De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot

Steven C Greenway, Alexandre C Pereira, Jennifer C Lin, Steven R DePalma, Samuel J Israel, Sonia M Mesquita, Emel Ergul, Jessie H Conta, Joshua M Korn, Steven A McCarroll, Joshua M Gorham, Stacey Gabriel, David M Altshuler, Maria de Lourdes Quintanilla-Dieck, Maria Alexandra Artunduaga, Roland D Eavey, Robert M Plenge, Nancy A Shadick, Michael E Weinblatt, Philip L De Jager, David A Hafler, Roger E Breitbart, Jonathan G Seidman & Christine E Seidman

Tetralogy of Fallot (TOF), the most common severe congenital heart malformation, occurs sporadically, without other anomaly, and from unknown cause in 70% of cases. Through a genome-wide survey of 114 subjects with TOF and their unaffected parents, we identified 11 de novo copy number variants (CNVs) that were absent or extremely rare (<0.1%) in 2,265 controls. We then examined a second, independent TOF cohort (n = 398) for additional CNVs at these loci. We identified CNVs at chromosome 1q21.1 in 1% (5/512, P = 0.0002, OR = 22.3) of nonsyndromic sporadic TOF cases. We also identified recurrent CNVs at 3p25.1, 7p21.3 and 22q11.2. CNVs in a single subject with TOF occurred at six loci, two that encode known (NOTCH1, JAG1) disease-associated genes. Our findings predict that at least 10% (4.5–15.5%, 95% confidence interval) of sporadic nonsyndromic TOF cases result from de novo CNVs and suggest that mutations within these loci might be etiologic in other cases of TOF.

The combination of a malpositioned aorta that overrides both ventricles, ventricular septal defect, pulmonary stenosis (which obstructs blood flow into the lungs) and right ventricular hypertrophy (Fig. 1) defines TOF. The most prevalent form of cyanotic heart disease, TOF occurs in one of 3,000 live births and accounts for 10% of all serious congenital heart disease. With recent advances in corrective surgery, early lethality from TOF is rare, but long-term sequelae, including arrhythmia, ventricular dysfunction and often lifelong disability, persist.

TOF can arise in the context of prenatal infections, exposure to teratogens or maternal illness, and from dominant mutations that...
usually alter gene dosage. Haploinsufficiency of cardiac transcription factor genes (NKX2–5, TBX1, TBX5 and GATA4) or the transmembrane receptors NOTCH1 and NOTCH2 and their ligand JAG1 can cause TOF, but, more commonly, mutations in these genes produce other heart malformations2–7. Cytogenetic abnormalities, including deletions of chromosome 22q11.2 (DiGeorge syndrome) or trisomy 21 (Down syndrome), account for 15% and 7% of TOF, respectively; however, these individuals usually have many noncardiac abnormalities8,9. Large de novo CNVs identified by array CGH occur with major congenital anomalies10,11; these typically include congenital heart disease (CHD) in half of all cases12. Far less is known about genes that cause sporadic and isolated CHD, particularly genes involved in complex malformations.

We hypothesized that de novo mutations that alter the dosage of genes involved in cardiac development might account for isolated TOF. We surveyed the genome of 121 TOF trios, each comprised of one proband and two unaffected parents, using the Affymetrix 6.0 array (Supplementary Fig. 1). CNVs identified in subjects with TOF but absent from parental samples using the algorithm Birdseye13 (putative de novo CNVs) were studied further. CNVs that corresponded to known copy number polymorphisms (CNPs)14 or that were smaller than 20 kb were discarded. To distinguish between TOF CNVs that shared CNVs at 10 unique loci from 114 TOF trios.

We considered whether the frequency of de novo CNVs differed between subjects with TOF and healthy subjects by analyzing 98 control trios, including 55 HapMap trios, genotyped using the Affymetrix 6.0 array (Supplementary Fig. 1). Using the same computational algorithm, we identified 20 putative de novo CNVs: 12 in HapMap trios and 8 in other control trios. Seven of the CNVs found in HapMap trios have been previously attributed to cell-line artifacts (chromosome 14:105829131–106116317 in NA10854, NA10838, NA06991, NA18857 and NA19154; chromosome 22:20777493–21581602 in NA12707 and NA19154)16, and two CNVs in HapMap trios and seven CNVs found in the other control trios fulfilled our criteria as CNPs. Four CNVs were validated by MLPA as being de novo in the control trios (Supplementary Table 2).

The frequency of rare de novo CNVs was greater in TOF trios than in control trios, but the difference was not statistically significant (11/114 versus 4/98, P = 0.18). Although de novo CNVs and pathogenesis seems to be related in schizophrenia17,18 and autism19, our trio study may be underpowered to detect a similar relationship in TOF. Alternatively, TOF mutations may be incompletely penetrant, a consideration that prompted us to assess whether inherited CNVs occurred at loci discovered by our de novo CNVs analyses. In three TOF trios, we identified CNVs at the 1q21.1, 3p25.1 and 7p21.3 loci that were inherited from unaffected parents (Table 1), a finding that supports the role of other genetic or environmental interactions in TOF.

To further evaluate the pathogenicity of CNVs, we used MLPA to assess nine loci in a second cohort of sporadic, nonsyndromic TOF cases (n = 398). Because the subjects in this validation cohort had previous chromosome 22q11.2 analyses, we excluded this locus from further study. We designed at least two unique synthetic oligonucleotide MLPA probes (Supplementary Table 3) to hybridize within each of the nine loci. MLPA studies demonstrated four more subjects with TOF having 1q21.1 CNVs (three duplications, one deletion; Supplementary Table 1). The boundaries of these CNVs were delineated by Affymetrix 6.0 array analyses. In combination with our initial genome-wide studies, we found a total of 17 CNVs at 10 loci in 512 individuals
with TOF (Table 1). CNVs at each of these loci were absent or very rare in 2,265 controls. CNVs at four loci (1q21.1, 3p25.1, 7p21.3 and 22q11.2) were found in at least two individuals with TOF. Because small CNVs would be predicted to escape detection by the array platform and detection algorithm used here, our data define a minimum estimate, approximately 10% (11/114), for the frequency of de novo CNVs in sporadic, isolated TOF. This was lower than the 25–30% frequency of de novo events seen in individuals with sporadic TOF, we sequenced exons and flanking splice sites in six genes. In 96 subjects with independent, overlapping regions of 3,756 deletions (data not shown). Syndromic heart malformations that occur with additional birth defects, but notably, identified causes for isolated TOF.

At chromosome 1q21.1, we found CNVs that were structurally complex in five TOF cases (Fig. 2a,b) (Supplementary Fig. 2). The shared duplicated segment in four subjects with TOF spans a small interval on chromosome 1q21 (chr1:144965244–145840510) that encompasses six known genes expressed in the human RVOT. Previously described duplications at this locus are associated with mental retardation (MR), macrocephaly (MaC) and other congenital phenotypes. Studies have identified multiple individuals (number of individuals in parentheses; only the minimal overlapping region between individuals is shown) that carry deletions at this locus with congenital heart disease other than TOF (CHD), MR, schizophrenia (SCZ), macrocephaly (MiC) or CHD and MR. (b) Plot showing normalized probe intensity measurements across 1q21 in the four subjects with TOF carrying a deletion (turquoise), the one subject with a duplication (green) and 273 copy number–neutral controls (gray lines) and summarized as mean (black) ± 2 × median absolute deviation (dark blue lines). Vertical dotted lines, boundaries of the 875,266-bp overlapping region. Absence of circles on the colored lines indicates an absence of probes owing to segmental duplications; there is a known CNP downstream of our region of interest. (c) A 12,380,330-bp duplication on chromosome 3 in a single individual with TOF (756) and an inherited duplication in this interval in a second individual (419), which narrowed the interval to a region (chr3:12605755–12781130) affecting RAF1. White bars, deletion; black bars, duplication; red bars, region of overlap between TOF CNVs. Chromosome position is indicated in Mb by the blue bar. All coordinates are based on build 36.1 of the human reference genome.
Table 2 Expression of TOF CNV genes in the human right ventricular outflow tract

| CNV     | Gene     | Tag count |
|---------|----------|-----------|
| 1q21.1  | PRKAB2   | 8 ± 3     |
|         | PDLA3P   | 0         |
|         | FM05     | 1 ± 1     |
|         | CHD1L    | 12 ± 6    |
|         | BCL9     | 11 ± 4    |
|         | ACP6     | 3 ± 2     |
|         | GJA5     | 8 ± 4     |
| 3p25.1  | RALI     | 99 ± 20   |
|         | TMEM40    | 1 ± 1     |
| 7p21.3  | No genes | 0         |
| 2q11.2  | TBX1     | 2 ± 1     |
| 9q34.3  | NOTCH1   | 22 ± 6    |
| 20q12.2 | JAG1     | 43 ± 8    |
| 2p23.3  | ASXL2    | 23 ± 5    |
|         | KIF3C    | 4 ± 1     |
|         | RAB10    | 70 ± 24   |
| 2p15    | BCL11A   | 15 ± 4    |
|         | PAPOLG   | 12 ± 4    |
|         | REL      | 2 ± 1     |
|         | PUS10    | 0         |
|         | PEX13    | 13 ± 4    |
|         | KIAA1841 | 10 ± 4    |
|         | AHSA2    | 30 ± 10   |
|         | USP34    | 33 ± 15   |
|         | SNORA7OB | 0         |
|         | XP01     | 40 ± 11   |
|         | FAM161A  | 0         |
|         | CCT4     | 59 ± 15   |
| 4q22.1  | PPM1K    | 17 ± 6    |
| 10q11.21| No genes | 0         |

CNV: cytogenetic location of TOF copy number variant identified in Table 1; Gene, name of reference gene within TOF CNV according to build 36.1 of the human reference genome; Tag count, number of sense tags mapped to reference gene from four separate human right ventricular outflow tract mRNA expression libraries, expressed as mean ± s.d. For loci containing a known disease-associated gene (bold), the expression profile for only that gene is shown.

Notably, there is no perfect correlation between 1q21 dosage and phenotype in any study so far (including ours). Two studies described mild (seven cases) or severe (five cases, not including TOF) cardiovascular malformations, but nine of these individuals had additional phenotypes: developmental or intellectual disabilities, dysmorphic craniofacial features or other congenital anomalies. In contrast, all subjects with TOF with 1q21.1 duplications significantly (P = 0.007) more common in TOF, occurring in 1% of cases studied here.

Two subjects with TOF shared overlapping duplications at the 3p25.1 locus. Whereas one duplication spanned 12 Mb, the other duplication affected only two genes, RAFL1 and TMEM40 (Fig. 2c). RAFL1 is expressed approximately 100-fold more highly than TMEM40 in the RVOT (Table 2). Gain-of-function point mutations of RAFL1 cause Noonan syndrome, a multisystem disorder with cardiac manifestations that rarely includes TOF but produces one of its components: hypertrophy, atrial or ventricular septal defects or pulmonary stenosis. Identification of CNVs at 3p25.1 in TOF cases prompted reevaluation for signs of Noonan syndrome (Table 3). Subtle craniofacial abnormalities were identified in one individual (subject 756), suggesting some phenotypic overlap between RAFL1 gain-of-function mutations and increases in RAFL1 copy number produced by the 12-Mb duplication. The smaller 3p25.1 CNV truncates and duplicates RAFL1, and further study is necessary to determine which alteration causes the TOF phenotype.

Table 3 Phenotypic data for individuals with TOF and 1q21.1 and 3p25.1 CNVs

| Proband | CNV     | CN | Age (y) | Sex | Extracardiac features | Development | Neuropsychiatric features |
|---------|---------|----|---------|-----|-----------------------|-------------|--------------------------|
| 749     | 1q21.1  | Gain| 21 (<1) | F   | None                  | Normal      | None                     |
| 201.670 | 1q21.1  | Gain| 9 (<1)  | F   | None                  | Normal      | None                     |
| 200.430 | 1q21.1  | Gain| 4 (<1)  | M   | None                  | Normal      | None                     |
| 200.250 | 1q21.1  | Gain| 19 (<1) | F   | None                  | Normal      | None                     |
| 3701    | 1q21.1  | Loss| 19 (<1) | M   | None                  | Normal      | None                     |
| 756     | 3p25.1  | Gain| 13 (<1) | M   | High-arched palate, malocclusion, subteliaural dysplasia | Normal      | Hyperactivity |
| 419     | 3p25.1  | Gain| 15 (<1) | F   | None                  | Normal      | None                     |

Proband, individual subject identification number; CNV, cytogenetic location of copy number variant; CN, copy number; Age, current age (and age at diagnosis) in years; Sex, female (F) or male (M).
malformations. Mice engineered to lack components of the NOTCH1 signaling pathway have TOF-like phenotypes. Our findings provide direct evidence for NOTCH1 mutations in TOF. JAG1 mutations are known to cause TOF in the context of Alagille syndrome and in isolation. The individual with the 4-Mb de novo deletion of JAG1 identified here has no clinical features of Alagille syndrome. The discovery of CNVs altering NOTCH1 and JAG1 underscores the need to assess gene dosage in mutation analyses of congenital heart disease–associated genes.

Four other loci were altered by CNVs in a single individual with TOE CNVs at these candidate TOF loci occurred at a low frequency similar to those of NOTCH1 and JAG1 mutations, emphasizing the genetic heterogeneity of TOF. Three genes are encoded at the 2p23.3 locus: RAB10, KIF3C and ASXL2. Although none have been previously implicated in cardiac development, each is expressed in the RVOT (Table 2). The expression of RAB10, which encodes a GTPase, is highest. Because KRAS, another GTPase, is activated by Noonan syndrome mutations, RAB10 is a promising candidate gene in TOF.

Pathogenicity of a 1.8-Mb de novo duplication at 2p15 found in one TOF case is likely, given its size and the absence of CNVs at this locus in all controls. Of 12 genes located in the duplicated interval, 9 are expressed in the human RVOT (Table 2) and none were previously implicated in cardiogenesis. The 4q22.1 deletion encompasses PPMIK, and its transcripts, encoding a phosphatase, are present in RVOT tissues. The deletion at 10q11.2 contains no known genes or coding or noncoding RNAs.

In summary, our studies defined seven new loci that substantially increase risk (odds ratio ≥ 8.9) for sporadic, nonsyndromic TOF. Although some loci are large (>100 kb) and three loci showed incomplete penetrance, expression data from human RVOT highlights a subset of genes that should be prioritized in future studies. Moreover, the data identified de novo CNVs at 10 loci, accounting for 10% of TOF in our case series, which explains sporadic presentation and defines genetic heterogeneity in this serious heart malformation.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/. Published online at http://www.nature.com/naturegenetics/

Note: Supplemental information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We gratefully acknowledge the participation of families. We also thank C. Souczek and M. Parkin for technical assistance and R. Geggel for supplying TOF images. This work was supported by grants from the Howard Hughes Medical Institute (C.E.S.), US National Institutes of Health (to C.E.S., J.G.S. and R.E.B. and to the Broad Institute (National Center for Research Resources)), Pediatric Scientist Development Program (S.C.G.) and Sarnoff Cardiovascular Research Foundation (J.C.L.). Multiple sclerosis controls were genotyped in collaboration with Affymetrix, Inc.

AUTHOR CONTRIBUTIONS

S.C.G., A.C.P., R.E.B., J.G.S. and C.E.S. designed the experiments. S.C.G., I.C.L., S.I.J. and J.M.G. performed the experiments. S.C.G., S.R.D., J.M.K., S.A.M., S.G., D.M.A., J.G.S. and C.E.S. were involved in genotyping and data analysis. E.E., J.H.C., A.C.P., M.M., M.L.Q.-D., M.A.A., R.D.E., R.M.P., N.A.S., M.E.W., E.L.D.J., D.A.H. and R.E.B. recruited subjects and collected DNA. S.C.G., I.G.S. and C.E.S. wrote the paper with input from all authors.

1. Ferencz, C. et al. Congenital heart disease: prevalence at livebirth. The Baltimore- Washington Infant Study. Am. J. Epidemiol. 121, 31–36 (1985).
2. Schott, J.J. et al. Congenital heart disease caused by mutations in the transcription factor NKX2–5. Science 281, 108–111 (1998).
3. Basson, C.T. et al. Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. Nat. Genet. 15, 30–35 (1997).
4. Eldadah, Z.A. et al. Familial tetralogy of Fallot caused by mutation in the jagged1 gene. Hum. Mol. Genet. 10, 163–169 (2001).
5. Tomita-Mitchell, A., Maslen, C.L., Morris, C.D., Garg, V. & Goldmuntz, E. GATA4 sequence variants in patients with congenital heart disease. J. Med. Genet. 44, 779–783 (2007).
6. McDaniel, R. et al. NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. Am. J. Hum. Genet. 79, 169–173 (2006).
7. Garg, V. et al. Mutations in NOTCH1 cause aortic valve disease. Nature 437, 270–274 (2005).
8. Goldmuntz, E. et al. Frequency of 22q11 deletions in patients with conotruncal defects. J. Am. Coll. Cardiol. 32, 492–498 (1998).
9. Danby, B. et al. Role of TBX1 in human de novo 12q13.2 syndrome. Lancet 362, 1366–1373 (2003).
10. Thienpont, B. et al. Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. Eur. Heart J. 28, 2778–2784 (2007).
11. Richards, A.A. et al. Cryptic chromosomal abnormalities identified in children with congenital heart disease. Pediatr. Res. 64, 358–363 (2008).
12. Loffredo, C.A. Epidemiology of cardiovascular malformations: prevalence and risk factors. Am. J. Med. Genet. 97, 319–325 (2000).
13. Korn, J.M. et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat. Genet. 40, 1253–1260 (2008).
14. McCarroll, S.A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat. Genet. 40, 1166–1174 (2008).
15. Sato, M. et al. The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28. Arthritis Res. Ther. 8, R57 (2006).
16. Redon, R. et al. Global variation in copy number in the human genome. Nature 444, 444–454 (2006).
17. Xu, B. et al. Strong association of de novo copy number mutations with sporadic schizophrenia. Nat. Genet. 40, 880–885 (2008).
18. Steffansson, H. et al. Large recurrent microdeletions associated with schizophrenia. Nature 455, 232–236 (2008).
19. Sebat, J. et al. Strong association of de novo copy number mutations with autism. Science 316, 445–449 (2007).
20. Christiansen, J. et al. Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. Circ. Res. 94, 1429–1435 (2004).
21. de Vries, B.B. et al. Diagnostic genome profiling in mental retardation. Am. J. Hum. Genet. 77, 606–616 (2005).
22. Brunetti-Pierri, N. et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat. Genet. 40, 1466–1471 (2008).
23. Mefford, H.C. et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. N. Engl. J. Med. 359, 1685–1699 (2008).
24. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 455, 237–241 (2008).
25. Walsh, T. et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320, 539–543 (2008).
26. Sharp, A.J. et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. Nat. Genet. 38, 1038–1042 (2006).
27. Razaque, M.A. et al. Germline gain-of-function RAI1 mutations cause Noonan syndrome. Nat. Genet. 39, 1013–1017 (2007).
28. Pandit, B. et al. Gain-of-function RAI1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. Nat. Genet. 39, 1007–1012 (2007).
29. Bassett, A.S. et al. Clinical features of 78 adults with 22q11 deletion syndrome. Am. J. Med. Genet. A 138, 307–313 (2005).
30. McDonald-McGinn, D.M. et al. Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: cast a wide FISHing net!. Genet. Med. 3, 23–29 (2001).
31. Niessen, K. & Karsan, A. Notch signaling in cardiac development. Circ. Res. 102, 1169–1181 (2008).
32. Krantz, I.D. et al. Spectrum and frequency of jagged1 (JAG1) mutations in Alagille syndrome patients and their families. Am. J. Hum. Genet. 62, 1361–1369 (1998).
33. Schubbert, S. et al. Germline KRAS mutations cause Noonan syndrome. Nat. Genet. 38, 331–336 (2006).
de novo control trios were analyzed using Birdseed version 1.5 with 97.8 ± 1.3% and Y chromosomes were excluded from analyses. Initially 121 TOF trios and mendelian errors present in each trio as calculated by PLINK were excluded. There was no evidence of nonpaternity (determined by the number of individuals) was identified, subjects were excluded from analyses. CNVs due to abnormalities were excluded. When 22q11.2 analyses (routinely obtained on all TOF cases with clinical features of developmental abnormalities in these studies, which were performed in accordance with institutional guidelines. TOF was diagnosed based on noninvasive imaging (two-dimen-
sional echocardiography and/or MRI) and/or invasive studies (cardiac cathe-
terization and/or surgery). TOF cases with clinical features of developmental syndromes, multiple major developmental anomalies or major cytogenetic abnormalities were excluded. When 22q11.2 analyses (routinely obtained on all Boston subjects and obtained when clinically indicated on Brazilian subjects) revealed microdeletion, subjects were excluded. TOF parents had neither significant congenital or cardiac disease. Control subjects were assembled from the BRASS cohort (n = 538), subjects with multiple sclerosis (n = 934) and healthy controls (n = 271) collected from Brigham & Women’s Hospital and unrelated parents of individuals with TOF (n = 228). Control trios were assembled from HapMap CEU and YRI subjects (n = 165) and South American controls (n = 129). Control subjects and trios had neither significant congenital nor cardiac disease.

Genotyping and identification of CNVs. Genotyping was performed using the Affymetrix Human Genome-Wide SNP Array 6.0 at the Broad Institute (TOF trios and controls) and Affymetrix (HapMap trios). Genotypes from X and Y chromosomes were excluded from analyses. Initially 121 TOF trios and 98 control trios were analyzed using Birdseed version 1.5 with 97.8 ± 1.3% and 99.7 ± 0.6% average call rates achieved, respectively (Supplementary Fig. 1).

There was no evidence of nonpaternity (determined by the number of mendelian errors present in each trio as calculated by PLINK) in the trios used. Genotype information on the CEU and YRI HapMap trios was obtained directly from Affymetrix. The Birdseye CNV-detection algorithm was used to identify de novo CNVs in trios using a confidence (lod) score of 10^10 for the proband to be copy number (CN) variable (CN = 0, 1, 3 or 4) and a lod score of 10^6 for the parents to be copy-number variable. This was done in order to maximize the probability that the child truly possessed a de novo CNV. CNVs in proband samples that were absent from both parental samples were considered putative de novo CNVs. Those putative de novo CNVs that corresponded to known copy number polymorphisms (CNPs) or that were smaller than 20 kb were discarded. De novo CNVs with ≥ 50% overlap with CNVs found in ≥ 0.1% of 2,265 control samples were designated CNPs. When an excessive number of de novo CNVs (> 3 s.d. above the mean number of CNVs per individual) was identified, subjects were excluded from analyses. CNVs due to cell-line artifacts (see main text) were identified in HapMap samples and were discarded. CNV locations are based on the March 2006 human reference sequence (US National Center for Biotechnology Information (NCBI) build 36.1). For non-trio controls, CNV calls were based on Birdseye using a lod score of 10^10 to be copy-number variable (CN = 0, 1, 3 or 4).

Multiplex ligation-dependent probe amplification (MLPA). At least two independent MLPA probes (Supplementary Table 3) corresponding to sequences encompassed by CNVs were designed to confirm each de novo CNV in samples from subjects with TOF and parents to identify new CNVs at these loci (excluding chromosome 22q11.2 deletions) in a screen of a second cohort of 398 subjects with TOF. Synthetic oligonucleotide probes with a final product size of 90–160 bp (including universal sequences) were designed from genomic sequences (March 2006 human reference sequence, NCBI build 36.1). Probe design sought to maximize unique hybridization using the BLAT program (http://genome.ucsc.edu), a Tm > 65 °C and G+C content 40%–60% according to IDT Oligoanalyzer 3.0 (http://www.idtdna.com/analyzer/ Applications/OligoAnalyzer/), absence of known SNPs (http://www.ncbi.nlm.nih.gov/SNP/) within the hybridizing region and to have no more than three cytosine or guanine bases flanking the ligation site. Probes were designed to differ by at least 4 bp in length to prevent overlapping mobility during electrophoresis. To allow ligation, downstream probes were 5'-phosphorylated and all probes used in these studies (sequences available upon request) were PAGE-purified after synthesis (IDT). DNA was purified from peripheral blood lymphocytes using routine phenol-chloroform extraction or from epithelial cells in saliva according to the manufacturer’s instructions (http://www.dna genetek.com/). DNA quality was assessed by agarose gel electrophoresis after denaturation, and only high quality DNA samples and MRC-Holland reagents were used for MLPA. MLPA reactions were performed as described, with a maximum of 20 probes (final concentration 2 nM) and 100 ng genomic DNA. After heating (95 °C) for 1 min, hybridization reactions continued for 16 h (60 °C). Hybridized probes were ligated using 1 U ligase-65 (MRC-Holland) for 15 min (54 °C) followed by ligase deactivation (98 °C for 5 min). Ligation product (5 μl) was added to PCR buffer, heated (60 °C) and then PCR reagents (2.5 nmol dNTPs, SALSA polymerase (MRC-Holland), 10 pmol universal primers 5'-FAM-GGGTCCCCCTAAAGGTTGA-3' and 5'-TCTAGATTTGACCTT GCTGGAC-3') were added to achieve a final 25 μl reaction. PCR was carried out for 33 cycles (95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min). Products were resolved by capillary electrophoresis on an Applied Biosystems 3730xl, and peaks were manually reviewed in GeneMapper 3.7 (Applied Biosystems). MLPA studies were performed in triplicate in at least two separate experiments.

Calculation of MLPA dosage quotient. Copy number was deduced from dosage quotient. A peak ratio was calculated by dividing probe peak area by the sum of all peak areas in each reaction. Each experimental peak ratio was divided by the average of control peak ratios to normalize for variation in probe signal strength. Control peak ratios were derived from probes hybridizing to unique, undeleted chromosomal locations on unrelated genes. The normalized peak ratio was divided by the average of each probe’s peak ratio among control samples (individuals without TOF or other CHD) to eliminate variation between samples.

Human right ventricular outflow tract mRNA expression libraries. Expression libraries were created from RNA isolated from right ventricular outflow tract tissue collected from four TOF cases at the time of primary surgical repair (mean patient age 2.6 ± 2 months). Tissue was snap-frozen in liquid nitrogen and maintained at −80 °C before processing. RNA extraction and library construction and amplification was carried out as previously described. Amplified library was sequenced on an Illumina Genome Analyzer. Each library generated more than 2,000,000 reads with a Chastity score (Illumina) > 2. Sense tags were assigned to cognate gene identities, and unique tags assigned to the same UniGene cluster or gene symbol were combined.

34. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575 (2007).
35. Stern, R.F. et al. Multiplex ligation-dependent probe amplification using a completely synthetic probe set. Biotechniques 37, 399–405 (2004).
36. Schouten, J.P. et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. 30, e57 (2002).
37. Yau, S.C., Bobrow, M., Mathew, C.G. & Abbs, S.J. Accurate diagnosis of carriers of deletions or that were smaller than 20 kb. J. Med. Genet. 33, 550–558 (1996).
38. Kim, J.B. et al. Polony multiplex analysis of gene expression (PMAGE) in mouse hypertrophic cardiomyopathy. Science 316, 1481–1484 (2007).