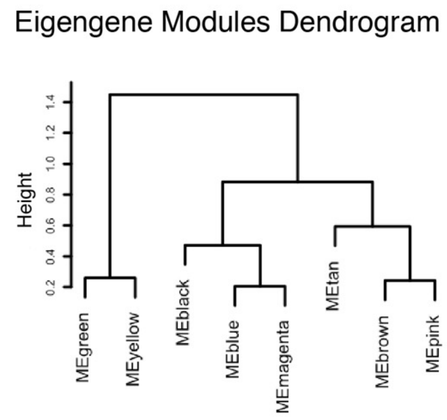
Table 1. Functional enrichment analysis according to DAVID for gene co-expression modules differentially expressed in ASD and controls

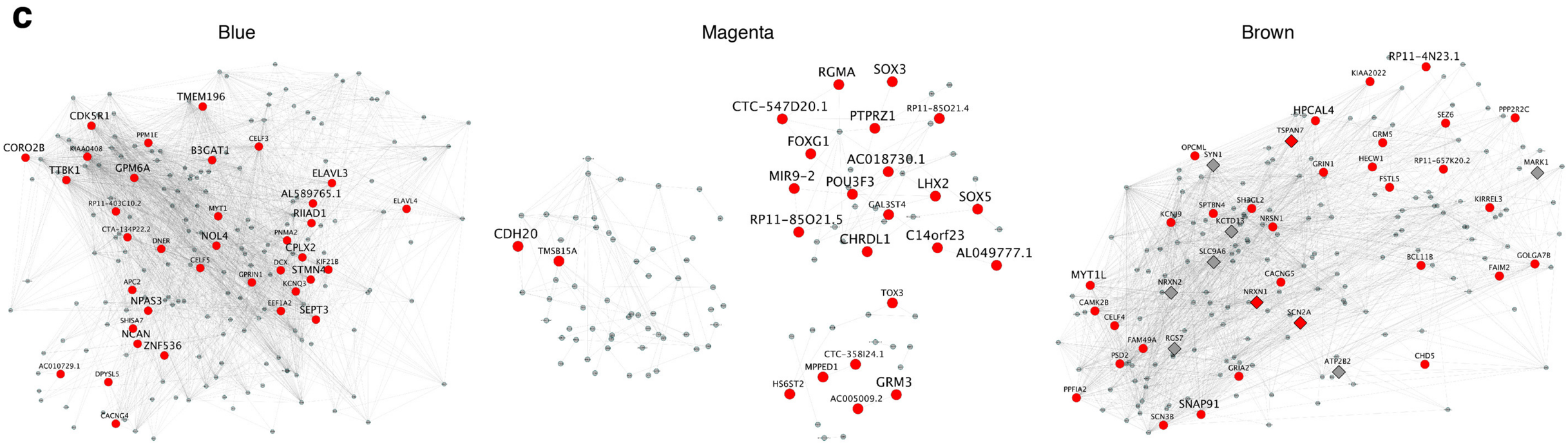
Module	GO term	FDR q-value
green	vasculature development	5.31E-07
	blood vessel development	7.71E-07
	angiogenesis	4.29E-06
	blood vessel morphogenesis	9.47E-06
yellow	lipid biosynthetic process	6.05E-07
	endoplasmic reticulum	9.62E-07
	cell fraction	1.87E-06
	extracellular region part	1.93E-06
blue	neuron differentiation	1.61E-11
	neuron projection	6.21E-11
	axonogenesis	1.53E-10
	cell morphogenesis involved in neuron differentiation	1.97E-10
magenta	regulation of transcription	1.22E-04
	DNA binding	2.30E-04
	transition metal ion binding	7.95E-04
	transcription	2.39E-03
brown	synaptic transmission	5.08E-14
	synapse part	4.25E-14
	ion channel activity	7.13E-14
	ion transport	2.00E-13
tan	gated channel activity	5.19E-02
	substrate specific channel activity	1.77E-01
	channel activity	2.85E-01
	passive transmembrane transporter activity	2.95E-01

Represented are the 4 top scoring GO terms from (full analysis is in Supplementary Table 6). Module: module ID; Gene Set Name: DAVID term ID; FDR q-value: FDR corrected p-value.

Figure 1


a
Eigengene Modules Boxplots
03-F
03-P
07-F
07-P
1120-F
1120-P
1123-F
1123-P
Green
Yellow
Blue
Magenta
Brown
Tan
F11
P11
F31
P31


b
Eigengene Modules Dendrogram
Height
MEgreen
MEyellow
MEblack
MEblue
MEmagenta
MEtan
MEbrown
MEpink


c
Blue
TMEM196
CDK5R1
PPM1E
CELF3
CORO2B
KIAA0408
B3GAT1
TTBK1
GPM6A
ELAVL3
AL589765.1
RP11-403C10.2
MYT1
RIIAD1
ELAVL4
CTA-134P22.2
DNER
NOL4
PNMA2
CPLX2
CELF5
DCX
KIF21B
STMN4
GPRIN1
KCNQ3
APC2
NPAS3
EEF1A2
SEPT3
SHISA7
NCAN
ZNF536
AC010729.1
DPYSL5
CACNG4
CDH20
TMSB15A
Magenta
RGMA
SOX3
CTC-547D20.1
PTPRZ1
RP11-85O21.4
FOXG1
AC018730.1
MIR9-2
POU3F3
LHX2
SOX5
GAL3ST4
RP11-85O21.5
CHRDL1
C14orf23
AL049777.1
TOX3
CTC-358I24.1
MPPED1
HS6ST2
GRM3
AC005009.2
Brown
RP11-4N23.1
KIAA2022
HPCAL4
PPP2R2C
SEZ6
TSPAN7
GRM5
OPCML
SYN1
GRIN1
HECW1
RP11-657K20.2
MARK1
FSTL5
SPTBN4
SH3GL2
KIRREL3
KCNJ9
KCTD13
NRSN1
SLC9A6
BCL11B
GOLGA7B
FAIM2
MYT1L
NRXN2
CACNG5
NRXN1
CAMK2B
SCN2A
CELF4
FAM49A
RGS7
PSD2
GRIA2
ATP2B2
CHD5
PPFIA2
SNAP91
SCN3B

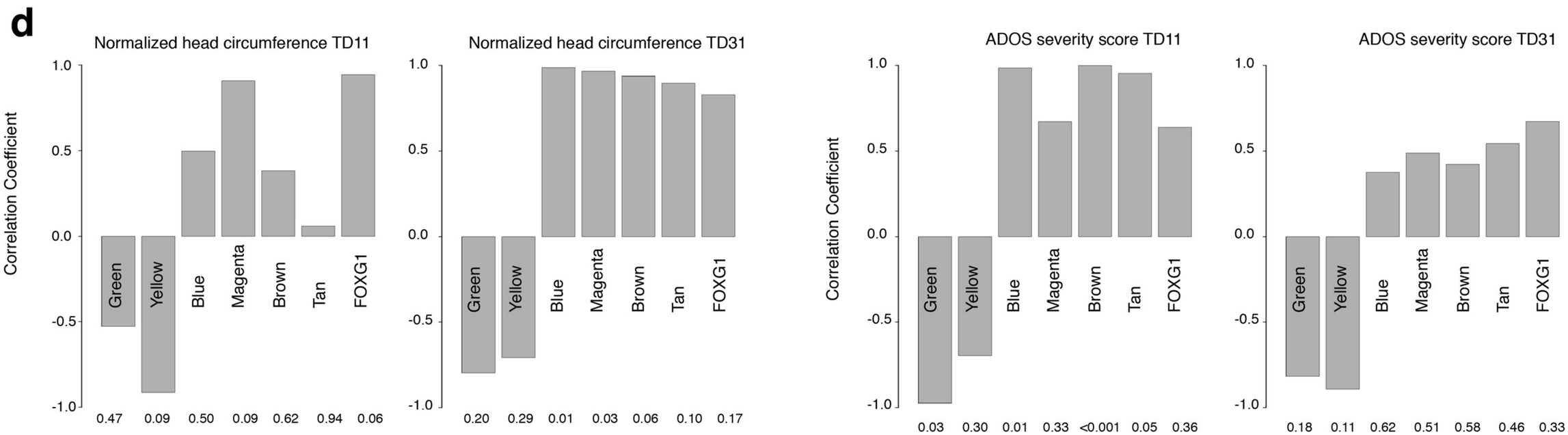

d
Normalized head circumference TD11
Normalized head circumference TD31
ADOS severity score TD11
ADOS severity score TD31
Correlation Coefficient
1.0
0.5
0.0
-0.5
-1.0
Green
Yellow
Blue
Magenta
Brown
Tan
FOXG1
0.47 0.09 0.50 0.09 0.62 0.94 0.06
0.20 0.29 0.01 0.03 0.06 0.10 0.17
0.03 0.30 0.01 0.33 <0.001 0.05 0.36
0.18 0.11 0.62 0.51 0.58 0.46 0.33

Figure 2

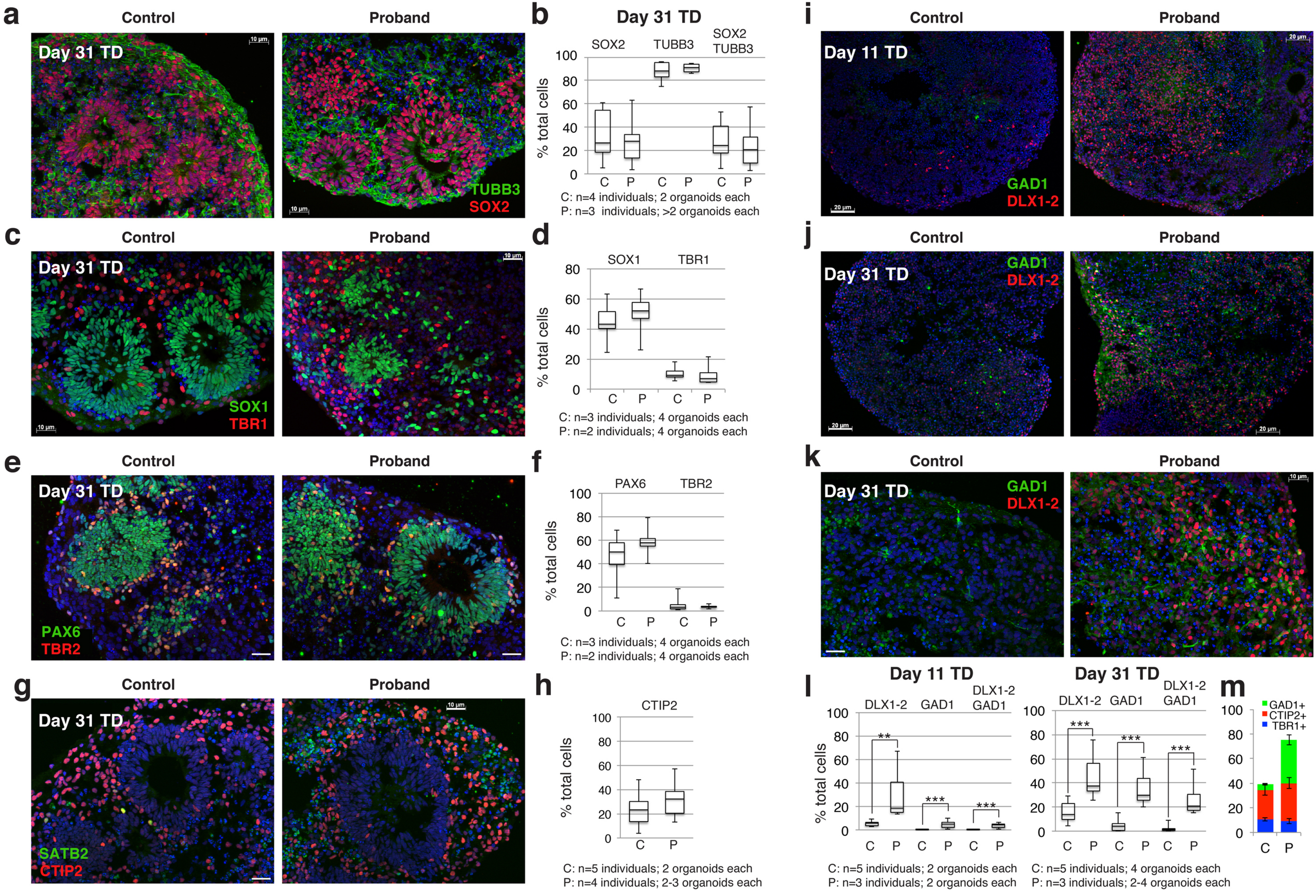
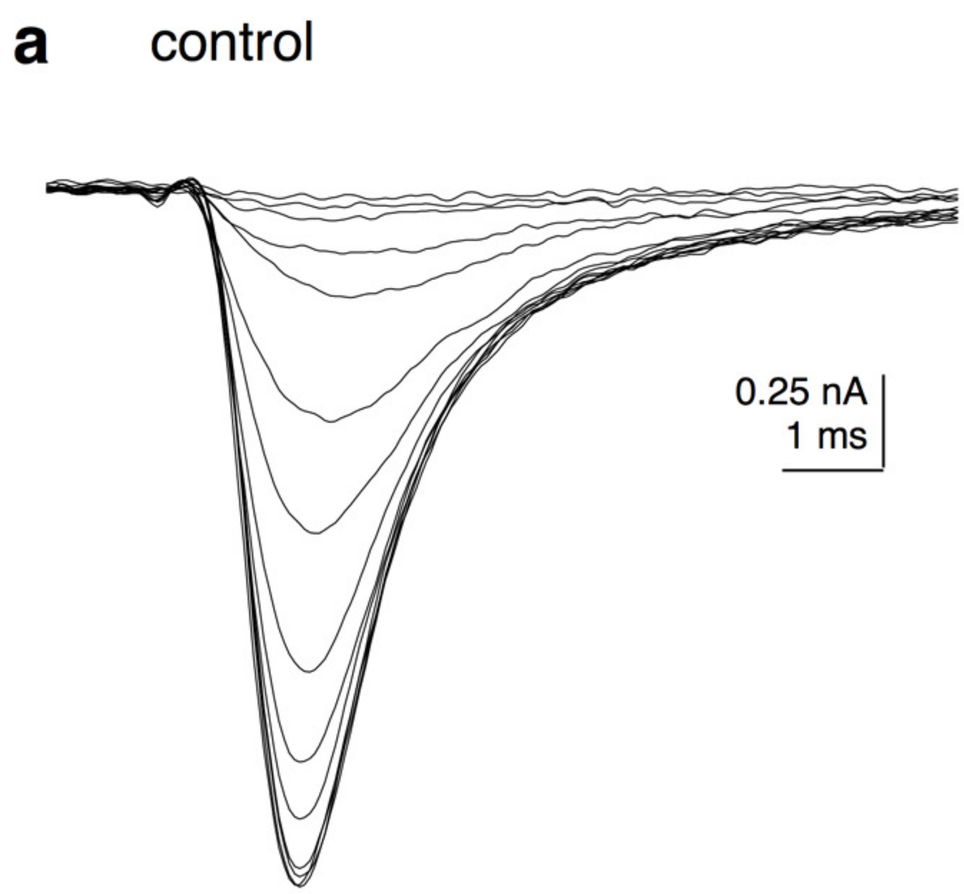
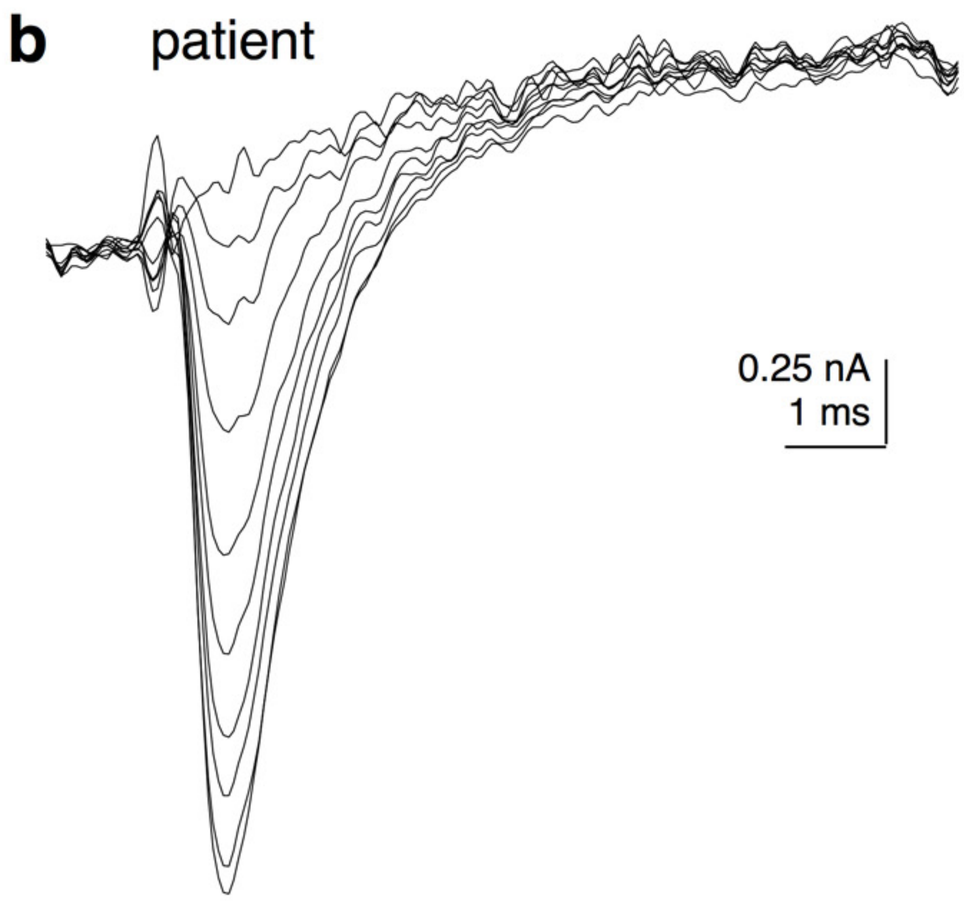
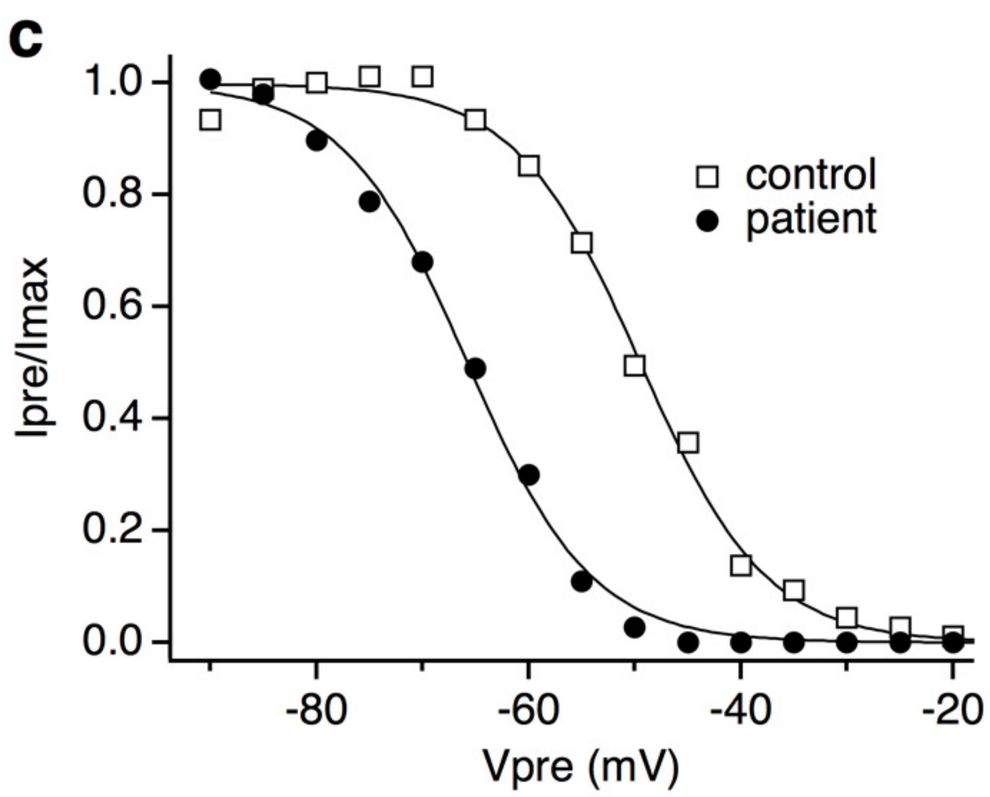
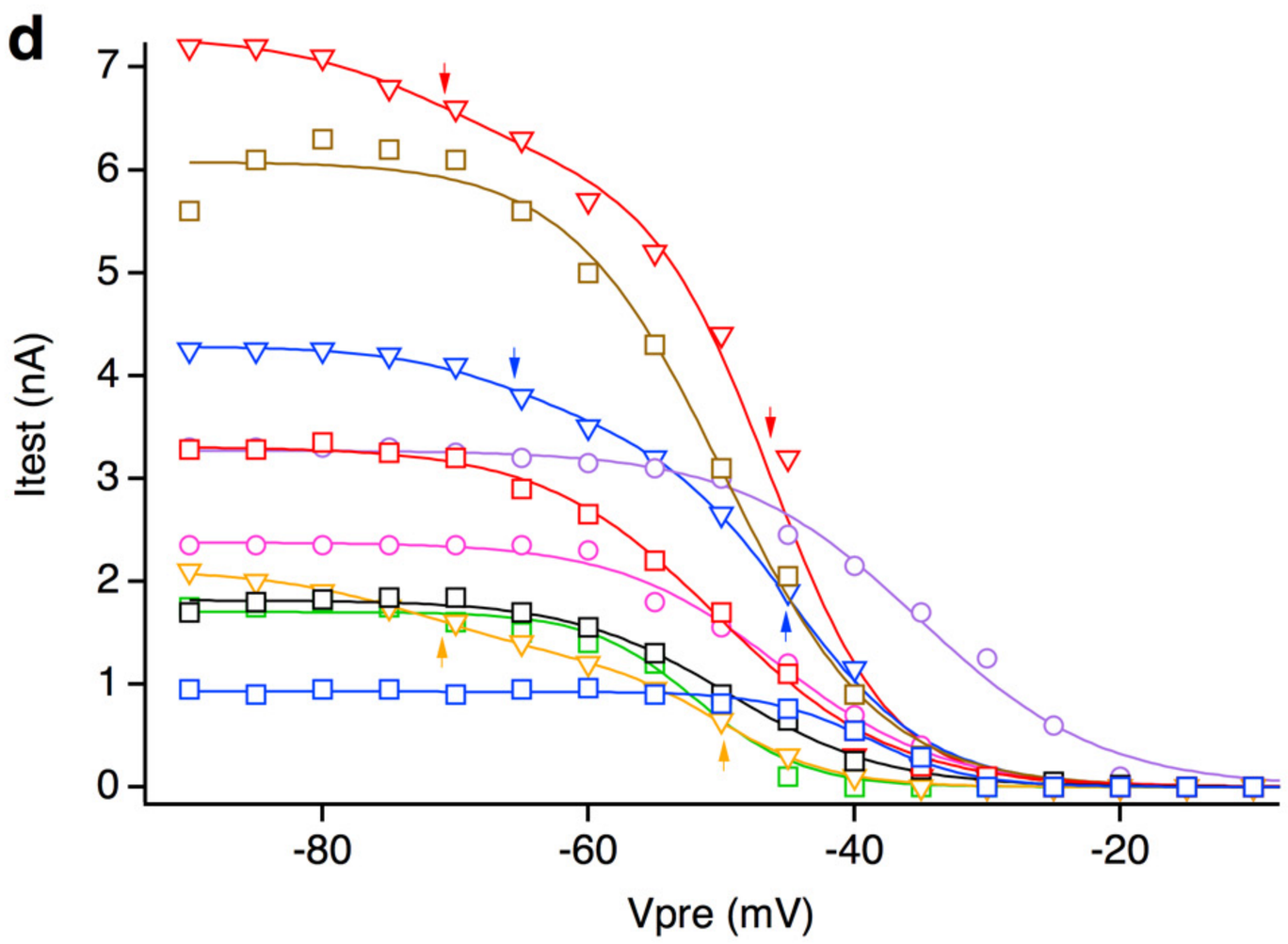

a
Control
Proband
Day 31 TD
10 µm
TUBB3
SOX2
b
Day 31 TD
SOX2
TUBB3
SOX2 TUBB3
% total cells
100
80
60
40
20
0
C P C P C P
C: n=4 individuals; 2 organoids each
P: n=3 individuals; >2 organoids each
i
Control
Proband
Day 11 TD
GAD1
DLX1-2
20 µm
c
Control
Proband
Day 31 TD
SOX1
TBR1
d
SOX1
TBR1
% total cells
C: n=3 individuals; 4 organoids each
P: n=2 individuals; 4 organoids each
j
Control
Proband
Day 31 TD
GAD1
DLX1-2
e
Control
Proband
Day 31 TD
PAX6
TBR2
f
PAX6
TBR2
% total cells
C: n=3 individuals; 4 organoids each
P: n=2 individuals; 4 organoids each
k
Control
Proband
Day 31 TD
GAD1
DLX1-2
g
Control
Proband
Day 31 TD
SATB2
CTIP2
h
CTIP2
% total cells
C: n=5 individuals; 2 organoids each
P: n=4 individuals; 2-3 organoids each
l
Day 11 TD
Day 31 TD
DLX1-2
GAD1
DLX1-2 GAD1
**
***
% total cells
C: n=5 individuals; 2 organoids each
P: n=3 individuals; 2 organoids each
C: n=5 individuals; 4 organoids each
P: n=3 individuals; 2-4 organoids each
m
GAD1+
CTIP2+
TBR1+

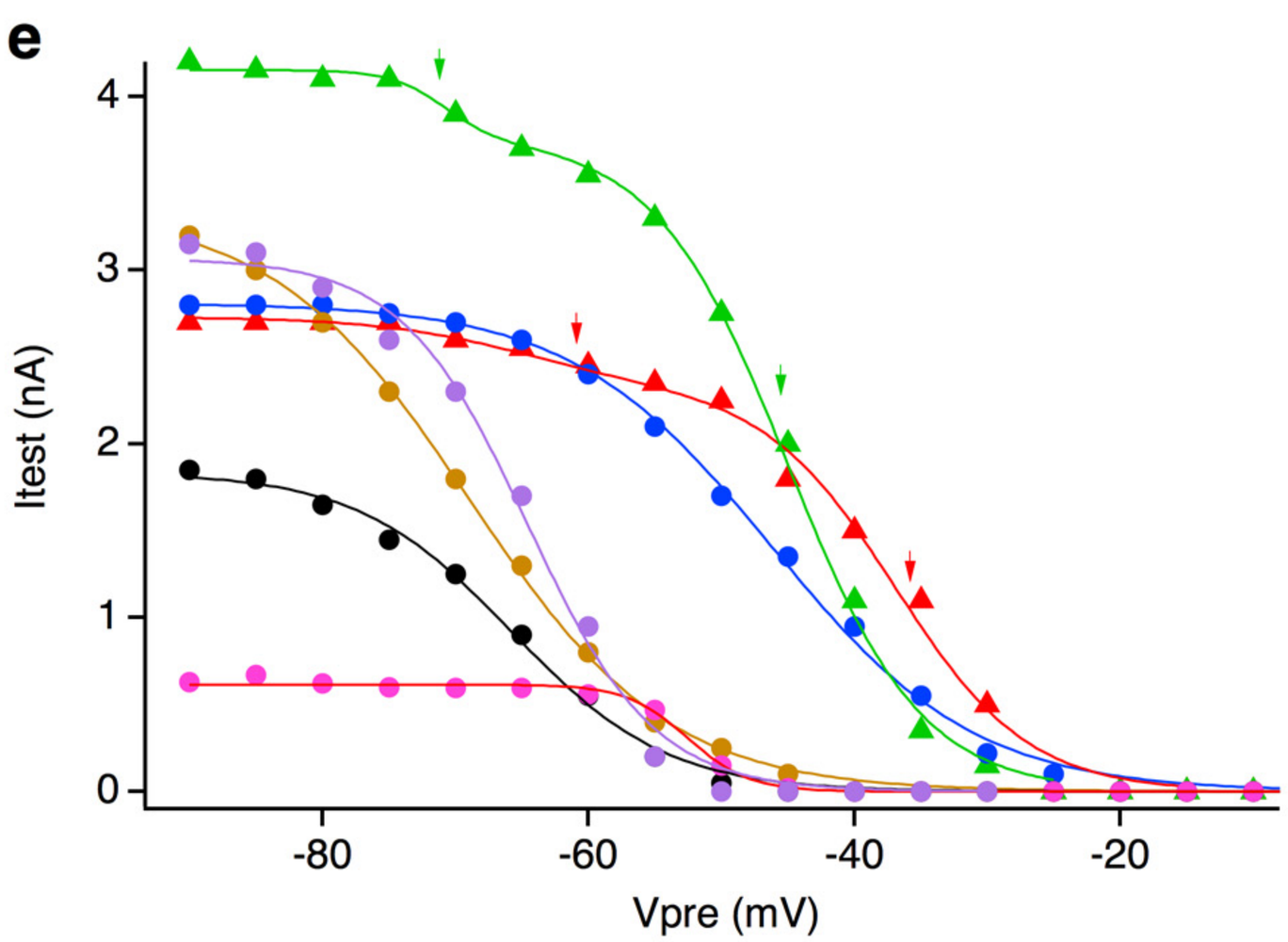
Figure 3


a
control
0.25 nA
1 ms


b
patient
0.25 nA
1 ms


c
1.0
0.8
0.6
0.4
0.2
0.0
Ipre/Imax
control
patient
-80
-60
-40
-20
Vpre (mV)


d
7
6
5
4
3
2
1
0
Itest (nA)
-80
-60
-40
-20
Vpre (mV)


e
4
3
2
1
0
Itest (nA)
-80
-60
-40
-20
Vpre (mV)

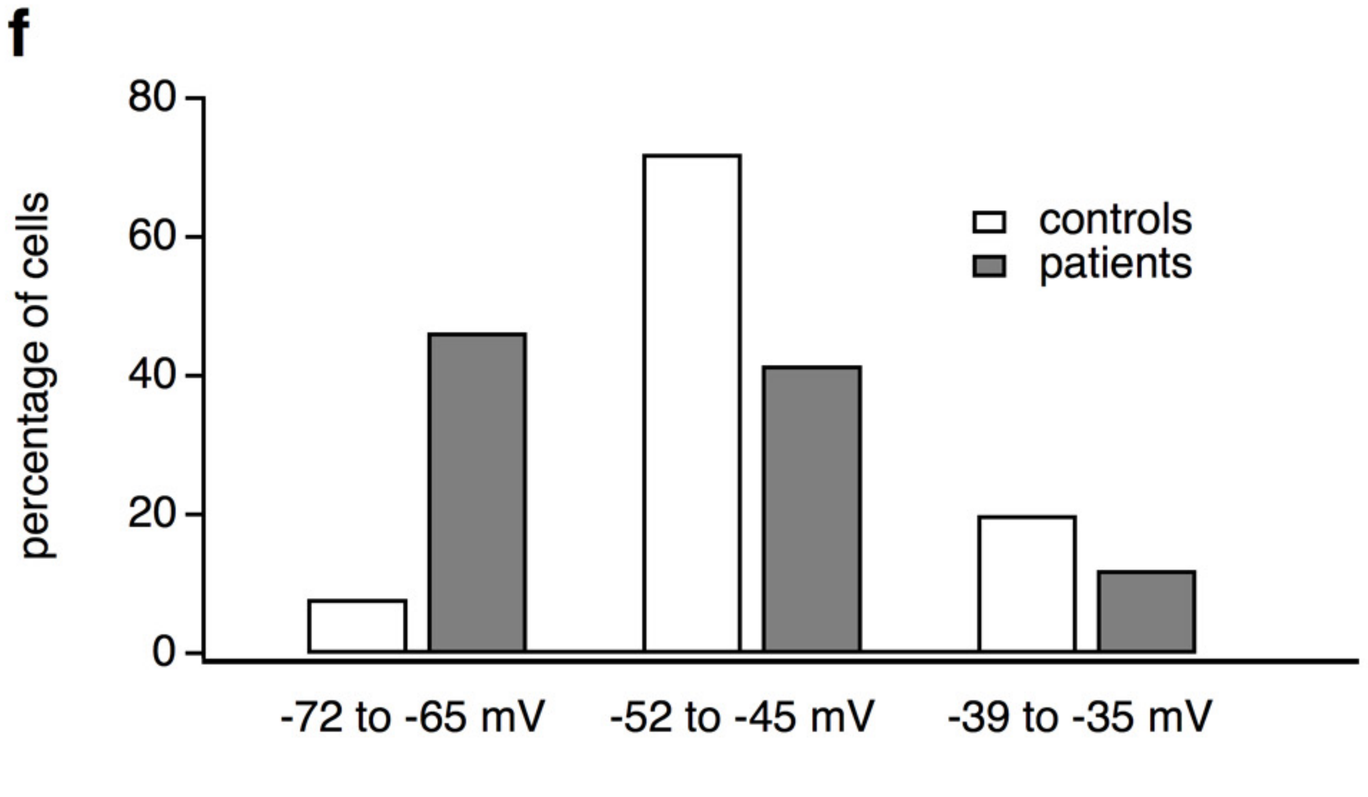
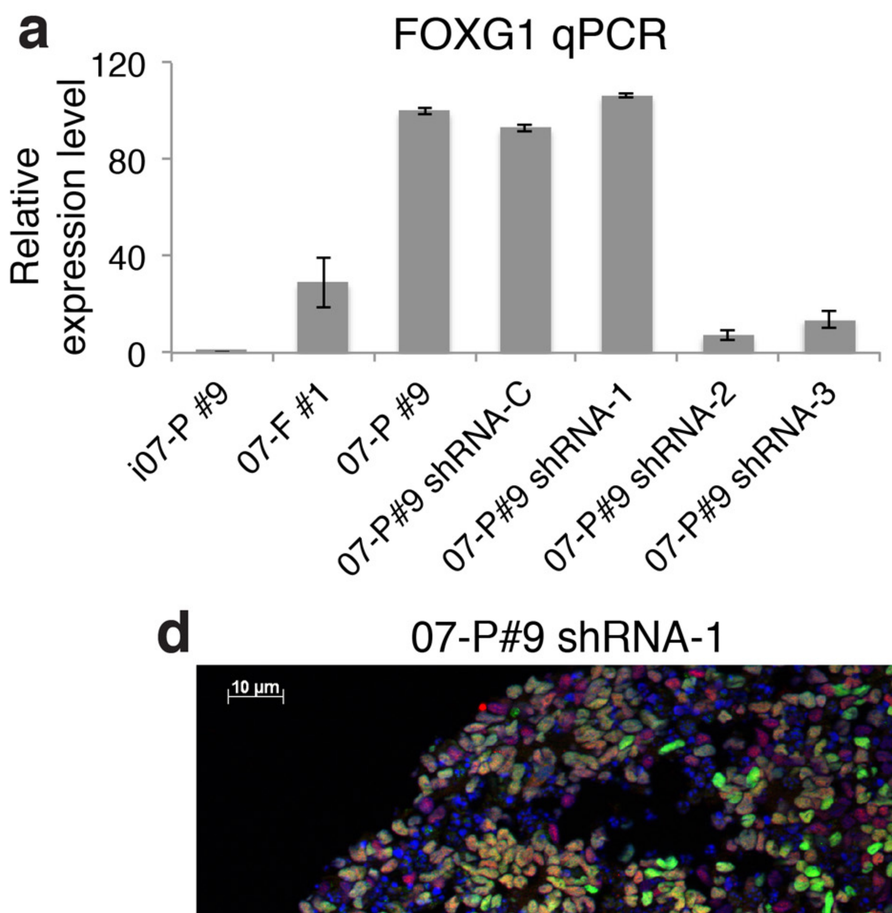
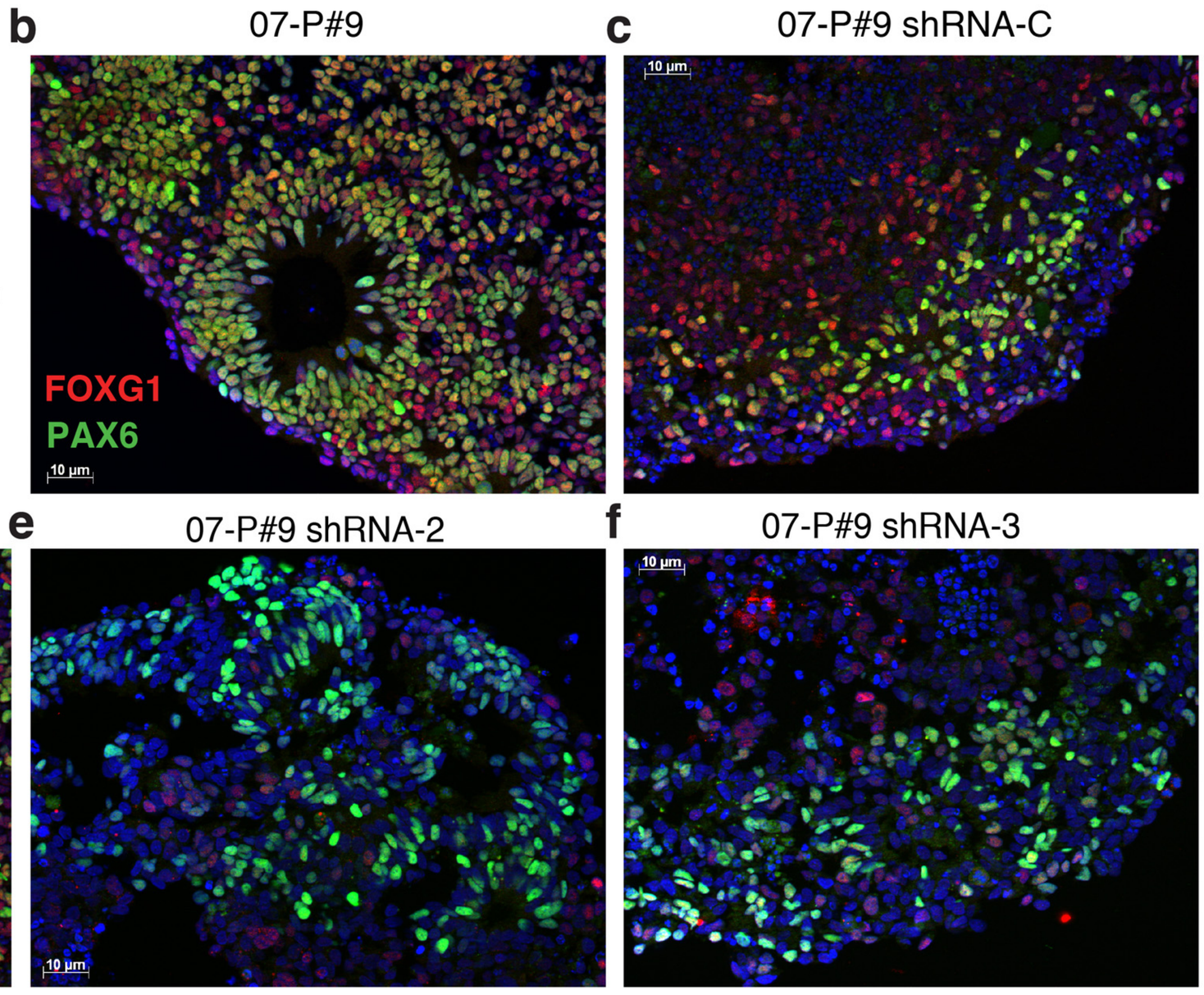
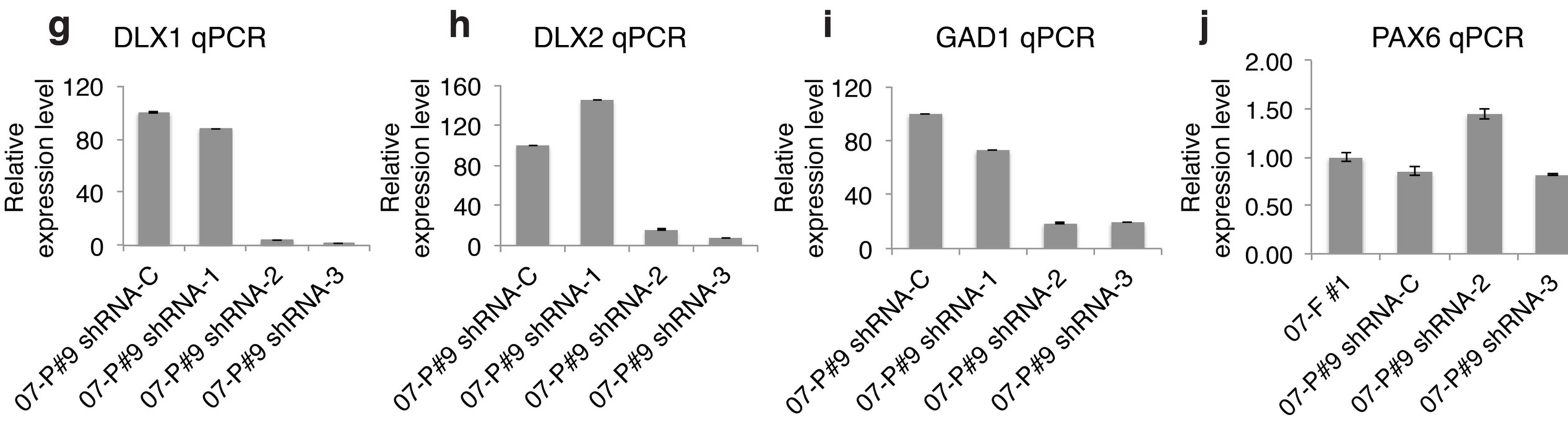
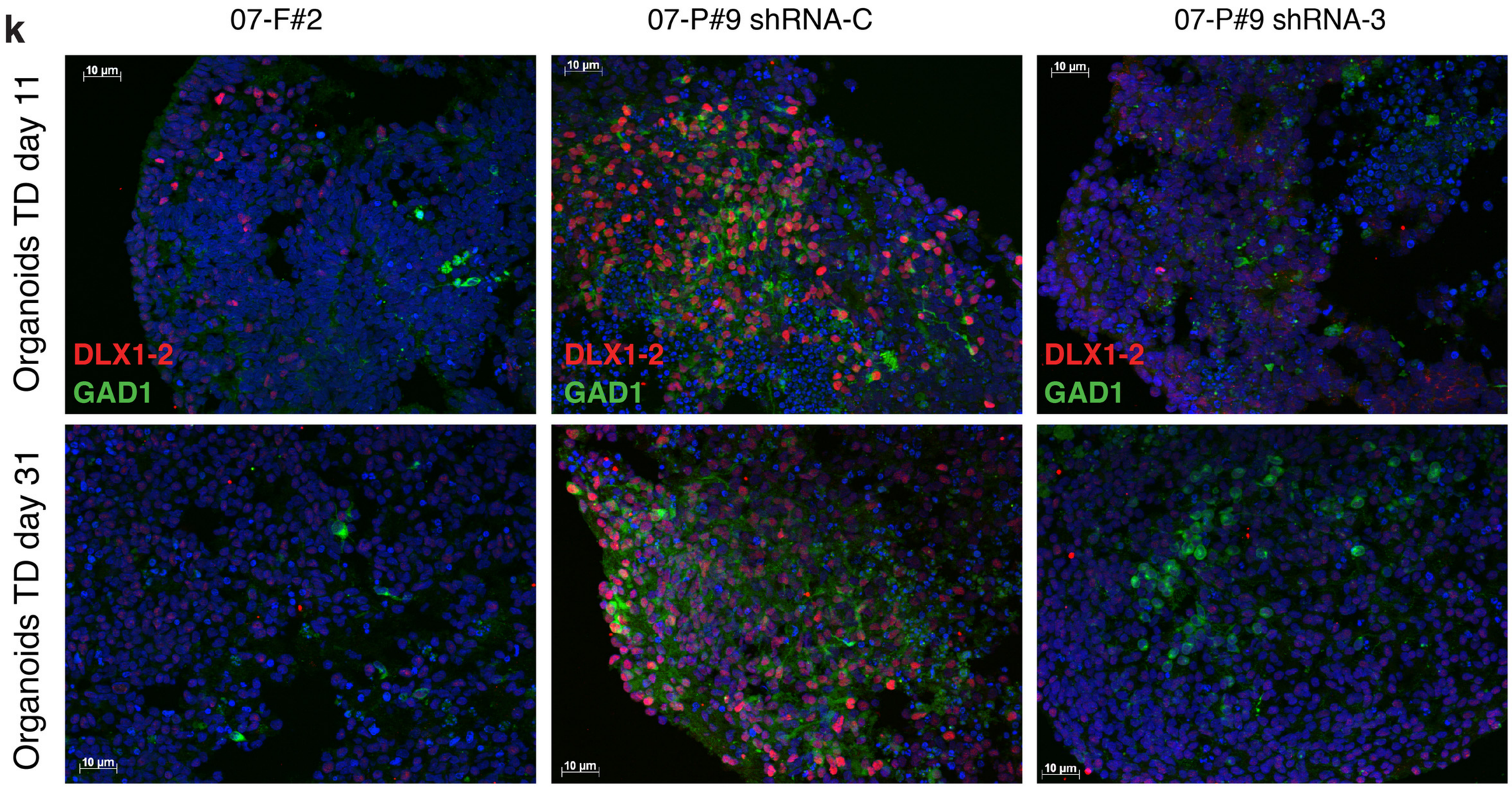

f
80
60
40
20
0
percentage of cells
controls
patients
-72 to -65 mV
-52 to -45 mV
-39 to -35 mV

Figure 4


a
FOXG1 qPCR
Relative expression level
120
80
40
0
i07-P #9
07-F #1
07-P #9
07-P#9 shRNA-C
07-P#9 shRNA-1
07-P#9 shRNA-2
07-P#9 shRNA-3


b
07-P#9
c
07-P#9 shRNA-C
d
07-P#9 shRNA-1
e
07-P#9 shRNA-2
f
07-P#9 shRNA-3
FOXG1
PAX6
10 µm
Organoids TD day 11


g
DLX1 qPCR
h
DLX2 qPCR
i
GAD1 qPCR
j
PAX6 qPCR
Relative expression level
07-P#9 shRNA-C
07-P#9 shRNA-1
07-P#9 shRNA-2
07-P#9 shRNA-3
07-F #1


k
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
Organoids TD day 11
Organoids TD day 31
DLX1-2
GAD1
10 µm

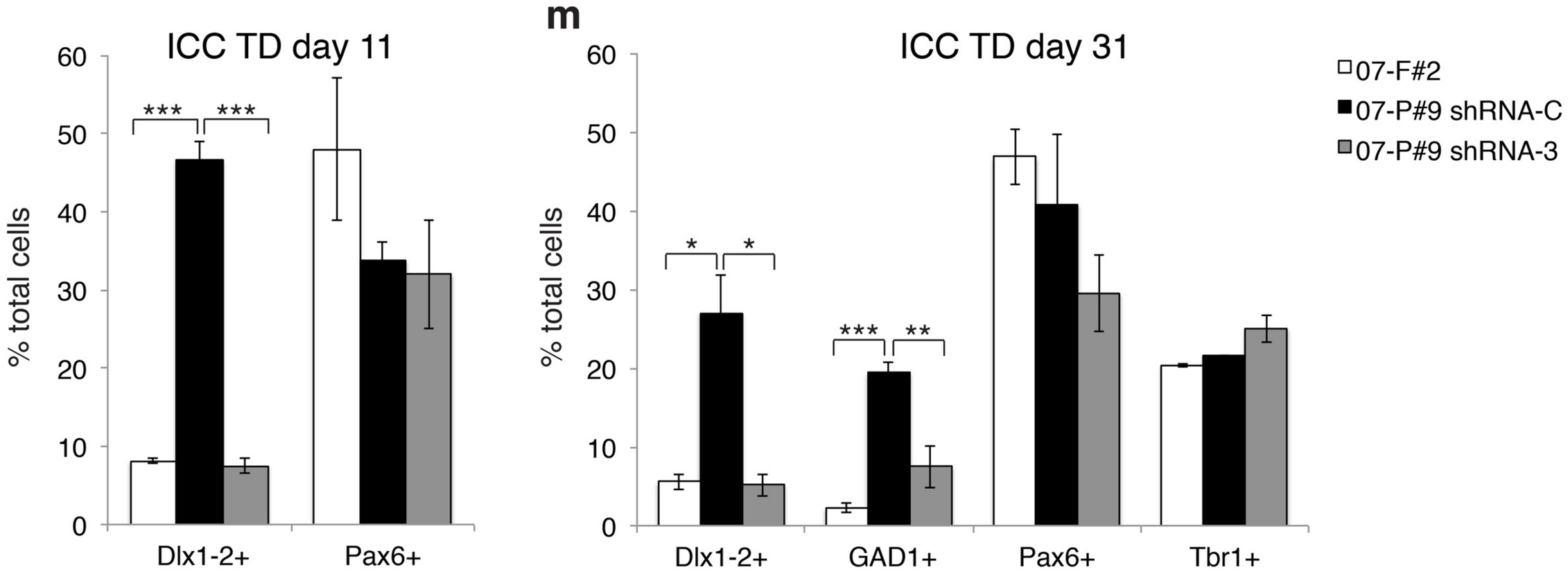
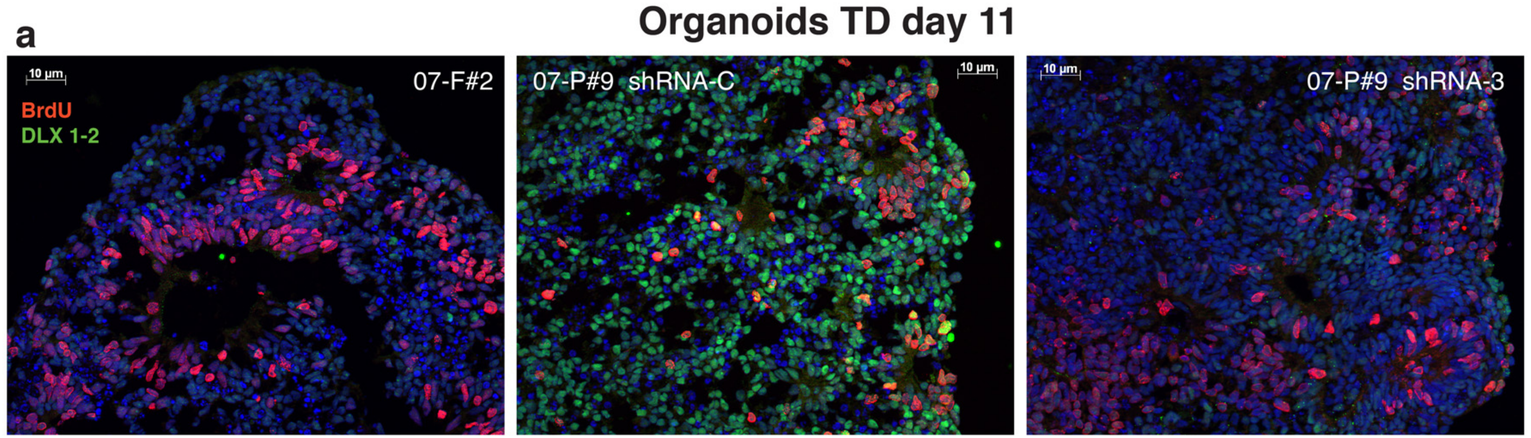
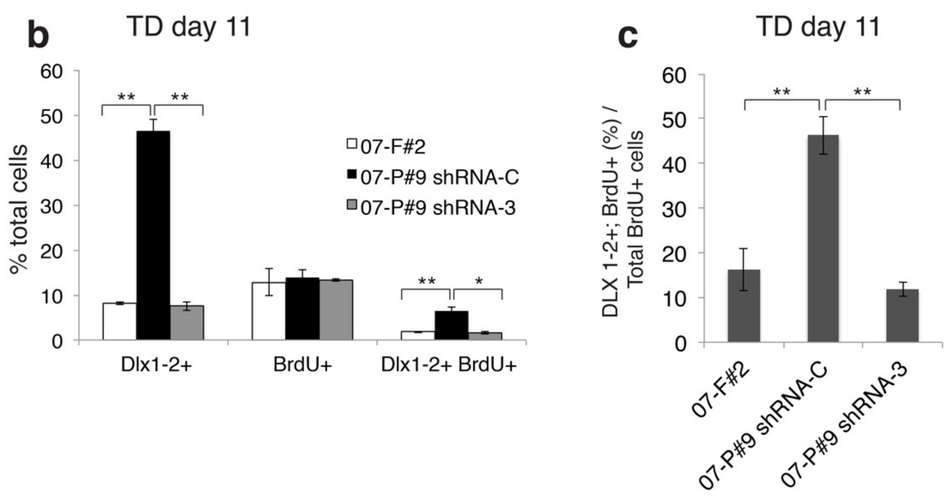
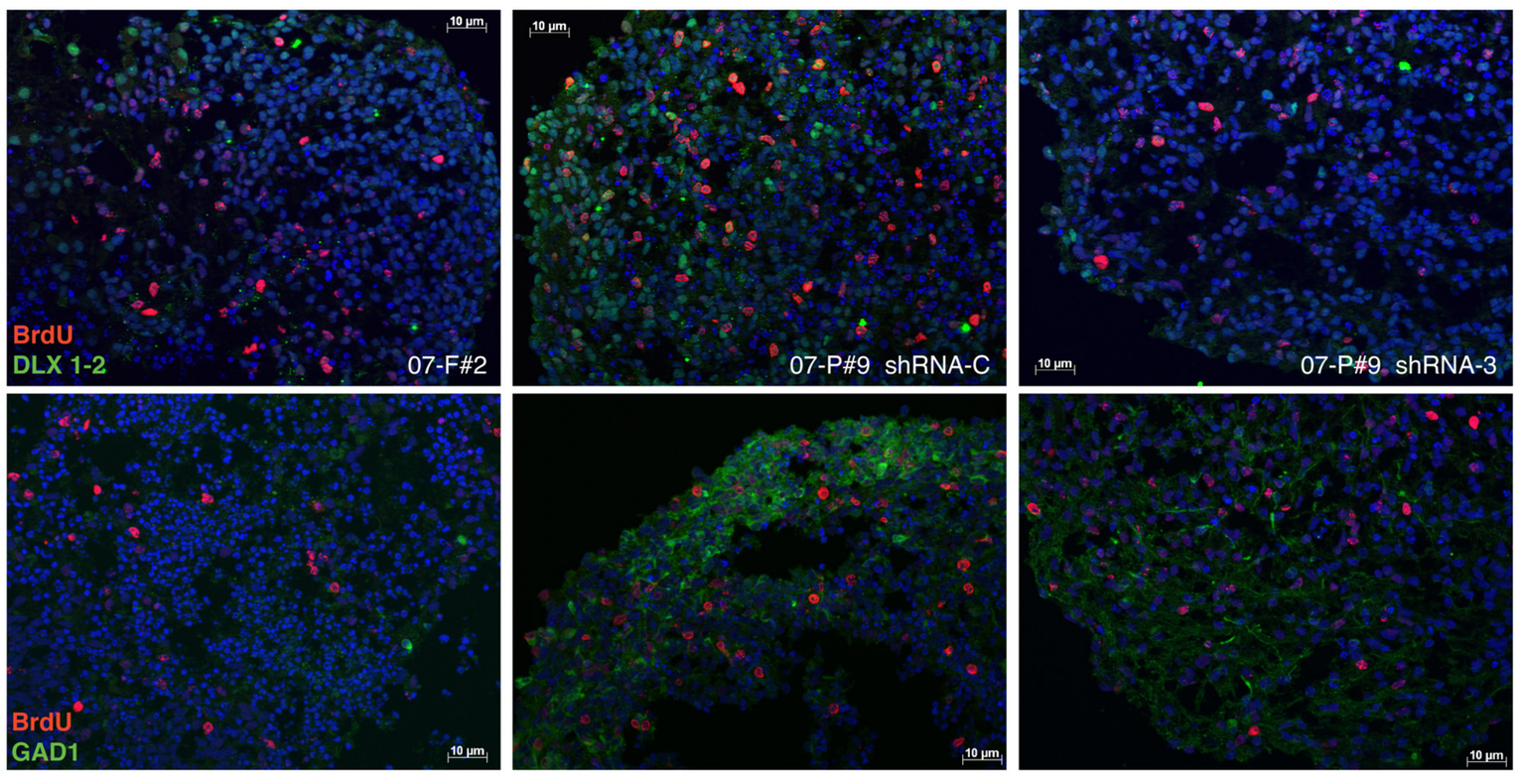
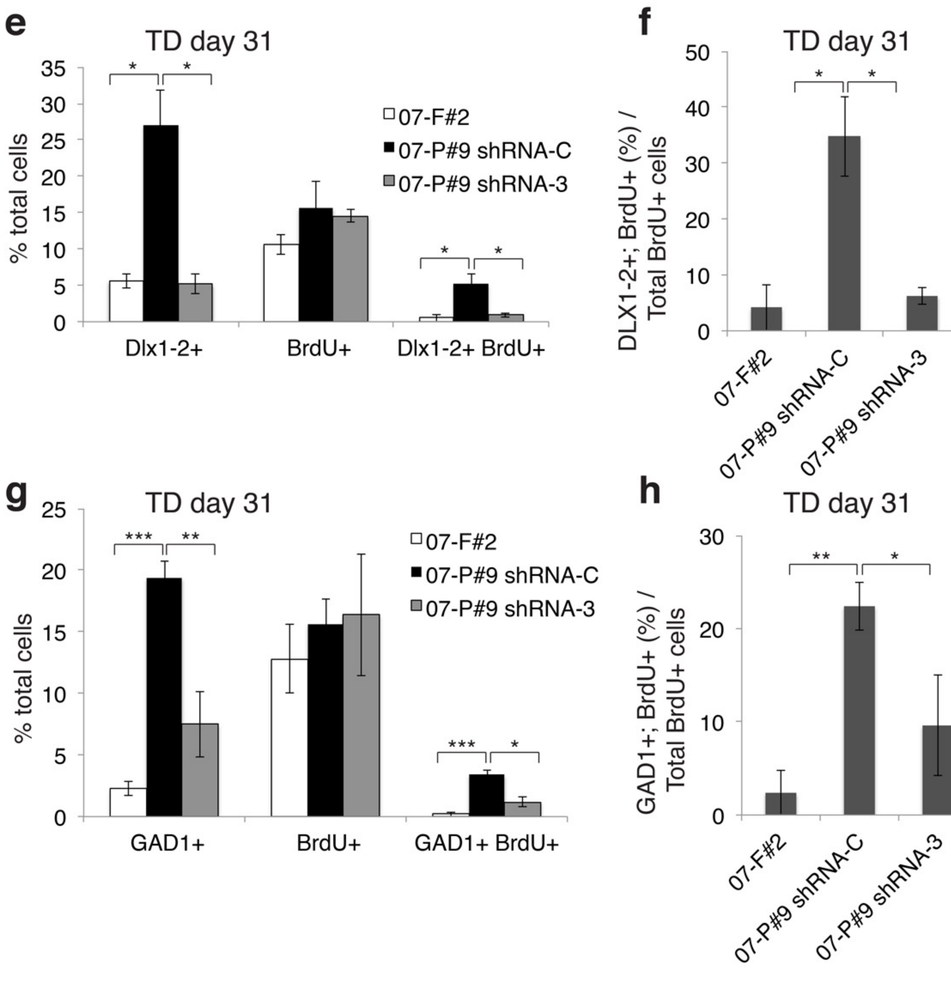

l
ICC TD day 11
% total cells
***
***
Dlx1-2+
Pax6+
m
ICC TD day 31
*
*
***
**
Dlx1-2+
GAD1+
Pax6+
Tbr1+
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3

Figure 5


a
Organoids TD day 11
10 µm
BrdU
DLX 1-2
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3


b
TD day 11
% total cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
Dlx1-2+
BrdU+
Dlx1-2+ BrdU+
c
TD day 11
DLX 1-2+; BrdU+ (%) / Total BrdU+ cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3


d
Organoids TD day 31
10 µm
BrdU
DLX 1-2
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
BrdU
GAD1


e
TD day 31
% total cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
Dlx1-2+
BrdU+
Dlx1-2+ BrdU+
f
TD day 31
DLX1-2+; BrdU+ (%) / Total BrdU+ cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
g
TD day 31
% total cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3
GAD1+
BrdU+
GAD1+ BrdU+
h
TD day 31
GAD1+; BrdU+ (%) / Total BrdU+ cells
07-F#2
07-P#9 shRNA-C
07-P#9 shRNA-3