Chromosome Res (2024) 32:6
https://doi.org/10.1007/s10577-024-09750-2

RESEARCH

A germline chimeric *KANK1-DMRT1* transcript derived from a complex structural variant is associated with a congenital heart defect segregating across five generations

Silvia Souza da Costa · Veniamin Fishman · Mara Pinheiro · Andre Rodrigueiro · Maria Teresa Sanseverino · Paulo Zielinsky · Claudia M. B. Carvalho · Carla Rosenberg · Ana Cristina Victorino Krepischi

Received: 11 December 2023 / Revised: 27 February 2024 / Accepted: 8 March 2024

Abstract Structural variants (SVs) pose a challenge to detect and interpret, but their study provides novel biological insights and molecular diagnosis underlying rare diseases. The aim of this study was to resolve a 9p24 rearrangement segregating in a family through five generations with a congenital heart defect (congenital pulmonary and aortic valvular stenosis and pulmonary artery stenosis), by applying a combined genomic analysis. The analysis involved multiple techniques, including karyotype, chromosomal microarray analysis (CMA), FISH, genome sequencing (GS), RNA-seq, and optical genome mapping (OGM). A complex 9p24 SV was hinted at by CMA results, showing three interspersed duplicated segments. Combined GS and OGM analyses revealed that the 9p24 duplications constitute a complex SV, on which a set of breakpoints matches the boundaries of the CMA duplicated sequences. The proposed structure for this complex rearrangement implies three duplications associated with an inversion of ~ 2 Mb region on chromosome 9 and a SINE element insertion at the more distal breakpoint. Interestingly, this genomic structure of rearrangement forms a chimeric transcript of the *KANK1/DMRT1 loci*, which was confirmed by both RNA-seq and Sanger sequencing on blood samples from 9p24 rearrangement carriers. Altogether with breakpoint

Responsible Editor: Amy Breman.

Silvia Souza da Costa and Veniamin Fishman contributed equally to this work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10577-024-09750-2.

S. S. da Costa · V. Fishman · M. Pinheiro · C. Rosenberg · A. C. V. Krepischi (✉)
Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil
e-mail: ana.krepischi@ib.usp.br

V. Fishman
The Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

A. Rodrigueiro
Uniscience do Brasil, São Paulo, Brazil

M. T. Sanseverino
Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil

M. T. Sanseverino
School of Medicine, Pontifícia Universidade Catolica do Rio Grande Do Sul, Porto Alegre, Brazil

P. Zielinsky
Department of Pediatrics and Childcare, Federal University of the Rio Grande do Sul, Porto Alegre, Brazil

C. M. B. Carvalho
Pacific Northwest Research Institute, Seattle, WA, USA

Published online: 19 March 2024

amplification and FISH analysis, this combined approach allowed a deep characterization of this complex rearrangement. Although the genotype–phenotype correlation remains elusive from the molecular mechanism point of view, this study identified a large genomic rearrangement at 9p24 segregating with a familial congenital heart defect, revealing a genetic biomarker that was successfully applied for embryo selection, changing the reproductive perspective of affected individuals.

Keywords Congenital pulmonary and aortic valvular stenosis · Pulmonary artery stenosis · Structural variation · Germline chimeric transcripts · *KANK1*

Introduction

Structural variations (SV) can have a significant impact on congenital human diseases (Schuy et al. 2022). Duplications at the short arm of chromosome 9 (9p) are a frequent autosomal alteration of the newborns (Temtamy et al. 2007; Guilherme et al. 2014; Cammarata-Scalisi 2019), with more than 200 cases reported up to now (Sams et al. 2022). In the majority of the cases, 9p duplication causes global developmental delay and a well-recognized spectrum of findings, such as craniofacial (mainly microcephaly and typical facial dysmorphisms) and hands/toes anomalies, accompanied by a broad range of less common additional varying features, including kidney abnormalities, other skeletal malformations, and congenital heart defects (Sams et al. 2022; Temtamy et al. 2007; Guilherme et al. 2014; Stagi et al. 2014; Morrissette et al. 2003, 9; Nakagawa et al. 1999). This phenotypic heterogeneity can be explained by the large number of genes that can be affected and the variable size of the duplicated regions. In addition, 9p duplications are mostly due to segregation of derivative chromosomes from balanced rearrangements, resulting in additional extra 9p chromosomal abnormalities, and only a few of them are de novo pure 9p duplications (Sams et al. 2022; Tkemaladze et al. 2023; Krepischi-Santos and Vianna-Morgante 2003).

Although 9p is a relatively gene-poor genomic segment, it contains more than 450 genes, some of them essential for human development. At least 50 of them were previously associated with human diseases. Several efforts were done to define specific *loci* within 9p responsible for each phenotypic manifestation (Sams et al. 2022; Wilson et al. 1985; Huret et al. 1988), resulting in the delimitation of a minimal critical subregion at 9p24-9p22 (Haddad et al. 1996; Fujimoto et al. 1998). However, there is still no consistent genotype–phenotype correlation (Cammarata-Scalisi 2019; Littooij et al. 2002).

It is noteworthy that both 9p deletions and duplications can be associated with congenital heart defects (CHD) (Morrissette et al. 2003; Nakagawa et al. 1999; Sams et al. 2022), implicating one or more *loci* for this pathology in the short arm of chromosome 9. CHD is the most common birth defect in newborns and a substantive cause of morbidity and mortality in infancy (Houyel and Meilhac 2021).

SVs, including duplications, can affect the expression of genes nearby breakpoints and even several hundred kilobases away (Kabirova et al. 2023). Here, we report a complex cryptic rearrangement at 9p24 comprising three duplications, which is segregating in a large pedigree in a dominant pattern through five generations, with 22 individuals affected by an isolated CHD (pulmonary artery and aortic stenosis). Using a combination of genomic approaches and transcriptomic analysis, we dissect the structure of this complex SV.

Results

We investigated here a 5-generation family with 22 individuals presenting with a phenotype of nonsyndromic pulmonary artery and aortic stenosis.

Clinical description

The proband was followed by a pediatric cardiologist from 2 years of age, with clinical and echocardiographic diagnosis of valvular aortic stenosis and both valvular and supravalvular pulmonary stenosis. At diagnosis, the maximal gradients were 25 mmHg through the aortic valve, 20 mmHg through the pulmonary valve, and 15 mmHg at the supravalvular pulmonary artery. The severity of the condition slowly progressed and at the age of 8 the gradients has risen to 34, 70, and 35 mmHg, respectively. He underwent percutaneous dilatation of the pulmonary valve, followed by a progressive clinical worsening of the

electrocardiographic, radiological, and echocardiographic signs, with an increase in aortic gradient to 114 mmHg at the age of 15. Cardiac surgery was then indicated, and he was submitted to implantation of a metallic aortic valve prosthesis coupled to supravalvular pulmonary artery angioplasty. Both immediate and late postoperative outcomes were uneventful, and he is still asymptomatic at 42 years of age, with near normal cardiac clinical, radiological, and echocardiographic examinations. The proband's offspring consisted of three children. The first baby had an early diagnosis of severe valvular aortic and valvular pulmonary stenosis requiring urgent cardiac surgery at 2 months of age; albeit technically successful, it was followed by clinical deterioration and rapid demise. The two other children were unaffected by the CHD and conceived by in vitro fertilization followed by embryo selection, based on the results of the genomic analysis described here.

G-banded karyotype, exome sequencing (ES), and chromosomal microarray analysis (CMA)

G-banded karyotype of the proband showed no abnormalities (Supplementary Fig. 1a), and exome sequencing did not detect any pathogenic variants. CMA at 180K resolution (Agilent Technologies) was performed in proband's DNA sample extracted from peripheral blood, revealing copy number variants (CNVs) of 9p24.3 sequences < 500 kb: two adjacent duplicated genomic segments, interspersed between a normal copy number segment (Supplementary Fig. 1b).

To further refine the breakpoints of the duplicated sequences, a high-resolution 9p microarray was used, which delimited the two 9p24.3 duplications and disclosed a third duplicated segment mapped to 9p24.2 (Fig. 1). The genomic coordinates of the three duplicated segments were as follows: **dup1**: arr[GRCh38] 9p24.3(346084_700056)×3, a 354,973 bp segment with partial duplication of *DOCK8* and *KANK1* sequences (OMIM morbid genes); **dup2**: arr[GRCh38] 9p24.3(788363_864099)×3, a 75,737 bp segment with partial *DMRT1* duplication (not a OMIM morbid gene); and **dup3**: arr[GRCh38] 9p24.2 (2271746_2351089)×3, a 79,344 bp region (no genes), centromeric to the *SMARCA2* gene.

Although the duplicated segments partially encompass *DOCK8*, *KANK1*, and *DMRT1* sequences, the resulting structure could not lead to gene disruption. Moreover, this region is covered by several overlapping CNVs (duplications and deletions) in control populations (DGV—http://dgv.tcag.ca/dgv/app/home; gnomAD—https://gnomad.broadinstitute.org/). In DGV gold standard, overlapping duplications are reported with frequencies ~ 0.1%; in gnomAD, there are scarce overlapping duplications, all of them with

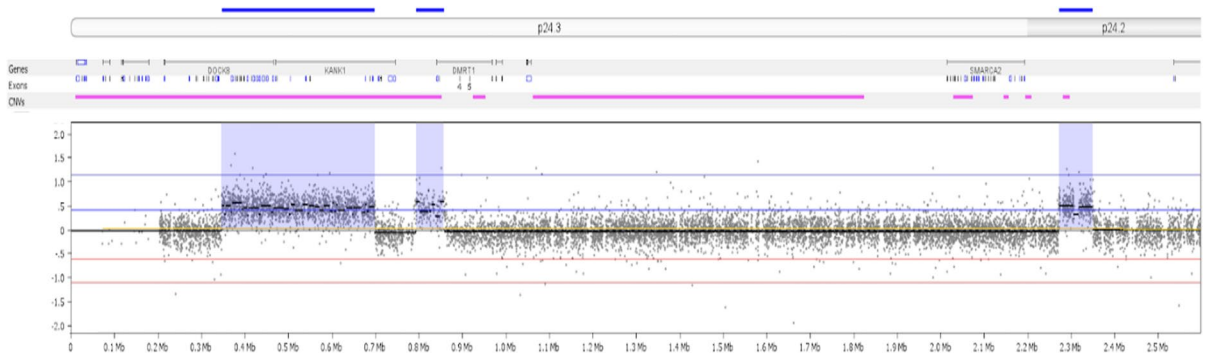

Fig. 1 Chromosomal microarray analysis (CMA) revealed a complex pattern of 9p24 duplications. The plot shows the copy number profile ($\log_2$ ratios, Y = axis) of the distal region of the short arm of chromosome 9 (9p24), with probes (black dots) depicted according to their genomic coordinates (from pter to the centromere, X-axis). To further refine the duplications breakpoints, we applied a CMA (array-CGH) based on a custom 44K (Agilent) platform covering at higher resolution the 9p sequences, which confirmed the presence of two adjacent 9p24.3 duplications (**dup1** and **dup2**) and disclosed a third (**dup3**) one at 9p24.3 (blue shadows and dark blue horizontal lines in the 9p24 ideogram). Above the CNV plot, regions with polymorphic CNVs are presented (pink horizontal lines), as well as genes mapped to the segment (black lines) with respective exons. Image extracted from Nexus Copy Number software (Bionano)

frequencies < 0.02%.Therefore, the 9p24 CNVs were classified as variants of unknown significance (VUS).

The proband is one of the 22 affected individuals of a large family with aortic and pulmonary artery stenosis transmitted in a dominant pattern through five generations (Fig. 2). To verify a possible association of the complex SV detected in the proband with the phenotype, 21 additional family members were evaluated by CMA (data not shown). The analysis revealed that the 9p24 rearrangement segregated with the CHD in all 11 affected individuals, while it was absent in all 11 normal relatives.

Genome sequencing (GS) and optical genome mapping (OGM)

To dissect the structure of the 9p chromosomal rearrangement discovered by CMA, we employed GS and OGM techniques.

OGM at 100X coverage was analyzed regarding the 9p24 rearrangement and other possible SVs. Although the duplications could be visualized in the copy number track, they were not called by either Access software pipelines (pipeline CNV, which detects copy number changes > 500 kb; and pipeline SV, which detects duplications > 30 kb). Analysis of the detected SVs showed the presence of three hybrid molecules at 9p24 (Fig. 3). These 9p24 hybrid molecules were manually analyzed based on the genomic coordinates at the breakpoints/junction sequences, identifying breakpoints and an inversion.

GS data analysis identified the three 9p24 duplicated segments, showed two breakpoints with discordant reads, and indicated that both homologs of chromosome 9 contain at least one copy of the concordant sequence around the breakpoints, as shown in Fig. 4a, b. OGM data confirmed all the breakpoints identified by GS analysis and revealed one additional junction, including an inverted segment.

Taken together, GS and OGM analyses confirmed the three 9p24 duplications reported by CMA, disclosed an SV complex pattern, and identified a partially overlapping set of breakpoints (Fig. 4), all of them matching the boundaries of the CMA duplicated sequences.

The combined evaluation of data allowed us to propose a structure for this complex rearrangement, which implies an inversion of ~ 2 Mb region on chromosome 9 with partial duplications at the breakpoint regions. Since this 2 Mb inversion is confirmed, this structure looks like a large inverted region with duplicated and nonduplicated sequences.

To validate the proposed 9p24 rearrangement structure, we amplified and sequenced by Sanger

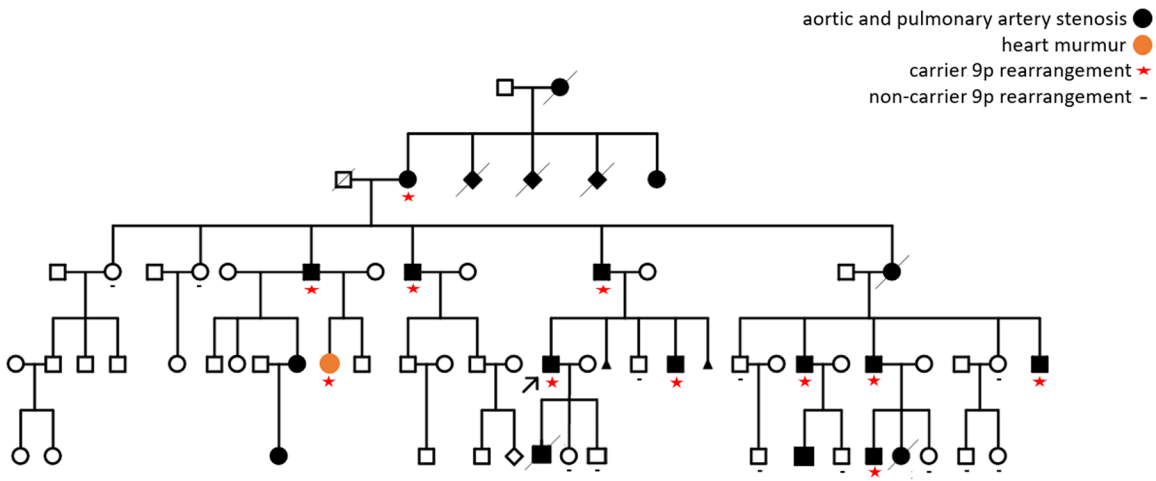

Fig. 2 Family pedigree showing five generations of individuals affected by an isolated congenital heart disease. The arrow indicates the proband (black symbols denoted affected individuals with pulmonary artery and aortic stenosis; orange is a female patient who was born only with heart murmur). All affected individuals evaluated by CMA were carriers of the 9p24 rearrangement (red asterisk), while evaluated normal family members were noncarriers (black minus symbol). The two alive affected individuals of the last generation were not tested. The two unaffected children of the proband were conceived by in vitro fertilization followed by embryo selection, based on the results of the genomic analysis described here

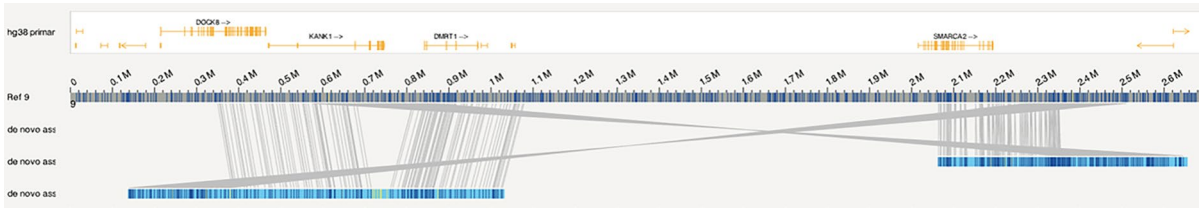

Fig. 3 Optical genome mapping reveals the structure of the 9p24 complex rearrangement. Single molecule view of proband DNA sample mapping to reference chromosome 9 with breakpoints mapped to *DOCK8*, *KANK1*, and *DMRT1*. GRCh38 reference chromosomes with OGM label patterns are shown on the top (Ref 9; gray with dark blue vertical bars), and assembled maps of hybrid molecules with label patterns are shown in light blue. Alignments between reference maps and hybrid molecules are shown as gray strings. Overlapping genes are depicted as orange bars. Image extracted from Access software

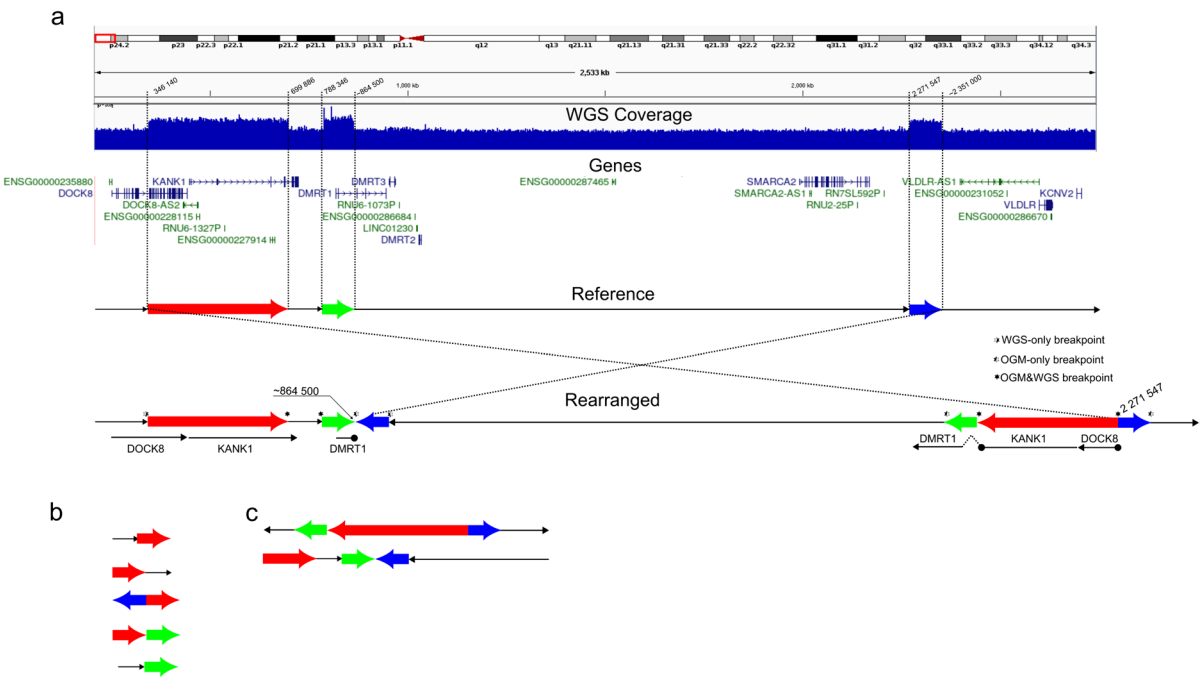

Fig. 4 Assembling the complex 9p24 chromosome rearrangement. **a** IGV screenshot showing GS read coverage, genes, and breakpoint coordinates in 9p regions. The reference sequence and proposed rearrangement structure are shown below the screenshot, with duplicated segments depicted as colored arrows (red for dup1, green for dup2, and blue for dup3). Schematic representation of the breakpoints (represented by *) identified using **b** GS and **c** OGM data

the **dup2-dup3** breakpoint, which was not covered by split-reads in the GS data. Due to the presence of homopolymer tracts within these regions, we were not able to obtain complete end-to-end sequence. Thus, we employed NGS to analyze the obtained amplicons. This analysis confirmed the junction between dup2 and inverted dup3 regions and revealed an insertion of a ~ 500 bp sequence between them that corresponds to a SINE element (Supplementary Fig. 2).

FISH analysis with BAC probes (Supplementary Fig. 3) mapped to 9p24.3 duplicated segment **dup1** showed that the additional 9p genomic copies were not inserted in other chromosomes nor moved to a distant region of chromosome 9. In addition, the rearranged chromosome 9 appeared to carry a segment

inserted between the two duplicated regions **dup1** and **dup2**; in interphase *nuclei,* using FISH probes mapped to distal duplicated 9p24.3 segments, an increased distance between the dup1 signals was observed in only the rearranged chromosome 9. We also performed FISH analysis with probes containing nonduplicated genomic sequences, one of them proximal (centromeric) to **dup3** and other within the inverted segment; in interphase *nuclei*, hybridization also showed an increased distance between these nonduplicated sequences. Taken together, this data further supported the existence of an inversion near the breakpoints.

At this point, 9p24 SV was considered a biomarker for the proband segregating with the phenotype, and this information was used for embryos selection in preimplantation genetic diagnosis (PDG), after fertilization in vitro (IFV). Two healthy noncarrier children were born after this procedure.

9p24 rearrangement results in a chimeric *KANK1-DMRT1* transcript

The proposed structure of the 9p24 rearrangement suggests that in the absence of *KANK1* termination sequence of transcription, the RNA polymerase starting at the *KANK1* promoter region should continue into the intergenic region upstream *KANK1*, and then terminate after *DMRT1* gene (Fig. 5, top panels).

This would result in the formation of a chimeric transcript including a 5′-prime fragment of *KANK1* (exons 1 and 2 of the MANE-annotated isoform of *KANK1* ENST00000382297.7) and a portion of the *DMRT1* gene. Moreover, according to the in silico prediction of GENA tool (Fishman et al. 2023), there is a pair of splice donor and acceptor sites located close to each other in a region between KANK1 and DMRT1; processing of this donor–acceptor pair results in a putative 48-bp exon, which can be included into the mature transcript sequence.

To study the structure of the transcripts originating from the *KANK1* promoter, we performed a RNA-seq experiment on blood samples from three 9p24 rearrangement carriers and three unrelated controls (Supplementary Fig. 4). In all samples, we observed high *DOCK8* expression and low-level expression of *KANK1*. In control samples, we detected no RNA-seq reads aligning to the *DMRT1* locus, consistent with the testis-specific expression pattern of this gene (Raymond et al. 1999). However, in the 9p24 SV carriers, the duplicated region including *DMRT1* gene was covered by RNA-seq reads. All reads mapped in this region support transcription from the forward strand; this strand should be observed if transcription starts from the *KANK1* gene promoter. Although the

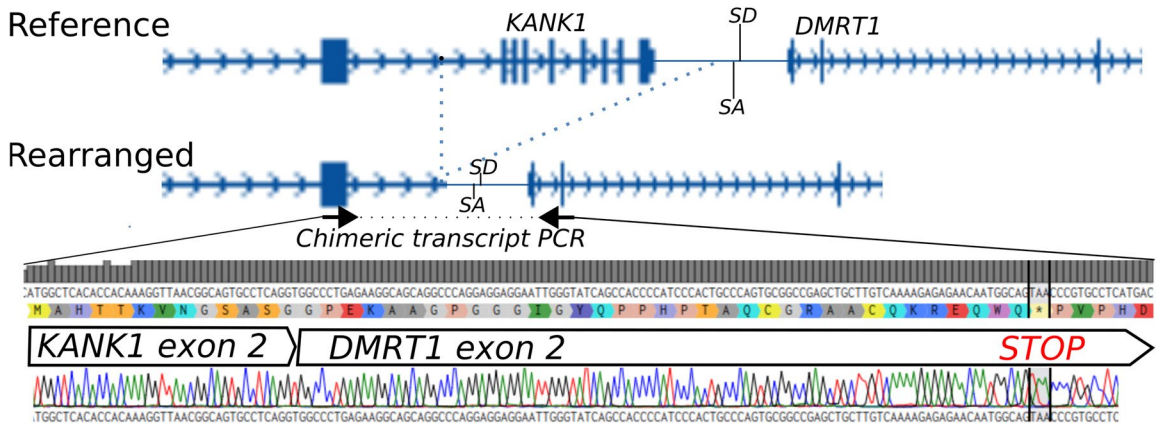

Fig. 5 Structure and sequencing data of the *KANK1-DMRT1* chimeric transcripts. Top panels show the structure of *KANK1-DMRT1* locus in the reference genome and the 9p24 rearrangement. SA and SD show positions of the in silico predicted splice acceptor (SA) and splice donor (SD) sites. Arrows indicate the positions of primers within *KANK1* exon 2 and *DMRT1* exon 2 that were used to amplify the chimeric transcript. The Sanger sequencing data of the obtained amplicons are shown at the bottom. The sequencing data shows the junction of *KANK1* exon 2 and *DMRT1* exon 2 without any extra sequence between them. According to this transcript structure, there is a premature stop codon within *DMRT1* exon 2 (depicted as an asterisk in the translation track)

coverage of this region is low, we were able to detect several RNA-seq reads suggesting the junction of the *KANK1* exon 2 with *DMRT1* exon 2.

To further validate this observation, we designed primers specific to *KANK1* exon 2 and *DMRT1* exon 2 (Fig. 5, arrows) and used blood cDNA obtained from a rearrangement carrier as a PCR template. This PCR reaction yields a product with a length equal to the combined size of *KANK1* exon 2 and *DMRT1* exon 2, and no product was detected in unrelated control samples. Sanger sequencing of the PCR product confirmed the presence of a chimeric transcript involving the junction of exons 2 from both *KANK1* and *DMRT1* (Fig. 5, bottom panel). Moreover, the amplified product did not contain any extra sequence between *KANK1* and *DMRT1* exons. Hence, the splice sites predicted by GENA in the intergenic sequence may be nonfunctional or lead to an alternative splice isoform, which might go undetected by PCR due to the substantial length of the resulting product. We note that both observed *KANK1-DMRT1* fusion and predicted transcript with extra exon contain premature stop-codons (Fig. 5), suggesting that the chimeric transcripts probably undergo nonsense-mediated RNA decay.

Altogether with breakpoint amplification and FISH analysis, this data confirms the proposed structure for the 9p24 complex SV, indicating that there is an additional copy of *KANK1*, which is truncated, and there is a read-through from this copy into the *DMRT1* locus.

Discussion

Complex rearrangements are still an underestimated cause of genetic diseases, and in some *loci*, they constitute up to 30% of the pathogenic CNVs (Schuy et al. 2022). Sensitivity of the available methods for SV detection is especially limited for resolving complex SVs involving multiple chromosomal segments. This study confirms the importance of a multiomics approach and a combination of different techniques like CMA, FISH, GS, OGM, and RNASeq to fully dissect a complex chromosomal rearrangement. CMA revealed the duplications, whereas GS/OGM allowed the refinement of the breakpoints, revealed the presence of an inversion, phasing of the multiple rearrangements *in cis*, and provided a framework for the proposal of genomic structure. Although the complex nature of the 9p24 SV was revealed by OGM, confirming breakpoints already detected by GS and revealing a new one, the duplicated segments were not called, which revealed a limitation of the system. FISH was crucial to show that the duplicated segments mapped on 9p24 and also to support the proposed structure of the rearrangement, with an inversion associated with duplications. Finally, RNA-seq and Sanger sequencing provided experimental evidence of the existence of chimeric *KANK1/DMRT1* transcripts, and in silico AI-based predictive tools assisted in analysis of the chimeric transcript structure.

Duplication/deletions restricted to the 9p24.3 cytoband, including *DOCK8* and *KANK1*, have been reported across multiple neurodevelopmental/psychiatric phenotypes (Capkova et al. 2021; Glessner et al. 2017). *DOCK8* biallelic mutations cause a recessive condition (https://omim.org/entry/243700); its disruption in heterozygosity was identified in a few patients with intellectual disability and/or seizures (Griggs et al. 2008), who were not further evaluated by the presence of additional pathogenic variants by exome analysis. This is the case for several reports of 9p24.3 CNV cases, and current data can only support a possible contribution to neurodevelopmental/psychiatric phenotypes in a multifactorial model. Therefore, an association of 9p24.3 heterozygous CNVs with clinical findings, as major variants with high impact, is still controversial. CNVs encompassing *DOCK8* or *KANK1* are detected in the general population at a relatively high frequency, and an eventual contribution to a congenital rare phenotype should be evaluated with caution. The absence of a neurodevelopment phenotype associated with the *DOCK8/KANK1* duplication (**dup1**) disclosed in our family is not surprising.

Haploinsufficiency of *DMRT1*, 2 and 3, mainly due to 9p24.3 deletions, was already associated with disorders of the sexual development, such as ambiguous external genitalia in males, as well as gonadal dysgenesis (OMIM #154230 46XY sex reversal 4; (Muroya et al. 2000; Shan et al. 2000; Livadas et al. 2003; Quinonez et al. 2013)). In the current case, there is involvement only of the *DMRT1* gene (**dup2**), and similar phenotypes are not present in the 9p24 SV carriers reported here. In association with

the duplications and inversion, we detected a non-reference (both GRCh38 and T2T) SINE insertion at one of the breakpoints disrupts one of the copies of the *DMRT1* gene. SINE is a transposable element, and its mobilization has long been associated with evolution and human diseases (Akrami and Habibi 2014; Pfaff et al. 2022). Several cases linked with SINE-VNTR-Alus rearrangements induce aberrant splicing patterns, and we cannot exclude the possibility that this insertion alters the *DMRT1* expression pattern. Copy number variants overlapping the short arm of chromosome 9 were already associated with CHD (Sams et al. 2022; Morrissette et al. 2003; Nakagawa et al. 1999) implicating one or more *loci* in this genomic region. The genetic landscape of CHD is complex, and an interesting emerging feature is that CHD mutations often alter gene/protein dosage (Fahed et al. 2013; Simmons and Brueckner 2017; Yasuhara and Garg 2021). There are several genes mapped to the short arm of chromosome 9 that have been reported in association with heart development or function, such as *KANK1* (Nguyen and Lee 2022; Botos et al. 2023; Hensley et al. 2016), *SMARCA2* (Lim et al. 2021; Wang et al. 2022), *IFT74* (Bakey et al. 2023), *PIGO* (Krawitz et al. 2012), *DNAI1* (Kennedy et al. 2007; Nakhleh et al. 2012), and *NFIB* (Rao and Goel 2020; Schanze et al. 2018). *KANK1* and *SMARCA2* are involved in the studied SV, respectively, in **dup1** and in the inversion.

SMARCA2 is not disrupted by the rearrangement, but it is included in the inverted segment. The haploinsufficiency of *SMARCA2* causes two dominant developmental conditions, namely, blepharophimosis-impaired intellectual development syndrome (OMIM #619293) and Nicolaides-Baraitser (OMIM #601358), with other clinical signs including CHD. However, as both conditions are associated with severe syndromic intellectual disability, it is not probable that its expression is disrupted by the rearrangement.

Regarding *KANK1*, deletion of the paternal allele was reported in one single family to cause the condition named cerebral palsy, spastic quadriplegic 2 (OMIM #612900); however, no following studies support this association. Indeed, chromosome 9 uniparental disomy is not related to imprinted syndromes (Elbracht et al. 2020), and clinical findings in UPD(9) are commonly attributed to homozygous variants in genes related to recessive conditions or residual trisomy in mosaic. Currently, there is no clinical evidence for haploinsufficiency or triplosensitivity of *KANK1* (*KANK1* curation results for Dosage Sensitivity). Notwithstanding, we have found evidence in literature proposing a role for *KANK1* in cardiac development (Nguyen and Lee 2022; Botos et al. 2023). *KANK* genes are scaffold proteins, bridging microtubules to focal adhesion sites (Botos et al. 2023; Pan et al. 2018). The Kank1 protein expression was shown to be widely distributed in various murine tissues, with relatively high levels in cardiac muscle (Nguyen and Lee 2022). In humans, the longest transcript (NM_015158) shows tissue specific expression, predominantly in the heart and kidney. In addition, it was found in an injury-specific gene regulatory network in a transcriptome analysis related to cardiac regeneration in zebrafish (Botos et al. 2023).

It is not clear how a complex SV involving three DNA segments was formed, with six breakpoints (two in each CNV) with three breakpoint junctions. At both sides flanking the dup2-dup3 breakpoint, we observed microhomologies of simple repeats composed of polyA/T sequences. However, insertion of a nonreference SINE element between dup2 and dup3 argues against nonallelic recombination caused by homology of these polyA/T sequences. Alternatively, the SINE insertion might be present in the ancestral chromosome on which the rearrangement took place or is an additional event occurring after SV has been formed. It is interesting to note that the transcriptome analysis detected the presence of chimeric transcripts encompassing *KANK1* and *DMRT1* exons, which was further confirmed by cDNA amplification and Sanger sequencing. The detection of chimeric transcripts may reinforce a modified product of *KANK1* as a candidate for the phenotype. The role of chimeric transcripts as cause of congenital defects is poorly explored (Zuccherato et al. 2016), in contrast to fusion transcripts commonly described as somatic events in cancer (Salokas et al. 2023). Only isolated cases were reported related to the detection of chimeric transcripts (gene fusions) as underlying molecular cause of developmental/neurological phenotypes (Boone et al. 2014; Ferrari et al. 2017). Recently, two studies employed an approach of detecting chimeric transcripts using RNA-seq data in rare congenital diseases, one of them with individuals with birth defects (Yamada et al. 2021; Oliver et al. 2019), leading to an increased diagnostic rate. Both in silico analysis

of potential splice variants and experimental evaluation of transcript isoforms in the current case suggest the presence of a premature stop-codon in the fusion transcript, which probably would undergo nonsense-mediated RNA decay. However, some level of escape from the nonsense-mediated RNA decay mechanism is evident by the validated chimeric transcript, and an eventual contribution of this fusion *KANK1-DMRT1* gene to the cardiac phenotype remains to be fully explored.

Considering the recent report of ultra-long-range interactions between active regulatory elements (Friman et al. 2023), distant 9p genes with normal copy number could be misregulated due to this 9p rearrangement, which makes the derivation of genotype-to-phenotype association relationships even more complicated. In particular, the study of this SV was crucial for genetic counseling and reproductive choices of the family. Even without the identification of the precise mechanism underlying the CHD phenotype, this study identified the SV as a biomarker that was used to identify embryos at risk and select for implantation those without the CHD risk. This strategy resulted in a healthy offspring for at least one couple.

Patients and methods

Patients and genomic samples

Written informed consent for this study was obtained from affected individuals or their parents. Genomic DNA samples were extracted from peripheral blood of 22 family members ($n = 11$ patients and $n = 11$ non-affected relatives), using standard procedures (phenol–chloroform followed by ethanol precipitation). RNA samples were obtained from peripheral blood of three male patients and three nonrelated male controls using the RNeasy Mini Kit (QIAGEN).

GTG-banded karyotype, FISH, and chromosomal microarray analysis (CMA)

Peripheral blood temporary culture (72 h) was performed in the presence of phytohemagglutinin, and GTG-banding was obtained according to standard methods. FISH analysis based on metaphase spreads and interphase preparations was performed using BAC clones, as previously described (A. C. V. Krepischi-Santos et al. 2009), with genomic sequences mapped to the short arm of chromosome 9.

Chromosome microarray analysis (array-CGH) was done according to the manufacturer's instructions, using a 180 K and a custom tiling-path high-resolution chromosome 9p oligonucleotide platform (Agilent Technologies (Grochowski et al. 2018)). Data were extracted and analyzed for copy number changes using the software Nexus Copy Number Discovery (Bionano) and analyzed as previously reported (Krepischi et al. 2022). CNV calls were based on at least three consecutive probes with aberrant $\log_2$ ratios and compared with the Database of Genomic Variants (http://projects.tcag.ca/variation/) aiming to exclude polymorphisms (frequency > 1%). Detected CNVs were classified according to five tiers of pathogenicity, following the guidelines suggested by (Riggs et al. 2020). The following public databases were consulted: University of California Santa Cruz (UCSC) (http://genome.ucsc.edu/), OMIM (https://omim.org/), ClinGen Dosage Sensitivity (https://search.clinicalgenome.org/kb/gene-dosage/cnv?page=1&size=25&search=), DECIPHER (https://www.deciphergenomics.org/), and PubMed (https://pubmed.ncbi.nlm.nih.gov/).

Genome sequencing (GS) analysis

GS data of six individuals (carriers of the 9p rearrangement, and three unrelated controls) was obtained. Briefly, genomic libraries were constructed with 1 µg of genomic DNA and sequenced on the Illumina HiSeq 2500 platform using 150 base paired end reads (~30× coverage). Reads were aligned to the GRCh38 human genome reference using the BWA algorithm (Li 2013) to generate the BAM files, and PCR duplicates were removed from further investigation by Picard tools (v.1.8, http://broadinstitute.github.io/picard/). The Genome Analysis Toolkit (GATK 3.7) (McKenna et al. 2010) was used to realign indels, recalibrate the bases, and call (Unified Genotyper) and recalibrate variants (VQSR).

For analysis, based on the alignment results, we computed tracks showing the depth of coverage (using deepTools bamCoverage) (Ramírez et al. 2016) and discordant read pairs (using samtools) (Danecek et al. 2021). These data were visualized using the IGV software. Breakpoint structures were assessed

based on the following filters: the presence of split-reads (≥ 4 in case, no reads in control), with matching supplementary segment sequence, and the presence of discordantly aligned mates (≥ 5 in case, ≤ 1 in control), with matching mate alignment coordinates. The orientation of the DNA segments was determined based on the alignment locations and strands. For two breakpoints where discordant reads were detected, we identified single-nucleotide variants (SNV) near the breakpoints. Analyzing SNV distribution in reads, we classified all pairs as follows: (1) concordant read pair with reference sequence, (2) discordant read pair with an alternative sequence, and (3) concordant read pair with an alternative sequence. These data indicate that both homologs of chromosome 9 contain at least one copy of the concordant sequence around the breakpoint, as shown in the first two lines in Fig. 4B.

Optical genome mapping (OGM) data analysis

OGM was conducted with ultra-high molecular weight DNA samples (> 150 kb) extracted from peripheral blood cells of the proband using the Bionano Prep SP Blood and Cell DNA Isolation kit (Bionano, San Diego, CA, USA). DNA labeling was performed using the DLS DNA Labeling Kit (Bionano, San Diego, CA, USA) to add fluorophores to the specific motif "CTTAAG," and the sample was run on the Saphyr chip to collect data on the Saphyr System (Bionano, San Diego, CA, USA) at 100× coverage. OGM data were analyzed using the De Novo Assembly pipeline, followed by CNV and SV pipelines, and visualized using the Bionano Access software.

RNA-seq analysis

Total RNA samples extracted from peripheral blood of three patients and three unrelated male controls were used to build cDNA libraries using the TruSeq®Stranded Total RNA LT-kit (with Ribo-Zero TM Gold) (Illumina, USA). Sequencing was performed on the NextSeq 500 platform Mid Output v2 Kit (150 cycles) (Illumina, USA). The FASTQ files were aligned against the ribosomal reference sequence (NCBI, 12/2017) using the BWA software [26] version 0.7.17-r1188, in MEM mode, with the standard parameters, except for the -t 4 parameters. Reads not aligned to ribosomal sequences went to the alignment step against the reference sequence of the human genome (version GRCh37-hg19) using the STAR software [27], version 2.6.1a_08-27. The annotation database (GTF file) used was the Ensembl file in version 87 in the same build as the human genome reference (GRCh37).

Genomic breakpoint and chimeric transcript sequencing

One of the breakpoints was amplified by PCR using proband's and control genomic DNA as template ((94 °C, 4 min; 94 °C, 30 s; 67 °C, 40 s; 72 °C, 2 min) × 14 cycles, decreasing the annealing temperature by 0.5 °C after each cycle; (94 °C, 30 s; 60 °C, 40 s; 72 °C, 2 min) × 23 cycles; 72 °C, 10 min; 4 °C). Primer sequences were 9pD3_2350050F:tgaggtcaagca tctttttatatg, 9pD2_863684F:cgtcagatttcggacccaca, and 9pD2_863969F: gtgcaagtttcccccGACTA. Libraries of the two amplicons were prepared using E6177S NEBNext ultra II FS library preparation with beads and E7710S NEBNext multiplex oligos for illumina set 3 (index CACTCA and CAGGCG) and sequenced on an Illumina MiSeq equipment with MiSeq Reagent Nano Kit v2 (300-cycles). The total covered genomic region was ~ 2 Mb. Sequence reads were aligned to the reference human genome (hg37).

The chimeric transcript was amplified with blood cDNA from a carrier using the following primer sequences: r9pF1 TGCATGACTCCTCACTCCTT and r9pR1 CTGCAGTGGTGGGGACAC.

Acknowledgements Our sincere thanks are due to the individuals of the family here described for providing consent to participate in this study. We would like to thank Mrs. Ligia Sumi VIeira and Ms. Galina Koksharova for the technical support.

Author contribution CR and ACVK conceived and designed the study. Material preparation and data collection were performed by SSC, VF, MP, AR, MTS, and PZ. SSC, VF, CMBC, CR, and ACVK analyzed data. The first draft of the manuscript was written by VF and ACVK, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The researchers from the Institute of Bioscience (University of São Paulo) have financial support of Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2013/08028–1) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 305806/2019–0). CMBC is supported by the National Institute of General Medical Sciences (NIGMS R01 GM132589). V.F. was partially supported by the budget project of Institute of Cytology and Genetics SB RAS (state project FWNR-2022–0019) and FAPESP fellowship (2022/11064–9).

Data availability No datasets were generated or analyzed during the current study.

Declarations

Ethics approval Approval was obtained from the ethics committee of University of São Paulo, Brazil. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study or their legal guardians, including publishing their data.

Competing interests The authors declare no competing interests.

References

Akrami SM, Habibi L (2014) Retrotransposons and pediatric genetic disorders: importance and implications. J Pediatr Genet 3(1):9–16. https://doi.org/10.3233/PGE-14081

Bakey Z, Cabrera OA, Hoefele J, Antony D, Kaman Wu, Stuck MW, Micha D et al (2023) IFT74 variants cause skeletal ciliopathy and motile cilia defects in mice and humans. PLoS Genet 19(6):e1010796. https://doi.org/10.1371/journal.pgen.1010796

Boone PM, Yuan Bo, Campbell IM, Scull JC, Withers MA, Baggett BC, Beck CR et al (2014) The Alu-rich genomic architecture of SPAST predisposes to diverse and functionally distinct disease-associated CNV alleles. Am J Hum Genet 95(2):143–161. https://doi.org/10.1016/j.ajhg.2014.06.014

Botos MA, Arora P, Chouvardas P, Mercader N (2023) Transcriptomic data meta-analysis reveals common and injury model specific gene expression changes in the regenerating zebrafish heart. Sci Rep 13(April):5418. https://doi.org/10.1038/s41598-023-32272-6

Cammarata-Scalisi F (2019) Trisomy 9p. a brief clinical, diagnostic and therapeutic description. Arch Argent Pediatr 117(5):e473–e476. https://doi.org/10.5546/aap.2019.eng.e473

Capkova Z, Capkova P, Srovnal J, Adamova K, Prochazka M, Hajduch M (2021) Duplication of 9p24.3 in three unrelated patients and their phenotypes, considering affected genes, and similar recurrent variants. Mol Genet Genomic Med 9(3):e1592. https://doi.org/10.1002/mgg3.1592

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A et al (2021) Twelve years of SAMtools and BCFtools. GigaScience 10(2):giab008. https://doi.org/10.1093/gigascience/giab008

Elbracht M, Binder G, Hiort O, Kiewert C, Kratz C, Eggermann T (2020) Clinical spectrum and management of imprinting disorders. Med Gen 32(4):321–334. https://doi.org/10.1515/medgen-2020-2044

Fahed AC, Gelb BD, Seidman JG, Seidman CE (2013) Genetics of congenital heart disease: the glass half empty. Circ Res 112 (4). https://doi.org/10.1161/CIRCRESAHA.112.300853. https://doi.org/10.1161/CIRCRESAHA.112.300853

Ferrari L, Scuvera G, Tucci A, Bianchessi D, Rusconi F, Menni F, Battaglioli E, Milani D, Riva P (2017) Identification of an atypical microdeletion generating the RNF135-SUZ12 chimeric gene and causing a position effect in an NF1 patient with overgrowth. Hum Genet 136(10):1329–1339. https://doi.org/10.1007/s00439-017-1832-5

Fishman V, Kuratov Y, Petrov M, Shmelev A, Shepelin D, Chekanov N, Kardymon O, Burtsev M (2023) GENA-LM: a family of open-source foundational models for long DNA sequences. bioRxiv. https://doi.org/10.1101/2023.06.12.544594

Friman ET, Flyamer IM, Marenduzzo D, Boyle S, Bickmore WA (2023) Ultra-long-range interactions between active regulatory elements. Genome Res 33(8):1269–1283. https://doi.org/10.1101/gr.277567.122

Fujimoto A, Lin MS, Schwartz S (1998) Direct duplication of 9p22–>p24 in a child with duplication 9p syndrome. Am J Med Genet 77(4):268–271

Glessner JT, Li J, Wang D, March M, Lima L, Desai A, Hadley D et al (2017) Copy number variation meta-analysis reveals a novel duplication at 9p24 associated with multiple neurodevelopmental disorders. Genome Med 9(1):106. https://doi.org/10.1186/s13073-017-0494-1

Griggs BL, Ladd S, Saul RA, DuPont BR, Srivastava AK (2008) Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. Genomics 91(2):195–202. https://doi.org/10.1016/j.ygeno.2007.10.011

Grochowski CM, Shen Gu, Yuan Bo, Tcw J, Brennand KJ, Sebat J, Malhotra D et al (2018) Marker chromosome genomic structure and temporal origin implicate a chromoanasynthesis event in a family with pleiotropic psychiatric phenotypes. Hum Mutat 39(7):939–946. https://doi.org/10.1002/humu.23537

Guilherme RS, Meloni VA, Perez ABA, Pilla AL, Paula MA, de Ramos A, Dantas G, Takeno SS, Kulikowski LD, Melaragno MI (2014) Duplication 9p and their implication to phenotype. BMC Med Genet 15(December):142. https://doi.org/10.1186/s12881-014-0142-1

Haddad BR, Lin AE, Wyandt H, Milunsky A (1996) Molecular cytogenetic characterisation of the first familial case of partial 9p duplication (P22p24). J Med Genet 33(12):1045–1047. https://doi.org/10.1136/jmg.33.12.1045

Hensley MR, Cui Z, Chua RFM, Simpson S, Shammas NL, Yang J-Y, Leung YF, Zhang GuangJun (2016) Evolutionary and developmental analysis reveals KANK genes were co-opted for vertebrate vascular development. Sci Rep 6(June):27816. https://doi.org/10.1038/srep27816

Houyel L, Meilhac SM (2021) Heart development and congenital structural heart defects. Annu Rev Genomics Hum Genet 22(August):257–284. https://doi.org/10.1146/annurev-genom-083118-015012

Huret JL, Leonard C, Forestier B, Rethoré MO, Lejeune J (1988) Eleven new cases of del(9p) and features from 80 cases. J Med Genet 25(11):741–749. https://doi.org/10.1136/jmg.25.11.741

Kabirova E, Nurislamov A, Shadskiy A, Smirnov A, Popov A, Salnikov P, Battulin N, Fishman V (2023) Function and evolution of the loop extrusion machinery in animals. Int J Mol Sci 24(5):5017. https://doi.org/10.3390/ijms24055017

Kennedy MP, Omran H, Leigh MW, Dell S, Morgan L, Molina PL, Robinson BV et al (2007) Congenital heart disease and other heterotaxic defects in a large cohort of patients with primary ciliary dyskinesia. Circulation 115(22):2814–2821. https://doi.org/10.1161/CIRCULATIONAHA.106.649038

Krawitz PM, Murakami Y, Hecht J, Krüger U, Holder SE, Mortier GR, Chiaie BD et al (2012) Mutations in PIGO, a member of the GPI-anchor-synthesis pathway, cause hyperphosphatasia with mental retardation. Am J Hum Genet 91(1):146–151. https://doi.org/10.1016/j.ajhg.2012.05.004

Krepischi ACV, Villela D, Souza S, da Costa PC, Mazzonetto JS, Migliavacca MP, Milanezi F et al (2022) Chromosomal microarray analyses from 5778 patients with neurodevelopmental disorders and congenital anomalies in Brazil. Sci Rep 12(1):15184. https://doi.org/10.1038/s41598-022-19274-6

Krepischi-Santos ACV, Vianna-Morgante AM (2003) Disclosing the mechanisms of origin of de novo short-arm duplications of chromosome 9. Am J Med Genet A 117A(1):41–46. https://doi.org/10.1002/ajmg.a.10634

Krepischi-Santos ACV, Rajan D, Temple IK, Shrubb V, Crolla JA, Huang S, Beal S et al (2009) Constitutional haploinsufficiency of tumor suppressor genes in mentally retarded patients with microdeletions in 17p13.1. Cytogenet Genome Res 125(1):1–7. https://doi.org/10.1159/000218743

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. https://doi.org/10.48550/ARXIV.1303.3997

Lim TB, Foo SYR, Chen CK (2021) The role of epigenetics in congenital heart disease. Genes 12(3):390. https://doi.org/10.3390/genes12030390

Littooij AS, Hochstenbach R, Sinke RJ, van Tintelen P, Giltay JC (2002) Two cases with partial trisomy 9p: molecular cytogenetic characterization and clinical follow-up. Am J Med Genet 109(2):125–132. https://doi.org/10.1002/ajmg.10322

Livadas S, Mavrou A, Sofocleous C, van Vliet-Constantinidou C, Dracopoulou M, Dacou-Voutetakis C (2003) Gonadoblastoma in a patient with Del(9)(P22) and sex reversal: report of a case and review of the literature. Cancer Genet Cytogenet 143(2):174–177. https://doi.org/10.1016/s0165-4608(02)00849-x

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K et al (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20(9):1297–1303. https://doi.org/10.1101/gr.107524.110

Morrissette JJD, Laufer-Cahana A, Medne L, Russell KL, Venditti CP, Kline R, Zackai EH, Spinner NB (2003) Patient with trisomy 9p and a hypoplastic left heart with a tricentric chromosome 9. Am J Med Genet A 123A(3):279–284. https://doi.org/10.1002/ajmg.a.20293

Muroya K, Okuyama T, Goishi K, Ogiso Y, Fukuda S, Kameyama J, Sato H et al (2000) Sex-determining gene(s) on distal 9p: clinical and molecular studies in six cases. J Clin Endocrinol Metab 85(9):3094–3100. https://doi.org/10.1210/jcem.85.9.6771

Nakagawa M, Kato H, Aotani H, Kondo M (1999) Ebstein's anomaly associated with trisomy 9p. Clin Genet 55(5):383–385

Nakhleh N, Francis R, Giese RA, Tian X, Li Y, Zariwala MA, Yagi H et al (2012) High prevalence of respiratory ciliary dysfunction in congenital heart disease patients with heterotaxy. Circulation 125(18):2232–2242. https://doi.org/10.1161/CIRCULATIONAHA.111.079780

Nguyen MT, Lee W (2022) Kank1 is essential for myogenic differentiation by regulating actin remodeling and cell proliferation in C2C12 progenitor cells. Cells 11(13):2030. https://doi.org/10.3390/cells11132030

Oliver GR, Tang X, Schultz-Rogers LE, Noemi Vidal-Folch W, Jenkinson G, Schwab TL, Gaonkar K et al (2019) A tailored approach to fusion transcript identification increases diagnosis of rare inherited disease. PLoS ONE 14(10):e0223337. https://doi.org/10.1371/journal.pone.0223337

Pan W, Sun K, Tang K, Xiao Q, Ma C, Cong Yu, Wei Z (2018) Structural insights into ankyrin repeat–mediated recognition of the kinesin motor protein KIF21A by KANK1, a scaffold protein in focal adhesion. J Biol Chem 293(6):1944–1956. https://doi.org/10.1074/jbc.M117.815779

Pfaff AL, Singleton LM, Kōks S (2022) Mechanisms of disease-associated SINE-VNTR-Alus. Exp Biol Med (Maywood, N.J.) 247(9):756–64. https://doi.org/10.1177/15353702221082612

Quinonez SC, Park JM, Rabah R, Owens KM, Yashar BM, Glover TW, Keegan CE (2013) 9p partial monosomy and disorders of sex development: review and postulation of a pathogenetic mechanism. Am J Med Genet A 161A(8):1882–1896. https://doi.org/10.1002/ajmg.a.36018

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T (2016) deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44(W1):W160-165. https://doi.org/10.1093/nar/gkw257

Rao ACA, Goel H (2020) Pathogenic nonsense variant in NFIB in another patient with dysmorphism, autism spectrum disorder, agenesis of the corpus callosum, and intellectual disability. Eur J Med Genet 63(12):104092. https://doi.org/10.1016/j.ejmg.2020.104092

Raymond CS, Kettlewell JR, Hirsch B, Bardwell VJ, Zarkower D (1999) Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Dev Biol 215(2):208–220. https://doi.org/10.1006/dbio.1999.9461

Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, Raca G et al (2020) Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med 22(2):245–257. https://doi.org/10.1038/s41436-019-0686-8

Salokas K, Dashi G, Varjosalo M (2023) Decoding oncofusions: unveiling mechanisms, clinical impact, and prospects for personalized cancer therapies. Cancers 15(14):3678. https://doi.org/10.3390/cancers15143678

Sams EI, Ng JK, Tate V, Hou Y-C, Cao Y, Antonacci-Fulton L, Belhassan K et al (2022) From karyotypes to precision genomics in 9p deletion and duplication syndromes. Human Genet Genomics Adv 3(1):100081. https://doi.org/10.1016/j.xhgg.2021.100081

Schanze I, Bunt J, Lim JWC, Schanze D, Dean RJ, Alders M, Blanchet P et al (2018) NFIB haploinsufficiency is associated with intellectual disability and macrocephaly. Am J Hum Genet 103(5):752–768. https://doi.org/10.1016/j.ajhg.2018.10.006

Schuy J, Grochowski CM, Carvalho CMB, Lindstrand A (2022) Complex genomic rearrangements: an underestimated cause of rare diseases. Trends Genet 38(11):1134–1146. https://doi.org/10.1016/j.tig.2022.06.003

Shan Z, Zabel B, Trautmann U, Hillig U, Ottolenghi C, Wan Y, Haaf T (2000) FISH Mapping of the sex-reversal region on human chromosome 9p in two XY females and in primates. Eur J Hum Genet 8(3):167–173. https://doi.org/10.1038/sj.ejhg.5200431

Simmons MA, Brueckner M (2017) The genetics of congenital heart disease…understanding and improving long term outcomes in congenital heart disease: a review for the general cardiologist and primary care physician. Curr Opin Pediatr 29(5):520–528. https://doi.org/10.1097/MOP.0000000000000538

Stagi S, Lapi E, Seminara S, Guarducci S, Pantaleo M, Giglio S, Chiarelli F, de Martino M (2014) Long-term auxological and endocrinological evaluation of patients with 9p trisomy: a focus on the growth hormone-insulin-like growth factor-I axis. BMC Endocr Disord 14(January):3. https://doi.org/10.1186/1472-6823-14-3

Temtamy SA, Kamel AK, Ismail S, Helmy NA, Aglan MS, El Gammal M, El Ruby M, Mohamed AM (2007) Phenotypic and cytogenetic spectrum of 9p trisomy. Genet Couns (Geneva, Switzerland) 18(1):29–48

Tkemaladze T, Bregvadze K, Papiashvili N, Gagua S, Abzianidze E (2023) A de novo chromosome 9p duplication in a female child with short stature and developmental delay. SAGE Open Med Case Rep 11(March):2050313X231160883. https://doi.org/10.1177/2050313X231160883

Wang YJ, Zhang X, Lam CK, Guo H, Wang C, Zhang S, Wu JC, Snyder M, Li J (2022) Systems analysis of de novo mutations in congenital heart diseases identified a protein network in the hypoplastic left heart syndrome. Cell Syst 13(11):895-910.e4. https://doi.org/10.1016/j.cels.2022.09.001

Wilson GN, Raj A, Baker D (1985) The phenotypic and cytogenetic spectrum of partial trisomy 9. Am J Med Genet 20(2):277–282. https://doi.org/10.1002/ajmg.1320200211

Yamada M, Suzuki H, Watanabe A, Uehara T, Takenouchi T, Mizuno S, Kosaki K (2021) Role of chimeric transcript formation in the pathogenesis of birth defects. Congenit Anom 61(3):76–81. https://doi.org/10.1111/cga.12400

Yasuhara J, Garg V (2021) Genetics of congenital heart disease: a narrative review of recent advances and clinical implications. Transl Pediatr 10(9):2366–86. https://doi.org/10.21037/tp-21-297

Zuccherato LW, Alleva B, Whiters MA, Carvalho CMB, Lupski JR (2016) Chimeric transcripts resulting from complex duplications in chromosome Xq28. Hum Genet 135(2):253–256. https://doi.org/10.1007/s00439-015-1614-x

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work, the authors used ChatGPT to revise the text correcting grammatical errors. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.