

# One CNV Discordance in *NRXN1* Observed Upon Genome-wide Screening in 38 Pairs of Adult Healthy Monozygotic Twins

Patrik K. E. Magnusson,<sup>1</sup> Donghwan Lee,<sup>2</sup> Xu Chen,<sup>1</sup> Jin Szatkiewicz,<sup>3</sup> Setia Pramana,<sup>4</sup> Shumei Teo,<sup>1</sup> Patrick F. Sullivan,<sup>1,3</sup> Lars Feuk,<sup>5</sup> and Yudi Pawitan<sup>1</sup>

<sup>1</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

<sup>2</sup>Department of Statistics, Ewha Womans University, Seoul, South Korea

<sup>3</sup>Department of Genetics, University of North Carolina, Chapel Hill, NC, USA

<sup>4</sup>Institute of Statistics, Jakarta, Indonesia

<sup>5</sup>Department of Immunology, Genetics and Pathology, Rudbeck Laboratory and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Monozygotic (MZ) twins stem from the same single fertilized egg and therefore share all their *inherited* genetic variation. This is one of the unequivocal facts on which genetic epidemiology and twin studies are based. To what extent this also implies that MZ twins share genotypes in adult tissues is not precisely established, but a common pragmatic assumption is that MZ twins are 100% genetically identical also in adult tissues. During the past decade, this view has been challenged by several reports, with observations of differences in post-zygotic copy number variations (CNVs) between members of the same MZ pair. In this study, we performed a systematic search for differences of CNVs within 38 adult MZ pairs who had been misclassified as dizygotic (DZ) twins by questionnaire-based assessment. Initial scoring by PennCNV suggested a total of 967 CNV discordances. The within-pair correlation in number of CNVs detected was strongly dependent on confidence score filtering and reached a plateau of  $r = 0.8$  when restricting to CNVs detected with confidence score larger than 50. The top-ranked discordances were subsequently selected for validation by quantitative polymerase chain reaction (qPCR), from which one single ~120kb deletion in *NRXN1* on chromosome 2 (bp 51017111–51136802) was validated. Despite involving an exon, no sign of cognitive/mental consequences was apparent in the affected twin pair, potentially reflecting limited or lack of expression of the transcripts containing this exon in nerve/brain.

■ **Keywords:** copy number variation, *NRXN1*, monozygotic twins, PennCNV, quantitative polymerase chain reaction

One of the key assumptions in genetic modeling of twin data is that MZ twins are clonal copies, sharing 100% of their inherited genetic materials and thus identical to each other from an inherited genetic perspective (Neale & Cardon, 1992). In such modeling, twice of the difference in MZ versus DZ correlation coefficients typically provides an estimate of the additive genetic contribution to the total trait variability. The striking physical similarities and the perfect immunological fit in blood and tissue transplantation between members of the same MZ pair, as well as lack of evidence for genotypic mismatches over and above technical error rates, have historically lent considerable support for assumption of MZ genetic similarity.

Even so, mutations that might occur in the first cell divisions after embryo fragmentation will be present system-

ically in subsequent stem cells as well as differentiated cells in the affected twin, but be absent in the co-twin. Such mutations would be detectable as MZ genotype differences, independent of tissue and age of the pair. In contrast, mutations occurring later after fragmentation will affect a more differentiated cell and thereby more likely also just a subset of cells of a specific tissue. Such mosaic MZ genotype

RECEIVED 11 September 2015; ACCEPTED 11 January 2016. First published online 22 February 2016.

ADDRESS FOR CORRESPONDENCE: Patrik K. E. Magnusson, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, 17177, Stockholm, Sweden. E-mail: [patrik.magnusson@ki.se](mailto:patrik.magnusson@ki.se)

differences could occur in any proportion of cells and will therefore be much harder to detect with confidence. In this article, we focus on systemic MZ discordances, present in all cells of the affected individual.

One motivation for searching for MZ genotype differences is that they might explain MZ phenotypic discordances and in turn provide a 'shortcut' to functional variants. Thus, identification of genetic mismatches in MZ twin pairs that are discordant for a disease might directly point to the genetic variant(s) responsible for the disease. CNVs, which consist of deletions or duplications larger than 1 kilo base pairs (kb) (McCarroll & Altshuler, 2007) have been found to sometimes occur discordantly within phenotypically concordant and discordant MZ pairs (Bruder et al., 2008). CNVs have also been implicated to be involved in psychiatric and cognitive problems such as autism, schizophrenia, and mental retardation (Bassett et al., 2010; Chung et al., 2014; Szatkiewicz et al., 2014). Here, we performed a systematic search for differences of CNVs within 38 adult MZ pairs, by using signal intensities from 700K SNP markers on the Illumina Omni Express platform.

## Materials and Methods

### Samples

The study base was TwinGene, a substudy of the Swedish Twin Registry (Magnusson et al., 2013), in which the zygosity was initially determined by utilizing self-report answers about physical similarity in childhood. Samples selected for genotyping included both members of presumed DZ pairs but only one member of each presumed MZ pair. This resulted in genotyping of DNA from peripheral blood samples from 9,835 supposedly unique genomes with Illumina Omni Express 700K SNP chip. The genotyping was conducted by the SNP&SEQ genotyping facility at Uppsala University, Sweden. Among the genotyped samples, we found 38 presumed DZ pairs sharing close to 100% of their SNP alleles, representing misclassifications of zygosity based on questionnaire data collected before genotyping. These 38 MZ pairs constituted the main study population of this report. The mean age was 61 ( $SD = 8$ ) and 79% were female. We utilized the remaining twin sample (complete DZ pairs and incomplete MZ and DZ pairs) as a background population for data normalization and for assessing the recurrence of discovered CNVs.

### CNV Detection

CNV detection was first done using PennCNV software (Wang et al., 2007) for the total background population of 9,835 subjects. Standard data normalization procedures and canonical genotype clustering files provided by Illumina were used to process the genotyping signals. The signal intensities, that is, the Log R ratio (LRR) and B allele frequency (BAF), for all markers for all samples were directly calculated and exported from the Illumina BeadStudio soft-

ware. A total of 728,816 SNP markers were mapped to the hg19 (build 37) human genome assembly.

To generate the required population frequency of B allele (PFB) file, 300 high-quality unrelated samples (only one twin per pair) were randomly chosen from the study and the `compile_pfb.pl` program was applied to their signal intensity files. These high-quality samples had to pass: (1) SNP-based quality control filters (subject missingness  $< 0.02$ , no autosomal heterozygosity deviation); (2) intensity-based quality control filters (probe intensity variance  $LRR\_SD < 0.15$  and absolute value of waviness factor  $< 0.04$ ); and (3) should not belong to the samples displaying excessive number of CNVs ( $> 95$ th percentile, i.e.,  $> 92$  CNVs). To generate the required GC model file for genomic wave adjustment, the `cal_gc_snp.pl` was applied using the UCSC GC annotation file (`gc5Base.txt.gz` from <http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/>). The GC model file specifies the GC content of the 1Mb genomic region surrounding each marker (500kb each side). Using the recommended parameters for the OmniExpress chip (i.e., `hhall.hmm`), the PFB, and the genomic wave adjustment routine, CNV calls were generated for the 22 autosomes. The data (sample-set, experiments, CNV calls, and regions) of the 38 MZ pairs were submitted to Database of Genomic Variants archive (DGVA, study ID: `estd225`, <http://www.ebi.ac.uk/dgva/data-download>).

### Validation Method — Taqman qPCR

We used the PennCNV confidence score to rank the CNVs chosen for validation. The score is a number representing the likelihood that there is a CNV at a particular region; a higher number indicates a greater probability. Eleven top-ranked CNVs (six discordant and five concordant) present in 10 MZ pairs were selected for validation by quantitative PCR-based TaqMan<sup>®</sup> Copy Number Assays (Table S1). To increase reliability, triplicate runs were undertaken for each CNV of interest in each specific pair. Samples from eight other MZ pairs (for which no evidence for discordance was evident based on PennCNV) were tested for each particular CNV as negative controls, plus one reference control (RNase P gene) and two no template controls for each individual (Figure S1). Taqman assays were performed on three 96-well plates and run on Applied Biosystems<sup>®</sup> 7900HT real-time PCR instrument with parameters as 95°C 10min, 40 cycles of 95°C 15s, and 60°C 60s. The signals were captured and analyzed by CopyCaller<sup>™</sup> Software.

## Results

### Within-Pair MZ Correlation in Number of Putative CNVs

PennCNV algorithm detected a total of 1,917 putative CNVs among the 76 individuals of the 38 complete MZ pairs. The distribution of number of scored CNVs per sample ranged from 12 to 153 (Table 1), with a mean of 25

**TABLE 1**  
Spearman Correlations in Number of CNV Within 38 MZ Pairs

Conf score threshold <sup>a</sup>	Mean of $N_{\text{CNV}}^b$	Min-max of $N_{\text{CNV}}^c$	Spearman $r^d$	$p$ value
no	25.2	12–153	0.30	.06
>10	22.8	7–152	0.28	.09
>20	9.1	2–61	0.55	.0003
>30	4.9	0–23	0.62	<.0001
>40	3.6	0–12	0.79	<.0001
>50	2.9	0–10	0.84	<.0001
>60	2.5	0–10	0.80	<.0001

Note: <sup>a</sup>Confidence score estimated from PennCNV; <sup>b</sup>The mean number of CNVs by confidence score threshold; <sup>c</sup>The minimum and maximum number of CNVs by confidence score threshold; <sup>d</sup>Within-MZ pair correlation (Spearman coefficients) in number of CNVs by confidence score thresholds.

( $SD = 20$ ). There was also pronounced variation within MZ twin-pairs. The number of unfiltered CNVs was not significantly correlated within MZ pairs ( $p = .06$ ), but the correlation increased when we applied confidence score filters. For CNVs with confidence score >20 the within-pair correlation became significant ( $r = 0.55$ ,  $p = .0003$ ); and the correlation reached a plateau of approximately  $r = 0.8$  after raising the confidence score threshold to >50. Since we expect high within-pair correlation in number of true CNVs in MZ twins a priori, this result indicates that PennCNV detection is highly susceptible to false positives and that a confidence score >50 is needed to get acceptable specificity.

### Detecting Candidates for True CNV Discordances

In order to maximize the sensitivity to pick up true CNV discordances, we checked all initial CNVs ( $n = 1,917$ ) for any evidence of CNV discordance between the members within pairs (instead of ranking on individual CNV trustworthiness per se). The CNVs were first categorized as: (1) concordant ( $n = 562$ ): the regions matching perfectly within twin-pair; (2) overlapped ( $n = 388$ ): the regions overlapping but not matching perfectly within twin-pair; (3) discordant ( $n = 967$ ): the regions non-overlapping within the twin-pairs.

After ranking the MZ discordant CNVs, six were found to have confidence scores >50 and to be supported by at least 11 SNP markers (Table S2), they were selected for qPCR validation. The top ranked MZ CNV discordance was a deletion on chromosome 2 (bp 51017111–51136802) with confidence score of 124 and supported by 41 markers. This CNV was also strongly supported by the BAF pattern (Figure 1A). The deletion spans 120kb within the gene *NRXN1* and covers a 5 bp exon (Figures S2, S3) included in *nrxn1-201* and *nrxn1-202* protein coding transcripts. Among all the 9,835 genotyped twins in the base population, 107 CNVs (by PennCNV) were located within the *NRXN1* gene. Out of these, 31 had a high confidence score >50; hence there was strong evidence of recurrent CNVs in this region, a notion that was further supported by a comparison with previously reported CNVs in the DGVA

(Figure S4). The evidence for the other five MZ CNV discordances came solely from the LRR statistic, but lacked support from the BAF patterns (see Figure 1B and Figure S5A–S13A).

### Validation by qPCR

The *NRXN1* deletion was clearly validated by the Taqman qPCR (Figure 2A). As expected, none of the other five highest-confidence CNV discordances — for which there was no support from the BAF pattern — was validated. For all the concordant high-confidence CNVs, clear-cut validation was obtained (see Figure 2B and Figure S5B–S13B).

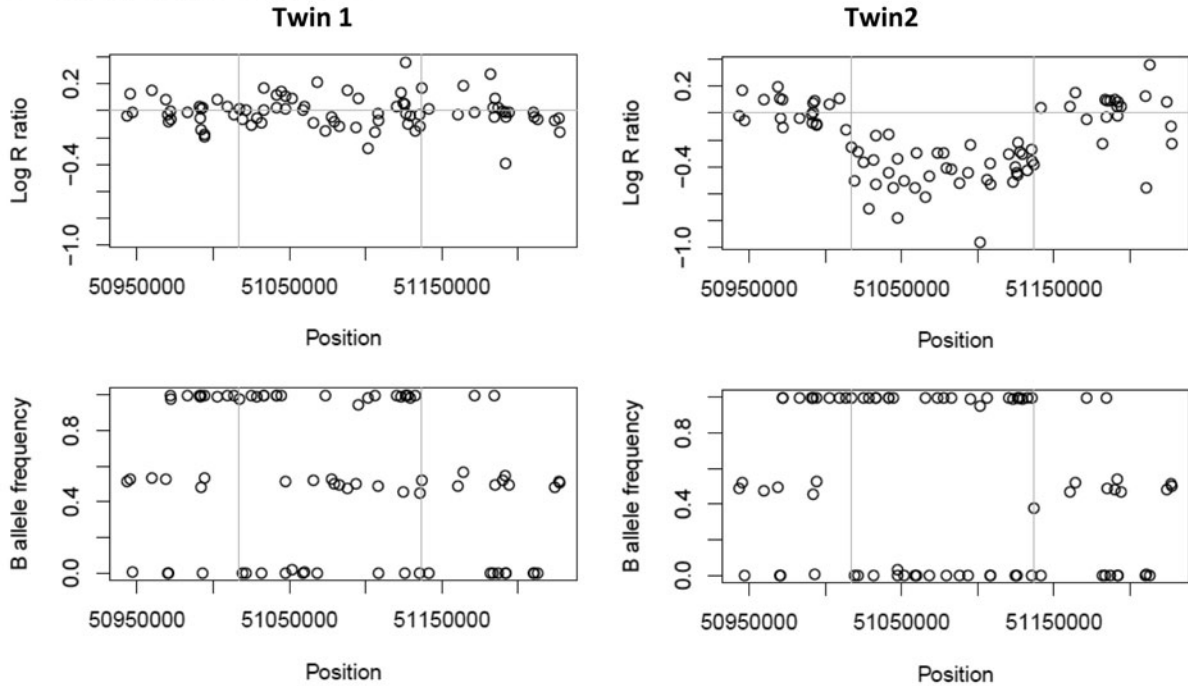
### Phenotypic Consequences of the *NRXN1* Deletion

*NRXN1* gene encodes neurexin 1, a neuronal adhesion molecule that plays an important role in synaptic function of brain (Kirov et al., 2009). Mutations in the gene, especially large chromosomal structural variations (e.g., CNVs) involving exons, have been reported to be associated with autism, schizophrenia, and other neurodevelopmental disorders (Ching et al., 2010; Curran et al., 2013; Kirov, 2015). The discordant *NRXN1* deletion region validated in this study was located within a part of the gene displaying recurrent CNVs reported in previous studies (Figure S4). In order to investigate whether the *NRXN1* CNV discordant twins displayed marked health discordance, we utilized data in the Swedish national patient registers (Ludvigsson et al., 2011). No neurological or psychiatric diagnoses had been recorded for any of the two members. Furthermore, both members of the pair have attained high university degree educations and have been employed in demanding occupations, speaking against neurological impact of the deletion.

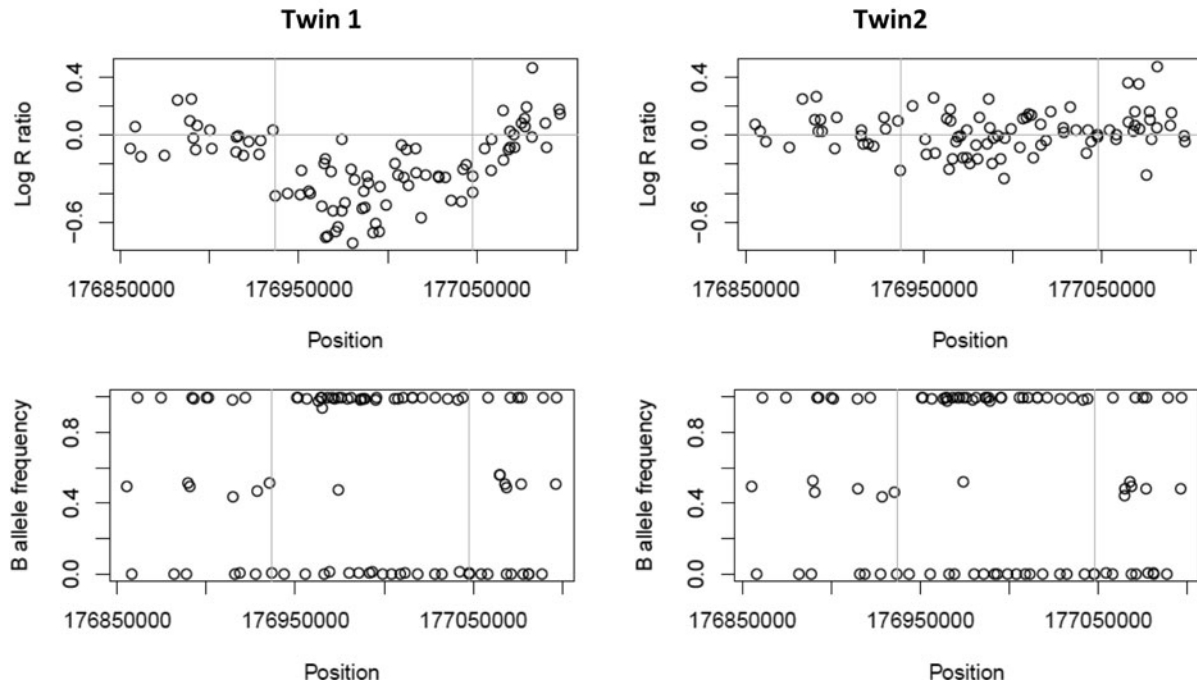
### Discussion

We searched for systemic CNV discordances within 38 MZ pairs and found one, consisting of a 120kb deletion within the *NRXN1* gene on chromosome 2. This was the only discordance that displayed convincing evidence from both LRR and BAF and that was subsequently validated by qPCR. In our base population of 9,835 subjects, we observed

**A. Top-ranked discordant CNV**



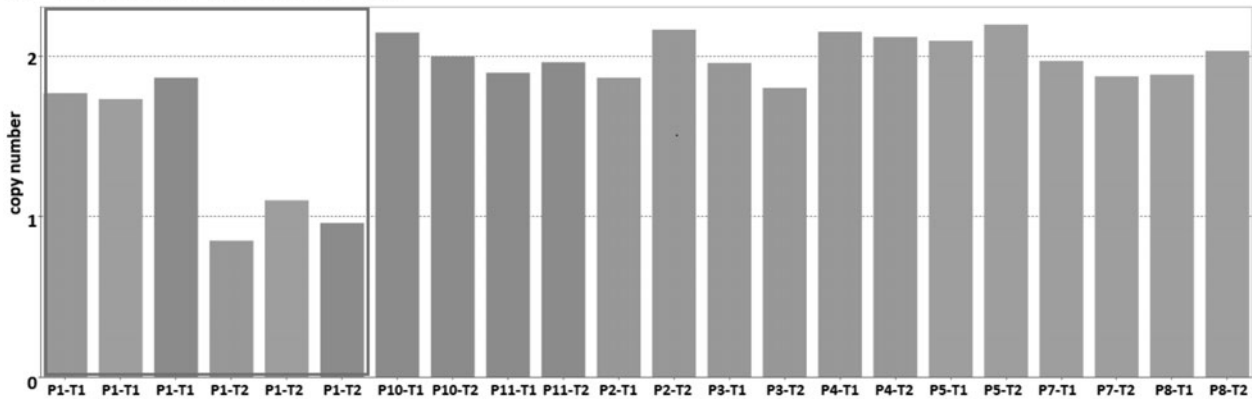
**B. Secondary-ranked discordant CNV**



**FIGURE 1**

Log R ratio (LRR) and B allele frequency (BAF) plot for the top- and secondary-ranked discordant CNV from PennCNV. (A) For the top-ranked discordant CNV (chr2: 51017111–51136802): located in the region of *NRXN1* gene, both LRR and BAF support a discordance (deletion in twin2); (B) For the secondary-ranked discordant CNV (chr2: 176936834–177047801): discordance is supported by LRR but not BAF.

## A. Top-ranked discordant CNV



## B. Secondary-ranked discordant CNV

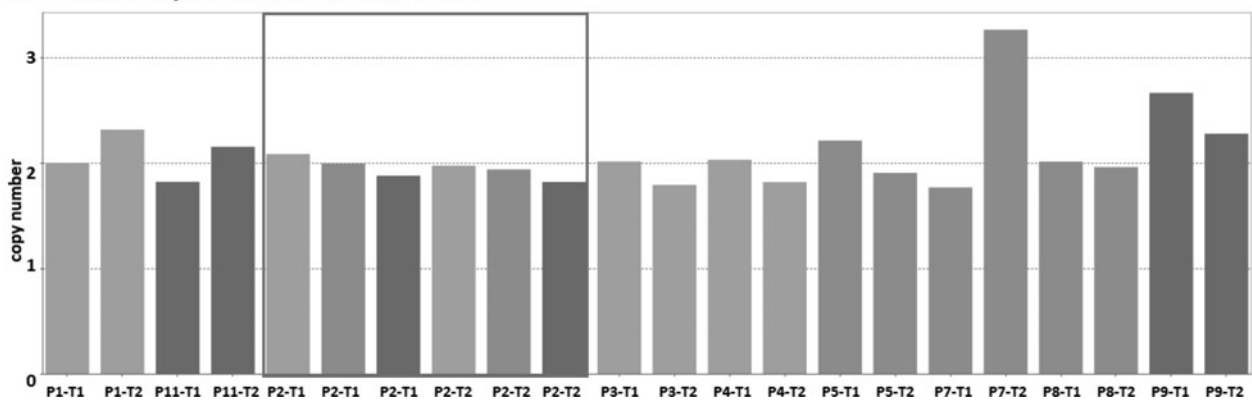


FIGURE 2

Taqman qPCR results for the validation attempts of the top- and secondary-ranked discordant CNV suggested by PennCNV. (A) In the top-ranked discordant CNV in *NRXN1*, the qPCR validates the deletion in twin 2; (B) Result from qPCR did not validate the secondary-ranked CNV discordance estimated from PennCNV. Both members possess the same normal two copies. Bars with the same pair-twin number indicate replicate runs (on 3 different 96-well plates). The boxes show the triplicate runs for the MZ pair in which the CNV implicated by PennCNV is tested. Bars outside the boxes are control samples for the particular CNV in question.

107 CNVs (by PennCNV) in the *NRXN1* gene, suggesting the gene is in a recurrent CNV region. Further, as judged by the distribution of previously reported CNVs within *NRXN1* in the DGVA, the discordant deletion we found is located in a part of the gene appearing to be often affected (Figure S4).

CNVs in the *NRXN1* gene have previously been shown to be associated with psychiatric outcomes (Kirov et al., 2009; Todarello et al., 2014) and *NRXN1* was recently ranked as number one out of 1,303 genes mapped to CNVs occurring in schizophrenia cases (Luo et al., 2014). Interestingly, a DZ twin-pair concordant for autism have recently been reported to carry bi-allelic *NRXN1* CNVs (Imitola et al., 2014). Even though the deletion verified as MZ discordant in the present study involved an *NRXN1* exon, there were no obvious phenotypic differences in the affected twin pair. The limited utilization of the exon (only included in two transcripts, *NRXN1*-201 and *NRXN1*-202 with expression limited to testis and liver) may explain the lack of

phenotypic consequence. It could also be that one copy suffices for full function.

The difficulty in finding and validating true systemic DNA differences within adult MZ twins illustrates the amazing fidelity of DNA replication, lending support to the common assumption of 100% shared genetics for MZ twins in quantitative genetic modeling from a general and practical point of view. However, as evident from validated CNV discordances, ontogenetic (developmental, after fertilization) mutations appearing completely systemic in blood do occur and may underlie phenotype discordances in specific individual cases.

Mutation rates between generations in humans are quite well established, but less is known about mitotic mutation rates. Since only mitotic post-twinning mutations will lead to MZ twin genetic discrepancies, the MZ twins are informative about this rate. However, careful consideration is needed for where and when the DNA differences in MZ pairs arise. When searching for systemic MZ

discordant mutations, we have to consider those that arise from early mitotic events during embryogenesis, in the blastula or morula stage. Nevertheless, non-systemic mutations may also appear as systemic when the analyses are restricted to single tissues. Such mutations may have occurred somewhat later in development (but before differentiation of the specific tissue(s) analyzed) or there may have been a drastic clonal expansion of the mutated cell (at the expense of non-mutated cells).

The literature provides a somewhat scattered picture of the occurrence of CNV discordances in MZ pairs. Among the most recent investigations, few validated MZ CNV discordances have been described. One study reported qPCR validation of two MZ CNV discordances in 1 out of 1,097 screened MZ pairs, both of which resided in 15q11.2 (Abdellaoui et al., 2015). Another study reported one MZ CNV discordance remaining after applying two different algorithms among 376 pairs (McRae et al., 2015).

Non-systemic differences may present as mosaic CNV differences in which only a fraction of cells in the sample harbor the CNV. In the literature describing such MZ CNV differences, the fraction of cells assumed to carry the CNV is typically optimized to fit the observed data, sometimes down to as low as 5% (Forsberg et al., 2012). Given the substantial level of noise in intensity data, such optimization is likely associated with increased type one error rates.

In the search for MZ genomic differences technical shortcomings are plentiful. After our failed attempts to validate MZ CNV differences that were only supported by probe intensity values, we concluded such CNVs are highly likely false. Here, we note that the statistically derived confidence score does not reflect the true false positive rate. The PenCNV algorithm, although generally recognized as one of the best algorithms for CNV detection from SNP array data, still has limitations that lead to false positive findings. This is well illustrated by a very modest duplicate concordance rate of 55–65% (Marenne et al., 2011; Zheng et al., 2012), indicating around 40% false positives. The tendency for a genome to get scored with many such false CNVs could in principle depend on characteristics of the inherited genome; for example, due to regions of extended homozygosity. However, we found no within-MZ pair correlation in number of unique CNVs (not shared with co-twin), providing no support for such a hypothesis. In this study, we have relied on CNVs estimated from SNP arrays. CNV calling based on next-generation sequencing is also highly challenging and not necessarily better than that based on SNP genotyping (Teo et al., 2012).

In conclusion, in a search for genomic differences within 38 MZ twin pairs, we found and validated one systemic CNV discordance occurring in the *NRXN1* gene, which is the first time MZ CNV discordance in *NRXN1* has been identified. Despite the *NRXN1* gene ranking among the top for displaying phenotypic consequences of CNVs, there was no sign of a phenotypic effect in the affected pair. Our

results also lend support to the overall view that MZ twin discordances for large systemic CNVs in blood are rare events.

## Acknowledgments

This work was supported by grants from the Swedish Research Council (grant number M-2005-1112); GenomEUtwin (grant numbers EU/QLRT-2001-01254, QLG2-CT-2002-01254); National Institutes of Health (grant number DK U01-066134); the Heart and Lung foundation (grant number 20070481). We thank the SNP&SEQ technology platform in Uppsala for excellent genotyping. The Swedish Twin Registry is financially supported by Karolinska Institutet. The authors report no conflict of interest.

## Supplementary Material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/thg.2016.5>.

## 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.
- Bassett, A. S., Scherer, S. W., & Brzustowicz, L. M. (2010). Copy number variations in schizophrenia: Critical review and new perspectives on concepts of genetics and disease. *American Journal of Psychiatry*, 167, 899–914.
- Bruder, C. E., Piotrowski, A., Gijsbers, A. A., Andersson, R., Erickson, S., Diaz de Stahl, T., . . . Dumanski, J. P. (2008). Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *American Journal of Human Genetics*, 82, 763–771.
- Ching, M. S., Shen, Y., Tan, W. H., Jeste, S. S., Morrow, E. M., Chen, X., . . . Children's Hospital Boston Genotype Phenotype Study Group. (2010). Deletions of *NRXN1* (neurexin-1) predispose to a wide spectrum of developmental disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 153B, 937–947.
- Chung, B. H., Tao, V. Q., & Tso, W. W. (2014). Copy number variation and autism: New insights and clinical implications. *Journal of the Formosan Medical Association*, 113, 400–408.
- Curran, S., Ahn, J. W., Grayton, H., Collier, D. A., & Ogilvie, C. M. (2013). *NRXN1* deletions identified by array comparative genome hybridisation in a clinical case series — Further understanding of the relevance of *NRXN1* to neurodevelopmental disorders. *Journal of Molecular Psychiatry*, 1, 4.
- Forsberg, L. A., Rasi, C., Razzaghiyan, 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.

- Imitola, J., Walleigh, D., Anderson, C. E., Jethva, R., Carvalho, K. S., Legido, A., & Khurana, D. S. (2014). Fraternal twins with autism, severe cognitive deficit, and epilepsy: Diagnostic role of chromosomal microarray analysis. *Seminars in Pediatric Neurology*, *21*, 167–171.
- Kirov, G. (2015). CNVs in neuropsychiatric disorders. *Human Molecular Genetics*, *24*, R45–49.
- Kirov, G., Rujescu, D., Ingason, A., Collier, D. A., O'Donovan, M. C., & Owen, M. J. (2009). Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophrenia Bulletin*, *35*, 851–854.
- Ludvigsson, J. F., Andersson, E., Ekblom, A., Feychting, M., Kim, J. L., Reuterwall, C., . . . Olausson, P. O. (2011). External review and validation of the Swedish national inpatient register. *BMC Public Health*, *11*, 450.
- Luo, X., Huang, L., Han, L., Luo, Z., Hu, F., Tieu, R., & Gan, L. (2014). Systematic prioritization and integrative analysis of copy number variations in schizophrenia reveal key schizophrenia susceptibility genes. *Schizophrenia Bulletin*, *40*, 1285–1299.
- Magnusson, P. K., Almqvist, C., Rahman, I., Ganna, A., Viktorin, A., Walum, H., . . . Lichtenstein, P. (2013). The Swedish Twin Registry: Establishment of a biobank and other recent developments. *Twin Research and Human Genetics*, *16*, 317–329.
- Marenne, G., Rodriguez-Santiago, B., Closas, M. G., Perez-Jurado, L., Rothman, N., Rico, D., . . . Malats, N. (2011). Assessment of copy number variation using the Illumina Infinium 1M SNP-array: A comparison of methodological approaches in the Spanish Bladder Cancer/EPICURO study. *Human Mutation*, *32*, 240–248.
- McCarroll, S. A., & Altshuler, D. M. (2007). Copy-number variation and association studies of human disease. *Nature Genetics*, *39*(Suppl.), S37–S42.
- 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.
- Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic.
- Szatkiewicz, J. P., O'Dushlaine, C., Chen, G., Chambert, K., Moran, J. L., Neale, B. M., . . . Sullivan, P. F. (2014). Copy number variation in schizophrenia in Sweden. *Molecular Psychiatry*, *19*, 762–773.
- Teo, S. M., Pawitan, Y., Ku, C. S., Chia, K. S., & Salim, A. (2012). Statistical challenges associated with detecting copy number variations with next-generation sequencing. *Bioinformatics*, *28*, 2711–2718.
- Todarello, G., Feng, N., Kolachana, B. S., Li, C., Vakkalanka, R., Bertolino, A., . . . Straub, R. E. (2014). Incomplete penetrance of NRXN1 deletions in families with schizophrenia. *Schizophrenia Research*, *155*, 1–7.
- 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.
- Zheng, X., Shaffer, J. R., McHugh, C. P., Laurie, C. C., Feenstra, B., Melbye, M., . . . Feingold, E. (2012). Using family data as a verification standard to evaluate copy number variation calling strategies for genetic association studies. *Genetic Epidemiology*, *36*, 253–262.