














Research Article

Variants in Candidate Genes for Phenotype Heterogeneity in Patients with the 22q11.2 Deletion Syndrome

Natalia Nunes ¹, Beatriz Carvalho Nunes ¹, Malú Zamariolli ¹,
Diogo Cordeiro de Queiroz Soares ², Leonardo Caires dos Santos ¹,
Anelisa Gollo Dantas ¹, Vera Ayres Meloni ¹, Sintia Iole Belangero ¹,
Vera Lúcia Gil-Da-Silva-Lopes ³, Chong Ae Kim ², and Maria Isabel Melaragno ¹

¹Genetics Division, Department of Morphology and Genetics, Universidade Federal de São Paulo, São Paulo, Brazil

²Genetics Unit, Instituto da Criança, Universidade de São Paulo, São Paulo, Brazil

³Department of Translational Medicine, School of Medical Sciences, University of Campinas, Campinas, São Paulo, Brazil

Correspondence should be addressed to Maria Isabel Melaragno; melaragno.maria@unifesp.br

Received 25 August 2023; Revised 29 February 2024; Accepted 20 March 2024; Published 30 March 2024

Academic Editor: Xiaoye Jin

Copyright © 2024 Natalia Nunes et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

22q11.2 deletion syndrome (22q11.2DS) is a microdeletion syndrome with a broad and heterogeneous phenotype, even though most of the deletions present similar sizes, involving ~3 Mb of DNA. In a relatively large population of a Brazilian 22q11.2DS cohort (60 patients), we investigated genetic variants that could act as genetic modifiers and contribute to the phenotypic heterogeneity, using a targeted NGS (Next Generation Sequencing) with a specific Ion AmpliSeq panel to sequence nine candidate genes (*CRKL*, *MAPK1*, *HIRA*, *TANGO2*, *PI4KA*, *HDAC1*, *ZDHHC8*, *ZFPM2*, and *JAM3*), mapped in and outside the 22q11.2 hemizygous deleted region. *In silico* prediction was performed, and the whole-genome sequencing annotation analysis package (WGA) was used to predict the possible pathogenic effect of single nucleotide variants (SNVs). For the *in silico* prediction of the indels, we used the genomic variants filtered by a deep learning model in NGS (GARFIELD-NGS). We identified six variants, 4 SNVs and 2 indels, in *MAPK1*, *JAM3*, and *ZFPM2* genes with possibly synergistic deleterious effects in the context of the 22q11.2 deletion. Our results provide the opportunity for the discovery of the co-occurrence of genetic variants with 22q11.2 deletions, which may influence the patients' phenotype.

1. Introduction

The 22q11.2 deletion syndrome (22q11.2DS) results from the loss of chromosome 22 DNA segments and is the most frequent microdeletion syndrome. The presence of blocks of repetitive and similar DNA (low copy repeats, LCRs) in the 22q11.2 region predisposes to nonallelic homologous recombination, resulting in greater instability in this genomic region [1]. Most of these deletions (~90%) comprise 3 Mb between the LCR22A and LCR22D, involving around 60 genes [1, 2]. The manifestation of the clinical signs of 22q11.2DS is broad and may affect different organs and systems, ranging from mild to severe. Even though most patients have similar-sized deletions, it can result in variable

phenotypes, even in cases of inherited deletions or monozygotic twins [3–5]. 22q11.2DS phenotypes include congenital cardiac malformations, velopharyngeal dysfunction, metabolic and immunological disorders, and behavioral and cognitive difficulties, with an increased incidence of depressive anxiety, attention disorders, and schizophrenia [1, 4, 6, 7]. It has already been shown that genes outside the deleted region might be associated with psychiatric, cardiac, and immunophenotypes in the 22q11.2DS [8, 9]. One of the hypotheses lies in the fact that allelic variation, such as SNVs and indels, within the nondeleted 22q11.2 allele or in other genes mapped outside the 22q11.2 region, could influence the phenotype outcome of the 22q11.2DS [10]. Driven by this hypothesis, the whole-genome sequencing of patients

with 22q11.2DS gave indications that the *HADC1* (Histidine Decarboxylase) and *ZFPM2* (Zinc Finger Protein, FOG Family Member 2) genes may act as genetic modifiers associated with cardiac defects [11–13]. Changes in the immune system, such as inflammation and T-cell-mediated immune response, were shown to be associated with schizophrenia [14–16]. Garber et al. [17] saw that a sub-population of T-cells, the Th17 cells, influenced the development and/or regulation of psychotic symptoms in 22q11.2DS. A combined dysfunction of the relationship between the *MAPK1* (Mitogen-Activated Protein Kinase 1) and *CRKL* (Like Proto-Oncogene, Adaptor Protein) genes also appears to be related to the syndrome's phenotypic variability. In particular, *CRKL* appears to be involved in the occurrence of cardiac abnormalities, mostly tetralogy of Fallot [6, 9, 18, 19]. Still, the genetic analysis of 22q11DS remains highly elusive and is complicated by the complex regulatory circuits of early embryonic formation as well as by phenotypic heterogeneity [20, 21]. The advances in NGS (Next Generation Sequencing) have significantly increased the possibilities of genetic analysis in general, improving the chance of detecting gene variants in a substantial proportion of patients.

Studies in the literature show that some genes in hemizyosity must contribute to the phenotype of patients. Hestand et al. [22] performed candidate gene sequencing in the 22q11.2 region in 127 patients. They suggested that nonsynonymous variants found in several genes associated with the syndrome's phenotypes, including *PI4KA* (Phosphatidylinositol 4-Kinase Alpha), could result in partially functional proteins [23–25]. In addition, the *HIRA* gene (Histone Chaperone Complex) was indicated to be necessary for efficient suppression of viral infection, being involved in the chromatinization of viral DNA, and participating in intrinsic antiviral immunity [26].

Targeted NGS with gene panels offers a unique opportunity to sequence multiple genes at a lower cost and with less effort and thus is an efficient tool for mutation screening in the clinical diagnostic setting [27, 28]. This approach was used by Heike et al. [29] to sequence the *TBX1* region in patients with 22q11.2DS to identify genetic variants in *TBX1* that could influence the phenotype, and by Pulignani et al. [30] to sequence the *ZFPM2* gene in patients with nonsyndromic congenital heart defects. It is believed that the investigation of mechanisms influencing the 22q11.2DS phenotypic heterogeneity can help to understand the developmental pathways of the clinical traits involved and shed light on the management challenges of these patients [1, 31, 32]. Nevertheless, to date, no 22q11.2DS study based on targeted NGS was carried out in the Brazilian population, which is essential since the Brazilian population is mixed, whereby requires care to estimate the allelic frequencies of polymorphisms in a representative way and is a challenge considering only international databases, given that Latino populations are underrepresented. This is the first study that aimed to perform a targeted NGS of a specific gene panel in a Brazilian population of 22q11.2DS.

2. Methods

2.1. Sample. The sample of the present study consists of 60 patients with ~3 Mb deletions in 22q11.2 (average age of 19 ± 3 years in the first clinical evaluation), recruited at the Medical Genetic Center of the Universidade Federal de São Paulo (UNIFESP), at the Child Institute of the Faculty of Medicine of the Universidade de São Paulo (FMUSP), and the University Hospital of the Universidade Estadual de Campinas (UNICAMP), all in Brazil. Parental analysis of the 22q11.2 deletion was performed for at least one parent for 27 patients. Among the sample of patients, cases with and without cardiac and immunological or psychiatric alterations were evaluated. The statistical power of the sample size was calculated using the tool G Power (Universität Düsseldorf, Germany), which resulted in a statistical power of 81% with an α of 0.02 and an OR >20.

2.2. Custom 22q11.2 Gene Panel for the Risk of Cardiac or Immune-Psychiatric Phenotypes. The choice of genes for the panel was performed based on the gene expression microarray data from a previous study carried out by our research group [8], additionally to the gathering information using the following online tools: PubMed [33], to consult the literature for articles and scientific reviews; GeneCards [34], to general aggregate data on the function and pathways in which the gene is involved; UCSC Genome Browser [35], to check the genomic coordinates, size, and number of exons and introns of the gene; and OMIM (Online Mendelian Inheritance in Man) [36], to find out if the gene has ever been linked to comorbidity. The following genes were selected for the study: *CRKL*, *MAPK1*, *HIRA*, *TANGO2*, *PI4KA*, *HDAC1*, *ZDHHC8*, *ZFPM2*, and *JAM3* (based on Table 1).

2.3. Targeted NGS. The gene panel was designed with the online tool Ion AmpliSeq Designer [37] to capture coding regions, splicing sites, and immediately adjacent intron sequences. The sequencing of selected genes was performed on the equipment Ion Torrent (Thermo Fisher). The construction of libraries was performed by the Ion AmpliSeq Library Kit 2.0–96 and quantified by the Ion Library Equalizer Kit (Thermo Fisher). The template was subjected to clonal amplification in micelles using the Hi-Q Ion OT Kit (Thermo Fisher) in Ion OneTouch 2 (Thermo Fisher). The template enrichment was performed on the Ion OneTouch ES (Thermo Fisher) and applied to the Ion 316 Chip (Thermo Fisher). The tools Ion Torrent Suite and Ion Reporter (Thermo Fisher) were used for the initial data analysis. The Ion Reporter and Torrent Suite software were initially used for sequence alignment, coverage number, and determination of samples' genotypes.

2.4. Analysis of Variants and Filtering. The annotation and filtering of variants were performed from the VCF file (Variant Call Format) generated by the Ion Reporter tool (Thermo Fisher) in the UNIX environment. Variants with

TABLE 1: Genes curated for the NGS-panel.

Gene	Position	Ref seq. no.	OMIM	Association
<i>HIRA</i>	22q11.21	NP_003316.3	600237	Cardiac anomalies
<i>CRKL</i>	22q11.21	NP_005198.1	602007	Cardiac anomalies
<i>MAPK1</i>	22q11.22	NM_138957.3	176948	Cardiac anomalies
<i>HDAC1</i>	1p35.2	NP_004955.2	601241	Cardiac anomalies
<i>ZFPM2</i>	8q23	NP_036214.2	603693	Cardiac anomalies
<i>JAM3</i>	11q25	NP_116190.3	606871	Cardiac anomalies
<i>TANGO2</i>	22q11.21	NP_690870.3	616830	Immunopsychiatric disorders
<i>ZDHHC8</i>	22q11.21	NP_037505.1	608784	Immunopsychiatric disorders
<i>PI4KA</i>	22q11.21	NP_477352.3	600286	Immunopsychiatric disorders

low sequencing coverage (<30x), according to [38], were excluded. Variants with minor allele frequency (MAF, Minor Allele Frequency) < 5% were selected from the gnomAD database [39] and the Brazilian reference cohort database with 609 human genome samples ABraOM (Brazilian Online Archive of Mutations) [40]. To assess the veracity of the identified indel, the GENOMIC VARIANTS FILTERING BY DEEP LEARNING MODELS IN NGS (GARFIELD-NGS) [41] was used, which rely on machine learning models to distinguish true positive from false positive indels call. To study the possible effect of the variants, *in silico* prediction was performed using the WGS (Whole Genome Sequencing Annotation) analysis package according to Liu et al. [42]. The package has an annotation pipeline for human genome sequencing studies, aggregating databases, and prediction tools of various types; for the prediction of pathogenicity of variants, the databases and scores that were considered are shown in Supplementary Figure 1 (Supplementary Figure 1).

2.5. SNP Burden. The SNP-set (sequence) Kernel Association Test (SKAT) [43], based on Fisher's method, was used to examine whether all sequenced variants together contribute to the risk of cardiac or immunopsychiatric phenotypes.

2.6. Variants Validation. Sanger sequencing was performed to validate variants with pathogenic potential obtained from *in silico* analyses. The primers were designed for amplification reactions in the Primer3 program [44, 45], and the quality was verified by the OligoAnalyzer™ Tool [46]. Amplification was performed using the PCR Master kit (Promega). The generated amplification products were purified with the QIAquick PCR Purification kit (Qiagen) and submitted to the sequencing reaction using the commercial kit BigDye version 3.1 (Applied Biosystems) and the Genetic Analyzer 3130xl equipment (Applied Biosystems).

3. Results

3.1. Parental Analysis. For 16 patients, parental analysis was performed for both parents; for 10 patients, it was performed only for the mother; and for one patient, it was performed only for the father. Among the patients for whom only the mother sample was available, three mothers presented the

typical 3 Mb deletion in the 22q11.2 region. Two of them were part of the 60 patients included in this study. No other parent had deletions in the 22q11.2 region.

3.2. Phenotyping and SNV Burden. A total of 60 individuals with the canonical ~3 Mb deletion between LCR22A and LDR22D were assessed for traditional 22q11.2 phenotypes. Our population of patients with 22q11.2DS exhibited cardiac and immunopsychiatric phenotypes consistent with previously published literature [1], being 32 (52%) patients with cardiac malformations and 28 (46%) patients with immunopsychiatric alterations.

Among these, 19 (31%) patients had both presented cardiac malformations together with immunopsychiatric alterations. The burden analysis (SKAT) did not show any association with the absolute number of variants sequenced in the genes and the phenotype groups assessed (cardiac malformations and immunopsychiatric alterations were compared with patients without cardiac alterations and without immunopsychiatric alterations, respectively) by Fisher's exact test ($p < 0.05$).

3.3. Variant's Prediction. A total of 2,923 variants were identified in the nine genes sequenced in 60 patients, with an average of 40 variants per patient, without the application of filters (Figure 1, Supplementary Table S1). All target genes had sequencing coverage above 85% of its extension (Supplementary Table S2). After the pipeline for filtering the variants with potential effects, four single nucleotide variants and two indel variants remained. Regarding SNVs, four variants were interpreted with a possible effect on the phenotype (Table 2): rs897688340, rs13058, rs41282607, mapped in *MAPK1*, and rs7936421, mapped in *JAM3* (Table 2). The coordinate identified is also a transcription factor binding site and a target for miRNA (*MAPK1* miR-14303p). Two SNVs were predicted to be an eQTL and a transcription binding site. The variant rs897688340 is mapped at the UTR3 region of the *MAPK1* gene. It was predicted as potentially deleterious by CADD and FATHMM-XL, and it was also predicted to be likely to affect binding by RegulomeDB. The other two variants in this same gene (rs13058 and rs41282607) are registered at dbSNP by rs897688340 and rs41282607, respectively, and both are predicted as deleterious by the FATHMM-XL and CADD tools. Finally,

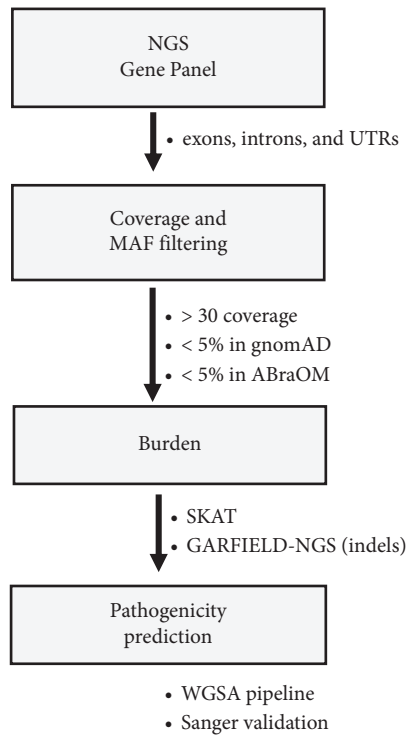


FIGURE 1: Data analysis pipeline.

the variant rs7936421, also associated in a genome-wide association study for cardiac valves, was the SNV in the *JAM3* gene, also located at an eQTL region, and predicted as deleterious by the CADD and FATHMM-XL tools. We also identified two indel variants that were highly predicted as pathogenic in all prediction tools consulted (Table 3). One of them, registered in the dbSNP with rs199956937 and mapped at the *ZFPM2* gene, was predicted to be at the target region of four miRNA (miR-130-3p; miR-17-5p; miR-143-5p; miR-340-5p). The other indel, identified in the *JAM3* gene, was registered in the dbSNP with rs3216140 and was predicted to be in the binding site of the transcription factor EZH2.

4. Discussion

In this study, we identified six genomic variants with possible effects on the phenotype of 22q11.2DS by a targeted NGS approach followed by appropriate filtering strategies. These variants will be discussed according to their prediction groups and the phenotypes of the patients in which they were identified (Table 2): [1] Expression Quantitative Trait Locus (eQTL), and Transcription Factor (TF) targets; [2] Hits on genomic association studies in a related phenotype (Table 2).

4.1. Expression Quantitative Trait Locus (eQTL) and Transcription Factor (TF) Targets. The analysis of the identified SNVs as potential eQTLs provided relevant results for three SNVs, whose genotypes could affect the expression of their related genes. The variants rs13058 and rs41282607 in

MAPK1 were found to be differentially associated in different tissues, including arteries, stomach, and thyroid, suggesting that these variants are likely to affect gene functions that are important for the body as a whole, which makes sense once they were identified in patients with cardiac malformations and immune-psychiatric alteration, respectively [47, 48]. We identified indels in the *ZFPM2* and *JAM3* genes that may be possibly pathogenic; they could be playing a role as a phenotype modifier in the patients. *ZFPM2* variants were found in three patients with cardiac malformations and had previously been associated with tetralogy of Fallot, a conotruncal heart defect commonly observed in patients with 22q11.2DS [12, 49]. The *ZFPM2* protein acts as a cofactor for dosage-sensitive GATA transcription factors during embryonic heart development in mouse models, and it is speculated that variants in this gene could lead to conotruncal heart defects [12, 50]. Regarding the *JAM3*, this gene is one of the candidate genes of the cardiac phenotype in patients with 11q25 haploinsufficiency which is an analogous syndrome to the 22q11.2DS [51, 52]. The indel variant in *JAM3* was identified in the binding site of transcription factor EZH2 which controls the methylation of H3K27 histone and can also act through methylation of nonhistone proteins, being a potential mechanism for EZH2-mediated gene activation, the perturbation of this pathway can lead to cardiac defects and was already established as a driven mechanism in some types of cancer [53–55].

4.2. Hits on Genomic Association Studies in a Related Phenotype. The SNV in the *JAM3* gene (rs7936421), mapped in chromosome 11, found in 16 patients with diverse phenotypes, and even with a higher frequency in the population, is of interest. Since this same SNV has already been associated in an association study for the cardiac phenotype in the GRASP database, this database includes variants that have been significantly associated in genome-wide association studies [56]. It is known that SNVs in eQTL can influence mRNA expression levels [57]. That is the case of the rs793642, which was predicted to be at an eQTL region in the artery aorta and muscle-skeletal tissues according to the GTEx database. Most importantly, we identified the same variants in more than one patient, suggesting that the co-occurrence of two or more rare variants may have an additive or synergistic deleterious effect with the deletion in 22q11.2. Accordingly, we can speculate that each variant alone could be tolerated. Still, when combined with another genetic event, as with the deletion, it would lead to the difference of penetrance of some of the syndrome phenotypes [28, 58]. Hestand et al. [22] studied 127 individuals with the 22q11.2DS using next-generation sequencing to sequence the genes in the 22q11.2 region of the intact allele and it was prepared a catalog of 22q11.2 hemizygous variation that could be used as a blueprint for future experiments to correlate 22q11.2DS variation with the phenotype in the Caucasian population. These authors provided insight into the phenotypic contributions of some genes in the region, but in our study, none of the variants reported were

TABLE 2: Identified SNVs and their predictions in the present sample ($n = 60$).

Variant Nomenclature*	NG_023054.2: g.109490C > T	NM_002745.5: c.*1037	NM_002745.5: c.*310	NC_000011.10: 134150005: T: C
Gene	MAPK1	MAPK1	MAPK1	JAM3
Position (hg19)	22: 22117480	22: 22117502	22: 22118229	11: 134019901
dbSNP (b151)**	rs897688340	rs13058	rs41282607	rs7936421
MAF (AbraOM)	Not identified in Brazillian population	0.050296	Not identified in Brazillian population	0.153285
MAF (gnomAD)	0.00000	0.050024	0.021495	0.162251
Location	UTR3	UTR3	UTR3	UTR3
CADD	32	31	24.3	17.07
FATHMM-XL	0.888171	0.887487	0.779486	0.803652
RegulomeDB	Likely to affect binding	Likely to affect binding	Less likely to affect binding	Likely to affect binding
SIFT	NA	NA	NA	NA
eQTLs e GTEX	NA	Artery tibial	Stomach/thyroid	Artery aorta/muscle skeletal
GRASP/trait	NA	Yes/microalbuminuria	NA	Yes/aortic valve
ORegAnno	NA	Transcription factor binding site	Transcription factor binding site	NA
miRNA-target	NA	NA	MAPK1: miR-217	NA
Patients	1 with cardiac malformation	1 with immuno/psychiatric alteration	1 with cardiac malformation	16 with variable phenotype

* According to the Human Genome Variation Society; ** Variant ID from dbSNP (b151); MAF: minor allele frequency, SIFT score: a score <0.5 is considered deleterious; Combined Annotation Dependent Depletion (CADD): a score of 20 means that a variant is amongst the top 1% of deleterious variant; FATHMM-XL: 0 to 1. Scores nearer 1 are more likely to be deleterious.

TABLE 3: Identified indels and their predictions in the present sample ($n = 60$).

Gene	<i>ZFPM2</i>	<i>JAM3</i>
dbSNP*	rs199956937	rs3216140
Position (hg19)	chr8: 106816289 C > CTT	chr11: 134014673 C > CCT
MAF (ABraOM)	Not identified in Brazilian population	0.032787
MAF (gnoMAD)	0.005	0.00003207
Genecanon	Damage	Damage
FATHMM-indel	Damage	Damage
SIFT-indel	09% closer to exon	42% closer to exon
RegulomeDB	NA	Binding site of transcription factor EZH2**
miRNA-target	miR-130-3p, miR-17-5p, miR-142-5p, miR-340-5p	NA
Patients	3 patients with congenital cardiac malformations	40 patients
Conclusion	Possibly pathogenic	Possibly pathogenic

*Variant ID from dbSNP (b151); MAF: minor allele frequency; **ChiP-Seq cluster from ENCODE with motifs.

identified. Importantly, since variant frequencies vary across populations, in our study, because we have a Brazilian sample, we also chose to use the Brazilian variant database ABraOM in the process of filtering and interpretation of the variants [22, 59]. One of the limitations of our study is the lack of functional assays performed on the large number of variants detected and the possibility that some variants that did not pass the filter pipeline could affect the phenotype. Secondly, other genes not targeted in this study may be responsible for the 22q11.2DS phenotypic variability. Finally, target NGS data processing methods are limited in detecting genomic structural variants (partial gene deletions or duplications) that have been implicated in the pathogenesis of 22q11.2DS. Despite these limitations, we provide relevant information about the genetic variants found in our cohort that may merit further studies to clarify the phenotypic heterogeneity in 22q11.2DS.

5. Conclusions

In conclusion, we performed targeted NGS in a cohort of 60 22q11.2DS Brazilian patients to investigate variants that could act as genetic modifiers. We identified six variants with possible deleterious effects in the context of the 22q11.2 deletion distributed in three genes: *MAPK1*, *JAM3*, and *ZFPM2*. These variants and genes could be related to cardiac malformations and immune-psychiatric alterations, both phenotypes present in 22q11.2DS. Moreover, the same variants could be identified in more than one patient, suggesting that the co-occurrence of two or more rare variants may have an additive or synergistic deleterious effect with the deletion in 22q11.2. To the best of our knowledge, this study was the first to apply a designed NGS target panel of 22q11.2DS-associated genes that include genes from the region deleted and outside the region, performed in a Brazilian population sample. Nevertheless, the studies that associated the selected genes with the

phenotypes studied were carried out in the majority of the Caucasian population, highlighting the relevance of studying their association in the Brazilian population with 22q11.2DS.

Data Availability

The genome data generated during this project will be made available upon request. Due to the sensitive nature of genomic information and in accordance with ethical guidelines, access to the data will be granted solely by contacting the corresponding author. Requests for the genomic data should include a brief description of the purpose and intended use of the data, along with the necessary assurances of data privacy and confidentiality. The corresponding author will assess the requests on a case-by-case basis and, if approved, provide the necessary data access and guidance to ensure its appropriate utilization.

Ethical Approval

The study was performed in accordance with the guidelines established by the Brazilian National Health Council.

Consent

The patients, their parents, and/or guardians signed an informed consent form according to the Research Ethics Committee of Universidade Federal de São Paulo, approved under number 1,156,489.

Disclosure

We certify that the submission is an original work.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All co-authors have seen and agreed with the manuscript's contents, and there is no financial interest to report.

Acknowledgments

This work was supported by the São Paulo Research Foundation, Brazil (FAPESP, grant #2019/21644-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) code 001, Brazil, and VLGSL was supported by the National Council for Scientific and Technological Development–CNPq #309782/2020-1.

Supplementary Materials

Figure S1: Prediction tools and scores used for pathogenicity prediction of variants. Table S1: Variants detected by the NGS panel before filtering. Table S2: Overall coverage per target gene of the NGS panel. Figure S1 shows the different prediction tools used for pathogenicity prediction of the different types of variants found on the patients. Missense variants with CADD >20, FATHMM >0.8, and SIFT <0.05 were considered possibly pathogenic. For nonmissense variants the same scores for CADD and FATHMM were used, as well as the RegulomDB score. For indels, we used GeneCanon and FATHMM-XL scores for pathogenicity prediction. Table S1 shows all the 2,923 variants identified in the nine target genes before filtering application. Location and identification of the variants are reported, as well as the reference and altered sequences. Table S2 shows the overall coverage per target gene of the NGS panel. All genes had an overall coverage >85%. Chromosomal location, number of amplicons, total of bases, total of covered, and missed bases and number of exons for each gene are also reported. (*Supplementary Materials*)

References

- [1] D. M. McDonald-McGinn, K. E. Sullivan, B. Marino et al., "22q11.2 deletion syndrome," *Nature Reviews Disease Primers*, vol. 1, Article ID 15071, 2015.
- [2] P. Kruszka, Y. A. Addissie, D. E. McGinn et al., "22q11.2 deletion syndrome in diverse populations," *American Journal of Medical Genetics, Part A*, vol. 173, no. 4, pp. 879–888, 2017.
- [3] E. Goldmuntz, "22q11.2 deletion syndrome and congenital heart disease," *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, vol. 184, no. 1, pp. 64–72, 2020.
- [4] J. L. Cohen, T. B. Crowley, D. E. McGinn et al., "22q and two: 22q11.2 deletion syndrome and coexisting conditions," *American Journal of Medical Genetics, Part A*, vol. 176, no. 10, pp. 2203–2214, 2018.
- [5] B. E. Morrow, D. M. McDonald-McGinn, B. S. Emanuel, J. R. Vermeesch, and P. J. Scambler, "Molecular genetics of 22q11.2 deletion syndrome," *American Journal of Medical Genetics, Part A*, vol. 176, no. 10, pp. 2070–2081, 2018.
- [6] R. Zemble, E. Luning Prak, K. McDonald, D. McDonald-McGinn, E. Zackai, and K. Sullivan, "Secondary immunologic consequences in chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome)," *Clinical Immunology*, vol. 136, no. 3, pp. 409–418, 2010.
- [7] M. Unolt, P. Versacci, S. Anaclerio et al., "Congenital heart diseases and cardiovascular abnormalities in 22q11.2 deletion syndrome: from well-established knowledge to new frontiers," *American Journal of Medical Genetics, Part A*, vol. 176, no. 10, pp. 2087–2098, 2018.
- [8] A. G. Dantas, M. L. Santoro, N. Nunes et al., "Downregulation of genes outside the deleted region in individuals with 22q11.2 deletion syndrome," *Human Genetics*, vol. 138, no. 1, pp. 93–103, 2019.
- [9] V. S. Aggarwal and B. E. Morrow, "Genetic modifiers of the physical malformations in velo-cardio-facial syndrome/DiGeorge syndrome," *Developmental Disabilities Research Reviews*, vol. 14, no. 1, pp. 19–25, 2008.
- [10] R. K. Jonas, C. A. Montojo, and C. E. Bearden, "The 22q11.2 deletion syndrome as a window into complex neuropsychiatric disorders over the lifespan," *Biological Psychiatry*, vol. 75, no. 5, pp. 351–360, 2014.
- [11] L. E. León, F. Benavides, K. Espinoza et al., "Partial microduplication in the histone acetyltransferase complex member KANSL1 is associated with congenital heart defects in 22q11.2 microdeletion syndrome patients," *Scientific Reports*, vol. 7, no. 1, p. 1795, 2017.
- [12] A. Pizzuti, A. Sarkozy, A. L. Newton et al., "Mutations of ZFPM2/FOG2 gene in sporadic cases of tetralogy of Fallot," *Human Mutation*, vol. 22, no. 5, pp. 372–377, 2003.
- [13] P. Zheng, L. M. Noroski, I. C. Hanson et al., "Molecular mechanisms of functional natural killer deficiency in patients with partial DiGeorge syndrome," *Journal of Allergy and Clinical Immunology*, vol. 135, no. 5, pp. 1293–1302, 2015.
- [14] M. Giacomelli, R. Kumar, A. Soresina et al., "Reduction of CRKL expression in patients with partial DiGeorge syndrome is associated with impairment of T-cell functions," *Journal of Allergy and Clinical Immunology*, vol. 138, no. 1, pp. 229–240.e3, 2016.
- [15] T. H. Shaikh, H. Kurahashi, S. C. Saitta et al., "Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis," *Human Molecular Genetics*, vol. 9, no. 4, pp. 489–501, 2000.
- [16] J. Steiner, R. Jacobs, B. Panteli et al., "Acute schizophrenia is accompanied by reduced T cell and increased B cell immunity," *European Archives of Psychiatry and Clinical Neuroscience*, vol. 260, no. 7, pp. 509–518, 2010.
- [17] M. Garber, M. Guttman, M. Clamp, M. C. Zody, N. Friedman, and X. Xie, "Identifying novel constrained elements by exploiting biased substitution patterns," *Bioinformatics*, vol. 25, no. 12, pp. i54–i62, 2009.
- [18] A. Bengoa-Alonso, M. Artigas-López, M. Moreno-Igoa, C. Cattalli, B. Hernández-Charro, and M. A. Ramos-Arroyo, "Delineation of a recognizable phenotype for the recurrent LCR22-C to D/E atypical 22q11.2 deletion," *American Journal of Medical Genetics, Part A*, vol. 170, no. 6, pp. 1485–1494, 2016.
- [19] D. C. Bittel, S. Yu, H. Newkirk et al., "Refining the 22q11.2 deletion breakpoints in DiGeorge syndrome by aCGH," *Cytogenetic and Genome Research*, vol. 124, no. 2, pp. 113–120, 2009.
- [20] A. Baldini, "The 22q11.2 deletion syndrome: a gene dosage perspective," *The Scientific World Journal*, vol. 6, pp. 1881–1887, 2006.
- [21] E. A. Lindsay, F. Vitelli, H. Su et al., "Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice," *Nature*, vol. 410, no. 6824, pp. 97–101, 2001.

- [22] M. S. Hestand, B. A. Nowakowska, E. Vergaelen et al., “A catalog of hemizygous variation in 127 22q11 deletion patients,” *Hum Genome Var*, vol. 3, no. 1, Article ID 15065, 2016.
- [23] Z.-Y. Sun, J. Wei, L. Xie et al., “The CLDN5 locus may be involved in the vulnerability to schizophrenia,” *European Psychiatry*, vol. 19, no. 6, pp. 354–357, 2004.
- [24] B. J. Jungerius, M. L. C. Hoogendoorn, S. C. Bakker et al., “An association screen of myelin-related genes implicates the chromosome 22q11 PIK4CA gene in schizophrenia,” *Molecular Psychiatry*, vol. 13, no. 11, pp. 1060–1068, 2008.
- [25] N. M. Williams, B. Glaser, N. Norton et al., “Strong evidence that GNBIL is associated with schizophrenia,” *Human Molecular Genetics*, vol. 17, no. 4, pp. 555–566, 2008.
- [26] T. S. Rai, M. Glass, J. J. Cole et al., “Histone chaperone HIRA deposits histone H3.3 onto foreign viral DNA and contributes to anti-viral intrinsic immunity,” *Nucleic Acids Research*, vol. 45, no. 20, pp. 11673–11683, 2017.
- [27] X. Guo, M. Delio, N. Haque et al., “Variant discovery and breakpoint region prediction for studying the human 22q11.2 deletion using BAC clone and whole genome sequencing analysis,” *Human Molecular Genetics*, vol. 25, no. 17, pp. 3754–3767, 2016.
- [28] E. E. Mlynarski, M. Xie, D. Taylor et al., “Rare copy number variants and congenital heart defects in the 22q11.2 deletion syndrome,” *Human Genetics*, vol. 135, no. 3, pp. 273–285, 2016.
- [29] C. L. Heike, J. R. Starr, M. J. Rieder et al., “Single nucleotide polymorphism discovery in TBX1 in individuals with and without 22q11.2 deletion syndrome,” *Birth Defects Research Part A: Clinical and Molecular Teratology*, vol. 88, no. 1, pp. 54–63, 2009.
- [30] S. Pulignani, C. Vecoli, A. Borghini, I. Foffa, L. Ait-Ali, and M. G. Andreassi, “Targeted next-generation sequencing in patients with non-syndromic congenital heart disease,” *Pediatric Cardiology*, vol. 39, no. 4, pp. 682–689, 2018.
- [31] B. Crowley, M. Ruffner, D. M. McDonald McGinn, and K. E. Sullivan, “Variable immune deficiency related to deletion size in chromosome 22q11.2 deletion syndrome,” *American Journal of Medical Genetics, Part A*, vol. 176, no. 10, pp. 2082–2086, 2018.
- [32] Y. Zhao, A. Diacou, H. R. Johnston et al., “Complete sequence of the 22q11.2 allele in 1,053 subjects with 22q11.2 deletion syndrome reveals modifiers of conotruncal heart defects,” *The American Journal of Human Genetics*, vol. 106, no. 1, pp. 26–40, 2020.
- [33] PubMed, “PubMed,” 2023, <https://www.ncbi.nlm.nih.gov/pubmed/>.
- [34] GeneCards, “GeneCards,” 2023, <https://www.genecards.org/>.
- [35] UCSC Genome Browser, “UCSC Genome Browser,” 2023, <https://genome.ucsc.edu/>.
- [36] Online Mendelian Inheritance in Man, “Online mendelian inheritance in man,” 2023, <https://www.omim.org/>.
- [37] Ion, “Ion AmpliSeq™ Designer,” 2023, <https://www.ampliseq.com/>.
- [38] G. Millat, V. Chanavat, and R. Rousson, “Evaluation of a new NGS method based on a custom AmpliSeq library and Ion Torrent PGM sequencing for the fast detection of genetic variations in cardiomyopathies,” *Clinica Chimica Acta*, vol. 433, pp. 266–271, 2014.
- [39] M. Lek, Exome Aggregation Consortium, K. J. Karczewski et al., “Analysis of protein-coding genetic variation in 60,706 humans,” *Nature*, vol. 536, no. 7616, pp. 285–291, 2016.
- [40] M. S. Naslavsky, M. O. Scliar, G. L. Yamamoto et al., “Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil,” *Nature Communications*, vol. 13, no. 1, p. 1004, 2022.
- [41] V. Ravasio, M. Ritelli, A. Legati, and E. Giacomuzzi, “GARFIELD-NGS: genomic vARiants Filtering by dEep Learning moDels in NGS,” *Bioinformatics*, vol. 34, no. 17, pp. 3038–3040, 2018.
- [42] X. Liu, S. White, B. Peng et al., “WGSA: an annotation pipeline for human genome sequencing studies,” *Journal of Medical Genetics*, vol. 53, no. 2, pp. 111–112, 2016.
- [43] T. Hasegawa, K. Kojima, Y. Kawai, K. Misawa, T. Mimori, and M. Nagasaki, “AP-SKAT: highly-efficient genome-wide rare variant association test,” *Bone Marrow Concentrate Genomics*, vol. 17, no. 1, p. 745, 2016.
- [44] A. Untergasser, I. Cutcutache, T. Koressaar et al., “Primer3—new capabilities and interfaces,” *Nucleic Acids Research*, vol. 40, no. 15, p. e115, 2012.
- [45] T. Koressaar and M. Remm, “Enhancements and modifications of primer design program Primer3,” *Bioinformatics*, vol. 23, no. 10, pp. 1289–1291, 2007.
- [46] OligoAnalyzer, “OligoAnalyzer,” 2023, <https://www.idtdna.com/calc/analyzer>.
- [47] A. C. Nica and E. T. Dermitzakis, “Expression quantitative trait loci: present and future,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, no. 1620, Article ID 20120362, 2013.
- [48] Y. Wang, B. He, Y. Zhao et al., “Comprehensive cis-regulation analysis of genetic variants in human lymphoblastoid cell lines,” *Frontiers in Genetics*, vol. 10, p. 806, 2019.
- [49] Z. Zhao, Y. Zhan, W. Chen, X. Ma, W. Sheng, and G. Huang, “Functional analysis of rare variants of GATA4 identified in Chinese patients with congenital heart defect,” *Genesis*, vol. 57, no. 11–12, p. e23333, 2019.
- [50] J. Lu, T. A. McKinsey, H. Xu, D. Wang, J. A. Richardson, and E. N. Olson, “FOG-2, a heart- and brain-enriched cofactor for GATA transcription factors,” *Molecular and Cellular Biology*, vol. 19, no. 6, pp. 4495–4502, 1999.
- [51] T. Mattina, C. S. Perrotta, and P. Grossfeld, “Jacobsen syndrome,” *Orphanet Journal of Rare Diseases*, vol. 4, no. 1, p. 9, 2009.
- [52] M. Ye, R. Hamzeh, A. Geddis, N. Varki, M. B. Perryman, and P. Grossfeld, “Deletion of JAM-C, a candidate gene for heart defects in Jacobsen syndrome, results in a normal cardiac phenotype in mice,” *American Journal of Medical Genetics, Part A*, vol. 149, no. 7, pp. 1438–1443, 2009.
- [53] A. Sav, F. Rotondo, L. V. Syro, M. A. Altinoz, and K. Kovacs, “Selective molecular biomarkers to predict biologic behavior in pituitary tumors,” *Expert Review of Endocrinology and Metabolism*, vol. 12, no. 3, pp. 177–185, 2017.
- [54] M. DuPage, G. Chopra, J. Quiros et al., “The chromatin-modifying enzyme Ezh2 is critical for the maintenance of regulatory T cell identity after activation,” *Immunity*, vol. 42, no. 2, pp. 227–238, 2015.
- [55] V. Sharma, P. B. Malgulwar, S. Purkait et al., “Genome-wide ChIP-seq analysis of EZH2-mediated H3K27me3 target gene profile highlights differences between low- and high-grade astrocytic tumors,” *Carcinogenesis*, vol. 38, no. 2, pp. 152–161, 2017.
- [56] H. M. Phillips, G. L. Renforth, C. Spalluto et al., “Narrowing the critical region within 11q24–qter for hypoplastic left heart and identification of a candidate gene, JAM3, expressed during cardiogenesis,” *Genomics*, vol. 79, no. 4, pp. 475–478, 2002.

- [57] S. Carrasco Pro, K. Bulekova, B. Gregor, A. Labadorf, and J. I. Fuxman Bass, "Prediction of genome-wide effects of single nucleotide variants on transcription factor binding," *Scientific Reports*, vol. 10, no. 1, Article ID 17632, 2020.
- [58] S. E. Bergen, A. Ploner, D. Howrigan et al., "Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia," *American Journal of Psychiatry*, vol. 176, no. 1, pp. 29–35, 2019.
- [59] M. S. Naslavsky, G. L. Yamamoto, T. F. Almeida et al., "Exomic variants of an elderly cohort of Brazilians in the ABraOM database," *Human Mutation*, vol. 38, no. 7, pp. 751–763, 2017.