Integrative Variation Analysis Reveals that a Complex Genotype May Specify Phenotype in Siblings with Syndromic Autism Spectrum Disorder

Viviane Neri de Souza Reis[1]*, **João Paulo Kitajima**[2], **Ana Carolina Tahira**[1], **Ana Cecília Feio-dos-Santos**[1], **Rodrigo Ambrósio Fock**[3], **Bianca Cristina Garcia Lisboa**[1], **Sérgio Nery Simões**[4], **Ana C. V. Krepischi**[5], **Carla Rosenberg**[5], **Naila Cristina Lourenço**[5], **Maria Rita Passos-Bueno**[5], **Helena Brentani**[1]

1 LIM23-Institute of Psychiatry, University of São Paulo School of Medicine, São Paulo, Brazil,
2 Bioinformatics, Mendelics Análise Genômica, São Paulo, Brazil, 3 Department of Morphology and Genetics, Federal University of São Paulo, São Paulo, Brazil, 4 Department of Informatics, Federal Institute of Espírito Santo, Serra, Brazil, 5 Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of Sao Paulo, São Paulo, Brazil

* viviane.nsreis@gmail.com

Abstract

It has been proposed that copy number variations (CNVs) are associated with increased risk of autism spectrum disorder (ASD) and, in conjunction with other genetic changes, contribute to the heterogeneity of ASD phenotypes. Array comparative genomic hybridization (aCGH) and exome sequencing, together with systems genetics and network analyses, are being used as tools for the study of complex disorders of unknown etiology, especially those characterized by significant genetic and phenotypic heterogeneity. Therefore, to characterize the complex genotype-phenotype relationship, we performed aCGH and sequenced the exomes of two affected siblings with ASD symptoms, dysmorphic features, and intellectual disability, searching for *de novo* CNVs, as well as for *de novo* and rare inherited point variations—single nucleotide variants (SNVs) or small insertions and deletions (indels)—with probable functional impacts. With aCGH, we identified, in both siblings, a duplication in the 4p16.3 region and a deletion at 8p23.3, inherited by a paternal balanced translocation, t(4, 8) (p16; p23). Exome variant analysis found a total of 316 variants, of which 102 were shared by both siblings, 128 were in the male sibling exome data, and 86 were in the female exome data. Our integrative network analysis showed that the siblings' shared translocation could explain their similar syndromic phenotype, including overgrowth, macrocephaly, and intellectual disability. However, exome data aggregate genes to those already connected from their translocation, which are important to the robustness of the network and contribute to the understanding of the broader spectrum of psychiatric symptoms. This study shows the importance of using an integrative approach to explore genotype-phenotype variability.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder [1] characterized by early-onset social and communication impairment, as well as restrictive stereotypic movements and behaviors [2]. Approximately 70% of ASD patients have at least one clinical or psychiatric comorbidity, whereas 48% have two or more [3]. In addition to presenting a heterogeneous phenotype, ASD is multifactorial disorder with a complex genetic architecture [4–7]. Various reviews have provided in-depth analyses of the degree of genetic and phenotypic association between ASD and genetic syndromes [8–12]. Miles et al. [13] defined complex ASD as that observed in a subset of individuals with evidence of abnormality occurring at the beginning of morphogenesis, manifesting as dysmorphic abnormalities but not associated with a known genetic syndrome. Therefore, syndromic ASD is represented by complex cases with or without an association with a known genetic syndrome.

Although common variations (with a minor allele frequency > 1%) are the largest genetic component of ASD, accounting for 40% of the risk [14,15], their individual effect is small [16]. Whole-genome screening for *de novo* and inherited copy number variations (CNVs) have revealed that rare variants are more common in ASD probands than in healthy controls and unaffected family members [17–22]. Comparisons of syndromic and non-syndromic cases of ASD have revealed that clinically relevant CNVs are more common in syndromic probands [23,24]. The pathogenicity of *de novo* CNVs has been related to CNV length, larger CNVs having been shown to have a greater impact [17,25–27]. However, recurrent CNVs have been associated with an increased risk of other neurodevelopmental disorders, such as schizophrenia and epilepsy [28]. Various authors have attempted to improve the understanding of the nature of CNVs and to elucidate their contribution to specific phenotypes [29,30]. In patients with developmental disorders, Andrews et al. [30] found that, if protein-protein interactions are taken into account, patients with genic (especially *de novo*) CNVs, disrupting the same cluster of functionally related genes, present phenotypes that are more similar than would be expected. Whole-exome sequencing studies have reported that *de novo* non-synonymous single nucleotide variants (SNVs) and small insertions and deletions (indels) are strongly associated with ASD [31–34]. Iossifov et al. [35] showed that almost all gene-disrupting mutations occur opposite wild-type alleles, that 13% of *de novo* missense mutations account for 12% of diagnoses, and that 43% of likely gene-disrupting *de novo* mutations account for 9% of diagnoses. A whole-genome sequencing study of 85 quartet families (two parents and two ASD-affected siblings) found that 69.4% of the affected siblings carried different ASD-relevant mutations (*de novo* and rare inherited SNVs or structural variants) and that sibling pairs with discordant mutations usually showed more clinical variability than did those who shared a risk variant [36].

The prevalence of CNVs suggests that chromosomal microarrays should be used in the genetic analysis of syndromic ASD or familial cases of the disorder, especially when the patient has an intellectual disorder [37]. Patients with localized or generalized syndromic overgrowth should also be tested, given that the condition can be caused by subtle genomic rearrangements [38]. It is of note that even in the presence of large pathogenic CNVs, increasing heterogeneity and complexity of a phenotype usually imply a greater number of variants. Assuming that different ASD phenotypes could be the product of multiple biological pathways, possibly affecting the same brain mechanisms, more studies should focus on systems genetics and network analyses [39,40]. This underscores the importance of gaining a better understanding of the role that sets of variant genes play in phenotype differences, which could be achieved through comparative studies of affected individuals with shared and distinct phenotypic effects of rare alleles [40]. Therefore, to characterize the complex genotype-phenotype relationship,

we performed an analysis of CNVs and rare point variations by conducting array comparative genomic hybridization (aCGH) and exome sequencing in two syndromic siblings with intellectual disability (ID) who shared various dysmorphic features, although they showed significant clinical differences in terms of the psychiatric phenotype.

Materials and Methods

Clinical assessment of the subjects

Two siblings (one male and one female) with syndromic ASD and their parents volunteered to participate in an ongoing exome research study at the Autism Spectrum Disorder Program Clinic of the Psychiatric Institute of the University of São Paulo School of Medicine. The study was approved by the Research Ethics Committee for the Analysis of Research Projects of the University of São Paulo School of Medicine *Hospital das Clínicas* (Ruling no. 0692–11). All subjects gave written informed consent. The siblings underwent evaluation by specialized clinicians. The diagnosis of ASD was based on the criteria established in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, as well as on the results of morphological and neurological examinations. The Childhood Autism Rating Scale (CARS) was used in order to determine the severity of the ASD symptoms. The results of karyotyping, screening for innate metabolism errors (testing for amino acids and saccharides in urine and serum), and the test for fragile X syndrome were all negative.

CNV analysis

We performed aCGH on a commercial whole-genome 180K platform containing approximately 180,000 oligonucleotide probes distributed throughout the human genome (Agilent Technologies, Santa Clara, CA, USA). A pool of healthy human DNA (Promega, Madison, WI, USA) was used as a reference in the experiments, and the procedure was performed as recommended by the manufacturer. Scanned images were processed using Feature Extraction 10.7.3.1 software (Agilent Technologies). We called copy number alterations on the basis of the $\log_2$ ratio of the Cy3/Cy5 intensities in a given genomic segment, > 0.3 or < -0.3 for gains or losses, respectively, using the aberration detection statistical algorithm aberration detection method 2, with a sensitivity threshold of 6.7. Because we were interested in pathogenic CNVs, we selected only *de novo* genic CNVs that were larger than 100 kb and were considered uncertain, likely pathogenic, or pathogenic according to ClinVar similarity regarding the CNV region (NCBI Variant Viewer, https://www.ncbi.nlm.nih.gov/variation/view/).

Imbalances identified with aCGH were validated by multiple ligation probe amplification (MLPA) with two probe sets (SALSA MLPA P036-E1 human telomere-3 and SALSA MLPA P070-B2 human telomere-5; MRC-Holland, Amsterdam, the Netherlands). Fluorescence *in situ* hybridization (FISH) was also use in order to determine whether the rearrangement had been inherited as a balanced translocation. The FISH was performed according to standard protocols.

Exome sequencing

We extracted DNA from whole blood and quantified it with a fluorometer (Qubit; Life Technologies Corporation, Carlsbad, CA, USA). The exome was captured and enriched from genomic DNA using an exome enrichment kit (Nextera; Illumina Inc., San Diego, CA, USA), representing 62 Mbp of the human genome (hg19 build). Sequencing was performed with a high-resolution optical imaging system (HiScan; Illumina Inc.).

Fragments were aligned to the hg19 reference genome through the use of Burrows-Wheeler alignment [41], and variant calling was performed with a software package for the analysis of sequence data (Genome Analysis Toolkit; Broad Institute, Cambridge, MA, USA) [42–44]. A total of 800 informative single nucleotide polymorphisms were analyzed with the whole-genome association analysis toolset (PLINK; http://pngu.mgh.harvard.edu/~purcell/plink/) to confirm the familial relationship between samples [45]. For genotype calls (in all members of the quartet, we used only reads with a Phred score $\geq$ 30 [46], bases with $\geq$ 20 reads, and a genotype quality score $\geq$ 20 [33].

SNV analysis

Variant annotation was performed with an annotation program (ANNOVAR, http://www.openbioinformatics.org/annovar/) [47]. To select rare variants (minor allele frequency $\leq$ 0.01) and novel variants, we employed the following databases: 1000 Genomes Project (http://www.1000genomes.org/); the National Heart, Lung, and Blood Institute GO Exome Sequencing Project (6500 exomes; http://evs.gs.washington.edu/EVS/); and the Exome Aggregation Consortium (http://exac.broadinstitute.org/). Variants with a depth of coverage of less than 20× reads in all members of the family were excluded in order to avoid false positives. To identify *de novo* variations, we accepted candidates only at loci that were homozygous for the reference allele in the parents. For the gene analysis, we selected only nonsynonymous SNVs, small indels, and splicing site variations. If present, we selected the variations that were considered conserved (score $\geq$ 2) in the Genomic Evolutionary Rate Profiling++ algorithm [48] were selected, as were those that were considered deleterious or possibly deleterious in the Polymorphism Phenotyping v2 HDIV algorithm [49].

Integration of CNVs and exome variants: network analysis

We performed enrichment analyses with a web-based gene set analysis toolkit (WebGestalt; http://bioinfo.vanderbilt.edu/webgestalt/) [50], using a list of the brain-expressed genes among the CNVs. For those analyses, we considered only genes expressed in at least one sample from any brain area with positive expression according to the BrainSpan Atlas of the Developing Human Brain [51].

For the integrative approach, network analysis was performed with the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; http://string-db.org/), the minimum interaction confidence score being 0.4 (medium confidence). The STRING database comprises known, predicted protein interactions, as well as direct (physical) and indirect (functional) associations, derived from genomic context, high-throughput experiments, conserved coexpression, and text mining [52]. For prioritization, we used measures of centrality, such as degree, bridgeness, and brokering, Degree is the number of edges of a vertex (node), and vertices with a very high degree are known as hubs [53,54]. Bridgeness is a measure of the extent to which a node or edge is located between connected regions [55], and brokering is when a broker gene acts as a hub that connects many genes that do not otherwise connect with each other [56].

Results

Phenotypic similarities and differences between the siblings

The two patients evaluated were siblings born to a nonconsanguineous couple. There was no family history of ASD, other psychiatric disorders, ID, or malformations.

The male sibling had his first physician visit at 11 years of age (in 1992). At the time, the mother complained that the child had spoken individual words only after 2 years of age and had stopped producing any words at all by 4 years of age. He had difficulty learning new things at school, had no social interests, and was very attached to routines. The mother also stated that she had had an uneventful pregnancy and a vaginal delivery without complications. She could not remember the birth weight or length. However, she was able to remember that he had taken his first steps at 2 years of age. He was diagnosed with ID in 1992 and with ASD in 2003. He began to show irritability, mainly in the morning, and sometimes became autoaggressive or exhibited outwardly directed aggressive behavior. Since his first physician visit (in 1992), he had been under pharmacological treatment with antipsychotic agents.

The female sibling had her first physician visit at 7 years of age (also in 1992). At the time, the mother complained that the girl had difficulties learning new things at school and that her social skills were extremely inadequate. She had taken her first steps at 18 months of age, and she had not begun to produce speech until after 2 years of age. The mother reported having had abdominal pain and vaginal bleeding during pregnancy and that it had been a cesarean delivery (birth weight: 4140 g; birth length: 51 cm). The patient was diagnosed with ID in 1992 and with ASD in 2003. Since her first physician visit, she exhibited psychomotor agitation, mainly in the mornings. She was occasionally aggressive, although she was an affectionate and gentle child most of the time. She liked playing with other children. However, she sometimes exhibited periods of inappropriate, hypersocial behavior and periods during which she could not stop talking and presented sleep difficulties. Since 1993 (when she was 8 years of age), she had been under pharmacological treatment with antipsychotic agents.

In the most recent medical evaluation (S1 Table), the male sibling met the criteria for a diagnosis of ASD according to the Diagnostic and Statistical Manual of Mental Disorders, fifth edition [2], whereas the female sibling was diagnosed with ID and syndromic features. The CARS interview with the parents resulted in a score of 33.5 for the male sibling and 24.0 for the female sibling (details of the psychiatric examination and neurological evaluation of the siblings are provided in S1 Table).

aCGH and MLPA analyses

The syndromic presentation suggested the existence of a chromosomal alteration, which was investigated with aCGH (Fig 1). In the two siblings, there were 12 *de novo* CNVs (S2 Table): 8 in the female and 4 in the male. Filtering revealed that both siblings had 2 terminal chromosomal alterations: a 4.6-Mb duplication at 4p16.2-p16.3 (chr4:71552–4673343; hg19); and a 752-kb deletion at 8p23.3 (chr8:176,814–928,886; hg19). The MLPA analysis confirmed the presence of the 4p16 duplication and the 8p23 deletion (S1 Fig).

FISH analysis

The FISH analysis confirmed the copy number alterations detected in the siblings (4p16.3 duplication and 8p23.1 deletion, as shown in Fig 2A and Fig 2B, respectively), and the chromosomal imbalances were derived from a translocation between chromosomes 4 and 8 [der(8)t (4;8)(p16.3;p23.1)]. The FISH analysis performed in the parents revealed that the father (Fig 2C) was a carrier of the balanced translocation t(4;8)(p16.3;p23.1).

Phenotype and chromosomal alterations correlation

Various cases of translocations between 4p16.6 and 8p23.1 have been described, the first having been reported by Goggin et al. [57]. Table 1 summarizes the characteristics observed in our
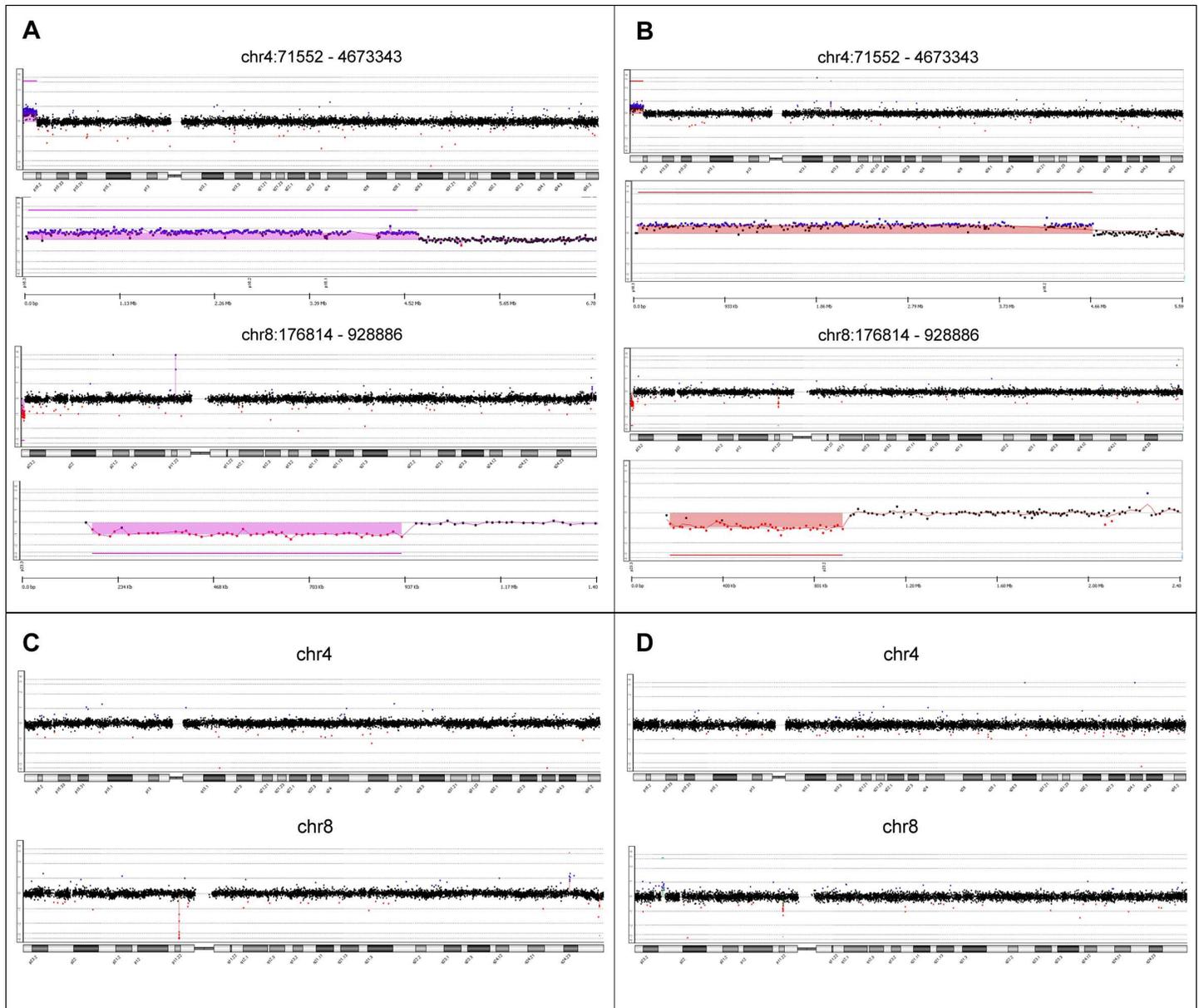
Fig 1. Chromosomal alterations detected with aCGH analysis in both siblings. A and B present the aCGH profiles of chromosomes 4 and 8, showing the 4p16.3 duplication and the 8p23.3 deletion in the male and female siblings, respectively; C and D present the normal copy number profiles of chromosomes 4 and 8 of the father and mother, respectively.

patients compared with those of other reported cases, indicating what is typically found in cases of pure 4p duplication and pure 8p deletion [57–66].

Genes in chromosomal alterations

To explore the phenotypes associated with the chromosomal alterations detected, we conducted an enrichment analysis based on the 69 genes present in the 4p duplication, expressed in brain tissue and present in the WebGestalt database (of the 80 RefSeq genes present in the region according to the hg19 build from the UCSC Genome Browser). The
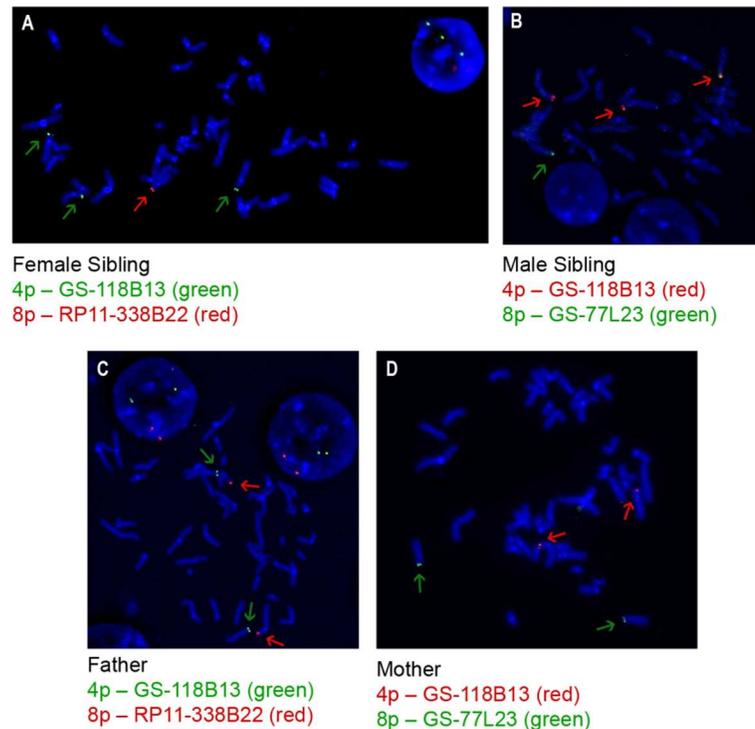
Fig 2. FISH images. (A) image showing that, in the female sibling, there were three chr4p copies, marked in green with a GS-118B13 probe (green arrow), and one chr8p copy, marked in red with a RP11-338B22 probe (red arrow); (B) image showing that, in the male sibling, there were also three chr4p copies, marked in red with a GS-118B13 probe (red arrow), and one chr8p copy, marked in green with a GS-77L23 probe (green arrow); (C) image showing that, in the father, there was a balanced translocation, t(4;8); and (D) image showing that, in the mother, the chromosomes were normal.

disease enrichment analysis (S3 Table) revealed that 5 genes (*WHSC1*, *FGFR3*, *SH3BP2*, *LETM1*, and *STX18*) were found to be overrepresented in craniofacial abnormalities (adjusted *p*-value = 0.0002), musculoskeletal abnormalities (adjusted *p*-value = 0.0009) and congenital abnormalities (adjusted p-value = 0.0040). Other genes, such as *GAK*, *HTT*, *TMEM175*, *FAM193A*, and *DGKQ*, were mainly related to movement disorders (adjusted *p*-value = 0.0001) and basal ganglia disease (adjusted *p*-value = 3.73e-05). The gene ontology analysis revealed enrichment of fibroblast growth factor-activated receptor activity (*FGFR3* and *FGFRL1* genes, adjusted *p*-value = 0.0082) and rough endoplasmic reticulum (*NAT8L*, *LRPAP1*, and *HGFAC* genes, adjusted *p*-value = 0.0445).

The 8p23.1 deletion affects 6 brain-expressed genes (from a total of 8 RefSeq genes according to the hg19 build from the UCSC Genome Browser): *RPL23AP53*, *ZNF596*, *FAM87A*, *FBXO25*, *TDRP*, and *ERICH1*. As expected, there was no significant functional enrichment.

Using the STRING, we searched possible connections described in the literature and protein-protein interaction databases connecting all the CNV genes. There were 61 genes represented in the STRING (S2 Fig), with 32 connected proteins (i.e., 52.45%, 32/61) and 29 interactions, representing enrichment of interactions (*p*-value = 3.09e-14). In addition, genes overrepresented in craniofacial abnormalities, musculoskeletal abnormalities, syndromic abnormalities, or congenital abnormalities were all presented in the biggest component of the network formed (the component with the most number of connecting genes). The genes

Table 1. Summary of the characteristics observed in our patients, compared with those of other cases described in the literature.

Characteristic	Sibling 1	Sibling 2	Skrlec et al. [58]	Mau et al. [59]	Partington et al. [60] Female (1)	Female (2)	Goggin et al. [57]	Literature on dup(4p)*	Literature on del(8p)*
Sex	Male	Female	Male	Female	Female (1)	Female (2)	Male	NA	NA
Age at evaluation	33 years	29 years	5 months	2 years, 10 months	15 years	16 years	NA	NA	NA
Cytogenetic alterations	46,XY, der (8)t(4;8) (p16.3; p23.1)	46,XX, der (8)t(4;8) (p16.3; p23.1)	46,XY,der(8)t(4;8) (p16.1;p23.1).ish der (8)t(4;8) (D8-S504-, WHCR+, D8Z2+)dn	46,XX,der (8) t(4;8) (p12; p23.1)	46,XX,der (8) t(4;8) (p16.2; p23.1)	46,XX,der (8) t(4;8) (p16.2; p23.1)	46,XY,der (8) t(4;8) (p12; p23)	NA	NA
Neonatal period									
Birth height (cm)			52	44				NA	NA
Birth weight (g)			3360	2450				NA	NA
Birth OFC (cm)			35	31				NA	NA
Anthropometric									
Height (cm)	190	183		84.5				NA	NA
Weight (kg)	95.8	91.5		11,2				NA	NA
OFC (cm)	60	59		46,5				NA	NA
Arm span (cm)	194	185		NA				NA	NA
Dysmorphic features									
Microcephaly	-	-		+			+	+	+
Macrocephaly	+	+				+			
Round face	-	-		+					
Long face	+	+							
Short neck	-	-					+	+	+
Low anterior hairline	-	-		+					
Low posterior hairline	+	+		+				+	
Arched eyebrows	+	+		+				+	
Long eyelashes	+	+		+				+	
Synophrys	-	-		+				+	
Hirsutism	-	-	+	+				+	
Microphthalmia	-	-						+	+
Epicanthic folds	-	-							+
Hypertelorism	-	-						+	+
Wide nasal bridge	+	+	+	+			+	+	+
Anteverted nares	-	-	+	+					
Long philtrum	-	-		+				+	
Thin lips	-	+		+					+
Down-turned corners of the mouth	-	-		+					
Micrognathia or Retrognathia	-	-	+					+	+
Prognathism	-	-							
Widely spaced teeth	-	-		+				+	
Crowded teeth	+	+							
Low-set ears	+	+	+	+				+	+
Malformed ears	+	+		+				+	+

(*Continued*)

Table 1. (*Continued*)

Characteristic	Sibling 1	Sibling 2	Skrlec et al. [58]	Mau et al. [59]	Partington et al. [60]		Goggin et al. [57]	Literature on dup(4p)*	Literature on del(8p)*
Widely spaced nipples	+	NA						+	+
Large hands	+	+			+	+			
Short tapering fingers	-	-						+	
Proximal implantation of thumbs	-	-		+					+
Arachnodactyly	-	-						+	
Clinodactyly	+	+	+					+	
Large feet	+	+			+	+			
Proximal implantation of toes	-	-	+						
Syndactyly	-	-						+	
Transverse palmar crease	-	-		+				+	
Hypoplastic nails	-	+		+					+
Rectus abdominis diastasis	-	-	+						
Cryptorchidism	-	-	+					+	
Hypospadias	-	-						+	+
Malformative features									
Cardiopathy			AVSD					+	+
Genitourinary anomalies			Vesicoureteral reflux	Small clitoris				+	+
Hypoplasia of the corpus callosum			+						
Coloboma of the iris or retina	-	-					Both		
Other									
Behavior features									
Intellectual disability	+	+	+	+	+	+	+		+
Seizures	-	-		-				+	+
Stereotypic movements	+	-		+					
Hyperactive	-	-		+					+
Obsessional behavior	-	-			+				
Feeding difficulties	-	-		+				+	+
Recurrent infections	-	-		+					
Other	-	-		NHS					

NA: not applicable; OFC: occipitofrontal circumference; AVSD: atrioventricular septal defect; NHS: neonatal hypocalcemic seizure.

*[61–66]

Table 2. Details of the hemizygous variations found in the exome of the male sibling.

Chr	Start	End	Ref	Alt	Gene	Variation	1000G	ESP6500	ExAC	dbSNP ID	Polyphen2	GERP++
X	19032163	19032163	G	T	*ADGRG2*	Nonsyn. SNV	0.0013	0.0017	0.0006	rs150013899	0.745	3.37
X	15376167	15376167	G	C	*FIGF*	Nonsyn. SNV	NA	NA	NA	NA	0.804	3.30
X	150906999	150906999	A	C	*CNGA2*	Nonsyn. SNV	0.0079	0.0085	0.0025	rs142562634	0.981	3.69

Chr: chromosome; Ref: reference allele; Alt: altered allele; 1000G: 1000 Genomes Project (database); ESP6500: National Heart, Lung, and Blood Institute 6500 exome GO Exome Sequencing Project (database); ExAC: Exome Aggregation Consortium (database); dbSNP: Single Nucleotide Polymorphism Database; Polyphen2: Polymorphism Phenotyping v2 (algorithm); GERP++: Genomic Evolutionary Rate Profiling++ (algorithm); Nonsyn.: nonsynonymous; SNV: single nucleotide variant; NA: not applicable.

doi:10.1371/journal.pone.0170386.t002

ZNF596 and *FBXO25* from 8p23.1 appeared to be connected, although not within the biggest component of the network (S2 Fig).

Exome sequencing

We sequenced 45 million reads per sample (4.5 Gb). To avoid miscalling variants [67], we removed polymerase chain reaction duplicates, resulting in 32 million reads (raw data submitted at the European Nucleotide Archive, accession number PRJEB15585). The mean coverage on target was 35×, and 44.63% of the targets showed at least 20× coverage (S4 Table).

After the SNV filtering process, only 1 *de novo* variant was identified, disrupting the *TUBAL3* gene in the exome of the female sibling. *TUBAL3* encodes the tubulin alpha chain-like 3 protein and occurs in exon 4, promoting the replacement of glutamate with lysine at position 174 and considered deleterious according to the Polymorphism Phenotyping v2 and Genomic Evolutionary Rate Profiling++ algorithms. In the male sibling data, we identified 3 rare X-chromosome variants (*ADGRG2*, *CNGA2*, and *FIGF*), as can be seen in Table 2.

As for the deleterious heterozygous variants inherited from one of the parents (Fig 3), we identified a total of 312 variants (S5 Table): in the male sibling, there were 125 exclusive variants (109 non-synonymous, 8 small indels, 5 stop-gains, 1 stop-loss, and 2 splicing sites); in the female sibling, there were 85 exclusive variants (81 nonsynonymous SNVs, 2 small indels,
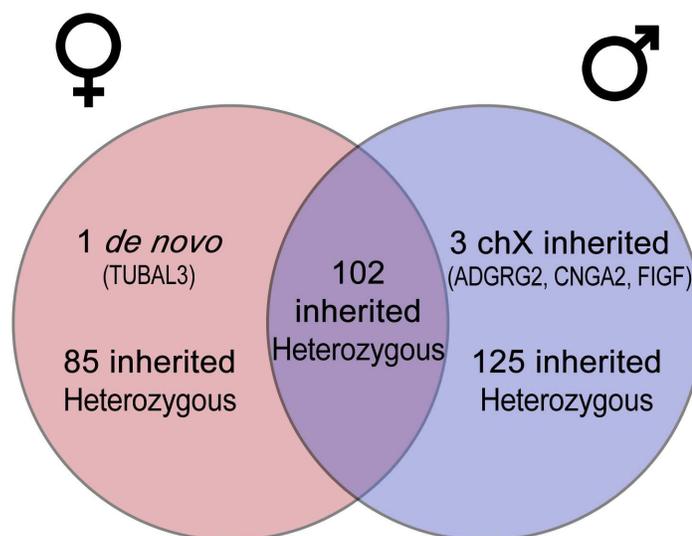

Fig 3. Summary diagram of the variations found in the two siblings.

doi:10.1371/journal.pone.0170386.g003

and 2 stop-gains); and the siblings shared 102 variants (96 nonsynonymous SNVs, 5 frameshift indels, and 1 non-frameshift indel).

Integrative network analysis of CNVs and exome variants

Using the combined brain-expressed gene list from the male sibling (70 genes present in the CNVs, 81 genes with variants shared by both siblings, and 105 genes with exclusive variants), we searched all connections using STRING and found enrichment of interactions (number of nodes: 236; number of edges: 129; average node degree: 1.09; clustering coefficient: 0.811; expected number of edges: 86 and protein-protein interaction enrichment *p*-value: 1.19e-05). Using the combined brain-expressed gene list from the female sibling (70 genes present in the CNVs, 81 genes with variants shared by both sibling; and 71 genes with exclusive variants), we searched all connections with STRING and again found enrichment of interactions (number of nodes: 204; number of edges: 101; average node degree: 0.99; clustering coefficient: 0.812; expected number of edges: 58; and protein-protein interaction enrichment p-value: 1.89e-07). Figs 4 and 5 show the biggest connected component for the male sibling and female sibling, respectively.

We selected genes connected in the biggest component of each sibling network analysis: 82 connecting genes for the male sibling (26 located in the 4p duplication) and 68 connecting genes for the female sibling (26 in the 4p duplication). The biggest connecting component in the networks of both siblings shared 46 genes. Therefore, 36 genes appeared to be connected exclusively in the network of the male, compared with 22 in the network of the female. After calculating the centrality measures for each network (S6 Table for the male and S7 Table for the female), we identified the broker and bridge genes (5% higher brokering score for the brokers, and 5% higher bridging scores for the bridges), as detailed in Table 3. There were 5 broker genes in the male (*CAD*, *FGFR3*, *PDE6B*, and *FYN*) and 3 in the female (*FLNC*, *ITGB4*, and *FYN*). There were 3 bridge genes in the biggest connected component of the network of the female: *CTBP1*, *HTT* (present in the translocation but connected only in the female network), and *SPTBN4* (exclusive, i.e. altered only in the female). *IDUA*, *NCKIPSD*, *FIGF*, and *NPR1* are bridge genes in the network of the male, the last 2 being exclusive. The *de novo* variant *TUBAL3* did not appear to be connected in the network of the female. Only one hemizygous variant gene (the bridge gene *FIGF*) was connected in the network of the male.

Discussion

Our major goal in this paper was to assess genetic findings in siblings that shared syndromic features and ID, a secondary goal being to explain the differences in the presentation of their psychiatric symptoms. According to Miles et al. [13,68], the male sibling could be classified as having complex autism (autism with ID and syndromic features). Both siblings presented irritability, agitation, and aggressive behavior, symptoms that are present in cases of ASD, as well as in those of ID only. Even among individuals with ASD, these symptoms are more common in those with ID [69].The male sibling had an attachment to repetitive behaviors, did not make eye contact, showed no capacity for joint attention, had no social interests, and had no communicative ability, all of which were symptoms that were not present in the female sibling. Their social impairments and alterations in social skills differed, and the female sibling had been described as presenting periods of inappropriate, hypersocial behavior and periods during which she could not stop talking and had sleep difficulties.

The duplication at 4p16.3 and deletion at 8p23.1, both of which were found in both siblings, are located in a recurrent translocation region in olfactory receptor gene clusters commonly correlated with non-allelic homologous recombination [61,70,71]. The FISH analysis showed

Fig 4. Male sibling STRING network analysis. Biggest component of the connecting brain-expressed genes present in the unbalanced translocation, altered by the shared and exclusive rare inherited SNV and indels found in the male sibling.

that these structural variations were inherited from a paternal balanced translocation. A terminal deletion within 8p is associated with hyperactivity, as well as with ID and impulsivity, whereas 4p terminal duplications are associated with microcephaly, motor retardation with seizures, and postnatal growth alterations [62–66]. Although translocation between chromosomes 4 and 8 is common, only a few cases of partial trisomy 4p and partial monosomy 8p have been reported [57–60,70]. The descriptions provided in the literature do not allow a typical phenotype of this type of dysmorphic syndrome to be assumed, because the distribution of dysmorphisms among patients is heterogeneous and the observed facial dysmorphisms are not specific.

Enrichment analysis confirmed that the gene list from the 4p16.3 region is enriched with genes associated with craniofacial and musculoskeletal abnormalities, which could explain the dysmorphic facial and growth features observed in the siblings evaluated here. Genes present in the 4p16.3 duplication and 8p23.3 deletion are involved in the cell cycle and cell differentiation,

Fig 5. Female sibling STRING network analysis. Biggest component of the connecting brain-expressed genes present in the unbalanced translocation, altered by the shared and exclusive rare inherited SNV and indels found in the female sibling.

Table 3. Comparison of broker and bridge genes in the biggest connected components of the networks of the two siblings.

Sib	Gene	Betweenness	Bridgeness	Brokering	Closeness	Clustering	Degree	Bridge or broker?*	Variation type	Exclusive?[†]
M	*CAD*	1816.83	15.939	0.167	0.379	0.033	14	Broker	Nonsyn. SNV	Yes
M	*FGFR3*	1232.00	74.089	0.074	0.333	0.000	6	Broker	4p Dup	No
M	*FYN*	1471.83	35.496	0.118	0.355	0.044	10	Broker	Nonsyn. SNV	No
M	*PDE6B*	346.67	13.867	0.074	0.250	0.000	6	Broker	4p Dup	No
M	*POLR3A*	907.83	45.305	0.085	0.345	0.143	8	Broker	Nonsyn. SNV	No
F	*FLNC*	224.52	15.273	0.078	0.236	0.133	6	Broker	Nonsyn. SNV	No
F	*FYN*	1155.52	14.664	0.176	0.288	0.015	12	Broker	Nonsyn. SNV	No
F	*ITGB4*	224.47	12.827	0.078	0.242	0.133	6	Broker	Nonsyn. SNV	Yes
M	*FIGF***	133.00	159.600	0.025	0.263	0.000	2	Bridge	Nonsyn. SNV	Yes
M	*IDUA*	159.50	239.250	0.025	0.265	0.000	2	Bridge	4p Dup	No
M	*NCKIPSD*	158.00	131.667	0.025	0.266	0.000	2	Bridge	Nonsyn. SNV	No
M	*NPR1*	227.17	121.156	0.025	0.272	0.333	3	Bridge	Nonsyn. SNV	Yes
F	*CTBP1*	396.83	264.556	0.030	0.223	0.000	2	Bridge	4p Dup	Yes
F	*HTT*	416.33	297.381	0.030	0.239	0.000	2	Bridge	4p Dup	Yes
F	*SPTBN4*	567.50	283.750	0.045	0.262	0.000	3	Bridge	Nonsyn. SNV	Yes

Sib: Sibling; M: Male; F: Female; Nonsyn.: nonsynonymous; Dup: duplication.

*bridging and brokering scores: 95th percentile.

[†]Connected in biggest network of only one sibling.

**Hemizygous variation.

doi:10.1371/journal.pone.0170386.t003

and some of these genes are likely to be more specifically related to the observed dysmorphic features in the siblings, such as *FGFR3* associated with chondrogenic and osteogenic events [72], and well known to be related to skeletal dysplasia [73]. Although the similarity of traits between the siblings could be explained by their unbalanced translocation, it is of note that cases in the literature sharing similar chromosomal abnormalities have not presented the same psychiatric symptoms.

In keeping with the hypotheses of network medicine proposed by Barabási et al. [53], the results from our network analysis of the siblings show that genes within the unbalanced translocation and the probable deleterious exome variants appeared to be connected in their biggest component. In addition, some genes that were disconnected only from the genes within the unbalanced translocation in our network analysis (S1 Fig) became connected when the exome results were included in the analysis. *MXD4*, *SH3BP2*, *ZFYVE28*, *LRPAP1* and *RGS12* appeared in both networks, whereas *LYAR* and *NOP14* became connected only in the network of the male sibling, *ADD1*, *ADRA2C*, *HTT*, and *CTBP1* being connected only in that of the female sibling.

According to Lim et al. [74], 5% of ASD cases can be affected by homozygous inherited rare variations, with loss of function, or hemizygous X-chromosome variations in males, and that deleterious mutations in X chromosome are more common in cases than in controls. In the present study, *FIGF* is the only gene that was connected in the biggest component of the network of the male sibling. In addition to being related to cell proliferation [74,75], probably contributing to the syndromic phenotype seen in our male patient, *FIGF* expression is modulated by activation of the Wnt signaling pathway [75] and is also related to dopaminergic neuron differentiation [76]. The Wnt pathway is one of the main pathways that have been associated with exome variants in ASD [77].

Pavlopoulos et al. [78] found that different types of biological networks can be modeled using graph theory, and that some of the nodes in a network may have topological properties

that stand out in comparison with the others. The *CAD* gene, a variant of which was observed only in the male sibling evaluated here, is an important and highly connected broker gene. *CAD* encodes a protein associated with the enzymatic activities of the first three enzymes in the pyrimidine biosynthesis pathway and is regulated by the mitogen-activated protein kinase cascade [79]. In ASD patients, *CAD* has been found to be altered by 1 deleterious SNV and present in 4 CNVs [4]; it is also one of the genes identified by Samocha et al. [80] as being enriched with *de novo* loss-of-function mutations in ASD patients and not in individuals without ASD.

Exclusively in our female patient, a variant of the *ITGB4* gene was a broker in the network. *ITGB4* encodes an integrin-associated transmembrane glycoprotein receptor, expressed on epithelial cells, and in neural cells such as astrocytes, Schwann cells, neurons, and neural stem cells [81]. In the central nervous system, beta4 integrin is expressed primarily in arterioles. Although a lack of endothelial beta4 integrin has no effect on vascular development, its absence in the hypoxic central nervous system disrupts transforming growth factor-beta–mediated signaling and results in defective arteriolar remodeling [82].

Among the exclusive variant genes within the network of the female sibling evaluated in the present study, *SPTBN4* functions as a bridge. It encodes the myelinated neuron-rich cytoskeleton protein beta-spectrin [83], which is important to the clustering of voltage-gated sodium channel axon initial segments and the nodes of Ranvier [84], being required for *ANK3* function [85]. Although not a broker or a bridge, *ANK3* was also connected in the network of the female sibling. *ANK3* encodes the protein ankyrin-G, also located at the nodes of Ranvier and the axon initial segment, a protein responsible for the generation of action potentials [86]. *ANK3* has been associated with autism susceptibility [87], as well as with increased risk for and modulation of working memory deficits in bipolar disorder and schizophrenia [88–90].

Another important female exclusive variant gene, *RELN*, encodes reelin, an extracellular matrix protein that activates a signaling pathway for neuronal migration and synaptic plasticity during brain development, as well as appearing to be involved in schizophrenia, autism, bipolar disorder, and other neuropsychiatric disorders [91–95]. In addition, a recent genome-wide association study of schizophrenia revealed that the strongest association was with a marker within *RELN*, and only in women [96]. Furthermore, Goes et al. [91] found evidence that genetic variation in *RELN* is associated with susceptibility to bipolar disorder, particularly in females.

Some of the altered genes shared by the two siblings evaluated in the present study functioned as hubs in only one of the siblings, differentially contributing to the stability of the gene network and probably the phenotypes. For instance, *IDUA* (present in duplication 4p) and *NCKIPSD* (present in a shared nonsynonymous SNV) were bridge genes only in the network of the male sibling.

The only *de novo* exome variant, *TUBAL3* (found in the female sibling), was not connected in the network. According to Veltman & Brunner [97], approximately 74 novel SNVs arise per genome per generation. Given the fact that deleterious *de novo* variants increase disease risk, various computational methods have been proposed to predict impact and gene prioritization in human disease [98–100]. However, even though we could not exclude *de novo TUBAL3* variation from our analysis, the integrative approach we used showed that it is not important to the understanding of the phenotypic differences between the two siblings.

Adding exome genes to the network of each sibling showed that, as stated by Andrews et al. [30], the more connected the genes were, the more severe was the clinical presentation. The male sibling had a network that was more well connected, and, with the inclusion of the exome, more genes from the translocation were included. Accordingly, his CARS score was indicative of greater severity and genes with deleterious variations exclusive to his network are clearly

related to ASD. The female had some psychiatric symptoms such as excess sociability, periods of sleep difficulties, and periods of talkativeness, a presentation similar to that of a mood disorder. Notably, the variants found only in the female sibling, although having been associated with autism, have also been shown to be strongly associated with other psychotic disorders, which probably explains her broader phenotype.

Conclusions

One of the limitations of the current study is the sample size, which was small for an integrative analysis, given that we accessed only the exome data of the two affected siblings. Therefore, we chose to leave the inherited CNV out of the analysis and focus on the variations that were the most likely to have deleterious effects, such as *de novo* CNVs, as well as *de novo* and rare inherited nonsynonymous SNVs. Nevertheless, with this analysis of only one family, we were able to show the importance of exploring the genotype-phenotype variability beyond the effect of one SNV or CNV.

Our results show that the translocation shared between the siblings could explain their similar syndromic phenotype, such as their overgrowth, macrocephaly, and ID. However, our network analysis showed that exome data aggregate genes to those already connected from their shared CNV, which are important to the robustness of the network and possibly contribute to the understanding of the broader spectrum of psychiatric symptoms.

Supporting Information

S1 Fig. MLPA results. (A) female sibling, (B) male sibling, (C) father, and (D) mother. In A and B, the dot over the 1.3 line of the peak ratio value represents the 4p16.3 duplication, and the dot under the 0.7 line of the peak ratio value represents the 8p23.3 deletion.
(TIF)

S2 Fig. STRING network of genes present in the chromosomal alteration der(8)t(4;8) (p16.3;p23.1) in the siblings. The biggest connected component is composed of 17 proteins with 20 interactions (i.e., 24.6%, 17 of the 69 brain-expressed genes), 6 connected components composed of 2 proteins with 1 interaction and 1 connected component with 3 proteins and 2 interactions, resulting in 15 proteins (i.e., 21.7%, 14 of the 69 brain-expressed genes) and 8 interactions.
(TIF)

S1 Table. Summary of the medical history and clinical evaluation of the probands.
(DOCX)

S2 Table. List of *de novo* CNVs found with 180K aCGH in the siblings.
(DOCX)

S3 Table. Results of WebGestalt disease enrichment analysis with at least 4 genes from the 4p16.3 duplication gene list (69 brain-expressed genes)
(DOCX)

S4 Table. Summary of exome sequencing data quality of each member of the family.
(DOCX)

S5 Table. List of deleterious variants after filtering in the male and female sibling exome data.
(XLSX)

S6 Table. Gene list with centrality measures of the biggest connected component of the network of the male sibling.
(XLSX)

S7 Table. Gene list with centrality measures of the biggest connected component of the network of the female sibling.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: VNSR HB.

Data curation: VNSR ACT.

Formal analysis: VNSR ACF ACT SNS RAF.

Funding acquisition: VNSR HB.

Investigation: VNSR ACF BCGL RAF.

Methodology: VNSR HB.

Project administration: VNSR HB.

Resources: HB ACVK CR NCL MRPB.

Software: JPK ACT SNS.

Supervision: HB.

Validation: BCGL ACVK CR NCL MRPB.

Visualization: VNSR ACT RAF.

Writing – original draft: VNSR.

Writing – review & editing: VNSR HB ACT JPK ACVK CR MRPB.

References

1. Geschwind DH, Levitt P. Autism spectrum disorders: developmental disconnection syndromes. Curr Opin Neurobiol. 2007; 17: 103–111. doi: 10.1016/j.conb.2007.01.009 PMID: 17275283

2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th Edition) [Internet]. 5th ed. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. Washington D.C: American Psychiatric Association; 2013.

3. Fombonne E. The Prevalence of Autism. JAMA. 2003; 289: 87–9. PMID: 12503982

4. Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, Witherspoon K, et al. Excess of rare, inherited truncating mutations in autism. Nat Genet. Nature Publishing Group; 2015; 47: 582–588. doi: 10.1038/ng.3303 PMID: 25961944

5. Ronemus M, Iossifov I, Levy D, Wigler M. The role of de novo mutations in the genetics of autism spectrum disorders. Nat Rev Genet. Nature Publishing Group; 2014; 15: 133–41. doi: 10.1038/nrg3585 PMID: 24430941

6. Devlin B, Scherer SW. Genetic architecture in autism spectrum disorder. Curr Opin Genet Dev. Elsevier Ltd; 2012; 22: 229–37. doi: 10.1016/j.gde.2012.03.002 PMID: 22463983

7. State MW, Levitt P. The conundrums of understanding genetic risks for autism spectrum disorders. Nat Neurosci. Nature Publishing Group; 2011; 14: 1499–506. doi: 10.1038/nn.2924 PMID: 22037497

8. Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. Nat Rev Genet. 2008; 9: 341–55. doi: 10.1038/nrg2346 PMID: 18414403

9. Benvenuto A, Manzi B, Alessandrelli R, Galasso C, Curatolo P. Recent Advances in the Pathogenesis of Syndromic Autisms. Int J Pediatr. 2009; 2009: 1–9.

10. Moss J, Howlin P. Autism spectrum disorders in genetic syndromes: implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. J Intellect Disabil Res. 2009; 53: 852–73. doi: 10.1111/j.1365-2788.2009.01197.x PMID: 19708861

11. Zafeiriou DI, Ververi A, Dafoulis V, Kalyva E, Vargiami E. Autism Spectrum Disorders: The Quest for Genetic Syndromes. Am J Med Genet B Neuropsychiatr Genet. 2013;

12. Persico AM, Napolioni V. Autism genetics. Behav Brain Res. Elsevier B.V.; 2013; 251: 95–112. doi: 10.1016/j.bbr.2013.06.012 PMID: 23769996

13. Miles JH, Takahashi TN, Bagby S, Sahota PK, Vaslow DF, Wang CH, et al. Essential versus complex autism: definition of fundamental prognostic subtypes. Am J Med Genet A. 2005; 135: 171–80. doi: 10.1002/ajmg.a.30590 PMID: 15887228

14. Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey a J, et al. Common genetic variants, acting additively, are a major source of risk for autism. Mol Autism. Molecular Autism; 2012; 3: 9. doi: 10.1186/2040-2392-3-9 PMID: 23067556

15. Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, et al. Genetic heritability and shared environmental factors among twin pairs with autism. Arch Gen Psychiatry. 2011; 68: 1095–102. doi: 10.1001/archgenpsychiatry.2011.76 PMID: 21727249

16. Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, et al. Individual common variants exert weak effects on the risk for autism spectrum disorderspi. Hum Mol Genet. 2012; 21: 4781–92. doi: 10.1093/hmg/dds301 PMID: 22843504

17. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. Science. 2007; 316: 445–9. doi: 10.1126/science.1138659 PMID: 17363630

18. Marshall C, Noor A, Vincent J. Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet. 2008; 82: 477–488. doi: 10.1016/j.ajhg.2007.12.009 PMID: 18252227

19. Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature. 2009; 459: 569–73. doi: 10.1038/nature07953 PMID: 19404257

20. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, et al. Functional impact of global rare copy number variation in autism spectrum disorders. Nature. 2010; 466: 368–72. doi: 10.1038/nature09146 PMID: 20531469

21. Levy D, Ronemus M, Yamrom B, Lee Y, Leotta A, Kendall J, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron. Elsevier Inc.; 2011; 70: 886–97. doi: 10.1016/j.neuron.2011.05.015 PMID: 21658582

22. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron. 2011; 70: 863–85. doi: 10.1016/j.neuron.2011.05.002 PMID: 21658581

23. Jacquemont M, Sanlaville D, Redon R, Raoul O, Cormier-Daire V, Lyonnet S, et al. Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. J Med Genet. 2006; 43: 843–849. doi: 10.1136/jmg.2006.043166 PMID: 16840569

24. Bremer A, Giacobini M, Eriksson M, Gustavsson P, Nordin V, Fernell E, et al. Copy number variation characteristics in subpopulations of patients with autism spectrum disorders. Am J Med Genet B Neuropsychiatr Genet. 2011; 156: 115–24. doi: 10.1002/ajmg.b.31142 PMID: 21302340

25. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell. Elsevier Inc.; 2012; 148: 1223–41. doi: 10.1016/j.cell.2012.02.039 PMID: 22424231

26. Cooper GM, Coe BP, Girirajan S, Rosenfeld J a, Vu TH, Baker C, et al. A copy number variation morbidity map of developmental delay. Nat Genet. Nature Publishing Group; 2011; 43: 838–46. doi: 10.1038/ng.909 PMID: 21841781

27. Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, et al. Population analysis of large copy number variants and hotspots of human genetic disease. Am J Hum Genet. 2009; 84: 148–61. doi: 10.1016/j.ajhg.2008.12.014 PMID: 19166990

28. Grayton HM, Fernandes C, Rujescu D, Collier D a. Copy number variations in neurodevelopmental disorders. Prog Neurobiol. Elsevier Ltd; 2012; 99: 81–91. doi: 10.1016/j.pneurobio.2012.07.005 PMID: 22813947

29. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. Neuron. Elsevier Inc.; 2011; 70: 898–907. doi: 10.1016/j.neuron.2011.05.021 PMID: 21658583

30. Andrews T, Honti F, Pfundt R, Leeuw N de, Hehir-Kwa J, Silfhout AV, et al. The clustering of functionally related genes contributes to CNV-mediated disease. Genome Res. 2015; 25: 802–813. doi: 10.1101/gr.184325.114 PMID: 25887030

31. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet. Nature Publishing Group; 2011; 43: 585–9. doi: 10.1038/ng.835 PMID: 21572417

32. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science (80-). 2012; 338: 1619–22. doi: 10.1126/science.1227764 PMID: 23160955

33. Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, et al. De novo gene disruptions in children on the autistic spectrum. Neuron. Elsevier Inc.; 2012; 74: 285–99. doi: 10.1016/j.neuron.2012.04.009 PMID: 22542183

34. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey a J, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature. 2012; 485: 237–41. doi: 10.1038/nature10945 PMID: 22495306

35. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. Nature. Nature Publishing Group; 2014; 515: 216–221. doi: 10.1038/nature13908 PMID: 25363768

36. Yuen RKC, Thiruvahindrapuram B, Merico D, Walker S, Tammimies K, Hoang N, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. Nat Med. 2015; 21: 185–191. doi: 10.1038/nm.3792 PMID: 25621899

37. Manning M, Hudgins L. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. Genet Med. 2010; 12: 742–5. doi: 10.1097/GIM.0b013e3181f8baad PMID: 20962661

38. Malan V, Chevallier S, Soler G, Coubes C, Lacombe D, Pasquier L, et al. Array-based comparative genomic hybridization identifies a high frequency of copy number variations in patients with syndromic overgrowth. Eur J Hum Genet. 2010; 18: 227–32. doi: 10.1038/ejhg.2009.162 PMID: 19844265

39. Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. Nat Rev Genet. Nature Publishing Group; 2015; 16: 441–458. doi: 10.1038/nrg3934 PMID: 26149713

40. Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. Science (80-). 2015; 349: 1489–1494. doi: 10.1126/science.aaa8954 PMID: 26404826

41. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics. Oxford University Press; 2010; 26: 589–595. Available: http://www.ncbi.nlm.nih.gov/pubmed/20080505 doi: 10.1093/bioinformatics/btp698 PMID: 20080505

42. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297–1303. doi: 10.1101/gr.107524.110 PMID: 20644199

43. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43: 491–8. doi: 10.1038/ng.806 PMID: 21478889

44. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Current Protocols in Bioinformatics. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2013. p. 11.10.1–11.10.33.

45. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–75. doi: 10.1086/519795 PMID: 17701901

46. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. Oxford University Press; 2009; 25: 1754–1760. Available: http://www.pubmedcentral.nih.gov/

articlerender.fcgi?artid=2705234&tool=pmcentrez&rendertype=abstract doi: 10.1093/bioinformatics/btp324 PMID: 19451168

47. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38: e164. doi: 10.1093/nar/gkq603 PMID: 20601685

48. Davydov E V., Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010; 6.

49. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: A Database of Human Non-synonymous SNVs and Their Functional Predictions and Annotations. Hum Mutat. 2013;

50. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res. 2013; 1–7.

51. Miller J a, Ding S-L, Sunkin SM, Smith K a, Ng L, Szafer A, et al. Transcriptional landscape of the prenatal human brain. Nature. 2014; 508: 199–206. doi: 10.1038/nature13185 PMID: 24695229

52. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, et al. STRING 8—A global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 2009; 37: 412–416.

53. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011; 12: 56–68. doi: 10.1038/nrg2918 PMID: 21164525

54. Cho Y-R, Zhang A. Identification of functional hubs and modules by converting interactome networks into hierarchical ordering of proteins. BMC Bioinformatics. 2010; 11 Suppl 3: S3.

55. Xue-Qi C, Fu-Xin R, Hua-Wei S, Zi-Ke Z, Tao Z. Bridgeness: a local index on edge significance in maintaining global connectivity. J Stat Mech Theory Exp. 2010; 2010: P10011.

56. Cai JJ, Borenstein E, Petrov DA. Broker Genes in Human Disease. Genome Biol Evol. Oxford University Press; 2010; 2: 815–825. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2988523&tool=pmcentrez&rendertype=abstract doi: 10.1093/gbe/evq064 PMID: 20937604

57. Goggin M, O'Keefe M, Hughes H. Two cases of coloboma associated with unbalanced translocations. Br J Ophthalmol. 1993; 77: 122–123. PMID: 8435414

58. Skrlec I, Wagner J, Puseljic S, Heffer M, Stipoljev F. De Novo Case of a Partial Trisomy 4p and a Partial Monosomy 8p. Coll Antropol. 2014; 38: 319–323. PMID: 24851635

59. Mau UA, Backsch C, Schaudt H, Trefz FK, Kaiser P. Three-year-old girl with partial trisomy 4p and partial monosomy 8p with resemblance to Brachmann-de Lange syndrome—Another locus for Brachmann-de Lange syndrome on 4p? Am J Med Genet. 2000; 91: 180–184. PMID: 10756338

60. Partington MW, Fagan K, Soubjaki V, Turner G. Translocations involving 4p16.3 in three families: deletion causing the Pitt-Rogers-Danks syndrome and duplication resulting in a new overgrowth syndrome. J Med Genet. 1997; 34: 719–728. PMID: 9321756

61. Giglio S, Calvari V, Gregato G, Gimelli G, Camanini S, Giorda R, et al. Heterozygous submicroscopic inversions involving olfactory receptor-gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. Am J Hum Genet. 2002; 71: 276–285. doi: 10.1086/341610 PMID: 12058347

62. Bosse K, Eggermann T, Van Der Ven K, Raff R, Engels H, Schwanitz G. Unbalanced translocation 8;Y (45,X,dic(Y;8)(q11.23;p23.1)): Case report and review of terminal 8p deletions. Ann Genet. 2004; 47: 191–197. doi: 10.1016/j.anngen.2004.02.004 PMID: 15183753

63. Cyr AB, Nimmakayalu M, Longmuir SQ, Patil SR, Keppler-Noreuil KM, Shchelochkov O a. A novel 4p16.3 microduplication distal to WHSC1 and WHSC2 characterized by oligonucleotide array with new phenotypic features. Am J Med Genet Part A. 2011; 155: 2224–2228.

64. Burnside RD, Pappas JG, Sacharow S, Applegate C, Hamosh A, Gadi IK, et al. Three cases of isolated terminal deletion of chromosome 8p without heart defects presenting with a mild phenotype. Am J Med Genet Part A. 2013; 161: 822–828.

65. Hemmat M, Hemmat O, Anguiano A, Boyar FZ, El Naggar M, Wang J-C, et al. Genotype-phenotype analysis of recombinant chromosome 4 syndrome: an array-CGH study and literature review. Mol Cytogenet. 2013; 6: 17. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3648413&tool=pmcentrez&rendertype=abstract doi: 10.1186/1755-8166-6-17 PMID: 23639048

66. Schönewolf-Greulich B, Ravn K, Hamborg-Petersen B, Brøndum-Nielsen K, Tümer Z. Segregation of a 4p16.3 duplication with a characteristic appearance, macrocephaly, speech delay and mild intellectual disability in a 3-generation family. Am J Med Genet Part A. 2013; 161: 2358–2362.

67. Altmann A, Weber P, Bader D, Preuss M, Binder EB, Müller-Myhsok B. A beginners guide to SNP calling from high-throughput DNA-sequencing data. Hum Genet. 2012; 131: 1541–54. doi: 10.1007/s00439-012-1213-z PMID: 22886560

68. Miles JH, Takahashi TN, Hong J, Munden N, Flournoy N, Braddock SR, et al. Development and validation of a measure of dysmorphology: useful for autism subgroup classification. Am J Med Genet A. 2008; 146A: 1101–16. doi: 10.1002/ajmg.a.32244 PMID: 18383511

69. Wing L, Gould J. Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. J Autism Dev Disord. 1979; 9: 11–29. Available: http://www.ncbi.nlm.nih.gov/pubmed/155684 PMID: 155684

70. Maas NMC, Van Vooren S, Hannes F, Van Buggenhout G, Mysliwiec M, Moreau Y, et al. The t(4;8) is mediated by homologous recombination between olfactory receptor gene clusters, but other 4p16 translocations occur at random. Genet Couns. 2007; 18: 357–65. Available: http://www.ncbi.nlm.nih.gov/pubmed/18286816 PMID: 18286816

71. Ou Z, Stankiewicz P, Xia Z, Breman AM, Dawson B, Wiszniewska J, et al. Observation and prediction of recurrent human translocations mediated by NAHR between nonhomologous chromosomes. Genome Res. 2011; 21: 33–46. doi: 10.1101/gr.111609.110 PMID: 21205869

72. Delezoide AL, Benoist-Lasselin C, Legeai-Mallet L, Le Merrer M, Munnich A, Vekemans M, et al. Spatio-temporal expression of FGFR 1, 2 and 3 genes during human embryo-fetal ossification. Mech Dev. 1998; 77: 19–30. PMID: 9784595

73. Krejci P. The paradox of FGFR3 signaling in skeletal dysplasia: Why chondrocytes growth arrest while other cells over proliferate. Mutation Research—Reviews in Mutation Research. 2014. pp. 40–48.

74. Lim ET, Raychaudhuri S, Sanders SJ, Stevens C, Sabo A, Macarthur DG, et al. Rare complete knock-outs in humans: population distribution and significant role in autism spectrum disorders. Neuron. Elsevier Inc.; 2013; 77: 235–42. doi: 10.1016/j.neuron.2012.12.029 PMID: 23352160

75. Orlandini M, Semboloni S, Oliviero S. Beta-catenin inversely regulates vascular endothelial growth factor-D mRNA stability. J Biol Chem. 2003; 278: 44650–6. doi: 10.1074/jbc.M304255200 PMID: 12920128

76. Schwartz CM, Tavakoli T, Jamias C, Park S-S, Maudsley S, Martin B, et al. Stromal factors SDF1α, sFRP1, and VEGFD induce dopaminergic neuron differentiation of human pluripotent stem cells. J Neurosci Res. 2012; 90: 1367–81. doi: 10.1002/jnr.23064 PMID: 22535492

77. Krumm N, O'Roak BJ, Shendure J, Eichler EE. A de novo convergence of autism genetics and molecular neuroscience. Trends Neurosci. Elsevier Ltd; 2013; 37: 1–11

78. Pavlopoulos G a, Secrier M, Moschopoulos CN, Soldatos TG, Kossida S, Aerts J, et al. Using graph theory to analyze biological networks. BioData Min. 2011; 4: 10. doi: 10.1186/1756-0381-4-10 PMID: 21527005

79. Graves LM, Guy HI, Kozlowski P, Huang M, Lazarowski E, Pope RM, et al. Regulation of carbamoyl phosphate synthetase by MAP kinase. Nature. 2000; 403: 328–332. doi: 10.1038/35002111 PMID: 10659854

80. Samocha KE, Robinson EB, Sanders SJ, Stevens C, Sabo A, McGrath LM, et al. A framework for the interpretation of de novo mutation in human disease. Nat Genet. Nature Publishing Group; 2014; advance on: 944–950. doi: 10.1038/ng.3050 PMID: 25086666

81. Su L, Lv X, Miao J. Integrin β4 in Neural Cells. NeuroMolecular Med. 2008; 10: 316–321. doi: 10.1007/s12017-008-8042-1 PMID: 18516507

82. Welser-Alves J V., Boroujerdi A, Tigges U, Wrabetz L, Feltri ML, Milner R. Endothelial β4 integrin is predominantly expressed in arterioles, where it promotes vascular remodeling in the hypoxic brain. Arterioscler Thromb Vasc Biol. 2013; 33: 943–53. doi: 10.1161/ATVBAHA.112.300566 PMID: 23471230

83. Berghs S, Aggujaro D, Dirkx R, Maksimova E, Stabach P, Hermel JM, et al. betaIV spectrin, a new spectrin localized at axon initial segments and nodes of ranvier in the central and peripheral nervous system. J Cell Biol. 2000; 151: 985–1002. PMID: 11086001

84. Lacas-Gervais S, Guo J, Strenzke N, Scarfone E, Kolpe M, Jahkel M, et al. βIVΣ1 spectrin stabilizes the nodes of Ranvier and axon initial segments. J Cell Biol. 2004; 166: 983–990. doi: 10.1083/jcb.200408007 PMID: 15381686

85. Smith KR, Kopeikina KJ, Fawcett-Patel JM, Leaderbrand K, Gao R, Sch??rmann B, et al. Psychiatric Risk Factor ANK3/Ankyrin-G Nanodomains Regulate the Structure and Function of Glutamatergic Synapses. Neuron. 2014; 84: 399–415. doi: 10.1016/j.neuron.2014.10.010 PMID: 25374361

86. Iqbal Z, Vandeweyer G, van der Voet M, Waryah AM, Zahoor MY, Besseling J a, et al. Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders. Hum Mol Genet. 2013; 22: 1960–70. doi: 10.1093/hmg/ddt043 PMID: 23390136

87. Bi C, Wu J, Jiang T, Liu Q, Cai W, Yu P, et al. Mutations of ANK3 identified by exome sequencing are associated with Autism susceptibility. Hum Mutat. 2012; 33: 1–4.

88. Delvecchio G, Dima D, Frangou S. The effect of ANK3 bipolar-risk polymorphisms on the working memory circuitry differs between loci and according to risk-status for bipolar disorder. Am J Med Genet Part B Neuropsychiatr Genet. 2015; 168: 188–196.

89. Zhang C, Cai J, Zhang J, Li Z, Guo Z, Zhang X, et al. Genetic modulation of working memory deficits by ankyrin 3 gene in schizophrenia. Prog Neuro-Psychopharmacology Biol Psychiatry. Elsevier Inc.; 2014; 50: 110–115.

90. Hori H, Yamamoto N, Teraishi T, Ota M, Fujii T, Sasayama D, et al. Cognitive effects of the ANK3 risk variants in patients with bipolar disorder and healthy individuals. J Affect Disord. Elsevier; 2014; 158: 90–96. doi: 10.1016/j.jad.2014.02.008 PMID: 24655771

91. Goes FS, Willour VL, Zandi PP, Belmonte PL, MacKinnon DF, Mondimore FM, et al. Sex-specific association of the reelin gene with bipolar disorder. Am J Med Genet Part B Neuropsychiatr Genet. 2010; 153: 549–553.

92. Chang BS, Duzcan F, Kim S, Cinbis M, Aggarwal A, Apse KA, et al. The role of RELN in lissencephaly and neuropsychiatric disease. Am J Med Genet Part B Neuropsychiatr Genet. 2007; 144: 58–63.

93. Fatemi SH, Earle J a, McMenomy T. Reduction in Reelin immunoreactivity in hippocampus of subjects with schizophrenia, bipolar disorder and major depression. Mol Psychiatry. 2000; 5: 654–663, 571. PMID: 11126396

94. Skaar D a, Shao Y, Haines JL, Stenger JE, Jaworski J, Martin ER, et al. Analysis of the RELN gene as a genetic risk factor for autism. Mol Psychiatry. 2005; 10: 563–571. doi: 10.1038/sj.mp.4001614 PMID: 15558079

95. Haas C a, Dudeck O, Kirsch M, Huszka C, Kann G, Pollak S, et al. Role for reelin in the development of granule cell dispersion in temporal lobe epilepsy. J Neurosci. 2002; 22: 5797–5802. PMID: 12122039

96. Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, Craddock NJ, et al. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. PLoS Genet. 2008; 4.

97. Veltman J a, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. Nature Publishing Group; 2012; 13: 565–75. doi: 10.1038/nrg3241 PMID: 22805709

98. Itan Y, Casanova J-L. Can the impact of human genetic variations be predicted? Proc Natl Acad Sci U S A. 2015; 112: 11426–7. doi: 10.1073/pnas.1515057112 PMID: 26351682

99. Miosge LA, Field MA, Sontani Y, Cho V, Johnson S, Palkova A, et al. Comparison of predicted and actual consequences of missense mutations. Proc Natl Acad Sci U S A. 2015; 112: E5189–98. doi: 10.1073/pnas.1511585112 PMID: 26269570

100. Wu J, Li Y, Jiang R. Integrating multiple genomic data to predict disease-causing nonsynonymous single nucleotide variants in exome sequencing studies. PLoS Genet. 2014; 10: e1004237. doi: 10.1371/journal.pgen.1004237 PMID: 24651380