Published in final edited form as:

Am J Med Genet A. 2014 February ; 0(2): 397–406. doi:10.1002/ajmg.a.36291.

Identification of Novel Candidate Gene Loci and Increased Sex Chromosome Aneuploidy among Infants with Conotruncal Heart Defects

Kazutoyo Osoegawa^{1,*}, David M. Iovannisci¹, Bin Lin¹, Christina Parodi¹, Kathleen Schultz¹, Gary M. Shaw², and Edward J. Lammer¹

¹Center for Genetics, Children's Hospital Oakland Research Institute, Children's Hospital Research Center Oakland, Oakland CA 94609 USA

²Department of Pediatrics, Stanford University School of Medicine, Stanford CA 94305 USA

Abstract

Congenital heart defects are common malformations, affecting 4–8 per 1,000 total births. Conotruncal defects are an important pathogenetic subset of congenital heart defects, comprising nearly 20 percent of the total. Although both environmental and genetic factors are known to contribute to the occurrence of conotruncal defects, the causes remain unknown for most. To identify novel candidate genes/loci, we used array comparative genomic hybridization to detect chromosomal microdeletions/duplications. From a population base of 974,579 total births born during 1999–2004, we screened 389 California infants born with tetralogy of Fallot or d-transposition of the great arteries. We found that 1.7% (5/288) of males with a conotruncal defect had sex chromosome aneuploidy, a seven-fold increased frequency (relative risk = 7.0; 95% confidence interval 2.9–16.9). We identified eight chromosomal microdeletions/duplications for conotruncal defects. From these duplications and deletions, we found five high priority candidate genes (*GATA4*, *CRKL*, *BMPRIA*, *SNAI2* and *ZFXH4*). This is the initial report that sex chromosome aneuploidy is associated with conotruncal defects among boys. These chromosomal microduplications/deletions provide evidence that *GATA4*, *SNAI2* and *CRKL* are highly dosage sensitive genes involved in outflow tract development. Genome wide screening for copy number variation can be productive for identifying novel genes/loci contributing to nonsyndromic common malformations.

Keywords

Conotruncal heart defects; array comparative genomic hybridization (array-CGH); sex chromosome aneuploidy; microdeletions/duplications; candidate genes

INTRODUCTION

Congenital heart defects are the most common major birth defect affecting 4–8 per 1,000 total births, and are the foremost cause of birth defects-associated infant mortality [O'Malley et al., 1996; Reller et al., 2008]. Both environmental and genetic factors appear to contribute to etiologies of congenital heart defects [Jenkins et al., 2007]. Complex molecular and cell

Correspondence: Kazutoyo Osoegawa, 5700 Martin Luther King Jr. Way Oakland, CA 94609 USA, Tel: 510-450-7911; Fax: 510-450-7910; kosoegawa@chori.org.

Disclosures

The authors declare no competing financial interests.

biological pathways are involved in heart development. Interruption of these pathways during heart development likely instigates atypical heart development, but little is known about the underlying mechanisms. Considering the wide anatomical spectrum of congenital heart defect phenotypes, it is of great importance to group pathogenetically related malformations when investigating etiologies of congenital heart defects. Conotruncal defects represent one pathogenetic group [Clark 1996] and they comprise approximately 20 percent of congenital heart defect cases, or 7.3 per 10,000 total births [Ferencz et al., 1985; O'Malley et al., 1996]. Conotruncal defects are well-ascertained by birth defects registries because their severity brings the infants to diagnosis shortly after birth. Conotruncal defects include tetralogy of Fallot, d-transposition of the great arteries, truncus arteriosus, double outlet right ventricle, and subaortic conoventricular septal defects. We chose to investigate subjects who were born with either tetralogy of Fallot or d-transposition of the great arteries, because they comprise >70% of conotruncal defects.

The human genome project has provided rapid, cost effective, high-resolution and high-throughput tools for studying human genetics. Array comparative genomic hybridization (array-CGH) has been a powerful tool for identifying disease-associated microdeletions/duplications [Vissers et al., 2004]. We demonstrated the utility of array-CGH for identifying candidate gene loci for cleft lip and palate and showed that screening cleft subjects using array-CGH complemented results uncovered by other genetic studies, such as linkage analyses or genome wide association studies [Osoegawa et al., 2008].

We identified 389 infants from a population-based registry who were born with conotruncal heart defects between 1999 and 2004. We previously reported on the clinical cytogenetic testing results of this study population [Lammer et al., 2009]. In this report we employed array-CGH as a useful strategy to discover candidate gene loci likely to be specifically associated with conotruncal heart defects.

MATERIALS AND METHODS

Study population

Research protocols for this study were approved by the California Committee for the Protection of Human Subjects as well as the Institutional Review Boards at Children's Hospital and Research Center Oakland and Stanford University. We used data derived from a population-based case-control study, California Study of Birth Defects Causes, described in detail elsewhere [Lammer et al., 2009; Shaw et al., 2010]. The eligible study population was defined by pregnant women who resided in Los Angeles, San Francisco, or Santa Clara counties, and who delivered between July 1999 and June 2004. Using active surveillance carried out by staff of California Birth Defects Monitoring Program (CBDMP), we identified every infant (diagnosed within 1 year after birth) and fetus born with tetralogy of Fallot (TOF) or with d-transposition of the great arteries (d-TGA). Anatomic and physiologic features of each case were confirmed by review of all echocardiographic, cardiac catheterization, surgical, or autopsy reports. Subjects with double outlet right ventricle or with atrioventricular septal defects that accompanied TOF or d-TGA were excluded. TOF cases included those with pulmonary valve stenosis or atresia. Reports of chromosomal karyotypes and FISH analyses were obtained by reviewing prenatal diagnostic reports, in-patient medical records, and logs of cytogenetic labs. We matched study subjects with their residual newborn screening bloodspots, which was the source of genomic DNA for our array-CGH investigations.

DNA isolation

Genomic DNA (~10 ng) was isolated from newborn screening blood spots, amplified using a whole genome amplification (WGA) procedure (GenomePlex, Sigma), and purified using Sephadex G-50 spin column. DNA concentration was determined using a NanoDrop spectrophotometer (Thermo Scientific). Genomic DNA isolated from one healthy male was used as the reference genome, and was amplified using the same WGA procedure as the test DNA samples.

Array comparative genomic hybridization (Array-CGH)

We used custom microarrays containing >32,000 bacterial artificial chromosome (BAC) clones (32k BAC clone set) for array-CGH [Jonsson et al., 2007; Krzywinski et al., 2004]. We previously showed reproducibility identifying microdeletions/duplications using WGA-DNA and genomic DNA from the same subjects [Osoegawa et al., 2008]. Test and reference WGA-DNA was labeled with Cy3 and Cy5 using the random priming procedure (Invitrogen: BioPrime aCGH Genomic Labeling System). The BAC array-CGH was performed as previously described [Osoegawa et al., 2008]. We used DNA from five subjects who had common 3-Mb chromosome 22q11.21 microdeletions to monitor the quality control of the array-CGH investigations.

For microdeletions/duplications less than 500-kb, we confirmed the deletions/duplications using dye swap experiments to exclude false positive results. For confirmation of each microdeletion/duplication, we used either 135k oligonucleotide microarrays (Roche Nimblegen), or 1 million oligonucleotide microarrays (Agilent Technologies), according to the manufacturers' recommended protocols. The results from these oligonucleotide microarrays also helped to narrow their chromosomal "break points".

Gene prioritization

To prioritize candidate genes from the deleted or duplicated loci, we first used the web-based Endeavour gene prioritization program available from Congenital Heart Defects (CHD) Wiki web site (URL: http://homes.esat.kuleuven.be/~bioiuser/chdwiki/index.php/Main_Page) [Barriot et al., 2010]. The Endeavour gene prioritization program prioritizes candidate genes based on their similarity to known genes (training gene set) involved in specific disease phenotypes [Aerts et al., 2006]. Conotruncal heart defects are considered to arise from faulty outflow tract development, likely involving cells originating from the second heart field and/or from cranial neural crest cells [Kirby 2002]. We, therefore, chose "second heart field" and "neural crest cells" for selection of reference genes for gene prioritization. We used default settings for gene prioritization. We performed the second gene prioritization analyses from the Endeavour web site (URL: <http://homes.esat.kuleuven.be/~bioiuser/endeavour/tool/endeavourweb.php>) using another training gene set: 36 nonsyndromic genes that have been associated with a variety of congenital heart defects and listed in CHD Wiki. To complement the results from the Endeavour analyses, we used the ToppGene gene prioritization program using the same set of 36 nonsyndromic congenital heart defect genes (URL: <http://toppgene.cchmc.org/prioritization.jsp>) [Chen et al., 2009].

DNA sequencing

We amplified exons from the *TBX1*, *FGF8* and *CRKL* genes, and sequenced PCR products using the Sanger DNA sequencing approach. The target specific PCR primers used are listed in Supplemental TABLE I and PCR conditions are described in Supplemental Methods (see supporting information online).

RESULTS

Study population

Our study population was a base of 968,885 live-born infants and 5,694 stillborn plus electively terminated fetuses, totaling 974,579 births (TABLE I). Of these, we ascertained 290 infants/fetuses who were born with tetralogy of Fallot (TOF) and 194 with d-transposition of the great arteries (d-TGA) (TABLE II). The frequencies of TOF and d-TGA are calculated to be 3.0 (290/974,579) and 2.0 (194/974,579) per 10,000 total births, respectively. Of the 290 infants/fetuses with TOF from among our study subjects, 10 were identified with chromosomal aneuploidies by clinical karyotyping and excluded. Of the remaining 280, 87 (31%) were clinically tested for the common ~3Mb 22q11 microdeletion using fluorescence *in situ* hybridization (FISH), of which 28 showed the microdeletion. After excluding these 28, a TOF study population of 252 remained. Of the 194 subjects with d-TGA, 40 (21%) were tested for the 22q11 microdeletion using FISH, but no microdeletion was identified. We successfully matched and retrieved newborn bloodspots for 225 (89%) subjects with TOF and 164 (85%) with d-TGA (TABLE II).

Sex chromosomal aneuploidy

We found two sex chromosome aneuploidies (47,XXY and 47,XXY) by array-CGH in two males who were born with d-TGA (Supplemental Figure 1). These boys were not analyzed by standard karyotyping during their first year. Using array-CGH, we were not able to exclude the possibility that one or both subjects has mosaicism of 46,XY/47,XXY or 46,XY/47,XXY. From karyotypes performed before enrollment in our investigation, three males with TOF were found to have sex chromosome aneuploidy: one with 47,XXY and two with mosaic 47,XXY/46,XY karyotypes [Lammer et al., 2009]. Thus, we found five (1.7%) of 288 eligible male infants with conotruncal defects had sex chromosome aneuploidy (TABLE II). To interpret the significance of this observation, the only appropriate comparison group is an unselected newborn population who were karyotyped independent of birth outcomes. This is necessary because the clinical detection of male sex chromosome aneuploidy is low and as few as 10% of affected males are diagnosed. Morris et al., [2008] summarized data from 16 newborn cohort studies, totaling 71,405 karyotyped newborn males, and 177 male infants with 47,XXY including mosaic 46,XY/47,XXY or with 47,XXY including mosaic 46,XY/47,XXY. Combined, that yields a frequency of 2.5/1000 for any male sex chromosome aneuploidy. The frequency we observed (5/288) is seven-fold greater (relative risk = 7.0; 95% confidence interval 2.9–16.9). We ascertained 153 male and 137 female infants who were born with TOF and 135 male and 59 female infants born with d-TGA (TABLE II). The increased frequency of male sex chromosome aneuploidy among infants born with TOF or d-TGA may explain why, in part, many birth defects registries have observed that d-TGA and TOF are more common among males.

Chromosomal microduplications

TABLE III summarizes microdeletions/microduplications found by array-CGH. We found three microduplications involving chromosome 8, two on the p arm and one on the q arm. The first duplication involved a 3.9-Mb segment at chromosome 8p23.1 from a male subject who was born with TOF (Subject C in TABLE III, Figure 1A). Interestingly, the duplicated region contains 26 genes, including duplication of *GATA4*, a transcription factor known to play an important role in cardiac development [Garg et al., 2003]. The second 2.2-Mb microduplication was identified at chromosome 8p23.2 from a subject who was born with d-TGA (Subject D in TABLE III, Figure 1B). This duplicated region has few genes and only includes the 5' part of *CUB and Sushi multiple domains 1 (CSMD1)* which is known to be a tumor suppressor gene [Ma et al., 2009]. Figure 1 shows the alignment of these two 8p23 duplications with their relationship to the nearby defensin gene cluster, shown in Figure 1C.

This cluster is known to be highly copy number variable, with 2 to 7 copies of the defensin components found in the general population [Hollox et al., 2003]. The duplicated segments of the two identified 8p23 microduplications, however, do not overlap with the defensin gene cluster region, as shown in Figure 1A–C. The duplicated segments are localized distal and proximal, respectively, to the defensin gene cluster. Figure 1D shows the third microduplication that we discovered on the long arm at chromosome 8q21.11, from Subject E who was born with TOF. Only one gene is located within the duplicated segment, *zinc finger homeobox 4 (ZFHX4)*, which may be involved in neural and muscle differentiation [Hemmi et al., 2006].

Chromosomal microdeletions

We identified five microdeletions by array-CGH. Two were at chromosome 22q11: (1) a 600-kb deletion at 22q11.21 of a male who was born with TOF (Subject F in TABLE III, Figure 2B), and (2) a more telomeric 800-kb deletion at 22q11.22 from a male who was born with d-TGA (Subject G, Figure 2C). There are eight low copy repeats (LCRs) at chromosome 22q11 and Figure 2 shows the six most centromeric of the eight LCRs, denoted A–F [Shaikh et al., 2007]. The common 3-Mb 22q11 microdeletion occurs via non-allelic homologous recombination between LCR-A and LCR-D [Shaikh et al., 2000]. The microdeletion of subject F represents loss of the most telomeric 600-kb of the common 3-Mb 22q11 deletion (Figure 2). The second 22q11 microdeletion (subject G) represents loss of the 800-kb segment that is ~1.3-Mb distal to the telomeric end of the common 3-Mb 22q11 deletion at LCR-D. We used previously published PCR primer sets to investigate whether this 800-kb deletion might have occurred by non-allelic homologous recombination via the low copy repeats (LCRs) as previously described [Shaikh et al., 2007]. We designed a long range PCR amplification using forward primer (BEfor2) residing within LCR-E and a reverse primer (BFrev), located telomeric of LCR-F (see Figure 2). The distance between LCR-E and LCR-F is approximately 800-kb, it is, therefore, not possible to amplify a fragment using these primers. It is only possible to obtain a PCR product if these PCR primers were brought into proximity to each other by deletion of the intervening 800kb chromosomal segment between LCR-E and LCR-F. Figure 3 (lane 2) shows that we successfully amplified a 7-kb PCR amplicon from subject G, whereas there was no amplification from our reference DNA that lacks the deletion (see lane 3). This evidence suggests that subject G had an 800-kb deletion that most likely occurred by non-allelic homologous recombination between LCR-E and LCR-F, as previously described [Shaikh et al., 2007].

Three additional microdeletions were found among subjects with TOF: a terminal 2-Mb deletion of chromosome 1q44 from a female (Subject H, Supplemental Figure 2A), a 7.3-Mb deletion at chromosome 10q23.1 from a male subject (Subject I, Supplemental Figure 2B) and 1.0-Mb deletion at chromosome 8q11.21 from a female subject (Subject J, Supplemental Figure 2C).

Candidate gene prioritization

We prioritized candidate genes found within the deleted or duplicated loci using the Endeavour gene prioritization program. A total of 14 and 45 genes were selected as training genes for second heart field and cranial neural crest cells, respectively. For each deletion/duplication locus, TABLE III lists the highest priority candidate gene. Endeavour prioritized *GATA binding protein 4 (GATA4)*, *V-Crk Sarcoma Virus CT10 Oncogene Homolog (Avian)-Like (CRKL)*, *bone morphogenetic protein receptor type IA precursor (BMPRIA)* and *snail homolog 2 (SNAI2 or snail-related zinc-finger transcription factor, SLUG)* as the highest candidate genes. *CUB and Sushi multiple domains 1 (CSMD1)* and *zinc finger homeobox 4 (ZFHX4)* are the only genes located within 2.2-Mb at chromosome 8p23.2 and 0.8-Mb

chromosome 8q21.11, respectively. *CSMD1* may not be considered a promising candidate gene, because the duplicated region only includes the 5' part of *CSMD1*, and so this duplication is unlikely to lead to increased levels of *CSMD1* mRNA or protein. The 800-kb deletion at chromosome 22q11.22 overlaps with a 500-kb known copy number variable (CNV) region. Endeavour analysis pointed to *suppressor of hairy wing homolog 2* [*SUHW2*, or *zinc finger protein 280B* (*ZNF280B*)] within the 800-kb region at chromosome 22q11.22, but this gene resides within the 500-kb CNV region. Because copy number variation of *SUHW2* can be found in the general population, *SUHW2* may also not be a promising candidate gene. There are 57 genes within the 2.0-Mb deletion at chromosome 1q44. The top ranked candidate genes using Endeavour were different for training sets from second heart field (*AHCTF1*) vs. neural crest cells (*ZNF496*).

The 14 and 45 training genes for second heart field and neural crest cells include genes that have not been investigated in humans. These genes have been listed as potential candidate genes only because they are expressed in second heart field and neural crest cells based on model organism studies. We carried out gene prioritization using 36 genes listed under the category “genes having non-syndromic congenital heart defects” as a training gene set [Breckpot et al., 2012a]. The Endeavour analyses using these three training sets essentially lead to the same sets of candidate genes. To assess the reproducibility of the Endeavour results, we performed additional gene prioritization tests with the ToppGene gene prioritization program using the same training set of 36 genes [Chen et al., 2009]. *CRKL*, *BMPRIA* and *SNAI2* genes were selected as the highest ranking candidate genes within the deleted regions, while *SUHW2* was selected as the second highest priority gene within the 800-kb deletion at chromosome 22q11.22. In summary *GATA4*, *CRKL*, *BMPRIA*, *SNAI2* and *ZFHX4* genes were ranked as the highest likely candidate genes using two independent gene prioritization programs.

DNA sequencing of infant (Subject F) with 600-kb deletion

We found a 600-kb deletion at chromosome 22q11.21 that represents the most telomeric portion of the common 3-Mb deletion from subject F. The deleted segment does not contain *T-box1* (*TBX1*) which resides toward the centromeric end of the 3-Mb 22q11 region and encodes a critical transcriptional regulator of second heart field development. Many investigators have considered that *TBX1* is the gene within the 3-Mb region that causes conotruncal defects or DiGeorge anomaly [Xu et al., 2004]. Within the distal 600-kb deletion, Endeavor prioritized *CRKL* as a candidate gene. Because *TBX1* is also known to interact with *FGF8* during heart development, it was essential to rule out the possibility that TOF was caused by accompanying nonsynonymous mutations of *TBX1*, *FGF8*, or the single copy of *CRKL* [Lindsay et al., 2001]. We sequenced PCR products containing exons from *TBX1*, *FGF8* and *CRKL* using the Sanger method and found no nonsynonymous mutations. This strongly suggests that haploinsufficiency of *CRKL* can cause conotruncal defects independent of the contributions of *TBX1*.

DISCUSSION

We postulated that mutations in genes pertinent to the development of the outflow tract may contribute to the formation of conotruncal heart defects. Among mutations, chromosomal duplications/deletions often provide clues of candidate genes and loci [Osoegawa et al., 2008; Vissers et al., 2004]. In this large population-based study of 389 California infants born with TOF or d-TGA, we found five of 288 male infants (1.7%) had sex chromosome aneuploidy, a seven-fold higher frequency compared to the expected [Morris et al., 2008]. A previous investigation of 230 German children with TOF noted a single subject with 47,XXX karyotype, but no male with extra sex chromosomes [Rauch et al., 2010]. We might

have missed 47,XXX karyotypes, because it is difficult to distinguish 46,XX from 47,XXX by array-CGH using male reference DNA. It is also possible that their results differ from ours, because their study population was hospital-based, included more syndromic cases, and did not include infants with d-TGA [Rauch et al., 2010]. In contrast, we used unbiased, population-based case recruitment that ascertained every infant with a conotruncal defect from a birth cohort of 974,579. Other investigations have reported TOF in an infant with 48,XXYY karyotype, and congenital heart defects, mostly patent ductus arteriosus, observed in ~14% boys with 49,XXXXY karyotype [Karsh 1975; Meschede et al., 1995; Tartaglia et al., 2011]. Because we observed two different types of sex chromosome aneuploidies (47,XXY and 47,XYY), we speculate that higher dosage of gene(s) in pseudoautosomal regions on X and Y chromosomes may increase the risk for conotruncal heart defects. Pseudoautosomal regions are the only regions that are “duplicated” in these two different sex chromosome aneuploidies.

Many studies show that d-TGA and TOF are more common among males than females, as we confirmed in our study population [Lary and Paulozzi 2001; Tennant et al., 2011]. Among boys, sex chromosome aneuploidy was the next most frequent abnormality (1.7%) after the 3Mb microdeletion at 22q11 (12%). Our observation may explain why, in part, d-TGA and TOF are more common among males. This is the first report linking higher birth prevalence among males with conotruncal defects with increased risk of sex chromosomal aneuploidy. Because our study is population-based, our observations cannot be attributed to referral bias.

We found 12% of infants/fetuses with TOF had a 3-Mb 22q11 microdeletion, while no infant with d-TGA did [Lammer et al., 2009]. Various studies show that there are eight low copy repeats (LCRs) at 22q11 and that nonhomologous allelic recombination between these repeats creates a variety of microdeletions and duplications (Figure 2) [Breckpot et al., 2012a; Rauch et al., 2005; Shaikh et al., 2000; Shaikh et al., 2007]. The common 3-Mb 22q11 microdeletion results from aberrant recombination between LCR-A and LCR-D [Shaikh et al., 2000]. Within this common 3-Mb deleted segment, *TBX1* has been considered the most likely haploinsufficient gene that induces the cardiac defects and other features of the DiGeorge anomaly [Merscher et al., 2001; Yagi et al., 2003]. Point mutations of *TBX1* are also associated with congenital heart defects [Xu et al., 2004]. Three infants with conotruncal defects have been reported with microdeletions of only the most telomeric 600-kb of the common 3-Mb 22q11.21 deletion (Figure 2) [Breckpot et al., 2012a; Kurahashi et al., 1997; van Engelen et al., 2010]. Our Subject F represents the fourth subject with TOF who has this same 600-kb deletion that most likely occurred via nonallelic homologous recombination between LCR-C and LCR-D [Shaikh et al., 2007]. Of the thirteen genes that lie within this 600-kb segment, only *CRKL* is expressed in the developing heart, and Endeavor ranked it as the most plausible candidate gene. A cranial neural crest cell conditional knockout of *Crkl* induced DiGeorge anomaly in mouse embryos [Moon et al., 2006]. *CRKL* is part of a signaling pathway that includes *Focal adhesion kinase (FAK)* and *MAP kinases [MAPK3 (ERK1) and MAPK1 (ERK2)]*, that transmits information from cell surface receptors to cytoplasmic and nuclear effectors [Vallejo-Illarramendi et al., 2009]. The pathway can be initiated by integrins, secreted fibroblast growth factors (particularly *FGF8*) bound to their receptor dimers, and transforming growth factor-beta. The embryonic phenotypes of murine neural crest conditional knockouts of *Fak*, *Crkl*, or *Erk2* are remarkably similar and show the features of human DiGeorge anomaly [Vallejo-Illarramendi et al., 2009]. These conditional knockouts permit normal cranial neural crest cell migration, but the post-migratory cells do not differentiate normally [Vallejo-Illarramendi et al., 2009]. These experimental observations indicate that this Fak-Crkl-Erk signaling pathway is critically important for integrating postmigratory cranial neural crest cells into the developing outflow tract, and that perturbation of multiple components of the

pathway leads to cardiac and other malformations that are characteristic of the human DiGeorge anomaly.

We found four other novel microdeletions and three microduplications, none of which represents a polymorphic copy number variant region. We identified an 800-kb microdeletion at chromosome 22q11.22 that likely resulted via nonallelic homologous recombination between LCR-E and LCR-F, deleting 12 genes. Subject G had isolated d-TGA, whereas the only published subject who had the same microdeletion had no congenital heart defect, but displayed developmental delay, speech apraxia, short stature, hypertelorism, and pectus excavatum [Shaikh et al., 2007].

We identified a 3.9-Mb duplication at chromosome 8p23.1 from Subject C with TOF, which creates a third copy of the transcription factor, *GATA4*, which plays a key role in heart development [Garg et al., 2003]. Nonsynonymous *GATA4* mutations have been identified from subjects with various congenital heart defects, including conotruncal defects [Nemer et al., 2006; Tomita-Mitchell et al., 2007]. Further, while microdeletions that delete *GATA4* have been associated with a variety of heart defects, including conotruncal [Barber et al., 2010; Pehlivan et al., 1999; Wat et al., 2009], ours represents only the second report of a *GATA4* duplication with a conotruncal defect, also TOF [Barber et al., 2008]. The subject with TOF reported by Barber et al., [2008] also had adrenal insufficiency, microcephaly and cerebral palsy, unlike our subject, who has nonsyndromic TOF. *GATA4* appears to be extremely dosage sensitive, because both duplication and deletion of the gene can cause pathogenetically similar heart phenotypes. Nonsynonymous *GATA4* mutations have been found in multiple families with atrial septal defects, rather than the more complex heart defects observed in subjects with 8p23.1 microdeletions [Wat et al., 2009]. This suggests that a second modifying gene may influence risk of cardiac defects. While we consider that *GATA4* gene is the most biologically plausible candidate gene, the 3.9-Mb duplication created three copies of *SRY-box 7 (SOX7)*, which the Endeavour program selected as the second highest ranking candidate gene. *Sox7* and its orthologous gene are expressed in mouse and human adult heart [Takash et al., 2001], and in the early cardiogenic region of *Xenopus* embryos [Zhang et al., 2005]. Knockdown of *SOX7* by injection of morpholinos displayed partial inhibition of cardiogenesis in *Xenopus* [Zhang et al., 2005]. It is conceivable that the combined copy number variation of *SOX7* and *GATA4* may induce more severe congenital heart defects.

We found a novel 1-Mb deletion that includes *SNAI2* (Subject J, with TOF). A child with a *de novo* 8q11.2->q13.3 duplication including *SNAI2* was associated with syndromic TOF [Perez-Mancera et al., 2006]. While heterozygous deletion of *SNAI2* resulted in human piebaldism [Sanchez-Martin et al., 2003], and a homozygous deletion has been detected in two individuals with Waardenburg disease type 2, ours is the first report of a deletion of *SNAI2* with a congenital heart defect [Sanchez-Martin et al., 2002]. *SNAI2* is expressed in various tissues including migratory neural crest cells and outflow tract [Jiang et al., 1998; Oram et al., 2003]. *SNAI2* is thus a plausible candidate gene for conotruncal defects and deserves further investigation.

We identified a large 7.3-Mb deletion at chromosome 10q23.1 from a subject with TOF (TABLE III: Subject I). Similar 7.3-Mb deletions were previously associated with cognitive impairment, autism, hyperactivity, and possibly psychiatric disease, and are likely mediated via non-allelic homologous recombination between low copy repeats [Alliman et al., 2010; Balciuniene et al., 2007; van Bon et al., 2011]. Congenital heart defects have been described in 3 out of 11 cases among those with this 7.3-Mb deletion, but no conotruncal defects [Alliman et al., 2010; van Bon et al., 2011]. The Endeavor program prioritized *BMPRIA* as the highest candidate among genes within the deleted region. A subject with syndromic

atrioventricular septal defect and a much smaller 22kb microdeletion that deleted only *BMPRIA* implicates this gene in heart development [Breckpot et al., 2012b]. To our knowledge ours is the first report that the 7.3-Mb deletion can cause TOF.

Two recent studies reported copy number variants from subjects with TOF [Greenway et al., 2009; Soemedi et al., 2012]. These copy number variants did not overlap with the microdeletions/duplications found in our study. Copy number variations from these studies together with our findings, however, emphasize the growing importance of genome wide assays that can detect microdeletions/duplications among subjects with conotruncal heart defects.

In summary, we found an increased risk of conotruncal defects among boys with sex chromosome aneuploidy and eight chromosomal microduplications/deletions providing evidence that *GATA4*, *SNAI2* and *CRKL* are dosage sensitive genes critical to normal outflow tract development. Array-CGH is a fruitful and a complementary approach to linkage analysis, genome wide association and whole exome sequencing for identifying novel candidate genes/loci associated with common nonsyndromic congenital malformations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Wei Yang (Stanford University) for contributing to data analysis, Drs. Johan Staaf and Åke Borg (Lund University) for production of BAC array-CGH slides. We thank the California Department of Public Health, Maternal Child and Adolescent Health Division for providing surveillance data for this study by the California Birth Defects Monitoring Program. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the California Department of Public Health. This research was supported by grants from National Institute of Health/National Heart Lung and Blood Institute HL092330, HL085859, HL077708 and from the Centers for Disease Control U50/CCU913241.

References

- Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, Tranchevent LC, De Moor B, Marynen P, Hassan B, Carmeliet P, Moreau Y. Gene prioritization through genomic data fusion. *Nat Biotechnol.* 2006; 24:537–44. [PubMed: 16680138]
- Alliman S, Coppinger J, Marcadier J, Thiese H, Brock P, Shafer S, Weaver C, Asamoah A, Leppig K, Dyack S, Morash B, Schultz R, Torchia BS, Lamb AN, Bejjani BA. Clinical and molecular characterization of individuals with recurrent genomic disorder at 10q22.3q23.2. *Clin Genet.* 2010; 78:162–8. [PubMed: 20345475]
- Balciuniene J, Feng N, Iyadurai K, Hirsch B, Charnas L, Bill BR, Easterday MC, Staaf J, Oseth L, Czapansky-Beilman D, Avramopoulos D, Thomas GH, Borg A, Valle D, Schimmenti LA, Selleck SB. Recurrent 10q22-q23 deletions: a genomic disorder on 10q associated with cognitive and behavioral abnormalities. *Am J Hum Genet.* 2007; 80:938–47. [PubMed: 17436248]
- Barber JC, Bunyan D, Curtis M, Robinson D, Morlot S, Dermitzel A, Liehr T, Alves C, Trindade J, Paramos AI, Cooper C, Ocraft K, Taylor EJ, Maloney VK. 8p23.1 duplication syndrome differentiated from copy number variation of the defensin cluster at prenatal diagnosis in four new families. *Mol Cytogenet.* 2010; 3:3. [PubMed: 20167067]
- Barber JC, Maloney VK, Huang S, Bunyan DJ, Cresswell L, Kinning E, Benson A, Cheetham T, Wyllie J, Lynch SA, Zwolinski S, Prescott L, Crow Y, Morgan R, Hobson E. 8p23.1 duplication syndrome; a novel genomic condition with unexpected complexity revealed by array CGH. *Eur J Hum Genet.* 2008; 16:18–27. [PubMed: 17940555]

- Barriot R, Breckpot J, Thienpont B, Brohee S, Van Vooren S, Coessens B, Tranchevent LC, Van Loo P, Gewillig M, Devriendt K, Moreau Y. Collaboratively charting the gene-to-phenotype network of human congenital heart defects. *Genome Med.* 2010; 2:16. [PubMed: 20193066]
- Breckpot J, Thienpont B, Bauters M, Tranchevent LC, Gewillig M, Allegaert K, Vermeesch JR, Moreau Y, Devriendt K. Congenital heart defects in a novel recurrent 22q11.2 deletion harboring the genes CRKL and MAPK1. *Am J Med Genet A.* 2012a; 158A:574–80. [PubMed: 22318985]
- Breckpot J, Tranchevent LC, Thienpont B, Bauters M, Troost E, Gewillig M, Vermeesch JR, Moreau Y, Devriendt K, Van Esch H. BMPR1A is a candidate gene for congenital heart defects associated with the recurrent 10q22q23 deletion syndrome. *Eur J Med Genet.* 2012b; 55:12–6. [PubMed: 22067610]
- Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* 2009; 37:305–11.
- Clark EB. Pathogenetic mechanisms of congenital cardiovascular malformations revisited. *Semin Perinatol.* 1996; 20:465–72. [PubMed: 9090774]
- Ferencz C, Rubin JD, McCarter RJ, Brenner JJ, Neill CA, Perry LW, Hepner SI, Downing JW. Congenital heart disease: prevalence at livebirth. The Baltimore-Washington Infant Study. *Am J Epidemiol.* 1985; 121:31–6. [PubMed: 3964990]
- Garg V, Kathiriyi IS, Barnes R, Schluterman MK, King IN, Butler CA, Rothrock CR, Eapen RS, Hirayama-Yamada K, Joo K, Matsuoka R, Cohen JC, Srivastava D. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature.* 2003; 424:443–7. [PubMed: 12845333]
- Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH, Korn JM, McCarroll SA, Gorham JM, Gabriel S, Altshuler DM, de Quintanilla-Dieck ML, Artunduaga MA, Eavey RD, Plenge RM, Shadick NA, Weinblatt ME, De Jager PL, Hafler DA, Breitbart RE, Seidman JG, Seidman CE. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet.* 2009; 41:931–5. [PubMed: 19597493]
- Hemmi K, Ma D, Miura Y, Kawaguchi M, Sasahara M, Hashimoto-Tamaoki T, Tamaoki T, Sakata N, Tsuchiya K. A homeodomain-zinc finger protein, ZFHX4, is expressed in neuronal differentiation manner and suppressed in muscle differentiation manner. *Biol Pharm Bull.* 2006; 29:1830–5. [PubMed: 16946494]
- Hollox EJ, Armour JA, Barber JC. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am J Hum Genet.* 2003; 73:591–600. [PubMed: 12916016]
- Jenkins KJ, Correa A, Feinstein JA, Botto L, Britt AE, Daniels SR, Elixson M, Warnes CA, Webb CL. Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation.* 2007; 115:2995–3014. [PubMed: 17519397]
- Jiang R, Lan Y, Norton CR, Sundberg JP, Gridley T. The Slug gene is not essential for mesoderm or neural crest development in mice. *Dev Biol.* 1998; 198:277–85. [PubMed: 9659933]
- Jonsson G, Staaf J, Olsson E, Heidenblad M, Vallon-Christersson J, Osoegawa K, de Jong P, Oredsson S, Ringner M, Hoglund M, Borg A. High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization. *Genes Chromosomes Cancer.* 2007; 46:543–558. [PubMed: 17334996]
- Karsh RB. Congenital heart disease in 49, XXXXY syndrome. *Pediatrics.* 1975; 56:462–4. [PubMed: 1161403]
- Kirby ML. Molecular embryogenesis of the heart. *Pediatr Dev Pathol.* 2002; 5:516–43. [PubMed: 12297889]
- Krzywinski M, Bosdet I, Smailus D, Chiu R, Mathewson C, Wye N, Barber S, Brown-John M, Chan S, Chand S, Cloutier A, Girm N, Lee D, Masson A, Mayo M, Olson T, Pandoh P, Prabhu AL, Schoenmakers E, Tsai M, Albertson D, Lam W, Choy CO, Osoegawa K, Zhao S, de Jong PJ, Schein J, Jones S, Marra MA. A set of BAC clones spanning the human genome. *Nucleic Acids Res.* 2004; 32:3651–60. [PubMed: 15247347]

- Kurahashi H, Tsuda E, Kohama R, Nakayama T, Masuno M, Imaizumi K, Kamiya T, Sano T, Okada S, Nishisho I. Another critical region for deletion of 22q11: a study of 100 patients. *Am J Med Genet.* 1997; 72:180–5. [PubMed: 9382139]
- Lammer EJ, Chak JS, Iovannisci DM, Schultz K, Osoegawa K, Yang W, Carmichael SL, Shaw GM. Chromosomal abnormalities among children born with conotruncal cardiac defects. *Birth Defects Res A Clin Mol Teratol.* 2009; 85:30–5. [PubMed: 19067405]
- Lary JM, Paulozzi LJ. Sex differences in the prevalence of human birth defects: a population-based study. *Teratology.* 2001; 64:237–51. [PubMed: 11745830]
- Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, Baldini A. Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature.* 2001; 410:97–101. [PubMed: 11242049]
- Ma C, Quesnelle KM, Sparano A, Rao S, Park MS, Cohen MA, Wang Y, Samanta M, Kumar MS, Aziz MU, Naylor TL, Weber BL, Fakhrazadeh SS, Weinstein GS, Vachani A, Feldman MD, Brose MS. Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues. *Cancer Biol Ther.* 2009; 8:907–16. [PubMed: 19276661]
- Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, Xavier RJ, Demay MB, Russell RG, Factor S, Tokooya K, Jore BS, Lopez M, Pandita RK, Lia M, Carrion D, Xu H, Schorle H, Kobler JB, Scambler P, Wynshaw-Boris A, Skoultschi AI, Morrow BE, Kucherlapati R. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell.* 2001; 104:619–29. [PubMed: 11239417]
- Meschede D, Nekarda T, Kececioğlu D, Loser H, Vogt J, Miny P, Horst J. Congenital heart disease in the 48,XXYY syndrome. *Clin Genet.* 1995; 48:100–2. [PubMed: 7586648]
- Moon AM, Guris DL, Seo JH, Li L, Hammond J, Talbot A, Imamoto A. Crkl deficiency disrupts Fgf8 signaling in a mouse model of 22q11 deletion syndromes. *Dev Cell.* 2006; 10:71–80. [PubMed: 16399079]
- Morris JK, Alberman E, Scott C, Jacobs P. Is the prevalence of Klinefelter syndrome increasing? *Eur J Hum Genet.* 2008; 16:163–70. [PubMed: 18000523]
- Nemer G, Fadlalah F, Usta J, Nemer M, Dbaibo G, Obeid M, Bitar F. A novel mutation in the GATA4 gene in patients with Tetralogy of Fallot. *Hum Mutat.* 2006; 27:293–4. [PubMed: 16470721]
- O'Malley CD, Shaw GM, Wasserman CR, Lammer EJ. Epidemiologic characteristics of conotruncal heart defects in California, 1987–1988. *Teratology.* 1996; 53:374–7. [PubMed: 8910983]
- Oram KF, Carver EA, Gridley T. Slug expression during organogenesis in mice. *Anat Rec A Discov Mol Cell Evol Biol.* 2003; 271:189–91. [PubMed: 12552634]
- Osoegawa K, Vessere GM, Utami KH, Mansilla MA, Johnson MK, Riley BM, L'Heureux J, Pfundt R, Staaf J, van der Vliet WA, Lidral AC, Schoenmakers EF, Borg A, Schutte BC, Lammer EJ, Murray JC, de Jong PJ. Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation. *J Med Genet.* 2008; 45:81–6. [PubMed: 17873121]
- Pehlivan T, Pober BR, Brueckner M, Garrett S, Slaugh R, Van Rheeden R, Wilson DB, Watson MS, Hing AV. GATA4 haploinsufficiency in patients with interstitial deletion of chromosome region 8p23.1 and congenital heart disease. *Am J Med Genet.* 1999; 83:201–6. [PubMed: 10096597]
- Perez-Mancera PA, Gonzalez-Herrero I, Maclean K, Turner AM, Yip MY, Sanchez-Martin M, Garcia JL, Robledo C, Flores T, Gutierrez-Adan A, Pintado B, Sanchez-Garcia I. SLUG (SNAI2) overexpression in embryonic development. *Cytogenet Genome Res.* 2006; 114:24–9. [PubMed: 16717446]
- Rauch A, Zink S, Zweier C, Thiel CT, Koch A, Rauch R, Lascorz J, Huffmeier U, Weyand M, Singer H, Hofbeck M. Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *J Med Genet.* 2005; 42:871–6. [PubMed: 15831592]
- Rauch R, Hofbeck M, Zweier C, Koch A, Zink S, Trautmann U, Hoyer J, Kaulitz R, Singer H, Rauch A. Comprehensive genotype-phenotype analysis in 230 patients with tetralogy of Fallot. *J Med Genet.* 2010; 47:321–31. [PubMed: 19948535]
- Reller MD, Strickland MJ, Riehle-Colarusso T, Mahle WT, Correa A. Prevalence of congenital heart defects in metropolitan Atlanta, 1998–2005. *J Pediatr.* 2008; 153:807–13. [PubMed: 18657826]

- Sanchez-Martin M, Perez-Losada J, Rodriguez-Garcia A, Gonzalez-Sanchez B, Korf BR, Kuster W, Moss C, Spritz RA, Sanchez-Garcia I. Deletion of the SLUG (SNAI2) gene results in human piebaldism. *Am J Med Genet A*. 2003; 122A:125–32. [PubMed: 12955764]
- Sanchez-Martin M, Rodriguez-Garcia A, Perez-Losada J, Sagrera A, Read AP, Sanchez-Garcia I. SLUG (SNAI2) deletions in patients with Waardenburg disease. *Hum Mol Genet*. 2002; 11:3231–6. [PubMed: 12444107]
- Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet*. 2000; 9:489–501. [PubMed: 10699172]
- Shaikh TH, O'Connor RJ, Pierpont ME, McGrath J, Hacker AM, Nimmakayalu M, Geiger E, Emanuel BS, Saitta SC. Low copy repeats mediate distal chromosome 22q11.2 deletions: sequence analysis predicts breakpoint mechanisms. *Genome Res*. 2007; 17:482–91. [PubMed: 17351135]
- Shaw GM, Carmichael SL, Yang W, Lammer EJ. Periconceptional nutrient intakes and risks of conotruncal heart defects. *Birth Defects Res A Clin Mol Teratol*. 2010; 88:144–51. [PubMed: 20063270]
- Soemedi R, Wilson IJ, Bentham J, Darlay R, Topf A, Zelenika D, Cosgrove C, Setchfield K, Thornborough C, Granados-Riveron J, Blue GM, Breckpot J, Hellens S, Zwolinski S, Glen E, Mamasoula C, Rahman TJ, Hall D, Rauch A, Devriendt K, Gewillig M, JOS, Winlaw DS, Bu'Lock F, Brook JD, Bhattacharya S, Lathrop M, Santibanez-Koref M, Cordell HJ, Goodship JA, Keavney BD. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am J Hum Genet*. 2012; 91:489–501. [PubMed: 22939634]
- Takash W, Canizares J, Bonneaud N, Poulat F, Mattei MG, Jay P, Berta P. SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. *Nucleic Acids Res*. 2001; 29:4274–83. [PubMed: 11691915]
- Tartaglia N, Ayari N, Howell S, D'Epagnier C, Zeitler P. 48,XXYY, 48,XXXY and 49,XXXXY syndromes: not just variants of Klinefelter syndrome. *Acta Paediatr*. 2011; 100:851–60. [PubMed: 21342258]
- Tennant PW, Samarasekera SD, Pless-Mulloli T, Rankin J. Sex differences in the prevalence of congenital anomalies: a population-based study. *Birth Defects Res A Clin Mol Teratol*. 2011; 91:894–901. [PubMed: 21987467]
- Tomita-Mitchell A, Maslen CL, Morris CD, Garg V, Goldmuntz E. GATA4 sequence variants in patients with congenital heart disease. *J Med Genet*. 2007; 44:779–83. [PubMed: 18055909]
- Vallejo-Illarramendi A, Zang K, Reichardt LF. Focal adhesion kinase is required for neural crest cell morphogenesis during mouse cardiovascular development. *J Clin Invest*. 2009; 119:2218–30. [PubMed: 19587446]
- van Bon BW, Balciuniene J, Fruhman G, Nagamani SC, Broome DL, Cameron E, Martinet D, Roulet E, Jacquemont S, Beckmann JS, Irons M, Potocki L, Lee B, Cheung SW, Patel A, Bellini M, Selicorni A, Ciccone R, Silengo M, Vetro A, Knoers NV, de Leeuw N, Pfundt R, Wolf B, Jira P, Aradhya S, Stankiewicz P, Brunner HG, Zuffardi O, Selleck SB, Lupski JR, de Vries BB. The phenotype of recurrent 10q22q23 deletions and duplications. *Eur J Hum Genet*. 2011; 19:400–8. [PubMed: 21248748]
- van Engelen K, Topf A, Keavney BD, Goodship JA, van der Velde ET, Baars MJ, Snijder S, Moorman AF, Postma AV, Mulder BJ. 22q11.2 Deletion Syndrome is under-recognised in adult patients with tetralogy of Fallot and pulmonary atresia. *Heart*. 2010; 96:621–4. [PubMed: 20357389]
- Visser LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet*. 2004; 36:955–7. [PubMed: 15300250]
- Wat MJ, Shchelochkov OA, Holder AM, Breman AM, Dagli A, Bacino C, Scaglia F, Zori RT, Cheung SW, Scott DA, Kang SH. Chromosome 8p23.1 deletions as a cause of complex congenital heart defects and diaphragmatic hernia. *Am J Med Genet A*. 2009; 149A:1661–77. [PubMed: 19606479]

- Xu H, Morishima M, Wylie JN, Schwartz RJ, Bruneau BG, Lindsay EA, Baldini A. Tbx1 has a dual role in the morphogenesis of the cardiac outflow tract. *Development*. 2004; 131:3217–27. [PubMed: 15175244]
- Yagi H, Furutani Y, Hamada H, Sasaki T, Asakawa S, Minoshima S, Ichida F, Joo K, Kimura M, Imamura S, Kamatani N, Momma K, Takao A, Nakazawa M, Shimizu N, Matsuoka R. Role of TBX1 in human del22q11.2 syndrome. *Lancet*. 2003; 362(9393):1366–73. [PubMed: 14585638]
- Zhang C, Basta T, Klymkowsky MW. SOX7 and SOX18 are essential for cardiogenesis in *Xenopus*. *Dev Dyn*. 2005; 234:878–91. [PubMed: 16193513]

Figure 1.

Array-CGH (Roche Nimblegen) profiles for a series of chromosome 8 duplication cases. A: 3.9-Mb duplication at chromosome 8p23.1. B: 2.2-Mb duplication at chromosome 8p23.2. C: Copy number variation (CNV: deletion) in defensin gene cluster region. The two duplication cases are localized distal and proximal, respectively, to the defensin gene region. D: 1.0 Mb duplication at chromosome 8q21.11.

Figure 2.

A series of chromosome 22q11.21–q11.22 deletion cases. The top BAC array-CGH profile indicates one of the typical 3-Mb deletion at chromosome 22q11.21 cases. A 600-kb deletion that overlaps with the typical deletion case at the distal end is shown in BAC array-CGH profile (second from the top: Subject F). The third array-CGH profile indicates a 800-kb deletion at chromosome 22q11.22 (Subject G). The relative position in the genome for low copy repeats (LCRs) denoted as A – F and the three deletion cases are depicted at the bottom. The relative position of *TBX1*, *CRKL* and *MAPK1* genes are shown in the figure.

Figure 3.

Long range PCR to narrow down deletion break points. A set of PCR primers previously published [Shaikh et al., 2007] was used to amplify 7-kb PCR fragments containing deletion break points. The PCR products were loaded into 0.7% agarose gel. Lane 1 contains 1-kb DNA ladder (New England Biolabs). Long range PCR was performed with DNA isolated from subject H (TABLE III) who has 800-kb deletion at chromosome 22q11.22 and control DNA who does not have the deletion. A 7-kb PCR product is observed from Subject H (Lane 2), but not seen from control DNA (Lane 3).

TABLE I

Study population

	Male	Female	Unknown	Total
Live-born	496,215	472,654	16	968,885
Stillborn	2,938	2,717	39	5,694
Total	499,153	475,371	55	974,579

TABLE II

Study group with d-transposition of the great arteries or tetralogy of Fallot

	d-TGA	TOF	Total	Retrieved d-TGA	Retrieved TOF	Retrieved Total
Male	135	153	288	115	116	231
Female	59	137	196	49	109	158
Total	194	290	484	164	225	389

d-TGA, d-transposition of the great arteries; TOF, tetralogy of Fallot; Retrieved, of 194 subjects with d-TGA and 290 with TOF, we retrieved newborn bloodspots for 164 with d-TGA and 225 with TOF.

TABLE III
Eight microdeletions/duplications among 389 subjects with conotruncal heart defects

Subject ID	Del/Dup	Cytogenetic band	Duplicated or deleted loci	Size Mb	Phenotype	Candidate gene
C	Dup	8p23.1	chr8:8127502-11849954	3.7	TOF	<i>GATA4</i>
D	Dup	8p23.2	chr8:3691799-5913749	2.2	d-TGA	
E	Dup	8q21.11	chr8:76874286-77683425	0.8	TOF	<i>ZFX4</i>
F	Del	22q11.21	chr22:20912000-21524524	0.6	TOF	<i>CRKL</i>
G	Del	22q11.22	chr22:22782875-23593344	0.8	d-TGA	
H	Del	1q44	chr1:246893732-248877423	2.0	TOF	
I	Del	10q23.1	chr10:81561361-88914011	7.3	TOF	<i>BMPRIA</i>
J	Del	8q11.21	chr8:48815290-49831818	1.0	TOF	<i>SNAI2</i>

d-TGA, d-transposition of the great arteries; TOF, tetralogy of Fallot; Del, deletion; Dup, duplication

The genomic DNA sequence positions for each deletion/duplication are indicated with hg19 UC Santa Cruz Genome Browser. Candidate gene column lists the highest ranked candidate genes.