Whole Exome Sequencing Uncovers the Genetic Complexity of Bicuspid Aortic Valve in Families with Early Onset Complications

Sara Mansoorshahi[1], Anji T Yetman[2], Malenka M Bissell[3], Yuli Y Kim[4], Hector Michelena[5], Dawn S Hui[6], Anthony Caffarelli[7], Maria G Andreassi[8], Ilenia Foffa[8], Dongchuan Guo[1], Rodolfo Citro[9], Margot De Marco[10], Justin T Tretter[11], Shaine A Morris[12], Simon C Body[13], Jessica X Chong[14], Michael J Bamshad[14], University of Washington Center for Rare Disease Research, BAVCon Investigators, EBAV Investigators, Dianna M Milewicz[1], Siddharth K Prakash[1]

[1]Department of Internal Medicine, University of Texas Health Science Center at Houston, Houston, Texas.

[2]Children's Hospital and Medical Center, University of Nebraska, Omaha, Nebraska.

[3]Department of Biomedical Imaging Science, Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, United Kingdom.

[4]Division of Cardiovascular Medicine, The Hospital of the University of Pennsylvania, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.

[5]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota.

[6]Department of Cardiothoracic Surgery, University of Texas Health Science Center San Antonio, Texas.

[7]Department of Cardiothoracic Surgery, Stanford University School of Medicine, Stanford, California.

[8]Consiglio Nazionale delle Richerche (CNR), Instituto di Fisiologia Clinica, Pisa, Italy.

[9]Cardio-Thoracic and Vascular Department, University Hospital "San Giovanni di Dio e Ruggi d'Aragona," Salerno, Italy.

[10]Department of Medicine, Surgery and Dentistry Schola Medica Salernitana, University of Salerno, Baronissi, Italy.

[11]Cleveland Clinic, Cleveland, Ohio.

[12]Department of Pediatrics, Division of Pediatric Cardiology, Baylor College of Medicine and Texas Children's Hospital, Houston, Texas

[13]Department of Anesthesiology, Boston University School of Medicine, Boston, Massachusetts.

[14]Division of Genetic Medicine, Department of Pediatrics, University of Washington, Seattle, WA, USA

Corresponding Author: Siddharth K. Prakash, MD, PhD

Department of Internal Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, 6431 Fannin Street, MSB 6.116, Houston, TX 77030, USA. Email: Siddharth.K.Prakash@uth.tmc.edu.

1

Key Words: Bicuspid Aortic Valve, Whole Exome Sequencing, Thoracic Aortic Aneurysm,

Congenital Heart Disease, Cardiovascular Genetics

Word count: 4704

Abstract

Bicuspid Aortic Valve (BAV) is the most common adult congenital heart lesion with an estimated population prevalence of 1%. We hypothesize that early onset complications of BAV (EBAV) are driven by specific impactful genetic variants. We analyzed whole exome sequences (WES) to identify rare coding variants that contribute to BAV disease in 215 EBAV families. Predicted pathogenic variants of causal genes were present in 111 EBAV families (51% of total), including genes that cause BAV (8%) or heritable thoracic aortic disease (HTAD, 17%). After appropriate filtration, we also identified 93 variants in 26 novel genes that are associated with autosomal dominant congenital heart phenotypes, including recurrent deleterious variation of *FBN2*, *MYH6*, channelopathy genes, and type 1 and 5 collagen genes. These findings confirm our hypothesis that unique rare genetic variants contribute to early onset complications of BAV disease.

Introduction

Congenital heart disease (CHD) affects 8.2 per 1000 live births worldwide (1). Bicuspid Aortic Valve (BAV) is the most common adult CHD with an estimated population prevalence of 0.5-2%. BAV is a major cause of aortic stenosis or thoracic aortic aneurysms and is associated with other left ventricular outflow lesions such as coarctation (2). The robust heritability of BAV was demonstrated in studies showing that primary relatives are ten-fold more likely to have a BAV than matched controls (3). While more than 50% of individuals with BAV undergo aortic valve surgery during their lifetimes, phenotypic expression of BAV varies widely, with only approximately 10% of patients with BAV requiring aortic valve surgery in the first two decades of life (4,5).

We previously found that rare genomic copy number (CNV) variants involving congenital heart disease (CHD) genes are enriched in probands with early onset complications of BAV (EBAV), defined as clinical presentation due to aortic stenosis, aortic regurgitation, a large thoracic aortic aneurysm, or aortic dissection prior to age 30 (6). We hypothesized that this subset of individuals may have an increased burden of pathogenic genetic variants that drive aortic complications related to BAV disease. CNV analysis showed that genetic contributions to EBAV overlap with other congenital syndromes that include BAV such as Velocardiofacial syndrome and 1q21 deletion syndrome. We identified more than 50 large, potentially pathogenic CNVs in 10% of EBAV cases that overlap with CHD genes that can cause BAV that were not present in a cohort of older sporadic BAV cases. These observations also validate our hypothesis that the genetic contribution to EBAV is distinct compared to cases with later onset BAV complications.

4

We also compared the EBAV cohort to a cohort that was independently ascertained due to early onset thoracic aortic aneurysms or acute aortic dissections (ESTAD). More than 50% of ESTAD patients have BAV, but the contribution of known causal BAV genes to TAD is not known. To ascertain the impact of functionally annotated WES variants in the major BAV genes *GATA4*, *NOTCH1*, *SMAD6*, and *ROBO4,* we compared the prevalence of rare variants in these genes between cohorts and to controls without TAD. The burden of rare BAV gene variants was significantly enriched in EBAV cases compared to ESTAD cases, even when matched for the presence of BAV. These findings confirmed our hypothesis that the EBAV cohort is enriched for highly penetrant rare gene variants in known BAV genes, but that these genes do not significantly contribute to BAV in cohorts with TAD. Consistent with our earlier observations, we concluded that young patients who present primarily due to BAV are a genetically distinct subgroup with implications for genetic testing and prognosis. The overall goal of this study was to identify genetic variants that predispose to the EBAV phenotype.

Methods

The study protocol was approved by the Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston. All participants signed a written consent form, and study procedures were conducted in compliance with the ethical standards of the relevant national guidelines on human experimentation (HHS regulations 45 CFR part 46) and with the Helsinki Declaration of 1975, as revised in 2008.

Individuals were included in this study if they presented due to complications of BAV disease prior to age 30. Early onset complications were defined as greater than mild aortic stenosis or aortic regurgitation, a large thoracic aortic aneurysm (Z-score > 4), or aortic

dissection. Individuals with known genetic mutations, genetic syndromes, or complex congenital heart disease were excluded.

The GenTAC cohort consists of whole exome sequencing (WES) data from 290 unrelated probands with BAV from the National Registry of Genetically Triggered Aortic Aneurysms and Other Cardiovascular Conditions (7). The JRRP cohort consists of WES data from 974 unrelated probands from the John Ritter Research Program in Aortic and Vascular Diseases at the University of Texas Health Science Center at Houston (8).

WES of the EBAV cohort was performed at the University of Washington Center for Rare Disease Research (UW-CRDR) using standard methods (see Supplemental Data). Variant filtration was performed using a customized version of seqr at the University of Washington Center for Rare Disease Research (v1.0-4cb7dfc3) with the following settings: de novo/dominant inheritance model; pathogenic, likely pathogenic, or VUS; nonsense, in frame, frameshift, or missense variants; REVEL > 0.5, CADD > 20, AF < 0.0001; pass variants only.

Missense variants were retained if they were predicted to be deleterious or disease-causing by at least 3 of polyphen, muttaster, SIFT, or FATHMM. Novel genes were prioritized if mutated in at least 3 unrelated EBAV probands. We filtered the final gene set to remove genes that cause recessive Mendelian phenotypes. The enrichment of rare variants was tabulated in 2x2 comparisons with selected deleterious rare variants in gnomAD version 2 data. gnomAD alleles were functionally annotated using FAVOR (https://favor.genohub.org).

Results

EBAV Cohort

We analyzed WES data from probands with early onset complication of bicuspid aortic valve disease (EBAV). The EBAV cohort consists of 215 families with 150 singletons, 16 duos, 34 trios, and 15 multiplex families. More than 95% of EBAV probands are from European ancestry populations. The mean age of EBAV probands at presentation was 18 years old, compared to the median age of 52 years in a recent meta-analysis of adult BAV cases (9)

HTAD, BAV, and CHD gene variants in the EBAV Cohort

We first identified rare gene variants known to cause congenital heart disease (CHD) in the EBAV cohort. The list of candidate genes included 29 HTAD genes that are on current clinical sequencing panels, as well as 12 BAV genes and 190 CHD genes that have strong cumulative evidence to cause BAV or related congenital malformations from human or animal model data (Supplemental Data) (10,11).

Missense and loss of function variants predicted to be harmful by our analysis were selected using the criteria described in the Methods.
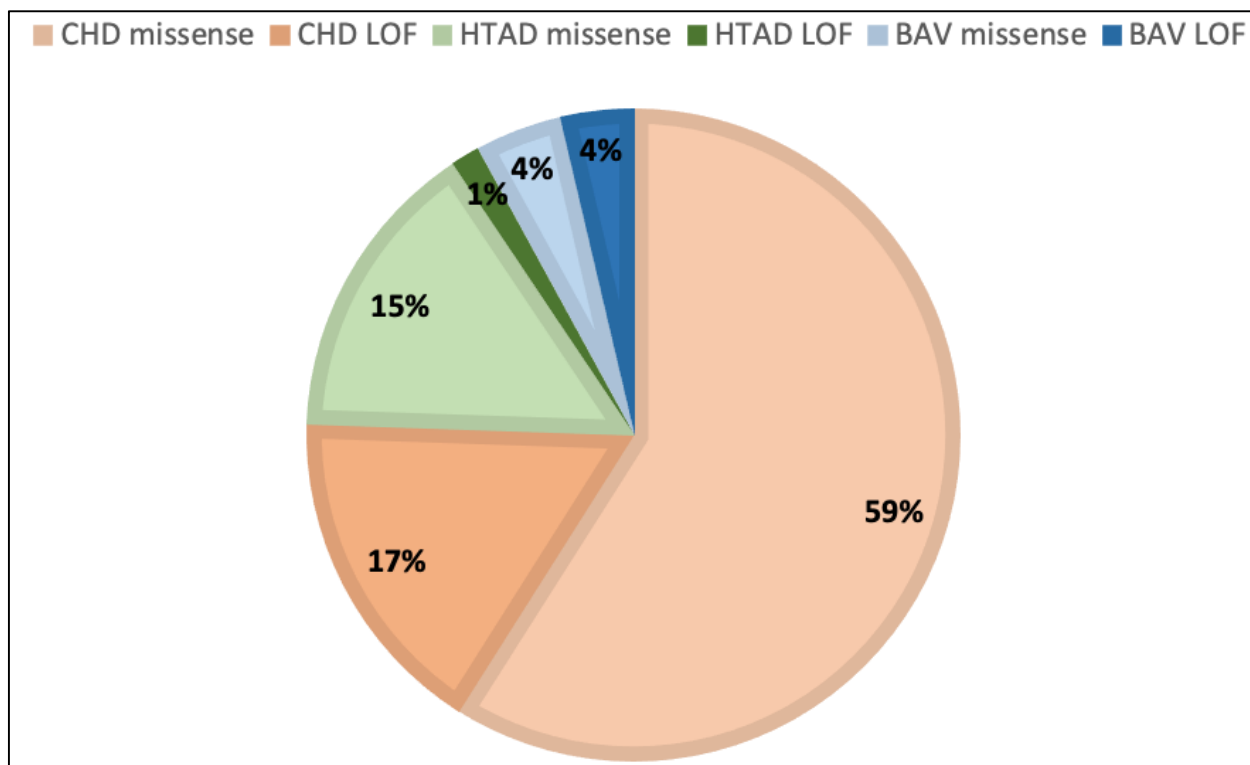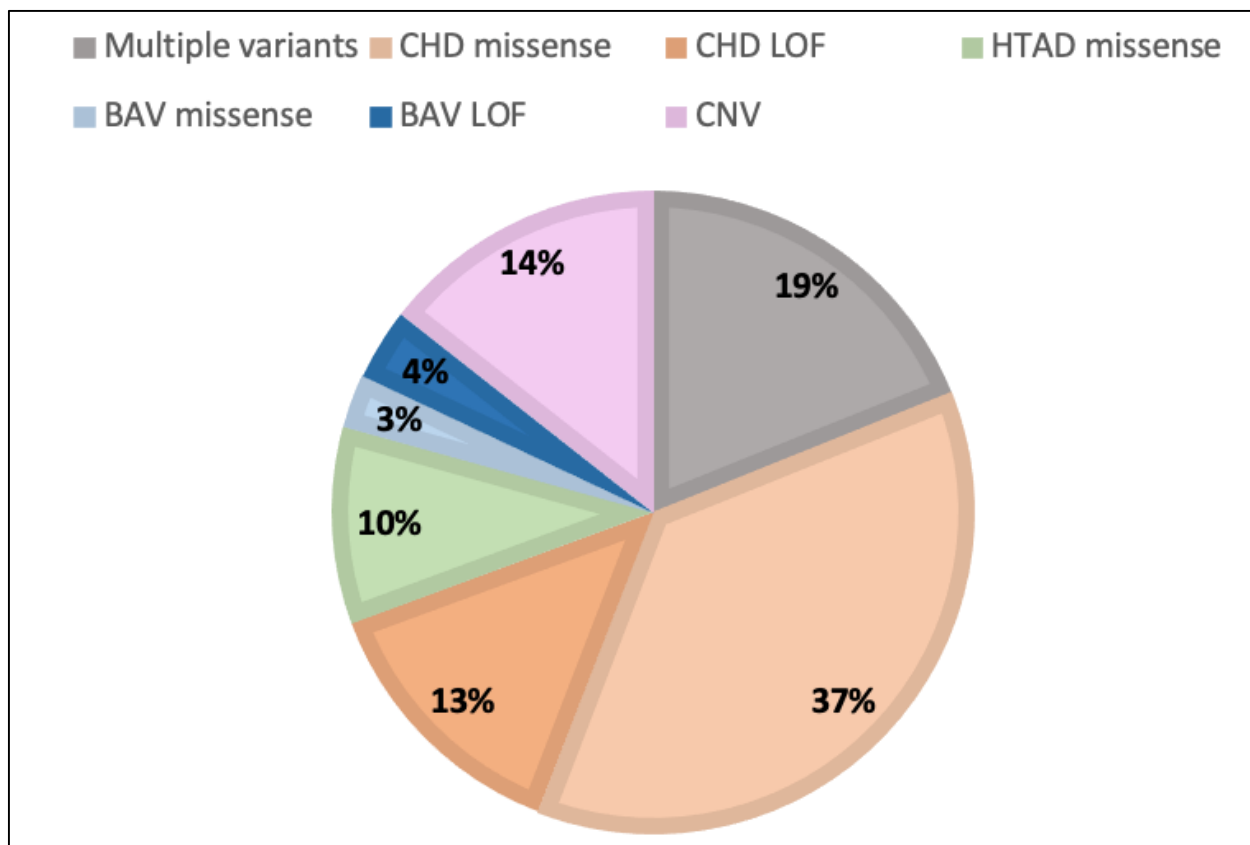
Figure 1: Distribution of predicted deleterious variants in 111 EBAV families and in candidate genes. A. Distribution of variants in EBAV families; **B.** Distribution of variants in candidate genes.

Predicted pathogenic variants of CHD genes were present in 111 (52%) EBAV families. In total, we discovered 138 alleles of 54 genes that cause dominant congenital heart phenotypes, primarily deleterious missense variants (75 probands). Predicted pathogenic variants of dominant CHD genes (including missense and LOF) comprised 77% of candidate variants, including 25 loss of function variants. We identified predicted pathogenic variants of BAV genes in 9 families (8%) and predicted pathogenic variants of HTAD genes in 19 families (17%). A total of 21 EBAV probands (19%) harbored deleterious rare variants in more than one candidate gene. Among all candidate genes, *MYH6* (6 missense, 1 LOF), *COL5A1* (5 missense), and *FBN2* (5 missense) harbored the most deleterious rare variants. The BAV genes with rare variants that passed all filters were *NOTCH1* (2 missense), *MIB1* (2 missense, 1 LOF) and *SMAD6* (2 missense, 2 LOF, 1 insertion). These observations suggest that genetic contributions to EBAV may differ significantly from the genetic landscape of later onset BAV disease.

Novel Candidate Genes in the EBAV Cohort

Next, using the same criteria, we filtered the WES data to identify genes with rare

deleterious variants in at least three unrelated EBAV probands (Figure 2).
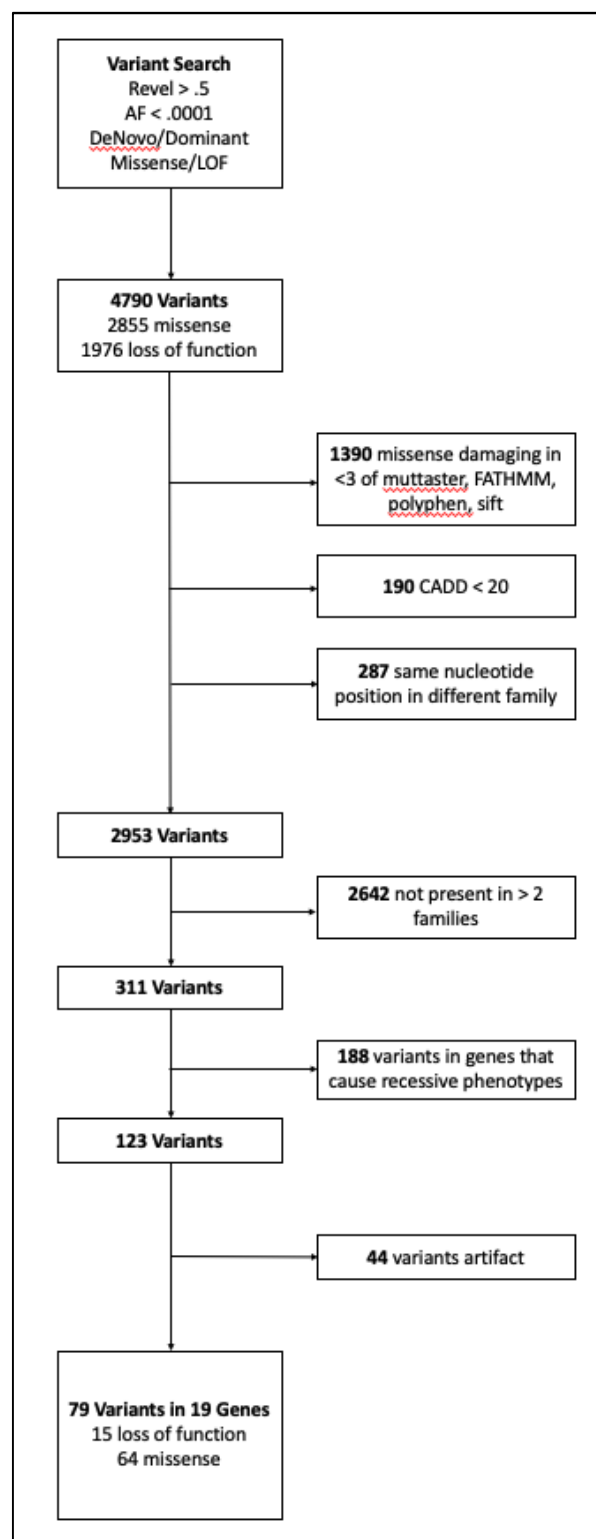
Figure 2: Filtration workflow to select candidate EBAV variants. The initial seqr project variant search generated 4790 candidates. Missense variants that were not rated as deleterious by at least three functional predictors, had normalized CADD scores < 20, and had the same base position in two different families were removed. We prioritized novel genes that had predicted deleterious variants in three or more families. We removed genes that are known to cause recessive or noncardiac phenotypes.

This selection strategy led to 79 variants in 19 genes associated with dominant cardiac phenotypes. The novel genes that are most highly enriched for deleterious rare variants are: *PKP2*, *MYH6*, *KCNH2, JAG1*, *KIF1A*, *TPTE2*, *COL1A1,* and *COL1A2* (Figure 3).
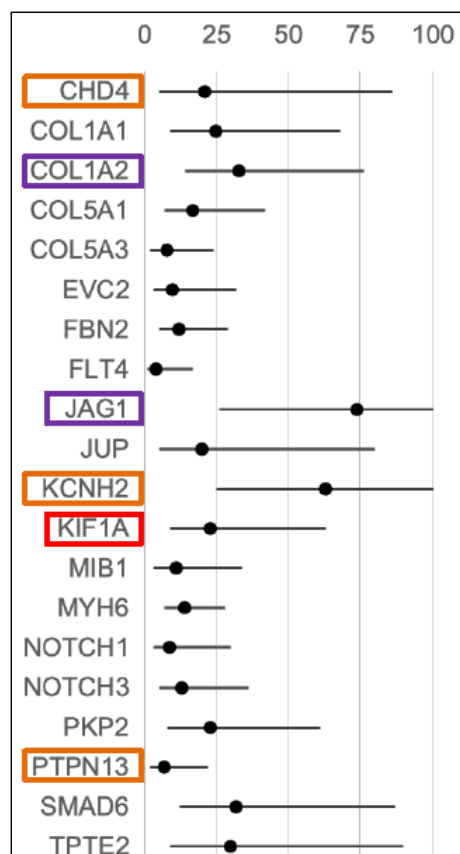
Figure 3. Enrichment of EBAV candidate gene variants. The number of rare deleterious loss of function and missense variants, selected according to the criteria defined in the Methods, were

compared to the number of variants with equivalent functional classifications from European ancestry populations in gnomAD v2. Circles: enrichment vs. gnomAD (0-100X); horizontal lines: 95% confidence intervals; orange boxes: segregates with BAV in more than one EBAV family; purple boxes: also enriched in JRRP and GenTAC datasets; red boxes: at least one *de novo* variant.

Loss of function variants in *PKP2* were observed in three families. One rare variant that causes a stop-gain mutation (c.169C>T, p.Gln57Ter) is not present in gnomAD and has a CADD score of 41. A stop-gain mutation (c.235C>T, p.Arg79Ter, CADD 37, gnomAD frequency 4×10^{-6}) and a frameshift mutation (c.2198_2202del, p.His733ProfsTer8, CADD 29, gnomAD frequency 1.6×10^{-5}) are predicted to be pathogenic by ClinVar and are rare in gnomAD. The PKP2 frameshift variant did not segregate with BAV in one of the three EBAV families.

Six families harbored harmful missense or loss of function variants in *MYH6*. A single splice donor mutation c.1410+1G>A has not been reported in gnomAD and is predicted to be pathogenic by ClinVar. Three of the five missense variants are reported as variants of uncertain significance, and one missense variant has conflicting interpretations of pathogenicity. Notably, four of the 5 missense variants are not in the myosin motor or actin-binding domains (Supplementary Data). All five missense variants are reported to be damaging by MutTaster, FATHMM, and SIFT, with CADD scores > 22 and REVEL > 0.65. In one multiplex family with an *MYH6* variant, the variant is not present in two affected relatives (father and brother) and does not segregate with disease. The other families with an *MYH6* variant do not have additional samples available to determine segregation.

12

Rare deleterious *KCNH2* variants were observed in five families, including one loss of

function variant and four missense variants. The loss of function frameshift variant c.33_34del

(p.Ala13SerfsTer34) is not present in gnomAD and is predicted to be likely pathogenic by

ClinVar. Three of four missense variants in *KCNH2* are not present in gnomAD, and the other

missense variant of uncertain significance is rare in gnomAD (8×10^{-6}). All four missense

variants have CADD > 20 and REVEL > 0.5. Three of the four missense variants are uniformly

predicted to be damaging by Polyphen, SIFT, MutTaster, and FATHMM.

Rare Type 1 or Type 5 collagen gene variants were present in 21 families, primarily

involving *COL1A1* (4), *COL1A2* (6), *COL5A1* (7), *COL5A2* (1) and *COL5A3* (4). All were

missense variants except for a single *COL5A3* frameshift variant (c.2999del,

p.Pro1000GlnfsTer144) that is present in gnomAD with a frequency of 3.18×10^{-5} and has a

CADD score of 27. The missense variants all had gnomAD frequencies $< 1.5 \times 10^{-5}$ and are

predicted to be damaging by Polyphen, SIFT, MutTaster, and FATHMM. In 6 out of 7 families

with available segregation data, the missense variants were observed in apparently unaffected

individuals (Supplemental Data). There were no other available samples to determine the

segregation of the other missense or loss of function *COL1* or *COL5* variants.

Rare deleterious *KIF1A* variants were observed in four families, with two loss of function

variants and two missense variants. The two predicted loss of function variants were a splice

acceptor variant (c.4257-2A>C) and an in-frame deletion (c.718_720del, p.Lys240del). Both

variants are not present in gnomAD and have CADD scores > 22. One missense variant

(c.4222C>G, p.Leu1408Val) is also not present in gnomAD and has variable evidence for

pathogenicity, with a REVEL score of 0.412 and a CADD score of 23.7. The second missense

variant (c.1862C>T, p.Thr621Met) is present in gnomAD with a frequency of 1×10^{-4} and is

13

reported as a variant of uncertain significance by ClinVar, but is predicted to be damaging by

SIFT and MutTaster, with a CADD score of 22 and a REVEL score of 0.504. We determined

that the splice acceptor variant is a *de novo* variant that is not present in two unaffected parents

or in two unaffected siblings.

Microarray analysis also identified a heterozygous subtelomeric terminal deletion of

2q37.3 in 11 EBAV probands with a conserved breakpoint approximately 1 Mb telomeric to

KIF1A (Supplemental Data). Given the distance between the breakpoint and the KIF1A

promoter, we concluded that this recurrent 200 kilobase microdeletion is unlikely to affect

KIF1A expression. No candidate genes with sequence variants that passed filtration are within

the deleted subtelomeric region.

Several other gene variants were observed in at least three EBAV families, but there were

no loss of function variants in these genes. There were *FBN2* missense variants in 5 families, and

FLNC, *NOTCH3*, or *PTPN12* missense variants in 4 families.

Developmental Functions of Candidate EBAV Genes

The GO processes that were most highly enriched in candidate EBAV genes that passed

all filters were circulatory system development, actin filament-based movement, extracellular

structure organization, multicellular organismal signaling, and cardiac development (Figure 4).

The most highly enriched individual genes in these functional categories were *MYH6*, *KIF1A*,

JAG1, MIB1, PKP2, and *SMAD6*, which all had predicted loss of function variants in the EBAV

cohort. These data highlight the pleiotropic roles of cardiac developmental genes in aortic valve

formation and cardiac function.
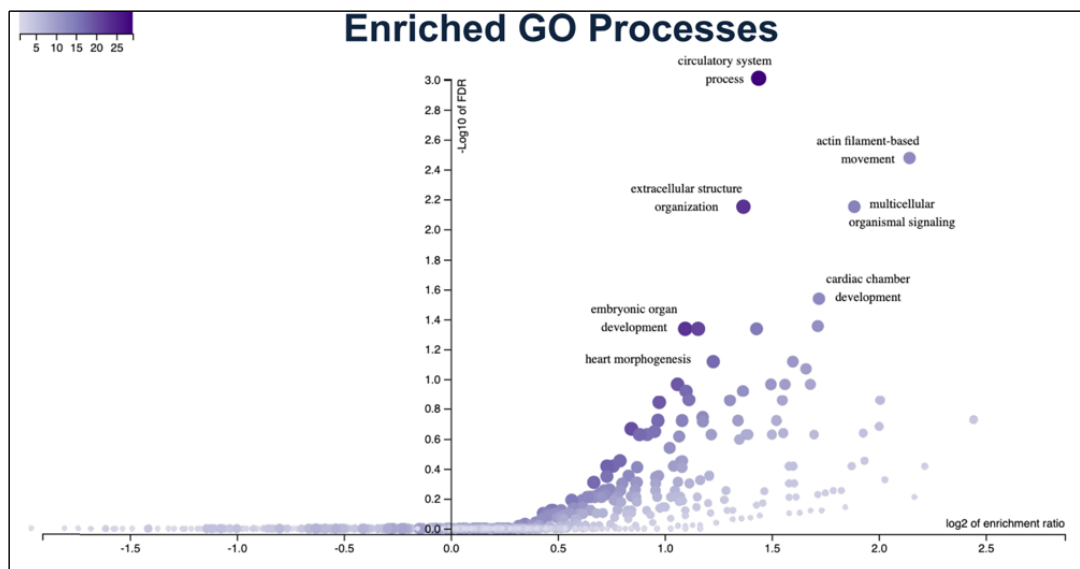
Figure 4. Gene ontology enrichment analysis of EBAV candidate genes. Plot was generated using Over-Representation Analysis of gene ontologies in WebGestalt 2019. The -log10 of false discovery rates (FDR) were plotted against the log2 of enrichment ratios.

CHD Phenotypes of EBAV Cases with Rare Variants

In total, we identified at least one candidate gene mutation in 111 families (51%). Nineteen of these probands (17%) also presented with at least one other congenital lesion besides BAV. The most common congenital lesions were coarctation of the aorta in 12 families (63%) and mitral valve prolapse in two families. Seven probands presented with abnormal electrocardiographic findings (6%), primarily conduction delays ranging from first degree AV block to left bundle branch block. Several EBAV families with coarctation have variants in the same genes. Coarctation was observed in 2 of 5 families with variants in *COL1A1*, 2 of 4 families with variants in *FBN2*, and 2 of 4 families with variants in *MEGF8*. EBAV probands with rare sequence variants in candidate genes were not diagnosed at significantly younger ages

and were not more likely to be diagnosed with thoracic aortic aneurysms or to undergo valve or aortic surgeries.

Replication in TAD Datasets

Rare deleterious variants of several EBAV candidate genes are enriched in the EBAV cohort compared to gnomAD and are also present in two other TAD cohorts, JRRP and GenTAC: *COL1A2*, *COL5A1*, *FBN2*, *MIB1*, *MYH6*, and *NOTCH3* (Table 1). Rare deleterious variants of *JAG1*, *SMAD6*, *PKP2*, *EVC2*, *KCNH2*, and *KIF1A,* and *TPTE2* are also enriched in EBAV probands and are present in one other TAD cohort. Notably, we observed recurrent predicted loss of function variants of *MYH6*, *MIB1*, *JAG1*, *EVC2*, and *TPTE2*. *FLNC* harbored the largest number of rare variants but was not significantly enriched in EBAV probands. These observations suggest that genes that cause BAV may also contribute to TAD and are consistent with the increased prevalence of BAV in cohorts that present due to TAD.

Gene	EBAV (n=215)	GenTAC (n=200)	JRRP (n=974)
NOTCH1	2	0	6
MIB1	3*	3*	6*
JAG1	4	0	4*
PKP2	3*	1	0
MYH6	6*	2*	9
KCNH2	5*	0	1
COL5A3	4*	0	8
KIF1A	4*	0	0
COL1A2	6	2	6
FBN2	5	1	12
FLNC	4	4	14
NOTCH3	4	3	4
SCN10A	3	8	3

SMAD6	5*	2	0
COL1A1	4	2	3
COL5A1	7	2	8

Table 1: Rare variants of candidate genes in EBAV, GenTAC, and JRRP cohorts. *Indicates at least one loss of function allele in that dataset.

Segregation of Candidate Variants

Some candidate genes showed limited evidence for segregation in EBAV families (Supplemental Data). For example, *KCNH2* segregates in two duo families, and *PTPN13* segregates with BAV in one duo family and in one trio family. *CACHA1H*, encoding a voltage gated calcium channel subunit, segregated with BAV in a family with three affected individuals and one unaffected relative. *A de novo* variant of KIF1A was observed in a single family with 5 unaffected individuals (2 parents and 2 siblings) and 1 affected individual. Additionally, a *de novo* variant of *SCN8A* was observed in a single family with 3 unaffected individuals (2 parents and 1 sibling) and 1 affected individual. *CHD4*, *FBN2*, *NOTCH3*, and *JUP* segregated with BAV in single small families.

Other candidate genes showed evidence for reduced penetrance or lack of segregation. Notably, we observed that *SMAD6* missense and insertion alleles did not segregate with BAV in two families, although *SMAD6* loss of function alleles did segregate with BAV. *COL5A1* variants are present in two apparently unaffected individuals in separate families (50% of carriers). Similarly, only 3 of 7 *COL1A2* variant carriers (43%) in two families were diagnosed with BAV. The penetrance of *JAG1* (50%) was also reduced in two families with available data. In three trio families, *MYH6*, *COL5A3* and *PKP2* variants were all observed in one apparently unaffected parent. Incomplete segregation may suggest reduced penetrance of these variants or could imply that additional genetic mutations may be required to cause BAV in some families.

17

Discussion

In this study, we analyzed WES data from 215 families (350 total participants) to identify genetic variants that cause early onset complications of BAV disease. We found that more than one-third of EBAV families have at least one variant in a likely pathogenic gene. In addition to known BAV and HTAD gene mutations, we observed recurrent deleterious variants in *MYH6*, *FBN2*, *KCNH2, PKP2*, type 1 and 5 collagens, and *KIF1A*.

We originally hypothesized that BAV and TAD may be caused by overlapping genes because BAV is enriched in some rare syndromic causes of TAD, and mutations in HTAD genes have been discovered in rare BAV cases with complex valvulo-aortopathy presentations (12,13) However, the results of this study reinforce our previous observations that there is minimal genetic overlap between young individuals who present primarily due to early onset valve complications and those who present due to early onset TAD (14). We did not identify any mutations of HTAD genes that cause Loeys-Dietz syndrome, which is associated with early onset TAD and BAV, and a total of three causal mutations in non-syndromic HTAD genes (a glycine substitution in FBN1 and loss of function mutations in SMAD4 and FOXE3). Correspondingly, we observed almost no thoracic aortic dissections in EBAV families, consistent with the low rates of aortic dissection in many BAV cohorts (15).

MYH6 encodes the alpha heavy-chain subunit of cardiac myosin. A recurrent missense mutation of MYH6 (p. Arg721Trp) was previously identified in 125 probands who had both coarctation of the aorta and BAV (16). In addition, MYH6 variants are implicated in a wide spectrum of congenital heart malformations, including hypoplastic left heart syndrome (HLHS), Shone complex, ostium secundum atrial septal defects, atrial fibrillation, and dilated cardiomyopathy (17–19). Depending on the cohort studied, MYH6 mutations are found in up to

10% of HLHS cases and predict poor clinical outcomes (19–22). BAV frequently co-occurs with HLHS and co-segregates with HLHS in families. BAV and HLHS are also caused by similar chromosomal abnormalities (Turner, Velocardiofacial, and Down syndromes). Therefore, MYH6 mutations most likely cause BAV in the context of other left-sided congenital lesions and may modify the clinical prognosis of BAV patients (19).

SMAD6 encodes an inhibitor of bone morphogenic protein signaling and is highly expressed in cardiac valves and the aortic root (23,24). SMAD6 mutations have been described in 3-9% of individuals with BAV and TAA, including predicted loss of function mutations that are similar to the frameshifts and indel that we identified in 3 EBAV probands (25,26). The SMAD6 mutant phenotype frequently includes other left ventricular obstructive lesions, such as coarctation and HLHS, as well as conotruncal defects, and these variable phenotypes are recapitulated in mice with knockout of the SMAD6 analog Madh6 (23,26–28). There is some evidence that SMAD6 functions downstream of NOTCH1 signaling to regulate endothelial cell motility and adhesion (28). Notably, other candidate genes (JAG1, MIB1) also participate in this pathway. Loss of this function during cardiac development may predispose to BAV (28).

Mutations of fibrillin and collagen genes can cause overlap syndromes with features of Ehlers-Danlos syndrome (EDS) hypermobility, cardiac valvulopathies, or TAD. We identified recurrent rare and predicted pathogenic variants of FBN2 and type 1 and type 5 collagens in the EBAV cohort. COL1A1 or COL1A2 mutations are the major cause of osteogenesis imperfecta (OI) with variable aortic and mitral valve disease and features of EDS (29,30). Up to half of COL1 mutation carriers develop cardiac complications, primarily aortic regurgitation due to a dysplastic or bicuspid aortic valve (31). COL1 mutations also predispose to sudden death due to TAD or intracranial aneurysms and dissections (32). Type 1 and Type 5 collagens are highly

expressed in the developing aortic valve and in the adventitia of blood vessels (33). Type 5

collagens form heterotypic fibrils with type 1 collagens and are required for the assembly of type

1 collagen fibers (33,34). Approximately 50% of classical EDS cases are caused by mutations of

COL5A1 or COL5A2 (35,36). COL5A1 has been implicated as the causal gene in several

families with EDS features who also present due to arterial dissections (32,37–39). COL5A2

mutations and copy number variants are enriched in cohorts with cervical arterial dissections and

may also predispose to TAD (40–43). Notably, low penetrance BAV has been described as a

feature of classical EDS, and rare mutations of COL5A1 and COL5A2 have been observed in

other BAV cohorts (44–46). Mutations of FBN2 encoding fibrillin 2, which cause congenital

contractural arachnodactyly with incompletely penetrant TAD, were also identified as a potential

modifier for BAV disease based on low penetrance variants in a younger subgroup of cases with

more aggressive complications (45). We hypothesize that deleterious variants of extracellular

matrix components such as collagens or fibrillins probably do not cause BAV but may accelerate

the onset and severity of BAV disease, explaining their enrichment in EBAV cases.

Mutations of KCNH2 cause several types of cardiac conduction abnormalities

predisposing to arrhythmias and sudden cardiac death, including Long QT syndrome, Romano-

Ward syndrome, Short QT syndrome, Brugada syndrome, and familial atrial fibrillation (47).

KCNH2 is one of several genes with recurrent missense and loss of function mutations in the

EBAV cohort that encode components of cardiac myocyte ion channels, along with HCN4,

KCNQ1, SCN5A, SCN8A, and SCN10A. KCNQ1, SCN5A, SCN8A, and SCN10A mutations

can also cause sudden cardiac death by disrupting potassium or sodium channel activities (48–

50). HCN4 mutations are implicated in familial sinus bradycardia, left ventricular

noncompaction cardiomyopathy (LVNC), and TAD (51–53). Intriguingly, LVNC was also

identified in families with KCNH2, KCNQ1 and SCN5A mutations (52–55). The mechanistic connection between conduction abnormalities and LVNC in channelopathy mutants is not currently known, but abnormal cardiac and vascular smooth muscle contraction during heart morphogenesis may contribute to impaired ventricular remodeling (56). The prevalence of LVNC is generally increased in BAV cohorts, and mutation of KCNJ2, another potassium channel subunit, causes a syndromic presentation of BAV with extracardiac features in addition to arrhythmias (57). BAV and TAD also co-segregate with LVNC in some families with mutations in NKX2-5, TBX20 or MIB1, another Notch pathway regulator that was convincingly linked to BAV in recent studies and is recurrently mutated in the EBAV cohort (52,58,59). Our findings provide further evidence for a genetic link between cardiomyopathies, BAV, conduction abnormalities, and TAD that may contribute to early onset complications of BAV disease.

Heterozygous truncating mutations of PKP2 are a common cause of arrhythmogenic right ventricular cardiomyopathy (ARVC) but have also been linked to LVNC in rare cases (52,60,61). Two other genes with loss of function variants in EBAV probands, JUP and MYH7, are also moderately or definitively linked to LVNC (52,53). PKP2 and JUP encode components of the cardiac desmosome, but more broadly anchor an intercalated disk scaffold called the connexome, which regulates intracellular signaling in cardiomyocytes (62). Disruption of the connexome alters the activity of sodium channels including SCN5A and contributes to arrhythmias in ARVC patients with PKP2 or JUP mutations (61,62). Gene therapies that restore myocardial expression of JUP or PKP2 can rescue structural heart and conduction ARVC phenotypes in mice by preventing connexome attrition (63). Components of the connexome are required in second heart field precursors for normal heart development (61). It is not known if

21

the molecular functions of the connexome are also required for outflow tract or aortic valve formation. However, the prevalence of BAV is not known to be enriched in ARVC patients.

KIF1A and SHANK3 are enriched in neural tissues and are associated with neurodevelopmental disorders such as autism, epilepsy, or spastic paraplegia (64). KIF1A encodes a kinesin-3 family microtubular motor required for organelle transport. SHANK3 encodes a scaffold protein that anchors actin filaments to the cell membrane and regulates cell motility (65). However, both genes are also highly expressed in the developing heart and have been implicated in cardiac malformations (66). Overexpression of the Drosophila *kif1A* ortholog results in a hypertrophic cardiomyopathy phenotype with cardiac valve malformations (67). Balanced translocations involving the KIF1A gene were observed in HLHS (67). Mutation of a mouse Kif1a paralog causes situs inversus with left-sided cardiac lesions including BAV (68). Mutations of SHANK3 cause Phelan-McDermid syndrome, which is characterized by a variable phenotype of neurodevelopmental abnormalities, syndromic facial features, and cardiac abnormalities including BAV (69).

Conclusion

In whole exome sequences of 215 probands with early onset complications of BAV disease, we identified likely pathogenic variants of CHD genes in 82 EBAV families (38%), including predicted pathogenic mutations that are known to cause BAV and HTAD. We identified variants in 27 other candidate genes that are associated with dominant cardiac phenotypes, including loss of function mutations in *PKP2*, *MYH6*, *KCNH2, KIF1A*, and *KMT2D*. Predicted pathogenic variants in *COL1A2*, *COL5A1*, *FBN2*, *MIB1*, *MYH6*, and *NOTCH3* were also observed in a GenTAC BAV cohort and in cohorts with familial or early onset HTAD. The

22

findings from this study corroborate our hypothesis that genetic variants promoting the early onset of BAV disease are distinct and overlap with genetic variants that cause other congenital left-sided cardiac lesions. Genetic testing for predictive variants in these candidate genes may eventually prove useful to identify patients who may benefit from more aggressive surveillance or therapies to prevent BAV complications.

Limitations

Genetic and phenotypic data on many EBAV families were incomplete, and the sample size of the EBAV cohort was insufficient for subgroup analysis. These factors likely decreased penetrance estimates and reduced sensitivity to detect genotype-phenotype correlations. Some contributory genetic variants may have been filtered out when recurrent rare variants were prioritized during the analysis. The limited evidence for familial segregation suggests that some variants are unlikely to cause BAV but modify the onset of clinically apparent disease. Future functional studies may be needed to validate variant effects and define the predictive utility of genetic testing.

Acknowledgments

The EBAV Investigators are: Shaine A. Morris, Rita Milewski, Giuseppe Limongelli, Allesandro Della Corte, Laura Perrone, Yuli Y. Kim, Hector Michelena, Maria G. Andreassi, Arturo

23

Evangelista, Denver Sallee, Anji Yetman, Kim McBride, Eduardo Bossone, Rodolfo Citro,

Dawn S. Hui, Malenka M. Bissell, Andrea Ballotti, Ilenia Foffa, Margot De Marco, Anthony

Caffarelli, Rita Weise, Julie DeBacker, Laura Muino Mosquera, Robbin Cohen, Laura Dos

Subira, Justin T. Tretter, Anna Sabe Rotes, Martina Caiazza, Lamia Ait Ali, Francesca

Pluchinotta, and Simon C. Body.

References

1. Liu Y, Chen S, Zühlke L, Black GC, Choy M kit, Li N, et al. Global birth prevalence of congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies. Int J Epidemiol. 2019 Apr;48(2):455–63.

2. Michelena HI, Prakash SK, Della Corte A, Bissell MM, Anavekar N, Mathieu P, et al. Bicuspid Aortic Valve. Circulation. 2014 Jun 24;129(25):2691–704.

3. Glotzbach JP, Hanson HA, Tonna JE, Horns JJ, McCarty Allen C, Presson AP, et al. Familial Associations of Prevalence and Cause-Specific Mortality for Thoracic Aortic Disease and Bicuspid Aortic Valve in a Large-Population Database. Circulation. 2023 Aug 22;148(8):637–47.

4. Kahveci G, Bayrak F, Pala S, Mutlu B. Impact of Bicuspid Aortic Valve on Complications and Death in Infective Endocarditis of Native Aortic Valves. Tex Heart Inst J. 2009;36(2):111–6.

5. Tripathi A, Wang Y, Jerrell JM. Population-based treated prevalence, risk factors, and outcomes of bicuspid aortic valve in a pediatric Medicaid cohort. Ann Pediatr Cardiol. 2018;11(2):119–24.

6. Prakash SK, Yetman A, Bissell MM, Kim YY, Michelena H, Hui DS, et al. Recurrent genomic copy number variants implicate new candidate genes for early onset bicuspid aortic valve disease. J Am Coll Cardiol. 2019 Mar 12;73(9_Supplement_1):620–620.

7. Roman MJ, Pugh NL, Devereux RB, Eagle KA, Holmes K, LeMaire SA, et al. Aortic Dilatation Associated With Bicuspid Aortic Valve: Relation to Sex, Hemodynamics, and Valve Morphology (the National Heart Lung and Blood Institute-Sponsored National Registry of Genetically Triggered Thoracic Aortic Aneurysms and Cardiovascular Conditions). Am J Cardiol. 2017 Oct 1;120(7):1171–5.

8. Bhave SA, Guo DC, Angelov S, Bamshad MJ, Nickerson DA, Milewicz D, et al. Human SMAD4 Genomic Variants Identified in Individuals with Heritable and Early-Onset Thoracic Aortic Disease. Cardiogenetics. 2021 Sep;11(3):132–8.

9. Masri A, Svensson LG, Griffin BP, Desai MY. Contemporary natural history of bicuspid aortic valve disease: a systematic review. Heart Br Card Soc. 2017 Sep;103(17):1323–30.

10. Zhu W, Lo CW. Insights into the genetic architecture of congenital heart disease from animal modeling. Zool Res. 2023 May 18;44(3):577–90.

11. Xie Y, Jiang W, Dong W, Li H, Jin SC, Brueckner M, et al. Network assisted analysis of de novo variants using protein-protein interaction information identified 46 candidate genes for congenital heart disease. PLoS Genet. 2022 Jun;18(6):e1010252.

12. Bravo-Jaimes K, Prakash SK. Genetics in bicuspid aortic valve disease: Where are we? Prog Cardiovasc Dis. 2020;63(4):398–406.

13. Michelena HI, Vallabhajosyula S, Prakash SK. Nosology Spectrum of the Bicuspid Aortic Valve Condition: Complex-Presentation Valvulo-Aortopathy. Circulation. 2020 Jul 21;142(3):294–9.

14. Musfee FI, Guo D, Pinard AC, Hostetler EM, Blue EE, Nickerson DA, et al. Rare deleterious variants of NOTCH1, GATA4, SMAD6, and ROBO4 are enriched in BAV with early onset complications but not in BAV with heritable thoracic aortic disease. Mol Genet Genomic Med. 2020 Oct;8(10):e1406.

15. Yang LT, Ye Z, Wajih Ullah M, Maleszewski JJ, Scott CG, Padang R, et al. Bicuspid aortic valve: long-term morbidity and mortality. Eur Heart J. 2023 Nov 14;44(43):4549–62.

16. Bjornsson T, Thorolfsdottir RB, Sveinbjornsson G, Sulem P, Norddahl GL, Helgadottir A, et al. A rare missense mutation in MYH6 associates with non-syndromic coarctation of the aorta. Eur Heart J. 2018 Sep 7;39(34):3243–9.

17. MYH6 gene: MedlinePlus Genetics [Internet]. [cited 2024 Jan 1]. Available from: https://medlineplus.gov/genetics/gene/myh6/

18. Posch MG, Waldmuller S, Müller M, Scheffold T, Fournier D, Andrade-Navarro MA, et al. Cardiac Alpha-Myosin (MYH6) Is the Predominant Sarcomeric Disease Gene for Familial Atrial Septal Defects. PLoS ONE. 2011 Dec 14;6(12):e28872.

19. Theis JL, Hu JJ, Sundsbak RS, Evans JM, Bamlet WR, Qureshi MY, et al. Genetic Association Between Hypoplastic Left Heart Syndrome and Cardiomyopathies. Circ Genomic Precis Med. 2021 Feb;14(1):e003126.

20. Zuo JY, Chen HX, Liu ZG, Yang Q, He GW. Identification and functional analysis of variants of MYH6 gene promoter in isolated ventricular septal defects. BMC Med Genomics. 2022 Oct 8;15(1):213.

21. Tomita-Mitchell A, Stamm KD, Mahnke DK, Kim MS, Hidestrand PM, Liang HL, et al. Impact of MYH6 variants in hypoplastic left heart syndrome. Physiol Genomics. 2016 Dec 1;48(12):912–21.

22. Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet. 2017 Nov;49(11):1593–601.

23. Galvin KM, Donovan MJ, Lynch CA, Meyer RI, Paul RJ, Lorenz JN, et al. A role for Smad6 in development and homeostasis of the cardiovascular system. Nat Genet. 2000 Feb;24(2):171–4.

24. Dickel DE, Barozzi I, Zhu Y, Fukuda-Yuzawa Y, Osterwalder M, Mannion BJ, et al. Genome-wide compendium and functional assessment of in vivo heart enhancers. Nat Commun. 2016 Oct 5;7:12923.

25. Gillis E, Kumar AA, Luyckx I, Preuss C, Cannaerts E, van de Beek G, et al. Candidate Gene Resequencing in a Large Bicuspid Aortic Valve-Associated Thoracic Aortic Aneurysm Cohort: SMAD6 as an Important Contributor. Front Physiol. 2017 Jun 13;8:400.

26. Luyckx I, Verstraeten A, Goumans MJ, Loeys B. SMAD6-deficiency in human genetic disorders. Npj Genomic Med. 2022 Nov 21;7(1):1–11.

27. Parker LE, Landstrom AP. Genetic Etiology of Left-Sided Obstructive Heart Lesions: A Story in Development. J Am Heart Assoc. 2021 Jan 19;10(2):e019006.

28. Ruter DL, Liu Z, Ngo KM, X S, Marvin A, Buglak DB, et al. SMAD6 transduces endothelial cell flow responses required for blood vessel homeostasis. Angiogenesis. 2021 May;24(2):387–98.

29. Maioli M, Gnoli M, Boarini M, Tremosini M, Zambrano A, Pedrini E, et al. Genotype-phenotype correlation study in 364 osteogenesis imperfecta Italian patients. Eur J Hum Genet EJHG. 2019 Jul;27(7):1090–100.

30. Takeda R, Yamaguchi T, Hayashi S, Sano S, Kawame H, Kanki S, et al. Clinical and molecular features of patients with COL1-related disorders: Implications for the wider spectrum and the risk of vascular complications. Am J Med Genet A. 2022 Sep;188(9):2560–75.

31. Ashournia H, Johansen FT, Folkestad L, Diederichsen ACP, Brixen K. Heart disease in patients with osteogenesis imperfecta - A systematic review. Int J Cardiol. 2015 Oct 1;196:149–57.

32. Prendergast A, Ziganshin BA, Papanikolaou D, Zafar MA, Nicoli S, Mukherjee S, et al. Phenotyping Zebrafish Mutant Models to Assess Candidate Genes Associated with Aortic Aneurysm. Genes. 2022 Jan 10;13(1):123.

33. Roulet M, Ruggiero F, Karsenty G, LeGuellec D. A comprehensive study of the spatial and temporal expression of the col5a1 gene in mouse embryos: a clue for understanding collagen V function in developing connective tissues. Cell Tissue Res. 2007 Feb;327(2):323–32.

34. Wenstrup RJ, Florer JB, Brunskill EW, Bell SM, Chervoneva I, Birk DE. Type V collagen controls the initiation of collagen fibril assembly. J Biol Chem. 2004 Dec 17;279(51):53331–7.

35. Malfait F, Wenstrup RJ, De Paepe A. Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. Genet Med Off J Am Coll Med Genet. 2010 Oct;12(10):597–605.

36. De Paepe A, Nuytinck L, Hausser I, Anton-Lamprecht I, Naeyaert JM. Mutations in the COL5A1 gene are causal in the Ehlers-Danlos syndromes I and II. Am J Hum Genet. 1997 Mar;60(3):547–54.

37. Richer J, Hill HL, Wang Y, Yang ML, Hunker KL, Lane J, et al. A Novel Recurrent COL5A1 Genetic Variant Is Associated With a Dysplasia-Associated Arterial Disease

Exhibiting Dissections and Fibromuscular Dysplasia. Arterioscler Thromb Vasc Biol. 2020 Nov;40(11):2686–99.

38. Chen P, Yu B, Li Z, Chen Y, Sun Y, Wang DW. COL5A1 Variants Cause Aortic Dissection by Activating TGF-β-Signaling Pathway. J Am Heart Assoc. 2021 Jun;10(11):e019276.

39. Monroe GR, Harakalova M, van der Crabben SN, Majoor-Krakauer D, Bertoli-Avella AM, Moll FL, et al. Familial Ehlers-Danlos syndrome with lethal arterial events caused by a mutation in COL5A1. Am J Med Genet A. 2015 Jun;167(6):1196–203.

40. Park AC, Phan N, Massoudi D, Liu Z, Kernien JF, Adams SM, et al. Deficits in Col5a2 Expression Result in Novel Skin and Adipose Abnormalities and Predisposition to Aortic Aneurysms and Dissections. Am J Pathol. 2017 Oct;187(10):2300–11.

41. Erhart P, Körfer D, Dihlmann S, Qiao JL, Hausser I, Ringleb P, et al. Multiple Arterial Dissections and Connective Tissue Abnormalities. J Clin Med. 2022 Jun 7;11(12):3264.

42. Traenka C, Kloss M, Strom T, Lyrer P, Brandt T, Bonati LH, et al. Rare genetic variants in patients with cervical artery dissection. Eur Stroke J. 2019 Dec;4(4):355–62.

43. Grond-Ginsbach C, Chen B, Pjontek R, Wiest T, Jiang Y, Burwinkel B, et al. Copy number variation in patients with cervical artery dissection. Eur J Hum Genet EJHG. 2012 Dec;20(12):1295–9.

44. Ritelli M, Venturini M, Cinquina V, Chiarelli N, Colombi M. Multisystemic manifestations in a cohort of 75 classical Ehlers-Danlos syndrome patients: natural history and nosological perspectives. Orphanet J Rare Dis. 2020 Jul 31;15(1):197.

45. Wu B, Li J, Wang Y, Cheng Y, Wang C, Shu X. Recurrent germline mutations as genetic markers for aortic root dilatation in bicuspid aortic valve patients. Heart Vessels. 2021 Apr;36(4):530–40.

46. Chen S, Jin Q, Hou S, Li M, Zhang Y, Guan L, et al. Identification of recurrent variants implicated in disease in bicuspid aortic valve patients through whole-exome sequencing. Hum Genomics. 2022 Sep 7;16(1):36.

47. KCNH2 gene: MedlinePlus Genetics [Internet]. [cited 2024 Jan 1]. Available from: https://medlineplus.gov/genetics/gene/kcnh2/

48. Abou Ziki MD, Seidelmann SB, Smith E, Atteya G, Jiang Y, Fernandes RG, et al. Deleterious protein-altering mutations in the SCN10A voltage-gated sodium channel gene are associated with prolonged QT. Clin Genet. 2018 Apr;93(4):741–51.

49. Park DS, Fishman GI. Nav-igating through a complex landscape: SCN10A and cardiac conduction. J Clin Invest. 2014 Apr 1;124(4):1460–2.

50. Hu D, Barajas-Martínez H, Pfeiffer R, Dezi F, Pfeiffer J, Buch T, et al. Mutations in SCN10A are responsible for a large fraction of cases of Brugada syndrome. J Am Coll Cardiol. 2014 Jul 8;64(1):66–79.

51. Hanania HL, Regalado ES, Guo DC, Xu L, Demo E, Sallee D, et al. Do HCN4 Variants Predispose to Thoracic Aortic Aneurysms and Dissections? Circ Genomic Precis Med. 2019 Dec;12(12):e002626.

52. Rojanasopondist P, Nesheiwat L, Piombo S, Porter GA, Ren M, Phoon CKL. Genetic Basis of Left Ventricular Noncompaction. Circ Genomic Precis Med. 2022 Jun;15(3):e003517.

53. Wang C, Hata Y, Hirono K, Takasaki A, Ozawa SW, Nakaoka H, et al. A Wide and Specific Spectrum of Genetic Variants and Genotype-Phenotype Correlations Revealed by Next-Generation Sequencing in Patients with Left Ventricular Noncompaction. J Am Heart Assoc. 2017 Aug 30;6(9):e006210.

54. Caiffa T, Tessitore A, Leoni L, Reffo E, Chicco D, D'Agata Mottolese B, et al. Long QT syndrome and left ventricular non-compaction in a family with KCNH2 mutation: A case report. Front Pediatr [Internet]. 2022 [cited 2024 Jan 1];10. Available from: https://www.frontiersin.org/articles/10.3389/fped.2022.970240

55. Kharbanda M, Hunter A, Tennant S, Moore D, Curtis S, Hancox JC, et al. Long QT syndrome and left ventricular noncompaction in 4 family members across 2 generations with KCNQ1 mutation. Eur J Med Genet. 2017 May;60(5):233–8.

56. Teng G, Zhao X, Lees-Miller JP, Belke D, Shi C, Chen Y, et al. Role of Mutation and Pharmacologic Block of Human KCNH2 in Vasculogenesis and Fetal Mortality. Circ Arrhythm Electrophysiol. 2015 Apr;8(2):420–8.

57. Shimizu W, Horie M, Priori S. Phenotypic Manifestations of Mutations in Genes Encoding Subunits of Cardiac Potassium Channels. Circ Res. 2011 Jun 24;109(1):97–109.

58. Tessler I, Albuisson J, Piñeiro-Sabarís R, Verstraeten A, Kamber Kaya HE, Siguero-Álvarez M, et al. Novel Association of the NOTCH Pathway Regulator MIB1 Gene With the Development of Bicuspid Aortic Valve. JAMA Cardiol. 2023 Aug 1;8(8):721–31.

59. Ross SB, Singer ES, Driscoll E, Nowak N, Yeates L, Puranik R, et al. Genetic architecture of left ventricular noncompaction in adults. Hum Genome Var. 2020 Oct 15;7(1):1–8.

60. PKP2 gene: MedlinePlus Genetics [Internet]. [cited 2024 Jan 11]. Available from: https://medlineplus.gov/genetics/gene/pkp2/

61. Agullo-Pascual E, Cerrone M, Delmar M. Arrhythmogenic cardiomyopathy and Brugada syndrome: diseases of the connexome. FEBS Lett. 2014 Apr 17;588(8):1322–30.

62. Cerrone M, Montnach J, Lin X, Zhao YT, Zhang M, Agullo-Pascual E, et al. Plakophilin-2 is required for transcription of genes that control calcium cycling and cardiac rhythm. Nat Commun. 2017 Jul 24;8(1):106.

63. Bradford WH, Zhang J, Gutierrez-Lara EJ, Liang Y, Do A, Wang TM, et al. Plakophilin 2 gene therapy prevents and rescues arrhythmogenic right ventricular cardiomyopathy in a mouse model harboring patient genetics. Nat Cardiovasc Res. 2023 Dec;2(12):1246–61.

64. Nair A, Greeny A, Rajendran R, Abdelgawad MA, Ghoneim MM, Raghavan RP, et al. KIF1A-Associated Neurological Disorder: An Overview of a Rare Mutational Disease. Pharmaceuticals. 2023 Jan 19;16(2):147.

65. SHANK3 SH3 and multiple ankyrin repeat domains 3 [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2024 Jan 1]. Available from: https://www.ncbi.nlm.nih.gov/gene/85358

66. Wang Y, Xu Y, Guo W, Fang Y, Hu L, Wang R, et al. Ablation of Shank3 alleviates cardiac dysfunction in aging mice by promoting CaMKII activation and Parkin-mediated mitophagy. Redox Biol. 2022 Dec;58:102537.

67. Akasaka T, Ocorr K, Lin L, Vogler G, Bodmer R, Grossfeld P. Overexpression of Kif1A in the Developing Drosophila Heart Causes Valvar and Contractility Defects: Implications for Human Congenital Heart Disease. J Cardiovasc Dev Dis. 2020 Jun 2;7(2):22.

68. Marszalek JR, Ruiz-Lozano P, Roberts E, Chien KR, Goldstein LS. Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. Proc Natl Acad Sci U S A. 1999 Apr 27;96(9):5043–8.

69. Costales JL, Kolevzon A. Phelan-McDermid Syndrome and SHANK3: Implications for Treatment. Neurother J Am Soc Exp Neurother. 2015 Jul;12(3):620–30.