DOI:
10.1016/j.biopsych.2020.06.014

Document Version
Peer reviewed version

Atypical neurogenesis in induced pluripotent stem cell (iPSC) from autistic individuals

Dwaipayan Adhya[1,3,*], Vivek Swarup[2,*], Roland Nagy[3], Lucia Dutan[3], Carole Shum[3], Eva P. Valencia-Alarcón[3], Kamila Maria Jozwik[4], Maria Andreina Mendez[5], Jamie Horder[5], Eva Loth[5], Paulina Nowosiad[3], Irene Lee[6], David Skuse[6], Frances A. Flinter[7], Declan Murphy[5], Grainne McAlonan[5], Daniel H. Geschwind[2,9], Jack Price[3,8], Jason Carroll[4], Deepak P. Srivastava[3,8]§†, & Simon Baron-Cohen[1]§

[1]Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, CB2 8AH UK.

[2]Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA.

[3]Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK, SE5 9NU, UK.

[4]Cancer Research UK Cambridge Institute, Cambridge CB2 0RE, UK.

[5]Department of Forensic and Neurodevelopmental Sciences, Sackler Institute for Translational Neurodevelopment, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE5 8AF, UK.

[6]Behavioural and Brain Sciences Unit, Population Policy Practice Programme, Great Ormond Street Institute of Child Health, University College London, London WC1N 1EH, UK.

[7]Department of Clinical Genetics, Guy's & St Thomas' NHS Foundation Trust, London, UK.

[8]MRC Centre for Neurodevelopmental Disorders, King's College London, London, UK.

[9]Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA.

§ Joint senior authors

† Corresponding author: Deepak P. Srivastava, deepak.srivastava@kcl.ac.uk

* Joint first authors

Short title: Atypical neurogenesis in autism iPSC-derived neurons

Keywords:

Autism, neural precursors, neural progenitor cells, cortical differentiation, midbrain differentiation, neurodevelopment, functional genomics.

Abstract

Background: Autism is a heterogenous collection of disorders with a complex molecular underpinning. Evidence from *post-mortem* brain studies have indicated that early prenatal development may be altered in autism. Induced pluripotent stem cells (iPSCs) generated from autistic individuals with macrocephaly also indicate prenatal development as a critical period for this condition. But little is known about early altered cellular events during prenatal stages in autism.

Methods: IPSCs were generated from 9 unrelated autistic individuals without macrocephaly and with heterogeneous genetic backgrounds, and 6 typically developing, control, individuals. IPSCs were differentiated towards either cortical or midbrain fates. Gene expression and high throughput cellular phenotyping was used to characterise iPSCs at different stage of differentiation.

Results: A subset of autism-iPSC cortical neurons were RNA-sequenced to reveal autism-specific signatures similar to *post-mortem* brain studies, indicating a potential common biological mechanism. Autism-iPSCs differentiated towards a cortical fate displayed impairments in the ability to self-form into neural rosettes. In addition, autism-iPSCs demonstrated significant differences in rate of cell type assignment of cortical precursors, and dorsal and ventral forebrain precursors. These cellular phenotypes occurred in the absence of alterations in cell proliferation during cortical differentiation, differing from previous studies. Acquisition of cell fate during midbrain differentiation was not different between control- and autism-iPSCs.

Conclusions: Taken together, our data indicate that autism-iPSCs diverge from control-iPSCs at a cellular level during early stage of neurodevelopment. This suggests that unique developmental differences associated with autism may be established at early prenatal stages.

Introduction

Autism spectrum conditions (henceforth autism) are a genetically heterogeneous spectrum of neurodevelopmental conditions (1-3). Autism is characterised by impairments in social-communicative behaviours as well as repetitive behaviours. Symptoms of autism cannot be detected until twelve to eighteen months of age (4). However, there is debate surrounding the origins of autistic symptoms. It is now well recognised that genetic factors play a key role in the emergence of autism (1, 2). Increasing evidence indicate that perturbation during critical periods of prenatal development may be key for the emergence of autism (5). Consistent with this idea, autism *post-mortem* brain studies have identified dysregulation of putative prenatal gene expression pathways (6). Thus, early prenatal development may be a critical period for the emergence of cellular pathophysiology associated with autism (6).

The use of induced pluripotent stem cell (iPSC) has made it possible to study prenatal cellular behaviour in autism (7-11). IPSC-neurons contain the same genetic information as the individuals from whom they were derived, which influences cellular behaviour. IPSCs generated from autistic individuals, comorbid for macrocephaly, show significant cellular/molecular anomalies during prenatal-equivalent periods (12-14). These iPSCs demonstrated atypical neural differentiation when fated towards a cortical lineage, and an imbalance in excitatory and inhibitory receptor activity (12, 13). Using the same collection of iPSCs, an acceleration in neuronal maturation was found to be dependent on early cortical neural precursor (NPC) development and circumventing this stage did not recapitulate altered neuronal development. Alterations in gene expression network dynamics during early neurodevelopment also accompanied these effects (14). These studies highlight that the cellular and molecular phenotypes associated with autism may start during prenatal brain development. A critical aspect of this studies is that atypical neural differentiation was associated with higher

cell proliferation (12-14). However, as the autistic participants in these studies also had macrocephaly, it is unclear whether the observed abnormal development was in part due to this comorbidity. Moreover, as macrocephaly is present only in a subset of autistic individuals it is not known if atypical development can be generalised to autistic individuals without macrocephaly. Finally, as most studies have predominantly focused on the development of forebrain/cortical neurons, it is yet to be tested whether atypical development can also be observed in NPCs fated towards a different lineage.

In this study, we generated iPSCs from autistic individuals without macrocephaly from 3 independent patient cohorts to capture a wider population of autistic individuals. Initial RNA-sequencing studies indicated that early neurodevelopment may be affected. To further investigate the source of atypical gene expression, we undertook extensive cellular phenotyping experiments. The goal of this study was to understand if there was a fundamental difference between typical and autistic prenatal neurodevelopment, focussing primarily on early neuroectodermal structures and cell types that constitute the developing cerebral cortex.

Materials and Methods

Further information can be found in Supplemental Information.

Induced pluripotent stem cells

Participants were recruited and methods carried out in accordance to the 'Patient iPSCs for Neurodevelopmental Disorders (PiNDs) study' (REC No 13/LO/1218). Informed consent was obtained from all subjects for participation in the PiNDs study. Ethical approval for the PiNDs study was provided by the NHS Research Ethics Committee at the South London and Maudsley (SLaM) NHS R&D Office. Autistic participants were selected based on ADOS, ADI-R scores,

while typical controls were selected from the population if they had no diagnosis of any psychiatric condition. Fifteen iPSC lines (autism: 9; control: 6) were generated from hair keratinocytes as previously described (15, 16). Details on all participants can be found in **Supplementary Results** and **Supplementary Tables S1, S2, S3**. Two clones per iPSCs were used in all experiments; pluripotency of all iPSCs was determined by immunocytochemistry (**Supplementary Table S4 and Supplementary Figure 1**).

Neuronal differentiation

IPSCs were differentiated cortical neurons using a dual SMAD inhibition protocol which recapitulates of key hallmarks of corticogenesis (10, 16). IPSCs were differentiated to midbrain floorplate precursors using established protocols (7, 8).

RNA-sequencing

RNA-sequencing was performed from a subset of our cohort, on 2 clones from each participant (ASDM1, 004ASM, 010ASM, CTRM1, CTRM2, CTRM3), and each clone had 2 technical replicates. Poly(A) containing mRNA was purified and libraries prepared using TruSeq Stranded mRNA kit (Illumina). Unstranded libraries with a mean fragment size of 150bp were constructed, and underwent 50bp single ended sequencing on an Illumina HiSeq 2500 machine. Bioinformatics analysis was performed using C++ and R based programs.

Immunocytochemistry

Differentiated iPSCs were fixed in 4% paraformaldehyde and processed as previously described (16). Briefly, fixed cells were permeabilized in 0.1% Triton-X-100/PBS, and blocked in 4% normal goat serum in PBS. Primary antibodies (**Supplementary Table S5)** were incubated overnight at 4°C. Nuclei were identified by staining with DAPI. High content screening (HCS) was performed using an Opera Phenix High-Content Screening System

(Perkin Elmer). Immunofluorescence was measured from known intracellular location of markers (e.g. nucleus or cytoplasm). Cell type analysis was performed using the Harmony Software (16). For Rate of Cell Type Assignment (deltaCTA), the percent positively stained cells appearing per day was estimated, which was then adjusted to the total number of positive cells appearing per day in one well of a 96-well plate, assuming each well had an average of 10^5 cells.

Statistics

Quantification of cell types was performed using the Harmony High Content Imaging and Analysis Software. Percentage of cells positive for desired marker versus total number of live cells was calculated. Eight independent experimental replicates of 2 clones per individual was used at every stage to account for variability associated with iPSC differentiation. Independent 2-group t-test was used to check significant difference between autism and control using p-value ≤ 0.05. One way ANOVA was performed to investigate in-group variance. All statistical analysis was performed on R statistical software.

Results

Neurodevelopmental gene expression signatures in autism-iPSC-derived neurons

IPSC cells were generated from 9 autistic and 6 typical control individuals from 3 independent cohorts (**Supplementary Results**; **Supplementary Tables S1-3**). To understand if iPSCs derived from individuals diagnosed with autism but without macrocephaly also displayed atypical cortical differentiation with altered cell proliferation as previously reported (12, 13, 17), cells were differentiated towards a cortical fate. We focused on three distinct developmental stages (**Figure 1A**): (1) Day 9: early neural precursor stage, when stem cells

form new precursor cells which self-organise into neural tube-like structures known as neural rosettes with a directional apical-basal arrangement; (2) Day 21: late neural precursor stage, a period during which neural progenitor cells begin forming layers from the apical surface and are primed for differentiation into neurons as they move outwards; and (3) Day 35: immature cortical neurons, a stage at which precursors become post-mitotic and adopt a deep layer neuronal identity (**Figure 1B)**. We initially sought to confirm whether day 35 neurons from autism-iPSCs showed a similar transcriptomic profile as that seen in *post-mortem* brain studies (6, 9, 18). For this analysis, we chose participants with no familial history of autism or known deletions in autism-associated genes to reduce genetic bias. Analysis of differential gene expression (**Figure 1C; Supplementary Table S6**) confirmed distinct transcriptomic profiles of control- and autism-iPSCs, and a high enrichment for genes identified in autism *post-mortem* brain studies, but not schizophrenia or cancer (see **Supplemental Results**). These data suggested that differences between autism and typically controls may appear during prenatal development.

Marked alteration in rosette structures in autism without proliferative differences in precursor pools

Differentiation of iPSCs towards a neuronal fate first results in the generation of neuroepithelium cells, which self-organise into structures known as 'neural rosettes' (10). These structures display apical-basal polarity similar to neural tubes (10, 19). They are thought to play a key role in determining cortical neurogenesis and thus generation of distinct cell fates (10, 19, 20). As our RNASeq data indicated that early neurodevelopment maybe affected in autism, we reasoned that this may be reflected by an alteration in neural rosette formation. We examined rosette formation at day 9 in control- and autism-iPSCs. Control-iPSCs robustly

formed structures identifiable as neural rosettes, with an inner lumens identified by ZO-1 staining. Neural progenitor cells self-organised radially around the inner lumen, typical of cells adopting an apical-basal polarity organisation (**Figure 1D**). Conversely, autism-iPSCs showed significant anomalies in lumen formation and establishment of apical-basal polarity (**Figure 1D**). Using a high content screening (HCS) approach, we assessed rosette structure to identify consistent alterations in rosette morphology between iPSC lines. All six control-iPSC line formed rosettes similar in structure (average diameter: 0.066-0.091mm; **Figure 1E, Supplementary Table S7**). Conversely, of the 9 autism-iPSCs, 6 formed rosettes with a smaller diameter (0.05-0.06mm); 2 did not form any rosette structures at all (026ASM and 004ASM; both clones); while 010ASM formed rosettes with diameters similar to controls (0.07mm) (**Figure 1E, Supplementary Table S7**). Autism-iPSC lines also formed more rosettes per 100 cells (**Figure 1F, Supplementary Table S7**). Anomalous formation of rosettes was recapitulated in day 30 3D cortical spheroids (**Supplementary Figure S4A**), with fewer complete rosettes observed in autism-iPSC spheroids (**Supplementary Figure S4B**). One explanation for these observed morphological differences could be that autism-iPSCs have altered levels of cell proliferation. All control- and autism-iPSCs had similar rates of cell proliferation at each developmental stage examined (**Figure 1G**). Together, these data show that autism-iPSCs form anomalous rosettes independent of alterations in cell proliferation.

Divergence from typical development in autism occurs at a precursor cell stages during cortical differentiation

Abnormal rosette proliferation observed in autism-iPSCs could indicate premature or atypical neuronal differentiation in autism-iPSCs. To investigate this possibility, we assayed cortically differentiating iPSCs at critical stages of cortical differentiation - day 9, 21 and 35 (**Figure 1A,**

B) for fundamental developmental markers using a HCS based approach. First, we assessed the expression of Pax6 and Tuj1 in control- and autism-iPSCs (**Figure 2A**). Pax6 is a robust marker for neural precursors of cortical lineage (21), while Tuj1 is a robust pan-neuronal and neural precursor marker (22). Eight independent experimental replicates using 2 clones per line were assayed at every stage (**Figure 2B**). At day 9, Pax6 and Tuj1 was expressed in majority of control-iPSC cells (**Figure 2B, Table 1**). On day 21, both markers remained highly expressed (**Figure 2B**, **Table 1**). We also measure the Rate of Cell Type Assignment (dCTA) as an independent way to compare how quickly cell identity was being acquired or lost between developmental stages. In control-iPSCs, Pax6 dCTA was 13 cells/day between day 9 and day 21, while for Tuj1, dCTA=159 cells/day (**Figure 2C**). In contrast in the autism group, Pax6 and Tuj1 day 9 expression was lower than in controls (**Figure 2B, Table 1**). Assessment of cell identity acquisition in autism-iPSCs showed that Pax6 dCTA was 317 cells/day and Tuj1 dCTA=368 cells/day. These values were higher than those observed following the differentiation of control-iPSCs. However, despite this increased rate of cell identity acquisition, Pax6 and Tuj1 expression was still significantly lower in autism-iPSCs at day 21 compared to control-iPSCs (**Figure 2C, Table 1**). As expected, variability was observed throughout the differentiation protocol between experimental replicates. However, this variability was more pronounced in the autism-iPSCs. ANOVA revealed greater overall spread of data points and higher F-values in the majority of parameters assessed during differentiation of autism-iPSCs. Of note, individual clones from each line behaved in a similar manner indicating that the use of multiple clones was not the source of variability (**Supplementary Figure S5A, S5C, Supplementary Table S8**). Moreover, both non-syndromic and syndromic samples appeared to behave similarly (**Supplementary Figure S6A, C, D**). These data showed that control iPSC-derived precursors expressed Pax6 and Tuj1 early during differentiation, while autism-iPSCs display lower Pax6 and Tuj1 expression at the equivalent stage. Beyond

this stage the rate of acquisition of Pax6 and Tuj1 was higher in autism-iPSCs, and the difference between control and autism was substantially reduced at day 21.

Altered development of forebrain precursor lineages in autism-iPSCs independent of cell proliferation

Previous iPSC studies have linked an imbalance in GABA-glutamatergic progenitor cells and neuronal function with a macrocephaly associated cell proliferation phenotype (13, 17). Thus, we were interested in establishing whether a similar imbalance in the presence of GABA-glutamatergic progenitor cells could be observed in our autism-iPSCs. We investigated the development of precursors expressing Emx1, known to be expressed in dorsal forebrain (glutamatergic) neurons and precursors (23-25), and Gad67, the rate limiting enzyme in the GABA synthesis pathway and known to be expressed in GABAergic cells (26, 27) (**Figure 3A**). At day 9, EMX1 expression was significantly higher in control compared to autism neural precursors (**Figure 3B, C, Table 1**). At day 21, EMX1 expression in both groups appeared to remain stable, with only minor reduction in control precursors (dCTA = -41 cells/day), as opposed to a minor increase (dCTA = +10 cells/day) in the autism group (**Figure 3C**). At this stage, control neural progenitors expressed EMX1 significantly higher than autism neural progenitors (**Figure 3B, C, Table 1**). In day 35 immature neurons, EMX1 expression in both control and autism neurons was reduced compared to day 9 and day 21 precursors; however the reduction was significantly more acute in the autism group (dCTA = -148 cells/day in control-iPSCs vs dCTA = -254 cells/day in autism-iPSCs) (**Figure 3C**). Gad67 expression in autism- and control-iPSCs followed an opposing trajectory. At day 9, Gad67 expression was significantly higher in the control precursors, while autism precursors displayed negligible expression (**Figure 3B, C, Table 1**). At day 21, Gad67 expression was reduced in the control

progenitors (dCTA = -68 cells/day), but significantly increased in autism neural progenitors (dCTA = +185 cells/day) (**Figure 3C**). Both control and autism progenitors had similar Gad67 expression at this stage (**Figure 3B, C**). However by day 35, Gad67 expression in autism neurons was higher than that in control neuron (control dCTA = -76 cells/day, autism dCTA = +176 cells/day) (**Figure 3C, D**). Similar to what we observed with Pax6 and Tuj1 expressing cells, Emx1 and Gad67 expression also showed conspicuous variability. Again, ANOVA revealed greater variability in majority of the parameters in autism lines; with no contribution of clones to the observed variability (**Supplementary Figure S5B, S5C, Supplementary Table S8**). Non-syndromic and syndromic samples behaved in a similar manner (**Supplementary Figure S6B-D**). Lastly, we examined the expression of TBR1, a transcription factor expressed in early born excitatory neurons (10, 28), in day 35 neurons. This revealed that differentiated control-iPSCs had higher levels of TBR1 positive cells than differentiated autism-iPSCs (**Figure 3E**). Taken together, these data showed significant differences in the determination of neuronal subtype identity of cortical lineage, in control- and autism-iPSCs.

Generation of midbrain floorplate progenitors reveal negligible differences between control- and autism-iPSCs

The differences in cell fate acquisition observed between control- and autism-iPSCs could be due to genetic differences between control- and autism-iPSCs. Alternatively, this variation could be due to stochastic fluctuations in activation of key transcription factors during differentiation, as reported during iPSC differentiation towards a cortical fate (29). However, these differences could also be due to an inherent abnormality in the ability of our autism-iPSCs to undergo neural differentiation. Therefore, we sought to determine whether both control- and autism-iPSCs differentiated efficiently into neural progenitor cells specific for

another neuronal linage; specifically mesencephalic dopamine (mDA) neurons. We chose this fate as mDAs are generated from midbrain floor plate progenitors (mFPPs) that arise from cells located on the ventral midline of the neural tube floor plate. The generation of mFPPs would, therefore, require a distinct set of factors compared to those needed for the generation of cortical precursor cells. While dysfunction in mDAs have been linked with Parkinson's disease and schizophrenia, there are fewer reports of dysfunction in these cells in autism. Therefore, we reasoned that generating mFPPs (7, 8), allow us to examine the differentiation capacity of our iPSCs . After 10 days of differentiation, nearly 100% of mFPPs from both control- and autism-iPSCs were positive for LMX1A, an essential transcription factor required for defining a midbrain identity (30) (**Figure 4A, B**). No difference was observed between control- and autism-iPSCs. Similarly, expression of the transcription factor FOXA2, which positively regulates neurogenic factors in dopaminergic precursor cells (31), did not differ between control and autism mFPPs (**Figure 4A, B**). Variability was also reduced in all the iPSC lines during midbrain differentiation (**Figure 4B**). Taken together, these data showed considerably reduced differences in midbrain lineage differentiation between control- and autism-iPSCs.

Hierarchical clustering reveals sub-grouping of study participants based on cellular phenotypes alone

The findings in this study indicates that there may be a link between autism and prenatal cortical development. Our HCS screening approach has allowed us to collect large data sets of multiple cellular readouts at a number of developmental time points for each iPSC line. Hierarchical clustering approaches allow for the identification of similar patterns between samples by placing them into cluster sets (32). Using this approach, we tested whether there was a relationship between atypical cortical neurogenesis and diagnosis, based on cellular

phenotypes from control- and autism-iPSCs. Data points from each iPSC-line was amalgamated into a heatmap (**Figure 5A**), and participants were ordered on the heatmap based on a mean linkage method. We then visualised the clustering in the form of an unrooted dendogram (**Figure 5B**), as participant in this study were unrelated (33). We discovered notable relationships between samples. First, the control and autism participants grouped separately. Within the autism cluster, the participants with *NRXN1* deletions (109NXM and 092NXF) grouped on the same branch. Three syndromic autism participants (109NXM, 092NXM, 245ASM) did not group together with the non-syndromic participants. Lastly, the two autism samples 004ASM and 010ASM seemed to group on the same branch based on not only the cellular data points but also gene expression patterns in **Figure 1C, S3B**. The individual patterns that emerged out of this unbiased analysis suggests that there is a potential that cellular phenotypes could reflect nature of autism diagnosis. Further studies using larger collections of deeply phenotype iPSCs as well as more detailed cellular readouts are needed in order to further understand whether such an association is robust over independent cohorts.

Discussion

In this study, we investigated whether iPSCs generated from autistic individuals display differences during the prenatal cortical development. Previous studies have indicated that prenatal development is a critical period for the emergence of phenotypes associated with autism (12-14). However, these studies utilized iPSCs generated from autistic individuals comorbid with macrocephaly, making it unclear whether the observed cellular effects were due to autism or altered brain size. In this study, we studied iPSCs generated from a heterogeneous group of autistic individuals without macrocephaly, recruited from three independent cohorts. Thus, we were able to test whether altered cellular identities occurred during differentiation of

autism-iPSCs towards at cortical fate, and if this was detectable from an early developmental stage. This collection included 4 autistic individuals with uncharacterised genetic background and 5 autistic individuals with CNVs in high risk autism loci.

First we found that autism-iPSCs generated atypical neural rosettes, indicating an alteration in neural differentiation. Consistent with this, autism-iPSCs showed significant differences in development of early neural progenitor cells. This effect persisted at a late precursor cell stage although to a lesser degree. No differences in proliferative capacities was observed between control- and autism-iPSCs indicating this this was not the cause of altered neurogenesis in autism-iPSCs. Examination of cortical neuron subtypes revealed a divergence in the development of dorsal forebrain or excitatory precursors and ventral forebrain or inhibitory precursors from an early stage of development. Conversely, control- and autism-iPSCs demonstrated the same ability to into mFFP cells. This indicates that atypical neurogenesis predominately impacts the development of cortical linages in autism-iPSCs. Finally, based on all the temporal cortical data points acquired in this study, the participants grouped separately into controls and autism, with further unbiased branching within the autism cohort. Together, these data suggests that unique developmental differences associated with autism may be established at early prenatal stages.

We were particularly interested in modelling divergent patterns of development in the autistic cortex. We used a cortical differentiation protocol that recapitulates cortical precursor development from iPSCs, and yielded primarily excitatory cortical neurons (10). This enabled us to study early stages of neural development, when neural rosette begin forming (day 9), equivalent to neural tube closure (approximately 4 weeks of gestation) (34). We found marked anomalies in rosette morphology in 3 out of 9 autism-iPSCs (004ASM, 026ASM, 245ASM) resulting in either malformation or negligible neural rosette formation. In 010ASM, neural

346 precursors were visibly dissociated from the rosette-structure, while in 092NXF, 109NXM,
347 ASDM1, 132ASM and 289ASM cells appeared elongated and lumen formation was also
348 affected. Further studies are required to elucidate the mechanisms responsible for the altered
349 rosette structures and formation observed. Disruption of neural rosettes has been found to
350 promote premature neurogenesis (35, 36). This may explain the high rate of Pax6+ and Tuj1+
351 precursor generation between day 9 and day 21 in autism-iPSCs. It could also explain divergent
352 precursor subtype assignment during early development, which we observed through opposing
353 trajectory of Gad67 expressing cells in control- and autism-iPSCs. We noted that the
354 appearance of Gad67+ cells in our cultures was surprising as SMAD inhibition is known to
355 drive stem cells towards a dorsal forebrain lineage, while GABAergic neurons are known to be
356 generated from a ventral forebrain lineage (37). However, low numbers of GABAergic cells
357 are known to be generated using the SMAD inhibition protocol (38, 39), and appearance of
358 Gad67+ cells and their dysregulation in our study may be a result of dysregulated molecular
359 mechanisms associated with atypical precursor subtype assignment.

360 It is of interest that in the current study the phenotypic changes occurred without the
361 presence of proliferative differences between control- and autism-iPSCs. This suggested that
362 cell type and structural anomalies previously reported using autism iPSCs (12, 13) may be
363 independent of macrocephaly associated cell proliferation alterations. Alterations in rosette
364 formation may also contribute to the switching of precursor identity seen during development
365 in autism-iPSCs. Further investigation into temporal precursor cell type specification will be
366 needed to understand the mechanisms and types of cells involved. Notably, iPSC studies of
367 non-syndromic autism remain underpowered. Nevertheless, the reports of neurodevelopmental
368 differences between autism- and control-iPSCs are robust (13, 14, 40). Although our cohort
369 size would be considered inadequate for a study into non-syndromic autism, it is comparable
370 to recent iPSC-based psychiatric studies (12-14, 41). To achieve effect size in our study, we

have used multiple clones for each iPSC-line. In addition, we utilized a HCS screening of 'cellomic' cell-based phenotyping approach (15, 16, 42), recording thousands of data points from each iPSC-line.

Another consideration we faced during cellular phenotyping of iPSCs being differentiated towards a cortical fate was the high degree of variability between experimental replicates. This variability is due in part to stochastic fluctuations in transcription factor activation during cortical differentation (29, 43). We observed that out of the 10 temporal data points recorded, 7 showed a greater degree of variability in autism-iPSCs. To rule out whether this variability was due to an iPSC-related abnormal artefact, we differentiated both control- and autism-iPSCs towards a mesencephalic fate. Following this protocol iPSCs from either control of autistic individuals behaved similarly and demonstrated reduced variability. This suggests that the variability observed in this study is specific to the cortical differentiation rather than an iPSC-related artefact. Moreover, these data indicate that alteration during early stage of development associated with autism may occur in a region specific manner.

In this study, we used iPSCs generated from independent cohorts and from individuals with autism but without macrocephaly. Using unbiased methods, we identify that differentiation of autism-iPSCs towards a cortical but not a mesencephalic fate, results in atypical neurogenesis characterised by premature maturation and abnormal specification of neural progenitor cells. These effects occur in the absence of altered proliferative activity between control- and autism-iPSCs. Identification of these cellular/molecular phenotypes enabled us to find common cellular pathways in a cohort having heterogeneous genetic background. In future, similarly designed studies will help identify which cellular pathways underlie these phenotypes, and may help to improve diagnosis and develop a greater understanding of the origins of autism.

Disclosures: The authors report no biomedical financial interests or potential conflicts of interest.

Acknowledgments

We gratefully acknowledge the participants in this study. This study was supported by grants from the European Autism Interventions (EU-AIMS) and AIMS-2-TRIALS: the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115300, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution (JP, SBC, DPS, DM, GM); StemBANCC: support from the Innovative Medicines Initiative joint undertaking under grant 115439-2, whose resources are composed of financial contribution from the European Union [FP7/2007-2013] and EFPIA companies' in-kind contribution (JP, DPS); MATRICS: the European Union's Seventh Framework Programme (FP7-HEALTH-603016) (DPS, JP); the Wellcome Trust ISSF Grant (No. 097819) and the King's Health Partners Research and Development Challenge Fund – a fund administered on behalf of King's Health Partners by Guy's and St Thomas' Charity (Grant R130587) awarded to DPS; an Independent Investigator's Award from the Brain and Behavior Foundation (formally National Alliance for Research on Schizophrenia and Depression (NARSAD); Grant No. 25957) to DPS, and Seed funding from Medical Research Council, UK (MR/N026063/1) awarded to DPS; the Mortimer D Sackler Foundation to DM; the Autism Research Trust, the Chinese University of Hong Kong, the Templeton World Charitable Foundation, and Wellcome Trust awarded to SBC; and a doctoral fellowship from the Jawaharlal Nehru Memorial Trust awarded to D.A. The funding organizations had no role in the design and conduct of the study, in the collection, management, analysis and interpretation of the data, or in the preparation, review or approval of the manuscript. We are grateful to Debbie Spain and Suzanne Coghlan for participant recruitment,

to Rosy Watkins, Hema Pramod, Rupert Faraway, Pooja Raval, Kate Sellers, Michael Deans and Rodrigo Rafagnin for assistance during the study, and to Aicha Massrali, Arkoprovo Paul, Bhismadev Chakrabarti, Michael Lombardo, Rick Livesey and Mark Kotter for valuable discussions. We thank the Wohl Cellular Imaging Centre (WCIC) at the IoPPN, Kings College, London for help with microscopy. A non-peer review version of this manuscript was posted on the pre-print server BioRxiv: doi.org/10.1101/349415.

Ethics, consent and permissions

Informed consent from participants have been taken before recruitment: Patient iPSCs for Neurodevelopmental Disorders (PiNDs) study' (REC No 13/LO/1218).

Consent to publish

We have obtained consent to publish from the participant to report individual patient data.

Availability of data and materials

Sequence data have been uploaded on synapse.org. Synapse ID: syn8118403, DOI: doi:10.7303/syn8118403

Authors' contribution

DA, JP, JC, DPS, SBC conceived the study and wrote the first draft. VS, DHG conceived and developed bioinformatics analysis framework and analysis. DA, PN, CS, KJ responsible for sample preparation. GM was responsible for ethics application. GM, MAZ, JH, IL, DS and DM responsible for recruiting and collecting hair samples from individuals with autism and controls. All co-authors contributed to study concept, design, and writing of the manuscript. All authors read and approved the final manuscript.

Figure Legends

Figure 1: Differentiation of iPSCs into cortical lineage reveals gene expression and neural rosette formation differences between control and autism. (A) Study design and differentiation time points used in this study. **(B)** Differentiation of control and autism iPSCs generate precursor markers Ki67, Nestin and Pax6 and neuronal markers TBR1 and MAP2. **(C)** Differential gene expression and hierarchical clustering reveals significant differences between control and autism samples (biological replicates for each sample labelled 1 and 2). **(D)** Day 9 neural rosette morphology from all participants in this study. **(E)** Rosette diameter violin plot (horizontal lines show mean rosette diameter, blue: control, red-dashed: autism). **(F)** Number of rosettes per 100 cells (horizontal lines show mean rosette number, blue: control, red-dashed: autism). **(G)** Proliferation during cortical differentiation at day 0, day 9, day 21, day 35 (dashed lines are control samples, colour key on top right corner).

Figure 2: Atypical cortical differentiation of autism iPSCs. (A) At day 9 and day 21 precursor cell stages, both control and autism iPSCs expressed Pax6 and Tuj1. **(B)** Quantification of Pax6+ and Tuj1+ cells of individual participants (% cells positive per

experimental replicate) showed significant differences between autism and control. **(C)** Mean values of % positive cells over time show significant difference between control and autism at both day 9 and day 21, as well as significant difference in rate of appearance of markers. Histogram shows normal distribution of experimental data points, and demonstrates variability between control and autism. (LV: Lentivirus reprogramming method used for generating these iPSCs; s: Participants with syndromic autism)

Figure 3: Atypical differentiation into dorsal and ventral forebrain precursors in autism. **(A)** EMX1 was expressed at day 9, day 21 and day 35 in both control and autism groups. Gad67 expression in both groups was time dependant, it decreased over time in in controls, while increased over time in autism. **(B)** Quantification of EMX1+ and Gad67+ cells (% cells positive per experimental replicate) showed significant differences between autism and control. **(C)** Mean values of % positive cells over time show significant difference between control and autism at every time point, except for Gad67 at day 21 precursor stage. **(D)** Histogram shows normal distribution of experimental data points and clear difference in distribution of data points between groups. **(E)** Control and autism iPSCs also expressed TBR1 at day 35 of cortical differentiation, and TBR1 expression was marginally higher in control vs autism. (LV: Lentivirus reprogramming method used for generating these iPSCs; s: Participants with syndromic autism)

Figure 4: Efficient differentiation of control and autism iPSCs towards a midbrain fate. **(A)** Both control and autism iPSCs expressed LMX1A and FOXA2 when differentiated into a midbrain floor plate precursor (mFPP) cells. **(B)** Differences between control and autism iPSCs expressing LMX1A or FOXA2 was near negligible.

Figure 5: Hierarchical clustering of cellular data using mean linkage method. (A) All controls and autism iPSC lines were grouped based on % positive values for Pax6, Tuj1, EMX1, Gad67 at day 9, 21, 35 cortical differentiation, and LMX1A and FOXA2 at day 11 of midbrain differentiation. Controls and autism participants were grouped separately using this unsupervised learning method. **(B)** Unrooted phylogenetic tree showing relatedness of individual participants based on cellular phenotypes. Syndromic samples branched separately to non-syndromic samples*. NRXN1 deletion samples grouped together on the same branch*†. 004ASM and 010ASM which grouped on the same branch (shown with dashed lines) also grouped similarly based on gene expression data shown in Fig 1C, S3B.

References

1. Berg JM, Geschwind DH (2012): Autism genetics: searching for specificity and convergence. *Genome Biol*. 13:247.
2. Bourgeron T (2015): From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat Rev Neurosci*. 16:551-563.
3. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. (2012): Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 485:246-250.
4. Johnson CP, Myers SM, American Academy of Pediatrics Council on Children With D (2007): Identification and evaluation of children with autism spectrum disorders. *Pediatrics*. 120:1183-1215.

5. Schubert D, Martens GJ, Kolk SM (2015): Molecular underpinnings of prefrontal cortex development in rodents provide insights into the etiology of neurodevelopmental disorders. *Mol Psychiatry*. 20:795-809.

6. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, et al. (2013): Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*. 155:1008-1021.

7. Fedele S, Collo G, Behr K, Bischofberger J, Muller S, Kunath T, et al. (2017): Expansion of human midbrain floor plate progenitors from induced pluripotent stem cells increases dopaminergic neuron differentiation potential. *Sci Rep*. 7:6036.

8. Kriks S, Shim JW, Piao J, Ganat YM, Wakeman DR, Xie Z, et al. (2011): Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature*. 480:547-551.

9. Pasca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Pasca AM, et al. (2011): Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nat Med*. 17:1657-1662.

10. Shi Y, Kirwan P, Livesey FJ (2012): Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc*. 7:1836-1846.

11. Lancaster MA, Renner M, Martin CA, Wenzel D, Bicknell LS, Hurles ME, et al. (2013): Cerebral organoids model human brain development and microcephaly. *Nature*. 501:373-379.

12. Marchetto MC, Belinson H, Tian Y, Freitas BC, Fu C, Vadodaria KC, et al. (2016): Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Mol Psychiatry*.

13. Mariani J, Coppola G, Zhang P, Abyzov A, Provini L, Tomasini L, et al. (2015): FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell*. 162:375-390.

14. Schafer ST, Paquola ACM, Stern S, Gosselin D, Ku M, Pena M, et al. (2019): Pathological priming causes developmental gene network heterochronicity in autistic subject-derived neurons. *Nat Neurosci*. 22:243-255.

15. Kathuria A, Nowosiad P, Jagasia R, Aigner S, Taylor RD, Andreae LC, et al. (2018): Stem cell-derived neurons from autistic individuals with SHANK3 mutation show morphogenetic abnormalities during early development. *Mol Psychiatry*. 23:735-746.

16. Shum C, Dutan L, Annuario E, Warre-Cornish K, Taylor SE, Taylor RD, et al. (2020): Delta(9)-tetrahydrocannabinol and 2-AG decreases neurite outgrowth and differentially affects ERK1/2 and Akt signaling in hiPSC-derived cortical neurons. *Mol Cell Neurosci*. 103:103463.

17. Marchetto MC, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, et al. (2010): A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell*. 143:527-539.

18. Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, et al. (2016): Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature*.

19. Lancaster MA, Huch M (2019): Disease modelling in human organoids. *Dis Model Mech*. 12.

20. Edri R, Yaffe Y, Ziller MJ, Mutukula N, Volkman R, David E, et al. (2015): Analysing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors. *Nat Commun*. 6:6500.

553 21. Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W, et al. (2015): Dissecting
554 neural differentiation regulatory networks through epigenetic footprinting. *Nature*. 518:355-
555 359.
556 22. Tischfield MA, Baris HN, Wu C, Rudolph G, Van Maldergem L, He W, et al. (2010):
557 Human TUBB3 mutations perturb microtubule dynamics, kinesin interactions, and axon
558 guidance. *Cell*. 140:74-87.
559 23. Zhang W, Peterson M, Beyer B, Frankel WN, Zhang ZW (2014): Loss of MeCP2 from
560 forebrain excitatory neurons leads to cortical hyperexcitation and seizures. *J Neurosci*.
561 34:2754-2763.
562 24. Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JL, Jones KR (2002): Cortical
563 excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing
564 lineage. *J Neurosci*. 22:6309-6314.
565 25. Costa MR, Muller U (2014): Specification of excitatory neurons in the developing
566 cerebral cortex: progenitor diversity and environmental influences. *Front Cell Neurosci*. 8:449.
567 26. Lazarus MS, Krishnan K, Huang ZJ (2015): GAD67 deficiency in parvalbumin
568 interneurons produces deficits in inhibitory transmission and network disinhibition in mouse
569 prefrontal cortex. *Cereb Cortex*. 25:1290-1296.
570 27. Azim E, Jabaudon D, Fame RM, Macklis JD (2009): SOX6 controls dorsal progenitor
571 identity and interneuron diversity during neocortical development. *Nat Neurosci*. 12:1238-
572 1247.
573 28. Gaspard N, Bouschet T, Hourez R, Dimidschstein J, Naeije G, van den Ameele J, et al.
574 (2008): An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature*. 455:351-
575 357.
576 29. Stumpf PS, Smith RCG, Lenz M, Schuppert A, Muller FJ, Babtie A, et al. (2017): Stem
577 Cell Differentiation as a Non-Markov Stochastic Process. *Cell Syst*. 5:268-282 e267.

30. Deng Q, Andersson E, Hedlund E, Alekseenko Z, Coppola E, Panman L, et al. (2011): Specific and integrated roles of Lmx1a, Lmx1b and Phox2a in ventral midbrain development. *Development*. 138:3399-3408.

31. Ferri AL, Lin W, Mavromatakis YE, Wang JC, Sasaki H, Whitsett JA, et al. (2007): Foxa1 and Foxa2 regulate multiple phases of midbrain dopaminergic neuron development in a dosage-dependent manner. *Development*. 134:2761-2769.

32. Fliri AF, Loging WT, Thadeio PF, Volkmann RA (2005): Biological spectra analysis: Linking biological activity profiles to molecular structure. *Proc Natl Acad Sci U S A*. 102:261-266.

33. Wilkinson M, McInerney JO, Hirt RP, Foster PG, Embley TM (2007): Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends Ecol Evol*. 22:114-115.

34. Elkabetz Y, Panagiotakos G, Al Shamy G, Socci ND, Tabar V, Studer L (2008): Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev*. 22:152-165.

35. Grabiec M, Hribkova H, Varecha M, Stritecka D, Hampl A, Dvorak P, et al. (2016): Stage-specific roles of FGF2 signaling in human neural development. *Stem Cell Res*. 17:330-341.

36. Hribkova H, Grabiec M, Klemova D, Slaninova I, Sun YM (2018): Calcium signaling mediates five types of cell morphological changes to form neural rosettes. *J Cell Sci*. 131.

37. Birey F, Andersen J, Makinson CD, Islam S, Wei W, Huber N, et al. (2017): Assembly of functionally integrated human forebrain spheroids. *Nature*. 545:54-59.

38. Crompton LA, Byrne ML, Taylor H, Kerrigan TL, Bru-Mercier G, Badger JL, et al. (2013): Stepwise, non-adherent differentiation of human pluripotent stem cells to generate basal forebrain cholinergic neurons via hedgehog signaling. *Stem Cell Res*. 11:1206-1221.

602 39. Pauly MG, Krajka V, Stengel F, Seibler P, Klein C, Capetian P (2018): Adherent vs.
603 Free-Floating Neural Induction by Dual SMAD Inhibition for Neurosphere Cultures Derived
604 from Human Induced Pluripotent Stem Cells. *Front Cell Dev Biol*. 6:3.
605 40. Muotri AR (2016): The Human Model: Changing Focus on Autism Research. *Biol*
606 *Psychiatry*. 79:642-649.
607 41. Flaherty E, Zhu S, Barretto N, Cheng E, Deans PJM, Fernando MB, et al. (2019):
608 Neuronal impact of patient-specific aberrant NRXN1alpha splicing. *Nat Genet*. 51:1679-1690.
609 42. Wang C, Ward ME, Chen R, Liu K, Tracy TE, Chen X, et al. (2017): Scalable
610 Production of iPSC-Derived Human Neurons to Identify Tau-Lowering Compounds by High-
611 Content Screening. *Stem Cell Reports*. 9:1221-1233.
612 43. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, et al. (2007):
613 Nanog safeguards pluripotency and mediates germline development. *Nature*. 450:1230-1234.

614

Figure 1

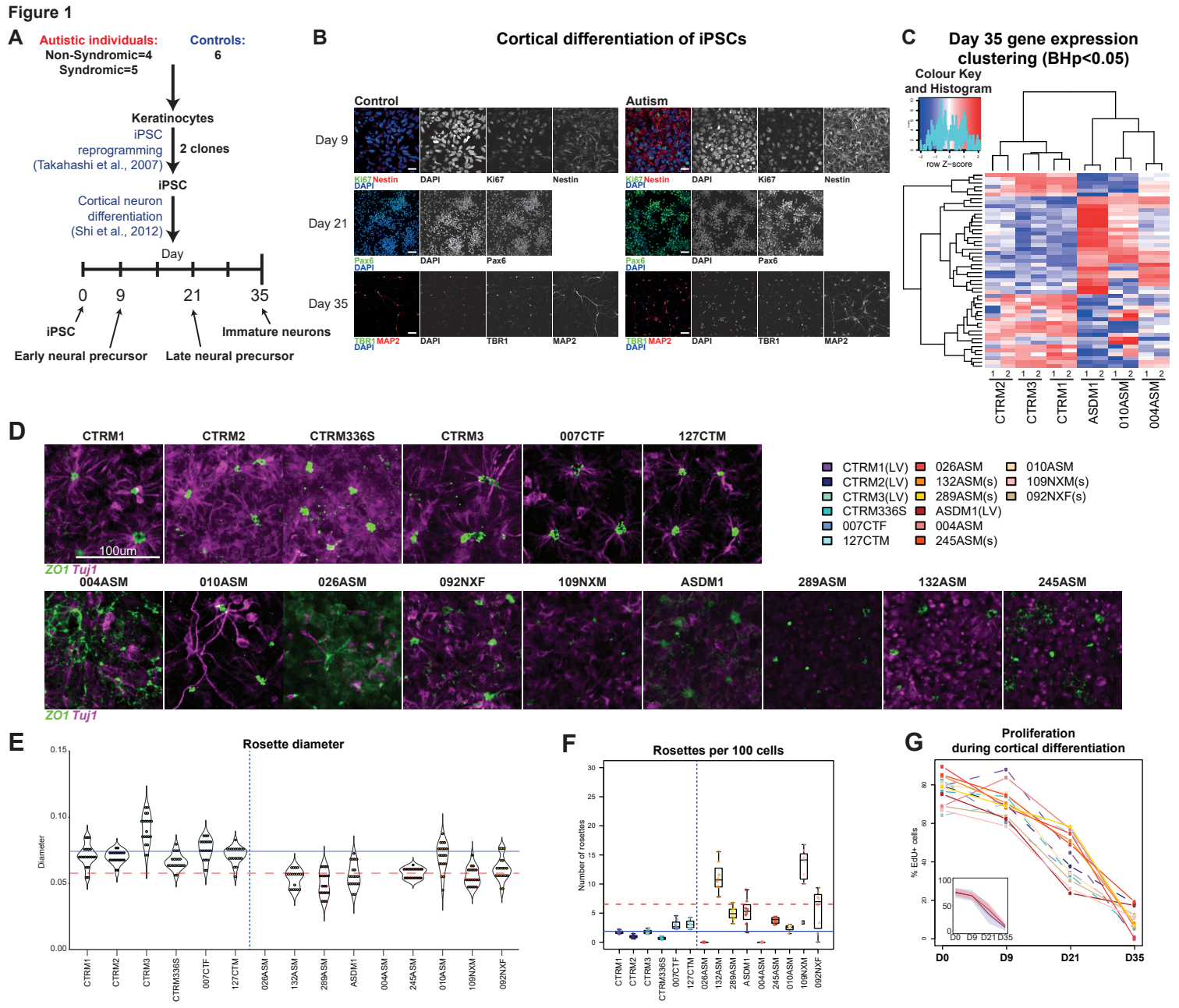

A
Autistic individuals:
Non-Syndromic=4
Syndromic=5
Controls:
6
Keratinocytes
iPSC reprogramming (Takahashi et al., 2007)
2 clones
iPSC
Cortical neuron differentiation (Shi et al., 2012)
Day
0
9
21
35
iPSC
Early neural precursor
Late neural precursor
Immature neurons
B
Cortical differentiation of iPSCs
Control
Autism
Day 9
Day 21
Day 35
Ki67 Nestin DAPI
DAPI
Ki67
Nestin
Pax6 DAPI
Pax6
TBR1 MAP2 DAPI
TBR1
MAP2
C
Day 35 gene expression clustering (BHp<0.05)
Colour Key and Histogram
row Z-score
CTRM2
CTRM3
CTRM1
ASDM1
010ASM
004ASM
D
CTRM1
CTRM2
CTRM336S
CTRM3
007CTF
127CTM
100um
ZO1 Tuj1
004ASM
010ASM
026ASM
092NXF
109NXM
ASDM1
289ASM
132ASM
245ASM
CTRM1(LV)
CTRM2(LV)
CTRM3(LV)
CTRM336S
007CTF
127CTM
026ASM
132ASM(s)
289ASM(s)
ASDM1(LV)
004ASM
245ASM(s)
010ASM
109NXM(s)
092NXF(s)
E
Rosette diameter
Diameter
0.15
0.10
0.05
0.00
CTRM1
CTRM2
CTRM3
CTRM336S
007CTF
127CTM
026ASM
132ASM
289ASM
ASDM1
004ASM
245ASM
010ASM
109NXM
092NXF
F
Rosettes per 100 cells
Number of rosettes
30
25
20
15
10
5
0
G
Proliferation during cortical differentiation
% EdU+ cells
80
60
40
20
0
100
50
0
D0 D9 D21 D35
D0
D9
D21
D35

Figure 2

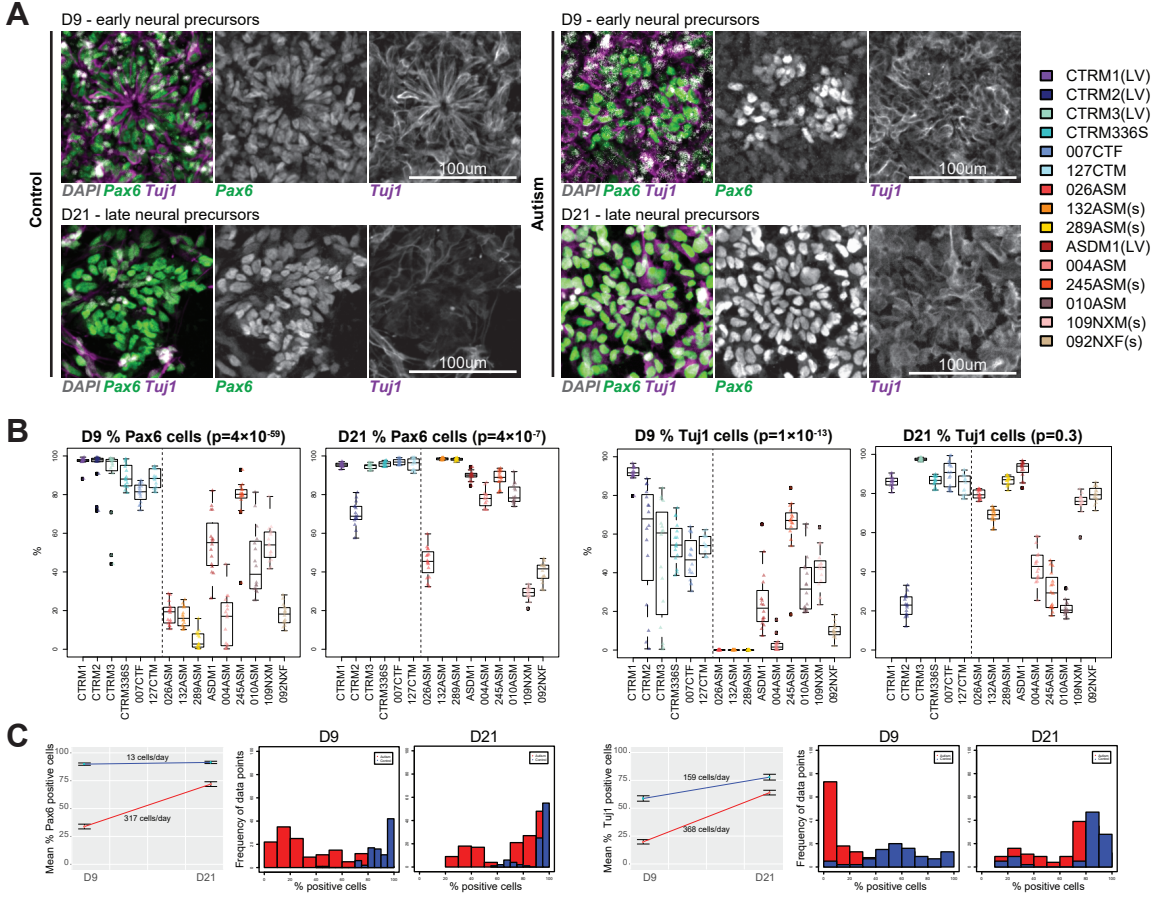

A
Control
Autism
D9 - early neural precursors
D21 - late neural precursors
DAPI Pax6 Tuj1
Pax6
Tuj1
100um
CTRM1(LV)
CTRM2(LV)
CTRM3(LV)
CTRM336S
007CTF
127CTM
026ASM
132ASM(s)
289ASM(s)
ASDM1(LV)
004ASM
245ASM(s)
010ASM
109NXM(s)
092NXF(s)
B
D9 % Pax6 cells (p=4×10-59)
D21 % Pax6 cells (p=4×10-7)
D9 % Tuj1 cells (p=1×10-13)
D21 % Tuj1 cells (p=0.3)
%
CTRM1
CTRM2
CTRM3
CTRM336S
007CTF
127CTM
026ASM
132ASM
289ASM
ASDM1
004ASM
245ASM
010ASM
109NXM
092NXF
C
Mean % Pax6 positive cells
13 cells/day
317 cells/day
D9
D21
Frequency of data points
% positive cells
Mean % Tuj1 positive cells
159 cells/day
368 cells/day

Figure 3

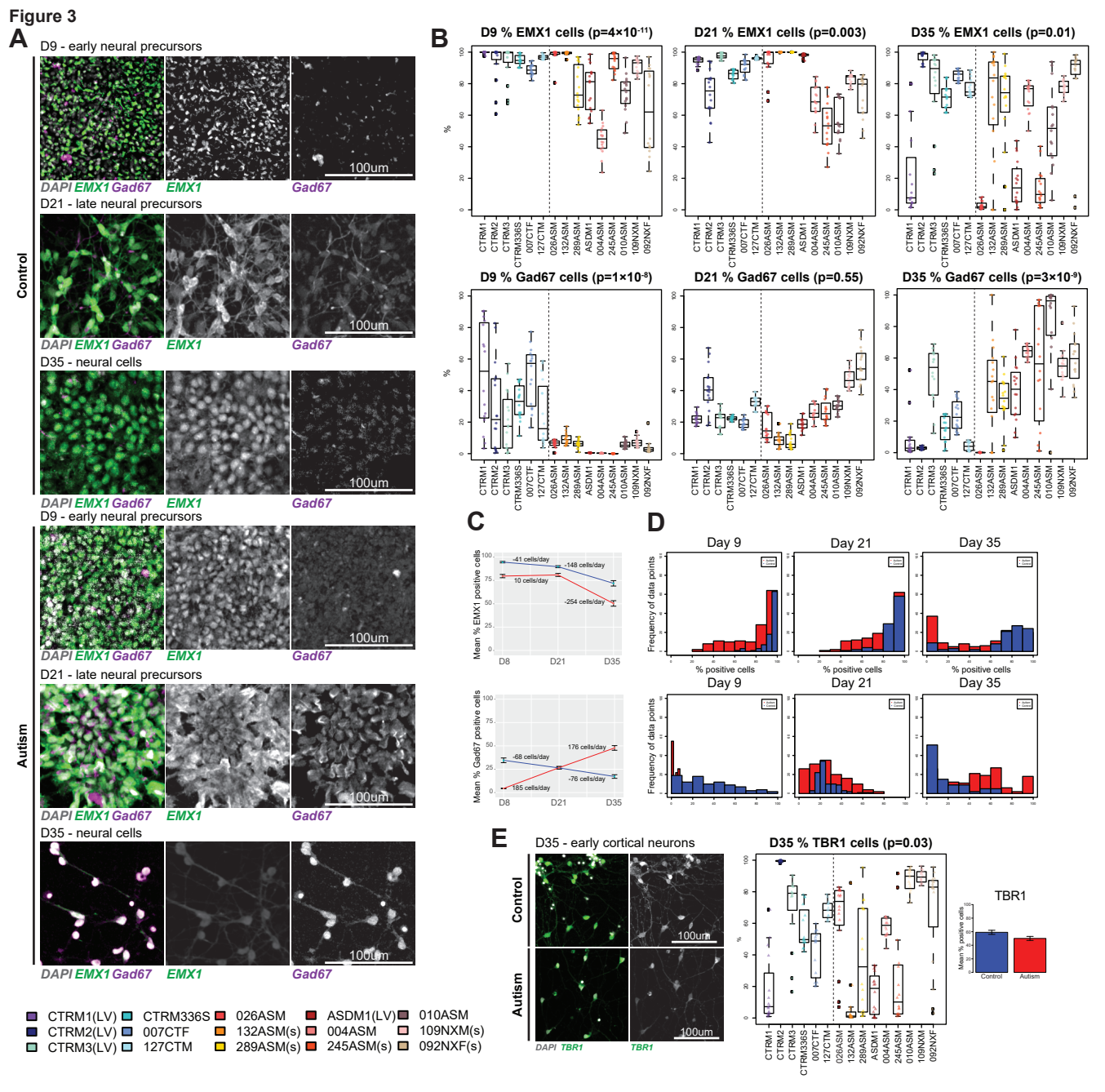

A
D9 - early neural precursors
DAPI EMX1 Gad67
EMX1
Gad67
100um
D21 - late neural precursors
D35 - neural cells
Control
Autism
B
D9 % EMX1 cells (p=4×10-11)
D21 % EMX1 cells (p=0.003)
D35 % EMX1 cells (p=0.01)
D9 % Gad67 cells (p=1×10-8)
D21 % Gad67 cells (p=0.55)
D35 % Gad67 cells (p=3×10-9)
%
CTRM1 CTRM2 CTRM3 CTRM336S 007CTF 127CTM 026ASM 132ASM 289ASM ASDM1 004ASM 245ASM 010ASM 109NXM 092NXF
C
Mean % EMX1 positive cells
-41 cells/day
10 cells/day
-148 cells/day
-254 cells/day
Mean % Gad67 positive cells
-68 cells/day
185 cells/day
176 cells/day
-76 cells/day
D8
D21
D35
D
Day 9
Day 21
Day 35
Frequency of data points
% positive cells
E
D35 - early cortical neurons
Control
Autism
DAPI TBR1
TBR1
D35 % TBR1 cells (p=0.03)
Mean % positive cells
Control
Autism
CTRM1(LV) CTRM2(LV) CTRM3(LV)
CTRM336S 007CTF 127CTM
026ASM 132ASM(s) 289ASM(s)
ASDM1(LV) 004ASM 245ASM(s)
010ASM 109NXM(s) 092NXF(s)

Figure 4

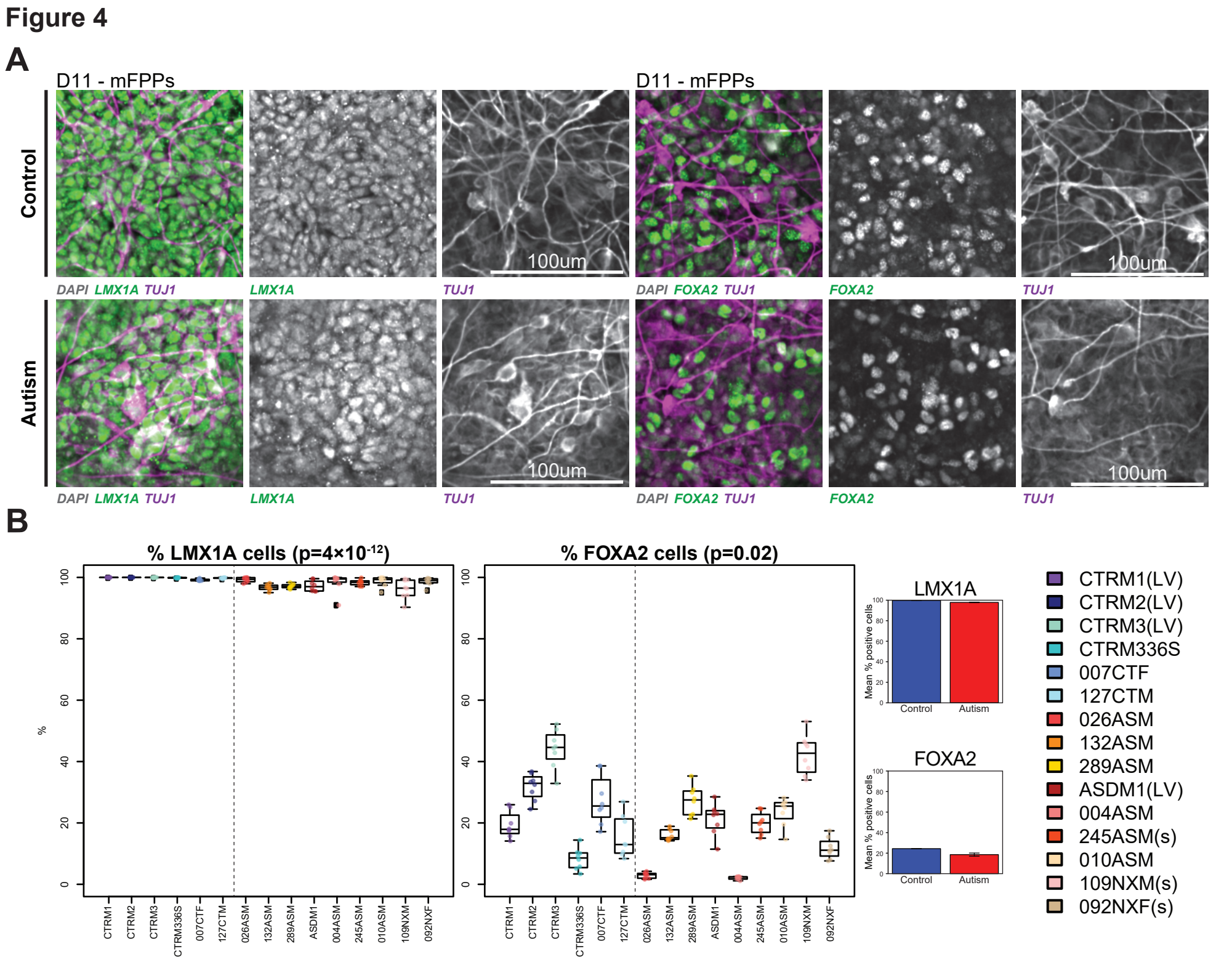
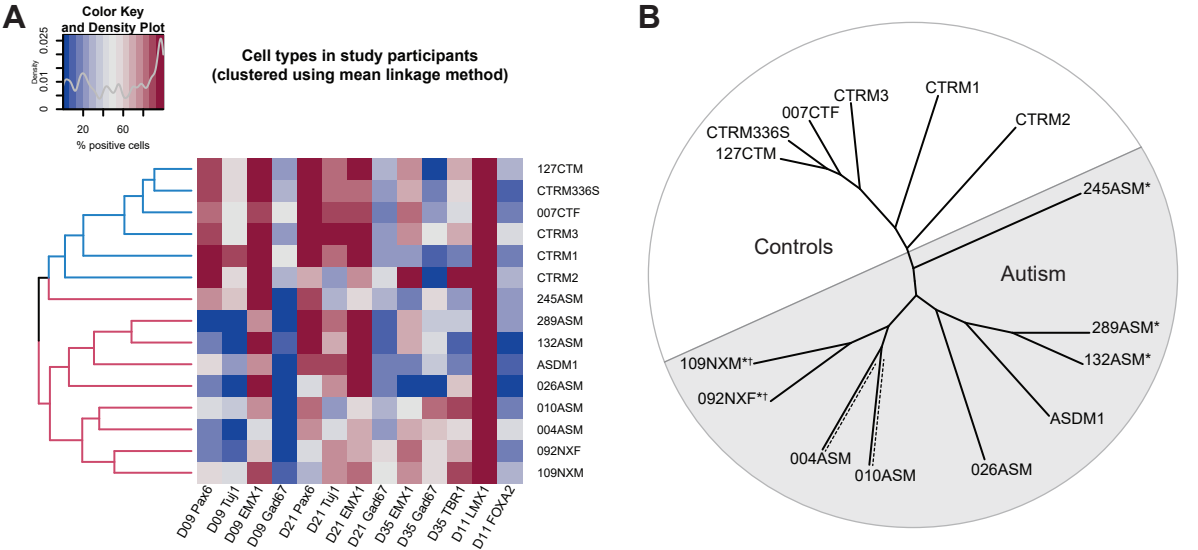

A
D11 - mFPPs
D11 - mFPPs
Control
Autism
100um
DAPI LMX1A TUJ1
LMX1A
TUJ1
DAPI FOXA2 TUJ1
FOXA2
TUJ1
B
% LMX1A cells (p=4×10-12)
% FOXA2 cells (p=0.02)
%
CTRM1
CTRM2
CTRM3
CTRM336S
007CTF
127CTM
026ASM
132ASM
289ASM
ASDM1
004ASM
245ASM
010ASM
109NXM
092NXF
LMX1A
FOXA2
Mean % positive cells
Control
Autism
CTRM1(LV)
CTRM2(LV)
CTRM3(LV)
CTRM336S
007CTF
127CTM
026ASM
132ASM
289ASM
ASDM1(LV)
004ASM
245ASM(s)
010ASM
109NXM(s)
092NXF(s)

Figure 5


A
Color Key
and Density Plot
Density
0.025
0.01
0
20
60
% positive cells
Cell types in study participants
(clustered using mean linkage method)
127CTM
CTRM336S
007CTF
CTRM3
CTRM1
CTRM2
245ASM
289ASM
132ASM
ASDM1
026ASM
010ASM
004ASM
092NXF
109NXM
D09 Pax6
D09 Tuj1
D09 EMX1
D09 Gad67
D21 Pax6
D21 Tuj1
D21 EMX1
D21 Gad67
D35 EMX1
D35 Gad67
D35 TBR1
D11 LMX1
D11 FOXA2
B
CTRM3
CTRM1
007CTF
CTRM2
CTRM336S
127CTM
245ASM*
Controls
Autism
289ASM*
132ASM*
109NXM*†
092NXF*†
ASDM1
004ASM
010ASM
026ASM

Atypical neurogenesis in induced pluripotent stem cell (iPSC) from autistic individuals

Dwaipayan Adhya[1,3,*], Vivek Swarup[2,*], Roland Nagy[3], Lucia Dutan[3], Carole Shum[3], Eva P. Valencia-Alarcón[3], Kamila Maria Jozwik[4], Maria Andreina Mendez[5], Jamie Horder[5], Eva Loth[5], Paulina Nowosiad[3], Irene Lee[6], David Skuse[6], Frances A. Flinter[7], Declan Murphy[5], Grainne McAlonan[5], Daniel H. Geschwind[2,9], Jack Price[3,8], Jason Carroll[4], Deepak P. Srivastava[3,8]§†, & Simon Baron-Cohen[1]§

Supplementary Info

Extended experimental procedures

Study participants and neuronal differentiation

Keratinocytes were collected from autistic participants, and typical controls without an autism diagnosis (Ethics approved, 13/LO/1218) as part of a larger European studies (EU-AIMS, STEMBANCC). All participants were Caucasian, while controls were selected if they did not have diagnosis of any psychiatric conditions. These were reprogrammed into iPSCs using previously described methods [1, 2]. IPS cells were cultured in E8 medium (Life Technologies) with E8 supplement (Life Technologies). Cell type quantification and proliferation assays were set up on 96-well plates. 2 clones were selected from each participant, and each clone had 8 technical replicates. RNA-sequencing was performed on 2 clones from each participant, and each clone had 2 technical replicates. This design was maintained at all stages of neural differentiation recorded. Induction of neurons of cortical lineage was established using a modified dual SMADi protocol[3]. Once the cell culture reached 95% confluence, neural induction was initiated by changing the culture medium to support neural induction, neurogenesis and neuronal differentiation. A combination of N2- and B27-containing media with additives was used, henceforth called 'neuralising medium'. N2 medium consisted of DMEM/F12 (Sigma), N2 (Gibco). B27 medium consisted of Neurobasal (Invitrogen), B27 (Gibco). Neuralising medium was supplemented with 'dual SMADi' 1 μM Dorsomorphin (Tocris), 500 ng/ml human Noggin-CF chimera (R&D Systems) – inhibitors of WNT pathway, BMPs and SMAD, and 10 μM SB431542 (Tocris) – inhibitor of TGFβ signaling. Noggin and dorsomorphin supresses embryonic development thereby inducing neural differentiation pathways, while SB431542 mediates loss of pluripotency. Midbrain floorplate precursors were differentiated from all participants till day 11 using previously established protocols[4, 5]. To generate cortical spheroids, iPSCs were first treated with KOSR-

based hiPSC media to form embryoid bodies. Neural induction was performed using dorsomorphin and SB431542. After 6 days of neural induction, spheroids were transferred into neural maintenance media till day 30, as per established protocols[6].

Cortical Spheroids

Cortical spheroids were generated using methods published by Pasca *et al.* (2015). Human iPSCs were passaged at high density (80 % confluent) as previously described. The cells were then suspended using a cell lifter in KOSR media, 10 µM ROCKi and SMAD inhibitors: 5 µM Dorsomorphin and 10 µM SB431542. The cultures were incubated in a 5% CO2 incubator and left undisturbed for 48-hours period to promote formation of embryoid bodies. Starting the second day and until the fourth day media is changed daily with fresh KOSR media supplemented with 5 µM Dorsomorphin and 10 µM SB431542. By the fifth day embryo bodies were clearly visible. At this stage, till 25th day the culture media used was B27 minus Vitamin A media plus 20 ng/mL of recombinant human EGF and recombinant human bFGF. From day 25, B27 minus Vitamin A media with 20 ng/mL of recombinant human BDNF and recombinant human NT3 was used. Media was changed every 24 hours during the first 15 days and once every 48 hours thereafter. The spheroids were harvested at day 30.

Immunocytochemistry

Cultures were fixed in 4% formaldehyde followed by ice-cold 100% methanol and processed for immunofluorescence staining, confocal microscopy and high throughput imaging. Secondary antibodies used for primary antibody detection were species-specific Alexa-dye conjugates (Invitrogen). We used the following primary antibodies to Ki67 (Thermo Fisher PA5-16785), Nestin (Millipore MAB5326), Pax6 (BioLegend 901301), TBR1 (Abcam ab31940), MAP2 (Abcam ab92434), Emx1 (ThermoFisher PA5-35373), Gad67 (Abcam ab26116), Tuj1 (BioLegend 801201), CD44 (R&D Systems MAB7045), LMX1A (Abcam

ab139726), FOXA2 (Invitrogen 701698). Quantification was performed on the Perkin Elmer Harmony Software v4.9, which is based on the CellProfiler high throughput image analysis system[7]. Cell nuclei were first identified based on DAPI staining. For nuclear proteins, only the nuclear area was selected. For cytoplasmic protein, the area around the nucleus was selected. Thresholds of fluorescent intensity was selected after background subtraction. Threshold for each probe remained unchanged for every sample imaged. Antibodies, dilutions used and fluorescence threshold information in **Supplementary Table S5**.

Cortical spheroids were first washed in PBS and then fixed in 4% formaldehyde in 4% sucrose-PBS for 45 minutes. Following fixation, spheroids were washed in PBS and stored in sucrose 30% (weight/volume); sucrose sinking improves their preservation. After sucrose sinking, spheroids were permeabilised in 2% normal goat serum (NGS) in 0.1% Triton X100 in PBS for 60 minutes. Spheroids were incubated in a solution of permeabilization-blocking solution containing the primary antibodies for 48 hours (**Supplementary Table S5**). The samples were washed three times for three minutes in PBS and incubated with permeabilisation-blocking solution containing the secondary antibodies and HOECHST (nuclear staining) for two hours and kept in darkness to prevent bleaching of the fluorophores. Cortical spheroids were mounted as whole tissues (without sectioning), due to their small size (1-5 mm). Cortical spheroids were imaged using a Lecia SP5 confocal microscope using a 40x oil immersion objective. Imagines were acquired as Z stacks at resolution 1024x1024px and employing multiple channels: 405 (DAPI/HOECHST, blue), 488nm (green), 561nm (red) and 633nm (far-red). A line average of three was used in each channel during point scanning to prevent random background signal due to light scattering.

EdU labelling

IPSCs at specific stages of differentiation were labelled with EdU (5-ethynyl-2′-deoxyuridine) using the Click-iT EdU Assay (Invitrogen). Cells were incubated with EdU for 4 hours at 37°C, then an additional 4 hours with EdU-free media. After incubation, labelled cells were fixed and prepared for detection using the Click-iT reaction cocktail. Nuclei were labelled using Hoechst 33342. Number of EdU-labelled cells were recorded as a percentage over total number of live nuclei. Imaging and analysis were done using the Opera Phenix HCS and Harmony Analysis Software.

RNA isolation and sequencing

RNA from 2 technical replicates was extracted using 2 clones from each participant (total: 4 samples per participant). TRIzol (Thermo Fischer) method was used, replacing chloroform with 1-bromo-3-chloropropane (BCP; Sigma). To remove genomic DNA during processing, turbo DNase (Thermo Fischer) was used. RNA concentration was quantified using Ribogreen assay (Invitrogen).

Starting with 500ng of total RNA, poly(A) containing mRNA was purified and libraries were prepared using TruSeq Stranded mRNA kit (Illumina). Unstranded libraries were constructed and underwent 50bp single ended sequencing on an Illumina HiSeq 2500 machine. To analyse iPSC mRNA-seq data, the raw reads were mapped to the human genome GRCh37.75 (UCSC version hg19) using STAR: RNA-seq aligner[8]. Aligned reads were sorted using samtools[9], while biases were removed using Picard tools (Broad Institute). Quality control was performed using Picard tools (Broad Institute) and QoRTs [10]. Gene expression levels were quantified using an union exon model with HTSeq [11], which uses uniquely aligned reads. Only the genes with >10 reads and expressed in 80% of the samples, were kept. The resulting read counts were log2 transformed and GC content, gene length, and library size normalised using the cqn package [12] in R.

mRNA weighted co-expression network analysis

Co-expression network analysis was performed using the R library, WGCNA[13]. We wanted to investigate autism-specific iPSC-neuronal culture co-expressed genes (or modules). Biweighted mid-correlations were calculated for all pairs of genes, then a signed similarity matrix was created. In the signed network, the similarity between genes reflects the sign of the correlation of their expression profiles. The signed similarity matrix was then raised to power β to emphasize strong correlations on an exponential scale. The resulting matrix (known as adjacency matrix) was then transformed into a topological overlap matrix. Since we are primarily interested in exploring co-expressed genes conserved across our cohort, we created consensus networks correlated to autism as previously published [14, 15]. After scaling for each individual network (consensus scaling quantile = 0.2), a soft thresholding power of 14 was chosen (as it was the smallest threshold that resulted in a scale-free R^2 fit of 0.8) (Fig S5). The consensus network was created by using a topological overlap matrix (TOM) to calculate the component-wise minimum values for topological overlap. Using dissTOM = 1 – TOM as distance measure, genes were hierarchically clustered. Modules were then assigned using a dynamic tree-cutting algorithm (cutreeHybrid, using default parameters except deepSplit = 4, cutHeight = 0.999, minModulesize = 100, dthresh=0.1 and pamStage = FALSE).

Resulting modules of co-expressed genes were used to calculate module eigengenes (MEs; or 1st principal component of the module). MEs were correlated to biological traits, in this case autism, to find disease-specific modules. Module hubs were defined by calculating module membership (kME) values which are the Pearson correlation between each gene and corresponding ME, and genes with kME < 0.7 were removed from the module. Network visualisation was done using iGraph package in R [16]. Differentially expressed genes, and gene module assignments in **Supplementary Table S6**.

Enrichment analysis for gene sets

Two types of gene set enrichments were performed. For autism-correlated module enrichment, logistic regression was performed using already published gene modules [14, 15, 17] to control for gene length and gene expression level. A two-sided Fisher exact test with 95% confidence interval was performed for cell-type enrichment analysis using published human brain dataset [18].

Module genes were characterised using GO Elite (version 1.2.5) [19] using total expressed genes as background. GO Elite uses a Z-score approximation of hypergeometric distribution to assess term enrichment, and removes redundant GO or KEGG terms to give a concise output. 10,000 permutations were used, and required at least 10 genes to be enriched in a given pathway at a Z-score of at least 2. Only biological process and molecular function categories are reported. Pipeline schematic in **Supplementary Figure S2**.

Supplementary Results

Participant overview

Participants were recruited from the Longitudinal European Autism Project (LEAP)[17]; Brain and Body Genetic Recourse Exchange (BBGRE) studies[18]; or the Social Communication Disorders Clinic at Great Ormond Street Institute of Child Health (GOS-ICH). Of the autistic participants, eight were male and one was female (**Supplementary Table S1**). The four participants from the LEAP cohort were diagnosed with non-syndromic, while participants from BBRGE and GOS-ICH cohorts were diagnosed with syndromic autism (**Supplementary Table S2**). Syndromic participants from GOS-ICI had deletions type CNVs in the 1p21.3 and 8q21.12 regions, with *DYPD* and *PTBP2* and the *AXL* genes of note in each region respectively.

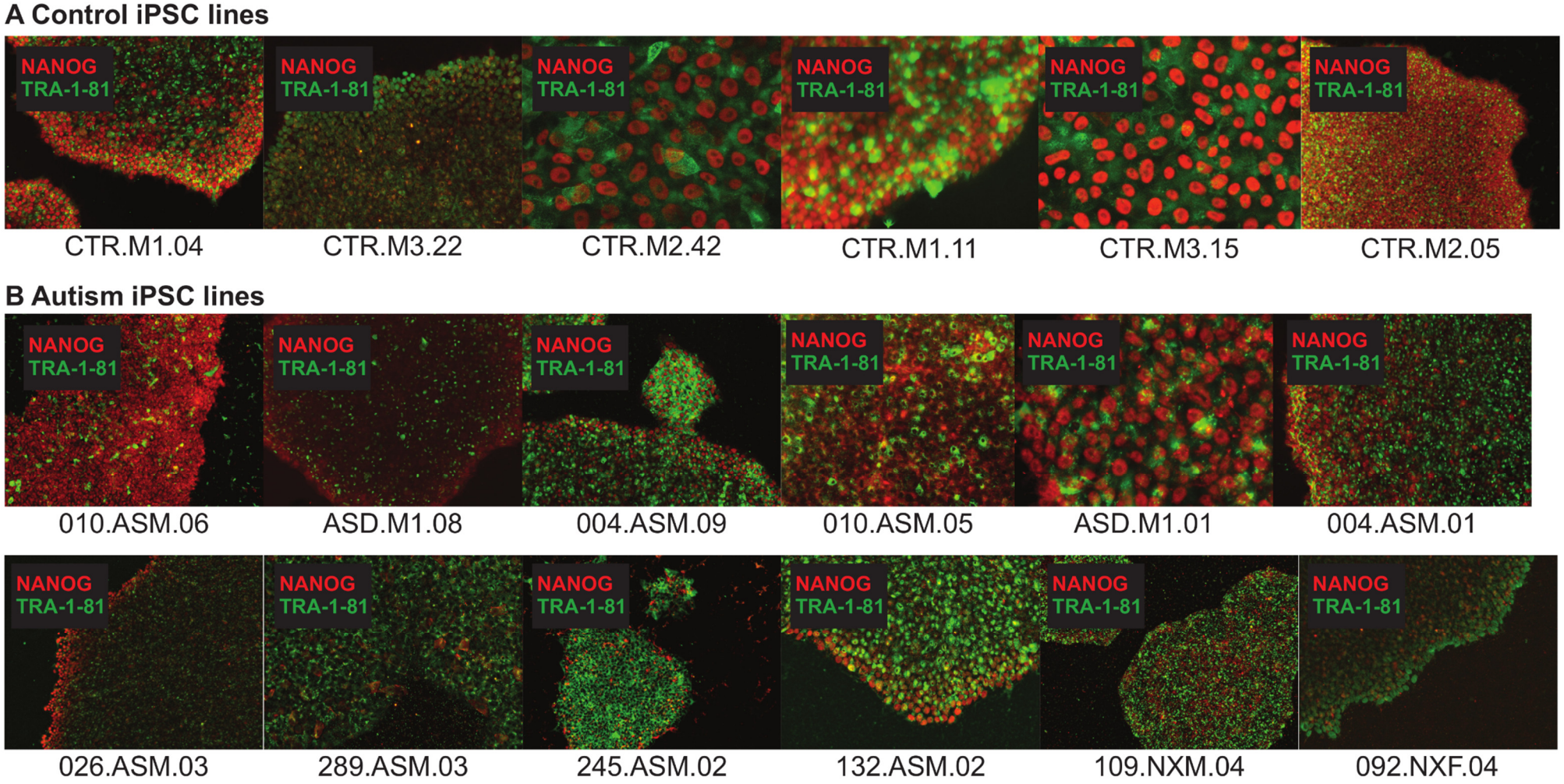
Of the syndromic participants from BBGRE, two syndromic participants had deletion type CNVs in the 2p16.3 region (*NRXN1*), while the third carried a duplication in the 3p chromosomal region[19] (**Supplementary Table S3**).

Transcriptomic analysis of iPSCs reveal enrichment of gene modules associated with autism

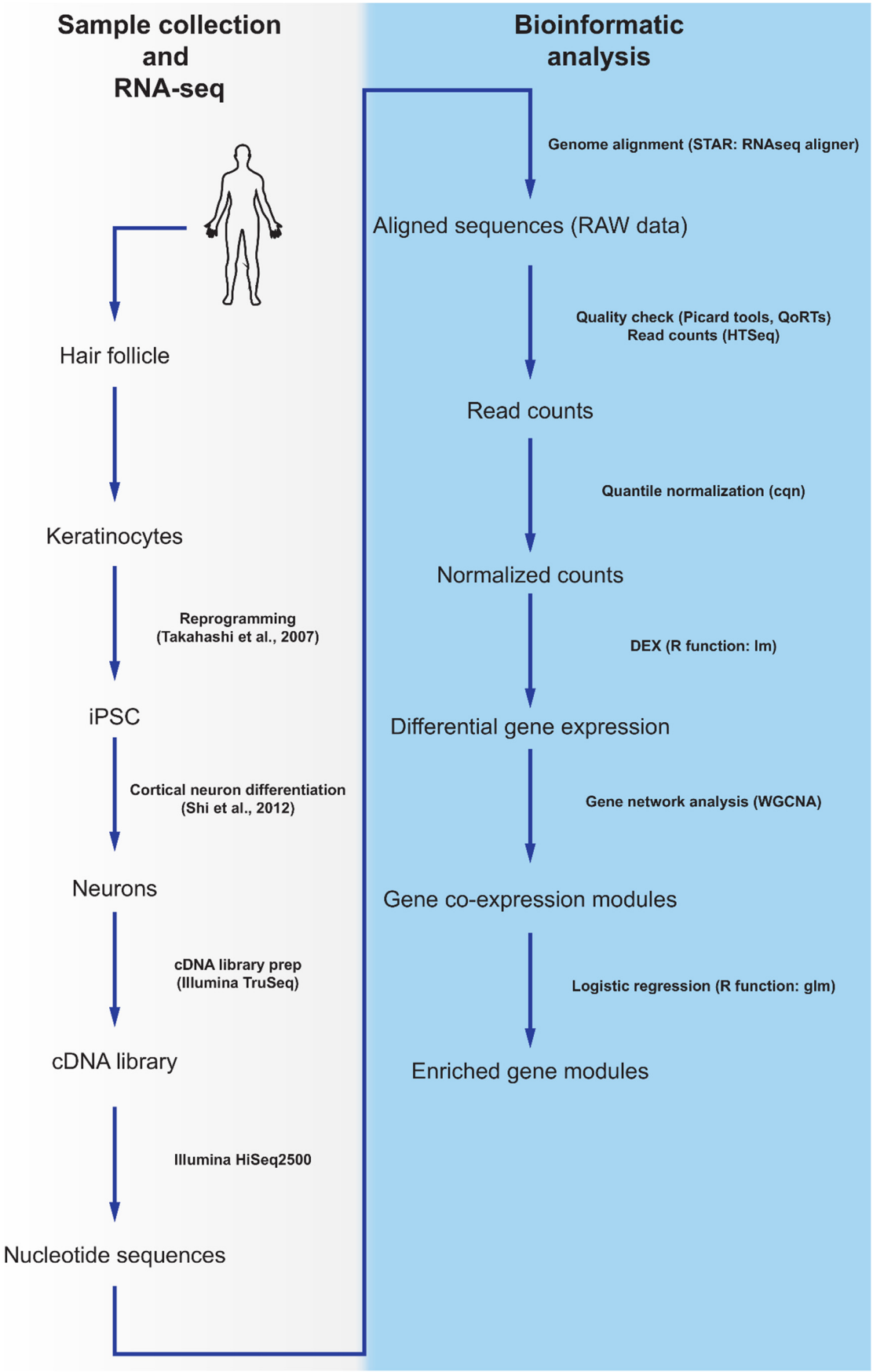
As *post-mortem* brain studies of adult brains have identified prenatal gene expression pathways as being altered in autism[6, 21, 22], we were interested in determining if we could observe similar altered gene expression networks in our cohort of iPSCs. Day 35 neurons were generated from three control and three non-syndromic autism-iPSCs. We chose participants with no familial history of autism or known deletions in autism-associated genes to reduce genetic bias that could drive atypical gene expression. Using an adapted bioinformatics pipeline[6, 22], we analysed gene expression pathways and assess its relatedness to autism (**Supplementary Figure S2**). Principle component analysis revealed distinct separation between the control- and autism-iPSC neurons (**Supplementary Figure S3A**). Differential gene expression (DEG) and hierarchical clustering grouped control- and autism-iPSC neurons on different branches (**Supplementary Figure S3B, Figure 1C; Supplementary Table S6**). Weighted gene co-expression analysis (WGCNA) revealed 11 gene modules significantly altered in autism-iPSC neurons (**Supplementary Figure S3C; Supplementary Table S6**). The three most upregulated and three most downregulated gene modules were strongly enriched respectively in autism *post-mortem* brain gene modules (**Supplementary Figure S3D**). These gene modules showed little to no enrichment in schizophrenia or cancer gene modules (**Supplementary Figure S3E**) indicating that the gene expression patterns in our samples were autism-specific. From this we concluded that altered gene expression in adult autism brains

was also found in prenatal neurons generated from iPSCs, and that gene expression patterns were specific to autism.

Supplementary Figures

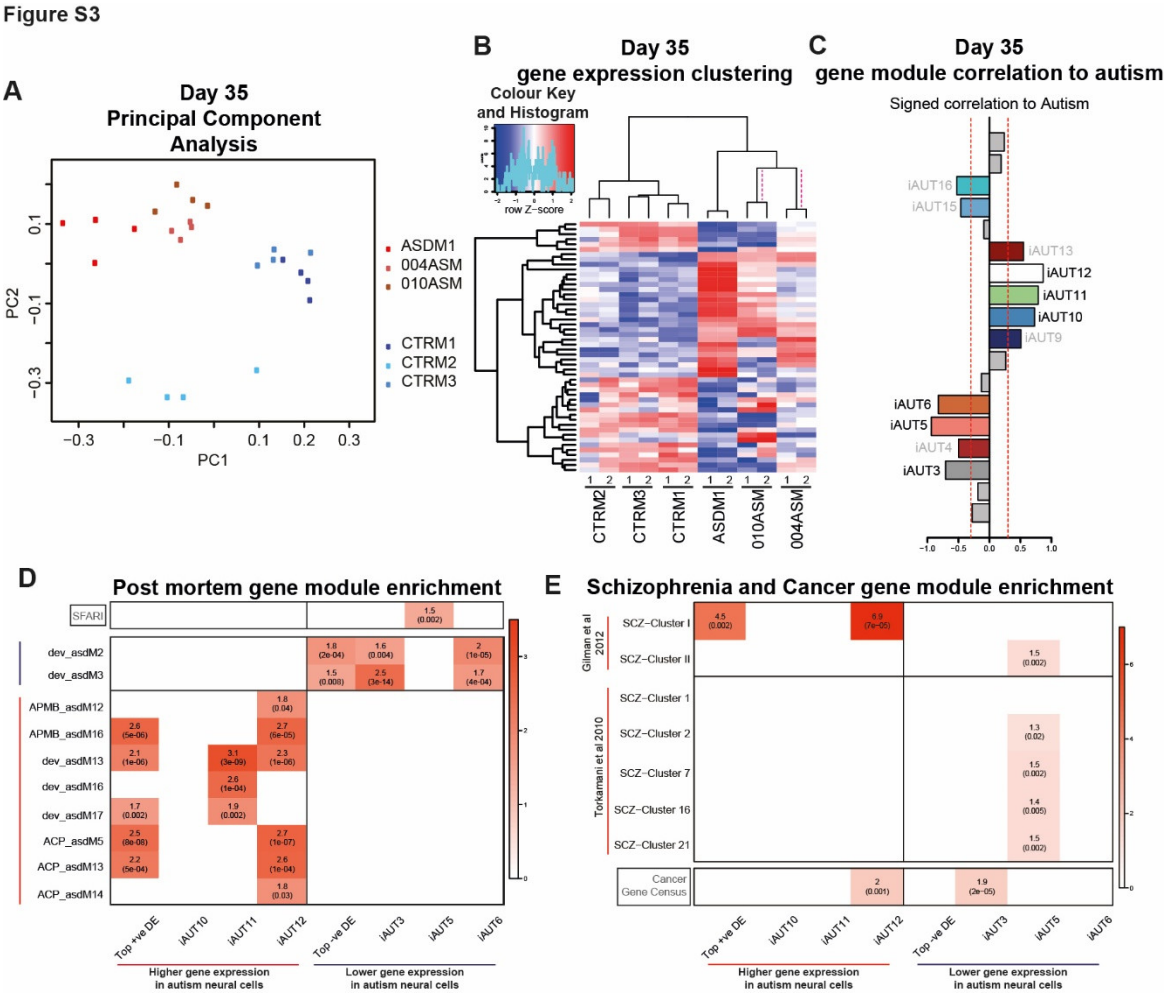


Supplementary Figure S1: Quality control images of iPSC lines. Pluripotency of all iPSC lines were determined by positive staining for stem cell markers NANOG and TRA-1-81.



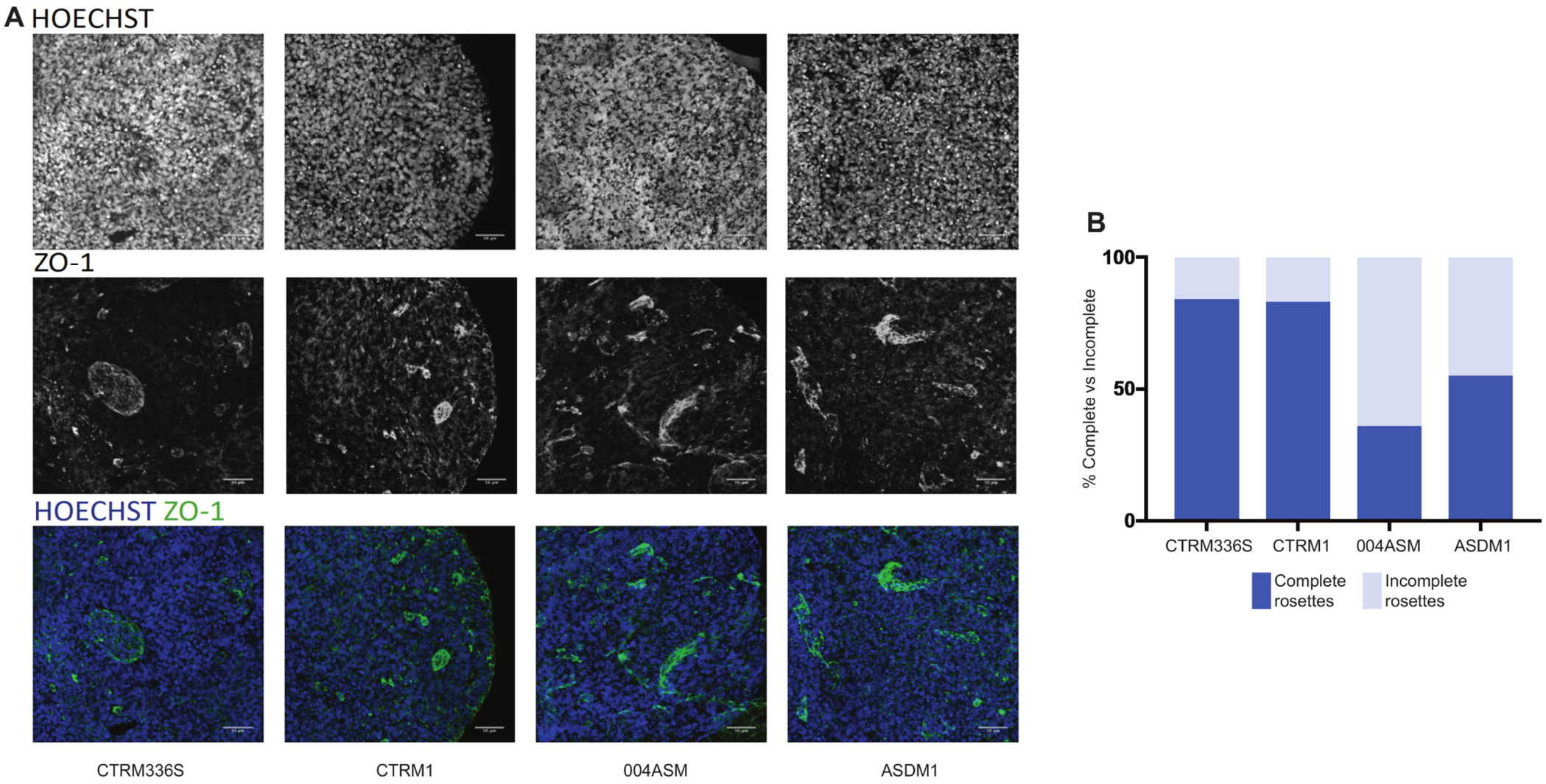
Supplementary Figure S2: Bioinformatics pipeline. Analysis pipeline for RNASeq.

Figure S3

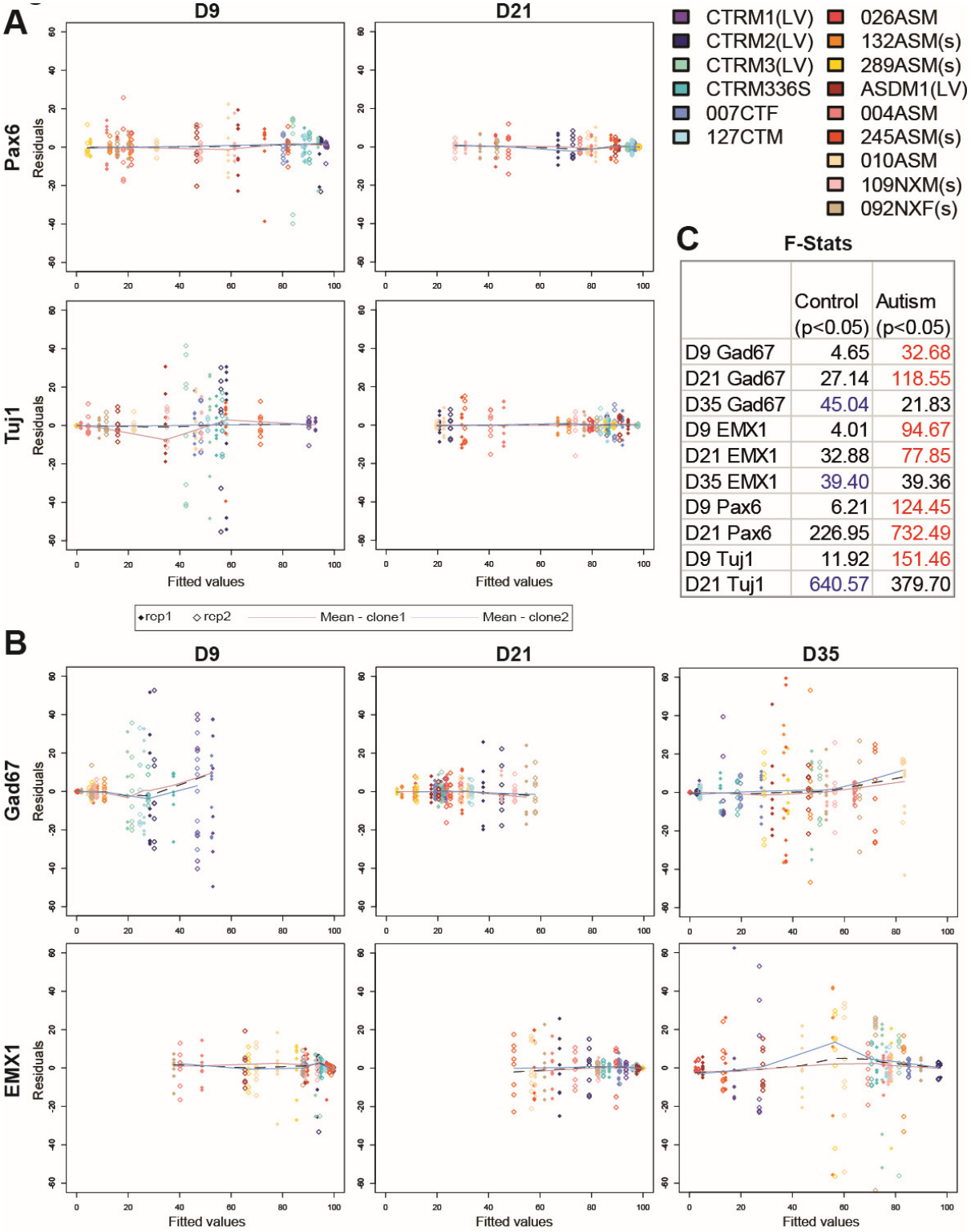


Supplementary Figure S3: RNA-sequencing analysis of day 35 cortical neurons. (A) Principal component analysis showing distinct transcriptomic profiles of autism and typical day35 cells. (B) Gene count from iPSC biological replicates (1, 2) of individual participants and clustering based on Z-scores of top 50 differentially expressed genes. Clustering of 010ASM, 004ASM indicated using magenta dashed lines. (C) WGCNA reveals 11 gene modules significantly correlated to autism (top 3 positively correlated and top 3 negatively correlated modules enrichment shown; greyed module enrichment not shown). (D) Gene module enrichment reveals positively correlated (red) modules are enriched in corresponding positively correlated *post-mortem* gene modules, while negatively correlated (blue) modules

are enriched in negatively correlated *post-mortem* gene modules. (E) Gene modules do not show sufficient enrichment in *post-mortem* gene modules from schizophrenia studies or cancer gene sets.



Supplementary Figure S4: Day 7 3D spheroids from contro-l and autism-iPSCs demonstrate deficiency in neural rosette formation. (A) Day 30 spheroids were immunostained for ZO1 (green) and Hoescht (Blue). **(B)** Percent complete and incomplete rosettes.



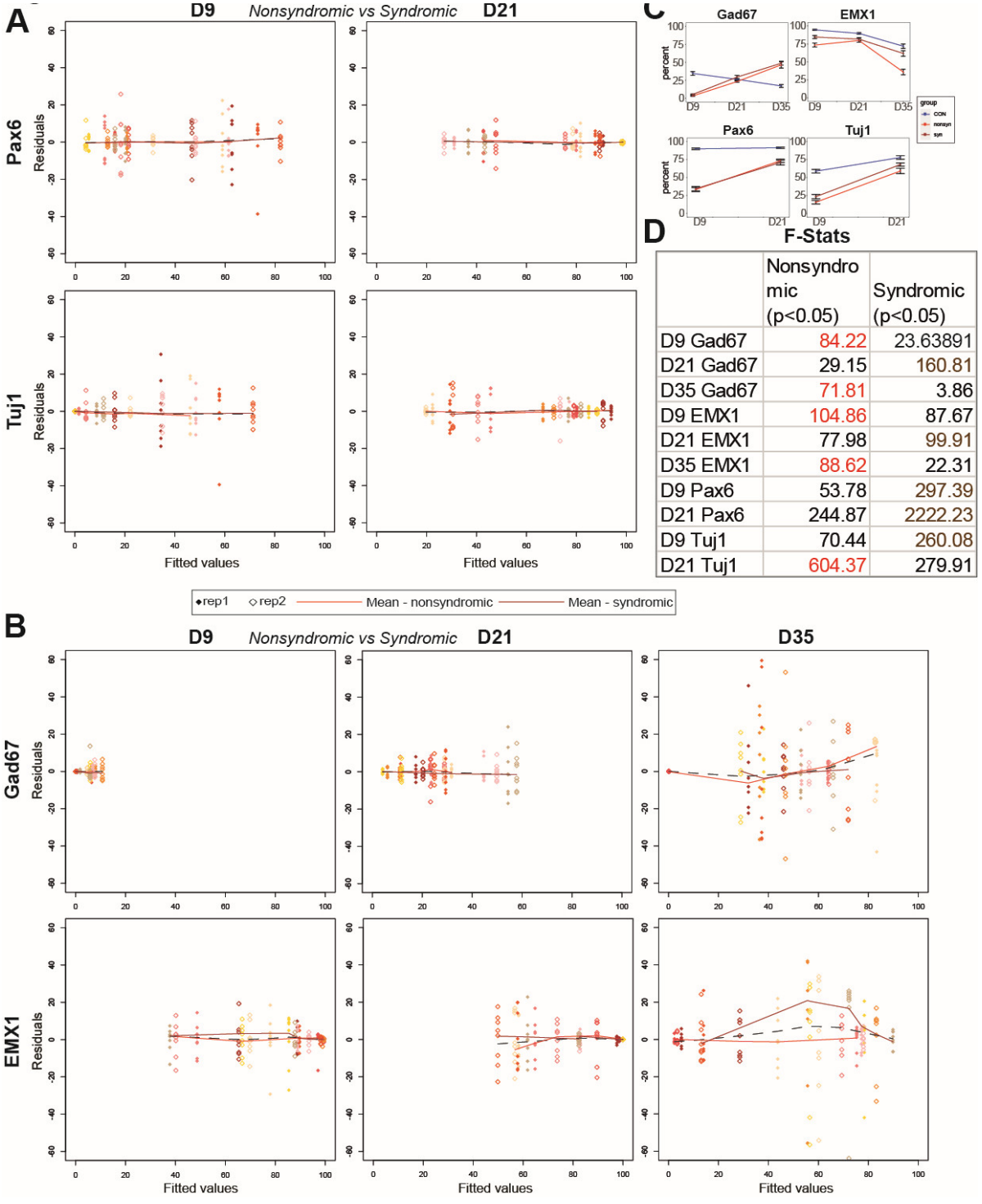
C

F-Stats

	Control (p<0.05)	Autism (p<0.05)
D9 Gad67	4.65	32.68
D21 Gad67	27.14	118.55
D35 Gad67	45.04	21.83
D9 EMX1	4.01	94.67
D21 EMX1	32.88	77.85
D35 EMX1	39.40	39.36
D9 Pax6	6.21	124.45
D21 Pax6	226.95	732.49
D9 Tuj1	11.92	151.46
D21 Tuj1	640.57	379.70

Supplementary Figure S5: ANOVA plots to demonstrate changes in variance of data points due to clones from individual participants. (A) Two-way ANOVA (Lines~Cells+Clones) Residuals vs Fitted values plots when observing cortical differentiation. **(B)** Two-way ANOVA (Lines~Cells+Clones) Residuals vs Fitted values plots demonstrate

spread of values of individual data points across all samples when observing dorsal vs ventral forebrain differentiation. **(C)** F-values show degree of variance within the control and autism groups. All parameters measured show significant variance ($p < 0.05$) across both groups, and 7 out of 10 parameters show greater variance in the autism group.



	Nonsyndromic (p<0.05)	Syndromic (p<0.05)
D9 Gad67	84.22	23.63891
D21 Gad67	29.15	160.81
D35 Gad67	71.81	3.86
D9 EMX1	104.86	87.67
D21 EMX1	77.98	99.91
D35 EMX1	88.62	22.31
D9 Pax6	53.78	297.39
D21 Pax6	244.87	2222.23
D9 Tuj1	70.44	260.08
D21 Tuj1	604.37	279.91

Supplementary Figure S6: ANOVA plots to demonstrate changes in variance of data points due to type of diagnosis (non-syndromic vs syndromic) from individual participants. (A) Two-way ANOVA (Lines~Cells+Syndromic) Residuals vs Fitted values plots when observing non-syndromic vs syndromic cortical differentiation in autistic samples.

(B) Two-way ANOVA (Lines~Cells+Syndromic) Residuals vs Fitted values plots demonstrate spread of values of individual data points across all autistic samples when observing non-syndromic vs syndromic dorsal vs ventral forebrain differentiation. **(C)** Mean values of % positive cells in controls, nonsyndromic and syndromic autism plotted over time. **(D)** F-values show degree of variance within the non-syndromic and syndromic autism groups. All parameters measured show significant variance ($p < 0.05$) across both groups, and each group shows greater variance compared to the other in equal number of parameters, 5 out of 10.

Table S1: General information on participants									
Donor No.	Cohort	Ethics	iPSC ID	Autism Diagnosis	Comorbidity	Head size	Sex	Age (years)	Ethnicity
1	EU-AIMS LEAP	PINDS	ASDM1	Yes		Normal	M	34	Caucasion
2	EU-AIMS LEAP	PINDS	004ASM	Yes		Normal	M	26	Caucasion
3	EU-AIMS LEAP	PINDS	010ASM	Yes	ADHD/Depression	Normal	M	27	Caucasion
4	EU-AIMS LEAP	PINDS	026ASM	Yes		Normal	M	16	Caucasion
5	GOS-ICH	PINDS	289ASM	Yes		Normal	M	19	Caucasion
6	GOS-ICH	PINDS	132ASM	Yes	ADHD	Normal	M	8	Caucasion
7	BBGRE	PINDS	245ASM	Yes		Normal	M	15	Caucasion
8	BBGRE	PINDS	092NXF	Yes		Normal	F	10	Caucasion
9	BBGRE	PINDS	109NXM	Yes		Microcephaly	M	9	Caucasion
1	StemBANCC	PINDS	CTRM1	No		Normal	M	58	Caucasion
2	StemBANCC	PINDS	CTRM2	No		Normal	M	33	Caucasion
3	StemBANCC	PINDS	CTRM3	No		Normal	M	37	Caucasion
4	StemBANCC	PINDS	CTRM336S	No		Normal	M	37	Caucasion
5	EU-AIMS LEAP	PINDS	127CTM	No		Normal	M	56	Caucasion
6	EU-AIMS LEAP	PINDS	007CTF	No		Normal	F	27	Caucasion

Table S2: Diagnosis details of autistic participants												
Donor no.	Cohort	iPSC ID	Autism Diagnosis	ADOS				ADI-R				
				Social Interaction	Communication	Imagination	Stereotypic behaviour	PD_B	PD_C	PD_CS	PD_D	BDI
1	EU-AIMS LEAP	ASDM1	Yes	4	4	1	3					22
2	EU-AIMS LEAP	004ASM	Yes	8	4	1	2	25	19	NA	4	3
3	EU-AIMS LEAP	010ASM	Yes	5	5	1	2	3	3	NA	3	25
4	EU-AIMS LEAP	026ASM	Yes	9	10	NA	9	23	22	NA	1	
5	GOS-ICH	289ASM	Yes; syndromic	3	3	0	1	9	5.7	2.3	1.3	
6	GOS-ICH	132ASM	Yes; syndromic	10	9	NA	9	15.8	14	7.5	5	
7	BBGRE*	245ASM	Yes; syndromic									
8	BBGRE*	092NXF	Yes; syndromic									
9	BBGRE*	109NXM	Yes; syndromic									
*Autism diagnosis was confirmed by referring physicians for BBGRE cohort participants. No further information available.												

Table S3: Genetic data of autistic participants			
Donor no.	Cohort	iPSC ID	Known Deletions/Duplications/SNVs
1	EU-AIMS LEAP	ASDM1	
2	EU-AIMS LEAP	004ASM	
3	EU-AIMS LEAP	010ASM	
4	EU-AIMS LEAP	026ASM	
5	GOS-ICH	289ASM	SNV- Chr19:41759516; C->T; AXL gene; STOP_GAINED; LoF, Synaptic transmission (Schizophrenia). CNV VOUS: 8q21.12 to q21.13 del - 8;79,886,962- 80,149,513.
6	GOS-ICH	132ASM	CNV Clinical abnormality - 1p21.3 del - 1:96,953,361-97,711,563 (758,201bp, DPYD, PTBP2). VOUS: 3q25.33 to q26.1 del- 3:160,629,302- 160,727,203 (97,900bp, PPM1L); 8p11.23 Dup- 8:37,124,969- 37,186,580 (61,611bp); 15q26.1 dup- 15:92,818,328- 92,875,717 (57,388bp); 17q24.1 del- 17:63,288,669- 63,375,080 (86,410bp);
7	BBGRE	245ASM	Complete duplication of CNTN6; paternal x3 chr3:4,354,703-4,532,449 (partail dup SETMAR; complete duplication of SUMF1)
8	BBGRE	092NXF	Paternally inherited duplication in long arms of chromosome 1 – likely to be benign – 1q21.1 (144,679,874 – 145,747,269) x3. De novo deletion of ~200kb in short arm of chromosome 2 – 2p16.3 (50,806,991 – 51,013,685) x1
9	BBGRE	109NXM	Maternally inherited deletion ~60kb in short arm of chromosome 2 – 2p16.3 (50,888,852 – 50,947,729) x1

	Table S4: Reprogramming method and iPSC validation				
Donor no.	**iPSC ID**	**Reprogramming method**	**hiPSC validation**	**Diagnosis**	**Diagnosis type**
1	ASDM101	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP	Autism	Nonsyndromic
	ASDM108	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP	Autism	Nonsyndromic
2	004ASM01	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
	004ASM09	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
3	010ASM05	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
	010ASM06	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
4	026ASM03	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
	026ASM02	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
5	132ASM02	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
	132ASM01	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
6	289ASM03	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
	289ASM01	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Nonsyndromic
7	245ASM02	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic
	245ASM04	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic

8	109NXM04	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic
	109NXM03	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic
9	092NXF04	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic
	092NXF09	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP	Autism	Syndromic
Controls iPSCs					
1	CTRM104	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
	CTRM111	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
2	CTRM205	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
	CTRM242	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
3	CTRM315	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
	CTRM322	Constitutive Polycistronic Lentivirus Reprogramming Kit*	ICC; cytoSNP		
4	CTRM336 S	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		

	CTRM337S	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		
5	127CTM04	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		
	127CTM10	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		
6	007CTF10	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		
	007CTF01	CytoTune-iPS Sendai Reprogramming Kit	ICC; cytoSNP		

*Note: The constitutive polycistronic lentivirus kit uses a single vector which has a negligible risk of insertional mutagenesis and viral reactivation

Table S5: Fluorescent thresholds for antibodies used in cell type analysis				
Antibody	**Supplier**	**Cat. No.**	**Dilution**	**Fluorescent intensity threshold (a.u.)**
Pax6	BioLegend	901301	1:300	3500
Tuj1	BioLegend	801201	1:500	2500
Emx1	ThermoFisher	PA5-35373	1:100	4000
Gad67	Abcam	ab26116	1:1000	600
TBR1	Abcam	ab31940	1:200	1000
CD44	R&D Systems	MAB7045	1:50	600
LMX1A	Abcam	ab139726	1:100	1000
FOXA2	Invitrogen	701698	1:300	3000

Table S7: Morphological details of Neural Rosettes						
Participant ID	**Average diameter (mm)**	**T-test ($p=1\times10^{-22}$)**	**(mm)**	**Mean rosettes number (%)**	**T-test ($p=6\times10^{-11}$)**	**(%)**
CTRM1	0.070855781	**Mean (control)**	0.0742113	1.714263246	**Mean (control)**	1.87176
CTRM2	0.070123668			0.980711145		
CTRM3	0.090845642			1.785184347		
CTRM336S	0.066494935			0.65878026		
007CTF	0.07566226			2.981324875		
127CTM	0.071285499			3.110279078		
026ASM	0	**Mean (autism)**	0.04042801	0	**Mean (autism)**	6.52623
132ASM	0.053815315			11.14960454		
289ASM	0.049887732			4.940428375		
ASDM1	0.055942962			5.255250823		
004ASM	0			0		
245ASM	0.057726738			3.7669185		
010ASM	0.070457893			2.453844441		
109NXM	0.055274512			12.55877341		
092NXF	0.060892681			5.558788098		

Supplementary Table S8: Variability in data points due to clones. Data points showing significant variation due to clones highlighted in red.

	Clones F-stat					
	Combined	*p-val*	Control	*p-val*	Autism	*p-val*
D9 Gad67	0.279	*0.598*	0.478	*0.4912*	1.568	*0.2129*
D21 Gad67	13.355	*0.0003*	1.554	*0.216*	14.505	*0.0002*
D35 Gad67	5.435	*0.0207*	4.817	*0.03*	0.634	*0.428*
D9 EMX1	0.269	*0.605*	0.025	*0.8751*	0.291	*0.591*
D21 EMX1	1.635	*0.202*	1.239	*0.2688*	0.663	*0.417*
D35 EMX1	1.095	*0.2965*	0.515	*0.475*	0.603	*0.4388*
D9 Pax6	11.116	*0.001*	0.598	*0.441*	13.65	*0.0003*
D21 Pax6	0.683	*0.41*	1.473	*0.2283*	0.062	*0.8035*
D9 Tuj1	4.19	*0.0419*	0.496	*0.483*	10.69	*0.0014*
D21 Tuj1	0.099	*0.7538*	0.04	*0.842*	0.06	*0.8072*
D35 TBR1	0.091	*0.763*	1.459	*0.231*	1.244	*0.267*

Supplementary References

1. Aasen T, Izpisua Belmonte JC. Isolation and cultivation of human keratinocytes from skin or plucked hair for the generation of induced pluripotent stem cells. *Nat Protoc* 2010; **5**(2)**:** 371-382.

2. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; **131**(5)**:** 861-872.

3. Shi Y, Kirwan P, Livesey FJ. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc* 2012; **7**(10)**:** 1836-1846.

4. Fedele S, Collo G, Behr K, Bischofberger J, Muller S, Kunath T *et al.* Expansion of human midbrain floor plate progenitors from induced pluripotent stem cells increases dopaminergic neuron differentiation potential. *Sci Rep* 2017; **7**(1)**:** 6036.

5. Kriks S, Shim JW, Piao J, Ganat YM, Wakeman DR, Xie Z *et al.* Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* 2011; **480**(7378)**:** 547-551.

6. Pasca AM, Sloan SA, Clarke LE, Tian Y, Makinson CD, Huber N *et al.* Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat Methods* 2015; **12**(7)**:** 671-678.

7. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 2006; **7**(10)**:** R100.

8. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**(1)**:** 15-21.

9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**(16)**:** 2078-2079.

10. Hartley SW, Mullikin JC. QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics* 2015; **16:** 224.

11. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; **31**(2)**:** 166-169.

12. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* 2012; **13**(2)**:** 204-216.

13. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9:** 559.

14. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V *et al.* Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* 2013; **155**(5)**:** 1008-1021.

15. Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ *et al.* Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* 2016.

16. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal, Complex Systems* 2006; **1695**(5)**:** 1-9.

17. Pasca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Pasca AM *et al.* Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nat Med* 2011; **17**(12)**:** 1657-1662.

18. Zhang Y, Sloan SA, Clarke LE, Caneda C, Plaza CA, Blumenthal PD *et al.* Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* 2016; **89**(1)**:** 37-53.

19. Zambon AC, Gaj S, Ho I, Hanspers K, Vranizan K, Evelo CT *et al.* GO-Elite: a flexible solution for pathway and ontology over-representation. *Bioinformatics* 2012; **28**(16)**:** 2209-2210.