






Article

MicroRNA and MicroRNA-Target Variants Associated with Autism Spectrum Disorder and Related Disorders

Anthony Wong ¹, Anbo Zhou ¹, Xiaolong Cao ¹ , Vaidhyanathan Mahaganapathy ¹ , Marco Azaro ¹, Christine Gwin ¹, Sherri Wilson ¹, Steven Buyske ² , Christopher W. Bartlett ^{3,4}, Judy F. Flax ¹, Linda M. Brzustowicz ^{1,5}  and Jinchuan Xing ^{1,5,*} 

¹ Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA; aw449@scarletmail.rutgers.edu (A.W.); zhouanbo@gmail.com (A.Z.); atps@outlook.com (X.C.); vaidhyanathan.m@rutgers.edu (V.M.); azaro@hginj.rutgers.edu (M.A.); gwin@dls.rutgers.edu (C.G.); sherriw@hginj.rutgers.edu (S.W.); judyflax17@gmail.com (J.F.F.); lbrz@hginj.rutgers.edu (L.M.B.)

² Department of Statistics, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA; buyske@stat.rutgers.edu

³ The Steve & Cindy Rasmussen Institute for Genomic Medicine, Battelle Center for Computational Biology, Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH 43215, USA; christopher.bartlett@nationwidechildrens.org

⁴ Department of Pediatrics, College of Medicine, The Ohio State University, Columbus, OH 43210, USA

⁵ Human Genetics Institute of New Jersey, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

* Correspondence: jinchuan.xing@rutgers.edu



Citation: Wong, A.; Zhou, A.; Cao, X.; Mahaganapathy, V.; Azaro, M.; Gwin, C.; Wilson, S.; Buyske, S.; Bartlett, C.W.; Flax, J.F.; et al. MicroRNA and MicroRNA-Target Variants Associated with Autism Spectrum Disorder and Related Disorders. *Genes* **2022**, *13*, 1329. <https://doi.org/10.3390/genes13081329>

Academic Editor: Andreas G. Chiochetti

Received: 3 July 2022

Accepted: 22 July 2022

Published: 26 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Autism spectrum disorder (ASD) is a childhood neurodevelopmental disorder with a complex and heterogeneous genetic etiology. MicroRNA (miRNA), a class of small non-coding RNAs, could regulate ASD risk genes post-transcriptionally and affect broad molecular pathways related to ASD and associated disorders. Using whole-genome sequencing, we analyzed 272 samples in 73 families in the New Jersey Language and Autism Genetics Study (NJLAGS) cohort. Families with at least one ASD patient were recruited and were further assessed for language impairment, reading impairment, and other associated phenotypes. A total of 5104 miRNA variants and 1,181,148 3' untranslated region (3' UTR) variants were identified in the dataset. After applying several filtering criteria, including population allele frequency, brain expression, miRNA functional regions, and inheritance patterns, we identified high-confidence variants in five brain-expressed miRNAs (targeting 326 genes) and 3' UTR miRNA target regions of 152 genes. Some genes, such as *SCP2* and *UCGC*, were identified in multiple families. Using Gene Ontology overrepresentation analysis and protein–protein interaction network analysis, we identified clusters of genes and pathways that are important for neurodevelopment. The miRNAs and miRNA target genes identified in this study are potentially involved in neurodevelopmental disorders and should be considered for further functional studies.

Keywords: whole-genome sequencing; miRNA; autism spectrum disorder; family cohort; 3' UTR; neurodevelopmental disorder

1. Introduction

Autism spectrum disorder (ASD) is a multifaceted neurodevelopmental disorder (NDD) identified by neurodevelopmental impairment, deficits in social communication, and repetitive patterns of behavior [1]. Published studies by the Center for Disease Control and Prevention (CDC) and the New Jersey Autism Study showed alarming rates of ASD prevalence in children, i.e., 1 in 59 nationally [2]. The high prevalence of ASD has made it essential to study the biological mechanism of, and identify predictive biomarkers for, the disorder.

ASD is highly heritable, with a recurrence rate around 20% among family members [3]. Previous studies of the genetic etiology of ASD have identified a large number of genetic risk factors that contribute to the disorder. These risk factors can be broadly divided into three classes: (1) small genetic variants (i.e., single nucleotide variants (SNVs) and insertion/deletion variants (indels)) in the protein coding regions and non-coding regions (~50% of ASD cases) [4,5]; (2) rare copy-number variants (5–10% of ASD cases) and chromosomal abnormalities (~5% of ASD cases); and (3) ASD-related genetic syndromes (~10% of cases) [4]. At the gene level, nearly 1000 genes are implicated in ASD. Although the genome-wide genetic risk based on common variants can be estimated for a patient, due to the heterogeneity of ASD, each identified mutation contributes only a small percentage of known cases, and in each patient, the combination of variants that resulted in ASD is mostly undetermined.

MicroRNAs (miRNAs) are a type of short, non-coding genes that are involved in post-transcriptional gene regulation. A single miRNA can bind to several hundred target sites in the 3' untranslated regions (3' UTR) of mRNAs and thus control a large variety of molecular pathways that could be involved in ASD pathophysiology. Indeed, miRNAs are expressed in different brain regions and have crucial functions during brain development [6]. Previous studies of dysregulated miRNAs identified several candidates [7,8]. For example, a study using mouse neurons discovered that miR-137 and miR-132 regulate two autism-related genes: *PTEN* and *MeCP2* [7]. A study of model mice with Angelman Syndrome discovered down-regulation of miR-708 in the brain, which led to aberrant Ca²⁺ signaling [8]. Thus, miRNA dysfunction could play a role in explaining the complex and heterogeneous nature of ASD and related NDDs.

While there are previous whole-genome sequencing (WGS) studies of miRNA mutations and their target regions in ASD patients [9,10], few if any studies have quantified the relationship between putative genes in ASD and other associated NDDs in a familial setting. In this study, we utilized WGS data from the New Jersey Language and Autism Genetics Study (NJLAGS) to examine the contribution of miRNAs to ASD and related disorders. NJLAGS is designed to detect genetic variation related to both ASD and language impairment [11]. More than 150 families were recruited in NJLAGS and this cohort provides the architecture to assess the contribution of both segregating and *de novo* variants to ASD genetic etiology. Here, we identified high-confidence SNVs and indels in miRNA and miRNA target regions from the NJLAGS cohort. We propose that mutations that affect miRNAs and their mRNA transcript binding sites could cause abnormal gene expression for ASD implicated genes and contribute to ASD and related NDDs in NJLAGS families.

2. Materials and Methods

2.1. Family Selection and Phenotyping

NJLAGS collected data from ASD probands and their family members for genetic analysis. Each individual was subjected to measures of ASD, oral language impairment (LI), written language impairment (RI), and social responsiveness (SRS). The assessment and diagnosis details were described previously [11]. Briefly, ASD diagnoses were based on the Autism Diagnostic Interview (ADI-R), Autism Diagnostic Observation Schedule (ADOS), and the Diagnostics and Statistical Manual-IV (DSM-IV), or DSM-5, depending on the recruitment date (see [11] for details). LI is defined as a score of less than 85 on the CELF-4 or a history of language/reading difficulties with at least a score of one standard deviation below their peers on at least 60% of oral language subtests. RI is defined as a score of one standard deviation below the mean of 60% on all reading tests. SRS is measured with the Social Responsiveness Scale. The cut-off for a dichotomous trait social deficit (SRS-DT) was a score of 54 for males and 45 for females which was equivalent to the T-score > 60 criteria used to identify mild impairment in children [11]. ADHD affection status for ASD probands and their family members was determined based on the relevant questionnaires, as described previously [12]. This study was approved by the Institutional Review Board at Rutgers, State University of New Jersey (IRB number: 13-112Mc).

2.2. miRNA Variant Identification

A total of 272 individuals were subject to WGS (Table S1). The details of sequencing and variant identification are described previously [12]. The SNVs and indels were obtained from a previous study [12] and are available in the National Institute of Mental Health Data Archive (NDA) under experiments C1932 and C2933.

The genomic positions of human miRNA genes (genome assembly GRCh37) were obtained from miRBase v22 [13]. The database includes 1917 miRNA genes in the human genome [14]. For variants within a known miRNA, ADmiRE [15,16] was used to classify the variants by their locations in miRNA functional regions (i.e., seed, mature, etc.). The seed region annotation indicates a variant lies within the core miRNA/mRNA binding region, while the mature region indicates a variant lies outside of the seed region but still within a functionally mature miRNA that is incorporated into the RNA-induced silencing complex. The arm and loop regions form the remaining hairpin of the precursor miRNA structure.

2.3. MiRNA Binding Site Identification

TargetScanHuman [17] was used to predict potential miRNA target genes [18]. Binding sites are categorized as broadly conserved across vertebrate species, conserved across mammals, or poorly conserved by aligning sequences across 28 different species [19]. Broadly conserved and conserved binding sites were considered high-confidence candidates and selected for downstream analysis. Using this data, 3' UTR variants were overlapped with TargetScanHuman conserved binding sites to identify 3' UTR variants that could disrupt miRNA-mRNA binding. To characterize the gene networks regulated by candidate miRNAs, miRNA predicted targets with a context++ score (CS) of < -0.4 was used to select sites with high targeting efficacy [18].

2.4. Characterizing Variants by Inheritance Modes

GEMINI v0.20.1 [20] is a computational framework that identifies variants following autosomal dominant, autosomal recessive, or *de novo* mode of inheritance [21]. Both pedigree and variant call information was loaded into GEMINI using the following command:

```
gemini load -v input.vcf -p input.ped
```

Variants were then identified as being autosomal dominant, autosomal recessive, or *de novo* for each phenotype by executing:

```
gemini autosomal_dominant/autosomal_recessive/de_novo input.db
```

GEMINI strict inheritance rules were applied, where at least a full trio (i.e., both parents are present for a proband) is required to infer the inheritance pattern.

2.5. Reference Population Dataset

WGS data from 15,708 (gnomAD v.2.1.1 non-neuro genome, Broad Institute, Cambridge, MA, USA) and 76,156 (gnomAD v.3 non-neuro, Broad Institute, Cambridge, MA, USA) unrelated individuals without a neurological condition, and whole-exome sequencing data from 125,748 unrelated individuals without a neurological condition (gnomAD v.2.1.1 non-neuro exome, Broad Institute, Cambridge, MA, USA) were obtained from the Genome Aggregation Database (gnomAD) [22] and used as reference populations. Variants that failed gnomAD quality control metrics (i.e., with RF, AC0, InbreedingCoeff, etc. in the "Filters" field) were excluded. Variant alternative allele frequencies (AF) in gnomAD populations were used to identify rare alleles in the general population. An AF cutoff of 1% for dominant and recessive and 0.1% for *de novo* variants in all datasets was applied for miRNA and 3' UTR variants. For variants which do not have AF values in all gnomAD datasets, the AF cutoffs were applied based on available databases. Variants that are not identified in gnomAD databases were instead filtered by cohort AF among the NJLAGS samples (1% for dominant and recessive and absent in all other samples for *de novo* variants).

2.6. Brain Expression Data

Expressions of miRNA and protein-coding genes in the brain were extracted from the Gene Tissue Expression project (GTEx) [23], the BrainSpan Atlas of the Developing Human Brain project [24], and the Human Developmental Biology Resource (HDBR) [25]. Detailed data processing procedures were described previously [26]. We selected genes and miRNAs that have a transcript per million (TPM) value greater than 5 in any of the datasets.

2.7. Gene Ontology Enrichment Analysis and Protein–Protein Interaction Network Analysis

Gene Ontology (GO) enrichment analysis was performed using ConsensusPathDB [27]. GO terms with a false-discovery rate (FDR) < 0.05 were considered enriched and terms with less than 600 total genes within the group were selected to increase the specificity of the enrichment results.

A protein–protein interaction network was built to investigate the potential interactions between candidate miRNA target genes and 3' UTR genes. ConsensusPathDB [27], STRING [28], and GIANT_v2 [29] were used to generate a list of known gene interactions. Detailed data processing procedures were described previously [26,30]. Edges with at least two database interactions were considered high-confidence interactions for the network. Additional known NDD genes collected from previous studies (Table S2, [26]) were included in the network analysis. Isolated and self-directed interactions were removed from the network.

3. Results

3.1. Cohort Description

We obtained WGS variants of 272 individuals from 73 families (Table S1). 65 males and 18 females were diagnosed with ASD (Table 1). All 73 families contained at least one individual diagnosed with ASD, LI, or RI, and 59 and 47 families contained at least one individual diagnosed with SRS and ADHD. Across all disorders, 35 and 67 families fit a dominant or a recessive/*de novo* mode of inheritance, respectively (Table 1). Due to a lack of ASD affected parents in the cohort of sequenced individuals, no families met the criteria for dominant inheritance pattern for ASD.

Table 1. Summary of Samples and Families.

	Patients	Male	Female	Families	Dominant	Recessive/ <i>de novo</i>
ASD	83	65	18	73	0	67
LI	117	86	31	73	9	58
RI	134	96	38	73	9	42
SRS	83	60	23	59	25	43
ADHD	63	43	20	47	10	24
All Samples	272	166	106	73	35	67

The first three columns are the total number of affected individuals, male, and female affected individuals, respectively. Families indicate the total number of families that contain at least one affected individual with the respective phenotype. Dominant and Recessive/*de novo* are the number of families that meet the criteria for the specific mode of inheritance.

3.2. miRNA Variant Identification

Out of 25,987,740 variants, 5104 variants overlapped miRNA genes and were annotated for the miRNA regions (e.g., seed, mature, etc.). Using the GEMINI program, we categorized the variants by their inheritance pattern for each phenotype. A total of 632 variants matched one of the segregation patterns (dominant, recessive, or *de novo*) (Figure 1, see Table S3 for details). Among them, we selected the 193 miRNA variants in seed and mature regions for downstream analysis, because variants in mature miRNAs and seed regions are most likely to affect the miRNA function. Among the 193 variants, 110

are within miRNAs expressed in the brain (see Methods for detail). To select rare variants, we applied AF filtering of less than 1% for dominant and recessive variants and 0.1% for *de novo* variants in individuals in the gnomAD database. After filtering, five unique miRNA variants from four families meet the AF cutoff under the dominant mode of inheritance (Table 2).

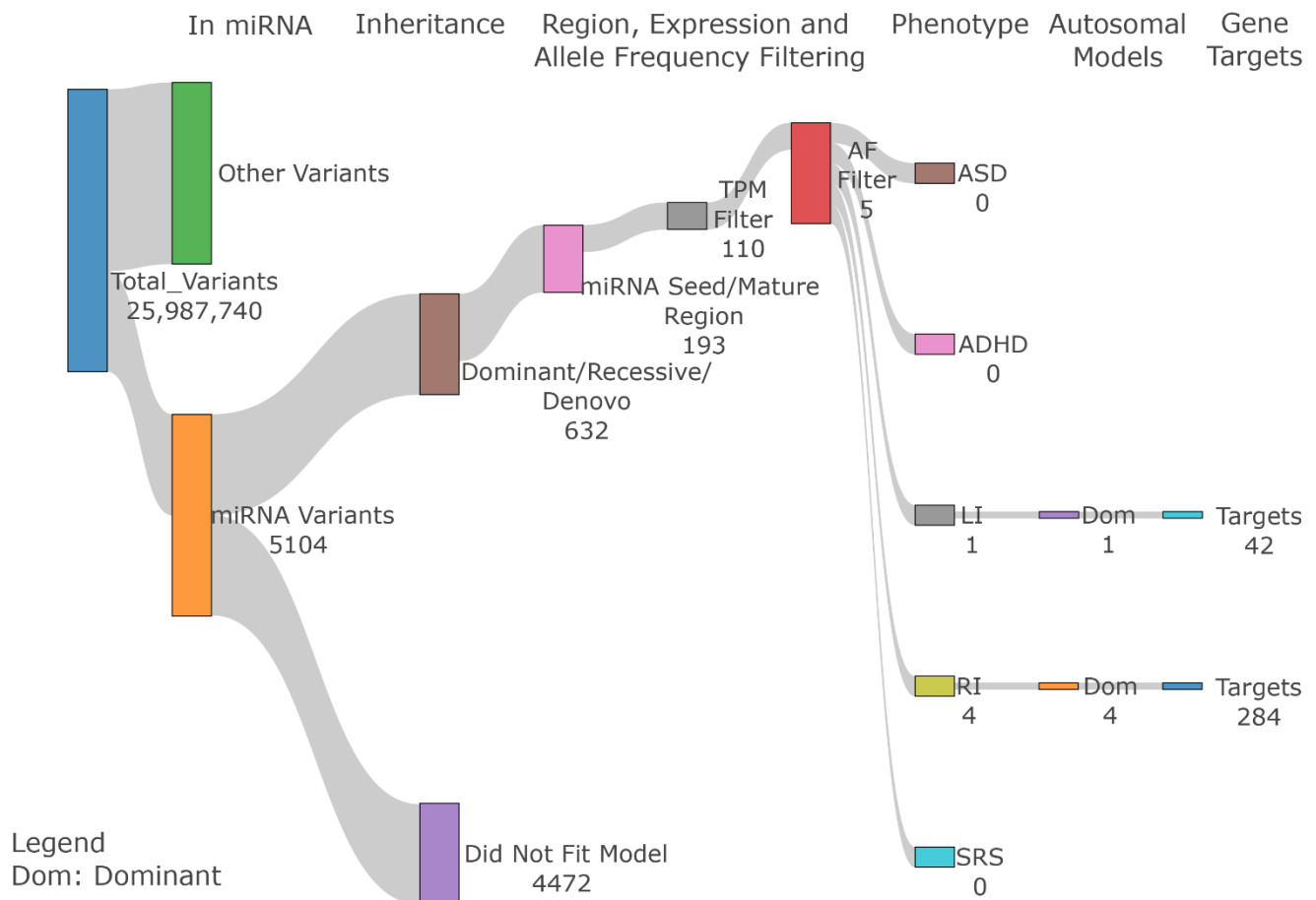


Figure 1. miRNA variant filtering steps. After selecting all variants within miRNAs, we filtered variants based on their inheritance pattern, miRNA region annotation, brain expression pattern, and population allele frequency to select final candidate variants.

Out of the five variants, two variants, rs371749301 and rs550720421, were identified in family FAM58 and segregated with the RI phenotype. Two additional variants (rs200279579, rs761222509) segregated with RI in families FAM5 and FAM66 and one segregated with LI (rs565141718) in family FAM13, respectively (Table 2).

Table 2. miRNA candidate variants.

miRNA	Region	Target Genes	chr	pos	rsID	Ref	Alt	AF_2_exome	AF_2_genome	AF_3	Family	Inheritance	Phenotype
hsa-miR-6780a-3p	seed	179	chr17	40860121	rs200279579	A	G	1.23×10^{-4}	9.42×10^{-5}	1.26×10^{-4}	FAM5	dominant	RI
hsa-miR-1225-5p	mature	43	chr16	2140269	rs371749301	C	A	1.48×10^{-5}	4.72×10^{-5}	2.00×10^{-4}	FAM58	dominant	RI
hsa-miR-2277-3p	mature	37	chr5	92956416	rs550720421	G	T	4.62×10^{-4}	4.24×10^{-4}	8.83×10^{-4}	FAM58	dominant	RI
hsa-miR-548j-5p	mature	42	chr22	26951249	rs565141718	C	T	3.95×10^{-4}	4.70×10^{-5}	1.41×10^{-4}	FAM13	dominant	LI
hsa-miR-100-5p	mature	26	chr11	122022992	rs761222509	G	A	1.93×10^{-5}	N/A	7.42×10^{-6}	FAM66	dominant	RI

miRNA: HGNC Symbol. Region: miRNA region. Target Genes: the number of miRNA target genes predicted by TargetScanHuman. chr: Chromosome. pos: Variant position. rsID: DBSNP150 reference SNV number. Ref: Reference allele. Alt: Alternate Allele. AF_2_exome: Population AF from gnomAD v2 non-neuro exome database. AF_2_genome: Population AF from gnomAD v2 non-neuro genome database. AF_3: Population AF from gnomAD v3_non_neuro database. Family: Affected Family ID. Inheritance: Inheritance mode for the affected family. Phenotype: Phenotype of the affected individual.

3.3. 3' UTR miRNA Binding Site Variant Identification

In addition to miRNA variants, variants in miRNA binding sites in the 3' UTR of miRNA target genes can also affect miRNA regulation. Next, we identified variants in the miRNA binding sites. In the dataset, 1,181,148 variants were identified within 3' UTR regions. These variants were subject to similar variant filtering criteria as miRNA variants (Figure 2). Using GEMINI, 155,802 variants fell within one of the three inheritance pattern categories. TargetsCanHuman was then used to identify 1092 variants that overlapped predicted miRNA conserved target sites (context++ score < -0.4). Applying an AF filter of 1% for dominant and recessive variants and 0.1% for *de novo* variants resulted in 153 unique variants in 152 brain-express genes (Table 3), see Supplemental Table S4 for details). A total of 6, 9, 18, and 7 families containing variants segregate in the dominant pattern for ADHD, LI, RI, and SRS, respectively (Figure 3A). Because some patients have multiple diagnoses, some variants meet certain segregation pattern in multiple phenotypes (Table S4). Therefore, 29, 6, and 10 unique families contain variants segregating in dominant, recessive, and *de novo* patterns, respectively (Figure 3).

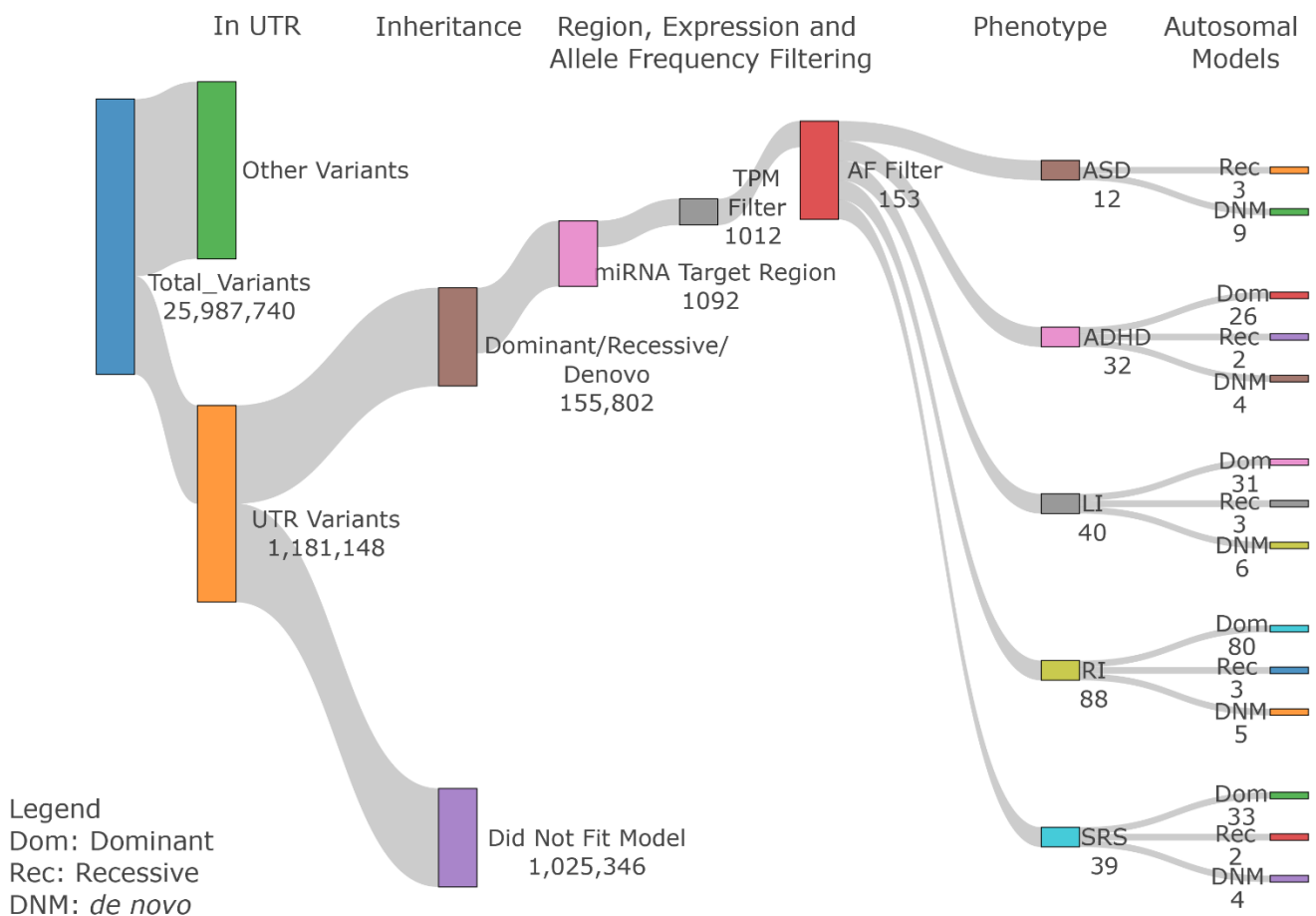
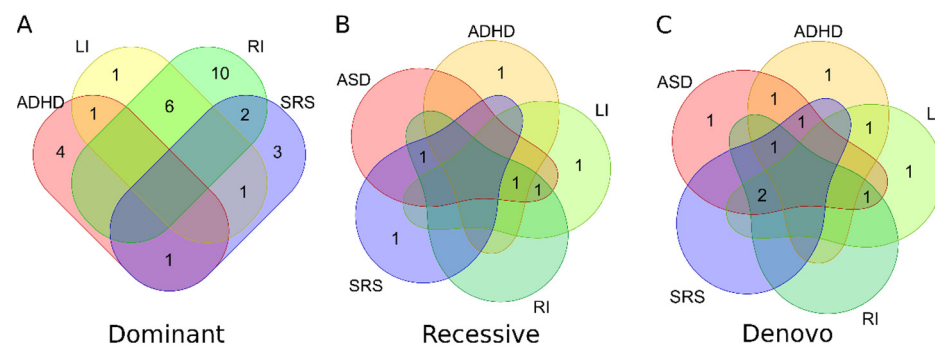


Figure 2. 3' UTR variant filtering steps. After selecting all variants within 3' UTR regions, we filtered variants based on their inheritance pattern, miRNA target region annotation, brain expression pattern, and population allele frequency to select final candidate variants.

Table 3. Summary of 3'UTR Candidate Genes. Total number of genes found in each phenotype and inheritance pattern.

	ASD	ADHD	LI	RI	SRS	Total Unique
Dominant (AF < 1%)	0	26	30	79	33	138
Recessive (AF < 1%)	3	2	3	3	2	5
<i>de novo</i> (AF < 0.1%)	9	4	6	5	4	10
Total Unique	12	32	39	87	39	152

**Figure 3.** Number of families with 3' UTR variants. Venn Diagrams of families with 3' UTR variants for each inheritance pattern: (A) Dominant; (B) Recessive; (C) *de novo*. The value indicates the number of families with candidate variants for the phenotype(s).

3.4. Candidate Gene Analysis

To understand the functional impact of variants in miRNAs and their binding sites on genes and molecular pathways, we first identified target genes of segregating miRNA variants. We selected genes that are expressed in brain tissues for their biological relevance (see Methods for detail). RI had four miRNAs with 284 target genes, and LI had one miRNA with 42 target genes (Table S5). Next, we selected brain-expressed genes containing segregating variants in the 3' UTR miRNA binding sites. A total of 12, 32, 39, 87, and 39 unique genes were selected for ASD, ADHD, LI, RI, and SRS, respectively (Table 3 and Table S5).

Four genes, *ONECUT2*, *OSBP*, *SCP2*, and *TUB*, appeared as targets of candidate miRNAs and contained variants in their 3' UTR miRNA target region (Table 4). For example, Family FAM5 carried a mutation within the seed region of hsa-miR-6780a-3p (rs200279579; chr17:40860121-A-G) which targets *SCP2* (Sterol Carrier Protein 2). Family FAM2 contained a mutation within the *SCP2* 3' UTR region (rs182947399; chr1:53516762-A-T) which affects the *SCP2*/miR-150-5p binding axis. Both mutations segregate in a dominant pattern with the RI phenotype.

Two candidate genes, *RBM24* and *UGCG*, overlapped multiple phenotypes in more than one family (Table 4). *UGCG* (UDP-Glucose Ceramide Glucosyltransferase) is a key gene for the biosynthesis of glycosphingolipids. The dominant variant (rs201977317; chr9:114695431-GA-G) appeared in family FAM5 for RI and families FAM5, FAM36, and FAM37 for SRS. The importance of *UGCG* and the role of glycosphingolipids in brain development has been reported previously [31,32].

Table 4. Top Candidate Genes.

Gene		miRNA Variant							3' UTR Variant						
Gene	Max Brain Expression	Other NDD	pLI	Previous Studies	miRNA	Inheritance	Family	Phenotype	chr	pos	rsID	AF_gnomAD	Inheritance	Family	Phenotype
<i>ONECUT2</i>	66	NA	0.91	[33]	hsa-miR-6780a-3p	dominant	FAM5	RI	chr18	55155836	rs147208471	7.27×10^{-4}	dominant	FAM56	SRS
<i>OSBP</i>	61.03	NA	1.00	[34]	hsa-miR-6780a-3p	dominant	FAM5	RI	chr11	59343007	rs149325846	3.61×10^{-3}	dominant	FAM14; FAM66	RI
<i>SCP2</i>	65.43	ASD_Low	0.00	[35]	hsa-miR-6780a-3p	dominant	FAM5	RI	chr1	53516762	rs182947399	1.33×10^{-3}	dominant	FAM2	ADHD
<i>TUB</i>	187	NA	0.00	[36,37]	hsa-miR-6780a-3p	dominant	FAM5	RI	chr11	8123523	rs1379616749	3.12×10^{-4}	dominant	FAM70	RI
<i>RBM24</i>	49.40	NA	0.09	[38]					chr6	17292448	rs914886490	2.60×10^{-4}	<i>de novo</i>	FAM5; FAM59	ADHD,LI
<i>UGCG</i>	86	NA	0.94	[31,32]					chr9	114695431	rs201977317	4.54×10^{-3}	dominant	FAM5; FAM36; FAM37	RI,SRS

Other NDD: genes implicated in NDDs from previous studies (Table S2). pLI: the probability the gene is loss-of-function intolerant. AF_gnomAD: maximum variant AF in the three gnomAD databases. Other headers are the same as in Table 2.

3.5. Gene Ontology Terms Early Forebrain Patterning and Apoptotic Process Involved in Development Are Enriched

To explore the possible shared relationship and affected pathways between our two sets of genes, we performed a GO over-representation analysis to identify neurodevelopmental terms in each phenotype (Table S6). Each gene list showed enrichment in pathways and terms related to brain structure and development. For example, “Forebrain dorsal/ventral pattern formation” (GO:0021798, $q = 0.0038$) is a small category consisting of five genes, three of which (*GSX2*, *SIX3*, *NKX2-1*) appear in our RI candidate genes. Other top terms include “apoptotic process involved in development” (GO:1902742, $q = 0.003$) and “actin cytoskeleton” (GO:0015629, $q = 0.038$) in LI and ASD candidate genes, respectively, and may suggest the importance of the regulation of apoptosis and actin in the context of neurodevelopment.

3.6. Protein–Protein Interactions

A protein–protein interaction (PPI) network can help identify shared etiology among families despite the heterogeneity of single gene variants among individuals. Therefore, we constructed a PPI network for both miRNA target genes and 3' UTR genes. Seventeen out of 327 miRNA target genes and 18 out of 152 3' UTR genes have high-confidence interactions within a single network (Figure 4, Table S7). An enrichment analysis on the 35 connected genes showed enrichment in terms such as “regulation of neuron death” (GO:1901214, $q = 1.21 \times 10^{-4}$). While these genes are not directly connected, it demonstrates a control of neuronal development from many different pathways that our candidate genes may be involved in. One of our top terms was “calcium-mediated signaling” (GO:0019722, $q = 2.23 \times 10^{-5}$) and overlapped 6 genes. Other terms are listed in Table S8. In addition, five genes, *MTOR*, *AGO1*, *EP300*, *XPO1*, and *PPP3CA*, have been implicated in ASD by the SFARI database. *AGO1* (Argonaute RISC Component 1), is the protein directly involved in the miRNA-mRNA binding complex and lends further credence to miRNA involvement in ASD and related NDDs.

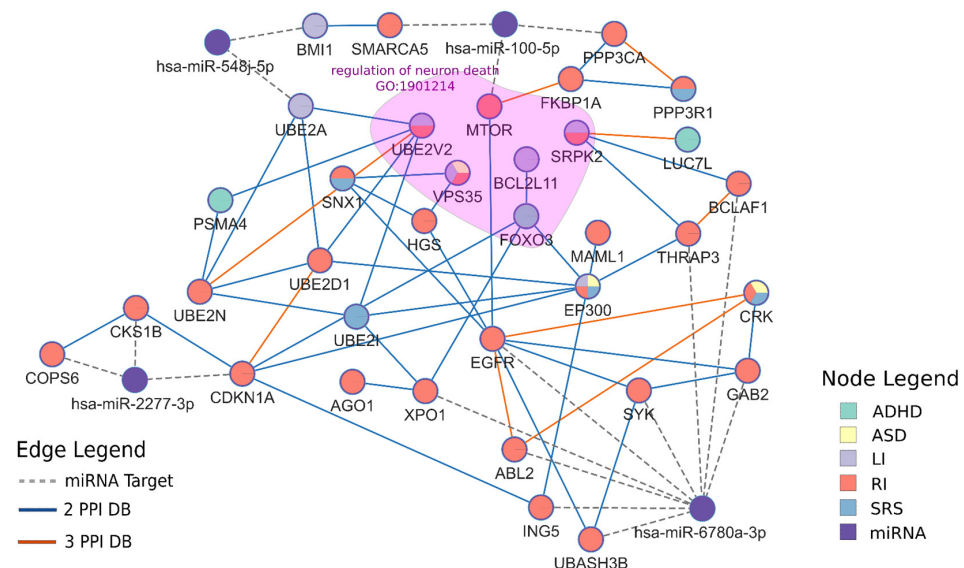


Figure 4. Candidate gene PPI network. The network includes both miRNA target genes and 3' UTR genes (Table S9). PPI edges are included if at least two databases show evidence of interaction. The highlighted region included genes enriched in regulation of neuron death (GO:1901214).

We then constructed additional networks between our candidate genes and known NDD genes to further explore the effect of our genes in NDDs (Figure 5, Table S10). The ADHD network (Figure 5B) shows a tightly clustered set of calcium voltage gated and glutamate receptor genes, both of which have been shown to be possible mechanisms linked to ASD and ADHD [39]. The RI and LI PPI networks contain genes involved in chromatin

and transcription factor binding and cell cycle control (Figure 5C,D). Common genes that appear in multiple networks include *MAPK3* and *PTEN*. *MAPK3* is located within the chromosome 16p11.2 band that is associated with 16p11.2 deletion syndrome, characterized by intellectual disability and developmental delay [40]. *PTEN* has been reported to be involved in a miRNA regulatory network and mutations in *PTEN* are present in 20% of children with both ASD and macrocephaly [7,41]. Candidate genes also interact with *NTRK1* and *NTRK3*, Neurotrophic Receptor Tyrosine Kinases, which interact with nerve growth factors.

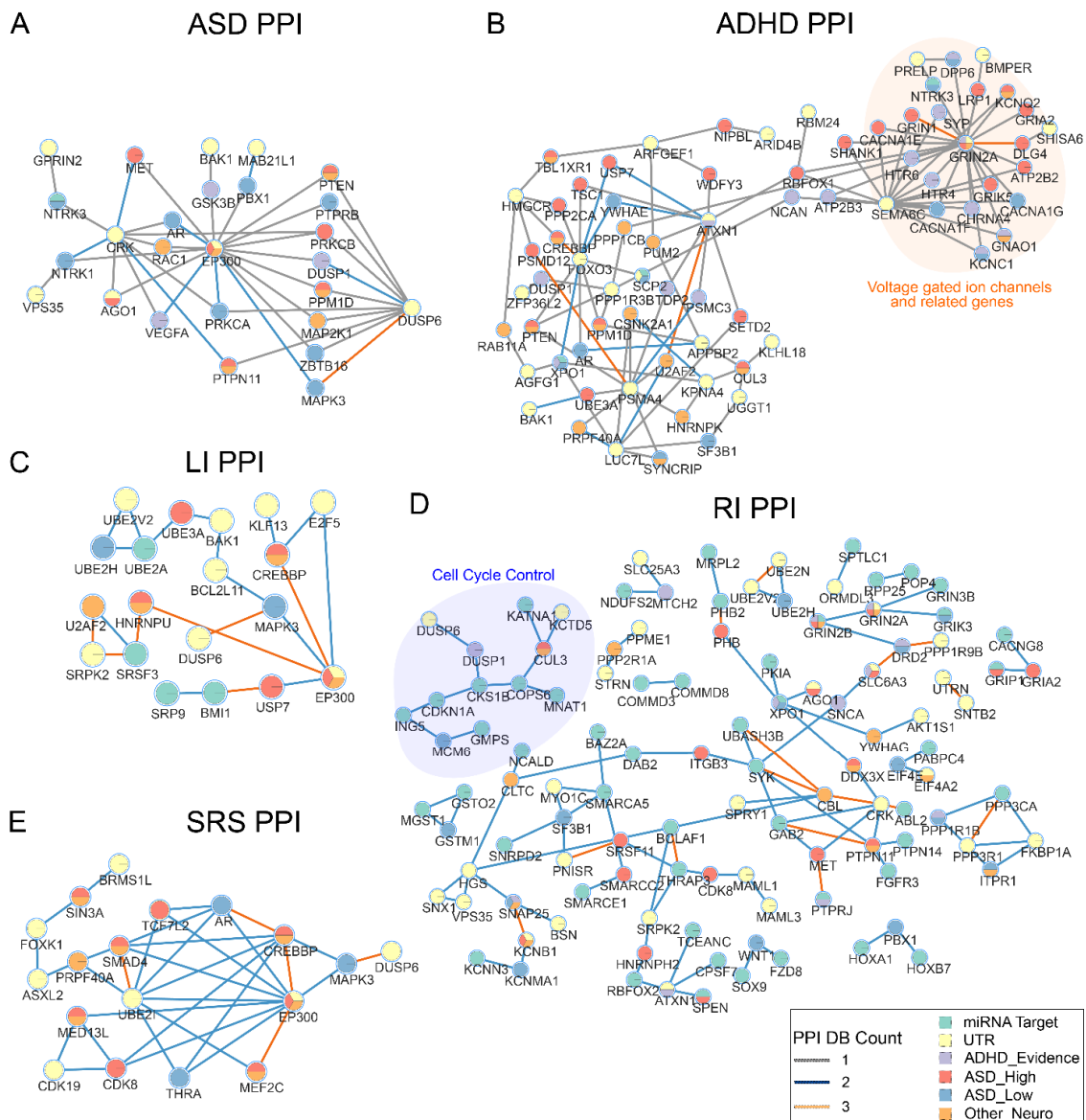


Figure 5. Candidate gene interaction network with known NDD genes. Known genes implicated in ASD, ADHD, and other NDDs (Table S2) were included in the PPI network with candidate genes in this study. (A) ASD candidates PPI. (B) ADHD candidates PPI. (C) LI candidates PPI. (D) RI candidates PPI. Nodes with degree greater than 13 were hidden from the network for clarity. (E) SRS candidates PPI. Isolated nodes were removed from the networks. Connections between non-candidate genes were not shown. Edges with only one protein–protein database evidence were not shown in (C,E). Highlighted regions included genes in GO terms.

4. Discussion

The genetic etiology of ASD has been studied extensively; however, many ASD patients have no known genetic cause [4]. Due to the role of miRNA in brain development, variants in miRNAs and their binding targets in 3' UTRs could contribute genetically to the emergence of ASD. In this study, we sought to identify candidate genes affected by miRNA/mRNA interaction variants by leveraging the WGS data generated from families collected by NJLAGS.

By studying segregating variants that were also rare in the general population, we identified five miRNA and 153 3' UTR high-confidence variants in 39 families across the five phenotypes. Although the large number of variants prevented us from performing experimental validation, all five miRNA candidate variants and the vast majority of the 3' UTR variants (146 out of 153) were either reported as high-quality variants in the gnomAD project or in multiple samples in our dataset. Combining with our stringent genotype segregation screening, we expect that most of the candidate variants are authentic. These variants are located in the seed/mature region of miRNAs or in conserved 3' UTR miRNA target sites and have a high probability of disrupting the complementary base pairing in the miRNA-mRNA complex. Six genes (Table 4) have high-confidence variants in multiple families, making them strong candidates as risk genes. *TUB* (TUB Bipartite Transcription Factor) and *SCP2* have been studied in individuals with neuronal disorders [35,36]. *TUB* is part of the Tubby signaling pathways that are important during neurodevelopment [37]. A *TUB* frameshift mutation in three siblings was associated with retinal dystrophy and obesity [36]. *SCP2* is responsible for mediating the transfer of common phospholipids, cholesterol, and gangliosides, and is implicated in Leukoencephalopathy with Dystonia and Motor Neuropathy [35]. For the remaining genes, *in vivo* studies provide evidence of their role in neuronal development. *ONECUT2* (One Cut Homeobox 2) is known to regulate early retinal progenitor cells [33]. A study of N2A mice neuroblast cell lines discovered that the overexpression of *OSBP* (Oxysterol Binding Protein) decreased miR-124 mediated neurite growth [34]. *RBM24* (RNA Binding Motif Protein 24) is an RNA-binding protein that is responsible for skeletal myogenesis, heart development, vertebrae sensory organ differentiation, and embryonic germ layer formation [38]. *UGCG* (UDP-Glucose Ceramide Glucosyltransferase) is important in the neural differentiation process [31]. *Ugcg* inhibition in embryonic mouse cell lines was found to decrease neural cell marker proteins *GFAP* and *MAP-2* [31]. While the six candidate genes appear in different processes, possibly due to the heterogeneity of the disorders, previous studies have shown that each gene plays important roles during neuronal development, and the disruption of their expression regulation could have contributed to the disorders in these families.

As expected with the heterogeneity of ASD, our study found hundreds of genes that were unique to single families. To further elucidate the contribution of these candidate genes, we conducted functional enrichment analyses and constructed PPI networks with known NDD genes. The results showed that our miRNA target genes are involved in processes which are important to neurodevelopment. For example, the ASD phenotype showed enrichment of the GO term "actin cytoskeleton", which is a component of dendritic spine formation and plasticity [42]. A study of stem cells from human exfoliated deciduous teeth of 13 ASD patients showed impaired actin polymerization in ASD patients and demonstrated a possible mechanism of NDDs [43]. Three genes from our RI group, *GSX2*, *SIX3*, and *NKX2-1*, are known to be expressed in the early forebrain and are responsible for forebrain dorsal-ventral pattern formation during early cortical development [44–46], an important developmental stage for ASD [47]. Additional functional analysis of our 37 genes in a PPI network (Figure 4) resulted in an enrichment of regulation of neuron death. Neuronal apoptosis is an important mechanism during neurodevelopment and has been implicated in ASD. For example, a comparison of neuron counts of seven autistic and six control children found 67% more neurons in the prefrontal cortex in autistic children than the mean. Additionally, a failure of subplate apoptosis in the prenatal brain is proposed as one reason for increase neuron counts in ASD children [48]. A look into our phenotype

subgraphs (Figure 5) reveals that our candidate genes target some key autism risk genes. For example, *PTEN* appears to be strongly connected in our network and has appeared as a high confidence risk gene for individuals with ASD and macrocephaly [49–51]. Overall, our findings as well as previous studies suggest that our discovered variants affect genes that have an important impact on neurogenesis.

Our study had a few limitations. First, it had a relatively small cohort size. This necessitates further study with a larger sample size for increased statistical power. We also only examined variants with a simple segregation pattern and limited our analysis to interactions among variants. With a large sample size, methods incorporating variants with incomplete penetrance (i.e., not following a strict segregation pattern) could further improve the power of the analysis. Expanding the region of analysis beyond miRNA seed regions paired with tools to predict gain or loss of binding sites such as SBSA [52] could also aid in identifying additional pathogenic variants in future works.

Psychiatric disorders in general have a large polygenic component [53], with variants distributed throughout the genome with varying effect sizes [54]. The genetic overlap between psychiatric disorders is considerable but not total. The importance of understanding the degree and specificity of overlap between disorders and closely related phenotypes is multifaceted, including implications of homogeneity in recruiting subjects for clinical research, understanding shared genetic risk, and for biological studies across disease contexts. Here, we focused on ASD, ADHD, LI, RI, and the SRS. Evaluating the relationship between ASD, LI, and RI was among the main goals of the NJLAGS study, with recruiting choices intended to identify families with high genetic burden for both ASD and language impairments (both spoken, LI, and written, RI) [11,55–58]. The SRS is a quantitative scale that captures a key component of the ASD phenotype, since social skills are a pillar of diagnosis [59]. To date, the relationship between language (LI/RI) and ASD has shown to be complex [11,58], with this study further showing indications of the polygenic nature of that relationship. ADHD is well known to be genetically related to RI, but again, when applied to families selected for ASD, the relationship is complex and requires further studies [60].

5. Conclusions

In conclusion, our study identified a number of high-confidence segregating variants that could affect the miRNA/mRNA regulation pathway. Many candidate genes and pathways play important roles in neuronal development and could contribute to the etiology of ASD and associated disorders. Previous studies have identified non-coding regulatory mutations associated with ASD. While our study did not replicate the findings from these studies [9,10], our results support the miRNA regulation pathway as an important contributor to the etiology of ASD and related NDDs. Future studies with experimental functional validation could further elucidate the roles of our identified genes in the etiology of ASD and related NDDs.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes13081329/s1>, Table S1: Summary of Family Pedigrees; Table S2: Previously implicated NDD Genes; Table S3: Summary of miRNA variants; Table S4: Summary of 3′-UTR variants; Table S5: Summary of Target Genes; Table S6: Candidate Gene GO Analysis; Table S7: Summary of 35 Genes in Figure 4; Table S8: GO Analysis of Genes in Figure 4; Table S9: Summary of Network Interactions in Figure 4; Table S10: Summary of Network Interactions in Figure 5.

Author Contributions: Conceptualization, L.M.B. and J.X.; formal analysis, A.W., A.Z., X.C., V.M., C.W.B. and J.X.; data curation, A.W., A.Z., V.M., C.W.B., M.A., C.G., S.W., S.B., J.F.F., L.M.B. and J.X.; visualization, A.W. and J.X.; writing—original draft preparation, A.W. and J.X.; writing—review and editing, A.W., A.Z., X.C., C.W.B., J.F.F., L.M.B. and J.X.; supervision, L.M.B. and J.X.; project administration, M.A., C.G., S.W., J.F.F., L.M.B. and J.X.; funding acquisition, L.M.B. and J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by NIMH grants R01 MH-070366 and RC1 MH-088288 to LMB; New Jersey Governor’s Council for Medical Research and Treatment of Autism grant CAUT12APS006 to LMB and CAUT19APL028 to J.X. and L.M.B.

Institutional Review Board Statement: The study is approved by the Institutional Review Board at Rutgers, The State University of New Jersey (IRB number: 13-112Mc).

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw sequencing reads, variants, and genotypes for all samples are available in the National Institute of Mental Health Data Archive (NDA) under experiments C1932 and C2933.

Acknowledgments: We thank the patients who participated in and contributed to this study. We gratefully acknowledge access to the HPC facilities and support of the computational STEM and bioinformatics scientists from the Office of Advanced Research Computing at Rutgers University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Washington, DC, USA, 2013.
2. Shaw, K.A.; Maenner, M.J.; Bakian, A.V.; Bilder, D.A.; Durkin, M.S.; Furnier, S.M.; Hughes, M.M.; Patrick, M.; Pierce, K.; Salinas, A.; et al. Early Identification of Autism Spectrum Disorder Among Children Aged 4 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR Surveill. Summ.* **2021**, *70*, 1–14. [[CrossRef](#)] [[PubMed](#)]
3. Ozonoff, S.; Young, G.S.; Carter, A.; Messinger, D.; Yirmiya, N.; Zwaigenbaum, L.; Bryson, S.; Carver, L.J.; Constantino, J.N.; Dobkins, K.; et al. Recurrence risk for autism spectrum disorders: A Baby Siblings Research Consortium study. *Pediatrics* **2011**, *128*, e488–e495. [[CrossRef](#)]
4. Devlin, B.; Scherer, S.W. Genetic architecture in autism spectrum disorder. *Curr. Opin. Genet. Dev.* **2012**, *22*, 229–237. [[CrossRef](#)]
5. Gaugler, T.; Klei, L.; Sanders, S.J.; Bodea, C.A.; Goldberg, A.P.; Lee, A.B.; Mahajan, M.; Manaa, D.; Pawitan, Y.; Reichert, J.; et al. Most genetic risk for autism resides with common variation. *Nat. Genet.* **2014**, *46*, 881–885. [[CrossRef](#)] [[PubMed](#)]
6. Kosik, K.S. The neuronal microRNA system. *Nat. Rev. Neurosci.* **2006**, *7*, 911–920. [[CrossRef](#)] [[PubMed](#)]
7. Lyu, J.W.; Yuan, B.; Cheng, T.L.; Qiu, Z.L.; Zhou, W.H. Reciprocal regulation of autism-related genes MeCP2 and PTEN via microRNAs. *Sci. Rep.* **2016**, *6*, 20392. [[CrossRef](#)]
8. Vatsa, N.; Kumar, V.; Singh, B.K.; Kumar, S.S.; Sharma, A.; Jana, N.R. Down-Regulation of miRNA-708 Promotes Aberrant Calcium Signaling by Targeting Neuronatin in a Mouse Model of Angelman Syndrome. *Front. Mol. Neurosci.* **2019**, *12*, 35. [[CrossRef](#)]
9. Williams, S.M.; An, J.Y.; Edson, J.; Watts, M.; Murigneux, V.; Whitehouse, A.J.O.; Jackson, C.J.; Bellgrove, M.A.; Cristino, A.S.; Claudianos, C. An integrative analysis of non-coding regulatory DNA variations associated with autism spectrum disorder. *Mol. Psychiatry* **2019**, *24*, 1707–1719. [[CrossRef](#)]
10. Toma, C.; Torrico, B.; Hervas, A.; Salgado, M.; Rueda, I.; Valdes-Mas, R.; Buitelaar, J.K.; Rommelse, N.; Franke, B.; Freitag, C.; et al. Common and rare variants of microRNA genes in autism spectrum disorders. *World J. Biol. Psychiatry* **2015**, *16*, 376–386. [[CrossRef](#)]
11. Bartlett, C.W.; Hou, L.; Flax, J.F.; Hare, A.; Cheong, S.Y.; Fermano, Z.; Zimmerman-Bier, B.; Cartwright, C.; Azaro, M.A.; Buyske, S.; et al. A genome scan for loci shared by autism spectrum disorder and language impairment. *Am. J. Psychiatry* **2014**, *171*, 72–81. [[CrossRef](#)]
12. Zhou, A.; Cao, X.; Mahaganapathy, V.; Azaro, M.; Gwin, C.; Wilson, S.; Buyske, S.; Bartlett, C.W.; Flax, J.F.; Brzustowicz, L.M.; et al. Analysis of Common Genetic Risk Factors in ASD and ADHD Co-occurring Families. *medRxiv* **2022**. [[CrossRef](#)]
13. miRBase v22. Available online: <https://www.mirbase.org/ftp.shtml> (accessed on 1 October 2020).
14. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. miRBase: From microRNA sequences to function. *Nucleic Acids Res.* **2019**, *47*, D155–D162. [[CrossRef](#)] [[PubMed](#)]
15. Oak, N.; Ghosh, R.; Huang, K.L.; Wheeler, D.A.; Ding, L.; Plon, S.E. Framework for microRNA variant annotation and prioritization using human population and disease datasets. *Hum. Mutat.* **2019**, *40*, 73–89. [[CrossRef](#)] [[PubMed](#)]
16. ADmiRE. Available online: <https://github.com/nroak/ADmiRE> (accessed on 1 October 2020).
17. TargetScanHuman Release 7.2. Available online: http://www.targetscan.org/vert_72/ (accessed on 1 October 2020).
18. Agarwal, V.; Bell, G.W.; Nam, J.W.; Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **2015**, *4*, e05005. [[CrossRef](#)] [[PubMed](#)]

19. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105. [[CrossRef](#)]
20. GEMINI v0.20.1. Available online: <https://gemini.readthedocs.io/en/latest/> (accessed on 14 December 2020).
21. Paila, U.; Chapman, B.A.; Kirchner, R.; Quinlan, A.R. GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. *PLoS Comput. Biol.* **2013**, *9*, e1003153. [[CrossRef](#)]
22. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)]
23. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **2020**, *369*, 1318–1330. [[CrossRef](#)]
24. Miller, J.A.; Ding, S.L.; Sunkin, S.M.; Smith, K.A.; Ng, L.; Szafer, A.; Ebbert, A.; Riley, Z.L.; Royall, J.J.; Aiona, K.; et al. Transcriptional landscape of the prenatal human brain. *Nature* **2014**, *508*, 199–206. [[CrossRef](#)]
25. Lindsay, S.J.; Xu, Y.; Lisgo, S.N.; Harkin, L.F.; Copp, A.J.; Gerrelli, D.; Clowry, G.J.; Talbot, A.; Keogh, M.J.; Coxhead, J.; et al. HDBR Expression: A Unique Resource for Global and Individual Gene Expression Studies during Early Human Brain Development. *Front. Neuroanat.* **2016**, *10*, 86. [[CrossRef](#)]
26. Cao, X.; Zhang, Y.; Abdulkadir, M.; Deng, L.; Fernandez, T.V.; Garcia-Delgar, B.; Hagstrom, J.; Hoekstra, P.J.; King, R.A.; Koesterich, J.; et al. Whole-exome sequencing identifies genes associated with Tourette’s disorder in multiplex families. *Mol. Psychiatry* **2021**, *26*, 6937–6951. [[CrossRef](#)] [[PubMed](#)]
27. Herwig, R.; Hardt, C.; Lienhard, M.; Kamburov, A. Analyzing and interpreting genome data at the network level with Consensus-PathDB. *Nat. Protoc.* **2016**, *11*, 1889–1907. [[CrossRef](#)] [[PubMed](#)]
28. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **2021**, *49*, D605–D612. [[CrossRef](#)] [[PubMed](#)]
29. Wong, A.K.; Krishnan, A.; Troyanskaya, O.G. GIANT 2.0: Genome-scale integrated analysis of gene networks in tissues. *Nucleic Acids Res.* **2018**, *46*, W65–W70. [[CrossRef](#)] [[PubMed](#)]
30. Sun, S.; Miller, M.; Wang, Y.; Tyc, K.M.; Cao, X.; Scott, R.T.; Tao, X., Jr.; Bromberg, Y.; Schindler, K.; Xing, J. Predicting embryonic aneuploidy rate in IVF patients using whole-exome sequencing. *Hum. Genet.* **2022**. [[CrossRef](#)]
31. Jung, J.U.; Ko, K.; Lee, D.H.; Ko, K.; Chang, K.T.; Choo, Y.K. The roles of glycosphingolipids in the proliferation and neural differentiation of mouse embryonic stem cells. *Exp. Mol. Med.* **2009**, *41*, 935–945. [[CrossRef](#)]
32. Jennemann, R.; Sandhoff, R.; Wang, S.; Kiss, E.; Gretz, N.; Zuliani, C.; Martin-Villalba, A.; Jager, R.; Schorle, H.; Kenzelmann, M.; et al. Cell-specific deletion of glucosylceramide synthase in brain leads to severe neural defects after birth. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12459–12464. [[CrossRef](#)]
33. Goetz, J.J.; Martin, G.M.; Chowdhury, R.; Trimarchi, J.M. Onecut1 and Onecut2 play critical roles in the development of the mouse retina. *PLoS ONE* **2014**, *9*, e110194. [[CrossRef](#)]
34. Gu, X.; Li, A.; Liu, S.; Lin, L.; Xu, S.; Zhang, P.; Li, S.; Li, X.; Tian, B.; Zhu, X.; et al. MicroRNA124 Regulated Neurite Elongation by Targeting OSBP. *Mol. Neurobiol.* **2016**, *53*, 6388–6396. [[CrossRef](#)]
35. Ferdinandusse, S.; Kostopoulos, P.; Denis, S.; Rusch, H.; Overmars, H.; Dillmann, U.; Reith, W.; Haas, D.; Wanders, R.J.; Duran, M.; et al. Mutations in the gene encoding peroxisomal sterol carrier protein X (SCPx) cause leukoencephalopathy with dystonia and motor neuropathy. *Am. J. Hum. Genet.* **2006**, *78*, 1046–1052. [[CrossRef](#)]
36. Borman, A.D.; Pearce, L.R.; Mackay, D.S.; Nagel-Wolfrum, K.; Davidson, A.E.; Henderson, R.; Garg, S.; Waseem, N.H.; Webster, A.R.; Plagnol, V.; et al. A homozygous mutation in the TUB gene associated with retinal dystrophy and obesity. *Hum. Mutat.* **2014**, *35*, 289–293. [[CrossRef](#)] [[PubMed](#)]
37. Ikeda, A.; Nishina, P.M.; Naggert, J.K. The tubby-like proteins, a family with roles in neuronal development and function. *J. Cell Sci.* **2002**, *115*, 9–14. [[CrossRef](#)] [[PubMed](#)]
38. Grifone, R.; Shao, M.; Saquet, A.; Shi, D.L. RNA-Binding Protein Rbm24 as a Multifaceted Post-Transcriptional Regulator of Embryonic Lineage Differentiation and Cellular Homeostasis. *Cells* **2020**, *9*, 1891. [[CrossRef](#)] [[PubMed](#)]
39. Breitenkamp, A.F.; Matthes, J.; Herzig, S. Voltage-gated Calcium Channels and Autism Spectrum Disorders. *Curr. Mol. Pharmacol.* **2015**, *8*, 123–132. [[CrossRef](#)] [[PubMed](#)]
40. Park, S.M.; Park, H.R.; Lee, J.H. MAPK3 at the Autism-Linked Human 16p11.2 Locus Influences Precise Synaptic Target Selection at Drosophila Larval Neuromuscular Junctions. *Mol. Cells* **2017**, *40*, 151–161. [[CrossRef](#)]
41. Busch, R.M.; Srivastava, S.; Hogue, O.; Frazier, T.W.; Klaas, P.; Hardan, A.; Martinez-Agosto, J.A.; Sahin, M.; Eng, C.; Developmental Synaptopathies Consortium. Neurobehavioral phenotype of autism spectrum disorder associated with germline heterozygous mutations in PTEN. *Transl. Psychiatry* **2019**, *9*, 253. [[CrossRef](#)]
42. Sekino, Y.; Kojima, N.; Shirao, T. Role of actin cytoskeleton in dendritic spine morphogenesis. *Neurochem. Int.* **2007**, *51*, 92–104. [[CrossRef](#)]
43. Griesi-Oliveira, K.; Suzuki, A.M.; Alves, A.Y.; Mafra, A.; Yamamoto, G.L.; Ezquina, S.; Magalhaes, Y.T.; Forti, F.L.; Sertie, A.L.; Zachi, E.C.; et al. Actin cytoskeleton dynamics in stem cells from autistic individuals. *Sci. Rep.* **2018**, *8*, 11138. [[CrossRef](#)]
44. Yun, K.; Potter, S.; Rubenstein, J.L. Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. *Development* **2001**, *128*, 193–205. [[CrossRef](#)]

45. Domene, S.; Roessler, E.; El-Jaick, K.B.; Snir, M.; Brown, J.L.; Velez, J.I.; Bale, S.; Lacbawan, F.; Muenke, M.; Feldman, B. Mutations in the human *SIX3* gene in holoprosencephaly are loss of function. *Hum. Mol. Genet.* **2008**, *17*, 3919–3928. [[CrossRef](#)]
46. Goulburn, A.L.; Alden, D.; Davis, R.P.; Micallef, S.J.; Ng, E.S.; Yu, Q.C.; Lim, S.M.; Soh, C.L.; Elliott, D.A.; Hatzistavrou, T.; et al. A targeted *NKX2.1* human embryonic stem cell reporter line enables identification of human basal forebrain derivatives. *Stem Cells* **2011**, *29*, 462–473. [[CrossRef](#)] [[PubMed](#)]
47. Garcia-Forn, M.; Boitnott, A.; Akpinar, Z.; De Rubeis, S. Linking Autism Risk Genes to Disruption of Cortical Development. *Cells* **2020**, *9*, 2500. [[CrossRef](#)] [[PubMed](#)]
48. Courchesne, E.; Mouton, P.R.; Calhoun, M.E.; Semendeferi, K.; Ahrens-Barbeau, C.; Hallet, M.J.; Barnes, C.C.; Pierce, K. Neuron number and size in prefrontal cortex of children with autism. *JAMA* **2011**, *306*, 2001–2010. [[CrossRef](#)] [[PubMed](#)]
49. Butler, M.G.; Dasouki, M.J.; Zhou, X.P.; Talebizadeh, Z.; Brown, M.; Takahashi, T.N.; Miles, J.H.; Wang, C.H.; Stratton, R.; Pilarski, R.; et al. Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline *PTEN* tumour suppressor gene mutations. *J. Med. Genet.* **2005**, *42*, 318–321. [[CrossRef](#)]
50. Orrico, A.; Galli, L.; Buoni, S.; Orsi, A.; Vonella, G.; Sorrentino, V. Novel *PTEN* mutations in neurodevelopmental disorders and macrocephaly. *Clin. Genet.* **2009**, *75*, 195–198. [[CrossRef](#)]
51. Varga, E.A.; Pastore, M.; Prior, T.; Herman, G.E.; McBride, K.L. The prevalence of *PTEN* mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genet. Med.* **2009**, *11*, 111–117. [[CrossRef](#)]
52. Jiang, L.; Guo, F.; Tang, J.; Yu, H.; Ness, S.; Duan, M.; Mao, P.; Zhao, Y.Y.; Guo, Y. SBSA: An online service for somatic binding sequence annotation. *Nucleic Acids Res.* **2022**, *50*, e4. [[CrossRef](#)]
53. Maier, R.M.; Visscher, P.M.; Robinson, M.R.; Wray, N.R. Embracing polygenicity: A review of methods and tools for psychiatric genetics research. *Psychol. Med.* **2018**, *48*, 1055–1067. [[CrossRef](#)]
54. Giangrande, E.J.; Weber, R.S.; Turkheimer, E. What Do We Know About the Genetic Architecture of Psychopathology? *Annu. Rev. Clin. Psychol.* **2022**, *18*, 19–42. [[CrossRef](#)]
55. Bruni, M.; Flax, J.F.; Buyske, S.; Shindhelm, A.D.; Witton, C.; Brzustowicz, L.M.; Bartlett, C.W. Behavioral and Molecular Genetics of Reading-Related AM and FM Detection Thresholds. *Behav. Genet.* **2017**, *47*, 193–201. [[CrossRef](#)]
56. Flax, J.; Gwin, C.; Wilson, S.; Fradkin, Y.; Buyske, S.; Brzustowicz, L. Social (Pragmatic) Communication Disorder: Another name for the Broad Autism Phenotype? *Autism* **2019**, *23*, 1982–1992. [[CrossRef](#)] [[PubMed](#)]
57. Hare-Harris, A.E.; Mitchel, M.W.; Myers, S.M.; Mitchel, A.D.; King, B.R.; Ruocco, B.G.; Martin, C.L.; Flax, J.F.; Brzustowicz, L.M. Within-task variability on standardized language tests predicts autism spectrum disorder: A pilot study of the Response Dispersion Index. *J. Neurodev. Disord.* **2019**, *11*, 21. [[CrossRef](#)] [[PubMed](#)]
58. Bartlett, C.W.; Flax, J.F.; Fermano, Z.; Hare, A.; Hou, L.; Petrill, S.A.; Buyske, S.; Brzustowicz, L.M. Gene × gene interaction in shared etiology of autism and specific language impairment. *Biol. Psychiatry* **2012**, *72*, 692–699. [[CrossRef](#)] [[PubMed](#)]
59. Constantino, J.N. Deconstructing autism: From unitary syndrome to contributory developmental endophenotypes. *Int. Rev. Psychiatry* **2018**, *30*, 18–24. [[CrossRef](#)]
60. Erbeli, F.; Rice, M.; Paracchini, S. Insights into Dyslexia Genetics Research from the Last Two Decades. *Brain Sci.* **2021**, *12*, 27. [[CrossRef](#)]