

Article

A Reassessment of Copy Number Variations in Congenital Heart Defects: Picturing the Whole Genome

Ilse Meerschaut ^{1,2} , Sarah Vergult ¹ , Annelies Dheedene ¹, Björn Menten ¹ , Katya De Grootte ² , Hans De Wilde ², Laura Muiño Mosquera ^{1,2} , Joseph Panzer ², Kristof Vandekerckhove ², Paul J. Coucke ¹, Daniël De Wolf ^{2,3}  and Bert Callewaert ^{1,*} 

- ¹ Center for Medical Genetics, Ghent University Hospital, Belgium and Department of Biomolecular Medicine, Ghent University, 9000 Ghent, Belgium; ilse.meerschaut@ugent.be (I.M.); sarah.vergult@ugent.be (S.V.); annelies.dheedene@ugent.be (A.D.); bjorn.menten@ugent.be (B.M.); laura.muinosmosquera@uzgent.be (L.M.M.); paul.coucke@ugent.be (P.J.C.)
- ² Department of Pediatric Cardiology, Ghent University Hospital, 9000 Ghent, Belgium; katya.degrootte@uzgent.be (K.D.G.); hans.dewilde@uzgent.be (H.D.W.); joseph.panzer@uzgent.be (J.P.); kristof.vandekerckhove@uzgent.be (K.V.); daniel.dewolf@uzgent.be (D.D.W.)
- ³ Department of Pediatric Cardiology, Brussels University Hospital, 1090 Brussels, Belgium
- * Correspondence: bert.callewaert@ugent.be; Tel.: +32-9-332-3603



Citation: Meerschaut, I.; Vergult, S.; Dheedene, A.; Menten, B.; De Grootte, K.; De Wilde, H.; Muiño Mosquera, L.; Panzer, J.; Vandekerckhove, K.; Coucke, P.J.; et al. A Reassessment of Copy Number Variations in Congenital Heart Defects: Picturing the Whole Genome. *Genes* **2021**, *12*, 1048. <https://doi.org/10.3390/genes12071048>

Academic Editors: Livia Garavelli and Stefano Giuseppe Caraffi

Received: 8 June 2021
Accepted: 5 July 2021
Published: 8 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Copy number variations (CNVs) can modulate phenotypes by affecting protein-coding sequences directly or through interference of gene expression. Recent studies in cancer and limb defects pinpointed the relevance of non-coding gene regulatory elements such as long non-coding RNAs (lncRNAs) and topologically associated domain (TAD)-related gene-enhancer interactions. The contribution of such non-coding elements is largely unexplored in congenital heart defects (CHD). We performed a retrospective analysis of CNVs reported in a cohort of 270 CHD patients. We reviewed the diagnostic yield of pathogenic CNVs, and performed a comprehensive reassessment of 138 CNVs of unknown significance (CNV-US), evaluating protein-coding genes, lncRNA genes, and potential interferences with TAD-related gene-enhancer interactions. Fifty-two of the 138 CNV-US may relate to CHD, revealing three candidate CHD regions, 19 candidate CHD genes, 80 lncRNA genes of interest, and six potentially CHD-related TAD interferences. Our study thus indicates a potential relevance of non-coding gene regulatory elements in CNV-related CHD pathogenesis. Shortcomings in our current knowledge on genomic variation call for continuous reporting of CNV-US in international databases, careful patient counseling, and additional functional studies to confirm these preliminary findings.

Keywords: copy number variations; congenital heart defects; protein-coding genes; non-coding elements; topology associated domains

1. Introduction

Copy number variations (CNVs) are DNA segments of one kilobase (kb) or larger, which are present at a variable copy number in comparison with the reference genome [1,2]. The introduction of chromosomal microarray analysis (CMA) in the late 1990s allowed the identification of small (below three megabases (Mb)) chromosomal imbalances down to resolutions of 100 kb. With the advent of massive parallel sequencing technologies, shallow whole genome sequencing (sWGS) entered the diagnostic field as a highly accurate, faster and cheaper alternative to CMA analysis.

CNVs can be directly causative of human diseases such as neurodevelopmental disorders and congenital malformations, or may contribute to the multifactorial risk for variable neurodevelopmental conditions including autism and intellectual disability, psychiatric disease, and congenital anomalies [2–8]. However, CNVs are also an important source of normal genetic variation [1,2], hence estimation of their pathogenicity is challenging.

Current clinical practice classifies CNV pathogenicity mainly based on size, parental inheritance, population frequency, known genotype–phenotype correlations, and protein-coding gene content of the CNV [9–12]. Despite these classification tools, many CNVs are classified as a CNV of unknown significance (CNV-US), with some of these CNV-US possibly being causative for the phenotype. In multifactorial disorders, in particular, common and often inherited variants might render a small effect to the phenotype, thus complicating the interpretation of CNV pathogenicity using the current criteria.

Moreover, recent advances in the identification of the molecular basis of limb defects and cancer revealed alternative CNV-associated disease mechanisms, affecting non-coding gene regulatory elements such as long non-coding RNAs (lncRNAs), enhancer elements, and disruption of the three-dimensional (3D) topologically associated domain (TAD) organization of the genome [13–16].

Congenital heart defects (CHD) are the most prevalent congenital malformation in live born children, and can occur isolated or be part of a syndromic constellation. Both isolated CHD (ICHD) and syndromic CHD (SCHD) can result from chromosomal aberrations including CNVs, single nucleotide variants, or multifactorial disease mechanisms [17–24]. So far, over 400 pathophysiological relevant CHD genes have been identified, including cardiogenic transcription factors such as NKX2.5 and members of the GATA or T-box families, cell signaling molecules involved in wnt, nodal, bmp, smoothed, notch, hippo, transforming growth factor β , fibroblast growth factor, ras, and vascular endothelial growth factor signaling, and chromatin-modifying genes [23–25]. Well-established CHD-related CNVs often affect such major cardiac developmental genes. Even though CNV analysis is relatively embedded in the diagnostic evaluation of CHD [8,26], studies reporting on the diagnostic value remain scarce and report yields between 12.8% and 18.5% [26].

Together, current knowledge on the molecular basis of CHD indicates an important contribution of transcriptional regulation in CHD. In line with this, a contribution of non-coding gene regulatory elements, as observed for limb defects and cancer, is conceivable. We sought preliminary evidence for such mechanisms in CHD performing a comprehensive retrospective analysis of reported CNV-US in a large in-house CHD cohort, focusing on potential interference with expression of pathophysiological relevant genes. More specifically, we evaluated recurrent CNV-US regions, the protein-coding gene content, lncRNA genes, and positional effects affecting TAD structures.

2. Materials and Methods

2.1. Selection and Characterization of the Cohort

Patients with a CHD born between January 2012 and July 2018 were recruited from the departments of pediatric cardiology and medical genetics of the Ghent University Hospital. CHD was defined as any defect in the structure of the heart or the great vessels present at birth. Acquired heart defects were excluded. For patients with atrial septal defects or patent ductus arteriosus, only those requiring a therapeutic intervention (surgical correction or percutaneous intervention) were retained.

Patients with available molecular karyotyping were included. Those patients with CHD and an underlying diagnosis (i.e., explanatory aneuploidy, a known monogenic condition, or prenatal conditions such as medication use, toxins or infection) were excluded.

The cohort was subdivided into SCHD and ICHD, based on the available clinical information. SCHD were defined as CHD that are associated with at least one additional major malformation and/or multiple minor physical anomalies and/or intellectual disability. Major malformations were considered malformations having a significant medical, social or cosmetic consequence, requiring medical intervention (e.g., neural tube defects, cleft palate, diaphragmatic hernia). In contrast, minor anomalies pose no significant health problem and have limited social or cosmetic consequences (e.g., preauricular tags, clinodactyly). Per the definition, we classified all patients with isomerism as SCHD cases.

This retrospective study was approved by the Ethics Committee of Ghent University Hospital (EC UZG 2016/0133).

2.2. Molecular Karyotyping

Molecular karyotyping was performed by CMA or sWGS. CMA was carried out using array comparative genome hybridization (arrayCGH) using a 180 k oligonucleotide array (Agilent Technologies, Santa Clara, CA, USA) with a genome-wide resolution of 100 kb. The reported breakpoints of the CNV regions identified by CMA represent the affected oligos, and thus the minimal affected CNV regions. sWGS was performed on HiSeq3000 sequencer (Illumina, San Diego, CA, USA) with an average genome-wide coverage of $0.1\times$ to $1\times$. The minimal resolution of sWGS is 100 kb. For library preparation the NEXTflex Rapid DNA sequencing kit (Perkin Elmer, Waltham, MA, USA) was used.

CNV data evaluation and interpretation was performed using in-house developed software tools arrayCGHbase [27], Vivar [28] and WisecondorX [29]. CNVs that passed the quality control were further interpreted based on their size, overlap with known disease loci (the human morbid map), presence in control populations (Database of Genomic Variants), gene content, and additional relevant information from an extended literature search. Known pathogenic CNVs, CNV-US ≥ 100 kb, and CNV-US < 100 kb containing one or more protein-coding genes were reported (irrespective of the performed CNV analysis technique), and thus further evaluated in this retrospective study. All CNVs were mapped to the reference genome GRCh37(hg19).

2.3. Descriptive Statistics and Study of CNV-US

We performed a descriptive statistical analysis of the clinical characteristics of the patients and the results of their CNV analysis, in the total cohort, as well as in the SCHD and ICHD subgroups. Reported CNVs were listed and subdivided into pathogenic CNVs and CNV-US, based on the list of ClinGen Pathogenic CNV regions (https://dosage.clinicalgenome.org/pathogenic_region.shtml, accessed on 28 February 2021) [30]. CNVs for which a causal relationship with CHD is unclear were considered CNV-US.

CNV-US were evaluated for:

1. Overlap with ClinGen Pathogenic CNV regions (see above) or ClinGen Dosage Sensitive regions (<https://search.clinicalgenome.org/kb/gene-dosage>, accessed on 28 February 2021) [30].
2. Overlap with other CNV-US within the study cohort.
3. The protein-coding gene content based on NCBI RefSeq Select genes (Updated Annotation Release 105.20190906), extracted from the UCSC Genome Browser using the Table Browser [31].
4. Non-coding gene-regulatory elements contained within the CNV-US. LncRNA genes were extracted from the lncipedia high confidence set (hg19) (<https://www.lncipedia.org/>, accessed on 8 December 2020) [32]. Human VISTA enhancer elements were downloaded from the VISTA enhancer database (<https://enhancer.lbl.gov/>, accessed on 7 December 2020) [33].
5. Interference with the genomic TAD structure, based on the TAD boundaries file that was provided by J. Dixon to the developers of ClinTAD (<https://www.clintad.com/>, accessed on 4 January 2021) and publicly available on github [34,35]. This TAD boundaries file was generated using H1 human embryonic stem cells, chromosome build GRCh37, a bin size of 40 kb, and a window size of 2 Mb. To evaluate the potential disruption of TAD-related gene-enhancer interactions, we considered the protein-coding gene content (NCBI RefSeq Select genes) and the human VISTA enhancer elements within the relevant TAD domains.

2.4. Data Browsing

2.4.1. Dosage Sensitivity of Protein-Coding Genes

Dosage sensitivity of affected protein-coding genes was evaluated using ClinGen Dosage Sensitivity Curations for haploinsufficiency and triplosensitivity (<https://search.clinicalgenome.org/kb/gene-dosage>, accessed on 31 March 2020) [30], and haploinsufficiency parameters: %HI scores (Decipher Downloads—Haploinsufficiency Predictions

version 3) [36], pLI scores and upper bounds of observed/expected (o/e) confidence interval (gnomAD version 2.2.1 download—pLOF Metrics by Gene) [37]. Cutoffs used as a positive argument for haploinsufficiency were a %HI score < 10%, a pLI score > 0.90 and/or an upper bound of the o/e confidence interval < 0.35.

2.4.2. Expression Data of Protein-Coding and lncRNA Genes

Expression data of genes during development of the human heart were obtained from a publicly available RNA-seq time-series dataset covering the development of seven organs, including the heart (<http://www.ebi.ac.uk/>, accession number E-MTAB-6814, accessed on 12 December 2020) [38]. All genes with an expression threshold of two transcripts per million (TPM) in heart tissue in at least one developmental stage between “four weeks post conception” and “neonate”, were considered as being expressed in developing heart tissue.

2.4.3. Protein-Coding Gene Function Annotation

Protein-coding gene function was explored, evaluating the presence of the protein-coding genes in gene lists that might link them to heart development and/or CHD pathogenesis. These gene lists include an in-house CHD panel of 471 known or potential CHD genes (<https://www.cmgg.be/assets/bestanden/GENPANEL-CHD.pdf>, accessed on 1 April 2020), 1638 transcription factors [39], and 3430 genes linked to CHD-related Gene Ontology terms (wnt signaling pathway GO:0016055, nodal signaling pathway GO:0038092, bmp signaling pathway GO:0030509, smoothened signaling pathway GO:0007224, notch signaling pathway GO:0007219, hippo signaling GO:0035329, transforming growth factor β receptor signaling pathway GO:0007179, fibroblast growth factor receptor signaling pathway GO:0008543, ras protein signal transduction GO:0007265, vascular endothelial growth factor signaling pathway GO:0038084, sarcomere GO:0030017, cilium GO:0005929, histone modification GO:0016570, and chromatin remodeling GO:0006338) (Gene Ontology Release version 2020-03-24) [40]. These gene lists can be found in Supplementary Table S1.

3. Results

3.1. Study Cohort and Clinical Characteristics

The study cohort consisted of 290 CHD patients. Twenty patients (19 SCHD patients and one ICHD patient) were excluded because alternative genetic tests revealed a genetic defect explanatory for the CHD phenotype. These included trisomy 13, trisomy 21, monosomy X, mosaic trisomy 12, derivative 18, derivative 22, 8p deletion/duplication syndrome, marker chromosome 22 (cat eye syndrome), a 10 Mb 13q33.2 deletion, five patients with Noonan (-like) syndrome, three patients with CHARGE syndrome, one patient with Kabuki syndrome, one patient with Cornelia de Lange syndrome, and one patient with an *ELN* mutation.

The remaining 270 CHD patients were included in this study. Hence, the cohort including 87 SCHD patients (32.2%) and 183 ICHD patients (67.8%). The sex ratio was 158 boys (58.5%) to 112 girls (41.5%) (56 boys to 31 girls in SCHD; 102 boys to 81 girls in ICHD). The heart defects are listed in Supplementary Table S2. It is of note that all 37 patients with transposition of the great arteries were ICHD patients.

In the SCHD group, two patients harbored a chromosomal defect unrelated to the CHD: a triple X and a mosaic i(Y) (p10) and were included in the further analysis. In the ICHD group, two patients harbored a chromosomal defect unrelated to the CHD: XYY and ish der(Y) and were included in the further analysis.

3.2. CNV Analyses

CNV analysis was performed by CMA in 227 (84.1%) patients and by sWGS in 43 (15.9%) patients

Seventeen patients of the SCHD group (19.5%) harbored a pathogenic CNV that was regarded causative for the CHD, including seven proximal 22q11.2 deletions (MIM 188400), one central 22q11.2 deletion, one distal 22q11.2 deletion (MIM 611867), one proximal

22q11.2 duplication (MIM 608363), four 7q11.23 deletions (MIM 194050), one 17q21.31 deletion (MIM 610443), one 17p11.2 duplication (MIM 610883), and one 4p16.3p15.32 deletion (MIM 194190) (see Supplementary Table S3). These pathogenic CNVs were all reported as pathogenic in the original lab reports, and hence not reclassified based on this study. Thirty-six patients with SCHD (41.4%) harbored a total of 47 CNV-US (min. 1, max. 4, median 1). One girl presented with a Xp22.31 deletion. Deletions of this region are associated with X-linked ichthyosis in males, but have not yet been associated with CHD. Therefore, we considered it as a CNV-US in the further analysis. The remaining 34 patients with SCHD (39.1%) had a normal CNV analysis.

None of the patients in the ICHD group had a CNV that was considered causative for the CHD. Two patients had a CNV for which the association with CHD remains under debate, a proximal 16p11.2 microdeletion and a 16p13.11 microdeletion, and were therefore considered as CNV-US in the further analysis. As such, 69 of the 183 patients with ICHD (37.7%) harbored a total of 91 CNV-US (min. 1, max. 3, median 1). The remaining 114 patients with ICHD (62.3%) had a normal CNV analysis.

The included patients and the results of their CNV analyses are depicted in Figure 1.

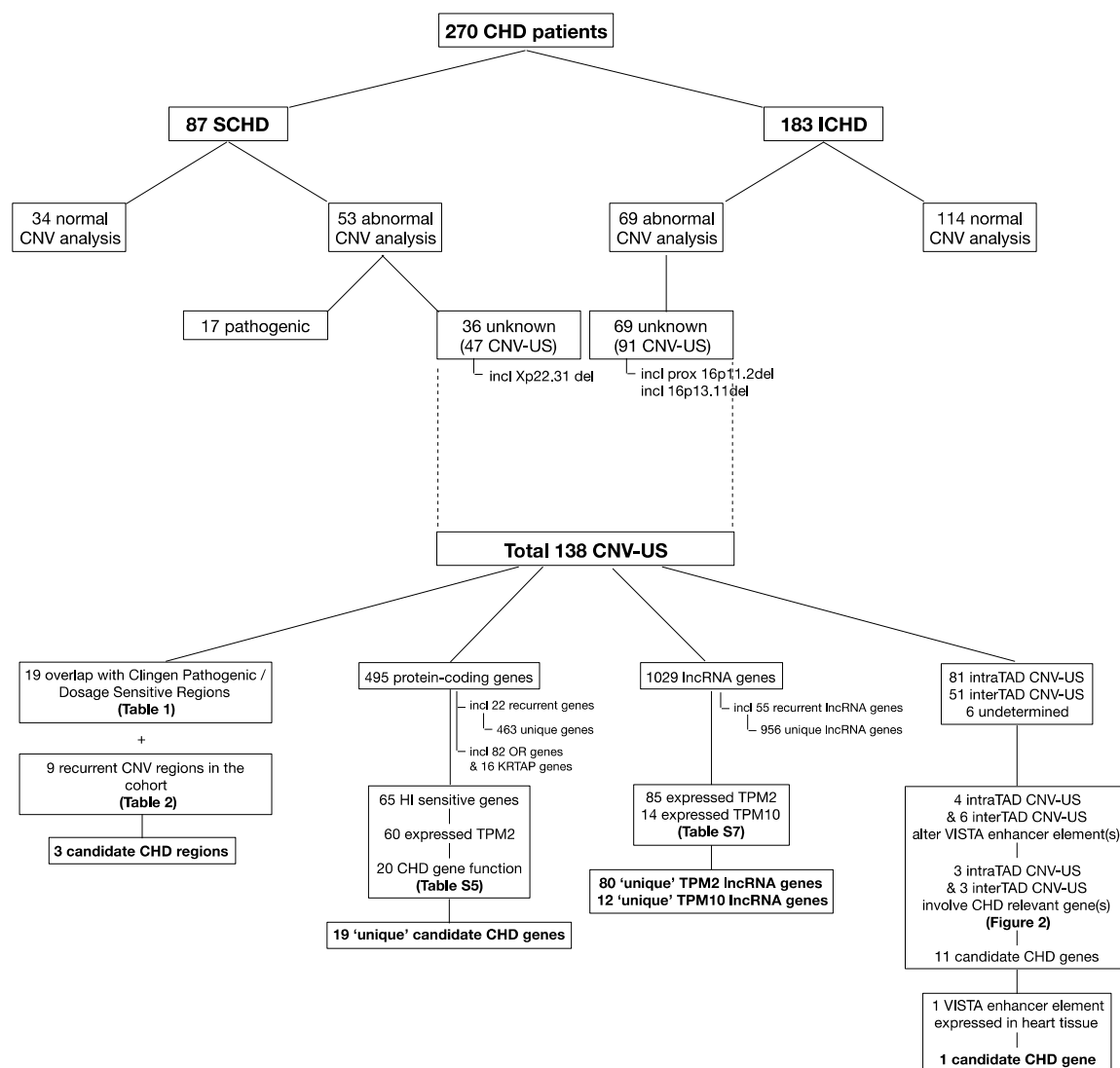


Figure 1. Structured overview of the study results. The upper part of the figure gives an overview of the included patients and the result of the copy number variation (CNV) analyses. The lower part of the figure depicts the flow and the final results of the explorative analyses performed in the CNV of unknown significance (CNV-US). CHD = congenital heart defects; SCHD = syndromic CHD; ICHD = isolated CHD; HI = haploinsufficiency; TPM = transcripts per million; TAD = topologically associated domain.

3.3. Comprehensive Reassessment of CNV-US

3.3.1. CNV Descriptives: CNV Type, Size and Parental Inheritance

The total study cohort contained 138 CNV-US for a comprehensive reevaluation, 47 CNV-US in the SCHD subgroup and 91 CNV-US in the ICHD subgroup. An overview of the CNV-US is given in Supplementary Table S4.

The CNV-US include 42 copy number losses (30.4%) (41 deletions and 1 mosaic deletion) and 96 copy number gains (69.6%) (90 duplications and six triplications). The proportion of copy number losses versus copy number gains is comparable between the SCHD and ICHD groups, respectively, 13 losses (27.7%) (including one mosaic deletion) and 34 gains (72.3%) (33 duplications and one triplication) in the SCHD group and 29 losses (31.9%) and 62 gains (68.1%) (57 duplications and five triplications) in the ICHD group.

The average length of the reported CNV-US is 417.76 kb (min. 6.66 kb, max. 5.08 Mb). The copy number gains are on average 394.99 kb (min. 6.66 kb, max. 2.06 Mb) and the copy number losses are on average 314.56 kb (min. 18.81 kb, max. 5.08 Mb). In the SCHD subgroup the average length of CNV-US is 409.05 kb (min. 19.83 kb, max. 2.06 Mb). In the ICHD subgroup the average length of CNV-US is 422.25 kb (min. 6.66 kb—max. 5.08 Mb).

For the 105 patients harboring at least one CNV-US, parental DNA was available for segregation analysis for 66 patients (62.9%). As such, parental inheritance was shown for 81 of the 138 CNV-US (58.7%). Six CNV-US (4.3%) occurred de novo. For the remaining 51 CNV-US (37%) information on parental inheritance was unavailable. The de novo CNV-US occurred in two patients with SCHD and four patients with ICHD, and included three copy number losses (two germline deletions and one mosaic deletion) and three copy number gains (two duplications and one triplication). The segregation data are included in Supplementary Table S8.

It is of note that X-linked CNVs in females may be particularly challenging to interpret due to possibly skewed X-inactivation.

3.3.2. Overlap with Known Dosage Sensitive Regions and Recurrence in the Study Cohort

In addition to the previously mentioned 16p11.2, 16p13.11, and Xp22.31 deletions, 16 other CNV-US show (partial) overlap with ClinGen Pathogenic and/or Dosage Sensitive CNV regions (Table 1). The contribution of 15q11.2 [41,42], 16p11.2 (MIM 611913), and 16p13.11 [43] microdeletions to CHD are debated. For the microdeletions in 2q13 [44] and 10q22.3q23.2 [45] and the duplication in 17p11.2 (MIM 610883), the suggested CHD candidate genes (respectively *BCL2L11*, *BMPRI1A*, and *RAI1*) are not present in these CNV-US regions.

Nine CNV regions show recurrence within the CHD cohort, of which two (Xp22.31 and 15q11.2 recurrent region) overlap with one of the ClinGen Pathogenic or Dosage Sensitive regions (Table 2). Based on the smallest region of overlap (SRO) of the recurrent CNV-US regions, the occurrence of comparable CNVs in Decipher (<https://decipher.sanger.ac.uk/>, accessed on 22 March 2021) [36] and/or the Database of Genomic Variants (DGV) (<http://dgv.tcag.ca/dgv/app/home>, accessed on 22 March 2021) [46], we retained the following three SRO as potential interesting candidate regions for CHD: chr9:107409509-107729796, chr21:47591379-47671404, and chrX/Y:61091-437220.

Table 1. Overlap of CNV-US with known ClinGen Pathogenic or Dosage Sensitive CNV regions.

CNV-US	Known Dosage Sensitive CNV Regions		Interpretation	
Chr1:145388355-145832995dup	1q21.1 recurrent (TAR) region (BP2-BP3, proximal) (ISCA-37428)	99.5% overlap	TS score 1	No definite link with CHD
Chr2:60998688-61093639dup	2p15p16.1 region (ISCA-37408)	2.8% overlap	TS score 1	No link with CHD
Chr2:112650001-112740000del	2q13 recurrent region (ISCA-37496)	5.3% overlap	HI score 2	Linked to CHD; excl CHD candidate gene <i>BCL2L11</i>
Chr5:1005001-1290000dup	5p15 terminal (Cri du chat syndrome) region (ISCA-37390)	2.5% overlap	TS score 2	No clear link with CHD
Chr10:84054763-84073574del	10q22.3q23.2 recurrent region (LCR-3/4-flanked) (ISCA-37424)	0.3% overlap	HI score 3	Linked to CHD; excl CHD candidate gene <i>BMPRI1A</i>
Chr10:88004601-88065186del	10q22.3q23.2 recurrent region (LCR-3/4-flanked) (ISCA-37424)	0.9% overlap	HI score 3	Linked to CHD; excl CHD candidate gene <i>BMPRI1A</i>
Chr15:22755001-23085000del	15q11.2 recurrent region (BP1-BP2) (ISCA-37448)	97.7% overlap	HI score 2	Link with CHD under debate
Chr15:22765628-23167699del	15q11.2 recurrent region (BP1-BP2) (ISCA-37448)	100% overlap	HI score 2	Link with CHD under debate
Chr15:22765628-23208842dup	15q11.2 recurrent region (BP1-BP2) (ISCA-37448)	100% overlap	TS score 40	-
Chr15:22765628-23208842dup	15q11.2 recurrent region (BP1-BP2) (ISCA-37448)	100% overlap	TS score 40	-
Chr15:24005491-24470088dup	15q11q13 recurrent (PWS/AS) region (BP2-BP3, Class 1) (ISCA-37404)	10.0% overlap	TS score 3	No clear link with CHD
Chr16:14968855-16292181del	16p13.11 recurrent region (BP2-BP3) (ISCA-37415)	100% overlap	HI score 3	Link with CHD under debate
Chr16:29656684-30197290del	16p11.2 recurrent region (proximal, BP4-BP5) (ISCA-37400)	98.3% overlap	HI score 3	Link with CHD under debate
Chr17:15257416-15482813dup	17p12 recurrent (HNPP/CMT1A) region (ISCA-37436)	12.5% overlap	TS score 3	No link with CHD
Chr17:18148172-18662098dup	17p11.2 recurrent (SMS/PLS) region (ISCA-37418)	15.1% overlap	TS score 3	Linked to CHD; excl CHD candidate gene <i>RAI1</i>
Chr17:58372095-58588996dup	17q23.1q23.2 recurrent region (ISCA-37501)	10.0% overlap	TS score 2	Link with CHD unclear
ChrX:6467006-8131751del *	Xp22.31 recurrent region (ISCA-37417)	99.3% overlap	HI score 3	No link with CHD
ChrX:6467006-8131751dup *	Xp22.31 recurrent region (ISCA-37417)	99.3% overlap	TS score 40	-
ChrX:7515001-8130000dup	Xp22.31 recurrent region (ISCA-37417)	36.5% overlap	TS score 40	-

HI = haploinsufficiency; TS = triplosensitivity; score 1 = little evidence; score 2 = emerging evidence; score 3 = sufficient evidence; score 40 = dosage sensitivity unlikely; X-chromosomal CNV-US occurring in females are marked with *. All CNV-US were mapped to reference genome GRCh37 (hg19).

Table 2. Recurrent CNV-US regions in the study cohort.

CNV-US	Smallest Region of Overlap (SRO) (Protein-Coding Genes)	Interpretation
Chr4:135455435-137460949dup Chr4:135700662-135829279dup	Chr4:135700662-135829279 (no protein-coding genes)	One nearly identical duplication in Decipher, classified 'likely benign'. Two partially overlapping duplications in the same boundaries in DGV.
Chr7:11221210-12462629dup Chr7:12300173-12462629del	Chr7:12300173-12462629 (<i>VWDE</i>)	Two comparable deletions in Decipher, both classified 'CNV-US'. Multiple comparable CNVs in DGV.
Chr9:195001-405000dup Chr9:210001-540000dup	Chr9:210001-405000 (<i>DOCK8</i>)	Multiple comparable CNVs in Decipher, most classified 'CNV-US'. Multiple comparable CNVs in DGV.
Chr9:107409506-107729796dup Chr9:107409509-107769094dup	Chr9:107409509-107729769 (<i>OR13D1, NIPSNAP3A, NIPSNAP3B, ABCA1</i>)	Four comparable duplications in Decipher, three classified 'CNV-US', one classified 'likely pathogenic'. No comparable CNVs in DGV.
Chr15:22755001-23085000del Chr15:22765628-23167699del Chr15:22765628-23208842dup Chr15:22765628-23208842dup	Chr15:22765628-23085000 (<i>CYFIP1, NIPA1, NIPA2, TUBGCP5</i>)	Known ClinGen Pathogenic and Dosage Sensitive CNV region (ISCA-37448).
Chr21:43014314-48090258del Chr21:47591379-47671404dup	Chr21:47591379-47671404 (<i>SPATC1L, LSS, MCM3AP</i>)	No comparable CNVs in Decipher. No comparable CNVs in DGV.
ChrX:61091-437220del * ChrY:61091-819199del	ChrX/Y:61091-437220 (<i>PLCXD1, GTPBP6, PPP2R3B</i>)	No comparable CNVs in Decipher. No comparable CNVs in DGV.
ChrX:6467006-8131751dup * ChrX:6467006-8131751del * ChrX:7515001-8130000dup	ChrX:7515001-8130000 (<i>VCX, PNPLA4</i>)	Multiple comparable CNVs in Decipher, all but one classified 'CNV-US'. Few comparable CNVs in DGV.
ChrX:130610000-130950000dup ChrX:130631863-130960558dup *	ChrX:130631863-130950000 (<i>OR13H1</i>)	Multiple comparable CNVs in Decipher, most classified 'CNV-US' or 'benign'. No comparable CNVs in DGV.

Decipher = Decipher database of CNV entries (<https://decipher.sanger.ac.uk/>, accessed on 22 March 2021); DGV = Database of Genomic Variants (<http://dgv.tcag.ca/dgv/app/home>, accessed on 22 March 2021); Potential candidate CHD regions are marked in bold. X-chromosomal CNV-US occurring in females are marked with *. All CNV-US were mapped to reference genome GRCh37 (hg19).

3.3.3. Protein-Coding Gene Content

Based on the protein-coding gene content, the 138 CNV-US can be subdivided into 116 coding CNV-US and 22 non-coding CNV-US. The 116 coding CNV-US affected a total of 495 NCBI RefSeq Select genes, ranging from 1 to 79 NCBI RefSeq Select genes per CNV-US (mean 4.27, standard deviation 9.21, median 2). Of these, 162 genes were entirely deleted by a CNV-US (seven genes resided in a mosaic deletion), 227 were duplicated, 9 were triplicated, and 97 genes were disrupted by a CNV breakpoint and thus partially deleted (25, including one mosaic), duplicated (67) or triplicated (5). Twenty-two genes were recurrently affected in two or more CNV-US, resulting in 463 uniquely affected genes. In the further results, we considered the recurrent genes as separate counts, because these were sometimes disrupted differently in different CNV-US. The 82 olfactory receptor (*OR*) genes and the 16 keratin associated protein (*KRTAP*) genes affected in the CNV-US, were disregarded for further evaluation in the context of CHD.

Evaluation of dosage sensitivity indicated that for 65 genes haploinsufficiency may be harmful. These included 24 deleted genes and 41 genes disrupted by a CNV breakpoint (10 partial gene deletions, 28 partial gene duplications and 3 partial gene triplications). Especially deletions and partial deletions are more likely to be disruptive than duplications or triplications. None of the duplicated genes are known to be triplosensitive. Of these haploinsufficiency sensitive genes, 60 genes (23 deleted genes and 37 breakpoint genes) show expression in developing human heart, with 20 genes (9 deleted genes (including one recurrent gene) and 11 breakpoint genes) (*ARFGEF2*, *AUTS2*, *CHAMP1*, *CHD8*, *CYFIP1* (2X), *FERMT2*, *ITCH*, *KMT2C*, *MAPK3*, *MAZ*, *MYH11*, *NASP*, *NPAS2*, *PIK3C3*, *PKNOX1*, *TIA1*, *TJP1*, *TRIM28*, *ZBTB21*) having a function that can be related to heart development or CHD pathogenesis. These genes are present in 18 different CNV-US, of which three occurred de novo (containing *AUTS2*, *CHD8* and *TRIM28*). These genes and their potential link to heart development and/or CHD (including the most relevant results from an additional Pubmed search looking for associations of these genes with (congenital) heart disease and/or heart development) are represented in Supplementary Table S5 [47–63].

An overview of the dosage sensitivity, expression in developing human heart, and gene function related to heart development or CHD of all 495 NCBI RefSeq Select genes in the CNV-US is given in Supplementary Table S6.

3.3.4. lncRNAs

The 138 CNV-US contained a total of 1029 lncRNA genes, with a range from 0 to 193 lncRNA genes per CNV-US (mean 7.46, standard deviation 17.25, median 4). Of these, 425 resided in deletions (eight in the mosaic deletion), 576 in duplications, and 28 in triplications. Fifty-five lncRNA genes appear in more than one CNV-US, resulting in a total of 956 uniquely affected lncRNA genes. In the further results, the recurrent lncRNA genes were all counted separately.

Evaluation of expression in developing human heart tissue (after conversion to ENSG IDs) indicated that 85 of all lncRNA genes in the CNV-US (one lncRNA gene occurred three times and three lncRNA genes occurred twice, thus 80 unique lncRNA genes) reach at least 2 TPM in the developing human heart, and 14 lncRNA genes (one lncRNA gene occurred three times, thus 12 unique lncRNA genes) even reach 10 TPM. The 85 lncRNA genes with expression in developing human heart are present in 38 of the 138 CNV-US. An overview of these lncRNA genes is given in Supplementary Table S7.

3.3.5. Enhancers and Interference with TAD-Related Gene-Enhancer Interactions

Of the 138 CNV-US, 81 are intraTAD CNVs (25 deletions, 52 duplications and four triplications) and 51 are interTAD CNVs (13 deletions, 37 duplications and one triplication). Six CNV-US occur in chromosome regions where no clear TAD regions/boundaries are determined (four at a telomere end of a chromosome and two at the Y chromosome).

Two intraTAD deletions and two intraTAD duplications alter the dosage of a VISTA enhancer element (elements 1, 1660, 92, and 742), thereby potentially impacting the ex-

pression of target gene(s) within the TAD. Four of these genes are considered potentially CHD-relevant, occurring in two intraTAD deletions (*FOXF1*, *IRF8*, *AUTS2*) and in one intraTAD duplication (*ZFHX4*). A schematic representation of these latter three CNV-US and their potential impact on these gene-enhancer interactions is given in Figure 2. It should be noted, however, that none of the VISTA enhancer elements involved here are shown to be expressed in the heart (evaluated in transgenic mice at embryonic day 11.5 (E11.5)).

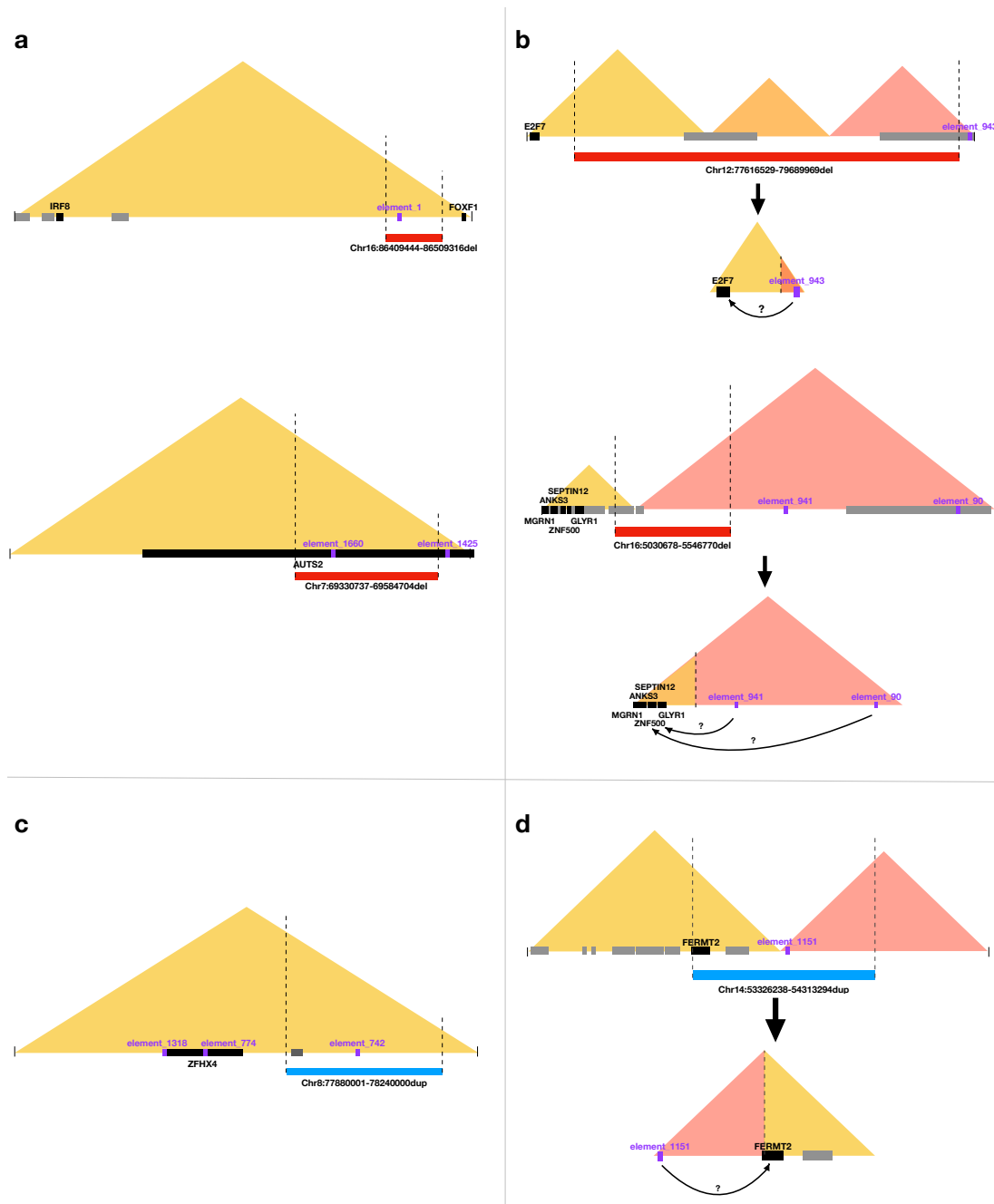


Figure 2. Schematic representation of the intraTAD deletions (a), interTAD deletions (b), intraTAD duplications (c), and interTAD duplications (d), potentially altering gene-enhancer interactions relevant for the congenital heart defects (CHD). The deletions are shown as red bars. The duplications are shown as blue bars. The triangles reflect the topologically associated domain (TAD) structure of the locus. CHD candidate protein-coding genes are depicted in black, other protein-coding genes within the TAD are depicted in grey, VISTA enhancer elements are depicted in purple. It is of note that only for VISTA element 1151 expression in the heart was confirmed (in transgenic mice at E11.5).

Four interTAD deletions and two interTAD duplications result in the formation of a neoTAD, involving a VISTA enhancer element (elements 943, 90, 941, 799, 305, 852, 1151, 1320, and 184), thereby potentially altering the expression of target gene(s) within the neoTAD. Seven of these genes are of interest in relation to CHD pathogenesis. They are related to two interTAD deletions (*E2F2*, *ANKS3*, *GLYR1*, *MGRN1*, *SEPTIN12*, *ZNF500*) and one interTAD duplication (*FERMT2*). However, the latter can only be if the duplication occurred in tandem. A schematic representation of the latter three neoTADs and the potential novel gene-enhancer interactions is given in Figure 2. It is of note that only one of the VISTA enhancer elements (element 1151) was shown to be expressed in the heart (evaluated in transgenic mice at E11.5).

An overview of the relevant NCBI RefSeq Select genes in the CNV-US related TADs and their gene function related to heart development or CHD is added to Supplementary Table S6.

A comprehensive overview of the final results of the explorative re-analysis of the CNV-US is given in Figure 1. Figure 3 depicts all CNV-US with their relevant results on a structured chromosome map.

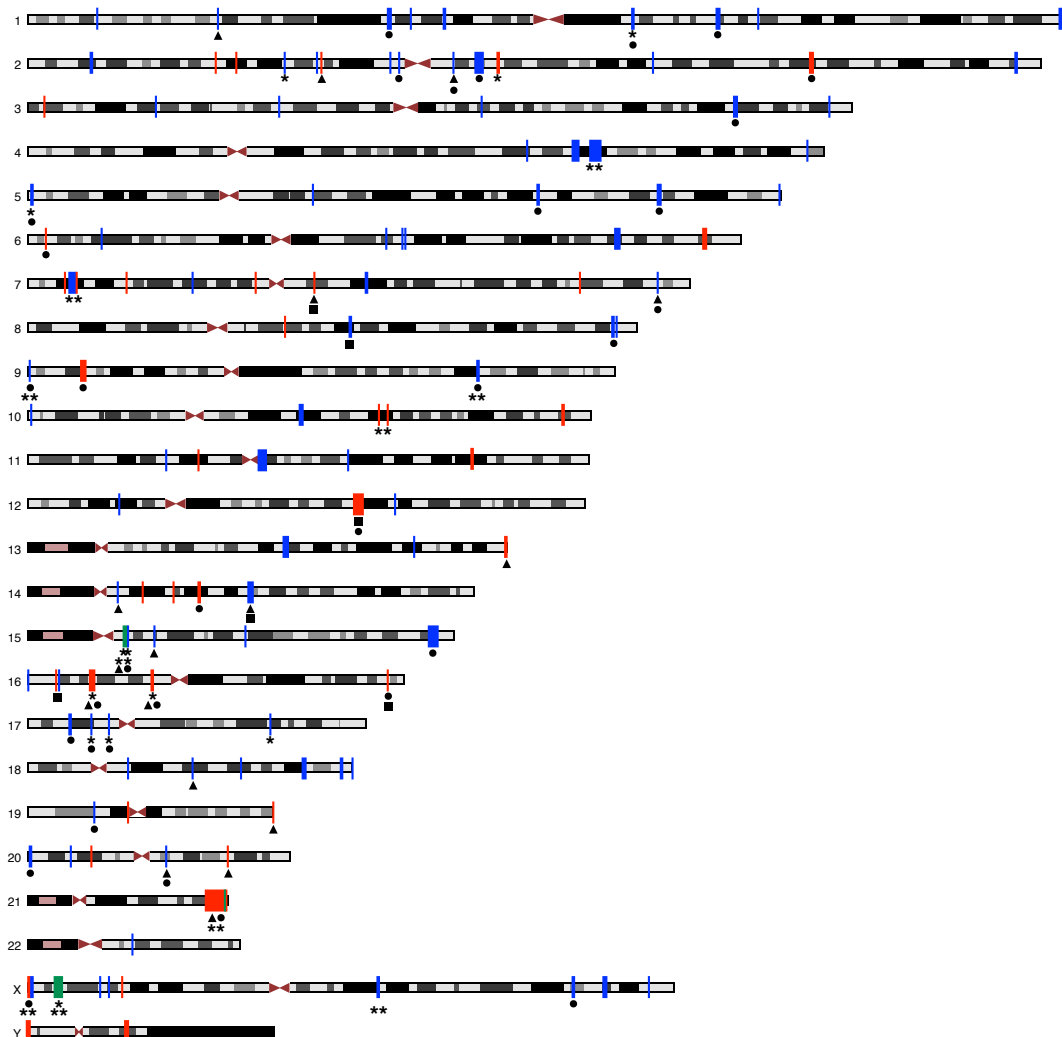


Figure 3. Overview of the copy number variations (CNV) of unknown significance (CNV-US) and the relevant results from our comprehensive re-analysis on a schematic chromosome map. Deletions are red, duplications are blue, regions containing both deletions and duplications are green. Depicted are CNV-US overlapping with a pathogenic region (*), recurrent regions in the study cohort (**), CNV-US containing one or more candidate congenital heart defect (CHD) protein-coding genes (▲), CNV-US containing one or more lncRNA genes expressed in developing human heart (●), and CNV-US potentially disrupting topologically associated domain (TAD)-related gene-enhancer interactions potentially relevant for CHD (■).

An overview of the protein-coding gene content, lncRNA genes, VISTA enhancer elements, and TAD interference for all studied CNV-US is given in Supplementary Table S8.

4. Discussion

In this retrospective study, we reviewed reported CNVs in an in-house cohort of 270 patients with CHD, and re-evaluated the CNV-US considering preliminary evidence of alternative non-coding regulatory elements and the 3D genomic TAD structures.

The number of SCHD was lower in our cohort than in previously reported series, which could explain the lower yield for pathogenic CNVs (6.8%) than previously reported (12.8% to 18.5%) [26]. The 22q11.2 deletion and Williams-Beuren syndrome were the most frequent SCHD entities.

The CNV-US were nearly equally distributed between the SCHD and ICHD subgroup, with copy number gains more than twice as frequent as copy number losses. Over 20% of the CNV-US was larger than 500 kb, which clearly surpasses the number of large CNVs (>500 kb) in the general population (8%) [12]. Twenty-nine of the reported CNV-US (21%) are <100 kb (three having a borderline size of 98.0 kb, 99.4 kb, and 99.9 kb). According to the in-house reporting protocol CNVs < 100 kb are only reported when containing gene-related sequences, though four of these small CNV-US did not contain any protein-coding sequences at all. At least six CNV-US (4.3%) in our cohort occurred de novo. This number might be underestimated, since segregation data were incomplete for over one third of patients. Warburton et al. reported 8% to 12.7% de novo CNVs in a cohort of probands with conotruncal defects or hypoplastic left heart disease, which was considerably higher than the approximately 2% de novo CNVs in a control cohort of trios unaffected with CHD obtained from the Simons Simplex Collection [64,65]. Altogether, this suggests that de novo CNVs may contribute to cardiac phenotypes, but does not preclude the contribution of inherited CNVs, especially in the view of multifactorial disease mechanisms, subtle subclinical phenotypes such as patent foramen ovale, bicuspid aortic valves or spontaneously closed septal defects, or incomplete penetrance.

Our in silico analysis yielded interesting findings for 52 of the 138 CNV-US. We retained three candidate CHD regions from the SRO: chr9:107409509-107729796, chr21:47591379-47671404, and chrX/Y:61091-437220. Evaluation of the protein-coding gene content, based on dosage sensitivity, expression during heart development, and gene function, yielded 19 candidate CHD genes: *ARFGEF2*, *AUTS2*, *CHAMP1*, *CHD8*, *CYFIP1*, *FERMT2*, *ITCH*, *KMT2C*, *MAPK3*, *MAZ*, *MYH11*, *NASP*, *NPAS2*, *PIK3C3*, *PKNOX1*, *TIA1*, *TJP1*, *TRIM28*, *ZBTB21*. At this time, an exhaustive literature search could not provide conclusive evidence for a certain (causal) relationship of these genes with CHD. lncRNA gene analysis confirmed FENRR as a CHD candidate gene, which has already previously been linked to heart development [66]. For most other lncRNA genes, their functional role remains to be uncovered, precluding suggestions regarding their role in CHD pathogenesis. We further identified six interesting potential interferences with TADs and associated candidate CHD genes. It must be emphasized that only one of the involved VISTA enhancer elements was shown to be expressed in mouse heart at E11.5. Unfortunately, extensive data of expression at other embryonic time points are not available. Moreover, our analysis is restricted to the limited set of experimentally validated VISTA enhancers, and did not evaluate many other enhancers elements that have been reported by initiatives such as ENCODE (<https://www.encodeproject.org/>) and ROADMAP (<http://www.roadmapgenomics.org/>).

Altogether, our explorative in silico analysis raises interest with regards to different potential CNV pathogenic mechanisms interfering with gene expression of protein-coding genes, either through direct copy number alterations of the coding sequence or through interference with lncRNA genes or the 3D genome structure. Our findings ask for experimental validation in future studies, including in vivo and in vitro model systems, expression studies and chromosome conformation capture based analyses such as Hi-C [67,68]. The latter technique has the advantage of allowing both the detection and the

interpretation of CNVs [69]. With time, a more complete annotation of both the coding and non-coding genome, and a better understanding of the complex gene regulatory landscape, will aid in estimating the phenotypic consequences of both de novo and (combinations of) inherited variants with smaller individual effects [68].

In patients with SCHD, the diagnostic yield of 19.5% underscores the need for continued molecular karyotyping in a clinical setting. In patients with ICHD, the diagnostic yield remains disputable, mainly due to the large amount of CNV-US identified. In this study, we identified 52 CNV-US of interest, of which 16 CNV-US occurred in 14 SCHD cases (16.1%) and 36 CNV-US occurred in 32 ICHD cases (17.5%). Our *in silico* analysis shows that the mechanisms affecting non-coding sequences, if confirmed, might substantially contribute to CHD, and will likely increase the diagnostic yield of molecular karyotyping in both SCHD and ICHD. Meanwhile, counseling CNV-US remains challenging. At least, in view of the significantly improved reproductive fitness of patients with CHD, it could be advised to reinterpret the CNV-US at childbearing age. In addition, our data indicate that reporting CNV-US in international databases remains valuable, even if they only contain non-coding sequences.

5. Conclusions

We performed a comprehensive reassessment of CNV-US in a large CHD cohort, identifying novel candidate genes for CHD and proving preliminary data for different potentially pathogenic CNV mechanisms interfering with the expression of protein-coding genes, warranting functional validation. Our data call for a reinterpretation of CNV-US in the context of CHD.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12071048/s1>, Table S1: CHD-related gene panels, Table S2: Heart defects in the study cohort, Table S3: Pathogenic CNVs in the study cohort, Table S4: CNV-US in the study cohort and subgroups, Table S5: Candidate CHD protein-coding genes in CNV-US, Table S6: Overview of the characteristics of protein-coding genes in CNV-US and TADs of interest, Table S7: LncRNA genes of interest in CNV-US, Table S8: Overview of the studied characteristics per CNV-US.

Author Contributions: Conceptualization, I.M., S.V. and B.C.; data curation, I.M., A.D. and B.M.; formal analysis, I.M. and S.V.; funding acquisition, P.J.C. and B.C.; investigation, I.M., S.V. and B.C.; methodology, I.M., S.V. and B.C.; project administration, I.M. and B.C.; resources, I.M., A.D., B.M., K.D.G., H.D.W., L.M.M., J.P., K.V., D.D.W. and B.C.; supervision, B.C. and P.J.C.; validation, I.M., S.V. and B.C.; visualization, I.M., S.V. and L.M.M.; writing—original draft, I.M.; writing—review and editing, I.M., S.V., A.D., B.M., K.D.G., H.D.W., L.M.M., J.P., K.V., P.J.C., D.D.W. and B.C. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by grant G028415N of the Research Foundation Flanders (FWO) to B.C. and P.J.C., Methusalem grant BOFMET2015000401 from the Special Research Fund of Ghent University, and starting grant BOF16/STA/045 from the Special Research Fund of Ghent University Hospital to B.C.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Ghent University Hospital (EC UZG 2016/0133, Date of Approval 3 February 2016).

Informed Consent Statement: Patient consent was waived due to the retrospective study design and the de-identification of the patients.

Data Availability Statement: The data presented in this study are available in the manuscript or in the Supplementary Materials, or can be obtained from the authors upon written request to the corresponding author.

Acknowledgments: This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from <https://deciphergenomics.org/about/stats> and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Feuk, L.; Carson, A.R.; Scherer, S.W. Structural variation in the human genome. *Nat. Rev. Genet.* **2006**, *7*, 85–97. [[CrossRef](#)]
2. Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shapero, M.H.; Carson, A.R.; Chen, W.; et al. Global variation in copy number in the human genome. *Nature* **2006**, *444*, 444–454. [[CrossRef](#)]
3. Vissers, L.E.; de Vries, B.B.; Osoegawa, K.; Janssen, I.M.; Feuth, T.; Choy, C.O.; Straatman, H.; van der Vliet, W.; Huys, E.H.; van Rijk, A.; et al. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am. J. Hum. Genet.* **2003**, *73*, 1261–1270. [[CrossRef](#)] [[PubMed](#)]
4. Shaw-Smith, C.; Redon, R.; Rickman, L.; Rio, M.; Willatt, L.; Fiegler, H.; Firth, H.; Sanlaville, D.; Winter, R.; Colleaux, L.; et al. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J. Med. Genet.* **2004**, *41*, 241–248. [[CrossRef](#)] [[PubMed](#)]
5. Devriendt, K.; Vermeesch, J.R. Chromosomal phenotypes and submicroscopic abnormalities. *Hum. Genom.* **2004**, *1*, 126–133. [[CrossRef](#)] [[PubMed](#)]
6. Menten, B.; Maas, N.; Thienpont, B.; Buysse, K.; Vandesompele, J.; Melotte, C.; de Ravel, T.; Van Vooren, S.; Balikova, I.; Backx, L.; et al. Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: A new series of 140 patients and review of published reports. *J. Med. Genet.* **2006**, *43*, 625–633. [[CrossRef](#)]
7. Miller, D.T.; Adam, M.P.; Aradhya, S.; Biesecker, L.G.; Brothman, A.R.; Carter, N.P.; Church, D.M.; Crolla, J.A.; Eichler, E.E.; Epstein, C.J.; 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. [[CrossRef](#)]
8. Southard, A.E.; Edelmann, L.J.; Gelb, B.D. Role of copy number variants in structural birth defects. *Pediatrics* **2012**, *129*, 755–763. [[CrossRef](#)] [[PubMed](#)]
9. Buysse, K.; Delle Chiaie, B.; Van Coster, R.; Loeys, B.; De Paepe, A.; Mortier, G.; Speleman, F.; Menten, B. Challenges for CNV interpretation in clinical molecular karyotyping: Lessons learned from a 1001 sample experience. *Eur. J. Med. Genet.* **2009**, *52*, 398–403. [[CrossRef](#)]
10. Gijbbers, A.C.; Schoumans, J.; Ruivenkamp, C.A. Interpretation of array comparative genome hybridization data: A major challenge. *Cytogenet. Genome Res.* **2011**, *135*, 222–227. [[CrossRef](#)]
11. Breckpot, J.; Thienpont, B.; Arens, Y.; Tranchevent, L.C.; Vermeesch, J.R.; Moreau, Y.; Gewillig, M.; Devriendt, K. Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenet Genome Res.* **2011**, *135*, 251–259. [[CrossRef](#)]
12. Nowakowska, B. Clinical interpretation of copy number variants in the human genome. *J. Appl. Genet.* **2017**, *58*, 449–457. [[CrossRef](#)]
13. Qiu, M.T.; Hu, J.W.; Yin, R.; Xu, L. Long noncoding RNA: An emerging paradigm of cancer research. *Tumour Biol.* **2013**, *34*, 613–620. [[CrossRef](#)]
14. Valton, A.L.; Dekker, J. TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* **2016**, *36*, 34–40. [[CrossRef](#)]
15. Lupiáñez, D.G.; Kraft, K.; Heinrich, V.; Krawitz, P.; Brancati, F.; Klopocki, E.; Horn, D.; Kayserili, H.; Opitz, J.M.; Laxova, R.; et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **2015**, *161*, 1012–1025. [[CrossRef](#)]
16. Spielmann, M.; Mundlos, S. Looking beyond the genes: The role of non-coding variants in human disease. *Hum. Mol. Genet.* **2016**, *25*, R157–R165. [[CrossRef](#)] [[PubMed](#)]
17. Thienpont, B.; Mertens, L.; de Ravel, T.; Eyskens, B.; Boshoff, D.; Maas, N.; Fryns, J.P.; Gewillig, M.; Vermeesch, J.R.; Devriendt, K. Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. *Eur. Heart J.* **2007**, *28*, 2778–2784. [[CrossRef](#)] [[PubMed](#)]
18. Erdogan, F.; Larsen, L.A.; Zhang, L.; Tümer, Z.; Tommerup, N.; Chen, W.; Jacobsen, J.R.; Schubert, M.; Jurkatis, J.; Tzschach, A.; et al. High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J. Med. Genet.* **2008**, *45*, 704–709. [[CrossRef](#)]
19. Richards, A.A.; Santos, L.J.; Nichols, H.A.; Crider, B.P.; Elder, F.F.; Hauser, N.S.; Zinn, A.R.; Garg, V. Cryptic chromosomal abnormalities identified in children with congenital heart disease. *Pediatric Res.* **2008**, *64*, 358–363. [[CrossRef](#)] [[PubMed](#)]
20. Greenway, S.C.; Pereira, A.C.; Lin, J.C.; DePalma, S.R.; Israel, S.J.; Mesquita, S.M.; Ergul, E.; Conta, J.H.; Korn, J.M.; McCarroll, S.A.; et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat. Genet.* **2009**, *41*, 931–935. [[CrossRef](#)] [[PubMed](#)]
21. Breckpot, J.; Thienpont, B.; Peeters, H.; de Ravel, T.; Singer, A.; Rayyan, M.; Allegaert, K.; Vanhole, C.; Eyskens, B.; Vermeesch, J.R.; et al. Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. *J. Pediatric* **2010**, *156*, 810–817.e811–817.e814. [[CrossRef](#)] [[PubMed](#)]

22. Soemedi, R.; Wilson, I.J.; Bentham, J.; Darlay, R.; Töpf, A.; Zelenika, D.; Cosgrove, C.; Setchfield, K.; Thornborough, C.; Granados-Riveron, J.; et al. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am. J. Hum. Genet.* **2012**, *91*, 489–501. [[CrossRef](#)] [[PubMed](#)]
23. Fahed, A.C.; Gelb, B.D.; Seidman, J.G.; Seidman, C.E. Genetics of congenital heart disease: The glass half empty. *Circ. Res.* **2013**, *112*, 707–720. [[CrossRef](#)] [[PubMed](#)]
24. Zaidi, S.; Brueckner, M. Genetics and Genomics of Congenital Heart Disease. *Circ. Res.* **2017**, *120*, 923–940. [[CrossRef](#)]
25. Andersen, T.A.; Troelsen Kde, L.; Larsen, L.A. Of mice and men: Molecular genetics of congenital heart disease. *Cell. Mol. Life Sci.* **2014**, *71*, 1327–1352. [[CrossRef](#)]
26. Geng, J.; Picker, J.; Zheng, Z.; Zhang, X.; Wang, J.; Hisama, F.; Brown, D.W.; Mullen, M.P.; Harris, D.; Stoler, J.; et al. Chromosome microarray testing for patients with congenital heart defects reveals novel disease causing loci and high diagnostic yield. *BMC Genom.* **2014**, *15*, 1127. [[CrossRef](#)]
27. Menten, B.; Pattyn, F.; De Preter, K.; Robbrecht, P.; Michels, E.; Buysse, K.; Mortier, G.; De Paepe, A.; van Vooren, S.; Vermeesch, J.; et al. arrayCGHbase: An analysis platform for comparative genomic hybridization microarrays. *BMC Bioinform.* **2005**, *6*, 124. [[CrossRef](#)]
28. Sante, T.; Vergult, S.; Volders, P.J.; Kloosterman, W.P.; Trooskens, G.; De Preter, K.; Dheedene, A.; Speleman, F.; De Meyer, T.; Menten, B. ViVar: A comprehensive platform for the analysis and visualization of structural genomic variation. *PLoS ONE* **2014**, *9*, e113800. [[CrossRef](#)]
29. Raman, L.; Dheedene, A.; De Smet, M.; Van Dorpe, J.; Menten, B. WisecondorX: Improved copy number detection for routine shallow whole-genome sequencing. *Nucleic Acids Res.* **2019**, *47*, 1605–1614. [[CrossRef](#)]
30. Rehm, H.L.; Berg, J.S.; Brooks, L.D.; Bustamante, C.D.; Evans, J.P.; Landrum, M.J.; Ledbetter, D.H.; Maglott, D.R.; Martin, C.L.; Nussbaum, R.L.; et al. ClinGen—the Clinical Genome Resource. *N. Engl. J. Med.* **2015**, *372*, 2235–2242. [[CrossRef](#)]
31. O’Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)] [[PubMed](#)]
32. Volders, P.J.; Anckaert, J.; Verheggen, K.; Nuytens, J.; Martens, L.; Mestdagh, P.; Vandesompele, J. LNCipedia 5: Towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* **2019**, *47*, D135–D139. [[CrossRef](#)] [[PubMed](#)]
33. Visel, A.; Minovitsky, S.; Dubchak, I.; Pennacchio, L.A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* **2007**, *35*, D88–D92. [[CrossRef](#)]
34. Dixon, J.R.; Selvaraj, S.; Yue, F.; Kim, A.; Li, Y.; Shen, Y.; Hu, M.; Liu, J.S.; Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **2012**, *485*, 376–380. [[CrossRef](#)]
35. Spector, J.D.; Wiita, A.P. ClinTAD: A tool for copy number variant interpretation in the context of topologically associated domains. *J. Hum. Genet.* **2019**, *64*, 437–443. [[CrossRef](#)] [[PubMed](#)]
36. Firth, H.V.; Richards, S.M.; Bevan, A.P.; Clayton, S.; Corpas, M.; Rajan, D.; Van Vooren, S.; Moreau, Y.; Pettett, R.M.; Carter, N.P. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **2009**, *84*, 524–533. [[CrossRef](#)] [[PubMed](#)]
37. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)] [[PubMed](#)]
38. Cardoso-Moreira, M.; Halbert, J.; Valloton, D.; Velten, B.; Chen, C.; Shao, Y.; Liechti, A.; Ascensão, K.; Rummel, C.; Ovchinnikova, S.; et al. Gene expression across mammalian organ development. *Nature* **2019**, *571*, 505–509. [[CrossRef](#)]
39. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R.; Weirauch, M.T. The Human Transcription Factors. *Cell* **2018**, *172*, 650–665. [[CrossRef](#)]
40. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)] [[PubMed](#)]
41. Burnside, R.D.; Pasion, R.; Mikhail, F.M.; Carroll, A.J.; Robin, N.H.; Youngs, E.L.; Gadi, I.K.; Keitges, E.; Jaswaney, V.L.; Papenhausen, P.R.; et al. Microdeletion/microduplication of proximal 15q11.2 between BP1 and BP2: A susceptibility region for neurological dysfunction including developmental and language delay. *Hum. Genet.* **2011**, *130*, 517–528. [[CrossRef](#)]
42. Vanlerberghe, C.; Petit, F.; Malan, V.; Vincent-Delorme, C.; Bouquillon, S.; Boute, O.; Holder-Espinasse, M.; Delobel, B.; Duban, B.; Vallee, L.; et al. 15q11.2 microdeletion (BP1-BP2) and developmental delay, behaviour issues, epilepsy and congenital heart disease: A series of 52 patients. *Eur. J. Med. Genet.* **2015**, *58*, 140–147. [[CrossRef](#)]
43. Hannes, F.D.; Sharp, A.J.; Mefford, H.C.; de Ravel, T.; Ruivenkamp, C.A.; Breuning, M.H.; Fryns, J.P.; Devriendt, K.; Van Buggenhout, G.; Vogels, A.; et al. Recurrent reciprocal deletions and duplications of 16p13.11: The deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *J. Med. Genet.* **2009**, *46*, 223–232. [[CrossRef](#)]
44. Riley, K.N.; Catalano, L.M.; Bernat, J.A.; Adams, S.D.; Martin, D.M.; Lalani, S.R.; Patel, A.; Burnside, R.D.; Innis, J.W.; Rudd, M.K. Recurrent deletions and duplications of chromosome 2q11.2 and 2q13 are associated with variable outcomes. *Am. J. Med Genet. Part A* **2015**, *167A*, 2664–2673. [[CrossRef](#)] [[PubMed](#)]

45. van Bon, B.W.; Balciuniene, J.; Fruhman, G.; Nagamani, S.C.; Broome, D.L.; Cameron, E.; Martinet, D.; Roulet, E.; Jacquemont, S.; Beckmann, J.S.; et al. The phenotype of recurrent 10q22q23 deletions and duplications. *Eur. J. Hum. Genet. EJHG* **2011**, *19*, 400–408. [[CrossRef](#)]
46. MacDonald, J.R.; Ziman, R.; Yuen, R.K.; Feuk, L.; Scherer, S.W. The Database of Genomic Variants: A curated collection of structural variation in the human genome. *Nucleic Acids Res.* **2014**, *42*, 986–992. [[CrossRef](#)]
47. Beunders, G.; van de Kamp, J.; Vasudevan, P.; Morton, J.; Smets, K.; Kleefstra, T.; de Munnik, S.A.; Schuurs-Hoeijmakers, J.; Ceulemans, B.; Zollino, M.; et al. A detailed clinical analysis of 13 patients with AUTS2 syndrome further delineates the phenotypic spectrum and underscores the behavioural phenotype. *J. Med. Genet.* **2016**, *53*, 523–532. [[CrossRef](#)] [[PubMed](#)]
48. Szot, J.O.; Cuny, H.; Blue, G.M.; Humphreys, D.T.; Ip, E.; Harrison, K.; Sholler, G.F.; Giannoulatou, E.; Leo, P.; Duncan, E.L.; et al. A Screening Approach to Identify Clinically Actionable Variants Causing Congenital Heart Disease in Exome Data. *Circ. Genom. Precis. Med.* **2018**, *11*, e001978. [[CrossRef](#)] [[PubMed](#)]
49. Batsukh, T.; Pieper, L.; Koszucka, A.M.; von Velsen, N.; Hoyer-Fender, S.; Elbracht, M.; Bergman, J.E.; Hoefsloot, L.H.; Pauli, S. CHD8 interacts with CHD7, a protein which is mutated in CHARGE syndrome. *Hum. Mol. Genet.* **2010**, *19*, 2858–2866. [[CrossRef](#)]
50. Shanks, M.O.; Lund, L.M.; Manni, S.; Russell, M.; Mauban, J.R.; Bond, M. Chromodomain helicase binding protein 8 (Chd8) is a novel A-kinase anchoring protein expressed during rat cardiac development. *PLoS ONE* **2012**, *7*, e46316. [[CrossRef](#)] [[PubMed](#)]
51. Wu, J.C.; Tsai, R.Y.; Chung, T.H. Role of catenins in the development of gap junctions in rat cardiomyocytes. *J. Cell. Biochem.* **2003**, *88*, 823–835. [[CrossRef](#)] [[PubMed](#)]
52. Toyofuku, T.; Yabuki, M.; Otsu, K.; Kuzuya, T.; Hori, M.; Tada, M. Direct association of the gap junction protein connexin-43 with ZO-1 in cardiac myocytes. *J. Biol. Chem.* **1998**, *273*, 12725–12731. [[CrossRef](#)]
53. Rhee, D.Y.; Zhao, X.Q.; Francis, R.J.; Huang, G.Y.; Mably, J.D.; Lo, C.W. Connexin 43 regulates epicardial cell polarity and migration in coronary vascular development. *Development* **2009**, *136*, 3185–3193. [[CrossRef](#)]
54. Zhu, L.; Vranckx, R.; Khau Van Kien, P.; Lalande, A.; Boisset, N.; Mathieu, F.; Wegman, M.; Glancy, L.; Gasc, J.M.; Brunotte, F.; et al. Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. *Nat. Genet.* **2006**, *38*, 343–349. [[CrossRef](#)]
55. Lu, C.C.; Liu, M.M.; Clinton, M.; Culshaw, G.; Argyle, D.J.; Corcoran, B.M. Developmental pathways and endothelial to mesenchymal transition in canine myxomatous mitral valve disease. *Vet. J.* **2015**, *206*, 377–384. [[CrossRef](#)]
56. Lorenz, K.; Schmitt, J.P.; Schmitteckert, E.M.; Lohse, M.J. A new type of ERK1/2 autophosphorylation causes cardiac hypertrophy. *Nat. Med.* **2009**, *15*, 75–83. [[CrossRef](#)] [[PubMed](#)]
57. Himeda, C.L.; Ranish, J.A.; Hauschka, S.D. Quantitative proteomic identification of MAZ as a transcriptional regulator of muscle-specific genes in skeletal and cardiac myocytes. *Mol. Cell. Biol.* **2008**, *28*, 6521–6535. [[CrossRef](#)]
58. Jaber, N.; Dou, Z.; Chen, J.S.; Catanzaro, J.; Jiang, Y.P.; Ballou, L.M.; Selinger, E.; Ouyang, X.; Lin, R.Z.; Zhang, J.; et al. Class III PI3K Vps34 plays an essential role in autophagy and in heart and liver function. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2003–2008. [[CrossRef](#)] [[PubMed](#)]
59. Gan, L.; Schwengberg, S.; Denecke, B. Transcriptome analysis in cardiomyocyte-specific differentiation of murine embryonic stem cells reveals transcriptional regulation network. *Gene Expr. Patterns* **2014**, *16*, 8–22. [[CrossRef](#)] [[PubMed](#)]
60. Wang, Y.; Singh, A.R.; Zhao, Y.; Du, T.; Huang, Y.; Wan, X.; Mukhopadhyay, D.; Wang, Y.; Wang, N.; Zhang, P. TRIM28 regulates sprouting angiogenesis through VEGFR-DLL4-Notch signaling circuit. *FASEB J.* **2020**, *34*, 14710–14724. [[CrossRef](#)] [[PubMed](#)]
61. Lu, F.I.; Wang, Y.T.; Wang, Y.S.; Wu, C.Y.; Li, C.C. Involvement of BIG1 and BIG2 in regulating VEGF expression and angiogenesis. *FASEB J.* **2019**, *33*, 9959–9973. [[CrossRef](#)] [[PubMed](#)]
62. Arrington, C.B.; Dowse, B.R.; Bleyl, S.B.; Bowles, N.E. Non-synonymous variants in pre-B cell leukemia homeobox (PBX) genes are associated with congenital heart defects. *Eur. J. Med. Genet.* **2012**, *55*, 235–237. [[CrossRef](#)] [[PubMed](#)]
63. Sailani, M.R.; Makrythanasis, P.; Valsesia, A.; Santoni, F.A.; Deutsch, S.; Popadin, K.; Borel, C.; Migliavacca, E.; Sharp, A.J.; Durliaux Sail, G.; et al. The complex SNP and CNV genetic architecture of the increased risk of congenital heart defects in Down syndrome. *Genome Res.* **2013**, *23*, 1410–1421. [[CrossRef](#)]
64. Levy, D.; Ronemus, M.; Yamrom, B.; Lee, Y.H.; Leotta, A.; Kendall, J.; Marks, S.; Lakshmi, B.; Pai, D.; Ye, K.; et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **2011**, *70*, 886–897. [[CrossRef](#)] [[PubMed](#)]
65. Warburton, D.; Ronemus, M.; Kline, J.; Jobanputra, V.; Williams, I.; Anyane-Yeboah, K.; Chung, W.; Yu, L.; Wong, N.; Awad, D.; et al. The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum. Genet.* **2014**, *133*, 11–27. [[CrossRef](#)] [[PubMed](#)]
66. Grote, P.; Wittler, L.; Hendrix, D.; Koch, F.; Währisch, S.; Beisaw, A.; Macura, K.; Bläss, G.; Kellis, M.; Werber, M.; et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* **2013**, *24*, 206–214. [[CrossRef](#)]
67. Zhang, F.; Lupski, J.R. Non-coding genetic variants in human disease. *Hum. Mol. Genet.* **2015**, *24*, R102–R110. [[CrossRef](#)]
68. D’Haene, E.; Vergult, S. Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. *Genet. Med.* **2021**, *23*, 34–46. [[CrossRef](#)]
69. Melo, U.S.; Schöpflin, R.; Acuna-Hidalgo, R.; Mensah, M.A.; Fischer-Zirnsak, B.; Holtgrewe, M.; Klever, M.; Türkmen, S.; Heinrich, V.; Pluym, I.D.; et al. Hi-C Identifies Complex Genomic Rearrangements and TAD-Shuffling in Developmental Diseases. *Am. J. Hum. Genet.* **2020**, *106*, 872–884. [[CrossRef](#)]