



# Exome Sequencing Identifies Rare Variants in Multiple Genes in Atrioventricular Septal Defect

## Citation

D'Alessandro, L. C., S. Al Turki, A. K. Manickaraj, D. Manase, B. J. Mulder, L. Bergin, H. C. Rosenberg, et al. 2018. "Exome Sequencing Identifies Rare Variants in Multiple Genes in Atrioventricular Septal Defect." *Genetics in medicine : official journal of the American College of Medical Genetics* 18 (2): 189-198. doi:10.1038/gim.2015.60. <http://dx.doi.org/10.1038/gim.2015.60>.

## Published Version

doi:10.1038/gim.2015.60

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:37298209>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Published in final edited form as:

*Genet Med.* 2016 February ; 18(2): 189–198. doi:10.1038/gim.2015.60.

## Exome Sequencing Identifies Rare Variants in Multiple Genes in Atrioventricular Septal Defect

Lisa C.A. D'Alessandro<sup>#1</sup>, Saeed Al Turki<sup>#2,3</sup>, Ashok Kumar Manickaraj<sup>1</sup>, Dorin Manase<sup>1</sup>, Barbara J.M. Mulder<sup>4</sup>, Lynn Bergin<sup>5</sup>, Herschel C. Rosenberg<sup>6</sup>, Tapas Mondal<sup>7</sup>, Elaine Gordon<sup>8</sup>, Jane Lougheed<sup>9</sup>, John Smythe<sup>10</sup>, Koen Devriendt<sup>11</sup>, UK10K Consortium, Shoumo Bhattacharya<sup>12</sup>, Hugh Watkins<sup>12</sup>, Jamie Bentham<sup>13</sup>, Sarah Bowdin<sup>14</sup>, Matthew E. Hurles<sup>2,\*\*</sup>, and Seema Mital<sup>1,\*\*</sup>

<sup>1</sup>Division of Cardiology, Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, Ontario, M5G 1X8, Canada <sup>2</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK <sup>3</sup>Department of Pathology, King Abdulaziz Medical City, Riyadh, 11426, Saudi Arabia <sup>4</sup>Heart Center, Academic Medical Center, Amsterdam, The Netherlands <sup>5</sup>Division of Cardiology, Department of Medicine, London Health Sciences Centre, London, Ontario, N6A 5A5, Canada <sup>6</sup>Department of Paediatrics, London Health Sciences Centre, London, Ontario, N6A 5W9, Canada <sup>7</sup>Department of Pediatrics, Hamilton Health Sciences Centre, Hamilton, Ontario, L8N 3Z5, Canada <sup>8</sup>Division of Cardiology, Department of Medicine, Hamilton Health Sciences Centre, Hamilton, Ontario, L8N 3Z5, Canada <sup>9</sup>Division of Cardiology, Department of Pediatrics, Children's Hospital of Eastern Ontario, Ottawa, Ontario, K1H 8L1, Canada <sup>10</sup>Department of Pediatrics, Kingston General Hospital, Kingston, Ontario, K7L 2V7, Canada <sup>11</sup>Centre for Human Genetics, Katholieke Universiteit Leuven, 3000 Leuven, Belgium <sup>12</sup>Radcliffe Department of Medicine & Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK <sup>13</sup>Department of Cardiology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02459, USA <sup>14</sup>Division of Genetics, Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, Ontario, M5G 1X8, Canada

<sup>#</sup> These authors contributed equally to this work.

### Abstract

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*\*Correspondence to: Seema Mital, Professor of Pediatrics, University of Toronto, Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G 1X8, Canada, Tel: 416-813-7418, Fax: 416-813-5857, [seema.mital@sickkids.ca](mailto:seema.mital@sickkids.ca); Matthew E. Hurles, The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA UK, Tel: +44 (0) 1223 495377, Fax: +44 (0) 1223 494919, [meh@sanger.ac.uk](mailto:meh@sanger.ac.uk).

#### Consortia

Members of the UK10K Rare Diseases Cohorts Working Group are: Matthew Hurles (co-chair), David R. FitzPatrick (co-chair), Saeed Al-Turki, Carl Anderson, Ines Barroso, Philip Beales, Jamie Bentham, Shoumo Bhattacharya, Keren Carss, Krishna Chatterjee, Sebhattin Cirak, Catherine Cosgrove, Allan Daly, Jamie Floyd, Chris Franklin, Marta Futema, Steve Humphries, Shane McCarthy, Hannah Mitchison, Francesco Muntoni, Alexandros Onoufriadis, Victoria Parker, Felicity Payne, Vincent Plagnol, Lucy Raymond, David Savage, Peter Scambler, Miriam Schmidts, Robert Semple, Eva Serra, Jim Stalker, Margriet van Kogelenberg, Parthiban Vijayarangakannan, Klaudia Walter, and Gretta Wood.

**Conflict of Interest:** The authors have no conflicts of interest to declare.

**Purpose**—The genetic etiology of atrioventricular septal defect (AVSD) is unknown in 40% cases. Conventional sequencing and arrays have identified the etiology in only a minority of non-syndromic individuals with AVSD.

**Methods**—Whole exome sequencing was performed in 81 unrelated probands with AVSD to identify potentially causal variants in a comprehensive set of 112 genes with strong biological relevance to AVSD.

**Results**—A significant enrichment of rare and rare/damaging variants was identified in the gene set, compared with controls (odds ratio 1.52, 95% confidence interval 1.35–1.71,  $p = 4.8 \times 10^{-11}$ ). The enrichment was specific to AVSD probands compared with a non-AVSD cohort with tetralogy of Fallot (odds ratio 2.25, 95% confidence interval 1.84–2.76,  $p = 2.2 \times 10^{-16}$ ). Six genes (*NIPBL*, *CHD7*, *CEP152*, *BMPR1a*, *ZFPM2* and *MDM4*) were enriched for rare variants in AVSD compared to controls, including three syndrome-associated genes (*NIPBL*, *CHD7*, *CEP152*). The findings were confirmed in a replication cohort of 81 AVSD probands.

**Conclusion**—Mutations in genes with strong biological relevance to AVSD, including syndrome-associated genes, can contribute to AVSD even in those with isolated heart disease. The identification of a gene set associated with AVSD will facilitate targeted genetic screening in this cohort.

### Keywords

whole exome sequencing; targeted sequencing; atrioventricular septal defect; endocardial cushion defect; atrioventricular canal defect; congenital heart disease; Cornelia de Lange syndrome; CHARGE syndrome

### Introduction

Atrioventricular septal defect (AVSD) is a relatively rare disorder, representing 7% of all congenital heart defects (CHD).<sup>1</sup> Although AVSD is most often associated with trisomy 21 or heterotaxy, approximately 40% of patients have neither of these associations. Within this etiologically diverse group of patients, approximately 30% have an identifiable genetic diagnosis relating to a chromosomal or single gene disorder, 10% have non-diagnostic dysmorphic features or extracardiac anomalies, and the remainder have no discernible extracardiac anomalies (non-syndromic AVSD).<sup>1,2</sup> To date, targeted sequencing,<sup>3–5</sup> and interrogation for large copy number variants<sup>6</sup> have identified the etiology in only a minority of non-syndromic individuals with AVSD. Traditional gene discovery strategies are limited in AVSD due to locus heterogeneity, possible reduced penetrance and relatively small numbers of cases and families. We recently reported results of an unsupervised analysis of whole exome sequencing (WES) data in AVSD probands that identified rare causal variants in *NR2F2.7* To identify additional genetic causes in this cohort, we analyzed for potentially causal variants in all known genes with biological relevance to AVSD.

## Methods

### Study Population

Patients were prospectively enrolled in an Ontario province-wide Heart Centre Biobank registry after informed consent. The study was approved by the Research Ethics Boards of the Hospital for Sick Children and all other participating sites.<sup>8,9</sup> The primary cohort included unrelated probands with AVSD of Caucasian race. Principal component analysis was performed to confirm overlap with HapMap individuals of Caucasian ancestry (Figure S1). Patients with an identified chromosomal or syndromic disorder, or a *situs* anomaly were excluded. Detailed cardiac and extra-cardiac features were assessed through review of medical records, imaging, and dysmorphology assessment, when possible.

Data for replication analysis (n = 81) were obtained from WES of 44 AVSD probands (GO-CHD, Oxford registry; Leuven, Belgium) provided by the Wellcome Trust Sanger Institute (WTSI) (previously described)<sup>7</sup>, and 37 AVSD probands from the CONCOR registry DNA-bank of the University Cardiology Institute Netherlands (ICIN) and the Ontario Heart Centre Biobank Registry who underwent targeted sequencing. All probands in the replication cohort were unrelated, non-syndromic and Caucasian. Sixty-four unrelated Caucasian probands from the Heart Centre Biobank Registry with isolated Tetralogy of Fallot (TOF) underwent WES as a non-AVSD CHD cohort.

Variant calls from WES of 4300 individuals of European American ancestry from the NHLBI Exome Sequencing Project Exome Variant Server<sup>10</sup> (EVS) served as controls (“EVS Controls”). Additional controls included whole genome sequencing data from 40 unaffected parents of children with autism spectrum disorders (S. Walker, S.W. Scherer, personal communication), and targeted genotyping in 97 unaffected Caucasian population-matched Ontario controls without heart disease.

### Whole Exome Sequencing

WES was performed using the Agilent SureSelect Target Enrichment Kit (V3 50Mb) (Agilent Technologies Inc., Santa Clara, CA) for sequence capture and Illumina HiSeq2500 (75 bp paired-end reads) for sequencing (Illumina Inc., San Diego, CA) to a target depth of 100X, as previously published.<sup>7</sup> The non-AVSD cohort was sequenced using the Agilent SureSelect Target Enrichment Kit (V4 51 Mb) (Agilent Technologies Inc., Santa Clara, CA) for sequence capture and Illumina HiSeq (90bp paired-end reads) for sequencing (Illumina Inc. San Diego, CA) (BGI Genomics, Philadelphia, PA) to a target depth of 50X.

### AVSD Gene List

A systematic review of the literature was undertaken to curate all genes associated with AVSD identified in previous human and animal studies (Table S1). This generated a list of 112 AVSD genes including 47 associated with human syndromes, 30 with animal models, 20 with both human and animal references, 9 with isolated AVSD in humans, and 6 reported as cardiac modifiers in trisomy 21. Of these 112 genes, 58 were associated with syndromic CHD (“syndromic” genes).

## Variant Filtering

Variants in these 112 genes were filtered for quality, leaving only calls with a passing GATK score (99% accuracy in read determination)<sup>11</sup> and read-depth coverage of  $\geq 30$ . Non-synonymous, non-synonymous splice, stop-gain and stop-loss variants were prioritized. The variants were annotated with Annovar<sup>12</sup> to obtain the minor allele frequency (MAF) for each variant from 1000Genomes<sup>13</sup> and EVS.<sup>10</sup> **Rare variants**, defined as  $MAF < 0.01$  in 1000G and EVS European American (EA) ancestry, were prioritized, and included **novel variants** i.e. variants not present in either 1000Genomes, EVS datasets or in the Database of Single Nucleotide Polymorphisms (dbSNP).<sup>14</sup> Variants were annotated with SIFT<sup>15</sup> and PolyPhen2<sup>16</sup> and any variant predicted to be pathogenic in at least one program was called **damaging**. **Evolutionary conservation** was defined as PhyloP score  $> 0$  (range -14 to +6; sites predicted to be conserved are assigned positive scores)<sup>17</sup> and/or PhastCONS  $\geq 0.8$  (range 0-1 where 1 indicates a highly conserved locus).<sup>18</sup>

## Mutation Burden Analysis

The overall burden of rare (including novel) non-synonymous variants and rare synonymous variants in the AVSD, EVS and non-AVSD CHD cohorts was assessed across all 112 genes. The EVS does not provide sample level information; therefore for consistency the burden analysis assumes that each genotype represents an independent sample. Variant data from EVS Control and non-AVSD CHD cohorts were filtered in the same manner as the case data. To assess mutation burden, Fisher's exact test was performed across all 112 genes for the number of individuals with and without rare and rare/damaging non-synonymous and rare synonymous variants in cases and controls. A control burden analysis was also performed using a list of 11 autism-spectrum disorder genes<sup>19</sup> (refer to Figure S2).

## Coverage Analysis

To ensure that a difference in mutation burden across these 112 genes was not secondary to differences in coverage of genomic regions between cohorts, coverage analysis was performed in cases and controls. Consensus coding sequence exon coordinates for all exons in 112 genes were extracted from University of California Santa Cruz Table Browser<sup>20</sup> and the mean depth per coding base pair was calculated across all BAM files as previously described.<sup>21</sup> The percentage of bases covered to a depth of  $\geq 20$  for 112 genes was calculated and the distribution was compared between AVSD, EVS and non-AVSD CHD cohorts using Student's T-Test (two-tailed). A nucleotide-by-nucleotide coverage analysis was then undertaken to verify adequate case and control coverage at each locus corresponding to variant location. Variants were excluded from the mutation burden calculation if the average depth of coverage in either cohort was  $< 20$ .

## Variant Validation

All prioritized variants were verified by Sanger's di-deoxy nucleotide sequencing. The amplified products were subjected to di-deoxy nucleotide PCR amplification and were sequenced on 3730xl DNA Analyzer (Applied Biosystems, Life Technologies, Grand Island, NY) (Table S2.1). The sequence chromatograms were aligned to primer target regions using

CodonCode Aligner (CodonCode Corporation, Centerville, MA). Parental genotypes were evaluated if parental DNA was available.

## Replication

Genes enriched for rare (including novel) and rare/damaging non-synonymous variants in AVSD cases vs. EVS controls on mutation burden analysis were selected for targeted sequencing in the replication cohort using Agilent Haloplex Custom Target Enrichment (Agilent Technologies Inc., Santa Clara, CA). A custom probe set was designed using Agilent SureDesign web-based software to target all coding exons for the prioritized genes (earray.chem.agilent.com/suredesign). Sequencing (100 bp paired-end reads, targeted at >200X coverage) was performed using Illumina HiSeq2500 (Illumina Inc., San Diego, CA) at The Centre for Applied Genomics (TCAG) in The Hospital for Sick Children. Agilent SureCall software was used to generate BAM files and perform variant calling. Variants were annotated and filtered using the same protocol as for the primary cohort. To assess the population frequency of the recurring variant in *MDM4* [8:204518457 A>C (K324Q, rs41299595)], genotyping was performed in 97 Caucasian controls without heart disease from Ontario using a custom-designed TaqMan<sup>®</sup> SNP Assays (Life Technologies Corporation) (Table S2.2). Samples were analyzed using the ViiA<sup>™</sup> 7 Real-Time PCR System with ViiA<sup>™</sup>7 software.

## Results

Of 260 individuals with AVSD enrolled in the Heart Centre Biobank Registry, 83 probands were eligible for inclusion and underwent WES (Table 1). Variants in 112 genes from 81 unrelated probands (2 probands did not sequence) were filtered to identify 163 unique rare variants, of which 91 were predicted to be damaging (Figure S3).

### The AVSD Cohort was enriched for rare and rare/damaging variants

A mutation burden analysis across all 112 genes was conducted to assess for enrichment of rare variants in the AVSD cohort compared with EVS Controls. There was no difference between the two cohorts in the frequency of rare synonymous variants (1.6% vs. 1.6%; Odds Ratio 1.01,  $p = 0.9$ ). The frequency of rare non-synonymous variants however, was higher in the AVSD cohort (3.2%) compared to the EVS Controls (2.1%) (OR 1.52,  $p = 4.8 \times 10^{-11}$ ), as was the frequency of rare/damaging variants in the AVSD (1.5%) versus EVS Controls (1.2%) (OR 1.26,  $p = 0.01$ ) (Figure 1, Table S3.1). We also observed an enrichment of rare non-synonymous variants in the AVSD cohort vs. EVS controls when comparing only the subset of 58 syndromic genes (OR 1.32,  $p=0.001$ ) (Table S3.1). There was no enrichment of rare or rare damaging non-synonymous variants in the AVSD cohort compared with the EVS cohort when the same analysis was conducted with a control set of 11 autism-spectrum genes (Figure S2).

To assess whether the enrichment was specific to AVSD or also seen in other forms of CHD, mutation burden was compared to a non-AVSD CHD cohort (TOF). Again, there was no significant difference between the cohorts in the frequency of rare synonymous variants (1.6% vs. 1.4%; OR 1.12,  $p = 0.3$ ) (Figure 1, Table S3.2). However, the AVSD cohort was

enriched for rare non-synonymous variants in the 112 genes (3.2% vs. 1.4%; OR 2.33,  $p = 2.2 \times 10^{-14}$ ) and rare damaging non-synonymous variants (1.5% vs. 0.7%; OR 2.11,  $p = 2.1 \times 10^{-6}$ ) compared with the non-AVSD CHD cohort. Enrichment of variants in the AVSD vs. non-AVSD CHD cohort was also seen when analysis was limited to only 58 syndromic genes (Table S3.2)

To ensure that enrichment of variants in the AVSD cohort was not related to differences in coverage between cohorts, we compared coverage across the 112 genes in the AVSD cohort versus EVS Control and non-AVSD CHD cohorts. The mean percentage of bases covered to a depth of 20 for all 112 genes was similar between the AVSD ( $83.0 \pm 22\%$ ) versus the non-AVSD CHD cohort ( $81.8 \pm 14.6\%$ ,  $p = 0.63$ ) but lower than the EVS controls ( $89.7 \pm 18.9\%$ ,  $p = 0.02$ ). Therefore the higher mutation burden in AVSD cases was not due to underreporting of variants in controls or in non-AVSD CHD cohort (Figure S4). The enrichment for rare and rare/damaging variants in AVSD cases remained significant when analysis was corrected for nucleotide-by-nucleotide coverage by excluding variants that did not meet the depth cut-offs in both cohorts (Tables S3.1, 3.2).

### Genotype-phenotype associations

As reported previously, *NR2F2* was significantly enriched for both rare and rare/damaging variants.<sup>7</sup> Besides *NR2F2*, six genes showed significant enrichment for rare and novel non-synonymous variants, with *MDM4* remaining significant after correction for multiple testing across 112 genes. These genes, *NIPBL*, *CHD7*, *CEP152*, *BMPR1a*, *ZFPM2* and *MDM4*, were prioritized for validation by Sanger sequencing, replication, and determination of inheritance patterns (Table 2; Figure S5). Across the 6 prioritized genes, 23 probands had rare/novel damaging variants (Tables 3, 4 and S4). Given that many of these genes are syndrome-associated, patients were further assessed for extra-cardiac anomalies. Genotype-phenotype associations for individual genes are described below.

**NIPBL**—Six AVSD probands (7.4%) had rare non-synonymous variants in *NIPBL* compared with 2.3% in EVS controls (OR 3.3,  $p=0.02$ ) (Table 2). The M1318V and S2471T variants were novel, highly conserved and predicted to be damaging (Tables 3 and S4). The N105D and N393K variants were each previously seen in a single individual among 4300 EVS EA controls, and not in the other control datasets. Both variants are conserved and the N105D variant is predicted to be damaging. The A179T variant (AVSD\_65) was also identified in one individual in the replication cohort (Table S5) and in one non-AVSD CHD proband. Although it has been previously reported in 29 individuals in the EA EVS, and in 1000G and dbSNP, it is listed in the Human Gene Mutation Database as a potentially disease causing variant (Table S4). An additional variant I314V was identified in the replication cohort (Table S5), which has not been previously reported in public datasets. *NIPBL* mutations are known to be associated with Cornelia de Lange syndrome (CDL). However, none of the AVSD probands with *NIPBL* variants had clinical characteristics of CDL (Table 4) although 2 probands had associated semilunar valve anomalies (AVSD\_33 and AVSD\_72), which are commonly associated with CDL.

**CHD7**—Nine AVSD probands (11.1%) had rare non-synonymous variants in *CHD7* compared with 5.0% of EVS controls (OR = 2.3, p=0.04) (Table 2). The T93A, Y1325H, A2105V and S2383C variants were all novel or exceptionally rare (Y1325H seen in 1 EVS individual). All were conserved and the A2105V, Y1325H and S2383C variants were predicted to be damaging. The M340V and M2527L variants seen in probands and in the replication cohort (Table S5) were also seen in the EVS dataset (62 and 43 individuals) and were predicted to be benign. Three additional rare non-synonymous variants were identified in the replication cohort: A1950T, P2083S and L935F (Table S5). None of the variants were seen in 40 unaffected parents from an Ontario-based cohort of families with autism spectrum disorders. *CHD7* mutations are known to be associated with CHARGE syndrome. However, six of nine probands with *CHD7* variants had isolated cardiac disease. One proband had dysmorphic features and extra-cardiac anomalies but none of the diagnostic features of CHARGE syndrome. The other two probands had single minor anomalies, unrelated to CHARGE syndrome (Table 4).

**CEP152**—Rare non-synonymous variants in *CEP152* were seen in 9.7% of AVSD cases compared with 4.3% in EVS controls (OR = 2.4, p=0.03) (Table 2). The L581I variant was novel, while the T1524A, L1105V, W960R, I294V, and S85R variants were all rare. Only the W960R variant was highly conserved and predicted to be damaging. Two additional variants were identified in the replication cohort, G181D and R115Q (Table S5). None of the variants were seen in 40 unaffected parents from an Ontario-based cohort of families with autism spectrum disorders. All variants were heterozygous and none of the probands had features consistent with Seckel syndrome, dwarfism or microcephaly, known to be associated with defects in *CEP152* (Table 4).

**BMPR1a**—Rare non-synonymous variants in *BMPR1a* were identified with in 3.7% of AVSD cases compared with 0.7% in EVS controls (OR = 5.3, p = 0.02) (Table 2). All three variants (R478H, D429V and P481S) were exceptionally rare and predicted to be damaging (Table 4). Two probands (AVSD\_2 and AVSD\_17) had isolated cardiac disease, and AVSD\_57 also had learning and psychiatric disabilities and cervical spine anomalies. AVSD\_17 and AVSD\_57 both had a left superior vena cava to coronary sinus (Table 4).

**ZFPM2**—Rare non-synonymous variants in *ZFPM2* were seen in 13.6% of AVSD cases compared with 6.7% of EVS controls (p = 0.03, OR = 2.2). Two variants (P361S and G885S) were exceptionally rare in EVS and absent in the other control datasets. Both these variants were conserved and the P361S variant was predicted to be damaging. Six of the 9 AVSD probands with *ZFPM2* variants had a second variant in one of the 6 priority genes, including 2 probands with 2 heterozygous *ZFPM2* variants. Of these, AVSD\_53 had associated anomalies including pulmonary atresia with major aortopulmonary collaterals, and multiple extracardiac anomalies including a diaphragmatic hernia (P361S and E30G). The E30G, D98N, M544I and V631I variants in the AVSD cohort and the S210T variant in the non-AVSD TOF cohort were also identified in the replication cohort, along with D1051Y, a novel variant. Additionally, 5 individuals in the non-AVSD CHD cohort had *ZFPM2* variants, none of which recurred in the AVSD cohort.



**MDM4**—Rare non-synonymous variants in *MDM4* were seen in 7.4% of AVSD cases compared with 1.2% of EVS controls (OR = 5.0,  $p = 3.2 \times 10^{-4}$ ). One variant, K324Q was seen in 5 (unrelated) probands (MAF = 0.061), and also in 1 unaffected autism control (S. Walker and S.W. Scherer personal communication) and 40 EVS individuals (MAF = 0.0047). As this single variant was seen with a high frequency in our cohort, genotyping was undertaken in 97 unaffected controls. The K324Q variant was seen at a lower frequency in controls (2/97 or 2.1%) compared to AVSD cases (6.1%) ( $p = 0.003$ ). The P338A variant identified in AVSD\_29 was seen in 4 EVS individuals, but not in 1000G or in unaffected autism controls. An additional three variants were identified in the replication cohort (Table S5). All the probands with *MDM4* variants had isolated cardiac disease.

### Compound mutations

Of the 34 probands with variants in the 6 prioritized genes, 8 probands had more than one rare or rare/damaging non-synonymous variant in more than one gene, of which 6 had a second mutation in *ZFPM2* as detailed above. Additionally, AVSD\_15 had variants in *CHD7*, *NIPBL* and *CEP152* and AVSD\_72 had variants in *CEP152* and *NIPBL*. Interestingly, AVSD\_15 and AVSD\_72 shared the same *NIPBL* variant (N393K), which was seen only in one individual in the EVS.

### Discussion

We conducted a comprehensive gene analysis using next-generation sequencing in a cohort of unrelated AVSD probands to identify genomic variants associated with AVSD. The strength of our approach included a comprehensive gene selection methodology identifying genes with biological associations to AVSD, the majority of which have not been systematically examined in a large cohort of patients with AVSD. There are three important findings from our study.

First, genes with known biological associations to AVSD were enriched for rare and rare/damaging non-synonymous variants compared to healthy controls. Mutation burden analysis at the gene level suggested an association between six genes (*NIPBL*, *CHD7*, *CEP152*, *BMPR1a*, *ZFPM2* and *MDM4*) and AVSD of which the association with *MDM4* remained significant after multiple testing. We used stringent methods to ensure that the results remained true when analysis was adjusted for gene-by-gene and nucleotide-by-nucleotide coverage between cases and controls.

Second, we demonstrated enrichment of variants in 112 genes not only in AVSD cases compared to EVS controls, but also in comparison to a non-AVSD cohort of TOF probands. This enrichment suggests that at least some of the associations were specific to AVSD and not to CHD in general and that there may be gene-specific cardiac phenotypes related to unique genetic contributions to cardiac morphogenesis.<sup>22,23</sup> While a recent paper analyzed all CHD phenotypes together to identify *de novo* mutations,<sup>24</sup> our results suggest the importance of studying probands with distinct cardiac phenotypes in order to improve the specificity of the observed associations and improve the predictive ability for clinical applications like prenatal and reproductive counselling in the future.

Finally, “syndromic” genes were enriched for mutations despite the absence of classic syndromic features. These mutations were seen in patients with either isolated cardiac disease or subtle manifestations that did not meet diagnostic criteria for a clinical syndrome. In particular, three of the top-ranked associated genes were “syndromic” genes i.e. *NIPBL*, *CHD7* and *CEP152* suggesting that a search for mutations should include “syndromic” genes regardless of clinical phenotype.

An important “syndromic” gene was *NIPBL*, a causal gene for autosomal dominant CDL (OMIM 122470), a developmental disorder. Approximately 50% CDL patients have *NIPBL* variants, typically with milder phenotypes in those with missense mutations.<sup>25,26</sup> Thirty percent of CDL patients have CHD, including septal defects, semilunar valve anomalies or AVSD.<sup>1,27</sup> To our knowledge, isolated CHD has not previously been reported with *NIPBL* mutations.<sup>28</sup> In our cohort, *NIPBL* was enriched for rare and rare-damaging variants. Yet, none of the probands had features of CDL syndrome (Table 4), although we cannot rule out subtle manifestations, as we were unable to re-contact all the probands for dysmorphology assessment. These findings suggest a contributory role for *NIPBL* not only in CHD associated with CDL syndrome, but also in cases with isolated AVSD.

Another “syndromic” gene enriched for variants was *CHD7*, the only gene known to cause CHARGE syndrome (OMIM 214800), a rare autosomal dominant condition with a prevalence of up to 1:8500.<sup>29</sup> This was not unique to AVSD but was also seen in the TOF cohort, a lesion type known to be associated with CHARGE syndrome. Approximately 75% of individuals with CHARGE syndrome have CHD.<sup>29</sup> Mutations in *CHD7* are identified in over 90% of patients with the CHARGE phenotype and incomplete penetrance has not previously been reported.<sup>30</sup> *CHD7* was also recently reported as part of the H3K4me gene mutations associated with CHD.<sup>24</sup> In our study, none of the 9 individuals with *CHD7* mutations had features of CHARGE syndrome.

*CEP152*, another syndromic gene, is involved in regulation of cellular response to DNA damage in Seckel syndrome.<sup>31</sup> Seckel syndrome (OMIM 210600) is a rare autosomal recessive disorder characterized by severe growth retardation, microcephaly and mental retardation, as well as partial AVSD and other forms of CHD.<sup>33,33</sup> *CEP152* was enriched for rare non-synonymous heterozygous variants in AVSD probands, none of whom had features of Seckel syndrome. While at a variant level, a causal role for *CEP152* appears less compelling, its role as a phenotype modifier cannot be excluded. Together, these findings reiterate that absence of obvious syndromic features should not preclude a search for mutations in “syndromic” genes.

Three patients had three exceptionally rare or novel variants in *BMPR1a* (Table 3 and Table S4) that were highly conserved and predicted damaging. One proband (AVSD\_17) had unbalanced AVSD with coarctation, the same cardiac phenotype as described previously in a patient with a 10q22q23 deletion encompassing *BMPR1a*.<sup>34</sup> *Bmpr1a* (*Alk3*) knockout mice show embryonic lethality between E10.5 and E11.5 with failure of endocardial cushion formation, failure of AV valve and adjacent septal formation, consistent with a specific and critical role in the formation of the AV canal.<sup>22,23</sup> Our findings are consistent with prior human and animal studies that implicate *BMPR1a* in the development of isolated AVSD.

*ZFPM2* is a zinc-finger nuclear protein expressed in the developing heart and known to alter *GATA4* transcriptional activity.<sup>35,36</sup> Knock-out mouse models demonstrate cardiac defects including AVSD,<sup>36,37</sup> DORV, TOF and coronary artery anomalies. *ZFPM2* variants have also been reported in patients with conotruncal defects, but not with AVSD.<sup>38,39</sup> The enrichment of variants in *ZFPM2* in 9 probands in our primary AVSD cohort and 8 probands in the replication cohort is a new finding. Nonetheless, it is unclear if *ZFPM2* variants play a causal or a modifier role in AVSD since the variants frequently co-segregated with variants in other candidate genes (Table S4), and although rare, were seen in multiple EVS cases. Also, the E30G variant was not specific to AVSD and has not been associated with functional effects on *GATA4*-mediated transcription and is therefore likely to be a benign polymorphism.<sup>40</sup> This is in contrast to the majority (4 of 5) of the variants identified in the non-AVSD TOF cohort, which were novel (Table S4).

Six AVSD probands had variants in *MDM4*, a p53 inhibitor, with *MDM4* showing the strongest association with AVSD compared to the other five genes. *MDM4* is involved in regulating endocardial EMT and subsequent proliferation of mesenchyme in the endocardial cushions. A mouse model heterozygous for *Mdm2* and *Mdm4* had impaired atrioventricular valvuloseptal development with significantly reduced mesenchymal cell density at the endocardial cushions.<sup>41</sup> Two rare/damaging and conserved variants were identified, one of which was a recurring variant (K324Q) seen in five probands. The variant frequency was higher than controls suggesting that this variant may be potentially disease causing in AVSD cases.

There were some limitations to our study. Functional validation of individual variants was beyond the scope of this study; however all variants were in genes with known association with CHD. We did not assess for structural variation, frame-shift variants, or non-coding regulatory variants. A paucity of parental samples limited our ability to ascertain transmission of the variants or perform segregation analysis. Additionally, local control data was limited; our control population was derived from public datasets therefore we corrected for two limitations, coverage and population stratification.

In summary, a biology-guided approach allowed us to identify enrichment of variants in genes that play a central role in cardiac development in general and in AV canal development in particular. There is compelling data to support a contributory role for *NIPBL*, *CHD7*, *BMPR1a*, and *MDM4* gene variants whereas the role of *ZFPM2* variants and of *CEP152* requires further investigation. The enrichment of variants in “syndromic” genes highlights the importance of not excluding these genes when investigating patients with non-syndromic phenotypes. Targeted search for variation in lesion-specific genes within next-generation sequencing data has the potential to reduce time and cost for bioinformatics analysis and increase the specificity and predictive accuracy of the observed associations.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

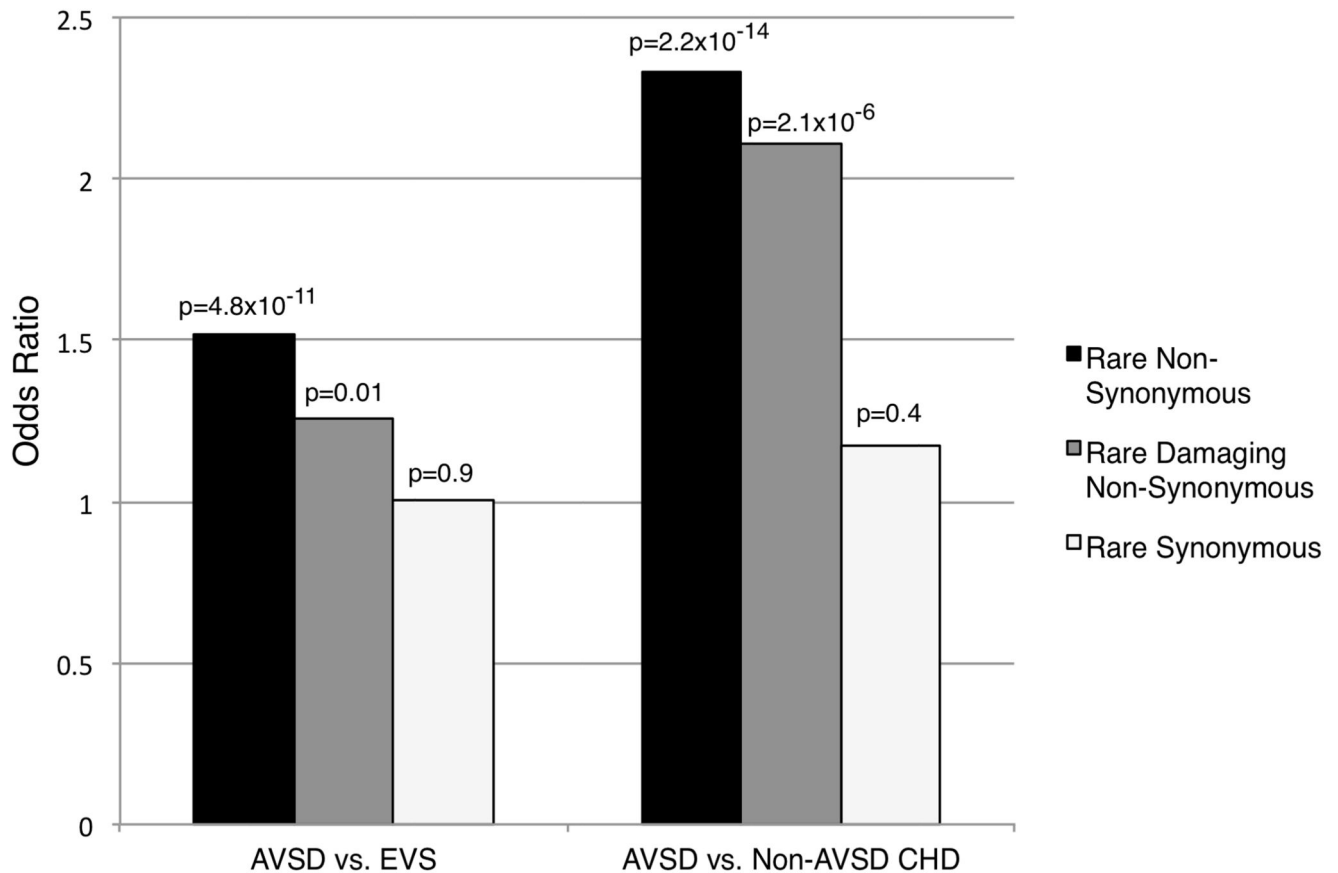
The authors acknowledge the SickKids Labatt Family Heart Centre Biobank Registry for DNA samples; Mina Safi, Tanya Papaz and Carly Ogaki for patient recruitment and follow-up; Elizabeth Uleryk, Director SickKids Hospital Library and Archives, for assistance with systematic literature review; The SickKids Centre for Applied Genomics for targeted resequencing and genotyping; Christian Marshall for annotation of variants with HGMD; Susan Walker and Stephen W. Scherer for providing variant calls in unaffected parents from an Ontario cohort with autism spectrum disorders. The work was supported by funding from the Heart and Stroke Foundation of Ontario Chair in Cardiovascular Science (S.M.), Heart and Stroke Foundation of Canada Research Fellowship (A.K.M.), McLaughlin Centre Accelerator Grant in Genomic Medicine (SM), SickKids Labatt Family Heart Centre Innovation Funds (S.M., L.D., A.K.M., D.M.), the Roma and Marvin Auerback Endowment funds (S.M.), Wellcome Trust (grant number WT098051) (M.H., S.A.), British Heart Foundation (grant numbers CH/09/003/26631 and RG/10/17/28553) (S.B.).

## References

1. Ferencz, C., Loffredo, CA., Correa-Villasenor, A., Wilson, PD. Atrioventricular Septal Defects With and Without Down Syndrome. Genetic & Environmental Risk Factors of Major Cardiovascular Malformations: the Baltimore-Washington Infant Study 1981-1989. Ferencz, C., Loffredo, CA., Correa-Villasenor, A., Wilson, PD., editors. Vol. 1997. Armonk: Futura Publishing Co., Inc.; 1997. p. 103-122.
2. Digilio MC, Marino B, Toscana A, et al. Atrioventricular canal defect without Down syndrome: a heterogeneous malformation. *Am J Med Genet.* 1999; 85:140–146. [PubMed: 10406667]
3. Robinson SW, Morris CD, Goldmuntz E, et al. Missense mutations in CRELD1 are associated with cardiac atrioventricular septal defects. *Am J Hum Genet.* 2003; 72:1047–1052. [PubMed: 12632326]
4. Sarkozy A, Lepri F, Marino B, et al. Additional evidence that PTPN11 mutations play only a minor role in the pathogenesis of non-syndromic atrioventricular canal defect. *Am J Med Genet A.* 2006; 140A:1970–1972.
5. Smith KA, Joziassse IC, Chocron S, et al. Dominant-negative ALK2 allele associates with congenital heart defects. *Circulation.* 2009; 119:3062–3069. [PubMed: 19506109]
6. Priest JR, Girirajan S, Vu TH, et al. Rare copy number variants in isolated sporadic and syndromic atrioventricular septal defects. *J Med Genet A.* 2012; 158A(6):1279–84.
7. Al Turki S, Manickaraj AK, Mercer CL, et al. Rare Variants in NR2F2 Cause Congenital Heart Defects in Humans. *Am J Hum Genet.* 2014; 94:574–585. [PubMed: 24702954]
8. Papaz T, Safi M, Manickaraj AK, et al. Factors influencing participation in a population-based biorepository for childhood heart disease. *Pediatrics.* 2012; 130(5):e1198–205. [PubMed: 23045559]
9. Fung A, Manlhiot C, Naik S, et al. Impact of prenatal risk factors on congenital heart disease in the current era. *J Am Heart Assoc.* 2013; 2:e000064. [PubMed: 23727699]
10. Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP). Seattle, WA: Available at: <http://evs.gs.washington.edu/EVS/> v.2013Jun3 [Accessed 2013 Aug 22]
11. De Pristo M, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43:491–498. [PubMed: 21478889]
12. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38:e164–e164. [PubMed: 20601685]
13. The 1000 Genomes Project Consortium. A map of human genome variation from population scale sequencing. *Nature.* 2010; 467:1016–1073. [v.1000g2012Feb].
14. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* 2001; 29:308–311. [Build 138 Phase I]. [PubMed: 11125122]
15. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols.* 2009; 4:1073–1081. [PubMed: 19561590]
16. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Meth.* 2010; 7:248–249.

17. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Research*. 2010; 20:110–121. [PubMed: 19858363]
18. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research*. 2005; 15:1034–1050. [PubMed: 16024819]
19. Yuen RKC, Thirubahindrapuram B, Merico D, Walker S, Tammimies K, Hoang N, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nature Medicine*. 2015 [Advance online publication].
20. Karolchik D, Hinrichs AS, Furey TS, et al. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res*. 2004; 32:D493–D496. [PubMed: 14681465]
21. Manase, D., D'Alessandro, LCA., Manickaraj, AK., Mital, S. High throughput exome coverage and capture of clinically relevant cardiac genes. Presented at the Annual Meeting of the American Society of Human Genetics; October 2013; Boston, MA, USA.
22. Park C, Lavine K, Mishina Y, et al. Bone morphogenetic protein receptor 1A signaling is dispensable for hematopoietic development but essential for vessel and atrioventricular endocardial cushion formation. *Development*. 2006; 133:3473–3484. [PubMed: 16887829]
23. Kaneko K, Li X, Zhang X, Lamberti JJ, Jamieson SW, Thistlethwaite PA. Endothelial expression of bone morphogenetic protein receptor type 1a is required for atrioventricular valve formation. *Ann Thorac Surg*. 2008; 85:2090–2098. [PubMed: 18498827]
24. Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature*. 2014; 498:220–223.
25. Gillis LA, McCallum J, Kaur M, et al. NIPBL Mutational Analysis in 120 Individuals with Cornelia de Lange Syndrome and Evaluation of Genotype-Phenotype Correlations. *Am J Hum Genet*. 2004; 75:610–623. [PubMed: 15318302]
26. Selicorni A, Colli AM, Passarini A, et al. Analysis of congenital heart defects in 87 consecutive patients with Brachmann-de Lange syndrome. *Am J Med Genet A*. 2009; 149A:1268–1272. [PubMed: 19449412]
27. Chatfield KC, Schrier SA, Li J, et al. Congenital heart disease in Cornelia de Lange syndrome: Phenotype and genotype analysis. *Am J Med Genet A*. 2012; 158A:2499–2505. [PubMed: 22965847]
28. Deardorff MA, Clark DM, Krantz ID. Cornelia de Lange Syndrome. *GeneReviews*. 2005 (Last Update: October 27, 2011). Available at: [www.ncbi.nlm.nih.gov/books/NBK1104/](http://www.ncbi.nlm.nih.gov/books/NBK1104/).
29. Corsten-Janssen N, Kerstjens-Frederikse WS, duMarchie Sarvassd GJ, et al. The Cardiac Phenotype in Patients with a CHD7 Mutation. *Circ Cardiovasc Genet*. 2013; 6:248–254. [PubMed: 23677905]
30. Lalani SR, Hefner MA, Belmont JW, Davenport SLH. CHARGE Syndrome. *GeneReviews*. 2006 (Last Update: February 2, 2012). Available at: [www.ncbi.nlm.nih.gov/books/NBK1117/](http://www.ncbi.nlm.nih.gov/books/NBK1117/).
31. Kalay E, Yigit G, Aslan Y, Brown KE, Pohl E. CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nature*. 2011; 43:23–26.
32. Ucar B, Kilic Z, Dinleyici EC, Yakut A, Dogruel N. Seckel syndrome associated with atrioventricular canal defect: a case report. *Clin Dysmorphol*. 2004; 13:53–55. [PubMed: 15127771]
33. Rappen U, von Brenndorff AI. [Cardiac symptoms in 2 patients with Seckel syndrome]. *Monatsschr Kinderheilkd*. 1993; 141:584–586. [PubMed: 8413337]
34. Breckpot J, Tranchevent L-C, Thienpont B, et al. BMPR1A is a candidate gene for congenital heart defects associated with the recurrent 10q22q23 deletion syndrome. *Eur J Med Genet*. 2012; 55:12–16. [PubMed: 22067610]
35. Svensson EC, Tufts RL, Polk CE, Leiden JM. Molecular cloning of FOG-2: a modulator of transcription factor GATA-4 in cardiomyocytes. *Proc Natl Acad Sci USA*. 1999; 96:956–961. [PubMed: 9927675]
36. Tevosian SG, Deconinck AE, Tanaka M, Schinke M. FOG-2, a Cofactor for GATA Transcription Factors, Is Essential for Heart Morphogenesis and Development of Coronary Vessels from Epicardium. *Cell*. 2000; 101:729–739. [PubMed: 10892744]

37. Katz SG, Williams A, Yang J, et al. Endothelial lineage-mediated loss of the GATA cofactor Friend of GATA 1 impairs cardiac development. *Proc Natl Acad Sci USA*. 2003; 100:14030–14035. [PubMed: 14614148]
38. Huang X, Niu W, Zhang Z, et al. Identification of novel significant variants of ZFPM2/FOG2 in non-syndromic Tetralogy of Fallot and double outlet right ventricle in a Chinese Han population. *Mol Biol Rep*. 2014; 41:2671–2677. [PubMed: 24469719]
39. De Luca A, Sarkozy A, Ferese R, et al. New mutations in ZFPM2/FOG2 gene in tetralogy of Fallot and double outlet right ventricle. *Clin Genet*. 2010; 80:184–190. [PubMed: 20807224]
40. Pizzuti A, Sarkozy A, Newton AL, et al. Mutations of ZFPM2/FOG2 gene in sporadic cases of tetralogy of Fallot. *Hum Mutat*. 2003; 22:372–377. [PubMed: 14517948]
41. Zhang Q, He X, Chen L, et al. Synergistic regulation of p53 by Mdm2 and Mdm4 is critical in cardiac endocardial cushion morphogenesis during heart development. *J Pathol*. 2012; 228:416–428. [PubMed: 22821713]



**Figure 1. Mutation burden analysis across 112 genes with biological relevance to AVSD.**

All rare, non-synonymous variants in the gene were collapsed for each individual to allow a patient level analysis. Odds ratio (OR) for the number of individuals with rare non-synonymous (blue), rare/damaging non-synonymous (red) and synonymous (green) variants across 112 genes in the AVSD cohort compared to EVS controls and non-AVSD CHD cohort with p values are shown. The AVSD cohort showed enrichment of rare non-synonymous variants (OR 1.5,  $p = 4.7 \times 10^{-11}$ ) and rare damaging non-synonymous variants (OR 1.3,  $p = 0.01$ ) compared to EVS cohort. The AVSD cohort showed a similar enrichment of rare non-synonymous variants (OR 2.3,  $p = 2.2 \times 10^{-14}$ ) and rare damaging non-synonymous variants (OR 2.1,  $p = 2.1 \times 10^{-6}$ ) compared to the non-AVSD CHD cohort. There was no difference between the cohorts in rare synonymous variants (AVSD vs. EVS, OR 1.0,  $p = 0.9$ ; and AVSD vs. non-AVSD CHD, OR = 1.2  $p = 0.2$ ).

**Table 1**  
**Primary Cohort Characteristics (n = 81 unrelated probands)**

Male (n) (%)	43 (53.1%)	
Mean Age at Screening (years)	14.3 ± 13	
<b>ATRIOVENTRICULAR SEPTAL DEFECT TYPE</b>	<b>n</b>	<b>%</b>
Partial	34	42.0
Complete	23	28.4
Unbalanced	11	13.6
Intermediate	10	12.3
Unknown	3	3.7
<b>OTHER ASSOCIATED CARDIAC LESIONS</b>	<b>n</b>	<b>%</b>
None	49	60.5
Left Sided Lesion	18	22.2
Arterial or Venous Anomaly	6	7.4
Right Sided Lesion	4	4.9
Conotruncal	3	3.7
Unknown	1	1.2
<b>EXTRACARDIAC ANOMALIES</b>	<b>18</b>	<b>22.2</b>
<b>FAMILY HISTORY OF CONGENITAL HEART DISEASE</b>	<b>10</b>	<b>12.3</b>



Table 2

## Gene Prioritization by Mutation Burden Analysis

Mutation burden analysis was conducted for 112 genes with biological relevance to AVSD. The top 6 ranked genes are shown.

GENE	CASES n = 81		CONTROLS n = 4300*		Odds Ratio	95% Confidence Interval	Fisher's Exact p-value (two-tailed)	Fisher's Exact p-value with nucleotide-by- nucleotide coverage correction <sup>†</sup>
	With Mutation	Without Mutation	With Mutation	Without Mutation				
<i>NIPBL</i>	Rare	6	75	4198	3.3	1.1 – 7.8	0.02	0.01
	Rare Damaging	3	78	27	6.1	1.2 – 20.4	0.008	0.02
<i>CHD7</i>	Rare	9	72	3940	2.3	1.0 – 4.7	0.04	0.07
	Rare Damaging	3	78	58	2.7	0.5 – 8.6	0.02	0.3
<i>CEP152</i>	Rare	8	73	3925	2.4	1.0 – 5.1	0.03	0.03
	Rare Damaging	4	77	70	3.0	0.7 – 8.3	0.08	0.05
<i>BMPRIa</i>	Rare	3	78	4269	5.3	1.0 – 17.5	0.02	0.02
	Rare Damaging	3	78	27	6.1	1.2 – 20.4	0.008	0.02
<i>ZFPM2</i>	Rare	11	70	3872	2.2	1.0 – 4.2	0.03	0.03
	Rare Damaging	8	73	189	2.3	0.9 – 4.9	0.05	0.05
<i>MDM4</i>	Rare	6	75	4232	5.0	1.7 – 11.9	3.2 x 10 <sup>-4</sup>	0.002
	Rare Damaging	6	75	53	6.4	2.2 – 15.2	1.8 x 10 <sup>-5</sup>	6.8 x 10 <sup>-4</sup>

\* The EVS does not provide sample level information, therefore for consistency the burden analysis assumes that each genotype serves as an independent sample. The total number of patients for the EVS cohort per gene was the median of the total number at each position with a rare variant in that gene.

<sup>†</sup> Refer to Table 3 for variant level data

**Table 3**  
**Rare Non-Synonymous Variants in Prioritized AVSD Genes**

Gene	Genomic Position	Base Change	Amino Acid Change	Rare or Novel*	Predicted Damaging?*	Conserved?*	Proband ID
NIPBL	5:36958288	A>G	N105D	Rare	Yes	Yes	AVSD_88
NIPBL	5:36962301	G>A	A179T	Rare	No	Yes	AVSD_65
NIPBL	5:36976188	T>G	N393K	Novel	No	Yes	AVSD_15 AVSD_72
NIPBL	5:37006555	A>G	M1318V	Novel	Yes	Yes	AVSD_33
NIPBL	5:37058993	T>A	S2471T	Novel	Yes	Yes	AVSD_25
CHD7	8:61654268	A>G	T93A	Novel	No	Yes	AVSD_15
CHD7	8:61655009	A>G	M340V	Rare	No	Yes	AVSD_13 AVSD_50 AVSD_66 AVSD_71
CHD7	8:61748826	T>C	Y1325H	Rare	Yes	Yes	AVSD_79
CHD7	8:61765598	C>T	A2105V	Novel	Yes	Yes	AVSD_26
CHD7	8:61768745 <sup>+</sup>	C>G	S2383C	Novel	Yes	Yes	AVSD_49
CHD7	8:61769418	A>C	M2527L	Rare	No	Yes	AVSD_73
CEP152	15:49030841	T>C	T1524A	Rare	No	No	AVSD_19
CEP152	15:49048132	G>C	L1105V	Rare	No	No	AVSD_20
CEP152	15:49048567	A>G	W960R	Rare	Yes	Yes	AVSD_63 AVSD_64 AVSD_72
CEP152	15:49064725	G>T	L581I	Novel	Yes	No	AVSD_41
CEP152	15:49076311	T>C	I394V	Rare	No	No	AVSD_15
CEP152	15:49089864	A>C	S85R	Rare	No	No	AVSD_38
BMPR1a	10:88681396	A>T	D429V	Novel	Yes	Yes	AVSD_17
BMPR1a	10:88683223	G>A	R478H	Rare	Yes	Yes	AVSD_57
BMPR1a	10:88683231	C>T	P481S	Novel	Yes	Yes	AVSD_2
ZFPM2	8:106431420	A>G	E30G	Rare	Yes	Yes	AVSD_53 AVSD_74
ZFPM2	8:106456600	G>A	D98N	Rare	Yes	Yes	AVSD_25 AVSD_10 AVSD_24
ZFPM2	8:106813787	C>T	P361S	Rare	Yes	Yes	AVSD_53
ZFPM2	8:106813942	G>A	M544I	Rare	No	Yes	AVSD_38 AVSD_46
ZFPM2	8:106814597	G>A	V631I	Rare	Yes	Yes	AVSD_74 AVSD_50
ZFPM2	8:106815359	G>A	G885S	Rare	No	Yes	AVSD_45
MDM4	1:204518457	A>C	K324Q	Rare	Yes	Yes	AVSD_10 AVSD_39 AVSD_80 AVSD_85 AVSD_87
MDM4	1:204518499	C>G	P338A	Rare	Yes	Yes	AVSD_29

\* Genomic position for human genome assembly 37/build 105. \*\* Refer to the Supplementary Table 4 for additional detail + With the exception of this variant, all variants in Table 3 were covered to a depth of  $\geq 30$  in the AVSD and EVS cohorts (nucleotide-by-nucleotide coverage assessment)

**Table 4**  
**Clinical characteristics of AVSD Probands with Rare Non-Synonymous Variants in**  
**Prioritized AVSD Genes**

Proband ID	Sex	AVSD Type	Other Cardiac Lesions	Extracardiac anomaly	Gene	Variant	Validated?	Transmitted?
AVSD_15*	M	Partial	Double outlet LAVV	None	NIPBL	5:26976188 T>G	o	Paternal
AVSD_72*	M	Complete	BAV (partial fusion)	migraines, extra help in reading/math	NIPBL	5:26976188 T>G	Yes	Unknown
AVSD_88	F	Partial	None	None	NIPBL	5:36958288 A>G	o	Unknown
AVSD_65	F	Partial	None	None	NIPBL	5:36962301 G>A	Yes	Unknown
AVSD_33	M	Unbalanced	DORV, PS	None	NIPBL	5:37006555 A>G	Yes	Unknown
AVSD_25*	M	Intermediate	None	None	NIPBL	5:37058993 T>A	Yes	Unknown
AVSD_15*	M	Partial	Double outlet LAVV	None	CHD7	8:61654268 A>G	o	Maternal
AVSD_13	F	Complete	CoA, hypoplastic arch, PDA	LD	CHD7	8:61655009 A>G	Yes	Unknown
AVSD_50*#	M	Complete	None	None	CHD7	8:61655009 A>G	Yes	Maternal †
AVSD_66#	F	Partial	None	None	CHD7	8:61655009 A>G	Yes	Unknown
AVSD_71	M	Partial	None	hernia	CHD7	8:61655009 A>G	Yes	Unknown
AVSD_79	F	Intermediate	RSCA, vertebral from AA	None	CHD7	8:61748826 T>C	Yes	Paternal
AVSD_26	F	Partial	None	None	CHD7	8:61765598 C>T	Yes	Maternal
AVSD_49	M	Intermediate	None	None	CHD7	8:61768745 C>G	Yes	Not Paternal
AVSD_73	M	Unbalanced	HLV & AA, CoA, LPV stenosis	sagittal synostosis, hypospadias, cryptorchidism, DD, IUGF, nephrocalcinosis	CHD7	8:61769418 A>C	Yes	Unknown
AVSD_19	F	Complete	Secundum ASD	LD	CEP152	15:49030841 T>C	Yes	Paternal
AVSD_20	M	Partial	None	None	CEP152	15:49048132 G>C	Yes	Unknown
AVSD_63	F	Intermediate	Secundum ASD, multiple VSD	None	CEP152	15:49048567 A>G	o	Unknown
AVSD_64	M	Complete	None	Hirschsprung	CEP152	15:49048567 A>G	Yes	Maternal
AVSD_72*	M	Complete	BAV (partial fusion)	migraines, extra help in reading/math	CEP152	15:49048567 A>G	Yes	Unknown
AVSD_41	M	Complete	None	None	CEP152	15:49064725 G>T	Yes	Unknown
AVSD_15*	M	Partial	Double outlet LAVV	None	CEP152	15:49076311 T>C	o	Paternal
AVSD_38*#	M	Partial	None	None	CEP152	15:49089864 A>C	Yes	Unknown
AVSD_17	F	Unbalanced	CoA, LSVC to CS	None	BMPR1a	10:88681396 G>T	Yes	Maternal
AVSD_57#	M	Complete	multiple VSD, LSVC to CS	LD, psychiatric, cervical spine anomalies	BMPR1a	10:88683223 G>A	Yes	Unknown
AVSD_2	F	Complete	None	None	BMPR1a	10:88683231 C>T	Yes	Unknown
AVSD_53*	F	Complete	PA/MAPCAS, LSVC to CS	Bilateral coloboma, bicornuate uterus, Bockdalek diaphragmatic hernia, midline spleen, hydrocephalus	ZFPM2	8:106431420 C>G	Yes	Unknown

Proband ID	Sex	AVSD Type	Other Cardiac Lesions	Extracardiac anomaly	Gene	Variant	Validated?	Transmitted?
AVSD_74*#	F	Partial	None	None	ZFPM2	8:106431420 C>G	Yes	Unknown
AVSD_10*	M	Partial	PDA	None	ZFPM2	8:106456600 G>A	Yes	Unknown
AVSD_24	M	Complete	None	None	ZFPM2	8:106456600 G>A	Yes	Paternal
AVSD_25*	M	Intermediate	None	None	ZFPM2	8:106456600 G>A	Yes	Unknown
AVSD_53*	F	Complete	PA/MAPCAS, LSVC to CS	Bilateral coloboma, bicornuate uterus, Bockdalek diaphragmatic hernia, midline spleen, hydrocephalus	ZFPM2	8:106813787 C>T	Yes	Unknown
AVSD_38*#	M	Partial	None	None	ZFPM2	8:106813942 G>A	Yes	Unknown
AVSD_46	M	Complete	None	None	ZFPM2	8:106813942 G>A	Yes	Unknown
AVSD_50*#	M	Complete	None	None	ZFPM2	8:106814597 G>A	Yes	Paternal
AVSD_74*#	F	Partial	None	None	ZFPM2	8:106814597 G>A	Yes	Unknown
AVSD_45	F	Partial	PDA	Congenital rubella syndrome, epilepsy, hearing impairment, blindness, psychiatric disorder	ZFPM2	8:106815359 G>A	Yes	Unknown
AVSD_10*	M	Partial	PDA	None	MDM4	1:204518457 A>C	Yes	Unknown
AVSD_39	F	Complete	LVOTO	None	MDM4	1:204518457 A>C	Yes	Paternal
AVSD_80	M	Unknown	LVOTO	None	MDM4	1:204518457 A>C	Yes	Unknown
AVSD_85	M	Complete	LSVC to CS, RAA	None	MDM4	1:204518457 A>C	Yes	Unknown
AVSD_87	F	Partial	PDA	None	MDM4	1:204518457 A>C	Yes	Unknown
AVSD_29	F	Partial	PDA	None	MDM4	1:204518499 C>G	Yes	Paternal

Genomic position for human genome assembly 37/build 105. All variants are heterozygous. \* probands with > 1 variant, o sample/validation not available, +Inherited from mother also affected with AVSD, # of CHD family history

Abbreviations: M – male, F – female; AA - aortic arch, ASD - atrial septal defect, BAV - bicuspid aortic valve, CoA - coarctation of the aorta, DORV - double outlet right ventricle, HLV - hypoplastic left center ventricle, LAVV - left center atrioventricular valve, LPV - left center pulmonary vein, LSVC to CS - left center superior vena cava to the coronary sinus, LVOTO - left center ventricular outflow tract obstruction, PA/ MAPCAS - pulmonary atresia with major aortopulmonary collaterals, PDA - patent ductus arteriosus, RAA - right aortic arch, RSCA - right subclavian artery, VSD - ventricular septal defect