1-1-2019

Cellular and molecular characterization of multiplex autism in human induced pluripotent stem cell-derived neurons

Emily M A Lewis

Kesavan Meganathan

Dustin Baldridge

Paul Gontarz

Bo Zhang

See next page for additional authors

Authors

Emily M A Lewis, Kesavan Meganathan, Dustin Baldridge, Paul Gontarz, Bo Zhang, Azad Bonni, John N Constantino, and Kristen L Kroll

RESEARCH   **Open Access**

Cellular and molecular characterization of multiplex autism in human induced pluripotent stem cell-derived neurons

Emily M. A. Lewis[1†], Kesavan Meganathan[1†], Dustin Baldridge[2], Paul Gontarz[1], Bo Zhang[1], Azad Bonni[3], John N. Constantino[4] and Kristen L. Kroll[1*]

Abstract

Background: Autism spectrum disorder (ASD) is a neurodevelopmental disorder with pronounced heritability in the general population. This is largely attributable to the effects of polygenic susceptibility, with inherited liability exhibiting distinct sex differences in phenotypic expression. Attempts to model ASD in human cellular systems have principally involved rare de novo mutations associated with ASD phenocopies. However, by definition, these models are not representative of polygenic liability, which accounts for the vast share of population-attributable risk.

Methods: Here, we performed what is, to our knowledge, the first attempt to model multiplex autism using patient-derived induced pluripotent stem cells (iPSCs) in a family manifesting incremental degrees of phenotypic expression of inherited liability (absent, intermediate, severe). The family members share an inherited variant of uncertain significance (VUS) in *GPD2*, a gene that was previously associated with developmental disability but here is insufficient by itself to cause ASD. iPSCs from three first-degree relatives and an unrelated control were differentiated into both cortical excitatory (cExN) and cortical inhibitory (cIN) neurons, and cellular phenotyping and transcriptomic analysis were conducted.

Results: cExN neurospheres from the two affected individuals were reduced in size, compared to those derived from unaffected related and unrelated individuals. This reduction was, at least in part, due to increased apoptosis of cells from affected individuals upon initiation of cExN neural induction. Likewise, cIN neural progenitor cells from affected individuals exhibited increased apoptosis, compared to both unaffected individuals. Transcriptomic analysis of both cExN and cIN neural progenitor cells revealed distinct molecular signatures associated with affectation, including the misregulation of suites of genes associated with neural development, neuronal function, and behavior, as well as altered expression of ASD risk-associated genes.

Conclusions: We have provided evidence of morphological, physiological, and transcriptomic signatures of polygenic liability to ASD from an analysis of cellular models derived from a multiplex autism family. ASD is commonly inherited on the basis of additive genetic liability. Therefore, identifying convergent cellular and molecular phenotypes resulting from polygenic and monogenic susceptibility may provide a critical bridge for determining which of the disparate effects of rare highly deleterious mutations might also apply to common autistic syndromes.

Keywords: Multiplex autism, iPSC modeling, Neurodevelopment, Cortical excitatory neurons, Cortical inhibitory neurons, Transcriptomics, Gene networks

* Correspondence: kkroll@wustl.edu
†Emily M. A. Lewis and Kesavan Meganathan contributed equally to this work.
[1]Department of Developmental Biology, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, USA
Full list of author information is available at the end of the article

Background

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with a complex and poorly understood etiology [1–3]. Behavioral and imaging studies have been valuable for defining deficits in affected individuals and characterizing alterations at the level of the brain. However, we are extremely limited in our ability to acquire or experimentally manipulate human brain tissue from living patients or post-mortem brain slices. This has hampered efforts to study cellular and molecular abnormalities that accompany ASD, both during and after fetal and post-natal development. Notably, both the relative integrity of brain structures in affected individuals and the diversity of ASD genetics suggest that convergent mechanisms that contribute to affectation in ASD may operate at the level of the cell [3, 4]. These may be identifiable in experimental models derived from affected individuals. In particular, ASD appears to frequently involve abnormal development and/or function of two major classes of neurons in the cerebral cortex, glutamatergic excitatory projection neurons (cExNs) and GABAergic inhibitory interneurons (cINs) [3, 5, 6]. In vitro differentiation-based models of these neuronal cell types can identify cellular and molecular deficits associated with ASD and provide a tractable platform to screen for pharmacologic agents that can rescue these deficits.

In recent years, such cellular models of ASD have been generated, either by deriving induced pluripotent stem cell lines (iPSCs) from affected individuals or by using CRISPR/Cas9-based gene editing to engineer ASD-associated mutations into wild-type PSCs [7–19]. Most of these studies have focused on syndromic forms of ASD, or on monogenic, de novo cases, where causality is attributed to the mutation of a single ASD-linked gene [10, 11, 13–16]. These forms of ASD are attractive for cellular modeling, as they streamline study design and reduce many potential confounding variables. Other studies have included individuals with an unknown genetic cause of ASD, but with subject selection based upon a shared phenotypic characteristic, such as macrocephaly [18–20]. Together, these models have been informative, revealing both cellular and molecular alterations associated with affectation. These include shared and model-specific disruptions of gene expression in ASD-derived neurons, frequently involving altered expression of genes in key developmental signaling pathways, and genes that control cellular proliferation and growth [7, 8, 13–20]. In addition, differences were observed in neural precursor cell (NPC) proliferation and differentiation [9, 19], neurogenesis, [9, 11, 18, 20], synaptogenesis [8, 10, 17, 18], or functional neuronal activity [7–9, 11, 16, 17]. Altered expression of ASD genes, in which a mutation is linked to ASD causation or risk, is also frequently observed [7, 8, 13–16, 18–20].

These cellular modeling studies have revealed potential contributors to affectation and, in some cases, have identified targets amenable to pharmacological rescue in vitro [10, 18]. However, they do not encompass the range of contributors to ASD burden in the general population: no single gene mutation accounts for more than 1% of overall ASD cases with predicted monogenic causality [21], while the majority of genetic risk appears to be polygenic or idiopathic [2, 22–28]. Polygenic ASD risk can involve both common and rare variation in protein-coding and non-coding regions of the genome, which may act in a combinatorial manner [2, 27, 29, 30]. Furthermore, ASD exhibits pronounced heritability in families (estimated at 50–90% [31]), none of which can be accounted for by de novo (germline) events. Even within each multiplex ASD family, there is often a considerable range in the extent of affectation among individuals, and in most multiplex autism families, a single causative gene mutation cannot be identified [31].

While ASD burden in the general population predominantly involves polygenic or idiopathic risk, heritability, and variable affectation [2, 21, 27, 29–31], these genetically complex forms of the disorder have been largely neglected in cellular modeling studies. Therefore, we deemed it important to determine whether cellular modeling of these complex but prevalent forms of ASD could also reveal affectation-related deficits. To test this, we focused on a multiplex ASD family with variable affectation among family members. We generated iPSC lines from three first degree relatives (a male and a female with differing degrees of affectation and their unaffected mother), as well as an unrelated unaffected female. We used these lines to perform differentiation into both cExN and cIN neural progenitors and/or neurons. Models from the affected individuals exhibited compromised cellular responses to differentiation cues and had disrupted gene expression profiles. This included altered expression of many ASD-associated genes, genes with roles related to behavior, cognition, and learning, and genes involved in nervous system development and function, including cell adhesion molecules and ion channels.

This is, to our knowledge, the first cellular modeling study of multiplex ASD, including graded affectation among family members. This work demonstrates that even genetically complex forms of ASD have discernable cellular and molecular abnormalities that track with affectation, some of which overlap those identified in prior modeling of syndromic, monogenic, and de novo forms of ASD. Therefore, this novel study design highlights the potential for cellular modeling to identify convergent hallmarks across the broad diversity and genetic complexity of pathways to affectation.

Lewis *et al. Molecular Autism*   (2019) 10:51

Page 3 of 23

Methods

Phenotyping of the multiplex family

The nuclear family consisted of working professional non-consanguineous parents, whose firstborn child was a daughter with DSM-5 autism spectrum disorder (ASD), Level I (requiring support, meeting DSM-5 criteria for Asperger syndrome) who was very high functioning and ultimately attended college, followed by a pair of monozygotic twin boys with ASD, Level III—one more fluently verbal than the other but both severely impaired and requiring very substantial support (see below)—followed by a third son with very subtle autistic traits and predominantly affected by attention deficit hyperactivity disorder, which improved substantially with stimulant medication treatment. Trio exome sequencing (ES) of one of the twins and his parents revealed a variant of uncertain significance (VUS) in *GPD2*, which was inherited by all of the children from the mother, who is of above-average intelligence with no dysmorphism and no history of developmental problems. All pregnancies were uncomplicated, except for the post-natal hospital course of the twins.

The daughter was born at term with no complications or dysmorphia. Her language and motor development were typical and she was able to read at an early age. By age five she was reading at a fifth-grade level. She has been described as talented in writing and drawing. According to her parents, she exhibited social oddities from an early age, mainly in communication, and has somewhat intense/restricted interests in fantasy games. She has strong language abilities, and currently attends a 4-year college, but at times uses odd phrases and the rhythm of her speech includes irregular pauses. She has described feeling alienated and "different", and was the victim of bullying in middle school, with few close friends. In late adolescence, she developed major depressive disorder with moderate severity, which brought her to first psychiatric contact. She is cognizant of some degree of social awkwardness, which leads to feelings of anxiety and self-consciousness. The social anxiety inhibits her from activities such as eating in the cafeteria and pursuing job opportunities for which she is otherwise well-qualified. She has a history of becoming emotionally dysregulated and overwhelmed in times of stress, which has led to self-injurious behaviors. She has had ongoing struggles with depressive decompensation and suicidal ideation. She has above-average intelligence but has struggled academically in college due to depression and anxiety. She is medically healthy with the exception of supraventricular tachycardia secondary to atrioventricular node reentry, which was treated with ablation and resulted in subsequent normalization of her electrocardiogram.

The twin boys were born at 35 weeks, had breathing problems at birth, and spent 10 days in the newborn intensive care unit. Neither child has any dysmorphic features nor congenital medical abnormalities, and brain imaging studies were negative. Likewise, neither child has a history of confirmed seizures; however, there are concerns for possible absence epilepsy. There is no history of abnormal neurological examination or macro-/microcephaly. Development of both siblings was delayed, but neither had appreciable regression. The more severely affected twin (designated as the affected proband, AP, and from whom the induced pluripotent stem cell (iPSC) model of severe ASD affectation was acquired) began to exhibit delays in development by 9 months of age. He was speaking single words at 14 months and was ultimately diagnosed with autism at 3.5 years old. Research confirmation of the diagnosis was obtained using the Autism Diagnostic Observation Schedule. Compared to his twin brother, he has had more perseverative interests on odd objects. Psychological testing at the age of nine revealed an IQ of 65 using the Leiter International Performance Scale. Now in late adolescence, he has the ability to engage in reciprocal and meaningful verbal exchanges, although his language is often echolalic and repetitive. He is socially motivated and develops superficial friendships with peers. Functionally, he is able to complete most self-care, dress himself, prepare food and feed himself, and count money. He participates in a vocational program at school and is able to complete rudimentary tasks assigned to him. His monozygotic twin was also diagnosed with ASD at 3.5 years old and is less severely affected, but still requires significant support. Selected clinical characteristics of the family members studied are summarized in Table 1.

Genotyping of the multiplex family

Standard trio ES was performed by the clinical diagnostic laboratory GeneDx for the unaffected mother (UM), the AP, and the unaffected father. As described by GeneDx, exons were captured and sequenced on an Illumina platform with at least 100 base pair read length, followed by alignment to human genome build GRCh37/hg19 and subsequent variant identification. GeneDx's custom-developed tool, Xome Analyzer, was used for variant analysis, as described [32]. This process involves comparing the sequences of each individual to a number of resources, such as published reference sequences, other family members, and control individuals, including the 1000 Genomes database, NHLBI Exome Sequencing Project, ExAC, gnomAD, OMIM, PubMed, and Clinvar. Variant annotations include evolutionary conservation scores, results of in silico prediction tools, and references from the published literature. A phenotype-based approach is also used to generate candidate gene lists. This

Table 1 Selected clinical characteristics and mutational status of several individuals in the multiplex ASD pedigree

Clinical characteristic	AP	IS	UM
Age of ASD diagnostic confirmation	3.5 years	18 years	N/A
Social responsiveness scale-2	83 T	72 T	56 T (spouse-report)
Depression and anxiety	Yes	Yes	No
Seizure history	No	No	No
Developmental delay	Yes	No	No
Eye contact	Poor	Fair	Good
Repetitive behavior	Yes	No	No
Abnormal sensory sensitivities	Yes	No	No
IQ	65	102 (Raven)	108 (Raven)
Speech delay	Yes	No	No
ASD	Severe	Moderate	No
Intellectual disability	Yes	No	No
Mutation location	GPD2	GPD2	GPD2
	chr2:157352686 (hg19) G>A	chr2:157352686 (hg19) G>A	chr2:157352686 (hg19) G>A
	p.G78E, c.233G>A	p.G78E, c.233G > A	p.G78E, c.233G>A

information was interpreted by GeneDx experts according to the American College of Medical Genetics and Genomics Guidelines. ES was performed by The Genome Technology Access Center (GTAC) at Washington University on the intermediate phenotype sister (IS), and variant-specific testing was performed by GeneDx on the trait-affected brother (TB) to confirm the presence of the identified *GPD2* variant.

iPSC generation
iPSC lines were generated by the Genome Engineering and iPSC Center (GEiC) at Washington University. Biomaterials for reprogramming were only available from the UM, IS, and AP. Briefly, renal epithelial cells were isolated and cultured from fresh urine samples and were reprogrammed using a CytoTune-iPS 2.0 Sendai Reprogramming kit (Thermo Fisher Scientific), following the manufacturer's instructions. At least three clonal iPSC lines were derived for each subject, and one or two of these clonal lines (clones 1 and 2) were used for all experimentation involving the UM, IS, and AP. The UC line was previously derived by the GEiC, and one clonal line was available for use in all experiments involving this cell line. All clones (clones 1 and 2) used in experiments were assessed for karyotypic abnormalities by the Washington University School of Medicine Cytogenetics and Molecular Pathology Laboratory, and were also characterized for pluripotency by immunocytochemistry (ICC) and RT-qPCR. Each statistically significant experimental finding reported here was made in experiments that used two different clonal lines per individual (except for the UC, where only one clonal line was available), with at least three independent biological replicate

experiments performed per clonal line. Statistical comparisons were made by one-way ANOVA or unpaired *t* test. Documentation of the clone used for each replicate experiment, the replicates performed, and the statistics for each finding is detailed in Additional file 1: Table S1.

iPSC maintenance and differentiation
iPSC lines were grown under feeder-free conditions on Matrigel (Corning) in mTeSR1 (STEMCELL Technologies). cExN and cIN differentiation of iPSCs was performed using previously described protocols [33]. Briefly, for cExN differentiation, iPSCs were dissociated to single cells with Accutase (Life Technologies) and 40,000 cells were seeded in V-bottom 96-well non-adherent plates (Corning). Plates were spun at 200×*g* for 5 min to generate embryoid bodies (EBs) and were incubated in 5% CO_2 at 37 °C in cExN differentiation medium with 10 μM Y-27632 (Tocris Biosciences). cExN differentiation medium components include Neurobasal-A (Life Technologies), 1X B-27 supplement (without Vitamin A) (Life Technologies), 10 μM SB-431542 (Tocris Biosciences), 100 nM LDN-193189 (Tocris Biosciences). On day four of differentiation, EBs were transferred from V-bottom plates to Poly-L-Ornithine- (20 μg/ml) and laminin- (10 μg/ml) coated plates. Media (without Y-27632) was replenished every other day, and on day 12, Neural Rosette Selection reagent (STEMCELL Technologies) was used to select neural progenitor cells (NPCs) from within neural rosettes, per the manufacturer's instructions. cExN NPCs were grown as a monolayer using cExN differentiation media for up to 15 passages. cIN differentiation media contained the same components as cExN differentiation media, while also including 1 μM

Lewis *et al. Molecular Autism* (2019) 10:51

Page 5 of 23

Purmorphamine (Calbiochem) and 2 μM XAV-939 (Tocris Biosciences). EBs were generated as described for cExNs. At day four of differentiation, the EBs were transferred to non-adherent plates and were placed on an orbital shaker (80 rpm) in an incubator with 5% CO_2 at 37 °C. The media was replenished every other day and, at day ten, EBs were transferred to Matrigel- and laminin- (5 μg/ml) coated plates. Y-27632 was included in media until day eight of differentiation. On day 12 of differentiation, NPCs were dissociated with Accutase and maintained as a monolayer for up to 15 passages. For both cIN and cExN NPC growth analysis, an equal number of cells were seeded on Matrigel- and laminin- (5 μg/ml) coated plates and total cells were counted 4 days later when the cells reached 70–80% confluence.

For differentiation of cExN NPCs into neurons for maturation, 40,000 cells per well were seeded in V-bottom 96-well non-adherent plates. Plates were spun at 200×*g* for 5 min and incubated in 5% CO_2 at 37 °C in maturation medium with Y-27632. Maturation medium components include Neurobasal-A, 1X B-27 supplement (without Vitamin A), 200 μM cAMP (Sigma), 200 μM Ascorbic acid (Sigma), and 20 ng BDNF (PeproTech). After 2 days, EBs were transferred to Matrigel- and laminin- (5 μg/ml) coated plates and media was replenished every other day (without Y-27632). On day 12 of neuronal differentiation and maturation, cells were dissociated with Accutase and seeded in an eight-well chamber for ICC.

For neurosphere size measurement analysis, *P* values *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ were determined by one-way ANOVA.

Sanger sequencing
DNA was isolated from cell lines using the PureLink Genomic DNA Kit (Invitrogen). Primers were designed to amplify a 248 base pair region of *GPD2* flanking the identified point mutation (forward primer: AAGCAG CAGACTGCATTTCA, reverse primer: CACCATGGCA CACACTTACC). Sanger sequencing was performed on this PCR amplified fragment using either the forward or reverse primer. CodonCode Aligner software was used to analyze sequencing results.

Immunocytochemistry (ICC) and immunoblotting
For ICC, cells were plated on eight-well chamber slides coated with Matrigel and laminin (5 μg/mL). After 1 day, cells were washed once with PBS without calcium and magnesium (PBS - Ca^{2+}/Mg^{2+}) and fixed in 4% paraformaldehyde for 20 min, followed by washing with PBS + Ca^{2+}/Mg^{2+}. Cells were blocked with blocking buffer (10% donkey serum, 1% BSA, and 0.1% TritonX-100 in PBS + Ca^{2+}/Mg^{2+}) for at least 1 h and incubated with primary antibodies overnight (Additional file 1: Table

S1) in antibody dilution buffer (1% donkey serum, 1% BSA, and 0.1% TritonX-100 in PBS + Ca^{2+}/Mg^{2+}). After overnight incubation, cells were washed three times with wash buffer (0.1% Triton X-100 in PBS + Ca^{2+}/Mg^{2+}). Cells were incubated with corresponding secondary antibodies (Additional file 1: Table S1), along with DAPI (1 mg/mL; ThermoFisher Scientific), diluted in antibody dilution buffer for 1 h. Following secondary antibody incubation, cells were washed twice with wash buffer and once with PBS + Ca^{2+}/Mg^{2+}. Slides were mounted with Prolong Gold anti-fade agent (Life Technologies). Images were obtained using a spinning-disk confocal microscope (Quorum) with MetaMorph software and were processed using ImageJ. For immunoblotting, cell lysate was extracted and 30 μg of protein was used per lane. Antibodies used are listed in Additional file 1: Table S1.

FACS analysis
For each experiment, approximately one million cells were pelleted, washed with PBS – Ca^{2+}/Mg^{2+}, resuspended in PBS – Ca^{2+}/Mg^{2+} and fixed by adding 70% ice-cold ethanol dropwise while vortexing. Cells were stained with 10 μg/mL propidium iodide (Sigma) and 200 μg/mL RNase A (Fisher Scientific) in FACS buffer (PBS – Ca^{2+}/Mg^{2+}, 0.2% BSA, 1 mM EDTA). FACS was performed on single-cell suspensions and the cell cycle analysis function of FlowJo was used to analyze cell cycle composition for each sample, based on propidium iodide staining to detect DNA content in each cell. *P* values *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ were determined by one-way ANOVA.

RNA-seq and RT-qPCR
Total RNA was collected from iPSC-derived day 12 cExN and cIN NPCs using the NucleoSpin RNA II kit (Takara) per the manufacturer's instructions. RNA was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific), and the integrity of RNA was confirmed with an Agilent Bioanalyzer 2100 to ensure a RIN value above eight. RNA-sequencing (RNA-seq) library preparation and Illumina sequencing were performed by the GTAC at Washington University. Single-end 50 base pair reads were obtained using an Illumina HiSeq 3000 sequencer, obtaining an average of ~30 million uniquely aligned reads per sample. For RT-qPCR, 1 μg total RNA was reverse transcribed using iScript Reverse Transcription Supermix (Bio-Rad). Equal quantities of cDNA were used as a template for RT-qPCR, using the Applied Biosystems Fast Real-Time quantitative PCR system. RPL30 mRNA levels were used as endogenous controls for normalization. *P* values *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ were determined by an unpaired *t* test.

Bioinformatics and IPA analyses

RNA-seq data analysis was performed as described in [33] to curate differentially expressed gene (DEG) lists. In summary, RNA-seq reads were aligned to the human genome (assembly hg38) with STAR version 2.5.4b [34]. Gene counts were derived from the number of uniquely aligned unambiguous reads by Subread:featureCount, version 1.6.3, with GENCODE gene annotation (V27) [35, 36]. All gene-level transcript counts were then imported into the R/Bioconductor package DESeq2 [37]. Genes expressed below a CPM of 1.0 in more than half the samples were excluded from further analysis. DEG cutoffs were set at a log2-fold change of > 1.0 and a Benjamini and Hochberg FDR of < 0.05.

To uncover the biological significance of DEGs, network analysis was performed with the data interpretation tool Ingenuity Pathway Analysis (IPA) (Qiagen). IPA's Ingenuity Knowledge Base uses network-eligible DEGs to generate networks and to define connections between one and more networks. Based on the number of eligible DEGs, IPA defines network scores as inversely proportional to the probability of finding the network and defines significant networks ($P \leq 0.001$). Within each network, red symbols indicate upregulated genes and green symbols indicate downregulated genes, where the color intensity represents the relative degree of differential expression.

Co-expression and variance analysis

For this correlation analysis, the read counts matrix for DEGs in either the cIN or cExN samples was log2-CPM transformed. A similarity matrix for these genes was created by calculating the Euclidian distance among the genes from the log2-CPM matrix in R. The WGCNA R package "adjacency.fromSimilarity" function with arguments power=12, type='signed' was used to create an adjacency matrix from the similarity matrix [38]. Hierarchical clustering was performed using the hclust function in R on the distance matrix transformation of the adjacency matrix using the ward.D2 method. The dendrogram of this gene tree was split into $k = 3$ clusters using the cutree function in the stats package of R. The read counts matrix for all genes in either the cIN or cExN samples was CPM transformed. Genes expressed at < 1 CPM in over half of the samples were excluded. Average CPM among the four sample types was calculated for each gene, and genes with expression ratios of < 1.25 between the highest and lowest expressed sample type were excluded from further analysis. The filtered gene CPM matrix was log2 transformed, and the Pearson correlation matrix for the remaining genes was calculated using the cor function from the stats package in R. A gene was considered to be correlated to a DEG cluster if it had a Pearson correlation coefficient > 0.7

for over half the genes in the DEG cluster. Genes with expression correlated to a cluster of genes from the DEG clustering were merged with those genes and fed into the ToppFun GO analysis of ToppGene [39]. For each correlated cluster, up to five terms with the most significant Benjamini and Hochberg FDR values were retained for each biological process, cellular component, and molecular function GO terms. The top three pathway and disease terms were also retained for each cluster. To assess the impact of different covariates on expression, the R/Bioconductor package variancePartition was utilized [40]. The parameters assessed were cell type (cExN vs. cIN), subject (UC, UM, IS, AP), age (young vs. old), or sex (male vs. female).
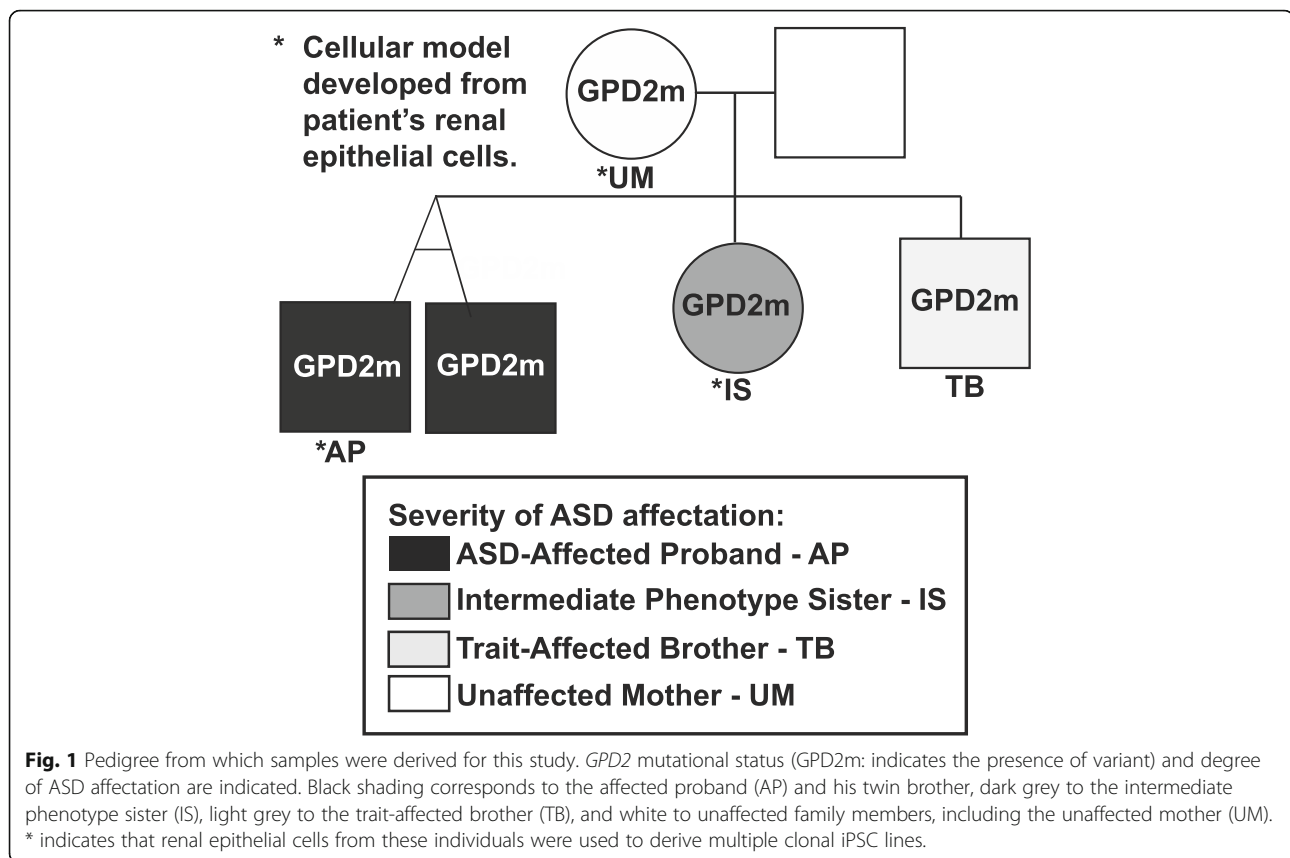
Results

Phenotyping and genotyping of the multiplex family

The multiplex autism spectrum disorder (ASD) pedigree selected for study (Fig. 1; Table 1) underwent clinical phenotyping and genotyping (see Methods). From this pedigree, the individuals selected for iPSC line derivation and modeling included the affected proband (AP), his sister, who has an intermediate phenotype (IS), and their unaffected mother (UM) (indicated in Fig. 1). As described in the Methods, a non-synonymous single nucleotide variant in the GPD2 gene was identified in the UM and all of the children (chr2:157352686 (hg19) G>A, NM_001083112.2 c.233G>A, p.G78E). This variant is not present in the father. The variant is in exon three of the GPD2 gene, within the region encoding the flavin adenine dinucleotide (FAD)-binding domain of the GPD2 protein (Additional file 2: Figure S1A). The GeneDx interpretation of this variant states that it was not observed in approximately 6500 individuals of European and African American ancestry in the NHLBI Exome Sequencing Project and that it is evolutionarily conserved. In addition, in silico analysis predicts that this variant is probably damaging to the protein structure and function. Overall, GeneDx designated this as a variant of uncertain significance (VUS) following the American College of Medical Genetics criteria. Subsequent mutation-specific testing and Sanger sequencing identified this same variant in the AP, IS, and UM, while it was absent in an unrelated, unaffected control (UC) (Additional file 2: Figure S1B). Given the differential affectation of members of this pedigree carrying this variant, it was apparent that this inherited VUS was insufficient to cause ASD by itself, but may have contributed to polygenic risk/liability in this family.

Generation of subject-derived iPSC models and directed differentiation into cortical excitatory neurons

Two clonal iPSC lines were derived from the UM, IS, and AP, and were characterized in parallel with a single,

Fig. 1 Pedigree from which samples were derived for this study. GPD2 mutational status (GPD2m: indicates the presence of variant) and degree of ASD affectation are indicated. Black shading corresponds to the affected proband (AP) and his twin brother, dark grey to the intermediate phenotype sister (IS), light grey to the trait-affected brother (TB), and white to unaffected family members, including the unaffected mother (UM). * indicates that renal epithelial cells from these individuals were used to derive multiple clonal iPSC lines.

clonal iPSC line derived from the UC. When grown in stem cell maintenance media, there were minimal differences observed in the expression of pluripotency markers between these iPSC lines, as assessed by RT-qPCR (Additional file 2: Figure S1C) and immunocytochemistry (ICC) (Additional file 2: Figure S1D). In addition, no differences in GPD2 protein levels were detected in iPSCs by ICC (Additional file 2: Figure S1D) or Western blotting (Additional file 2: Figure S1E). All cell lines used in this study were also shown to be karyotypically normal (Figure S1F). Finally, we used FACS analysis of propidium iodide (PI)-stained iPSCs to assess cell cycle progression and detected no observable differences between these iPSC lines, which had similar percentages of cells in each stage of the cell cycle (Additional file 2: Figure S1G-H).

In the cortex, as a result of their abnormal development, imbalances in glutamatergic excitatory neurons (cExN) and GABAergic inhibitory interneurons (cINs) are thought to contribute to neurodevelopmental disorders including ASD [3, 5, 41]. We, therefore, differentiated iPSCs derived from the UC, and from three family members (the UM, IS, and AP), in parallel into either cExN or cIN neural progenitor cells (NPCs) and/or neurons, to determine if we observed any alterations in the in vitro development of either or both of these neural

cell types. We performed 12 days of cExN differentiation (4 days as embryoid bodies (EBs) in V-bottom plates, followed by 8 days with the EBs plated for two-dimensional (2-D) culture; Fig. 2a). At all time points assessed during this differentiation, the IS and AP lines generated significantly smaller neurospheres than the UC and UM. The UM neurospheres were also slightly smaller than those of the UC (Fig. 2b, c).

To identify whether an increase in apoptosis and/or a decrease in proliferation could be contributing to these differences in neurosphere size, we performed FACS analysis of PI-stained cells at day four of differentiation and found that the IS and AP neurospheres had a significantly higher fraction of sub-G1 (apoptotic) cells, compared to neurospheres derived from the UM and UC lines (< 2 N DNA content; Fig. 2d, e, g). There was a corresponding decrease in the percentage of cells in the G1 phase of the cell cycle in the IS and AP neurospheres (2 N DNA content; Fig. 2d, f–g). However, neurospheres from all lines had similar percentages of cells in the S and G2/M phases of the cell cycle, suggesting that their cell cycle characteristics and rates of progression were otherwise similar (S phase and 4 N DNA content; Fig. 2d, g). To determine whether induction of neural differentiation was a stressor that was contributing to this increase in apoptosis in the IS and AP line-derived neurospheres,
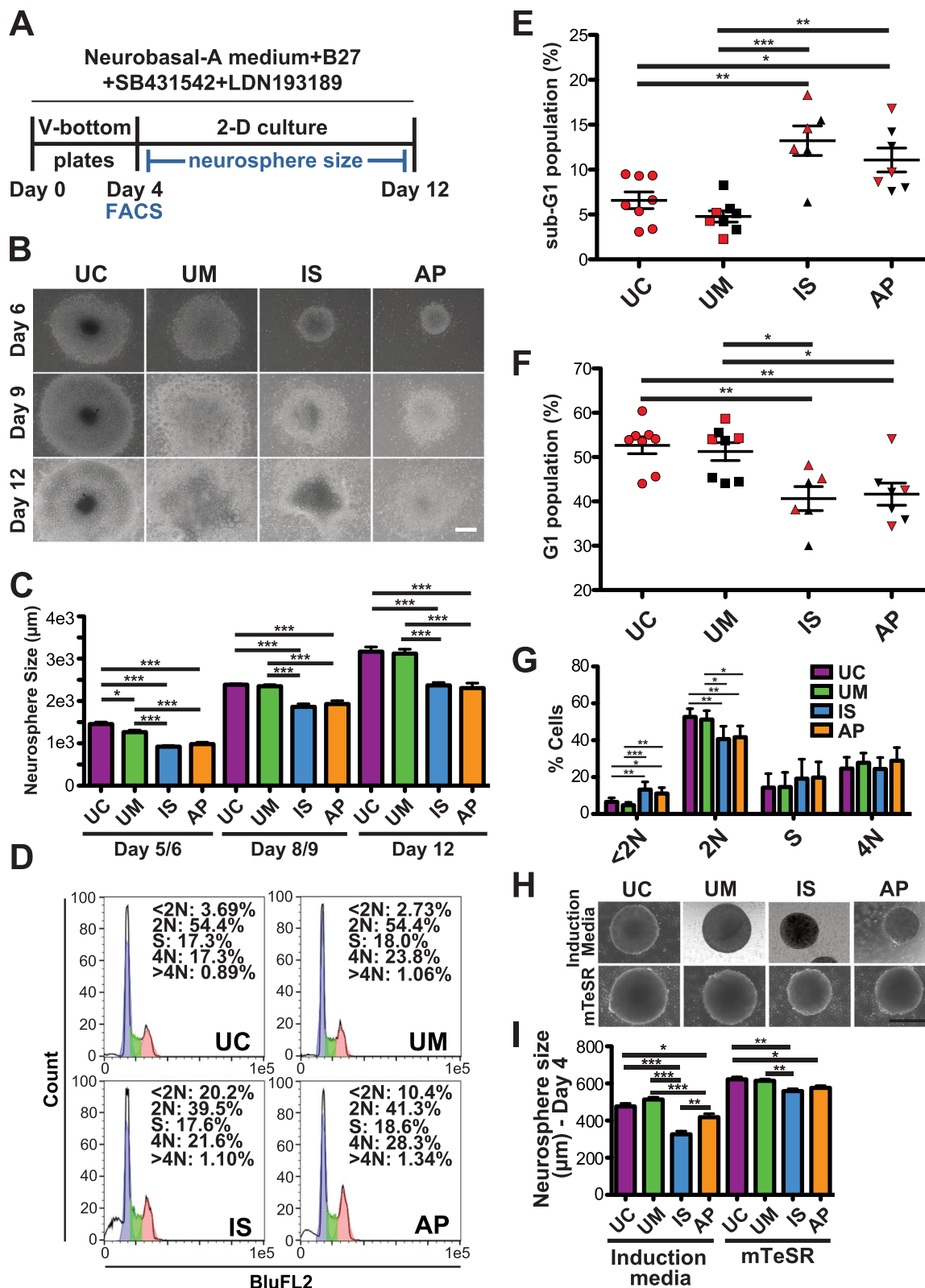
Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 Characterization of iPSC lines during differentiation into cExN NPCs. **a** Differentiation scheme, including timeline and small molecules used. **b, c** iPSCs derived from an unrelated, unaffected control (UC), as well as the UM, IS, and AP were differentiated for 12 days to generate cExN NPCs. Neurosphere size at several time points is shown in (**b**) and quantified in (**c**) (mean ± SEM; scale bar = 500 μm; *n* = 16 biological replicates, encompassing two different clonal lines from each subject, and one clonal line for the UC). **d-g** At day four of differentiation, cells were stained with propidium iodide and FACS analysis of DNA content was performed. **d** Representative FACS plots. In (**e**) <2 N (sub-G1) and (**f**) 2 N (G1) cells are quantified, with values shown for each replicate. **g** shows mean values for all cell cycle stages for each cell line (mean ± SEM; *n* = 3 or more biological replicates for each subject, encompassing two different clonal lines from each subject, and one clonal line for the UC). **h-i** iPSCs were cultured in either neural induction media or in mTeSR stem cell media, and EB size was analyzed at day four of differentiation. Representative images are shown in (**h**) (scale bar = 500 μm), with quantification in (**i**) (mean ± SEM; *n* = 3 or more biological replicates for each subject, encompassing two different clonal lines from each subject, and one clonal line for the UC). *P* values **P* < 0.05, ***P* < 0.01, ****P* < 0.001 were determined by one-way ANOVA and all other pairwise comparisons had a non-significant *P* value (*P* ≥ 0.05). Red and black data points denote experiments performed with clone 1 or with clone 2, respectively

we compared sphere size after culturing spheres from each line either in stem cell maintenance media (mTeSR) or in neural induction media. In general, sphere size was larger for all cell lines when kept in mTeSR media rather than neural induction media, while the differences in sphere size for the IS and AP versus the UC and UM was much less pronounced in mTeSR, relative to differences seen under neural induction conditions (Fig. 2h, i). These data suggest that, by comparison with the UC and UM, the IS and AP lines have a slightly elevated propensity to undergo apoptosis upon dissociation and sphere formation, while this is exacerbated by induction of neural differentiation.

We next maintained these four lines as NPCs after neural rosette selection at day 12 and then subjected them to PI staining and FACS analysis (Additional file 2: Figure S2A). Unlike the results from earlier time points, the cExN NPCs showed little difference in cell cycle across the four lines (Additional file 2: Figure S2B-C), nor in the rate of growth or apoptosis over the course of culture for 4 days (Additional file 2: Figure S2D). However, morphological analysis by bright-field imaging indicated a possible adhesion defect in the IS NPCs, as indicated by uneven growth on the cell culture plate surface (Additional file 2: Figure S2E). At the NPC stage, GPD2 protein levels remained similar across the four lines, as was shown for iPSCs (Additional file 2: Figure S2F).

Finally, to determine if the NPCs derived from the affected individuals exhibited an altered capacity to differentiate into cExN neurons, NPCs from the four lines were further differentiated for 12 days as shown (Additional file 2: Figure S3A) and subjected to ICC. No apparent differences between the four lines were observed in the expression of NPC markers (PAX6, NESTIN, and SOX1) or markers of immature (TUJ1) and mature cExN neurons (VGLUT, MAP2) (Additional file 2: Figure S3B). Furthermore, there were no observable differences between the lines in the fraction of cells expressing Ki-67, a marker of cell proliferation, or cleaved Caspase-3, a marker of apoptosis (Additional file 2: Figure S3B).

Differentiation of subject-derived iPSCs into cortical interneuron progenitors

We also characterized cellular phenotypes of these four lines during differentiation into cIN NPCs, to define any differences between the development of this neural cell type in lines derived from affected and unaffected individuals. The differentiation scheme to produce cIN NPCs is outlined in Fig. 3a. On day five of differentiation in this scheme, neurospheres derived from the IS line were smaller than those of the UM. Conversely, the AP line-derived neurospheres were slightly larger than the UM line neurospheres (Additional file 2: Figure S4).

After dissociation on day 12 of differentiation, we assessed the cell cycle of the cIN NPCs using FACS of PI-stained cells (Fig. 3b–e). The IS and AP cIN NPCs had an increased sub-G1 cell population, compared to the UC and UM NPCs, an indication of increased apoptosis in the cells from the affected individuals (Fig. 3c). Correspondingly, there was a decrease in the proportion of cells in the G1 phase of the cell cycle (Fig. 3d). However, no differences were observed in frequencies of cells in the S and G2/M phases of the cell cycle between lines, suggesting that these lines had similar proliferation rates (Fig. 3b, e). This result was supported by analysis of NPC cell counts after 4 days of growth, which revealed a significant reduction in the number of AP NPCs, as well as a slight reduction in the number of IS NPCs, compared to the UM NPCs (Fig. 3f). These reductions in NPC number may have resulted from the increased NPC apoptosis detected in our PI FACS analysis. The AP NPCs also exhibited altered morphology that could indicate impaired adhesion capacity relative to the control UM/UC lines, which could also have contributed to the reduction in the number of AP NPCs persisting in the culture after 4 days of growth (Fig. 3g).

Transcriptomic differences in neural progenitor cells derived from affected individuals versus controls

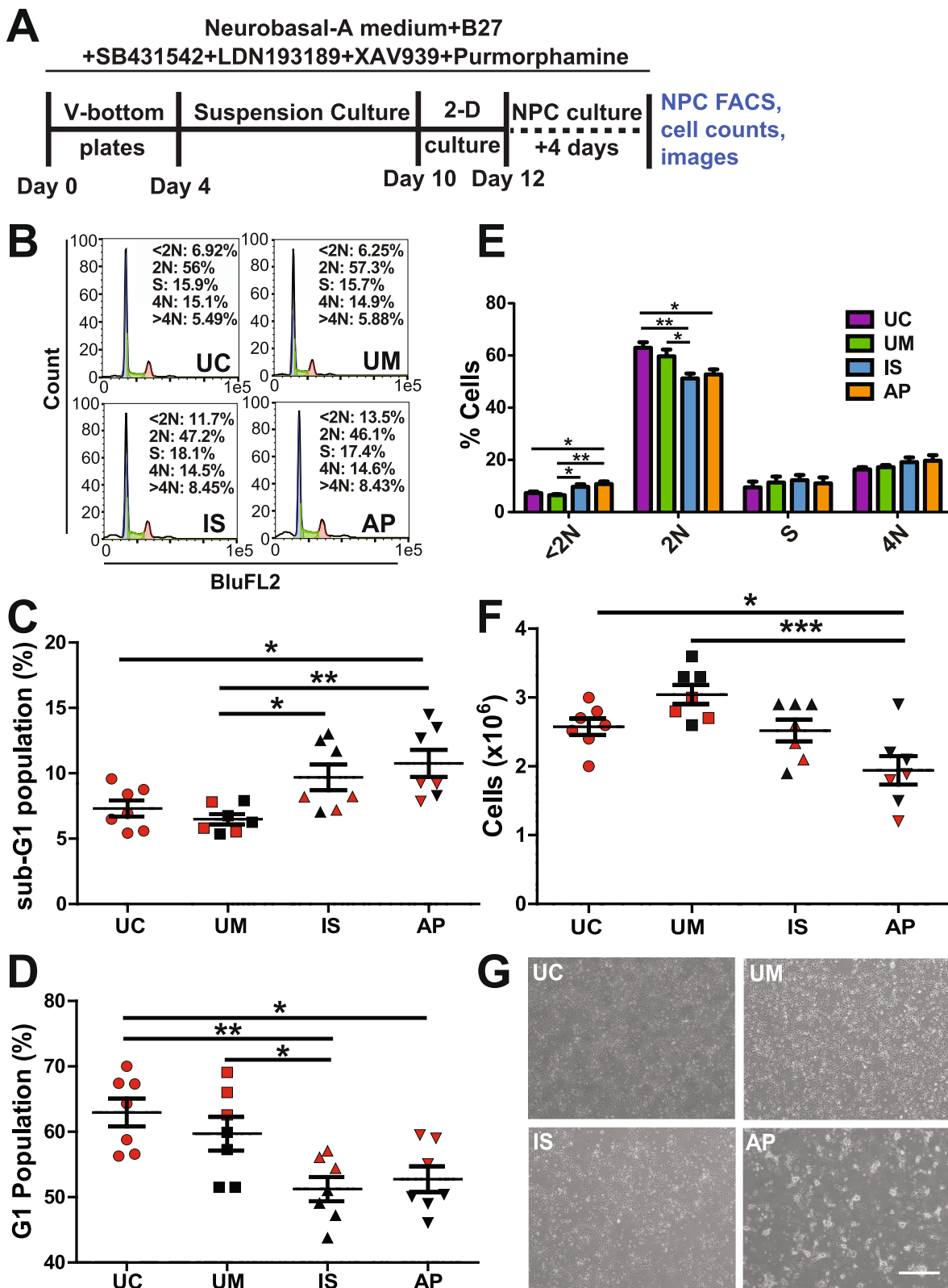To investigate which classes of genes could be differentially expressed in neural cells from the affected

Fig. 3 (See legend on next page.)
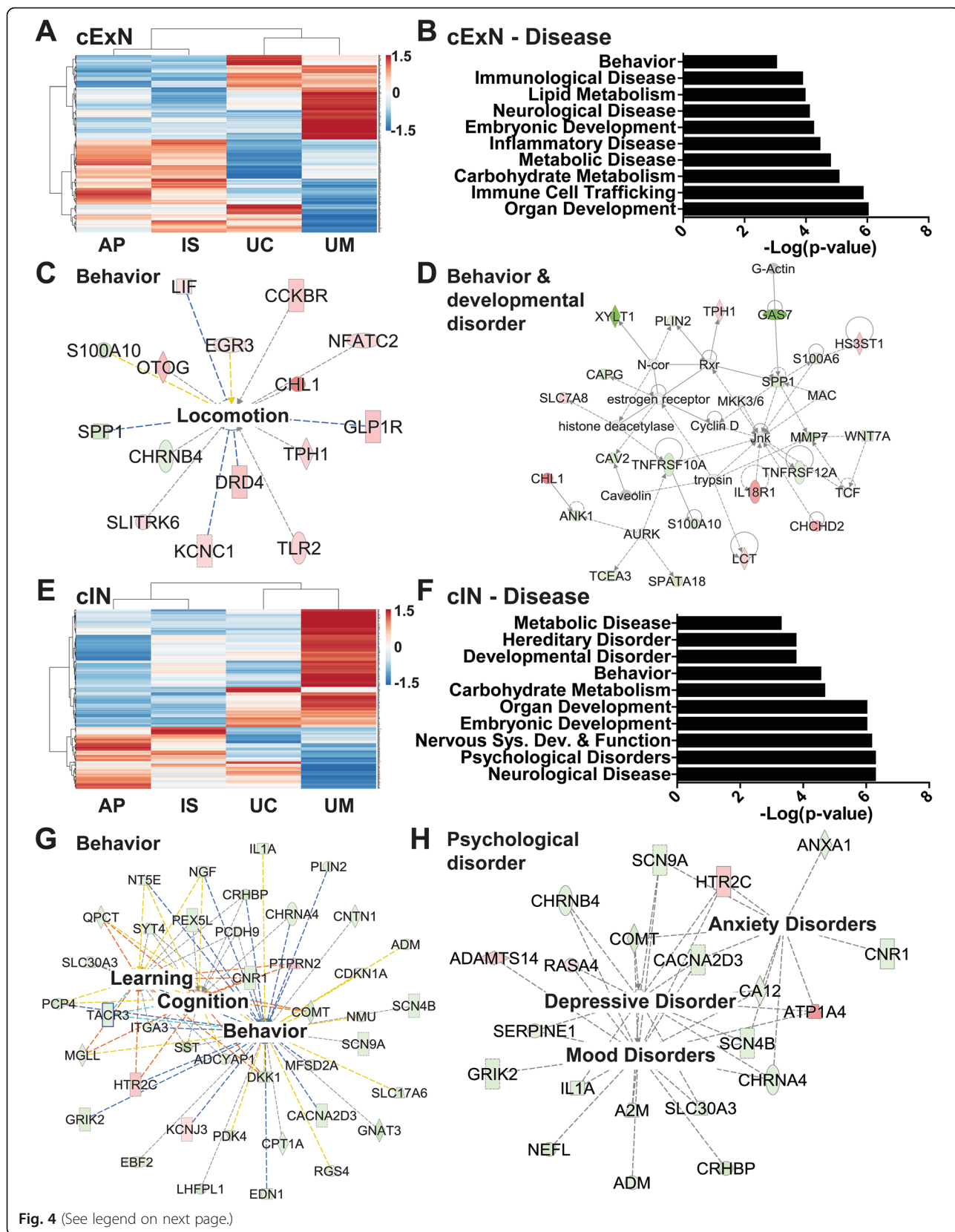
(See figure on previous page.)

Fig. 3 Characterization of iPSC-derived cIN NPCs. **a** Differentiation scheme, including timeline and small molecules used. **b-d** After 12 days of differentiation, cIN NPCs were stained with propidium iodide and analyzed by FACS for DNA content. (**b**) Representative FACS plots. In C and D, respectively, < 2 N (sub-G1) and 2 N (G1) cells were quantified, with values shown for each replicate (n = 3 or more biological replicates from each of two clonal lines for each subject, and one clonal line for the UC). **e** Mean percentages of cells in each cell cycle stage are shown (n = 3 or more biological replicates from each of two clonal lines for each subject). **f-g** cIN NPCs were plated in equal numbers for each sample and counted after four days of culture. Data are quantified in (**f**) and representative images are shown in (**g**) (n = 3 or more biological replicates from each of two clonal lines for each subject, and one clonal line for the UC). P values *P < 0.05, **P < 0.01, ***P < 0.001 were determined by one-way ANOVA and all other pairwise comparisons had a non-significant P value (P ≥ 0.05). Red and black data points denote experiments performed with clone 1 or with clone 2, respectively

individuals, by comparison with the unaffected controls, we performed RNA-seq analysis on both cExN and cIN NPCs on day 12 of differentiation for all four subject-derived lines. Four biological replicates were analyzed for each sample type and were clustered by principal component analysis (PCA) of processed reads (Additional file 2: Figure S5A). We defined genes that were significantly differentially expressed genes (DEGs) in pairwise comparisons of these four sample types for either cExN or cIN NPCs, selecting DEGs with a log2-fold difference between sample types of > 1 and a Benjamini and Hochberg FDR of < 0.05 (Additional file 3: Table S2). In a within-family comparison of the UM, IS, and AP samples, greater numbers of DEGs were obtained in the cIN NPC pairwise comparisons versus the numbers of DEGs obtained for cExN NPC pairwise comparisons (Additional file 2: Figure S5B-C). These data indicate that the cIN samples from the affected individuals (IS/AP) exhibit more transcriptomic differences from the UM control than the affected individual-derived cExN samples.

We focused first on identifying classes of genes that were differentially expressed in NPCs derived from the affected individuals, by comparison with unaffected controls. To do this, we defined the subset of DEGs that were similarly expressed in samples from both affected individuals (AP/IS) but that differed in expression by comparison with the unaffected mother (UM) sample. Relative expression is also shown for the UC, for a full cross-sample comparison. Four hundred fifty-two and 437 DEGs for the cExN and cIN NPC samples met these criteria, respectively. Hierarchical clustering and visualization of the relative expression of these DEGs across the four sample types are shown for the cExN NPCs (Fig. 4a, Additional file 4: Table S3). We next used the Ingenuity Pathway Analysis (IPA) to assess the potential biological significance of these genes. For the 452 DEGs in the cExN NPCs described above, the most significant function- and disease-related gene ontology (GO) terms included 'behavior', 'neurological disease', and 'embryonic development' (Fig. 4b). Network analysis using IPA revealed several interesting networks of DEGs related to these GO terms, including networks related to 'locomotion' (from DEGs within the 'behavior' GO term) and 'behavior and

developmental disorder' (Fig. 4c, d). Within the 'locomotion' network, most genes were upregulated in the affected individuals compared to the controls, including genes relating to neural adhesion and ion channels (Fig. 4c and Additional file 5: Table S4). Genes with known roles in NPCs or neurons, as well as stress-related genes were present in the larger 'behavior and developmental disorder' network (Fig. 4d and Additional file 5: Table S4). Interestingly, another network comprising genes from the GO term 'neurological disease' is related to 'inflammation of central nervous system' (Additional file 2: Figure S6A and Additional file 5: Table S4).

IPA analysis of the cIN DEGs also revealed several interesting classes of genes that were differentially expressed between the affected participants and unaffected control NPC-derived samples. Hierarchical clustering and visualization of the relative expression of DEGs across the four sample types for the cIN NPCs are shown in Fig. 4e (Additional file 4: Table S3). The top GO terms included 'developmental disorder', 'behavior', 'nervous system development and function', 'psychological disorders', and 'neurological disease' (Fig. 4f). Within the term 'behavior', a network that includes 'learning'-, 'cognition'-, and 'behavior'-related genes was identified (Fig. 4g, Additional file 5: Table S4). The network related to the GO term 'psychological disorder' includes genes related to 'anxiety disorders', 'mood disorders', and 'depressive disorder' (Fig. 4h). The 'nervous system development and function' network includes genes involved in the 'quantity of neurons' and 'quantity of synapse', as well as cell adhesion genes (Additional file 2: Figure S6B and Additional file 5: Table S4). Finally, a 'neurological' network included a number of genes also present in the other networks (Additional file 2: Figure S6C). We further assessed these affectation-linked DEGs by gene co-expression analysis and hierarchical clustering, as described in the methods (Additional file 2: Figure S7). Gene ontology analysis of co-expressed gene clusters identified neuron-related (e.g. neuron projection, axon, synapse) and neurological disease-related terms (e.g. mental depression, autistic disorder) as top GO terms (Additional file 6: Table S5). Taken together, this analysis of DEGs in both cExN and cIN NPCs shows evidence of altered expression of a number of neurological and psychological disease-relevant gene classes in the AP- and IS-

Lewis *et al. Molecular Autism* (2019) 10:51

Page 12 of 23

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Transcriptomic analysis of genes with differential expression in affected subject-derived NPCs, relative to unaffected controls. **a** Hierarchical clustering of RNA-seq data from the cExN NPCs identified differentially expressed genes (DEGs) with shared expression in the AP and IS that differed from that in the UM. Relative expression is also shown for the UC, for a full cross-sample comparison ($P < 0.05$, fold-change > 2; $n = 4$ biological replicates from one clonal line for each subject, and one clonal line for the UC). **b-d** Ingenuity Pathway Analysis (IPA) of these cExN NPC DEGs defined disease- and function-associated gene ontology (GO) terms and identified gene networks associated with the (**c**) 'behavior' and (**d**) 'behavior and developmental disorder' GO terms. **e** Hierarchical clustering of RNA-seq data for the cIN NPCs identified DEGs with shared expression in the AP and IS, that differed from that in the UM. Relative expression is also shown for the UC, for a full cross-sample comparison ($P < 0.05$, fold-change > 2; $n = 4$ biological replicates from one clonal line for each subject, and one clonal line for the UC). **f-h** IPA analysis of these cIN NPC DEGs defined (**f**) disease- and function-associated GO terms and identified gene networks associated with (**g**) 'behavior' and (**h**) 'psychological disorder'. Within each network, red symbols indicate upregulated genes and green symbols indicate downregulated genes, while color intensity indicates the relative degree of differential expression

derived lines, relative to lines derived from the UM and/or UC.

Within-family comparison identifies a transcriptome signature specific to neural progenitor cells derived from the ASD-affected proband

As differences in genetic background can confound differential gene expression analysis [42], we also performed a pairwise, within-family data comparison of DEGs that distinguish the UM-, IS-, and AP-derived samples, focusing on DEGs specific to the AP that could contribute to the greater degree of affectation observed. Using pairwise comparisons of DEGs, we defined 190 genes which were uniquely differentially expressed in cExN NPCs from the AP (Fig. 5a). The top GO terms associated with these AP-specific DEGs included 'psychological disorders', 'behavior', 'nervous system development and function', 'developmental disorder', and 'neurological disease' terms (Fig. 5b). Within the 'behavior' term, network analysis showed genes related to 'memory' and 'learning' to be dysregulated (Fig. 5c and Additional file 5: Table S4). Within the 'nervous system development and function' term, a network of dysregulated genes related to 'differentiation of neurons' was identified (Fig. 5d, Additional file 5: Table S4).

A similar analysis was performed on the cIN samples, revealing 384 DEGs unique to the AP samples in the within-family comparison (Fig. 5e). IPA analysis identified classes of DEGs related to the GO terms 'psychological disorders', 'developmental disorder', 'neurological disease', 'behavior', and 'nervous system development and function' (Fig. 5f). Within the 'behavior' disease term, a network of genes related to 'behavior' and 'cognition' was identified (Fig. 5g and Additional file 5: Table S4). Within the 'nervous system development and function' term, a network of genes related to 'development of neurons' and 'synaptic transmission' had altered expression in the AP versus the IS/UM-derived samples (Fig. 5h and Additional file 5: Table S4). Together, this analysis identified AP-unique DEGs in both cExN and cIN NPCs, many of which are broadly related to neural development, as well as to specific aspects of ASD, such as behavioral alterations. These

gene expression changes, therefore, correlate with and may contribute to, the severity of affectation in the AP.

Comparison of differentially expressed genes with ASD-associated genes and validation

The Simons Foundation Autism Research Initiative (SFARI) [43] maintains a database of genes that are mutated to cause or that contribute to ASD risk. We compared our DEGs to these ASD-related genes, to assess whether their dysregulated expression could contribute to affectation in these individuals. Of the 584 unique DEGs in the cExN differentiation scheme that were either specific to the AP (Fig. 5a) or that had similar expression in the AP and IS that differed from that seen in the UM (Fig. 4a), 30 (5.1%) were SFARI ASD genes (Fig. 6a; Additional file 6: Table S5). For the corresponding cIN NPC comparison, 48 of 692 unique DEGs (6.9%) were ASD genes in the SFARI Gene database (Fig. 6b; Additional file 6: Table S5). Based upon the 1019 genes present in the SFARI Gene database [43] and the total of 27,731 genes with > 0.1 RPKM average expression across all cExN and cIN samples, the number of AP- and IS-specific DEGs that are ASD genes is significantly greater than would be expected by chance (hypergeometric distribution, $P = 3.67 \times 10^{-3}$ and 3.49×10^{-7} for cExN and cIN data, respectively). A subset of these are associated with syndromic ASD (*NTNG1, ALDH1A3, DMD, EBF3, PRODH,* and *RNF135*) and/or are linked with ASD with the highest confidence (SFARI gene scores 1–2: *KAT-NAL2, MYT1L, CACNA2D3, GRIA1, SCN9A,* and *CNTN4*). Comparison with the 465 genes in the Geisinger Developmental Brain Disorder Gene Database also revealed recurrent association of some of these genes with ASD and/or intellectual disability, but infrequent association with other neurological disorders (Additional file 6: Table S5) [44]. Interestingly, only one (for cExN) or none (for cIN) of the DEGs overlapped with high-confidence genes unique to adult-onset psychiatric disorders (e.g., not also associated with neurodevelopmental disorders) identified in the PsyGeNET database (Additional file 6: Table S5) [45, 46]. Therefore, it is possible that misregulated expression and
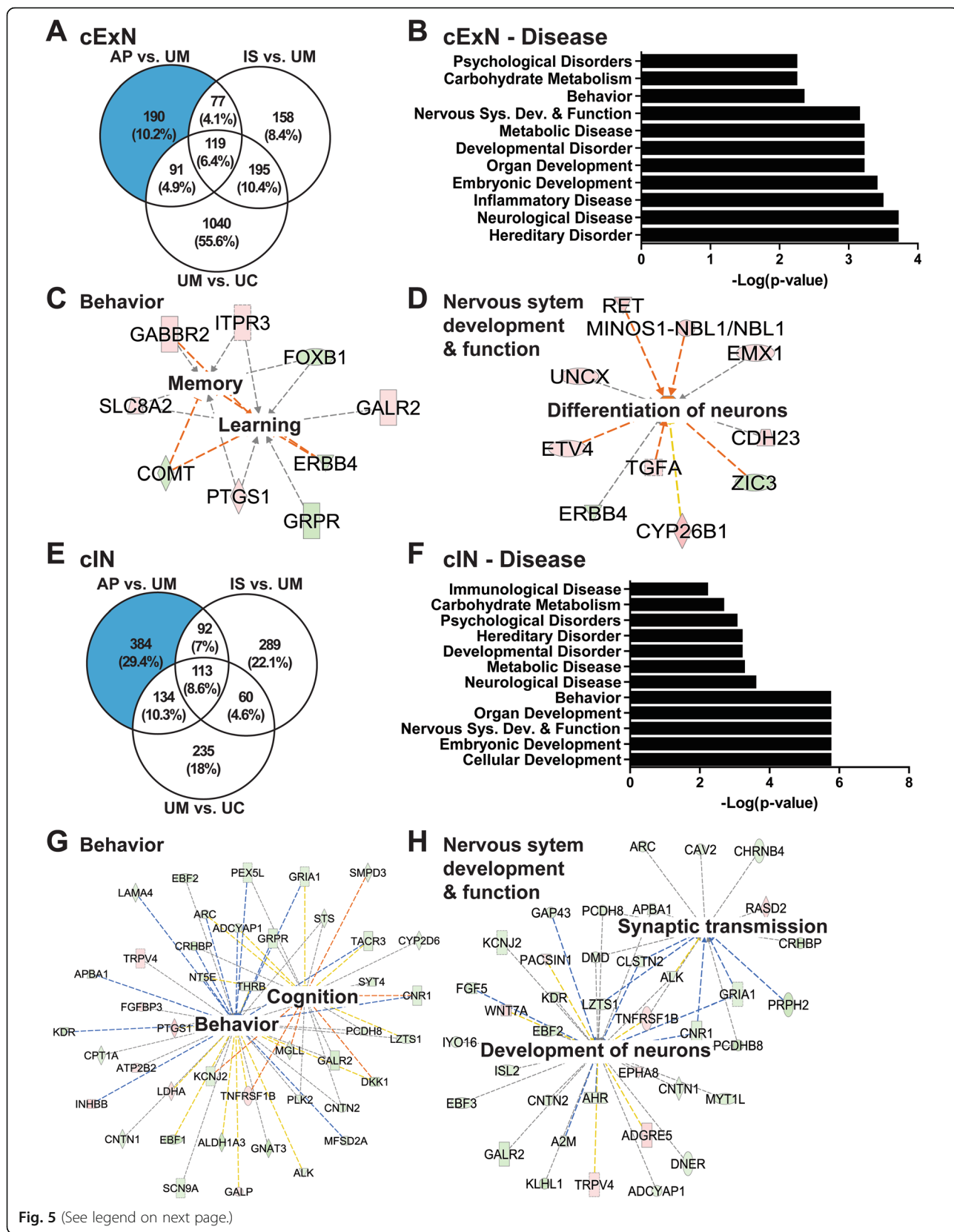
Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Within-family analysis of transcriptomic signatures specific to the affected proband-derived samples. **a** Venn diagram for the cExN NPCs, showing the DEGs from pairwise comparisons of different samples, including numbers of overlapping DEGs. The blue shaded portion of the Venn diagram indicates DEGs unique to the AP, not shared by the IS or UM. **b-d** Ingenuity Pathway Analysis (IPA) of the AP-unique DEGs in cExN NPCs defined (**b**) class and function-associated GO terms and identified gene networks associated with (**c**) 'behavior' and (**d**) 'nervous system development and function'. **e-h** IPA analysis of the AP-unique DEGs in cIN NPCs determined (**f**) class and function-associated GO terms and identified gene networks associated with (**g**) 'behavior' and (**h**) 'nervous system development and function'. Within each network, red symbols indicate upregulated genes and green symbols indicate downregulated genes, where the color intensity represents the relative degree of differential expression

consequently the function of ASD genes contributed to the disruption of neural development and/or continues to contribute to altered neurological function in the IS and AP.

We validated the differential expression of a subset of the DEGs described above by RT-qPCR analysis, isolating RNA from NPCs derived from the second set of iPSC clones that differed from those used for the RNA-seq experiments. Expression changes of DEGs selected from the cExN NPC RNA-seq data (Fig. 7a) were robustly recapitulated in these experiments (Fig. 7b). Genes from the 'behavior and developmental disorder', from other identified networks, and genes involved in neurodevelopment were evaluated (Additional file 5: Table S4). We also derived cIN NPC RNA from the second set of iPSC clones and validated the corresponding RNA-seq data for a subset of the DEGs (Fig. 7c). Differential expression was assessed for SFARI ASD genes [43], and for genes encoding transcription factors, ion channels, and cell adhesion molecules (Fig. 7d, e and Additional file 5: Table S4). In addition, we validated a subset of DEGs in cINs which also had differential expression in cExN NPCs (Fig. 7e).

Differential gene expression in iPSC disease modeling studies involving female iPSC lines can be confounded by erosion of X-inactivation, including alteration of sex chromosome-linked gene expression [47–50]. Therefore, we also assessed XIST expression levels, as an indicator of potential X-erosion in our female lines. We observed that XIST RNA expression was reduced in the UC but not in either clonal line derived from the UM or IS, evidence of potential X-erosion in the UC line (Additional file 2: Figure S8A-B). However, as few of our DEGs are X-linked (Additional file 4: Table S3), this does not appear to be a major contributor to the differential gene expression identified in this study. Finally, we performed variancePartition analysis on these RNA-seq data to quantify contributions of multiple sources of variation to the differential gene expression obtained. The cell type (cExN or cIN) or subject from whom the sample was derived were the main contributors to differential gene expression, while the age and sex of the subject were only minor contributors to the total variance observed in both cExN and cIN RNA-seq datasets (Additional file 2: Figure S8C). Together, these analyses revealed that

relative to unaffected individuals, samples from affected individuals exhibited altered expression of classes of genes involved in behavior, learning, cognition, mood disorders, and neurodevelopment, including perturbed ASD gene expression, suggesting that these differences could contribute to aberrant neural development or function in the affected individuals.

Discussion

In recent years, the genetic structure of autism spectrum disorder (ASD) risk in the general population has been clarified. This work has confirmed that while, in some cases, deleterious, single gene variants are significant contributors to ASD, the vast proportion of population attributable risk is polygenic [2, 51]. Furthermore, this risk is highly heritable, and individuals within a multiplex family typically exhibit variable degrees of affectation [31]. Here, we modeled cellular and molecular correlates of ASD within one such multiplex family, performing cortical neural differentiation of iPSCs derived from several family members with differential affectation. In this family, both polygenic liability and a shared variant of uncertain significance (VUS) may contribute to risk. In cells derived from the affected individuals, we identified compromised responses to differentiation cues and altered gene expression profiles during iPSC differentiation into cortical excitatory (cExN) and inhibitory (cINs) neurons, compared to related and unrelated unaffected controls. This work demonstrates that iPSC-based modeling can be used to characterize these more genetically complex but prevalent forms of ASD, in addition to modeling simplex and monogenic forms, which have been the focus of most studies to date. Moreover, these data provide information on physiologic and transcriptomic signatures of multiplex autism, with which cellular models derived from other families and other combinations of inherited susceptibility factors can be compared in future work.

Our phenotypic analysis of these four iPSC-based models of cortical neural development included assays conducted in the stem cells, during neural specification, in the proliferating NPCs, and during neuronal differentiation. During cExN NPC specification and during cIN NPC propagation, models from both affected individuals exhibited elevated fractions of cells with sub-G1 DNA
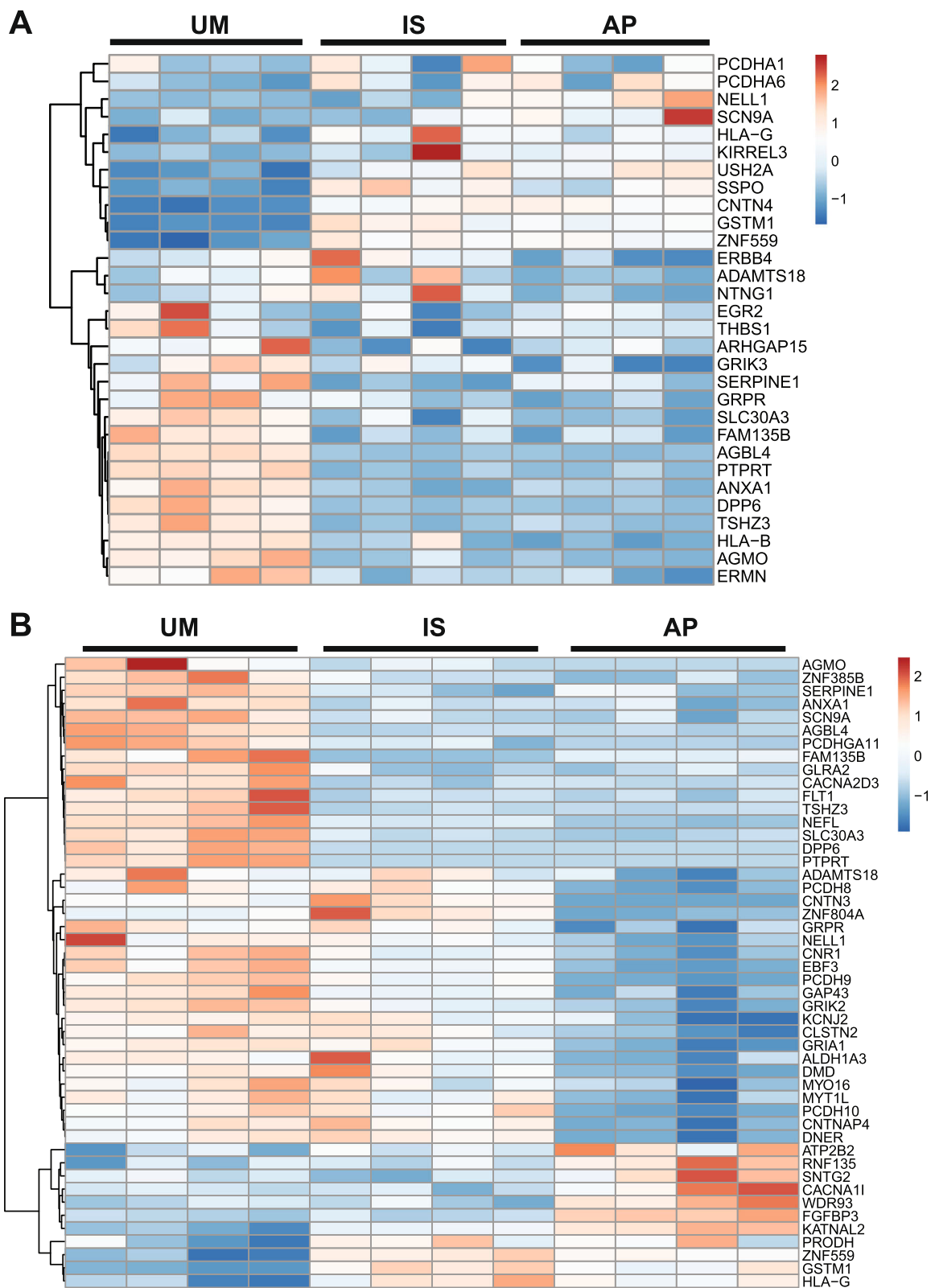
Fig. 6 Hierarchical clustering of DEGs that are also ASD genes in the SFARI autism gene database. **a** Relative gene expression for the cExN NPC samples. **b** Relative gene expression for the cIN NPC samples. Data from four biological replicates from one clonal line are shown for each sample type
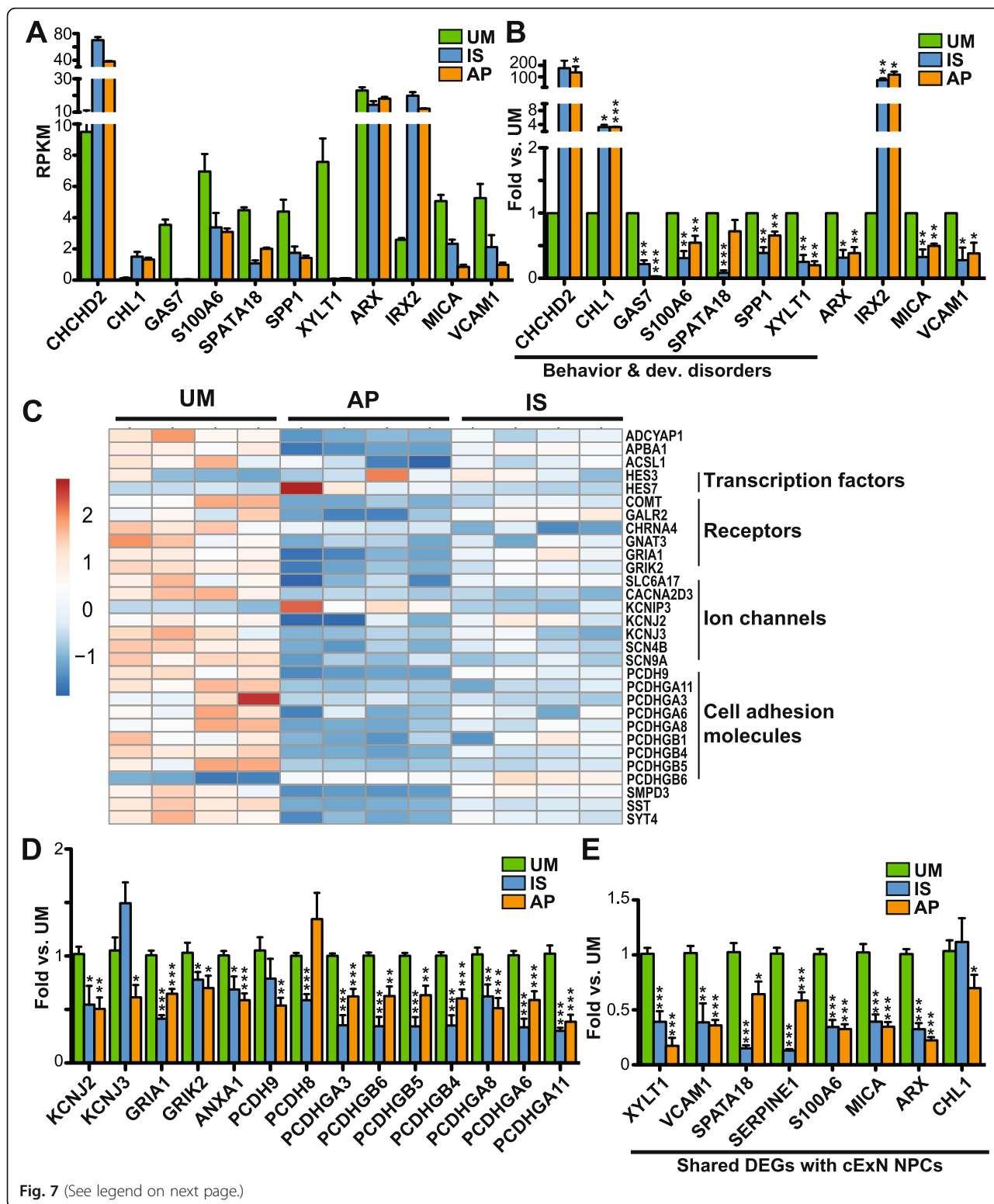
Fig. 7 (See legend on next page.)

content, relative to control-derived models. These data suggest that models derived from the affected individuals are less resistant to stressors, such as induction of differentiation, with these stressors increasing the propensity for cells to undergo apoptosis. While the molecular trigger for the induction of apoptosis here is unclear, expression of stress and apoptosis-related genes, such as *CHCHD2, ANXA1,* and *SPATA18* are dysregulated in these models [52–56]. These findings are reminiscent of some observations made in prior work, in which schizophrenia subject-derived iPSCs exhibited reduced neurosphere size [57] and increased apoptosis was observed in Williams-Syndrome iPSC-derived NPCs [58]. Interestingly, few studies report cellular alterations observed prior to the NPC stage, often focusing predominantly on phenotypes seen in NPCs and mature neurons [7–19]. While this may reflect a lack of earlier phenotypic changes in some models, our findings highlight the importance of tracking neurodevelopmental alterations from their earliest onset. A recent report underscores the value of using cellular modeling approaches that aim to recapitulate some aspects of in vivo neurodevelopment [20]. This study found that direct conversion of iPSCs into neurons masked ASD-associated cellular phenotypes, which were observable during the directed differentiation of iPSCs [20].

In our study, transcriptomic analysis of neural progenitor cells revealed dysregulated expression in affected individuals compared to controls of gene networks related to behavior, psychological disorders, and neuronal development and disease. Genes encoding transcription factors were among the neurodevelopment-related genes with reduced expression in both affected individuals. For example, ARX is required for normal telencephalic development and is associated with syndromic autism and other neurodevelopmental disorders [59], while EMX1 and FOXB1 also play important roles in neural development [60–62]. Behavioral misregulation is a key trait of ASD, and gene networks related to the GO term 'behavior' exhibited dysregulated expression in both affected individuals. Genes in these networks include *COMT, ADCYAP1, CNR1, HTR2C, GRIK2,* and *RGS4,* all of which are implicated in behavior-related phenotypes in humans and/or mice [63–71]. ASD genes were also dysregulated in these affected individuals, relative to controls.

Mutation of these genes in other individuals is implicated in autism risk or causation. These include adhesion-related genes (*PCDHA1, PCHDHA6, PCDHGA11, PCDH8, PCDH9, PCDH10* [72–76], *KIRREL3* [77], *CNTN3, CNTN*4 [78], *CNTNAP4* [79], and *THBS1* [80]), receptor and channel genes (*CACNA2D3* and *SCN9A* [81], *GRIK2* and *GRIK3* [82–85], *KCNJ2* [86, 87], and *GRIA1* [88, 89]), and genes associated with central nervous system development and axon guidance (*ERBB4* [90, 91], *NTNG1* [92], *TSHZ3* [93], *EBF3* [94], *MYT1L* [95, 96], and *ANXA1* [54–56]). Altered expression of ASD-associated genes has also been observed in cellular models derived from affected individuals in other studies [8, 13–16, 19, 97]. Therefore, these findings suggest that misregulated expression of suites of ASD-associated genes may contribute to risk or affectation, and may do so by altering neurodevelopment and/or neuronal function in these affected individuals.

A unique aspect of this study is the use of iPSC-based directed differentiation into both cExNs and cINs, enabling us to identify neural cell-type-specific alterations associated with affectation. Although DEGs identified in affected individuals in both neural cell types were associated with many similar functions and diseases (e.g., behavior), the specific DEGs obtained often varied by cell type. For example, cIN DEGs included many more ASD-associated genes and protocadherin genes, the latter of which control neuronal migration, axonal growth, and synapse formation [74, 75]. Human post-mortem cortical tissue from individuals with ASD has been shown to exhibit disrupted expression of cIN-associated genes, evidence that this cell type may commonly be disrupted in affected individuals in vivo [98]. These findings suggest that extending cellular modeling studies to multiple disease-relevant neuronal cell types, including cINs, may reveal additional neurodevelopmental disruptions related to affectation.

To define cellular and molecular perturbations commonly related to affectation, we compared our findings to other studies that modeled ASD by directed differentiation of iPSCs into cExNs. We identified subsets of overlapping DEGs in comparisons with studies involving idiopathic autism cases vs. controls (26 shared DEGs [7]), syndromic ASD involving macrocephaly (31 shared DEGs [19]), and modeling of mutation of the syndromic ASD gene *CHD8* (32 shared DEGs [13]) (Additional file

7: Table S6). Data for such comparisons is limited at present because iPSC-based models have been generated for a relatively small number of individuals and mutations, and these almost exclusively characterize cExNs or cerebral organoids [7, 13, 18, 19].

The multiplex pedigree studied here was subjected to clinical exome sequencing, as it was hypothesized that a single, shared, genetic contributor might mediate autism risk and differential phenotypic expression in this family. In this sequencing analysis, a thread of shared genetic liability among all children was a VUS in the ASD- and ID-associated gene, GPD2 [99–101], which was inherited from their mother. However, there is variable ASD expressivity among these individuals, ranging from absent, to intermediate, to severe. In addition, both males and females in the pedigree are variably affected, indicating the presence of other significant contributors to variation in severity of affectation within this family. This observation is consistent with recent evidence that genetic liability for ASD is prevalently polygenic, and that, even in multiplex pedigrees where a significant monogenic contributor has been identified, additional polygenic risk can contribute to affectation [51]. Moreover, this multiplex family was prototypic in reflecting the most severe form of affectation occurring in a male.

We hypothesized that it might be possible to identify graded cellular phenotypes that correlated with the level of severity of phenotypic expression. In general, we instead observed many cellular and molecular alterations that were shared by the cellular models derived from the affected individuals, while not being observed in those derived from the unaffected individuals. However, we did define some proband-specific DEGs, not present in the less severely affected sister, many of which relate to behavior and nervous system development. A subset of these DEGs had graded expression, exhibiting intermediate expression levels in the intermediate phenotype sister, between her unaffected mother and her severely affected brother. These findings suggest that both the degree of dysregulation of expression and the number and identity of DEGs within these networks may contribute to the level of affectation. While further experimentation might reveal additional graded phenotypes, particularly in mature neurons, ex vivo cellular modeling cannot recapitulate many aspects of fetal and post-natal neurodevelopment that may have been perturbed to contribute to the graded affectation observed in these individuals.

Limitations
This work highlights several considerations for ongoing scientific efforts to model this complex but prevalent form of ASD in future studies. First, since the unique characteristics of any multiplex ASD pedigree present challenges for cellular modeling, it is important to control for sex and variation in affectation in subject and family selection, study design, and analysis. Related to this point is the importance of modeling affected females in such studies. Most ASD cellular modeling to date has been restricted to affected males [7–10, 13, 14, 16, 18, 19], given the increased prevalence of ASD among males, and the fact that constraint to a single sex simplifies some modeling considerations. In particular, sex chromosome dosage effects do not need to be accounted for in male cells, while female-derived iPSC models cannot currently recapitulate the process of random X-chromosome inactivation that occurs in developing somatic tissues, including the brain [102, 103]. However, the transcriptomic differences that we observed here were not driven by sex chromosome-linked gene expression: very few DEGs in any potential pairwise sample comparison (whether between same or opposite sex models) were sex chromosome-linked and/or potential contributors to sex-biased gene expression in the human brain [104, 105]. Therefore, this work supports the feasibility of identifying DEGs associated with affectation by cellular model cross-comparisons, even when these models are derived from both female and male subjects.

Furthermore, studying both unaffected and affected male and female individuals within a multiplex family, as in this pedigree, necessitates consideration of how the so-called female protective effect may contribute to affectation [106–108]. Multiplex ASD families often exhibit differences in phenotypic expression of additive genetic liability, some of which appear to be related to sex. While the most affected individual in this family is male, we cannot fully differentiate the contributors to variable phenotypic expression of ASD in this pedigree, which also includes an unaffected male and affected and unaffected females. However, this experimental system could identify affectation-linked phenotypes that may have involved a sex-based contributor, since both the cellular and behavioral phenotypes of the females studied here differed from that of the affected male. Future studies in strategically-selected samples such as these provide an opportunity to assess how sex influences phenotypic expression of ASD genetic liability; if such mechanisms exist at the level of the cell, this approach may yield insights into sex-specific interventions.

Another consideration for iPSC-based modeling of ASD is genetic background, which can be a confounding variable for cross-comparisons [42]. In this pedigree, ASD risk was polygenic, such that it was not possible to engineer a correction of a single genome variant to create pairs of isogenic mutant versus wild-type iPSC lines with an identical genetic background for study. In such cases, modeling of first degree relatives may serve as the best control, and modeling of multiple related

individuals with varying affectation provides additional opportunities for identifying potential contributors to these differences in affectation. Including unrelated controls and performing comparisons with other studies can further highlight which phenotypic and transcriptomic alterations track with affectation, even by comparison with models derived from individuals with an unrelated genetic background.

Conclusions

In summary, this work used robust schemes for differentiation of cortical neurons from iPSCs to model cellular and molecular signatures associated with multiplex ASD in a family reflecting varying degrees of affectation. Even in this prevalent, complex form of ASD, involving heritability, polygenic etiology, and variable affectation, we could identify affectation-linked cellular and molecular alterations of neurodevelopment, some of which overlapped those defined in other iPSC-based studies of monogenic, syndromic, and de novo ASD. As more cellular models of ASD are characterized, these data can be harnessed in the search for convergent and divergent contributors to impairment across the genetically complex and multi-factorial pathways that give rise to ASD.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13229-019-0306-0.

Additional file 1: Table S1. (A) Antibodies used in immunocytochemistry and immunoblotting experiments, (B) Statistical analysis data for all panels, (C) Clonal line usage for each experiment and biological replicate experiments conducted to obtain data reported in each panel.

Additional file 2: Figures S1-S8. Figure legends provided in file.

Additional file 3: Table S2. DEGs from pairwise comparisons of the four sample types for cExN or cIN NPCs. (A) Summary of DEGs obtained across all pairwise comparisons, with Log2-fold change (FC) and adjusted *P* value/FDR for each gene/comparison, and average RPKM values across sample types for all genes shown. (B-M) Individual pairwise dataset comparisons, including the log2FC and standard error calculations (lfcSE), *P* value and adjusted *P* value/FDR, and averaged RPKM values for all sample types.

Additional file 4: Table S3. Hierarchical clustering of genes that had similar expression in the AP and IS, which differed from expression in the UM +/− the UC. Relative expression of DEGs across the four sample types for cExN and cIN NPCs was assessed using ClustVis. Order of genes in the cluster and unit variance scaled relative expression values are indicated. Genes located on the X-chromosome are indicated.

Additional file 5: Table S4. Information about selected DEGs within IPA networks.

Additional file 6: Table S5. (A) Clusters of co-expressed DEGs in cExN and cIN NPCs, with (B-G) associated ToppGene GO-, disease-, and pathway-associated terms. (H-I) DEGs associated with ASD. DEGs specific to the AP (Fig. 5A, E) and DEGs with similar expression in the AP and IS that differed from expression in the UM +/− UC (Fig. 4a, e) were visualized by hierarchical clustering in Fig. 6 and were compared with ASD-associated genes in the SFARI gene database [43], with the Geisinger Developmental Brain Disorder Genes Database [44], and with adult-onset psychiatric disorder-associated genes from PsyGeNET [45, 46]. RPKM values for each gene and sample are shown. The SFARI gene database

indicates how each gene is associated with ASD (genetic category, gene score, and number of reports). The Geisinger database indicates the pattern of inheritance of mutations in each gene and the number of reports linking each gene to intellectual disability (ID), ASD, epilepsy (EP), attention-deficit/hyperactivity disorder (ADHD), schizophrenia (SCZ), or bipolar disorder (BD). The PsyGeNET database indicates genes associated with the disorders shown, with 'unique 4 or 5 abstract' indicating higher confidence associations.

Additional file 7: Table S6. cExN DEG comparison to other studies.

Abbreviations
AP: Affected proband; ASD: Autism spectrum disorder; cExNs: Cortical excitatory neurons; cINs: Cortical inhibitory neurons; DEG: Differentially expressed gene; EBs: Embryoid bodies; ES: Exome sequencing; GEiC: Genome engineering and iPSC Center; GO: Gene ontology; GTAC: Genome Technology Access Center; ICC: Immunocytochemistry; IPA: Ingenuity Pathway Analysis; iPSCs: Induced pluripotent stem cells; IS: Intermediate phenotype sister; NPCs: Neural progenitor cells; PCA: Principal component analysis; PI: Propidium iodide; RT-qPCR: Reverse transcription and quantitative PCR; SFARI: Simons Foundation Autism Research Initiative; UC: Unaffected control; UM: Unaffected mother; VUS: Variant of uncertain significance

Lewis *et al. Molecular Autism*   (2019) 10:51

Page 21 of 23

Author details
[1]Department of Developmental Biology, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, USA. [2]Department of Pediatrics, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, USA. [3]Department of Neuroscience, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, USA. [4]Department of Psychiatry, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, USA.

References
1. Wing L, Gould J. Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. J Autism Dev Disord. 1979;9(1):11–29.
2. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, et al. Most genetic risk for autism resides with common variation. Nat Genet. 2014; 46(8):881–5.
3. Donovan AP, Basson MA. The neuroanatomy of autism - a developmental perspective. J Anat. 2017;230(1):4–15.
4. Ecker C, Bookheimer SY, Murphy DG. Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. Lancet Neurol. 2015;14(11):1121–34.
5. Zikopoulos B, Barbas H. Altered neural connectivity in excitatory and inhibitory cortical circuits in autism. Front Hum Neurosci. 2013;7:609.
6. Canitano R, Pallagrosi M. Autism Spectrum Disorders and Schizophrenia Spectrum Disorders: Excitation/Inhibition Imbalance and Developmental Trajectories. Front Psychiatry. 2017;8:69.
7. DeRosa BA, El Hokayem J, Artimovich E, Garcia-Serje C, Phillips AW, Van Booven D, et al. Convergent Pathways in Idiopathic Autism Revealed by Time Course Transcriptomic Analysis of Patient-Derived Neurons. Sci Rep. 2018;8(1):8423.
8. Liu X, Campanac E, Cheung HH, Ziats MN, Canterel-Thouennon L, Raygada M, et al. Idiopathic Autism: Cellular and Molecular Phenotypes in Pluripotent Stem Cell-Derived Neurons. Mol Neurobiol. 2017;54(6):4507–23.
9. Aksoy I, Utami KH, Winata CL, Hillmer AM, Rouam SL, Briault S, et al. Personalized genome sequencing coupled with iPSC technology identifies GTDC1 as a gene involved in neurodevelopmental disorders. Hum Mol Genet. 2017;26(2):367–82.
10. Griesi-Oliveira K, Acab A, Gupta AR, Sunaga DY, Chailangkarn T, Nicol X, et al. Modeling non-syndromic autism and the impact of TRPC6 disruption in human neurons. Mol Psychiatry. 2015;20(11):1350–65.
11. Deshpande A, Yadav S, Dao DQ, Wu ZY, Hokanson KC, Cahill MK, et al. Cellular Phenotypes in Human iPSC-Derived Neurons from a Genetic Model of Autism Spectrum Disorder. Cell Rep. 2017;21(10):2678–87.
12. Kathuria A, Nowosiad P, Jagasia R, Aigner S, Taylor RD, Andreae LC, et al. Stem cell-derived neurons from autistic individuals with SHANK3 mutation show morphogenetic abnormalities during early development. Mol Psychiatry. 2018;23(3):735–46.
13. Wang P, Lin M, Pedrosa E, Hrabovsky A, Zhang Z, Guo W, et al. CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in neurodevelopment. Mol Autism. 2015;6:55.
14. Wang P, Mokhtari R, Pedrosa E, Kirschenbaum M, Bayrak C, Zheng D, et al. CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in cerebral organoids derived from iPS cells. Mol Autism. 2017;8:11.
15. Woodbury-Smith M, Deneault E, Yuen RKC, Walker S, Zarrei M, Pellecchia G, et al. Mutations in RAB39B in individuals with intellectual disability, autism spectrum disorder, and macrocephaly. Mol Autism. 2017;8:59.
16. Deneault E, White SH, Rodrigues DC, Ross PJ, Faheem M, Zaslavsky K, et al. Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons. Stem Cell Reports. 2018;11(5):1211–25.
17. Russo FB, Freitas BC, Pignatari GC, Fernandes IR, Sebat J, Muotri AR, et al. Modeling the Interplay Between Neurons and Astrocytes in Autism Using Human Induced Pluripotent Stem Cells. Biol Psychiatry. 2018;83(7):569–78.
18. Marchetto MC, Belinson H, Tian Y, Freitas BC, Fu C, Vadodaria K, et al. Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. Mol Psychiatry. 2017;22(6):820–35.
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):375–90.
20. Schafer ST, Paquola ACM, Stern S, Gosselin D, Ku M, Pena M, et al. Pathological priming causes developmental gene network heterochronicity in autistic subject-derived neurons. Nat Neurosci. 2019.
21. Christensen DL, Baio J, Van Naarden BK, Bilder D, Charles J, Constantino JN, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years--Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. MMWR Surveill Summ. 2016;65(3):1–23.
22. Constantino JN, Todorov A, Hilton C, Law P, Zhang Y, Molloy E, et al. Autism recurrence in half siblings: strong support for genetic mechanisms of transmission in ASD. Mol Psychiatry. 2013;18(2):137–8.
23. Gronborg TK, Schendel DE, Parner ET. Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. JAMA Pediatr. 2013;167(10):947–53.
24. Constantino JN. Recurrence rates in autism spectrum disorders. Jama. 2014; 312(11):1154–5.
25. Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P. Sibling recurrence and the genetic epidemiology of autism. Am J Psychiatry. 2010;167(11): 1349–56.
26. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, et al. Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. Pediatrics. 2011;128(3):e488–95.
27. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014;515(7526):209–15.
28. Mefford HC, Batshaw ML, Hoffman EP. Genomics, intellectual disability, and autism. N Engl J Med. 2012;366(8):733–43.
29. Turner TN, Eichler EE. The Role of De Novo Noncoding Regulatory Mutations in Neurodevelopmental Disorders. Trends Neurosci. 2018.
30. Williams SM, An JY, Edson J, Watts M, Murigneux V, Whitehouse AJO, et al. An integrative analysis of non-coding regulatory DNA variations associated with autism spectrum disorder. Mol Psychiatry. 2018.
31. Griesi-Oliveira K, Sertie AL. Autism spectrum disorders: an updated guide for genetic counseling. Einstein (Sao Paulo). 2017;15(2):233–8.
32. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016;18(7):696–704.
33. Meganathan K, Lewis EMA, Gontarz P, Liu S, Stanley EG, Elefanty AG, et al. Regulatory networks specifying cortical interneurons from human embryonic stem cells reveal roles for CHD2 in interneuron development. Proc Natl Acad Sci U S A. 2017;114(52):E11180–9.
34. 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.
35. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22(9):1760–74.
36. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 2013;41(10):e108.
37. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
38. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
39. Chen J, Xu H, Aronow BJ, Jegga AG. Improved human disease candidate gene prioritization using mouse phenotype. BMC Bioinformatics. 2007;8:392.
40. Hoffman GE, Schadt EE. variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics. 2016;17(1):483.
41. Lunden JW, Durens M, Phillips AW, Nestor MW. Cortical interneuron function in autism spectrum condition. Pediatr Res. 2019;85(2):146–54.
42. Hoekstra SD, Stringer S, Heine VM, Posthuma D. Genetically-Informed Patient Selection for iPSC Studies of Complex Diseases May Aid in Reducing Cellular Heterogeneity. Front Cell Neurosci. 2017;11:164.
43. SFARI gene Database. https://www.sfari.org/resource/sfari-gene/. 08-22-2018 database release, exported 10-17-2018.
44. Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, Martin CL. A Cross-Disorder Method to Identify Novel Candidate Genes for Developmental Brain Disorders. JAMA Psychiatry. 2016;73(3):275–83.

45. Gutierrez-Sacristan A, Bravo A, Portero-Tresserra M, Valverde O, Armario A, Blanco-Gandia MC, et al. Text mining and expert curation to develop a database on psychiatric diseases and their genes. Database (Oxford). 2017;2017.

46. Gutierrez-Sacristan A, Grosdidier S, Valverde O, Torrens M, Bravo A, Pinero J, et al. PsyGeNET: a knowledge platform on psychiatric disorders and their genes. Bioinformatics. 2015;31(18):3075–7.

47. Booth HDE, Wessely F, Connor-Robson N, Rinaldi F, Vowles J, Browne C, et al. RNA sequencing reveals MMP2 and TGFB1 downregulation in LRRK2 G2019S Parkinson's iPSC-derived astrocytes. Neurobiol Dis. 2019;129:56–66.

48. DeBoever C, Li H, Jakubosky D, Benaglio P, Reyna J, Olson KM, et al. Large-Scale Profiling Reveals the Influence of Genetic Variation on Gene Expression in Human Induced Pluripotent Stem Cells. Cell Stem Cell. 2017; 20(4):533–46 e7.

49. Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K. Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell. 2012;10(5):595–609.

50. Salomonis N, Dexheimer PJ, Omberg L, Schroll R, Bush S, Huo J, et al. Integrated Genomic Analysis of Diverse Induced Pluripotent Stem Cells from the Progenitor Cell Biology Consortium. Stem Cell Reports. 2016;7(1):110–25.

51. Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J, et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nat Genet. 2017;49(7):978–85.

52. Wang DB, Kinoshita C, Kinoshita Y, Morrison RS. p53 and mitochondrial function in neurons. Biochim Biophys Acta. 2014;1842(8):1186–97.

53. Liu Y, Clegg HV, Leslie PL, Di J, Tollini LA, He Y, et al. CHCHD2 inhibits apoptosis by interacting with Bcl-x L to regulate Bax activation. Cell Death Differ. 2015;22(6):1035–46.

54. Solito E, McArthur S, Christian H, Gavins F, Buckingham JC, Gillies GE. Annexin A1 in the brain--undiscovered roles? Trends Pharmacol Sci. 2008; 29(3):135–42.

55. Correia CT, Conceicao IC, Oliveira B, Coelho J, Sousa I, Sequeira AF, et al. Recurrent duplications of the annexin A1 gene (ANXA1) in autism spectrum disorders. Mol Autism. 2014;5(1):28.

56. Parente L, Solito E. Annexin 1: more than an anti-phospholipase protein. Inflamm Res. 2004;53(4):125–32.

57. Toyoshima M, Akamatsu W, Okada Y, Ohnishi T, Balan S, Hisano Y, et al. Analysis of induced pluripotent stem cells carrying 22q11.2 deletion. Transl Psychiatry. 2016;6(11):e934.

58. Chailangkarn T, Trujillo CA, Freitas BC, Hrvoj-Mihic B, Herai RH, Yu DX, et al. A human neurodevelopmental model for Williams syndrome. Nature. 2016; 536(7616):338–43.

59. Friocourt G, Poirier K, Rakic S, Parnavelas JG, Chelly J. The role of ARX in cortical development. Eur J Neurosci. 2006;23(4):869–76.

60. Zhang Y, Hoxha E, Zhao T, Zhou X, Alvarez-Bolado G. Foxb1 Regulates Negatively the Proliferation of Oligodendrocyte Progenitors. Front Neuroanat. 2017;11:53.

61. Takebayashi-Suzuki K, Kitayama A, Terasaka-Iioka C, Ueno N, Suzuki A. The forkhead transcription factor FoxB1 regulates the dorsal-ventral and anterior-posterior patterning of the ectoderm during early Xenopus embryogenesis. Dev Biol. 2011;360(1):11–29.

62. Kobeissy FH, Hansen K, Neumann M, Fu S, Jin K, Liu J. Deciphering the Role of Emx1 in Neurogenesis: A Neuroproteomics Approach. Front Mol Neurosci. 2016;9:98.

63. Hattori S, Takao K, Tanda K, Toyama K, Shintani N, Baba A, et al. Comprehensive behavioral analysis of pituitary adenylate cyclase-activating polypeptide (PACAP) knockout mice. Front Behav Neurosci. 2012;6:58.

64. Shintani N, Hashimoto H, Tanaka K, Kawagishi N, Kawaguchi C, Hatanaka M, et al. Serotonergic inhibition of intense jumping behavior in mice lacking PACAP (Adcyap1-/-). Ann N Y Acad Sci. 2006;1070:545–9.

65. Gogos JA, Morgan M, Luine V, Santha M, Ogawa S, Pfaff D, et al. Catechol-O-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. Proc Natl Acad Sci U S A. 1998;95(17):9991–6.

66. Qayyum A, Zai CC, Hirata Y, Tiwari AK, Cheema S, Nowrouzi B, et al. The Role of the Catechol-o-Methyltransferase (COMT) GeneVal158Met in Aggressive Behavior, a Review of Genetic Studies. Curr Neuropharmacol. 2015;13(6):802–14.

67. Bowers ME, Ressler KJ. Sex-dependence of anxiety-like behavior in cannabinoid receptor 1 (Cnr1) knockout mice. Behav Brain Res. 2016; 300:65–9.

68. Paes LA, Torre OHD, Henriques TB, de Mello MP, Celeri E, Dalgalarrondo P, et al. Association between serotonin 2C receptor gene (HTR2C) polymorphisms and psychopathological symptoms in children and adolescents. Braz J Med Biol Res. 2018;51(8):e7252.

69. Martin CB, Ramond F, Farrington DT, Aguiar AS Jr, Chevarin C, Berthiau AS, et al. RNA splicing and editing modulation of 5-HT(2C) receptor function: relevance to anxiety and aggression in VGV mice. Mol Psychiatry. 2013;18(6): 656–65.

70. Shaltiel G, Maeng S, Malkesman O, Pearson B, Schloesser RJ, Tragon T, et al. Evidence for the involvement of the kainate receptor subunit GluR6 (GRIK2) in mediating behavioral displays related to behavioral symptoms of mania. Mol Psychiatry. 2008;13(9):858–72.

71. Stratinaki M, Varidaki A, Mitsi V, Ghose S, Magida J, Dias C, et al. Regulator of G protein signaling 4 [corrected] is a crucial modulator of antidepressant drug action in depression and neuropathic pain models. Proc Natl Acad Sci U S A. 2013;110(20):8254–9.

72. Chen J, Yu S, Fu Y, Li X. Synaptic proteins and receptors defects in autism spectrum disorders. Front Cell Neurosci. 2014;8:276.

73. Tsai NP, Huber KM. Protocadherins and the Social Brain. Biol Psychiatry. 2017;81(3):173–4.

74. Washbourne P, Dityatev A, Scheiffele P, Biederer T, Weiner JA, Christopherson KS, et al. Cell adhesion molecules in synapse formation. J Neurosci. 2004;24(42):9244–9.

75. Wang X, Weiner JA, Levi S, Craig AM, Bradley A, Sanes JR. Gamma protocadherins are required for survival of spinal interneurons. Neuron. 2002;36(5):843–54.

76. Bruining H, Matsui A, Oguro-Ando A, Kahn RS, Van't Spijker HM, Akkermans G, et al. Genetic Mapping in Mice Reveals the Involvement of Pcdh9 in Long-Term Social and Object Recognition and Sensorimotor Development. Biol Psychiatry. 2015;78(7):485–95.

77. Martin EA, Muralidhar S, Wang Z, Cervantes DC, Basu R, Taylor MR, et al. Correction: The intellectual disability gene Kirrel3 regulates target-specific mossy fiber synapse development in the hippocampus. Elife. 2016;5.

78. Mohebiany AN, Harroch S, Bouyain S. New insights into the roles of the contactin cell adhesion molecules in neural development. Adv Neurobiol. 2014;8:165–94.

79. Karayannis T, Au E, Patel JC, Kruglikov I, Markx S, Delorme R, et al. Cntnap4 differentially contributes to GABAergic and dopaminergic synaptic transmission. Nature. 2014;511(7508):236–40.

80. Adams JC, Lawler J. The thrombospondins. Cold Spring Harb Perspect Biol. 2011;3(10):a009712.

81. Rubinstein M, Patoway A, Stanaway IB, McCord E, Nesbitt RR, Archer M, et al. Association of rare missense variants in the second intracellular loop of NaV1.7 sodium channels with familial autism. Mol Psychiatry. 2018;23(2):231–9.

82. Soto D, Altafaj X, Sindreu C, Bayes A. Glutamate receptor mutations in psychiatric and neurodevelopmental disorders. Commun Integr Biol. 2014; 7(1):e27887.

83. Contractor A, Mulle C, Swanson GT. Kainate receptors coming of age: milestones of two decades of research. Trends Neurosci. 2011;34(3):154–63.

84. Kim SA, Kim JH, Park M, Cho IH, Yoo HJ. Family-based association study between GRIK2 polymorphisms and autism spectrum disorders in the Korean trios. Neurosci Res. 2007;58(3):332–5.

85. Pravetoni M, Wickman K. Behavioral characterization of mice lacking GIRK/ Kir3 channel subunits. Genes Brain Behav. 2008;7(5):523–31.

86. Guglielmi L, Servettini I, Caramia M, Catacuzzeno L, Franciolini F, D'Adamo MC, et al. Update on the implication of potassium channels in autism: K(+) channelautism spectrum disorder. Front Cell Neurosci. 2015;9:34.

87. Binda A, Rivolta I, Villa C, Chisci E, Beghi M, Cornaggia CM, et al. A Novel KCNJ2 Mutation Identified in an Autistic Proband Affects the Single Channel Properties of Kir2.1. Front Cell Neurosci. 2018;12:76.

88. Ramanathan S, Woodroffe A, Flodman PL, Mays LZ, Hanouni M, Modahl CB, et al. A case of autism with an interstitial deletion on 4q leading to hemizygosity for genes encoding for glutamine and glycine neurotransmitter receptor sub-units (AMPA 2, GLRA3, GLRB) and neuropeptide receptors NPY1R, NPY5R. BMC Med Genet. 2004;5:10.

89. Gu X, Zhou L, Lu W. An NMDA Receptor-Dependent Mechanism Underlies Inhibitory Synapse Development. Cell Rep. 2016;14(3):471–8.

90. Perez-Garcia CG. ErbB4 in Laminated Brain Structures: A Neurodevelopmental Approach to Schizophrenia. Front Cell Neurosci. 2015;9:472.

91. Yau HJ, Wang HF, Lai C, Liu FC. Neural development of the neuregulin receptor ErbB4 in the cerebral cortex and the hippocampus: preferential expression by interneurons tangentially migrating from the ganglionic eminences. Cereb Cortex. 2003;13(3):252–64.

92. Yaguchi K, Nishimura-Akiyoshi S, Kuroki S, Onodera T, Itohara S. Identification of transcriptional regulatory elements for Ntng1 and Ntng2 genes in mice. Mol Brain. 2014;7:19.

93. Caubit X, Gubellini P, Andrieux J, Roubertoux PL, Metwaly M, Jacq B, et al. TSHZ3 deletion causes an autism syndrome and defects in cortical projection neurons. Nat Genet. 2016;48(11):1359–69.

94. Sleven H, Welsh SJ, Yu J, Churchill MEA, Wright CF, Henderson A, et al. De Novo Mutations in EBF3 Cause a Neurodevelopmental Syndrome. Am J Hum Genet. 2017;100(1):138–50.

95. Blanchet P, Bebin M, Bruet S, Cooper GM, Thompson ML, Duban-Bedu B, et al. MYT1L mutations cause intellectual disability and variable obesity by dysregulating gene expression and development of the neuroendocrine hypothalamus. PLoS Genet. 2017;13(8):e1006957.

96. Stevens SJ, van Ravenswaaij-Arts CM, Janssen JW, Klein Wassink-Ruiter JS, van Essen AJ, Dijkhuizen T, et al. MYT1L is a candidate gene for intellectual disability in patients with 2p25.3 (2pter) deletions. Am J Med Genet A. 2011; 155A(11):2739–45.

97. Germain ND, Chen PF, Plocik AM, Glatt-Deeley H, Brown J, Fink JJ, et al. Gene expression analysis of human induced pluripotent stem cell-derived neurons carrying copy number variants of chromosome 15q11-q13.1. Mol Autism. 2014;5:44.

98. Wang P, Zhao D, Lachman HM, Zheng D. Enriched expression of genes associated with autism spectrum disorders in human inhibitory neurons. Transl Psychiatry. 2018;8(1):13.

99. Nguyen NH, Brathe A, Hassel B. Neuronal uptake and metabolism of glycerol and the neuronal expression of mitochondrial glycerol-3-phosphate dehydrogenase. J Neurochem. 2003;85(4):831–42.

100. Barge-Schaapveld DQ, Ofman R, Knegt AC, Alders M, Hohne W, Kemp S, et al. Intellectual disability and hemizygous GPD2 mutation. Am J Med Genet A. 2013;161A(5):1044–50.

101. Daoud H, Gruchy N, Constans JM, Moussaoui E, Saumureau S, Bayou N, et al. Haploinsufficiency of the GPD2 gene in a patient with nonsyndromic mental retardation. Hum Genet. 2009;124(6):649–58.

102. Sahakyan A, Kim R, Chronis C, Sabri S, Bonora G, Theunissen TW, et al. Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. Cell Stem Cell. 2017;20(1):87–101.

103. Dandulakis MG, Meganathan K, Kroll KL, Bonni A, Constantino JN. Complexities of X chromosome inactivation status in female human induced pluripotent stem cells-a brief review and scientific update for autism research. J Neurodev Disord. 2016;8:22.

104. Tukiainen T, Villani AC, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. Nature. 2017;550(7675):244–8.

105. Gershoni M, Pietrokovski S. The landscape of sex-differential transcriptome and its consequent selection in human adults. BMC Biol. 2017;15(1):7.

106. Werling DM, Geschwind DH. Understanding sex bias in autism spectrum disorder. Proc Natl Acad Sci U S A. 2013;110(13):4868–9.

107. Gockley J, Willsey AJ, Dong S, Dougherty JD, Constantino JN, Sanders SJ. The female protective effect in autism spectrum disorder is not mediated by a single genetic locus. Mol Autism. 2015;6:25.

108. Jacquemont S, Coe BP, Hersch M, Duyzend MH, Krumm N, Bergmann S, et al. A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. Am J Hum Genet. 2014;94(3):415–25.

Publisher's Note