Received: 30 September 2022 | Revised: 16 December 2022 | Accepted: 25 January 2023

DOI: 10.1002/ajmg.a.63145

ORIGINAL ARTICLE

AMERICAN JOURNAL OF medical genetics PART A WILEY

Phenotypic heterogeneity in 22q11.2 deletion syndrome: Copy Number Variants as genetic modifiers for congenital heart disease in a Brazilian cohort

M. Zamariolli[1] | A. G. Dantas[1] | N. Nunes[1] | M. Moysés-Oliveira[2] | I. C. Sgardioli[3] | D. C. Q. Soares[4] | V. L. Gil-Da-Silva-Lopes[3] | C. A. Kim[4] | M. I. Melaragno[1]

[1]Genetic Division, Universidade Federal de São Paulo, São Paulo, Brazil

[2]Sleep Institute, São Paulo, Brazil

[3]Department of Translational Medicine, Area of Medical Genetics and Genomic Medicine, Universidade Estadual de Campinas, Campinas, Brazil

[4]Genetics Unit, Instituto da Criança, Universidade de São Paulo, São Paulo, Brazil

Correspondence
M. I. Melaragno, Genetics Division, Department of Morphology and Genetics, Universidade Federal de São Paulo, Rua Botucatu, 740, CEP 04023-900, São Paulo, SP, Brazil.
Email: melaragno.maria@unifesp.br

Funding information
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Numbers: 2020/11241-2, 2019/21644-0

Abstract

The clinical heterogeneity in 22q11.2 deletion syndrome (22q11.2DS) underlies complex genetic mechanisms including variants in other regions of the genome, known as genetic modifiers. Congenital heart disease (CHD) is one of the most relevant phenotypes in the syndrome and copy number variants (CNVs) outside the 22q11.2 region could play a role in its variable expressivity. Since those described loci account for a small proportion of the variability, the CNV analysis in new cohorts from different ancestry-based populations constitutes a valuable resource to identify a wider range of modifiers. We performed SNP-array in 117 Brazilian patients with 22q11.2DS, with and without CHD, and leveraged genome-wide CNV analysis. After quality control, we selected 50 CNVs in 38 patients for downstream analysis. CNVs' genetic content and implicated biological pathways were compared between patients with and without CHD. CNV-affected genes in patients with CHD were enriched for several functional terms related to ubiquitination, transcription factor binding sites and miRNA targets, highlighting the complexity of the phenotype's expressivity. Cardiac-related genes were identified in both groups of patients suggesting that increasing risk and protective mechanisms could be involved. These genes and enriched pathways could indicate new modifiers to the cardiac phenotype in 22q11.2DS patients.

KEYWORDS
22q11.2 deletion syndrome, congenital heart disease (CHD), copy-number variant (CNV), genetic modifiers

1 | INTRODUCTION

The 22q11.2 deletion syndrome (22q11.2DS) (OMIM #188400; #192430) is the most frequent microdeletion in humans with a prevalence of 1:2148 (Blagojevic et al., 2021). Most individuals with 22q11.2DS carry the typical ~3 Mb deletion on chromosome 22 (79%). However, smaller nested deletions are described in approximately 21% of patients (Blagojevic et al., 2021) and atypical deletions of varying sizes can also be present in the minority of patients (Michaelovsky et al., 2012). These deletions are a result of non-allelic homologous recombination events between the different low copy repeats (LCRs) in the 22q11.2 region (Shaikh, 2000).

The syndrome is characterized by a great range of clinical features with variable expressivity including congenital heart disease (CHD), dysmorphic facial features, palatal anomalies, immune deficiencies, different neuropsychiatric disorders, and cognitive impairment (McDonald-McGinn et al., 2015).

In the 22q11.2DS, cardiac defects have been reported in 64% of the patients (Campbell et al., 2018). CHD are primarily comprised of conotruncal cardiac defects including tetralogy of Fallot, interrupted aortic arch type B, and ventricular septal defect (Peyvandi et al., 2013). The variable expressivity of this phenotype is not fully understood, and it does not seem to be related to race, sex, parental origin of the deletion (Goldmuntz et al., 2009; Swaby et al., 2011) or even with the 22q11.2 deletion size (Rozas et al., 2019). An influence of genetic factors has been suggested since patients with 22q11.2DS and CHD more frequently present relatives with isolated cardiac defects when compared to 22q11.2DS patients with normal cardiac development (Swaby et al., 2011).

CNVs outside of the 22q11.2 region have been described to act as genetic modifiers in the syndrome (León et al., 2017; Mlynarski et al., 2015, 2016). It has been suggested that common CNVs may not be individually pathogenic, but in the presence of the 22q11.2 deletion, as the primary alteration, they may influence expressivity of cardiac phenotype (León et al., 2017; Mlynarski et al., 2015, 2016). This combined action of the 22q11.2 deletion and other CNVs may be due to the participation of affected genes in common regulatory pathways as was seen for the modifier *KANSL1* gene and the 22q11.2 *CRKL* gene (León et al., 2017). In addition, rare CNVs identified in patients with 22q11.2DS and CHD are enriched for genes related to cardiac development (Mlynarski et al., 2016).

Although some genes and regions have already been described as modifiers of the cardiac phenotype in the 22q11.2 deletion, they explain only a small proportion of its incomplete penetrance, suggesting that other modifiers for this condition remain undiscovered. Therefore, here we investigated a cohort of individuals with 22q11.2DS to identify new potential genetic modifiers and pathways possibly involved in the expressivity of the cardiac phenotype.

2 | MATERIALS AND METHODS

2.1 | Editorial policies and ethical considerations

Written informed consent was obtained from patients or their parents as approved by the local ethics committee.

2.2 | Cohorts

Blood samples were obtained from 117 Brazilian patients diagnosed with 22q11.2DS. The 22q11.2 microdeletion was previously identified by multiplex ligation-dependent probe amplification (MLPA). Cardiac anatomy was evaluated by echocardiogram and patients were divided in two groups: with and without CHD.

CEL files for 339 Brazilian control individuals from BIPMED consortium (Brazilian Initiative on Precision Medicine, https://bipmed.org/) (Rocha et al., 2020) previously genotyped with Affymetrix SNP 6.0 arrays were also used. These individuals had no history of birth defects, oral cleft, rheumatological diseases, or cancer in three generations.

2.3 | SNP-array and quality control in the 22q11.2DS cohort

DNA samples from patients with 22q11.2DS were genotyped with Axiom™ Precision Medicine Diversity Array (Axiom PMDA, Affymetrix) according to manufacturer's protocol at Affymetrix's core facility (Santa Clara, CA, USA). Quality control values were calculated with Affymetrix Power Tools v. 2.11.3 (Affymetrix) pipeline. Samples with Dish Quality Control (QC) less than 0.82 or QC call rate below 97% were excluded according to manufacturer's guidelines. Samples from patients with insufficient cardiac phenotype information were excluded. Identity by descent (IBD) analysis was performed to exclude duplicate or related samples with PLINK 1.9 toolset (www.cog-genomics.org/plink/1.9/) (Chang et al., 2015). Only patients with deletions within LCR-A and LCR-D were included, including patients with A-D, A-B and C-D deletions.

Quality control values for individuals from the BIPMED consortium were also calculated with Affymetrix Power Tools v. 2.11.3 (Affymetrix) pipeline. Samples with contrast QC values below 0.4 were excluded according to manufacturer's instructions.

2.4 | CNV detection

Extraction of signal intensity data from the raw CEL files was performed with Affymetrix Power Tools v. 2.11.3 (Affymetrix) pipeline (Appendix S1).

Autosomal CNV detection was performed with an adapted pipeline of PennCNV software for both cohorts. The pipeline takes log *R* ratio (LRR) and B-allele frequency (BAF) values for each probe and sample, calls CNVs with PennCNV software (Wang et al., 2007) and computes a quality score (QS) based on several quality metrics and CNV features (such as length and number of probes). This QS estimates the probability of a CNV identified by PennCNV to be a true positive call (Macé et al., 2016). Adjacent CNVs with gaps shorter than 20% of the total length were merged according to default PennCNV parameters. Details of parameters and reference files used for CNV detection are found in Appendix S1.

For subsequent analysis of CNVs outside of the 22q11.2 region, we selected CNVs with a QS ≥ |0.5| (Auwerx et al., 2022). After this selection, CNVs with an overlap of at least 50% with centromere and telomere regions were excluded.

2.5 | Copy number variable regions definition and content

As CNVs called on SNP-array platforms may present variations in breakpoints, we clustered CNVs with 50% reciprocal overlap into copy number variable regions (CNVRs) with R package "CNVgears" (CNVgears, Simone Montalbano, https://github.com/SinomeM/CNVgears, 2020). This was done separately for duplications and deletions. CNVR boundaries were defined by the maximum spam of CNVs in the cluster (Zarrei et al., 2015).

CNVRs in subjects with 22q11.2DS were annotated with the NCBI RefSeq gene set downloaded from the UCSC Table Browser of the hg38 build (http://genome.ucsc.edu/cgi-bin/hgTables?command=start). An overlap of at least 80% of the gene was needed for annotation. Olfactory receptor genes were excluded from further analysis.

2.6 | Brazilian reference cohort

To assess CNV frequency on a Brazilian reference cohort, CNVRs were generated separately for deletions and duplications in the BIPMED consortium samples. The frequency for each CNVR was calculated as: (number of individuals with the CNVR/total number of assessed individuals) × 100 (see details in Appendix S1).

2.7 | CNV frequency

We defined rare CNVs as those with a frequency less than 1% in three population cohorts: BIPMED, Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home) (MacDonald et al., 2014) and The Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/) (Collins et al., 2020) (Appendix S1). CNVs were considered common when their frequency was 1% or above in at least one of those cohorts (Appendix S1).

2.8 | Functional terms enrichment analysis

Different gene sets, for each group (22q11.2DS with CHD × 22q11.2DS without CHD) were individually submitted for enrichment analysis using the online tool "Gene Set Enrichment Analysis" (http://www.gsea-msigdb.org/gsea/msigdb/index.jsp) (Appendix S1). False discovery rate (FDR) procedure was applied as the correction method and an adjusted-p value of 0.05 was adopted for significance. Enrichment analysis was separately performed for seven gene collections from the Molecular Signatures Database (MsigDB v7.2): Gene Ontology (GO) biological process (n = 7479 genes); GO cellular component (n = 996 genes); GO molecular function (n = 1708 genes); Human Phenotype Ontology (HPO genes) (n = 4813 genes); Cell type signature gene sets (n = 673 genes); microRNA targets (n = 2598 genes) and all transcription factor targets (n = 1133 genes).

2.9 | Cardiac gene sets enrichment analysis

Three sets of cardiac-related genes were compiled from previously published work. *List 1* represents genes expressed in developing mouse heart and pharyngeal arches (n = 12,165 genes). Genes were ranked based on their expression levels and the top 25% of genes with the highest expression in mouse pharyngeal arch at day E9.5 (Mlynarski et al., 2016) and mouse developing heart at day E9.5 (Mlynarski et al., 2016) and at day E14.5 (Homsy et al., 2015; Zaidi et al., 2013) were selected. *List 2* is the list of genes previously associated to congenital heart defects (n = 1387). Genes associated with the term "abnormal heart morphology" were obtained from the HPO database (https://hpo.jax.org/app/). Genes affected by de novo SNVs, damaging missense and LoF variants, in a cohort of individuals with CHD were selected (Homsy et al., 2015). *List 3* is a list of genes expressed in human fetal heart (n = 411) obtained from published datasets (Iruretagoyena et al., 2014; Wang et al., 2017).

We compared the frequency in which genes affected by CNVs in the groups 22q11.2DS CHD and 22q11.2DS without CHD appeared in each list via Fisher's exact test ($p < 0.05$).

2.10 | Protein–protein interaction

A protein–protein interaction (PPI) analysis was conducted to address protein interactions between the coding genes encompassed by the 22q11.2 typical ~3 Mb deletion (N = 50 genes) and the coding genes contained in the cardiac gene sets that were identified as affected by the detected CNVs outside of the 22q11.2 region (N = 22 genes) (Section 2.9). NCBI RefSeq genes contained in the 22q11.2 region (chr22:18,147,233_21,555,711, GRCh38/hg38) were downloaded from the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables?command=start). Gene names were converted into their respective Ensembl protein ID with the online tool BioMart (Ensembl release 107—July 2022). PPI analysis was performed in String DB (https://string-db.org/, v11.5) with experiments and databases as interactive sources and minimal interaction score of 0.7.

2.11 | Statistics

Fisher exact test or the Wilcoxon rank sum test were used for CNV comparison between groups with and without CHD as appropriate. Statistical analyses were done in R version 4.0.2.

3 | RESULTS

After excluding individuals that did not pass quality control or meet inclusion criteria, 95 unrelated patients with 22q11.2DS remained, 56 with CHD and 39 with normal cardiac anatomy. Evaluation of 22q11.2 deletion sizes revealed 89 patients with a deletion between LCR-A and LCR-D (55 with CHD and 34 without CHD), 5 patients with deletions between LCR-A and LCR-B (1 with CHD and 4 without CHD) and one patient without CHD with a deletion between LCR-C and LCR-D. There was no association between 22q11.2 deletion category and CHD phenotype (p = 0.07, two-tailed Fisher's exact test). Although, the number of individuals with nested deletions in our cohort may be insufficient to verify a statistical difference, the lack of association is consistent with previous studies (León et al., 2017; Mlynarski et al., 2015, 2016; Rozas et al., 2019).

TABLE 1 CNVRs and cohort and populational frequency.

Patient ID	Chromosome	CNVR start	CNVR end	CNV type	Patients group	Cohort frequency	BIPMED (%)	gnomAD (%)	DGV (%)	CNV frequency
1	1	78121656	78972463	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
2	1	143611765	143908256	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0	3.65	Common
3	1	247666923	248126273	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0.18	0.24	Rare
36	1	109687833	109698628	Deletion	22q11.2DS CHD	1.05% (1/95)	0	0	15.4	Common
20	2	87165034	87743968	Duplication	22q11.2DS CHD	1.05% (1/95)	1.44	0	1.41	Common
15	2	110631098	112316999	Deletion	22q11.2DS without CHD	1.05% (1/95)	0	0.01	0	Rare
31	4	69877120	70369163	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0	0	Rare
38	4	188413838	190017927	Deletion	22q11.2DS without CHD	1.05% (1/95)	0	0	0.02	Rare
31	5	13789193	14182211	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0.08	0.08	Rare
32	6	167934049	168196919	Duplication	22q11.2DS without CHD	1.05% (1/95)	0.96	0.64	2.61	Common
33	7	57189818	57781917	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0.04	0	Rare
14	7	76555666	76980627	Duplication	22q11.2DS without CHD	1.05% (1/95)	0.96	0	1.67	Common
34	8	2483981	2725305	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0.1	0.49	Rare
22	8	17963420	18228429	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
35	9	135255502	135417825	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0.15	0.56	Rare
4	11	35717887	36289127	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
37	11	89197899	92112768	Deletion	22q11.2DS CHD	1.05% (1/95)	0	0	0	Rare
5	13	32340234	32341197	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
6	14	18544151	19954640	Duplication	22q11.2DS without CHD	2.11% (2/95)	4.31	0	4.2	Common
7	14	18544151	19954640	Duplication	22q11.2DS CHD	2.11% (2/95)	4.31	0	4.2	Common
8	14	18689021	19951082	Duplication	22q11.2DS CHD	2.11% (2/95)	2.87	0	4.2	Common
9	14	18689021	19951082	Duplication	22q11.2DS without CHD	2.11% (2/95)	2.87	0	4.2	Common
10	15	31724822	32327926	Duplication	22q11.2DS CHD	1.05% (1/95)	0.48	0	0.69	Rare
11	16	6713902	6797227	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
12	16	87727265	87947734	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
13	17	43092614	43104264	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	0	0	Rare
14	17	43092931	43093694	Duplication	22q11.2DS without CHD	3.16% (3/95)	0	0	0	Rare
5	17	43092931	43093694	Duplication	22q11.2DS without CHD	3.16% (3/95)	0	0	0	Rare
15	17	43092931	43093694	Duplication	22q11.2DS without CHD	3.16% (3/95)	0	0	0	Rare
16	19	855587	1293917	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0	0	Rare
17	19	11098953	11120428	Duplication	22q11.2DS CHD	2.11% (2/95)	0	0	0	Rare
18	19	11098953	11120428	Duplication	22q11.2DS CHD	2.11% (2/95)	0	0	0	Rare
14	19	40844677	40860295	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	1.14	0.73	Common
19	19	42792104	43003003	Duplication	22q11.2DS without CHD	1.05% (1/95)	0	2.06	0.88	Common
21	21	45386899	45730952	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0	0	Rare

TABLE 1 (Continued)

Patient ID	Chromosome	CNVR start	CNVR end	CNV type	Patients group	Cohort frequency	BIPMED (%)	gnomAD (%)	DGV (%)	CNV frequency
3	22	21959377	22219884	Duplication	22q11.2DS without CHD	1.05% (1/95)	0.96	0	0.45	Rare
22	22	25279622	25521694	Duplication	22q11.2DS without CHD	1.05% (1/95)	2.87	0	3.99	Common
23	22	25280972	25512027	Duplication	22q11.2DS without CHD	2.11% (2/95)	2.87	0	3.99	Common
19	22	25280972	25512027	Duplication	22q11.2DS without CHD	2.11% (2/95)	2.87	0	3.99	Common
24	22	42115069	42134934	Duplication	22q11.2DS CHD	3.16% (3/95)	0	0	22.32	Common
19	22	42115069	42134934	Duplication	22q11.2DS without CHD	3.16% (3/95)	0	0	22.32	Common
1	22	42115069	42134934	Duplication	22q11.2DS without CHD	3.16% (3/95)	0	0	22.32	Common
21	22	42126599	42134934	Duplication	22q11.2DS CHD	5.26% (5/95)	0	0	22.32	Common
11	22	42126599	42134934	Duplication	22q11.2DS without CHD	5.26% (5/95)	0	0	22.32	Common
25	22	42126599	42134934	Duplication	22q11.2DS without CHD	5.26% (5/95)	0	0	22.32	Common
26	22	42126599	42134934	Duplication	22q11.2DS without CHD	5.26% (5/95)	0	0	22.32	Common
27	22	42126599	42134934	Duplication	22q11.2DS without CHD	5.26% (5/95)	0	0	22.32	Common
28	22	42130482	42143437	Duplication	22q11.2DS CHD	1.05% (1/95)	0	0	22.32	Common
29	22	42130967	42134934	Duplication	22q11.2DS without CHD	2.11% (2/95)	0	0	0	Rare
30	22	42130967	42134934	Duplication	22q11.2DS CHD	2.11% (2/95)	0	0	0	Rare

Abbreviations: CHD, congenital heart disease; CNV, copy number variant; CNVR, copy number variable region.

3.1 | CNVs outside of the 22q11.2

A total of 50 CNVs (QS ≥ |0.5|) outside of the 22q11.2 deletion region were identified in 38 out of the 95 patients, 46 duplications (22 common and 24 rare) and 4 deletions (1 common and 3 rare) (Table 1). When comparing patients with at least one CNV, there was no significant difference in the number or length of detected CNVs between individuals with CHD and those with normal cardiac phenotype (Table 2). Noteworthy, there was significant association in the number of patients carrying CNVs outside of the 22q11.2 deletion region with cardiac phenotype ($p = 0.033$, two-tailed Fisher's exact test, Table 2).

After CNVR definition and genetic content annotation, we identified 129 distinct genes (106 duplicated and 23 deleted) in the 22q11.2DS CHD group and 102 genes (68 duplicated and 34 deleted) in the 22q11.2DS without CHD group (Table S1). A total of 17 genes were identified in both groups.

3.2 | Functional terms enrichment

Pathway analysis revealed a great number of enriched terms for the gene sets generated from CNVs identified in the 22q11.2DS CHD group (Table S2). In this group, enriched terms such as ubiquitin like protein ligase activity, transition metal ion binding, predicted targets from miR-2355-3p, and several transcription factors binding sites were observed. On the other hand, a considerably smaller number of terms was verified for the gene sets generated from CNVs identified in the 22q11.2DS without CHD group. These terms were only from the cell type signature and transcription factor target collections (Table S3). There were no overlapping terms between groups. It is worth noticing that in this analysis, it is possible for one functional term to be enriched by the genes contained within a single CNV or by genes contained in different CNVs affecting distinct individuals.

3.3 | Cardiac gene sets

To address if CNVs in individuals with 22q11.2DS and CHD affected genes involved in cardiac development and malformation, we examined three lists of cardiac-related genes. No significant differences were observed between groups when comparing all the affected genes or only the genes exclusively affected in each of the groups. There was also no statistically significant difference when comparing the number of patients with CNVs affecting at least one of the genes. Even though we found no statistically significant differences, 25 cardiac-related genes were identified encompassed by CNVs in individuals with 22q11.2DS (Table 3).

We evaluated whether the proteins encoded by any of these genes interacted with protein products of the 22q11.2 region to examine if this could be a potential mechanism involved in the penetrance of cardiac defects in the 22q11.2DS. At a high confidence score threshold of 0.7, MED16 protein was found to interact with MED15, the latter being coded by a gene localized in the 22q11.2

TABLE 2 Number and length of autosomal CNVs outside of the 22q11 region.

	CHD	Without CHD	*p* value
Number of patients with CNVs	17	21	0.033[a]
Median of CNVs per subject (IQR)	1 (0.75)	1 (0.75)	0.064[b]
Median CNV length (bp) per subject (IQR) (bp)	349,857 (585,057)	262,870 (708,207)	0.75[b]

Abbreviations: CHD, congenital heart disease; CNV, copy number variant; IQR: interquartile range.
[a]Two-tailed Fisher's exact test.
[b]Wilcoxon rank-sum test. Only considered individuals with at least one CNV.

TABLE 3 CNVRs and cardiac-related genes.

CNVR	CNV type	Patients group	CNV frequency	Number individuals	Genes[a]
chr1:109687833_109698628	Deletion	22q11.2DS CHD	Common	1	*GSTM1*
chr11:89197899_92112768	Deletion	22q11.2DS CHD	Rare	1	*CHORDC1*
chr19:855587_1293917	Duplication	22q11.2DS CHD	Rare	1	*ARID3A*, ***ATP5F1D***, *CIRBP*, *CNN2*, ***GPX4***, *MED16*, *MIDN*, *POLR2E*, *R3HDM4*, *STK11*, *TMEM259*, *WDR18*
chr21:45386899_45730952	Duplication	22q11.2DS CHD	Rare	1	***COL18A1***, *SLC19A1*
chr5:13789193_14182211	Duplication	22q11.2DS CHD	Rare	1	***DNAH5***
chr11:35717887_36289127	Duplication	22q11.2DS without CHD	Rare	1	*LDLRAD3*
chr16:87727265_87947734	Duplication	22q11.2DS without CHD	Rare	1	*SLC7A5*
chr2:110631098_112316999	Deletion	22q11.2DS without CHD	Rare	1	*ANAPC1*, *BCL2L11*, ***BUB1***
chr22:21959377_22219884	Duplication	22q11.2DS without CHD	Rare	1	*TOP3B*
chr4:188413838_190017927	Deletion	22q11.2DS without CHD	Rare	1	*FRG1*
chr8:17963420_18228429	Duplication	22q11.2DS without CHD	Rare	1	*ASAH1*

Abbreviations: CHD, congenital heart disease; CNV, copy number variant; CNVR, copy number variant region.
[a]Cardiac-genes identified in each CNVR, non-cardiac-related genes are not displayed. Genes highlighted in bold were identified in the list of genes previously associated to congenital heart defects (List 2). The remaining genes were identified in the list of genes expressed in developing mouse heart and pharyngeal arches (List 1).

region, with evidence of known interactions from curated databases and experiments.

4 | DISCUSSION

The aim of our study was to identify new CNVs, outside of the 22q11.2 region, that can act as potential genetic modifiers for the cardiac phenotype in the 22q11.2DS. The study of modifying CNVs for CHD in the 22q11.2DS by Mlynarski et al. (2015) and León et al. (2017) revealed common CNVs involving the *SLC2A3* and *KANSL1* genes, respectively, which were not identified as affected by CNVs in our study. The lack of replication concerning genetic modifiers in the 22q11.2DS highlights the complexity and heterogeneity of the genetic factors involved.

Since there is great variability in the identification of CNVs depending on the SNP-array platform and the detection tools applied, false positive results may be obtained (Haraksingh et al., 2017). Thus, with the aim of obtaining highly reliable results, we applied a pipeline that calculates the probability of a CNV being identified by two other independent tools (Macé et al., 2016). The choice of this stringent CNV filtering is likely the cause of the low number of CNVs identified (50 CNVs in total).

Even though we found a significantly larger number of individuals from the group 22q11.2DS without CHD carrying at least one CNV, a considerably larger number of enriched terms were found for the group of patients with CHD, suggesting greater biological connectivity of genes affected by CNVs in this group. Some of the identified terms in the 22q11.2DS CHD group are related to ubiquitination, a process to maintain protein homeostasis on cells (Zhang et al., 2020) that has been related to cardiac pathogenesis (Li et al., 2018) and to tetralogy of Fallot (You et al., 2020). These enrichments were driven by CNV-affected genes from the TRIM family, which has been associated to different cardiac functions and conditions, such as heart failure and arrhythmia (Zhang et al., 2020).

In addition, we have identified one enriched term for miRNA targets and a larger number of functional terms referring to targets of different transcription factors. Although the specific miRNA, miR-2355-3p, has not yet been related to heart development or cardiac pathogenesis, the role of miRNAs in regulation of gene expression during cardiovascular development and disease has been described (Liu & Olson, 2010). Heart development is also controlled by a

conserved network of transcription factors (Olson, 2006) and, therefore, abnormalities involving these transcription factors could lead to improper development and consequently CHD. One of the transcription factors enriched in our work, *ZSCAN5B*, is expressed in developing mouse heart(Sun et al., 2016). Altogether, these findings may indicate the complexity of regulatory mechanisms involved in the penetrance of the cardiac phenotype in the 22q11.2.

It is also worth noting that, in the 22q11.2DS with CHD group, enriched terms were observed only for genes affected by rare CNVs and not for genes affected by common CNVs. This suggests that a disturbance in these biological pathways can be damaging, since they do not appear enriched in CNVs that are more frequent in the general population, reinforcing the biological relevance of these findings.

Among the 25 cardiac-related genes affected by CNVs, 20 of them had been identified in studies of gene expression related to cardiac development in mice (Homsy et al., 2015; Mlynarski et al., 2016). Thus, CNVs involving these genes may have affected their expression, contributing to the appearance of the cardiac phenotype in conjunction with the deletion of the 22q11.2 region or acting in a protective way for individuals without CHD.

To further investigate these genes in the context of the 22q11.2DS, we performed a PPI analysis that revealed interaction between the proteins encoded by *MED16*, a gene expressed in developing mouse heart and pharyngeal arches (List 1, see Section 2.9), and *MED15*, deleted in the 22q11.2DS. Both are components of Mediator, a multiprotein complex conserved among eukaryotes that plays a role in transcriptional regulation (Ansari & Morse, 2013). In *Saccharomyces cerevisiae*, both components are required for the recruitment of Mediator by Hsf1 (heat shock factor 1) in a non-redundant manner (Kim & Gross, 2013). A synthetic lethality interaction between them has also been shown in yeast (Balciunas et al., 2003). This observation reinforces the idea of an interaction between genes deleted in the 22q11.2 region and genes affected by genetic variants elsewhere in the genome driving the variable penetrance and expressivity of a phenotype.

In conclusion, this work reinforces that exploring different cohorts, with various backgrounds, and genetic variant types allows for the identification of new genes or regions relevant to the cardiac phenotype. Further investigations involving these regions in the context of the 22q11.2DS, will enable a greater understanding of the candidates for genetic modifiers and their relationship with the 22q11.2 deletion region.

AUTHOR CONTRIBUTIONS

M. Zamariolli contributed to study design, analysis, and interpretation of the data and wrote the article. M. Moysés-Oliveira, A.G. Dantas and N. Nunes contributed to study design and critically revised the manuscript. D.C.Q. Soares, V.L. Gil-Da-Silva-Lopes and C.A. Kim contributed to data collection and performed clinical evaluation of patients in the 22q11.2DS cohort. I.C. Sgardioli contributed to data collection. M.I. Melaragno contributed to study design, writing and critically revised the manuscript. All authors revised the manuscript and approved the final version.

ACKNOWLEDGMENTS

We thank the patients and their families for the participation in the study. We also thank Dr. Paulo Paiva for maintenance of server and support for computational analysis. Maria Isabel Melaragno received financial support from Fundação de Amparo à Pesquisa do Estado de São Paulo [2018/20618-2, M. Zamariolli; 2019/21644-0, M.I. Melaragno] and from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

FUNDING INFORMATION

This work was supported from grant 2018/20618-2 and 2019/21644-0, São Paulo Research Foundation (FAPESP) and from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

M. I. Melaragno https://orcid.org/0000-0002-4344-9698

REFERENCES

Ansari, S. A., & Morse, R. H. (2013). Mechanisms of mediator complex action in transcriptional activation. *Cellular and Molecular Life Sciences*, *70*(15), 2743–2756. https://doi.org/10.1007/s00018-013-1265-9

Auwerx, C., Lepamets, M., Sadler, M. C., Patxot, M., Stojanov, M., Baud, D., Mägi, R., Porcu, E., Reymond, A., Kutalik, Z., Esko, T., Metspalu, A., Milani, L., Mägi, R., & Nelis, M. (2022). The individual and global impact of copy-number variants on complex human traits. *The American Journal of Human Genetics*, *109*(4), 647–668. https://doi.org/10.1016/j.ajhg.2022.02.010

Balciunas, D., Hallberg, M., Björklund, S., & Ronne, H. (2003). Functional interactions within yeast mediator and evidence of differential subunit modifications. *Journal of Biological Chemistry*, *278*(6), 3831–3839. https://doi.org/10.1074/jbc.M206946200

Blagojevic, C., Heung, T., Theriault, M., Tomita-Mitchell, A., Chakraborty, P., Kernohan, K., Bulman, D. E., & Bassett, A. S. (2021). Estimate of the contemporary live-birth prevalence of recurrent 22q11.2 deletions: A cross-sectional analysis from population-based newborn screening. *CMAJ Open*, *9*(3), E802–E809. https://doi.org/10.9778/cmajo.20200294

Campbell, I. M., Sheppard, S. E., Crowley, T. B., McGinn, D. E., Bailey, A., McGinn, M. J., Unolt, M., Homans, J. F., Chen, E. Y., Salmons, H. I., Gaynor, J. W., Goldmuntz, E., Jackson, O. A., Katz, L. E., Mascarenhas, M. R., Deeney, V. F. X., Castelein, R. M., Zur, K. B., Elden, L., … McDonald-McGinn, D. M. (2018). What is new with 22q? An update from the 22q and You Center at the Children's Hospital of Philadelphia. *American Journal of Medical Genetics Part A*, *176*(10), 2058–2069. https://doi.org/10.1002/ajmg.a.40637

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., Khera, A. V., Lowther, C., Gauthier, L. D., Wang, H., Watts, N. A., Solomonson, M., O'Donnell-Luria, A., Baumann, A., Munshi, R., Walker, M., Whelan, C. W., Huang, Y., Brookings, T., … Talkowski, M. E. (2020). A structural variation reference for medical and population genetics. *Nature*, *581*(7809), 444–451. https://doi.org/10.1038/s41586-020-2287-8

Goldmuntz, E., Driscoll, D. A., Emanuel, B. S., McDonald-McGinn, D., Mei, M., Zackai, E., & Mitchell, L. E. (2009). Evaluation of potential modifiers of the cardiac phenotype in the 22q11.2 deletion syndrome. *Birth Defects Research Part A: Clinical and Molecular Teratology*, *85*(2), 125–129. https://doi.org/10.1002/bdra.20501

Haraksingh, R. R., Abyzov, A., & Urban, A. E. (2017). Comprehensive performance comparison of high-resolution array platforms for genome-wide copy number variation (CNV) analysis in humans. *BMC Genomics*, *18*(1), 321. https://doi.org/10.1186/s12864-017-3658-x

Homsy, J., Zaidi, S., Shen, Y., Ware, J. S., Samocha, K. E., Karczewski, K. J., DePalma, S. R., McKean, D., Wakimoto, H., Gorham, J., Jin, S. C., Deanfield, J., Giardini, A., Porter, G. A., Kim, R., Bilguvar, K., Lopez-Giraldez, F., Tikhonova, I., Mane, S., … Chung, W. K. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science*, *350*(6265), 1262–1266. https://doi.org/10.1126/science.aac9396

Iruretagoyena, J. I., Davis, W., Bird, C., Olsen, J., Radue, R., Teo Broman, A., Kendziorski, C., Splinter BonDurant, S., Golos, T., Bird, I., & Shah, D. (2014). Metabolic gene profile in early human fetal heart development. *MHR: Basic Science of Reproductive Medicine*, *20*(7), 690–700. https://doi.org/10.1093/molehr/gau026

Kim, S., & Gross, D. S. (2013). Mediator recruitment to heat shock genes requires dual Hsf1 activation domains and mediator tail subunits Med15 and Med16. *Journal of Biological Chemistry*, *288*(17), 12197–12213. https://doi.org/10.1074/jbc.M112.449553

León, L. E., Benavides, F., Espinoza, K., Vial, C., Alvarez, P., Palomares, M., Lay-Son, G., Miranda, M., & Repetto, G. M. (2017). Partial microduplication in the histone acetyltransferase complex member KANSL1 is associated with congenital heart defects in 22q11.2 microdeletion syndrome patients. *Scientific Reports*, *7*(1), 1795. https://doi.org/10.1038/s41598-017-01896-w

Li, J., Johnson, J. A., & Su, H. (2018). Ubiquitin and ubiquitin-like proteins in cardiac disease and protection. *Current Drug Targets*, *19*(9), 989–1002. https://doi.org/10.2174/1389450117666151209114608

Liu, N., & Olson, E. N. (2010). MicroRNA regulatory networks in cardiovascular development. *Developmental Cell*, *18*(4), 510–525. https://doi.org/10.1016/j.devcel.2010.03.010

MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014). The database of genomic variants: A curated collection of structural variation in the human genome. *Nucleic Acids Research*, *42*-(Database issue), D986–D992. https://doi.org/10.1093/nar/gkt958

Macé, A., Tuke, M. A., Beckmann, J. S., Lin, L., Jacquemont, S., Weedon, M. N., Reymond, A., & Kutalik, Z. (2016). New quality measure for SNP array based CNV detection. *Bioinformatics*, *32*(21), 3298–3305. https://doi.org/10.1093/bioinformatics/btw477

McDonald-McGinn, D. M., Sullivan, K. E., Marino, B., Philip, N., Swillen, A., Vorstman, J. A. S., Zackai, E. H., Emanuel, B. S., Vermeesch, J. R., Morrow, B. E., Scambler, P. J., & Bassett, A. S. (2015). 22q11.2 deletion syndrome. *Nature Reviews Disease Primers*, *1*, 15071. https://doi.org/10.1038/nrdp.2015.71

Michaelovsky, E., Frisch, A., Carmel, M., Patya, M., Zarchi, O., Green, T., Basel-Vanagaite, L., Weizman, A., & Gothelf, D. (2012). Genotype-phenotype correlation in 22q11.2 deletion syndrome. *BMC Medical Genetics*, *13*(1), 122. https://doi.org/10.1186/1471-2350-13-122

Mlynarski, E. E., Sheridan, M. B., Xie, M., Guo, T., Racedo, S. E., McDonald-McGinn, D. M., Gai, X., Chow, E. W. C., Vorstman, J., Swillen, A., Devriendt, K., Breckpot, J., Digilio, M. C., Marino, B., Dallapiccola, B., Philip, N., Simon, T. J., Roberts, A. E., Piotrowicz, M., … Emanuel, B. S. (2015). Copy-number variation of the glucose transporter gene SLC2A3 and congenital heart defects in the 22q11.2 deletion syndrome. *The American Journal of Human Genetics*, *96*(5), 753–764. https://doi.org/10.1016/j.ajhg.2015.03.007

Mlynarski, E. E., Xie, M., Taylor, D., Sheridan, M. B., Guo, T., Racedo, S. E., McDonald-McGinn, D. M., Chow, E. W. C., Vorstman, J., Swillen, A., Devriendt, K., Breckpot, J., Digilio, M. C., Marino, B., Dallapiccola, B., Philip, N., Simon, T. J., Roberts, A. E., Piotrowicz, M. M., … Emanuel, B. S. (2016). Rare copy number variants and congenital heart defects in the 22q11.2 deletion syndrome. *Human Genetics*, *135*(3), 273–285. https://doi.org/10.1007/s00439-015-1623-9

Olson, E. N. (2006). Gene regulatory networks in the evolution and development of the heart. *Science (New York, NY)*, *313*(5795), 1922–1927. https://doi.org/10.1126/science.1132292

Peyvandi, S., Lupo, P. J., Garbarini, J., Woyciechowski, S., Edman, S., Emanuel, B. S., Mitchell, L. E., & Goldmuntz, E. (2013). 22q11.2 deletions in patients with conotruncal defects: Data from 1,610 consecutive cases. *Pediatric Cardiology*, *34*(7), 1687–1694. https://doi.org/10.1007/s00246-013-0694-4

Rocha, C. S., Secolin, R., Rodrigues, M. R., Carvalho, B. S., & Lopes-Cendes, I. (2020). The Brazilian initiative on precision medicine (BIPMed): Fostering genomic data-sharing of underrepresented populations. *NPJ Genomic Medicine*, *5*(1), 42. https://doi.org/10.1038/s41525-020-00149-6

Rozas, M. F., Benavides, F., León, L., & Repetto, G. M. (2019). Association between phenotype and deletion size in 22q11.2 microdeletion syndrome: Systematic review and meta-analysis. *Orphanet Journal of Rare Diseases*, *14*(1), 195. https://doi.org/10.1186/s13023-019-1170-x

Shaikh, T. H. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: Genomic organization and deletion endpoint analysis. *Human Molecular Genetics*, *9*(4), 489–501. https://doi.org/10.1093/hmg/9.4.489

Sun, Y., Zhang, H., Kazemian, M., Troy, J. M., Seward, C., Lu, X., & Stubbs, L. (2016). ZSCAN5B and primate-specific paralogs bind RNA polymerase III genes and extra-TFIIIC (ETC) sites to modulate mitotic progression. *Oncotarget*, *7*(45), 72571–72592. https://doi.org/10.18632/oncotarget.12508

Swaby, J.-A. M., Silversides, C. K., Bekeschus, S. C., Piran, S., Oechslin, E. N., Chow, E. W. C., & Bassett, A. S. (2011). Complex congenital heart disease in unaffected relatives of adults with 22q11.2 deletion syndrome. *The American Journal of Cardiology*, *107*(3), 466–471. https://doi.org/10.1016/j.amjcard.2010.09.045

Wang, B., You, G., & Fu, Q. (2017). Human fetal heart specific coexpression network involves congenital heart disease/defect candidate genes. *Scientific Reports*, *7*(1), 46760. https://doi.org/10.1038/srep46760

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F. A., Hakonarson, H., & Bucan, M. (2007). PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research*, *17*(11), 1665–1674. https://doi.org/10.1101/gr.6861907

You, G., Zu, B., Wang, B., Fu, Q., & Li, F. (2020). Identification of miRNA–mRNA–TFs regulatory network and crucial pathways involved in tetralogy of Fallot. *Frontiers in Genetics*, *11*, 552. https://doi.org/10.3389/fgene.2020.00552

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J. D., Romano-Adesman, A., Bjornson, R. D., Breitbart, R. E., Brown, K. K., Carriero, N. J., Cheung, Y. H., Deanfield, J., DePalma, S., Fakhro, K. A., Glessner, J., Hakonarson, H., Italia, M. J., Kaltman, J. R., … Lifton, R. P. (2013). De novo mutations in histone-modifying genes in congenital heart disease. *Nature*, *498*(7453), 220–223. https://doi.org/10.1038/nature12141

Zarrei, M., MacDonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nature Reviews Genetics*, *16*(3), 172–183. https://doi.org/10.1038/nrg3871

Zhang, J., Li, X., Hu, W., & Li, C. (2020). Emerging role of TRIM family proteins in cardiovascular disease. *Cardiology*, *145*(6), 390–400. https://doi.org/10.1159/000506150

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zamariolli, M., Dantas, A. G., Nunes, N., Moysés-Oliveira, M., Sgardioli, I. C., Soares, D. C. Q., Gil-Da-Silva-Lopes, V. L., Kim, C. A., & Melaragno, M. I. (2023). Phenotypic heterogeneity in 22q11.2 deletion syndrome: Copy Number Variants as genetic modifiers for congenital heart disease in a Brazilian cohort. *American Journal of Medical Genetics Part A*, 1–9. https://doi.org/10.1002/ajmg.a.63145