



Article

# Next Generation Sequencing of 134 Children with Autism Spectrum Disorder and Regression

Jiani Yin <sup>1,2,3</sup>, Chun-An Chun <sup>1,2</sup>, Nikolay N. Zavadenko <sup>4</sup> , Natalia L. Pechatnikova <sup>4</sup>,  
Oxana Yu. Naumova <sup>5,6,7</sup>, Harsha V. Doddapaneni <sup>8</sup>, Jianhong Hu <sup>8</sup> , Donna M. Muzny <sup>8</sup>,  
Christian P. Schaaf <sup>1,2,9</sup> and Elena L. Grigorenko <sup>1,6,7,\*</sup>

<sup>1</sup> Molecular and Human Genetics Department, Baylor College of Medicine, Houston, TX 77030, USA; yinjiani900129@gmail.com (J.Y.); chunanc@bcm.edu (C.-A.C.); schaaf@bcm.edu (C.P.S.)

<sup>2</sup> Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX 77030, USA

<sup>3</sup> Department of Neurology, Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

<sup>4</sup> Neurology, Neurosurgery and Medical Genetics, Department of Pediatrics, Pirogov Russian National Research Medical University, 117997 Moscow, Russia; zavadenko@mail.ru (N.N.Z.); Funny85@bk.ru (N.L.P.)

<sup>5</sup> Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX 77024, USA; oksana.yu.naumova@gmail.com

<sup>6</sup> Vavilov Institute of General Genetics, Russian Academy of Sciences, 119991 Moscow, Russia

<sup>7</sup> Laboratory of Translational Developmental Sciences, Saint Petersburg State University, 199034 Saint Petersburg, Russia

<sup>8</sup> Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA; doddapan@bcm.edu (H.V.D.); Jianhong.Hu@bcm.edu (J.H.); donnam@bcm.edu (D.M.M.)

<sup>9</sup> Institute of Human Genetics, Heidelberg University, 69120 Heidelberg, Germany

\* Correspondence: Elena.Grigorenko@bcm.edu; Tel.: +1-713-743-7983

Received: 18 June 2020; Accepted: 22 July 2020; Published: 25 July 2020



**Abstract:** Approximately 30% of individuals with autism spectrum disorder (ASD) experience developmental regression, the etiology of which remains largely unknown. We performed a complete literature search and identified 47 genes that had been implicated in such cases. We sequenced these genes in a preselected cohort of 134 individuals with regressive autism. In total, 16 variants in 12 genes with evidence supportive of pathogenicity were identified. They were classified as variants of uncertain significance based on ACMG standards and guidelines. Among these were recurring variants in *GRIN2A* and *PLXNB2*, variants in genes that were linked to syndromic forms of ASD (*GRIN2A*, *MECP2*, *CDKL5*, *SCN1A*, *PCDH19*, *UBE3A*, and *SLC9A6*), and variants in the form of oligogenic heterozygosity (*EHMT1*, *SLC9A6*, and *MFSD8*).

**Keywords:** autism; developmental regression; exon capture and sequencing; variant classification; ACMG standards and guidelines

## 1. Introduction

About thirty percent of individuals with ASD experience developmental regression [1]. These individuals commonly manifest cognitive decline, loss of language skills, seizures, and difficulty in managing behavior [1]. Symptoms can manifest as early as infancy or early childhood. ASD and developmental regression are involved in a few neurological conditions with known etiologies, such as Rett syndrome, Landau–Kleffner syndrome, or Christianson syndrome [1,2]. In addition, genetic variants in isolated cases of autism and developmental regression have been reported. For example, an R426H homozygous variant in the *ADSL* gene was identified in three affected individuals of the same family [3]. A translocation t(14;21)(q21.1; p11.2), disrupting the expression of *LRFN5*, was

reported in a severely autistic girl who manifested regression of skills after 2 years old [4]. A female who manifested neuronal ceroid lipofuscinosis with Rett-like onset was found to bear compound heterozygous variants of *MFSD8* and a heterozygous variant in *CLN5* [5]. A boy with autism and two episodes of neurodevelopmental regression was reported to have a frameshifting variant in *TMLHE* [6]. However, beyond these aforementioned cases, the genetic and pathophysiologic underpinnings of developmental regression remain largely unknown, and it remains challenging to provide a prognosis for individuals with regressive autism.

Next generation sequencing has proven to be a powerful way to identify variants in many neurological disorders [7–13]. To explore the genetic predisposition to a regressive form of autism, we performed targeted exon capturing and the sequencing of candidate genes. We recruited a cohort of 141 individuals diagnosed with ASD and developmental regression and screened them for variants in genes that had previously been reported in the manifestation of these traits.

## 2. Materials and Methods

### 2.1. Patient Recruitment

The criteria for recruitment included a diagnosis of ASD, a history of severe developmental regression, and no known cause of either at the time of recruitment. Developmental regression was defined by a loss of acquired functions, as documented by a pediatrician or child neurologist or psychiatrist in the child's medical records and a referral for a consultation for developmental regression to a specialized central medical center in Moscow, Russia. Altogether, 141 patients were recruited from 2011 to 2013. All experiments were conducted in accordance with the IRB's requirements from St. Petersburg State University and Yale University.

### 2.2. Literature Search for Candidate Genes

Given the limited number of variants reported in cases with regressive autism, we searched for variants that were associated with developmental regression both with and without autism. We searched for the terms "autism AND regression" or "developmental regression" in the PubMed database. Then, we read the abstracts and results to confirm the presence of autism and developmental regression in the reported cases. In these cases, gene names were noted if variants were found or if the genes were involved in a copy number variation. A total of 1659 abstracts were reviewed by August 2015, which resulted in a list of 47 genes (Table 1) [2–6,14–25].

**Table 1.** Gene list for variant screening.

	Mode of Inheritance
<i>ATXN2</i>	x-linked
<i>EHMT1</i>	x-linked
<i>CDKL5</i>	x-linked
<i>PCDH19</i>	x-linked
<i>TMLHE</i>	x-linked
<i>MECP2</i>	x-linked
<i>SLC9A6</i>	x-linked
<i>ATP13A4</i>	Autosomal dominant
<i>ATP13A5</i>	Autosomal dominant
<i>CDH13</i>	Autosomal dominant
<i>CDH4</i>	Autosomal dominant
<i>CDH9</i>	Autosomal dominant

Table 1. Cont.

	<b>Mode of Inheritance</b>
<i>GRIN2A</i>	Autosomal dominant
<i>HSBP1</i>	Autosomal dominant
<i>LRFN5</i>	Autosomal dominant
<i>MACROD2</i>	Autosomal dominant
<i>MBD5</i>	Autosomal dominant
<i>MDGA2</i>	Autosomal dominant
<i>RELN</i>	Autosomal dominant
<i>SCN1A</i>	Autosomal dominant
<i>SGSH</i>	Autosomal dominant
<i>SHANK3</i>	Autosomal dominant
<i>STXBP1</i>	Autosomal dominant
<i>TCF4</i>	Autosomal dominant
<i>UBE3A</i>	Autosomal dominant
<i>ADSL</i>	Autosomal recessive
<i>ALDH5A1</i>	Autosomal recessive
<i>FOXP1</i>	Autosomal recessive
<i>MFSD8</i>	Autosomal recessive
<i>PLCB1</i>	Autosomal recessive
<i>PRSS12</i>	Autosomal recessive
<i>SCN8A</i>	Autosomal recessive
<i>MTHFR</i>	Autosomal recessive
<i>DIAPH3</i>	Autosomal recessive
<i>FGF12</i>	Involved in copy number variations
<i>HRASLS</i>	Involved in copy number variations
<i>OPA1</i>	Involved in copy number variations
<i>SCN2A</i>	Involved in copy number variations
<i>ATXN7</i>	Involved in copy number variations
<i>CDH6</i>	Involved in copy number variations
<i>CNTNAP2</i>	Involved in copy number variations
<i>CTNNA3</i>	Involved in copy number variations
<i>GRIN2B</i>	Involved in copy number variations
<i>HDAC10</i>	Involved in copy number variations
<i>KCNQ2</i>	Involved in copy number variations
<i>KIF26B</i>	Involved in copy number variations
<i>PLXNB2</i>	Involved in copy number variations

### 2.3. Library Preparation

Genomic DNA was isolated from whole frozen EDTA-blood (1–2 mL) using the phenol chloroform extraction method [26,27]. DNA samples were constructed into Illumina paired-end pre-capture libraries using Beckman robotic workstations (Biomek FX and FXp models). In total, 250 ng of DNA per sample was used as input for the library preparation.

Whole Genome Amplification was completed on 47 samples, the total DNA amount of which was less than 250 ng. Briefly, DNA in a 70  $\mu$ L volume was sheared into fragments of approximately 500–700 base pairs in a Covaris plate with a E220 system (Covaris, Inc., Woburn, MA, USA). The fragmented DNA was purified using 0.8 $\times$  AMPure XP (Beckman, Cat. No. A63882) followed by end-repair (NEBNext End-Repair Module; Cat. No. E6050L), A-tailing (NEBNext<sup>®</sup> dA-Tailing Module; Cat. No. E6053L), and ligation of the Illumina multiplexing PE adaptors with 96 different 9-bp barcode sequences using ExpressLink<sup>™</sup> T4 DNA Ligase (a custom product from LifeTech). Pre-capture Ligation Mediated-PCR (LM-PCR) was performed for 6–8 cycles using the Library Amplification Readymix containing KAPA HiFi DNA Polymerase (Kapa Biosystems, Inc., Cat # KK2612). Universal primers, IMUX-P1.0 and IMUX-P3.0, were used in PCR amplification. AMPure XP beads (0.8 $\times$ ) were used to purify reaction contents after each enzymatic reaction. Libraries were quantified and their sizes were estimated using a Fragment Analyzer capillary electrophoresis system (Advanced Analytical Technologies, Inc., Ames, IA, USA). The complete protocol and oligonucleotide sequences are accessible from the HGSC website ([https://www.hgsc.bcm.edu/sites/default/files/documents/Protocol-Illumina\\_Whole\\_Exome\\_Sequencing\\_Library\\_Preparation-KAPA\\_Version\\_BCM-HGSC\\_RD\\_03-20-2014.pdf](https://www.hgsc.bcm.edu/sites/default/files/documents/Protocol-Illumina_Whole_Exome_Sequencing_Library_Preparation-KAPA_Version_BCM-HGSC_RD_03-20-2014.pdf)).

Seven samples failed to pass quality control; correspondingly, we proceeded with 134 samples.

#### 2.4. Sequence Capture and Sequencing on HiSeq 2000

For capture enrichment, pools of libraries (~104 ng/library and 48 libraries/pool) were prepared, and 5.0  $\mu$ g/pool was used for in-solution hybridization to custom capture a design [150825\_HG19\_10bp\_Mende\_chr\_EZ\_HX3] that covered all exons and 10 bp intron-exon boundary regions on each side of the exons of the 47 genes (Nimblegen).

Human COT1 DNA and full-length blocking oligonucleotides (Sigma-Aldrich Inc., St. Louis, MO, USA) were added into the hybridization to block repetitive genomic sequences and the adaptor sequences. Hybridization was done at 56 °C O/N. Post-capture LM-PCR amplification was performed using the Library Amplification Readymix containing KAPA HiFi DNA Polymerase (Kapa Biosystems, Inc., Cat # KK2612) with 12 cycles of amplification. After the final AMPure XP bead purification, quantity and size of the capture library was analyzed using an Agilent Bioanalyzer 2100 DNA Chip 7500. The enriched library pools (48 libraries/pool or 96 per lane) were sequenced on a HiSeq 2000 to generate 2  $\times$  100 bp reads. On average, for these 134 samples, including 47 samples that were whole genome amplified, 397 Mb of mapped sequence data (295 Mb unique with 537.38 $\times$  sequence coverage) across the length of the targeted region were generated. On average, 96% of bases in the design were covered at a  $\geq$  20 $\times$  depth.

#### 2.5. Primary Data Analysis

The initial sequence analysis was performed using the HGSC Mercury analysis pipeline [28,29]. In summary, the bcl files produced on-instrument were first transferred into the HGSC analysis infrastructure by the HiSeq Real-time Analysis module and Mercury; we then ran the vendor's primary analysis software (CASAVA) to de-multiplex the pooled samples and generate sequence reads and base-call confidence values (qualities). Reads were then mapped to the GRCh37 Human reference genome (<http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/>) using the Burrows-Wheeler aligner (BWA, <http://bio-bwa.sourceforge.net/>). In this step, the resulting BAM (binary alignment/map) files underwent quality recalibration using GATK (<http://www.broadinstitute.org/gatk/>) BAM sorting, duplicate read marking, and realignment to improve indel discovery. BAM files were used by the ATLAS 2 suite (<https://www.hgsc.bcm.edu/software/atlas-2>) to call single nucleotide variants (SNV) and insertion/deletion variants and produce a variant call file (VCF). Finally, annotation data were added to the VCF using a suite of annotation tools (Cassandra) that brings together frequency, function, and other relevant information, using AnnoVar with UCSC and RefSeq gene models, as well as a host of other internal and external data resources (<https://www.hgsc.bcm.edu/software/cassandra>).

## 2.6. Sanger Sequencing

Mutation validation was performed with bidirectional sequencing of the selected sample sites. PCR reactions were prepared using 5 ng of genomic DNA, 0.4  $\mu$ M oligonucleotide primers, and 0.7 $\times$  Qiagen Multiplex Master Mix (Cat. no. 206145) containing HotStar Taq, buffer, and polymerase. Two reactions were completed for each sample site, one with and one without, using the QSOL PCR additive provided in the kit. Touchdown PCR was performed with the following parameters: 98 °C for 5 min and 10 cycles of 98 °C for 30 s, 72 °C for 30 s (decreasing by 1 °C per cycle), and 72 °C for 1 min. The reaction then continued with 30 cycles of 98 °C for 30 s, 63 °C for 30 s, and 72 °C for 1 min, followed by a final extension at 72 °C for 5 min. The PCR products were purified with a 1:15 dilution of Exo-SAP, diluted by 0.6 $\times$ , and then cycle-sequenced for 25 cycles using a 1/32nd dilution of a BigDye Terminator v3.1 reaction mix (Applied Biosystems, Cat. No. 4337456). Finally, the reactions were precipitated with ethanol, resuspended in 0.1 mM EDTA, and analyzed on AB 3730xl sequencing instruments using the Rapid36 run module and 3xx base-caller. SNPs were identified and visually validated with a Mutation Surveyor.

## 3. Results

### 3.1. Targeted Sequencing of Regressive Autism Candidate Genes

We sequenced 47 candidate genes in 134 individuals with autism and severe developmental regression using targeted capture and sequencing. Candidate genes were selected based on at least one case of developmental regression, with or without autism, reported in the literature prior to August 2015, with either a single nucleotide variant in the gene or copy number variation involving the gene (Methods section “Literature search for candidate genes”). A custom library was designed to capture all exons and intron-exon boundary regions of the 47 genes.

Initial variant calling yielded 2121 variants. We filtered these variants based on the following criteria: (1) Nonsense, missense, frameshift, or splice-region variant type (Synonymous, inframe, and variants in the UTR region were excluded); (2) a minor allele frequency lower than  $2 \times 10^{-4}$ , or 1 in 5000 individuals, in the ExAC database or 1000 Genomes; (3) prediction to be damaging or possibly damaging by SIFT and/or PolyPhen; and (4) 10 or fewer instances of calls among these 134 individuals (variants called more than 10 times were considered likely false positives). After filtering, 130 variants remained to be validated. Of those, we were able to confirm 41 of the filtered variants using Sanger sequencing, among which 16 unique variants had evidence supportive of pathogenicity based on the ACMG criteria. All 16 variants were classified as “uncertain significance” based on the ACMG standards and guidelines for the interpretation of sequence variants [30]. Among the remaining 89 variants, 65 were false positives, and another 34 were not confirmed due to insufficient DNA amount or failure of PCR amplification of the variant region. The clinical and genetic data of the 15 individuals carrying variants of uncertain significance, as well as evidence supportive of pathogenicity of the variants, are summarized in Table 2.

**Table 2.** Summary of the clinical and genetic data of individuals with pathogenic variants.

Indivi-Dual ID	Sex	Onset Age of Regression	Variants	ACMG Classify-Cation	Seizures	Language Impair-Ment	Motor Impair-Ment	Additional Notes
090	M	36 months	<i>GRIN2A</i> c.28C > A p.Leu10Met	PM2	No	Yes	No	
113	M	36 months	<i>GRIN2A</i> c.28C > A p.Leu10Met	PM2	No	Yes	Yes	Stereotypies
031	M	39 months	<i>GRIN2A</i> c.904G > T p.Ala302Ser	PM2	Yes	Yes	No	
086	M	36 months	<i>PLXNB2</i> c.742C > T p.Arg248Cys	PP3	Yes	Yes	Yes	
110	F	14 months	<i>PLXNB2</i> c.742C > T p.Arg248Cys	PP3	Yes	Yes	Yes	Stereotypies, hand tremor
037	F	6 months	<i>SCN1A</i> c.4852 + 1G > T	PM2	Yes	Yes	Yes	Convulsions
121	M	24 months	<i>SCN1A</i> c.3269G > C p.Ser1090Thr	PM2	Yes	Yes	Yes	Stereotypies, attention deficit
051	F	16 months	<i>PCDH19</i> c.121A > G p.Asn41Asp	PM2	Yes	Yes	Yes	Facial dysmorphology
073	F	6 months	<i>MECP2</i> c.439A > G p.Lys147Glu	PS1	Yes	Yes	Yes	Loss of vision, hand tremor
139	F	12 months	<i>CDKL5</i> c.1387A > G p.Lys463Glu	PP3	No	Yes	Yes	Sleep disturbances, obesity, severe temper tantrums, stereotypies, aggression

Table 2. Cont.

Indivi-Dual ID	Sex	Onset Age of Regression	Variants	ACMG Classify-Cation	Seizures	Language Impair-Ment	Motor Impair-Ment	Additional Notes
029	F	18 months	<i>UBE3A</i> c.2439C > G p.Asp813Glu	PM2	Yes	Yes	Yes	Ataxia
030	F	Not available	<i>SLC9A6</i> c.171C > G p.Ile57Met	PP3	Yes	Yes	Yes	Spastic quadriparesis and ataxia, sleep disturbances
104	M	Not available	<i>EHMT1</i> c.989A > T p.Lys330Met & <i>SLC9A6</i> c.1777C > G p.Leu593Val	PM2 (for <i>EHMT1</i> ) & PP3 (for <i>SLC9A6</i> )	Not available	Not available	Not available	
041	M	12 months	<i>EHMT1</i> c.1513G > A p.Gly505Ser & <i>MFSD8</i> c.353A > G p.Asn118Ser	PP3 for both	No	Yes	Yes	
137	M	11 months	<i>SCN2A</i> c.5549A > G p.Asp1850Gly & <i>MTHFR</i> c.136C > T p.Arg46Trp	PM2 for both	No	Yes	Yes	

### 3.2. Recurrent Variants Identified in Our Regressive Autism Cohort

In our sequenced cohort of 134 individuals with autism and regression, we identified two recurrent variants, *GRIN2A* c.28C > A (p.Leu10Met) and *PLXNB2* c.742C > T (p.Arg248Cys). The variant in *GRIN2A* was novel and was predicted to be damaging using both SIFT and PolyPhen. Heterozygous *GRIN2A* germline variants are considered to be major genetic causes of Landau Kleffner syndrome, which is characterized by speech impairment and focal epilepsy [31]. Two unrelated males in our cohort were found to harbor the variant *GRIN2A* c.28C > A. Both were referred for evaluation at 3 years of age and had language development arrest and regression. In addition, one of the males (ID113) had hydrocephaly and showed an increase in emotional negativity and stereotypies. The other male (ID090) experienced sleep disturbance.

The *PLXNB2* gene is important for brain development [32–34] but has not been associated with any disorder. The deletion of *PLXNB2* and other genes was reported in an individual with autism and speech impairment [35]. Unrelated individuals ID086 (male) and ID110 (female) had seizures and regression and were found to harbor variants in *PLXNB2*. The male was referred at the age of 14 years; retrospectively, he experienced the onset of febrile seizures at the age of 3 years, which intensified and occurred regularly 2–3 times a night. The girl was referred at the age of 10 years; retrospectively, febrile seizures were reported with onset at the age of 1 year. In both children, the progression of seizures was coupled with the regression of language and cognitive functions. Recently, the female was diagnosed with Angelman Syndrome using the PRC test based on the allelic methylation differences at the *SNRPN* locus (15q11.2). In both cases, *PLXNB2* c.742C > T was the only candidate variant from our sequencing analysis. This variant was found in the ExAC database at a rate of  $4.1 \times 10^{-5}$  and was predicted to be damaging using SIFT, PolyPhen, and MutationTaster. Due to the lack of a proper control cohort, we were not able to assess the frequency of this variant in the normal population of the same ethnicity.

### 3.3. Variants in Genes Associated with Syndromic Forms of ASD

Rett syndrome, caused by mutations in *MECP2*, is one well-established cause of developmental regression and autism, which accounts for roughly 0.5% of such cases. We identified a missense variant c.403A > G (p.Lys147Glu) in *MECP2* while sequencing one female (ID073), regardless of excluding individuals known to carry pathogenic variants in *MECP2*. This girl was referred at the age of 7 years. Her developmental history indicated the presence of microcephaly and the onset of seizures at the age of 6 months, followed by arrested development and a rapid loss of acquired skills. This variant in *MECP2* was absent from the ExAC and gnomAD databases, was predicted to be damaging, and was reported as pathogenic or with uncertain significance in two separate records (ClinVar#143564). We believe that autism and the developmental regression of ID073 are likely attributed to this variant.

Genetic variants in *CDKL5* cause a disorder that is characterized by early onset seizures and subsequent significant developmental delay. Here, we identified an individual (ID139) who carried a missense variant c.1387A > G (p.Asp813Glu) in *CDKL5* on the X chromosome. This female was referred at the age of 3 years. Her developmental history indicates the manifestation of arrested development at the age of 1 year. At 3 years of age, the family reported sleeping difficulties; at 8 years old, the family reported obesity, a lack of motor coordination, and profound hyperactivity. The variant in *CDKL5* was predicted to be deleterious using SIFT and MutationTaster and was present in the ExAC database at a rate of  $2 \times 10^{-5}$ . This variant appears to be a plausible cause of autism and developmental regression in ID139.

Genetic variants in *GRIN2A* are a major cause of Landau Kleffner syndrome. Patients ID090 and ID113 carried the same variant in *GRIN2A* (see above). In addition, ID031 carried a different variant, *GRIN2A* c.904G > T (p.Ala302Ser), which was absent from the ExAC database and present in the gnomAD database at a rate of  $7.98 \times 10^{-6}$ . This variant was predicted to be deleterious. This boy was referred at the age of 3 years, 3 months after the onset of seizures and developmental arrest. His seizures were classified as atypical absences, myoclonic, axial tonic, and generalized tonic-clonic. The observation of three individuals with probable pathogenic variants in *GRIN2A* suggests that the



dysfunction of *GRIN2A* may account for a substantial (2.2%) proportion of cases in this sample of individuals with regressive autism.

Genetic variants in *SCN1A* have been associated with Dravet syndrome and other severe epileptic encephalopathies [24,36]. Patients with Dravet syndrome commonly present severe forms of epilepsy starting within the first year of life and the deterioration of intellectual development at around 2 years of age. The female ID037 and male ID121 carried novel splice-site and missense variants in *SCN1A*. Both individuals had developmental regression. The female's seizures manifested at the age of 3 months as febrile seizures, which were followed by focal motor seizures in the right hand at 4.5 months and generalized tonic-clonic seizures at the age of 6 months; she was resistant to antiepileptic drugs. The male also manifested attention deficit, extreme hyperactivity, severe expressive speech delay with limited receptive language skills, stereotypic and repetitive movements, and cranial dysmorphism (brachycephaly).

Mutations in *PCDH19* on the X chromosome cause epilepsy and intellectual disability preferentially in females, which is also known as Dravet-like syndrome [16]. Few males are affected, and it has been postulated that in hemizygous transmitting males, *PCDH11Y* on the Y chromosome likely compensates for the function of *PCDH19* [37]. We identified a female individual, ID051, who carried a missense variant c.121A > G (p.Asn41Asp) in *PCDH19*. This variant is absent in the ExAC database and present in the gnomAD database at a rate of  $6.03 \times 10^{-6}$ . The patient manifested focal epilepsy with pseudoabsences, macrocephaly, cranial dysmorphism, decreased muscular tone, and macrosomia. Her developmental history indicates the onset of regression at the age of 16 months, first in the cognitive and then in the motor domains. Due to phenotypic presentation with macrosomia, Sotos Syndrome was considered in the differential diagnosis, but genetic testing did not detect any aberrations in the *NSD1* gene. Identification of the variant in *PCDH19* provides a possible molecular diagnosis to this patient and suggests the disturbance of shared pathways in regressive autism and Dravet-like syndrome.

The malfunction of *UBE3A* on the maternal allele is known to cause Angelman syndrome, which is characterized by intellectual disability, ataxia, severe speech impairment, and other neurological and behavioral features. Psychomotor regression is rarely seen or reported among individuals with Angelman syndrome, but there was some evidence for a linkage of psychomotor regression with *UBE3A* and its adjacent regions on chromosome 15 [17]. We identified an individual (ID029) who carries a novel missense variant c.2439C > G (p.Asp813Glu) in *UBE3A*. This variant was absent from the ExAC and gnomAD databases and was predicted to be damaging using SIFT and PolyPhen. The girl was referred at the age of 4 years with absence epilepsy and developmental arrest, intellectual disability, severe speech impairment, and problems with active movements (quadripareisis) and balance (ataxia). The history indicated a delay in motor and intellectual development but with progress in development until the age of 18 months, when she started to fall down frequently due to clumsiness and developed a spastic gait. Specific facial features and excitability were interpreted as suggestive of Angelman Syndrome. An allele-specific DNA methylation analysis of the promoter region of the *SNRPN* gene (chromosome 15q11.2) was performed and the outcome was normal.

Christianson syndrome is caused by genetic variants in *SLC9A6* on the X chromosome. About 50% of patients with Christianson syndrome experience developmental regression at some point and commonly have intellectual disability, microcephaly, absent speech, ataxia, and seizures [2]. We identified one female (ID030) who carried a missense variant c.171C > G (p.Ile57Met) in *SLC9A6*. At the age of 11 months, she developed flexor infantile spasms and later developed generalized tonic-clonic seizures. From 2 years of age, developmental regression was noted; in addition, spastic quadripareisis and ataxia developed. Evidence for the pathogenicity of this variant has been conflicting. The variant is present in the ExAC database at a rate of  $4.12 \times 10^{-5}$  and was predicted to be damaging using SIFT and MutationTaster. It was classified in one other case with Christianson syndrome as likely benign and in another case with unspecified conditions as of uncertain significance (ClinVar# 379120). We would like to bring attention to this variant, as we observed a second case of an individual with a severe neurological phenotype and the presence of this rare genetic variant in *SLC9A6*.

### 3.4. Individuals Carrying Multiple Variants with Evidence Supportive of Pathogenicity

We also identified two individuals who carried variants with evidence supportive of pathogenicity in two different genes. A male (ID104) was found to have a heterozygous missense variant c.989A > T (p.Lys330Met) in *EHMT1* and a missense variant c.1777C > G (p.Leu593Val) in *SLC9A6*. Limited clinical information was available about this male. The variant in *EHMT1* was absent from the ExAC and gnomAD databases. De novo variants in *EHMT1* have been reported in individuals with autism, but developmental regression has not been reported [24]. Using SIFT and PolyPhen, the c.1777C > G variant in *SLC9A6* was predicted to be damaging, but a different variant at the same amino acid, c.1777C > T (p.Leu593Phe), was found in the ExAC database at a rate of  $8.24 \times 10^{-6}$ .

A male (ID041), unrelated to ID104, carried heterozygous missense variants c.1513G > A (p.Gly505Ser) in *EHMT1* and c.353A > G (p.Asn118Ser) in *MFSD8*. He was seen at 7 years and 10 months and, at that time, was severely developmentally delayed in multiple domains (motor, cognitive, and language). Using Polyphen and MutationTaster, variants were predicted to possibly be damaging, but they were present in the ExAC database (*EHMT1* c.1513G > A at a rate of  $4.95 \times 10^{-5}$ , *MFSD8* c. 353A > G at a rate of  $8.24 \times 10^{-6}$ ). As stated above, a heterozygous variant in *EHMT1* may not result in developmental regression. On the other hand, variants in *MFSD8* have been reported to cause autism and developmental regression, but only in an autosomal recessive condition [5]. *EHMT1* encodes a methyltransferase that specifically methylates histone H3K9me1 and H3K9me2 in euchromatin. *MFSD8* likely encodes a protein that transports molecules across the lysosomal membrane, with unknown specificity. This case suggests that ASD with regression may stem from a combinatorial effect of a partial loss of functions in both pathways.

## 4. Discussion

Autism spectrum disorder is a genetically heterogeneous and common developmental disorder that affects communication and behavior, with approximately one-third of individuals manifesting a regression of acquired skills. Various factors that may cause developmental regression include genetic predispositions; alteration in mechanisms critical at certain developmental stages, such as synaptic pruning [38]; infections or environmental triggers; and a combinatorial effect of genetic predisposition and environmental triggers.

Given a lack of appreciation for the genetic landscape of regressive autism, we conducted a complete literature search for genes that were reported in the comorbidity of developmental regression with or without autism prior to August 2015, which formed a list of 47 genes. Using a preselected cohort of 134 individuals who had a diagnosis of autism and developmental regression, we assessed the likelihood of identifying variants in these 47 genes. It should be noted that even in this fairly homogeneous subtype of autism, clinical heterogeneity was substantial. Co-occurrence between seizures and developmental regression has been well-established [39], which was also observed in our cohort. In this study, we regard seizures, developmental regression, intellectual disability, autism, and other clinical manifestations all as traits under the spectrum of neurological disorders. Overall, we found variants with evidence supportive of pathogenicity in 15 individuals.

Interestingly, from our cohort with specific clinical manifestations, which was not large in size, we identified two recurrent variants, *GRIN2A* c.28C > A and *PLXNB2* c.742C > T, which are rare variants in the general population. This finding supports the role of *GRIN2A* and *PLXNB2* in the molecular etiology of autism and developmental regression. We also identified variants in genes that are linked to neurological syndromes, such as *MECP2*, *CDKL5*, *GRIN2A*, *SCN1A*, *PCDH19*, *UBE3A*, and *SLC9A6*. Developmental regression is frequently seen in patients with Rett Syndrome, Landau Kleffner Syndrome, and Christianson Syndrome. However, it has rarely been reported in other neurological syndromes, such as Dravet or Dravet-like syndrome and Angelman syndrome. Our findings provide further evidence that genetic risk factors of these syndromes also predispose one to developmental regression and autism.

Some individuals were found to carry variants in multiple genes, which is in accordance with the idea of oligogenic heterozygosity in the molecular etiology of neurological disorders [40]. Genes that contributed in such a combinatorial manner in our study include *SCL9A6*, *EHMT1*, and *MFSD8*, which are involved in a variety of biological processes, such as cellular homeostasis, histone modification, and amino acid metabolism. Investigations into how oligogenic events increase one's risk for ASD and regression will unravel new disease mechanisms.

The ACMG criteria we used in the study to define likely pathogenic variants was stringent, and we wondered whether there were variants in these 12 genes in the individuals without significant findings. Upon re-examination of those individuals, we found an additional 29 non-intronic and non-synonymous variants absent in the ExAC database in 15 individuals. The predicted consequence of these variants included the 5' UTR variant (1 variant), splice region variant (5 variants), frameshift (7 variants), in-frame deletion or insertion (2 variants), missense (13 variants), and stop-gain (1 variant). Given the genetic heterogeneity underlying ASD and developmental regression, these variants could contribute to genetic predisposition in the disease conditions.

The candidate approach that we used to explore the genetic risk factors in autism and developmental regression allowed rapid screening for variants throughout the samples, and the subsequent variant analysis process was relatively straightforward. However, this study has several limitations: (1) the candidate gene approach limits discovery to genes already described in their respective phenotypic contexts; (2) the lack of parental samples prevented us from determining inheritance vs. de novo status; (3) our study had a relatively small sample size. Future studies are needed to further investigate the genetic landscape of autism with regression, and genome-wide studies will be required to identify novel disease genes for this disease mechanism.

## 5. Conclusions

Our study provides additional insights into the relationship between autism and developmental regression by interrogating a set of genes that are associated with this comorbidity. The application of next generation sequencing in a sample of 134 children with autism and developmental regression highlights the roles of single gene variants in *MECP2*, *CDKL5*, *GRIN2A*, *SCN1A*, *PCDH19*, *UBE3A*, and *SLC9A6*, as well as oligogenic events (variants in *EHMT1* and *SCL9A6*, *EHMT1* and *MFSD8*) that may disrupt multiple biological processes.

**Author Contributions:** J.Y., C.P.S., and E.L.G. designed the study. J.Y. performed the literature survey. N.N.Z., N.L.P., O.Y.N., and E.L.G. acquired clinical information and genomic DNA from all of the 141 subjects. J.Y. and C.-A.C. prepared the samples for sequencing. H.V.D., J.H., and D.M.M. supervised the next generation sequencing and Sanger sequencing. J.Y. and C.P.S. interpreted the results. J.Y., C.-A.C., N.N.Z., O.Y.N., E.L.G., and C.P.S. drafted and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was in part supported by the Baylor College of Medicine Human Genomic Sequencing Center (NIH/NHGRI UM1HG008898). Schaaf was generously supported by the Joan and Stanford Alexander Family. The processing of the specimens was supported, in part, by the grant from the Russian Foundation for Basic Research (Award No. 17-29-02384, PI: ELG).

**Acknowledgments:** We thank Joy Jayaseelan, Shannon Dugan-Perez, Harsha Doddapaneni, Jianhong Hu, William Salerno, Donna Muzny, and Richard A. Gibbs at the Baylor College of Medicine Human Genome Sequencing Center for their contributions to this study. We thank Megan E. Rech for proofreading the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Ethics Statements:** Individual informed consent was obtained from all participants aged 18 and older and from the parents of participants below age 18. All procedures were conducted in accordance with the IRB's requirements from St. Petersburg State University (Russia) and Yale University (USA).

## References

1. Al Backer, N.B. Developmental regression in autism spectrum disorder. *Sudan. J. Paediatr.* **2015**, *15*, 21–26. [[PubMed](#)]
2. Pescosolido, M.F.; Stein, D.M.; Schmidt, M.; El Achkar, C.M.; Sabbagh, M.; Rogg, J.M.; Tantravahi, U.; McLean, R.L.; Liu, J.S.; Poduri, A.; et al. Genetic and phenotypic diversity of NHE6 mutations in Christianson syndrome. *Ann. Neurol.* **2014**, *76*, 581–593. [[CrossRef](#)]
3. Ederly, P.; Chabrier, S.; Ceballos-Picot, I.; Marie, S.; Vincent, M.F.; Tardieu, M. Intrafamilial variability in the phenotypic expression of adenylosuccinate lyase deficiency: A report on three patients. *Am. J. Med. Genet. Part A* **2003**, *120*, 185–190. [[CrossRef](#)] [[PubMed](#)]
4. De Bruijn, D.R.H.; Van Dijk, A.H.A.; Pfundt, R.; Hoischen, A.; Merckx, G.F.M.; Gradek, G.A.; Lybæk, H.; Stray-Pedersen, A.; Brunner, H.G.; Houge, G. Severe Progressive Autism Associated with Two de novo Changes: A 2.6-Mb 2q31.1 Deletion and a Balanced t(14;21)(q21.1;p11.2) Translocation with Long-Range Epigenetic Silencing of LRFN5 Expression. *Mol. Syndromol.* **2010**, *1*, 46–57. [[CrossRef](#)] [[PubMed](#)]
5. Craiu, D.; Dragostin, O.; Dica, A.; Hoffman-Zacharska, D.; Gos, M.; Bastian, A.E.; Gherghiceanu, M.; Rolfs, A.; Nahavandi, N.; Craiu, M.; et al. Rett-like onset in late-infantile neuronal ceroid lipofuscinosis (CLN7) caused by compound heterozygous mutation in the MFSD8 gene and review of the literature data on clinical onset signs. *Eur. J. Paediatr. Neurol.* **2015**, *19*, 78–86. [[CrossRef](#)] [[PubMed](#)]
6. Ziats, M.N.; Comeaux, M.S.; Yang, Y.; Scaglia, F.; Elsea, S.H.; Sun, Q.; Beaudet, A.L.; Schaaf, C.P. Improvement of regressive autism symptoms in a child with TMLHE deficiency following carnitine supplementation. *Am. J. Med. Genet. Part A* **2015**, *167*, 2162–2167. [[CrossRef](#)]
7. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **2012**, *485*, 237–241. [[CrossRef](#)]
8. Neale, B.M.; Kou, Y.; Liu, L.; Ma’Ayan, A.; Samocha, K.E.; Sabo, A.; Lin, C.F.; Stevens, C.; Wang, L.S.; Makarov, V.; et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **2012**, *485*, 242–245. [[CrossRef](#)]
9. O’Roak, B.J.; Vives, L.; Girirajan, S.; Karakoc, E.; Krumm, N.; Coe, B.P.; Levy, R.; Ko, A.; Lee, C.; Smith, J.D.; et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **2012**, *485*, 246–250. [[CrossRef](#)]
10. Goes, F.S.; Pirooznia, M.; Parla, J.S.; Kramer, M.; Ghiban, E.; Mavruk, S.; Chen, Y.C.; Monson, E.T.; Willour, V.L.; Karchin, R.; et al. Exome Sequencing of Familial Bipolar Disorder. *JAMA Psychiatry* **2016**, *73*, 590–597. [[CrossRef](#)]
11. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **2014**, *511*, 421–427. [[CrossRef](#)] [[PubMed](#)]
12. Yin, J.; Schaaf, C.P. Autism genetics—An overview. *Prenat. Diagn.* **2017**, *37*, 14–30. [[CrossRef](#)] [[PubMed](#)]
13. Yin, J.; Oleson, D.; Schaaf, C.P. Next Generation Sequencing in Autism Spectrum Disorder. *OBM Genet.* **2018**, *2*, 014. [[CrossRef](#)]
14. Berghuis, B.; De Kovel, C.G.; van Iterson, L.; Lamberts, R.J.; Sander, J.W.; Lindhout, D.; Koeleman, B.P. Complex SCN8A DNA-abnormalities in an individual with therapy resistant absence epilepsy. *Epilepsy Res.* **2015**, *115*, 141–144. [[CrossRef](#)] [[PubMed](#)]
15. Boutry-Kryza, N.; Labalme, A.; Ville, D.; de Bellescize, J.; Touraine, R.; Prieur, F.; Dimassi, S.; Poulat, A.L.; Till, M.; Rossi, M.; et al. Molecular characterization of a cohort of 73 patients with infantile spasms syndrome. *Eur. J. Med. Genet.* **2015**, *58*, 51–58. [[CrossRef](#)]
16. Depienne, C.; Bouteiller, D.; Keren, B.; Cheuret, E.; Poirier, K.; Trouillard, O.; Benyahia, B.; Quelin, C.; Carpentier, W.; Julia, S.; et al. Sporadic Infantile Epileptic Encephalopathy Caused by Mutations in PCDH19 Resembles Dravet Syndrome but Mainly Affects Females. *PLoS Genet.* **2009**, *5*, e1000381. [[CrossRef](#)]
17. Guffanti, G.; Lievers, L.S.; Bonati, M.T.; Marchi, M.; Geronazzo, L.; Nardocci, N.; Estienne, M.; Larizza, L.; Macciardi, F.; Russo, S. Role of UBE3A and ATP10A genes in autism susceptibility region 15q11–q13 in an Italian population: A positive replication for UBE3A. *Psychiatry Res.* **2011**, *185*, 33–38. [[CrossRef](#)]
18. Meyer, A.; Kossow, K.; Gal, A.; Steglich, C.; Mühlhausen, C.; Ullrich, K.; Bräulke, T.; Muschol, N. The Mutation p.Ser298Pro in the sulphamidase gene (SGSH) is associated with a slowly progressive clinical phenotype in mucopolysaccharidosis type IIIA (Sanfilippo A Syndrome). *Hum. Mutat.* **2008**, *29*, 770. [[CrossRef](#)]

19. Ngoh, A.; McTague, A.; Wentzensen, I.M.; Meyer, E.; Applegate, C.; Kossoff, E.H.; Batista, D.A.; Wang, T.; Kurian, M.A. Severe infantile epileptic encephalopathy due to mutations in PLCB1: Expansion of the genotypic and phenotypic disease spectrum. *Dev. Med. Child Neurol.* **2014**, *56*, 1124–1128. [[CrossRef](#)]
20. Noh, G.J.; Graham, J.M. 2q23.1 microdeletion of the MBD5 gene in a female with seizures, developmental delay and distinct dysmorphic features. *Eur. J. Med. Genet.* **2012**, *55*, 354–357. [[CrossRef](#)]
21. Philippe, A.; Craus, Y.; Rio, M.; Bahi-Buisson, N.; Boddart, N.; Malan, V.; Bonnefont, J.P.; Robel, L. Case report: An unexpected link between partial deletion of the SHANK3 gene and Heller’s dementia infantilis, a rare subtype of autism spectrum disorder. *BMC Psychiatry* **2015**, *15*, 256. [[CrossRef](#)]
22. Prasad, A.N.; Rupa, C.A.; Prasad, C. Methylenetetrahydrofolate reductase (MTHFR) deficiency and infantile epilepsy. *Brain Dev.* **2011**, *33*, 758–769. [[CrossRef](#)]
23. Ramocki, M.B.; Chapieski, L.; McDonald, R.O.; Fernandez, F.; Malphrus, A.D. Spinocerebellar ataxia type 2 presenting with cognitive regression in childhood. *J. Child Neurol.* **2008**, *23*, 999–1001. [[CrossRef](#)]
24. Willemsen, M.H.; Rensen, J.H.M.; de Valk, H.V.S.L.; Hamel, B.C.J.; Kleefstra, T. Adult Phenotypes in Angelman- and Rett-Like Syndromes. *Mol. Syndromol.* **2012**, *2*, 217–234. [[CrossRef](#)] [[PubMed](#)]
25. Yamakawa, Y.; Nakazawa, T.; Ishida, A.; Saito, N.; Komatsu, M.; Matsubara, T.; Obinata, K.; Hirose, S.; Okumura, A.; Shimizu, T. A boy with a severe phenotype of succinic semialdehyde dehydrogenase deficiency. *Brain Dev.* **2012**, *34*, 107–112. [[CrossRef](#)]
26. Fan, H.; Gulley, M.L. DNA extraction from paraffin-embedded tissues. *Methods Mol. Med.* **2001**, *49*, 1–4. [[PubMed](#)]
27. Killeen, A.A. (Ed.) *Molecular Pathology Protocols*; Springer Science & Business Media: Berlin, Germany, 2001.
28. Challis, D.; Yu, J.; Evani, U.S.; Jackson, A.R.; Paithankar, S.; Coarfa, C.; Milosavljevic, A.; Gibbs, R.A.; Yu, F. An integrative variant analysis suite for whole exome next-generation sequencing data. *BMC Bioinform.* **2012**, *13*, 8. [[CrossRef](#)] [[PubMed](#)]
29. Reid, J.G.; Carroll, A.; Veeraghavan, N.; Dahdouli, M.; Sundquist, A.; English, A.; Bainbridge, M.; White, S.; Salerno, W.; Buhay, C.; et al. Launching genomics into the cloud: Deployment of Mercury, a next generation sequence analysis pipeline. *BMC Bioinform.* **2014**, *15*, 30. [[CrossRef](#)]
30. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [[CrossRef](#)]
31. Lesca, G.; Rudolf, G.; Bruneau, N.; Lozovaya, N.; Labalme, A.; Boutry-Kryza, N.; Salmi, M.; Tsintsadze, T.; Addis, L.; Motte, J.; et al. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat. Genet.* **2013**, *45*, 1061–1066. [[CrossRef](#)]
32. Aldinger, K.A.; Kogan, J.; Kimonis, V.; Fernandez, B.; Horn, D.; Klopocki, E.; Chung, B.; Toutain, A.; Weksberg, R.; Millen, K.J.; et al. Cerebellar and posterior fossa malformations in patients with autism-associated chromosome 22q13 terminal deletion. *Am. J. Med Genet. Part A* **2013**, *161*, 131–136. [[CrossRef](#)] [[PubMed](#)]
33. Saha, B.; Ypsilanti, A.R.; Boutin, C.; Cremer, H.; Chedotal, A. Plexin-B2 regulates the proliferation and migration of neuroblasts in the postnatal and adult subventricular zone. *J. Neurosci.* **2012**, *32*, 16892–16905. [[CrossRef](#)] [[PubMed](#)]
34. Worzfeld, T.; Puschel, A.W.; Offermanns, S.; Kuner, R. Plexin-B family members demonstrate non-redundant expression patterns in the developing mouse nervous system: An anatomical basis for morphogenetic effects of Sema4D during development. *Eur. J. Neurosci.* **2004**, *19*, 2622–2632. [[CrossRef](#)]
35. Lesca, G.; Rudolf, G.; Labalme, A.; Hirsch, E.; Arzimanoglou, A.; Genton, P.; Motte, J.; de Saint Martin, A.; Valenti, M.P.; Boulay, C.; et al. Epileptic encephalopathies of the Landau-Kleffner and continuous spike and waves during slow-wave sleep types: Genomic dissection makes the link with autism. *Epilepsia* **2012**, *53*, 1526–1538. [[CrossRef](#)]
36. Rojo, D.C.; Hamiwka, L.; McMahon, J.M.; Dibbens, L.M.; Arsov, T.; Suls, A.; Stöberg, T.; Kelley, K.; Wirrell, E.; Appleton, B.; et al. De novo SCN1A mutations in migrating partial seizures of infancy. *Neurology* **2011**, *77*, 380–383. [[CrossRef](#)] [[PubMed](#)]

37. Dibbens, L.M.; Tarpey, P.S.; Hynes, K.; Bayly, M.A.; Scheffer, I.E.; Smith, R.; Bomar, J.; Sutton, E.; Vandeleur, L.; Shoubridge, C.; et al. X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nat. Genet.* **2008**, *40*, 776–781. [[CrossRef](#)]
38. Thomas, M.S.; Knowland, V.C.; Karmiloff-Smith, A. Mechanisms of developmental regression in autism and the broader phenotype: A neural network modeling approach. *Psychol. Rev.* **2011**, *118*, 637–654. [[CrossRef](#)] [[PubMed](#)]
39. Jeste, S.S.; Tuchman, R. Autism Spectrum Disorder and Epilepsy: Two Sides of the Same Coin? *J. Child Neurol.* **2015**, *30*, 1963–1971. [[CrossRef](#)]
40. Schaaf, C.P.; Sabo, A.; Sakai, Y.; Crosby, J.; Muzny, D.; Hawes, A.; Lewis, L.; Akbar, H.; Varghese, R.; Boerwinkle, E.; et al. Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum. Mol. Genet.* **2011**, *20*, 3366–3375. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).