CNV Analysis in Monozygotic Twin Pairs Discordant for Urorectal Malformations

Friederike Baudisch,1* Markus Draaken,2,3* Enrika Bartels,2 Eberhard Schmiedeke,4 Soyhan Bagci,5 Peter Bartmann,5 Markus M. Nöthen,2,3 Michael Ludwig,1 and Heiko Reutter2,5
1Department of Clinical Chemistry and Clinical Pharmacology, University of Bonn, Bonn, Germany
2Institute of Human Genetics, University of Bonn, Bonn, Germany
3Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany
4Department of Pediatric Surgery and Urology, Center for Child and Adolescent Health, Hospital Bremen-Mitte, Bremen, Germany
5Department of Neonatology, Children’s Hospital, University of Bonn, Bonn, Germany

Early post-twinning mutational events can account for discordant phenotypes in monozygotic (MZ) twin pairs. Such mutational events may comprise genomic alterations of different sizes, ranging from single nucleotides to large copy-number variations (CNVs). Anorectal malformations (ARM) and the bladder extrophy-epispadias complex (BEEC) represent the most severe end of the urorectal malformation spectrum. Recently, CNV studies in patients with sporadic ARM and the BEEC have identified de novo events that occur in specific chromosomal regions. We hypothesized that early arising, post-twinning CNVs might contribute to discordance in MZ twin pairs with ARM or the BEEC; knowledge of such CNVs might help to identify additional chromosomal regions involved in the development of these malformations. We investigated four discordant MZ twin pairs (three ARM and one BEEC) using molecular karyotyping arrays comprising 1,140,419 markers with a median marker spacing of 1.5 kb. Filtering the coding regions for possible disease-causing post-twinning de novo CNVs present only in the affected twin, but not in the unaffected twin or the parents, identified a total of 136 CNVs. These 136 CNVs were then filtered against publicly available databases and finally re-evaluated visually. No potentially causative CNV remained after applying these filter criteria. Our results suggest that post-twinning CNV events that affect coding regions of the genome did not contribute to the discordant phenotypes in MZ twin pairs that we investigated. Possible causes for the discordant phenotypes include changes in regulatory elements or smaller genetic changes within coding regions which may be detectable by whole-exome sequencing.

Keywords: twin pairs, monozygotic discordant, urorectal malformations, CNV analysis
TABLE 1
Phenotypes of the Affected Twins Investigated

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Head and neck</td>
<td>Head</td>
<td>Plagiocephalus</td>
<td>Hypoplastic right-sided mandible</td>
<td>Hypoplastic right-sided zygomatic bone</td>
<td>Auricular tags, bilateral</td>
</tr>
<tr>
<td></td>
<td>Face</td>
<td></td>
<td></td>
<td></td>
<td>Epidermoid cyst in left iris</td>
</tr>
<tr>
<td></td>
<td>Ears</td>
<td></td>
<td></td>
<td></td>
<td>Epibulbar dermoid left</td>
</tr>
<tr>
<td></td>
<td>Eyes</td>
<td></td>
<td></td>
<td></td>
<td>Asymmetric labial angles</td>
</tr>
<tr>
<td>Chest</td>
<td>Chest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen</td>
<td>Abdominal features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitourinary</td>
<td>Bladder</td>
<td></td>
<td></td>
<td></td>
<td>Rectourethral fistula</td>
</tr>
<tr>
<td></td>
<td>Bladder</td>
<td></td>
<td></td>
<td></td>
<td>Distended bladder</td>
</tr>
<tr>
<td></td>
<td>Classical exstrophy of the bladder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skeletal</td>
<td>Spine</td>
<td></td>
<td></td>
<td></td>
<td>Inherent cervical scoliosis</td>
</tr>
<tr>
<td>Neurologic</td>
<td>Behavioral psychiatric manifestation</td>
<td></td>
<td></td>
<td></td>
<td>Dysplasia of the os sacrum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aplasia of os coccygis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Vertebral fusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Attention-deficit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hyperactivity disorder</td>
</tr>
</tbody>
</table>

extrophy-epispadias complex (BEEC). There is substantial evidence that genetic factors contribute to the development of these malformations; in particular, CNV studies in sporadic patients have identified de novo events that occur in specific chromosomal regions (Bartels et al., 2011; Boocock & Donnai, 1987; Boyadjiev et al., 2004; Cuschieri & EUROCAT Working Group, 2001; Draeken et al., 2010; Falcone et al., 2007; Keppler-Noreuil, 2001; Ludvig et al., 2009; Marcelis et al., 2011; Schramm et al., 2011a, 2011b; Shapiro et al., 1984). We hypothesized that CNVs affecting coding regions that occur after twinning might contribute to discordance in MZ twin pairs with ARM or the BEEC. To test this hypothesis, we investigated four discordant MZ twin pairs (three ARM and one BEEC) using array-based molecular karyotyping.

Materials and Methods
Patients and DNA Isolation
In this study we included four discordant MZ twin pairs who were contacted and recruited through the German self-help organizations for patients with ARM (SoMA e.V.) and BEEC (Selbsthilfegruppe Blasenekstrophie/Epispadi e.V.) as well as through the Department of Neonatology, Children’s Hospital, University of Bonn, Bonn, Germany and the Department of Pediatric Surgery and Urology, Center for Child and Adolescent Health, Hospital Bremen-Mitte, Bremen, Germany. Patients with ARM were classified according to the Krikenbeck Classification System (Holschneider et al., 2005). Patients with the BEEC were classified according to Gearhart (2002). Written informed consent was obtained from all families prior to study entry. The Ethics Committee of the Medical Faculty of the University of Bonn approved the study. For each twin pair, blood or saliva samples were obtained from both twins and their parents for the purpose of DNA extraction and molecular-genetic analyses. Isolation of genomic DNA was carried out using the Chemagic Magnetic Separation Module I (Chemagen, Baesweiler, Germany) or, in the case of saliva samples, the Oragene DNA Kit (DNA Genotek Inc., Kanata, Canada).

Description of Twin Pairs
Twin pair 1. Twin pair 1 was 5-year old at the time of recruitment. These male twins were the second and third of four children of unrelated German parents. The affected twins presented with classic bladder exstrophy (CBE; Table 1). The mother had one intrauterine fetal dismissal due to fetal polycystic kidney disease. The family history was otherwise unremarkable.

Twin pair 2. Twin pair 2 was recruited during the newborn period. These male twins were the first children of a Peruvian mother and a German father. The affected twin presented with ARM, dilated lower and upper urinary tract, and distension of the abdominal wall, resembling features of prune belly syndrome (PBS) (Table 1). The family history was unremarkable.

Twin pair 3. Twin pair 3 was 9-year old at the time of recruitment. These male twins were the only children of their unrelated German parents. The affected twin presented with hemifacial microsomia and pre-auricular pits, resembling...
features of Goldenhar syndrome and ARM (Table 1). The family history was unremarkable.

**Twin pair 4.** Twin pair 4 was 13-year old at the time of recruitment. These male twins were the first two of four children of their unrelated German parents. The affected twin presented with ARM and fused vertebra of the thoracic spine. Besides his congenital anomalies, he developed an attention-deficit hyperactivity disorder during childhood (Table 1). The third pregnancy of the mother ended in fetal dismiss during the 13th week of gestation, with signs of severe intrauterine growth retardation.

History of pregnancy in all four twin pairs was uneventful for the ARM- and BEEC-associated environmental risk factors, for example, maternal diabetes, uterine vascular pathology, and infertility treatment (Reutter et al., 2011; Zwink et al., 2011).

**Monozygosity Testing**

We tested all twin pairs for their zygosity status before proceeding with molecular karyotyping. Analysis was performed using a PowerPlex® 16 System according to the manufacturer’s recommendations (Promega, Madison, WI, USA), allowing co-amplification and three-color detection of 16 loci.

**Paternity Testing**

The GenomeStudio (version V2011.1, http://www.illumina.com/) genotyping module was used for paternity testing.

**Array-Based Molecular Karyotyping**

HumanOmn1-Quad v1 Chip (Illumina, Inc., San Diego, CA, USA), which contains 1,140,419 markers with a median marker spacing of 1.5 kb, was used for molecular karyotyping. All analyses were performed according to the manufacturer’s protocol. A DNA sample was considered to have failed if fewer than 95% of loci were generated on the corresponding BeadChip. CNVs were predicted using the program QuantiSNP (v2.2, www.well.ox.ac.uk/QuantiSNP/), which uses an Objective Bayes Hidden-Markov model for the estimation. The log R ratio (LRR) represents a measure of the magnitude of combined fluorescence-intensity signals. The B-allele frequency (BAF) denotes the relative ratio of fluorescence signals from one allelic probe compared with another. Duplications can be identified by an increase of LRR and the occurrence of four clusters in BAF (at 0, ~0.33, ~0.67, and 1). Consequently, a deletion is characterized by a decrease of LRR and lack of heterozygosity (at 0.5) in BAF.

**CNV Filtering**

Filtering of CNV data was carried out using Cartagenia Bench™ software (released January 2011, Cartagenia, Leuven, Belgium). We excluded (1) CNV calls with a log Bayes factor below 7, (2) all CNV regions with fewer than five aberrant markers, and (3) all CNV calls with a frequency of over 5% in our internal control cohort (n = 531 healthy controls; Table 2). As we included small CNVs comprising ‘5 aberrant’ markers only, a certain number of remaining CNVs comprised less than three rs-markers; for example, 2 rs- and 3 cnvi-markers = ‘intensity only probes’. As cnvi-markers do not carry any genotype information, we excluded CNVs comprising less than three rs-markers later through visual inspection in filter step (7). In filter step (4) the remaining CNVs were filtered for being present only in the affected twin, but not in the unaffected twin or the parents, and in filter step (5) their gene content was filtered using the UCSC genome browser (Dreszer et al., 2012; http://genome.ucsc.edu) against RefSeqGene (coding regions only, http://www.ncbi.nlm.nih.gov/refseq/rs g/). Next, in step (6) we filtered for the presence of CNVs against the following publicly available databases: Database of Genomic Variants (DGV, v10; Iafrate et al., 2004), and Database of Chromosomal Imbalance and Phenotype in Humans, using Ensembl Resources (DECIPHER, v5,1; Firth et al., 2009). CNVs determined to exist in either of these databases, with at least 10 full overlapping reports in DGV and at least five full overlapping reports in DECIPHER, were excluded. However, we did not exclude CNV overlapping in at least five reports in DECIPHER if the associated phenotypes included any phenotype investigated here, for example, ARM. In step (7) the remaining CNVs were visually re-evaluated using the GenomeStudio (GS) genotyping module (GS, version V2011.1, http://www.illumina.com/). GS Data Analysis
Software visualizes and analyzes data generated by all of Illumina's platforms. GS supports the primary analysis of the microarray-based data generated by the iScan System and BeadXpress Reader. The results of genotyping are displayed in the form of LRR and BAF as described above (see section ‘Array-Based Molecular Karyotyping’). As mentioned above, through visual inspection of the remaining CNVs we excluded all CNVs comprising less than three rs-markers, for example, two rs- and three cnvi-markers.

In addition, we used the available data to search for possible disease-causing de novo CNVs with incomplete penetrance. For this analysis, we used identical filter criteria, except that criterion (4) was modified to filter CNVs present in both twins but not in parents.

Quantitative Polymerase Chain Reaction (qPCR)
Reactions for qPCR were performed on an ABI Prism 7900HT Fast Real-Time PCR System (Applied Biosystems, Carslbad, CA, USA) using SYBR Green for detection. Each assay included DNA from four controls (two male and two female samples) and the probands’ DNA at a final concentration of 20 ng/mL; reactions were run in triplicate. Reaction mixtures (10 μL) contained 0.2 μmol of each primer and 5 μL of Power SYBR Green PCR Master Mix (Applied Biosystems), with cycling conditions as follows: Initiation, 50 °C for 2 min; denaturation, 95 °C for 10 min; followed by 40 cycles at 95 °C for 15 s; and a combined annealing and extension step at 60 °C for 60 s. The threshold cycle (Ct) values were normalized using the Ct-value of three reference genes (BNC1, CTR, and RPP38) for each DNA sample. Relative quantification was done using the comparative Ct method (Livak & Schmittgen, 2001).

Results
Screening for CNVs present only in the affected twin, but not in the unaffected twin or the parents (complete penetrance). The numbers of CNVs remaining after various filtering steps are given in Table 2. After applying steps (1), (2), and (3), 125, 142, 174, and 126 CNVs remained in twin pairs 1, 2, 3, and 4 respectively. After applying step (4), 44, 70, 62, and 40 CNVs remained. After applying filtering steps (5), (6), and (7) no potentially causative CNV remained.

Screening for CNVs present in both twins but not the parents (incomplete penetrance). Filtering for possible disease-causing de novo CNVs present in both twins but not in the parents yielded, after step (4), nine CNVs for twin pair 1, eight CNVs for twin pair 2, 31 CNVs for twin pair 3, and 14 CNVs for twin pair 4. After applying filter steps (5), (6), and (7), one CNV remained in twin pair 3 (Table 2, column B). This CNV on chromosome 16p13.3 was suggested to have a size of around 14 kb and overlapped with exons 7–9 of human Rab11 family-interacting protein 3 (RAB11FIP3) gene. However, qPCR performed on the twins and their parents revealed two copies in all individuals and thus did not confirm the presence of a CNV.

Discussion
Our array-based analysis did not identify disease-causing CNVs in any of the four MZ discordant twin pairs that we investigated. This suggests that CNVs affecting coding regions that arise as early post-twinning mutational events are not a frequent cause of discordance among MZ twins with urorectal malformations. It may be the case, however, that we have missed true causative CNVs because our filter criteria excluded CNVs located outside of coding regions and would therefore have missed CNVs in promoter or enhancer regions. It is also possible that the arrays we used, with a median inter-marker spacing of 1.5 kb, missed smaller CNVs, or that mutational events were present involving small DNA changes (e.g., single-nucleotide substitutions) that are not detectable per se using an SNP-based array approach. These changes might well be detectable by systematic sequencing approaches such as exome- or genome-wide sequencing, and these methods therefore represent an important direction for future studies of discordant twin pairs. Furthermore, there could be somatic mosaicism for causal mutation (if any . . .), which remained undetected because we investigated DNA from peripheral lymphocytes. The investigation of this would require investigation of DNA from affected tissue.

In addition to genomic differences, there is also growing evidence that implicates epigenetic events occurring early in embryogenesis in the development of phenotypic discordance in MZ twin pairs (Kaminsky et al., 2009; Yamazawa et al., 2008), and such alterations can be detected in principle with appropriate experimental methods. For ethical reasons, however, investigations of tissue-specific differences in methylation in urorectal malformations are not possible because the unaffected twins are not available for tissue sampling. Retrospective evaluation of pregnancy history in the investigated discordant twin pairs did not identify any possible environmental risk factor. Nevertheless, an environmental contribution cannot be ruled out completely (Reutter et al., 2011; Zwink et al., 2011).

Acknowledgments
We thank all patients and their parents for their participation in this study. We thank the German self-help organizations for people with anorectal malformations (SoMA e.V.) and with exstrophy-epispidias complex (SHG Blasenekstrophe/Epispidiae e.V.). Grant sponsor: German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF; grant no. 01GM08107). Enrika Bartels, is supported by the University of Bonn, BONFOR (grant no. O-149.0099). Enrika Bartels, Markus M. Nöthen, Michael Ludwig, Heiko Reutter, and Markus Draaken are members of the Network for the Systematic
Investigation of the Molecular Causes, Clinical Implications and Psychosocial Outcome of Congenital Urorectal Malformations (CURE-Net), which is supported by a research grant from the BMBF.

References


...
