Copy Number Variation Analysis of 100 Twin Pairs Enriched for Neurodevelopmental Disorders

Sofia Stamouli,[1,2] Britt-Marie Anderlid,[3,4] Charlotte Willfors,[1,2] Bhooma Thiruvahindrapuram,[5] John Wei,[5] Steve Berggren,[1,2] Ann Nordgren,[3,4] Stephen W. Scherer,[5,6] Paul Lichtenstein,[7] Kristiina Tammimies,[1,2,*] and Sven Bölte[1,2,*]

[1]*Center of Neurodevelopmental Disorders (KIND), Department of Women's and Children's Health, Karolinska Institutet, Solna, Sweden*
[2]*Center for Psychiatry Research, Stockholm County Council, Stockholm, Sweden*
[3]*Department of Molecular Medicine and Surgery, Center of Molecular Medicine, Karolinska Institutet, Solna, Sweden*
[4]*Department of Clinical Genetics, Karolinska University Hospital, Solna, Sweden*
[5]*The Centre for Applied Genomics, Program in Genetics and Genome Biology, the Hospital for Sick Children, Toronto, Ontario, Canada*
[6]*Department of Molecular Genetics and the McLaughlin Centre, University of Toronto, Toronto, Canada*
[7]*Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden*

Hundreds of penetrant risk loci have been identified across different neurodevelopmental disorders (NDDs), and these often involve rare (<1% frequency) copy number variations (CNVs), which can involve one or more genes. Monozygotic (MZ) twin pairs are long thought to share 100% of their genomic information. However, genetic differences in the form of postzygotic somatic variants have been reported recently both in typically developing (TD) and in clinically discordant MZ pairs. We sought to investigate the contribution of rare CNVs in 100 twin pairs enriched for NDD phenotypes with a particular focus on postzygotic CNVs in MZ pairs discordant for autism spectrum disorder (ASD) using the Illumina Infinium PsychArray. In our sample, no postzygotic de novo CNVs were found in 55 MZ twin pairs, including the 13 pairs discordant for ASD. We did detect a higher rate of CNVs overlapping genes involved in disorders of the nervous system, such as a rare deletion affecting *HNRNPU*, in MZ pairs discordant and concordant for ASD in comparison with TD pairs ($p = .02$). Our results are in concordance with earlier findings that postzygotic de novo CNV events are typically rare in genomic DNA derived from saliva or blood, and suggests that the discordance of NDDs in our sample of twins is not explained by discordant CNVs. Still, studies investigating postzygotic variation in MZ discordant twins using DNA from different tissues and single cells and higher resolution genomics are needed in the future.

❚ **Keywords:** monozygotic twins, neurodevelopmental disorders, autism spectrum disorder, copy number variant, postzygotic variants, discordant twin pairs, *HNRNPU*

Neurodevelopmental disorders (NDDs), such as autism spectrum disorder (ASD) and attention-deficit hyperactivity disorder (ADHD), are a group of early onset conditions affecting approximately 10–15% of the population (Boyle et al., 2011). NDDs are characterized by alterations in developmental trajectories causing impairments in personal, social, academic, and occupational functioning (American Psychiatric Association, 2013). There is substantial co-occurrence of different NDDs, suggesting a common etiology for the disorders (Thapar et al., 2017). NDDs have a strong genetic component, and the familial risk is seen between different NDDs and psychiatric disorders (Faraone et al., 2017; Ghirardi et al., 2017; Sullivan et al., 2012). The etiological overlap has also been observed in molecular genetic studies. Many of the same risk genes and variants, identified through genome analyses of copy number variants (CNVs) and sequence level variants are found across NDDs (Torres et al., 2016). These include large deletions and duplications of 16p11.2 and 22q11.2, duplication of 7q11.23 and 15q11-13 as well as smaller CNVs

1

affecting specific genes such as *NRXN1* (OMIM#600565) and *SHANKs* (Kirov, 2015). Additionally, new NDD risk genes have been implicated using sequencing-based studies, mainly through identification of de novo mutations (DNMs) (De Rubeis et al., 2014; Iossifov et al., 2014; Tammimies et al., 2015; Yuen et al., 2017).

In addition to germline DNMs, somatic DNMs contribute to human disorders (Acuna-Hidalgo et al., 2016), including NDDs (Lim et al., 2017; Yuen et al., 2016). It has been hypothesized that postzygotic DNMs could cause discordant phenotypes in monozygotic (MZ) twin pairs that are otherwise genetically identical. However, the majority of studies testing this hypothesis have reported no or limited evidence for the postzygotic de novo CNVs in MZ discordant twin pairs for psychiatric disorders (Bloom et al., 2013; Castellani et al., 2014) and neurodevelopmental problems (Abdellaoui et al., 2015; Ehli et al., 2012; Laplana et al., 2014). In rare occasions, postzygotic CNVs can also be found in MZ pairs without discordant phenotypes (Abdellaoui et al., 2015; Magnusson et al., 2016; McRae et al., 2015).

Here, we analyze the presence of postzygotic CNVs in a Swedish twin sample from the Roots of Autism and ADHD Study in Sweden (RATSS; Bölte et al., 2014). The sample is enriched for twin pairs discordant for NDDs, including 13 MZ twin pairs discordant for ASD. Beyond rare case reports (Laplana et al., 2014), there have not been studies focused on ASD-discordant MZ pairs and the contribution of postzygotic CNVs. Additionally, we characterized the rate of shared rare CNVs in genes implicated earlier in brain development among the twin pairs by grouping them as concordant or discordant for NDD and ASD and comparing with typically developing (TD) twin pairs.

Materials and Methods

Participants and Clinical Assessments
All the participants included in this study are part of the RATSS (Bölte et al., 2014). The recruitment and procedure for the sample collection have been described in detail elsewhere (Bölte et al., 2014; Willfors et al., 2017). Overview of the study design and samples used for the analyses are shown in Figure 1. Within the RATSS, twins undergo detailed diagnostic and psychological assessments, including standardized diagnostic tools for NDDs: a psychosocial and anamnestic interview, the Autism Diagnostic Interview-Revised (ADI-R), the Autism Diagnostic Observation Schedule Second Edition (ADOS-2), the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS), or the Diagnostic Interview for ADHD in Adults (DIVA). Additionally, the Wechsler Intelligence Scales for Children or Adults, Fourth Edition (WISC-IV) or the Leiter-revised scales in combination with the Peabody Picture Vocabulary Test, Third Edition, and the parent-rated Adaptive Behavior Assessment Scale, 2nd Edition (ABAS-

2), were used to evaluate adaptive, cognitive, and verbal abilities and for diagnostic assessments. Also, a medical and family history of the twins was obtained by parental interviews and by collecting existing medical records.

The given NDD diagnoses are based on the DSM-5 criteria (American Psychiatric Association, 2013) by incorporating the different assessments followed by a consensus among three experienced clinicians within the RATSS study. Using the Child and Adolescent Twin Study in Sweden cohort, which is also a recruitment cohort for RATSS, it has been demonstrated that in the majority of ASD-discordant MZ pairs, the co-twin has another NDD diagnosis (Lundstrom et al., 2015). Therefore, in this study, we used two strategies to categorize the twins, based on their NDD diagnoses. First, we only considered whether any of the twins had ASD diagnosis and categorized the pair as discordant or concordant for ASD. Thereafter, we considered all NDD diagnoses and did an additional categorization as NDD discordant (affected twin has any NDD diagnosis and co-twin TD) and NDD concordant (both twins have at least one NDD diagnosis). If no NDD diagnoses were present in any of the twins, the pair was categorized as TD.

Figure 1 shows the overview of the study details and number of samples within our analysis. A total of 208 twins from 104 twin pairs (a triplet was counted as one MZ female pair and one dizygotic (DZ) brother) were included, collected within RATSS between 2011 and 2014. In three families, samples from two twin pairs were included. Additionally, 75 parents from 37 of these families were analyzed. After the quality control of genotyping and CNV data, the sample set used for the CNV analysis comprised of 69 MZ (including one female MZ pair from the triplet) and 31 DZ twin pairs (the male sibling from the triplet was excluded from the analysis), corresponding to 97 families. A total of 26 pairs were discordant and 11 pairs concordant for ASD (Table 1). In the twin pairs discordant for ASD, 69.2% (18/26) of the affected twins and 19.2% (5/26) of the co-twins had at least one additional NDD diagnosis. For the NDD category, 34 pairs were classified as discordant and 35 pairs as concordant. Twenty-seven twin pairs were classified as typically developed. Mean age, IQ, and the number of other DSM-5 NDD diagnoses for the twin pairs in each category are shown in Table 1.

DNA Samples and Genotyping
Saliva samples were collected from the twins and their parents using the Oragene•DNA OG-500 tubes (DNA Genotek, Inc., Ottawa, Ontario, Canada) during the RATSS study visit or at home. The extraction of genomic DNA from the saliva samples was done automatically using the Autopure LS system and Puregene DNA purification kit (Gentra Systems, Minneapolis, MN, USA) or Chemagen DNA saliva kit (PerkinElmer, Waltham MA, USA). A total of 200 ng of DNA was used for the genotyping using the Infinium PsychArray-24 v1.1 (Illumina Inc., San
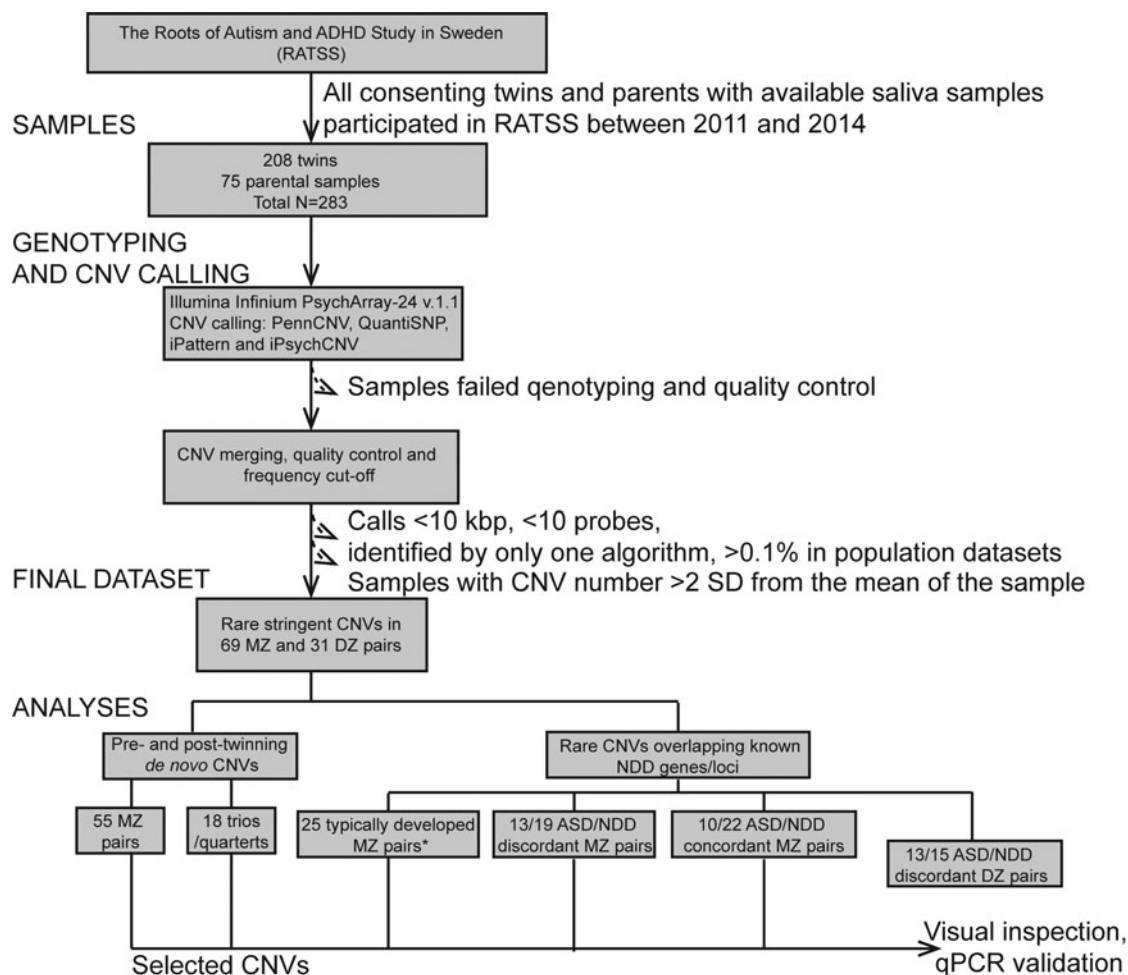
FIGURE 1

Overview of the study including the numbers of monozygotic (MZ) and dizygotic (DZ) pairs included in each step of the analysis. The dashed arrows indicate samples and copy number variants (CNVs) were excluded from the analysis. *Three twin pairs were excluded as psychiatric disorders were present in the pairs.

Diego, California, USA), which contains a total of ~593,260 SNP markers, including 50,000 specific genetic markers for neuropsychiatric disorders, according to the manufacturer's protocol. Samples with SNP call rate lower than 98% were excluded from further investigations. The zygosity of the twin pairs was confirmed using the genotype data by estimating the identity by descent using the PLINK/1.07 (Purcell et al., 2007) after quality control and removal of SNPs with minor allele frequency <0.05 within the samples.

CNV Calling

We derived the CNVs in our data by using four different algorithms: PennCNV (Wang et al., 2007), QuantiSNP (Colella et al., 2007), iPattern (Pinto et al., 2010), and iPsychCNV (prerelease beta v1.0; Bertalan et al., 2015). Although there is no established procedure for CNV calling using the intensity data from Infinium PsychArray, the combination of multiple CNV detection algorithms is a

'gold standard' approach to detect reliable CNV calls and has been widely used with other Illumina arrays (Pinto et al., 2014). The called CNVs were annotated based on GRCh37/hg19 genome build. For the downstream analyses, we used only stringent CNVs, which were called by iPattern, and at least one additional algorithm or by PennCNV together with iPattern or/and iPsychCNV with 50% reciprocal overlap between the algorithms. We further excluded variants that were <10 kb in size, covered by <10 consecutive probes and had more than 50% overlap with centromeres or telomeres. After obtaining the final set of stringent CNV calls, we removed individuals with excessive CNV calls defined by >2 standard deviations from the mean CNV count in the whole sample. Out of the 283 individuals genotyped, a total 231 calls proceeded for further analysis after the QC of genotyping and CNV calling (Figure 1). The majority of the samples that failed QC were parental saliva samples collected at home. For 55 MZ and

TABLE 1

Phenotypic Characteristics of 66 Monozygotic and 30 Dizygotic Twins

MZ twin pairs	Typically developed	Concordant ASD	Discordant ASD		Concordant NDD	Discordant NDD	
Number of pairs	25	10	13		22	19	
Female pairs (%)	12 (48%)	5 (50%)	5 (38.5%)		7 (31.8%)	9 (47.4%)	
Male pairs (%)	13 (52%)	5 (50%)	8(61.5%)		15 (68.2%)	10 (52.6%)	
Age mean (range)	15.6 (10–22)	15.5 (9–28)	14 (9–20)		15 (9–28)	13 (9–20)	
			Proband	Co-twin		Proband	Co-twin
IQ mean (range)	101.8 (81–126)	88.1 (58–142)	88.3 (40–138)	96.8 (65–121)	86.7 (58–142)	93.8 (40–138)	98.8 (77–121)
Cases with ASD (%)	–	20 (100%)	13 (100%)	–	20 (45.5%)	10 (52.6%)	–
Cases with ADHD (%)	–	11 (55%)	2 (15.4%)	1 (7.7%)	28 (63.6%)	9 (47.4%)	–
Cases with ID (%)	–	3 (15%)	3 (23.1%)	1 (7.7%)	11 (25%)	3 (15.8%)	–
Cases with other NDD diagnoses*	–	4 (20%)	1 (7.7%)	–	15 (34.1%)	6 (31.6%)	–
DZ twin pairs	Typically developed	Concordant ASD	Discordant ASD		Concordant NDD	Discordant NDD	
Number of pairs	2	1	13		13	15	
Female pairs (%)	–	–	2 (15.4%)		5 (38.5%)	5 (33.3%)	
Male pairs (%)	1 (50%)	1 (100%)	8 (61.5%)		7 (53.8 %)	7 (44.7%)	
Opposite sex pairs (%)	1 (50%)	–	3 (23.1%)		1 (7.7%)	3 (20%)	
Age mean (range)	13.5 (11–16)	14 (14–14)	14.1 (10–20)		14.2 (8–19)	14.6 (8–25)	
			Proband	Co-twin		Proband	Co-twin
Mean IQ (range)	125 (115–138)	97 (83–111)	83.5 (42–120)	98.2 (65–124)	95.3 (65–130)	88.1 (42–120)	101.9 (77–124)
Cases with ASD (%)	–	2 (100%)	13 (100%)	–	2 (7.7%)	10 (66.7%)	—
Cases with ADHD (%)	–	1 (50%)	8 (61.5%)	2 (15.3%)	20 (77%)	10 (66.7%)	–
Cases with ID (%)	–	–	3 (23.1 %)	1 (7.7%)	2 (7.7%)	2 (13.3%)	–
Cases with other NDD diagnoses*	–	–	1 (7.7%)	–	3 (11.5%)	4 (26.7%)	–

Note: *Additional NDD diagnosis include specific learning disorder, tic disorder, and speech sound disorder.

19 DZ pairs, both twins were successfully genotyped. In 18 families both parents were genotyped.

Population-based control data sets described by Lionel et al. (2014) were used to restrict the analyses to rare CNVs found in frequency <0.01% in controls. Additionally, we computed the frequency of CNV events using PLINK/1.07 (Purcell et al., 2007) based on two reference sets: the total cohort including both parents and the twins, and the parents only. We removed variants present in >3 copies in the parents.

Pre-Twinning and Post-Twinning De Novo CNV Detection

The pre-twinning de novo CNVs were analyzed in 18 families in which DNA was available for both parents. The rare stringent CNVs shared by a twin pair or present in one twin in the twin pairs and not found in the parents or the whole population stringent and non-stringent CNV calls were considered as putative pre-twinning de novo variants. Similarly, the postzygotic de novo CNV events were analyzed using the stringent CNV set from the 55 complete MZ twin pairs. Rare stringent CNVs found only in one twin were further filtered using the population frequencies as well as using the non-stringent calls in the co-twins/families. All the surviving unique CNVs that were not detected in any of the non-stringent calls in the co-twin or the whole population dataset were considered as post-twinning de novo variants. The putative pre- and post-twinning de novo CNV calls were visually inspected by plotting the Log R Ratio (LRR) and B allele frequency values of the twins and the parents. If both LRR and B allele frequency showed a clear pattern of the called copy number (CN) and no deviation from normal CN were found in the other twin or parents, the CNV proceeded for experimental validation.

Rare CNVs Overlapping Genes Implicated in Neurodevelopmental Disorders

Next, we characterized rare CNVs for the MZ and DZ twins groups based on the concordance/discordance for ASD and NDD. We included all MZ twin pairs for which at least one twin was successfully genotyped and each twin in DZ pairs. Among the all rare CNVs, we identified CNVs that overlapped known genomic disorders, earlier implicated in ASD/intellectual disability genes (Pinto et al., 2014), genes curated in the SFARI gene database (accession date September 2016; Banerjee-Basu & Packer, 2010) and genes associated with abnormality of the nervous system in the human phenotype ontology (HP:0000707) hereafter termed as HPO-NS (Groza et al., 2015). Thereafter, the CNVs were also evaluated based on the criteria provided by the American College of Medical Genetics and Genomics (ACMG; Kearney et al., 2011). Three group Fisher's exact test (FET) were used to test whether any significant changes among the MZ twin pairs within different discordancy groups. McNe-

mar's test was used among the DZ discordant pairs to test for any intra-pair significant differences. All statistical analyses were performed with R (3.2.3).

Experimental Validation of Selected CNVs

Putative pre- and post-zygotic de novo and NDD associated CNVs that passed our filtering criteria and visual inspection were subjected to experimental validation using quantitative PCR (qPCR). For this, Taqman® probes (Thermo Fisher Scientific, Waltham MA, USA) within the genomic region of interest for the validation experiment were chosen and tested together with a reference control probe targeting the *TERT* gene (OMIM#187270; Thermo Fisher Scientific). Also, the CN state of each region was tested in the twin pair, parental DNA if available, and at least one unrelated control DNA sample. Quadruplicates of each sample were analyzed on StepOne™ (Thermo Fisher Scientific) system according to manufacturer's recommendations. The CN of the target in each sample was analyzed using the CopyCaller™ software (Thermo Fisher Scientific).

Results

Pre-Twinning and Post-Twinning De Novo CNVs

No validated pre-twinning de novo CNVs were found in the 18 families in which both parents were successfully genotyped (Table S1). As our next step, we aimed to identify putative postzygotic de novo CNV events in the 55 complete MZ twin pairs, of which 11 pairs were classified as discordant for ASD or 18 pairs as discordant for NDDs. First, we calculated the concordance rate of each CNV detection algorithm when using the <10 kb size and <10 consecutive probe filter within the twin pairs (Table S2). The average concordance rates ranged between 13.6% for iPsychCNV and 65.4% for iPattern. In our final stringent CNV calls, the concordance rate was 64.2% with 16 of 55 twin pairs (29.1%) sharing all their CNVs.

For CNVs larger than 100 Kb the mean concordance rate among MZ twins was 73.6% (median 84.2%) with 24 of 55 twin pairs (43.6%) sharing all their CNVs consistent with the prior literature (Abdellaoui et al., 2015). After removing the common variants, variants called by lower stringency in the co-twin or present in additional individuals in our dataset (see Materials and Methods, Figure 1), a total of eight non-shared rare stringent putative postzygotic CNV calls were found in four MZ twin pairs in our analysis. Following a visual inspection, five of them proceeded to qPCR validation. None of the five CNVs were validated to be nonshared; two were found in both twins, and three CNVs were false positives (see Table S1).

Rate of Rare CNVs Affecting Genes Previously Implicated in NDDs or Nervous System Disorders

In total, 28 CNVs in 22 twin pairs were found to overlap the pre-defined gene sets, including ASD/ID genes (Pinto et al.,

ASD/ID/SFARI

HPO − Disorders of Nervous System
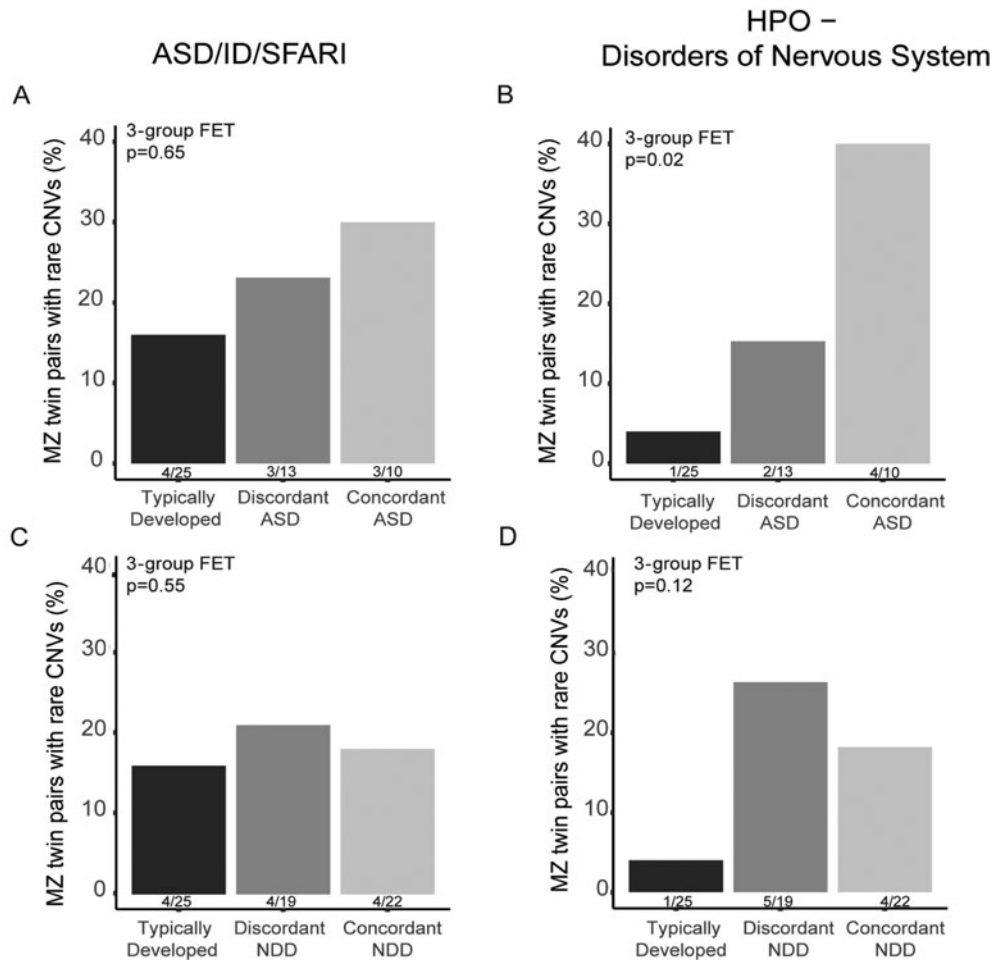
FIGURE 2

The rate of rare CNVs affecting genes earlier implicated in neurodevelopmental disorders (ASD/ID/SFARI) or abnormality of the nervous system in the Human Phenotype Ontology database (HPO-NS) in MZ twin pairs. The percentage of MZ twins in typically developing, ASD-discordant and -concordant pairs with CNVs overlapping genes in ASD/ID/SFARI gene lists (A) and in HPO-NS (B). The percentage of typically developing, NDD-discordant and -concordant pairs with CNVs overlapping genes in ASD/ID/SFARI gene lists (C) and in HPO-NS (D). Statistical significance was tested between the three twin pair groups using Fisher Exact Test, *p*-value for three group comparison is reported in the figure.

2014) in combination with SFARI genes (Banerjee-Basu & Packer, 2010) and/or HPO-NS genes. From these 28 CNVs, nine were targeted for experimental validation using qPCR (Additional file 2, Supplementary Table S3) and all of these CNVs were validated. We calculated the proportion of twin pairs or twins in discordant DZ pairs having at least one CNV in these gene categories within each twin category. We found an increased rate of MZ twin pairs discordant and concordant for ASD having CNVs affecting both NDD genes (Figure 2A) and HPO-NS (Figure 2B) in comparison with TD MZ pairs. However, only the HPO-NS comparison reached statistical significance (three-group comparison 2-sided FET $p = .02$). As expected, the largest difference was found between the frequency in TD pairs compared with concordant ASD pairs (OR $= 14.4$, 95% CI [1.2, 814.6]). For MZ twins categorized based on the NDD diagnosis, we did not observe any significant differences for the NDD or

HPO-NS (three-group comparison, 2-sided FET $p = .92$, and $p = .11$, respectively). To note, the proportion of twin pairs with CNVs for both NDD concordant and discordant pairs was higher than in the TD pairs for the HPO-NS genes although not significant (Figure 2D). For DZ twin pairs, no differences were found between ASD/NDD affected twins and their co-twins (Supplementary Table S4).

Rare CNVs with Putative Clinical Significance Found in the Sample

Next, we investigated the rare CNVs overlapping the ASD/ID and HPO-NS gene lists to categorize them based on the guidelines from the ACMG (Kearney et al., 2011) and evidence from the earlier literature. Two CNVs in two twin pairs among the 10 MZ pairs concordant for ASD (20.0%) and one additional CNV among the other MZ pairs concordant for NDD (13.6%, 3 out of 22 pairs), were

classified as clinically significant. A male MZ twin pair concordant for ASD and ADHD (Family ID 12) had a paternally inherited ~9.0 Mb duplication at 12q11-12q13.11 (GRCh37/hg19, chr12:37,995,364-47,072,741) affecting 33 genes. Proximal deletions of 12q12 had earlier been implicated in patients with NDDs and dysmorphic features (Adam et al., 2010). Another MZ male pair (Family ID 23) concordant for ASD and ID and discordant for tic disorder had a ~44.3 kb deletion on chromosome 1q44 (GRCh37/hg19, chr1:244,997,953-245,042,312) affecting *COX20* (OMIM#614698) and *HNRNPU* (OMIM#602869), as well as *HNRNPU-AS1* long non-coding antisense RNA. We were not able to conclude whether the deletion was of de novo origin as paternal DNA sample was not available. Recently, multiple reports have described cases with mutations affecting *HNRNPU* demonstrating its role in NDDs, including severe intellectual disability and early onset seizures (Bramswig et al., 2017; Leduc et al., 2017; Yates et al., 2017). Additionally, a maternally inherited 18.8 kb deletion on chr8p23.2 (GRCh37/hg19, chr8:3,063,043-3,081,903) disrupting *CSMD1* (OMIM#608397) was detected in a MZ male pair (Family ID 11) concordant for ADHD and discordant for tic disorder. Recently, both inherited and de novo variants in *CSMD1* have been found in ASD patients (Cukier et al., 2014; Krumm et al., 2015). Krumm et al. (2015) demonstrated a depletion of CNVs in controls within the 3′ genomic loci of *CSMD1* and overlap of genomic variants in cases. The deletion detected in the twins is within the region the study highlighted.

Among the discordant MZ pairs, we detected a maternally inherited 0.95 Mb gain on chromosome 6q13-14.1 (GRCh37/hg19, chr6:75,682,566-76,635,743) possibly contributing to the phenotypes seen in the male twin pair (Family ID 29). Both large deletions and duplications at chromosome 6q encompassing this region have been implicated in NDD associated phenotypes (Kaminsky et al., 2011; Van Esch et al., 2010), and no similar size overlapping CNVs have been detected in DGV. The affected twin has ASD with full-scale IQ of 70, dyslexia, and earlier indicated mild macrocephaly (current head circumference 54.5 cm, 84th percentile). The co-twin has mild macrocephaly (56 cm, 98th percentile) and a reported mild language delay described as language impairments specific for ASD, that is, reversal use of pronouns and regular use of idiosyncratic language, as well as mild understanding and articulation problems. Lack of social reciprocity, pretend, and fantasy play is reported for both twins. However, the impairments in the co-twin are overall reported as previous during development and at a sub-threshold level, and were not observed at the time of the assessment, in comparison to the affected twin. The co-twin also has a slightly higher IQ of 77.

In the DZ pairs, we identified a CNV affecting *SYNE1* (OMIM#608441) that is likely pathogenic in a male twin diagnosed with ASD and ID from an opposite-sex DZ pair (Family ID 31). We verified the 39 kb gain on chr6q25.2 (GRCh37/hg19, chr6:152,712,302-152,751,762) using qPCR and showed it to be de novo (Figure S1). His TD female sibling, mother, and father (DNA sample obtained for qPCR validation) had normal CN for the region. *SYNE1*, which encodes for the Nesprin-1 protein, has been implicated in ASD through rare mutations (De Rubeis et al., 2014; Jiang et al., 2013; Krumm et al., 2015; O'Roak et al., 2011) but is also involved in spinocerebellar ataxia (Wiethoff et al., 2016) and Emery–Dreifuss muscular dystrophy (Zhang et al., 2007).

Discussion

We aimed to investigate the presence of postzygotic de novo CNVs that could explain the discordance of a variety of NDD phenotypes, especially ASD, in MZ pairs by genotyping saliva-derived DNA using the Illumina Infinium PsychArray that was designed for studies of psychiatric disorders. We did not identify any validated postzygotically derived unique CNVs in the 55 MZ pairs analyzed, thus indicating that these do not contribute to the discordancy of NDDs in our sample. Our results are in accordance with the earlier studies showing that postzygotic de novo CNVs are rare when analyzing genomic DNA from blood or saliva samples (Abdellaoui et al., 2015; Bloom et al., 2013; Ehli et al., 2012; Laplana et al., 2014; Magnusson et al., 2016). Earlier literature suggests that a lower level of mosaicism is masked in blood-derived DNA (Erlich, 2011; King et al., 2017). Here, we restricted our analysis to saliva DNA and CNV level variants. However, we cannot exclude the possibility that other types or tissue-specific postzygotic somatic variations contribute to the discordancy. In non-twin cohorts, somatic sequence level mutations are found in 7.5% of affected children and shown to contribute to both ASD and ID diagnoses (Acuna-Hidalgo et al., 2015; Freed & Pevsner, 2016; Lim et al., 2017). There is no clear estimate of the rate of somatic CNVs, and although our sample is enriched for discordant twins that in theory should have a higher rate of such events, one obvious limitation of our study is the small sample size. Further studies investigating the rate of postzygotic CNV events in discordant twins should be done in larger twin cohorts with detailed phenotype information.

Additionally, somatic CNVs and sequence level variation can be found in human neurons from healthy individuals using single-cell genomics (Lodato et al., 2015; McConnell et al., 2013). Therefore, stochastic genetic variations at a neuronal level could contribute to disrupted brain development and underlie phenotype differences connected to brain disorders in MZ twins. As recent literature has implicated the role of somatic DNMs in NDDs, a better understanding of their contribution and origin in NDDs is needed as it can inform better estimations of the recurrence rates of the disorders and thus genetic counseling.

We also sought to understand the contribution of shared rare CNVs in the twin pairs with a specific focus on the ASD and NDD-discordant twins. As expected, we identified a higher rate of rare CNVs, especially in genes earlier implicated in NDDs and development of the nervous system in ASD-concordant pairs (Figure 2). The rate of putatively clinical significant CNVs in our twins concordant for NDD (13.6%) was similar to the reported rates in ASD cohorts ranging from ~10% to 28% (Ho et al., 2016; Shen et al., 2010; Tammimies et al., 2015). Our data also suggest that there could be an increased rate of the CNVs affecting genes involved in alterations of the nervous system also in the ASD/NDD-discordant MZ pairs (Figure 2B, D) demonstrating that CNVs with incomplete penetrance could contribute to the NDD phenotypes seen in the twins. Many of the most common CNVs found in NDD cases have variable expressivity as well as incomplete penetrance (Carter & Scherer, 2013). More severe phenotypes connected to the recurrent genomic disorders and rare CNVs are most likely due to additional genetic variants, epigenetic modifications, environmental exposures, and their combined effects. For instance, additional rare CNVs were more common in carriers of the recurrent 16p11.2 deletions with more severe NDD phenotypes (Girirajan et al., 2010). The benefit of MZ discordant twin pairs design is the possibility to pinpoint additional factors such as non-shared environmental factors that contribute to the etiology of NDDs alone or as modifying the effect of rare CNVs. Indeed, we have identified non-shared medical and environmental factors that are contributing to the phenotypes using this same sample of twins (Arora et al., 2017; Willfors et al., 2017).

Multiple studies have highlighted the need for a cross-disorder approach in genetic studies as earlier research has shown the clinical and genetic overlap between different NDDs (Lee et al., 2013; Thapar et al., 2017). We used two different approaches, either clinical ASD diagnosis only or the umbrella category of NDD, to define the concordance and discordance of the twins for our analysis of shared CNVs. Using this approach, we demonstrate a rising degree of the rate of CNVs affecting specific genes from TD to ASD-discordant and -concordant pairs. When using the NDD concept, we did not find any differences among the NDD discordant and concordant pairs despite an elevated rate seen in comparison with the TD pairs. The lack of difference could be due to small sample size but also the effect of the milder phenotypes such as learning disabilities included in the concordant NDD group lowering the proportion of more severely affected twin pairs in the concordant category. It has been implicated that CNV burden is not increased in milder NDDs (Girirajan et al., 2011).

In conclusion, we add to the growing literature that postzygotic de novo CNV events in MZ twin pairs are rare and do not explain the discordance of NDD phenotypes in the studied twin pairs. We suggest that shared CNVs overlapping genes essential for brain development increase the risk for NDDs even in discordant MZ pairs. Further studies using larger twin cohorts, additional sources of DNA, and better variant detection resolution such as whole genome sequencing (Yuen et al., 2017) are needed to confirm our results and overcome the limitations of this study. Additionally, studies should aim at investigating the contribution of other risk or resilience factors contributing to the phenotypic outcomes in genetic risk backgrounds.

Declarations

Ethics Approval and Consent to Participate

Consent for Publication

Availability of Data and Material

Competing Interests

Funding

Authors' Contributions

S.S., S.Bö., and K.T conceived and designed the study. B-M.A., C.W., S.Be., and A.N. recruited, diagnosed, and examined the twins. S.S., B.T., J.W., and K.T. processed the microarray data. S.S. and K.T. analyzed the data and drafted the manuscript. All authors contributed to acquisition of data, interpretation of results and critical discussion and approved the final version of the manuscript.

Acknowledgments

Supplementary material

Supplementary material is available on the Cambridge Core website.

To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2017.69

References

Abdellaoui, A., Ehli, E. A., Hottenga, J. J., Weber, Z., Mbarek, H., Willemsen, G., … Boomsma, D. I. (2015). CNV concordance in 1,097 MZ twin pairs. *Twin Research and Human Genetics, 18*, 1–12.

Acuna-Hidalgo, R., Bo, T., Kwint, M. P., van de Vorst, M., Pinelli, M., Veltman, J. A., … Gilissen, C. (2015). Postzygotic point mutations are an underrecognized source of de novo genomic variation. *American Journal of Human Genetics, 97*, 67–74.

Acuna-Hidalgo, R., Veltman, J. A., & Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biology, 17*, 241.

Adam, M. P., Mehta, A., Knight, L., Hall, D. E., & Rossi, M. R. (2010). A family with a 1.17 Mb deletion of 12q12: Refining genotype-phenotype correlation. *American Journal of Medical Genetics Part A, 152*, 2394–2398.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Arlington, VA: American Psychiatric Publishing.

Arora, M., Reichenberg, A., Willfors, C., Austin, C., Gennings, C., Berggren, S., … Bolte, S. (2017). Fetal and postnatal metal dysregulation in autism. *Nature Communications, 8*, 15493.

Banerjee-Basu, S., & Packer, A. (2010). SFARI Gene: An evolving database for the autism research community. *Disease Models & Mechanisms, 3*, 133–135.

Bertalan, M., Weinsheimer, S., Sparsø, T., Mazin, W., & Werge, T. (2015). IPSYCHCNV: A robust method for copy number variation detection on dried blood spots. *European Neuropsychopharmacology, 27*, S107–S138.

Bloom, R. J., Kahler, A. K., Collins, A. L., Chen, G., Cannon, T. D., Hultman, C., & Sullivan, P. F. (2013). Comprehensive analysis of copy number variation in monozygotic twins discordant for bipolar disorder or schizophrenia. *Schizophrenia Research, 146*, 289–290.

Bölte, S., Willfors, C., Berggren, S., Norberg, J., Poltrago, L., Mevel, K., … Lichtenstein, P. (2014). The Roots of Autism and ADHD Twin Study in Sweden (RATSS). *Twin Research and Human Genetics, 17*, 164–176.

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., … Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics, 127*, 1034–1042.

Bramswig, N. C., Ludecke, H. J., Hamdan, F. F., Altmuller, J., Beleggia, F., Elcioglu, N. H., … Wieczorek, D. (2017). Heterozygous HNRNPU variants cause early onset epilepsy and severe intellectual disability. *Human Genetics, 136*, 821–834.

Carter, M. T., & Scherer, S. W. (2013). Autism spectrum disorder in the genetics clinic: A review. *Clinical Genetics, 83*, 399–407.

Castellani, C. A., Awamleh, Z., Melka, M. G., O'Reilly, R. L., & Singh, S. M. (2014). Copy number variation distribution in six monozygotic twin pairs discordant for schizophrenia. *Twin Research and Human Genetics, 17*, 108–120.

Colella, S., Yau, C., Taylor, J. M., Mirza, G., Butler, H., Clouston, P., … Ragoussis, J. (2007). QuantiSNP: An objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Research, 35*, 2013–2025.

Cukier, H. N., Dueker, N. D., Slifer, S. H., Lee, J. M., Whitehead, P. L., Lalanne, E., … Pericak-Vance, M. A. (2014). Exome sequencing of extended families with autism reveals genes shared across neurodevelopmental and neuropsychiatric disorders. *Molecular Autism, 5*, 1.

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., … Buxbaum, J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature, 515*, 209–215.

Ehli, E. A., Abdellaoui, A., Hu, Y., Hottenga, J. J., Kattenberg, M., van Beijsterveldt, T., … Davies, G. E. (2012). De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on attention problems. *European Journal of Human Genetics, 20*, 1037–1043.

Erlich, Y. (2011). Blood ties: Chimerism can mask twin discordance in high-throughput sequencing. *Twin Research and Human Genetics, 14*, 137–143.

Faraone, S. V., Ghirardi, L., Kuja-Halkola, R., Lichtenstein, P., & Larsson, H. (2017). The familial co-aggregation of attention-deficit/hyperactivity disorder and intellectual disability: A register-based family study. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*, 167–174.e1.

Freed, D., & Pevsner, J. (2016). The contribution of mosaic variants to autism spectrum disorder. *PLoS Genetics, 12*, e1006245.

Ghirardi, L., Brikell, I., Kuja-Halkola, R., Freitag, C. M., Franke, B., Asherson, P., … Larsson, H. (2017). The familial co-aggregation of ASD and ADHD: A register-based cohort study. *Molecular Psychiatry*. (http://dx.doi.org/10.1038/mp.2017.17).

Girirajan, S., Brkanac, Z., Coe, B. P., Baker, C., Vives, L., Vu, T. H., … Eichler, E. E. (2011). Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genetics, 7*, e1002334.

Girirajan, S., Rosenfeld, J. A., Cooper, G. M., Antonacci, F., Siswara, P., Itsara, A., … Eichler, E. E. (2010). A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nature Genetics, 42*, 203–209.

Groza, T., Kohler, S., Moldenhauer, D., Vasilevsky, N., Baynam, G., Zemojtel, T., … Robinson, P. N. (2015). The human phenotype ontology: Semantic unification of common and rare disease. *American Journal of Human Genetics, 97*, 111–124.

Ho, K. S., Wassman, E. R., Baxter, A. L., Hensel, C. H., Martin, M. M., Prasad, A., … Butler, M. G. (2016). Chromosomal microarray analysis of consecutive individuals with autism spectrum disorders using an ultra-high resolution chromosomal microarray optimized for neurodevelopmental disorders. *International Journal of Molecular Sciences, 17*, 2070.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature, 515*, 216–221.

Jiang, Y. H., Yuen, R. K., Jin, X., Wang, M., Chen, N., Wu, X., … Scherer, S. W. (2013). Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *American Journal of Human Genetics, 93*, 249–263.

Kaminsky, E. B., Kaul, V., Paschall, J., Church, D. M., Bunke, B., Kunig, D., … Martin, C. L. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genetics in Medicine, 13*, 777–784.

Kearney, H. M., Thorland, E. C., Brown, K. K., Quintero-Rivera, F., South, S. T., & Working Group of the American College of Medical Genetics Laboratory Quality Assurance, C. (2011). American college of medical genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genetics in Medicine, 13*, 680–685.

King, D. A., Sifrim, A., Fitzgerald, T. W., Rahbari, R., Hobson, E., Homfray, T., … Hurles, M. E. (2017). Detection of structural mosaicism from targeted and whole-genome sequencing data. *Genome Research, 27*, 1704–1714.

Kirov, G. (2015). CNVs in neuropsychiatric disorders. *Human Molecular Genetics, 24*, R45–R49.

Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., … Eichler, E. E. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics, 47*, 582–588.

Laplana, M., Royo, J. L., Aluja, A., Lopez, R., Heine-Sunyer, D., & Fibla, J. (2014). Absence of substantial copy number differences in a pair of monozygotic twins discordant for features of autism spectrum disorder. *Case Reports in Genetics, 2014*, Article ID 516529.

Leduc, M. S., Chao, H. T., Qu, C., Walkiewicz, M., Xiao, R., Magoulas, P., … Yang, Y. (2017). Clinical and molecular characterization of de novo loss of function variants in HNRNPU. *American Journal of Medical Genetics Part A, 173*, 2680–2689.

Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., Perlis, R. H., … Wray, N. R. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics, 45*, 984–994.

Lim, E. T., Uddin, M., De Rubeis, S., Chan, Y., Kamumbu, A. S., Zhang, X., … Walsh, C. A. (2017). Rates, distribution, and implications of post-zygotic mosaic mutations in autism spectrum disorder. *Nature Neuroscience, 20*, 1217–1224.

Lionel, A. C., Tammimies, K., Vaags, A. K., Rosenfeld, J. A., Ahn, J. W., Merico, D., … Scherer, S. W. (2014). Disruption of the ASTN2/TRIM32 locus at 9q33.1 is a risk factor in males for autism spectrum disorders, ADHD and other neurodevelopmental phenotypes. *Human Molecular Genetics, 23*, 2752–2768.

Lodato, M. A., Woodworth, M. B., Lee, S., Evrony, G. D., Mehta, B. K., Karger, A., … Walsh, C. A. (2015). Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science, 350*, 94–98.

Lundstrom, S., Reichenberg, A., Melke, J., Rastam, M., Kerekes, N., Lichtenstein, P., … Anckarsater, H. (2015). Autism spectrum disorders and coexisting disorders in a nationwide Swedish twin study. *Journal of Child Psychology and Psychiatry, 56*, 702–710.

Magnusson, P. K., Lee, D., Chen, X., Szatkiewicz, J., Pramana, S., Teo, S., … Pawitan, Y. (2016). One CNV Discordance in NRXN1 observed upon genome-wide screening in 38 pairs of adult healthy monozygotic twins. *Twin Research and Human Genetics, 19*, 97–103.

McConnell, M. J., Lindberg, M. R., Brennand, K. J., Piper, J. C., Voet, T., Cowing-Zitron, C., … Gage, F. H. (2013). Mosaic copy number variation in human neurons. *Science, 342*, 632–637.

McRae, A. F., Visscher, P. M., Montgomery, G. W., & Martin, N. G. (2015). Large autosomal copy-number differences within unselected monozygotic twin pairs are rare. *Twin Research and Human Genetics, 18*, 13–18.

O'Roak, B. J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J. J., Girirajan, S., … Eichler, E. E. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nature Genetics, 4*, 585–589.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., … Scherer, S. W. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *American Journal of Human Genetics, 94*, 677–694.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., … . Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature, 466*, 368–372.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*, 559–575.

Shen, Y., Dies, K. A., Holm, I. A., Bridgemohan, C., Sobeih, M. M., Caronna, E. B., … Miller, D. T. (2010). Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics, 125*, e727–e735.

Sullivan, P. F., Magnusson, C., Reichenberg, A., Boman, M., Dalman, C., Davidson, M., … Lichtenstein, P. (2012). Family history of schizophrenia and bipolar disorder as risk factors for autism. *Archives of General Psychiatry, 69*, 1099–1103.

Tammimies, K., Marshall, C. R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A. C., … Fernandez, B. A. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA, 314*, 895–903.

Thapar, A., Cooper, M., & Rutter, M. (2017). Neurodevelopmental disorders. *Lancet Psychiatry, 4*, 339–346.

Torres, F., Barbosa, M., & Maciel, P. (2016). Recurrent copy number variations as risk factors for neurodevelopmental disorders: Critical overview and analysis of clinical implications. *Journal of Medical Genetics, 53*, 73–90.

Van Esch, H., Rosser, E. M., Janssens, S., Van Ingelghem, I., Loeys, B., & Menten, B. (2010). Developmental delay and connective tissue disorder in four patients sharing a common microdeletion at 6q13-14. *Journal of Medical Genetics, 47*, 717–720.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F., … Bucan, M. (2007). PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research, 17*, 1665–1674.

Wiethoff, S., Hersheson, J., Bettencourt, C., Wood, N. W., & Houlden, H. (2016). Heterogeneity in clinical features and disease severity in ataxia-associated SYNE1 mutations. *Journal of Neurology, 263*, 1503–1510.

Willfors, C., Carlsson, T., Anderlid, B. M., Nordgren, A., Kostrzewa, E., Berggren, S., … Bölte, S. (2017). Medical history of discordant twins and environmental etiologies of autism. *Translational Psychiatry, 7*, e1014.

Yates, T. M., Vasudevan, P. C., Chandler, K. E., Donnelly, D. E., Stark, Z., Sadedin, S., … Balasubramanian, M. (2017). De novo mutations in HNRNPU result in a neurodevelopmental syndrome. *American Journal of Medical Genetics: Part A, 173*, 3003–3012.

Yuen, R. K. C., Merico, D., Bookman, M., Howe, J. L., Thiruvahindrapuram, B., Patel, R. V., … Scherer, S. W. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nature Neuroscience, 20*, 602–611.

Yuen, R. K. C., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., … Scherer, S. W. (2016). Genome-wide characteristics of de novo mutations in autism. *NPJ Genomic Medicine, 1*, 16027.

Zhang, Q., Bethmann, C., Worth, N. F., Davies, J. D., Wasner, C., Feuer, A., … Shanahan, C. M. (2007). Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Human Molecular Genetics, 16*, 2816–2833.