CNV Concordance in 1,097 MZ Twin Pairs

Abdel Abdellaoui,[1] Erik A. Ehli,[2] Jouke-Jan Hottenga,[1] Zachary Weber,[2] Hamdi Mbarek,[1] Gonneke Willemsen,[1] Toos van Beijsterveldt,[1] Andrew Brooks,[3] Jim J. Hudziak,[4] Patrick F. Sullivan,[5] Eco J. de Geus,[1,6,7] Gareth E. Davies,[2] and Dorret I. Boomsma[1,6,7]

[1]*Department of Biological Psychology, VU University Amsterdam, Amsterdam, the Netherlands*
[2]*Avera Institute for Human Genetics, Avera McKennan Hospital & University Health Center, Sioux Falls, South Dakota, USA*
[3]*Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA*
[4] *College of Medicine, University of Vermont, Burlington, Vermont, USA*
[5]*Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA*
[6]*Neuroscience Campus Amsterdam, Amsterdam, the Netherlands*
[7]*EMGO+ Institute for Health and Care Research, Amsterdam, the Netherlands*

Monozygotic (MZ) twins are genetically identical at conception, making them informative subjects for studies on somatic mutations. Copy number variants (CNVs) are responsible for a substantial part of genetic variation, have relatively high mutation rates, and are likely to be involved in phenotypic variation. We conducted a genome-wide survey for post-twinning de novo CNVs in 1,097 MZ twin pairs. Comparisons between MZ twins were made by CNVs measured in DNA from blood or buccal epithelium with the Affymetrix 6.0 microarray and two calling algorithms. In addition, CNV concordance rates were compared between the different sources of DNA, and gene-enrichment association analyses were conducted for thought problems (TP) and attention problems (AP) using CNVs concordant within MZ pairs. We found a total of 153 putative post-twinning de novo CNVs >100 kb, of which the majority resided in 15q11.2. Based on the discordance of raw intensity signals a selection was made of 20 de novo CNVs for a qPCR validation experiments. Two out of 20 post-twinning de novo CNVs were validated with qPCR in the same twin pair. The 13-year-old MZ twin pair that showed two discordances in CN in 15q11.2 in their buccal DNA did not show large phenotypic differences. From the remaining 18 putative de novo CNVs, 17 were deletions or duplications that were concordant within MZ twin pairs. Concordance rates within twin pairs of CNV calls with CN ≠ 2 were ∼80%. Buccal epithelium-derived DNA showed a slightly but significantly higher concordance rate, and blood-derived DNA showed significantly more concordant CNVs per twin pair. The gene-enrichment analyses on concordant CNVs showed no significant associations between CNVs overlapping with genes involved in neuronal processes and TP or AP after accounting for the source of DNA.

■ **Keywords:** monozygotic twins, copy number variation, de novo mutations, thought problems, attention problems

MZ twins have long been assumed to be genetically identical, which is an important assumption for twin studies, where phenotypic correlations between MZ twins and dizygotic twins are compared in order to estimate the relative contribution of genes and environment in human traits (Boomsma et al., 2002). MZ twins are, in fact, genetically identical at conception, but can accumulate mutations after the zygote splits, making MZ twins informative for the study of somatic mutations. Post-twinning point mutations have been reported (Kondo et al., 2002; Reumers et al., 2012; Sakuntabhai et al., 1999; Vadlamudi et al., 2010; Ye et al., 2013), but are expected to be scarcer than post-twinning de novo CNVs. CNVs, the most studied type of structural variant, are segments of DNA ranging from 1 kb to sev-

eral Mb that differ in copy number (CN) across different members of the species. CNVs have a higher mutation rate than single nucleotide polymorphisms (SNPs) and affect larger segments of the genome (Itsara et al., 2010; Lupski, 2007; van Ommen, 2005). Even though post-twinning de novo CNVs are expected to be rare, they can potentially

aid in finding causal variants for genomic disorders. After Bruder et al. (2008) demonstrated the existence of CNV discordance in MZ twins, many studies followed that tried to find CNV discordances that might be explanatory for phenotypic MZ discordances.

Table 1 shows an overview of studies attempting to detect CNV differences between MZ twins since 2008. Forsberg et al. (2012) conducted the largest study of this kind to date, examining 159 MZ pairs, and validated five post-twinning mutations >1 Mb and five <1 Mb, all found in the older twin pairs of their sample (>60 years old). An estimate of the post-twinning mutation rate for CNVs is difficult to make with this design, since it is likely to depend on the age of the twins, with older individuals having an increased chance for somatic mutations (Forsberg et al., 2012; Ye et al., 2013), and likely also depends on tissue (Piotrowski et al., 2008). The majority of studies looking for CNV discordances in MZ twins did not detect reproducible post-twinning CNV mutations, indicating that relatively large CNV discordances between MZ twins are a considerably rare phenomenon, or are at least hard to detect, even among phenotypically discordant twins.

Most studies on post-twinning de novo CNVs first scan the entire genome using genome-wide microarray technology, making them only sensitive for relatively large CNVs (>10–100 kb), and then validate suggestive signals with additional and more sensitive molecular assays, such as qPCR. In practice, CNVs have been considered relatively noisy when using currently available genome-wide microarray technologies, and qPCR has shown to be effective in validating CNV signals from microarray data (Weaver et al., 2010; Zhang et al., 2011).

We conducted a genome-wide scan for post-twinning de novo CNVs (>100 kb) in 1,097 unselected MZ twin pairs with a wide age range (0–79 years old). DNA was extracted from blood for about half of the samples, which included the majority of adult subjects >18 years old, and the other half of the samples (mainly children) had their DNA extracted from buccal swabs (see Figure 1). CNVs are measured with the Affymetrix 6.0 microarray, and after stringent Quality Control (QC), a selection of post-twinning de novo CNV candidates is made for qPCR replication. Phenotypic data based on extensive longitudinal questionnaires (Boomsma et al., 2006) were available to be examined for twin pairs with validated post-twinning mutations. In addition, we examined the concordance rates of CNV calls within MZ twin pairs and compared those between the different sources of DNA. Finally, we selected CNVs concordant within MZ pairs to conduct gene-enrichment tests in order to test whether CNV events impacting gene sets involved in neuronal processes are associated with TP or AP. The TP and AP scales measure heritable constructs (Abdellaoui et al., 2008; 2012; Derks et al., 2009) that are predictive for schizophrenia (Kasius et al., 1997; Morgan & Cauce, 1999) and Attention Deficit Hyperactivity Disorder

(ADHD) (Derks et al., 2006) respectively, for which CNVs have been shown to be a risk factor (Cook Jr., & Scherer, 2008; Stefansson et al., 2014; Williams et al., 2010).

Methods

Participants

The 1,097 MZ twin pairs included in this study were registered with the Netherlands Twin Registry (NTR) (van Beijsterveldt et al., 2013, Willemsen et al., 2013), and were not selected based on phenotypic information. SNPs from the Affymetrix 6.0 microarray confirmed that all twins were indeed MZ. The mean age of the twins was 25.04 (SD = 15.86), and ranged from 0 to 79 years old (see Figure 1). DNA was extracted from blood for 1,163 twins (mean age = 35.53, SD = 13.24), and from buccal epithelium for 1,031 twins (mean age = 13.11, SD = 8.39). There were 566 pairs in which both twins had their DNA extracted from blood, 500 pairs in which both twins had their DNA extracted from buccal epithelium, and 31 pairs where one had DNA from blood and the other from buccal epithelium. Methods for buccal and blood collection and genomic DNA extraction have been described previously (Willemsen et al., 2010).

CNV Calling

Data from 1,097 MZ twin pairs were extracted from a dataset containing a total of 13,188 samples that were genotyped on the Affymetrix Human Genome-Wide SNP 6.0 Array according to the manufacturer's protocol. This array contains 906,600 SNP and 940,000 CN probes. Of the CN probes, 800,000 are evenly spaced across the genome and the rest across 3,700 known CNV regions. SNPs were called using Affymetrix Powertool, and were used during the QC stage and to confirm the zygosity of MZ twins. CNVs were called with the Birdsuite (Korn et al., 2008) and PennCNV (Wang et al., 2007) algorithms.

For Birdsuite 1.5.5, the Affymetrix Powertool (APT-1.10.2, plug-in to Birdsuite 1.5.5) was used for plate-wise normalization. This algorithm searches for consistent evidence for CNVs across multiple neighboring probes. Information from neighboring probes is integrated into a CN call (0, 1, 2, 3 or 4) for the segment covered by the probes using a hidden Markov model (HMM)-based algorithm. A logarithm of the odds ratio (LOD)-score was generated for each CNV segment, indicating the likelihood of a CNV relative to no CNV in the region. CNV segments were only included if they had a LOD-score >10. We followed the recommendation from the manual in creating batches (http://www.broadinstitute.org/science/programs/medical-and-population-genetics/birdsuite/birdsuite-faq), and processed a maximum of 96 samples per batch. If the plate of origin was known, samples from the same plate were included in the same batch, resulting in 178 batches. Samples where the plate of origin was not known (~3%) were randomly distributed across five batches.

TABLE 1

List of Studies Searching for Post-Twinning De Novo CNVs

Study	N (pairs)	Phenotype	Age	Platform	Number of post-twinning de novo CNVs replicated with qPCR
Bruder et al., 2008	19	Unselected (N = 10); discordant for neurodegenerative disease (N = 9)	Unselected: 37–60 Discordant: >60	32 k BAC array & Illumina HumanHap 300 Duo beadchip	3 with qPCR replication + many supported only by 32 k BAC array & Illumina beadchip
Baranzini et al., 2010	3	Discordant for multiple sclerosis	19, 39, 56	Affymetrix 6.0, Illumina GAIIx	0
Lasa et al., 2010	1	Discordant for breast cancer	61	Agilent Human Genome CGH 244 k microarray	0
Ono et al., 2010	3	Discordant for schizophrenia	μ = 53	Affymetrix 6.0	0
Jakobsen et al., 2011	1	Discordant for cleft lip	Adult	CGH & Affymetrix 5.0	0
Maiti et al., 2011	2	Discordant for schizophrenia	43, 53	Affymetrix 6.0	0 (*but several without qPCR replication*)
Pamphlett & Morahan, 2011	5	Healthy (N = 2), discordant for amyotrophic lateral sclerosis (N = 3)	35–74	Affymetrix 6.0	Two in blood in one discordant pair, seven in hair in one discordant pair (*no qPCR replication performed*)
Sasaki et al., 2011	1	Discordant for multiple system atrophy	67	Agilent SurePrint G3 Human CNV 400 K Microarray; Illumina 57 k whole-genome CNV beadchip	One replicated by non-qPCR method
Breckpot et al., 2012	6	Discordant for congenital heart defects	0 (N = 5), 4 (N = 1)	OGT CytoSure Syndrome Plus array	Three (all in one affected twin)
Ehli et al., 2012	50	Attention Problems: concordant affected (N = 17), concordant unaffected (N = 22), discordant (N = 11)	<13	Affymetrix 6.0	Two (one in concordant unaffected, one in discordant pair)
Forsberg et al., 2012	159	Unselected	<60 (N = 81), >60 (N = 78)	Illumina 1M-Duo, Illumina Omni-Express, Nimblegen 135 K, Nimblegen 720 K	5 > 1 Mb, 5 < 1 Mb
Halder et al., 2012	1	Discordant for DiGeorge syndrome with major cardiac malformation and cardiac failure	0	Fluorescence in situ hybridization (FISH) & Illumina HumanCytoSNP-12 BeadChip	1
Solomon et al., 2012	2	Discordant for VACTERL association-type congenital malformations	Not reported	Illumina Omni1-Quad SNP array; Agilent SureSelect Human All Exon	0
Veenma et al., 2012	11	Discordant for esophageal atresia (N = 7), discordant for congenital diaphragmatic hernia (N = 4)	Not reported	Illumina HumanCytoSNP-12	0
Baudisch et al., 2013	4	Discordant for anorectal malformations (N = 3) or for bladder exstrophy-epispadias complex (N = 1)	0, 5, 9, 13	HumanOmni1-Quad v1	0
Bloom et al., 2013	5	Discordant for schizoaffective disorder (N = 2), bipolar disorder (N = 2) or schizophrenia (N = 1)	>60	Roche Nimblegen 2.1 Mprobe CGH array	0
Furukawa et al., 2013	1	Discordant for systemic lupus erythematosus	38	Illumina Human 610-Quad BeadChips	0
Miyake et al., 2013	1	Discordant for Rett syndrome	13	Illumina HumanOmni2.5-Quad, Illumina HiSeq2000	0
Laplana et al., 2014	1	Discordant for autism spectrum disorder	24	Agilent 400 K CGH array	0
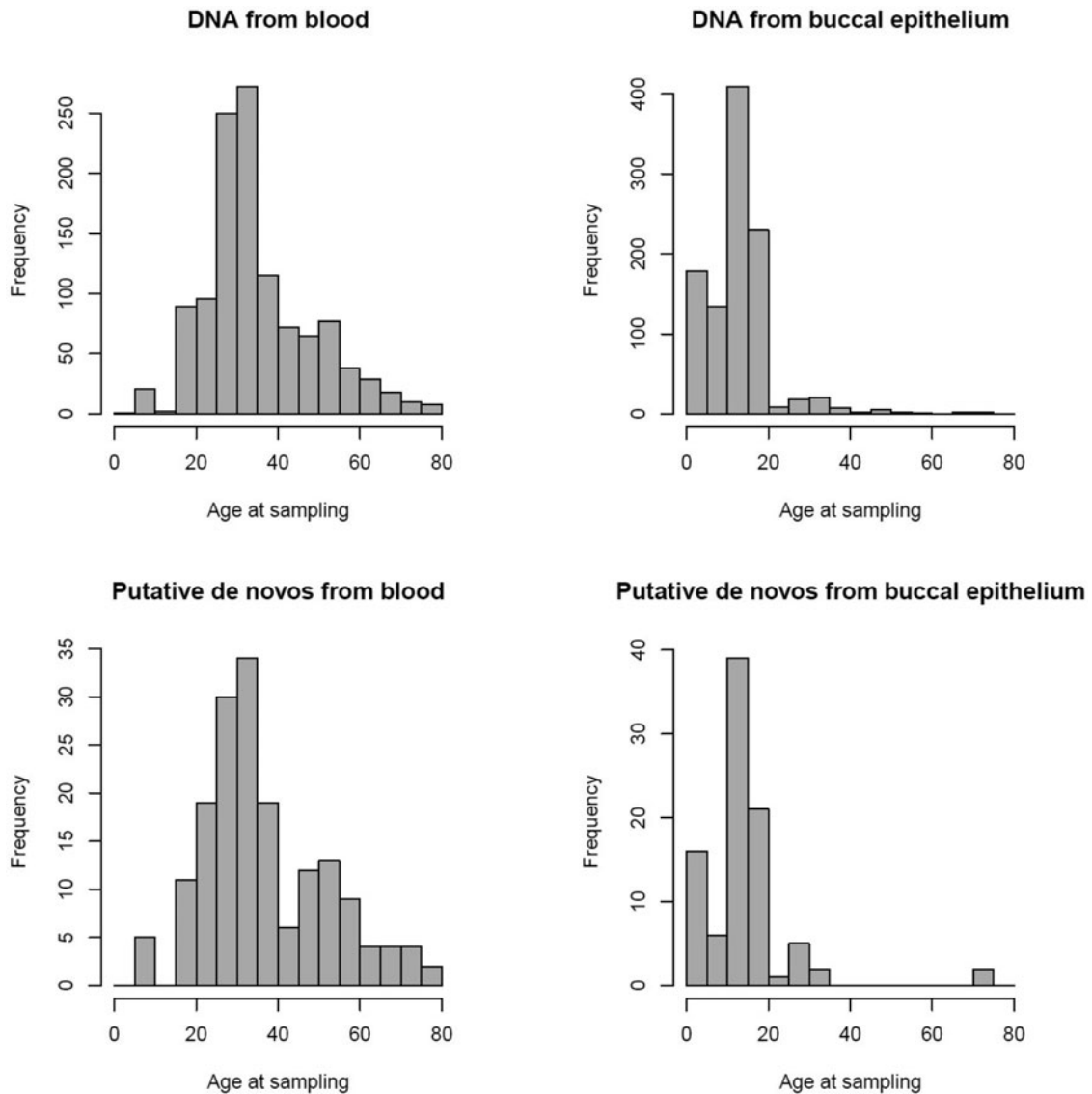
FIGURE 1

Age distribution of twins per tissue type for all samples (above) and the 152 putative de novo CNVs (below).

PennCNV was used to call genotypes, extract allele-specific signal intensities, cluster canonical genotypes, and finally generate a standard input file including log-R ratio (LRR) values and the 'B allele' frequency (BAF) for each marker in each individual. PennCNV uses a HMM-based approach for kilobase-resolution detection of CNVs. We followed the recommendation from the manual in creating batches (http://www.openbioinformatics.org/penncnv/penncnv˙tutorial˙affy˙gw6.html), and processed as many samples per calling batch as possible, resulting in four batches (one batch including all twins and duplicates with $N = 4{,}182$, and three batches with $N = 3{,}002$ per batch).

The CN calls of Birdsuite and PennCNV were compared with a script written in Perl. CN segments were only in-

cluded in further analyses if the following conditions were met: (1) the CN calls agreed between both algorithms, (2) the overlapping part of the segments from both algorithms was > 100 kb, and (3) the segment was not in a centromere. Calls were also included if the CN call in Birdsuite was equal to the expected CN (CN $= 2$) and the segment was not present in the PennCNV output, since PennCNV only gives the CN state when the CN deviates from the expected CN, and Birdsuite gives CN states for all segments. Since calling algorithms can produce artificially split CNV calls, adjacent CNV calls were merged after manual inspection of LRR and BAF plots, if the gap in between was $\leq 50\%$ of the entire length of the newly merged CNV.

Individuals were excluded from CNV calling if they had: (1) contrast QC < 0.4 (CQC, a quality metric from

Affymetrix representing how well allele intensities separate into clusters); (2) SNP missingness > 10%; (3) had excess genome-wide heterozygosity/inbreeding levels (F, as calculated in PLINK (Purcell et al., 2007) on an LD-pruned set, must be greater than -0.10 and smaller than 0.10); (4) if they had >50 CNVs with CN $\neq$ 2. After QC, 12,559 samples remain with a mean CQC of 2.17 (datasets are considered problematic if the mean CQC is smaller than 1.70).

Identifying Putative Post-Twinning De Novo CNVs

CN calls of complete MZ twin pairs passing QC ($N =$ 1,097, mean CQC = 2.25) were analyzed to detect possible post-twinning de novo CNV events. Segments with CN differences between MZ twins were extracted with a purpose written Perl script, which compares segments with the same start and end positions between twins, as well as overlapping segments.

As an additional quality control, LRR and BAF plots were created for the putative de novo CNV segments and were visually inspected by AA and EE. CNVs with LRR and BAF plots that showed the strongest discordance were chosen for qPCR validation candidates.

qPCR Validation for Putative Post-Twinning De Novo CNVs

Calibrator sample selection. We selected a sample with CN = 2 in the regions included in the qPCR experiments as a calibrator sample, which was used to calibrate the qPCR assay to what a signal from CN = 2 should look like. Calibrator samples were selected using Affymetrix 6.0 and next generation sequence data from the partially overlapping NTR-GoNL (Boomsma et al., 2014) database (total overlap between the NTR-Affymetrix 6 and GoNL dataset = 81 samples). For these 81 individuals, we first selected samples that showed CN = 2 in Birdsuite and no call from PennCNV in the candidate regions. From this set, we then selected samples that showed no CN calls in the GoNL sequence data for two CNV calling algorithms, CNVnator (Abyzov et al., 2011) and DWAC-seq (http://tools.genomes.nl/dwac-seq.html), since these algorithms, like PennCNV, only make calls when CN $\neq$ 2. After visual inspection of the LRR & BAF plots for the remaining samples, we then selected one calibrator sample with CN = 2 for the qPCR experiments.

CNV Confirmation by qPCR. Samples identified as possible carriers of post-twinning de novo CNVs ($N =$ 20 MZ pairs) were removed from -20˚C storage at the Avera Institute for Human Genetics, quantitated using Qubit 2.0 Broad Range Assay (Life Technologies, Carlsbad, CA), and normalized to 5ng/µl. Proposed CNVs were validated using qPCR. Four TaqMan Copy Number Assays (see Table 3) were run on a Viia7 real-time PCR machine (Life Technologies, Carlsbad, CA). TaqMan Copy Number Reference Assay RNase P (Life Technologies, Carlsbad, CA) was used as an internal reference because it is known to exist in two copies in a diploid genome. The copy number assay reporter was FAM and the RNase P reference assay reporter was VIC. All four assays were performed on genomic DNA and run in 384 well PCR plates, with individual reaction volumes of 10 µl. Each sample was run with four replicates for accuracy. The four assay plates each contained the respective CNV candidates along with one non-template control sample and one calibrator sample (CN = 2). Using ViiA7 Software v1.2, the Ct threshold was set to manual with a value of 0.2 and auto-baseline was selected to 'ON'. PCR conditions included an initial hold at 95°C for 10 min, and then 95°C for 15 s followed by 60°C for 1 min, together repeated for 40 cycles.

Data generated from the four CNV assays were analyzed with CopyCaller Software v2.0 (Life Technologies, Carlsbad, CA). Ct values from both the copy number assay and the reference assay were exported as (.txt) files to CopyCaller. Analysis settings incorporated a calibrator sample with CN = 2. Comparative Ct ($\Delta\Delta$Ct) relative quantitation analysis was performed and sample copy numbers were called using the software algorithm. The $\Delta\Delta$Ct analysis method first determines the difference in Ct value (ΔCt) between the target regions and the reference assay, then it determines the difference between those ΔCt values and the calibrator sample ($\Delta\Delta$Ct). With this information, the CopyCaller Software generates both a calculated and a predicted CN value.

Statistical Analyses

CNV discordance within MZ pairs. Pearson correlations of LOD-scores between co-twins and Pearson correlations of the number of probes between co-twins were computed in IBM SPSS Statistics 21. A chi-squared test was conducted in order to test whether the putative de novo CNVs were associated with the source of DNA. The difference in age between samples that showed a putative de novo CNV and the rest of the samples was tested with a t test for blood and buccal epithelium separately.

CNV concordance within MZ pairs. We performed chi-squared tests in IBM SPSS Statistics 21 to test whether CNV calls with CN $\neq$ 2 were equally concordant within MZ pairs for three groups of twin pairs: twins pairs with DNA from blood, twin pairs with DNA from epithelium, and twin pairs were one twin had DNA from blood and the other from epithelium. A CNV was regarded as concordant if the overlap between MZ pairs was > 100 kb. A total of 4,415 deletions and 3,037 duplications were included in these analyses. It was also tested whether the total number of concordant CNVs with CN $\neq$ 2 per twin pair differed between these three groups of twin pairs with a one-way ANOVA.

Concordant CNVs versus psychiatric symptoms. A gene-enrichment test was performed in PLINK (Purcell et al., 2007, Raychaudhuri et al., 2010) in order to test whether CNV events impacting gene sets involved in neuronal processes were associated with TP or AP. TP and AP were measured longitudinally with the adult self-report questionnaires (Achenbach & Rescorla, 2003), which is part of the Achenbach System of Empirically Based Assessment. The maximum TP and AP scores over four measurement time points was used for the gene-enrichment analyses. We randomly selected one twin per MZ pair, unless one twin had a missing phenotype, in which case the twin with the non-missing phenotype was selected (TP $N = 674$, AP $N = 461$). The gene sets involved in neuronal processes were downloaded from the Molecular Signature Database, and were derived from the GO Biological Process Ontology (http://www.geneontology.org/GO.process.guidelines.shtml), and included genes involved in the generation of neurons (83 genes), neuron development (61 genes), neuronal differentiation (76 genes), and neuron apoptosis (17 genes).

Results

CNV Discordance within MZ Pairs

There were 556 CNV segments that showed a CN discordance between MZ twins >100 kb. The LOD-scores from the Birdsuite calls (a quality metric indicating the likelihood of a CNV relative to no CNV in the region) showed a significant negative correlation within twin pairs ($r = -0.247$, $p = 3.5 \times 10^{-9}$), as did the number of probes encompassing the CNV ($r = -0.248$, $p = 3.1 \times 10^{-9}$), indicating systematic quality differences in CNV calls within twin pairs. More than 70% of these calls ($N = 400$) showed an overlap of <10% between twins from the same twin pair (note that the overlap was >100 kb). This indicates that many CN discordances may be caused by inaccurate CNV breakpoint estimates and/or a quality difference in CN calls. After only including CNVs with an overlap between twins of >10%, 153 putative de novo CNVs >100 kb remained, of which the correlations of the LOD-scores and number of probes between co-twins were no longer significant ($r = -0.029$, $p = .724$, and $r = -0.036$, $p = .654$, respectively). Of these 153 CNVs, more than half ($N = 90$; 58.8%) were from chromosome 15q11.2, ranging from bp positions 18,466,953 to 20,776,822 (build 36). LRR and BAF plots were generated for both twins for all 153 CNVs. These LRR and BAF plots were inspected manually in order to select putative de novo CNVs suited for qPCR replication. Twenty CNVs were chosen based on discordance in the LRR&BAF plots (inspected by AA and EE), of which 19 were in chromosome 15q11.2, and were followed up with qPCR validation experiments.

Two CNVs in the same twin pair showed a CNV discordance in the qPCR experiments for two CNVs in 15q11.2 (~350 kb in 18,491,920–18,841,578, and ~280 kb in 19,090,388–19,369,260; see Table 2 and Figure 2). The twin pair was 13 years old at the time of sampling, and their DNA was extracted from a buccal epithelium sample. They do not show large phenotypic differences with respect to overall health, behavior, (birth) length, (birth) weight, or other physical appearance in longitudinal parental and self-report questionnaires from age 1 to 21. The twin with CN = 3 for both CNVs (twin 2 in Table 2 and Figure 2) did perform better in school and finished high school two levels higher than the twin with CN = 1 and CN = 2, consistent with their CITO (http://www.cito.nl/) score difference (10 points higher for twin 2). Of the remaining 18 non-replicated de novo CNVs, 17 were due to a failure to detect a CNV with CN $\neq$ 2 in one of the twins (Table 2).

The remaining 133 putative de novo CNVs were not independent from the source of DNA, $\chi^2(2) = 7.91$, $p = .019$. Post-hoc tests showed that this was due to de novo CNVs being found significantly more often in DNA from blood than in DNA from buccal epithelium; 65.5% had blood-derived DNA; $\chi^2 (1) = 7.77$, $p = .005$. As nearly all young twins were done on buccal epithelium, and adult samples in blood, we checked whether the age difference might have contributed to the overrepresentation of blood-derived samples. For both blood and buccal epithelium samples, samples with a putative de novo CNV showed a higher average age than the rest of the samples from the same source without a putative de novo CNV (38.63 vs. 35.70 for blood; 14.84 vs. 12.67 for buccal epithelium samples), but these differences were not significant ($p = .119$ and $p = .294$ respectively).

CNV Concordance within MZ Pairs

Figure 3 shows the percentage of CNVs that were concordant between MZ pairs for deletions and duplications for each source of DNA. The percentages were ~80% for all three groups: DNA from blood for both twins, epithelium for both twins, and one twin from blood and one from epithelium. A one-way ANOVA showed that the small differences in concordance rates between different sources of DNA were significant; deletions: $\chi^2(2) = 8.69$, $p = .013$; duplications: $\chi^2(2) = 20.24$, $p = 4 \times 10^{-5}$. Post-hoc tests showed these differences to be significant between blood and buccal epithelium-derived samples only (deletions: $p = .012$; duplications: $p = 7 \times 10^{-6}$), with buccal epithelium-derived samples showing a slightly higher concordance rate. Note that there were very few twin pairs where one twin had his/her DNA extracted from blood and the other from epithelium ($N = 31$ pairs), which likely makes a comparison between this group and the other two groups underpowered.

There was also a significant difference between the different sources in the total number of concordant CNVs per twin, $F(1, 1,094) = 7.24$, $p = .001$. A post-hoc test

TABLE 2

CNV Calls from Affymetrix 6.0 and qPCR Experiments for 20 Putative De Novo CNVs

ID	Source	Age	chr	start	end	size	CN call Affy 6	qPCR results CN Predicted	CN Calculated	Min CN	Max CN
1–1	BLOOD	18	15	18,527,876	18,700,552	172,676	4	3	3.14	2.80	3.46
1–2	BLOOD	18	15	18,527,876	18,700,552	172,676	2	3	3.38	3.23	3.53
2–1	BUCCAL	13	15	18,846,002	19,379,050	533,048	3	4	3.77	3.37	4.11
2–2	BUCCAL	13	15	18,846,002	19,379,050	533,048	2	4	3.72	3.47	3.92
3–1[a]	**BUCCAL**	**13**	**15**	**18,491,920**	**18,841,578**	**349,658**	**1**	**1**	**1.07**	**0.96**	**1.12**
3–2[a]	**BUCCAL**	**13**	**15**	**18,491,920**	**18,841,578**	**349,658**	**2**	**3**	**2.54**	**2.42**	**2.67**
3–1[b]	**BUCCAL**	**13**	**15**	**19,090,388**	**19,369,260**	**278,872**	**1**	**2**	**1.54**	**1.47**	**1.64**
3–2[b]	**BUCCAL**	**13**	**15**	**19,090,388**	**19,369,260**	**278,872**	**3**	**3**	**3.29**	**2.77**	**3.60**
4–1	BUCCAL	12	15	18,700,552	19,090,596	390,044	3	3	3.46	3.13	3.65
4–2	BUCCAL	12	15	18,700,552	19,090,596	390,044	2	4	3.76	3.50	4.00
5–1	BUCCAL	13	15	18,816,444	19,090,596	274,152	3	3	2.54	2.39	2.62
5–2	BUCCAL	13	15	18,816,444	19,090,596	274,152	2	2	2.43	2.39	2.46
6–1	BLOOD	34	15	18,544,080	18,841,578	297,498	2	1	1.20	1.08	1.27
6–2	BLOOD	34	15	18,544,080	18,841,578	297,498	1	1	1.20	1.12	1.31
7–1	BLOOD	29	15	18,522,250	18,810,700	288,450	1	1	1.11	1.08	1.15
7–2	BLOOD	29	15	18,522,250	18,810,700	288,450	2	1	1.08	1.01	1.13
8–1	BLOOD	50	15	19,103,576	19,438,545	334,969	2	4	3.64	3.60	3.72
8–2	BLOOD	50	15	19,103,576	19,438,545	334,969	3	4	3.66	3.56	3.89
9–1	BLOOD	25	15	19,107,947	19,466,195	358,248	2	3	3.48	3.38	3.75
9–2	BLOOD	25	15	19,107,947	19,466,195	358,248	3	3	3.47	3.34	3.60
10–1	BLOOD	40	15	18,866,712	19,488,342	621,630	1	2	1.76	1.60	1.96
10–2	BLOOD	38	15	18,866,712	19,488,342	621,630	2	2	2.04	1.98	2.09
11–1	BLOOD	57	15	18,522,250	18,906,699	384,449	2	3	3.46	3.32	3.57
11–2	BLOOD	57	15	18,522,250	18,906,699	384,449	3	3	3.36	3.25	3.54
12–1	BLOOD	30	15	18,850,041	19,555,435	705,394	3	3	3.17	2.99	3.26
12–2	BLOOD	30	15	18,850,041	19,555,435	705,394	2	3	3.35	3.14	3.60
13–1	BLOOD	32	15	18,491,920	18,841,578	349,658	2	3	3.30	3.09	3.43
13–2	BLOOD	32	15	18,491,920	18,841,578	349,658	3	3	3.28	3.19	3.32
14–1	BLOOD	29	15	18,824,589	19,387,323	562,734	2	3	3.01	2.94	3.09
14–2	BLOOD	29	15	18,824,589	19,387,323	562,734	3	3	3.04	2.87	3.16
15–1	BUCCAL	11	15	19,066,371	19,4883,42	421,971	3	4	3.68	3.60	3.75
15–2	BUCCAL	11	15	19,066,371	19,488,342	421,971	2	4	3.62	3.51	3.73
16–1	BUCCAL	9	15	19,103,576	19,466,195	362,619	4	5	5.01	4.74	5.38
16–2	BUCCAL	9	15	19,103,576	19,466,195	362,619	3	5	4.54	4.40	4.73
17–1	BLOOD	34	15	18,977,400	19,338,198	360,798	3	3	3.11	2.93	3.25
17–2	BLOOD	34	15	18,977,400	19,338,198	360,798	2	3	3.15	3.07	3.26
18–1	BLOOD	32	15	19,066,371	19,488,342	421,971	2	4	3.54	3.37	3.69
18–2	BLOOD	32	15	19,066,371	19,488,342	421,971	3	4	3.62	3.54	3.66
19–1	BUCCAL	13	9	43,255,666	43,720,905	465,239	3	3	2.95	2.69	2.85
19–2	BUCCAL	13	9	43,255,666	43,720,905	465,239	2	3	3.07	3.17	3.32

Note: Note: [a,b]See Figures 2a and 1b, respectively for LRR & BAF plots for these two CNVs. Bold type = CNV discordance in qPCR experiments.

TABLE 3

TaqMan Copy Number Assay Names and Chromosome Locations

Copy number assay name	Chromosome location (hg 19)
Hs04452190_cn	Chr15:20,602,985
Hs03908328_cn	Chr15:20,394,238
Hs03900971_cn	Chr15:20,549,958
Hs04450511_cn	Chr15:20,863,835
Hs03724389_cn	Chr9:43,505,886

showed that this was due to blood-derived samples showing significantly more CNVs per twin (mean = 2.88, SD = 1.74) than epithelium-derived samples (mean = 2.48, *SD* = 2.01; *p* = .001). Twin pairs discordant for source of DNA showed 3.19 CNVs per twin on average (SD = 1.76). The difference in the total number of concordant CNVs between sources was also significant when analyzing deletions and duplications separately; deletions: $F(2, 1,094) = 4.21$, $p = .015$; duplications: $F(2, 1,094) = 5.04$, $p = .007$.

Concordant CNVs versus Psychiatric Symptoms

CNVs that were concordant within MZ pairs were tested for association with AP (ADHD symptoms) and TP (schizo-obsessive symptoms) using the gene-enrichment test in PLINK (Purcell et al., 2007, Raychaudhuri et al., 2010). The enrichment was tested for all genes, and gene-sets involved in generation of neurons (83 genes), neuron development (61 genes), neuronal differentiation (76 genes), and neuron apoptosis (17 genes).

The only significant association was observed between AP and the gene set involved in neuronal apoptosis ($p = 4 \times 10^{-39}$). This association disappeared after permutations. Permutations (10 k) were performed within four clusters based on gender and source of DNA.

Discussion

We searched for post-twinning de novo CNV mutations >100 kb in ~1,100 unselected MZ twin pairs using the Affymetrix 6.0 microarray. CNVs were called using two
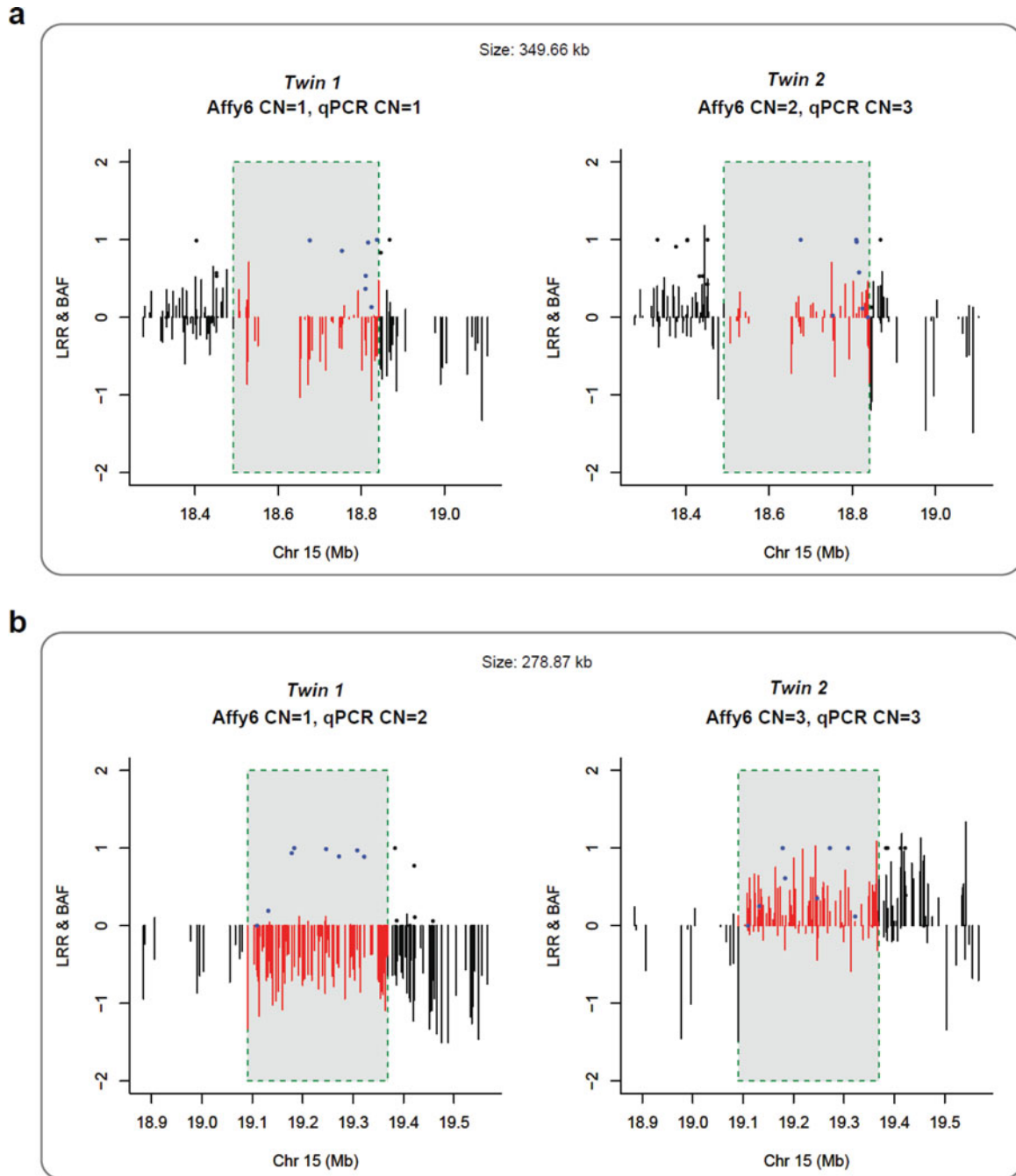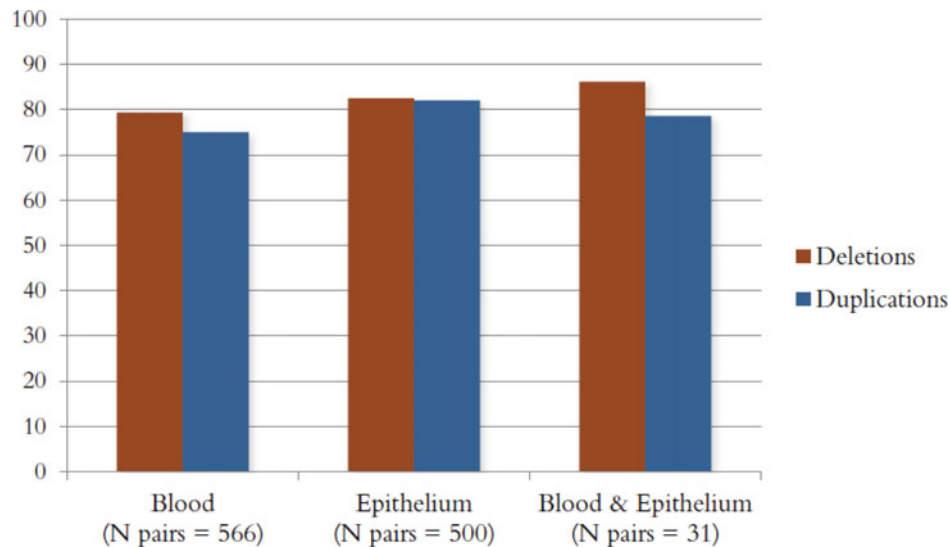
FIGURE 2

(Colour online) Log-R Ratio (LRR) for CNV probes & B allele frequency (BAF) for SNP probes of the two validated post-twinning mutations in the same 13-year-old twin pair (see Table 1 for bp positions and more details on the qPCR results for a and b, respectively). LRR is shown in vertical bars and BAF in solid points. The LRR & BAF values are shown in color in the region of the post-twinning de novo CNV (red and blue, respectively), and in black in the flanking regions.

algorithms, which resulted in 153 putative de novo CNVs, of which the majority came from the 15q11.2 region. Twenty candidates, of which 19 were from 15q11.2, were selected for qPCR replication based on visual inspection of 153 LRR and BAF plots. Two were validated, suggesting the remaining 133 putative de novo mutations also likely contain a substantial proportion of false positives. The large major-ity of non-replicated de novo CNVs (17 out of 18) are due to a failure to detect a CNV with CN ≠ 2 in one of the twins (Table 2). The significant overrepresentation of blood-derived samples among the remaining 133 putative somatic mutations may be explained by quality differences between blood- and buccal epithelium-derived samples, but may partly be explained by true mutations that increase

FIGURE 3

(Colour online) The percentage of CNVs that was concordant within MZ pairs for three groups: DNA from blood for both twins, epithelium for both twins, and one twin from blood and one from epithelium.

with age, as (1) blood-derived samples were predominantly adult as opposed to buccal-derived samples, (2) carriers of putative de novo CNVs from both blood and buccal epithelium showed a higher average age than the rest of the samples from the same tissue (although non-significant), and (3) previous studies have shown that de novo mutations increase with age (Forsberg et al., 2012, Kong et al., 2012, Ye et al., 2013).

Two post-twinning CNVs in 15q11.2 were replicated in a young MZ twin pair that showed no large phenotypic differences. CNVs in 15q11.2 have been associated with Prader–Willi and Angelman syndromes (Donlon, 1988), schizophrenia (Stefansson et al., 2008), behavioral disturbances (Doornbos et al., 2009), developmental and language delay (Burnside et al., 2011), epilepsy (de Kovel et al., 2010), and more recently with decreased fecundity, dyslexia, dyscalculia, and brain structure changes that are associated with schizophrenia and dyslexia (Stefansson et al., 2014). The 15q11.2 region is one of the genomic regions rich in segmental duplications (Zody et al., 2006), which makes CNVs in these regions harder to detect and therefore more likely to contain false positives, but also means this region is enriched for CNVs and more prone to de novo CNV mutations through non-allelic homologous recombination (Redon et al., 2006).

MZ twins provide the opportunity for an extra QC step for the relatively noisy microarray CNV data. About 80% of CNV calls were concordant between MZ pairs. It was difficult to judge which source of DNA is more suitable for CNV detection, as buccal epithelium-derived DNA showed a significantly higher concordance rate between MZ pairs, but blood-derived DNA allowed us to pick up significantly more concordant CNVs per twin pair. It was clear, however, that it is important to account for the source of DNA in association analyses, as a highly significant association between AP and CNVs affecting genes involved in neuronal apoptosis disappeared after accounting for source of DNA. Besides a relatively small sample size, another reason for not replicating associations with psychiatric symptoms may be false negative CNV calls in one of the twins. Since nearly all discordant CNVs that were included in the qPCR experiments (17 out of 18, excluding the replicated de novo CNVs) showed either a deletion or duplication in both twins, it is likely that a substantial part of the CNVs that showed a discordance within the MZ pairs reflect true CNV events (i.e., events with CN $\neq$ 2) that were missed by the CNV calling algorithm(s) in one of the twins. In other words, even though the confidence level of CNV calls is increased when only including concordant (i.e., replicated) CNVs, it may also result in missing true CNV calls.

In short, this study confirms the importance of qPCR replication when attempting to detect large post-twinning de novo CNVs and shows the importance of accounting for the source of DNA in studies using microarray CNV data. It is not clear yet why the 15q11.2 region is over-represented among CNVs discordant within twin pairs, since these may also reflect true post-twinning de novo CNVs. Association studies may also benefit from qPCR validation and genetic duplicates, as the large majority of discordant CNVs that were followed up with qPCR validation experiments turned out to be deletions or duplications that were concordant within MZ twin pairs.

Acknowledgments

References

Abdellaoui, A., Bartels, M., Hudziak, J. J., Rizzu, P., Van Beijsterveldt, T. C., & Boomsma, D. I. (2008). Genetic influences on thought problems in 7-year-olds: A twin-study of genetic, environmental and rater effects. *Twin Research and Human Genetics*, *11*, 571–578.

Abdellaoui, A., de Moor, M. H., Geels, L. M., van Beek, J. H., Willemsen, G., & Boomsma, D. I. (2012). Thought problems from adolescence to adulthood: Measurement invariance and longitudinal heritability. *Behavior Genetics*, *42*, 19–29.

Abyzov, A., Urban, A. E., Snyder, M., & Gerstein, M. (2011). CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Research*, *21*, 974–984.

Achenbach, T., & Rescorla, L. (2003). *Manual for the ASEBA adult forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Baranzini, S. E., Mudge, J., van Velkinburgh, J. C., Khankhanian, P., Khrebtukova, I., Miller, N. A., . . . Kim, R. W. (2010). Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature*, *464*, 1351–1356.

Baudisch, F., Draaken, M., Bartels, E., Schmiedeke, E., Bagci, S., Bartmann, P., . . . Reutter, H. (2013). CNV analysis in monozygotic twin pairs discordant for urorectal malformations. *Twin Research and Human Genetics*, *16*, 802–807.

Bloom, R. J., Kähler, 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.

Boomsma, D. I., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, *3*, 872–882.

Boomsma, D. I., De Geus, E. J., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J.-J., . . . Bartels, M. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics*, *9*, 849–857.

Boomsma, D. I., Wijmenga, C., Slagboom, E. P., Swertz, M. A., Karssen, L. C., Abdellaoui, A., . . . van Dijk, F. (2014). The genome of the Netherlands: Design, and project goals. *European Journal of Human Genetics*, *22*, 221–227.

Breckpot, J., Thienpont, B., Gewillig, M., Allegaert, K., Vermeesch, J., & Devriendt, K. (2012). Differences in copy number variation between discordant monozygotic twins as a model for exploring chromosomal mosaicism in congenital heart defects. *Molecular Syndromology*, *2*, 81–87.

Bruder, C. E., Piotrowski, A., Gijsbers, A. A., Andersson, R., Erickson, S., Diaz de Ståhl, T., . . . Poplawski, A. (2008). Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *The American Journal of Human Genetics*, *82*, 763–771.

Burnside, R. D., Pasion, R., Mikhail, F. M., Carroll, A. J., Robin, N. H., Youngs, E. L., . . . Papenhausen, P. R. (2011). Microdeletion/microduplication of proximal 15q11. 2 between BP1 and BP2: A susceptibility region for neurological dysfunction including developmental and language delay. *Human Genetics*, *130*, 517–528.

Cook, Jr. E. H., & Scherer, S. W. (2008) Copy-number variations associated with neuropsychiatric conditions. *Nature*, *455*, 919–923.

de Kovel, C. G., Trucks, H., Helbig, I., Mefford, H. C., Baker, C., Leu, C., . . . Ostertag, P. (2010). Recurrent microdeletions at 15q11. 2 and 16p13. 11 predispose to idiopathic generalized epilepsies. *Brain*, *133*, 23–32.

Derks, E. M., Hudziak, J. J., & Boomsma, D. I. (2009). Genetics of ADHD, hyperactivity, and attention problems. In Y. K. Kim (Ed.), *Handbook of behavior genetics* (pp. 361–378). New York: Springer.

Derks, E. M., Hudziak, J. J., Dolan, C. V., Ferdinand, R. F., & Boomsma, D. I. (2006). The relations between DISC-IV DSM diagnoses of ADHD and multi-informant CBCL-AP syndrome scores. *Comprehensive Psychiatry*, *47*, 116–122.

Donlon, T. (1988). Similar molecular deletions on chromosome 15q11. 2 are encountered in both the Prader–Willi and Angelman syndromes. *Human Genetics*, *80*, 322–328.

Doornbos, M., Sikkema-Raddatz, B., Ruijvenkamp, C. A., Dijkhuizen, T., Bijlsma, E. K., Gijsbers, A. C., . . . Kerstjens-Frederikse, W. (2009). Nine patients with a microdeletion 15q11. 2 between breakpoints 1 and 2 of the Prader–Willi critical region, possibly associated with behavioural disturbances. *European Journal of Medical Genetics*, *52*, 108–115.

Ehli, E. A., Abdellaoui, A., Hu, Y., Hottenga, J. J., Kattenberg, M., van Beijsterveldt, T., . . . Scheet, P. (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.

Forsberg, L. A., Rasi, C., Razzaghian, H. R., Pakalapati, G., Waite, L., Thilbeault, K. S., . . . Dumanski, J. P. (2012). Age-related somatic structural changes in the nuclear genome of human blood cells. *American Journal of Human Genetics*, *90*, 217–228.

Furukawa, H., Oka, S., Matsui, T, Hashimoto, A., Arinuma, Y., Komiya, A., . . . Tohma, S. (2013). Genome, epigenome and transcriptome analyses of a pair of monozygotic twins discordant for systemic lupus erythematosus. *Human Immunology*, *74*, 170–175.

Halder, A., Jain, M., Chaudhary, I., & Varma, B. (2012). Chromosome 22q11. 2 microdeletion in monozygotic twins with discordant phenotype and deletion size. *Molecular Cytogenetics*, *5*, 13.

Itsara, A., Wu, H., Smith, J. D., Nickerson, D. A., Romieu, I., London, S. J., . . . Eichler, E. E. (2010). De novo rates and selection of large copy number variation. *Genome Research*, *20*, 1469–1481.

Jakobsen, L. P., Bugge, M., Ullmann, R., Schjerling, C. K., Borup, R., . . . Tommerup, N. (2011). 500 K SNP array analyses in blood and saliva showed no differences in a pair of monozygotic twins discordant for cleft lip. *American Journal of Medical Genetics Part A*, *155*, 652–655.

Kasius, M. C., Ferdinand, R. F., Berg, H., & Verhulst, F. C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *38*, 625–632.

Kondo, S., Schutte, B. C., Richardson, R. J., Bjork, B. C., Knight, A. S., Watanabe, Y., . . . Murray, J. C. (2002). Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nature Genetics*, *32*, 285–289.

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P, Magnusson, G., . . . Jonasdottir, A. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature*, *488*, 471–475.

Korn, J. M., Kuruvilla, F. G., McCarroll, S. A., Wysoker, A., Nemesh, J., Cawley, S., . . . Darvishi, K. (2008). Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature Genetics*, *40*, 1253–1260.

Laplana, M., Royo, J. L., Aluja, A., López, 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*, vol. 2014, Article ID 516529, 9 pages.

Lasa, A., y Cajal, T. R., Llort, G., Suela, J., Cigudosa, J., Cornet, M., . . . Baiget, M. (2010). Copy number variations are not

modifiers of phenotypic expression in a pair of identical twins carrying a BRCA1 mutation. *Breast Cancer Research and Treatment, 123*, 901–905.

Lupski, J. R. (2007). Genomic rearrangements and sporadic disease. *Nature Genetics*, *39*, S43–S47.

Maiti, S., Kumar, K. H. B. G., Castellani, C. A., O'Reilly, R., & Singh, S. M. (2011). Ontogenetic de novo copy number variations (CNVs) as a source of genetic individuality: studies on two families with MZD twins for schizophrenia. *PLoS One*, *6*, e17125.

Miyake, K., Yang, C., Minakuchi, Y., Ohori, K., Soutome, M, Hirasawa, T, . . . Itoh, M. (2013). Comparison of genomic and epigenomic expression in monozygotic twins discordant for Rett syndrome. *PLoS One*, *8*, e66729.

Morgan, C. J., & Cauce, A. (1999). Predicting DSM-III-R Disorders from the Youth Self-Report: Analysis of data from a field study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*, 1237–1245.

Ono, S., Imamura, A., Tasaki, S., Kurotaki, N., Ozawa, H., Yoshiura, K.-I., . . . Okazaki, Y. (2010). Failure to confirm CNVs as of etiological significance in twin pairs discordant for schizophrenia. *Twin Research and Human Genetics*, *13*, 455–460.

Pamphlett, R., & Morahan, J. M. (2011). Copy number imbalances in blood and hair in monozygotic twins discordant for amyotrophic lateral sclerosis. *Journal of Clinical Neuroscience*, *18*, 1231–1234.

Piotrowski, A., Bruder, C. E., Andersson, R., Diaz de Stahl, T., Menzel, U., Sandgren, J., . . . Dumanski, J. P. (2008). Somatic mosaicism for copy number variation in differentiated human tissues. *Human Mutation*, *29*, 1118–1124.

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

Raychaudhuri, S., Korn, J. M., McCarroll, S. A., Altshuler, D., Sklar, P., Purcell, S., . . . Consortium, IS. (2010). Accurately assessing the risk of schizophrenia conferred by rare copy-number variation affecting genes with brain function. *PLoS Genetics*, *6*, e1001097.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., . . . Chen, W. (2006). Global variation in copy number in the human genome. *Nature*, *444*, 444–454.

Reumers, J., De Rijk, P., Zhao, H., Liekens, A., Smeets, D., Cleary, J., . . . Del-Favero, J. (2012). Optimized filtering reduces the error rate in detecting genomic variants by short-read sequencing. *Nature Biotechnology*, *30*, 61–68.

Sakuntabhai, A., Ruiz-Perez, V., Carter, S., Jacobsen, N., Burge, S., Monk, S., . . . Hovnanian, A. (1999). Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease. *Nature Genetics*, *21*, 271–277.

Sasaki, H., Emi, M., Iijima, H., Ito, N., Sato, H., Yabe, I., . . . Matsubara, K. (2011). Copy number loss of (src homology 2 domain containing)-transforming protein 2 (SHC2) gene: Discordant loss in monozygotic twins and frequent loss in patients with multiple system atrophy. *Molecular Brain*, *4*, 24.

Solomon, B., Pineda-Alvarez, D., Hadley, D., Hansen, N., Kamat, A., Donovan, F., . . . Mullikin, J. (2012). Exome sequencing and high-density microarray testing in monozygotic twin pairs discordant for features of VACTERL association. *Molecular Syndromology*, *4*, 27–31.

Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., . . . Doyle, O. M. (2014). CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature*, *505*, 361–366.

Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O. P., Ingason, A., Steinberg, S., . . . Buizer-Voskamp, J. E. (2008). Large recurrent microdeletions associated with schizophrenia. *Nature*, *455*, 232–236.

Vadlamudi, L., Dibbens, L. M., Lawrence, K. M., Iona, X., McMahon, J. M., Murrell, W., . . . Berkovic, S. F. (2010). Timing of de novo mutagenesis — A twin study of sodium-channel mutations. *The New England Journal of Medicine*, *363*, 1335–1340.

van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., . . . Schutte, N. (2013). The Young Netherlands Twin Register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics*, *16*, 252–267.

van Ommen, G.-J. B. (2005). Frequency of new copy number variation in humans. *Nature Genetics*, *37*, 333–334.

Veenma, D., Brosens, E., de Jong, E., van de Ven, C., Meeussen, C., Cohen-Overbeek, T., . . . Tibboel, D. (2012). Copy number detection in discordant monozygotic twins of congenital diaphragmatic hernia (CDH) and esophageal atresia (EA) cohorts. *European Journal of Human Genetics*, *20*, 298–304.

Wang, K., Li, M., Hadley, D., Lium, 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.

Weaver, S., Dube, S., Mir, A., Qin, J., Sun, G., Ramakrishnan, R., . . . Livak, K. J. (2010). Taking qPCR to a higher level: analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. *Methods*, *50*, 271–276.

Willemsen, G., De Geus, E. J., Bartels, M., Van Beijsterveldt, C., Brooks, A. I., Estourgie-van Burk, G. F., . . . Kluft, K. (2010). The Netherlands Twin Register biobank: A resource for genetic epidemiological studies. *Twin Research and Human Genetics*, *13*, 231–245.

Willemsen, G., Vink, J. M., Abdellaoui, A., den Braber, A., van Beek, J. H., Draisma, H. H., . . . van Lien, R. (2013). The Adult Netherlands Twin Register: Twenty-five years of survey and biological data collection. *Twin Research and Human Genetics*, *16*, 271–281.

Williams, N. M., Zaharieva, I., Martin, A., Langley, K., Mantripragada, K., . . . Gudmundsson, O. O. (2010). Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: A genome-wide analysis. *The Lancet*, *376*, 1401–1408.

Ye, K., Beekman, M., Lameijer, E. W., Zhang, Y., Moed, M. H., van den Akker, E. B., . . . Slagboom, P. E. (2013). Aging as accelerated accumulation of somatic variants: Whole-genome sequencing of centenarian and middle-aged monozygotic twin pairs. *Twin Research and Human Genetics*, *16*, 1026–1032.

Zhang, D., Qian, Y., Akula, N., Alliey-Rodriguez, N., Tang, J., Gershon, E. S., . . . Liu, C. (2011). Accuracy of CNV detection from GWAS data. *PLoS One*, *6*, e14511.

Zody, M. C., Garber, M., Sharpe, T., Young, S. K., Rowen, L., O'Neill, K., . . . Cuomo, C. A. (2006). Analysis of the DNA sequence and duplication history of human chromosome 15. *Nature*, *440*, 671–675.