

# Targeted DNA Sequencing from Autism Spectrum Disorder Brains Implicates Multiple Genetic Mechanisms

Alissa M. D’Gama,<sup>1,2,3</sup> Sirisha Pochareddy,<sup>4</sup> Mingfeng Li,<sup>4</sup> Saumya S. Jamuar,<sup>1,2,3,5,6</sup> Rachel E. Reiff,<sup>1,2,3</sup> Anh-Thu N. Lam,<sup>1,2,3</sup> Nenad Sestan,<sup>4</sup> and Christopher A. Walsh<sup>1,2,3,\*</sup>

<sup>1</sup>Division of Genetics and Genomics, Manton Center for Orphan Disease, and Howard Hughes Medical Institute, Boston Children’s Hospital, Boston, MA 02115, USA

<sup>2</sup>Departments of Pediatrics and Neurology, Harvard Medical School, Boston, MA 02115, USA

<sup>3</sup>Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

<sup>4</sup>Departments of Neuroscience, Genetics, and Psychiatry, and Kavli Institute for Neuroscience, Yale School of Medicine, New Haven, CT 06510, USA

<sup>5</sup>Department of Paediatrics, KK Women’s and Children’s Hospital, Singapore

<sup>6</sup>Paediatrics Academic Clinical Programme, Duke-NUS Graduate School of Medicine, Singapore

\*Correspondence: [christopher.walsh@childrens.harvard.edu](mailto:christopher.walsh@childrens.harvard.edu)

<http://dx.doi.org/10.1016/j.neuron.2015.11.009>

## SUMMARY

Single nucleotide variants (SNVs), particularly loss-of-function mutations, are significant contributors to autism spectrum disorder (ASD) risk. Here we report the first systematic deep sequencing study of 55 postmortem ASD brains for SNVs in 78 known ASD candidate genes. Remarkably, even without parental samples, we find more ASD brains with mutations that are protein-altering (26/55 cases versus 12/50 controls,  $p = 0.015$ ), deleterious (16/55 versus 5/50,  $p = 0.016$ ), or loss-of-function (6/55 versus 0/50,  $p = 0.028$ ) compared to controls, with recurrent deleterious mutations in *ARID1B*, *SCN1A*, *SCN2A*, and *SETD2*, suggesting these mutations contribute to ASD risk. In several cases, the identified mutations and medical records suggest syndromic ASD diagnoses. Two ASD and one Fragile X premutation case showed deleterious somatic mutations, providing evidence that somatic mutations occur in ASD cases, and supporting a model in which a combination of germline and/or somatic mutations may contribute to ASD risk on a case-by-case basis.

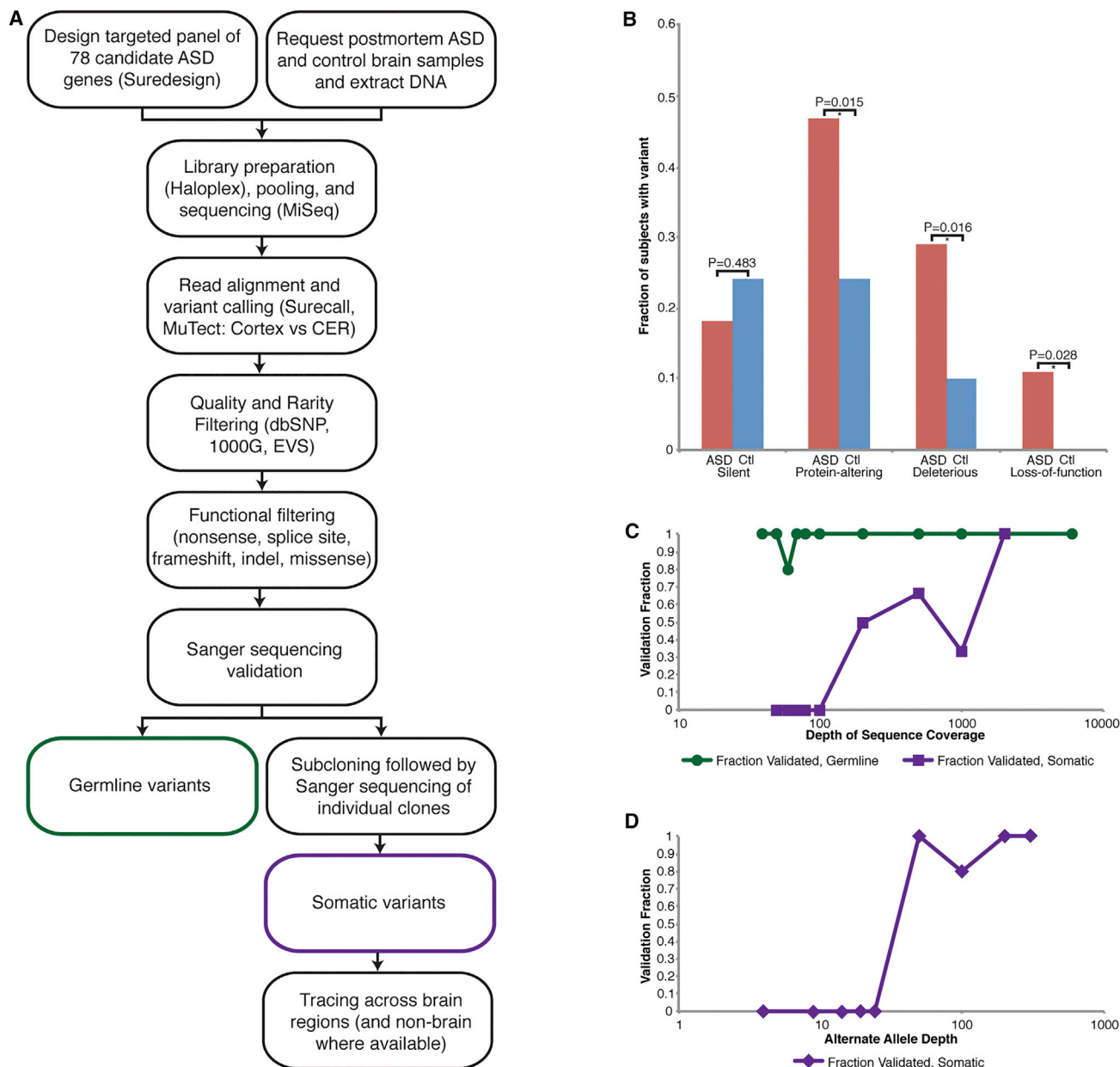
## INTRODUCTION

Autism spectrum disorder (ASD) affects  $\geq 1\%$  of American children and is characterized by (1) deficits in social interaction and social communication and (2) restricted and repetitive behaviors, interests, or activities (Wingate et al., 2014). ASD is genetically heterogeneous, with contributions from heritable mutations—including monogenic conditions, common variants, and rare recessive mutations—as well as from de novo mutations (Huguet et al., 2013). Whereas genetic studies of ASD risk have analyzed DNA from blood, saliva, or cell lines (De Rubeis et al., 2014; Ios-

sifov et al., 2014; O’Roak et al., 2011), postmortem ASD brain samples have been utilized mostly to study neuropathology, transcriptomics, and splicing (Chen et al., 2015; Irimia et al., 2014). Apart from a set of postmortem autism brain samples analyzed for copy number variants (CNVs) (Wintle et al., 2011), we know little of the underlying genetics of these postmortem autism brains. Identifying pathogenic mutations in these brains would aid genotype-phenotype correlation and interpretation of other alterations. In addition, somatic mutations—de novo mutations occurring postfertilization—are of growing interest in neurodevelopmental disorders that have similarly high de novo mutation rates to ASD, or with similar or overlapping symptoms, and have been reported in case studies with ASD (Poduri et al., 2013). Targeted deep sequencing allows detection of germline and somatic mutations in developmental brain disorders, especially when brain tissue is directly analyzed (D’Gama et al., 2015; Lee et al., 2012; Poduri et al., 2012; Rivière et al., 2012) and sequence coverage is very high (Jamuar et al., 2014). We used targeted deep sequencing of brain tissue from ASD patients and neurotypical controls to investigate genetic mechanisms underlying ASD risk in these patients, including possible contributions of somatic mutations.

## RESULTS

We performed ultradeep sequencing on DNA from postmortem cerebral cortical and/or cerebellar brain samples from 55 ASD cases, 50 neurotypical controls, and 17 cases with suspected ASD or related diagnoses (Figure 1A; see Table S1 available online). In order to achieve the high coverage necessary to sensitively detect somatic mosaic mutations, we used a targeted panel of 78 ASD candidate genes showing dominant or X-linked inheritance (see Supplemental Experimental Procedures) (Basu et al., 2009; Betancur, 2011; O’Roak et al., 2012). We achieved 547 $\times$  average coverage, with 95.7% of the analyzable target regions covered by  $\geq 100$  reads and 48.2% covered by  $\geq 1,000$  reads, providing sensitivity to detect somatic mutations in these



### Figure 1. Targeted Deep Sequencing Identifies Germline and Somatic Mutations in ASD Candidate Genes in Postmortem ASD Brain

(A) Study schematic. DNA extracted from ASD and control brain samples was sequenced at high coverage using a targeted panel of candidate ASD genes. After variant calling and filtering, protein-altering variants were validated using Sanger sequencing and subcloning, and somatic variants were traced across the brain and nonbrain tissues, when available.

(B) Fraction of cases and controls harboring synonymous, protein-altering, deleterious (subset of protein-altering), or loss-of-function (subset of deleterious) variants. p values were calculated using a two-tailed Fisher's exact test.

(C) Validation of germline and somatic variants according to total depth of sequence coverage. Essentially all germline variants validated; somatic variants were more likely to validate at higher total depth ( $\geq 100\times$ ).

(D) Validation of somatic variants according to alternate allele depth of coverage. Somatic variants were more likely to validate at alternate allele depth  $\geq 25\times$ . Each data point represents the validation fraction for all variants with (B) total or (C) alternate allele depth from the previous data point to that data point. See also [Table S1](#) and [Tables S3–S5](#).

genes to a level of  $\approx 5\%$  alternate allele frequency (AAF). We identified  $408 \pm 69$  SNVs and indels per sample and validated all rare, high-quality, protein-altering variants (missense,

nonsense, splice site, indel, and frameshift) using Sanger sequencing; any potential somatic variants were also validated with subcloning followed by Sanger sequencing of individual

**Table 1. Deleterious Mutations in ASD and Fragile X Cases**

Case	Diagnosis	Gene	Mutation	cDNA	Protein	Type
797	Autism	<i>ADNP</i>	Ms	c.3281G>T	p.G1094V	Germline
AN16641	Autism	<i>ARID1B</i>	Ms	c.4237C>T	p.P1413S	Germline
AN17515	Autism	<i>ARID1B</i>	Ms	c.1463C>G	p.P488R	Germline
5340	Autism/Parkinson's Disease	<i>ATP10A</i>	Ms	c.622C>T	p.R208W	Germline
AN16641	Autism	<i>CACNA1C</i>	Ms	c.3416G>A	p.R1139H	Germline
5006	Fragile X, premutation	<i>CACNA1C</i>	St	c.2T>C	(start loss)	Somatic
5278	Autism	<i>CACNA1H</i>	Ms	c.5909C>G	p.S1970C	Germline
UK28768	Autism	<i>CHD8</i>	Ms	c.2230G>A	p.V744I	Germline
5297	Autism	<i>GRIN2A</i>	Ms	c.2650G>A	p.D884N	Germline
797	Autism	<i>PQBP1</i>	Ms	c.731C>T	p.P244L	Germline
AN01570	Autism	<i>PTEN</i>	Ns	c.195C>A	p.Y65*	Germline
5278	Autism	<i>SCN1A</i>	Sp	c.602+1G>A	N/A	Somatic
AN16115	Autism	<i>SCN1A</i>	Ms	c.4319C>T	p.A1440V	Germline
4849	Autism	<i>SCN2A</i>	Ms	c.2021C>A	p.T674K	Germline
AN04166	Autism	<i>SCN2A</i>	Ms	c.5230G>A	p.G1744R	Germline
AN08043	Autism	<i>SCN2A</i>	Sp	c.4551+1G>C	N/A	Germline
AN09714	Autism	<i>SCN2A</i>	Ns	c.4543C>T	p.R1515*	Germline
1349	Autism	<i>SEMA5A</i>	Ns	c.586C>T	p.R196*	Germline
UK20244	Autism	<i>SETD2</i>	Ms	c.4871C>G	p.S1624C	Somatic
AN00090	Autism/Angelman Syndrome	<i>SETD2</i>	Ms	c.1463A>G	p.Y488C	Germline
AN16641	Autism	<i>SLC6A8</i>	Fs	c.205_206delG	p.A69Pfs*28	Germline

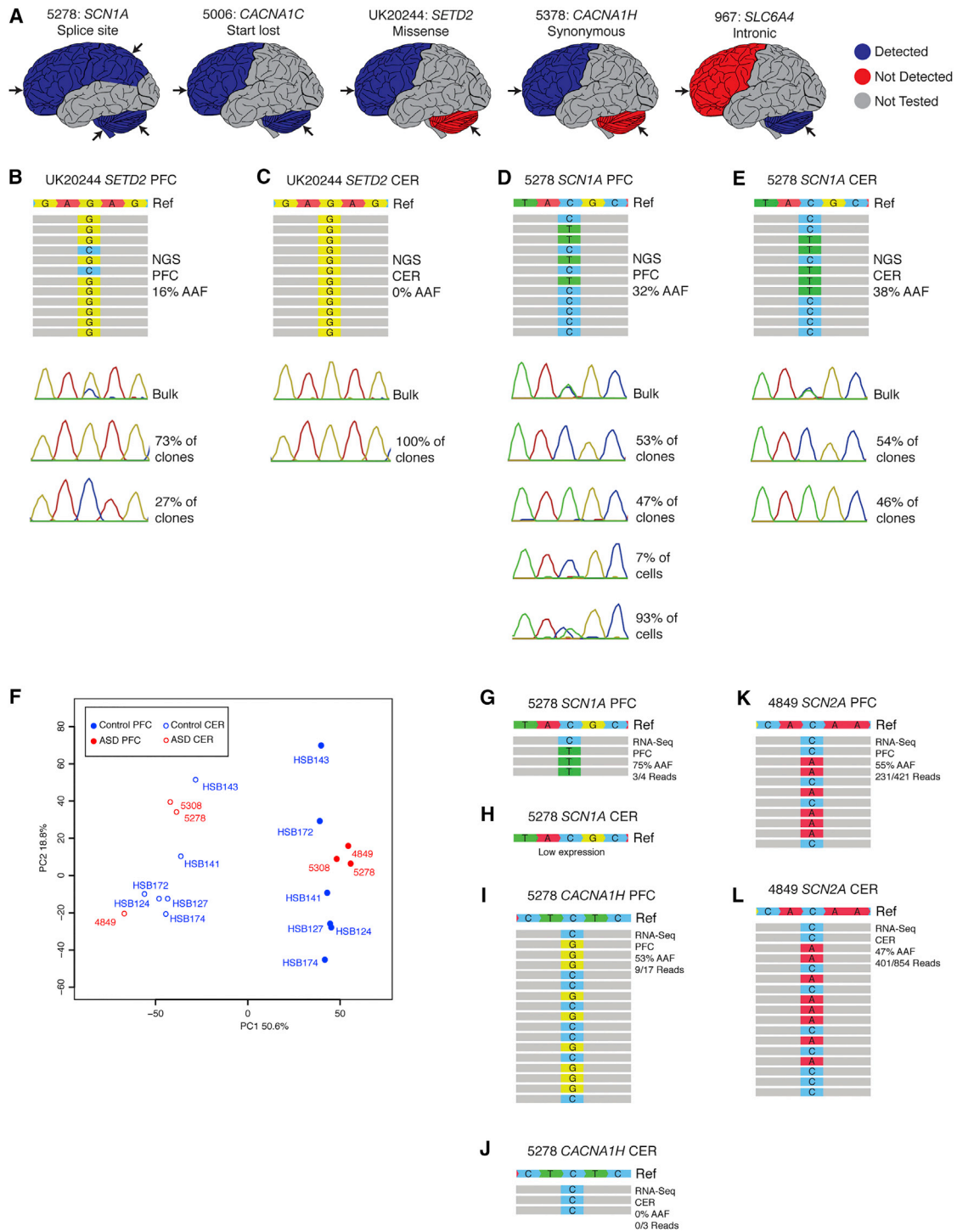
Details on the mutations detected in ASD and Fragile X cases that are predicted to be loss-of-function or deleterious by at least three prediction tools. Case AN00090 also has a 15q deletion consistent with Angelman Syndrome (Wintle et al., 2011). Del, deletion; Fs, frameshift; Ms, missense; Ns, nonsense; Sp, splicing; St, start lost. See also Figure S1 and Table S2.

clones, to determine mosaicism using a technology independent of the discovery platform.

Remarkably, despite the modest sample of ASD brains available for sequencing, and the absence of familial samples for testing segregation, damaging mutations were strikingly more common in ASD brains than control brains. Overall, we identified 34 protein-altering variants in 25 genes in ASD cases and 15 protein-altering variants in 13 genes in neurotypical controls (Table S2). Compared to controls, significantly more ASD cases harbored protein-altering variants (26/55 versus 12/50,  $p = 0.015$ ), deleterious variants (16/55 versus 5/50,  $p = 0.016$ )—i.e., loss-of-function and missense predicted to be deleterious by at least three prediction tools—and loss-of-function variants (nonsense, splice site, and frameshift, 6/55 versus 0/50,  $p = 0.028$ ), whereas functionally neutral variants were equally common in cases and controls (10/55 versus 12/50,  $p = 0.483$ , two-tailed Fisher's exact tests) (Figure 1B; Table S3). The excess of protein-altering variants in ASD cases compared to controls was driven by the deleterious and particularly the loss-of-function variants, consistent with previous reports (Sanders et al., 2012). Recurrent loss-of-function and/or deleterious missense variants occurred in four genes (*ARID1B*, *SCN1A*, *SCN2A*, and *SETD2*) in ASD cases (Table 1), consistent with previous Whole Exome Sequencing (WES) studies (De Rubeis et al., 2014; Iossifov et al., 2014). We also identified protein-altering variants in cases with suspected ASD or related diagnoses (Table S2). Although parental samples are not available to test for the de novo nature of these variants, the large excess of deleterious var-

iants in ASD brains strongly suggests that most of the deleterious and loss-of-function variants contributed to ASD pathogenesis.

To identify de novo somatic variants, we validated all rare, high-quality, protein-altering variants with AAF <40% (see Supplemental Experimental Procedures). Although NGS identified many potential somatic variants, extensive validation showed some to be false positives or germline and not bona fide somatic (Tables S4 and S5). Potential somatic variants with a higher total read depth ( $\geq 100$ , Figure 1C) and higher alternate allele read depth ( $\geq 25$ , Figure 1D) were more likely to validate. We identified protein-altering somatic variants in this 78 gene targeted panel in two ASD cases (Figure 2A; Table S6), with no protein-altering somatic variants identified in control brains even after extensive validation. Case UK20244, a 16-year-old male with ASD, harbors a somatic missense variant in *SETD2*, predicted to be deleterious, present in frontal cortex but undetectable in cerebellum (Figure 2B and 2C), which may have contributed to his ASD symptoms and death. Case 5278, a 15-year-old female with ASD, harbors a somatic splice site mutation in *SCN1A*, present across brain and in nonbrain tissues at relatively high AAF (Figures 2D and 2E). We confirmed the mosaic nature of this high AAF mutation by fluorescent-activated nuclear sorting (FANS), using antisera to NeuN to separate neuronal and nonneuronal nuclei from the prefrontal cortex, followed by multiple displacement amplification of single cell genomes (Dean et al., 2001; Evrony et al., 2012). Genotyping genome-amplified DNA from single nuclei confirmed that the mutation was mosaic, with  $93\% \pm 1.6\%$  of cells (corrected for allelic



**Figure 2. Regional Distribution of Somatic Mutations and Transcriptomic Analysis of ASD Brains**

(A) Protein-altering somatic mutations in ASD (5278, UK20244), suspected ASD (967), Social Anxiety Disorder (5378), or Fragile X premutation (5006) brain. Blue indicates brain regions tested where the mutation was detected, red indicates brain region tested where the mutation was not detected, and gray indicates regions not tested. Arrows show approximate location of the sample tested.

(B–E) Detection and validation of protein-altering somatic mutations in ASD brains. Each panel depicts representative NGS reads, Sanger sequencing trace from bulk brain tissue, and Sanger sequencing traces from individual clones following subcloning for case UK20244, mutation in *SETD2*, (B) PFC, and (C) CER; case 5278, mutation in *SCN1A*, (D) PFC, and (E) CER. In addition, (D) depicts single nuclei Sanger sequencing traces for case 5278, mutation in *SCN1A*, PFC.

(legend continued on next page)

**Table 2. Syndromic ASD Cases Suggested by Genotype-Phenotype Correlation**

Case	Gender	Age	Cause of Death	Mutation	Associated Syndrome	Clinical Symptoms
797	M	9	Drowning	<i>ADNP</i> Ms, <i>PQBP1</i> Ms	HVDAS, RENS1, ID (Redin et al., 2014)	Language and developmental delay, ADD, mild hypotonia
AN01570	F	18	SUDEP	<i>PTEN</i> LOF	Cowden syndrome (Tan et al., 2011), autism with macrocephaly	Large brain size noted at autopsy, epilepsy
AN16641	M	9	SUDEP	<i>ARID1B</i> Ms, <i>CACNA1C</i> Ms, <i>SLC6A8</i> LOF	MRD12 (Timothy syndrome), CCDS1	FH seizures and speech delay, febrile seizures from 3 years, language delay, hyperactivity, ID, multiplex developmental disorder
5278	F	15	Drowning, seizure disorder	<i>CACNA1H</i> Ms, <i>SCN1A</i> LOF	Epilepsy susceptibility; Dravet syndrome, GEFSP2, <i>FEB3A</i> (Fujiwara et al., 2003; Depienne et al., 2010; Allen et al., 2013)	Progressed from febrile seizures in infancy to nonfebrile and generalized tonic-clonic seizures
AN16115	F	11	Drowning, seizure disorder	<i>SCN1A</i> Ms	Dravet syndrome (Gaily et al., 2013), GEFSP2, <i>FEB3A</i>	FH seizures and neuropsychiatric disease, progressed from febrile seizures in infancy to febrile, tonic-clonic, and potentially myoclonic seizures, mild ID, hyperlexia
AN09714	M	60	Pancreatic cancer	<i>SCN2A</i> LOF ( <i>SGSM3</i> Ms predicted to be benign)	EIEE11, BFIS3	Seizures from 3 years old, likely episode of status epilepticus, delayed motor milestones, ID

Details on cases in which medical records and syndromes associated with mutated genes suggest syndromic ASD diagnoses. ADD, attention deficit disorder; BFIS3, seizures, benign familial infantile, 3; CCDS1, cerebral creatine deficiency syndrome 1; EIEE11, epileptic encephalopathy, early infantile, 11; FEB3A, febrile seizures, familial, 3A; FH, family history; GEFSP2, epilepsy, generalized, with febrile seizures plus, type 2; HVDAS, Helsmoortel-van der Aa syndrome; ID, intellectual disability; LOF, loss-of-function; MRD12, mental retardation, autosomal dominant 12; Ms, missense; RENS1, Renpenning syndrome; and SUDEP, sudden unexplained death in epilepsy.

dropout; see [Experimental Procedures](#)) harboring the mutation and the remainder not carrying the mutation (Figure 2D). This mutation has been previously reported in epileptic encephalopathies (Allen et al., 2013; Depienne et al., 2010; Fujiwara et al., 2003), and indeed case 5278 suffered generalized epilepsy in addition to ASD. While we focused our attention on protein-altering somatic variants, we also validated a synonymous variant in *CACNA1H* present in frontal cortex but not detected in cerebellum of case 5378, a 22-year-old male with Social Anxiety Disorder, and an intronic variant in *SLC6A4*, predicted to be benign, present in cerebellum but not detected in frontal cortex of case 967, a 32-year-old male with suspected autism (Figure 2A; Table S6).

Review of available medical records and OMIM identified consistent genotype-phenotype correlations in many cases with deleterious and loss-of-function variants (Table 2). Case 797, a 9-year-old male with autism, harbors missense mutations in *ADNP* and *PQBP1* predicted to be deleterious. Mutations (mostly loss-of-function) in both genes are associated with intellectual disability (ID) syndromes, and the same *PQBP1* missense mutation we identified was recently reported as potentially causative in two brothers with ID (Redin et al., 2014). While medical

records did not indicate non-nervous system symptoms, which are commonly seen in the associated syndromes, case 797 had language and developmental delay, attention deficit disorder, and mild hypotonia, suggesting the missense mutations may contribute to a milder form of the associated syndromes. Case AN01570, an 18-year-old female with autism, harbors a nonsense mutation in *PTEN* previously reported in Cowden syndrome (Tan et al., 2011). She had her first seizure at age 8 and was diagnosed with epilepsy, and her autopsy records noted her large brain size (2,100 g at 18 years old; head circumference measurement was not available), suggestive of autism with macrocephaly associated with *PTEN* mutations. Case AN16641, a 9-year-old male with autism, harbors a frameshift mutation in *SLC6A8* and missense mutations in *ARID1B* and *CACNA1C* predicted to be deleterious. He has a family history of seizures and speech delay, and developed febrile seizures at almost 3 years of age. He was noted to be hyperactive and was diagnosed with language delay, ID, and a multiplex developmental disorder. His symptoms suggest a diagnosis of cerebral creatine deficiency syndrome 1, associated with mutations in *SLC6A8*, and potential contribution from *ARID1B*, associated with ID.

(F) PCA plot of samples analyzed by RNA sequencing.

(G–L) Detection of mutations in RNA-seq data. Each panel depicts representative RNA-seq reads and read counts for case 5278, mutation in *SCN1A*, (G) PFC and (H) CER; case 5278, mutation in *CACNA1H*, (I) PFC and (J) CER; case 4849, mutation in *SCN2A*, (J) PFC and (L) CER. CER, cerebellum; PFC, prefrontal cortex. See also Table S6 and Table S7.

Three cases harbor mutations in ion channel genes associated with seizure disorders, and presented with symptoms suggestive of epilepsy. Case 5278, a 15-year-old female with autism, harbors a missense mutation in *CACNA1H* predicted to be deleterious and a somatic splicing mutation in *SCN1A*, previously reported in epileptic encephalopathies, as mentioned above (Allen et al., 2013; Depienne et al., 2010; Fujiwara et al., 2003). After febrile seizures in early infancy, she developed nonfebrile and generalized tonic-clonic seizures, suggestive of generalized epilepsy with febrile seizures plus type 2 (GEFSP2). Case AN16115, an 11-year-old female with autism, harbors a missense mutation in *SCN1A* predicted to be deleterious; a mutation that leads to a different amino acid at the same position was previously reported in Dravet syndrome (Gaily et al., 2013). She had a family history of seizures and neuropsychiatric disease, and was diagnosed with mild ID and hyperlexia. She had her first febrile seizure at 4.5 months old, and developed febrile, tonic-clonic, and potentially myoclonic seizures, consistent with a *SCN1A* mutation. Case AN09714, a 60-year-old male with autism and a nonsense mutation in *SCN2A*, showed delayed motor milestones and ID. His first seizure was at age 3, and at 15 he was hospitalized with pneumonia and experienced seizures every 5 min for a 12 hr period, suggestive of an *SCN2A*-associated syndrome like epileptic encephalopathy, early infantile, 11 (EIEE11). These three cases highlight the complex relationship between ASD and seizures (Jeste and Geschwind, 2014). Based on available medical records, almost half of the ASD cases had seizures, while there was no evidence of seizures in controls.

Fragile X accounts for ~2% of ASD cases (Huguet et al., 2013), and the Fragile X mutation is subject to somatic expansion (Lokanga et al., 2013), so we also tested postmortem brain samples from ASD and ten Fragile X cases for expansion of the CGG repeat in the 5' UTR of *FMR1*. We did not detect unsuspected expansion in any ASD cases, but did find mosaic expansion in three Fragile X mutation and premutation cases (Figure S1). Case 1421, a 69-year-old male with Fragile X, harbored a mosaic full mutation in prefrontal cortex. Case 5006, an 85-year-old male diagnosed with a premutation, harbored a mosaic full mutation in both prefrontal cortex and cerebellum. We also detected a somatic null mutation in *CACNA1C* in case 5,006 that disrupts the start methionine, present in both prefrontal cortex and cerebellum (Table S6), suggesting that this mutation might contribute to pathogenesis. Case 4664, a 71-year-old male diagnosed with a premutation and an unspecified neurological disorder, harbored a mosaic full mutation in prefrontal cortex not detected in cerebellum. These data show that Fragile X premutations and mutations can show somatic instability in the brain, with potentially distinct effects in different brain regions.

To test if mutations were detectable in RNA and hence affect the transcriptome, we performed RNA-sequencing (RNA-seq) on total RNA with acceptable RNA integrity (RIN  $\geq$  5) extracted from prefrontal cortex and cerebellar samples of ASD cases with protein-altering somatic mutations and/or sodium channel mutations. The ASD cases were compared to matched (age and sex) neurotypical controls (Table S1). After quality control, batch correction, and quantile normalization (see Supplemental Experimental Procedures), the expression values for protein coding genes were calculated. Consistent with previous findings

(Kang et al., 2011), principal component analysis (PCA) confirmed that both ASD and control PFC and CER samples have distinctive transcriptional profiles. PCA also revealed that the ASD PFC and CER samples clustered based on their regional identity and not the disease state (Figure 2F), indicating that changes in the transcriptome associated with ASD are smaller than global transcriptional differences between the two analyzed regions.

RNA-seq data analysis confirmed the presence of several germline and somatic mutations in the ASD cases (Figures 2G–2L). The germline *CACNA1H* and somatic *SCN1A* mutations were detected in case 5278, and the germline *SCN2A* mutation was detected in case 4849. Low gene expression, low read depth in RNA-seq, and/or potential destabilization of RNA by the mutation might contribute to instances where mutant alleles were not detected (e.g., Figure 2H). To further elucidate transcriptomic differences, we identified genes that exhibited the highest fold differences in expression between ASD and control samples (see Supplemental Experimental Procedures). As each case sample had distinct mutations and there were no biological replicates, each case sample was compared with two control samples and only genes that had an absolute fold change of 2 and were changed in the same direction between case and both control samples were considered (Table S7). GO enrichment analysis identified neuron and synapse-related annotation clusters (data not shown), but larger sample sizes would be required to fully characterize global effects of these mutations on transcriptome.

## DISCUSSION

Our analysis provides a unique resource of genetically characterized, publically available human postmortem ASD brains for functional studies, and supports a heterogeneous model of ASD risk, where germline and/or de novo somatic mutations contribute on a case-by-case basis. In many cases, we identify a single variant predicted to be deleterious or loss-of-function, including four independent variants in *SCN2A*, currently one of the strongest ASD candidate genes (Hoischen et al., 2014). Similarly to previous studies (O'Roak et al., 2011), we also identify several ASD cases with multiple germline variants predicted to be deleterious. As discussed above, case AN16641 harbors a frameshift mutation in *SLC6A8* and missense variants in *ARID1B* and *CACNA1C* predicted to be deleterious, and case 797 harbors missense variants in *ADNP* and *PQBP1* predicted to be deleterious. In addition, case AN00090 harbors a germline missense variant in *SETD2* predicted to be deleterious and a previously identified 15q deletion consistent with her diagnosis of Angelman syndrome (Wintle et al., 2011). Finally, we identify a case with a somatic mutation and a case with a germline and somatic variant: case UK20244 harbors a somatic missense variant in *SETD2* predicted to be deleterious, and case 5278 harbors a somatic splice site mutation in *SCN1A* and a germline missense variant in *CACNA1H* predicted to be deleterious. These cases underscore the complex genetic architecture of ASD, and the need for detailed phenotypic information to aid interpretation of identified genetic variants.

Despite targeting only the best-validated ASD genes, we identified two ASD and one Fragile X premutation case with protein-altering somatic mutations, providing evidence that somatic mutations can occur in ASD, though the overall contribution of somatic mutations to ASD risk cannot yet be defined. We observed a somatic mutation that likely occurred early in embryonic development, present across the brain and nonbrain tissues (*SCN1A*), as well as somatic mutations that likely occurred later, after separation of the three germ layers and even after separation of the forebrain, midbrain, and hindbrain. Hence, somatic mutations that are undetectable in current clinical assays using blood-derived DNA could be important mediators of disease expression or penetrance and modify germline variants. The abnormal “patches” observed in the prefrontal and temporal cortex of some autism cases (Casanova et al., 2013; Stoner et al., 2014) may represent visible consequences of somatic mutations, and it will be important to understand how the cell type and brain region in which a somatic mutation occurs modifies its effect on phenotype. With our small sample size and focused gene list, we did not detect a significant burden of rare protein-altering somatic mutations in ASD cases compared to controls ( $p = 0.5$ , two-tailed Fisher’s exact test). Somatic mutations may occur in genes not targeted by our panel, since we only studied genes implicated in ASD via germline mutations. Other more severe mutations, such as chromosome trisomies or activating *AKT3* mutations, can be tolerated in the mosaic state while being lethal in the germline (Poduri et al., 2012, 2013). Larger sample sizes examined with high-coverage ( $\geq 200\times$ ) sequencing will be needed to further assess the full contribution of somatic mutations to ASD and other neuropsychiatric diseases, like schizophrenia, where germline risk mutations have been slower to emerge.

## EXPERIMENTAL PROCEDURES

### Human Tissue and DNA Samples

Research performed on deidentified postmortem human specimens was approved by the institutional review boards of Boston Children’s Hospital and Yale School of Medicine. Case details are available in Table S1.

### DNA Extraction and Targeted Sequencing

Brain DNA was extracted using the QIAamp DNA Mini kit (QIAGEN) and libraries prepared using a custom Haloplex Target Enrichment Kit (Agilent) according to the manufacturer’s protocols. We performed sequencing on MiSeq sequencers (Illumina). Further details on DNA sequencing and analysis are available in the Supplemental Experimental Procedures. Targeted sequencing data is deposited in NDAR Collection 2029.

### Variant Validation

Rare and protein-altering (nonsynonymous, nonsense, splice site, frameshift, and insertion-deletion) variants in the target genes were validated using Sanger sequencing. For potential somatic variants, the original DNA was amplified using polymerase chain reaction (PCR), subcloned into a TOPO TA vector (Invitrogen), and transformed into TOP10 chemically competent *E. coli* cells (Invitrogen); multiple clones were isolated and sequenced to quantify the degree of mosaicism independently of the NGS platform.

### Fragile X Analysis

CGG repeat assay was performed using the AmpliX FMR1 PCR Reagents (Asuragen) and analyzed using GeneMapper Software (Life Technologies) according to the manufacturer’s protocol.

### Single Nuclei Analysis

Isolation of single neuronal and nonneuronal nuclei using fluorescence-activated nuclear sorting (FANS) with antisera to NeuN, and single-cell, whole-genome-amplification using multiple displacement amplification (MDA) (Dean et al., 2001) were performed as described previously (Evrony et al., 2012). Single nuclei from the prefrontal cortex of case 5278 were isolated, amplified, and sequenced for the splicing mutation in *SCN1A* using PCR (GoTaq Hot Start DNA Polymerase, Promega). Sequencing analysis to calculate the number of cells with the mutation, taking into account allelic dropout, was performed as described previously (Evrony et al., 2012).

### RNA Extraction and Sequencing

Total RNA was extracted using mirVana kit (Ambion) with some modifications to the manufacturer’s protocol, and we prepared libraries using TruSeq Stranded Total RNA with Ribo-Zero Gold Kit (Illumina) according to the manufacturer’s protocol. We performed sequencing on HiSeq 2000 sequencer (Illumina). Further details on RNA sequencing and analysis are available in the Supplemental Experimental Procedures.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes one figure, seven tables, and Supplemental Experimental Procedures and can be found with this article at <http://dx.doi.org/10.1016/j.neuron.2015.11.009>.

## AUTHOR CONTRIBUTIONS

A.M.D., S.S.J., and C.A.W. designed DNA experiments. A.M.D., S.S.J., and A.-T.N.L. obtained tissue samples and extracted DNA, prepared sequencing libraries, analyzed sequencing data, and performed validation. A.M.D. and R.E.R. isolated single nuclei and performed MDA and genotyping. S.P., M.L., and N.S. designed RNA experiments, and S.P. and M.L. extracted RNA, prepared sequencing libraries, and analyzed sequencing data. A.M.D. and C.A.W. wrote the manuscript, and all coauthors edited the manuscript.

## ACKNOWLEDGMENTS

We thank Aldo Rozzo, R. Sean Hill, and Jennifer N. Partlow for logistical assistance; Timothy W. Yu and Annapurna Poduri for help with panel design; Madelen Diaz for help with DNA extractions; Jian Wang and Yiping Shen for MiSeq sequencing support; the Dana Farber Cancer Institute Flow Cytometry Core; the Biopolymers Facility at Harvard Medical School; Boston Children’s Hospital IDDRC (P30 HD18655); Stela Filipovic-Sadic of Asuragen, Inc. for help with Fragile X analysis; and Robert Johnson, Jane Pickett, and Carolyn Sloan for help obtaining samples. Postmortem human tissue was obtained from the NIH NeuroBioBank at the University of Maryland, Baltimore, MD; the Harvard Brain Tissue Resource Center, Belmont, MA; The Centre for Applied Genomics Genome Resource Facility, The Hospital for Sick Children, Toronto, ON; the Oxford Brain Bank, Oxford, UK (supported by the Medical Research Council, Brains for Dementia Research and the NIHR Oxford Biomedical Research Center); Autism BrainNet, whose predecessor was the Autism Tissue Program (sponsored by the Simons Foundation and Autism Speaks), and tissue collection in the Sestan lab. A.M.D. was supported by the NIGMS (T32GM007753) and NRSA (5T32 GM007226-39) and is a Stuart H.Q. & Victoria Quan Fellow at Harvard Medical School. R.E.R. was supported by the NIGMS (T32GM007753). S.P., M.L., and N.S. were supported by the National Institute of Mental Health (P50MH106934, U01MH103339, and U01MH106874). C.A.W. was supported by the National Institute of Mental Health (R01MH083565 and U01MH106883), the Simons Foundation (178093), and the Manton Center for Orphan Disease Research. C.A.W. is an Investigator of the Howard Hughes Medical Institute.

Received: July 8, 2015

Revised: September 16, 2015

Accepted: October 27, 2015

Published: December 2, 2015

## REFERENCES

- Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., et al.; Epi4K Consortium; Epilepsy Phenome/Genome Project (2013). De novo mutations in epileptic encephalopathies. *Nature* 501, 217–221.
- Basu, S.N., Kollu, R., and Banerjee-Basu, S. (2009). AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* 37, D832–D836.
- Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res.* 1380, 42–77.
- Casanova, M.F., El-Baz, A.S., Kamat, S.S., Dombroski, B.A., Khalifa, F., Elnakib, A., Soliman, A., Allison-McNutt, A., and Switala, A.E. (2013). Focal cortical dysplasias in autism spectrum disorders. *Acta Neuropathol. Commun.* 1, 67.
- Chen, J.A., Peñagarikano, O., Belgard, T.G., Swarup, V., and Geschwind, D.H. (2015). The emerging picture of autism spectrum disorder: genetics and pathology. *Annu. Rev. Pathol.* 10, 111–144.
- D’Gama, A.M., Geng, Y., Couto, J.A., Martin, B., Boyle, E.A., LaCoursiere, C.M., Hossain, A., Hatem, N.E., Barry, B., Kwiatkowski, D.J., et al. (2015). Mammalian target of rapamycin pathway mutations cause hemimegalencephaly and focal cortical dysplasia. *Ann. Neurol.* 77, 720–725.
- De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al.; DDD Study; Homozygosity Mapping Collaborative for Autism; UK10K Consortium (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 515, 209–215.
- Dean, F.B., Nelson, J.R., Giesler, T.L., and Lasken, R.S. (2001). Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 11, 1095–1099.
- Depienne, C., Trouillard, O., Gourfinkel-An, I., Saint-Martin, C., Bouteiller, D., Graber, D., Barthez-Carpentier, M.A., Gautier, A., Villeneuve, N., Dravet, C., et al. (2010). Mechanisms for variable expressivity of inherited SCN1A mutations causing Dravet syndrome. *J. Med. Genet.* 47, 404–410.
- Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496.
- Fujiwara, T., Sugawara, T., Mazaki-Miyazaki, E., Takahashi, Y., Fukushima, K., Watanabe, M., Hara, K., Morikawa, T., Yagi, K., Yamakawa, K., and Inoue, Y. (2003). Mutations of sodium channel alpha subunit type 1 (SCN1A) in intractable childhood epilepsies with frequent generalized tonic-clonic seizures. *Brain* 126, 531–546.
- Gaily, E., Anttonen, A.K., Valanne, L., Liukkonen, E., Träskelin, A.L., Polvi, A., Lommi, M., Muona, M., Eriksson, K., and Lehesjoki, A.E. (2013). Dravet syndrome: new potential genetic modifiers, imaging abnormalities, and ictal findings. *Epilepsia* 54, 1577–1585.
- Hoischen, A., Krumm, N., and Eichler, E.E. (2014). Prioritization of neurodevelopmental disease genes by discovery of new mutations. *Nat. Neurosci.* 17, 764–772.
- Huguet, G., Ey, E., and Bourgeron, T. (2013). The genetic landscapes of autism spectrum disorders. *Annu. Rev. Genomics Hum. Genet.* 14, 191–213.
- Iossifov, I., O’Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515, 216–221.
- Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B., O’Hanlon, D., et al. (2014). A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* 159, 1511–1523.
- Jamuar, S.S., Lam, A.T., Kircher, M., D’Gama, A.M., Wang, J., Barry, B.J., Zhang, X., Hill, R.S., Partlow, J.N., Rozzo, A., et al. (2014). Somatic mutations in cerebral cortical malformations. *N. Engl. J. Med.* 371, 733–743.
- Jeste, S.S., and Geschwind, D.H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat. Rev. Neurol.* 10, 74–81.
- Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature* 478, 483–489.
- Lee, J.H., Huynh, M., Silhavy, J.L., Kim, S., Dixon-Salazar, T., Heiberg, A., Scott, E., Bafna, V., Hill, K.J., Collazo, A., et al. (2012). De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nat. Genet.* 44, 941–945.
- Lokanga, R.A., Entezam, A., Kumari, D., Yudkin, D., Qin, M., Smith, C.B., and Usdin, K. (2013). Somatic expansion in mouse and human carriers of fragile X pre-mutation alleles. *Hum. Mutat.* 34, 157–166.
- O’Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C., et al. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.* 43, 585–589.
- O’Roak, B.J., Vives, L., Fu, W., Egerton, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K., et al. (2012). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338, 1619–1622.
- Poduri, A., Evrony, G.D., Cai, X., Elhosary, P.C., Beroukhi, R., Lehtinen, M.K., Hills, L.B., Heinzen, E.L., Hill, A., Hill, R.S., et al. (2012). Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron* 74, 41–48.
- Poduri, A., Evrony, G.D., Cai, X., and Walsh, C.A. (2013). Somatic mutation, genomic variation, and neurological disease. *Science* 341, 1237758.
- Redin, C., Gérard, B., Lauer, J., Herenger, Y., Müller, J., Quartier, A., Masurel-Paulet, A., Willems, M., Lesca, G., El-Chehadeh, S., et al. (2014). Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J. Med. Genet.* 51, 724–736.
- Rivière, J.B., Mirzaa, G.M., O’Roak, B.J., Beddaoui, M., Alcantara, D., Conway, R.L., St-Onge, J., Schwartzentruber, J.A., Gripp, K.W., Nikkel, S.M., et al.; Finding of Rare Disease Genes (FORGE) Canada Consortium (2012). De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* 44, 934–940.
- Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237–241.
- Stoner, R., Chow, M.L., Boyle, M.P., Sunkin, S.M., Mouton, P.R., Roy, S., Wynshaw-Boris, A., Colamarino, S.A., Lein, E.S., and Courchesne, E. (2014). Patches of disorganization in the neocortex of children with autism. *N. Engl. J. Med.* 370, 1209–1219.
- Tan, M.H., Mester, J., Peterson, C., Yang, Y., Chen, J.L., Rybicki, L.A., Milas, K., Pederson, H., Remzi, B., Orloff, M.S., and Eng, C. (2011). A clinical scoring system for selection of patients for PTEN mutation testing is proposed on the basis of a prospective study of 3042 probands. *Am. J. Hum. Genet.* 88, 42–56.
- Wingate, M., Kirby, R.S., Pettygrove, S., Cunniff, C., Schulz, E., Ghosh, T., Robinson, C., Lee, L.C., Landa, R., Constantino, J., et al. (2014). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 Sites, United States, 2010. *MMWR Surveill. Summ.* 28, 1–21.
- Wintle, R.F., Lionel, A.C., Hu, P., Ginsberg, S.D., Pinto, D., Thiruvahindrapduram, B., Wei, J., Marshall, C.R., Pickett, J., Cook, E.H., and Scherer, S.W. (2011). A genotype resource for postmortem brain samples from the Autism Tissue Program. *Autism Res.* 4, 89–97.