Article

Neuregulin 2 Is a Candidate Gene for Autism Spectrum Disorder

Wei-Hsien Chien *, Chia-Hsiang Chen, Min-Chih Cheng, Yu-Yu Wu, Susan Shurfen Gau *

Posted Date: 28 March 2024

doi: 10.20944/preprints202403.1755.v1

Keywords: autism spectrum disorders; neuregulin 2 (NRG2); insertion/deletion polymorphism (indel)

Article

Neuregulin 2 Is a Candidate Gene for Autism Spectrum Disorder

Wei-Hsien Chien [1,*], Chia-Hsiang Chen [2], Min-Chih Cheng [3] and Yu-Yu Wu [2], Susan Shur-Fen Gau [4,5,*]

[1] Department of Occupational Therapy, College of Medicine, Fu Jen Catholic University, No.510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 242062, Taiwan (R.O.C.)
[2] Department of Psychiatry, Chang Gung Memorial Hospital- Linkou Medical Center, Tao-Yuan, Taiwan
[3] Department of Psychiatry, Yuli Branch, Taipei Veterans General Hospital, Yuli, Taiwan
[4] Department of Psychiatry, National Taiwan University Hospital and College of Medicine, No. 7, Chung-Shan South Road, Taipei 10002, Taiwan
[5] Graduate Institute of Brain and Mind Sciences and Graduate of Clinical Medicine, National Taiwan University, Taipei, Taiwan
* Correspondence: wchien58@gmail.com or 082394@mail.fju.edu.tw (W.-H.C.); gaushufe@ntu.edu.tw (S.S.-F.G.); Tel.: +886-2-2905-2031 (W.-H.C.); +886-2-23123456 ext. 66802 (S.S.-F.G.); Fax: +886-2-2904-6743 (W.-H.C.); +886-2-23812408 (S.S.-F.G.)

Abstract: Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder with heterogeneous and complex genetic underpinnings. Our previous microarray gene-expressed profiling identified significantly different neuregulin-2 (NRG2) gene expressions between ASD patients and controls. Thus, we aimed to clarify whether NRG2 is a candidate gene associated with ASD. The study consisted of two stages. First, we used real-time quantitative PCR in 20 ASDs and 20 controls to confirm the microarray gene expression profiling results. The average NRG2 gene expression level in patients with ASD (3.23±2.80) was significantly lower than that of controls (9.27±4.78, $p < 0.001$). Next, we resequenced all exons of NRG2 in other 349 ASD subjects, who were compared with super-controls from the Taiwan Biobank using case-control association analysis to determine the existence of mutations or single-nucleotide polymorphisms (SNPs). We identified three SNPs, including two noncoding ones, IVS3 + 13A > G (rs889022) and IVS9 + 32T > A (rs182642591), and one insertion/deletion (indel) SNP, named indel GCCCGGC, rs1323957797, in the NRG2 coding region. The genotype and allele frequency distribution of rs1323957797 significantly differed between ASDs and controls ($p < 0.0001$). Our findings showed significant differences in genotype and allele frequency distribution of the indel SNP, rs1323957797, located in the EGF-like domain region at the C-terminal of the NRG2 precursor protein. Together with our findings, the indel SNP mutation was significantly more common in ASDs than controls, leading to lower NRG2 expression than controls, influencing NRG2 function. Our study's findings suggest that NRG2 is a susceptibility gene for ASD.

Keywords: autism spectrum disorders; neuregulin 2 (NRG2); insertion/deletion polymorphism (indel)

ClinicalTrials.gov number: NCT00494754; NCT02228876; C105098

1. Introduction

According to the Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5), autism spectrum disorder (ASD) is defined as a childhood-onset neurodevelopmental disorder with two essential features: persistent deficits in social communication and interactions, and restricted, repetitive patterns of behavior, interests, or activities [1]. Although the prevalence of ASD in Taiwan was estimated to be 1% in 2017 [2], the figure has markedly increased in the past decade. Recent data

 doi:10.20944/preprints202403.1755.v1

showed that around 1 in 36 persons (2.8%) aged 8 years have been diagnosed with ASD come from 11 communities in the Autism and Developmental Disabilities Monitoring Network in the United States published in 2023 CDC's Morbidity and Mortality Weekly Report [3], and that males are more frequently affected than females, with a male-to-female ratio of approximately 4.3:1 [4,5]. The heritability of ASD is estimated to be >90%, suggesting a strong genetic component [6]. Recently, large-scale gene discovery efforts have shown that ASD is not a simple Mendelian disorder [7–9]. Conversely, the genetic underpinnings of ASD are heterogeneous and complex, involving multiple genes, gene–gene and gene–environment interactions [10]. The genetic risk factors for ASD identified so far range from common variants conferring a small clinical effect to rare mutations with a high clinical outcome [11]. Recently, the application of large-scale exome sequencing has led to the identification of hundreds of novel ASD-associated genes enriched in common genetic signaling pathways such as synaptic development, plasticity, and signaling [12]. The high heterogeneity of ASD may account for the varied clinical presentations of patients with ASD. Despite these advances, there are likely more ASD-associated genes to be discovered.

Microarray-based gene expression profiling allows simultaneous measurement of hundreds to thousands of gene transcripts, useful for large-scale gene discovery [13]. Comparative gene expression profiling analysis of lymphoblastoid cell lines (LCL) identified shared pathways among different forms of autism [14]. Similarly, altered pathways in neural development and steroid biogenesis were detected in a total gene expression profiling analysis of LCL in autistic patients and their unaffected siblings [15]. Previously, our group conducted comparative gene expression profiling of LCL in ASD patients and control subjects to identify differentially expressed genes associated with ASD [16]. 187 differentially expressed probe sets, including 131 transcripts (50 up-regulated and 81 down-regulated), were detected between cases and controls. We verified one of the selected differentially expressed genes, FOXP1, using real-time quantitative polymerase chain reaction (RT-qPCR) in a sample of ASD patients and control subjects. The average expression level of FOXP1 in ASD was significantly higher than that of the controls [16].

Neuregulin 2 (NRG2) emerged as a notable differentially expressed gene during the comparative total gene expression profiling analysis in our previous study [16]. NRG2 belongs to the NRG gene family, including NRG1–NRG4. NRGs serve as cell–cell ligands for receptor tyrosine kinases belonging to the ErbB family, which comprises four homologous type I receptor tyrosine kinases known as EGFR (epidermal growth factor receptor), ErbB1, ErbB2, ErbB3, and ErbB4 [17,18]. The neuregulin–ErbB signaling network is involved in several processes in both the developing and adult brain [19]. NRGs specifically play a role in synaptic plasticity, promoting neuronal migration and differentiation, regulating the expression of neurotransmitter receptors, and influencing glial proliferation, survival, and differentiation [19,20]. Through previous genetic and functional analytic studies, several variants of NRG1 and its neuronal receptor ErbB4 were found to be associated with schizophrenia and its endophenotypes [21–23]. Moreover, studies in human induced pluripotent stem cells from affected subjects demonstrated that the NRG–ErbB pathway is related to psychiatric disorders [24–26].

NRG2, encoding a novel member of the NRG family, induces the growth and differentiation of epithelial, neuronal, glial, and other cell types through interaction with ErbB receptors while sharing a similar genomic structure with NRG1 [27]. Like NRG1, NRG2 can bind ErbB4 and ErbB3 but has no active kinase domain [21]. NRG2 is expressed in the embryonic heart, lung, bladder and the developing nervous system [28,29] but confined to the cerebellum (granule cells and Purkinje cells), dentate gyrus (granule cells), and olfactory bulb (granule cells) in the adult brain [28–30]. Recently, Vullhorst et al. reported that NRG2, but not NRG1, is a major functional ErbB4 ligand in the postnatal brain controlling N-methyl-D-aspartate (NMDA) receptor function in cortical interneurons associated with cognitive deficits in psychiatric disorders [31]. Recently, Yan et al. found higher dopamine levels in the dorsal striatum but lower levels in the medial prefrontal cortex (mPFC) in NRG2 knockout mice (KO). The pattern of dopamine expression was similar to schizophrenia, and animals showed behavioral abnormalities relevant to psychiatric disorders, such as impaired social behavior and cognitive function [32]. Further, a genome-wide association analysis of disturbances in

 doi:10.20944/preprints202403.1755.v1

the electroencephalography early gamma-frequency band between schizophrenia and control subjects found significant differences in several markers of the NRG2 and KALRN genes involved in neuronal development and the NRG–ErbB signaling pathway [33]. Furthermore, several studies have shown that NRG1, NRG2, or genes of the NRG–ErbB network contribute to the etiology of psychiatric disorders [34–36]. Nevertheless, the function of NRG2 is not well known, and its association with complex neurodevelopmental disorders like ASD remains unclear.

NRG2 (Gene ID 9542) is located at chromosome 5q31.2, an ASD candidate gene region, as shown by genetic linkage studies [37,38]. In addition, several case reports indicated that cases with ASD have de novo translocations of chromosome 5q and chromosomes 1q, 4q, and 18q [39–41]. Furthermore, Phillippi et al. reported a strong and consistent association between two Single-Nucleotide Polymorphisms (SNPs) within the paired-like homeodomain transcription factor 1 (PITX1) on chromosome 5q31 and autism [42].

Based on the above, we suggest that NRG2 might be a susceptibility gene for ASD. To test this hypothesis, we first used RT-qPCR to assess NRG2 expression in a sample of 20 patients with ASD and 20 controls to confirm the microarray gene expression analysis results. Next, we conducted systemic sequencing of all exons of NRG2 in another independent sample of ASD patients. We compared them with super-controls of the Taiwan Biobank using a case-control association study to determine the existence of mutations or SNPs in NRG2 associated with ASD.

2. Results

2.1. Differential Expression of NRG2 in ASD Patients and Controls

The differential NRG2 expression was examined first between 20 patients with ASD and 20 healthy controls using RT-qPCR. The average NRG2 expression in the 20 ASD cases (3.23 ± 2.80) was significantly lower than that of control subjects (9.27 ± 4.78, $p < 0.001$). The distribution and mean NRG2 transcript levels of LCL in 20 ASD patients and 20 controls are shown in Figure 1.

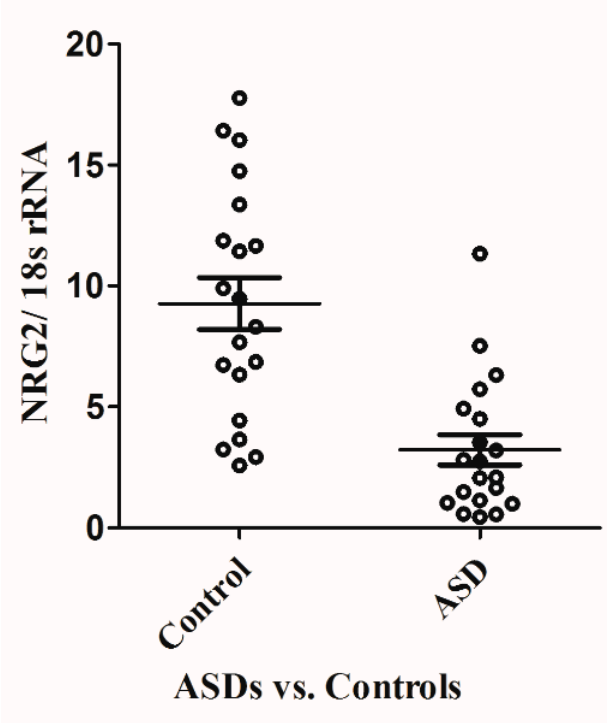

Figure 1. Scatter plot of the NRG2 mRNA level as normalized by 18S rRNA in 20 patients with Autism spectrum disorder (ASD) and 20 control subjects. Horizontal line indicates the mean of NRG2 transcript. The average expression level of NRG2 gene transcript in patients with ASD (3.23 ± 2.80) was significantly lower than that of control subjects (9.27 ± 4.78), $p < 0.001$.

 doi:10.20944/preprints202403.1755.v1

2.2. Three Single-Nucleotide Polymorphisms Were Identified in NRG2

We screened 349 patients by sequencing all exons and splicing sites of the NRG2 gene. No missense or disease-causing mutations were identified in NRG2. During resequencing of the E9 amplicon, two samples failed among the total of 301 ASD subjects, excluding the 48 cases whose DNA had been used off; in resequencing of the E10 amplicon, 36 samples failed among 300 ASD cases, excluding the 49 cases without DNA; resequencing of E3 amplicon, 22 samples failed in 231 ASD excluding 66 samples without DNA and 30 cases whose DNA had been run out, similarly. Three single-nucleotide polymorphisms (dbSNPs) were identified in our subjects. The genotype results of the three SNPs for all ASD cases are in the supplementary table. All were two noncoding SNPs, IVS3 + 13A > G (rs889022) and IVS9 + 32T > A (rs182642591), and one indel SNP, rs1323957797, in the NRG2 cording region. The locations of these SNPs are listed in Figure 2.

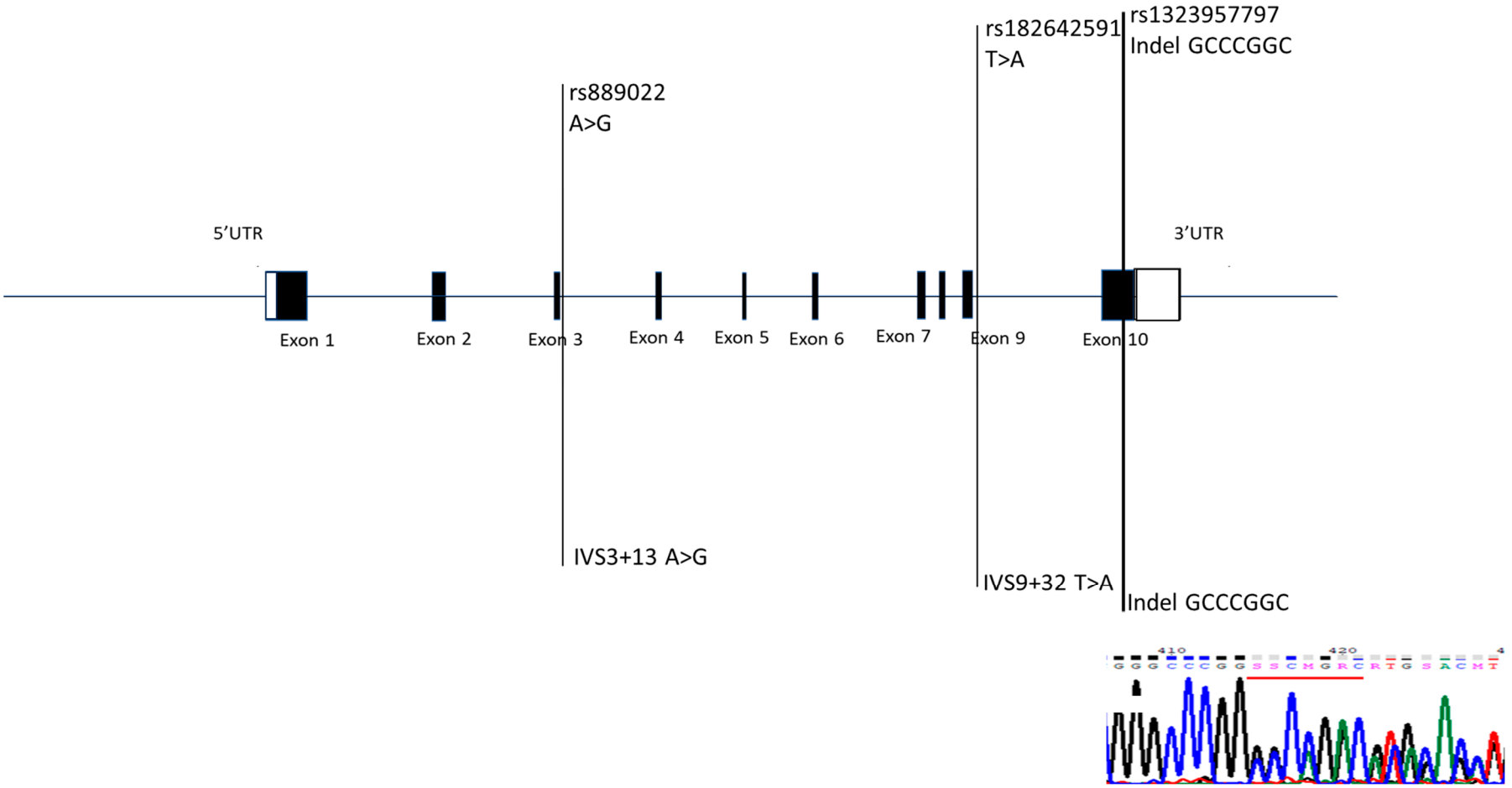

Figure 2. Schematic genomic structure of NRG2, and locations of the variants identified in this study. The genomic structure of human NRG2 gene were available from the NCBI (https://www.ncbi.nlm.nih.gov/). NRG2 (Gene ID: 9542) comprised ten exons, the genomic size is 196,519 bp on chromosome 5q31.2. (RefSeq: NM_004883.3). dbSNPs of NRG2 amplicons were available from NCBI (https://www.ncbi.nlm.nih.gov/) in the upper region. The three SNPs, IVS3+13, IVS9+32, and Indel GCCCGC, we identified were in the lower region. The locations of these SNPs are listed in figure.

doi:10.20944/preprints202403.1755.v1

 doi:10.20944/preprints202403.1755.v1

2.3. Differences in Allele and Genotype Frequencies between Super-Controls and ASD Patients

All mutations and variants identified were compared with super-controls from the Taiwan Biobank (https://taiwanview.twbiobank.org.tw/index). The genotype and allele frequencies of ASD patients and controls from Taiwan Biobank are illustrated in Table 3. The rs889022 SNP was found in 14.7% (34/231) of ASD patients and 20.2% (302/1495) of control subjects ($p = 0.08$), and the rs182642591 was found in 0.3% (1/ 299) of ASD patients and 1.5% (22/1495) of controls ($p = 0.11$). There were no significant differences in genotype or allele frequency distributions of rs889022 and rs182642591 between ASD patients and controls. The genotype distributions of these two SNPs did not deviate from Hardy–Weinberg equilibrium in ASD patients. However, the insertion/ deletion SNP (indel GCCCGGC) was also dbSNP, named rs1323957797, located in Pos.139848469 (Ref. Version: GRCh38) of the NRG2 coding region. The genotype and allele frequency distribution of the indel SNP was significantly differed between ASD patients and controls. Genotype C/IndelGCCCGGC + indelGCCCGGC/ indelGCCCGGC) was found in 11.7% (31/ 264) of ASD subjects and 0.2% (3/1495) of controls ($p < 0.0001$). Similarly, the allele frequency distribution was significantly different between ASD cases and controls (5.9% v.s. 0.1%, $p < 0.0001$) (Table 3).

2.4. Functional Prediction of the Indel GCCCGGC Variant of NRG2

The indel GCCCGGC variant at position 661–663 aa of the amino acid sequence of NRG2 named rs1323957797 was detected in 31 ASD patients. All patients carrying this indel variant were heterozygotes. Similarly, all controls in the TaiwanBiobank were heterozygotes. The functional prediction of this indel GCCCGGC variant was performed using PolyPhen-2, SIFT, and UniProt (Table 4). Based on SIFT or UniProt, this variant might have a neutral or likely benign effect, whereas the impact in PolyPhen-2 analysis was unknown.

2.5. Power Analysis

A post-hoc power analysis showed that with a total sample size of 299 participants, excluding about 50 cases without DNA, we had a power of 0.32 to detect a small effect (0.1) and 0.998 to see a medium effect (0.3) of the genotype distributions of IVS3 + 13, IVS9 + 32, and the indelGCCCGGC variant in exon 10 at an alpha level of 0.05. To analyze allele frequencies of these three variants, we had a power of 0.32 to detect a small effect (0.1) and 0.998 to see a medium effect (0.3) at an alpha level of 0.05.

3. Discussion

In our study, we found a significantly lower average NRG2 expression in ASD patients (3.23 ± 2.80) than in controls (9.27 ± 4.78, $p < 0.001$; Figure 1). Previously, several linkage analysis or association studies pointed to a genome region encompassing the NRG2 locus to be associated with psychiatric disorders [43–45]. In addition, many studies have pointed out that 5q containing the NRG2 gene may also be a candidate region for ASD [37–42]. Recently, Yan L. et al. reported that NRG2-KO mice presented a pattern of dopamine that was similar to schizophrenia and also showed various behavioral abnormalities relevant to psychiatric disorders, such as impaired social behavior and cognitive function in several behavioral tests [32]. Our findings align with prior research indicating that NRG2 could be a susceptibility gene for ASD, and the reduced NRG2 expression or its loss may be associated with the development of ASD.Furthermore, the indel GCCCGGC variant corresponded to amino acids 661–663 of NRG2, which are part of the EGF-like domain region at the C-terminal of the NRG2 precursor protein as predicted in silico (UniProt). Previous studies have shown that the EGF-like domain regulates the biological activity of NRG2 [19,28], as it is the region that binds to receptors and can regulate processes such as cell proliferation, differentiation, and survival. The function of the protein C-terminal domain includes regulation of protein stability or function or interactions with other proteins, such as binding domains or regions mediating protein–protein interactions. Recently, Czarnek and Bereta reported the proteolytic degradation mechanism

 doi:10.20944/preprints202403.1755.v1

of NRG2 protein; the C-terminal part contains a site for γ-secretase. They are being cleaved by ADAM10 or BACE2 for degradation. The interaction domains in the C-terminal region of NRG2 lie at 712–713 amino acids. This domain interacts with other neurotrophic factors; these interactions might help regulate physiological and pathological processes such as neural development and neurodegeneration [46].

Our study found that the indel GCCCGGC variant at 661–663 amino acids also in the C-terminal region of NRG2 was significantly more common in ASD patients than in controls. Consequently, we suggested the indel SNP might affect the function of NRG2 in the NRG–ErbB signaling pathway and be associated with ASD.

This study still has some limitations. First, we identified two intronic SNPs (rs889022 and rs182642591) in nearly 300 ASD Taiwanese patients, but we did not find a significant association with ASD in our autistic samples. Population-based case-control association studies of complex conditions, like ASD with clinical and genetic heterogeneity, need a large sample size to identify risk genes with small effects. In this study, we recruited 349 ASD subjects excluding about 50 cases without DNA with limited power (0.32) to detect small effects in our post-hoc analysis. Therefore, further research should involve a larger patient population with this disorder. Second, despite the significantly lower NRG2 expression level in ASD patients than in 20 controls, the regulatory mechanism underlying its potential influence on ASD remains unknown. We could not address causality in this study. Future studies should focus on the regulatory mechanism governing the effects of lower NRG2 expression on ASD. Finally, the indel GCCCGGC variant located in the EGF-like domain at the C-terminal region of NRG2 was significantly higher than the control group ($p < 0.0001$). However, how this indel SNP mediates protein–protein interaction in the NRG–ErbB pathway remains unknown. Consequently, functional analysis of NRG2 and the influence of this indel SNP on its function might help us understand the physiological and pathological mechanism underlying its contribution to ASD.

4. Materials and Methods

4.1. Subjuects

The study was conducted in two stages. First, an RT-qPCR analysis in ASD cases and controls. NRG2 transcription levels were compared between 20 patients with ASD and 20 healthy controls. The 20 autistic patients were all males aged 9.1 ± 3.2 years (4–18), and the 20 male controls were 37.15 ± 16.90 years (18–67) recruited among the cohort of ASD patients receiving regular medical checkups in a local medical center. Second, 349 additional ASD patients were recruited for exonic resequencing analysis, including 308 male patients (15.26 ± 5.23 years old) and 41 female patients (14.88 ± 4.90 years old). All the patients with ASD were recruited from the Children Mental Health Center, Department of Psychiatry, Taiwan University Hospital (NTUH), Taipei, Taiwan, and the Department of Child Psychiatry, Chang Gung Memory Hospital (CGMH), Kwei-Shan, Taiwan. The clinical diagnosis of ASD made by qualified child psychiatrists was confirmed by interviewing the caregivers using the Chinese version of the Autism Diagnostic Interview-Revised [47,48] and patients using the Chinese version of the Autism Diagnostic Observation Scale [49]. Patients with known chromosomal abnormalities or associated medical conditions were excluded from the study. Finally, the super-controls from the Han Chinese Cell and Genome Bank established by the Institute of Biomedical Sciences, Academia Sinica in Taiwan [50] were used for case-control association analysis. The study protocol was approved by the Research Ethics Committee of NTUH, CGMH, and Fu Jen Catholic University; written informed consent was obtained from the participants and their parents after the procedures were fully explained and before study implementation, including blood sample collection for subsequent laboratory investigations.

 doi:10.20944/preprints202403.1755.v1

4.2. Preparation of the Lymphoblastoid Cell Line and cDNA

Immortalized LCLs were established from each subject by transforming lymphocytes with Epstein-Barr virus following the procedures described elsewhere [51]. The cDNA of 20 ASD participants and 20 controls was prepared using SuperScriptTM II RNase H- Reverse Transcriptase (Invitrogen Life Technologies). The details have been described in our previous report [51].

4.3. Real-Time Quantitative PCR

RT-qPCR was performed using the SYBR Green method and implemented in the StepOne Plus Real-Time PCR System according to the manufacturer's protocol (Applied Biosystems, Forster City, CA, USA). Detailed procedures can be found in our previous report [51]. The relative standard curve method was used to quantify mRNA expression (User Bulletin #2 ABI PRISM 7700 sequence detection system). This method used serial dilutions of known RNA amounts from a reference sample (pooled from 40 LCLs of male controls) to generate an external standard curve. The relative amount was calculated using linear regression analysis from their respective common angles for each unknown piece. The mRNA expression level of NRG2 was normalized by the 18S rRNA. The 18S rRNA reference gene was measured using pre-developed TaqMan assay reagents 18S rRNA MGB according to the manufacturer's protocol (Applied Biosystems). All experiments were performed in two independent sets to ensure reproducibility. Primer sequences for PCR amplification were designed using online Primer3 software (http://bioinfo.ut.ee/primer3/) and listed in Table 1.

Table 1. Sequences of primers used in the real-time quantitative PCR (RT-qPCR) experiments in NRG2.

Gene	Forward primer	Backward primer	Ta
NRG2	CCACAGACCATGTCATCAGG	CCGACTGGGAGTCAGAAGTC	60

Ta, annealing temperature. NRG2 [Gene ID: 9542; GenBank: NM_004883.3].

4.4. DNA Purification

Genomic DNA was prepared from peripheral blood using the Puregene DNA purification system (Gentra Systems Inc. Minneapolis, MI) according to the manufacturer's protocol or from saliva using the Oragene DNA self-collection kit (DNA Genotek, Ottawa, Ontario, Canada) following the manufacturer's instructions.

4.5. PCR Amplification

The genomic sequences of the human neuregulin 2 gene were available from the NCBI (https://www.ncbi.nlm.nih.gov/). NRG2 (Gene ID: 9542) comprises 10 exons with a genomic size of 196,519 bp (RefSeq: NM_004883.3). Optimal PCR primer sequences were generated to amplify each NRG2 exon using Primer3 (http://WWW-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) (Table 2). In the standard reaction, genomic DNA (75 ng) was amplified in a reaction volume of 15 μL containing 0.75 μM each of sense and antisense primers, 0.15 mM of dNTP, 50 mM of KCl, 1.5 mM of MgCl2,0.1% vol/vol of Triton X-100, 10 mM of Tris–HCl (pH 9.0), and 1 U Taq polymerase. PCR conditions consisted of an initial denaturation at 95ºC for 5 min, followed by 30 cycles of 95ºC for 1 min, the optimal annealing temperature of detected amplicons for 1 min, and 72ºC for 1 min. PCR was performed with a PTC-200 DNA engine (MJ Research, Watertown, MA). Each amplicon was subjected to direct sequencing using the ABI Prism™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit Version 3.1 and ABI auto sequencer 3730 (Perkin-Elmer Applied Biosystem, Foster City, CA), according to the manufacturer's protocol.

 doi:10.20944/preprints202403.1755.v1

Table 2. Primer sequences, optimal annealing temperature (Ta) and size of PCR products of the NRG2 gene.

Amplicon	Forward	Reverse	Ta (°C)	Size (bp)
Exon 1.1	TTACGCTGTTTCCGGTTTTC	TGGTCCTGCACTGACTTGAG	60	458
Exon 1.2	GCTTCTCCATGCTGCTCTTC	ttcttctctccaaccccaac	58	586
Exon 2	acagtggcccttactctcca	ctggttccatgggtgagtct	63	372
Exon 3	agggaatctccttcccatct	gttgagtgcgagatggatca	63	356
Exon 4	gagatgattcctggggccta	acttctgacccagcatctcc	60	250
Exon 5	ccaagtgcctgacttggttt	tgcacccagaagctttctaa	63	244
Exon 6	aaggggtctctgcaccacta	acattcttggaggcccatc	67	238
Exon 7-8	gaagttcatcgttggcgagt	ggtgtgctgtgattcctgtg	67	786
Exon 9	gagtggagaagggcattgag	atggagatgaggctctttgg	67	468
Exon 10	gttatgcccgcgtaacagat	CCCAGATGAGCATACAGCAA	60	1343

Ta, annealing temperature. NRG2 [Gene ID: 9542; GenBank: NM_004883.3].

Table 3. Genotype and allele frequencies of variants of the NRG2 gene identified in ASD patients vs controls (from Taiwan biobank).

SNP	Location	groups	n	Genotype			P value	Allele		P value
				A/A	A/G	G/G		A	G	
rs889022	IVS3+13 A>G	ASD	231	197 (85.3%)	30 (13.0%)	4 (1.7%)	0.08	424 (91.8%)	38 (8.2%)	0.10
		control	1495	1193 (79.8%)	284 (19.0%)	18 (1.2%)		2670 (89.3%)	320 (10.7%)	
				T/T	T/A	A/A		T	A	
rs182642591	IVS9+32 T>A	ASD	299	298 (99.7%)	1 (0.3%)	0	0.11	597 (99.8%)	1 (0.2%)	0.11
		control	1495	1473 (98.5%)	22 (1.5%)	0		2968 (99.3%)	22 (0.7%)	
				C/C	C/indelGCCCGGC	indelGCCCGGC/indelGCCCGGC		C	GCCGGGC	
rs1323957797	IntelGCCCGGC	ASD	264	233 (88.3%)	31 (11.7%)	0	<0.0001	497 (94.1%)	31 (5.9%)	<0.0001
		controls	1495	1492 (99.8%)	3 (0.2%)	0		2987 (99.9%)	3 (0.1%)	

The rs1323957797 was also named rs753251234, rs933769137, rs1284917723, rs1226370164 located at chromosome 5: 139,848,469 (Ref. Version: GRCh38).

Table 4. Variants of the NRG2 gene identified in ASD patients and controls and their functional predictions.

location	a.a positon	variants	Consequence type	In silico analysis			Case	Control
				SIFT	PolyPhen-2	UniProt		
Exon 10	661-663	IndelGCCCGGC	Frameshift	neutral	unknown	Likely Benign	31/248 (12.5%)	3/1495 (0.2%)

doi:10.20944/preprints202403.1755.v1

 doi:10.20944/preprints202403.1755.v1

4.6. Direct PCR Autosequencing

After PCR amplification, aliquots of PCR products were processed using the PCR Pre-sequencing Kit (USB Corp. Cleveland, OH) to remove residual primers and dNTPs according to the manufacturer's instructions. The purified PCR products were subjected to direct sequencing using the ABI Prism™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit Version 3.1 and the ABI autosequencer 3730 (Perkin-Elmer Applied Biosystem, Foster City, CA) according to the manufacturer's protocol.

4.7. Statistical Analysis

Differences in allele and genotype frequencies of NRG2 between patients and controls were assessed by the chi-square test or Fisher's exact test where appropriate. Assessment of pair-wise linkage disequilibrium (D') of NRG2 SNPs was implemented using SHEsis computer program [52]. A p-value of <0.05 was considered statistically significant. Moreover, post-hoc power analysis was performed using the computer program G*Power [53].

4.8. In Silico Analysis

Prediction of potential deleterious effects of NRG2 SNPs was performed using the software tools PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/), and UniProt (https://www.uniprot.org/).

5. Conclusions

This study identified three SNPs in NRG2, comprising two noncoding SNPs and one indel in the coding region. In addition, we found significantly lower average NRG2 expression in patients with ASD than controls. However, there were no significant differences in the genotype or allele frequency distributions of the first two noncoding SNPs between ASD patients and controls. In contrast, the genotype and allele frequency distribution of the indel SNP showed significant differences between ASD cases and controls. In conclusion, the higher presence of the indel SNP in ASD patients, potentially leading to decreased NRG2 expression or altered function, suggests that NRG2 may be a candidate gene associated with ASD in Taiwan. This finding might be valuable implications for comprehending the etiology, diagnosis, and treatment of ASD in the Taiwanese population. Moreover, our research offers an effective approach for studying or exploring the genes associated with complex diseases, such as schizophrenia and ASD.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Authors contributions: The contributions of the authors are as follows: Wei-Hsien Chien designed the study and wrote the protocol. Susan Shur-Fen Gau trained the clinical research team and supervised the research execution. Susan Shur-Fen Gau and Yu-Yu Wu were responsible for the ADI-R training and interviews. Susan Shur-Fen Gau and Yu-Yu Wu, conducted clinical diagnosis and helped recruit the patients (in the order of number of patients recruitment). Wei-Hsien Chien, Men-Chih Cheng, and Chia-Hsiang Chen ran the experimental works and analyzed the data. Wei-Hsien Chien wrote the draft, and all the authors approved the manuscript.

Funding: This work was supported by grants from the National Science and Technology Council (NSC96-3112-B-002-033, NSC97-3112-B-002-009, NSC98-3112-B-002-004, NSC 99-3112-B-002-036, MOST 103-2314-B-002-055-MY3, MOST 109-2327-B-002 -004, MOST 110-2327-B-002 -006, MOST 111-2327-B-002 -009 to SS Gau and NSC 107-2635-B-030-002 to WH Chien), Taiwan.

Institutional Review Board Statement: The Institutional Review Board/ Research Ethics Committee of both Hospitals and Fu Jen Catholic university approved the study protocol. Written informed consent were obtained from the parents of patients after the procedures were fully explained. Each individual obtained the written informed permission prior to collecting the peripheral blood or saliva for subsequent laboratory investigations. (ClinicalTrials.gov number, NCT00494754, NCT02228876, C105098.).

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

 doi:10.20944/preprints202403.1755.v1

Data Availability Statement: The raw data are available upon request of the corresponding author. All SNPs information on controls were from Taiwan Biobank (https://taiwanview.twbiobank.org.tw/index). Primers are designed using online Primer3 software (http://bioinfo.ut.ee/primer3/). Prediction of potential deleterious effects of SNPs performed using the software tools were PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) and SIFT (http://sift.jcvi.org/) and UniProt (https://www.uniprot.org/).

Acknowledgments: We thank all the subjects, families, and research colleagues who participated in this study.

Conflicts of Interest: All authors declared no conflicts of interest and were responsible for their authorship in this article. The funders had no role in the study's design, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. America Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM-V). In. Washington, DC: *American Psychiatric Association Press Inc; 2013*.
2. Chen, Y.L.; Chen, W.J.; Lin, K.C.; Shen, L.J.; Gau, S.S. Prevalence of DSM-5 mental disorders in a nationally representative sample of children in Taiwan: methodology and main findings. *Epidemiol. Psychiatr. Sci.* ***2019****, 29, e15.*
3. Maenner, M.J.; Warren, Z.; Williams, A.R.; Amoakohene, E.; Bakian, A.V.; Bilder, D.A.; Durkin, M.S.; Fitzgerald, R.T.; Furnier, S.M.; Hughes, M.M.; et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *MMWR Surveill Summ* ***2023****, 72(2):1-14.*
4. Fombonne, E. Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J. Autism. Dev. Disord.* ***2003****, 33(4):365-382.*
5. Chakrabarti, S.; Fombonne, E. Pervasive developmental disorders in preschool children: confirmation of high prevalence. *Am. J. Psychiatry* ***2005****, 162(6):1133-1141.*
6. Lichtenstein, P.; Carlstrom, E.; Rastam, M.; Gillberg, C.; Anckarsater, H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am. J. Psychiatry* ***2010****, 167(11):1357-1363.*
7. Gupta, A.R.; State, M.W. Recent advances in the genetics of autism. *Biol. Psychiatry* ***2007****, 61(4):429-437.*
8. Holt, R.; Monaco, A.P. Links between genetics and pathophysiology in the autism spectrum disorders. *EMBO Mol. Med.* ***2011****, 3(8):438-450.*
9. Genovese, A.; Butler, M.G. Clinical Assessment, Genetics, and Treatment Approaches in Autism Spectrum Disorder (ASD). *Int. J. Mol. Sci.* ***2020,*** *21(13).*
10. Eapen, V. Genetic basis of autism: is there a way forward? *Curr. Opin. Psychiatr.* ***2011****, 24(3):226-236.*
11. State, M.W.; Levitt, P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nat. Neurosci.* ***2011****, 14(12):1499-1506.*
12. Ebrahimi-Fakhari, D.; Sahin, M. Autism and the synapse: emerging mechanisms and mechanism-based therapies. *Curr. Opin. Neurol.* ***2015****, 28(2):91-102.*
13. Schena, M.; Shalon, D.; Heller, R.; Chai, A.; Brown, P.O.; Davis, R.W. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U. S. A.* ***1996****, 93(20):10614-10619.*
14. Nishimura, Y.; Martin, C.L.; Vazquez-Lopez, A.; Spence, S.J.; Alvarez-Retuerto, A.I.; Sigman, M.; Steindler, C.; Pellegrini, S.; Schanen, N.C.; Warren, S.T.; et al. Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways. *Hum. Mol.* ***2007****, 16(14):1682-1698.*
15. Hu, V.W.; Nguyen, A.; Kim, K.S.; Steinberg, M.E.; Sarachana, T.; Scully, M.A.; Soldin, S.J.; Luu, T.; Lee, N.H. Gene expression profiling of lymphoblasts from autistic and nonaffected sib pairs: altered pathways in neuronal development and steroid biosynthesis. *PloS One* ***2009****, 4(6): e5775.*
16. Chien, W.H.; Gau, S.S.; Chen, C.H.; Tsai, W.C.; Wu, Y.Y.; Chen, P.H.; Shang, C.Y. Increased gene expression of FOXP1 in patients with autism spectrum disorders. *Mol. Autism* ***2013****, 4(1):23.*
17. Falls, D.L. Neuregulins: functions, forms, and signaling strategies. *Exp Cell Res* ***2003****, 284(1):14-30.*
18. Olayioye, M.A.; Neve, R.M.; Lane, H.A.; Hynes, N.E. The ErbB signaling network: receptor heterodimerization in development and cancer. *EMBO J.* ***2000****, 19(13):3159-3167.*
19. Benzel, I.; Bansal, A.; Browning, B.L.; Galwey, N.W.; Maycox, P.R.; McGinnis, R.; Smart, D.; St Clair, D.; Yates, P.; Purvis, I. Interactions among genes in the ErbB-Neuregulin signalling network are associated with increased susceptibility to schizophrenia. *Behav. Brain. Funct.* ***2007****, 3:31.*
20. Buonanno, A.; Fischbach, G.D. Neuregulin and ErbB receptor signaling pathways in the nervous system. *Curr. Opin. Neurobiol.* ***2001****, 11(3):287-296.*
21. Buonanno, A. The neuregulin signaling pathway and schizophrenia: from genes to synapses and neural circuits. *Brain Res Bull* ***2010****, 83(3-4):122-131.*
22. Mei, L.; Nave, K.A. Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases. *Neuron* ***2014****, 83(1):27-49.*

 doi:10.20944/preprints202403.1755.v1

23. Mostaid, M.S.; Lloyd, D.; Liberg, B.; Sundram, S.; Pereira, A.; Pantelis, C.; Karl, T.; Weickert, C.S.; Everall, I.P.; Bousman, C.A. Neuregulin-1 and schizophrenia in the genome-wide association study era. *Neurosci. Biobehav. Rev.* ***2016****, 68:387-409.*
24. Neddens, J.; Fish, K.N.; Tricoire, L.; Vullhorst, D.; Shamir, A.; Chung, W.; Lewis, D.A.; McBain, C.J.; Buonanno, A. Conserved interneuron-specific ErbB4 expression in frontal cortex of rodents, monkeys, and humans: implications for schizophrenia. *Biol. Psychiatry* ***2011****, 70(7):636-645.*
25. Shamir, A.; Kwon, O.B.; Karavanova, I.; Vullhorst, D.; Leiva-Salcedo, E.; Janssen, M.J.; Buonanno, A. The importance of the NRG-1/ErbB4 pathway for synaptic plasticity and behaviors associated with psychiatric disorders. *J Neurosci* ***2012****, 32(9):2988-2997.*
26. Brennand, K.J.; Simone, A.; Jou, J.; Gelboin-Burkhart, C.; Tran, N.; Sangar, S.; Li, Y.; Mu, Y.; Chen, G.; Yu, D.; et al. Modelling schizophrenia using human induced pluripotent stem cells. *Nat.* ***2011****, 473(7346):221-225.*
27. Ring, H.Z.; Chang, H.; Guilbot, A.; Brice, A.; LeGuern, E.; Francke, U. The human neuregulin-2 (NRG2) gene: cloning, mapping and evaluation as a candidate for the autosomal recessive form of Charcot-Marie-Tooth disease linked to 5q. *Hum. Genet.* ***1999****, 104(4):326-332.*
28. Busfield, S.J.; Michnick, D.A.; Chickering, T.W.; Revett, T.L.; Ma, J.; Woolf, E.A.; Comrack, C.A.; Dussault, B.J.; Woolf, J.; Goodearl, A.D.; et al. Characterization of a neuregulin-related gene, Don-1, that is highly expressed in restricted regions of the cerebellum and hippocampus. *Mol. Cell. Biol.* ***1997****, 17(7):4007-4014.*
29. Carraway, K.L.; Weber, J.L.; Unger, M.J.; Ledesma, J.; Yu, N.; Gassmann, M.; Lai, C. Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine kinases. *Nat.* ***1997****, 387(6632):512-516.*
30. Chang, H.; Riese, D.J.; Gilbert, W.; Stern, D.F.; McMahan, U.J. Ligands for ErbB-family receptors encoded by a neuregulin-like gene. *Nat.* ***1997****, 387(6632):509-512.*
31. Vullhorst, D.; Mitchell, R.M.; Keating, C.; Roychowdhury, S.; Karavanova, I.; Tao-Cheng, J.H. Buonanno A: A negative feedback loop controls NMDA receptor function in cortical interneurons via neuregulin 2/ErbB4 signalling. *Nat. Commun.* ***2015****, 6:7222.*
32. Yan, L.; Shamir, A.; Skirzewski, M.; Leiva-Salcedo, E.; Kwon, O.B.; Karavanova, I.; Paredes, D.; Malkesman, O.; Bailey, K.R.; Vullhorst, D.; et al. Neuregulin-2 ablation results in dopamine dysregulation and severe behavioral phenotypes relevant to psychiatric disorders. *Mol. Psychiatry* ***2018****, 23(5):1233-1243.*
33. Konte, B.; Leicht, G.; Giegling, I.; Pogarell, O.; Karch, S.; Hartmann, A.M.; Friedl, M.; Hegerl, U.; Rujescu, D.; Mulert, C. A genome-wide association study of early gamma-band response in a schizophrenia case-control sample. *World J. Biol. Psychiatry* ***2018,*** *19(8):602-609.*
34. Seshadri, S.; Kamiya, A.; Yokota, Y.; Prikulis, I.; Kano, S.; Hayashi-Takagi, A.; Stanco, A.; Eom, T.Y.; Rao, S.; Ishizuka, K.; et al. Disrupted-in-Schizophrenia-1 expression is regulated by beta-site amyloid precursor protein cleaving enzyme-1-neuregulin cascade. *Proc. Natl. Acad. Sci. U. S. A.* ***2010****, 107(12):5622-5627.*
35. Mostaid, M.S.; Lee, T.T.; Chana, G.; Sundram, S.; Shannon, Weickert. C.; Pantelis, C.; Everall, I.; Bousman, C. Peripheral Transcription of NRG-ErbB Pathway Genes Are Upregulated in Treatment-Resistant Schizophrenia. *Front. Psychiatry* ***2017****, 8:225.*
36. Diaz-Moran, S.; Palencia, M.; Mont-Cardona, C.; Canete, T.; Blazquez, G.; Martinez-Membrives, E.; Lopez-Aumatell, R.; Sabariego, M.; Donaire, R.; Moron, I.; et al. Gene expression in hippocampus as a function of differential trait anxiety levels in genetically heterogeneous NIH-HS rats. *Behav. Brain Res.* ***2013****, 257:129-139.*
37. Ma, D.Q.; Cuccaro, M.L.; Jaworski, J.M.; Haynes, C.S.; Stephan, D.A.; Parod, J.; Abramson, R.K.; Wright, H.H.; Gilbert, J.R.; Haines, J.L.; et al. Dissecting the locus heterogeneity of autism: significant linkage to chromosome 12q14. *Mol. Psychiatry* ***2007****, 12(4):376-384.*
38. Philippi, A.; Roschmann, E.; Tores, F.; Lindenbaum, P.; Benajou, A.; Germain-Leclerc, L.; Marcaillou, C.; Fontaine, K.; Vanpeene, M.; Roy, S.; et al. Haplotypes in the gene encoding protein kinase c-beta (PRKCB1) on chromosome 16 are associated with autism. *Mol. psychiatry* ***2005****, 10(10):950-960.*
39. Neves-Pereira, M.; Muller, B.; Massie, D.; Williams, J.H.; O'Brien, P.C.; Hughes, A.; Shen, S.B.; Clair, D.S.; Miedzybrodzka, Z. Deregulation of EIF4E: a novel mechanism for autism. *J. Med. Genet.* ***2009****, 46(11):759-765.*
40. Vincent, J.B.; Noor, A.; Windpassinger, C.; Gianakopoulos, P.J.; Schwarzbraun, T.; Alfred, S.E.; Stachowiak, B.; Scherer, S.W.; Roberts, W.; Wagner, K.; et al. Characterization of a de novo translocation t(5;18)(q33.1;q12.1) in an autistic boy identifies a breakpoint close to SH3TC2, ADRB2, and HTR4 on 5q, and within the desmocollin gene cluster on 18q. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* ***2009,*** *150B(6):817-826.*
41. Della Monica, M.; Lonardo, F.; Faravelli, F.; Pierluigi, M.; Luquetti, D.V.; De Gregori, M.; Zuffardi, O.; Scarano, G. A case of autism with an interstitial 1q deletion (1q23.3-24.2) and a de novo translocation of chromosomes 1q and 5q. *Am. J. Med. Genet. A.* ***2007****, 143A(22):2733-2737.*
42. Philippi, A.; Tores, F.; Carayol, J.; Rousseau, F.; Letexier, M.; Roschmann, E.; Lindenbaum, P.; Benajjou, A.; Fontaine, K.; Vazart, C.; et al. Association of autism with polymorphisms in the paired-like homeodomain transcription factor 1 (PITX1) on chromosome 5q31: a candidate gene analysis. *BMC Med. Genet.* ***2007****, 8:74.*

43. Schwab, S.G.; Eckstein, G.N.; Hallmayer, J.; Lerer, B.; Albus, M.; Borrmann, M.; Lichtermann, D.; Ertl, M.A.; Maier, W.; Wildenauer, D.B. Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sib-pair linkage analysis. *Mol. Psychiatry* ***1997**, 2(2):156-160.*
44. Sklar, P.; Pato, M.T.; Kirby, A.; Petryshen, T.L.; Medeiros, H.; Carvalho, C.; Macedo, A.; Dourado, A.; Coelho, I.; Valente, J.; et al. Genome-wide scan in Portuguese Island families identifies 5q31-5q35 as a susceptibility locus for schizophrenia and psychosis. *Mol. Psychiatry* ***2004**, 9(2):213-218.*
45. Arcos-Burgos, M.; Castellanos, F.X.; Pineda, D.; Lopera, F.; Palacio, J.D.; Palacio, L.G.; Rapoport, J.L.; Berg, K.; Bailey-Wilson, J.E.; Muenke, M. Attention-deficit/hyperactivity disorder in a population isolate: linkage to loci at 4q13.2, 5q33.3, 11q22, and 17p11. *Am. J. Hum. Genet.* ***2004**, 75(6):998-1014.*
46. Czarnek, M.; Bereta, J. Proteolytic Processing of Neuregulin 2. *Mol. Neurobiol.* ***2020**, 57(4):1799-1813.*
47. Lord, C.; Rutter, M.; Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* ***1994**, 24(5): 659-685.*
48. Gau, S.S.; Lee, C.M.; Lai, M.C.; Chiu, Y.N.; Huang, Y.F.; Kao, J.D.; Wu, Y.Y. Psychometric properties of the Chinese version of the social communication questionnaire. *Res. Autism Spectr. Disord.* ***2011**, 5, 809-818.*
49. Chang, J.C.; Lai, M.C.; Chien, Y.L.; Cheng, C.Y.; Wu, Y.Y.; Gau, S.S. Psychometric properties of the Mandarin version of the Autism Diagnostic Observation Schedule-Generic. *J. Formos. Med. Assoc.* ***2023**, 122, 574-583.*
50. Pan, W.H.; Fann, C.S.; Wu, J.Y.; Hung, Y.T.; Ho, M.S.; Tai, T.H.; Chen, Y.J.; Liao, C.J.; Yang, M.L.; Cheng, A.T.; et al. Han Chinese cell and genome bank in Taiwan: purpose, design and ethical considerations. *Hum. Hered.* ***2006**, 61(1):27-30.*
51. Huang, C.H.; Chen, M.L.; Tsai, Y.L.; Tsai, M.T.; Chen, C.H. Elevated adrenomedullin mRNA in lymphoblastoid cells from schizophrenic patients. *Neuroreport.* ***2004**, 15(9):1443-1446.*
52. Shi, Y.Y.; He, L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* ***2005**, 15(2):97-98.*
53. Buchner, A. F.F.; Erdfelder, E. G-Power: A priori, post hoc and Compromise Power Analyses for the Macintosh, Version 2.1.1. *Trier: University of Trier;* ***1996**.*