Journal of the American Heart Association

ORIGINAL RESEARCH

A Multicenter Analysis of Abnormal Chromosomal Microarray Findings in Congenital Heart Disease

Benjamin J. Landis, MD; Lindsey R. Helvaty, BS; Gabrielle C. Geddes, MD; Jiuann-Huey Ivy Lin, MD, PhD; Svetlana A. Yatsenko, MD; Cecilia W. Lo, PhD; William L. Border, MD; Stephanie Burns Wechsler, MD; Chaya N. Murali, MD; Mahshid S. Azamian, MD, MPH; Seema R. Lalani, MD; Robert B. Hinton, MD; Vidu Garg, MD; Kim L. McBride, MD, MS; Jennelle C. Hodge, PhD; Stephanie M. Ware, MD, PhD

BACKGROUND: Chromosomal microarray analysis (CMA) provides an opportunity to understand genetic causes of congenital heart disease (CHD). The methods for describing cardiac phenotypes in patients with CMA abnormalities have been inconsistent, which may complicate clinical interpretation of abnormal testing results and hinder a more complete understanding of genotype–phenotype relationships.

METHODS AND RESULTS: Patients with CHD and abnormal clinical CMA were accrued from 9 pediatric cardiac centers. Highly detailed cardiac phenotypes were systematically classified and analyzed for their association with CMA abnormality. Hierarchical classification of each patient into 1 CHD category facilitated broad analyses. Inclusive classification allowing multiple CHD types per patient provided sensitive descriptions. In 1363 registry patients, 28% had genomic disorders with well-recognized CHD association, 67% had clinically reported copy number variants (CNVs) with rare or no prior CHD association, and 5% had regions of homozygosity without CNV. Hierarchical classification identified expected CHD categories in genomic disorders, as well as uncharacteristic CHDs. Inclusive phenotyping provided sensitive descriptions of patients with multiple CHD types, which occurred commonly. Among CNVs with rare or no prior CHD association, submicroscopic CNVs were enriched for more complex types of CHD compared with large CNVs. The submicroscopic CNVs that contained a curated CHD gene were enriched for left ventricular obstruction or septal defects, whereas CNVs containing a single gene were enriched for conotruncal defects. Neuronal-related pathways were over-represented in single-gene CNVs, including top candidate causative genes *NRXN3*, *ADCY2*, and *HCN1*.

CONCLUSIONS: Intensive cardiac phenotyping in multisite registry data identifies genotype–phenotype associations in CHD patients with abnormal CMA.

Key Words: chromosomal microarray ■ congenital heart disease ■ conotruncal defects ■ genomics ■ neurodevelopment

Congenital heart disease (CHD) is a major cause of mortality and morbidity from infancy to adulthood. As genetic testing technologies have advanced, so has the understanding of the genetic underpinnings of CHD.[1] Chromosomal microarray analysis (CMA) is a genome-wide technique that identifies intervals of genomic gains or losses, referred to as copy number variants (CNVs), as well as regions of homozygosity (ROH), and has been recommended as a first-tier test for patients with neurodevelopmental disorders and congenital anomalies.[2] CMA has been integrated into routine practice at many pediatric cardiac centers for infants

Correspondence to: Stephanie M. Ware, MD, PhD, Indiana University School of Medicine, 1044 West Walnut St, R4-W227, Indianapolis, IN 46202. Email: stware@iu.edu

This manuscript was sent to Kevin F. Kwaku, MD, PhD, Associate Editor, for review by expert referees, editorial decision, and final disposition.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.123.029340

For Sources of Funding and Disclosures, see page xxx.

JAHA is available at: www.ahajournals.org/journal/jaha

RESEARCH PERSPECTIVE

What Is New?

- Novel genotype-phenotype associations were established for genomic disorders allowing identification of distinct congenital heart defect subtypes for certain reciprocal gains and losses of genetic information.
- Large copy number variants were associated with simple septal defects, copy number variants containing curated congenital heart defect genes were associated with left ventricular obstructive lesions, and single-gene copy number variants were associated with conotruncal defects.

What Question Should Be Addressed Next?

- Single-gene copy number variants were enriched for genes related to neuronal development and cell–cell adhesion, and included novel candidate genes *NRXN3*, *ADCY2*, and *HCN1*, highlighting the need for additional studies to define neuronal genes' roles in heart development and neurodevelopmental outcome.

Nonstandard Abbreviations and Acronyms

APVR	anomalous pulmonary venous return
AVSD	atrioventricular septal defect
CMA	chromosomal microarray
CNV	copy number variant
CTD	conotruncal defect
HLHS	hypoplastic left heart syndrome
LVOTO	left ventricular obstructive lesion
Mb	million base pairs
OMIM	Online Mendelian Inheritance in Man
pLI	probability of loss-of-function intolerant
ROH	region of homozygosity
TOF	tetralogy of Fallot

with severe CHD.[3] CMA may establish a diagnosis in patients suspected to have a genomic disorder commonly associated with CHD, such as Williams (7q11.23 deletion; MIM #194050) or 22q11.2 deletion (#192430) syndromes. CMA provides greater sensitivity in detection of smaller CNVs than traditional fluorescence in situ hybridization and reveals genomic alterations in patients with atypical phenotypes.[4] The discovery of CHD association for more recently described genomic disorders, such as chromosome 1q21.1 duplication and deletion syndromes (MIM #612475, MIM #612474), has been facilitated by CMA.[5] In the clinical setting, CMA may identify CNVs that are suspected to cause a patient's CHD; however, gaps in current knowledge may limit the interpretation. Larger numbers of patients are required to characterize these CNVs, which will foster improved clinical interpretation and patient management. Also, pathophysiological insight may be gained by establishing CHD causality of CNVs and their associated cardiac phenotypes in humans.

The objectives of the Cytogenomics of Cardiovascular Malformations Consortium are to identify genomic regions that cause or increase susceptibility to CHD, correlate the findings with clinical phenotypes, and solidify the CHD associations with more recently characterized CNVs.[6] This multisite, cross-disciplinary collaboration has created a comprehensive registry of patients with CHD and abnormal clinical CMA. In this study, we analyze the genetic abnormalities identified by CMA in 1363 patients and correlate them with detailed description and systematic classification of cardiac phenotypes.

METHODS

Transparency and Openness

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Registry Overview and Organization

The Consortium is a collaborative group of medical geneticists, clinical molecular and cytogeneticists, basic scientists, and pediatric cardiologists from 9 pediatric clinical centers: Riley Hospital for Children at Indiana University Health (Indiana University School of Medicine, IUSM), Nationwide Children's Hospital (Ohio State University), Texas Children's Hospital (Baylor College of Medicine), Children's Hospital of Pittsburgh, Cincinnati Children's Hospital Medical Center, Primary Children's Hospital (University of Utah), Children's Wisconsin (Medical College of Wisconsin), Children's Healthcare of Atlanta (Emory School of Medicine), and Advocate Children's Hospital (Chicago Medical School).

Patient Cohort and Eligibility

The study was approved by each clinical center's Institutional Review Board and utilized a waiver of informed consent. Cases were included in the registry if the clinical laboratory reported an abnormal finding on CMA and the patient had an abnormal echocardiogram. Clinical laboratories used by Consortium sites had Clinical Laboratory Improvement Amendments certification. The CMA abnormalities included CNVs

interpreted as variants of uncertain significance, likely pathogenic or pathogenic variants, and ROHs. Patients who had a normal CMA interpretation were excluded.

Data Collection

Data were collected by medical chart reviews at the individual sites. Demographic data included sex, ethnicity, and race. Race and ethnicity were reported as included in the electronic medical record using hospital-determined categories, including "other," which was frequently based on patient self-report. These data were not always reported and could not be verified. A medical diagnosis list and corresponding *International Classification of Diseases, Ninth Revision/ International Classification of Diseases, Tenth Revision* (*ICD9/ICD10*) codes were recorded. The cytogenetic data elements collected from clinical CMA reports were previously described.[6] The echocardiography report (the earliest available complete echocardiogram) from each registry patient was sent to the Consortium's central hub at IUSM for systematic phenotyping and data entry. All data were stored in a REDCap database.[7]

Centralized CMA Data Processing

The CMA results were organized and annotated at the Consortium's central hub (IUSM) using the UCSC Genome Browser platform (https://genome.ucsc.edu). Genomic coordinates for CNV and ROH regions were recorded using the GRCh37 (hg19) genome assembly. The Lift Genome Annotations (ucsc.edu) tool was used to convert findings reported in other versions of the reference genome. The hgTables tool was utilized to generate a list of genes in each abnormal genomic interval based on the NCBI RefSeq Track.[8] Genes were annotated for association with human disease using the OMIM compendium (4341 genes), accessed on June 30, 2022. Genes were annotated for association with CHD using the manually curated CHDgene resource (http://chdgene.victorchang.edu.au; 139 CHD genes), accessed on June 30, 2022. Each gene was annotated for the likelihood that loss-of-function is not tolerated based on probability of loss-of-function intolerant (pLI) scores (9) downloaded from the Genome Aggregation Database (http://gnomad.broadinstitute.org), accessed on April 5, 2020. Genes were also annotated based on the level of RNA expression in developing mouse heart at embryonic day (e)14.5 or brain at e9.5, which were acquired from published data.[9]

Centralized Cardiac Phenotyping and Classification of CHD Types

The cardiac abnormalities documented in echocardiography reports were recorded for each registry patient. The cardiac phenotype definitions were developed from a modified classification approach utilized by the National Birth Defects Prevention Study,[10] as previously described.[6] The specific cardiac abnormalities were recorded as "Level 1" diagnoses. Each Level 1 diagnosis belongs within 1 broader category of CHD type ("Level 3" categories). For example, a Level 1 diagnosis of aortic valve stenosis is within the Level 3 category of left ventricular obstructive lesion (LVOTO). All Level 1 diagnoses and the corresponding Level 3 categories (Figure S1) were entered for each patient at the time of cardiac phenotyping. At the time of entry, patients can have >1 Level 1 diagnosis and >1 Level 3 category, constituting an inclusive approach to CHD classification. Additional cardiac phenotype information was obtained from the individual study sites when initial echocardiography reports were incomplete or inconclusive. Two investigators (LRH or BJL) performed all cardiac data entry in order to ensure consistency in phenotype extraction and classification. All cases with laterality defects, double outlet right ventricle, or pulmonary atresia, or other phenotypically complex cases, were reviewed by a board-certified pediatric cardiologist (BJL). For Level 1 diagnoses that are established compilations of lesions such as tetralogy of Fallot (TOF), rules were adopted in order to reduce redundancy in data entry, which is further explained in Data S1.

In addition to the prespecified assignment of a Level 3 CHD category for each Level 1 diagnosis at the time of entry, the Level 1 diagnosis information was also utilized to assign 1 overall CHD type to each patient. This postdata entry hierarchical classification was accomplished using a tiered structure similar to Oyen's prior use of the National Birth Defects Prevention Study classification system.[11] The hierarchy utilized for the present study is shown in Figure S2. In this method, each patient had 1 hierarchical category of CHD. For specific analyses, CHD was also sorted as those likely to create a univentricular versus biventricular physiology for the patient. CHD presumed to create univentricular physiology were the following: hypoplastic left heart syndrome (HLHS), tricuspid atresia, mitral atresia, single ventricle/double inlet left ventricle, and unbalanced complete atrioventricular septal defect (AVSD).

Statistical Analysis

Genetic Classification of the Cohort

The cohort was subgrouped genetically into 3 major groups according to CMA abnormality (Figure 1). Genetic Group I was defined as patients with a CMA abnormality for 1 of the known, well-characterized CHD-associated genomic disorders as defined by Consortium investigators (listed in Table 1). Group II consisted of patients who had 1 or more CNVs that do not cause a Group I disorder. Group II patients were

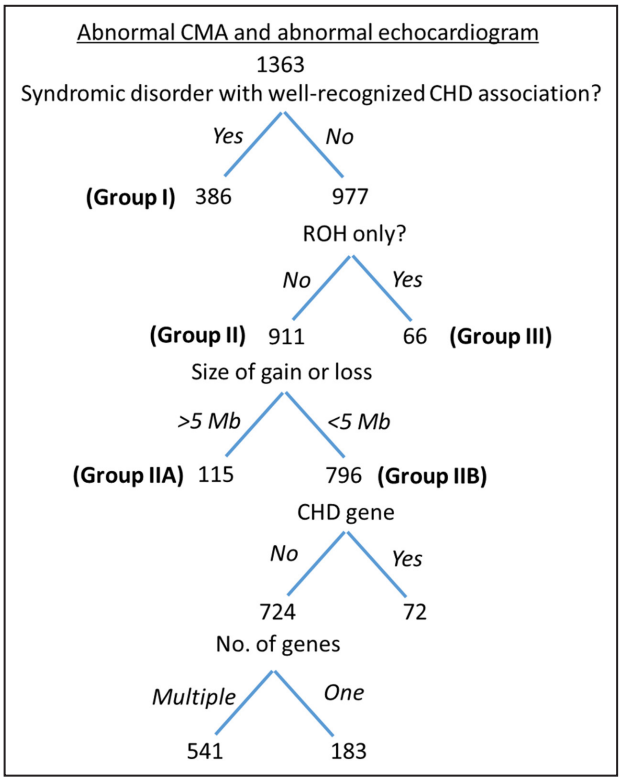

Figure 1. Genetic groups in registry patients with CMA abnormalities.
Group I includes patients with CMA abnormalities for one of the known and well-characterized genomic disorders associated with congenital heart disease (CHD) (Table 1). Group II includes patients with 1 or more CNVs that do not cause a Group I disorder. Group III includes patients who had only regions of homozygosity reported (ROH). Group IIA includes patients with 1 or more CNVs larger than 5 million base pairs (Mb). Group IIB includes patients who only had submicroscopic CNV(s) smaller than 5 Mb. CMA indicates chromosomal microarray; and CNVs, copy number variants.

further subgrouped by whether the patient had 1 or more CNVs >5 million base pairs (Mb) (Group IIA) or only had submicroscopic CNV(s) <5 Mb in size (Group IIB). This is the approximate size threshold for detection of CNVs by standard chromosome analysis. Thus, the CNVs >5 Mb category included aneuploidies and larger structural rearrangements. Meanwhile, patients who had 1 or multiple ROHs as their only CMA abnormality were placed into Group III. ROH inclusion was based on individual clinical laboratories reporting practices and all ROH reported in Group III patients are >1 Mb.

Cardiac Phenotype Comparisons

The frequencies of CHD phenotypes were analyzed graphically with heatmaps generated using the gplots package in R (version 4.1.2). Pearson's χ^2 test (all expected cell counts in contingency table ≥5) or Fisher's exact test (any expected cell counts <5) was used to test for significant differences in the proportions of hierarchical CHD phenotypes, and a Bonferroni correction was utilized for determining statistical significance when performing multiple comparisons. To compare the overall distribution of hierarchical CHD phenotypes between groups, Fisher's exact test utilized a Monte Carlo simulation of 1×10^6 replicates. Estimated and Bonferroni-corrected P values <0.05 were considered statistically significant. Statistical analyses were performed using functions in base R (version 4.1.2).

Table 1. Genetic Group I Patients With Genomic Disorders

Genomic disorder	Phenotype MIM number* [PMID]	Number of patients in registry (N=386)
Microdeletion or microduplication		
22q11.2 deletion syndrome	188400; 192430	166†
7q11.23 deletion (Williams syndrome)	194050	43
1q21.1 duplication	612475	14
22q11.2 duplication	608363	11
1q21.1 deletion	612474	10
1p36 deletion syndrome	607872; 619343	10
8p23.1 deletion	[20969981]	12
8p23.1 duplication	[17940555]	9
16p11.2 deletion	611913	9
22q11.2 distal deletion	611867	7
11q terminal deletion (Jacobsen syndrome)	147791	6
16p11.2 duplication	614671	6
7q11.23 duplication	609757	4
Aneuploidy or large chromosomal deletion		
Trisomy 21 (Down syndrome)	NA	40†
Monosomy X (Turner syndrome)	NA	22
Trisomy 18 (Edward's syndrome)	NA	6
Trisomy 13 (Patau syndrome)	NA	5
Monosomy 5p (Cri du chat syndrome)	123450	4†
Monosomy 4p (Wolf-Hirschhorn syndrome)	194190	2

NA indicates not available; and PMID, PubMed reference number.
*Acquired from Online Catalog of Human Genes and Genetic Disorders (OMIM); updated on June 20, 2022.
†Count includes a patient with additional copy number variant (CNV) that is associated with a separate Group I disorder. One patient with 22q11.2 deletion and 1 patient with monosomy 5p also had 8p23.1 duplication. One patient with Down syndrome also had 22q11.2 duplication. Further analysis focused on the primary CNV.

Bootstrapping Enrichment Analysis of Curated CHD Genes in CNVs

The list of 139 curated CHD genes was procured as described above. Patients with only submicroscopic CNVs (Genetic Group IIB) were selected for CHD gene

enrichment analysis because aneuploidies and large structural chromosomal abnormalities that include many genes would limit sensitivity and specificity. The frequency of CHD genes observed in registry CNVs was calculated. The probability for the observed frequency of CHD genes was empirically tested; 10 000 random gene lists, each containing 139 genes, were generated using the Random Gene Set Generator application (www.molbiotools.com). For each random gene list, the frequency that the genes in the list were observed in Group IIB CNVs was calculated. An empirical *P* value estimating the likelihood of CHD gene frequency in Group IIB CNVs was calculated as the fraction of the 10 000 random gene lists that contained genes that were more frequent in CNVs than was observed for CHD genes.

Gene Set Enrichment Analyses

In silico gene set enrichment analysis for annotated pathways, gene ontology biological processes, and disease annotation were completed using ToppFun.[12,13] The threshold for statistical significance was defined by *P* value divided by the number of enrichment categories tested <0.05 (Bonferroni correction).

RESULTS

Overview of Cytogenomics of Cardiovascular Malformations Registry Patients

This study included 1363 registry patients who had an abnormal echocardiogram and at least 1 abnormal (uncertain or pathogenic) finding on CMA identified as part of their clinical care. Basic demographics are shown in Table S1. Proportions of male and female patients were similar. There was a modest underrepresentation of Hispanic or Latino, Asian, and Native American patients relative to the United States population, but the representations were reflective of the regional demographics of the centers participating in the study.

The clinically reported CNVs included large cytogenetic abnormalities, such as aneuploidies and partial chromosome aneusomies resulting from unbalanced translocations, as well as submicroscopic deletions and duplications. ROHs were also reported. Each patient was classified into 1 of 3 genetic groups based on CMA findings (Figure 1). Genetic Group I included 386 patients (28%) with 19 genomic disorders that have well-recognized association with CHD (Table 1). Group II included 911 patients (67%) with CNVs that have less recognized or no prior known association with CHD. Group III included 66 patients (5%) with ROH only. In total, there were 919 copy number losses, 771 copy number gains, and 730 ROH in the registry. More than 1 CNV was reported in 332 patients, who constituted 24% of Group I and 26% of Group II patients (Table S2). The types of CHD in registry patients are summarized in hierarchical CHD categories in Figure 2. Conotruncal defect (CTD) was the most frequent hierarchical CHD category, comprising 29% of patients. Overall, the registry contains patients with genomic imbalances of variable size and variable level of prior evidence for CHD association. There was a relatively higher frequency of CTD, LVOTO, and other complex CHDs, and lower frequency of isolated septal defects than general CHD populations.[11,14,15]

Genotype–Phenotype Analysis of Registry Patients in Genetic Group I

The CHD categories that are well known to be associated with specific Group I genomic disorders were well represented in the hierarchical classification system (Figure 3A), in which each patient has a single CHD category assigned, and with inclusive classification, which allows for multiple Level 3 CHD categories per patient (Figure 3B). For example, CTD was frequent in 22q11.2 deletion, AVSD in trisomy 21, LVOTO in Turner or Jacobsen syndrome, arteriopathy in Williams syndrome, and aortopathy in 7q11.23 duplication. Inclusive classification (Figure 3B) underscores pleiotropy of these genomic disorders by showing that most patients (262 of 386, 68%) presented with >1 Level 3 CHD category. These included uncharacteristic types of CHD. For example, 3 (1.8%) patients with 22q11.2 deletion syndrome had the uncharacteristic CHD category of AVSD. More strikingly, 7 (16%) patients with Williams syndrome (7q11.23 deletion) had uncharacteristic CHD including CTD, AVSD, and anomalous pulmonary venous return (APVR; Table 2), which could not be attributed to variable sizes in 7q11.23 deletion or concomitant CNV or ROH.

Chromosome 1q21.1 Duplication and Deletion Syndromes

There were 24 patients with 1q21.1 duplications (N=14) or reciprocal deletions (N=10). Hierarchical classification identified that CTD, septal defect, and LVOTO were frequent and similarly represented between duplications and deletions (Figure 3A). Inclusive classification (Figure 3B) identified a relatively higher frequency of LVOTO in 1q21.1 duplications. APVR was associated with these syndromes, which was not evident in the hierarchical approach. Figure 4 shows the Level 1 diagnoses for CTD (4A) and LVOTO (4B). The CTD of TOF was enriched in 1q21.1 duplication (29%) and absent in 1q21.1 deletions (Figure 4A). The LVOTO of HLHS was present in 3 (21%) patients with 1q21.1 duplications versus none with 1q21.1 deletions (Figure 4B).

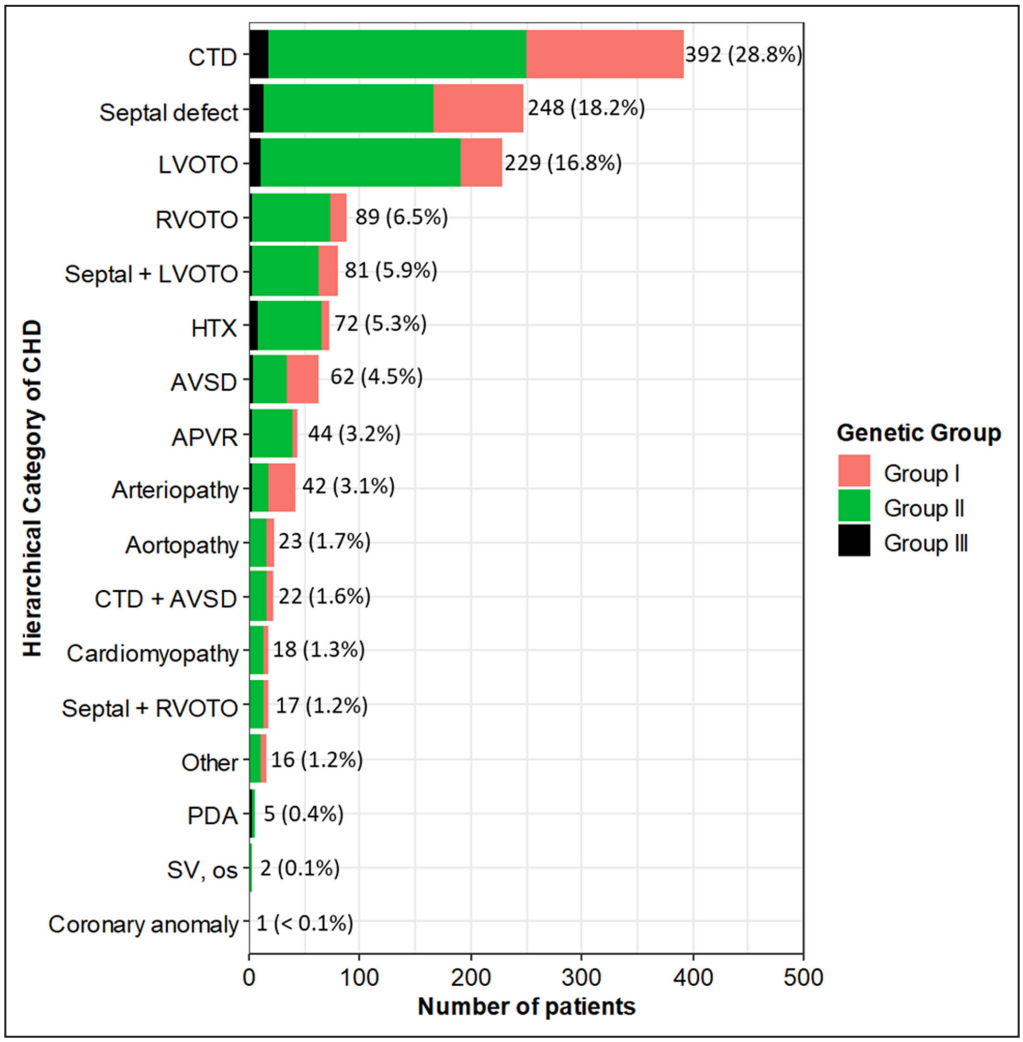

Figure 2. Hierarchical categories of CHD in registry patients.
The frequency of each CHD category is shown for all patients (N [% of total]) and for patients within Genetic Group I, II, or III. APVR indicates anomalous pulmonary venous return; AVSD, atrioventricular septal defect; CHD, congenital heart disease; CTD, conotruncal defect; HTX, heterotaxy; LVOTO, left ventricular obstructive lesion; PDA, patent ductus arteriosus; RVOTO, right ventricular obstructive lesion; and SV, os, single ventricle otherwise specified.

Chromosome 8p23.1 Deletions and Duplications

A total of 21 patients had 8p23.1 deletions (N=12) or 8p23.1 duplications (N=9). Hierarchical category of AVSD or AVSD+CTD was frequent in 8p23.1 deletion (50%) (Figure 3A). Inspecting Level 3 categories with inclusive classification identified the frequent combination of AVSD with LVOTO or right ventricular obstructive lesion in 8p23.1 deletions (4 of 6 patients with AVSD; all with semilunar valve defects). Inclusive classification identified a higher frequency of LVOTO (78%) in 8p23.1 duplication versus 8p23.1 deletion (33%). Whereas inclusive description was required to identify APVR in 1q21.1 disorders, conversely, it confirmed an absence of APVR in registry 8p23.1 deletions and duplications (Figure 3B). While LVOTO was frequent in 8p23.1 duplications or deletions, no patients had HLHS (Figure 4B).

Chromosome 16p11.2 Deletion and Duplication Syndromes

The notable difference between patients with 16p11.2 deletions (N=9) or duplications (N=6) was increased frequency of CTD in deletions (44%) (Figure 3A and 3B).

Chromosome 22q11.2 Duplication Syndrome

The 11 registry patients with 22q11.2 duplication (reciprocal to classic 22q11.2 deletion) had a relatively low frequency of CTD (Figure 3A and 3B). Six patients had right ventricular obstructive lesion or LVOTO independent of

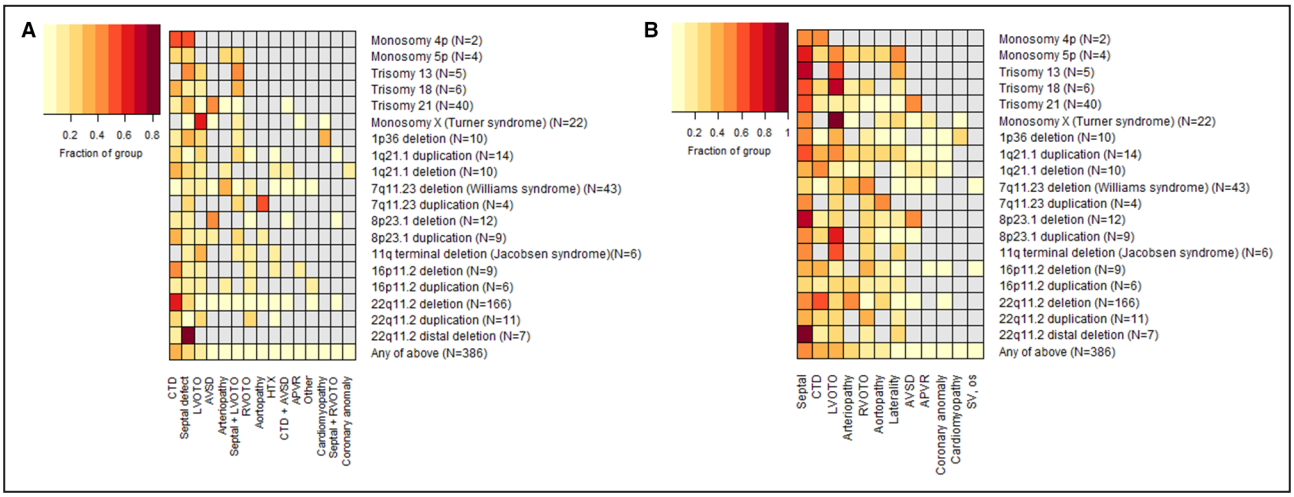

Figure 3. Distributions of CHD categories in Group I patients (N=386).
A, Results of hierarchical CHD classification method. **B**, Results of inclusive classification. Box colors indicate the fraction of patients within a genomic disorder that have the CHD type. Gray squares indicate that no patients had the CHD type. APVR indicates anomalous pulmonary venous return; AVSD, atrioventricular septal defect; CHD, congenital heart disease; CTD, conotruncal defect; HTX, heterotaxy; LVOTO, left ventricular obstructive lesion; RVOTO, right ventricular obstructive lesion; and SV, os, single ventricle otherwise specified.

CTD, all of which constituted significant atrioventricular valve involvement: 2 patients with Ebstein's anomaly, 2 with HLHS, 1 with tricuspid atresia, and 1 with a thickened, dysplastic tricuspid valve and thickened mitral valve. APVR was absent from all 184 patients with Group I disorders involving the 22q11.2 region (Figure 3B).

Chromosome 22q11.2 Distal Deletions

Hierarchical CHD of septal defect was observed in 6 of the 7 patients with distal 22q11.2 deletion (Figure 3A), which included muscular or perimembranous ventricular septal defect in 4 patients and secundum atrial septal defect without ventricular septal defect in 2. This indicates an enrichment of isolated septal defects compared with other 22q11.2 disorders.

Aggregating relatively large numbers of patients and applying consistent and detailed CHD phenotyping portrays the similarities and differences in CHD phenotypes among genomic disorders caused by CNVs at the 1q21.1, 8p23.1, 16p11.2, and 22q11.2 loci, which are summarized in Table 3.

Table 2. Novel or Rare Cardiac Abnormalities Present in 7 of 43 Registry Patients With Williams Syndrome (7q11.23 Deletion)

Patient	Novel or rare Level 3 CHD categories	Level 1 CHD diagnoses
1	CTD; laterality	d-TGA; CoA; aortic valve stenosis or hypoplasia; tricuspid atresia with VSD; RV hypoplasia; dextrocardia; persistent left SVC†
2*	CTD	Conoventricular VSD; CoA; peripheral PSs or hypoplasia; secundum ASD.†
3	CTD; SV	DORV with d-malposed great vessels; ASD, nos; SV, nos†
4	AVSD	Complete AVSD; peripheral PS or hypoplasia; secundum ASD†
5	CTD; AVSD	DORV; CoA; aortic valve stenosis or hypoplasia; subaortic stenosis or narrowing; mitral atresia; LV hypoplasia; common atrium; AVSD, os: common atrioventricular valve draining to RV†
6	APVR	PAPVR; peripheral PS or hypoplasia†
7	APVR	TAPVR; aortic valve stenosis or hypoplasia; muscular VSD; secundum ASD†

ASD indicates atrial septal defect; AVSD, atrioventricular septal defect; CHD, congenital heart defect; CTD, conotruncal defect; CoA, coarctation of the aorta; DORV, double outlet right ventricle; d-TGA, d-transposition of the great arteries; IVC, inferior vena cava; LV, left ventricle; nos, not otherwise specified; os, otherwise specified; PAPVR, partial anomalous pulmonary venous return; PS, pulmonary stenosis; RV, right ventricle; SV, single ventricle; SVC, superior vena cava; TAPVR, total anomalous pulmonary venous return; and VSD, ventricular septal defect.

*Patient 2 had concurrent 29.4-kb duplication at 3p25.1 (chr3:15097391-15126800x3).

†Level 1 CHD dianoses that belong to the novel or rare Level 3 categories.

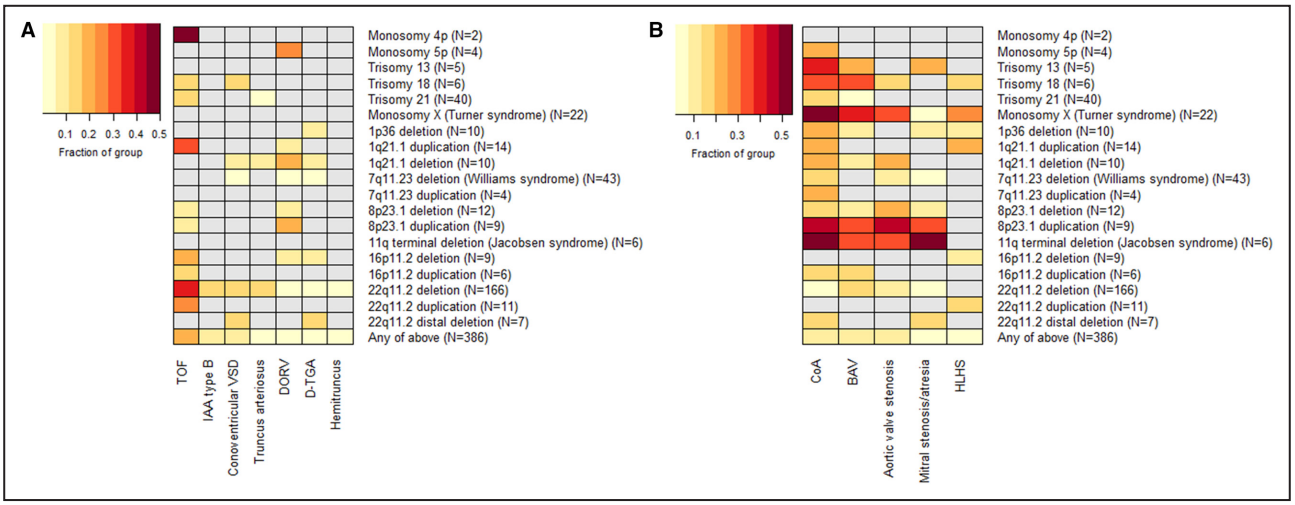

Figure 4. Distributions of Level 1 CHD diagnoses in Genetic Group I patients who have Level 3 diagnoses of a conotruncal defect (CTD) (A) and left ventricular obstructive lesion (LVOTO) (B).
BAV indicates bicuspid aortic valve; CHD, congenital heart disease; CoA, coarctation of the aorta; D-TGA, d-transposition of the great arteries; DORV, double outlet right ventricle; HLHS, hypoplastic left heart syndrome; IAA, interrupted aortic arch; TOF, tetralogy of Fallot; and VSD, ventricular septal defect.

Genetic Analysis of Registry Patients in Genetic Group II

Large Versus Submicroscopic CNV

Group II included 911 patients with CNVs that have rare or no known prior association with CHD. Reported imbalances included large CNVs (>5 Mb; N=115 patients; Group IIA) and submicroscopic CNVs (<5 Mb; N=796 patients; Group IIB; Figure 1). In Group IIA, the median CNV size was 15.504 (interquartile range: 8.833–35.290) Mb and the median total number of genes in CNVs per patient was 162 (interquartile range: 89–302). In Group IIB, the median CNV size was 0.308 (interquartile range: 0.139–0.671) Mb and the median total number of genes within CNVs per patient was 3 (interquartile range: 1–9). Furthermore, in Group IIB, 72 patients (9%) had CNV(s) containing curated CHD genes, 183 patients (23%) had CNV that involved only 1 gene (non-CHD gene), and 541 (68%) patients had CNV(s) comprising multiple genes, of which none is a curated CHD gene.

Submicroscopic CNVs Involving CHD Genes

There were a total of 1008 CNVs among the 796 Group IIB patients. CHD genes were present in 73/1008 (7.2%) CNVs, including 42 gains and 31 losses. In total, 34 CHD genes were located in these CNVs (Table 4). Bootstrapping analysis determined that CNVs in Group IIB were significantly enriched for CHD genes (P=0.002; Figure S3).

Genotype–Phenotype Analysis of Registry Patients in Genetic Group II

Comparison of CHD Phenotypes Between Group IIA and Group IIB Patients

The proportions of hierarchical CHD categories were significantly different between Group IIA and Group IIB (P<0.0001). The clearest differences were (1) an increased frequency of septal defect in Group IIA and (2) increased frequencies of CTD, LVOTO, and heterotaxy in Group IIB (Table 5). Indeed, only 1 Group IIA patient had heterotaxy. CHDs that were presumed to create univentricular physiology were significantly more frequent in Group IIB (17.5%) compared with Group IIA (3.5%). Patients with large CNVs had less complex CHD than those with submicroscopic CNVs.

Analysis of CHD Phenotypes in Group IIA and in Subclasses of Group IIB

The proportions of hierarchical CHD categories for Group IIA and subclasses of Group IIB are shown in Figure 5. Stark differences in the predominant CHD category (darker red shading) are evident between subgroups. In patients with monogenic CNV (N=183 patients), the categories of CTD (N=60, 33%) and heterotaxy (N=14, 8%) were frequent and septal defect (N=21, 11.5%) was infrequent compared with other subgroups (counts and fractions are specified in Table S3). Monogenic CNVs comprised 20% of Group II cases overall, but accounted for 26% of CTD, 25% of heterotaxy, and 31% of CTD+AVSD cases in the

Table 3. Summary of Cardiac Phenotype Findings for Copy Number Variants at 1q21.1, 8p23.1, 16p11.2, and 22q11.2 loci in Genetic Group I

Cardiac classification method	Hierarchical		Inclusive		
Genomic locus (number of patients)	Enriched in deletions	Enriched in duplications	Enriched in deletions: Level 3 diagnoses (Level 1 details)	Enriched in duplications: Level 3 diagnoses (Level 1 details)	Absent Level 3 diagnoses
1q21.1 (24)	CTD in 3 of 10 patients	LVOTO or Septal+LVOTO in 6 of 14 patients	CTD in 5 of 10 patients (none with TOF)	CTD in 5 of 14 patients (four with TOF) LVOTO in 7 of 14 patients (CoA or HLHS only)	Most categories were represented
8p23.1 (21)	AVSD in 5 of 12 patients	CTD in 3 of 9 patients	AVSD combined with LVOTO or RVOTO in 4 of 12 patients Septal defect in 10 of 12 patients	LVOTO in 7 of 9 patients (none with HLHS)	APVR
16p11.2 (15)	CTD in 4 of 9 patients	None	None	None	AVSD
22q11.2 duplication (11)	NA	LVOTO; RVOTO; CTD (each in 3 of 11 patients)	NA	LVOTO and RVOTO (mitral or tricuspid valve* anomaly in 6 of 11 patients)	AVSD APVR
22q11.2 distal deletion (7)	Septal in 6 of 7 patients	NA	Septal defect in 6 of 7 patients	NA	AVSD APVR

APVR indicates anomalous pulmonary venous return; AVSD, atrioventricular septal defect; CoA, coarctation of the aorta; CTD, conotruncal defect; HLHS, hypoplastic left heart syndrome; LVOTO, left ventricular obstructive lesion; NA, not applicable; RVOTO, right ventricular obstructive lesion; and TOF, tetralogy of Fallot.

*Tricuspid valve anomalies included 2 patients with isolated Ebstein's anomaly, 1 patient with tricuspid atresia, and 1 patient with a dysplastic tricuspid valve.

group. In contrast, the categories of LVOTO, Septal, or Septal+LVOTO were markedly enriched in patients with CHD gene containing CNVs, which together accounted for 60% of the patients with CHD gene CNVs in Group IIB. LVOTO was associated with the smooth muscle myosin gene *MYH11* (7 of 12 patients) and the vascular signaling CHD gene *NOTCH1* (3 of 6 patients), as well as other biological groups in Table 4.

Among Group IIB patients, presumed univentricular physiology was less frequent in patients with CHD gene CNVs (11%) compared with monogenic CNVs (20%) or other CNVs (18%; Table S3). Monogenic CNVs accounted for 25% of all Group II patients with presumed univentricular physiology. The marked phenotypic differences between monogenic CNVs and CHD gene CNVs suggests differences in the biological roles of the genes they comprise.

Submicroscopic CNVs Involving a Single Non-CHD Gene

Group IIB monogenic CNVs included 107 copy number losses and 76 copy number gains that involved a total of 141 distinct genes, annotated in Data S2. Twenty-three genes recurred in monogenic CNVs in the registry. Forty-six genes (33%) are associated with human disease in OMIM. Among the 111 monogenic CNV genes that were included in published embryonic mouse RNA sequencing data,[9] 37 genes (33%) had expression level ranking above the 50th percentile of all mouse genes in embryonic day (e)14.5 heart. Among 114 genes with Genome Aggregation Database pLI scores available, 30 genes (26%) had a pLI score >0.5. These annotations were used to prioritize candidate genes.

We identified 9 monogenic CNV genes that had e14.5 mouse heart expression level above the 50th percentile and pLI score above 0.5: *AP3B1, AUTS2*, *DMD*, *MAGT1*, *MBD5*, *PDE4D*, *PTPRM*, *TCF12*, and *TEAD1*. Reciprocal deletions or duplications involving *MBD5*, identified here in 4 patients with monogenic CNVs, cause a rare disorder (Chromosome 2q23.1 Deletion or Duplication Syndrome; MIM #156200) also known as *MBD5*-Associated Neurodevelopmental Disorder that has reported CHD association at very low penetrance.[16] The CHD in 2 registry patients included CTD (Table S4), a CHD phenotype that was not previously reported.[16] The other genes may be considered candidate causal genes.

Enrichment Analysis of Genes in Monogenic CNVs in Genetic Group II

Pathway analysis of 141 monogenic CNV genes identified enrichment for the neuronal system pathway (Reactome; Stable Identifier: R-HSA-112316). Eleven

Table 4. Curated CHD Genes (N=139) Grouped by their Biological Role and The Number of Occurrences in Group IIB

Chromatin regulation (19)	Cytoskeleton (13)	Vascular signaling (7)	RAS-MAPK (14)	TFs (5)	Transcription regulation (3)	TGF-beta pathway/ superfamily (2)	Left right pattern (1)	Cohesin (2)	RNA splicing or processing (3)	ECM (0)	Cilia (4)	PG or GAG (0)	NAD bio-synthesis (0)	GPI bio-synthesis (2)	Ubiqui-tination (0)	Other* (1)
EHMT1 (5)‡	*MYH11* (12)‡	*NOTCH1* (6)‡	*LZTR1* (9)‡	*FOXC1* (1)‡	*CITED2* (2)‡	*ACVR1* (1)‡	*PITX2* (1)‡	*ESCO2* (1)‡	*TXNL4A* (2)‡	*ADAMTS10*	*DNAH5* (3)‡	*B3GAT3*	*HAAO*	*PIGL* (2)‡	*TRAF7*	*KDR* (1)‡
KANSL1 (5)‡	*FLNA* (1)‡	*JAG1* (1)‡	*NF1* (2)‡	*GATA4* (1)†	*MED13L* (1)‡	*TAB2* (1)‡	*HAND1*	*NIPBL* (1)‡	*HNRNPK* (1)‡	*ELN*	*DNAH11* (1)‡	*CHST14*	*KYNU*	*PIGV*	*UBR1*	*ABL1*
CREBBP (3)‡	*ACTC1*	*DLL4*	*RAF1* (2)‡	*GATA5* (1)‡	*MED12*	*ACVR2B*	*HAND2*	*RAD21*	*SF3B4* (1)‡	*FBN1*	*EVC*	*GPC3*				*CDK13*
SMARCB1 (3)‡	*DOCK6*	*FLT4*	*PRKD1* (1)‡	*NKX2-5* (1)‡	*MEIS2*	*BMPR2*	*INVS*	*SMC1A*	*EFTUD2*	*LTBP2*	*EVC2*					*CRELD1*
ARID1B (1)‡	*MYBPC3*	*NOTCH2*	*BRAF*	*TBX20* (1)‡	*NSD1*	*GDF1*	*NODAL*	*SMC3*	*NONO*	*TLL1*	*NPHP3*					*FGFR2*
KMT2D (1)‡	*MYH6*		*CFC1*	*AFF4*	*PBX1*	*MAP3K7*	*ZIC3*		*RBFOX2*		*NPHP4*					*GJA1*
RERE (1)‡	*MYH7*		*HRAS*	*FOXC2*		*SMAD2*			*SMG9*		*PKD1L1*					*KDR*
ANKRD11			*KRAS*	*FOXH1*		*SMAD3*			*SON*							NUP188
ARID1A			*MAP2K1*	*FOXP1*		*SMAD4*										*STRA6*
BCOR			*MAP2K2*	*GATA6*		*SMAD6*										*TMEM260*
CHD4			*NRAS*	*GLI3*		*TGFBR1*										*WASHC5*
CHD7			*PTPN11*	*MESP1*		*TGFBR2*										
EP300			*RAB23*	*NR2F2*												
HDAC8			*RIT1*	*NKX2-6*												
KAT6A			*SHOC2*	*SALL1*												
KAT6B			*SOS1*	*SALL4*												
KDM6A				*TBX1*												
KMT2A				*TBX5*												
PRDM6				*TFAP2B*												
SETD5				*ZEB2*												
SMARCA4				*ZFPM2*												
SMARCE1																

There were 4 patients with CNV involving both *NOTCH1* and *EHMT1*. One patient had a CNV that included *FLNA* and a separate CNV that included *MYH11*. CHD indicates congenital heart defect; CNV, copy number variant; ECM, extracellular matrix; GAG, glycosaminoglycan; GPI, glycophosphatidylinositol; NAD, nicotinamide adenine dinucleotide; PG, proteoglycan; TF, transcription factor; and TGF, transforming growth factor.

*Other CHD gene biological roles are: *ABL1*: nonreceptor protein tyrosine kinase; *CDK13*: cell cycle regulation; *CRELD1*: member of subfamily of epidermal growth factor-related proteins; *FGFR2*: fibroblast growth factor signaling; *GJA1*: gap junctions; *KDR*: vascular endothelial growth factor receptor; *NUP188*: nuclear pore complex; *STRA6*: retinol transporter; *TMEM260*: not defined; *WASHC5*: endosomal processes.

†One patient with a small 245-kb duplication in 8p23.1 containing the CHD gene *GATA4* was included in Group IIB.

‡Genes present in CNV intervals in this study.

Table 5. Frequencies of Hierarchical Congenital Heart Disease Categories and Presumed Univentricular Physiology in Genetic Group II and Subgroups IIA and IIB

CHD category	All Group II	Group IIA: CNV>5 Mb	Group IIB: CNV<5 Mb	*P* value
CTD	25.5% (232/911)	16.5% (19/115)	26.8% (213/796)	0.019
LVOTO	19.8% (180/911)	13.0% (15/115)	20.7% (165/796)	0.053
Septal defect	16.7% (152/911)	30.4% (35/115)	14.7% (117/796)	0.000023*
RVOTO	7.9% (72/911)	7.0% (8/115)	8.0% (64/796)	0.69
Septal+LVOTO	6.6% (60/911)	10.4% (12/115)	6.0% (48/796)	0.075
HTX	6.3% (57/911)	0.9% (1/115)	7.0% (56/796)	0.0063
APVR	4.1% (37/911)	1.7% (2/115)	4.4% (35/796)	0.21
AVSD	3.3% (30/911)	6.1% (7/115)	2.9% (23/796)	0.089
Aortopathy	1.8% (16/911)	2.6% (3/115)	1.6% (13/796)	0.44
Arteriopathy	1.8% (16/911)	1.7% (2/115)	1.8% (14/796)	1.0
CTD+AVSD	1.8% (16/911)	0	2.0% (16/796)	0.25
Septal+RVOTO	1.5% (14/911)	2.6% (3/115)	1.4% (11/796)	0.40
Univentricular physiology	15.7% (143/911)	3.5% (4/115)	17.5% (139/796)	0.00016*

Hierarchical CHD categories that were present in <1.5% of all Group II patients are not shown. APVR indicates anomalous pulmonary venous return; AVSD, atrioventricular septal defect; CHD, congenital heart disease; CNV, copy number variant; CTD, conotruncal defect; HTX, heterotaxy; LVOTO, left ventricular obstructive lesion; and RVOTO, right ventricular obstructive lesion. specified.

*CHD categories with *P* value that reaches statistical significance after multiple hypothesis correction (α=0.0038).

monogenic CNV genes are in this pathway (Table 6; Table S5), and these accounted for 18 patients (10%) with monogenic CNVs. Eleven of these 18 patients (61%) had a hierarchical CHD of CTD, including 9 with TOF (50%). Eight of these 11 genes had similar expression levels between embryonic mouse heart and brain. The second significantly enriched pathway in monogenic CNV genes was Phosphodiesterase in Neuronal Function (MSigDB; BIOCARTA), which included 3 Neuronal System pathway genes (*ADCY2*, *ADCY8*, and *CHRNA7*) and the genes *PDE4D* and *GUCY1A2* (Table S5). In comparison to these enrichments, the 34 CHD genes identified in CNVs of Group IIB patients were not enriched for neuronal pathways and instead were enriched for heart development, cardiac progenitor differentiation, NOTCH signaling, and chromatin organization (Table S6). The results identify biological pathway differences, which appear to correlate with the overarching pattern of divergence in CHD phenotypes that was observed between these CNV subgroups (Figure 5).

Further enrichment analysis of monogenic CNV genes for gene ontology Biological Processes and Disease annotations identified enrichment for cell–cell adhesion (20 genes) and neurodevelopmental diseases including Intellectual Disability and Pervasive Development Disorder (Table S5). Eight of the 9 monogenic CNV genes that were initially prioritized for having relatively high embryonic heart expression and pLI >0.5 were found in at least 1 of the enriched biological and disease terms: *AP3B1, AUTS2*, *DMD*, *MAGT1*, *MBD5*, *PDE4D*, *PTPRM*, and *TCF12*.

Three monogenic CNV genes in the Neuronal System pathway recurred in Group IIB patients, display embryonic heart expression, and have pLI >0.99: *NRXN3* (N=3, including twin brothers with TOF), *ADCY2* (N=2), and *HCN1* (N=2). In all such cases the CNV was either a deletion or a duplication that terminated within the gene. These may present top candidate genes for CHD causality that warrant further investigation.

DISCUSSION

With an emphasis on the role of cardiac phenotyping in a large, multicenter cohort of patients with abnormal clinical CMA results, this study expands the CHD phenotypes to consider in established genomic disorders, identifies novel candidate genes in CNVs, and reveals a stratification in CHD categories between subclasses of CNVs.

The first organizing principle of cardiac phenotyping in the registry was that collection of detailed raw cardiac data using a consistent taxonomy would be required to create precise CHD descriptions, which has been inconsistent in genetic literature, and to draw comparison between different genetic causes. The second was that because genetic causes of CHD are heterogeneous and have variable expressivity, the structure of cardiac phenotype data should permit flexibility in how lesions are combined, for which there is currently no consensus, and depends upon the purpose and type of analysis. In this study, hierarchical classification of each patient into 1 CHD category facilitated broad descriptions and genotype–phenotype analyses between genetic subgroups. There were also examples where inclusive classification was important for its sensitivity, such as in describing prevalence of APVR in 1q21.1

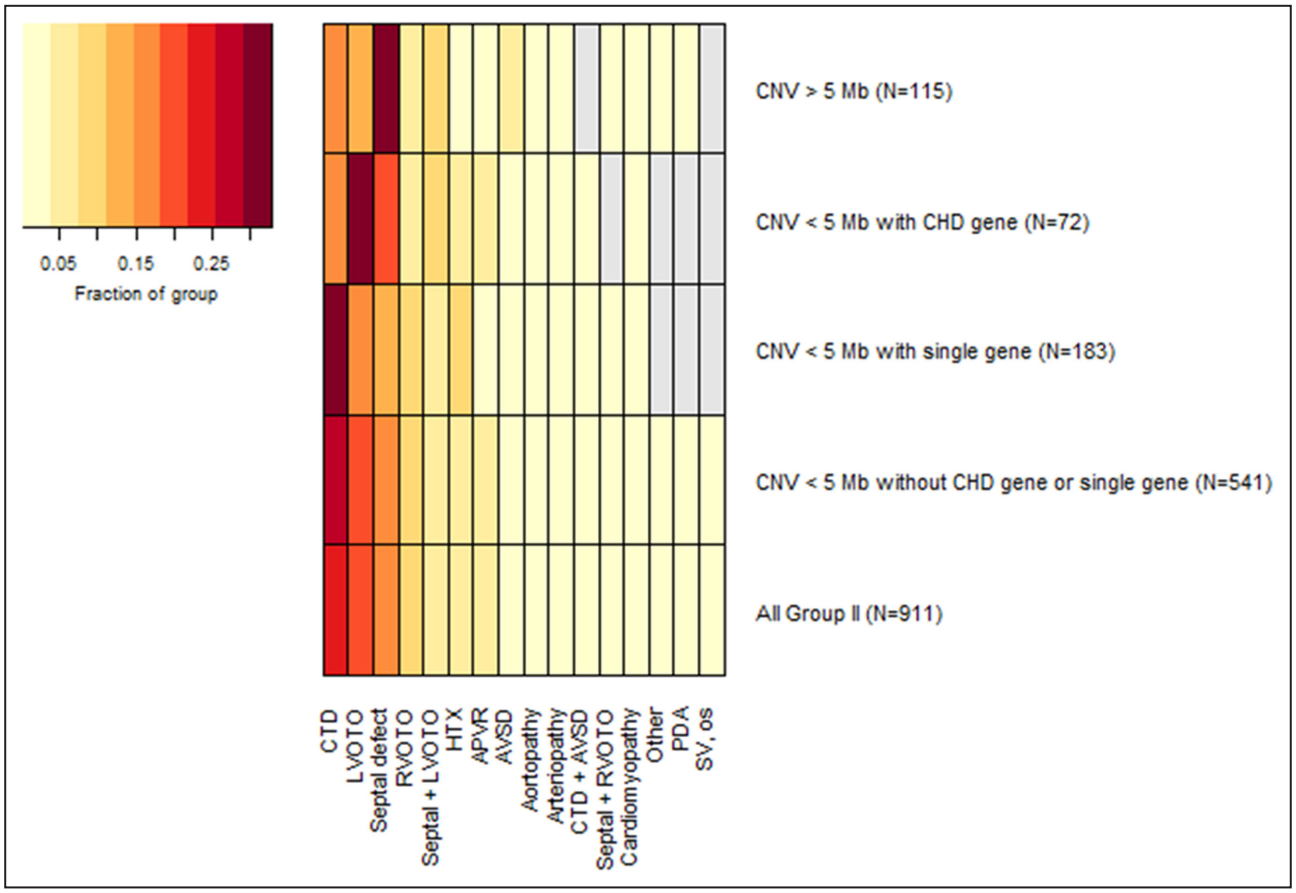

Figure 5. Distributions of hierarchical CHD categories in Group II patients (N=911).
Group IIA includes patients with CNV>5 Mb. Group IIB is further subdivided into patients with CNV involving a CHD gene; CNVs involving a single non-CHD gene; and CNVs that have multiple genes, none of which are CHD genes. APVR indicates anomalous pulmonary venous return; AVSD, atrioventricular septal defect; CHD, congenital heart disease; CNV, copy number variant; CTD, conotruncal defect; HTX, heterotaxy; LVOTO, left ventricular obstructive lesion; Mb, million base pairs; PDA, patent ductus arteriosus; RVOTO, right ventricular obstructive lesion; and SV, os, single ventricle otherwise specified.

imbalances and, conversely, determining that APVR was absent in registry patients with disorders involving 8p23.1 and 22q11.2. Inclusive classification facilitated identifying the frequent combination of AVSD with semilunar valve abnormalities in patients with 8p23.1 deletions, a corollary to the role of the involved gene *GATA4* in both atrioventricular septation and semilunar valve development.[17,18]

Cardiac Phenotypes in More Recently Described Genomic Disorders

The registry identifies well-known cardiac phenotypes for established CHD genomic disorders. Registry data also firmly establish associations that were suggested in previous reports, including enrichments of TOF in 1q21.1 duplications[19–22] and non-TOF CTD in 1q21.1 deletions,[19,23,24] and a notable frequency of HLHS in 22q11.2 duplications.[25] These associations are important to improve the differential diagnosis in infants with specific categories of CHD and subtle or no dysmorphic features.

Novel Cardiac Phenotypes in Williams Syndrome, 1q21.1 Duplication, and 22q11.2 Duplication

The registry has also led to novel phenotype observations in genomic disorders that have well-established CHD association. Novel phenotypes in patients with Williams syndrome included double outlet right ventricle, dextro-transposition of the great arteries, tricuspid atresia, unbalanced common atrioventricular valve, common atrium, and partial anomalous pulmonary venous return. The high prevalence of LVOTO (and specifically coarctation of the aorta or HLHS) in 1q21.1 duplications (50%) is novel; literature review identified only 1 prior case description of LVOTO.[26] The enrichment of atrioventricular valve abnormality in patients with 22q11.2 duplication is also novel. To our

Table 6. Genes in Monogenic Copy Number Variants That Belong to the Enriched Neuronal System Pathway

Genes	Gene name	Cytoband	No. of occurrences in monogenic CNVs (total/losses/gains)	RNA expression level in e14.5 mouse heart (percentile rank)*	RNA expression level in e9.5 mouse brain (percentile rank)*	pLI
ADCY2†	Adenylate cyclase 2†	5p15.31	1/1/0	28	31	0.999
ADCY8	Adenylate cyclase 8	8q24.22	1/0/1	19	23	0
CACNA2D3	Calcium voltage-gated channel auxiliary subunit alpha2delta3	3p21.1	2/0/2	24	31	0.578
CHRNA7	Cholinergic receptor nicotinic alpha 7 subunit	15q13.3	2/0/2	19	46	<0.001
DLG2	Discs large MAGUK scaffold protein	11q14.1	1/1/0	43	22	0.780
HCN1†	Hyperpolarization activated cyclic nucleotide gated potassium channel 1†	5p12	2/0/2	30	21	0.999
NRXN3†	Neurexin 3†	14q31.1	2/2/0	42	23	**1**
SLC1A1	Solute carrier family 1 member 1	9p24.2	1/1/0	42	29	<0.001
SLC22A2	Solute carrier family 22 member 2	6q25.3	1/1/0	0	14	0
SYT10	Synaptotagmin 10	12p11.1	3/0/3	0	17	<0.001
TSPAN7	Tetraspanin 7	Xp11.4	2/0/2	NR	NR	0.74577

CNVs indicates copy number variants; pLI, predicted intolerance to loss of function (gnomAD); and NR, not reported.
*RNA sequencing data of embryonic (e) day of mouse development published in Homsy et al.[9]
†Top candidate genes and gene names.

knowledge, this is the first report of isolated Ebstein's anomaly (present in 2 registry patients) and the first report of tricuspid atresia in 22q11.2 duplications. The tricuspid valve defects in registry patients may provide the first human CHD correlation with the hypoplastic right ventricle phenotype that was observed in mice with gain of function of *Tbx1*, a CHD gene within the human 22q11.2 locus.[27] Overall, the factors regulating phenotypic variability within genomic disorders, which may include environmental or other genetic factors, require additional study. The registry finds that many patients have multiple CMA abnormalities, which could contribute to phenotype heterogeneity.

Candidate Genes Identified in Monogenic CNVs

Analysis of monogenic CNVs identified novel candidate causal genes, which have a role in neuronal development and in the registry, were enriched for CTD phenotypes. The connection between heart and brain development aligns with results of exome analysis of a similar number of patients with CHD.[9] Neurocristopathies including 22q11.2 deletion and CHARGE syndrome are associated with neurodevelopmental abnormalities and enriched for CTD.[28] Cell adhesion, which was an enriched biological process in monogenic CNVs, is important for migration of cardiac neural crest cells to form the developing outflow tract[29] and therefore a potential mechanism of CTD development. The registry identifies *NRXN3*, *ADCY2*, and *HCN1* as particularly strong candidates. *NRXN3* encodes a neurexin protein that is important for cell adhesion. Its related gene *NRXN1*, which contained an intragenic deletion in 3 registry patients, has been previously associated with CHD.[30–32] *ADCY2* catalyzes cyclic adenosine monophosphate formation, and knockdown of its ortholog in zebrafish led to cardiac malformation.[33] *HCN1* encodes a cyclic adenosine monophosphate activated potassium/sodium channel and is highly expressed in the sinoatrial node.[34] The candidate genes identified in this study warrant further investigation.

Complexity of CHD and CNV Size

We observed an interesting separation of CHD categories when the CNVs were subclassified for size, involvement of a curated CHD gene, or involvement of a single gene. Patients with large (>5 Mb) CNVs in general presented less complex cardiac lesions, possibly due to reduced viability in fetuses with large CNVs and complex CHD. CHD gene–containing CNVs were phenotypically more similar to large CNVs in terms of increased frequency of septal defects and decreased frequencies of CTD, heterotaxy, and univentricular physiology, possibly because many established CHD genes are critical in early cardiac development or regulate transcription broadly, and decrease fetal viability when CHD is complex. Whether patients who are born

with complex CHD and large CNVs or CHD gene–containing CNVs have different survival compared with other genetic causes warrants further investigation.

The clinical implications of registry findings are genetic and cardiac. Enrichment of severe cardiac phenotypes in Group IIB patients supports performing CMA over karyotype in patients with complex CHD, consistent with current guidelines. Also, CMA can identify patients with incomplete or atypical cardiac presentations that may not have in the past prompted targeted testing such as fluorescence in situ hybridization for 7q11.23 deletion. Given evidence for dosage sensitivity of many CHD genes, sequencing panels should include copy number analysis. Whole genome sequencing interpretation should target copy number abnormalities, as well as sequence variants in transcriptional regulatory regions that may alter dosage, for these genes. Registry data suggest that targeted testing of CHD genes may have lower yields than CMA in patients with CHD that creates univentricular physiology.

The cardiac considerations emerging from registry data relate to clinical care and outcomes research. For example, the registry adds APVR as a phenotype to consider in patients with 1q21.1 duplications or deletions. Additional imaging to completely define pulmonary vein anatomy may in some cases be indicated, such as in Turner syndrome. The high frequency of LVOTO in patients with disorders of 1q21.1 or 8p23.1 supports cardiac screening of patients for occult left-sided CHD, such as bicuspid aortic valve, upon diagnosis. Many clinical outcomes studies in CHD have been complicated by variable genetic testing or reporting. The registry data suggest that patients with complex CHD, including those with univentricular physiology, are more likely to have submicroscopic CNVs than large CNVs. Recent studies indicate that CNVs may impact survival in patients with nonsyndromic presentations.[35,36] Taken together, these data point to a critical role for copy number analysis in the genetic classification of study cohorts with severe CHD.

Limitations

We were unable to analyze CNVs not included on clinical reports. Using only CNVs that were reported based on laboratory-established criteria helped to reduce noise and greatly supports the clinical utility of the overall findings. Parental testing data were not available for many patients, largely because such testing is not performed clinically in most cases. CMA data were only collected from liveborn infants. The registry is enriched for more severe CHD than general populations, likely because genetic testing is more common in these patients. Also, genomic disorders that were prevalent in the registry, such as 22q11.2 deletion, are commonly associated with relatively severe CHD. Although genetic testing practices likely vary between centers, the multicenter nature of this study supports generalizability. Geographically the centers were in the Southwest, South, and Midwest and replication in additional cohorts, including in populations with different racial and ethnic composition, is important. The consistent methodology for recording cardiac phenotypes reduces variability between centers in regard to cardiac reporting.

The registry includes only patients with CMA abnormality. CMA does not detect sequence variants. Prior studies indicated that CMA abnormalities are present in 10% to 20% of CHD, so the registry includes a fraction of the overall CHD population. Gene analyses in this study focused on single-gene and CHD gene containing CNVs, which constituted 28% of Group II patients. The identified candidate genes and CHD genes were present in a minority of Group IIB patients, indicating that a large number of causal genes remain to be identified. Future work will include analysis of recurrent loci and algorithmic genotype–phenotype analyses leveraging the registry's detailed raw cardiac data. Utilizing candidate genes identified by exome studies to determine if they were represented in CNV regions may also be a useful approach.[9]

In some patients, genetic testing may have been performed because of concurrent developmental delay, which could skew CMA results toward neurodevelopmental genes and pathways. This merits consideration; however, multiple consortium centers routinely perform CMA in neonates and infants with CHD as the primary indication. In general, standardizing the early genetic evaluation of infants with CHD will likely foster a more complete understanding of the clinical impact of abnormal testing results on later neurodevelopmental and survival outcomes.[35]

CONCLUSIONS

A cardiac phenotype-intensive analysis of a large number of patients with CMA abnormalities has generated novel cardiogenomic associations including expanding of CHD phenotypes in genomic disorders, novel candidate genes, and stratification of CHD phenotypes based on CNV size and the involved genes.

ARTICLE INFORMATION

Received January 7, 2023; accepted May 24, 2023.

Affiliations

Indiana University School of Medicine, Indianapolis, IN (B.J.L., L.R.H., G.C.G., J.C.H., S.M.W.); University of Pittsburgh, Pittsburgh, PA (J.-H.I.L., S.A.Y., C.W.L.); Emory University School of Medicine, Atlanta, GA (W.L.B., S.B.W.); Baylor College of Medicine, Houston, TX (C.N.M., M.S.A., S.R.L.); Texas Children's Hospital, Houston, TX (C.N.M., M.S.A., S.R.L.); Cincinnati Children's Hospital Medical Center, Cincinnati, OH (R.B.H.); Nationwide

Children's Hospital, The Ohio State University, Columbus, OH (V.G., K.L.M.); and University of Calgary, Calgary, Canada (K.L.M.).

Sources of Funding

This work was supported by the American Heart Association Transformational Award AHA 19TPA34850054 (SMW); and National Institutes of Health K23 award HL141667 (BJL).

Disclosures

None.

Supplemental Material

Data S1–S2
Tables S1–S6
Figures S1–S3
Data S2

REFERENCES

1. Pierpont ME, Brueckner M, Chung WK, Garg V, Lacro RV, McGuire AL, Mital S, Priest JR, Pu WT, Roberts A, et al. Genetic basis for congenital heart disease: revisited: a scientific statement from the American Heart Association. *Circulation*. 2018;138:e653–e711. doi: 10.1161/CIR.0000000000000606
2. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. 2010;86:749–764. doi: 10.1016/j.ajhg.2010.04.006
3. Geddes GC, Basel D, Frommelt P, Kinney A, Earing M. Genetic testing protocol reduces costs and increases rate of genetic diagnosis in infants with congenital heart disease. *Pediatr Cardiol*. 2017;38:1465–1470. doi: 10.1007/s00246-017-1685-7
4. Poirsier C, Besseau-Ayasse J, Schluth-Bolard C, Toutain J, Missirian C, Le Caignec C, Bazin A, de Blois MC, Kuentz P, Catty M, et al. A French multicenter study of over 700 patients with 22q11 deletions diagnosed using FISH or aCGH. *Eur J Hum Genet*. 2016;24:844–851. doi: 10.1038/ejhg.2015.219
5. Digilio MC, Bernardini L, Consoli F, Lepri FR, Giuffrida MG, Baban A, Surace C, Ferese R, Angioni A, Novelli A, et al. Congenital heart defects in recurrent reciprocal 1q21.1 deletion and duplication syndromes: rare association with pulmonary valve stenosis. *Eur J Med Genet*. 2013;56:144–149. doi: 10.1016/j.ejmg.2012.12.004
6. Hinton RB, McBride KL, Bleyl SB, Bowles NE, Border WL, Garg V, Smolarek TA, Lalani SR, Ware SM. Rationale for the cytogenomics of cardiovascular malformations consortium: a phenotype intensive registry based approach. *J Cardiovasc Dev Dis*. 2015;2:76–92. doi: 10.3390/jcdd2020076
7. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. doi: 10.1016/j.jbi.2019.103208
8. Navarro Gonzalez J, Zweig AS, Speir ML, Schmelter D, Rosenbloom KR, Raney BJ, Powell CC, Nassar LR, Maulding ND, Lee CM, et al. The UCSC genome browser database: 2021 update. *Nucleic Acids Res*. 2021;49:D1046–d57. doi: 10.1093/nar/gkaa1070
9. Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science*. 2015;350:1262–1266. doi: 10.1126/science.aac9396
10. Botto LD, Lin AE, Riehle-Colarusso T, Malik S, Correa A. Seeking causes: classifying and evaluating congenital heart defects in etiologic studies. *Birth Defects Res A Clin Mol Teratol*. 2007;79:714–727. doi: 10.1002/bdra.20403
11. Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. National time trends in congenital heart defects, Denmark, 1977–2005. *Am Heart J*. 2009;157:467–73.e1. doi: 10.1016/j.ahj.2008.10.017
12. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;3:W305–W311. doi: 10.1093/nar/gkp427
13. Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ. ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. *Nucleic Acids Res*. 2010;38:W96–W102. doi: 10.1093/nar/gkq418
14. Marelli AJ, Mackie AS, Ionescu-Ittu R, Rahme E, Pilote L. Congenital heart disease in the general population: changing prevalence and age distribution. *Circulation*. 2007;115:163–172. doi: 10.1161/CIRCULATIONAHA.106.627224
15. Ferencz C, Rubin JD, McCarter RJ, Brenner JI, Neill CA, Perry LW, Hepner SI, Downing JW. Congenital heart disease: prevalence at livebirth. The Baltimore-Washington Infant Study. *Am J Epidemiol*. 1985;121:31–36. doi: 10.1093/oxfordjournals.aje.a113979
16. Hodge JC, Mitchell E, Pillalamarri V, Toler TL, Bartel F, Kearney HM, Zou YS, Tan WH, Hanscom C, Kirmani S, et al. Disruption of MBD5 contributes to a spectrum of psychopathology and neurodevelopmental abnormalities. *Mol Psychiatry*. 2014;19:368–379. doi: 10.1038/mp.2013.42
17. Rajagopal SK, Ma Q, Obler D, Shen J, Manichaikul A, Tomita-Mitchell A, Boardman K, Briggs C, Garg V, Srivastava D, et al. Spectrum of heart disease associated with murine and human GATA4 mutation. *J Mol Cell Cardiol*. 2007;43:677–685. doi: 10.1016/j.yjmcc.2007.06.004
18. LaHaye S, Majumdar U, Yasuhara J, Koenig SN, Matos-Nieves A, Kumar R, Garg V. Developmental origins for semilunar valve stenosis identified in mice harboring congenital heart disease-associated GATA4 mutation. *Dis Model Mech*. 2019;12:1–13. doi: 10.1242/dmm.036764
19. Soemedi R, Topf A, Wilson IJ, Darlay R, Rahman T, Glen E, Hall D, Huang N, Bentham J, Bhattacharya S, et al. Phenotype-specific effect of chromosome 1q21.1 rearrangements and GJA5 duplications in 2436 congenital heart disease patients and 6760 controls. *Hum Mol Genet*. 2012;21:1513–1520. doi: 10.1093/hmg/ddr589
20. Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, Yuen T, Rickaby J, Thiruvahindrapuram B, Marshall CR, et al. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet*. 2012;8:e1002843. doi: 10.1371/journal.pgen.1002843
21. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH, Korn JM, McCarroll SA, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet*. 2009;41:931–935. doi: 10.1038/ng.415
22. Glessner JT, Bick AG, Ito K, Homsy J, Rodriguez-Murillo L, Fromer M, Mazaika E, Vardarajan B, Italia M, Leipzig J, et al. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. *Circ Res*. 2014;115:884–896. doi: 10.1161/CIRCRESAHA.115.304458
23. Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, Huang S, Maloney VK, Crolla JA, Baralle D, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med*. 2008;359:1685–1699. doi: 10.1056/NEJMoa0805384
24. Christiansen J, Dyck JD, Elyas BG, Lilley M, Bamforth JS, Hicks M, Sprysak KA, Tomaszewski R, Haase SM, Vicen-Wyhony LM, et al. Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. *Circ Res*. 2004;94:1429–1435. doi: 10.1161/01.RES.0000130528.72330.5c
25. Butensky A, de Rinaldis CP, Patel S, Edman S, Bailey A, McGinn DE, Zackai E, Crowley TB, McDonald-McGinn DM, Min J, et al. Cardiac evaluation of patients with 22q11.2 duplication syndrome. *Am J Med Genet A*. 2021;185:753–758. doi: 10.1002/ajmg.a.62032
26. Shanshen E, Rosenberg J, Van Bergen AH. Identification of novel congenital heart disease candidate genes using chromosome microarray. *Pediatr Cardiol*. 2018;39:148–159. doi: 10.1007/s00246-017-1741-3
27. Hasten E, McDonald-McGinn DM, Crowley TB, Zackai E, Emanuel BS, Morrow BE, Racedo SE. Dysregulation of TBX1 dosage in the anterior heart field results in congenital heart disease resembling the 22q11.2 duplication syndrome. *Hum Mol Genet*. 2018;27:1847–1857. doi: 10.1093/hmg/ddy078
28. Pauli S, Bajpai R, Borchers A. CHARGEd with neural crest defects. *Am J Med Genet C Semin Med Genet*. 2017;175:478–486. doi: 10.1002/ajmg.c.31584
29. Kirby ML, Hutson MR. Factors controlling cardiac neural crest cell migration. *Cell Adh Migr*. 2010;4:609–621. doi: 10.4161/cam.4.4.13489
30. Schaaf CP, Boone PM, Sampath S, Williams C, Bader PI, Mueller JM, Shchelochkov OA, Brown CW, Crawford HP, Phalen JA, et al. Phenotypic spectrum and genotype-phenotype correlations of NRXN1 exon deletions. *Eur J Hum Genet*. 2012;20:1240–1247. doi: 10.1038/ejhg.2012.95
31. Dabell MP, Rosenfeld JA, Bader P, Escobar LF, El-Khechen D, Vallee SE, Dinulos MB, Curry C, Fisher J, Tervo R, et al. Investigation of NRXN1 deletions: clinical and molecular characterization. *Am J Med Genet A*. 2013;161:717–731. doi: 10.1002/ajmg.a.35780

32. Ching MS, Shen Y, Tan WH, Jeste SS, Morrow EM, Chen X, Mukaddes NM, Yoo SY, Hanson E, Hundley R, et al. Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am J Med Genet B Neuropsychiatr Genet*. 2010;153B:937–947. doi: 10.1002/ajmg.b.31063
33. Izarzugaza JMG, Ellesøe SG, Doganli C, Ehlers NS, Dalgaard MD, Audain E, Dombrowsky G, Banasik K, Sifrim A, Wilsdon A, et al. Systems genetics analysis identifies calcium-signaling defects as novel cause of congenital heart disease. *Genome Med*. 2020;12:76. doi: 10.1186/s13073-020-00772-z
34. Fenske S, Krause SC, Hassan SI, Becirovic E, Auer F, Bernard R, Kupatt C, Lange P, Ziegler T, Wotjak CT, et al. Sick sinus syndrome in HCN1-deficient mice. *Circulation*. 2013;128:2585–2594. doi: 10.1161/CIRCULATIONAHA.113.003712
35. Landis BJ, Helm BM, Herrmann JL, Hoover MC, Durbin MD, Elmore LR, Huang M, Johansen M, Li M, Przybylowski LF, et al. Learning to crawl: determining the role of genetic abnormalities on postoperative outcomes in congenital heart disease. *J Am Heart Assoc*. 2022;11:e026369. doi: 10.1161/JAHA.122.026369
36. Boskovski MT, Homsy J, Nathan M, Sleeper LA, Morton S, Manheimer KB, Tai A, Gorham J, Lewis M, Swartz M, et al. De novo damaging variants, clinical phenotypes, and post-operative outcomes in congenital heart disease. *Circ Genom Precis Med*. 2020;13:e002836. doi: 10.1161/CIRCGEN.119.002836

SUPPLEMENTAL MATERIAL

Data S1

Supplemental Methods.

The Level 1 diagnoses of tetralogy of Fallot (TOF) and hypoplastic left heart syndrome (HLHS) are established compilations of cardiac defects. Any right-sided obstructive lesions located from the subpulmonary region to the distal pulmonary arteries were not recorded when TOF was present. Similarly, left-sided obstructive lesions from the mitral valve to proximal descending aorta were not individually recorded in HLHS. Patients with conotruncal defects (CTDs) that characteristically have an associated ventricular septal defect (VSD), including TOF, double outlet right ventricle, and truncus arteriosus, did not have a VSD entered separately unless there was a second anatomically distinct VSD. Similarly, patients with double inlet left ventricle who had a bulboventricular foramen functioning as an interventricular communication did not have VSD separately recorded. Absence of VSD in any patients with these particular diagnoses would be noted in the cardiac data entry.

Data S2

Supplemental Data Set (Excel File).

Annotations for Group IIB monogenic copy-number variants.

Table S1. Demographics of registry cohort (N=1363)

Characteristic	**N (%)**
Sex	
Male	660 (48)
Female	642 (47)
Not Reported/Unknown	61 (5)
Ethnicity	
Not Hispanic or Latino	956 (70)
Hispanic or Latino	110 (8)
Not Reported/Unknown	297 (22)
Race	
White	823 (60)
Black or African American	186 (14)
Asian	40 (3)
Native American or Alaska Native	5 (0.4)
Native Hawaiian or Other Pacific Islander	3 (0.2)
Other	58 (4)
Not Reported/Unknown	248 (18)

Table S2. Distribution of the number of CNVs that were reported in study patients.

Number of reported CNVs	Group I (N=386) Number of cases, (%)	Group II (N=911) Number of cases, (%)
1	292 (76)	673 (74)
2	79 (20)	203 (22)
3	14 (4)	28 (3)
4	0	5 (0.5)
5	1 (0.3)	2 (0.2)

Table S3. Frequency of Group IIB hierarchical CHD categories and presumed univentricular physiology classification among the subgroups of patients with submicroscopic CNVs.

CHD category	CNV with CHD gene (72 patients) N (%)	CNV with a single gene (183 patients) N (%)	CNV without CHD gene or single gene (541 patients) N (%)
CTD	13 (18.1)	60 (32.8)	140 (25.9)
LVOTO	23 (31.9)	33 (18.0)	111 (20.5)
Septal defect	14 (19.4)	21 (11.5)	82 (15.2)
RVOTO	4 (5.6)	17 (9.3)	43 (7.9)
Septal + LVOTO	6 (8.3)	13 (7.1)	27 (5.0)
HTX	3 (4.2)	14 (7.7)	39 (7.2)
APVR	3 (4.2)	6 (3.3)	26 (4.8)
Arteriopathy	2 (2.8)	3 (1.6)	9 (1.7)
CTD + AVSD	1 (1.4)	5 (2.7)	10 (1.8)
AVSD	1 (1.4)	5 (2.7)	17 (3.1)
Aortopathy	1 (1.4)	3 (1.6)	9 (1.6)
Septal + RVOTO	0	2 (1.1)	9 (1.6)
Cardiomyopathy	1 (1.4)	1 (0.5)	8 (1.5)
Other	0	0	7 (1.3)
PDA	0	0	2 (0.4)
SV, os	0	0	2 (0.4)
Univentricular physiology	8 (11.1)	36 (19.7)	95 (17.6)

Two registry patients had a monogenic CNV that contained a CHD gene. These were included in the CHD gene group.

CTD: conotruncal defect, LVOTO: left ventricular obstruction, RVOTO: right ventricular obstruction, HTX: heterotaxy, AVSD: atrioventricular septal defect, APVR: anomalous pulmonary venous return, PDA: patent ductus arterious, SV, os: single ventricle otherwise specified.

Table S4. Cardiac phenotypes in patients with 2q23.1 duplications or deletions (MIM #156200)

Patient	Sex	CMA abnormality (arr[hg19])	Level 3 Diagnoses	Level 1 Diagnoses	Hierarchical Category
1	M	2q23.1 (149177826-149359334) x3	CTD	TOF only	CTD
2	M	2q23.1 (149177826-149359334) x3	Septal; LVOTO	Secundum ASD; Mitral stenosis or hypoplasia	Septal defect
3	F	2q23.1 (149177826-149359334) x3	Septal	Secundum ASD	Septal defect
4	M	2q23.1 (148938816-149034418) x1	CTD	d-TGA with intact ventricular septum and no ventricular obstruction	CTD

ASD: atrial septal defect; CTD: conotruncal defect; F: Female; L: left ventricular obstruction; M: Male; d-TGA: D-transposition of the great arteries; TOF: tetralogy of Fallot

Table S5. Enrichment analysis results for 141 monogenic CNV genes.

Pathway (834 annotation terms tested; significance threshold: 6.00E-05)		GO: Biological Process (3271 annotation terms tested; significance threshold: 1.53E-05)		Disease (DisGeNET) (2964 annotation terms tested; significance threshold: 1.69E-05)	
Pathway (P value)	Genes	Biological Process (P value)	Genes	Phenotype term (P value)	Genes
Neuronal System* (3.07E-05)	*HCN1, SLC1A1, SYT10, ADCY2, ADCY8, CHRNA7, NRXN3, SLC22A2, TSPAN7, DLG2, CACNA2D3*	neuron cell-cell adhesion (2.65E-06)	*CNTN4, NRXN3, ASTN2, TNR*	Intellectual Disability (2.92E-09)	*CNTN4, PDE4D, TCF12, AFF2, ATR, DOCK8, PPM1D, MYT1, MAGT1, SLC1A1, CHRNA7, CNTN6, KANK1, NRXN3, TMLHE, ASL, RBFOX1, TSPAN7, MACROD2, AUTS2, ULK4, DLG2, PRKN, DMD, CTNND2, FTO, MBD5, TNR*
Phosphodiesterases in neuronal function** (3.60E-05)	*PDE4D, ADCY2, ADCY8, CHRNA7, GUCY1A2*	cell-cell adhesion (9.25E-06)	*CNTN4, PDE4D, DOCK8, ADAM9, OBSCN, GRID2, AP3B1, CCDC141, YES1, CLDN23, CNTN6, NRXN3, PTPRM, PCDH11X, DLG2, ASTN2, CTNND2, TNR, CDH4, ALOX5*	Neurodevelopmental Disorders (1.49E-08)	*CNTN4, DOCK8, PPM1D, CHRNA7, CNTN6, KANK1, NRXN3, RBFOX1, MACROD2, AUTS2, ULK4, DLG2, ASTN2, PRKN, DMD, MBD5, IMMP2L*
		neuron development (4.41E-06)	*CNTN4, TCF12, WDR36, THSD7A, MYOC, HCN1, GRID2, BBS5, CCDC141, CHRNA7, CNTN6, KANK1, NRXN3,*	Developmental delay (disorder) (4.52E-07)	*CNTN4, TCF12, AFF2, PPM1D, SHOX, GRID2, CHRNA7, CNTN6, FANCD2, KANK1, RBFOX1, AUTS2, DMD, TANGO2, FTO, MBD5*

			PTPRM, VAMP7, AUTS2, ULK4, DLG2, PRKN, DMD, CTNND2, TNR, CDH4, SEMA3E, SEZ6		
				Mental Retardation (6.80E-07)	*CNTN4, TCF12, AFF2, ATR, DOCK8, MYT1, SLC1A1, ASL, RBFOX1, TSPAN7, MACROD2, AUTS2, PRKN, DMD, CTNND2*
				Pervasive Development Disorder (1.44E-06)	*CNTN4, AFF2, SLC1A1, CHRNA7, NRXN3, RBFOX1, MACROD2, AUTS2, PRKN, CTNND2, MBD5, IMMP2L*
				Autistic Disorder (7.53E-06)	*CNTN4, CHRNA7, GHR, NRXN3, TMLHE, RBFOX1, MACROD2, ASTN2, PRKN, IMMP2L*
				Global developmental delay (8.84E-06)	*CNTN4, TCF12, AFF2, PPM1D, SHOX, GRID2, CHRNA7, CNTN6, FANCD2, RBFOX1, AUTS2, DMD, TANGO2, FTO, MBD5*

Statistical significance was defined as P values < 0.05 divided by the number of annotation terms that were tested.

*BioSystems: REACTOME

**MSigDB C2 BIOCARTA (v7.5.1)

Table S6. Pathways that are enriched among 34 curated CHD genes that were in Group IIB CNVs.

Pathway	Source	P value	Genes
Heart development	MSigDB C2 BIOCARTA (v7.5.1)	1.10E-09	*GATA4, NKX2-5, TBX20, NOTCH1, FOXC1, PITX2*
Cardiac progenitor differentiation	MSigDB C2 BIOCARTA (v7.5.1)	1.61E-07	*GATA4, NKX2-5, KDR, TBX20, NOTCH1*
Notch-mediated HES/HEY network	MSigDB C2 BIOCARTA (v7.5.1)	5.17E-06	*GATA4, KDR, CREBBP, NOTCH1*
NFAT and Hypertrophy of the heart (Transcription in the broken heart)	MSigDB C2 BIOCARTA (v7.5.1)	6.61E-06	*GATA4, RAF1, NKX2-5, CREBBP*
Thyroid hormone signaling pathway	BioSystems: KEGG	8.09E-06	*GATA4, RAF1, MED13L, CREBBP, NOTCH1*
BMP signaling in eyelid development	MSigDB C2 BIOCARTA (v7.5.1)	1.44E-05	*NOTCH1, FOXC1, PITX2*
22q11.2 copy number variation syndrome	MSigDB C2 BIOCARTA (v7.5.1)	1.46E-05	*RAF1, NKX2-5, LZTR1, FOXC1, PITX2*
Pathways affected in adenoid cystic carcinoma	MSigDB C2 BIOCARTA (v7.5.1)	1.86E-05	*RAF1, KANSL1, CREBBP, NOTCH1*
Angiogenesis	PantherDB	2.82E-05	*RAF1, KDR, PRKD1, NOTCH1, JAG1*
TFAP2 (AP-2) family regulates transcription of other transcription factors	BioSystems: REACTOME	3.41E-05	*CITED2, PITX2*
altered Notch signaling involving the main players	Pathway Ontology	3.41E-05	*NOTCH1, JAG1*
YAP1- and WWTR1 (TAZ)-stimulated gene expression	BioSystems: REACTOME	4.55E-05	*GATA4, NKX2-5, CREBBP*
Chromatin organization	BioSystems: REACTOME	4.89E-05	*EHMT1, SMARCB1, ARID1B, KMT2D, KANSL1, CREBBP*
Chromatin modifying enzymes	BioSystems: REACTOME	4.89E-05	*EHMT1, SMARCB1, ARID1B, KMT2D, KANSL1, CREBBP*
transforming growth factor-beta signaling	Pathway Ontology	5.68E-05	*NF1, CREBBP*

819 annotation terms were tested. P value significance threshold is 0.05/819 = 6.11E-05.

CONOTRUNCAL
"Truncus only"
"IAA, type B"
"IAA type B and Truncus"
"IAA, nos"
"D-TGA with intact ventricular septum and no OTO"
"D-TGA with intact ventricular septum and LVOTO"
"D-TGA with intact ventricular septum and RVOTO"
"D-TGA with VSD and no OTO"
"D-TGA with VSD and LVOTO"
"D-TGA with VSD and RVOTO"
"D-TGA, nos"
"D-TGA, os"
"TOF only"
"TOF with absent pulmonary valve"
"PA with VSD and TOF anatomy"
"DORV-TOF type"
"DORV-TGA type"
"DORV, os"
"DORV, nos"
"Conoventricular VSD"
"Hemitruncus"

HETEROTAXY
"Atrial Isomerism"
"Atrial Situs Inversus"
"Dextrocardia with CVM"
"Superior-Inferior Ventricles"
"Ventricular Inversion, L-Looped"
"LSVC or Bilateral SVC"
"Interrupted IVC"
"L-Sided IVC"
"L-TGA without OTO"
"L-TGA with RVOTO"
"L-TGA with LVOTO"
"Right-Sided Abdominal Aorta"
"Situs Inversus Totalis with CVM"
"Situs Inversus Totalis, No CVM"
"Dextrocardia with Normal Intracardiac Anatomy"
"Mesocardia"
"Levocardia with CVM"

LEFT VENTRICULAR OBSTRUCTIVE LESION (LVOTO)
"HLHS with intact ventricular septum"
"HLHS with VSD"
"HLHS with APVR"
"IAA, type A"
"IAA, type C"
"CoA with intact ventricular septum"
"CoA with VSD"
"Aortic Valve Stenosis or Hypoplasia"
"Subaortic Stenosis or Narrowing"
"BAV"
"Aortic Valve Malformation, os"
"Mitral Stenosis or Hypoplasia"
"Mitral Atresia"
"Mitral Malformation, os"
"Mitral Valve Prolapse"
"Left Ventricular Hypoplasia, non-HLHS"
"Aortic Atresia"

RIGHT VENTRICULAR OBSTRUCTIVE LESION (RVOTO)
"Tricuspid Atresia with intact ventricular septum"
"Tricuspid Atresia with VSD"
"Tricuspid Stenosis or Hypoplasia"
"Tricuspid Malformation, os"
"Ebstein's anomaly"
"Pulmonary Valve Stenosis or Hypoplasia"
"Pulmonary Valve Malformation, os"
"Subpulmonary Stenosis or Narrowing"
"Anatomic Peripheral PS or Hypoplasia"
"PA with intact ventricular septum"
"PA, nos"
"PA with VSD, not TOF Anatomy"
"PA with VSD, nos"
"Right Ventricular Hypoplasia"

SEPTAL DEFECT
"Perimembranous VSD"
"Muscular VSD"
"Supracristal VSD"
"VSD nos"
"VSD os"
"Secundum ASD"
"Sinus Venosus Defect"
"ASD, nos"
"ASD, os"
"Common Atrium"

ATRIOVENTRICULAR SEPTAL DEFECT
"Primum ASD"
"Inlet VSD"
"Complete balanced AVSD"
"Complete unbalanced RV dominant AVSD"
"Complete unbalanced LV dominant AVSD"
"Transitional AVSD"
"AVSD, os"
"AVSD, nos"
"Cleft Mitral Valve"

ANOMALOUS PULMONARY VENOUS RETURN
"TAPVR without OTO"
"TAPVR with RVOTO"
"TAPVR with LVOTO"
"PAPVR without OTO"
"PAPVR with RVOTO"
"PAPVR with LVOTO"
"Cor Triatriatum"

ARTERIOPATHY
"Right Aortic Arch"
"Double Aortic Arch"
"Aberrant Subclavian Artery"
"Pulmonary Artery Anomaly, nos"
"Pulmonary Artery Anomaly, os"
"Vascular Ring, nos"
"Aortopulmonary Window"
"Supravalvar AS"
"PDA, age greater than 1 yr"

SINGLE VENTRICLE
"DILV, L-malposition"
"DILV, D-malposition"
"DILV, normally related great vessels"
"DILV, nos"
"DIRV"
"SV, os"
"SV, nos"

AORTOPATHY
"Aortic Root Dilation"
"Ascending Aorta Dilation"
"Aortic Sinotubular Junction Dilation"

CORONARY ANOMALY
"ALCAPA"
"Coronary Anomaly, nos"
"Coronary Anomaly, os"

CARDIOMYOPATHY
"HCM"
"DCM"
"RCM"
"ARVC"
"LVNC"
"Cardiomyopathy, nos"

MYOCARDIAL
"LV trabeculations"

OTHER
Entered as free text

Figure S1. Cardiac phenotyping framework that was utilized at the time of registry phenotyping. Level 3 CHD category (bold headers) comprises a set of Level 1 cardiac phenotypes. Both Level 1 and Level 3 categories were recorded for each registry entry. The cardiac phenotype definitions were developed from a modified classification approach utilized by the National Birth Defects Prevention Study (NBDPS) [10].

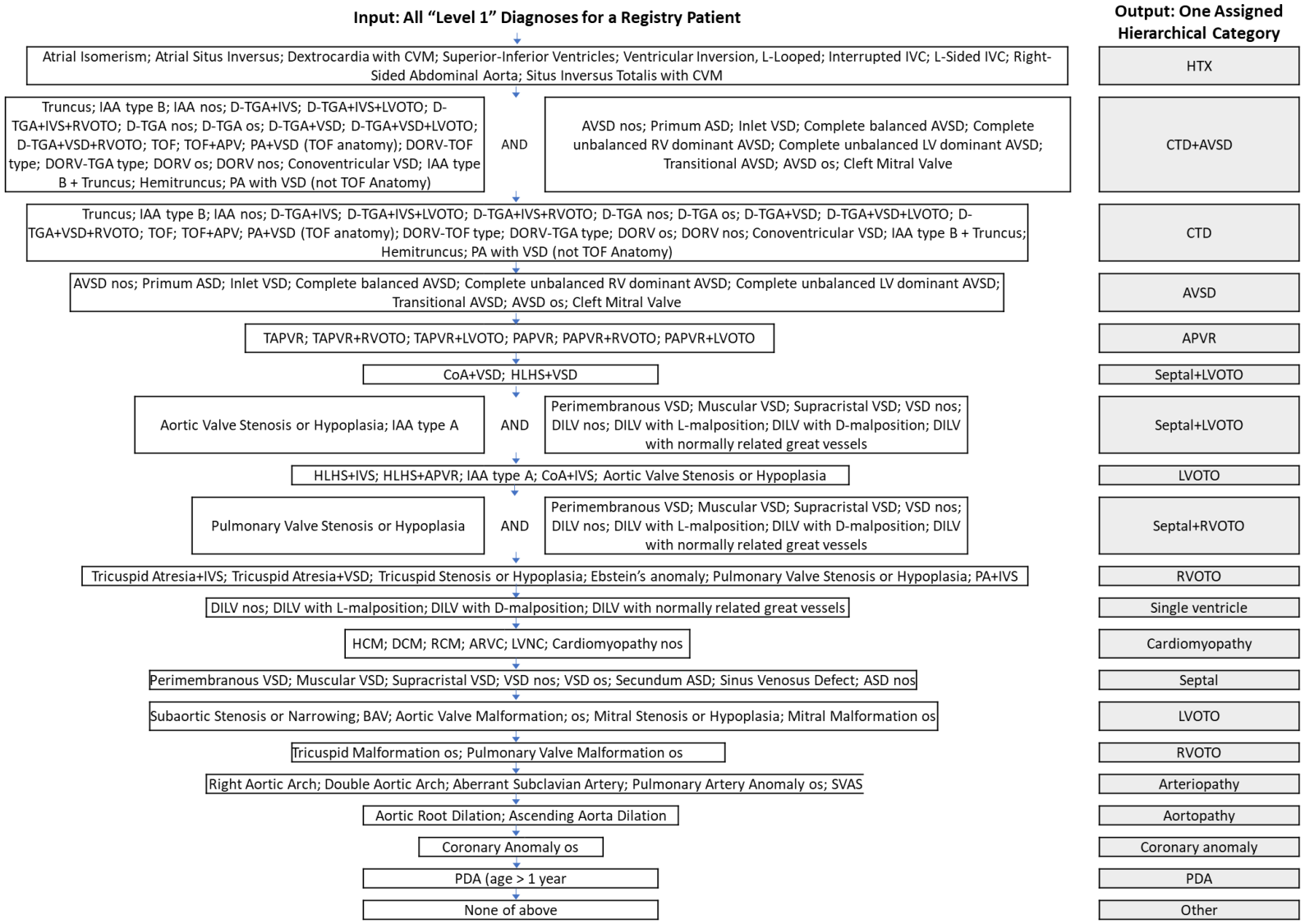

Figure S2. Schema for the algorithmic hierarchical classification of an overall CHD category based upon a registry patient's Level 1 diagnoses. Starting from top, a patient with a Level 1 diagnosis (white box) is assigned to the corresponding hierarchical category (gray box) and removed from further classification. The method was adapted from Oyen et al [11].

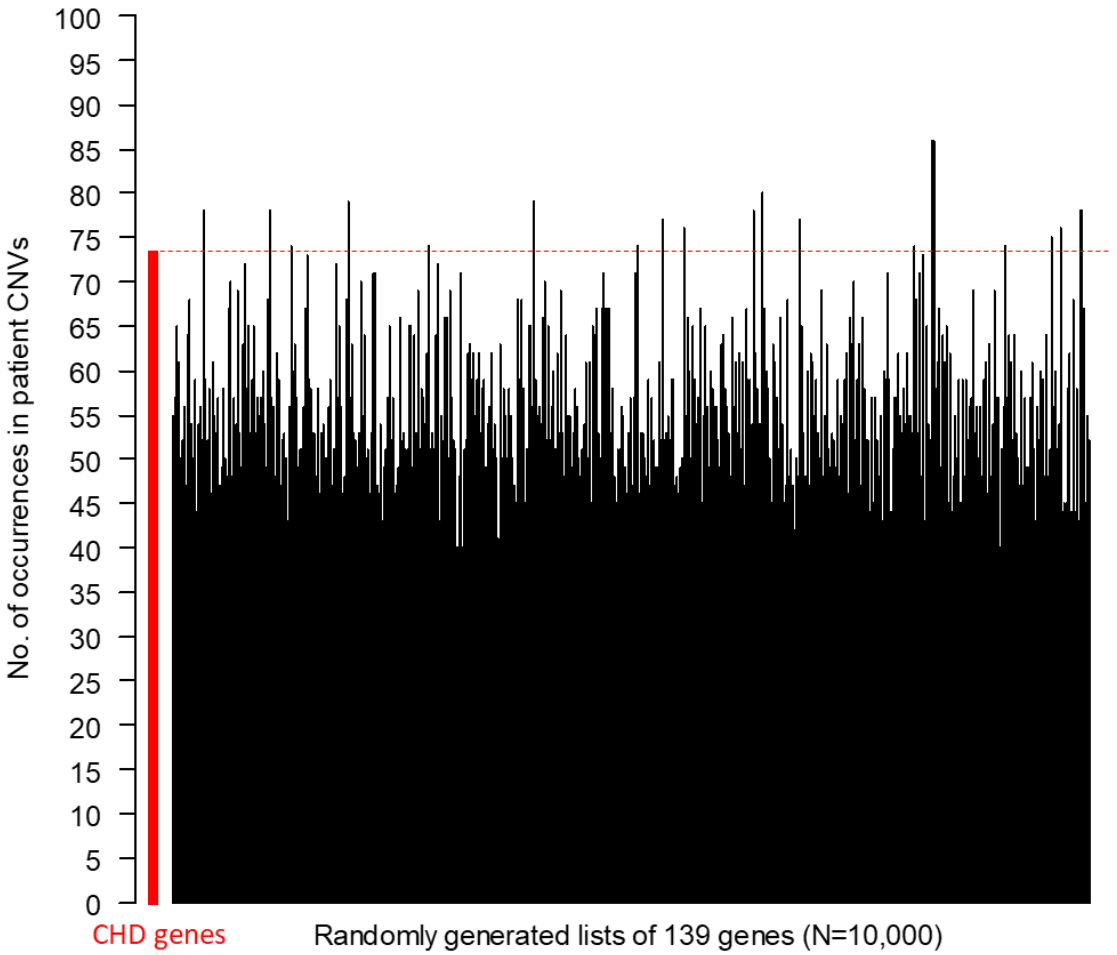

Figure S3. Permutations estimating the likelihood for the frequency that CHD genes were located in Group II with submicroscopic CNVs. In Genetic Group IIB patient data, CHD genes (N=139) occurred in 73 CNVs (red bar). In comparison, only 20 of 10,000 randomly generated 139-gene lists contained genes that occurred in Group IIB patient CNVs more than 73 times. These are shown as black bars that extend above the dashed horizontal red line. This analysis indicates significant enrichment of CHD genes in Group IIB CNVs based on an estimated p-value of 0.002.