International Journal of Academic Health and Medical Research (IJAHMR)
ISSN: 2643-9824
Vol. 4 Issue 12, December - 2020, Pages: 25-34

To Identify the Congenital Heart Defects and Structural Malformations of the Heart before Birth

Tayyaba Bashir, Hina Younis, Zahra Farhat Malik

Shaikh Khalifa Bin Zayed Al Nahyan Medical & Dental College, Lahore, Pakistan

Abstract: *The presented study identifies rare and common CNV risk factors for non-syndromic obstructive congenital heart defects. It is important to use the most up to date genomic surveillance technology to identify pathogenic structural variants. The presented study utilizes the Illumina Omni5-Exome array which contains ~4.5 million SNP probes which assisted in the discovery of rare and common CNV risk factors with better genomic resolution than previous studies. To assist in the evaluation of rare CNV risk factors, a CNV annotation pipeline for non-syndromic congenital heart defects was assembled. Furthermore, we interrogated the impact of CNVs on known topology associated domains in non-syndromic congenital heart defects and our control population. Lastly, the relationship between multiallelic CNVs and non- syndromic congenital heart defects risk was explored.*

Keywords: Congenital Heart Defects, Structural Malformations, Heart, before Birth

Introduction

Congenital heart defects structural malformations of the heart that develop before birth, are the most common type of birth defect in the world, affecting 0.4-1.3% of live births, leading to over 40,000 new cases a year in the United States alone (Ferencz et al., 1985; Khoshnood et al., 2012; Oyen et al., 2010; Reller, Strickland, Riehle-Colarusso, Mahle, & Correa, 2008; Tanner, Sabrine, & Wren, 2005; Wren et al., 2012; M. H. Wu et al., 2010). There is evidence to suggest a genetic basis for CHDs (Andersen, Troelsen Kde, & Larsen, 2014; Fahed, Gelb, Seidman, & Seidman, 2013; Gelb & Chung, 2014; Pierpont et al., 2007), but known genetic and environmental causes explain less than half of CHD cases (Gelb & Chung, 2014). Population studies investigating the etiological basis of congenital heart defects have demonstrated the prevalence of pathogenic copy number variants (CNVs) in both isolated CHDs (4-14% of cases) and CHDs with extra-cardiac defects (15-20% of cases)(Andersen et al., 2014; Fahed et al., 2013; Marian, 2014).

CNVs are genomic structural variations, a type of variant that alters a *segment* of DNA that is at least 50 bases in size. CNVs are genomic imbalances that increase or decrease the amount of genetic data within a particular genome through duplication or deletion. Copy number variation is ubiquitous in human populations, with recent studies reporting that 4.8-12% of the human genome in healthy individuals is copy number variable (Redon et al., 2006; J. Wu et al., 2012; Yim et al., 2010; Zarrei, MacDonald, Merico, & Scherer, 2015). The genetic data affected by CNVs varies in size and may range from small regulatory elements (50-5000 bases) to large genomic regions (>1000 kilobases) that may encompass *dozens* of genes. Many CNVs have multiple haplotypes that vary in the number of copies. A report by Sudmant and colleagues in *Science* identified CNVs in 159 human subjects with absolute copy number states ranging from 0- 48 copies per genome (Sudmant et al., 2010). It is clear that CNVs are an important component of human diversity and many have been identified as risk factors for disease.

Research has identified many pathogenic CNVs associated with CHDs and other developmental phenotypes, but there remains a need to conduct a large and focused association study of biallelic and multiallelic CNVs in left and right obstructive CHDs. Typically, CNV case-control association studies focus on rare or enriched CNVs in the case population that have a copy number state of 0 (homozygous deletion), 1 (heterozygous deletion) or 3 (heterozygous duplication). This approach has successfully identified many rare pathogenic CNVs but overlooks a segment of CNVs that are the largest source of variation in genomic content between individuals: multiallelic CNVs (Handsaker et al., 2015). Multiallelic CNVs arise through 2 or more duplication events resulting in a copy number state of 4 or more alleles in a single genome and are readily observed in case and control populations (Cantsilieris & White, 2013; Forni et al., 2015; Handsaker et al., 2015; Sudmant et al., 2010). Some multiallelic CNVs show measurable increases in gene expression with *increasing* copy number state (Handsaker et al., 2015). Similar to biallelic CNVs, multiallelic CNVs may increase CHD risk through altered gene expression of the contained alleles with increasing or decreasing copy number state. Therefore, we hypothesize that biallelic and multiallelic copy number variation contribute to the occurrence of isolated left- and right-sided obstructive heart defects.

To address this hypothesis, we will conduct a genome wide association study of Illumina SNP microarray genotyping data from a population of non- syndromic obstructive CHD cases and controls to identify biallelic and multiallelic CNVs that may contribute to disease risk.

Specific Aim 1:

- Identify copy number variants (CNVs) and copy number variable loci that are associated with the risk of right- and left-

sided obstructive congenital heart defects.

- Assemble a CNV calling pipeline to identify existing and to discover novel CNVs among 570 case infants and 828 controls enrolled through the National Birth Defects Prevention Study and genotyped on the Illumina® Infinium Omni5Exome-4 beadchip. Select CNVs and copy number variable loci associated with obstructive CHDs using case-control statistical methods.
- Explore CNVs statistically associated with CHD through analysis of their exonic content, rarity of CNVs, pathway analysis of enriched genes, and evaluation of CNV overlap with other syndromic loci and disease associated mutations.

Specific Aim 2: Identify multiallelic copy number variants (CNVs) that are associated with obstructive congenital heart defects..

- From the population of 570 non-syndromic CHD cases and 828 controls, we will identify potential multiallelic CNVs using a modified CNV calling pipeline.
- Identify multiallelic CNVs associated with obstructive CHD using case- control statistical methods.

Background

Congenital heart defects (CHDs) are the most prevalent birth defects, affecting 0.4-1.7% live births affecting 32-40,000 new cases per year in the United States (Figure 1) (Ferencz et al., 1985; Hoffman & Kaplan, 2002; Khoshnood et al., 2012; Leirgul et al., 2014; Øyen et al., 2009; Reller et al., 2008; Roger et al., 2012; Stephensen et al., 2004; Tanner et al., 2005; van der Linde et al., 2011; Wren et al., 2012; M. H. Wu et al., 2010). There is evidence to suggest that CHDs are influenced by genetic variation (Andersen et al., 2014; Fahed et al., 2013; Gelb & Chung, 2014; Pierpont et al., 2007), but known genetic and environmental causes explain less than half of CHD cases (Gelb & Chung, 2014; Samir Zaidi & Martina Brueckner, 2017). Copy number variants (CNVs) have been associated with congenital heart defect risk for a number of studied CHD phenotypes revealing a high prevalence of pathogenic CNVs in both isolated CHDs (4-14% of case infants) and CHDs with extra-cardiac defects (15-20% of case infants) (Andersen et al., 2014; Fahed et al., 2013; Marian, 2014).

Heart Embryology: The heart is the first functional organ in the developing human embryo. The developing heart undergoes a series of complex changes in the first 7 weeks following fertilization. On day 16 of development, cardiac progenitor cells at the cranial end of the embryo, adjacent to the primitive streak, migrate into the mesoderm to form the primary heart field. This primary heart field begins the process of vasculogenesis leading to the formation of blood islands, which produce the first red blood cells in the embryo. Vasculogenesis occurs in the cardiogenic region and also along the midline of the embryo, which will eventually develop into the heart and the aorta, respectively. On day 18, the heart tube is formed in the cardiogenic region which begins to fold into a primitive cardiac loop. On day 22, this early heart structure begins pulsing. By day 30, the heart will have formed trabeculae in the ventricles and will be poised for the formation of atrial and ventricular septa. From day 30 to day 49, the outflow tract, aortic arches, and septa develop into the mature infant heart (Sadler, 2015). The fast pace and complexity of this process leaves the heart vulnerable to genetic and environmental insults leading to several possible defects at each stage (Bruneau, 2008).

Because of the sequential staging, the precise temporal and spatial signaling involved, and the rapid nature of heart development, there are a number of processes that can go awry with nearly every structure of the heart susceptible to malformation (Figure 2) (Bruneau, 2008). Septation defects are the most common defects which include ventricular, atrial and atrioventricular septal defects (holes between the chambers of the hearts) (Leirgul et al., 2014).

Perhaps lending to their commonality, septal defects can occur through the disruption of two major cardiogenic processes: early differentiation of the primary heart field (days 16-18 of development) and endocardial cushion development (days 26-35 of development)(Sadler, 2015). The secondary heart field progenitor cells contribute to the development of the outflow tract of the heart and the region of the atria nearest these tracts. Therefore, insults impacting the secondary heart field can lead to malformations of the great vessels and outflow defects (Kern et al., 2010; Liang et al., 2014). Complex malformations can also occur that affect multiple heart structures. For example, Tetralogy of Fallot, which originates from an unequal division of the primitive pulmonary artery and aorta (conus cordis), causes several heart malformations including: a large ventricular septal defect, an overriding aorta, pulmonary stenosis and right ventricular hypertrophy. In malformations like Tetralogy of Fallot, there is an altered hemodynamic state (e.g. pulmonary stenosis) in the heart that will lead to some of the cardiovascular manifestations (e.g. right ventricular hypertrophy). These examples, while not exhaustive, illustrate the number of susceptible embryological processes and the range of possible structural malformations possible in the heart. The presented study is investigating obstructive congenital heart defects. These heart defects impact one or both of the outflow tracts of the heart, the aorta and pulmonary artery. Right sided obstructive heart defects included in the study include pulmonary atresia, tricuspid atresia, pulmonary valve stenosis and Tetralogy of Fallot. Left sided obstructive heart defects included in the study include coarctation of the aorta, aortic stenosis, hypoplastic left heart syndrome and interrupted aortic arch.

Epidemiology: Recent studies have estimated that CHDs are present in approximately 0.4-1.7% of all live births (Figure 1.0) (Ferencz et al., 1985; Hoffman & Kaplan, 2002; Khoshnood et al., 2012; Leirgul et al., 2014; Øyen et al., 2009; Reller et al., 2008; Stephensen et al., 2004; Tanner et al.,

A

Author	Location	Population Size (n)	Dates of Surveillance	Measured Prevalence per 1,000 livebirths
Ferencz et al. 1985	United States, Baltimore-Washington area	179,697	1981-1982	3.7
Stephenson et al. 2004	Iceland	44,013	1990-1999	17.0
Tanner et al. 2005	United Kingdom, Northern Health Region of England	521,619	1987-2001	5.7 (95% CI, 5.5-5.9)
Reller et al. 2008	United States, Atlanta, GA	398,140	1998-2005	8.14
Oyen et al. 2009	Denmark	354,308	2000-2005	11.3 (95% CI, 11.0-11.7)
Wu et al. 2010	Taiwan	238,143	2000-2006	13.1 (95% CI, 12.91-13.25)
Wren et al. 2012	United Kingdom, Northern Health Region of England	676,927	1987-2006	6.6 (95% CI, 6.4-6.8)
Khoshnood et al. 2012	France, Paris	317,538	2005-2008	9.0 (95% CI, 8.70-9.36)
Leirgul et al. 2014	Norway	945,413	1994-2009	13.32
Meta-Analysis				
Hoffman and Kaplan, 2002	Worldwide	NA	1946-1999	9.60 (95% CI, 6.02-10.57)
Van der Linde et al. 2011	Worldwide	24,091,867	1930-2009	9.2 (95% CI, 9.0-9.2) *

B

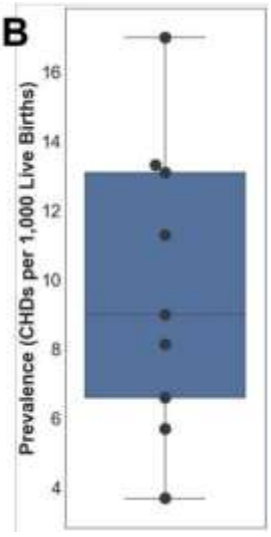

Figure 1.0: Measured congenital heart defect prevalence in 9 studies and 2 meta-analysis reports. (A) Table outlining the study, location, population size, dates of surveillance and the measured prevalence per 1,000 live births. The population reported in this table is reflective of the entire infant population studied and may include still births that were not used for the measurement of the reported prevalence in the table. (B) Boxplot visualizing the spread of measured CHD prevalence in multiple populations. The overall average prevalence of CHD across these multiple studies and meta-analysis reports is 9.76 CHDs per 1,000 live births.

The reported prevalence of the specific obstructive heart defects under investigation in this study range from 0.1 – 3.7 per 10,000 live births (Table 1.0) (Leirgul et al., 2014; Reller et al., 2008). The overall CHD prevalence in the Leirgul et al. and Reller et al. reports was 1.3% and 0.8% of live births, respectively, of which the obstructive heart defects included in our study make up nearly one third of the total reported cases.

Heart defect	**Live birth prevalence**	**Stillbirth I terminated pregnancies prevalence**
Right sided obstructive heart defects:		
Pulmonary atresia	0.4	1.9
Tricuspid atresia	0.5*	
onary valve stenosis	3.7	0.9
Tetrology of Fallot	2.6	14.2
Left sided obstructive heart defects		
Coarctation of the aorta	2.9	14.2
Aortic stenosis	3.0	2.8
astic left heart syndrome	1.6	85.4

Interrupted aortic arch	0.1	N/A**

Table 1.0: Prevalence of obstructive heart defects phenotypes in current study as reported by Leirgul et al. 2014. Prevalence reported per 10,000 births.
*Tricupsid atresia values not reported in the Leirgul et al. study, so values from Reller et al. 2008 reported.
**No incidence of interrupted aortic arch observed in the stillbirth/terminated pregnancy population (n = 10,542) during the study window (1994-2009).

Estimates of the prevalence of congenital heart defects of live births have been subject to a remarkable variability and warrants some discussion. In 2002, Hoffman and Kaplan sought to determine the causes for this variability and reviewed 62 reports of the CHD prevalence in live births and reported a range of 4 to 50 CHDs per 1,000 live births (Hoffman & Kaplan, 2002). Prevalence estimates in earlier studies were lower than estimates in more recent studies that utilized echocardiograms for the surveillance of CHD in the study population. The highest estimated prevalence of was reported in a study by Roguin and colleagues in which color echocardiography was performed on 1,053 consecutive neonates born to the Western Galilee Hospital-Nahariya, Israel. The population under investigation was known to have high rates of consanguinity, which initially prompted the study. Ventricular septal defects were found in 56 of the 1,053 infants in the study or an estimated prevalence of 53.2 per 1,000 live births in the Hiraishi study (Hiraishi et al., 1992; Roguin et al., 1995). Similarly, Ishikawa and colleagues in Japan performed two-dimensional color echocardiography on a population of 2,067 consecutive live births and identified 104 infants with a CHD (50.3 per 1,000 live births), approximately 5 times the average reported prevalence of CHDs (Figure 1.0) (Ishikawa, Iwashima, Ohishi, Nakagawa, & Ohzeki, 2011). Both the Roguin study and the Ishikawa study reported a large number of asymptomatic and mild CHDs. The tendency for mild phenotypes of CHDs like VSDs and atrial septal defects (ASDs) to be asymptomatic or asymptomatic at birth contributes to the variability in reported CHD birth prevalence measurements. This variability in CHD prevalence often begins to decrease if we consider symptomatic or moderate to severe CHDs (ex: mild to moderate aortic stenosis, moderate pulmonary stenosis, large ASD, acyanotic lesions). These moderate to severe CHDs are less likely to be missed in the post-natal population resulting in stabilization of prevalence variability with most researchers reporting ~1.5 cases per 1,000 live births (Hoffman & Kaplan, 2002; Pradat, Francannet, Harris, & Robert, 2003; van der Bom et al., 2011; Wren & O'Sullivan, 2001; Wren, Reinhardt, & Khawaja, 2008). Taken together, the prevalence of various CHDs depends on the screening mechanism, the demographics of the population investigated and severity of the CHDs.

A higher prevalence of congenital heart defects is observed in spontaneously aborted children, miscarriages, and infants born prematurely. Congenital birth defects typically arise during the first trimester, during organogenesis, and can negatively impact the fetus for the remainder of the pregnancy predisposing to adverse pregnancy outcomes. Insight into the prevalence of CHDs in live births and stillbirths is offered by Leirgul and colleagues who, in 2014, reported the prevalence of CHDs among 943,871 live births and 10,542 stillbirths / electively terminated infants in a population of 954,413 births. A total of 12,577 live births with CHDs (133.2 per 10,000) and 504 stillbirths / terminated pregnancies (478.1 per 10,000) with CHDs were reported. Overall, the prevalence of CHDs among stillbirths and terminated pregnancies in this report was 3.6 times greater than the live birth population. Several CHD phenotypes occurred at an increased rates in the stillbirth/aborted pregnancies (prevalence estimates reported per 10,000 live births / still births and terminated pregnancies): ventricular septal defects (VSD)(47.7 / 108.1), hypoplastic left heart syndrome (HLHS)(1.6 / 85.4), transposition of the great arteries (TGA) (3.3 / 21.8), double outlet right ventricle (DORV) (0.7 / 14.2), Ebstein anomaly (0.6 / 6.6), hypoplastic right heart syndrome (HRHS) (0.7 / 15.2), Tetralogy of Fallot (2.6 / 14.2) and coarctation of the aorta (CoA) (2.9 / 14.2)

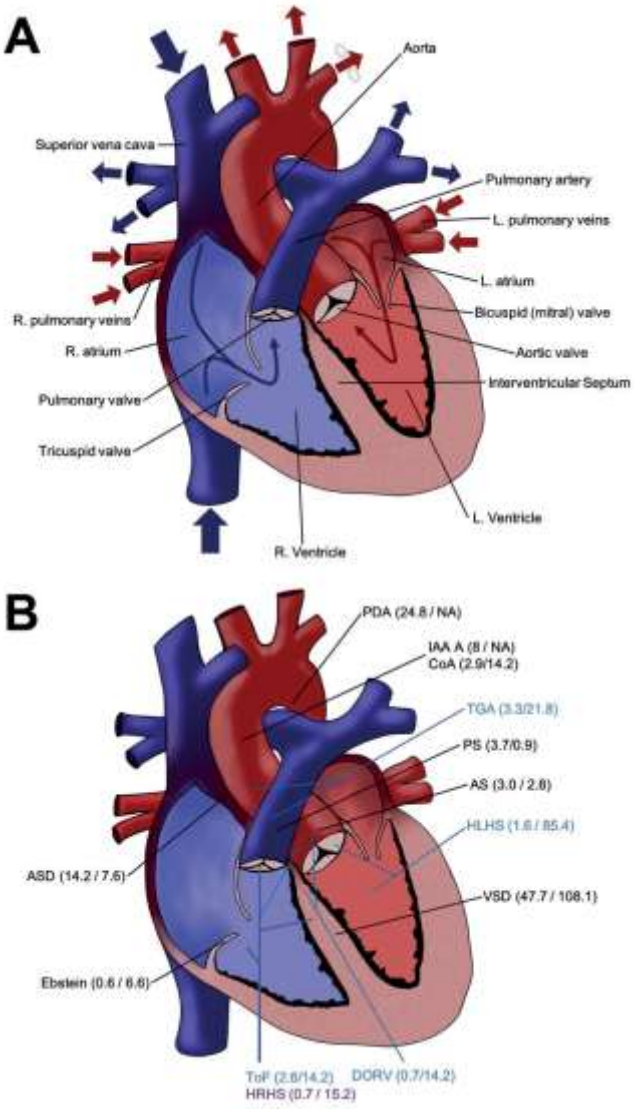

Figure 1.1: Overview of cardiac anatomy and structures impacted by congenital heart defects with live and still birth prevalence per 10,000 infants as reported by Leirgul et al. 2014. (A) Cardiac anatomy with structures carrying deoxygenated blood labeled in blue and structures carrying oxygenated blood labeled in red. Left (L.), right (R.) (B) Heart structures impacted by various congenital heart defects with the prevalence reported next to each in live births (n = 943,871) and still births/terminated pregnancies for medical reasons (n = 10,542) per 10,000 (Leirgul et al. 2014). The congenital heart defects listed are atrial septal defect (ASD), Ebstein anomaly, Tetralogy of Fallot (ToF), hypoplastic right heart syndrome (HRHS), double outlet right ventricle (DORV), ventricular septal defect (VSD), hypoplastic left heart syndrome (HLHS), aortic valve stenosis (AS), pulmonary valve stenosis (PS), transposition of the great arteries (TGA), interrupted aortic arch type A (IAA A), coarctation of the aorta (CoA) and patent ductus arteriosus (PDA). Some of the CHDs that impact multiple structures are labeled in light blue. Conditions that were not observed in the study population are marked with an NA. For clarity, HRHS which affects many of the same structures impacted by ToF, is labeled in purple. Illustrated by Joseph Levy. Figure idea adapted from Bruneau et al. 2008.

Birth defects and CHDs have remained a leading cause of neonatal and infant death in the United States despite the significant medical innovation over the past half century in both the diagnosis and surgical management of these defects (Petrini et al., 2002). Improvements in surgical repair has improved early survival rates and extended life expectancy ages of affected patients (Greutmann & Tobler, 2012; Marelli, Mackie, Ionescu-Ittu, Rahme, & Pilote, 2007; Tennant, Pearce, Bythell, & Rankin, 2010; van der Bom et al., 2011). Outcomes have improved substantially in some complex CHDs like Tetralogy of Fallot,

transposition of the great vessels and hypoplastic left heart syndrome where dramatic surgical improvements have been made over the past 3 decades (Greutmann et al., 2015). Surgical management of hypoplastic left heart syndrome (HLHS), for example, has been improved through development and maturation of techniques like the Norwood operation (Karamlou, Diggs, Ungerleider, & Welke, 2010), which has resulted in an estimated 70% of newborn patients with HLHS now reaching adulthood (Feinstein et al., 2012). The success of these surgeries is often complicated by genetic co-morbidities and mutations, such as SNPs in VEGFA or SOD2 (Kim et al., 2014). Continued genetic discoveries in CHD may help to advise physicians in the care of CHD patients, potentially extending the trend of progress in the management of CHDs.

From 1985 to 2000, it was reported that the prevalence of congenital heart defects has increased in both pediatric (+5.01 children with CHD per 1,000 population) and adult populations (+0.52 adults with CHD per 1,000 population) (Marelli et al., 2007). This shift in the overall population was most dramatic when considering severe CHD adults which more than doubled between 1985 and 2000 in the study cohort (Marelli et al., 2007). This increase in the prevalence of CHDs is largely due to the improvements in surgical and catheter-based managements made over the last 4 decades which have resulted in a growing adult population of CHD patients. In fact, a *NEJM* review article covering medical progress in CHD management in 2000 reported that the adult congenital heart defect population is expected to increase by ~5% each year (Brickner, Hillis, & Lange, 2000). This increase in the adult population of CHD has led to a greater need to understand the long-term consequences of these defects, reproductive fitness and the proper management of the normal conditions of aging in CHD cases. Understanding the complex genetic basis of CHD is essential to properly care for this growing patient population.

Genetic basis of CHD: Dr. Bruce Gelb published an article in 2015 outlining the history of our understanding the causes of CHD in which he described three eras of our understanding and progress in the genetic basis of CHDs: 1) preinterventional era, 2) interventional era, and 3) the molecular biology era (Gelb, 2015). The preinterventional era of CHD begins in the mid 1800s and ends with the first patent ductus arteriosus ligation in 1939 (Gross & Hubbard, 1984). The interventional era continues through the mid 19^{th} century and ends in 1986 with the first genetic clone of a mutated protein without reference to the specific protein, but by reliance on its genomic position (Royer- Pokora et al., 1986). During these periods, the genetic basis of CHD was hypothesized, garnered epidemiological and experimental support and is currently maturing through the support of modern genomic technology.

The genetic etiology of CHDs was first alluded to in 1858 by Thomas Peacock, MD, who suggested a *hereditary predisposition* to the defective development of the heart in his pioneering work, *On Malformations of the Human Heart* (Gelb, 2015; Peacock, 1858). This suggestion of a hereditary predisposition, like many of the initial observations alluding to a genetic basis for CHD, was based on the observation that CHD phenotypes had a tendency to aggregate within families. Other observations of familial recurrence and aggregation of CHD occurred during the preinterventional and interventional era by physicians like Maud Abbott, Maurice Campbell, and James Nora (Gelb, 2015). By the mid-twentieth century, our understanding of genetic mechanisms had advanced and the structure of DNA had been described. It was during this time that Dr. Helen Taussig, an American born cardiologist who would later found the field of pediatric cardiology, first suggested that CHD was driven by genetic defects (Taussig, 1965). Over the next several decades, the genetic basis of CHD was explored leading to the discovery of numerous genetic variants that increase the risk of CHD, including; aneuploidies, single nucleotide polymorphisms, and copy number variants (Fahed et al., 2013).

Family based studies in CHD: Population-based and familial aggregation studies are some of the earliest investigative tools that alluded to the genetic basis of CHD and can act as a foundation for the genetic basis of CHD risk (Loffredo et al., 2004; Oyen et al., 2009; Oyen et al., 2010). Perhaps the most substantial research analyzing the recurrence rates of CHDs in families originated out of a 2009 study that utilized Denmark's extensive resident registry and database which included >1.7 million persons and 18,708 patients with CHDs. Overall, the relative risk ratio (RRR) of any CHD in first, second, or third- degree relatives to a CHD patient, excluding patent ductus arteriosus, was calculated to be 3.5 (95% CI 3.2-3.8), 1.4 (95% CI 1.3-1.5), and 1.2 (95% CI 1.1-1.3), respectively. Familial clustering of CHDs is further affirmed when considering the recurrence rate of the same heart defect with estimated RRRs that ranged from 3.4 (95% CI 2.7-4.3) in isolated septal defects to *79.1* (95% CI 32.9-190.0) for heterotaxy with an overall recurrence risk ratio of 8.2 (95% CI 7.0- 9.6) for the same heart defect (Oyen et al., 2009). A more recent study from this group expanded on these findings showing that the recurrence of the same CHD in the younger sibling (full sibling) had a RRR of 5.6 (95% CI 4.6-6.9) with some specific CHD phenotypes having much higher RRRs including 53.2 (95% CI 7.5-

378) for heterotaxy, 117.8 (95% CI 37.7-368) for atrioventricular septal defects and 251.7 (95% CI 62.4-1014) for anomalous pulmonary venous return (Brodwall et al., 2017). The relative risk ratio variability reported by these two studies across multiple CHD phenotypes demonstrate the complex genetics underpinning CHD risk.

Twin studies further affirm to the complex inheritance of CHD. In 2005, Caputo et al. studied a population of 1,743 CHD patients with at least 1 sibling, and 66 pairs of dizygotic twins (Caputo et al., 2005). In this study, they found that twins had a much higher likelihood of recurrence with 9/66 dizygotic twins (13.6%) both having identified CHDs, compared to only 67 / 1743 non-twin siblings (3.8%) containing CHDs. The results are interesting but are limited by the small study populations and the need for CHD phenotype stratification. A recently published article by Brodwall et al. from the Øyen lab aimed to study CHD recurrence rates among 719,714 births in Norway during the years 1994- 2009 (Brodwall et al., 2017). Brodwall and colleagues

found that, among 297 twins with at least one CHD, 66 of the 297 (22.2%) twins had CHDs in both twins. Interestingly, they found that same-sex twins had an increased rate of recurrence with 51/207 same sex twins (24.6%, adjusted RRR 14.0, 95% CI 10.6-18.6) and
15/90 opposite sex twins (16.7%, adjusted RRR 11.9, 95% CI 7.1-19.9) having CHDs in both twins.

Copy number variants and CHD Risk: Copy number variants (CNVs) are unbalanced rearrangements of DNA that can increase or decrease the amount of genetic code within a particular genome, through either duplication or deletion. The genetic data affected by CNVs can be anything from small regulatory elements to dozens of genes, ranging from 50 bases to megabases (>1,000,000 bases) in size (Feuk, Carson, & Scherer, 2006; Zarrei et al., 2015). CNVs are part of a larger group of genomic variants called *structural variants* and represent the most frequent type of structural variation that occurs in the human genome (Feuk et al., 2006). Study of these heterogeneous and common variants is crucial for the understanding of the mechanisms of complex genetic disease.

The human genome project revealed several mechanisms of genetic variation in the human genome including single nucleotide polymorphism (SNP), transposable elements, GC content variability, CpG islands and recombination (Lander et al., 2001). Following the success of the human genome project, the International HapMap project began in 2002 to identify common patterns of human genetic variation leading to a catalogue of SNPs common to human populations (Deloukas & Bentley, 2004). Using the newly mapped genome and identified SNPs, the first genome wide single nucleotide polymorphism and comparative genomic hybridization arrays were developed which allowed for whole genome surveys in populations of patients. Utilization of these genome- wide arrays lead to the first reports of large copy-number variations across the human genome contributing to genetic diversity (Iafrate et al., 2004; Sebat et al., 2004). Since these first reports of large-scale copy number variation (Iafrate et al., 2004; Sebat et al., 2004), large scale case-control studies have been conducted to understand the impact of copy number variation on human physiological diversity and disease pathogenesis. For example, in an effort to create a CNV map of the human genome, Zarrei et al. selected 2,057,368 high- confidence CNVs reported on the Database of Genomic Variants to determine the extent of the healthy human genome that is copy number variable. They concluded that between 4.8% and 9.5% of the human genome is copy number variable and identified 99-107 protein coding genes that can be deleted in apparently healthy individuals (Zarrei et al., 2015). Diseases that are highly associated with copy number variants include structural birth defects (Cooper et al., 2011; Southard, Edelmann, & Gelb, 2012), autism spectrum disorders (Abrahams & Geschwind, 2008; Cooper et al., 2011; Gulsuner & McClellan, 2015; Kirov, 2015; Pinto et al., 2010; Ronemus, Iossifov, Levy, & Wigler, 2014; Sebat et al., 2007), cancers (Manier et al., 2016; Mitchell & Neal, 2015; H. Wang, Liang, Fang, & Xu, 2016; Zucman-Rossi, Villanueva, Nault, & Llovet, 2015), schizophrenia (Gulsuner & McClellan, 2015; Kirov, 2015), intellectual disability (Cooper et al., 2011), and many others (Weischenfeldt, Symmons, Spitz, & Korbel, 2013). CNVs have also been discovered to play both protective (de Smith, Trewick, & Blakemore, 2010) and destructive roles in developmental biology (Cooper et al., 2011).

Aneuploidy syndromes are a type of structural variation where, instead of only encompassing a small section of the genome, like copy number variants, they result from the duplication or deletion of an entire chromosome. Common aneuploidy syndromes all have significant risk of congenital heart defects.

Trisomy 21, the most common aneuploidy, occurs in 1.1-1.3 of every 1000 live births (Hassold et al., 1996; Morris & Alberman, 2009; Stoll, Alembik, Dott, & Roth, 1990). Approximately half of trisomy 21 patients have a CHD with atrioventricular canal defects, ventricular septal defects and atrial septal defects being the most common CHD in this population (Roizen et al., 2014; Tan et al., 2013). The detrimental effects of these aneuploidies as well as pathogenic CNVs cause pathologic effect through alteration of genetic dosage. Due to the size of the variant and the variation in phenotypic features observed in aneuploidy syndromes, determining the exact cause of the heart defect related outcomes has been difficult, but a number of causative loci within and outside the affected chromosome primarily effected by the aneuploidy are under investigation.

The genetic and environmental basis of CHDs is complex; many causative mutations have been identified, and, among these, CNVs have emerged as a common disease-driving variant. In a hallmark review of CHD genetics, Fahed and colleagues noted that variations in gene dosage play a principle role in CHD etiology (Fahed et al., 2013). Gene dosage is a term regularly used in CNV literature and refers to the quantitative gene transcription of the contained exonic content. Copy number variants can change gene dosage by physically altering the number of gene alleles present in a genome. The observation by Fahed et al. that gene dosage is a theme in the genetic basis of CHD is supported by reproducible associations between CHDs and microdeletions, microduplications, trisomies, monosomies, and transcription factor mutations in genes such as NXK2.5. Recent studies identifying pathogenic CNVs in Tetralogy of Fallot (Greenway et al., 2009; Silversides et al., 2012), heterotaxy (Rigler et al., 2015), pulmonary atresia (Xie et al., 2014), coarctation of the aorta (Moosmann et al., 2015), hypoplastic left heart syndrome (Glidewell et al., 2015; Payne, Chang, Koenig, Zinn, & Garg, 2012; Warburton et al., 2014), ventricular septal defects (An et al., 2016; J. Wang et al., 2011), grouped CHD phenotypes (Glessner et al., 2014; Soemedi et al., 2012; White et al., 2014) provide further evidence for the association between pathogenic CNVs and non-syndromic CHD risk. Taken together, Zaidi and colleagues estimate that pathogenic CNVs are identified in
~10% of patients with non-syndromic CHD (S. Zaidi & M. Brueckner, 2017).

As previously mentioned, a common mechanism of disease risk attributed to CNVs is by alteration of the *gene dosage* (Beckmann, Estivill, & Antonarakis, 2007; Fahed et al., 2013). Some genomic loci are sensitive to copy number change where duplication or deletion of the contained alleles corresponds to increases or decreases in gene expression (Handsaker et al., 2015; Ruderfer et al., 2016). Though most studies focus on single deletion and duplication states, many CNVs can exist in states with multiple duplications of a particular locus that results in >3 segregating alleles. Theses CNVS with > 3 segregating alleles are classified as multiallelic copy number variants (Usher & McCarroll, 2015).

Handsaker et al. recently demonstrated that these multiallelic copy number variants alter gene dosage of the included genes with higher copy number states leading to further increased gene expression directly related to the number of copies. While multiallelic CNVs were identified as a minority of the total CNVs observed in this population (~15%), they contributed to >85% of the gene copy variation between two individuals in the 1000 Genomes Project population (Handsaker et al., 2015). Multiallelic CNVs have not been the focus of a study in congenital defects population and their relationship to disease phenotypes is poorly characterized (Usher & McCarroll, 2015).

National Birth Defects Prevention Study: The National Birth Defects Prevention Study (NBDPS) is a large scale case-control study that sought to study genetic and environmental factors associated with non- syndromic birth defects in the United States (Yoon et al., 2001). Funding for the NBDPS was obtained through the Birth Defects Prevention Act of 1998 (Public Law 105-168) which established Centers of Excellence for Birth Defects Research and Prevention. The Centers for Disease Control and Prevention administered the funds of the Birth Defects Prevention Act. A funding announcement was released requesting proposals from institutions that had population-based birth defect registries. Arkansas was able to participate due the Arkansas Reproductive Health Monitoring System, which was founded in 1980 and is one of the oldest birth defect surveillance systems in the United States. Participating centers were expected to contribute at least 300 cases per year to the study and had to surveil approximately 35,000 births per year. Centers that participated in this large national study included Arkansas, California, Georgia, Iowa, Massachusetts, North Carolina, New Jersey, New York, Texas and Utah (Reefhuis et al., 2015). Eligibility for enrollment in the NBDPS began with pregnancies that ended on October 1st, 1997 and concluded with pregnancies that ended on December 31, 2011. The last maternal interviews were completed by December 31, 2013. During this more than 15 year active study period, the participating CBDRPs completed over 44,000 maternal interviews and collected over 69,000 DNA samples (Reefhuis et al., 2015). The proposed study uses samples acquired through the NBDPS.

Conclusion

The presented study identifies rare and common CNV risk factors for non-syndromic obstructive CHDs. It is important to use the most up to date genomic surveillance technology to identify pathogenic structural variants. The presented study utilizes the Illumina Omni5-Exome array which contains ~4.5 million SNP probes which assisted in the discovery of rare and common CNV risk factors with better genomic resolution than previous studies. To assist in the evaluation of rare CNV risk factors, a CNV annotation pipeline for non-syndromic CHD was assembled. Furthermore, we interrogated the impact of CNVs on known topology associated domains in non-syndromic CHD and our control population. Lastly, the relationship between multiallelic CNVs and non- syndromic CHD risk was explored.

References

Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet, 9*(5), 341-355. doi:10.1038/nrg2346

Andersen, T. A., Troelsen Kde, L., & Larsen, L. A. (2014). Of mice and men: molecular genetics of congenital heart disease. *Cell Mol Life Sci, 71*(8), 1327-1352. doi:10.1007/s00018-013-1430-1

Beckmann, J. S., Estivill, X., & Antonarakis, S. E. (2007). Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet, 8*(8), 639-646. doi:10.1038/nrg2149

Brickner, M. E., Hillis, L. D., & Lange, R. A. (2000). Congenital heart disease in adults. First of two parts. *N Engl J Med, 342*(4), 256-263. doi:10.1056/NEJM200001273420407

Gelb, B. D. (2015). History of Our Understanding of the Causes of Congenital Heart Disease. *Circ Cardiovasc Genet, 8*(3), 529-536. doi:10.1161/circgenetics.115.001058

Gelb, B. D., & Chung, W. K. (2014). Complex genetics and the etiology of human congenital heart disease. *Cold Spring Harb Perspect Med, 4*(7), a013953. doi:10.1101/cshperspect.a013953

Glessner, J. T., Bick, A. G., Ito, K., Homsy, J. G., Rodriguez-Murillo, L., Fromer, M., . . . Chung, W. K. (2014). 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, 115*(10), 884-896. doi:10.1161/CIRCRESAHA.115.304458

Glidewell, S. C., Miyamoto, S. D., Grossfeld, P. D., Clouthier, D. E., Coldren, C. D., Stearman, R. S., & Geraci, M. W. (2015). Transcriptional Impact of Rare and Private Copy Number Variants in Hypoplastic Left Heart Syndrome. *Clin Transl Sci, 8*(6), 682-689. doi:10.1111/cts.12340

Greenway, S. C., Pereira, A. C., Lin, J. C., DePalma, S. R., Israel, S. J., Mesquita, S. M., . . . Seidman, C. E. (2009). De novo copy number

Greutmann, M., & Tobler, D. (2012). Changing epidemiology and mortality in adult congenital heart disease: looking into the future. *Future Cardiol, 8*(2), 171-177. doi:10.2217/fca.12.6

Greutmann, M., Tobler, D., Kovacs, A. H., Greutmann-Yantiri, M., Haile, S. R., Held, L., . . . Colman, J. M. (2015). Increasing mortality burden among adults with complex congenital heart disease. *Congenit Heart Dis, 10*(2), 117-127. doi:10.1111/chd.12201

Karamlou, T., Diggs, B. S., Ungerleider, R. M., & Welke, K. F. (2010). Evolution of treatment options and outcomes for hypoplastic left heart syndrome over an 18-year period. *J Thorac Cardiovasc Surg, 139*(1), 119-126; discussion 126-117. doi:10.1016/j.jtcvs.2009.04.061

Kim, D. S., Kim, J. H., Burt, A. A., Crosslin, D. R., Burnham, N., McDonald- McGinn, D. M., . . . Jarvik, G. P. (2014). Patient genotypes impact survival after surgery for isolated congenital heart disease. *Ann Thorac Surg, 98*(1), 104-110; discussion 110-101. doi:10.1016/j.athoracsur.2014.03.017

Morris, J. K., & Alberman, E. (2009). Trends in Down's syndrome live births and antenatal diagnoses in England and Wales from 1989 to 2008: analysis of data from the National Down Syndrome Cytogenetic Register. *BMJ, 339*, b3794. doi:10.1136/bmj.b3794

Oyen, N., Poulsen, G., Boyd, H. A., Wohlfahrt, J., Jensen, P. K., & Melbye, M. (2009). Recurrence of congenital heart defects in families. *Circulation, 120*(4), 295-301. doi:10.1161/CIRCULATIONAHA.109.857987

Øyen, N., Poulsen, G., Boyd, H. A., Wohlfahrt, J., Jensen, P. K. A., & Melbye, M. (2009). National time trends in congenital heart defects, Denmark, 1977- 2005. *Am Heart J, 157*(3), 467-473.e461. doi:10.1016/j.ahj.2008.10.017

Oyen, N., Poulsen, G., Wohlfahrt, J., Boyd, H. A., Jensen, P. K., & Melbye, M. (2010). Recurrence of discordant congenital heart defects in families. *Circ Cardiovasc Genet, 3*(2), 122-128. doi:10.1161/CIRCGENETICS.109.890103

Payne, A. R., Chang, S. W., Koenig, S. N., Zinn, A. R., & Garg, V. (2012). Submicroscopic chromosomal copy number variations identified in children with hypoplastic left heart syndrome. *Pediatr Cardiol, 33*(5), 757- 763. doi:10.1007/s00246-012-0208-9

Roguin, N., Du, Z. D., Barak, M., Nasser, N., Hershkowitz, S., & Milgram, E. (1995). High prevalence of muscular ventricular septal defect in neonates. *J Am Coll Cardiol, 26*(6), 1545-1548. doi:10.1016/0735-1097(95)00358-4

Roizen, N. J., Magyar, C. I., Kuschner, E. S., Sulkes, S. B., Druschel, C., van Wijngaarden, E., . . . Hyman, S. L. (2014). A community cross-sectional survey of medical problems in 440 children with Down syndrome in New York State. *J Pediatr, 164*(4), 871-875. doi:10.1016/j.jpeds.2013.11.032

Ronemus, M., Iossifov, I., Levy, D., & Wigler, M. (2014). The role of de novo mutations in the genetics of autism spectrum disorders. *Nat Rev Genet, 15*(2), 133-141. doi:10.1038/nrg3585

Royer-Pokora, B., Kunkel, L. M., Monaco, A. P., Goff, S. C., Newburger, P. E., Baehner, R. L., . . . Orkin, S. H. (1986).

Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location. *Nature, 322*(6074), 32-38. doi:10.1038/322032a0

Ruderfer, D. M., Hamamsy, T., Lek, M., Karczewski, K. J., Kavanagh, D., Samocha, K. E., . . . Purcell, S. M. (2016). Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. *Nat Genet, 48*(10), 1107-1111. doi:10.1038/ng.3638

Stoll, C., Alembik, Y., Dott, B., & Roth, M. P. (1990). Epidemiology of Down syndrome in 118,265 consecutive births. *Am J Med Genet Suppl, 7*, 79-83. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2149980

Tanner, K., Sabrine, N., & Wren, C. (2005). Cardiovascular malformations among preterm infants. *Pediatrics, 116*(6), e833-838. doi:10.1542/peds.2005- 0397

Taussig, H. B. (1965). On the Evolution of Our Knowledge of Congenital Malformations of the Heart; the T. Duckett Jones Memorial Lecture. *Circulation, 31*, 768-777. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14278070

Warburton, D., Ronemus, M., Kline, J., Jobanputra, V., Williams, I., Anyane- Yeboa, K., . . . Levy, D. (2014). The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum Genet, 133*(1), 11-27. doi:10.1007/s00439-013-1353- 9

Weischenfeldt, J., Symmons, O., Spitz, F., & Korbel, J. O. (2013). Phenotypic impact of genomic structural variation: insights from and for human disease. *Nat Rev Genet, 14*(2), 125-138. doi:10.1038/nrg3373

Yoon, P. W., Rasmussen, S. A., Lynberg, M. C., Moore, C. A., Anderka, M., Carmichael, S. L., . . . Edmonds, L. D. (2001). The National Birth Defects Prevention Study. *Public Health Rep, 116 Suppl 1*, 32-40. doi:10.1093/phr/116.S1.32

Zaidi, S., & Brueckner, M. (2017). Genetics and Genomics of Congenital Heart Disease. *Circ Res, 120*(6), 923-940. doi:10.1161/CIRCRESAHA.116.309140

Zaidi, S., & Brueckner, M. (2017). Genetics and Genomics of Congenital Heart Disease. *Circ Res, 120*(6), 923-940. doi:10.1161/circresaha.116.309140

Zarrei, M., MacDonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat Rev Genet, 16*(3), 172- 183. doi:10.1038/nrg3871

Zucman-Rossi, J., Villanueva, A., Nault, J. C., & Llovet, J. M. (2015). Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology, 149*(5), 1226-1239 e1224. doi:10.1053/j.gastro.2015.05.061