Exome Sequencing and Congenital Heart Disease in Sub-Saharan Africa

Ekanem N. Ekure, MB, BCh[1,*], **Adebowale Adeyemo, MD**[2,*], **Hanhan Liu, PhD**[3], **Ogochukwu Sokunbi, MBBS**[1], **Nnenna Kalu, MBBS**[1], **Ariel F. Martinez, PhD**[4], **Babajide Owosela, BS**[4], **Cedrik Tekendo-Ngongang, MD**[4], **Yonit A. Addissie, BA**[4], **Akinsanya Olusegun-Joseph, MBBS**[5], **Desmond Ikebudu, BSc**[6], **Seth I. Berger, MD, PhD**[7], **Maximilian Muenke, MD**[4], **Zhe Han, PhD**[3], **Paul Kruszka, MD**[4]

[1]Department of Pediatrics, University of Lagos/Lagos University Teaching Hospital, Lagos, Nigeria

[2]National Human Genome Research Institute, National Human Genome Research Institute, National Institutes of Health, Bethesda

[3]Center for Precision Disease Modeling, University of Maryland School of Medicine, Baltimore, MD

[4]Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda

[5]Department of Medicine, College of Medicine, University of Lagos/Lagos University Teaching Hospital, Lagos, Nigeria

[6]Central Research Laboratory, College of Medicine, University of Lagos, Idi-Araba, Lagos, Nigeria

[7]Center for Genetic Medicine Research, Children's National Research Institute, Washington, DC

Abstract

Background ——Congenital heart disease (CHD) is the most common birth defect and affects roughly 1% of the global population. There have been a number of large CHD sequencing projects in developing countries, but none in sub-Saharan Africa. In this exome sequencing study, we recruited families from Lagos, Nigeria affected by structural heart disease.

Methods ——Ninety-eight participants with CHD and an average age of 3.6 years were recruited from Lagos, Nigeria. Exome sequencing was performed on probands, and parents when available. For genes of high interest, we conducted functional studies in *Drosophila* using a cardiac-specific RNA interference (RNAi)-based gene silencing system.

Results ——The three most common CHDs were tetralogy of Fallot (20%), isolated ventricular septal defect (14%), and transposition of the great arteries (8%). Ten percent of the cohort had pathogenic or likely pathogenic variants in genes known to cause CHD. In 64 complete trios, we

Correspondence: Paul Kruszka, MD, MPH, FACMG, National Institutes of Health Building, 35 Convent Drive, Room 1B203, Bethesda, MD 20892, Tel: 301-402-9654, paul.kruszka@nih.gov.
*contributed equally

found 34 *de novo* variants that were not present in the African population in the gnomAD (v3) database. Nineteen loss of function variants were identified using the genome wide distribution of selection effects for heterozygous protein truncating variants (s_{het}). Nine genes caused a significant mortality when silenced in the *Drosophila* heart, including 4 novel disease genes not previously associated with CHD (*UBB, EIF4G3, SREBF1,* and *METTL23*).

Conclusions ——This study identifies novel candidate genes and variants for CHD and facilitates comparisons with previous CHD sequencing studies in predominantly European cohorts. The study represents an important first step in genomic studies of CHD in under-studied populations.

Keywords

genetic heart disease; exome; exome sequencing; sub-Saharan Africa; Drosophila; Genetics; Congenital Heart Disease; Developmental biology

Introduction

Congenital heart disease (CHD) is the most common birth defect and affects roughly 1% of the global population.[1] Much progress has been made in elucidating the genetic component of CHD with a fourth of CHD being attributed to aneuploidies and copy number variants (CNVs).[2–4] The association of genetic syndromes with congenital heart disease has provided a genetic basis for CHD, as demonstrated by 50% of individuals with Down syndrome having congenital heart disease with most of these defects being atrioventricular septal defects[5] and 50%–80% of individuals with Noonan syndrome having structural heart defects.[6] The first non-syndromic CHD gene using linkage analysis was *NKX2–5*;[7] however, this technique is not useful in sporadic CHD as it requires informative pedigrees. Much of CHD is sporadic, and familial recurrence is very low, as shown in the Baltimore Washington Infants Study, where in the non-syndromic population of individuals with CHD, familial recurrence of CHD only occurred in 3–5%.[8] Many questions remain as the majority of CHD do not have an underlaying etiology and incomplete penetrance and variable expressivity are commonplace.

Over the last decade, there have been several large exome sequencing studies that offered evidence of a genetic basis to CHD. Three studies have used the US National Heart, Lung, and Blood Institute's (NHLBI) Pediatric Cardiac Genomics Consortium (PCGC) which consists of over 10,000 CHD probands and over 5,000 trios.[9–11] The initial two studies showed that roughly 10% of cases were attributable to *de novo* variants in genes that were highly expressed in the developing heart.[9] The Homsy et al. study also showed that there was an excess of *de novo* variants in individuals with CHD plus extracardiac congenital anomalies and neurodevelopment disorders but not isolated CHD.[9] In a separate cohort than PCGC, Sifrim et al. confirmed the findings of the PCGC studies and showed that in isolated CHD, there was an increase in damaging inherited variants.[12]

In common with most genetic and genomic studies, most of CHD studies have been carried out in European ancestry populations. The largest of the PCGC studies used a predominantly European descent cohort with 71.9% of European descent and only 6.6% were African

American.[10] Similarly, the Sifrim et al. (2016) study was also predominantly European (>80%).[12] The genetics of CHD is particularly understudied in sub-Saharan Africa, and with the exception of studies on genetic syndromes, there has been no wide scale sequencing projects for congenital heart disease in sub-Saharan Africa.

This study focuses on two questions. First, it evaluates the genetic basis of CHD in a sub-Saharan population for genes of known and syndromic importance. Second, it uses an exome sequencing approach to identify new genes associated with CHD with functional validation studies in a *Drosophila* model for genes of higher interest.

Methods

The methods for this study are available in the supplemental material. The data that supports the findings of this study are available from the corresponding author upon reasonable request. Patients and family members were consented to the National Institutes of Health (NIH) protocol Heart Genetics Study in Africa (13-HG-N207). Ethical approval for the studies was obtained from the Lagos University Teaching Hospital Health Research Ethics Committee and the National Human Genome Research Institute Institutional Review Board (IRB).

Results

Clinical

Age, gender, cardiac diagnosis, and extra-cardiac phenotype for each patient are listed in Supplemental Table 1. Ninety eight probands with echocardiography confirmed CHD were evaluated. Sixty-four (65%) of these probands also had their parents examined and sequenced (trios), another 13 probands had one parent evaluated and 21 were singletons. One trio was not analyzed due to a Mendelian error (participant 11791) with potential explanations including non-paternity or sample switch; this trio was analyzed as a mother-proband dyad. The average age was 43 months (3.6 years) and 55% of the study population were males (Table 1). Twenty three percent of individuals were classified as syndromic if other congenital anomalies and developmental delay were present (Table 1; Supplemental Table 1). The three most common CHDs were tetralogy of Fallot (20%), isolated ventricular septal defect (14%), and transposition of the great arteries (8%) (Table 1).

Copy number variants; XHMM analysis

Chromosomal microarray was completed on 66 (68%) of the probands. Individuals found to have contiguous deletion syndromes associated with CHD such as 22q11.2 deletion syndrome were excluded from this study. Only one participant (12160) with a CNV greater than 1Mb was included in the study. This individual had a 1.1 Mb duplication at 14q11.2 duplication (chr14:19327823–20420338×3; hg19) which did not contain genes known to cause CHD. In order to cover the remaining 32% of probands without chromosomal microarray, eXome-Hidden Markov Model (XHMM) CNV analysis was performed on all study participants (See Supplemental Methods).[13] The results of the XHMM analysis are shown in Supplemental Table 2. The largest CNV found with XHMM was a 3.8 Mb deletion that contained *GATA4* (study participant 12737; Supplemental Table 2); this individual had

not had chromosomal microarray. The remaining CNVs in Supplemental Table 2 did not contain genes known to be associated with CHD.

Candidate gene approach

The cohort was screened for a set of genes that are known to be associated with CHD in humans (http://chdgene.victorchang.edu.au).[14] Using ACMG-AMP guidelines,[15] we found pathogenic or likely pathogenic variants in 10 individuals (Supplemental Table 3); however, one individual had a heterozygous variant in autosomal recessive associated conditions, leaving 9 individuals (9%) with likely pathogenic or pathogenic in known CHD associated genes (Table 2), variants of unknown significance (VUS) in 26 cases and benign or likely benign in 53 individuals (Supplemental Table 3). Figure 1 shows extracardiac phenotypes in individuals with syndromic CHD. The majority of variants were dominant inheritance except for compound heterozygous variants in *NPHP4* (individual 12791) that were VUSs; given that this individual does not have eye or renal issues, these are more likely benign variants (Supplemental Table 3).

De novo variants

In the 64 trios, a total of 57 *de novo* variants (not including synonymous changes) were found. This *de novo* rate of 0.89 is similar to 0.80 in Sifrim et als study where 1098 non-synonymous *de novo* variants were found in 1,365 trios (p=0.08, chi-square). There were 34 *de novo* variants that were not in the African population in the gnomAD (v3) data base and had a CADD score greater than 20 (Supplemental Table 4). Of these 34 genes, 6 were pathogenic variants in known CHD genes. The remaining 28 variants were of high interest given their *de novo* inheritance and in silico damaging prediction. Four of these 28 variants were LOF variants in the genes *XPOT, CRTC1, USP26,* and *PHLDA1,* and the remainder were missense variants.

Protein truncating variants in genes intolerant to loss of function

Given that intolerance to LOF of one or both copies of a gene is related to disease, we evaluated all presumed LOF variants in this study cohort that were not tolerant to variation. Using the genome wide distribution of selection effects for heterozygous protein truncating variants developed by Cassa et al. (s_{het}), we were able to discriminate between genes with variants that were more likely to result in disease.[18] Our cut off was a fitness loss corresponding to s_{het} greater than 0.04 (4%). There were 19 total presumed LOF variants (splice variants, frameshift indels, stopgain) with s_{het} >0.04 (Supplemental Table 5). In the interest of minimizing false positive and increasing specificity, we further refined LOF variants to minor allele frequency in gnomAD (v3) Africa to $< 10^{-5}$ and a probability of loss of function intolerance score (pLI) of > 0.9 (Table 3).

Drosophila functional testing

For the human homologs of genes that were constrained ($s_{het} > 0.04$) and had LOF variants in our cohort, cardiac-specific RNA interference (RNAi) mediated gene silencing in *Drosophila* resulted in mortality indexes that are shown in Supplemental Table 5 in decreasing severity. The mortality index is a measure of developmental mortality attributable

to RNAi heart expression. *UBB* had the highest mortality index (100%) and has not been previously associated with human disease. The second highest mortality index in the fly was for *EIF4G3,* also a gene not previously associated with human disease that codes for a component of the eukaryotic translation initiation factor 4F required for recruitment of ribosomes to mRNA.[19] The third gene on the list is *CHD7.* Although this gene is well known to cause CHARGE syndrome and heart malformations, it provides reassurance and quality control for functional testing in the fly model.

Discussion

We present the first large scale next generation sequencing study of congenital heart disease in sub-Saharan Africa. In this study we exome sequenced 98 probands and their families using a three-pronged analytic approach. The first analysis or "prong" was of pathologic variants in genes known to cause CHD, the second analysis was *de novo* variants, and the third subgroup was LOF variants in contrained genes. In the last analysis of LOF in constrained genes, mortality indices of the fly with heart specific knockouts were evaluated.

Nine percent of this cohort had pathogenic or likely pathogenic variants in genes known to cause CHD (Table 2). A similar sized study found a two fold greater amount (22%) of pathogenic or likely pathogenic variants in genes known to cause CHD.[20] This study was 76% Caucasian; however, it is unlikely that ethnicity plays a large role in pathogenic variant burden as most cases of CHD are sporadic and autosomal dominant. A more likely explanation is ascertainment bias. In African CHD studies, there is a large fraction of severely affected individuals.[21] Shown in the present study, TOF alone makes up 20% of diagnoses in probands. Of special interest in this group, *MYH7* is usually associated with cardiomyopathy; however, one individual in the present study has a VSD and a likely pathogenic variant in *MYH7.* There have been a few other reports of *MYH7* associated with structural CHD, reinforcing this gene's association with structural heart disease.[22–24]

The second category was *de novo* variants that were absent in the African gnomAD (v3) population and had CADD scores above 20. There were 34 variants in this group; however, 6 of these variants (Table 2) were in genes already known to cause CHD, leaving 28 variants of high interest (Supplemental Table 3). Most of these 28 variants were missense variants and will need functional studies combined with additional unrelated affected individuals with variants in the same gene to be considered a genetic etiology for CHD. However, it is important to publish these variants and genes of interest as future publications will likely associate a fraction of these genes with CHD.

The most difficult aspect of next generation sequencing studies is assigning pathogenicity to novel genes and variants. In order to maximize our gene discovery efforts, our third prong or category were 19 variants that were loss of function variants in constrained genes (Supplemental Table 5). In the early stages of this project we used the genome wide distribution of selective effects for heterogenous protein coding variants developed from the ExAC database (Supplementary Table 5). Sixteen of these genes has homologous *Drosophila* genes (Supplementary Table 4). Although the evolutionary distance between the fruit fly and humans is large, *Drosophila* and human hearts share developmental networks.[25]

We found that 9 of 16 genes in this group resulted in elevated fly mortality after heart specific gene silencing; and 5 of these genes were previously known to cause CHD in humans, allowing for quality assurance of the assay. The remaining 4 genes are novel CHD genes of high interest (*UBB, EIF4G3, SREBF1,* and *METTL23*). Two of these genes, *EIF4G3* and *UBB* have not previously been associated with human disease. In *EIF4G3*, a canonical splice site variant was found in a 3 year old child with TOF. *EIF4G3* is among the 1p36 deletion syndrome deleted genes and 1p36 deletion syndrome is known to be associated with CHD with one study showing 71% of individuals to have CHD and 8% with TOF.[26] In *UBB,* we found a frameshift insertion in a child with TOF. The fly data is convincing for *UBB,* with a mortality index of 100%; however, more unrelated CHD cases in the future will be helpful for assigning causality. Interestingly, *SREBF1* was recently associated with ichthyosis follicularis with atrichia and photophobia syndrome (IFAP syndrome).[27] In this study, there were three variants in two unrelated families and an additional 9 simplex cases, and these variants were in residues (527, 528, and 530) that were involved in site-1-protease (S1P) dependent cleavage of SREBF1. The variant in our study (c.3387dupC; p.Gly1130Argfs*15) was located in a different domain. *METTL23* has also been associated with an autosomal recessive human disease, intellectual disability (MIM 615942). As this study progressed, more population data became available and we further refined our loss of function variants to pLI > 0.9 and MAF < 10^{-5} in the gnomAD (v3) database (Table 3). In Table 3, there are variants in 8 genesand 5 of these genes are known to cause CHD, again, providing excellent quality control.

Two genes in Table 3 did not affect fly mortality and one gene did not have an available fly line; however, these three genes, *PRKCB, FRYL,* and *XPOT* are still of interest given their constraint and LOF status in our cohort. Although not associated with human disease (not in OMIM; omim.org visited on June 11, 2020), loss of function variants in *FRYL* have been reported in other CHD cohorts. Sifirm et al. also associated a stop gain in *FRYL,* c.6121C>T p.(R2041*) with a syndromic individual with CHD.[12] *PRKCB* is an example of a gene that is interesting but we only have one patient with a variant in this gene. In the gnomAD database, there are no LOF variants even though 36.8 were expected [o/e 0 (0.0–0.8)]. Special mention is made of *ANAK* (Supplementary Table 4); the pLI = 0.9 for *AHNAK,* thus precluding its presence in Table 3. Although *AHNAK* is not associated with human disease in OMIM, Homsey et al. reported the *AHNAK* stop gain variant c.760G>T p.(G254*) in an individual with with aortic arch hypoplasia and aortic stenosis.[9]

There are a number of similarities and differences with this study compared to other next generation sequencing studies of cohorts of individuals affected by CHD. Beginning with differences, the first is the methodology of the current study in that we did not use a control cohort to conduct burden testing. Previous studies were able to show an increase in the burden of *de novo* variants in genes highly expressed in the developing heart.[9, 11, 12] However, unlike other next generation sequencing studies, the present study used functional studies to evaluate the highest interest genes. Given that all studies to date including the present study have a large number of interesting unaccompanied de novo missense variants, a high throughput functional study methodology is needed. Similar to the other CHD sequencing studies, autosomal dominant inheritance is the most common pattern.[9–12]

Another similarity is that the present study only found variants in genes known to cause CHD in a minority of probands (9%).

A key strength of the present study is the ancestry of the participants. Unlike previous CHD studies that have focused on individuals of European descent, this study is entirely of sub-Saharan Africans from Nigeria and with 239 exomes is, to date, amongst the largest exome sequencing studies for a clinical disorder in sub-Saharan Africa.[28] African populations are genetically more diverse than non-African populations due to their longer evolutionary history.[29] It is well known that individuals of African descent are underrepresented in genomic studies and reference databases[30] which may lead to misclassification of variants. [28, 31] The present study adds to the small but growing literature on genetic and genomic studies of complex disorders in African populations.

Limitations to this study include small sample size. Even though this is the first study in sub-Saharan Africa, larger studies and in more countries in Africa will be needed to more fully delineate the genetic architecture of CHD on the continent. Given the large number of variants identified by exome sequencing in this study, another limitation was that variants were not confirmed by an orthogonal method such as Sanger sequencing, leaving the possibility that a small fraction of variants are false positives. We have miminized the possibility of false positives with aggressive variant quality control and filtering (Supplemental Methods).

In conclusion, this study represents the first large scale CHD next generation sequence study in sub-Saharan Africa. Genes known to be associated with CHD and novel genes are identified in this study and compared to CHD studies of predominantly European descent. The study represents an important first step in genomic studies of CHD in under-studied populations and highlight the need for such studies to refine our knowledge of the genetic architecture of CHD and inform translation of such findings into clinical care.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

Nonstandard Abbreviations and Acronyms

TOF	Tetralogy of Fallot
AS	Aortic stenosis

ASD	Atrial septal defect
AVSD	Atrioventricular septal defect
BAV	Bicuspid aortic valve
DORV	Double outlet right ventricle
PDA	Patent ductus arteriosus
PPS	Peripheral pulmonary stenosis
PS	Pulmonary stenosis
RVOT	Right ventricular outflow obstruction
s_{het}	genome wide distribution of selection effects for heterozygous protein truncating
TGA	Transposition of great arteries
VSD	Ventricular septal defect

References:

1. van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, Roos-Hesselink JW. Birth prevalence of congenital heart disease worldwide: A systematic review and meta-analysis. J Am Coll Cardiol. 2011;58:2241–2247 [PubMed: 22078432]

2. Soemedi R, Wilson IJ, Bentham J, Darlay R, Topf A, Zelenika D, Cosgrove C, Setchfield K, Thornborough C, Granados-Riveron J, et al. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. Am J Hum Genet. 2012;91:489–501 [PubMed: 22939634]

3. Glessner JT, Bick AG, Ito K, Homsy J, Rodriguez-Murillo L, Fromer M, Mazaika E, Vardarajan B, Italia M, Leipzig J, et al. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. Circ Res. 2014;115:884–896 [PubMed: 25205790]

4. Zaidi S, Brueckner M. Genetics and genomics of congenital heart disease. Circ Res. 2017;120:923–940 [PubMed: 28302740]

5. Muenke M, Kruszka PS, Sable CA, Belmont JW. Congenital heart disease : Molecular genetics, principles of diagnosis and treatment. Basel; New York: Karger; 2015.

6. Allanson JE, Roberts AE. Noonan syndrome In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, eds. Generereviews((r)). Seattle (WA); 1993.

7. Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE, Seidman JG. Congenital heart disease caused by mutations in the transcription factor nkx2–5. Science. 1998;281:108–111 [PubMed: 9651244]

8. Calcagni G, Digilio MC, Sarkozy A, Dallapiccola B, Marino B. Familial recurrence of congenital heart disease: An overview and review of the literature. Eur J Pediatr. 2007;166:111–116 [PubMed: 17091259]

9. Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science. 2015;350:1262–1266 [PubMed: 26785492]

10. Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant MC, et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet. 2017;49:1593–1601 [PubMed: 28991257]

11. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, et al. De novo mutations in histone-modifying genes in congenital heart disease. Nature. 2013;498:220–223 [PubMed: 23665959]

12. Sifrim A, Hitz MP, Wilsdon A, Breckpot J, Turki SH, Thienpont B, McRae J, Fitzgerald TW, Singh T, Swaminathan GJ, et al. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. Nat Genet. 2016;48:1060–1065 [PubMed: 27479907]

13. Fromer M, Moran JL, Chambert K, Banks E, Bergen SE, Ruderfer DM, Handsaker RE, McCarroll SA, O'Donovan MC, Owen MJ, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. Am J Hum Genet. 2012;91:597–607 [PubMed: 23040492]

14. Szot JO, Cuny H, Blue GM, Humphreys DT, Ip E, Harrison K, Sholler GF, Giannoulatou E, Leo P, Duncan EL, et al. A screening approach to identify clinically actionable variants causing congenital heart disease in exome data. Circ Genom Precis Med. 2018;11:e001978 [PubMed: 29555671]

15. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the american college of medical genetics and genomics and the association for molecular pathology. Genet Med. 2015;17:405–424 [PubMed: 25741868]

16. Ekure EN, Addissie YA, Sokunbi OJ, Kruszka P, Muenke M, Adeyemo AA. Tuberous sclerosis in a patient from nigeria. Am J Med Genet A. 2019;179:1423–1425 [PubMed: 31140686]

17. Kruszka P, Porras AR, Addissie YA, Moresco A, Medrano S, Mok GTK, Leung GKC, Tekendo-Ngongang C, Uwineza A, Thong MK, et al. Noonan syndrome in diverse populations. Am J Med Genet A. 2017;173:2323–2334 [PubMed: 28748642]

18. Cassa CA, Weghorn D, Balick DJ, Jordan DM, Nusinow D, Samocha KE, O'Donnell-Luria A, MacArthur DG, Daly MJ, Beier DR, et al. Estimating the selective effects of heterozygous protein-truncating variants from human exome data. Nat Genet. 2017;49:806–810 [PubMed: 28369035]

19. Gingras AC, Raught B, Sonenberg N. Eif4 initiation factors: Effectors of mrna recruitment to ribosomes and regulators of translation. Annu Rev Biochem. 1999;68:913–963 [PubMed: 10872469]

20. Alankarage D, Ip E, Szot JO, Munro J, Blue GM, Harrison K, Cuny H, Enriquez A, Troup M, Humphreys DT, et al. Identification of clinically actionable variants from genome sequencing of families with congenital heart disease. Genet Med. 2019;21:1111–1120 [PubMed: 30293987]

21. Ekure EN, Kalu N, Sokunbi OJ, Kruszka P, Olusegun-Joseph AD, Ikebudu D, Bala D, Muenke M, Adeyemo A. Clinical epidemiology of congenital heart disease in nigerian children, 2012–2017. Birth Defects Res. 2018;110:1233–1240 [PubMed: 30230273]

22. van der Linde IHM, Hiemstra YL, Bokenkamp R, van Mil AM, Breuning MH, Ruivenkamp C, Ten Broeke SW, Veldkamp RF, van Waning JI, van Slegtenhorst MA, et al. A dutch myh7 founder mutation, p.(asn1918lys), is associated with early onset cardiomyopathy and congenital heart defects. Neth Heart J. 2017;25:675–681 [PubMed: 28864942]

23. Postma AV, van Engelen K, van de Meerakker J, Rahman T, Probst S, Baars MJ, Bauer U, Pickardt T, Sperling SR, Berger F, et al. Mutations in the sarcomere gene myh7 in ebstein anomaly. Circ Cardiovasc Genet. 2011;4:43–50 [PubMed: 21127202]

24. Hirono K, Hata Y, Ibuki K, Yoshimura N. Familial ebstein's anomaly, left ventricular noncompaction, and ventricular septal defect associated with an myh7 mutation. J Thorac Cardiovasc Surg. 2014;148:e223–226 [PubMed: 25444217]

25. Zhu JY, Fu Y, Nettleton M, Richman A, Han Z. High throughput in vivo functional validation of candidate congenital heart disease genes in drosophila. Elife. 2017;6

26. Battaglia A, Hoyme HE, Dallapiccola B, Zackai E, Hudgins L, McDonald-McGinn D, Bahi-Buisson N, Romano C, Williams CA, Brailey LL, et al. Further delineation of deletion 1p36 syndrome in 60 patients: A recognizable phenotype and common cause of developmental delay and mental retardation. Pediatrics. 2008;121:404–410 [PubMed: 18245432]

27. Wang H, Humbatova A, Liu Y, Qin W, Lee M, Cesarato N, Kortum F, Kumar S, Romano MT, Dai S, et al. Mutations in srebf1, encoding sterol regulatory element binding transcription factor 1, cause autosomal-dominant ifap syndrome. Am J Hum Genet. 2020

28. Bope CD, Chimusa ER, Nembaware V, Mazandu GK, de Vries J, Wonkam A. Dissecting in silico mutation prediction of variants in african genomes: Challenges and perspectives. Front Genet. 2019;10:601 [PubMed: 31293624]

29. Campbell MC, Tishkoff SA. African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet. 2008;9:403–433 [PubMed: 18593304]

30. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–291 [PubMed: 27535533]

31. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, Margulies DM, Loscalzo J, Kohane IS. Genetic misdiagnoses and the potential for health disparities. N Engl J Med. 2016;375:655–665 [PubMed: 27532831]

Figure 1.

A. Participant 12957 - Neurofibromatosis, type 1 (*NF1* c.888+1G>T). A large Café au-lait patch on the chest and others near the right axilla are noted. A midline sternotomy scar from a ventricular septal defect repair is also seen; B. Participant 11764 - Tuberous sclerosis (*TSC2* c.551T>A:p.Val184Asp). Multiple angiofibromas are seen on face;[16] C. Participant 12156 - Hadju-Cheney syndrome (*NOTCH2* c.7198C>T:p.Arg2400*). Notice pseudo-clubbing of fingernails, camptodactyly of the proximal interphalangeal joints of the 3rd and 4th digits, and clinodactyly of the 5th digits.

Table 1.

Cohort characteristics.

	Total participants	Non-syndromic participants	Syndromic participants
Number	98	75	23
Average age in months (standard deviation)	43 (51.4)	45 (50.9)	34 (54.1)
Sex (male)	55%	60%	40%
Tetralogy of Fallot	20	14	6
Isolated ventricular septal defect	14	11	3
Transposition of the great arteries	8	6	2
Atrioventricular septal defect	7	6	1
Left heart obstruction lesions	7	7	0
Isolated patent ductus arteriosus (PDA)	6	5	1
Tricuspid atresia	5	5	0
Combined septal and right ventricular outflow tract obstruction lesions	5	3	2
Combined septal and ductal lesions	4	3	1
Double outlet right ventricle	4	4	0
Pulmonary stenosis	3	2	1
Truncus arteriosus	3	2	1
Other single ventricle physiology hearts	3	2	1
Ebstein anomaly	2	2	0
Isolated atrial septal defect	1	0	1
Cardiomyopathy, PDA	1	1	0
Others- heterotaxy, anomalous origin of right pulmonary artery from the ascending aorta, Aortico-left ventricular tunnel, dextrocardia, left ventricle rhabdomyoma.	5	3	2

Table 2.

Congenital heart disease candidate genes with pathogenic or likely pathogenic variants

Participant	Variant	Inheritance	Syndrome	Sex	Age (months)	Echo diagnosis	Non-cardiac phenotype
12957	*NF1* c.888+1G>T	*de novo* (trio)	Neurofibromatosis, type 1	female	24	VSD	Café au lait spots, neurofibromas, developmental delay
12490	*JAG1* c.2103dupC; p.Cys702LeufsTer5	singleton	Alagille syndrome 1	male	84	AS	None
12156	*NOTCH2* c.7198C>T:p.Arg2400 [*]	*de novo* (trio)	Hadju-Cheney syndrome	male	10	Partial AVSD, PDA.	Flattened nasal bridge, micrognathia, impacted dentition and malocclusion, high arched palate, small midline cleft palate, low set ears, bathrocephaly and delayed closure of posterior fontanelle. Short stature, bone abnormalities, umbilical hernia, cryptorchidism.
12793	*CHD7* c.4480C>T:p.Arg1494 [*]	singleton	CHARGE syndrome	female	13	AS, BAVright aortic arch; PPS	Prominent forehead; developmental delay
12353	*KMT2A* c.3022_3023dupTC; p.Met1009Profs [*]11	*de novo* (trio)	Wiedemann-Steiner syndrome	female	2	VSD, PDA	none
11764	*TSC2* c.551T>A:p.Val184Asp	*de novo* (trio)	Tuberous sclerosis [*]	male	15	Cardiac rhabdomyoma	Skin lesions
12552	*MAP2K1* c.199G>A:p.Asp67Asn	*de novo* (trio)	Noonan syndrome [†]			PS, ASD	Developmental delay
12431	*MYH7* c.C4762T:p.Arg1588Cys	absent in mother; father not available	Non syndromic CHD	male	23	VSD	Hypertelorism; hooded eyelids; flattened nasal bridge
12160	*ACTC1* c.704C>T:p.Ser235Phe	*de novo* (trio)	Non syndromic CHD	male	36	VSD; Noncompaction cardiomyopathy	none

VSD - Ventricular septal defect; AS - Aortic stenosis; AVSD - Atrio-ventricular septal defect; AS - Aortic stenosis; BAV - Bicuspid aortic valve; PPS - Peripheral pulmonary stenosis; PDA - Patent ductus arteriosus; PS - Pulmonary stenosis; ASD – Atrial septal defect; CHD - Congenital heart disease.

[*] Reported in a previous publication[16]

[†] Reported in a previous publication[17]

Table 3.

Loss of function variants in variants with $s_{het} > 0.04$, pLI > 0.9 and minor allele frequency (MAF)_ in gnomAD (v3) < 10^{-5}. *Drosophila* heart specific knockout data included.

ID	Inheritance	Gene:variant	Phenotype	s_het	gnomAD_all (v3)	gnomAD_Africa (v3)	pLi	Fly gene name	Mortality Index
11759	de nov (trio)	*XPOT*:NM_007235:c.C1447T:p.R483*	DILV, hypoplastic RV, IAA, PDA, VSD	0.0579	6.98E-06	0	0.99711	none	not examined
12156	de nov (trio)	*NOTCH2*:NM_024408:c.7198C>T:p.R 2400X	chd, ostium primum asd, small pda, mild mr, mod pah	0.1227	0	0	1	Notch	37.6
12353	de nov (trio)	KMT2A:NM_001197104:c.3022_3023 dupTC; p.Met1009Profs*11	VSD, PDA	0.2168	0	0	1	trx	28
12419	duo (not found in mother)	*PRKCB*:NM_002738:c.1181T>A; p.L394*	TOF	0.3746	0	0	1	Pkc53E	0
12484	singleton	*FRYL*:NM_015030:c.8965_8969dupT CTCT; p.Arg2991Leufs*11	D-TGA, hypoplastic aorta, ASD, VSD.	0.0675	6.98E-06	2.38E-05	0.99727	fry	0
12490	singleton	*JAG1*:NM_000214:JAG1 c.2103dupC; p.Cys702LeufsTer5	AS	0.3136	0	0	1	Serrate	14.5
12793	singleton	*CHD7*:NM_017780:c.4480C>T; p.R1494*	AS, BAV, peripheral PS, AP window, RT aortic arch.	0.25703208	0	0	1	kismet	47.3
12795	de nov (trio)	*NF1*:NM_001128147:c.888+1G>T	Sub pulmonic VSD	0.06140851	0	0	0.90175	Nf1	7.6

TOF – Tetralogy of Fallot; AS - Aortic stenosis; BAV - Bicuspid aortic valve; PPS - Peripheral pulmonary stenosis; AVSD - Atrio-ventricular septal defect; PDA - Patent ductus arteriosus, VSD - Ventricular septal defect; PS - Pulmonary stenosis; d-TGA- Dextro-Transposition of great arteries; ASD – Atrial septal defect; s_{het} - genome wide distribution of selection effects for heterozygous protein truncating.