Atypical neurogenesis and excitatory-inhibitory progenitor generation in induced pluripotent stem cell (iPSC) from autistic individuals

Dwaipayan Adhya[1,3,*], Vivek Swarup[2,*], Roland Nagy[3], Carole Shum[3], Paulina Nowosiad[3], Kamila Maria Jozwik[4], Irene Lee[5], David Skuse[5], Frances A. Flinter[6], Grainne McAlonan[7], Maria Andreina Mendez[7], Jamie Horder[7], Declan Murphy[7], 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]Behavioural and Brain Sciences Unit, Population Policy Practice Programme, Great Ormond Street Institute of Child Health, University College London, London WC1N 1EH, UK.

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

[7]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.

[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

* Joint first authors

Short title: Atypical neurogenesis in autism iPSC-derived neurons

Key Words: Glutamate, GABA, cortex, corticogenesis, neural progenitor, immune pathways.

Abstract

Autism is a set of neurodevelopmental conditions with a complex genetic basis. Previous induced pluripotent stem cell (iPSC) studies with autistic individuals having macroencephaly have revealed atypical neuronal proliferation and GABA/glutamate imbalance, the latter also being observed in magnetic resonance spectroscopy (MRS) studies. Functional genomics of autism post mortem brain tissue has identified convergent gene expression networks. However, it is not clear whether the established autism phenotypes are observed in the wider autism spectrum. It also not known whether autism-associated *in vivo* gene expression patterns are recapitulated during *in vitro* neural differentiation. To examine this we have generated induced pluripotent stem cells (iPSCs) from a cohort of autistic individuals with heterogeneous backgrounds, which were differentiated into early and late neural precursors, and early neural cells using an *in vitro* model of cortical neurogenesis. We observed atypical neural differentiation of autism iPSCs compared with controls, and dynamic imbalance in GABA/glutamate cell populations over time. RNA-sequencing identified altered gene co-expression networks associated with neural maturation and GABA/glutamate imbalance, and these pathways correlated with pathways in post-mortem brains. Autism neural cells also recapitulated autism post mortem immune pathways, and found CD44, an autism-associated gene, to be predicted as a highly connected gene. In conclusion, our study demonstrates significant differences in neural differentiation between autism and control iPSCs including GABA/glutamate precursor imbalance, and significant preservation of atypical autism-associated gene networks observed in other model systems.

Introduction

Autism spectrum conditions (henceforth autism) are neurodevelopmental in nature, with a heterogeneous genetic background[1-3]. Autism is diagnosed on the basis of impaired social-communication, alongside unusually narrow and repetitive interests and activities[4]. The primary sensory cortex, association and frontal cortex, and parietal-occipital circuits[5, 6], as well as the medial prefrontal cortex, superior temporal sulcus, temporoparietal junction, amygdala, and fusiform gyrus[7, 8] have been shown to be affected in autism. Based on clinical criteria, autism is typically classified into syndromic and non-syndromic forms. Individuals carrying single gene mutations, copy number variations and/or chromosomal abnormalities, in addition to an autism diagnosis are usually classified as 'syndromic'[9]. Non-syndromic autism is characterized by individuals with a primary diagnosis of autism that is not associated with a mutation in a well-known genetic variant[9]. Exome sequencing studies and analysis of copy number variation have revealed hundreds of rare genomic mutations associated with non-syndromic autism[3, 10]. It is difficult to ascertain cellular and molecular mechanisms based solely on the varied genetic mutations found to be associated with autism. However, RNA sequencing of autism post mortem brains has revealed a greater convergence of cellular and molecular mechanisms such as altered synaptogenesis and immune activity associated with the condition[11, 12]. Post mortem brain tissue, however, is a scarce resource and RNA integrity may be susceptible to confounding factors such as anoxic-ischemic changes based on cause of death and post mortem interval[13]. Conditions of storage is also known to have a bearing on RNA integrity[14]. More confounding factors may be introduced in brains of donors with a history of illnesses, seizures and substance abuse[15]. In addition, studying post-mortem brains primarily obtained from adults does not provide any insight into developmental events associated with autism. There is also considerable ethical dilemma associated with human organ donation for research[14]. To tackle these challenges, there has been a shift towards development of induced

pluripotent stem cell (iPSC) models of autism, by reprogramming iPSCs from somatic cells such as skin fibroblasts or hair follicle keratinocytes, then differentiating them into neural cells[16, 17].

Studies using both 2D and 3D iPSC-cultures derived from autistic individuals with macrocephaly, have demonstrated atypical neural differentiation and increased cell proliferation of neural precursor cells (NPCs), and also an imbalance in excitatory (glutamate-producing) and inhibitory (GABA-producing) receptor activity[18, 19], and these cellular effects were found to correlate with enlarged brain size of participants. These observations strengthen the hypothesis that iPSC-based systems can recapitulate cellular phenotypes relevant for disease[18]. In 3D iPSC cultures derived from autistic individuals with macrocephaly, an overproduction of GABAergic neurons has been observed[19], while in the 2D cultures, alterations in excitatory/inhibitory (E/I) neural networks suggest decreased glutamatergic excitation[18]. Critically, there is increasing evidence from magnetic resonance spectroscopy (MRS) studies of autistic individuals, of abnormalities in levels of excitatory glutamate and inhibitory GABA metabolites[20, 21]. These reports appear to demonstrate a common trend consistent across various model systems, of a reduction of glutamate signalling versus GABA signalling[18, 19, 21]. This also opposes an existing hypothesis of GABA/glutamate signalling, which suggested increased glutamate signalling[22], but which was based on the co-occurrence of epilepsy in autism. However, most autistic individuals do not have seizures[23], and epilepsy cannot be explained as a simple consequence of glutamate overproduction. Imbalances in GABA-glutamate neuron markers have also been observed in autism post mortem brains[11, 24]. Two major neuronal phenotypes have been associated with autism so far, *(1)* atypical neural differentiation and cell proliferation, and *(2)* excitatory/inhibitory imbalances in neurons, and post mortem brain RNA sequencing studies have revealed a third non-neuronal phenotype – an

unusually high enrichment of immune pathways. However, not much is known about the mechanisms that underlie the emergence of these three cellular phenotypes.

In this study, we have generated neural cells of a cortical lineage from iPSCs generated from individuals with autism taking advantage of the ability of iPSCs to phenotypically recapitulate *in vivo* developmental processes[25]. We hypothesised that: *(1)* autism and control iPSC-derived neural precursors would show developmental differences, *(2)* there would be an imbalance in precursor pools destined towards glutamatergic vs GABAergic fate, *(3)* gene networks in autism iPSC-derived neural cells would mimic autism-associated gene networks identified in post mortem brain, and *(4)* there would be greater prevalence of non-neuronal/immune pathways in autism. Using a cortical neuron differentiation method, we differentiated iPSCs from our cohort into precursors and neural cells, and found that precursor cells from autism showed a significant delay in expression of neuronal markers. More importantly, we found a dynamic imbalance in GABA/glutamate fate of neural cells from autism iPSCs. We also found autism iPSC-derived neural cells to be enriched for gene networks previously identified in autism post mortem brains, some of which suggested atypical developmental pathways, excitatory/inhibitory imbalance, and immune-related pathways. In our pathway analyses, we also found CD44 to be a highly connected gene in autism associated with the immune system, and higher expression of CD44 in autism neural cells compared to control neural cells.

Study participants, Materials and Methods

Induced pluripotent stem cells

iPSCs (2 clones from each individual) were produced using keratinocytes from plucked hair follicles, from 9 autistic individuals – including six non-syndromic autistic individuals, one individual with 3p deletion syndrome and two autistic individuals with a mutation in the NRXN1 gene, and 3 typically developing individuals[26] (see extended experimental procedures). All autistic individuals had a primary diagnosis of autism. For clinical diagnosis of autistic individuals included in our RNA-sequencing based gene network analyses, see Table S1. Participants were recruited for this study under approval by NHS Research Ethics Committee (REC No 13/LO/1218); informed consent and methods were carried out in accordance to REC No 13/LO/1218.

Neuronal differentiation

We differentiated iPSC lines into cortical neurons using a well-established method based on dual SMAD inhibition; this results in the recapitulation of key hallmarks of corticogenesis and the generation of cortical neurons[25, 27]. iPSCs were differentiated till early neuron stage – day 35 (**Figure 1A**) (see extended experimental procedures).

Immunocytochemistry

Cortical differentiation of autism and control iPSCs were characterised using immunocytochemistry. iPSCs were differentiated till day 8, day 21 and day 35 and tagged with antibodies of appropriate markers associated with each developmental stage (see extended experimental procedures). Nuclei were stained using DAPI, and imaging was performed using a 40× objective on a Lecia SP5 confocal microscope (**Figure 1B**). High throughput imaging was performed at day 8, day 21 and day 35 of cortical differentiation on the Opera Phenix High-Content Screening System (Perkin Elmer), and cell type quantification was performed using the Harmony High Content Imaging and Analysis Software (Perkin Elmer).

Statistics

Quantification was performed using the Harmony High Content Imaging and Analysis Software (Perkin Elmer). Percentage of cells positive for desired marker versus total number of live cells (stained by DAPI) was calculated for every line and every stage. Independent 2-group t-test was used to check if there was significant difference between autism and control (p-value ≤ 0.05). A linear model fit was used to look at trajectory of marker expression from day 8 to day 35. All statistical analysis was performed on R software.

RNA-sequencing

Starting with 500ng of total RNA, poly(A) containing mRNA was purified and libraries were prepared using TruSeq Stranded mRNA kit (Illumina). Unstranded libraries with a mean fragment size of 150bp (range 100-300bp) were constructed, and underwent 50bp single ended sequencing on an Illumina HiSeq 2500 machine. Reads were mapped to the human genome GRCh37.75 (UCSC version hg19) using STAR: RNA-seq aligner[28]. Quality control was performed using Picard tools (Broad Institute) and QoRTs[29]. Gene expression levels were quantified using an union exon model with HTSeq[30].

Differential gene expression (DGE)

DGE analysis was performed using R statistical packages[31] with gene expression levels adjusted for gene length, library size, and G+C content (henceforth referred to as "Normalized FPKM"). A linear mixed effects model framework was used to assess differential expression in $\log_2$(Normalized FPKM). Autism diagnosis was treated as a fixed effect, while also using technical covariates accounting for RNA quality, library preparation, and batch effects as fixed effects in this model.

Weighted gene coexpression network analysis

The R package weighted gene coexpression network analysis (WGCNA)[32] was used to construct coexpression networks as previously shown[11]. Biweight midcorrelation was used to assess correlations between $\log_2$(Normalized FPKM). For module-trait analysis, 1st principal component of each module (eigengene) was related to an autism diagnosis in a linear mixed effects framework as above, replacing the expression values of each gene with the eigengene.

Gene sets

A SFARI autism associated gene set was compiled using the online SFARI gene database, AutDB, using "Gene Score" as shown previously[11]. We obtained dev_asdM2, dev_asdM3, dev_asdM13, dev_asdM16 and dev_asdM17 modules from an independent transcriptome analysis study using RNA-sequencing data from post mortem early developing brains[11]. Modules asdM12 and asdM16 were obtained from an autism post mortem gene expression study[12]. We obtained another three autism-associated modules: ACP_asdM5, dev_asdM13, ACP_asdM14 from an independent gene expression study profiling dysregulated cortical patterning genes in autism post mortem brain[24]. All three studies used WGCNA to identify modules of dysregulated genes in autism.

Gene set overrepresentation analysis

Enrichment analyses were performed either with logistic regression (all enrichments analyses in **Figures 5A, 6B, 6C, S3B**) or Fisher's exact test (cell type enrichment, **Figure 5B**). All GO enrichment analysis to characterize gene modules was performed using GO Elite[33] with 10,000 permutations. Molecular function and biological process terms were used for display purposes.

Protein-protein interaction analysis

Protein-protein interactions (PPI) of enriched modules were studied using DAPPLE web resource (http://www.broadinstitute.org/mpg/dapple/dappleTMP.php) which looks for

connectivity between genes using a large protein-protein interaction (PPI) network. To enable a robust evaluation of genes significantly connected to each other via protein-protein interactions, degree matched permutations were applied, which were controlled for biological and methodological biases in PPI databases used in this analysis[34].

Results

Cortical differentiation in autism iPSC lines diverge from typical development from an early precursor cell stage

Nine autistic individuals and three healthy individuals participated in this study, from whom we generated a total of 12 autism iPSC lines and 6 control iPSC lines. Of the nine autistic participants, eight were male, with one female. Six were diagnosed as having non-syndromic autism, while three were diagnosed with syndromic autism. Two non-syndromic participants had deletion type CNVs in the 1p21.3 and 8q21.12 regions respectively, with DYPD and PTBP2 being autism-associated genes affected in the former, while the latter also having a deletion in the AXL gene (**Supplementary Table S2**). We also detected a stop-gain mutation in the SHANK3 gene of another non-syndromic participant from exome analysis (**Supplementary Table S2**). Of the three syndromic participants, two syndromic participants had deletion type CNVs in the 2p16.3 region, in both affecting the NRXN1 gene, a well-established autism-associated gene, while the third had 3p deletion syndrome (**Supplementary Table S2**). Keratinocytes were extracted from hair follicles from the participants, and reprogrammed into iPSCs using the Yamanaka factors[35]. As autism is known to affect several regions of the cerebral cortex[5-8], the iPSC lines were differentiated into neural cells of a cortical

lineage using a previously described method[25, 27]. Three developmental stages were specifically studied (**Figure 1A**): (1) Day 8: early neural precursors, (2) Day 21: late neural precursors, (3) Day 35: cortical neural cells. Both control and autistic iPSCs efficiently differentiated using this method, producing neural cell expressing cellular markers and exhibiting cellular morphologies typical for each stage of corticogenesis (**Figure 1B)**.

Genomic characteristics of the autistic participants being heterogeneous, we first investigated basic neuronal differentiation markers in the autism and control iPSCs. Based on previous studies from independent cohorts[16, 18, 19], we hypothesised that irrespective of genomic backgrounds, autism iPSCs would demonstrate developmental differences compared to control iPSCs. To this end, we examined the developmental expression profile of Pax6 and Tuj1 in neural precursors in autistic and control iPSCs (**Figure 2A-C**, **Table 1**). Pax6 is a commonly used marker for identifying neural precursors of cortical lineage[36], while Tuj1 is a robust pan-neuronal and precursor marker[37]. As anticipated, control precursor cells on day 8 were highly positive for Pax6 (**Pax6** Control: 93.54545%) and Tuj1 (**Tuj1** Control: 65.17584%), and on day 21 both markers remained highly expressed (**Pax6**; Control: 86.66410%, **Tuj1**; Control: 68.68563%). (**Figure 2A** and **B**). However, in day 8 autism precursor cells, Pax6 and Tuj1 levels were significantly lower when compared to controls (**Pax6**; Control: 93.54545%, Autism: 33.88251%; $p=4\times10^{-59}$. **Tuj1**; Control: 65.17584%, Autism: 19.87218%; $p=1\times10^{-13}$) (**Figure 2A and B**). Remarkably, when we examined Pax6 and Tuj1 positive cells at day 21, we found that the number of autism precursor cells positive for both markers had increased to a level similar to that seen in control precursors. Interestingly, Pax6 levels still remained lower in autism line, but no significant differences in Tuj1 levels between autism and controls were observed (**Pax6**; Control: 86.66410%, Autism: 71.94075%; $p=4\times10^{-7}$. **Tuj1**; Control: 68.68563%, Autism: 64.00949%; p=0.3) (**Figure 2A-C**). These data showed atypical neurodevelopmental, possibly developmental delay, during neural differentiation in autism

iPSCs compared to controls. A heterogeneous genomic background did not seem to appreciably affect this phenotype.

Having established atypical neural differentiation in autism iPSCs, we also found evidence of an E/I imbalance phenotype during neural differentiation. Excitatory and inhibitory neurons are known to originate from different neuroectodermal lineages[38, 39]. We hypothesised that E/I imbalance, being an important cellular phenotype of autism, might be a result of atypical cell fate specification during neural differentiation. We tested this by observing the development of forebrain excitatory versus inhibitory precursors during neural differentiation. We undertook a time-dependant study of Emx1 and Gad67 expression in neural precursors at day 8 and day 21, and neural cells at day 35 of differentiation (**Figure 3A-D**, **Table 1**). Emx1 is a marker for dorsal telencephalon excitatory neurons and their precursors[39-41], while Gad67 is the rate limiting enzyme in the GABA synthesis pathway and a marker for inhibitory neurons and their precursors[42, 43]. At day 8, Emx1 appears to be expressed in majority of controls as well as autism precursors, although Emx1 expression was significantly higher in control than in autism lines (**Emx1**; Control: 95.69082%, Autism: 79.65836%; $p=4\times10^{-11}$) (**Figure 3B, C**). At day 21, Emx1 expression in control precursors appears to slightly reduce compared to day 8, while Emx1 expression in autism precursors appears to remain the same. At this stage both control and autism precursors expressing Emx1 were similar, although expression was significantly higher in control precursors (**Emx1**; Control: 88.5446%, Autism: 80.8861%; p=0.003) (**Figure 3B, C**). At day 35, the Emx1 expression in both control and autism neural cells was reduced compared to day 8 and day 21 precursors, however the reduction was significantly more acute in autism neural cells than in the control neural cells (**Emx1**; Control: 65.83102%, Autism: 50.35212%; p=0.01) (**Figure 3B, C**). The expression of Gad67 over time in both autism and controls follows a very different trajectory compared to Emx1. At day 8, a significant difference in expression of Gad67 was seen between control and autism precursors. A modest

number of control precursors expressed Gad67, while Gad67 expression in autism precursors was negligible (**Gad67**; Control: 33.223989%, Autism: 4.406441%; $p=1\times10^{-8}$) (**Figure 3B, C**). At day 21, average Gad67 expression in control precursors remains almost the same as on day 8. Conversely, Gad67 expression increased significantly between day 8 and 21 across the autism precursors. Thus, at this stage, no difference in Gad67 expression was found between control and autism precursors; control and autism precursors expressed Gad67 at a similar level across all lines (**Gad67**; Control: 28.04423%, Autism: 26.66252%; p=0.55) (**Figure 3B, C**). By day 35, Gad67 expression in autism neural cells overtook Gad67 expression in control neural cells: Gad67 expression was significantly higher in autism neural cells compared with control neural cells (**Gad67**; Control: 20.05228%, Autism: 47.78413%; $p=3\times10^{-9}$) (**Figure 3B, C**). Taken together, our data suggests a time-dependent reversal of forebrain excitatory, versus inhibitory neural differentiation in autism lines at early stages of neural differentiation. Although neural cells at day 35 indicated higher Gad67 expression in autism than controls, it was interesting to note that at day 8, it was the opposite when control precursors showed higher Gad67 expression than autism precursors. This strongly indicates that atypical neuronal cell fate specification may contribute to the pathophysiology of autism. To our knowledge this is the first time that neural differentiation and the GABA/glutamate phenotype has been investigated at an early precursor stage of development, and therefore, gives us a better insight into the developmental origins of cellular phenotypes associated with autism.

Atypical neurodevelopmental and immune pathways revealed in autism iPSC neural cells

Based on our cellular analyses, we established *(1)* a developmental delay during neural differentiation, *(2)* a dynamic GABA/glutamate imbalance, associated with autism. However, there has been criticism of the induced pluripotent stem cell technology, one of the arguments against it being insufficient recapitulation of *in vivo* and adult cellular phenotypes[44], thus,

making it unsuitable for studying many neuropsychiatric conditions. There has also been criticism of the study of non-syndromic autism using an iPSC model system[45], and the presence of a heterogeneous genetic background in non-syndromic autism has been speculated to likely introduce confounds. However, as autism is a neurodevelopmental condition, the iPSC method is more suited as a model system to look at gene pathways and cellular phenotypes at the earliest stages of neural differentiation. Also, the cortical differentiation method that we used in this study has been shown to robustly produce forebrain neural cells which are most affected in autism. Therefore, we looked to extend our findings, by using RNA-sequencing, and sought to develop a bioinformatics pipeline – based on established methods (**Figure 4A**), to investigate gene pathway information in studies with small autism iPSC cohorts. We hypothesised that: *(1)* autism neural cells from our cohort would recapitulate the autism gene pathways discovered in post-mortem brain tissue, *(2)* developmental gene pathways observed in similarly designed autism iPSC studies as ours, would be preserved.

To test this, we performed RNA-sequencing on neural cells from control and autism iPSCs, and based on transcriptome levels and differential gene expression, the control and autism samples were explicitly separated into two distinct clusters (**Figure 4B**, **Supplementary Figure S6C**). To reveal gene expression pathways enriched in autism iPSCs, we undertook signed weighted gene coexpression network analysis (WGCNA) and identified 11 coexpression modules significantly correlated to autism (labelled according to R-assigned colours, e.g., salmon, **Figure 4C**). We ranked the modules according to their module eigengene values (ME, the first principal component of the module) (**Figure 4D**, **Supplementary Figure S1**). Of the 11 modules, 5 modules were positively correlated in autism neural cells, while 6 modules were negatively correlated in autism neural cells. The modules were assigned consensus functions based on gene ontology (GO) terms. The top 3 positively correlated modules having higher MEs in autism – 'steelblue' (Cellular Metabolic Processes), 'lightgreen'

(Neural Development) and 'white' (Immune Activation), and the top 3 negatively correlated modules having lower MEs in autism – 'grey60' (Epigenetic Regulation), 'salmon' (Gene Regulation) and 'sienna3' (Chromosome Organisation) (**Figure 4C, D**), were also functionally most significant.

Of the positively correlated modules (Figure 4E), the 'steelblue' module was enriched for GO terms for metabolic functions associated with atypical cell proliferation (**Figure 4G**). The 'lightgreen' module was enriched for GO terms including regulation of cell-cell adhesion, cognition, calcium mediated signalling and regulation of dendrite maturation associated with neural development. One of the most interconnected genes of this module (also known as 'hub gene', based on correlation to ME) and also an autism associated gene was *GABRA4* – a subunit of the inhibitory GABA-A receptor[46] (**Figure 4H**). The 'white' module was enriched for cytokine binding, regulation of DNA damage response, positive regulation of apoptosis and negative regulation of neuronal death (**Figure 4I**). Of the negatively correlated modules (**Figure 4F**), the 'salmon' module was enriched for RNA methyltransferase activity, epigenetic regulation of gene expression and s-adenosylmethionine-dependant methyltransferase activity (**Figure 4J**). The 'sienna3' module was enriched for nucleic acid binding, regulation of RNA metabolic process and regulation of gene expression (**Figure 4K**), while the 'grey60' module was enriched for regulation of histone H3-K4 methylation, DNA binding and chromosome organisation (**Figure 4L**). *HTR7* ('salmon' module), *ROBO1* ('sienna3' module) and *SLITRK5* ('salmon' module) were autism-associated[i] genes enriched in negatively correlated modules in this study, suggesting a causal link between their dysregulated expression and autism. In addition, we performed protein-protein interaction (PPI) analysis of our gene modules and identified *CD44* – an autism associated gene, as a highly interconnected gene in the positively correlated 'white' (immune activation) module (**Figure 5A**). Incidentally, we found higher levels of day 35 autism neural cells expressing CD44 than controls, while there was negligible

CD44 expression in control neural cells as would be expected at this stage of differentiation (**CD44**; control: 0.3000217%, autism: 10.1067822%; $p=5\times10^{-14}$) (**Supplementary Figure S1E, F**). We used TBR1 expression as a control as day 35 neural cells are generally mostly early neurons expressing TBR1; no differences in cells expressing TBR1 were seen between control and autism neural cells (**TBR1**; control: 62.46833%, autism: 50.07018%; p=0.053) (**Supplementary Figure S1E, F**).

Autism post-mortem gene expression networks are highly enriched in autism iPSC neural cells

Next, we tested preservation of previously reported gene sets associated with autism and expression networks from similarly designed autism post mortem brain studies, in gene networks from our iPSC neural cells. First, we used a set of 155 autism associated candidate genes from a previous study[11] using the Simons Foundation Autism Research Initiative (SFARI) database to identify load of high impact autism associated genes in our gene modules. The SFARI list of autism genes is a database of genes collated according to the type of genetic variations from whole genome sequencing studies, rare genetic mutations and mutations causing syndromic forms of autism. It was first published in 2009[47] and an up-to-date reference for all known associated genes can be found at: https://gene.sfari.org/autdb/HG_Home.do. We mapped the SFARI autism associated genes with our gene networks, and found the SFARI genes to be enriched in the negatively correlated 'salmon' module (p=0.002; odds ratio [OR] = 1.5) (**Figure 5B**). This suggested downregulation of SFARI genes in our autism iPSC-derived neural cells. We then mapped 5 autism-associated developmental gene modules dysregulated in post mortem brains (APMB), from Parikshak et al., (2013) (dev_asdM2, dev_asdM3, dev_asdM13, dev_asdM16, dev_asdM17)[11], shown in Figure 5B. Of these 5 sets, dev_asdM2 and dev_asdM3 represent DNA-binding and transcriptional regulation and were downregulated in autism, while dev_asdM13, dev_asdM16 and dev_asdM17 represent later

phase neuronal functions and development of synaptic structure, and were upregulated in autism. The dev_asdM2 set was enriched in the top downregulated genes ('Top –ve DE', $p = 2\times10^{-4}$; OR = 1.8), as well as the 'grey60' (p = 0.004; OR = 1.6) and the 'sienna3' ($p = 10^{-5}$; OR = 2) modules. The dev_asdM3 set is enriched in the top downregulated genes ('Top –ve DE', p = 0.008; OR = 1.5), and the 'grey60' ($p = 3\times10^{-14}$; OR = 2.5) and 'sienna3' ($p = 4\times10^{-4}$; OR = 1.7) modules. The dev_asdM13 set is enriched in the top upregulated genes ('Top +ve DE', $p = 10^{-6}$; OR = 2.1), and the 'lightgreen' ($p = 3\times10^{-9}$; OR = 3.1) and 'white' ($p = 10^{-6}$; OR = 2.3) modules. The dev_asdM16 set is enriched in the 'lightgreen' module ($p = 10^{-4}$; OR = 2.6), and, the dev_asdM17 set is enriched in the top upregulated genes ('Top +ve DE', p = 0.002; OR = 1.7) and the 'lightgreen' module (p = 0.002; OR = 1.9). We then mapped two gene modules known to be upregulated in the temporal and frontal cortex of the adult autism brain – APMB_asdM12 (a synaptic function module) and APMB_asdM16 (an immune module) from Voineagu et al., (2011) [12] (**Figure 5B**). The APMB_asdM12 module was enriched in the 'white' module (p = 0.04; OR = 1.8), while the APMB_asdM16 was enriched in the top upregulated genes ('Top +ve DE', $p = 5\times10^{-6}$; OR = 2.6) and the 'white' module ($p = 6\times10^{-5}$; OR = 2.7) (**Figure 5B**). Gene sets associated with attenuated cortical patterning or ACP [24] were also mapped (**Figure 5B**), and suggested greater prediction of ACP in autism iPSC neural cells. Neural development and immune activity were two major autism associated cellular pathways that we found to be dysregulated in autism iPSC neural cells from our cohort. Gene expression networks between iPSC neural cells and post mortem brains were also highly preserved as enrichment of positively correlated and negatively correlated modules were mutually exclusive in our analysis (**Figure 5B**).

Gene expression networks from independent autism iPSC studies are moderately preserved in our autism iPSC neural cells

Gene networks identified in two previous autism iPSC studies[18, 19] were then mapped with gene networks in this study. As both studies used neural cells and tissue derived through differentiation of iPSCs, which were similarly designed as our study, we hypothesized that gene modules identified in them would be preserved in equivalent gene modules identified in our study (**Figure 5C**). The 'white', 'sienna3', 'grey60', 'lightgreen' and 'salmon' modules were moderately well preserved in the Mariani et al 2015 study using iPSC-derived cerebral organoids ('minibrains') ($2 < Z_{summary} < 10$; $p < 0.05$) (Fig 3b). While, the 'steelblue', 'lightgreen', 'salmon' and 'sienna3' were moderately preserved in the Marchetto et al 2016 study which used iPSC-derived neural precursors ($2 < Z_{summary} < 10$; $p < 0.05$). Preservation of gene modules with both iPSC studies strongly suggested convergent autism-associated gene networks in iPSC derived neural tissue.

Discussion

iPSCs can be differentiated into cortical neural cells and 3D tissue using methods that mimic corticogenesis, thus making it a powerful tool to study neurodevelopmental conditions. However there have been criticisms of using iPSCs to study autism due to the genetic heterogeneity of the condition. Nevertheless, studies using both iPSC neural cells as well as cerebral organoids from independent cohorts have demonstrated convergent cellular phenotypes relevant for the condition[18, 19]. Evidence of defects in neural development as well as a GABA/glutamate imbalance have been established as critical cellular phenotypes associated with autism. These phenotypes have also been observed in RNA-sequencing data from autism post mortem brains[11, 12], and GABA/glutamate imbalance has been consistently observed in MRS studies of autistic individuals[20, 21]. Interestingly, iPSC studies report a reduction of glutamate signalling versus GABA signalling, which was also true in the MRS

studies[18-21]. In this study, we had access to autistic individuals with a wide spectrum of clinical symptoms. This provided us a unique opportunity to test the hypothesis that the spectrum of autistic behavioural traits may be associated with convergent cellular traits, and thus a common developmental origin. We hypothesised that: *(1)* autism and control iPSC-derived neural precursors would show developmental differences, *(2)* there would be an imbalance in precursor pools destined towards glutamatergic vs GABAergic fate, *(3)* gene networks in autism iPSC-derived neural cells would mimic autism-associated gene networks identified in post mortem brain, and *(4)* there would be greater prevalence of non-neuronal/immune pathways in autism.

Despite a heterogeneous cohort of autistic individuals, we found significant delay in appearance of Pax6 and Tuj1 in early neural precursors from autism iPSC cohort. This demonstrated developmental differences associated with autism – a phenotype that is well established using different model systems and post mortem brains. Interestingly, in our study this was manifested in the form of developmental delay during early neurogenesis of autism iPSCs. This delay, however was not as apparent in the late neural precursor cells, during which autism precursors appear to be expressing these neuron developmental markers at a similar level as in the controls. In support of the GABA/glutamate imbalance theory, we found fewer day 35 autism neural cells expressing EMX1, a forebrain excitatory precursor and neuron marker, compared to neural cells from control lines. Conversely, more day 35 autism neural cells expressed Gad67, a marker for inhibitory precursors and neurons, compared to control neural cells. This was consistent with the prevalent GABA/glutamate or E/I imbalance phenotype observed in many autism studies[18-21]. However, at day 8 we observed the opposite phenotype, with more excitatory precursors and fewer inhibitory precursors in autism than controls, contrary to that observed at day 35. This suggested neuroectoderm cell fate specification abnormalities in autism iPSCs during early development.

First major criticism of the iPSC method has been whether iPSCs can recapitulate *in vivo* phenotypes and thus be a suitable cellular model for human diseases. However, we found enrichment of adult autism associated gene expression pathways in our autism iPSC neural cells. Synaptic function, vesicular transport, and neuronal projection, as well as, immune and inflammatory responses were adult autism pathways enriched in autism iPSC neural cells. There was also enrichment of gene pathways associated with attenuated cortical patterning (ACP), which suggests that typical patterns of transcriptional differences between different brain regions may be reduced in autism[24]. There was nevertheless considerable enrichment of autism-associated developmental pathways such as those involving synaptic plasticity, synaptic structure, and synaptic maturation genes. The second criticism of the iPSC method is with regards to its use in the study of a complex neuropsychiatric conditions with a heterogeneous genetic background, such as autism. However, upon further investigation we found high to moderate preservation of our gene modules identified with gene modules identified in independent autism iPSC studies using unrelated cohorts of participants[18, 19]. Both these aforementioned studies were designed slightly differently, with one differentiating autism iPSCs into neural precursors while the other differentiating them into cerebral organoids, and although the aim of both studies was to look at autism neural differentiation, different differentiation protocols can activate slightly different transcriptional pathways based on the chemical composition of growth media and factors they are exposed to. Be that as it may, it was still intriguing to observe strong enrichment of our gene expression pathways in autism post mortem brains as well as autism iPSC studies, thus providing validation of the iPSC system as suitable means to model a neurodevelopmental condition such as autism. In addition, atypical neural differentiation in autism was demonstrated by unusually high number of CD44 expressing cells at day 35, a gene that we also predicted to have a high number of protein-protein interactions in our autism cohort.

There has been a suggestion that the Wnt signalling pathway is dysregulated in autism, and that might be responsible for atypical proliferation of precursors in autism iPSCs[18]. Future studies will reveal if stabilising the Wnt pathway at an early stage of differentiation can recover some of the proliferative phenotypes and consequently salvage the autism-associated forebrain precursor fates and GABA/glutamate identities observed in our study. One possibility is that the proliferation and differentiation abnormalities detected in previous studies linked with certain specific comorbidities of autism such as macroencephaly[18, 19], is prevalent throughout the autism spectrum, and is more a result of atypical precursor cell fate determination at early neuroectodermal stages of brain development. Further studies on the nature of neuroectodermal cell fates in autism can explain the two major cellular phenotypes illustrated in this study, and provide basis for exploration of therapeutic interventions.

In summary, we undertook cellular and gene expression studies on iPSCs generated from our heterogeneous cohort of autistic individuals to test primarily two prevalent hypotheses associated with the condition: *(1)* developmental differences in autism neurons, *(2)* imbalance in GABA/glutamate. We differentiated iPSCs into neural precursors and neural cells and found that autism neural precursors demonstrate atypical neural differentiation, while both neural precursors as well as neural cells showed a dynamic GABA/glutamate cellular fate. We also discovered an immune component in our autism neural cells which was consistent with immune response pathways previously observed in autism post mortem brains. Our data supports the hypothesis that proliferation/differentiation abnormalities might be leading to these cellular phenotypes, and we further believe this might not be restricted to individuals demonstrating macroencephaly, but prevalent throughout the autism spectrum, due to the atypical differentiation of the neuroectoderm during early stages of brain development.

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 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), and Seed funding from Medical Research Council, UK (MR/N026063/1) awarded to DPS; 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); the European Union's Seventh Framework Programme (FP7-HEALTH-603016) (DPS, JP); the Mortimer D Sackler Foundation; the Autism Research Trust, the Chinese University of Hong Kong, 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.

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. Characterisation of iPSC from individuals with and without autism. (A) Schematic of iPSC generation process from keratinocytes, followed by cortical differentiation into neural cells. Early neural precursors (day 8), late neural precursors (day 21) and neural cells (day 35) were imaged. (B) Immunofluorescence staining to show morphological changes during development of autism and control iPSC-derived neural cells. Confirmation of $Ki67^+$ and $Nestin^+$ early neural precursor (day 8) (scale bar: 10µm), $Pax6^+$ late neural precursor (day 21) (scale bar: 10µm), and $TBR1^+$ and $MAP2^+$ neural cells (day 35) (scale bar: 10µm).

Figure 2. Evidence of atypical neural precursor populations in autism. (A) High throughput confocal imaging of iPSC-derived neural precursors from autism and control individuals showing $Pax6^+$ and $Tuj1^+$ cells during day 8 and day 21. (B) Quantification of $Pax6^+$ and $Tuj1^+$ cells shows significant differences between autism and control neural precursors expressing Pax6 and Tuj1. (C) Fitted linear regression line plots demonstrate trends in Pax6 and Tuj1 protein expression in autism and control iPSC-derived neural precursors (day 8, day 21).

Figure 3. Evidence of atypical excitatory-inhibitory neural development in autism. (A) High throughput confocal imaging of iPSC-derived neural cell differentiation from autism and control individuals showing $Emx1^+$ and $Gad67^+$ cells during day 8, day 21 and day 35 of differentiation. (B) Quantification of $Emx1^+$ and $Gad67^+$ cells shows significant differences between autism and control neural precursors (day 8, day 21) and neural cells (day 35) expressing Gad67 and EMX1. (C) Fitted linear regression line plots demonstrate trends in Gad67 and EMX1 protein expression in autism and control iPSC-derived precursors (day 8, day 21) and neural cells (day 35).

Figure 4. Transcriptome-wide gene co-expression network analysis in autism and control neurons. (A) Schematic of RNAseq experiments and analyses. (B) Gene expression in control and autism iPSC neural cells (day 35). Top 50 differentially expressed genes shown here. (C) Signed association of mRNA module eigengenes with autism. Modules with positive values indicate increased expression in autism iPSC-derived neural cells, while modules with negative values indicate decreased expression in autism iPSC-derived neural cells. Red dotted lines indicate Benjamini-Hochberg corrected p-values ($p<0.05$). R package designated module colour names are used. (D) Correlation network dendogram with consensus function of significant gene modules and their relationship with autism (red denotes a positive relationship). (E) Module eigengene values of positively correlated gene modules. (F) Module eigengene values of negatively correlated gene modules. (G) Coexpression network plot and GO term enrichment of the 'steelblue' module. Top 25 genes are shown. (H) Coexpression network plot and GO term enrichment of the 'lightgreen' module. Top 25 genes are shown. (I) Coexpression network plot and GO term enrichment of the 'white' module. Top 25 genes are shown. (J) Coexpression network plot and GO term enrichment of the 'salmon' module. Top 25 genes are shown. (K) Coexpression network plot and GO term enrichment of the 'sienna3' module. Top 25 genes are shown. (L) Coexpression network plot and GO term enrichment of the 'grey60' module. Top 25 genes are shown.

Figure 5. Predicted CD44 protein protein interaction in 'white' module, and autism-associated gene modules enrichment analysis. (A) Predicted protein-protein interaction for genes in the 'white' module having connections with CD44, the most interconnected gene in our analyses. (B) Enrichment of autism post mortem gene modules and SFARI autism associated gene list in gene modules from autism iPSC-derived neural cells. Odds ratio after logistic regression

have been shown. Only OR>1.5 has been shown (p-value in parenthesis). (C) Module preservation of gene modules from autism ‘minibrain’ and autism iPSC-derived NPCs in gene modules from this study.

Supplementary Info

1. Supplementary methods
2. Supplementary figure legends
3. Supplementary Figure S1
4. Supplementary Figure S2
5. Supplementary Figure S3
6. Supplementary Figure S4
7. Supplementary Figure S5
8. Supplementary Figure S6
9. Supplementary Figure S7
10. Supplementary Figure S8
11. Supplementary Figure S9
12. Supplementary table S1
13. Supplementary table S2

Keywords:

Autism, iPSC, precursors, neural cells, cortical differentiation, neurodevelopment, GABA-glutamate imbalance, post mortem brain, transcriptome, functional genomics, molecular pathways

References

1. Berg JM, Geschwind DH. Autism genetics: searching for specificity and convergence. *Genome Biol* 2012; **13**(7)**:** 247.

2. Bourgeron T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat Rev Neurosci* 2015; **16**(9)**:** 551-563.

3. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012; **485**(7397)**:** 246-250.

4. APA. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub, 2013.

5. Minshew NJ, Keller TA. The nature of brain dysfunction in autism: functional brain imaging studies. *Curr Opin Neurol* 2010; **23**(2)**:** 124-130.

6. Samson F, Mottron L, Soulieres I, Zeffiro TA. Enhanced visual functioning in autism: an ALE meta-analysis. *Hum Brain Mapp* 2012; **33**(7)**:** 1553-1581.

7. Dichter GS. Functional magnetic resonance imaging of autism spectrum disorders. *Dialogues Clin Neurosci* 2012; **14**(3)**:** 319-351.

8. Philip RC, Dauvermann MR, Whalley HC, Baynham K, Lawrie SM, Stanfield AC. A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neurosci Biobehav Rev* 2012; **36**(2)**:** 901-942.

9. Sztainberg Y, Zoghbi HY. Lessons learned from studying syndromic autism spectrum disorders. *Nat Neurosci* 2016; **19**(11)**:** 1408-1417.

10. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 2011; **43**(6)**:** 585-589.

11. 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.

12. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 2011; **474**(7351)**:** 380-384.

13. Lewis DA. The human brain revisited: opportunities and challenges in postmortem studies of psychiatric disorders. *Neuropsychopharmacology* 2002; **26**(2)**:** 143-154.

14. Kretzschmar H. Brain banking: opportunities, challenges and meaning for the future. *Nat Rev Neurosci* 2009; **10**(1)**:** 70-78.

15. Woolfenden S, Sarkozy V, Ridley G, Coory M, Williams K. A systematic review of two outcomes in autism spectrum disorder - epilepsy and mortality. *Dev Med Child Neurol* 2012; **54**(4)**:** 306-312.

16. Marchetto MC, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 2010; **143**(4)**:** 527-539.

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

19. Mariani J, Coppola G, Zhang P, Abyzov A, Provini L, Tomasini L *et al.* FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell* 2015; **162**(2)**:** 375-390.

20. Ajram LA, Horder J, Mendez MA, Galanopoulos A, Brennan LP, Wichers RH *et al.* Shifting brain inhibitory balance and connectivity of the prefrontal cortex of adults with autism spectrum disorder. *Transl Psychiatry* 2017; **7**(5)**:** e1137.

21. Horder J, Petrinovic MM, Mendez MA, Bruns A, Takumi T, Spooren W *et al.* Glutamate and GABA in autism spectrum disorder-a translational magnetic resonance spectroscopy study in man and rodent models. *Transl Psychiatry* 2018; **8**(1)**:** 106.

22. Rubenstein JL, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav* 2003; **2**(5)**:** 255-267.

23. Bolton PF, Carcani-Rathwell I, Hutton J, Goode S, Howlin P, Rutter M. Epilepsy in autism: features and correlates. *Br J Psychiatry* 2011; **198**(4)**:** 289-294.

24. 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.

25. 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.

26. 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.

27. Deans PJM, Raval P, Sellers KJ, Gatford NJF, Halai S, Duarte RRR *et al.* Psychosis Risk Candidate ZNF804A Localizes to Synapses and Regulates Neurite Formation and Dendritic Spine Structure. *Biol Psychiatry* 2017; **82**(1)**:** 49-61.

28. 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.

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

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

31. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

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

33. 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.

34. Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet* 2011; **7**(1)**:** e1001273.

35. 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.

36. Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W *et al.* Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* 2015; **518**(7539)**:** 355-359.

37. Tischfield MA, Baris HN, Wu C, Rudolph G, Van Maldergem L, He W *et al.* Human TUBB3 mutations perturb microtubule dynamics, kinesin interactions, and axon guidance. *Cell* 2010; **140**(1)**:** 74-87.

38. Corbin JG, Rutlin M, Gaiano N, Fishell G. Combinatorial function of the homeodomain proteins Nkx2.1 and Gsh2 in ventral telencephalic patterning. *Development* 2003; **130**(20)**:** 4895-4906.

39. Costa MR, Muller U. Specification of excitatory neurons in the developing cerebral cortex: progenitor diversity and environmental influences. *Front Cell Neurosci* 2014; **8:** 449.

40. Zhang W, Peterson M, Beyer B, Frankel WN, Zhang ZW. Loss of MeCP2 from forebrain excitatory neurons leads to cortical hyperexcitation and seizures. *J Neurosci* 2014; **34**(7)**:** 2754-2763.

41. Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JL, Jones KR. Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. *J Neurosci* 2002; **22**(15)**:** 6309-6314.

42. Lazarus MS, Krishnan K, Huang ZJ. GAD67 deficiency in parvalbumin interneurons produces deficits in inhibitory transmission and network disinhibition in mouse prefrontal cortex. *Cereb Cortex* 2015; **25**(5)**:** 1290-1296.

43. Azim E, Jabaudon D, Fame RM, Macklis JD. SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development. *Nat Neurosci* 2009; **12**(10)**:** 1238-1247.

44. Vadodaria KC, Amatya DN, Marchetto MC, Gage FH. Modeling psychiatric disorders using patient stem cell-derived neurons: a way forward. *Genome Med* 2018; **10**(1)**:** 1.

45. Muotri AR. The Human Model: Changing Focus on Autism Research. *Biol Psychiatry* 2016; **79**(8)**:** 642-649.

46. Roberts DS, Raol YH, Bandyopadhyay S, Lund IV, Budreck EC, Passini MA *et al.* Egr3 stimulation of GABRA4 promoter activity as a mechanism for seizure-induced up-regulation of GABA(A) receptor alpha4 subunit expression. *Proc Natl Acad Sci U S A* 2005; **102**(33)**:** 11894-11899.

47. Basu SN, Kollu R, Banerjee-Basu S. AutDB: a gene reference resource for autism research. *Nucleic Acids Res* 2009; **37**(Database issue)**:** D832-836.

48. 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.

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

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

51. Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol* 2011; **7**(1)**:** e1001057.

52. 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.

53. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**(14)**:** 1754-1760.

54. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc* 2015; **10**(10)**:** 1556-1566.

55. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 2010; **26**(16)**:** 2069-2070.

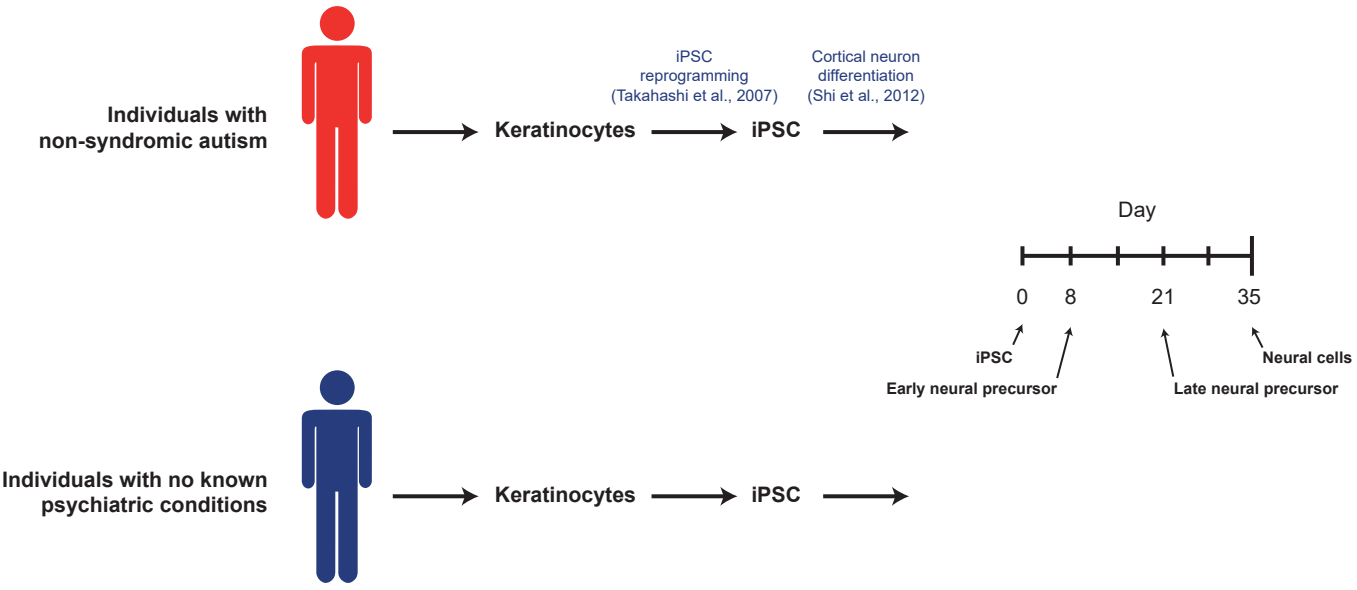
56. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; **31**(13)**:** 3812-3814.

57. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 2002; **30**(17)**:** 3894-3900.

[i] We use the term 'autism associated' genes instead of 'autism-risk' genes because some sections of the autism community have said that the term 'risk' paints a negative view of autism when autism entails disability, differences and even strengths.

Figure 1

A


Individuals with non-syndromic autism
Keratinocytes
iPSC reprogramming (Takahashi et al., 2007)
iPSC
Cortical neuron differentiation (Shi et al., 2012)
Day
0
8
21
35
iPSC
Early neural precursor
Late neural precursor
Neural cells
Individuals with no known psychiatric conditions
Keratinocytes
iPSC

B

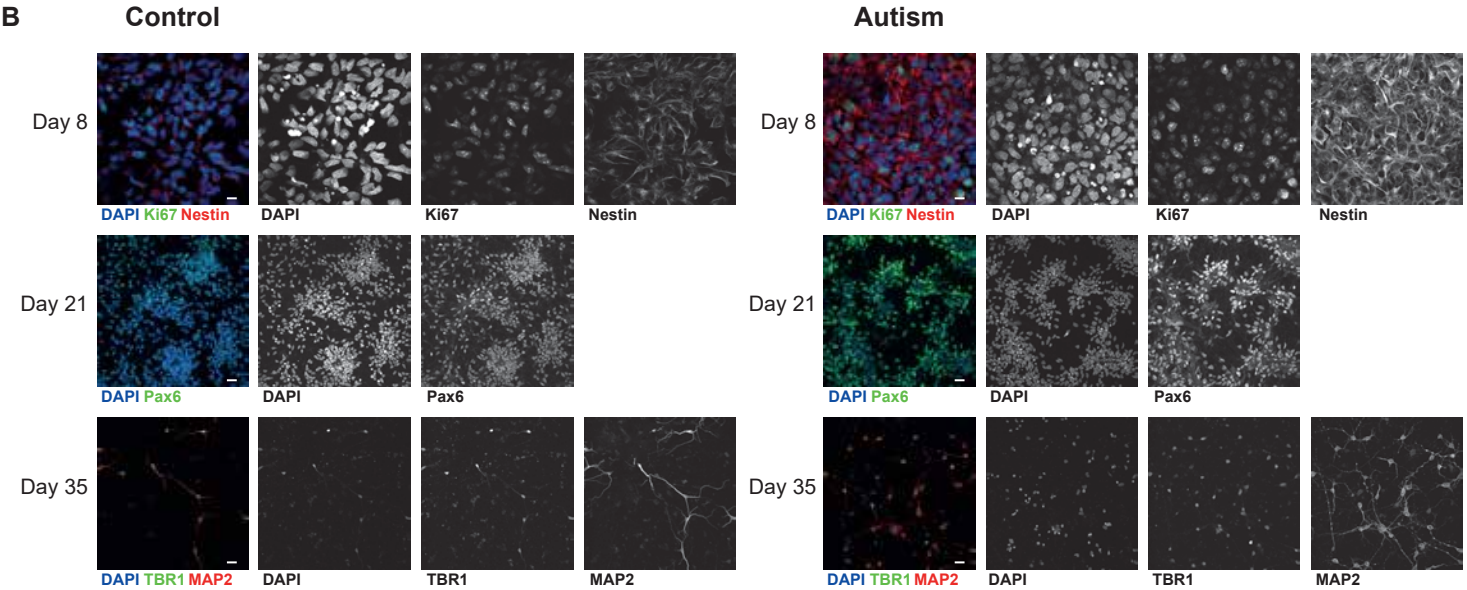

Control
Autism
Day 8
Day 21
Day 35
DAPI Ki67 Nestin
DAPI
Ki67
Nestin
DAPI Pax6
Pax6
DAPI TBR1 MAP2
TBR1
MAP2

Figure 2

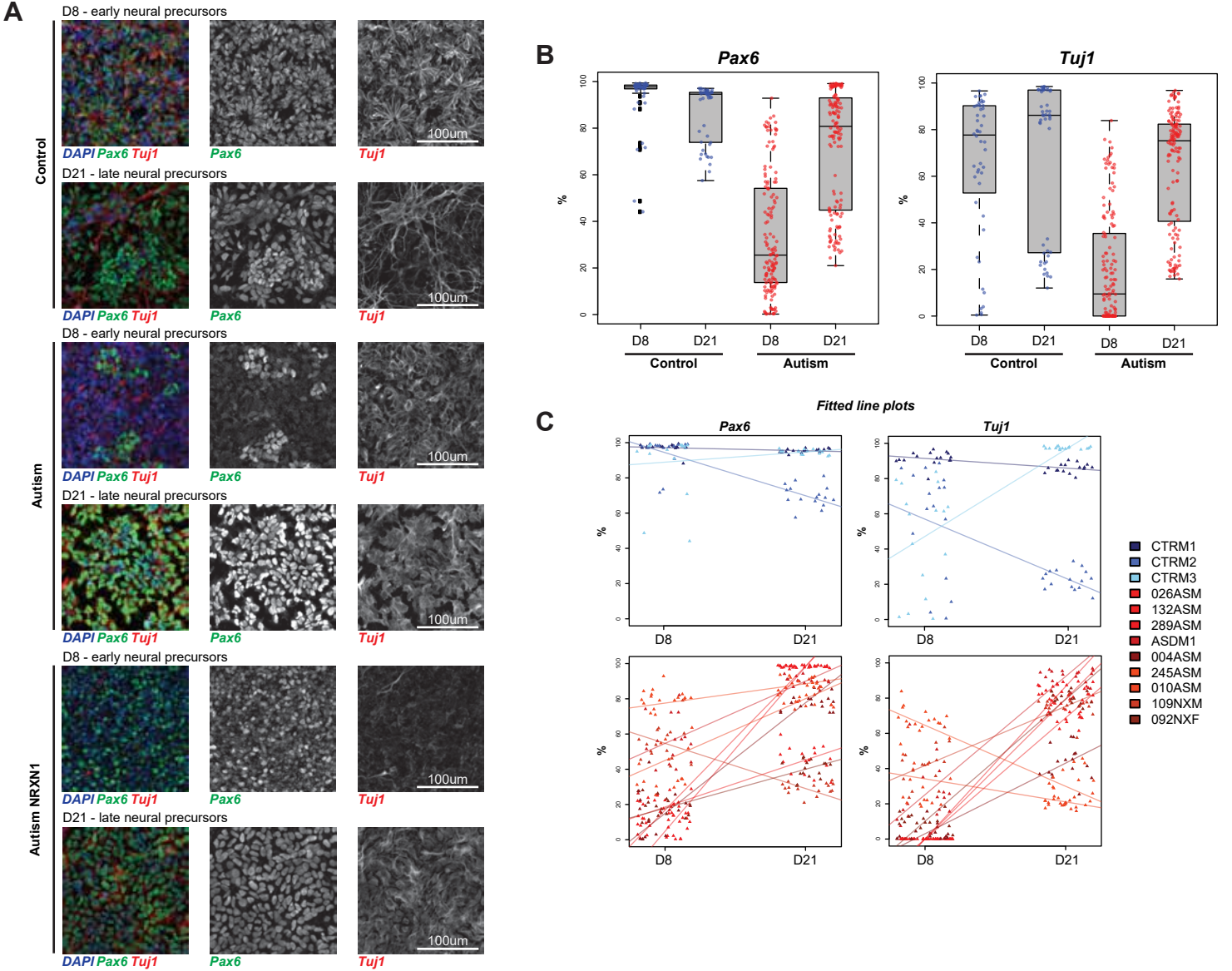

A
Control
D8 - early neural precursors
DAPI Pax6 Tuj1
Pax6
Tuj1
100um
D21 - late neural precursors
Autism
Autism NRXN1
B
Pax6
Tuj1
%
D8
D21
Control
Autism
C
Fitted line plots
CTRM1
CTRM2
CTRM3
026ASM
132ASM
289ASM
ASDM1
004ASM
245ASM
010ASM
109NXM
092NXF

Figure 3

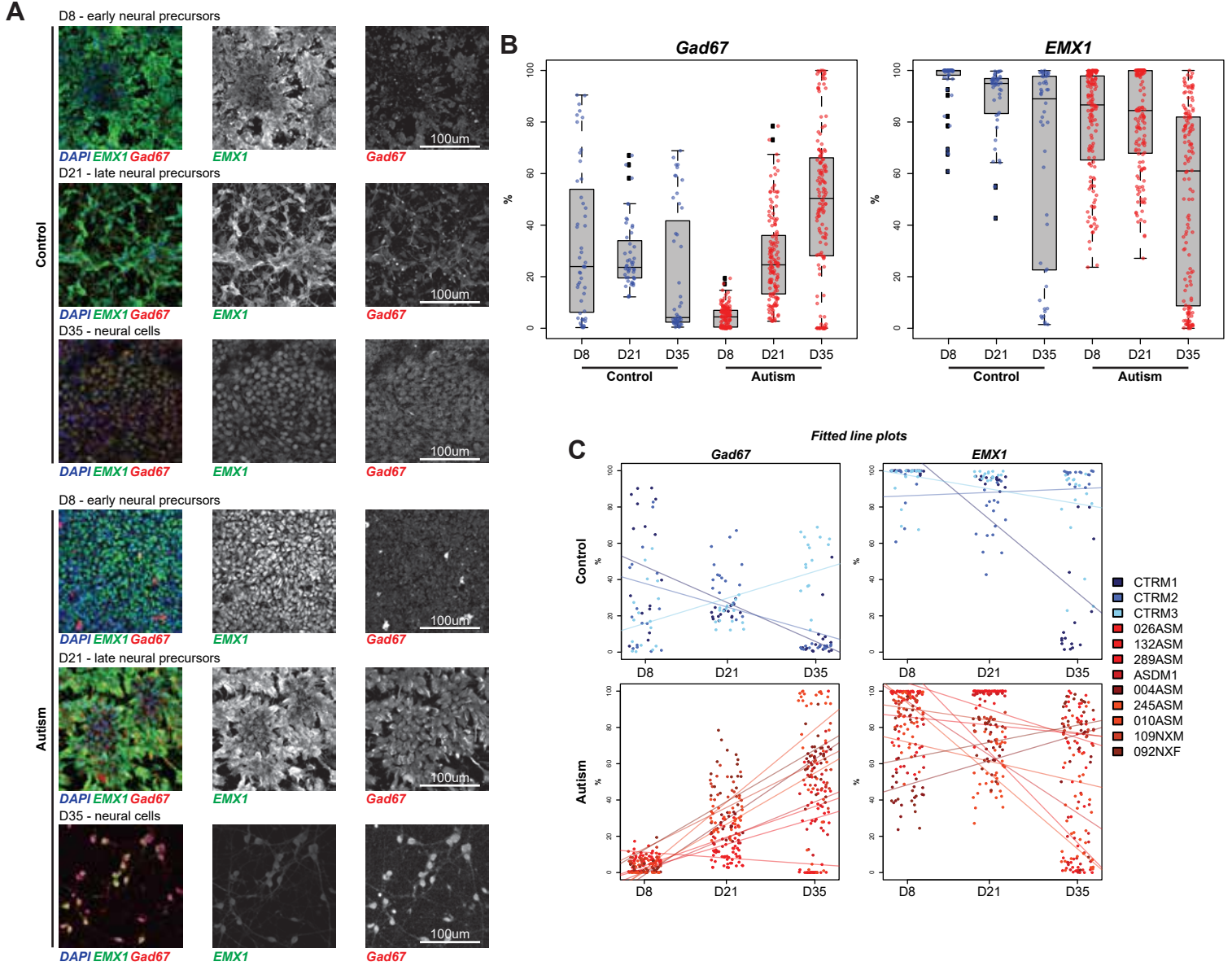

A
Control
D8 - early neural precursors
DAPI EMX1 Gad67
EMX1
Gad67
100um
D21 - late neural precursors
D35 - neural cells
Autism
B
Gad67
EMX1
%
D8
D21
D35
Control
Autism
C
Fitted line plots
CTRM1
CTRM2
CTRM3
026ASM
132ASM
289ASM
ASDM1
004ASM
245ASM
010ASM
109NXM
092NXF

Figure 2

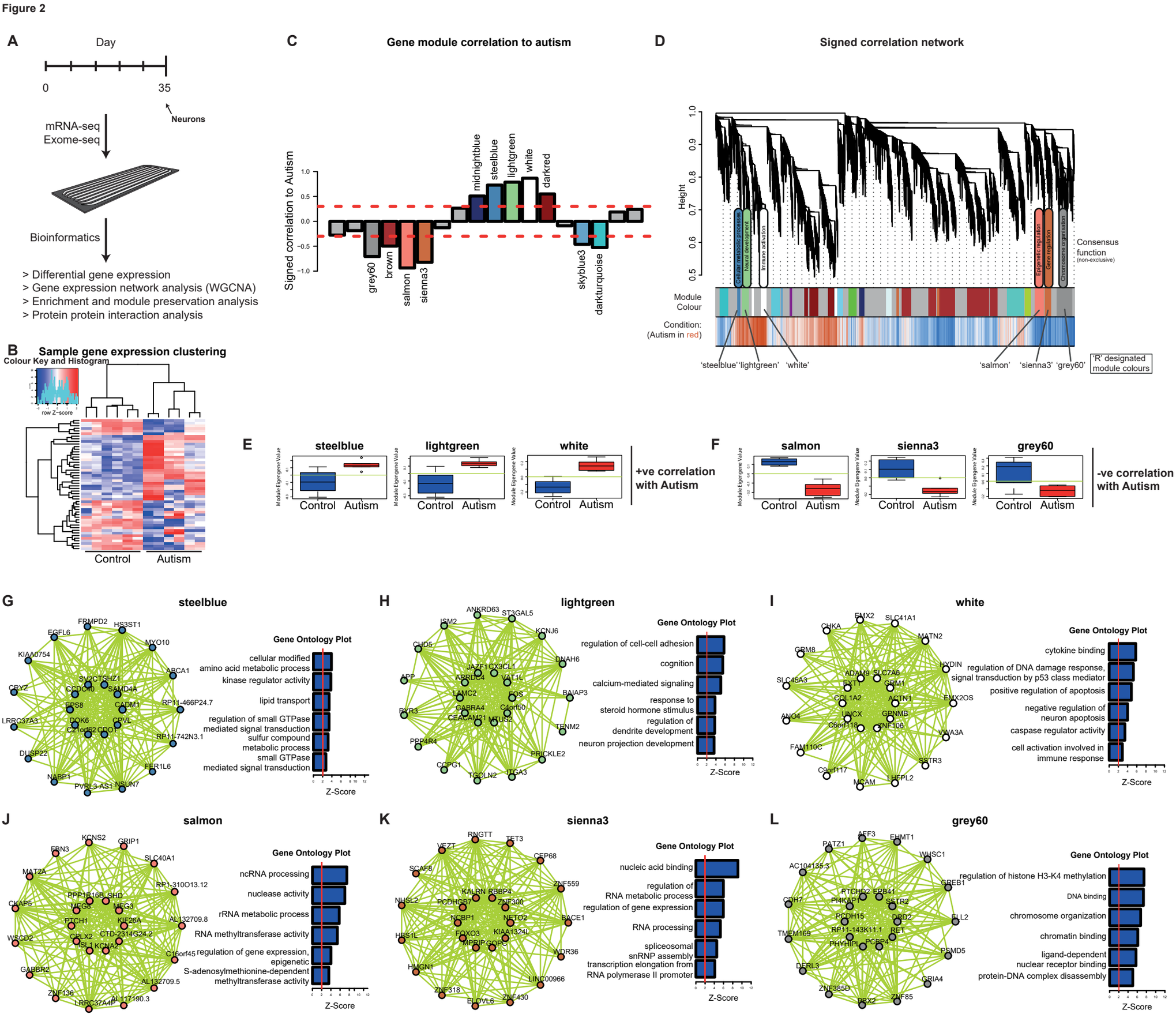

A
Day
0
35
Neurons
mRNA-seq
Exome-seq
Bioinformatics
> Differential gene expression
> Gene expression network analysis (WGCNA)
> Enrichment and module preservation analysis
> Protein protein interaction analysis
B
Sample gene expression clustering
Colour Key and Histogram
row Z-score
Control
Autism
C
Gene module correlation to autism
Signed correlation to Autism
midnightblue
steelblue
lightgreen
white
darkred
grey60
brown
salmon
sienna3
skyblue3
darkturquoise
D
Signed correlation network
Height
Cellular metabolic processes
Neural development
Immune activation
Epigenetic regulation
Gene regulation
Chromosome organisation
Consensus function (non-exclusive)
Module Colour
Condition: (Autism in red)
'steelblue' 'lightgreen' 'white'
'salmon' 'sienna3' 'grey60'
'R' designated module colours
E
steelblue
lightgreen
white
Control
Autism
+ve correlation with Autism
F
salmon
sienna3
grey60
Control
Autism
-ve correlation with Autism
G
steelblue
Gene Ontology Plot
cellular modified amino acid metabolic process
kinase regulator activity
lipid transport
regulation of small GTPase mediated signal transduction
sulfur compound metabolic process
small GTPase mediated signal transduction
Z-Score
H
lightgreen
Gene Ontology Plot
regulation of cell-cell adhesion
cognition
calcium-mediated signaling
response to steroid hormone stimulus
regulation of dendrite development
neuron projection development
Z-Score
I
white
Gene Ontology Plot
cytokine binding
regulation of DNA damage response, signal transduction by p53 class mediator
positive regulation of apoptosis
negative regulation of neuron apoptosis
caspase regulator activity
cell activation involved in immune response
Z-Score
J
salmon
Gene Ontology Plot
ncRNA processing
nuclease activity
rRNA metabolic process
RNA methyltransferase activity
regulation of gene expression, epigenetic
S-adenosylmethionine-dependent methyltransferase activity
Z-Score
K
sienna3
Gene Ontology Plot
nucleic acid binding
regulation of RNA metabolic process
regulation of gene expression
RNA processing
spliceosomal snRNP assembly
transcription elongation from RNA polymerase II promoter
Z-Score
L
grey60
Gene Ontology Plot
regulation of histone H3-K4 methylation
DNA binding
chromosome organization
chromatin binding
ligand-dependent nuclear receptor binding
protein-DNA complex disassembly
Z-Score

Figure 3

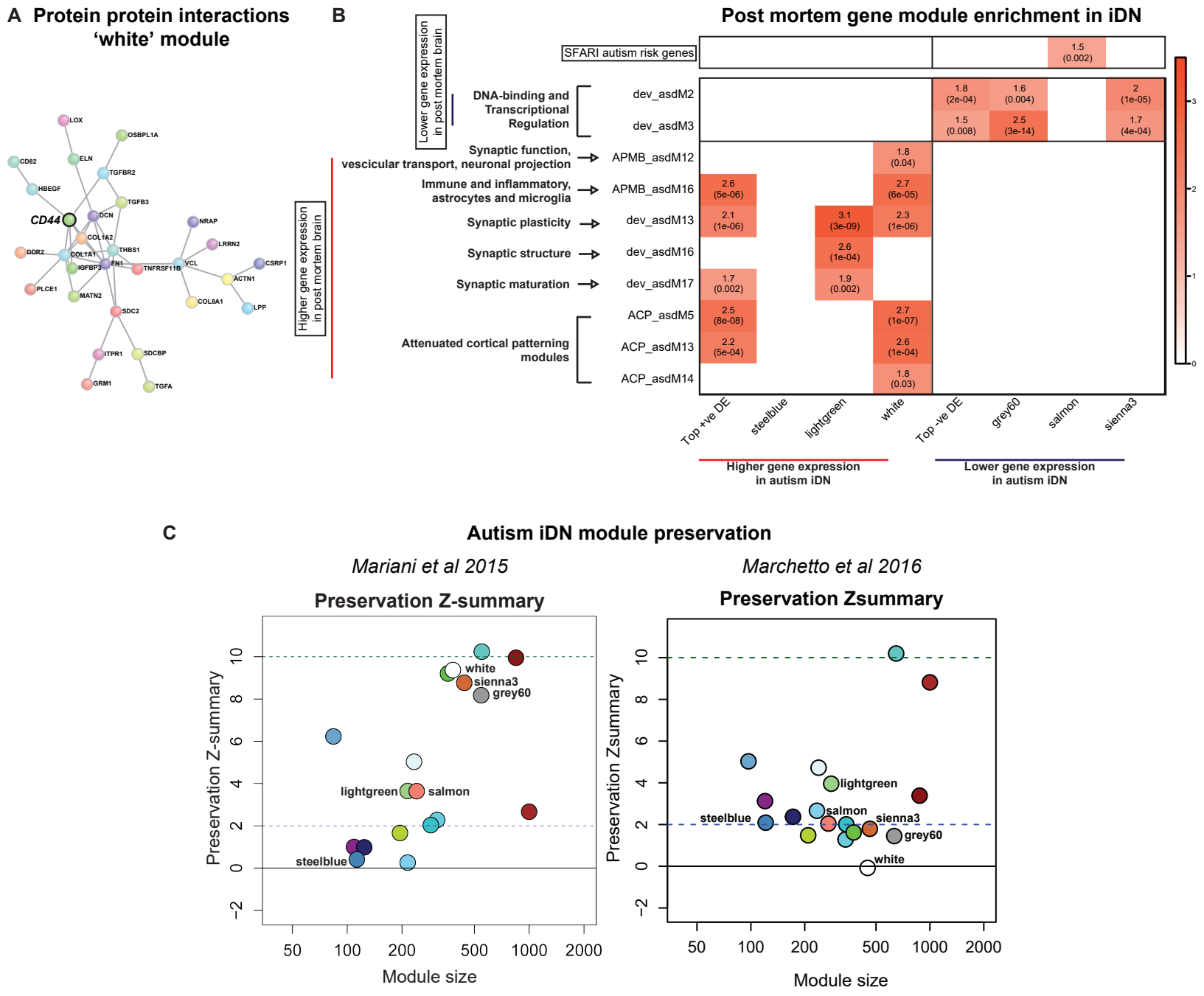

A Protein protein interactions 'white' module
LOX
OSBPL1A
CD82
ELN
TGFBR2
HBEGF
TGFB3
CD44
DCN
NRAP
COL1A2
LRRN2
DDR2
COL1A1
THBS1
CSRP1
IGFBP3
FN1
TNFRSF11B
VCL
ACTN1
PLCE1
MATN2
COL8A1
LPP
SDC2
ITPR1
SDCBP
GRM1
TGFA
Higher gene expression in post mortem brain
B
Post mortem gene module enrichment in iDN
Lower gene expression in post mortem brain
SFARI autism risk genes
1.5 (0.002)
DNA-binding and Transcriptional Regulation
dev_asdM2
1.8 (2e-04)
1.6 (0.004)
2 (1e-05)
dev_asdM3
1.5 (0.008)
2.5 (3e-14)
1.7 (4e-04)
Synaptic function, vescicular transport, neuronal projection
APMB_asdM12
1.8 (0.04)
Immune and inflammatory, astrocytes and microglia
APMB_asdM16
2.6 (5e-06)
2.7 (6e-05)
Synaptic plasticity
dev_asdM13
2.1 (1e-06)
3.1 (3e-09)
2.3 (1e-06)
Synaptic structure
dev_asdM16
2.6 (1e-04)
Synaptic maturation
dev_asdM17
1.7 (0.002)
1.9 (0.002)
Attenuated cortical patterning modules
ACP_asdM5
2.5 (8e-08)
2.7 (1e-07)
ACP_asdM13
2.2 (5e-04)
2.6 (1e-04)
ACP_asdM14
1.8 (0.03)
Top +ve DE
steelblue
lightgreen
white
Top -ve DE
grey60
salmon
sienna3
Higher gene expression in autism iDN
Lower gene expression in autism iDN
C
Autism iDN module preservation
Mariani et al 2015
Preservation Z-summary
Preservation Z-summary
Module size
white
sienna3
grey60
lightgreen
salmon
steelblue
Marchetto et al 2016
Preservation Zsummary
Preservation Zsummary
Module size
lightgreen
steelblue
salmon
sienna3
grey60
white

Table 1: Percent cells expressing neural differentiation markers. Independent 2-group t-test was performed between control and autism values for each time point ($p \leq 0.05$). Pax6 and Tuj1 expression at day 35 was not observed as there are zero Pax6 cells in terminally differentiated neurons, while all terminally differentiated cells of neuronal lineage express Tuj1 (β3-tubulin).

Marker	**Day 8 – early precursors (%)**			**Day 21 – late precursors (%)**			**Day 35 – Neural cells (%)**		
	Control	**Autism**	***p-value***	**Control**	**Autism**	***p-value***	**Control**	**Autism**	***p-value***
Pax6	93.54545	33.88251	4×10^{-59}	86.66410	71.94075	4×10^{-7}	-	-	-
Tuj1	65.17584	19.87218	1×10^{-13}	68.68563	64.00949	*0.3**	-	-	-
Emx1	95.69082	79.65836	4×10^{-11}	88.5446	80.8861	*0.003*	65.83102	50.35212	*0.01*
Gad67	33.223989	4.406441	1×10^{-8}	28.04423	26.66252	*0.55**	20.05228	47.78413	3×10^{-9}

*Not significant