CNV concordance in MZ twin pairs

Abdel Abdellaoui, Erik A. Ehli, Jouke-Jan Hottenga, Zachary Weber, Hamdi Mbarek, Gonneke Willemsen, Toos van Beijsterveldt, Andrew Brooks, Jim J. Hudziak, Patrick F. Sullivan, Eco J. de Geus, Kai Ye, P. Eline Slagboom, Gareth E. Davies, Dorret I. Boomsma

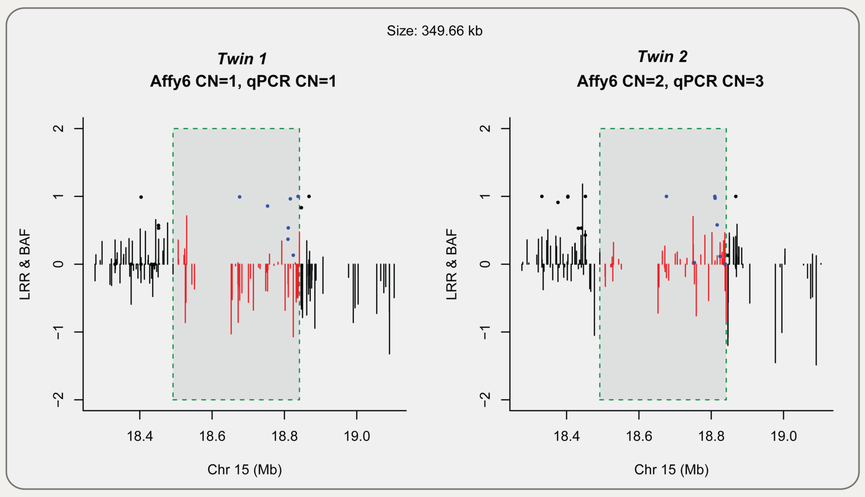
Introduction

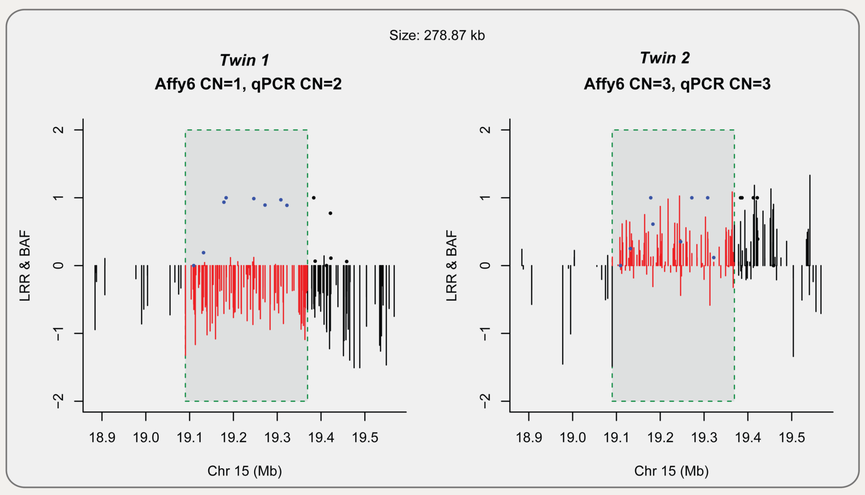
Monozygotic (MZ) twins are genetically identical at conception, making them informative subjects for studies on somatic mutations. Copy number variants (CNV) are responsible for a substantial part of genetic variation, have relatively high mutation rates, and have been associated with susceptibility to several psychiatric disorders. We conducted a genome-wide study for post-twinning *de novo* CNVs (i.e., not shared by co-twins). CNVs from 1,097 MZ pairs were measured in DNA from peripheral blood or buccal epithelium with the Affymetrix 6.0 platform.

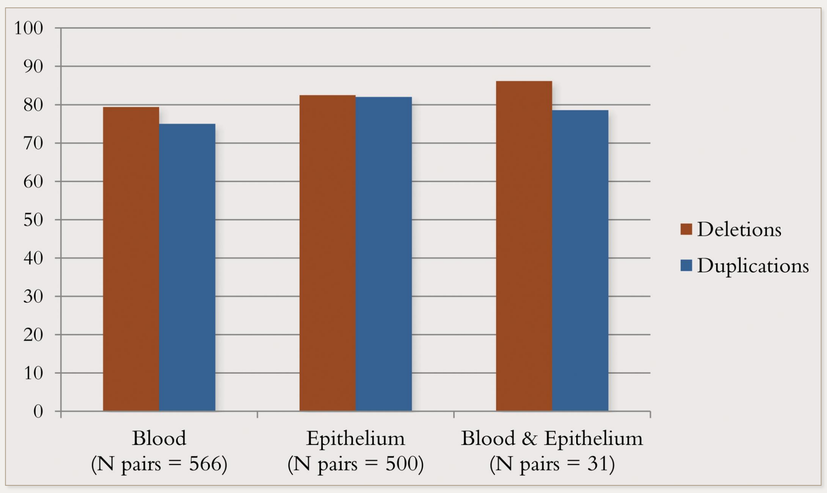
Post-twinning *de novo* CNVs

CNV segments were included in analyses if: 1) the CN calls agreed between 2 algorithms (Birdsuite & PennCNV), 2) the overlapping part of the segments from both algorithms was >100 kb, and 3) the segment was not in a centromere.

We found 153 putative post-twinning *de novo* CNVs >100 kb, of which the majority resided in the same unstable genomic region (15q11.2). Based on how well the raw intensity signals visually agreed with CNV calls, a selection was made of 19 *de novo* CNVs from 15q11.2 for qPCR validation experiments. Two post-twinning *de novo* CNVs were validated with qPCR in the same twin pair (Figure below). This 13-year old twin pair did not show large phenotypic differences.







CNVs concordant within MZ pairs

The Figure above shows the percentage of CNVs that were concordant between MZ pairs. The percentages were similar (~80%) for three groups: DNA from blood for both twins, epithelium for both twins, and one twin from blood and one from epithelium. There was however a significant difference in the nr of concordant CNVs. Post-hoc tests showed that this was due to blood-derived samples showing significantly more CNVs per twin ($\mu = 2.9$) than epithelium derived samples ($\mu = 2.5$; $p = .001$).

CNVs vs. psychiatric symptoms

The concordant CNVs were tested for association with Attention Problems (AP; ADHD symptoms) and Thought Problems (TP; schizo-obsessive symptoms) using the gene-enrichment test in Plink. 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).

There was a significant association between AP and the gene set involved in neuronal apoptosis ($p = 4\times10^{-39}$). This association disappeared after permutation tests. Permutations (10k) were performed within 4 clusters based on gender and source of DNA.

Conclusions

Two post-twinning *de novo* CNVs were detected in a healthy twin pair in 15q11.2, an unstable genomic region that has been associated with several disorders. MZ twins also provide the opportunity for an extra QC step for the relatively noisy micro-array CNV data. About 80% of CNV calls were concordant between MZ pairs in both blood-derived and epithelium-derived DNA, but we were able to detect significantly more CNVs in blood. CNV calling in more MZ twin pairs and relatives of twins (N total = ~14,000 Ss) will likely increase power for association studies and improve the overall quality of CNV calls.

Contact information: a.abdellaoui@vu.nl