Human gene copy number spectra analysis in congenital heart malformations

Aoy Tomita-Mitchell,1,2,3 Donna K. Mahnke,1,2,3 Craig A. Struble,8 Maureen E. Tuffnell,1,8 Karl D. Stamm,1,8 Mats Hidestrand,1,2,3 Susan E. Harris,4 Mary A. Goetsch,1,2,3 Pippa M. Simpson,6 David P. Bick,5 Ulrich Broeckel,7 Andrew N. Pelech,4 James S. Tweddell,4 and Michael E. Mitchell1

1Department of Surgery, Division of Cardiothoracic Surgery; 2Biotechnology and Bioengineering Center; 3Human and Molecular Genetics Center; 4Department of Pediatrics, Division of Cardiology; 5Department of Pediatrics, Division of Genetics; 6Department of Pediatrics, Division of Quantitative Health Sciences; 7Department of Pediatrics, Medical College of Wisconsin; and 8Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, Wisconsin

Submitted 27 January 2012; accepted in final form 6 February 2012

Tomita-Mitchell A, Mahnke DK, Struble CA, Tuffnell ME, Stamm KD, Hidestrand M, Harris SE, Goetsch MA, Simpson PM, Bick DP, Broeckel AN, Pelech AN, Tweddell JS, Mitchell ME. Human gene copy number spectra analysis in congenital heart malformations. Physiol Genomics 44: 518–541, 2012. First published February 7, 2012; doi:10.1152/physiolgenomics.00013.2012.—The clinical significance of copy number variants (CNVs) in congenital heart disease (CHD) continues to be a challenge. Although CNVs including genes can confer disease risk, relationships between gene dosage and phenotype are still being defined. Our goal was to perform a quantitative analysis of CNVs involving 100 well-defined CHD risk genes identified through previously published human association studies in subjects with anatomically defined cardiac malformations. A novel analytical approach permitting CNV gene frequency “spectra” to be computed over prespecified regions to determine phenotype-gene dosage relationships was employed. CNVs in subjects with CHD (n = 945), subphenotyped into 40 groups and verified in accordance with the European Paediatric Cardiac Code, were compared with two control groups, a disease-free cohort (n = 2,026) and a population with coronary artery disease (n = 880). Gains (≥200 kb) and losses (≥100 kb) were determined over 100 CHD risk genes and compared using a Barnard exact test. Six subphenotypes showed significant enrichment (P ≤ 0.05), including aortic stenosis (valvar), atrioventricular canal (partial), atrioventricular septal defect with tetralogy of Fallot, subaortic stenosis, tetralogy of Fallot, and truncus arteriosus. Furthermore, CNV gene frequency spectra were enriched (P ≤ 0.05) for losses at: FKBP6, ELN, GTF2IRD1, GATA4, CRKL, TBX1, ATRX, GCPS, BCOR, ZIC3, FLNA, and MID1; and gains at: PRKAB2, FMO5, CHD11, BCL9, ACP6, GJA5, HRAS, GATA6 and RUNX1. Of CHD subjects, 14% had causal chromosomal abnormalities, and 4.3% had likely causal (significantly enriched), large, rare CNVs. CNV frequency spectra combined with precision phenotyping may lead to increased molecular understanding of etiologic pathways.

congenital heart disease; copy number variation; genetics

STRUCTURAL CONGENITAL HEART DISEASE (CHD) is the most common form of congenital malformations, affecting 0.8% of live births (21). Other than infection, more children die from CHD in infancy than from all other forms of disease (25). In addition, it is estimated that at least 10% of early miscarriages are a consequence of severe cardiac malformations (10). The causes of congenital cardiac malformations are largely unknown. It is estimated that 18% are due to chromosomal causes or genetic structural abnormalities including trisomies (Trisomy 21, 13, and 18) as well as deletion syndromes; all of these are associated with significant disease risk for CHD (36). A small percentage of congenital cardiac malformations are disorders in which underlying single genes have been discovered such as TBX5 in Holt-Oram syndrome; JAG1 in Alagille syndrome; and PTPN11, SOS1, and KRAS in Noonan syndrome (36). Known environmental risk factors during pregnancy, such as maternal diabetes or prenatal exposure to drugs, viruses, and reduced folate intake account for a small percentage of CHD cases (16, 24). Although our understanding of molecular pathways in cardiac development has grown tremendously in the past few years, the etiology of human and clinically relevant CHD in the majority (~75%) of cases cannot yet be identified or explained (14, 16).

The widespread use of microarray-based genomic technologies over the past 5–6 yr have implicated copy number variants (CNVs) in numerous disorders such as neuropsychiatric diseases (49), craniofacial phenotypes, cancer, and congenital anomalies including CHD (7, 18, 35, 36). Relative to sequence variations such as single base-pair mutations or single nucleotide polymorphisms (SNPs), rare and large CNVs are hypothesized to confer higher disease risk as entire genes are deleted or duplicated (12, 31). However, poor reproducibility between microarray platforms and the lack of standardized analytical tools highlight the importance of careful filtering in CNV detection studies (37). Nondisease-related copy number polymorphisms (CNPs and/or common CNVs ≥1%) are abundant, as evidenced by the growing Database of Genomic Variants (DGV) (22, 57). Similar to the challenges in the sequence analysis of unique genetic variants, the discovery of rare etiologic CNVs remains a challenge, both because it is more difficult to detect a rare event over another event seen many times and because of the intrinsic low prior probability of there being such a variant at any particular location in the genome in any individual (28).

Recently, an algorithm to clinically interpret CNVs in patients with CHD was described (6). This approach is primarily based on gene content and overlap with known causal CHD syndromes, rather than on CNV inheritance and size (6). We employed a parallel approach in this study and utilized a strict criteria to define “likely causal” duplications or deletions, in well-established human CHD risk genes. We chose 100 CHD risk genes or regions that were supported by published observations in human studies as a means to identify potentially disease-relevant CNVs. A majority of these known CHD risk genes were previously described or could be identified through the CHD WIKI portal (1, 36). In
addition, genes associated with recognized causal chromosomal abnormalities in CHD were included, as well as recently identified candidate genes from association studies (see Table 1) (1, 42).

CHD consists of heterogenous anatomy with distinct phenotypic subtypes. The European Paediatric Cardiac Coding (EPCC) System (17) has been cross mapped with the Society of Thoracic Surgeons/ European Association of Cardiothoracic Surgery (STS/EACTS) coding system through the International Society for Nomenclature of Paediatric and Congenital Heart Disease in the creation of the International Pediatric and Congenital Cardiac Code (IPCCC). We characterized cardiac malformations by subphenotyping according to both the EPCC and the STS/EACTS coding systems. We compared 945 CHD cases with a publicly available cohort of 2,026 disease-free primarily pediatric individuals (40). Cases and controls were genotyped on different platforms; therefore, a second cohort of 880 control subjects genotyped on the same platform and within the same facility as the CHD cohort was included in the analysis.

This study represents a quantitative analysis of CNVs in a large population of subjects with precisely phenotyped cardiac malformations involving 100 candidate CHD risk genes. We hypothesized that large rare CNVs that were statistically enriched against two control cohorts would be causal. A strict algorithm was employed to determine if subphenotypes were enriched in gains and losses within 100 recognized CHD risk genes selected based on gene content compared with two control cohorts. Finally, a novel analytical approach, permitting CNV gene frequency spectra to be computed as a proportion of each cohort containing a gain or a loss over the above prespecified regions, was employed to determine phenotype–gene dosage relationships.

**METHODS**

**CHD Case Ascertainment and Confirmation**

This study was reviewed and approved in accordance to institutionally approved research [Institutional Review Board (IRB)] protocols by the Children’s Hospital of Wisconsin (CHW, Milwaukee, WI). Subjects were consented through the Congenital Heart Disease Tissue Bank (CHDTB) and the Wisconsin Pediatric Cardiac Registry (WPCR), IRB-approved research databases housed at CHW (20, 47). These two biobanks provide DNA samples from cases and family members, detailed maternal environmental exposure data, family history of CHD, and cardiac tissue discs.

**Inclusion criteria.** Structural congenital cardiac abnormalities, as identified within the IPCCC, included abnormalities of the following: the atria and atrial septum; atrioventricular valves or atrioventricular septum; cardiac position and connections; chest wall; conduction system; coronary arteries, arterial duct, pericardium, or arteriovenous fistulae; great veins; ventricles or ventricular septum; and ventriculoarterial valves or great arteries.

**Exclusion criteria.** All acquired forms of pediatric heart disease in the absence of CHD, and frequent nonpathologic structural variants when no other CHD is present, included: patent foramen ovale, patent ductus arteriosus (PDA) under 30 days of age, PDA in premature infants (<35 wk gestation) and mitral valve prolapse (in the absence of at least mild valve insufficiency).

Note: The presence of a known or suspected chromosomal abnormality or known sequence variant in a CHD risk gene did not preclude participation in the study. In addition, the presence or absence of known environmental exposures did not preclude participation in the study.

Anatomic cardiac malformations were carefully characterized by phenotyping and subphenotyping according to both the EPCC 2011 and the STS/EACTS 2011 coding systems. All phenotypes were initially reviewed by a coding specialist, a surgeon, and a cardiologist. All discrepancies were reconciled by review of source documents including operative notes, echocardiograms, and review of operative surgeon. Anatomic phenotypes and subphenotypes were reported using EPCC 2011 terms, and final confirmatory review of all cases was performed by a single pediatric cardiothoracic surgeon (17). In addition, information regarding additional diagnosis, accompanying conditions, demographics, and a limited number of genetic risk factors was obtained through the Herma Heart Center (HHC) cardiac database at CHW.

**Children’s Hospital of Philadelphia Control Cohort**

DNA samples analyzed in this study were obtained from the whole blood of healthy subjects routinely seen at primary care and well-child clinic practices within the Children’s Hospital of Philadelphia (CHOP) Health Care Network. Data using hg18/March 2006/build 36.1 genomic coordinates were downloaded from http://cnv.chop.edu/ (40). High-resolution mapping of copy number variations in 2,026 healthy individuals was performed using the Illumina HumanHap 550 BeadChip (Illumina, San Diego, CA) (40).

**Milwaukee Family Heart Study Control Cohort**

Control subjects were drawn from the Milwaukee Family Heart Study (MFHS) in accordance with Medical College of Wisconsin IRB protocols (MCW, Milwaukee, WI). Subjects were ascertained as a hospital-based cohort, referred to the catheterization laboratory for diagnostic coronary angiography. Inclusion criteria were the ability to consent and age >21 yr. The following were considered exclusion criteria: end-stage renal disease, current treatment for a malignancy, and a diagnosis of coronary artery disease or a myocardial infarction at age >69 yr. In addition, we excluded all participants with acute coronary syndrome and significant valvular disease. Individuals with a diagnosis of other cardiac structural abnormalities were excluded based on either the result of echocardiography prior to or as determined during the invasive cardiac procedure.

**Genomic DNA Extraction**

Genomic DNA for CHD and MFHS cohorts was obtained from peripheral blood using standard protocols for DNA isolation from Roche Diagnostics, Promega Biotech (Wizarden), and Qiagen (Gentra Puregene). Purified genomic DNA was resuspended in 1.0 mM Tris HCl pH 8.0 and 0.1 mM EDTA. DNA quality was tested by optical density 260/280 ratios, quantified by UV spectrophotometry using a Nanodrop 2000 (Thermo Scientific, Wilmington, DE). DNA stocks were stored at −20°C. Dilutions for microarray analysis were stored at 100 ng/μl at −20°C.

**CHD Risk Gene Prioritization and Selection**

Genes or regions with previously associated disease/syndrome variants as identified through the CHD WIKI website (searched 01/04/2011 and updated 07/28/2011) and/or supported by previously published observations in human studies were selected (1, 34, 36, 42, 48). These known CHD risk genes are outlined in Table 1.

Briefly, CHD WIKI offers an updated overview of genes implicated in human CHD, obtained by an OMIM search, and complemented with a study of the PubMed literature concerning mutation analysis of candidate genes for congenital heart defects (1). The level of support was defined by inheritance of the mutation (de novo or inherited and segregated with a phenotype) and the association of a...
<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Name</th>
<th>Cytoband</th>
<th>Gene Start</th>
<th>Gene Size</th>
<th>ABI CN Assay #</th>
<th>CHD WIKI</th>
<th>OMIM ID</th>
<th>PubMed ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP6</td>
<td>ACID PHOSPHATASE 6, LYSOPHOSPHATIDE</td>
<td>1q21.1</td>
<td>14558791</td>
<td>23467</td>
<td>Hs00320736_cn</td>
<td>61471</td>
<td>15117819, 19597493</td>
<td></td>
</tr>
<tr>
<td>ACTC1</td>
<td>ACTIN, ALPHA, CARDIAC MUSCLE</td>
<td>15q14</td>
<td>32867588</td>
<td>7631</td>
<td>NS</td>
<td>102540</td>
<td>17611253, 17947298</td>
<td></td>
</tr>
<tr>
<td>ACRV2B</td>
<td>ACTIVIN A RECEPTOR, TYPE IB</td>
<td>3p22.2</td>
<td>38470793</td>
<td>38844</td>
<td>NS</td>
<td>602730</td>
<td>20193066</td>
<td></td>
</tr>
<tr>
<td>ALDH1A2</td>
<td>ALDEHYDE DEHYDROGENASE 1 FAMILY, MEMBER A2</td>
<td>15q22.1</td>
<td>56032918</td>
<td>112280</td>
<td>NS</td>
<td>603687</td>
<td>19886994</td>
<td></td>
</tr>
<tr>
<td>ANKRDI</td>
<td>ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 1</td>
<td>1q23.3</td>
<td>92661836</td>
<td>9176</td>
<td>NS</td>
<td>609599</td>
<td>18273862, 20193066</td>
<td></td>
</tr>
<tr>
<td>ASXL2</td>
<td>ADDITIONAL SEX COMBS-LIKE 2</td>
<td>2p23.3</td>
<td>25815756</td>
<td>139060</td>
<td>612991</td>
<td>19597493</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATRX</td>
<td>ATR-X GENE</td>
<td>Xq21.1</td>
<td>76647011</td>
<td>28136</td>
<td>S</td>
<td>300032</td>
<td>20193066</td>
<td></td>
</tr>
<tr>
<td>BCL9</td>
<td>B-CELL CLL/LYMPHOMA 9</td>
<td>1q21.1</td>
<td>145479805</td>
<td>84834</td>
<td>Hs01608359_cn</td>
<td>602597</td>
<td>15117819, 19597493</td>
<td></td>
</tr>
<tr>
<td>BCOR</td>
<td>BCL6 COREPRESSOR</td>
<td>Xp11.4</td>
<td>39795442</td>
<td>46221</td>
<td>Hs02764783_cn</td>
<td>300485</td>
<td>15770227</td>
<td></td>
</tr>
<tr>
<td>BRAF</td>
<td>V-RAF MURINE VIRAL ONCOGENE HOMOLOG B1</td>
<td>7q34</td>
<td>140080281</td>
<td>190752</td>
<td>S</td>
<td>164757</td>
<td>1674404, 19206169</td>
<td></td>
</tr>
<tr>
<td>CBL</td>
<td>CAS-BR-M MURINE ECOTROPIC RETROVIRAL TRANSFORMING SEQUENCE HOMOLOG</td>
<td>11q23.3</td>
<td>118582199</td>
<td>101870</td>
<td>165360</td>
<td>15266616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFC1</td>
<td>CRYPTIC PROTEIN CHROMOMODomain HELICASE DNA-BINDING PROTEIN 1-LIKE</td>
<td>2q21.1</td>
<td>131066804</td>
<td>6748</td>
<td>NS, S</td>
<td>605194</td>
<td>11062482</td>
<td></td>
</tr>
<tr>
<td>CHD1L</td>
<td>CHROMOMODomain HELICASE DNA-BINDING PROTEIN 7</td>
<td>8q12.2</td>
<td>61753892</td>
<td>188129</td>
<td>Hs01604098_cn</td>
<td>608892</td>
<td>15300250</td>
<td></td>
</tr>
<tr>
<td>CHD7</td>
<td>CHROMOMODomain HELICASE DNA-BINDING PROTEIN 7</td>
<td>8q12.2</td>
<td>61753892</td>
<td>188129</td>
<td>Hs01604098_cn</td>
<td>608892</td>
<td>15300250</td>
<td></td>
</tr>
<tr>
<td>CITED2</td>
<td>CBP/p300-INTERACTING TRANSSACTIVATOR, WITH GLU/ASP-RICH C-TERMINAL DOMAIN</td>
<td>6q24.1</td>
<td>139735091</td>
<td>2387</td>
<td>Hs01362863_cn</td>
<td>602937</td>
<td>16287139</td>
<td></td>
</tr>
<tr>
<td>COL2A1</td>
<td>COLLAGEN, TYPE II, ALPHA-1</td>
<td>12q13.11</td>
<td>46653014</td>
<td>31538</td>
<td>Hs00560273_cn</td>
<td>120140</td>
<td>20193066</td>
<td></td>
</tr>
<tr>
<td>CRELD1</td>
<td>CYSTEINE-RICH PROTEIN WITH EGF-LIKE DOMAINS 1</td>
<td>3p25.3</td>
<td>9950505</td>
<td>11585</td>
<td>NS</td>
<td>607170</td>
<td>12632326</td>
<td></td>
</tr>
<tr>
<td>CRKL</td>
<td>V-CRK AVIAN SARCOMA VIRUS CT10 ONCOGENE HOMOLOG-LIKE</td>
<td>22q11.21</td>
<td>19601713</td>
<td>36177</td>
<td>Hs01301005_cn</td>
<td>602007</td>
<td>20494672</td>
<td></td>
</tr>
<tr>
<td>CSDE1</td>
<td>COLD-SHOCK DOMAIN-CONTAINING E1, RNA-BINDING EUCHROMATIC HISTONE METHYLTRANSFERASE 1</td>
<td>1p13.2</td>
<td>115061059</td>
<td>41135</td>
<td>S</td>
<td>191510</td>
<td>20193066</td>
<td></td>
</tr>
<tr>
<td>EHM1</td>
<td>EUCHROMATIC HISTONE METHYLTRANSFERASE 1</td>
<td>9q34.3</td>
<td>139725237</td>
<td>125162</td>
<td>Hs00150023_cn</td>
<td>607001</td>
<td>16826528, 20193066</td>
<td></td>
</tr>
<tr>
<td>ELN</td>
<td>ELASTIN ELLIS-VAN CREVELD SYNDROME</td>
<td>7q11.23</td>
<td>73080362</td>
<td>41810</td>
<td>Hs03073113_cn</td>
<td>130160</td>
<td>12952863</td>
<td></td>
</tr>
<tr>
<td>EVC</td>
<td>EVC2 GENE</td>
<td>4p16.1</td>
<td>5763824</td>
<td>103108</td>
<td>S</td>
<td>225500</td>
<td>12571802</td>
<td></td>
</tr>
<tr>
<td>EVC2</td>
<td>EVC2 GENE</td>
<td>4p16.1</td>
<td>5763824</td>
<td>103108</td>
<td>S</td>
<td>225500</td>
<td>12571802</td>
<td></td>
</tr>
<tr>
<td>FBW1</td>
<td>FIBRILLIN 1</td>
<td>3p25.3</td>
<td>9950505</td>
<td>11585</td>
<td>NS</td>
<td>607170</td>
<td>12632326</td>
<td></td>
</tr>
<tr>
<td>FKB1</td>
<td>FIBRILLIN 1</td>
<td>15q21.1</td>
<td>46487796</td>
<td>237414</td>
<td>S</td>
<td>134797</td>
<td>10441597, 18412115</td>
<td></td>
</tr>
<tr>
<td>FLNA</td>
<td>FIBRILLIN 1</td>
<td>7q11.23</td>
<td>72380235</td>
<td>30342</td>
<td>S</td>
<td>604839</td>
<td>12952863</td>
<td></td>
</tr>
<tr>
<td>FMO5</td>
<td>FLAVIN-CONTAINING MONOOXYGENASE 5</td>
<td>Xq28</td>
<td>153230093</td>
<td>26107</td>
<td>S</td>
<td>300017</td>
<td>17190868</td>
<td></td>
</tr>
<tr>
<td>FOXC1</td>
<td>FORKHEAD BOX C1</td>
<td>6p25.3</td>
<td>1555679</td>
<td>3449</td>
<td>S</td>
<td>601900</td>
<td>15654696</td>
<td></td>
</tr>
<tr>
<td>FOXH1</td>
<td>FORKHEAD BOX H1</td>
<td>8q24.3</td>
<td>145670316</td>
<td>2210</td>
<td>NS</td>
<td>603621</td>
<td>18532893</td>
<td></td>
</tr>
<tr>
<td>FOXL2</td>
<td>FORKHEAD TRANSCRIPTION FACTOR FOXL2</td>
<td>3q22.3</td>
<td>140145755</td>
<td>2736</td>
<td>Hs01045878_cn</td>
<td>605597</td>
<td>18642388, 20193066</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 1.—Continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Name</th>
<th>Cytoband</th>
<th>Gene Start</th>
<th>Gene Size</th>
<th>ABI CN Assay #</th>
<th>CHD WIKI</th>
<th>OMIM ID</th>
<th>PubMed ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXL2</td>
<td>FORKHEAD TRANSCRIPTION FACTOR FOXL2</td>
<td>3q22.3</td>
<td>140145755</td>
<td>2736</td>
<td>Hs01045878_cn</td>
<td>S</td>
<td>605597</td>
<td>18642388, 20193066</td>
</tr>
<tr>
<td>GATA4</td>
<td>GATA-BINDING PROTEIN 4</td>
<td>8p23.1</td>
<td>11599125</td>
<td>55793</td>
<td>Hs01321405_cn</td>
<td>S</td>
<td>600576</td>
<td>16025100</td>
</tr>
<tr>
<td>GATA6</td>
<td>GATA-BINDING PROTEIN 6</td>
<td>18p11.2</td>
<td>18003413</td>
<td>32812</td>
<td>Hs02615249_cn</td>
<td>NS</td>
<td>601656</td>
<td>19666519</td>
</tr>
<tr>
<td>GDF1</td>
<td>GROWTH DIFFERENTIATION FACTOR 1</td>
<td>19p13.11</td>
<td>18840360</td>
<td>27593</td>
<td>Hs07489748_cn</td>
<td>NS</td>
<td>602880</td>
<td>17924340</td>
</tr>
<tr>
<td>GJA1</td>
<td>GAP JUNCTION PROTEIN, ALPHA-1</td>
<td>6q22.31</td>
<td>121798443</td>
<td>14129</td>
<td></td>
<td>S</td>
<td>121014</td>
<td>11470490</td>
</tr>
<tr>
<td>GJA5</td>
<td>GAP JUNCTION PROTEIN, ALPHA-5</td>
<td>7q21.1</td>
<td>145694955</td>
<td>17153</td>
<td>Hs00957111_cn</td>
<td>S</td>
<td>121013</td>
<td>15117819</td>
</tr>
<tr>
<td>GPC3</td>
<td>GLYPICAN 3</td>
<td>Xq26.2</td>
<td>132497441</td>
<td>449891</td>
<td>Hs00702786_cn</td>
<td>S</td>
<td>300037</td>
<td>10232747, 20193066</td>
</tr>
<tr>
<td>GTF2IRD1</td>
<td>GTF2I REPEAT DOMAIN-CONTAINING PROTEIN 1</td>
<td>7q11.23</td>
<td>73506055</td>
<td>148793</td>
<td></td>
<td>NS</td>
<td>604318</td>
<td>12952863</td>
</tr>
<tr>
<td>HAND1</td>
<td>HEART- AND NEURAL CREST DERIVATIVES-EXPRESSION 1</td>
<td>5q33.2</td>
<td>153834724</td>
<td>3293</td>
<td></td>
<td>S</td>
<td>602406</td>
<td>10189962</td>
</tr>
<tr>
<td>HEY2</td>
<td>HAIRY/ENHANCER OF SPLIT-RELATED</td>
<td>6q22.31</td>
<td>126112424</td>
<td>11684</td>
<td></td>
<td>NS</td>
<td>604674</td>
<td>20193066</td>
</tr>
<tr>
<td>HOXA1</td>
<td>HOMEobox A1</td>
<td>7p15.2</td>
<td>27099138</td>
<td>3012</td>
<td>Hs00428080_cn</td>
<td>S</td>
<td>142955</td>
<td>16155570</td>
</tr>
<tr>
<td>HRAS</td>
<td>V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG</td>
<td>11p15.5</td>
<td>522241</td>
<td>3309</td>
<td>Hs01379757_cn</td>
<td>S</td>
<td>190020</td>
<td>17054105</td>
</tr>
<tr>
<td>ISL1</td>
<td>ISL LIM HOMEBOX 1</td>
<td>5q11.2</td>
<td>50714714</td>
<td>11606</td>
<td></td>
<td>S</td>
<td>600366</td>
<td>20520780</td>
</tr>
<tr>
<td>JAG1</td>
<td>JAGGED 1</td>
<td>20p12.2</td>
<td>10566331</td>
<td>36363</td>
<td></td>
<td>NS,S</td>
<td>601920</td>
<td>11152664</td>
</tr>
<tr>
<td>KIF3C</td>
<td>KINESIN FAMILY MEMBER 3C</td>
<td>2p23.3</td>
<td>26002958</td>
<td>55989</td>
<td></td>
<td>S</td>
<td>602845</td>
<td>19597493</td>
</tr>
<tr>
<td>KRAS</td>
<td>V-KI-RAS2 KIRSTEN RAT SARCOMA VIRAL ONCOGENE HOMOLOG</td>
<td>12p12.1</td>
<td>25249446</td>
<td>45765</td>
<td></td>
<td>S</td>
<td>190070</td>
<td>16474405, 16474404</td>
</tr>
<tr>
<td>LBR</td>
<td>LAMIN B RECEPTOR</td>
<td>1q42.12</td>
<td>223655826</td>
<td>27316</td>
<td></td>
<td>S</td>
<td>60024</td>
<td>20193066</td>
</tr>
<tr>
<td>LEFTY1</td>
<td>LEFT-RIGHT DETERMINATION FACTOR 1</td>
<td>1q42.12</td>
<td>224140604</td>
<td>2855</td>
<td></td>
<td>S</td>
<td>603037</td>
<td>10053005</td>
</tr>
<tr>
<td>LEFTY2</td>
<td>LEFT-RIGHT DETERMINATION FACTOR 2</td>
<td>1q42.12</td>
<td>224190295</td>
<td>4618</td>
<td></td>
<td>S</td>
<td>601877</td>
<td>10053005</td>
</tr>
<tr>
<td>MAP2K1</td>
<td>MITOGEN-ACTIVATED PROTEIN KINASE KINASE 1</td>
<td>15q22.31</td>
<td>64466264</td>
<td>104672</td>
<td></td>
<td>S</td>
<td>176872</td>
<td>18042262</td>
</tr>
<tr>
<td>MAP2K2</td>
<td>MITOGEN-ACTIVATED PROTEIN KINASE KINASE 2</td>
<td>19p13.3</td>
<td>4041319</td>
<td>33807</td>
<td></td>
<td>S</td>
<td>601263</td>
<td>18042262</td>
</tr>
<tr>
<td>MAP3K7IP2</td>
<td>MITOGEN-ACTIVATED PROTEIN KINASE KINASE 7</td>
<td>6q25.1</td>
<td>149680755</td>
<td>93687</td>
<td></td>
<td>S</td>
<td>602614</td>
<td>20493459</td>
</tr>
<tr>
<td>MAPK1</td>
<td>MITOGEN-ACTIVATED PROTEIN KINASE 1</td>
<td>q11.21−22q11.</td>
<td>20443946</td>
<td>108024</td>
<td>Hs02937892_cn</td>
<td>S</td>
<td>176948</td>
<td>21127295**</td>
</tr>
<tr>
<td>MED13L</td>
<td>MEDIATOR COMPLEX SUBUNIT 13-LIKE</td>
<td>12q24.21</td>
<td>114880763</td>
<td>318763</td>
<td></td>
<td>NS</td>
<td>608771</td>
<td>14638541</td>
</tr>
<tr>
<td>MGP</td>
<td>MATRIX GAMMA-CARBOXYGLUTAMIC ACID</td>
<td>12p12.3</td>
<td>14926093</td>
<td>4002</td>
<td></td>
<td>S</td>
<td>154870</td>
<td>9916809, 20193066</td>
</tr>
<tr>
<td>MID1</td>
<td>MIDLINE 1</td>
<td>Xp22.2</td>
<td>10373595</td>
<td>388135</td>
<td>Hs02158662_cn</td>
<td>S</td>
<td>300552</td>
<td>12833403, 20193066</td>
</tr>
<tr>
<td>MLL2</td>
<td>MYELOID/LYMPHOID OR MIXED LINEAGE LEUKEMIA 2</td>
<td>12q13.12</td>
<td>47699024</td>
<td>36350</td>
<td>Hs0784563_cn</td>
<td>S</td>
<td>602213</td>
<td>20711175</td>
</tr>
<tr>
<td>MYH11</td>
<td>MYOSIN, HEAVY CHAIN 11, SMOOTH MUSCLE</td>
<td>16p13.11</td>
<td>15704492</td>
<td>153896</td>
<td>Hs00358138_cn</td>
<td>S</td>
<td>160745</td>
<td>16444274</td>
</tr>
<tr>
<td>MYH6</td>
<td>MYOSIN, HEAVY CHAIN 6, CARDIAC MUSCLE, ALPHA</td>
<td>14q11.2</td>
<td>22921038</td>
<td>26284</td>
<td></td>
<td>NS</td>
<td>160710</td>
<td>15735645</td>
</tr>
<tr>
<td>MYH7</td>
<td>MYOSIN, HEAVY CHAIN 7, CARDIAC MUSCLE, BETA</td>
<td>14q11.2</td>
<td>22951786</td>
<td>22924</td>
<td></td>
<td>NS</td>
<td>160760</td>
<td>21604106, 18159245</td>
</tr>
<tr>
<td>NF1</td>
<td>NEUROFIBROMATOSIS, TYPE 1</td>
<td>17q11.2</td>
<td>26446120</td>
<td>282701</td>
<td></td>
<td>S</td>
<td>162200</td>
<td>11078559, 20193066</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Name</th>
<th>Cytoband</th>
<th>Gene Start</th>
<th>Gene Size</th>
<th>ABI CN Assay #</th>
<th>CHD WIKI</th>
<th>OMIM ID</th>
<th>PubMed ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKK2-5</td>
<td>NK2 HOMEOBOX 5</td>
<td>5q35.2</td>
<td>172591743</td>
<td>3125</td>
<td>NS</td>
<td>600584</td>
<td>9651244</td>
<td></td>
</tr>
<tr>
<td>NKK2-6</td>
<td>NK2, DROSOPHILA, HOMOLOG OF 6</td>
<td>8p21.1</td>
<td>23615909</td>
<td>3957</td>
<td>NS</td>
<td>611770</td>
<td>15649947</td>
<td></td>
</tr>
<tr>
<td>NODAL</td>
<td>NODAL, MOUSE, HOMOLOG OF</td>
<td>10q22.1</td>
<td>71862076</td>
<td>9353</td>
<td>NS</td>
<td>601265</td>
<td>19064609</td>
<td></td>
</tr>
<tr>
<td>NOTCH1</td>
<td>NOTCH, DROSOPHILA, HOMOLOG OF 1</td>
<td>9q34.3</td>
<td>138508716</td>
<td>51343</td>
<td>Hs00041764_cn</td>
<td>NS</td>
<td>190198</td>
<td>16025100, 19597493</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>NOTCH, DROSOPHILA, HOMOLOG OF 2</td>
<td>1p12</td>
<td>120255698</td>
<td>158101</td>
<td>S</td>
<td>600275</td>
<td>16773578</td>
<td></td>
</tr>
<tr>
<td>NPHP3</td>
<td>NPHROCYSTIN 3</td>
<td>3q22.1</td>
<td>133882143</td>
<td>41823</td>
<td>S</td>
<td>608002</td>
<td>19177160</td>
<td></td>
</tr>
<tr>
<td>NRAS</td>
<td>NEUROBLASTOMA RAS VIRAL ONCOGENE HOMOLOG</td>
<td>12q24.13</td>
<td>111340918</td>
<td>91182</td>
<td>S</td>
<td>176876</td>
<td>15117819</td>
<td></td>
</tr>
<tr>
<td>NSD1</td>
<td>NUCLEAR RECEPTOR-BINDING Su-var, ENHANCER OF ZESTE, AND TRITHORAX</td>
<td>5q35.2–5q35.3</td>
<td>176492685</td>
<td>167135</td>
<td>Hs00053100_cn</td>
<td>S</td>
<td>606681</td>
<td>15742365, 20193066</td>
</tr>
<tr>
<td>PDGFRa</td>
<td>PLATELET-DERIVED GROWTH FACTOR RECEPTOR, ALPHA</td>
<td>4q12</td>
<td>54790020</td>
<td>69149</td>
<td>NS</td>
<td>173490</td>
<td>20071345</td>
<td></td>
</tr>
<tr>
<td>PITX2</td>
<td>PAIRED-LIKE HOMEODOMAIN TRANSCRIPTION FACTOR 2</td>
<td>4q25</td>
<td>111758028</td>
<td>19929</td>
<td>NS</td>
<td>601542</td>
<td>16274491</td>
<td></td>
</tr>
<tr>
<td>PPM1K</td>
<td>PROTEIN PHOSPHATASE, PP2C DOMAIN-CONTAINING, 1K</td>
<td>4q22.1</td>
<td>89400555</td>
<td>24357</td>
<td>NS</td>
<td>611065</td>
<td>19597493</td>
<td></td>
</tr>
<tr>
<td>PRKAB2</td>
<td>PROTEIN KINASE, AMP-ACTIVATED, NONCATALYTIC, BETA-2</td>
<td>1q21.1</td>
<td>145093308</td>
<td>17445</td>
<td>Hs02605549_cn</td>
<td>NS</td>
<td>602741</td>
<td>15117819</td>
</tr>
<tr>
<td>PTPN11</td>
<td>PROTEIN-TYROSINE PHOSPHATASE, NONRECEPTOR-TYPE, 11</td>
<td>12q24.13</td>
<td>111340918</td>
<td>91182</td>
<td>S</td>
<td>176876</td>
<td>17515436</td>
<td></td>
</tr>
<tr>
<td>RAB10</td>
<td>RAS-ASSOCIATED PROTEIN RAB10</td>
<td>2p23.3</td>
<td>26110477</td>
<td>103305</td>
<td>NS</td>
<td>612672</td>
<td>19597493</td>
<td></td>
</tr>
<tr>
<td>RAF1</td>
<td>V-RAF-1 MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1</td>
<td>3p25.1</td>
<td>12600099</td>
<td>80601</td>
<td>Hs02645733_cn</td>
<td>164760</td>
<td>17603483, 19597493</td>
<td></td>
</tr>
<tr>
<td>RAI1</td>
<td>RETINOIC ACID-INDUCED GENE 1</td>
<td>17p11.2</td>
<td>17525511</td>
<td>129979</td>
<td>S</td>
<td>607642</td>
<td>16845274, 20193066</td>
<td></td>
</tr>
<tr>
<td>ROR2</td>
<td>RECEPTOR TYROSINE KINASE-LIKE ORPHAN RECEPTOR 2</td>
<td>9q22.31</td>
<td>93524704</td>
<td>227561</td>
<td>S</td>
<td>602337</td>
<td>20193066</td>
<td></td>
</tr>
<tr>
<td>RUNXI</td>
<td>RUNT-RELATED TRANSCRIPTION FACTOR 1</td>
<td>21q22.12</td>
<td>35081967</td>
<td>261498</td>
<td>NS</td>
<td>602275</td>
<td>19684605, 20193066</td>
<td></td>
</tr>
<tr>
<td>SALLA</td>
<td>SAL-LIKE 4</td>
<td>20q13.2</td>
<td>49833989</td>
<td>18466</td>
<td>Hs00139344_cn</td>
<td>S</td>
<td>607343</td>
<td>12843316</td>
</tr>
<tr>
<td>SEMA5A</td>
<td>SEMAPHORIN 5A</td>
<td>5p15.2</td>
<td>9088137</td>
<td>51096</td>
<td>Hs01709772_cn</td>
<td>S</td>
<td>609297</td>
<td>9464278</td>
</tr>
<tr>
<td>SH3PXD2B</td>
<td>SH3 AND PX DOMAINS-CONTAINING PROTEIN 2B</td>
<td>5q35.1</td>
<td>171693107</td>
<td>121025</td>
<td>S</td>
<td>613293</td>
<td>20137777</td>
<td></td>
</tr>
<tr>
<td>SHOC2</td>
<td>SUPPRESSOR OF CLEAR, C. ELEGANS, HOMOLOG OF</td>
<td>10q25.2</td>
<td>112713902</td>
<td>49511</td>
<td>S</td>
<td>607775</td>
<td>19684605, 20193066</td>
<td></td>
</tr>
<tr>
<td>SLC2A10</td>
<td>SOLUTE CARRIER FAMILY 2 (FACILITATED GLUCOSE TRANSPORTER), MEMEBER 10</td>
<td>20q13.12</td>
<td>44771685</td>
<td>26707</td>
<td>S</td>
<td>606145</td>
<td>16550171, 20193066</td>
<td></td>
</tr>
<tr>
<td>SOS1</td>
<td>SON OF SEVENLESS, DROSOPHILA, HOMOLOG 1</td>
<td>2p22.1</td>
<td>39062193</td>
<td>138915</td>
<td>S</td>
<td>182530</td>
<td>17143285</td>
<td></td>
</tr>
</tbody>
</table>

Continued...
variant in the investigated CHD population vs. a normal control population (1). A comprehensive list of 100 CHD risk genes was selected; the vast majority of these selected genes are known to be expressed in the human heart (3, 11, 43, 46, 50, 54). According to CHD WIKI, syndromic genes were defined as congenital heart defects that are associated with a second major malformation (i.e., renal defects, cleft palate, brain malformations), with developmental delay or mental handicap, and/or the presence of dysmorphism.

Genotyping

Genotyping for the CHD and MFHS control cohort was performed with the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA) as previously described (30, 47). All samples were run in the Advanced Genomics (AGEN) laboratory core at the Children’s Research Institute (CRI)/MCW (Milwaukee, WI). A reference genomic DNA control sample, ref 103, supplied by Affymetrix, was run with every batch of subjects (Santa Clara, CA).

CNV Analysis and Quality Control

The CHD subject cohort comprised 1,020 subjects consented through the CHDTB or WPCR. We evaluated the quality and suitability of the subject population for a genetic association study. The population was required to pass copy number analysis quality metrics as seen in Table 2.

CNV identification of study subjects required the processing of Affymetrix intensity (CEL) files using Genotyping Console version 3.0.2 (GTC) software as previously described (20, 47). CEL files of subjects with a median absolute pairwise difference \( >0.35 \) and a CNV segmentation count \( >250 \), indicative of poor DNA quality, were excluded from the study.

Table 1.—Continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Name</th>
<th>Cytoband</th>
<th>Gene Start</th>
<th>Gene Size</th>
<th>ABI CN Assay#</th>
<th>CHD WIKI</th>
<th>OMIM ID</th>
<th>PubMed ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOX7</td>
<td>SRY-BOX 7</td>
<td>8p23.1</td>
<td>10618687</td>
<td>6745</td>
<td>Hs00923277_cn</td>
<td>S</td>
<td>612202</td>
<td>19606479</td>
</tr>
<tr>
<td>STRA6</td>
<td>STIMULATED BY RETINOIC ACID 6, MOUSE, HOMOLOG OF</td>
<td>15q24.1</td>
<td>72258860</td>
<td>23885</td>
<td>Hs01994903_cn</td>
<td>S</td>
<td>610745</td>
<td>17273977</td>
</tr>
<tr>
<td>TBX1</td>
<td>T-BOX 1</td>
<td>22q11.21</td>
<td>18124225</td>
<td>26887</td>
<td>Hs01313390_cn</td>
<td>NS,S</td>
<td>602054</td>
<td>14585638</td>
</tr>
<tr>
<td>TBX20</td>
<td>T-BOX 20</td>
<td>7p14.3</td>
<td>35208566</td>
<td>51201</td>
<td>Hs04957392_cn</td>
<td>NS,S</td>
<td>606061</td>
<td>17668378, 19762328</td>
</tr>
<tr>
<td>TBX3</td>
<td>T-BOX 3</td>
<td>12q24.21</td>
<td>113592441</td>
<td>13911</td>
<td>S</td>
<td></td>
<td>601621</td>
<td>16892408</td>
</tr>
<tr>
<td>TBX5</td>
<td>T-BOX 5</td>
<td>12q24.21</td>
<td>113276117</td>
<td>54513</td>
<td>S</td>
<td></td>
<td>601620</td>
<td>11376442</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>TRANSMEMB</td>
<td>3p21.31</td>
<td>46594183</td>
<td>4773</td>
<td>NS</td>
<td></td>
<td>187395</td>
<td>18538293, 20193066</td>
</tr>
<tr>
<td>TERT</td>
<td>TELOMERASE REVERSE TRANSCRIPTASE</td>
<td>5p15.33</td>
<td>1306286</td>
<td>41876</td>
<td>Hs03078158_cn</td>
<td>S</td>
<td>187270</td>
<td></td>
</tr>
<tr>
<td>TFAP2B</td>
<td>TRANSCRIPTION FACTOR AP2-BETA</td>
<td>6p12.3</td>
<td>50894397</td>
<td>28888</td>
<td>Hs01355864_cn</td>
<td>NS,S</td>
<td>601601</td>
<td>10802654</td>
</tr>
<tr>
<td>TGFB2</td>
<td>TRANSFORMING GROWTH FACTOR-BETA RECEPTOR, TYPE II</td>
<td>3p24.1</td>
<td>30622997</td>
<td>87640</td>
<td>190182</td>
<td>15235604, 15731757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMEM40</td>
<td>TRANSMEMB</td>
<td>3p25.1</td>
<td>12750391</td>
<td>25417</td>
<td>Hs01878707_cn</td>
<td>S</td>
<td>19597493</td>
<td></td>
</tr>
<tr>
<td>VEGFA</td>
<td>VASCULAR ENDOTHELIAL GROWTH FACTOR A</td>
<td>6p21.1</td>
<td>43845930</td>
<td>16271</td>
<td>NS</td>
<td></td>
<td>192240</td>
<td>20420808</td>
</tr>
<tr>
<td>WHSC1</td>
<td>WHS CANDIDATE 1 GENE</td>
<td>4p16.3</td>
<td>1842920</td>
<td>110812</td>
<td>Hs02237093_cn</td>
<td>S</td>
<td>602952</td>
<td>9222965</td>
</tr>
<tr>
<td>ZEB2</td>
<td>ZINC FINGER E BOX-BINDING HOMEOBOX 2</td>
<td>2q22.3</td>
<td>144862052</td>
<td>132334</td>
<td>S</td>
<td></td>
<td>605802</td>
<td>11595972</td>
</tr>
<tr>
<td>ZFPM2</td>
<td>ZINC FINGER PROTEIN, MULTITYPE 2</td>
<td>8q23.1</td>
<td>106400322</td>
<td>485621</td>
<td>NS</td>
<td></td>
<td>603693</td>
<td>9927675, 10892744</td>
</tr>
<tr>
<td>ZIC3</td>
<td>ZINC FINGER PROTEIN OF CEREBELLUM 3</td>
<td>Xq26.3</td>
<td>136476011</td>
<td>5914</td>
<td>Hs02692150_cn</td>
<td>NS</td>
<td>300265</td>
<td>14681828, 10980576</td>
</tr>
</tbody>
</table>

NS, nonsyndromic; S, syndromic, NCBI Build 36.1/hg18. **Animal study.

Table 2. Quality control of CHD case and MFHS control cohorts and genotyping data

<table>
<thead>
<tr>
<th>Subjects, n</th>
<th>MFHS Control Cohort</th>
<th>CHD Case Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting subjects</td>
<td>950</td>
<td>1,020</td>
</tr>
<tr>
<td>Remaining subjects</td>
<td>880</td>
<td>958</td>
</tr>
</tbody>
</table>

QC Exclusions % Total

| MAPD QC | 3.05 |
| Segment QC | 4.32 |
| Consent QC | NA |
| Sex QC | NA |

Copy number analysis exclusions were as follows: median absolute pairwise difference (MAPD) quality control (QC) \( >0.35 \), number of copy number polymorphism (CNP) segments \( \geq 250 \), 1 subject with a status change to his/her consent, and sex tracking QC. Congenital heart disease (CHD) cases were reduced to a final \( n = 945 \) after inclusion and exclusion criteria were met.
Table 3. CHD case, CHOP, and MFHS control cohort demographics

<table>
<thead>
<tr>
<th></th>
<th>CHD Case Cohort</th>
<th>CHOP Control Cohort</th>
<th>MFHS Control Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>655</td>
<td>1,320</td>
<td>870</td>
</tr>
<tr>
<td>African American</td>
<td>92</td>
<td>694</td>
<td>5</td>
</tr>
<tr>
<td>Native American</td>
<td>14</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>26</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>945</td>
<td>2,026</td>
<td>880</td>
</tr>
<tr>
<td>Sex, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>44.02</td>
<td>36.59</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>55.87</td>
<td>63.41</td>
<td></td>
</tr>
<tr>
<td>Average age</td>
<td>4.03</td>
<td>67.00</td>
<td>65.66</td>
</tr>
</tbody>
</table>

CHOP, Children's Hospital of Philadelphia; MFHS, Milwaukee Family Heart Study.

A final number of 945 CHD subjects and 880 MFHS controls remained in the study after inclusion and exclusion criteria were met. As summarized in Table 3, the cases and controls were stratified according to age, sex, and race/ethnicity.

Copy number state of those subjects who passed quality control thresholds were determined with reference to the GenomeWideSNP 6.0 hapmap270 file and copy number calls were determined using the Affymetrix GTC segmentation algorithm. To reduce the presence of false positive CNVs, the segmentation algorithm parameters were set to identify only those regions larger than 25 kb comprising at least 25 contiguous markers. It has been shown that CNVs smaller than this are frequently false positive detection (40). In addition, all segments were monitored for degree of overlap with previously identified common CNVs, annotated by the DGV (22, 57).

Using a BED file format (chromosome, gene starting position, gene ending position, gene name), copy number information was drawn from custom gene regions (Table 1) extracted from the processed segment data.

A flowchart for copy number analysis is presented in Fig. 1. A multipurpose Access database (Microsoft, Redmond, WA) served as a central repository for the cohort demographic data as well as the entire experimental set of copy number variant data. Database tables were populated with copy number data from the GTC analysis, detailed demographic data, and the annotated 100 CHD risk gene list (Table 1). Demographic data for CHD cases and MFHS controls were obtained via clinical and consent verification methods. SQL query results included aggregate CNV counts by phenotype or region for both CHD and MFHS controls. Graphical representation of the query results was accomplished using Excel (Microsoft) and R (45). Supplemental Table S1 includes a complete summary of all CNV profiles over the 100 CHD risk gene list for each subject as well as phenotypic and demographic information.

Overall CNV burden. The total number of large CNVs throughout the genome was calculated by importing GTC segment files filtered by size (duplication ≥200 kb or deletion ≥100 kb) into an Access database. An external R program further filtered CNVs for all Build 36 annotated genes that did not occur as a CNP, defined as a normal variant (≥1%) in either CHOP or MFHS control cohorts.

Algorithm for likely causal CNV determination. A strict algorithm was employed to determine likely causal CNVs. Gains and losses were considered as potentially disease relevant if they fulfilled the following criteria: 1) size: duplication ≥200 kb or deletion ≥100 kb,

2) they did not occur as a CNP, defined as a normal variant (≥1%) in either CHOP or MFHS control cohort, and 3) CNV occurred over a gene region known to be associated with CHD (CHD 100 gene list). A final step was taken because the MFHS cohort was aged and significantly different from CHD cases. Sex chromosome degradation in peripheral blood appears to be an age-related phenomenon (19).

Studies have shown that a strong correlation exists between patient age and loss of the Y chromosome (52). Sex chromosome degradation is easily detected by the segment reports created by GTC because males have only one copy of Chr. X. To optimize the analysis of sex chromosomes, sex-matched references were employed; for X chromosome analysis, only females from all three cohorts were compared (55). Thus male MFHS controls were excluded from X chromosome results in all CNV analyses.

CNV frequency by phenotype. CNVs fulfilling criteria 1–3 were analyzed for enrichment by subphenotypes.

CNV frequency by gene region. CNV frequency “spectra” were computed as a proportion of each cohort containing a gain or a loss over the CHD associated gene list.

Complex CNV analysis. To determine if subjects carried multiple CNVs, large rare CNVs outside of and in addition to the defined set of 100 disease-related CHD genes were screened using criteria 1 and 2 (see Ref. 56).

Confirmatory Studies

CNVs that were identified in the CHD cases were confirmed by either karyotype, FISH analysis, or TaqMan CN real-time quantitative PCR assays (Applied Biosystems). CNVs for one case asterisked in Table 5 was difficult to confirm and is currently pending, due to inconclusive TaqMAN copy number results. A representative set of identified CNVs within the CHOP cohort were previously validated (40), whereas CNVs identified in the MFHS cohort as part of this study were not confirmed. As a means of secondary CNV confirmation of CHD cases, microarray analysis was performed by an independent lab on a number of the CHD study subjects (n = 34). TaqMan copy number reactions (Table 1) were run in triplicate on an ABI HT7900 instrument (Applied Biosystems) under the following cycling conditions:

1) size: duplication ≥200 kb or deletion ≥100 kb,
conditions: 50°C for 2 min, 95°C for 10 min, then 40 cycles of 95°C for 15 s followed by 60°C for 1 min. Typically ~20 ng of template genomic DNA was amplified in reaction volumes of 10 μl, as previously described (47). Copy number confirmations were assessed using a calibrator panel of six individuals with known copy number state over the gene of interest and analyzed using Copy Caller software version 1.0 (Applied Biosystems). If parents of subjects with confirmed CNVs were available, their DNA was analyzed to determine if CNVs were inherited or de novo, as noted in Table 5.

Statistical Analysis

Since the expected incidence is very small (typically <5%) tests based on a normality assumption would be incorrect, therefore a one-tailed Barnard exact test was used for all comparisons of proportions of CNVs (8). A P ≤ 0.05 without adjustment is used for significance. A custom R program was used to calculate the P value and checked using Cytel StatXact (Cytel, Cambridge, MA) (15). StatXact was also used to calculate power. With a sample of 810, and a CNV incidence of 4.3%, we would have at least 90% power to detect a significant difference from 0.0196 (the CNV incidence of the CHOP control CNV fraction) at an alpha = 0.05, power at least 80%. CHOP, Children’s Hospital of Philadelphia.

RESULTS

Phenotypes of CHD Study Subjects

Subjects diagnosed with congenital heart malformations (n = 945) and phenotyped in accordance with the EPCC terms were categorized into the 40 cardiac subphenotypes listed in Table 4 (17). The five largest phenotypes represented were as follows: hypoplastic left heart syndrome (HLHS) 14.8%, ventricular septal defect (VSD perimembranous) 7.7%, tetralogy of Fallot (TOF) 7.7%, coarctation of the aorta (CoA) 7.0%, and atrioventricular canal complete (AVC complete) 5.0%. The majority of subjects were represented by individual subphenotypes most of which contained <5.1% of the total CHD cohort.

Subjects With Recognized Causal Chromosomal Abnormalities

We ascribed 135 subjects to known CHD-related chromosomal abnormalities [T21 (n = 42), Turner (n = 8), William’s (n = 3), and XXX (n = 1)] (36, 44). The syndromes and their associated phenotypes were as follows: T21: aorto-pulmonary window with PDA n = 2; AVSD + TOF n = 5; ASD-SEC n = 4; AVC complete n = 80).
35; AVC intermediate n = 5; AVC partial n = 2; AVC unbalanced + AVSD with ventricular imbalance n = 1; other cardiac n = 1; pulmonary atresia (PA), IVS n = 1; subaortic stenosis n = 1; TOF n = 6; vascular ring + PA sling n = 1; VSD (inlet) n = 2; VSD (perimembranous) n = 13 and VSD (subarterial) n = 1, T18: TOF n = 1, 22qDS: DORV n = 1; IAA n = 4; mitral stenosis, subvalvar, parachute + mitral stenosis n = 1; other cardiac n = 1; PA, VSD n = 10; TOF n = 9; truncus arteriosus n = 12; vascular ring + PA sling n = 1 and VSD (perimembranous) n = 3, Turner: aortic stenosis (valvar) n = 1; CoA n = 4 and HLHS n = 2, mosaic Turner: CoA n = 1, William’s: supravalvar aortic stenosis and XXX: PA, IVS.

**CHD Case Reports**

Likely etiologic large, rare CNVs were identified in 35 CHD subjects. Table 5 summarizes the complete list of CHD subjects with CNVs over the known CHD risk gene regions (excluding the 135 subjects with known CHD-related chromosomal abnormalities). Three HLHS subjects (cases 16, 17, and 18) were studied for inheritance, a gain over FKBP6 was found to be a de novo event, a gain involving GATA4 and SOX7 was not present in one parent and the status of the other parent was unknown, and the MYH11 gain was inherited. Table 5 reports all of the known genes within each CNV segment, including our selected 100 CHD-associated genes.

**Statistical Analysis of CNVs**

Subphenotype analysis. The CHD cohort, even after excluding genes involved in the known CHD-related chromosomal abnormalities, was enriched in large, rare CNVs involving CHD risk genes, where 35 of 810 subjects carried such a CNV (P = 0.05 vs. both CHOP with 39 of 2,026 and MFHS with 14 of 880). Breaking this cohort into subgroups by specific phenotype often resulted in groups too small for statistical significance. Different subdivision schemes may achieve nominal significance. The entries in Table 6 where the frequency of CNV was significantly (P ≤ 0.05) different from the CHOP and MFHS cohorts are marked with a double asterisk. The CHD cohort, after excluding known causal chromosomal abnormalities, showed a frequency of CNV at 4.3%, and a power calculation is performed in Fig. 2 showing the difficulty in detecting a difference from the control’s 1.9%. For subgroups of 10–25 individuals, the power to detect a difference from 1.9% (CHOP) required a proportion of 30 and 17%, respectively. Phenotypes showing significant (P ≤ 0.05) enrichment of large CNV events were aortic stenosis (valvar), AV canal (partial), AVSD with TOF, subaortic stenosis, TOF, and truncus arteriosus. Although HLHS was the most common phenotype in the CHD case cohort, this phenotype did not demonstrate significant large rare CNV enrichment.

**CNV gene frequency analysis and gene enrichment.** In addition, CNV frequency “spectra” were computed as a proportion of each cohort containing a gain or a loss over 100 CHD genes of interest (Fig. 4). (Spectra for individual CHD subphenotypes with statistically higher CNV frequencies are represented in Fig. 5.) The frequency of genes with gain or loss was compared with both control cohorts and significantly enriched genes are listed in Table 7. In addition, Supplemental Table S1 includes a complete summary of all CNV profiles over the 100 CHD risk gene list for each CHD subject, and a heatmap (Supplemental Fig. S1) illustrates the clustering of various groups of multiple subjects who share contiguous blocks of deleted or duplicated genes.

Numerous genes were identified as significantly enriched (P ≤ 0.05 against both control cohorts), including losses, FKBP6, ELP, GTF2IRD1, GATA4, CRKL, TBX1, ATRX, GPC3, BCOR, ZIC3, FLNA and MIDD, and gains, PRKAB2, FMO5, CHD1L, BCL9, ACP6, GJA5, HRA5, GATA6, and RUNX1. These genes are identified in Table 7.

The authors recognize that syndromic forms of congenital heart disease are relatively well understood; therefore, genes in chromosomal abnormalities known to be causally related to CHD were intentionally kept on the 100 candidate CHD risk genes associated with congenital heart disease.

**Overall CNV Burden**

The total number of large CNVs (≥100 kb loss, ≥200 kb gain) throughout the genome were similar in both CHD and MFHS cohorts. When subjects with chromosomal abnormalities such as Trisomy 21 and 18, Turner, 22qDS, William’s, and XXX were excluded, a significant number of the CHD cohort, after excluding known causal chromosomal abnormalities, was enriched in large, rare CNVs involving CHD risk genes, where 35 of 810 subjects carried such a CNV (P = 0.05 vs. both CHOP with 39 of 2,026 and MFHS with 14 of 880). Breaking this cohort into subgroups by specific phenotype often resulted in groups too small for statistical significance. Different subdivision schemes may achieve nominal significance. The entries in Table 6 where the frequency of CNV was significantly (P ≤ 0.05) different from the CHOP and MFHS cohorts are marked with a double asterisk. The CHD cohort, after excluding known causal chromosomal abnormalities, showed a frequency of CNV at 4.3%, and a power calculation is performed in Fig. 2 showing the difficulty in detecting a difference from the control’s 1.9%. For subgroups of 10–25 individuals, the power to detect a difference from 1.9% (CHOP) required a proportion of 30 and 17%, respectively. Phenotypes showing significant (P ≤ 0.05) enrichment of large CNV events were aortic stenosis (valvar), AV canal (partial), AVSD with TOF, subaortic stenosis, TOF, and truncus arteriosus. Although HLHS was the most common phenotype in the CHD case cohort, this phenotype did not demonstrate significant large rare CNV enrichment.

**CNV gene frequency analysis and gene enrichment.** In addition, CNV frequency “spectra” were computed as a proportion of each cohort containing a gain or a loss over 100 CHD genes of interest (Fig. 4). (Spectra for individual CHD subphenotypes with statistically higher CNV frequencies are represented in Fig. 5.) The frequency of genes with gain or loss was compared with both control cohorts and significantly enriched genes are listed in Table 7. In addition, Supplemental Table S1 includes a complete summary of all CNV profiles over the 100 CHD risk gene list for each CHD subject, and a heatmap (Supplemental Fig. S1) illustrates the clustering of various groups of multiple subjects who share contiguous blocks of deleted or duplicated genes.

Numerous genes were identified as significantly enriched (P ≤ 0.05 against both control cohorts), including losses, FKBP6, ELP, GTF2IRD1, GATA4, CRKL, TBX1, ATRX, GPC3, BCOR, ZIC3, FLNA and MIDD, and gains, PRKAB2, FMO5, CHD1L, BCL9, ACP6, GJA5, HRA5, GATA6, and RUNX1. These genes are identified in Table 7.

The authors recognize that syndromic forms of congenital heart disease are relatively well understood; therefore, genes in chromosomal abnormalities known to be causally related to CHD were intentionally kept on the 100 candidate CHD risk genes associated with congenital heart disease.
<table>
<thead>
<tr>
<th>Subject</th>
<th>Subphenotype</th>
<th>100 CHD Gene Region</th>
<th>Exon(s)</th>
<th>LOSS_GAIN</th>
<th>Cytoband</th>
<th>CNV Start (Build 36, hg18)</th>
<th>CNV Size, kb</th>
<th>Markers, n</th>
<th>Inheritance</th>
<th>Gene Names on CNV Segment (100 CHD Genes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 AS (valvar)</td>
<td>ACP6 BCL9 CHD1L FMOS GJA5 PRKAB2</td>
<td>all Loss</td>
<td>1q21.1</td>
<td>144643813</td>
<td>1654</td>
<td>684</td>
<td>NBP1F1 FAM108A3 PRKAB2 FMOS CHD1L BCL9 ACPE6 GA5 GJA8 GPR89B NBP1F1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 AS (valvar)</td>
<td>CHD1L FMOS PRKAB2</td>
<td>all Gain</td>
<td>1q21.1</td>
<td>144943150</td>
<td>478</td>
<td>280</td>
<td>THOC3 FAM153B C5orf25 KIAA191 ARL10 HSPC1114 HGD2A CLTB FAK2 REF64 PCDH24 GPR11 SNCB EIE41B1 TSPAN17 UNC5A H3 UIMC1 ZNF466 GFRG4 NSD1 RAB24 PRELID1 MXD3 LMAN2 GRS3 SLIC3A1 PRN3 F12 GRK6 PRR7 DBN1 PDLIM7 DOK3 DD41 FLJ1404 TEM9 B4GALT7 DUSP22 IRF4 EXOC2 HUS1B FOXQ1 FOX2 FOX1 GMDS C6orf195 MYL4 MPV17L C16orf45 KIAA0430 NDE1 NSD1 RAB24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 AS (valvar)</td>
<td>FOXC1</td>
<td>all Loss</td>
<td>6p25.3–6p25.2</td>
<td>94649</td>
<td>2539</td>
<td>2130</td>
<td>GJA8 GPR89B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 ASD-SEC</td>
<td>MYH11</td>
<td>all Loss</td>
<td>16p13.11–16p12.3</td>
<td>15186307</td>
<td>2903</td>
<td>1521</td>
<td>BK</td>
<td>BRAH44</td>
<td>NEIL2</td>
<td></td>
</tr>
<tr>
<td>5 ASD-SV</td>
<td>GATA4</td>
<td>all Loss</td>
<td>8p23.1</td>
<td>11390744</td>
<td>304</td>
<td>213</td>
<td>SEPT5 GP1BB NBX1 TXNDR2 COMT ARVCF C22orf25 DGCR8 HTF9C RANBP1 ZDHHC8 RTN4R DGCR6L RIMBP3 ZNF74 SCARF2 KHL22 MED15 PS44A SERPIN1 SNAP29 CRKL AIFM3 LZTR1 THAP7 P2RX6 SLIC7A4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 AVC (partial)</td>
<td>GDF1</td>
<td>all Gain</td>
<td>19p13.11</td>
<td>18763592</td>
<td>378</td>
<td>163</td>
<td>SEPT5 GP1BB TXNDR2 COMT ARVCF C22orf25 DGCR8 HTF9C RANBP1 ZDHHC8 RTN4R DGCR6L RIMBP3 ZNF74 SCARF2 KHL22 MED15 PS44A SERPIN1 SNAP29 CRKL AIFM3 LZTR1 THAP7 P2RX6 SLIC7A4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 AVC unbalanced + AVSD with ventricular imbalance</td>
<td>MIDI</td>
<td>5’ UTR-i1 Gain</td>
<td>Xp22.2</td>
<td>10741630</td>
<td>509</td>
<td>265</td>
<td>MID1 HCCS ARHGAP6 AMELX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 AVSD with TOF</td>
<td>CRKL TBX1</td>
<td>all Gain</td>
<td>2q11.21</td>
<td>17953160</td>
<td>1838</td>
<td>1106</td>
<td>SEPT5 GP1BB TXNDR2 COMT ARVCF C22orf25 DGCR8 HTF9C RANBP1 ZDHHC8 RTN4R DGCR6L RIMBP3 ZNF74 SCARF2 KHL22 MED15 PS44A SERPIN1 SNAP29 CRKL AIFM3 LZTR1 THAP7 P2RX6 SLIC7A4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 CoA</td>
<td>ACP6 BCL9 CHD1L FMOS GJA5 PRKAB2</td>
<td>all Gain</td>
<td>1q21.1</td>
<td>144812585</td>
<td>1480</td>
<td>678</td>
<td>PRKAB2 FMOS CHD1L BCL9 ACPE6 GA5 GJA8 GPR89B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 CoA</td>
<td>NOTCH1</td>
<td>all Loss</td>
<td>9q34.3</td>
<td>138377108</td>
<td>229</td>
<td>105</td>
<td>NOTCH1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued...
Table 5.—Continued

<table>
<thead>
<tr>
<th>Subject</th>
<th>Subphenotype</th>
<th>100 CHD Gene Region</th>
<th>Exon(s)</th>
<th>LOSS_GAIN</th>
<th>Cytoband</th>
<th>CNV Start (Build 36, hg18)</th>
<th>CNV Size, kb</th>
<th>Markers, n</th>
<th>Inheritance</th>
<th>Gene Names on CNV Segment (100 CHD Genes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>DILV</td>
<td>HRAS</td>
<td>all</td>
<td>Gain</td>
<td>11p15.5</td>
<td>354390</td>
<td>256</td>
<td>62</td>
<td></td>
<td>B4GALT4 PKP3 SIGIRR TME16J PTDS2 RN1H HRAS LRPC56 C11orf35 RASS7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR SEMA5A</td>
</tr>
<tr>
<td>13</td>
<td>DORV</td>
<td>SEMA5A</td>
<td>i8-3’ UTR</td>
<td>Gain</td>
<td>5p15.31–5p15.2</td>
<td>7119715</td>
<td>2152</td>
<td>1769</td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR SEMA5A</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>EBSTEIN’S</td>
<td>FKBP6</td>
<td>5’ UTR-i8</td>
<td>Gain</td>
<td>7q11.23</td>
<td>72073034</td>
<td>330</td>
<td>28</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>15</td>
<td>HLHS</td>
<td>EHIIT1</td>
<td>all</td>
<td>Gain</td>
<td>9q34.3</td>
<td>139701521</td>
<td>264</td>
<td>142</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>16</td>
<td>HLHS</td>
<td>FKBP6</td>
<td>5’ UTR-i8</td>
<td>Gain</td>
<td>7q11.23</td>
<td>72052197</td>
<td>348</td>
<td>34</td>
<td>de novo</td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>17</td>
<td>HLHS</td>
<td>GATA4</td>
<td>all</td>
<td>Gain</td>
<td>8p23.1</td>
<td>11049252</td>
<td>1438</td>
<td>755</td>
<td>unknown</td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOX7</td>
<td>all</td>
<td>Gain</td>
<td>8p23.1</td>
<td>8055434</td>
<td>2992</td>
<td>2584</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>18</td>
<td>HLHS</td>
<td>MYH11</td>
<td>all</td>
<td>Gain</td>
<td>16p13.11</td>
<td>14846829</td>
<td>1414</td>
<td>640</td>
<td>inherited</td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>19</td>
<td>Other, Cardiac</td>
<td>CRKL TBX1</td>
<td>all</td>
<td>Gain</td>
<td>22q11.21</td>
<td>17161534</td>
<td>2634</td>
<td>1575</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>20</td>
<td>PA, VSD</td>
<td>ACP6 BCL9 CHD1L FMOS5 GJA5 PRKAB2</td>
<td>all</td>
<td>Gain</td>
<td>1q21.1</td>
<td>144812585</td>
<td>1480</td>
<td>678</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>21</td>
<td>Subaortic stenosis</td>
<td>CRKL</td>
<td>all</td>
<td>Gain</td>
<td>22q11.21</td>
<td>19093207</td>
<td>699</td>
<td>626</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>22</td>
<td>Subaortic stenosis</td>
<td>HRAS</td>
<td>all</td>
<td>Gain</td>
<td>11p15.5</td>
<td>339238</td>
<td>271</td>
<td>63</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
<tr>
<td>23</td>
<td>DORV</td>
<td>FOXL2 NPHP3</td>
<td>all</td>
<td>Gain</td>
<td>3q22.1–3q26.1</td>
<td>131972967</td>
<td>32134</td>
<td>19750</td>
<td></td>
<td>FOSJ1 NUP156 RASSF7 KIAA1542 IRF7 MUPCDH ADCY2 C5orf49 FASTKD3 MTRR</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Subject Subphenotype</th>
<th>100 CHD Gene Region</th>
<th>Exon(s)</th>
<th>LOSS_GAIN</th>
<th>Cytoband</th>
<th>CNV Start (Build 36, hg18)</th>
<th>CNV Size, kb</th>
<th>Markers, n</th>
<th>Inheritance</th>
<th>Gene Names on CNV Segment (100 CHD Genes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C3orf36 SLCO2A1 RYK AMOTL2 ANAPC13 CEP63 KY EPHB1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FAM PIK3CB FOXL2 C3orf72 LOC389515 MRPS22 COPB2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RBP2 BIP NMNAT3 CLSTN2 TRIM42 SLC25A36 SPSB4 ACPL2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ZBTB38 RASA2 RNF7 GRK7 ATPIB1 TDP2G K5XRN1 ATR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PLS1 TRPC1 POLOC2 PAQ9R SR140 CHST2 SLC9A9 C3orf58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PLD2 PLS4 PLSCR2 PLSCR1 PLSCR5 ZIC4 ZIC1 AGTR1 CPB1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CPA3 GYG1 HL3F HPS3 CP TM4SF18 TM4SF4 WWTR1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>COMMD2 RNF13 RNF13 PFN2 TSC2D2 SERPI EIF2A SELT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C3orf44 SIAH2 CLR1N1 CLR1N4 MEDI2L GPR171 P2RY14 GPR87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P2RY13 P2RY13 P2RY12 ACSF10 AADACL2 AADAC SUCNR1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MBNL1 TEMEM14E P2RY1 NTR2B LOCL15118 SGEF DHX36 GPR149</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MME PLCH1 C3orf33 SLC33A1 GMP5 KCNAB1 SSR3 TIPARP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LEKR1 CCNI VEPH1 PTX3 C3orf55 SHOX2 RSRC1 MLF1 GFM1 LXX</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RARRES1 MFSDD1 IQCJ SCHIP1 IL12A IFT80 SMC4 TRIM59 KPNAA3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ARL14 PPM11 BIGALNT1 NMD3 C3orf57 OTOL1 SI SLTRK3 BCH3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ZBBX SERPINC2 WDR49 PDCD10 SERPINC1 GOLM4 EVI1 EVI1 MOD1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ARPM1 MYNN LLRC34 LIRRQ4 LLRC31 SAMD7 SEC62 GPR160</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PHC3 PRKCI SKIL C3DHN11 SCI7A14 RPL2L1 EIF5A3 SLCS2A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TNIK PLD1 FNDC3B GSRN TNSF10 AADACL1 ECT2 SPATA16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NLGN1 NAAALD2 TBL1XR1 KC- NMB2 ZMAT3 PIK3CA KCNMBS3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ZNF695 MFN1 GBNI4 ACTL6A MRPL47 NDUF55 USP13 PEX5L</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TTC14 CDC39 FXR1 DNAJC19 SOX2 ATP11B DCUN1D1 MCC1C1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LAMP3 MCF2L3 B3GNT5 KLHL6 KLHL24 YEATS2 MA6P61 PARL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ABCC5 HTR3D HTR3C HTR3E</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Subject</th>
<th>Subphenotype</th>
<th>100 CHD Gene Region</th>
<th>Exon(s)</th>
<th>LOSS_GAIN</th>
<th>Cytoband</th>
<th>CNV Start (Build 36, hg18)</th>
<th>CNV Size, kb</th>
<th>Markers, n</th>
<th>Inheritance</th>
<th>Gene Names on CNV Segment (100 CHD Genes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>TOF</td>
<td>ACP6 BCL9 GJA5</td>
<td>all</td>
<td>Gain</td>
<td>1q21.1</td>
<td>145250193</td>
<td>1678</td>
<td>471</td>
<td></td>
<td>EIF2B5 DVL3 AP2M1 ABCF3 ALG3 ECE2 CAMK2N2 ECE2 PSM2D EEF4G1 FAM313A CLCN2 POLR2H THPO CHRD EPB3 MAGEF1 VPS8 C3orf70 EHHADH MAP3K13 TMEM41A LIP1 SENP2 IGF2BP2 C3orf65 SFRS10 ETV5 DGKG CRYGS TBCDD1 Dnajb11 AHS VETUB HRG KNG1 EIF4A2 RFC3 ADIPQ ST6GAL1 RPL39L RPT1 MASPI RPT4 SST RT2 BCL6 LPP TPRG1 TP63 LEPRE1 SENP2 IGF2BP2 C3orf65 SFRS10 ETV5 DGKG CRYGS TBCDD1 Dnajb11 AHS VETUB HRG KNG1 EIF4A2 RFC3 ADIPQ ST6GAL1 RPL39L RPT1 MASPI RPT4 SST RT2 BCL6 LPP TPRG1 TP63 LEPRE1</td>
</tr>
<tr>
<td>25</td>
<td>TOF</td>
<td>CRKL</td>
<td>all</td>
<td>Loss</td>
<td>22q11.21</td>
<td>18710744</td>
<td>1085</td>
<td>673</td>
<td></td>
<td>BCL9 ACP6 GJA5 GJA8 GPR89B GPR89C NBPF11 LOC728912 PPLA4 NBPF14 NBPF10 NBPF15 NBPF16 NBPF11 LOC728912 FAM108A PRKAB2 FMOS CHDIL RIMB3P ZNF74 SCARF2 KHLH2 MEDI15 PI4KA SERPIND1 SNAP29 CRKL AIFM3 LZTR1 THAP7 P2RX6 SLC7A4</td>
</tr>
<tr>
<td>26</td>
<td>TOF</td>
<td>HOXA1</td>
<td>all</td>
<td>Gain</td>
<td>7p15.2–7p15.1</td>
<td>26113744</td>
<td>4718</td>
<td>3324</td>
<td></td>
<td>TBX20 MYH11</td>
</tr>
<tr>
<td>27</td>
<td>TOF</td>
<td>MYH11</td>
<td>all</td>
<td>Gain</td>
<td>16p13.11</td>
<td>14805290</td>
<td>1455</td>
<td>642</td>
<td></td>
<td>KBTBD2 FKB89 NTSC3 RP3 BBSS BMPER NPSR1 D PYL1TL1 TBX20 HERPUD2 SEPT7 EDEF1 NOM1 NPIP PXD1C1 NTAI1 RNR3 MPV17L C16orf45 KIAA0430 NDE1 MYH11 C16orf63 ABCC1 ABCC6 NOM3</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Subject</th>
<th>Subphenotype</th>
<th>100 CHD Gene Region</th>
<th>Exon(s)</th>
<th>LOSS_GAIN</th>
<th>Cytoband</th>
<th>CNV Start (Build 36, hgl8)</th>
<th>CNV Size, kb</th>
<th>Markers, n</th>
<th>Inheritance</th>
<th>Gene Names on CNV Segment (100 CHD Genes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td>28*</td>
<td>TOF</td>
<td>TERT</td>
<td>all</td>
<td>Loss</td>
<td>5p15.33</td>
<td>80069</td>
<td>2948</td>
<td>1893</td>
<td></td>
<td>PLEKHG4B LOC589257 CCDC127 SDHA PDCD6 LOC1163-49 EXOC3 SLC9A3 CEP72 TPPP ZDHHC11 BRD9 TRIP13 NDK2 SLC12A7 SLC6A19 SLC6A18 TERT C1PTM1L SLC6A3 LPCAT1 MRPL36 NDUF66 IRX4 IRX2 C5orf38 UBE2L3 YDJC CCDC116 SDF2LI PPL2 YPEL1 MAPK1 PPM1F TOP3B</td>
</tr>
<tr>
<td>29</td>
<td>TRI-AT</td>
<td>MAPK1</td>
<td>all</td>
<td>Gain</td>
<td>22q11.21–22q1.122</td>
<td>20264556</td>
<td>447</td>
<td>243</td>
<td></td>
<td>SNP242 PREDI MXD3 LMAN2 RGS14 SLC34A1 PEN3 F12 GRK6 PR7 DBN1 PDLIM7 DOK3 DDX41 PL10404 TMEH9 B4GAL7</td>
</tr>
<tr>
<td>30</td>
<td>TRI-AT</td>
<td>NSD1</td>
<td>e24-3'</td>
<td>Gain</td>
<td>5q35.3</td>
<td>176652686</td>
<td>330</td>
<td>133</td>
<td></td>
<td>NSD1 RAB24 PREDI MXD3 LMAN2 RGS14 SLC34A1 PEN3 F12 GRK6 PR7 DBN1 PDLIM7 DOK3 DDX41 PL10404 TMEH9 B4GAL7</td>
</tr>
<tr>
<td>31</td>
<td>Truncus arteriosus</td>
<td>MAPK1</td>
<td>all</td>
<td>Loss</td>
<td>22q11.21–22q1.122</td>
<td>20055986</td>
<td>1237</td>
<td>863</td>
<td></td>
<td>HIC2 RIMBP3B RIMBP3C UBE2L3 YDJC CCDC116 SDF2LI PPL2 YPEL1 MAPK1 PPM1F TOP3B</td>
</tr>
<tr>
<td>32</td>
<td>Truncus arteriosus</td>
<td>GATA6</td>
<td>all</td>
<td>Gain</td>
<td>18q11.2</td>
<td>17749666</td>
<td>308</td>
<td>168</td>
<td></td>
<td>GATA6 NFAT2 ATP9A SALL4 ZF6P24 TSH22</td>
</tr>
<tr>
<td>33</td>
<td>VSD (perimembranous)</td>
<td>CRKL</td>
<td>all</td>
<td>Loss</td>
<td>20q11.32</td>
<td>49428074</td>
<td>1839</td>
<td>1357</td>
<td></td>
<td>PH4KA SERPIN1 SNAP29 CRKL AIFM3 LZR1 THAP2 P2RX6 SLC7A4</td>
</tr>
<tr>
<td></td>
<td>VSD (perimembranous)</td>
<td>GATA4 SOX7</td>
<td>all</td>
<td>Loss</td>
<td>8p23.1</td>
<td>8027361</td>
<td>4456</td>
<td>3349</td>
<td></td>
<td>ACP6 BCL9 CHD11 FMOS5 GJA5 PRKAB2</td>
</tr>
<tr>
<td>35</td>
<td>VSD (perimembranous)</td>
<td>GATA6</td>
<td>all</td>
<td>Gain</td>
<td>18q11.1–18q11.2</td>
<td>16795645</td>
<td>6118</td>
<td>3683</td>
<td></td>
<td>GATA6 NEIL2 FDFT1 CTSB CTSB DEFB137 DEFB134 DEFB130 ZNF705 DUB3 FAM861 BDF130</td>
</tr>
</tbody>
</table>

*Inconclusive TAQMAN results (see Subject 28). Boldface indicates confirmed genes. “Unknown” means one parental DNA was unavailable. “Other cardiac” phenotype (case 19) is double-chamber right ventricle (DCRV).
Table 6. CNV frequency by subphenotype

<table>
<thead>
<tr>
<th>Phenotype/Subphenotype</th>
<th>Subjects</th>
<th>Subjects with CNV Loss</th>
<th>Subjects with CNV Gain</th>
<th>Subjects with CNV Loss or Gain</th>
<th>Subjects</th>
<th>Subjects with CNV Loss</th>
<th>Subjects with CNV Gain</th>
<th>Subjects with CNV Loss or Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHOP Cohort</td>
<td>2,026</td>
<td>19 (0.94)</td>
<td>20 (0.99)</td>
<td>39 (1.92)</td>
<td>2,026</td>
<td>19 (0.94)</td>
<td>20 (0.99)</td>
<td>39 (1.92)</td>
</tr>
<tr>
<td>MFHS Cohort</td>
<td>880</td>
<td>3 (0.34)</td>
<td>11 (1.25)</td>
<td>14 (1.59)</td>
<td>880</td>
<td>3 (0.34)</td>
<td>11 (1.25)</td>
<td>14 (1.59)</td>
</tr>
<tr>
<td>CHD Cohort</td>
<td>945</td>
<td>66 (6.98)</td>
<td>110 (11.64)</td>
<td>172 (18.20)**</td>
<td>810</td>
<td>12 (1.48)</td>
<td>23 (2.84)</td>
<td>35 (4.32)**</td>
</tr>
<tr>
<td>Turner</td>
<td>8</td>
<td>8 (0.84)</td>
<td>1 (0.10)</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trisomy18 (T18)</td>
<td>1</td>
<td>0 (0.00)</td>
<td>1 (0.10)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trisomy21 (T21)</td>
<td>80</td>
<td>0 (0.00)</td>
<td>80 (8.35)</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams</td>
<td>3</td>
<td>3 (0.31)</td>
<td>0 (0.00)</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XXX</td>
<td>1</td>
<td>0 (0.00)</td>
<td>1 (0.10)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>42</td>
<td>42 (4.38)</td>
<td>1 (0.10)</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aorto-pulmonary window + PDA</td>
<td>5</td>
<td>0 (0.00)</td>
<td>2 (40.00)</td>
<td>2 (40.00)**</td>
<td>3</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>AVSD + TOF (AVSD + TOF)</td>
<td>7</td>
<td>0 (0.00)</td>
<td>6 (85.71)</td>
<td>6 (85.71)**</td>
<td>2</td>
<td>0 (0.00)</td>
<td>1 (50.00)</td>
<td>1 (50.00)**</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aortic Stenosis (Valvar)</td>
<td>31</td>
<td>3 (9.68)</td>
<td>1 (3.23)</td>
<td>4 (12.90)**</td>
<td>30</td>
<td>2 (6.67)</td>
<td>1 (3.33)</td>
<td>3 (10.00)**</td>
</tr>
<tr>
<td>Turner</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial Septal Defect Secundum (ASD-SEC)</td>
<td>47</td>
<td>1 (2.13)</td>
<td>4 (8.51)</td>
<td>5 (10.64)**</td>
<td>43</td>
<td>1 (2.33)</td>
<td>0 (0.00)</td>
<td>1 (2.33)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial Septal Defect Sinus Venosus (ASD-SV)</td>
<td>13</td>
<td>1 (7.69)</td>
<td>0 (0.00)</td>
<td>1 (7.69)</td>
<td>13</td>
<td>1 (7.69)</td>
<td>0 (0.00)</td>
<td>1 (7.69)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-V Canal Complete (AVC Complete)</td>
<td>48</td>
<td>0 (0.00)</td>
<td>35 (72.92)</td>
<td>35 (72.92)**</td>
<td>13</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-V Canal Intermediate (AVC Intermediate)</td>
<td>7</td>
<td>0 (0.00)</td>
<td>5 (71.43)</td>
<td>5 (71.43)**</td>
<td>2</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-V Canal Partial (AVC Partial)</td>
<td>17</td>
<td>1 (5.88)</td>
<td>3 (17.65)</td>
<td>4 (23.53)**</td>
<td>15</td>
<td>1 (6.67)</td>
<td>1 (6.67)</td>
<td>2 (13.33)**</td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-V Canal Unbalanced + AVSD with ventricular imbalance</td>
<td>14</td>
<td>0 (0.00)</td>
<td>2 (14.29)</td>
<td>2 (14.29)**</td>
<td>13</td>
<td>0 (0.00)</td>
<td>1 (7.69)</td>
<td>1 (7.69)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coarctation of the Aorta (CoA)</td>
<td>66</td>
<td>6 (9.09)</td>
<td>2 (3.03)</td>
<td>8 (12.12)**</td>
<td>61</td>
<td>1 (1.64)</td>
<td>1 (1.64)</td>
<td>2 (3.28)</td>
</tr>
<tr>
<td>Turner</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double Inlet Left Ventricle (DILV)</td>
<td>19</td>
<td>0 (0.00)</td>
<td>1 (5.26)</td>
<td>1 (5.26)</td>
<td>19</td>
<td>0 (0.00)</td>
<td>1 (5.26)</td>
<td>1 (5.26)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double Outlet Right Ventricle (DORV)</td>
<td>42</td>
<td>1 (2.38)</td>
<td>3 (7.14)</td>
<td>4 (9.52)**</td>
<td>41</td>
<td>0 (0.00)</td>
<td>2 (4.88)</td>
<td>2 (4.88)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebstein’s Anomaly (EBSTEINS)</td>
<td>9</td>
<td>0 (0.00)</td>
<td>1 (11.11)</td>
<td>1 (11.11)</td>
<td>9</td>
<td>0 (0.00)</td>
<td>1 (11.11)</td>
<td>1 (11.11)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypoplastic Left Heart Syndrome (HLHS)</td>
<td>140</td>
<td>2 (1.43)</td>
<td>5 (3.57)</td>
<td>7 (5.00)**</td>
<td>138</td>
<td>0 (0.00)</td>
<td>4 (2.90)</td>
<td>4 (2.90)</td>
</tr>
<tr>
<td>Turner</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrupted Aortic Arch (IAA)</td>
<td>11</td>
<td>4 (36.36)</td>
<td>0 (0.00)</td>
<td>4 (36.36)**</td>
<td>7</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitral Valve Stenosis (MS, subvalvar, parachute)</td>
<td>6</td>
<td>1 (16.67)</td>
<td>0 (0.00)</td>
<td>1 (16.67)</td>
<td>5</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Trisomy21</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22qDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other, Cardiac</td>
<td>18</td>
<td>1 (5.56)</td>
<td>2 (11.11)</td>
<td>3 (16.67)**</td>
<td>16</td>
<td>0 (0.00)</td>
<td>1 (6.25)</td>
<td>1 (6.25)</td>
</tr>
</tbody>
</table>

Continued
gene list to contrast with CNVs found elsewhere. For instance, haploinsufficiency of the genes associated with William’s Syndrome, FKBP6, ELN, and GTF2IRD1, identified the three William’s Syndrome patients in the study (1). Losses of the TBX1 and CRKL genes are associated with 22qDS and were observed in deleted subjects (32, 53). Turner syndrome subjects carrying losses on the chromosome X genes involving MID1, BCOR, ATRX, GPC3, ZIC3, and FLNA were identified, as well as a female subject (XXX) who was identified with gains over these chromosome X gene regions. In addition, duplications involving RUNX1 were primarily Trisomy 21 subjects.

Gains at 1q21.1 including PRKAB2, FMO5, CHD11, BCL9, ACP6, and GJA5 were significantly enriched in this study; however, losses that were observed in both control cohorts as well as the CHD cohort were not. Interestingly, gains at 1q21.1 were previously reported in isolated sporadic TOF (18). In our case cohort we observed one subject (case 24) with TOF (2 contiguous CNVs, 0.6 and 1.6 Mb), one subject (case 20) with PA-VSD (1.5 Mb), and another (case 10) with CoA (1.5 Mb). One complex subject (case 2) with AS valvar and Shone’s had a shorter gain (418 kb) involving only PRKAB2, FMO5, and CHD11 in conjunction with a 1.8 Mb gain at 5q35.2, which included the NSD1 gene.

Chromosome 8p23.1 deletions involving GATA4 were enriched and have been reported as a cause of complex congenital heart defects and diaphragmatic hernia (51). These included subjects with AVC partial (case 6, 3.8 Mb loss), VSD per-

<table>
<thead>
<tr>
<th>Table 6.—Continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenotype/Subphenotype</td>
</tr>
<tr>
<td>Subjects</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>22qDS</td>
</tr>
<tr>
<td>Pulmonary Atresia (PA)</td>
</tr>
<tr>
<td>-IVS-</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>XXX</td>
</tr>
<tr>
<td>22qDS</td>
</tr>
<tr>
<td>VSD-</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>Subaortic stenosis</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>Supravalvar AS</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>Williams</td>
</tr>
<tr>
<td>Tetralogy of Fallot (TOF)</td>
</tr>
<tr>
<td>Trisomy18</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>22qDS</td>
</tr>
<tr>
<td>Tricuspid Atresia (TRI-AT)</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>Truncus Arteriosus (TA)</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>22qDS</td>
</tr>
<tr>
<td>VASCULAR ring and PA sling</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>VSD inlet</td>
</tr>
<tr>
<td>VENTRICULAR Septal Defect (VSD perimembranous)</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
<tr>
<td>22qDS</td>
</tr>
<tr>
<td>VENTRICULAR Septal Defect (VSD subarterial)</td>
</tr>
<tr>
<td>Trisomy21</td>
</tr>
</tbody>
</table>

**Significance over both CHOP and MFHS controls (P ≤ 0.05). Four patients had both gains and losses but are only counted once in the column “Subjects with CNV Loss or Gain”. The following subphenotypes contained 0 subjects with a CNV and were therefore removed from the table: Arrhythmias (Congenital Heart Block, Long QT, WPW), 7; Cardiomyopathy (DILATED), 13; Cardiomyopathy (HYPERTROPHIC), 4; Chest Wall, 4; Coronary Arteries (COR ART), 10; L-TGA, 7; Dilated Ascending Aorta (MARFAN), 8; Partial Anomalous Pulmonary Venous Return (PAPVR), 12; Pulmonary Stenosis (Valvar), 9; Shone’s, 8; Total Anomalous Pulmonary Venous Connection (TAPVC; infracardiac, intracardiac, mixed, supracardiac), 15; Transposition of Great Arteries (IVS), 21; (VSD), 20; and Ventricular Septal Defect (VSD multiple + muscular), 10 (n = 161 total).
imembranous \textit{(case 34, 4.5 Mb loss)}, and ASD-SV \textit{(case 5, 304 kb loss)}.

Three subjects had gains involving the \textit{HRAS} gene. The first was found in a complex subject with coarctation of the aorta: in addition to a 284 kb duplication involving the \textit{HRAS} gene the subject had Turner syndrome. The remaining two gains \textit{(case 12, 256 kb; case 22, 271 kb)} were found in subjects with DILV and subaortic stenosis, respectively (Table 5). Cardiovascular malformations are known to be related to Ras/MAPK pathway syndromes, and previous literature findings have reported associations of \textit{HRAS} mutations in Costello Syndrome and with the subaortic stenosis phenotype (29). These gains involving \textit{HRAS} appear to expand phenotypes related to the Ras/MAPK pathway.

Enriched CNVs identified in Table 7 are previously reported or can be found in the Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources (DECIPHER) with the exception of the gains involving \textit{GATA6}. One of the three gains involving \textit{GATA6} was in a subject with Trisomy 18 with TOF. The remaining two subjects with CNV gains involving \textit{GATA6} were 1) a subject \textit{(case 31)} with truncus arteriosus with a complex CNV over two CHD genes of interest, a 308 kb gain including \textit{GATA6}, and a 1.2 Mb 22q11.2 distal deletion involving \textit{MAPK1} (losses in the distal region of 22q11.2 have previously been reported in subjects with truncus arteriosus) (2), and 2) a subject \textit{(case 35)} with VSD perimembranous with two neighboring 6.1 and 6.9 Mb gains involving a gain on \textit{GATA6}. Although sequence variants in \textit{GATA6} have been previously found to be associated with cardiac outflow tract defects (27), these gains have not been reported and suggest possible \textit{GATA6} triple sensitivity to conotruncal defects.

\textit{Collapsing groups of phenotypes by recognized causal chromosomal abnormalities.} To increase statistical power, a strategy for summing cohorts was employed; subphenotypes associated with T21, 22qDS, and Turner Syndrome (see Tables 4 and 6) were collapsed into three groups, respectively (33). We hypothesized collapsing subphenotypes into genetically related groups would increase power to detect additional related CNVs by phenotype. The three collapsed groups each demonstrated significant enrichment \textit{(P ≤ 0.05)} of additional CNVs compared with both control cohorts (see Table 8 - Enriched Syndrome Genes and Fig. 6 - Spectra). Large, rare CNVs were significantly more frequent \textit{(P ≤ 0.05)} in the groups of T21 subphenotypes and included gains involving \textit{GATA6} and \textit{RUNX1} and losses involving \textit{GATA4}, \textit{SOX7}, \textit{TBX1}, and \textit{CRKL}. Likewise, collapsing the HLHS, CoA, and AS (valvar) subphenotypes, which made up the Turner syndrome group, indicated significant gains involving the 1q21.1 gene regions, enriched losses involving the Chr. X genes, as well as gains involving \textit{GATA4}, \textit{SOX7}, \textit{EHMT1} \textit{(case 15)}, and \textit{HRAS} and losses involving \textit{FOXC1} \textit{(case 3)} and \textit{NOTCH1} \textit{(case 11)}. Although the T21 and 22qDS subclasses share some overlap of phenotypes (other cardiac, TOF, vascular ring/PA sling, and VSD perimembranous), it is interesting to note that the 22qDS grouping also included gains involving the 1q21.1 genes as well as \textit{GATA6} and \textit{RUNX1}. Significant CNV losses within the 22qDS subclasses involved \textit{TBX1} and \textit{CRKL}. All CNVs identified through the collapsed phenotypes are listed in Table 8 and are reported in DECIPHER.
Additional findings of note include a gain involving \textit{TBX20} and loss involving \textit{SALL4}. Three losses including \textit{TBX20} have been previously reported in subjects with CHD (ASD and VSDs) (26, 38). We identified a subject (case 26) with TOF with a 3.3 Mb gain involving \textit{TBX20} and an adjacent 4.7 Mb gain involving \textit{HOXA1}, which has been reported in DECI-PHER. Finally, we report a subject (case 32) with truncus arteriosus with a 1.8 Mb loss over the \textit{SALL4} gene, which has not been previously reported. This segment included a loss over \textit{NFATC2}, a regulator of cardiac transcription factors but was not included in our 100 gene list because likely causal variants have not previously been reported in humans in this gene (9).

\textbf{Distribution of CNVs by subject.} To characterize CHD study subjects with an approach more typically used in clinical genetics, CNVs were separated by size (whether or not they would be cytogenetically visible) and then the CHD WIKI site was employed to determine if remaining CNVs should be classified as involving a “syndromic” (two or more clinical features) or a “nonsyndromic” gene (1). Cytogenetically visible CNVs (category A) included chromosomal abnormalities \( \geq 3 \) Mbps. This category contained subjects with Trisomy 21, 18; Turner; and XXX syndrome and represented \( \approx 9\% \) of the CHD cohort. Category B, contributing \( 6\% \) to the overall CNV distribution, were those subjects with a CNV over a “syndromic-associated” CHD gene as reported by CHD WIKI (1).
This subset contained 22qDS subjects \((n = 42)\) with losses over the \(TBX1\) gene, William’s Syndrome subjects, all with a phenotype of supravalvar aortic stenosis \((n = 3)\) with losses over the \(ELN, GTF2IRD1,\) and \(FKBP6\) genes. The “nonsyndromic” segment \((\text{category C})\) representing 1% of the CHD cohort was also defined by the CHD WIKI portal. Six CHD case subjects, contributing 1% to the total, had a CNV over one of the 100 CHD-associated genes; however, their category was considered unknown. \(\text{Category E}\) represented individuals with no CNV over our predefined 100 CHD risk gene list. An individual could only fit into one category where \(D > A > B > C\) (see Fig. 7).

**Complex CNVs.** Four basic mechanisms are involved in the generation of a majority of CNVs: deletion, duplication, inversion, and related combinations \((56)\). We were interested if CHD subjects were at increased risk for carrying multiple CNVs. In the current study, 125 CHD subjects were defined as complex (\(\text{METHODS}\)). We identified 100 of those with known

### Table 8. Enriched syndrome genes

<table>
<thead>
<tr>
<th>Chr.</th>
<th>Gene</th>
<th>22q Like</th>
<th>T21 Like</th>
<th>Turner Like</th>
<th>CHOP</th>
<th>MFHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gain, %</td>
<td>Loss, %</td>
<td>Gain, %</td>
<td>Loss, %</td>
<td>Gain, %</td>
</tr>
<tr>
<td>1</td>
<td>ACP6</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>1</td>
<td>BCL9</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>1</td>
<td>CHD1L</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>1</td>
<td>FMO5</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>1</td>
<td>GATA4</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>1</td>
<td>PRKAB2</td>
<td>✓</td>
<td>0.67</td>
<td>0.33</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>6</td>
<td>FOXC1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>GATA4</td>
<td>0.00</td>
<td>0.33</td>
<td>0.00</td>
<td>✓</td>
<td>0.55</td>
</tr>
<tr>
<td>8</td>
<td>SOX7</td>
<td>0.00</td>
<td>0.33</td>
<td>0.00</td>
<td>✓</td>
<td>0.55</td>
</tr>
<tr>
<td>9</td>
<td>EHMT1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>9</td>
<td>NOTCH1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>11</td>
<td>HRAST</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>18</td>
<td>GATA6</td>
<td>✓</td>
<td>1.00</td>
<td>0.00</td>
<td>✓</td>
<td>0.55</td>
</tr>
<tr>
<td>21</td>
<td>RUNX1</td>
<td>✓</td>
<td>7.36</td>
<td>0.00</td>
<td>✓</td>
<td>22.25</td>
</tr>
<tr>
<td>22</td>
<td>CRKL</td>
<td>✓</td>
<td>0.67</td>
<td>13.38</td>
<td>1.10</td>
<td>✓</td>
</tr>
<tr>
<td>22</td>
<td>TBX1</td>
<td>✓</td>
<td>0.33</td>
<td>14.05</td>
<td>0.55</td>
<td>✓</td>
</tr>
<tr>
<td>X</td>
<td>ATRX</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>X</td>
<td>BCO</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>X</td>
<td>GPC3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>X</td>
<td>MID1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>X</td>
<td>ZIC3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Boldface indicates significant values.
Fig. 6. CNV frequency spectra of collapsed phenotypes by syndrome. Note that any gene showing 0% CNV frequency in all 3 cohorts was omitted from this figure due to space considerations. Vertical error bars drawn represent 1 SD in the estimated sampling distribution. Turner phenotypes, T21 phenotypes, and 22qDS phenotypes.
CHD-associated syndromes (T21, 59; T18, 1; 22qDS, 31; Turner, 6; William’s, 2; XXX Syndrome, 1). Of the remaining 25, 24 contained likely causal CNVs for CHD as outlined in Table 5, whereas one subject contained a nonconfirmed CNV over a CHD-associated gene. Three complex subjects had CNVs on different chromosomes over two of our CHD associated genes of interest: subjects 2, 31, and 33 (Table 5). In addition, two subjects from the CHD cohort were both syndromic with their additional CNV over a second gene of interest: a Turner syndrome subject had a gain involving the HRA5 gene and a 22qDS subject had an additional CNV involving a gain over the MAPK1 gene.

It is interesting to note that applying the “complex” criteria to the MFHS control cohort also identified 10 subjects from the controls that met the complex analysis requirements. These subjects had gains over the genes FKBP6, MYH11, TERT, TBX1, CRKL, SH3PXD2B, and losses over the 1q21.1 gene region and MID1.

DISCUSSION

CHD is a complex disease with demonstrated genetic etiology in a subset of patients. CNVs, viewed as an evolutionary driving force for new gene function resulting in improved survival and/or adaption to new environments and disease, contribute the largest component of natural human variation between any two individuals; indeed, CNVs contribute significantly more to inter-individual variation than SNPs (35, 39, 41). There is a broad range of CNV lengths. In this study we focused on large CNVs that can be detected with high accuracy and are relatively straightforward to confirm. It has previously been estimated that ~65~80% of individuals have a large CNV (>100 kb) and approximately three to seven CNV segments per individual (56). The average number of CNVs per subject in our CHD cohort supports these previous observations (Fig. 3). It is apparent that as CNV data continue to grow, the development of higher-resolution approaches will permit smaller CNV detection with better accuracy. This will potentially lead to additional disease association discoveries (23). However, data suggest that common CNVs (CNPs) are likely to be lower penetrance risk factors, whereas rare CNV variants are more likely to carry highly penetrant disease risk factors (13).

Significant challenges remain in CNV disease-association studies at both the platform and analysis levels (37). The relationship between phenotype and gene dosage is complex. Our study represents a comprehensive data curation and filtering of CNVs involving 100 recognized CHD risk genes detected in a large, anatomically phenotyped CHD population. We employed a strict algorithm to determine frequencies of CNVs involving regions that encompassed these CHD risk genes. The algorithm employed was very similar to a recent recommendation by Breckpot et al. (6) for determining if CNVs detected in CHD patients are clinically relevant; herein we performed a comparison against two different control populations and an analysis primarily based on known chromosomal abnormalities and gene content rather than more commonly used CNV detection approaches that prioritize by size. CNVs over these predefined gene regions were then used to search for relationships between cardiac phenotype and gene dosage.

The novel analytical approach described herein identified known causal chromosomal abnormalities (including T21, T18, 22qDS, Turner, William’s, and XXX Syndromes), which represent 14% of CHD subjects in this study, similar to previous observations (36). Overall, this descriptive study suggests that (after excluding well-established causal chromosomal abnormalities) large, rare CNVs in 100 well-defined CHD risk genes confers significant risk of CHD and is likely etiologic in 4.3% of CHD cases, similar to previous observations (6). Cardiac subphenotypes showing the most significant (P ≤ 0.05) enrichment of large CNV events were aortic stenosis (valvar), AV canal (partial), AVSD with TOF, subaortic stenosis, TOF, and truncus arteriosus. CNV frequency spectra analysis identified enriched genes (P ≤ 0.05): losses: FKBP6, ELN, GTF2IRD1, GATA4, CRKL, TBX1, ATRX, GPC3, BCO, ZIC3, FLNA, and MID1; and gains: PRKAB2, FM05, CHD1L, BCL9, ACP6, GJA5, HRAS, GATA6, and RUNX1. 1q21.1 gains were enriched in subjects with conotruncal defects and coarctation of the aorta. 8p23.1 losses were enriched in subjects with septal defects and gains involving HRAS were observed in subaortic stenosis and DILV. Cardiovascular malformations are known to be related to Ras/MAPK pathway syndromes and previous literature findings have reported associations of HRAS mutations resulting in increased hRAS signaling with the subaortic stenosis phenotype. Other common phenotypes occurring in patients with hRAS mutations (also known as Costello syndrome) are cardiac hypertrophy (usually typical hypertrophic cardiomyopathy) and arrhythmia (usually supraventricular tachycardia, especially chaotic atrial rhythm/multifocal atrial tachycardia or ectopic atrial tachycardia) (43). Although DILV sometimes are associated with pulmonary stenosis we have not found any previous reports of hRAS mutations linked to this phenotype. Thus, our data appear to expand phenotypes related to the Ras/MAPK pathway.

We hypothesized that CNV frequency spectra combined with detailed anatomic classes would define the impact of gene dosage in etiologic molecular pathways. One set of clues when searching for genetic causes of CHD is given by the enrichment of CHD cases in various recognized causal chromosomal
abnormalities such as T21, T18, 22qDS, Turner syndrome, and William’s Syndrome (36). For instance, three of four (75%) subjects in our study with supravalvar aortic stenosis had deletions involving FKBP6, ETVN, and GTF2IRD1 genes; all of these subjects had William’s syndrome. In Turner syndrome, the incidence of CHD can be as high as 50% and include phenotypes such as BAV, CoA, ASD-VSD partially anomalous pulmonary vena cava, and HLHS, but these data vary (4). The specific cause for CHD in patients with Turner syndrome is currently unknown; several genes have been implicated but for the most part do not quite match Turner syndrome phenotypes (4, 5). In the present study, eight Turner syndrome subjects were easily identified from the total CHD cohort by CNV frequency spectra analysis. In these subjects, copy number losses were present on all six of the Chr. X genes (MID1, BCOR, ATRX, GPC3, ZIC3, and FLNA) that were selected as CHD-associated from our list of 100 genes. Two out of eight Turner cases in the study had HLHS, five had coarctation of the aorta, and one had aortic stenosis (valvar). Turner syndrome-associated phenotype percentages for the CHD cohort were in good agreement with published reports (4).

To test for additional CNV gene enrichment with increased power, subphenotypes associated with T21, 22qDS, and Turner Syndrome were collapsed in these groups, respectively. The three collapsed groups of subphenotypes each demonstrated enrichment ($P \leq 0.05$) in additional CNVs compared with both control cohorts. Large, rare CNVs significantly increased ($P \leq 0.05$) in the groups of T21 subphenotypes included gains over GATA6 and RUNX1 and losses over GATA4, SOX7, TBX1, and CRKL. Likewise, collapsing the HLHS, CoA, and AS (valvar) subphenotypes that made up the Turner syndrome group indicated significant gains over the 1q21.1 gene regions, enriched losses over the Chromosome X genes, as well as gains over likely etiologic genes such as GATA4, SOX7, EMT1, and Hras and losses over FOXC1 and NOTCH1. Although the T21 and 22qDS collapsed groups share some overlap of phenotypes, it is interesting to note that the 22qDS grouping also included gains over the 1q21.1 genes, as well as GATA6 and RUNX1. Significant CNV losses within the 22qDS subclasses were over TBX1 and CRKL. Incorporating gene dosage with detailed phenotyping into current molecular cardiogenesis models may allow future models of development to fine tune and increase our understanding of etiologic pathways.

By narrowing our focus on a select set of 100 well-known CHD risk genes, we limited the study by design. We focused on large rare CNVs in the current study; therefore, smaller CNVs have not yet been examined. Furthermore, the number of subjects per phenotype was small because of detailed anatomic groupings; collapsing into fewer groups (with larger $n$) according to developmental models would increase power and may permit identification of additional enriched genes. An additional foreseeable limitation was that CNVs may have been enriched in genes, but because the analysis required statistical significance with two control cohorts (where CNVs were not confirmed and may have been inflated and manifested as false positives), the study may not have been sufficiently powered to detect smaller but true differences.

To our knowledge, this is the first paper to curate a large and diverse CHD population with regard to subphenotype and CNV frequency by gene region. This appears to be a useful approach to visualize and eventually, given sufficient numbers, to quantify relative risk of CNVs for specific subphenotypes. Broadening to encompass the entire genome and performing the copy number spectra analysis at higher resolution should identify additional candidate genes in CHD. The ability to quantify risk of particular cardiac malformations by gene dosage should offer insight into critical molecular pathways impacted during human cardiogenesis. Furthermore, overlaying CNV data and details of resulting cardiac phenotype with known functional pathways of cardiogenesis should lead to increased understanding of the molecular etiology of heart malformations.

ACKNOWLEDGMENTS

First and foremost, we thank the families of the HHC from the CHW. We also acknowledge the significant contributions from the WPCR. We thank M. Kroulikowski for excellent assistance with IRB contributions; the cardiothoracic surgical team including D. A. Jones, M. E. Barnes, J. L. Dunham-Ingle, T. A. Fehrenbacher, M. R. Madrakz, L. E. May, and R. G. Smith for research coordination; and the OR nursing staff for making this study possible. We gratefully acknowledge A. Klemm for assistance with the HHC database at CHW; the AGEN laboratory, especially R. B. Lorier, A. Mahter, A. Turner, B. Chirejmes, and J. Wittke for technical assistance in the genetic laboratory; E. Semina for assistance in validations; S. Ghanta, J. Schroeder, A. Stoddard for helpful bioinformatic discussions; H. Liang for Tissue Bank support; and J. M. Routes, J. Lough, D. Beard, and B. Aronow for very helpful discussions and editing of this manuscript.

GRANTS

This study was supported in part by National Institutes of Health (NIH) Grant NIH1R21HD-060309-01 (A. Tomita-Mitchell), the CRI (A. Tomita-Mitchell and M. E. Mitchell), the Department of Surgery (A. Tomita-Mitchell, J. S. Tweddell, M. E. Mitchell), and in part by NIH Grant NIH5R01HL-089655 (U. Broeckel).

DISCLOSURES

A. Tomita-Mitchell, M. E. Mitchell, and C. A. Struble have conflicts of interest (significant financial interest in Aria Diagnostics, a molecular diagnostics company). M. Hidestrand and M. A. Goetsch have conflicts of interest (significant financial interest in Aria Diagnostics, a molecular diagnostics company). M. Hidestrand and M. E. Mitchell), and in part by NIH Grant NIH5R01HL-089655 (U. Broeckel).

AUTHOR CONTRIBUTIONS


REFERENCES


5. Bondy CA, Cheng C. Monosomy for the X chromosome. Chromosoma 

Moreau Y, Gewillig M, Devriendt K. Challenges of interpreting copy 
number variation in syndromic and non-syndromic congenital heart 

Allegaert K, Vanhcole C, Eyskens B, Vermeesch JR, Gewillig M, 
Devriendt K. Array comparative genomic hybridization as a diagnostic 

8. Chan ISF. Exact tests of equivalence and efficacy with a non-zero lower 

Graef IA, Crabtree GR. A field of myocardial-endocardial NFAT sig-

10. Chinn A, Fitzsimmons J, Shepard TH, Fantel AG.

11. Cronk CE, Gangnon R, Cossette S, Mielcarek MA, Pelech AN.

12. Cooper GM, Zerr T, Kidd JM, Eichler EE, Nickerson DA.

13. Cronk CE, Gangnon R, Cossette S, Mielcarek MA, Pelech AN.

14. D'Armiento M, Olla C, Paladini D, Zannini M, Nitsch L.

15. Degli Esposti M, Hirstman JW, Hillier LW, Howard H, 
Hudson JM, Hughes D, Hunt S, Isaacs JD, Jain M, Jewell DP, 
Johnson T, Jolley JD, Jones IR, Jones LA, Kirov G, Landgford 
CF, Lango-Aleno H, Lathrop GM, Lee J, Lee KL, Lees C, Lewis K, 
Lindgren CM, Maisuria-Armer M, Maller J, Mansfield J, Martin P, 
Massey DC, McArdle WL, McCullogh P, McLaey KE, Mentzer A, 
Minnack ML, Morgan AE, Morf C, Myers S, Newman W, O'Leary 
ER, O'Bannon MC, Onishima A, Onyiah I, Ovington NR, Owen MJ, 
Palin K, Parnell K, Pernet D, Perry JR, Philips A, Pinto D, 
Pescott NJ, Prokopenko I, Quail MA, Rafiel S, Rayner A, 
Redon R, Reid DM, Renwick Ring SM, Robertson N, Russell E, 
St Clair D, Sambrook JG, Sanderson JD, Schuijlenburg H, Scott CE, 
Scott R, Seal S, Shaw-Hawkins S, Shields BM, Simmons MJ, Smythe DJ, 
Somaskantharajah E, Spanova K, Steer S, Stephens J, Stevens HE, 
Stone MA, Su Z, Symmons DP, Thompson J, Thomson W, 
Travers ME, Turnbull C, Valdesia A, Walker M, Walker NM, 
Wallace C, Warren-Perry M, Watkins NA, Webster J, Weeden MN, 
Wilson AG, Woodburn M, Wordsworth BP, Young AH, Zeggini E, 
Carter NP, Frayling TM, Lee C, McVeane G, Munroe PB, Palotie A, 
Sawyer S, Scherer SW, Strachan DP, Tyler-Smith C, Brown MA, 
Burton PR, Caulfield MJ, Compston A, Farrall M, Gough SC, Hall AS, 
Hattersley AT, Hill AV, Mathew CG, Pembrey M, Satsangi J, 
Stratton MR, Worthington J, Deloukas P, Duncanson A, Kwiatkowski 
DP, McCarthy MI, Ouwehand W, Parkes M, Rahman N, Todd JA, 
Samanji N, Donnelly P. Genome-wide association study of CNVs in 
16,000 cases of eight common diseases and 3,000 shared controls. Nature 


17. Erugler K. Barnard: Barnard’s Unconditional Test. R package version 

18. Ferencz C, Correa-Villasenor A, Loffredo CA, Wilson PD. Malforma-
tions of the cardiac outflow tract. In: Genetic and Environmental Risk 
Factors of Major Congenital Cardiovascular Malformations: The Baltimore-Wash-
59–102.

19. Franklin RC, Jacobs JP, Tchervenkov CI, Beland MJ. Bidirectional 
crossmap of the Short Lists of the European Paediatric Cardiac Code and the 
International Congenital Heart Surgery Nomenclature and Database 

SM, Ergul E, Conta TL, Korn JM, McCarroll SA, Gorham JM, 
Gabriel S, Altshuler DM, Quintanilla-Dieck M, Artunduaga MA, 
Eavey RD, Plenge RM, Shadick NA, Weinblatt ME, De Jager PL, 
Haller DA, Breithart RE, Seidman JG, Seidman CE. De novo copy 
number variants identify new genes and loci in isolated sporadic tetralogy 

chromosome loss and aging: in situ hybridization studies on human 

22. Harris S, Cranck C, Cassidy L, Simpson PM, Tomita-Mitchell A, 
Pelech AN. Exploring the environmental and genetic etiologies of con-
genital heart defects: the Wisconsin Pediatric Cardiac Registry. J Registry 


24. Iaffrate AF, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, 
Scherer SW, Lee C. Detection of large-scale variation in the human 

RM, Myers RM, Ridker PM, Chasman DI, Mefford H, Ying P, 
Nickerson DA, Eichler EE. Population analysis of large copy number 
variants and hotspots of human genetic disease. Am J Hum Genet 84: 

26. Jenkins KJ, Correa A, Feinstein J, Botto L, Brit AE, Daniels SR, 
Elison M, Warnes CA, Webb CL. Noninherited risk factors and con-
genital cardiovascular defects: current knowledge: a scientific state-
ment from the American Heart Association Council on Cardiovascular 

27. Kathiria IS, Srivastava D. Left-right asymmetry and cardiac looping: 
implications for cardiac development and congenital heart disease. Am J 

AD, Hayward C, Keoghan A, Macdonald P, Griffiths L, Fatkin D, 
Sholler GF, Zorn AM, Fenley MP, Winlans DW, Harvey RP. Mutations in 
cardiac T-box factor gene TBX20 are associated with diverse cardiac 
pathologies, including defects of septation and valvulogenesis and cardio-

human cardiac outflow tract defects by disrupting semaphorin-plxin 

30. Korn JM, Kuruviula FG, McCarroll SA, Wyssok A, Nemesh J, 
MM, Gabriel SB, Purcell S, Daly MJ, Altschuler D. Integrated genotype 
calling and association analysis of SNPs, common copy number polymor-

31. Lin AE, Alexander ME, Colan SD, Kerr B, Rauen KA, Noonan J, 
Baffa J, Hopkins E, Sol-Church K, Limongelli G, Dillgo MC, Marino 
B, Innes AM, Aoki Y, Silberbach M, Delrue MA, White SM, Hamilton 
RM, O’Connor W, Grossfeld PD, Smoot LB, Padera RF, Gripp KW. Clinical, 
pathological, and molecular analyses of cardiovascular abnor-
malities in Costello syndrome: a Ras/MAPK pathway syndrome. J Hum 

32. Magri C, Sacchetti E, Traversa M, Valsecchi P, Gardella R, Bonvicini 
C, Minelli A, Gennarelli M, Barlati S. New copy number variations in 

33. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter 
D, Lieber MT, McCarroll MA, O'Connell RJ, Chakravarti A, Cho JH, 
Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin 
M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, 
Gibson G, Haines J, Mackay TF, McCarroll SA, Visscher PM.


