Discovering Pathogenic Variants Associated with Tricuspid Atresia through Whole Exome Sequencing

By

Jason Andrew Hoggard

A Dissertation submitted to the Department of Biology and Biochemistry,

College of Natural Sciences and Mathematics

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

In Biochemistry

Chair of Committee: Preethi Gunaratne, PhD

Co-Chair of Committee: Robert Schwartz, PhD

Committee Member: Lisa D'Alessandro, M.D.

Committee Member: Shaine Morris, M.D.

Committee Member: Yu Liu, PhD

Committee Member: Bradley McConnell, PhD

University of Houston

December 2019

Dedication/Epigraph

I would like to thank my family for their support. Without their willingness to support me, and participate, this certainly would not have been possible. They recognized the significance of defining the genetic cause of my own condition which has caused myself and my family so many long days and nights, not only at the hospital but during my recovery at every step, not only to me but to others with tricuspid atresia. My family has always been willing to go to any length to keep me going and provide every bit of support that I needed to get past every obstacle that has come along during my life, especially when it involved my heart.

However, none of this would have been possible without Holly. She provided support and encouragement at every step of the way, even before I began my journey through graduate school. She was there to pick me up when I thought I could not go any further and when I thought I hit a dead-end. Holly always listened with a smile on her face when I was going through various difficulties I had during graduate school, whether it was when I was trying to understand my data or when I was feeling burnt-out and having a bad day. When I said I couldn't do something, she always told me "yes, you can" no matter how big the challenge was and she always believed in me.

Acknowledgements

This dissertation would not be possible without the countless contributions and support from so many individuals. My co-mentors, Dr. Preethi Gunaratne and Dr. Robert Schwartz, for taking a chance on me and this project. Dr. Lisa D'Alessandro and Dr. Shaine Morris for welcoming me into the cardiology department at Texas Children's Hospital, providing access to all the necessary records, and allowing me to attend so many lectures that have greatly benefited my growth as a student and a scientist. Dr. Yu Liu and Dr. Bradley McConnell for always providing guidance and much-needed insight at every turn. The members of the Gunaratne lab for being there, whether it be just to help brainstorm a solution to a problem I'm running into in my project or just being friends.

There are several individuals I worked with during these last few years that have helped me grow, not only as a person but also as a mentor. Ivy Bui and Himadri Gunarathna worked with me in various capacities, which allowed me to grow as a mentor. However, most of all, Dr. Dhanushya Amaratunga assisted me with countless hours of literature searches, repeatedly looking through the variant lists for any possible information that may have been overlooked, and working on the patient chart documentation for hours on end. Her assistance greatly sped up my progress and helped me identify key pieces of data that I would have otherwise overlooked. Her contribution cannot be overstated, as she was a vital part of completing my dissertation.

Abstract

Congenital heart disease (CHD) is the most common birth defect, present in 1/110 live births, and those considered critical (CCHD) require surgical intervention within the first year of life. A rare form of CCHD, tricuspid atresia, is present in 1/10,000 live births in the United States and accounts for approximately 1-3% of all CHD. It is characterized by the absence of the tricuspid, or right atrioventricular, valve and presents with additional phenotypes that are required to survive to birth. At this point, very few genetic studies have been conducted on this condition and the results have been very sparse. Currently 22q11 deletion (DiGeorge syndrome), 8p23 (GATA4 region), 4q31 (NFκB), and 3p (TGFBR2) have been found associated with the few tricuspid atresia patients that have been characterized through limited genetic testing.

In this study, a retrospective chart review was undertaken on the largest cohort of tricuspid atresia patients (n=234) and includes the first genetic testing outcome results for any tricuspid atresia retrospective review study. Following this, a family with various cardiac phenotypes including tricuspid atresia and bicuspid aortic valve was assessed via whole exome sequencing (WES) to discover pathogenic variants. Following the compilation of all genetic testing data from the literature, the retrospective review, and the family WES, a common pathway was identified that is disrupted in all subjects without a syndromic diagnosis. The pathway, beginning with TGF-β and RANKL signaling, involves the expression of *NFATC1* via NFκB activity and NFATC1 transcription factor function regulation by a complex including TAB2. WES in 342 patients with congenital cardiac left-sided lesions revealed extensive genetic heterogeneity. This is the only other study to screen a large cohort of patients with WES

and reported 28 candidate variants in 27 genes. Of these, 17 genes were not previously associated with CHD. Our study is the first to begin identifying a potential genetic etiology for tricuspid atresia which is a right-sided lesion.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Genes
GATA4 – GATA Binding Protein 4
IKKβ – Inhibitor Of Nuclear Factor Kappa B Kinase Subunit Beta
MYH6/7 – Myosin Heavy Chain 6/7
NFATC1 – Nuclear Factor Of Activated T Cells 1
NFκB – Nuclear Factor Kappa B
NKX2.5 – NK2 Homeobox 5
PWWP2B – PWWP Domain Containing 2B
RANK – Receptor Activator Of Nuclear Factor Kappa B
RANKL – Receptor Activator Of Nuclear Factor Kappa B Ligand
TAB2 – TGF-Beta Activated Kinase 1 (MAP3K7) Binding Protein 2
TAK1 – TGF-Beta Activated Kinase 1 (MAP3K7)
TGF-β – Transforming Growth Factor Beta
TGFBR1/2 – Transforming Growth Factor Beta Receptor 1/2
TRAF6 – TNF Receptor Associated Factor 6

Genetic Testing
CMA – Chromosomal microarray
FISH – Fluorescence *in situ* hybridization
MAF – Minor allele frequency
WES – Whole exome sequencing
WGS – Whole genome sequencing

Cardiac Phenotypes
ASD – Atrial septal defect
BAV – Bicuspid aortic valve
IVC – Inferior vena cava
LA – Left atrium
LV – Left ventricle
OFT – Outflow tract
PDA – Patent ductus arteriosus
PFO – Patent foramen ovale
RA – Right atrium
RV – Right ventricle
SVC – Superior vena cava
TGA – Transposition of the great arteries
VSD – Ventricular septal defect

Chapter 1: Introduction

Congenital heart disease (CHD) is a structural or functional abnormality of the heart present from birth. CHD is the leading cause of birth defect-related illness and death (*1*) and occurs in approximately 1/110 live births (*2*). However, 25% of these cases are considered critical (CCHD), such as tricuspid atresia. CCHD requires intervention within 1 year from birth to increase the chance of survival. While survival of CCHD varies by region, it is approximately 75% to 1 year and 69% to age 18 in the United States (*1, 3*).

Chromosomal defects and single-gene disorders can cause CHD (i.e., DiGeorge, Marfan's, Edwards, Loeys-Dietz Syndrome), but represent less than 20% of CHD (*4*). While there are some conditions for which genetic etiologies have become much clearer, like syndromic cases such as DiGeorge (chromosome 11q22) or Turner's syndrome (45X), many complex non-syndromic lesions still have unknown genetic etiologies. What makes defining genetic association with a specific phenotype more difficult is that CHD appears to be pleiotropic, meaning a variant in a single gene can influence a spectrum of phenotypes. This obviously presents a problem when trying to define a genetic etiology of specific phenotypes or genotypes within a population as multiple forms of CHD within a family can be the result of the same variant within a single gene (*5-8*). As such, many conditions still have unknown genetic etiologies.

Tricuspid atresia

Tricuspid atresia is a form of CHD in which the right atrioventricular (AV), or tricuspid, valve, which allows passage of blood from the right atrium to the right ventricle is absent (Figure 1). It is a rare form of CHD, occurring in roughly 1/10,000 live births in the United States and accounting for 1-3% of all CHD (*9*). The risk of recurrence of tricuspid atresia in siblings is low at approximately 1% (*10, 11*).

As blood enters the heart, returning from the body with poor oxygen saturation, it must be directed from the right atrium (RA) to the right ventricle (RV) and into the pulmonary artery (PA) for distribution to the lungs. However, in individuals with tricuspid atresia, this does not occur as expected. Due to the absence of a right AV valve, the blood must pass through to the left atrium (LA) via an atrial septal defect (ASD) or a patent foramen ovale (PFO). The deoxygenated blood then mixes with freshly oxygenated blood returning from the lungs to the LA, passing down to the left ventricle (LV). This causes a host of problems, including volume overload of the left side of the heart. With the presence of a ventricular septal defect (VSD), the blood can also pass from the LV to the right ventricle (RV), allowing distribution to both the aorta and the pulmonary artery. However, this is not sustainable in part due to the previously mentioned volume overload. This results in left atrial dilation and left ventricular hypertrophy, both of which can lead to heart failure when left untreated. An additional concern is the circulation of poorly oxygenated blood throughout the body as a result of the mixing of the blood in the left atrium. This results in cyanosis, which is observed as

a blue tint to the extremities, nose, tongue, and lips. If left untreated, it can cause additional organs to become oxygen-poor, leading to failure.

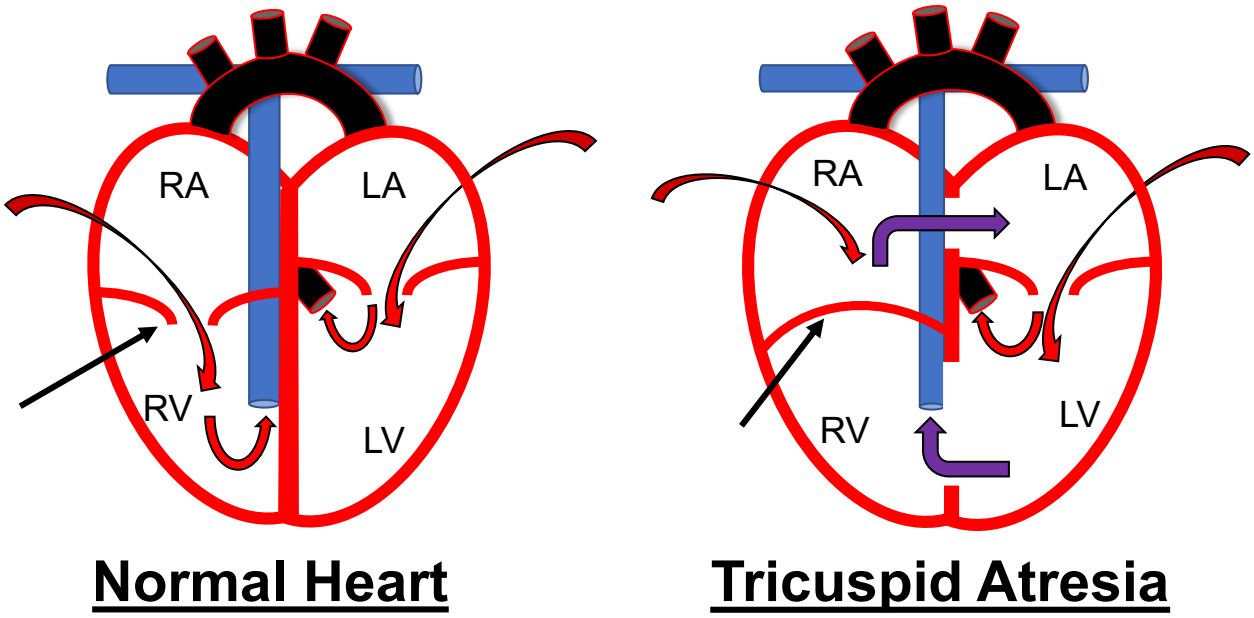

Figure 1. Tricuspid atresia. As opposed to a normal, unaffected heart, tricuspid atresia is characterized by the absence of the tricuspid valve (arrow). This prevents passage of deoxygenated blood from the right atrium to the right ventricle for distribution to the pulmonary artery. In these individuals, an atrial septal defect allows blood to enter the left atrium. A ventricular septal defect then allows the passage of blood to enter the right ventricle for distribution to the pulmonary arteries. This creates a situation in which oxygenated blood is mixing with deoxygenated blood, leading to cyanosis, as well as volume overload in the left side of the heart.

Associated lesions include the previously mentioned ASD and VSD, but also hypoplastic right ventricles due to the lack of blood flow entering the chamber, as well as patent ductus arteriosus (PDA) and, in approximately 30% of patients, transposition of the great arteries (TGA) in which the pulmonary artery draws blood from the left ventricle and the aorta draws from the right ventricle.

Currently, tricuspid atresia is classified into three main categories based on the presence, or absence, of conotruncal abnormalities (*12*). Type 1 classification indicates the presence of normally related great arteries, while Type 2 classifies those with D-transposition of the great arteries (D-TGA). Type 3 is reserved for L-TGA, congenitally corrected TGA, and malposition of the great arteries(*13, 14*). Within each of those categories are sub-classifications indicating the status of the pulmonary valve. Type A indicates pulmonary valve atresia, while Type B denotes pulmonary valve stenosis. Type C is reserved for those with normal pulmonary valves and unrestrictive VSDs.

Treatment of this condition generally follows a scheduled three-stage protocol in which the blood flow is redirected from the right atrium directly to the pulmonary arteries (*15*). Unfortunately, this is not curative. It is a palliative treatment with a finite length of functionality and most individuals develop severe complications later in life (*16*).

The first stage is referred to as a Blalock-Taussig shunt (BT Shunt), in which a conduit is utilized to connect an aortic artery branch to the pulmonary artery to redirect blood flow back into the PAs (*17*). This generally occurs within the first week of life. At this stage, depending on the size of the ASD, the patient's ASD may be enlarged via an atrial septectomy or atrial septostomy. Additionally, if a PDA is still present it will be ligated.

Following this, usually a few months later the BT shunt is taken down and the patient is treated once more with a bidirectional Glenn Shunt, or Hemi-Fontan, in which the superior vena cava (SVC) is ligated directly to the pulmonary artery (PA) bifurcation (*18*). This allows for volume unloading of the left side of the heart, reducing the potential for dilation and hypertrophy, as well as heart failure.

The final phase of palliative treatment comes a few years later in which a Fontan procedure is completed. During this, the inferior vena cava (IVC) is disconnected from the RA and ligated to the PA. There are multiple types of Fontan procedures, including extracardiac and lateral tunnel. A lateral tunnel Fontan utilizes an inter-atrial baffle to close the ASD and leaves the IVC connected to the RA; in this setup, the opening previously from the SVC to the RA is ligated directly to the PA, creating a functional tunnel for the blood to flow directly through the IVC to the RA and into the PA directly, without mixing into the LA. However, extracardiac Fontans involve the use of a conduit that connects the IVC directly to the PA bifurcation, while the SVC is also ligated to the PA (*19, 20*).

Unfortunately, many issues still exist with this treatment. These procedures may only last for up to 15-25 years prior to the onset of heart failure (*16*). At that point, the only treatment is to undergo a heart transplant. Additionally, as a result of the altered Fontan circulation, liver failure may also occur due to increased fibrosis within the liver (*21, 22*).

Up to this point, there has been no known genetic etiology for tricuspid atresia. Very few genetic studies exist, most of which are single-gene sequencing or

chromosomal microarrays (CMA). CMAs are a form of genetic testing that seeks to identify small and large gains or deletions within the chromosomes.

Currently within the literature, a few tricuspid atresia patients have been characterized in terms of limited genetic testing (Figure 2). Chromosomal abnormalities identified include 22q11 deletion (DiGeorge syndrome) (*23-25*), 8p23 (*GATA4* region)(*26*), 4q31 (*NFκB*)(*27*), and 3p (*TGFBR2*) (*28*). While very few have been found, genes of interest that potentially contain pathogenic variants include *NFATC1* (*29-32*), *NKX2.5* (*33*), and *MYH6* (*34*). Interestingly, mutations for *HEY2* came up negative in almost 40 patients across 2 studies (*35-37*).

Commonly used genetic tests in research and clinical practice

Chromosomal microarray (CMA) is the most commonly-utilized genetic test in clinical practices. CMA can be detected small and large deletions/duplications within chromosomal regions which may be playing a role in the manifestation of a phenotype. Karyotype, while not as prevalent, can be used to diagnose conditions such as Turner's Syndrome in which a full chromosome is missing. It can also be utilized to diagnose trisomy's, such as trisomy 20 which is associated with Down's syndrome. Gene panels are also very applicable when utilized properly. These can provide sequence information on a range of genes, from only a handful to 50 or more depending on the condition which is being investigated. There are several congenital heart disease panels which are utilized based on symptoms and phenotypes, both cardiac and extracardiac. WES/WGS is the most powerful tool available, however it is rarely utilized in a clinical manner due to the cost. It can uncover thousands of mutations that may be playing a

role in a disease state, allowing the clinician to identify a gene or handful of genetic variants.

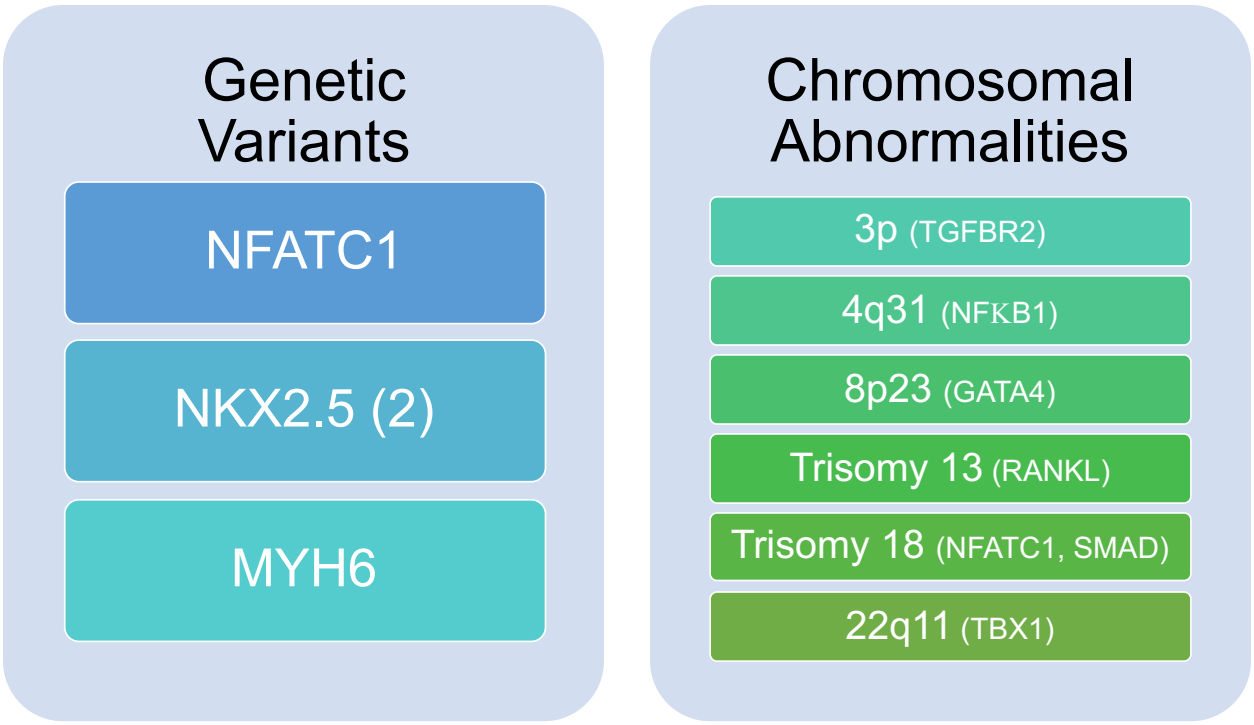

Figure 2. Known genetic and chromosomal abnormalities in tricuspid atresia patients. Shown in this figure are the genes in which variants have been identified in patients with tricuspid atresia (*NFATC1*, *NKX2.5*, and *MYH6*), as well as chromosomal regions that contained a deletion, duplication, or trisomy with the notable gene of interest within the affected region.

Chapter 2: Whole Genome Sequencing: Principles and Ethical Considerations

Introduction

Next generation sequencing (NGS) is a relatively new technology that provides the ability to rapidly sequence large quantities of different DNA fragments within a single reaction (*38*). NGS has enhanced basic science and clinical efforts in many aspects, including analysis of RNA sequences, observing epigenetic changes of an individual's genome, and identifying genes associated with various disease states (i.e. diabetes). Whole genome sequencing, (WGS), as well as whole exome sequencing (WES), is a growing application of NGS. WGS is usually expected to uncover approximately three to four million variants within a single human genome, while WES is more likely to yield 20,000 variants (*38, 39*). WGS/WES can be used in a clinical aspect, as well as for research purposes. Clinically, it can be utilized as a diagnostic tool for patients that display abnormal symptoms that might otherwise require an extensive battery of tests (*40*). It also can play a role in preventative care, by providing information about potential susceptibilities, in the form of genetic variants, which can allow the patient to begin seeking treatment or altering their lifestyle to reduce the effects of a condition that may arise later in life. For example, women that test positive for BRCA1/2 variants are more susceptible to developing breast cancer(*41*); this allows those women to begin seeking treatment options to combat the disease before it becomes a life-threatening situation.

Research scientists have been using WGS/WES to expand the base of knowledge regarding normal biological processes, as well as various pathological states such as cardiovascular disease and diabetes. By unraveling the genetic implications of

pathological states, as well as those that may further complicate those states, additional treatment options can potentially be developed. As a result, WGS/WES seeks to play a significant role in personalized medicine as it becomes more commonplace, allowing medical professionals to treat their patients with greater efficiency.

As WGS/WES uncovers previously unknown gene variants, more relationships between those variants and known disease states will be made. This, however, places an emphasis on unresolved ethical issues. The primary ethical issues associated with WGS/WES include informed consent, the return of results, privacy, and data sharing (Figure 3); this is not an exhaustive list, but does constitute the primary issues discussed in literature at this time. While no consensus exists on many issues, many recommendations have been made regarding the handling of several issues as they arise during clinical and research settings. Unfortunately, there is no federal legislation regarding the myriad issues surrounding the ethics of WGS/WES, there have been various regulations aimed at extending protections that include genetic data. However, those protections only go so far and apply in only a few contexts at this point. Despite this, progress has been made during group sessions and Presidential Commissions aimed at handling these issues.

Primary Ethical Issues

The primary issues to be discussed are informed consent, the return of results, privacy, and data sharing (Figure 3). The goal of this review is not to be exhaustive, but to cover the primary issues and the current state of consensus and recommendations from those within the field. Significant overlap exists among subtopics within each primary issue. These will be addressed in the proper context under each primary issue.

Informed Consent

Prior to beginning a research study or clinical test utilizing WGS/WES, the participant must go through an informed consent process. This process should cover what is relevant for the participant without overwhelming the individual with every single detail, as a concern exists regarding information overload that may cloud their understanding of the process (*42*). Informed consent is much different from a regular consent process in that the individual involved must be informed of the scope of the study, what is expected to be found, which results can and/or will be returned and how it will be returned, as well as the availability of counseling before and after the study. This allows the individual to make an informed decision regarding their participation and minimize any potential misunderstandings. Among the ethical issues surrounding informed consent, the participation of children/minors, secondary uses of data, withdrawal from the study, and third-party rights are heavily debated areas. Participation of children/minors, as well as others that cannot reasonably consent for themselves, requires informed consent; however, it occurs via “proxy consent” because a legal guardian must be involved in the process to provide consent, making the

decision in the best interest of the child participant. While the guardians must act in the best interest of the child, it does not make them an autonomous decision-maker for that child. In a clinical setting, it is highly recommended that children should only be involved

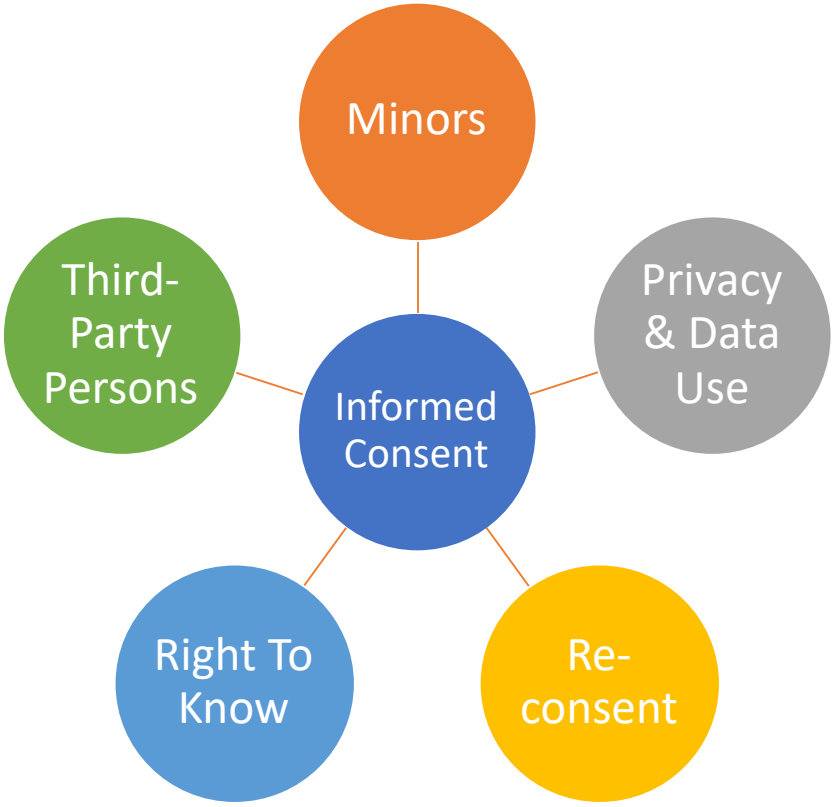

Figure 3. Primary areas of ethical concern. During the informed consent process, several primary ethical issues should be addressed with each potential human participant or with the participant's legal guardian.

when the purpose is diagnostic for a condition for which symptoms have already manifested (*43*). Preventative screening of an asymptomatic minor that is not reasonably expected to currently have a medical condition is strongly discouraged, as the individual still has the right to autonomy. In this scenario, the individual should be allowed to wait until an age at which they are then capable of undergoing the informed consent process and make their own decision. The current consensus within the field is that even if a legal guardian consents in this situation, the child still should not be involved and the clinician should override the wishes of the guardian to protect the rights of the child (*44*). As new findings come to light, some of which may indicate a potential susceptibility to developing a disease later in life, a child should not be forced to know about it without consent because knowledge of this type can have serious psychological consequences that cannot be reversed.

Minors participating in research studies should also be treated as an adult, however, there appear to be fewer perceived ethical hurdles involved especially if the results will not be returned. If a study is seeking to identify heritable traits associated with a specific condition and there is a reasonable expectation the child possesses those traits, it is justifiable to allow proxy consent by a legal guardian for the child to participate (*42*).

During the informed consent process, some individuals may be asked to provide consent for secondary uses of the data, or re-consent. While some members of the field believe this should be standard for all informed consent processes, there is currently not a consensus (*45*). By forcing, or coercing, an individual to decide on re-consent during the process for the original study, that individual's autonomy is violated by not allowing

them to know what future studies will be focusing on when utilizing the original data. It is the right of each participant to be informed of each use and have the ability to consent to, or withdraw from, participation if they do not feel comfortable with the manner in which their genetic information will be utilized. A potential exception could be if their genetic information is used solely as a reference genome for studying a disease already diagnosed and identified with a variant in the original study; using a person's genome as a reference for identification purposes may allow future investigators to set parameters for locating variants within other individual's genomes and can assist in furthering the knowledge of the field. However, all personal identifiers must be removed from the data prior to being used as reference material.

While the issues that have been focused on are those for which a consensus has not been established, the concept of the right to withdrawal is one in which the scientific community appears to be in full agreement at this time. Providing a participant with the right to withdrawal ensures that their right to autonomy is respected and also provides a bridge to building the trust of others outside of the scientific community, which is absolutely crucial for WGS/WES studies involving human subjects. Upon exercising their right to withdraw, all materials associated with that participant are expected to be destroyed – samples, data, etc. However, there are limitations to this which must be addressed during the informed consent process. Once a study has reached a certain point, such as sharing data to databases, it is almost impossible to retrieve and destroy the data (*46*). Other researchers may have already accessed and saved the data at this point and there may be no way to know if that has occurred and who may have done so. At this point, the investigators associated with the study should take measures to

ensure as much is destroyed as possible to continue to protect the participant's rights and respect their wishes.

When considering to participate in a research study or clinical exams, an obligation to third-party persons exists (*45*). Third-party persons, such as siblings, parents, children, and other close relatives, should be involved. The results of one person's genetic test may have significant effects on their relatives, as they share a certain amount of genetic material, including potential genetic variants associated with diseases. These individuals may not want to know about the possibility of possessing variants, whether they're associated with adult-onset or even congenital conditions. As such, participants and clinicians/investigators are obligated to involve third-party persons in the informed consent process (*45*). This is another area in which the scientific community has reached a consensus, as it protects the rights of those that may be affected by knowing the potential content within their genome, regardless of whether the effects may be psychological or social.

Additional complications may arise that must be addressed during the informed consent process prior to beginning any WGS/WES study. These challenges include the ability of an individual to understand and comprehend the information being given to them during the initial process, as well as developing biased or false expectations for what they may receive from the genetic testing. If an individual cannot understand the objective of the study or a test in a clinical setting, several issues may arise after completion of the study once the results are presented. It is absolutely crucial in each instance that every precaution is taken to ensure the individuals participating have a full understanding of what the WGS/WES process involves, what results they may receive

or that they may not receive any results, as well as the validity of the results and the impact of the findings.

Return of Results

Upon completion of a WGS/WES study or clinical test, results may be returned under certain circumstances. The results of a clinical test should be returned and incorporated into, or be accessible via, that individual's medical records (*40*). As these techniques generate an enormous amount of data, which may likely be useless to most of the population, it must be presented in a manner in which the person involved can understand what was found and the potential effects of the findings. Upon completion of a research study, it is widely accepted that the investigator has an obligation to return the results to the participant while funding for the study is still ongoing (*47*); however, this still does not guarantee results will be returned as there are many hurdles that must be overcome first, including validating the results and determining their clinical significance. After funding has ended, they no longer have an obligation to communicate the results, but this does not mean they cannot provide the results to the participant if they would like to.

WGS/WES research studies may involve communities within a population. In these studies, it is highly unlikely for results to be returned to any single individual. An exception is when the individual's results are easily identifiable and can be verified. Many within the field recommend that the results be returned to a healthcare provider (i.e. physician) instead of the participant, for integration into their medical records (*48*). For this to occur, the results must be validated and demonstrate clinical significance

(*45*). However, the meaning of “clinical significance” may continually evolve as more studies are conducted and more knowledge is gained, implicating more variants as clinically significant. Therefore, all identified variants could be maintained in an individual’s medical records even if they do not correspond to a gene currently associated with a specific condition.

Another hurdle in the return of results rests in the hands of the participant, who should be allowed to indicate their right to know or right not to know during the informed consent process. If an individual opts to not know the results, this decision should be respected (*49*). Some within the scientific community feel that a medical professional has the obligation to override that decision if a finding is clinically significant and should be acted upon (*50*) despite it conflicting with a person’s right to autonomy. The decision regarding the right to know or not know also applies to minors (*49*), as well as during prenatal testing (*42*).

Incidental findings are variants not relevant to the scope of the study that is uncovered during the process of analyzing the genome, may or may not be returned. This is determined primarily by two specific factors: right to know/not to know and the classification of the incidental findings. A system was proposed for classifying these incidental findings into three “bins”: Unknown or of no clinical significance, clinically valid but not directly actionable, and clinically actionable (*51*). Results classified as “unknown of or no clinical significance” are those that are not currently linked to a specific phenotype and will consist of the majority of the variants found within an individual’s genome. The classification of a variant as “clinically valid but not directly actionable” indicates a variant that is known to be associated with a specific condition, but does not

have a method of intervention developed to treat it at that time. And those classified as "clinically actionable" refer to those that are associated with a specific medical condition and have established treatment options, which represents the smallest group of variants found within a human genome. Variants associated with *BRCA1/2*, which is implicated in breast cancer, are an example of those classified as clinically actionable. This system provides a great method for determining how a variant should be classified, as well as a way to help determine if it should be returned. However, at this time it is not universally applied by all investigators and clinicians.

It is highly recommended that a professional with expertise and the proper qualifications to verify the findings should return the results to the participant. This is recommended so the results, as well as their implications, can be adequately communicated to the participants in a manner they can understand. As previously mentioned, the data produced by WGS/WES is essentially meaningless to most participants because they have no way to interpret it. Additionally, for research studies it is required that the lab returning results must be Clinical Laboratory Improvement Amendment (CLIA) certified (*52*); if they are not, the results must be validated by a CLIA-certified lab and be disseminated to the professional responsible for communicating the results to the participant. In addition to having the results validated by a CLIA-certified lab, some recommend investigators use a CLIA-certified lab from the beginning of the study if it is anticipated variants of clinical significance will be found (*52*).

Upon the return of results, it is strongly recommended that counseling be available to participants as well as any third-party persons that may be impacted (*53*).

The information communicated to the individual may leave a psychological effect, which a counselor may be able to help them cope. With the primary source of contention regarding visiting a counselor is the timing; some believe it should occur before and after the test (*48, 50*), while some feel it should only occur after the results have been returned.

Privacy Considerations

Due to the sensitive nature of the results obtained by WGS/WES, concerns regarding privacy represent a major ethical issue. During the informed consent process, the participant is to be allowed a choice of remaining anonymous or allowing their identity to be associated with their results. Upon the request of privacy, a participant's decision must be respected and maintained during the entirety of the study, as well as during data distribution. Not only would it be in violation of an individual's right to autonomy and their privacy, but it could also potentially damage the relationship between those within the public and scientific communities.

If an individual opts to remain anonymous, a key linking the sample to the identity-determining information should be established for their samples and data that few individuals have access. This allows the individual's privacy request to be respected, but also allows the investigator to return results by linking an individual to a specific set of results. In the event an individual opts to not receive any results, their samples and data should still be treated with the same sensitive nature.

Maintaining an individual's privacy can help protect against unauthorized access and misuse of data. By removing the ability to associate genomic data with an individual

or a family, it makes it much more difficult for the results to be used against them in a discriminatory manner, whether it be by future employers or some other aspect (*54*). The misuse of data in this manner is an obvious violation of an individual's rights. The previously mentioned coding system should have a very specific manner for keeping the identity-determining information safe to prevent theft or improper access; this can be through physical means, such as utilizing a locked storage area, or by maintaining the information on an encrypted or password-protected hard drive. The only individuals with access to this information should be the principal investigators of the study.

Data Sharing

Researchers have an obligation to share their results. The variants and data obtained via WGS/WES are no exception. While most variants may be clinically insignificant, they still provide utility to other scientists during future studies. The intent and scope of data sharing should be clearly stated during the consent process. Some believe a participant should be able to opt-in or out of data sharing, yet still participate in the study (*55*). While not sharing an individual's results is not detrimental to the study, there is still some debate regarding the moral obligations to share any and all data within the scientific community. Data sharing can assist with future diagnoses and statistical studies, as well as reduce the number of necessary participants in future studies due to the growing population of genomic data made available. Obtaining the data of genomes known to not contain a variant you are investigating may reduce the number of participants required as a control for a study, which will decrease the expense of a study. As previously mentioned, one complication associated with data

sharing is an individual's right to withdraw from a study. Once the data has been submitted to a database, public or private, it may be impossible to destroy the data and prevent its future use.

While most data are posted to public databases for others around the world to access and utilize for their own studies, one option that may ease the concerns of WGS/WES participants is that of only sharing to protected databases that require certain qualifications to access. In one group study, when presented with fewer options, participants were comfortable sharing their results via public databases; however, when presented with multiple data sharing options, those involved generally preferred private or encrypted databases due to privacy concerns (*55*).

Shared data still allows the right of privacy to be maintained, which helps prevent future users of the data from identifying the individual from which it came. However, it is still possible to identify an individual using their anonymous data (*56, 57*), as a person's genome is the ultimate identifier. While that may seem concerning, identification still requires intimate knowledge of other various details about the individual, so there are still many hurdles to overcome in order to identify the person. One drawback to anonymous data is that investigators of future studies that would like to contact the individual for more information or for consent to participate in an additional study will not be able to do so. A potential alternative is to contact the investigator of the original study and ask them to reach out to the participants; this would allow those participants to contact the new investigator if they are willing to be involved instead of being solicited for more information from an investigator they are not familiar with.

As with the informed consent process, third-party persons should still be involved in determining whether or not the data should be shared and to which type of database it should be disseminated. They should be informed of the various options, as well as the benefits and consequences of each option (*45*).

Additional Recommendations and Ethical Principles

Governance

Currently, there are no established federal regulations specifically intended for WGS/WES. Recommendations have been made to establish a national body to assist with key ethical issues on a case-by-case basis, as well as provide guidance for determining the clinical significance of obtained results and how they are classified (*58*). Formation of such a committee would provide a singular entity that could define the criteria for clinical significance, as well as assist with the development of federal regulations for ethical practices in WGS/WES. Investigators and even clinicians should be able to submit results to this body for further analysis and validation, as well as be given guidance for the dissemination of those results. Additionally, local experts should play a role in the process by assisting with studies prior to the completion of analysis; this would relieve the stress and burden of a national committee from having an abundance of requests that do not align with their specified function. While the current recommendation is for post-study guidance, national bodies could also be established to provide guidance for properly conducting a project and how to handle the various issues associated with informed consent, as well as other pre-study challenges.

To establish such national oversight committees would be, understandably, a monumental task that would require significant effort and insight from the scientific community. However, the formation of such a committee would also provide much-needed relief on the part of the investigators and Institutional Review Boards (IRBs) by establishing a set guideline for validating and reporting results, as well as disseminating those results to others within the field. Alternatively, a potential national committee could develop a protected database, available internationally, for sharing data. By developing such a database with significant oversight, participants of studies may feel more comfortable being involved in the study and allowing their results to be shared with others as it may provide a sense of security not felt with public databases. Additionally, the ability of a national body to validate findings may also provide an avenue to report the clinically significant findings to the appropriate medical professionals for efficient incorporation into a person's electronic medical record.

Role of IRBs

Under the Common Rule (Williams 2005), research studies being provided with federal funding are required to be reviewed by an Institutional Review Board (IRB) (*59*). IRBs are a critical entity involved in clinical trials, as well as research involving human subjects. IRBs have the authority to approve a study, require investigators to modify their studies, as well as deny approval of a study. They are in place to protect the rights of everyone involved, as well as ensure their safety, on a study-by-study basis. Due to the relatively new nature of WGS/WES studies and the lack of established federal guidelines, the duties of IRBs have become more challenging. For WGS/WES studies,

this must occur regardless of any intention to return, or not return, results to participants, as well as when using samples that have been previously collected even if they are anonymous (*46*).

Presidential Commission for Bioethics Principles

Principles established by a 2012 Presidential Commission, comprised of a group of experts within the genome sequencing field, include respect for persons, public beneficence, responsible stewardship, intellectual freedom and responsibility, democratic deliberation, and justice and fairness (*59*).

Respect for persons requires an investigator, along with anyone involved in a WGS/WES study, to uphold and respect a participant's wishes unless they are going to be harmful to others. This alludes to the right to autonomy of all participants involved, whether for research or in a clinical setting. The principle also aims to protect those who cannot act autonomously for themselves, such as minors, as well as those who are ill or otherwise incapable of understanding and making informed decisions. Special responsibilities must be upheld to ensure their safety and that their best interests are maintained in decision-making processes.

Public beneficence requires those involved in a study to act in the best interest of the public and the participants. Investigators must act to ensure public benefit while minimizing harm to all. This principle embraces our responsibility to pursue directives with the potential to improve the well-being of the public and any potential participants.

Demonstrating responsible stewardship demands to account for the interests and needs of individuals who are unable to represent themselves in the process of

promoting scientific advances. Those who cannot represent themselves, while outlined previously, also include future generations. It is not only the obligation of the scientific community but also governing bodies within this country, to uphold this principle to ensure the protection of every member of society in these situations.

Scientists, when acting responsibly, are given the right to intellectual freedom in their quests to advance science while promoting the good of the public. Allowing scientists to exercise their creative freedom is crucial to the design of novel studies. While scientists are given the freedom to exercise their creativity, they are concurrently obligated to the principle of intellectual responsibility. This requires that while conducting these studies, scientists are obligated to uphold the ideals of research, as well as ensuring no harm comes to others in the process. They are also responsible for upholding all appropriate policies and regulations during their studies.

Democratic deliberation is a collaborative method of decision making that involves openly debating different views and incorporating the participation of citizens within the community. This is recommended not only for the design and execution of research studies but also for the development of regulatory measures. Additionally, the involvement of those concerned with issues that may be raised as a result of scientific progress is also highly encouraged as it can alleviate some of their concerns and begin to build trust between the public and the scientific community. Investigators may develop a well-rounded approach to solving a problem when exercising this responsibility. This can potentially include accounting for hurdles they may not have foreseen themselves, while also helping those outside of the scientific community understand exactly what is taking place and what is hoped to be gained by the study.

Finally, the principle of justice and fairness is related to the manner in which not only benefits, but also burdens, of advancements will be distributed across society. The desire is to ensure investigators do what is necessary to make sure the burdens of their advances do not fall unequally on the shoulders of any group or individual. It also requires that benefits are distributed widely, as well as equally, among the scientific community and the public. For example, any advances made in WGS/WES that may decrease the expense of utilizing the technology should benefit not only the investigators that plan to use it but the patients that may seek to use the service for clinical purposes.

Federal Guidelines

The Health Insurance Portability and Accountability Act (HIPAA), established in 1996, sets forth regulations to define what constitutes protected health information (PHI) and its legal protections. Additionally, 18 identifying information markers were also established that apply to PHI. If PHI is stripped of these 18 identifiers, it is classified as "not identifiable". Additionally, WGS/WES data that has not been stripped of these identifiers are considered "protected health information" and is not covered under the HIPAA Privacy, Security, and Enforcement Rules, as well as the Common Rule (*59*). Upon removal of these identifiers, the data is no longer protected. Interestingly, research using data that no longer possesses these identifiers is no longer classified as human research and does not require IRB review.

In 2008, the Genetic Information Nondiscrimination Act (GINA) was enacted to provide an additional measure of protection against discrimination due to genetic

information by health insurers, which is an extension of HIPAA protection, as well as in employment decisions. However, GINA does not provide protection against discrimination in other contexts, such as life insurance or disability insurance (*59*). One significant limitation to GINA is that it only applies to individuals not displaying symptoms and those in which disease has not already manifested. Also, it does not provide protection from the access or use of genetic data; while your genetic data may be accessed, it cannot be used against you in the specific contexts defined by the regulation.

Conclusion

As with any new technology, whole genome and whole exome sequencing raise a host of ethical concerns as more uses and variant-disease associations come to light. These ethical issues include, but are not limited to, informed consent, the return of results, privacy of participants and patients, and sharing of data. These concerns are not only focused on the use of the technology, which can be applied in a clinical setting and a research setting but also the data which it provides and how it is stored and used. A lack of consensus among the scientific community muddies the direction and guidance provided for many within the field regarding these ethical issues. However, progress is continually being made as more recommendations have come to light due to group sessions involving those within the field.

Chapter 3: Assessing Genetic Testing Frequency and Outcomes in a Cohort of Unrelated Tricuspid Atresia Patients

Introduction

The purpose of the retrospective chart review was to classify and describe the detailed cardiac phenotype associated with tricuspid atresia (TA), as well as evaluate the state of genetic testing frequency and outcome in this population.

No studies exist evaluating the history of clinical genetic testing within a TA cohort, making this the first study to do so. One study evaluated 105 TA patients in terms of surgical procedures utilized and outcomes (*60*). An additional study looked at 225 TA patients that underwent Fontan to assess the overall impact on tricuspid atresia (*61*).

In terms of research genetic testing studies on TA cohorts, *HEY2* was screened in almost 40 patients, however, none presented with a mutation in the gene (*36*).

Methods

Inclusion Screening

In total, 285 patients from Texas Children's Hospital were screened for inclusion. Inclusion criteria required one echocardiogram report on record with a diagnosis of tricuspid atresia. Within the Texas Children's Hospital medical record system, I utilized the CardioIMS and EPIC systems to check for echocardiogram records and diagnoses. Approval for the study was granted under IRB H-41142 at Baylor College of Medicine.

Review and Documentation of Medical Records

Following inclusion screening, 234 patients remained. The patients' medical records were reviewed in CardioIMS and EPIC. Demographics, cardiac phenotype, surgical history, and genetic testing were recorded utilizing a REDcap database.

Demographics included sex, date of birth, deceased status, cause of death if deceased, and race. Patients that had not been seen by any department or clinic at Texas Children's Hospital within the last five years were considered "lost to follow-up" and were marked as "Unknown" for their deceased status.

Results

Beginning with inclusion screening, 285 patients, born between January 1, 1990 - March 2017, were assessed. Following this screening, 234 patients were for the study.

Demographics Data

56.8% (133) of the cohort was male, with 12.8% (30) overall deceased ranging from one week to 13 years of age (Table 1). Causes of death include respiratory failure, kidney failure, necrotizing enterocolitis, and surgical complications. The cohort was predominately white (180, 76.9%) with 55% being non-Hispanic.

Table 1. Demographics and primary diagnosis categories of tricuspid atresia cohort

	Type1		Type 2		Type 3/4		Cumulative	
	Total	**Percent**	**Total**	**Percent**	**Total**	**Percent**	**Total**	**Percent**
Overall	151	64.5%	52	22.2%	32	13.7%	235	100%
Male	74	49.0%	40	76.9%	19	59.4%	133	56.8%
Female	77	51.0%	12	23.1%	12	37.5%	101	43.2%
Deceased	18	11.9%	9	17.3%	4	12.5%	30	12.8%
Unknown	37	24.5%	5	9.6%	4	12.5%	46	19.7%

The cohort of 234 patients was grouped according to the Van Praagh tricuspid atresia classification. Patients considered “unknown” in terms of deceased status are those that have been lost to follow-up, or have not been seen in the hospital by any department within the last 5 years. Values provided indicate the overall number of patients in each category.

Cardiac Phenotype

Primary diagnosis categories include Type 1 (151, 64.5%), Type 2 (52, 22.2%), and Type 3/4 (32, 13.7%) (Tables 1 and 2). The Type 1 cohort consists of individuals with tricuspid atresia and normally-related great arteries. The Type 2 cohort is characterized by those with D-transposition of the great arteries (D-TGA). Type 3/4 cohort includes those with other abnormalities of the great arteries, such as L-transposition of the great arteries (L-TGA), double outlet right ventricle (DORV), malposition of the great arteries, anatomically corrected malposition of the great arteries, and truncus arteriosus.

The Type 1 cohort (151), four presented with dextrocardia (2.6%). The most common abnormality in the right-sided lesion group was secundum ASD (83, 55.0%), muscular VSD (66, 43.7%), pulmonary valve stenosis (53, 35.1%), and patent ductus arteriosus (53, 35.1%). Additionally, 34 individuals (22.5%) presented with pulmonary valve atresia. Patent foramen ovale, as opposed to secundum ASD, was present in 30 individuals (19.9%). It was the second most-common septal defect within this group. Interestingly, a few lesions which were not expected within this group were observed. One case of hypoplastic aortic arch (0.7%) was seen. Additionally, one case of cleft mitral valve (0.7%) and left ventricular non-compaction (0.7%) were also noted. Most notably, two individuals presented with bicuspid aortic valve (1.3%).

Within the Type 2 cohort (52), the most common associated lesions observed include secundum ASD (27, 51.9%), muscular VSD (30, 57.7%), and patent ductus arteriosus (30, 57.7%). In addition, PFO (12, 23.1%), coarctation of the aorta (17, 32.7%), hypoplastic aortic arch (3, 5.8%), interrupted aortic arch (3, 5.8%), and bicuspid

aortic valve (1, 1.9%) were also seen. Interestingly, pulmonary atresia (7, 13.5%) and pulmonary valve stenosis (6, 11.5%) were observed.

In the Type 3/4 cohort (32), six individuals presented with dextrocardia (19.4%). Interestingly, 22 had L-TGA (71.0%), as well as one with truncus arteriosus (3.2%) and another individual with anatomically corrected malposition of the great arteries (3.2%). Fifteen patients had PDA (48.4%). Four individuals had pulmonary valve atresia (12.9%) and four others had pulmonary valve stenosis (12.9%). Secundum ASD was present in 17 (54.8%). Muscular VSD was the most common type observed, present in 12 individuals (38.7%).

Surgical history

The most frequently documented surgical procedures include the Blaylock-Taussig (BT) shunt, bidirectional Glenn shunt, and the Fontan (Table 3). Within the Type 1 cohort, 72 individuals had a BT shunt (52.8% right-sided). One-hundred twelve bidirectional Glenn shunts (65.2% right-sided) were noted and 108 Fontan procedures (57.4% extracardiac, 41.7% non-fenestrated). Additionally, 31 patients within the group underwent atrial septectomies (20.5%), 24 with PA bands (15.9%), and 29 had oversewn pulmonary valves (19.2%). Only one patient (0.7%) was documented to undergo orthotopic heart transplant

The Type 2 cohort had a much more complex schedule of surgical procedures. Twenty patients (38.5%) were documented to undergo BT shunts (45% right-sided). Forty-two (80.8%) underwent a bidirectional Glenn shunt (59.5% right-sided) and 39 (75.0%) had Fontan procedures (76.9% extracardiac, 56.4% non-fenestrated).

Additionally, 21 patients (40.4%) each had an atrial septectomy and PA band. However, additional operations documented include DKS/Norwood procedures (16, 30.8%) and seven with arterial switch (13.5%). Two individuals (3.8%) underwent orthotopic heart transplant, with one requiring a re-transplant.

Within Type 3/4 cohort, 13 (40.6%) had a BT shunt (30.8% right-sided), 26 with bidirectional Glenn (65.4% right-sided), and 23 had a Fontan (73.9% extracardiac, 47.8% non-fenestrated). Thirteen individuals (40.6%) had an atrial septectomy and 12 (37.5%) were given a PA band. Additionally, one (3.1%) underwent orthotopic heart transplant.

Table 2. Cardiac phenotypes by diagnosis category

	Type 1 (n=151)		Type 2 (n=52)		Type 3/4 (n=31)	
	Total	Percent	Total	Percent	Total	Percent
Levocardia	137	90.7%	44	84.6%	21	67.7%
Dextrocardia	4	2.6%	5	9.6%	6	19.4%
Mesocardia	0	0.0%	0	0.0%	1	3.2%
D-TGA	0	0.0%	52	100.0%	0	0.0%
L-TGA	0	0.0%	0	0.0%	22	71.0%
Anatomically Corrected Malposition	0	0.0%	0	0.0%	1	3.2%
Truncus Arteriosus	0	0.0%	0	0.0%	1	3.2%
DORV	0	0.0%	0	0.0%	7	22.6%
Coarctation of Aorta	2	1.3%	17	32.7%	4	12.9%
Hypoplastic Aortic Arch	1	0.7%	3	5.8%	2	6.5%
IAA	0	0.0%	3	5.8%	0	0.0%
Aortic Dilation	61	40.4%	18	34.6%	10	32.3%
Aortic Valve Atresia	0	0.0%	1	1.9%	0	0.0%
Bicuspid Aortic Valve	2	1.3%	1	1.9%	1	3.2%
Hypoplastic Aortic Valve	1	0.7%	2	3.8%	0	0.0%
Pulmonary Atresia	34	22.5%	7	13.5%	4	12.9%
Pulmonary Stenosis	53	35.1%	6	11.5%	4	12.9%
Bicuspid Pulmonary Valve	2	1.3%	1	1.9%	0	0.0%
Truncal Valve	0	0.0%	1	1.9%	0	0.0%
Secundum ASD	83	55.0%	27	51.9%	17	54.8%
PFO	30	19.9%	12	23.1%	7	22.6%
Common Atrium	0	0.0%	0	0.0%	2	6.5%
Normal Atrial Septum	2	1.3%	4	7.7%	0	0.0%
Muscular VSD	66	43.7%	30	57.7%	12	38.7%
Inlet VSD	3	2.0%	2	3.8%	4	12.9%
Conoventricular VSD	12	7.9%	2	3.8%	1	3.2%
Normal Ventricular Septum	9	6.0%	1	1.9%	0	0.0%
Cleft Mitral Valve	1	0.7%	1	1.9%	0	0.0%
Patent Ductus Arteriosus	53	35.1%	30	57.7%	15	48.4%

Overall, those within the Type 1 group had fewer complex phenotypes, while the Type 2 and Type 3/4 groups were more complex, as expected. Interestingly, the Type 2 group, composed of those with D-transposition of the great arteries still possessed a surprising number of pulmonary valve abnormalities, which would be considered a left-sided lesion in these cases. Values provided indicate the number of patients with documentation of each phenotype.

Genetic Testing

Genetics testing was completed on 65 patients, while only 32 were evaluated by genetics. Of the Type 1 cohort, 30 (19.9%) underwent genetic testing. Twenty-three of those (15.2%) were tested via chromosomal microarray (CMA), however, only two (1.3%) were clinically significant and two others (1.3%) were of unknown clinical significance. Three abnormal WES were documented (2.0%) (Table 4). The genes of interest noted in the clinical report include *NFATC1*, *TGFB1*, *MYH6*, *VKORC1*, *FBN1*, and *BMP1*. Two individuals (7.4%) tested for individual genes showed variants in *MTHFR* and *NIPBL*. Three abnormal karyotypes (11.1%) were observed. Results include 47 XX trisomy 21, 46 XX with 10% mosaicism, and 46 XY with a terminal deletion on chromosome 10 (10q26.3 verified by CMA). One abnormal FISH was seen, resulting in partial tetrasomy 8q (46 XX with added 8q24.2).

The Type 2 cohort was tested more frequently, as 25 (48.1%) underwent genetic testing. Twenty-two CMAs (42.3%) were given, however only one (1.9%) was considered clinically significant; two others (3.8%) were of unknown clinical significance. One abnormal WES (1.9%) was noted. The primary genes found was *MYH6*. Three individuals (5.8%) had abnormal single-gene sequencing results, with *F5* and *HgbS* showing variants. One abnormal FISH (1.9%) was seen, resulting in maternal heterodisomy of chromosome 15.

Within the Type 3/4 cohort, ten individuals (32.3%) were genetically tested. Nine (29.0%) were given CMAs, however none were clinically significant. Two individuals (6.5%) had an abnormal WES, with *MYH6*, *ERCC6L2*, and *RTEL1* variants found in

both individuals. Those two individuals both had diagnoses of L-TGA. One (3.2%) tested positive for a variant *MTHFR*. There were no abnormal karyotypes or FISH.

Table 3. Surgical history for each tricuspid atresia type

	Type 1 (n=151)		Type 2 (n=52)		Type 3/4 (n=31)	
Surgical Procedure	**Total**	**Percent**	**Total**	**Percent**	**Total**	**Percent**
BT Shunt	72	47.7%	20	38.5%	13	40.6%
Right	38	52.8%	9	45.0%	4	30.8%
Left	7	9.7%	1	5.0%	0	0.0%
Bilateral	0	0.0%	0	0.0%	0	0.0%
Atrial Septectomy	31	20.5%	21	40.4%	13	40.6%
PA Band	24	15.9%	21	40.4%	12	37.5%
CoA end-to-end repair	1	0.7%	2	3.8%	1	3.1%
CoA arch advancement	0	0.0%	7	13.5%	3	9.4%
PDA ligation (surgical)	10	6.6%	3	5.8%	2	6.3%
BT shunt takedown	60	39.7%	16	30.8%	12	37.5%
Pulmonary artery band takedown	11	7.3%	6	11.5%	3	9.4%
Oversewing of pulmonary valve	29	19.2%	8	15.4%	6	18.8%
DKS/Norwood	1	0.7%	16	30.8%	3	9.4%
Glenn	112	74.2%	42	80.8%	26	81.3%
Right-Sided	73	65.2%	25	59.5%	17	65.4%
Left-Sided	2	1.8%	3	7.1%	0	0.0%
Bilateral	4	3.6%	0	0.0%	0	0.0%
Fontan	108	71.5%	39	75.0%	23	71.9%
Extracardiac	62	57.4%	30	76.9%	17	73.9%
Lateral Tunnel	22	20.4%	6	15.4%	5	21.7%
Atriopulmonary	2	1.9%	0	0.0%	0	0.0%
Fenestrated	39	36.1%	10	25.6%	7	30.4%
Non-Fenestrated	45	41.7%	22	56.4%	11	47.8%
Pacemaker	7	4.6%	4	7.7%	4	12.5%
Epicardial Dual	5	71.4%	3	75.0%	2	50.0%
Epicardial Atrial	2	28.6%	0	0.0%	0	0.0%
Epicardial Ventricular	0	0.0%	0	0.0%	1	25.0%
Pacemaker Revision	3	2.0%	5	9.6%	1	3.1%
Arterial Switch	0	0.0%	7	13.5%	2	6.3%
Heart Transplant	1	0.7%	2	3.8%	1	3.1%
Re-Transplant	0	0.0%	1	1.9%	0	0.0%

The surgical history was very straight-forward for the Type 1 group, but became much more complex in those in Type 2 and Type 3/4 groups. The primary documented procedures are the Bidirectional Glenn shunt and the Fontan. Values provided indicate the number of patients for which documentation of each surgical procedure was available.

Table 4. Clinical genetic testing outcomes in the tricuspid atresia cohort

	Type 1 (n=151)		Type 2 (n=52)		Type 3/4 (n=31)	
	Total	**Percent**	**Total**	**Percent**	**Total**	**Percent**
Genetics Evaluation	18	11.9%	9	17.3%	5	16.1%
Genetic Testing	30	19.9%	25	48.1%	10	32.3%
CMA	23	15.2%	22	42.3%	9	29.0%
Clinically Significant	2	1.3%	1	1.9%	0	0.0%
	Loss 10q26.3; Loss 1q41, 2q11.2 (USH2A – Usher Syndrome)		Gain 1q42.13 (OBSCN – Dilated Cardiomyopathy)			
Unknown Significance	2	1.3%	2	3.8%	0	0.0%
WES (Abnormal)	3	2.0%	1	1.9%	2	6.5%
Results	NFATC1, TGFB1, MYH6, VKORC1, BMP1, FBN1		MYH6		MYH6 (2), ERCC6L2 (2), RTEL1 (2)	
Gene Sequencing (Abnormal)	2	1.3%	3	5.8%	1	3.2%
Results	MTHFR, NIPBL		F5, Hgb S (2)		MTHFR	
Karyotype (Abnormal)	3	2.0%	0	0.0%	0	0.0%
Results	47XX Trisomy 21; 46XX 10% mosaicism; 46, XY, Deletion at 10q26.3 (terminal arm)					
FISH (Abnormal)	1	0.7%	1	1.9%	0	0.0%
Results	Partial Tetrasomy 8q (46, XX, add(8)(q24.2)		Maternal heterodisomy of Chromosome 15			

Genetic testing observed in this cohort was more infrequent than previously believed. While there were more Type 1 patients, the Type 2 and Type 3/4 groups were tested more frequently. The primary test utilized was a chromosomal microarray, while whole exome sequencing was only used in 7 patients.

Discussion

While the literature generally states transposition of the great arteries accounts for 25% of all tricuspid atresia cases, it was slightly higher at 31.9% in this cohort including D-TGA and L-TGA. Very few cardiac phenotype surprises arose. However, the notable exception is the presence of pulmonary valve defects in the Type 2 cohort – 25% had either pulmonary atresia or pulmonary valve stenosis. Interestingly, four cases of bicuspid aortic valve were observed. It was generally considered that left-sided lesions were mutually exclusive from right-sided lesions such as tricuspid atresia. While two cases out of 234 isn't a very large number (1.7%), it still indicates that this combination of phenotypes does exist.

As expected, most patients appeared to follow the traditional classification criteria. However, if re-classification were to be recommended, it would be to combine D-TGA with L-TGA. Reserve Type 3/4 for other, rarer phenotypes as a "miscellaneous bin" including lesions such as truncus arteriosus, DORV, and anatomically corrected malposition.

The aorta issues in the Type 2 cohort are a result of "no flow, no grow". What would normally be considered a left-sided lesion is actually right-sided due to the transposition. This creates a situation in which the PA is patent, as long as no LV OFT obstruction exists and the PV is "normal". However, what is interesting in these is that 25**.0**% of these individuals had some sort of pulmonary artery defect, stenosis or atresia. Interestingly, pulmonary valve stenosis and atresia among the Type 2 cohort were of particular interest as these would be considered left-sided lesions due to the altered anatomy of the heart in these individuals.

Interestingly, over time it appears as if the lateral tunnel Fontan procedure was phased out in favor of the extracardiac Fontan. From 1992 until 2003, 21 of 35 (60.0%) recorded Fontan types were lateral tunnel, with 13 being extracardiac (35.3%). Of those 34, 31 (91.2%) were fenestrated. However, beginning in 2000, the extracardiac Fontan began to be used more frequently, but it wasn't until June 2003 when it became the primary Fontan type utilized as the lateral tunnel was phased out. From 2003 through July 2019, 88 of 98 (89.7%) recorded Fontan procedures were extracardiac; the remaining 10 were lateral tunnel Fontans. Within that time span, 87 of the Fontan procedures conducted showed documented fenestration status. Of those 87, 70 (80.4%) were non-fenestrated, which fenestration also appearing to be phased out with the switch to the extracardiac Fontan.

A previous meta-analysis study demonstrated the superiority of the extracardiac Fontan for functional single ventricle patients in several areas including the occurrence of supraventricular arrhythmia and rate of protein-losing enteropathy, as well as not requiring a cardiopulmonary bypass (*62*).

Surprisingly, only 65 patients had undergone genetic testing, but even more significant is number with clinically significant results. Four clinically significant CMAs, six abnormal WES, six abnormal single-gene tests, three abnormal karyotypes, and two abnormal FISH. There were no overlapping regions among those with abnormal CMAs. However, there were two individuals with chromosome 1 (1q41 loss, 1q42.13 gain), chromosome 2 (2q21.1 gain, 2q11.2 loss), chromosome 3 (3p21.1 gain, 3q25.2 gain), and chromosome 10 (10p12.2 gain, 10q26.3 loss).

The most interesting results, however come from the patients that underwent WES, as *NFATC1*, *TGFB1*, *FBN1*, and *MYH6* have cardiac significance based on the available literature. *FBN1* is associated with Marfan syndrome, however some patients have tricuspid valve abnormalities such as prolapse and regurgitation due to degeneration of tricuspid valve leaflet tissue (*63*). Particularly noteworthy is the *NFATC1* variant due to the previously published study showing it in a tricuspid atresia patient and the mutations result in cytosolic accumulation of NFATC1 (*32*), failing to enter the nucleus to conduct its function. It has also been shown to be critical for cardiac valve formation(*29*). It is expressed specifically in endocardial endothelial cells of the primitive heart tube and is restricted to the atrioventricular cushion (AVC) and outflow tract (OFT) endothelial cells at the early stages of endocardial cushion formation (*64*).

TGFB1 is weakly expressed in cushion mesenchyme but strongly expressed in endocardium throughout valvulogenesis (*65-68*). *MYH6* (*34*) variants in this have been shown in multiple individuals in the literature with tricuspid atresia. Interestingly, *ERCC6L2* and *RTEL1* were possessed (in addition to *MYH6*) by two patients. Both of whom have L-TGA. Both are DNA helicase genes – one for telomerase (*RTEL1*), one for mitochondria (*ERCC6L2*).

A study conducting WES on 342 left-sided lesion patients revealed two novel candidate genes, *JARID2* and *SMURF* (*69*), which are of interest in relation to this tricuspid atresia cohort. It has been shown that SMURF1 is a downstream effector of TGFB activity, which is a pathway showing involvement in tricuspid atresia and has been previously shown to affect AV valve formation. *JARID2* is also of interest as it is

co-expressed with *NFATC1* and knockout mouse models develop double outlet right ventricle and VSD, both of which have been observed in tricuspid atresia patients.

Chapter 4: Whole Exome Sequencing to Identify Candidate Variants in a Family with Congenital Heart Disease

Introduction

Following the chart reviews, we selected for our study a family with recurrent congenital heart disease including tricuspid atresia and bicuspid aortic valve to identify any commonalities among genetic testing and associated lesions from the previously described tricuspid atresia cohort. Initially, it appeared as if only two affected individuals were in this family, along with one affected great-uncle (Figure 4). However, this changed after echocardiogram and electrocardiogram testing, as more lesions were revealed that had been previously undiagnosed (Figure 5).

The proband within this family has a history of tricuspid atresia with ASD, VSD, PDA, cleft mitral valve, and hypoplastic right ventricle (Figure 6). Additionally, the proband also was previously diagnosed with cleft lip and palate, as well as misshapen vertebrae and fused ribs. While no significant extracardiac lesions were identified, the father had a bicuspid aortic valve, effacement of the sinotubular junction, and ascending aorta dilation (Figure 7), while the sister had effacement of the sinotubular junction (Figure 8). Additionally, the sister also has an abnormally high-arched oral palate (Table 5). The sinotubular junction is a region connecting the aortic root with the ascending aorta. Generally, this region constricts slightly before opening up to the ascending aorta. However, in the two individuals in this family with effacement, this constriction is not present. It appears more as a congenital form of aortic dilation.

Bicuspid aortic valve disease is present when two of the three aortic valve leaflets are fused or only two leaflets are formed. This results in aortic insufficiency,

causing regurgitation into the left ventricle and acceleration of blood flow into the aorta. This can cause many issues later in life, including left ventricular hypertrophy and aortic rupture or dissection if left untreated. While tricuspid atresia is particularly rare, bicuspid aortic valve is much more common as it is present in 1-2% of the population(*70*).

WES was utilized as a genetic test of choice using whole venous blood as the source of DNA (Figure 9). The exome covers 2% of the genome, approximately 30,000 genes. However, 85% of all known disease-causing variants are in the exome. This makes it a good choice in terms of cost-effectiveness with the highest likelihood of identifying disease-associated variants. While this would also allow for the investigation of germline variants, tissue-specific somatic variants would remain unknown. Once the WES was completed, the data required bioinformatics analysis to provide a fully annotated list of variants for each individual. This included information such as minor allele frequency (MAF), protein function, expression patterns, and rsID information if known (Figure 10).

While preparing for this study, many ethical issues and current guidelines had to be considered. The issues of interest include informed consent, the participation of minors, third party rights, right to know, and data use. This required developing informed consent documents, along with HIPAA privacy waivers, and allowing each participant to determine whether they wanted to be informed of results or not, or if they would like to decide later.

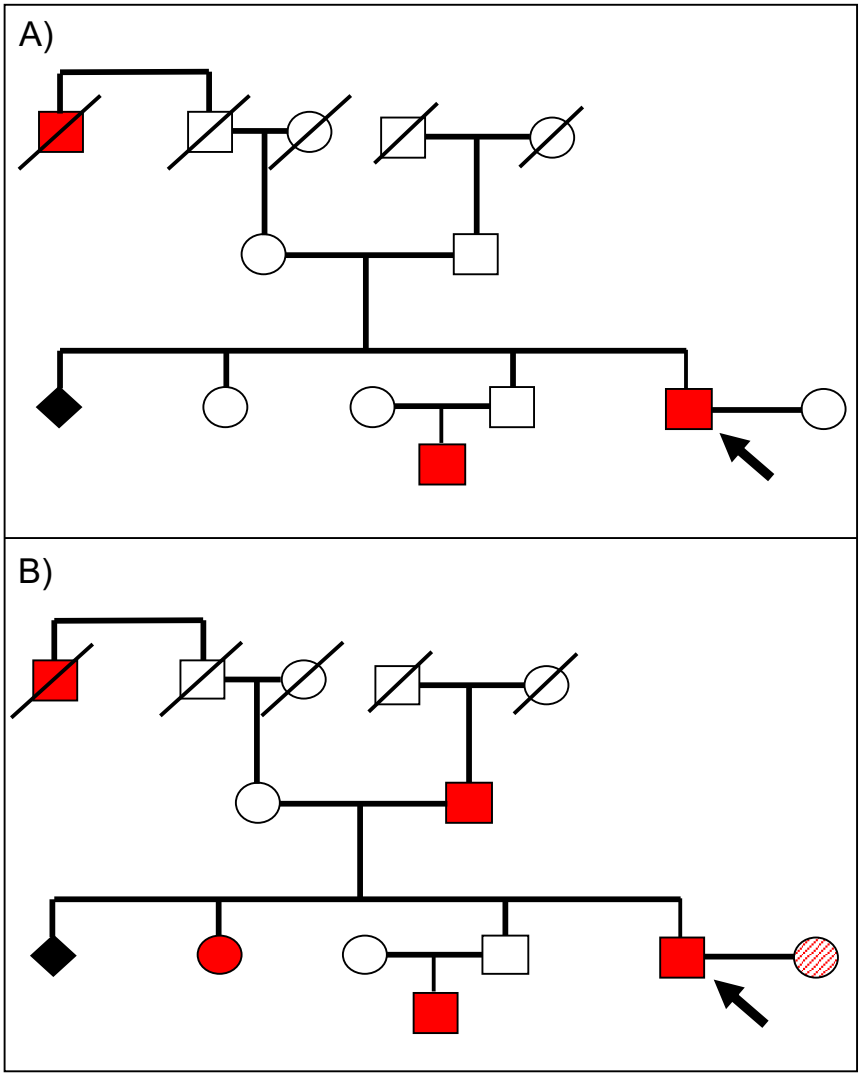

Figure 4. Previously undiagnosed cardiac lesions revealed in the index family. A) Prior to the scheduled echocardiogram and electrocardiogram for all involved family members, it was believed that only two living members were affected. **B)** Following echocardiogram and electrocardiogram, it was revealed that additional family members had previously undiagnosed cardiac structural abnormalities. The proband is represented by the arrow.

Table 5. Cardiac and extracardiac lesions in the index family members

Sample ID	Sex	Cardiac Structural and Electrophysiological Abnormality	Extracardiac Lesions
TA1.1	M	None	None
TA1.2	F	None	None
TA2.1	F	Left Bundle Branch Block	None
TA2.2	M	Bicuspid Aortic Valve Effacement of Sinotubular Junction Premature Ventricular Complexes	None
TA3.1	F	Effacement of Sinotubular Junction	High-Arched Oral Palate
TA3.2	M	None	None
TA3.3	M	Tricuspid Atresia Ventricular Septal Defect (Muscular) Atrial Septal Defect (Secundum) Cleft Mitral Valve Patent Ductus Arteriosus	Cleft Lip and Palate Fused Ribs “Butterfly” vertebrae
TA3.4	F	Subaortic Tissue Tag	None
TA3.5	F	None	None
TA4.1	M	Bicuspid Aortic Valve Ventricular Septal Defect (Muscular)	None

This table lists the observed structural and electrophysiological cardiac abnormalities in each member of the index family, as well as any known extracardiac defects.

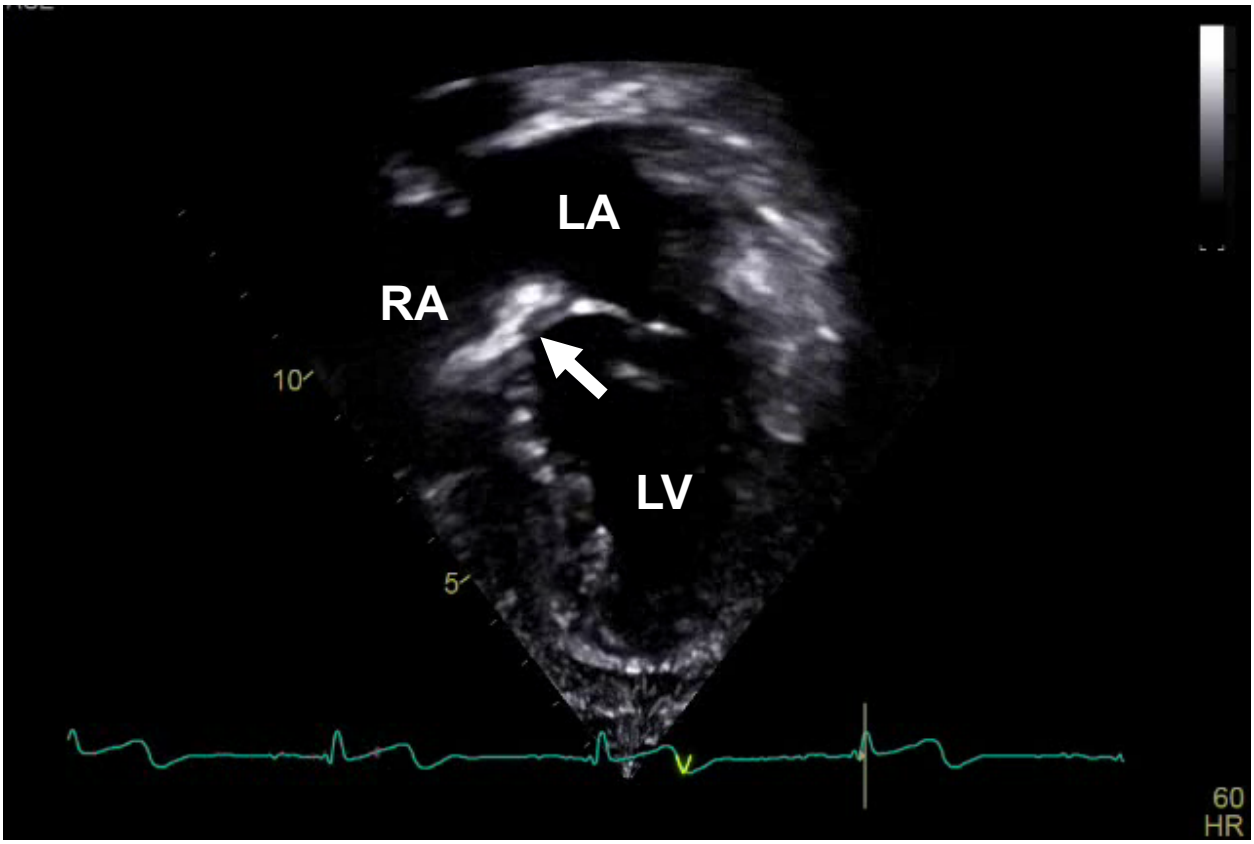

Figure 5. Tricuspid atresia in TA3.3. Apical 4-chamber view of the atretic tricuspid valve (arrow) and hypoplastic right ventricle. LV: Left Ventricle. LA: Left Atrium. RA: Right Atrium.

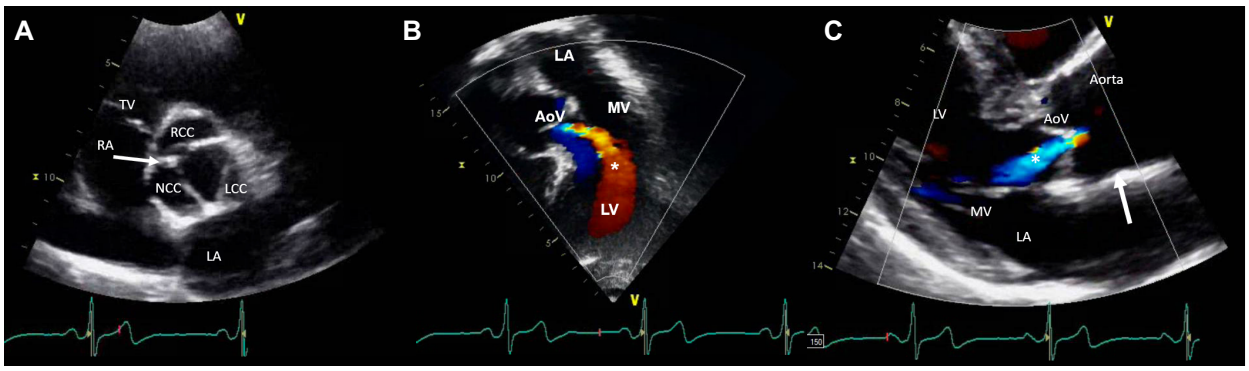

Figure 6. Bicuspid Aortic Valve in TA2.2. A) Parasternal short-axis view of the aortic valve in systole demonstrating thickening and fusion of the right-non-coronary commissure (arrow). **B)** Apical view with anterior angulation ("apical 5-chamber view"). Aortic insufficiency (*). **C)** Parasternal long-axis view with colour demonstrating aortic regurgitation (*). Effacement of the sinotubular junction and dilation (arrow) of the ascending aorta is also seen (#). AoV: Aortic Valve. MV: Mitral Valve. TV: Tricuspid Valve. LV: Left Ventricle. LA: Left Atrium. RCC: Right Coronary Commissure. NCC: Non-Coronary Commissure. LCC: Left Coronary Commissure.

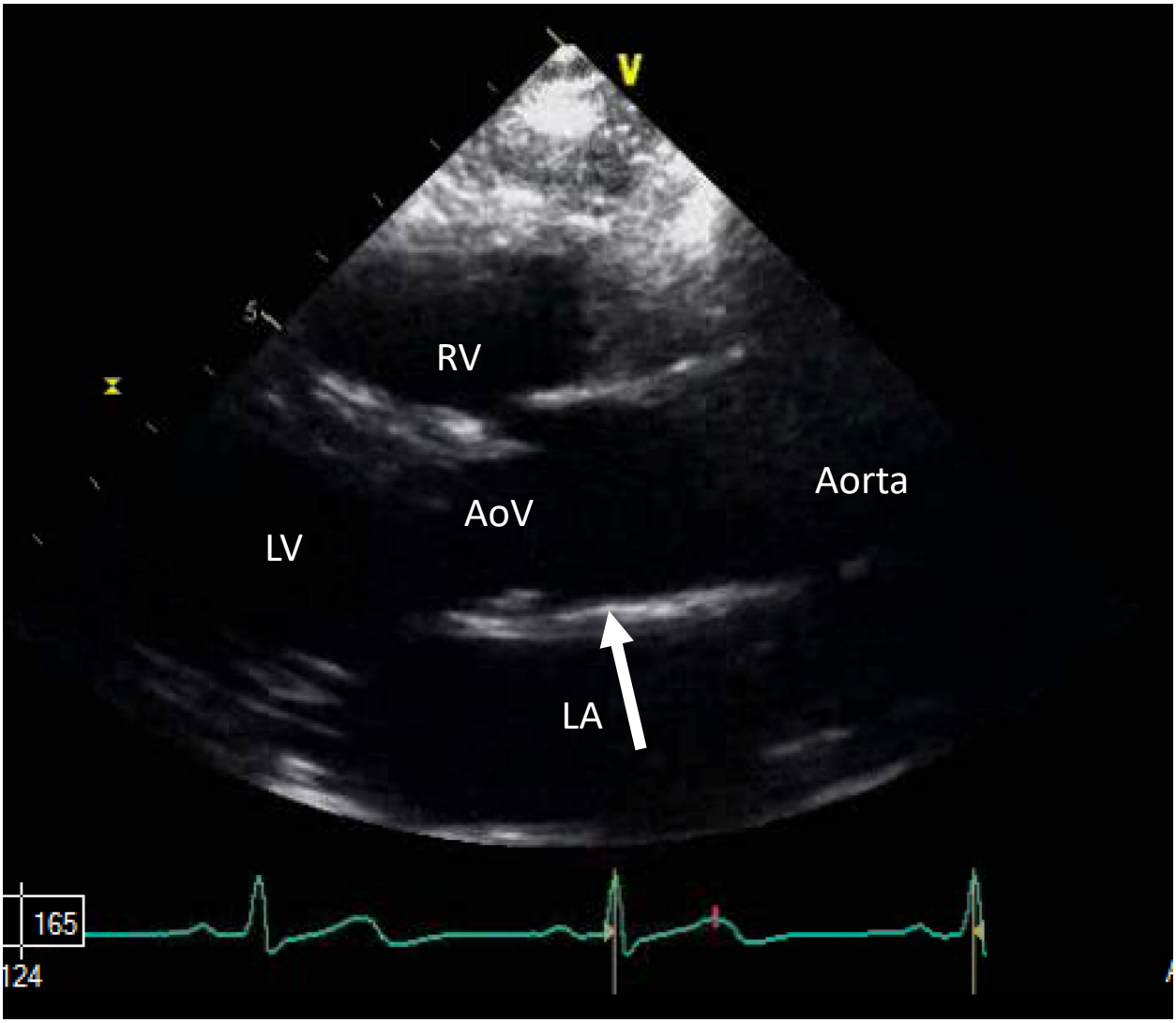

Figure 7. Sinotubular Junction in TA3.1. Parasternal long-axis view of the aortic valve with effacement of the sinotubular junction (arrow). The ascending aorta measures within normal limits. LA: Left Atrium. LV: Left Ventricle. RV: Right Ventricle. AoV: Aortic Valve

Methods

Identification of Subjects and Recruitment

The index family was recruited to the study and underwent informed consent prior to providing whole venous blood samples. The study was conducted under IRB STUDY00000103 (Genomic Basis for Tricuspid Atresia). A complete study echocardiogram and electrocardiogram were performed at TCH for each participant in the study under IRB H-41143 at Baylor College of Medicine.

Whole Exome Sequencing

The whole venous blood was provided to Baylor Genetics for DNA isolation and WES. This was conducted to avoid any potential bias or conflicts of interest. After the sequencing run, FASTQ files were generated. Variants need to be identified and annotated. While some of this information was provided initially, the data was re-analyzed two more times by independent bioinformaticians culminating with the final gene lists complete with all necessary annotations by the James Lupski lab at Baylor College of Medicine.

Variant Candidate Filtering

Candidate filtering began by isolating mutations identified in the proband, then narrowed down by MAF (≤ 0.001) using ExAC database data. This MAF was utilized based on guidelines and recommendations from the Cardiology Genetics department at Texas Children's Hospital, as it is expected that a rare condition such as tricuspid

atresia would only result from such a rare variant. The ExAC database was utilized for MAF because it possesses approximately 65,000 exomes. At this point, predictive algorithms were utilized to further narrow the mutations, CADD score (≥ 20), MutationTaster (D or A), and PolyPhen2 (Damaging/Deleterious) were then employed due to the reliability of their predictions (≥ 84% accurate).

Following this, the family was segregated into cohorts based on phenotypes (cardiac vs oral palate) and affected/unaffected status. This allowed for further narrowing of candidate mutations based on how certain phenotypes tracked within the family.

Literature Review of Candidate Variants

Lastly, an extensive literature review was conducted for each candidate variant that remained. In addition to publications, databases such as ClinVar, ExAC, and Ensembl were utilized to gain further insight into each variant. This provided additional predictions of possible protein product outcomes (i.e. nonsense mediated decay), as well as the known history of each variant in terms of previous studies or cases in which it has been reported.

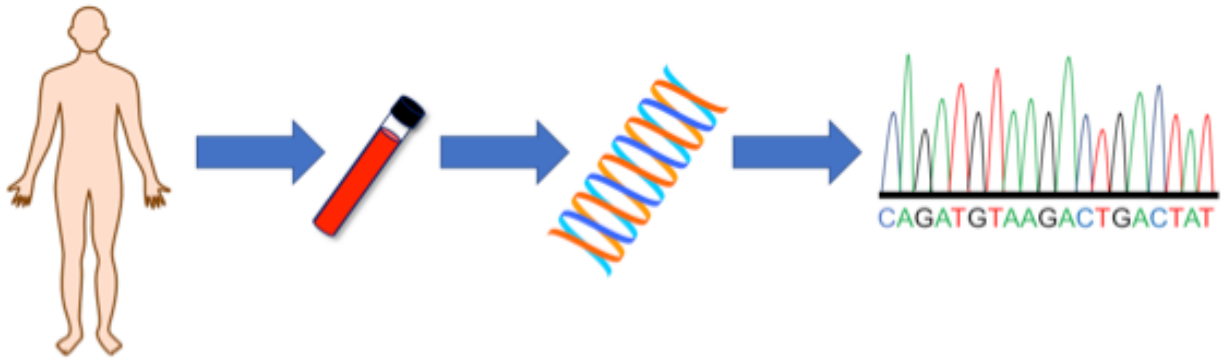

Figure 8. Whole exome sequencing workflow. Following informed consent, venous whole blood samples were acquired from each participant and utilized for DNA library preparation. WES was utilized due to the probability of identifying a genetic abnormality associated with the previously described cardiac lesions as 85% of known disease-related variants are located within the exome.

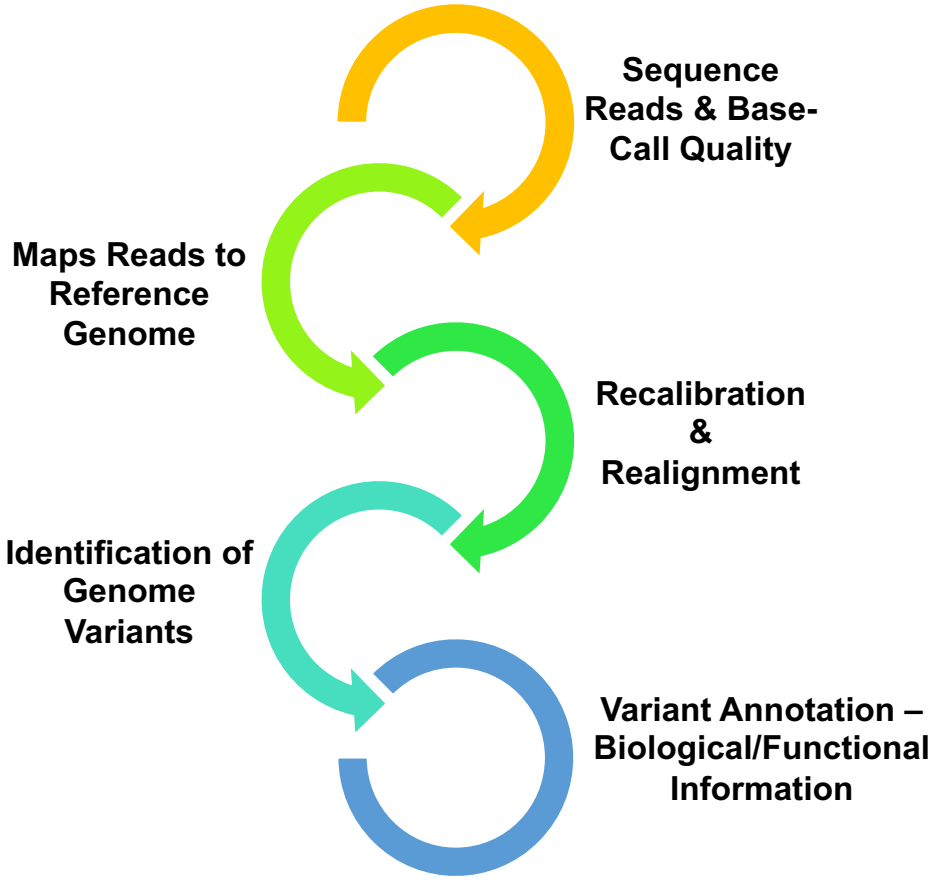

Figure 9. Bioinformatic analysis pipeline workflow. Following sequencing of the DNA samples, the generated reads are then subjected to bioinformatic processing. This process involves sequence reads, mapping to a reference genome (GRCh37/hg19), recalibration and realignment, identification of genomic variants (base-calling), and annotation of each variant with biological, functional, and potential pathogenicity information. This provides a list of single nucleotide variants and insertions/deletions within each individuals genome.

Results

The initial number of variants reported in the proband was 741. These were then narrowed down to only those with MAF ≤ 0.001, leaving 274 total variants. This MAF value was utilized based on the rarity of the condition (tricuspid atresia) and standards used by collaborating clinical institutions, requiring the variants of interest be more rare than in 1% of the population. However, this required further narrowing down to begin identifying the candidate variant, or variants, that may be associated with these conditions. To do this, the family was segregated into multiple cohorts to look at only those genes that track through the family in specific patterns. Additionally, intergenic and intronic variants were removed if they did not reside within an intron/exon splice junction location as we were looking for variants within the coding region or those that would affect post-transcriptional splicing. Finally, the three predictive algorithms were then utilized as additional variant prioritization.

CADD scores were utilized as one predictive algorithm. Variants with a CADD score below 20 were removed. Scores of ≥ 20 are in the top 1% of disease-associated variants, while scores of ≥ 30 indicate those within the top 0.1% of disease-associated variants. MutationTaster was also utilized and only variants that were scored as predicted to be disease-causing (D) and automatic disease-causing (A) were included. MutationTaster has a previously established accuracy rate of 88%(*71*). Polyphen2 was also utilized as a third predictive algorithm, including only variants with a Damaging/Deleterious designation. This algorithm has an accuracy rate of 84%.

Overall, for inclusion a variant must have a MAF ≤ 0.001, a CADD score ≥ 20, a MutationTaster prediction of D or A, and a PolyPhen2 prediction of

Damaging/Deleterious. If any of those 3 predicted any other result, indicating a potential for being benign, the variant was excluded.

Once these variants were segregated into the cohorts and all inclusion criteria were satisfied, that brought the list of potential candidate variants to 69 (Figure 11). The genes in which these variants resided were then subjected to extensive literature reviews.

Paternal Cohorts

Paternal cohort 1 had 6 mutations that were only present in the father, brother, and proband (Table 6). The variant of interest in this cohort was in the gene *TAB2*. *TAB2*, which encodes for TGF-Beta Activated Kinase 1 (MAP3K7) Binding Protein 2, is an activator of MAP3K7/TAK1. It forms a kinase complex with TRAF6, MAP3K7, and TAB1. This mutation, p.Ser450Thr (dbSNP rsID 141984528), is located slightly upstream of the coiled-coil domain and has a MAF of 0.0001 according to the ExAC database. The local site appears to be fully conserved across humans, rhesus, mouse, chicken, Xenopus, and zebrafish. Interestingly, it is predicted by Ensembl to result in nonsense mediated decay of the transcript and has a pLI score of 1.00.

Paternal cohort 2 had 8 mutations only present in the father, sister, and proband. Within this cohort, the remaining variant of interest resided within DHCR7. While this study was looking for variants associated with congenital heart disease and cardiac development, *DHCR7* primarily shows a strong association with oral palate malformations. *DHCR7*, which encodes for 7 dehydrocholesterol reductase, is involved in cholesterol biosynthesis. The mutation, c.385_412 del, is a 33-nucleotide deletion

that spans the intron-exon junction from exon 5 to the following intron. The local site is conserved in human, rhesus, mouse, and zebrafish. While no CADD score or predictive algorithms results were available, it was included due to being a rare frameshift deletion that is likely to result in nonsense mediated decay and has a literary history of association with craniofacial abnormalities.

Additionally, another paternal cohort contained another variant of interest in Factor 7. This variant was shared by the father, sister, brother, and proband. Factor 7 is of interest due to its role in prothrombin time (PT) as a blood clotting factor. The proband has presented previously with a prolonged PT, which has been shown in the literature to be caused by *F7* deficiencies.

Maternal Cohorts

The only maternal cohort with a variant of interest consisted of the maternal grandfather, mother, and proband (Table 7). The variant of interest is in the gene *P2RY2*, which is involved in Hippo/YAP signaling in cardiac development, regulating cell proliferation and migration in cardiac progenitors. The gene *P2RY2* encodes for Purinergic Receptor P2Y2 which is a G-protein-coupled receptor. The identified mutation, p.Asn285Asp, is located in extracellular domain 7, which is slightly upstream of the cytoplasmic domain. The local site is conserved in human, rhesus, and mouse, but not in zebrafish.

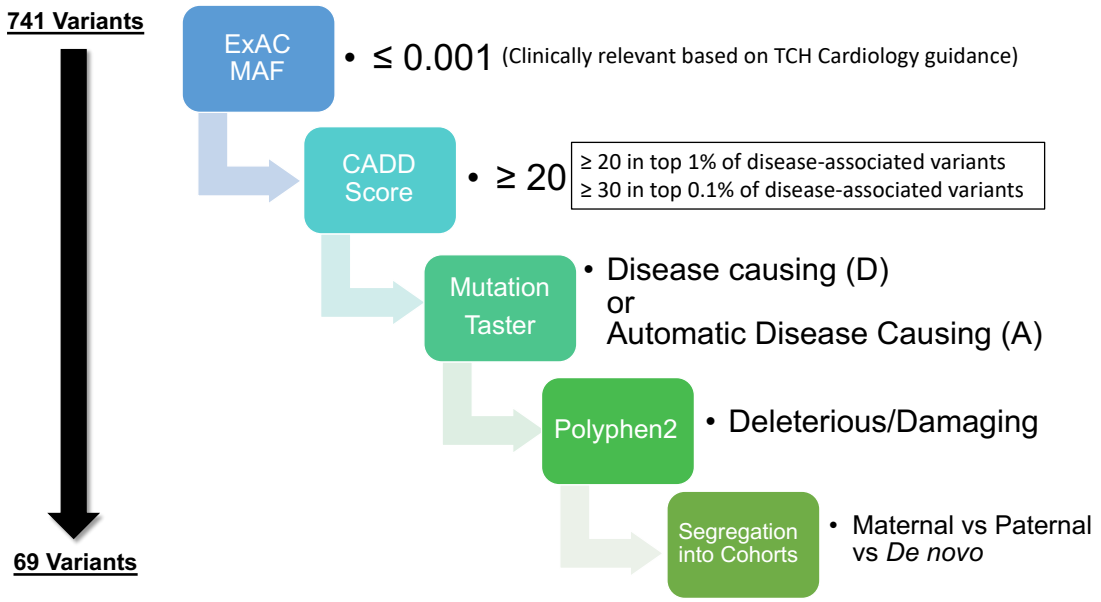

Figure 10. Candidate variant list generation workflow. Once the list of 741 variants in the proband was obtained, the list was narrowed in a step-wise process. This process included filtering by ExAC Database MAF, multiple predictive algorithms (CADD score, MutationTaster, and PolyPhen2), followed by segregation of variants into specific cohorts within the index family. This provided a final list of 69 genetic variants, which then required further investigation including extensive literature reviews for each gene.

Table 6. Candidate variants in the paternal cohorts

Cohort	Gene	Mutation Type	ExAC MAF	CMG MAF	CADD	MutationTaster	Polyphen 2
Paternal Cohort 1	BCAS3	Nonsynonymous SNV	0.00001	0.00018	23.1	D	D,D,P,D
	SYTL3	Nonsynonymous SNV	0.00002	0.00018	28.3	D	P,D
	TAB2	Nonsynonymous SNV	0.00010	0.00024	24.3	D	D,D
	HMCN1	Nonsynonymous SNV	0.00040	0.00073	24.6	D	D
	IQCA1	Frameshift Insertion	.	0.00018	.	n/a	n/a
	NUP88	Stop-gain SNV	.	0.00018	39	A	n/a
Paternal Cohort 2	BCCIP	Nonsynonymous SNV	0.00001	0.00018	28.1	D	D,D,D,D,D
	DHCR7	Frameshift Deletion	0.00006	0.00024	.	n/a	n/a
	HMX2	Nonsynonymous SNV	0.00020	0.00024	34	D	D
	COL6A1	Nonsynonymous SNV	0.00020	0.00024	26.8	D	D
	SNX25	Nonsynonymous SNV	0.00050	0.00128	28.6	D	D
	PDZK1IP1	Nonsynonymous SNV	0.00090	0.00049	29.9	D	D
	ATP8B2	Splicing	.	0.00018	.	n/a	n/a
	SUPT20H	Nonframeshift Deletion	.	0.00018	.	n/a	n/a
Paternal Cohort 3	BTBD11	Splicing	0.00001	0.00012	.	n/a	n/a
	C19ORF55	Nonframeshift Deletion	0.00003	0.00900	.	n/a	n/a
	DNAH3	Frameshift Deletion	0.00005	0.00018	.	n/a	n/a
	CCDC65	Nonframeshift Deletion	0.00005	0.00030	.	n/a	n/a
	AVIL	Nonsynonymous SNV	0.00020	0.00012	25.5	D	P,D
	NIT2	Frameshift Deletion	.	0.00012	.	n/a	n/a
Paternal Cohort 4	SVEP1	Nonsynonymous SNV	0.00002	0.00030	26	D	D
	TCF20	Nonframeshift Insertion	0.00003	0.00043	.	n/a	n/a
Paternal Cohort 5	C1ORF173	Nonsynonymous SNV	0.00002	0.00030	24.3	D	D,D
	PANK2	Nonsynonymous SNV	0.00002	0.00030	29.9	D	D
	F7	Nonsynonymous SNV	0.00004	0.00037	24.3	D	D,D,D
	CCNYL2	ncRNA Splicing	0.00130	0.00195	.	D	n/a
	ZC3H8	Nonsynonymous SNV	.	0.00024	27.3	D	D

This represents the composite list of all candidate variants of paternal origin in the proband once filtering was completed.

Table 7. Candidate variants in the maternal cohorts

Cohort	Gene	Mutation Type	ExAC MAF	CMG MAF	CADD	MutationTaster	Polyphen2
Maternal Cohort 1	PIPSL	ncRNA Exonic	0.00390	0.00402	.	n/a	n/a
	GOT1	Nonsynonymous SNV	.	0.00030	25.9	D	D
	VWF	Nonsynonymous SNV	.	0.00030	25.3	D	D
Maternal Cohort 2	UCMA	stopgain_SNV	0.00010	0.00043	36	A	n/a
	SH3BP5L	Nonsynonymous SNV	0.00060	0.00073	29.8	D	P,P,P,D
	OGDHL	Nonsynonymous SNV	0.00090	0.00103	33	D	D,D,D
	CCDC6	Nonframeshift Insertion	0.00170	0.00158	.	n/a	n/a
	HK1	Splicing	0.00200	0.00079	.	n/a	n/a
	C12ORF49	Nonframeshift Deletion	.	0.00018	.	n/a	n/a
	RABL2B	Nonsynonymous SNV	.	0.00018	26.3	D	D,D,D
	P2RY2	Nonsynonymous SNV	.	0.00018	26.1	D	D

This represents the composite list of candidate variants of maternal origin in the proband once filtering was completed.

Final Candidate Mutations

Following the analysis of each variant, including extensive literature reviews, the final list of candidate variants was brought down to 3 (Table 8). These three candidates match the MAF and predictive algorithm criteria and segregate based on specific patterns involving the proband, as well as have literature showing a pattern that either shows a high probability of involvement of the gene in cardiac development or oral palate development. The *TAB2* variant, possessed by the father, brother, and proband shows the most promise as the primary variant driving the cardiac phenotypes in the father and proband.

DHCR7 possessed a frameshift deletion in the father, sister, and proband may provide an explanation for the oral palate abnormalities within the sister and proband.

P2RY2, with a variant in the maternal grandfather, mother, and proband, may also play a minor role in the manifestation of CHD in the proband.

Finally, a nonsynonymous SNV in *PWWP2B* was identified as *de novo* in the proband.

Table 8. Final candidate variant list

Gene	Mutation Type	ExAC MAF	CMG MAF	CADD	MutationTaster	Polyphen2
TAB2	Nonsynonymous SNV	0.00010	0.00024	24.3	D	D,D
DHCR7	Frameshift Deletion	0.00006	0.00024	.	n/a	n/a
P2RY2	Nonsynonymous SNV	.	0.00018	26.1	D	D
PWWP2B	Nonsynonymous SNV	0.000032	0.000061	25.4	D	D

After reviewing all candidate variants of paternal and maternal origin, this is the final candidate variant list for the proband.

Discussion

Multiple attempts at identifying genes, then requiring additional bioinformatic analysis to gain necessary critical information following the identification of candidate mutations. The extensive literature review was undertaken to study each possible candidate. Assessments of their functions, expression patterns, signaling pathways, role in the development and developmental processes, as well as the potential outcome resulting in the specific variant were investigated.

After narrowing down the list of potential candidate variants, the 3 most-likely candidates based on segregation patterns and literature review were possessed in *TAB2*, *DHCR7*, *PWWP2B*, and *P2RY2*.

TAB2

TAB2 translocates from the membrane to the cytosol, following stimulation by IL-1, where it then activates TAK1. TAB2/TAK1 then forms a ternary complex with TRAF6. TAB2/TAK1/TRAF6 ternary complex formation leads to activation of NFκB and JNK signaling via MEKs and IKKß. Interestingly, dominant-negative *TAB2* prevents NFκB/JNK activation via IL-1 signaling (*72*).

TAB2/TAK1 signaling is activated by TNFα, TGF-β, RANKL, and BMP2/4/7 signaling. Following RANKL activation, TAB2/TAK1 stimulates the nuclear accumulation of NFATC1 and disruption of TAB2/TAK1 interaction prevents NFATC1 nuclear entry (*73*). This is very similar to what was observed in a previous study in which two

NFATC1 heterozygous mutations, identified in a tricuspid atresia patient, also caused cytosolic accumulation NFATC1 via nuclear entry failure (*32*).

Interestingly, BMP2/4/7 binds to BMPRII, stimulating TAB2/TAK1 interaction. This activates MAPK signaling, transactivating the expression of *NKX2.5* and *GATA4*, which are critical for cardiac development. At gestational week 5.5, the expression is primarily localized to the endocardial cushion of the outflow tract and the ventricular trabeculae (*74*). *TAB2* has also shown to play a role in endothelial cell migration, which is another critical process during cardiac development (*75*). A homozygous knockout in mice is embryonic lethal, while heterozygous knockout is postnatal lethal and incompletely penetrant (*76, 77*). The lack of full penetrance in the heterozygous model is relevant, as the pathologies associated with the presence of a heterozygous *TAB2* mutation appear to be incompletely penetrant in this family and are predicted to result in nonsense mediated decay.

Human patients with mutations and haploinsufficiency in *TAB2* have previously been shown to be associated with valve disease & outflow tract defects(*78*), as well as facial dysmorphisms and high oral palates. Some of the previously reported valve abnormalities include bicuspid aortic valve with aortic dilation, myxomatous and prolapsed mitral and tricuspid valves, and dysplastic tricuspid and pulmonary valves. Additionally, ventricular septal defect (VSD), atrial septal defect (ASD), hypoplastic aortic arch and coarctation of the aorta, along with patent ductus arteriosus (PDA) have also been reported in individuals with TAB2 disruptions. Electrophysiological disturbances such as premature ventricular contractions, atrial fibrillation, and supraventricular tachycardia have been observed. Several of these pathologies,

including bicuspid aortic valve with aortic dilation, premature ventricular complexes, ASD, VSD, PDA, and tricuspid valve abnormalities are present in two of the three individuals within this family that possess the observed TAB2 mutation

Interestingly, disruption of TAB2/TAK1 interaction creates an environment in which NFATC1 fails to enter the nucleus to exert its gene regulatory functions. As a result, it is retained with the cytosol and eventually degraded. This is of particular interest in the previously described tricuspid atresia patient who had 2 heterozygous variants within the *NFATC1* gene. These variants also caused a failure of nuclear entry, leading to retention within the cytosol and subsequent degradation. As *TAB2* plays a role in *NFATC1* signaling, it is possible that this is a potential player in the manifestation of tricuspid atresia.

DHCR7

DHCR7 is involved in cholesterol biosynthesis by catalyzing the conversion of 7-DHC to cholesterol. The alternate reaction catalyzes 7-DHC to Vitamin D3 (*79*). It appears to have a role in development, however, it's much more closely associated with craniofacial development as opposed to cardiac. DHCR7 knockout mice generally present with cleft lip and palate, however, it is not completely penetrant (*80-83*). DHCR7 mutations are generally associated with Smith-Lemli-Opitz syndrome (SLOS), which is an autosomal-recessive disease with multiple congenital anomalies, including cardiac defects and oral palate abnormalities (*84*).

DHCR7 is co-expressed with Sonic Hedgehog (SHH) during midline development and functions as a regulator of SHH and BMP expression. It is also known

to be expressed in palatal mesenchymal cells, where SHH and BMP2 are required for oral palate development. Interestingly, in the absence of DHCR7, SHH and BMP2 expression are down-regulated (*81*). Additionally, the accumulation of Vitamin D3 in the absence of DHCR7 activity inhibits SMO in the Sonic Hedgehog (SHH) pathway.

P2RY2

P2RY2 is a G-protein-coupled receptor that responds to extracellular purine and pyrimidine nucleotides (*85*). Its primary cardiac-associated function appears to be its function involving the Hippo pathway. P2YR2 activation protects cardiomyocytes from hypoxia in vitro and reduces post-ischemic myocardial damage *in vivo* (*86*) and it is upregulated in dystrophic hearts(*87*). Stimulation of P2YR2 by extracellular nucleotides enhances the proliferation and migration of human cardiac progenitors via YAP-mediated activation of Hippo pathway(*88*). The Hippo pathway is critical for cardiac development & regeneration and it inhibits cell proliferation, as well as promoting apoptosis. Additionally, it has been shown to regulate the fates of stem and progenitor cells and limit cell size (*89*). Due to its role in cardiovascular development and disease, P2RY2 has recently begun to be recognized as a potential therapeutic target (*90*).

Only one individual possessing this mutation has a cardiac phenotype (TA3.3), however, it is possible that in a polygenic model the combination of this mutation along with the additional observed mutations may have aided the manifestation of the additional severity and complexity of the phenotype observed in this individual. P2RY2 activates a phosphatidylinositol-calcium second messenger system and is involved in many cellular functions, such as proliferation, apoptosis, and inflammation.

PWWP2B

While not much is known about *PWWP2B*, it has been shown to be up-regulated in *MESP1*-expressing embryonic stem cells (ESCs) (*91*). This is of particular interest as the *MESP1*-expressing ESCs are those that can be differentiated into a cardiac fate. Other than that, the main knowledge about this gene can only be inferred from its paralog, *PWWP2A*, which binds chromatin near highly-expressed genes and appears to play a role in neural crest stem cell migration and differentiation, a process incredibly important to cardiac and craniofacial development.

Potential Hedgehog Pathway Effects

The combined mutations within this family all appear to play a role in critical SHH signaling processes during development that may be involved in the proper formation of the heart and craniofacial region. *SHH* signaling influences cellular migration and motility and is required for the second heart field (SHF) and oral palate development. It works in concert with other signaling pathways to confer cardiac cell identity and drives atrioventricular septation in the second heart field (*92*). Decreased interaction of TAB2 with TAK1 likely leads to decreased transactivation of *GATA4* via MAPK activation, following stimulation by *BMP* signaling. It is known that *GATA4* is required for activation of SHH in the second heart field via interaction with SMO. Decreased cholesterol biosynthesis from 7-DHC leads to the accumulation of Vitamin D3, which has been shown to bind and inhibit SMO, leading to decreased SHH activity.

Nephew

Potentially an isolated case of BAV with VSD as the genes observed within the WES does not appear to align with the affected individuals within the family. His mother is also healthy; however, her family's medical history is unknown at this point. It is possible this is the result of multiple variants (some from each parent) creating the genetic environment resulting in manifestation of phenotypes despite neither parent being affected.

Polygenic Model

In addition to the primary variant, *TAB2*, there are many other mutations playing a role that may also begin to explain penetrance issues seen in this family pedigrees like this. *TAB2* variant in 2 affected and 1 unaffected; it may be that *TAB2* sets the stage, but additional variants may still be "required" for the manifestation of cardiac phenotypes to arise. If there is a "backup mechanism", as exists with some genes/proteins, additional variants/disturbances may be necessary to propagate a deleterious effect. So, it is possible that additional mutations in these 2 affected individuals are causing the manifestation of their conditions. They have different sets of variants, which may lead to their unique diagnoses and complexities.

Chapter 5: Conclusion

Retrospective Chart Review

The overall cohort was 56.8% male, however, the Type 2 cohort (76.9%) had a surprisingly disproportionate number of males compared to the other cohorts. At this point, it is unknown why there is such a large disparity between the number of males and females in those groups. Additionally, the generally published overall mortality rate of 15% was slightly decreased in this cohort at 12.8%. However, it is possible that this number could go up if those lost to follow-up had been updated.

The primary diagnosis category, Type 1 cohort, was comprised of 64.5% (152) of patients while Type 2 cohort constituted 22.2% (52) of patients. Surprisingly, the Type 2 cohort had 13 individuals (25.0%) with pulmonary valve abnormalities (7 pulmonary atresia, 6 pulmonary stenosis). This is notable as it would be considered a left-sided lesion in this cohort due to the transposition of the great arteries. Overall, secundum ASD (127, 54.2%) and muscular VSD (108, 46.1%) were the most prevalent septal defects. PDA was present in 98 (41.8%) patients.

As expected, the most common procedures were the BT shunt, bidirectional Glenn, and Fontan. While the Type 1 cohort primarily had a more straightforward path to palliative surgical intervention, those in the Type 2 cohort underwent a much more complicated surgical schedule consisting of DKS/Norwood procedures, repair of coarctation of the aortic arch, and arterial switch.

While 22.2% of patients in the unrelated tricuspid atresia cohort were classified as Type 2, they constituted 38.5% of individuals that underwent genetic testing.

However, only one variant of interest was observed in this cohort (*MYH6*). *NFATC1*, *MYH6*, *FBN1*, and *BMP1* were identified in the Type 1 cohort, while two L-TGA patients in the Type 3/4 cohort both had *MYH6*, *ERCC6L2,* and *RTEL1* variants. Despite 40 patients undergoing CMA testing, there were no overlapping regions of gains or loss. While there were two individuals with chromosome 1 (1q41 loss, 1q42.13 gain), chromosome 2 (2q21.1 gain, 2q11.2 loss), chromosome 3 (3p21.1 gain, 3q25.2 gain), and chromosome 10 (10p12.2 gain, 10q26.3 loss).

Index Family Exome Study

Following the findings of the retrospective chart review, a family with recurring congenital heart disease was investigated for the potential candidate variants associated with tricuspid atresia and bicuspid aortic valve. After filtering 741 variants by ExAC MAF (≤ 0.001), CADD score (≥ 20), MutationTaster (probable disease-causing or automatic disease-causing), and PolyPhen2 (Damaging), and segregating into specific cohorts based on phenotypes, 69 variants remained.

The MAF utilized (≤ 0.001) was chosen as a result of the rarity of tricuspid atresia and through consultation with members of the cardiology genetics department at Texas Children's Hospital. The three predictive algorithms were used due to their higher frequency of accuracy of prediction. MutationTaster has been shown to be 88% accurate, while PolyPhen2 is 84% accurate; CADD scores do not appear to have a published accuracy at this point. Each variant was extensively investigated using literature, databases such as ClinVar and ExAC, and additional resources providing signaling pathways and GO analysis.

After the review was completed, three variants remained. A variant in *TAB2*, possessed by the father, brother, and proband is the most likely candidate to be associated with recurring congenital heart disease within this family. *DHCR7* was the second candidate gene, however it is associated most closely with oral palate abnormalities and was possessed by the father, sister, and proband. The proband is primarily of interest due to cleft lip and palate, while the sister was born with an abnormally high-arched palate, and an additional case of cleft palate exists on the paternal side of the family. A variant in *P2RY2* was possessed by the mother, brother, nephew, and proband. It is of interest as it has shown a role in cardiac development, particularly at the stage of cardiac progenitors prior to differentiation into the various cardiac cell fates. Lastly, a variant in *PWWP2B* was identified as de novo in the proband. While not much is known about this gene, it has been shown to be up-regulated in *MESP1*-expressing ESCs (*91*) and is a paralog of a gene (PWWP2A) that is involved in chromatin binding, as well as neural crest stem cell migration and proliferation.

A Potential Link Between Phenotypes

Based on the results of the chart review and the family WES, it is possible that BAV and tricuspid atresia can be linked, albeit very infrequently. While not common, 4 of 234 (1.7%) patients had BAV with tricuspid atresia, indicating that it does occur together. Additionally, the father and proband within the family study both had the same *TAB2* variant. So, it is plausible that these can result from the same variant within the same family due to the pleiotropic nature of CHD.

Despite the infrequency of occurrence together, this still indicates there may be a link that they can occur together or within the same family. This may be due to due to the pleiotropic nature of CHD or even additional variants playing a role. Proband has an abnormality, although considered clinically insignificant, on the same chromosome (15q21.3) as the patient in chart reviews with BAV & tricuspid atresia (15q11.2).

Additionally, both the aortic valve and tricuspid valve are formed from the second heart field, which is a pool of cardiac progenitor cells that also give rise to the right ventricle, atrial and ventricular septae, and the outflow tract. While these phenotypes have been previously considered mutually exclusive, due to the new findings from this study and the nature of the second heart field, it is very plausible they may be related genetically despite the currently infrequent observation together.

GO Analysis via Cytoscape

Of the genes affected in the unrelated tricuspid atresia cohort, the proband of the family WES study, and in the literature, all 10 genes show interconnection via GO analysis (*93*) (Figure 12). Some of the GO terms associated with cardiac processes include: cardiac morphogenesis, cardiac chamber morphogenesis, stem cell differentiation, embryonic organ morphogenesis, cardiomyocyte differentiation, regulation of epithelial-to-mesenchymal transformation, heart looping, embryonic heart tube morphogenesis, blood vessel development, cardiac atrium morphogenesis, ventricular septum development, cardiac septum development, endocardial cushion development, vascular endothelial growth factor production, OFT septum development, and cardiac right ventricle formation. Extracardiac processes include palate

development, embryonic cranial skeleton morphogenesis, skeletal system development, and bone remodeling.

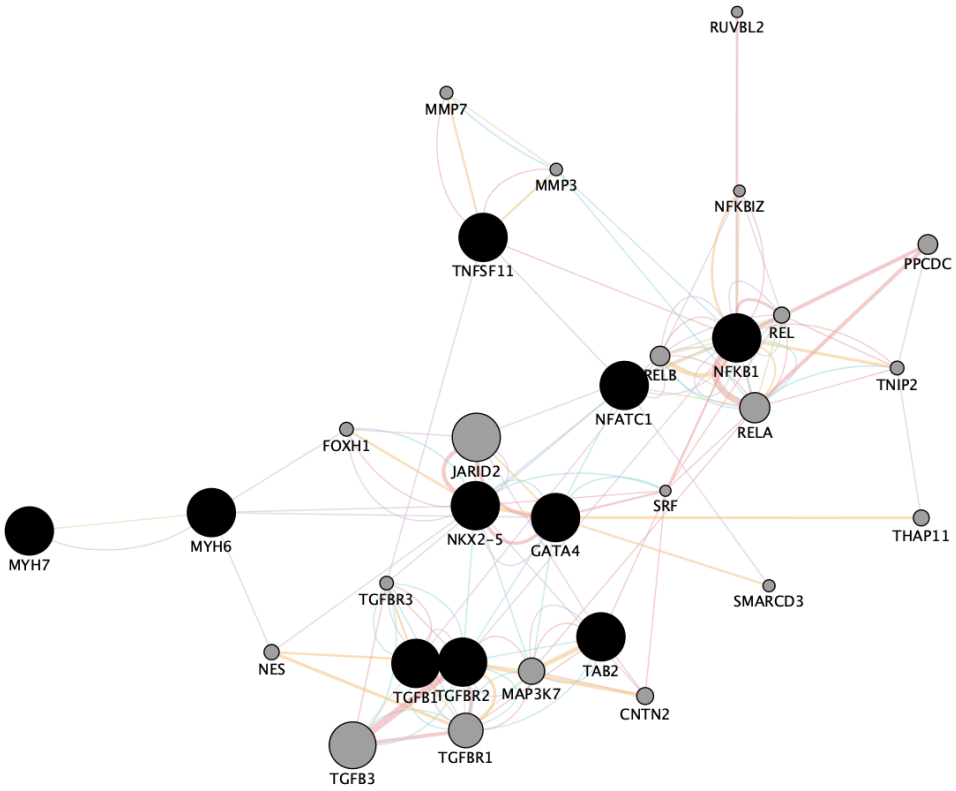

Figure 11. GO analysis network of candidate genes in tricuspid atresia patients. Once the final candidate gene list was compiled from the literature, retrospective chart review, and family exome study, GO analysis was run utilizing Cytoscape to assess for potential interaction networks involving more than one gene. This is the result showing the potential interactions and co-expression of each related candidate gene.

TGF-β and RANKL Signaling in Tricuspid Atresia

Mutations in *TAB2* and *NFATC1* result in cytosolic accumulation of NFATC1, preventing nuclear entry to proceed with its proper function (*32, 73*). However, the involvement of the remaining candidate genes from the literature and chart reviews was unclear. After performing GO analysis, their potential interaction became more apparent. Excitingly, all of the candidate genes fit within a single process during early development involving *TGF-β* and *RANKL* signaling (Figure 13).

TGF-β/RANKL signaling leads to the single pathway that is involved in 17 cases of tricuspid atresia from the literature, chart reviews, and the WES proband. Starting at *TGF-β* and *RANKL*, continuing to *GATA4* and *NKX2.5* and branching off with *NFκB* as well. TGFB1 then binds to TGFBR2, which then dimerizes with TGFBR1 on the cell surface. This leads to the activation of TAK1 (*94*), which then binds to TAB2 and TRAF6 forming a ternary complex (*95*). Alternatively, RANKL binds with its receptor, activating TRAF6 to form a trimeric complex with TAK1/TAB2. This ternary complex is then involved in multiple processes.

After formation, the ternary complex activates NFκB via stimulating dissociation of NFκB from its inhibitor, IKKβ. NFκB is then able to translocate to the nucleus and activate *NFATC1* expression (*96, 97*). At this point, NFATC1 exits the nucleus during translation/post-translational processing. The TAB2/TAK1/TRAF6 ternary complex then stimulates the entry of NFATC1 into the nucleus (*73*). NFATC1 and GATA4 then regulate the expression of MYH6, *MYH7* (*98*) and also directly targets *NKX2.5* for activation(*99*). The heavy involvement of *NFATC1* is of particular importance in this case, as *NFATC1* expression is limited to pro-valve endocardial cells in the AV cushion

and outflow tract during development (*100*) which are the cells from which the tricuspid and aortic valves are derived.

However, *TGF-β* signaling has already been implicated in Loeys-Dietz syndrome which involves aortopathy (*101*). The primary distinction observed between Loeys-Dietz and what is seen with *TGF-β* signaling in the tricuspid atresia patients previously described is that Loeys-Dietz appears to be more closely associated with SMAD signaling mechanism, while the tricuspid atresia patients' mutations mostly lie within a *SMAD*-independent pathway.

Interestingly, *NFATC1* mice also have osteoclastogenesis abnormalities(*97*). While not proven experimentally, it is possible *TAB2* mutations could lead to similar osteoclastogenesis problems (*102*) as the proband has skeletal abnormalities and GO analysis of these genes also show palate and skeletal development involvement.

Penetrance and Expressivity

The evidence indicating the *TAB2* variant being the primary cause of CHD in this family may seem problematic upon first glance, as the two affected individuals possessing this variant display different phenotypes while a third individual with this variant is completely unaffected. However, this may likely be explained by the concepts of penetrance and expressivity, more specifically reduced penetrance and variable expressivity (*103*).

Penetrance refers to the number of individuals within a population or cohort that possess the same variant and also have the associated phenotype. With reduced penetrance, there may be individuals within a group that possesses the disease-

associated variant but do not have the associated phenotype. Using this family as an example, the father, proband, and brother all have the TAB2 variant but only the father and proband are affected with the associated phenotype, congenital heart disease. The brother also possesses this variant but is completely unaffected. This would indicate, as is observed in the heterozygous animal models (*76*), that this *TAB2* variant exhibits incomplete or reduced penetrance as not all individuals possessing it are affected with the associated phenotype.

Additionally, expressivity is the degree of phenotypic expression associated with a particular variant. For individuals with the same disease-associated variant, not all may have an identical phenotype. This is referred to as variable expressivity. Again, using this family as an example of variable expressivity, the father and proband are both affected and possess the same *TAB2* variant. However, these individuals have different phenotypes of congenital heart disease. While the father has bicuspid aortic valve disease and effacement of the sinotubular junction, the proband has tricuspid atresia with ASD, VSD, and PDA. Both individuals' phenotypes fall under the same overall classification of congenital heart disease, but they are not the same phenotype.

It is likely that there are additional factors at play when encountering situations of reduced penetrance and variable expressivity. There may be environmental factors, modifier genes, allelic variation, and even the genetic environment playing a role as well. In this case, it is believed that this phenotype may be the result of polygenic inheritance in which multiple variants play a role in the manifestation of this phenotype and those possessing the single *TAB2* mutation are not necessarily going to be affected without also possessing other variants. As observed in the previously described WES

study of 342 left-sided lesion patients, complex inheritance is an emerging theme in families with CHD (*69*).

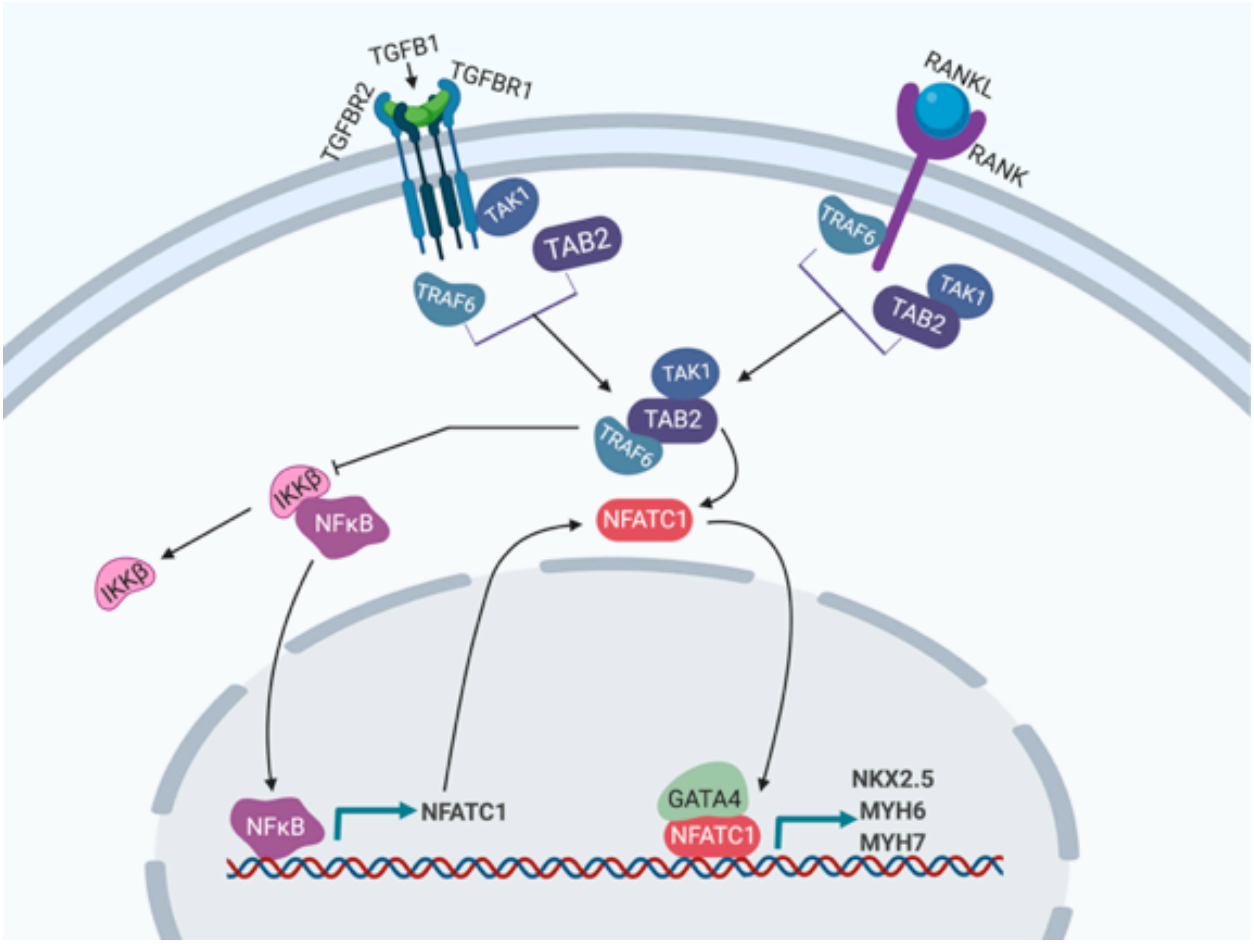

Figure 12. Signaling pathway involvement of tricuspid atresia candidate genes. Upon establishing the potential for signaling network involvement of multiple candidate genes, literature searches revealed the nature of the interactions. Beginning with stimulation by TGFB1 or RANKL, a process during early development is initiated in which TAB2/TAK1/TRAF6 release NFκB from inhibition, allowing it to activate expression of NFATC1. The TAB2 trimeric complex then stimulates NFATC1 to enter the nucleus, where it works in coordination with GATA4 to regulate the expression of Nkx2.5, MYH6, and MYH7, all of which are candidate genes in tricuspid atresia.

Polygenic CHD

Chromosomal defects and single-gene disorders can cause CHD but represent less than 20% of CHD (*4*). In the study of 342 left-sided lesion patients, the scope of genetic heterogeneity within similar forms of CHD was highlighted as 27 candidate genes were identified (*69*). This genetic heterogeneity is not limited to left-sided lesions and also presents an additional challenge when attempting to identify the genetic etiology of a single cardiac defect. Additionally, CHD can follow a very complex inheritance pattern, as many family pedigrees with recurring CHD may have unaffected offspring of affected parents, yet the children of those unaffected individuals are. This penetrance may possibly be explained by a polygenic model in which effector mutations are presently augmenting the potential for manifestation of a phenotype within an individual. The polygenic model indicates that it's not a single gene, but multiple variants playing a role to result in the manifestation of a specific phenotype. For instance, complex conditions such as tricuspid atresia that involve failure to remodel the AV cushion into a valve on the right side, along with failure to form the atrial and ventricular septae via proliferation would serve as strong candidates for such a model. As they involve multiple process failures, such as proliferation, migration, and apoptosis, it is possible that multiple genes controlling these processes possess variants that result in the developmental failures that result in the manifestation of the condition. This family is a good model for this model, as the father has bicuspid aortic valve with no septal defects, while one son was born with tricuspid atresia and another son was unaffected. Yet, the unaffected son produced an affected child with bicuspid aortic valve and ventricular septal defect. Not only is there an issue of penetrance, but those defects are

also likely caused by different processes – failure to proliferate and migrate in the ventricular septum, along with incomplete apoptosis during aortic valve remodeling. While TA3.2 is unaffected, along with TA3.5, their child is affected.

Chapter 6: Future Directions

Whole Exome Sequencing of Tricuspid Atresia Cohort

While this study has made significant toward classifying and characterizing tricuspid atresia, there is still much to be done to begin to unravel the genetic etiology of this complex condition. Currently, there are samples available for research-based genetic testing from approximately 50 unrelated tricuspid atresia patients at Texas Children's Hospital. These samples still must be matched with a medical record number and require deep cardiac phenotyping prior to sequencing to identify specific lesions within each individual that may affect how the sequencing results are segregated among the cohort. Following the deep phenotyping, these samples will be subjected to WES and subsequent genetic analysis similar to what was conducted in the Family WES study previously described. Common genes and pathways will be sought to hopefully begin to understand the genetic underpinnings of tricuspid atresia.

TAB2 had not previously been shown to be associated with tricuspid atresia, as it has been observed in other phenotypes such as BAV and VSD. However, as a novel candidate for tricuspid atresia, and its involvement in the previously described signaling pathway, it represents an exciting new target for those with tricuspid atresia and other right-sided lesions. Additionally, the involvement of *DHCR7* in syndromes that may be associated with CHD and its presence as a candidate within this family warrants further investigation in those with CHD as a possible target as well.

Functional Models of Candidate Variants

Following this, the genetic variants that remain as candidates from the family WES and tricuspid atresia cohort will be utilized to create cell models. Monogenic and polygenic models will be developed utilizing human induced pluripotent stem cells (iPSCs) and CRISPR/Cas9 technology to mimic the variants identified within the tricuspid atresia cohort and family (*104, 105*). The iPSC model will be pushed toward cardiac fate into cardiac progenitors. From this point, they will be further pushed toward various cardiac cell fates including cardiomyocytes, cardiac endothelial as seen in AV valves, smooth vascular cells typical of the great arteries, and pacemaker cells. These will be assessed for morphological, electrophysiological, functional (migration, contraction), and gene expression changes.

Once the primary candidates are narrowed down further with these functional cell models, the remaining few candidates will then be utilized to create mouse models in an attempt to mimic the tricuspid atresia phenotype (*105*). This will provide significant information regarding the role of each gene in cardiac development, as well as the pleiotropic nature of these variants as additional phenotypes are observed in the animal models. The mice will be utilized to investigate cardiac function, electrophysiological disturbances, and gross physiological structural defects within the hearts.

These studies, once complete will provide significant breakthroughs regarding tricuspid atresia and conotruncal abnormalities, including pulmonary atresia and transposition of the great arteries. It will also provide an incredible window into the role of these genes, many of which may not have been investigated in a cardiac model to this point.

Bibliography

1. M. E. Oster, K. A. Lee, M. A. Honein, T. Riehle-Colarusso, M. Shin *et al.*, Temporal trends in survival among infants with critical congenital heart defects. *Pediatrics* **131**, e1502-1508 (2013).

2. S. M. Gilboa, O. J. Devine, J. E. Kucik, M. E. Oster, T. Riehle-Colarusso *et al.*, Congenital heart defects in the united states: Estimating the magnitude of the affected population in 2010. *Circulation* **134**, 101-109 (2016).

3. C. Vecoli, Congenital heart disease the crossroads of genetics, epigenetics, and environment. *Current Genomics*, (2014).

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

5. C. B. Arrington, S. B. Bleyl, N. Matsunami, G. D. Bonnell, B. E. Otterud *et al.*, Exome analysis of a family with pleiotropic congenital heart disease. *Circ Cardiovasc Genet* **5**, 175-182 (2012).

6. G. M. Blue, D. Humphreys, J. Szot, J. Major, G. Chapman *et al.*, The promises and challenges of exome sequencing in familial, non-syndromic congenital heart disease. *Int J Cardiol* **230**, 155-163 (2017).

7. B. J. Landis, S. M. Ware, The current landscape of genetic testing in cardiovascular malformations: Opportunities and challenges. *Front Cardiovasc Med* **3**, 22 (2016).

8. C. Preuss, M. Capredon, F. Wunnemann, P. Chetaille, A. Prince *et al.*, Family based whole exome sequencing reveals the multifaceted role of notch signaling in congenital heart disease. *PLoS Genet* **12**, e1006335 (2016).

9. C. T. Mai, T. Riehle-Colarusso, A. O'Halloran, J. D. Cragan, R. S. Olney *et al.*, Selected birth defects data from population-based birth defects surveillance programs in the united states, 2005-2009: Featuring critical congenital heart defects targeted for pulse oximetry screening. *Birth Defects Research Part A: Clinical and Molecular Teratology* **94**, 970-983 (2012).

10. A. E. Lin, L. Rosti, Tricuspid atresia in sibs. *J Med Genet* **35**, 1055-1056 (1998).

11. J. W. Grant, Congenital malformations of the tricuspid valve in siblings. *Pediatr Cardiol* **17**, 327-329 (1996).

12. P. S. Rao, A unified classification for tricuspid atresia. *Am Heart J* **99**, 799-804 (1980).

13. P. S. Rao, Consensus on timing of intervention for common congenital heart diseases: Part i - acyanotic heart defects. *Indian J Pediatr* **80**, 32-38 (2013).

14. P. S. Rao, Consensus on timing of intervention for common congenital heart diseases: Part ii - cyanotic heart defects. *Indian J Pediatr* **80**, 663-674 (2013).

15. P. J. Czesniewicz, J. Kusa, Approaching the 50(th) anniversary of the first fontan procedure. What is the current state of treatment provided to patients with functional single ventricles? *Kardiochir Torakochirurgia Pol* **14**, 186-191 (2017).

16. Y. d'Udekem, A. J. Iyengar, J. C. Galati, V. Forsdick, R. G. Weintraub *et al.*, Redefining expectations of long-term survival after the fontan procedure: Twenty-five years of follow-up from the entire population of australia and new zealand. *Circulation* **130**, S32-38 (2014).

17. U. Kiran, S. Aggarwal, A. Choudhary, B. Uma, P. M. Kapoor, The blalock and taussig shunt revisited. *Ann Card Anaesth* **20**, 323-330 (2017).

18. R. Mainwaring, J. Lamberti, K. Uzark, R. Spicer, Bidirectional glenn: Is accessory pulmonary blood flow good or bad? *Circulation* **92**, 294-297 (1995).

19. K. Sughimoto, K. Okauchi, D. Zannino, C. P. Brizard, F. Liang *et al.*, Total cavopulmonary connection is superior to atriopulmonary connection fontan in preventing thrombus formation: Computer simulation of flow-related blood coagulation. *Pediatr Cardiol* **36**, 1436-1441 (2015).

20. M. Gewillig, S. C. Brown, The fontan circulation after 45 years: Update in physiology. *Heart* **102**, 1081-1086 (2016).

21. S. Camposilvan, O. Milanesi, G. Stellin, A. Pettenazzo, L. Zancan *et al.*, Liver and cardiac function in the long term after fontan operation. *Ann Thorac Surg* **86**, 177-182 (2008).

22. R. Romero, Liver in congenital heart disease: Implications of the fontan procedure for pediatric and adult liver specialists. *Clinical Liver Disease* **2**, 210-214 (2013).

23. B. Marino, M. C. Digilio, G. Novelli, A. Giannotti, B. Dallapiccola, Tricuspid atresia and 22q11 deletion. *Am J Med Genet* **72**, 40-42 (1997).

24. R. E. Ensenauer, A. Adeyinka, H. C. Flynn, V. V. Michels, N. M. Lindor *et al.*, Microduplication 22q11.2, an emerging syndrome: Clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* **73**, 1027-1040 (2003).

25. P. Hu, X. Q. Ji, C. Yang, J. J. Zhang, Y. Lin *et al.*, 22q11.2 microduplication in a family with recurrent fetal congenital heart disease. *European Journal of Medical Genetics* **54**, E433-E436 (2011).

26. M. J. Wat, O. A. Shchelochkov, A. M. Holder, A. M. Breman, A. Dagli *et al.*, Chromosome 8p23.1 deletions as a cause of complex congenital heart defects and diaphragmatic hernia. *Am J Med Genet A* **149A**, 1661-1677 (2009).

27. C. W. Yu, H. Chen, R. W. Baucum, A. M. Hand, Terminal deletion of the long arm of chromosome 4. Report of a case of 46, xy, del(4)(q31) and review of 4q- syndrome. *Ann Genet* **24**, 158-161 (1981).

28. A. Brand, R. M. Reifen, Y. Armon, E. Kerem, E. Horenstein *et al.*, Double mitral valve, complete atrioventricular canal, and tricuspid atresia in chromosomal 3p-syndrome. *Pediatr Cardiol* **8**, 55-56 (1987).

29. A. M. Ranger, M. J. Grusby, M. R. Hodge, E. M. Gravallese, F. C. de la Brousse *et al.*, The transcription factor nf-atc is essential for cardiac valve formation. *Nature* **392**, 186-190 (1998).

30. J. L. De la Pompa, Role of the nfatc transcription factor in morphogenesis of cardiac valves and septum. *Nature* **392**, (1998).

31. A. Yehya, R. Souki, F. Bitar, G. Nemer, Differential duplication of an intronic region in the nfatc1 gene in patients with congenital heart disease. *Genome* **49**, 1092-1098 (2006).

32. Z. Abdul-Sater, A. Yehya, J. Beresian, E. Salem, A. Kamar *et al.*, Two heterozygous mutations in nfatc1 in a patient with tricuspid atresia. *PLoS One* **7**, e49532 (2012).

33. B. Stallmeyer, H. Fenge, U. Nowak-Gottl, E. Schulze-Bahr, Mutational spectrum in the cardiac transcription factor gene nkx2.5 (csx) associated with congenital heart disease. *Clin Genet* **78**, 533-540 (2010).

34. J. T. Granados-Riveron, T. K. Ghosh, M. Pope, F. Bu'Lock, C. Thornborough *et al.*, Alpha-cardiac myosin heavy chain (myh6) mutations affecting myofibril formation are associated with congenital heart defects. *Hum Mol Genet* **19**, 4007-4016 (2010).

35. A. Fischer, B. Klamt, N. Schumacher, C. Glaeser, I. Hansmann *et al.*, Phenotypic variability in hey2 -/- mice and absence of hey2 mutations in patients with congenital heart defects or alagille syndrome. *Mamm Genome* **15**, 711-716 (2004).

36. A. Sarkozy, E. Conti, R. D'Agostino, M. C. Digilio, R. Formigari *et al.*, Zfpm2/fog2 and hey2 genes analysis in nonsyndromic tricuspid atresia. *Am J Med Genet A* **133A**, 68-70 (2005).

37. I. El-Rassy, J. Bou-Abdallah, S. Al-Ghadban, F. Bitar, G. Nemer, Absence of notch2 and hey2 mutations in a familial alagille syndrome case with a novel frameshift mutation in jag1. *Am J Med Genet A* **146A**, 937-939 (2008).

38. W. Pinxten, H. C. Howard, Ethical issues raised by whole genome sequencing. *Best Pract Res Clin Gastroenterol* **28**, 269-279 (2014).

39. B. Rabbani, M. Tekin, N. Mahdieh, The promise of whole-exome sequencing in medical genetics. *J Hum Genet* **59**, 5-15 (2014).

40. K. A. Johansen Taber, B. D. Dickinson, M. Wilson, The promise and challenges of next-generation genome sequencing for clinical care. *JAMA Intern Med* **174**, 275-280 (2014).

41. K. L. Smith, C. Isaacs, Brca mutation testing in determining breast cancer therapy. *Cancer J* **17**, 492-499 (2011).

42. W. Dondorp, B. Sikkema-Raddatz, C. de Die-Smulders, G. de Wert, Arrays in postnatal and prenatal diagnosis: An exploration of the ethics of consent. *Hum Mutat* **33**, 916-922 (2012).

43. B. Committee On, A. N. D. Committee On Genetics, A. N. D. American College Of Medical Genetics, S. Genomics, Ethical *et al.*, Ethical and policy issues in genetic testing and screening of children. *Pediatrics* **131**, 620-622 (2013).

44. A. L. McGuire, S. Joffe, B. A. Koenig, B. B. Biesecker, L. B. McCullough *et al.*, Point-counterpoint. Ethics and genomic incidental findings. *Science* **340**, 1047-1048 (2013).

45. A. L. McGuire, T. Caulfield, M. K. Cho, Research ethics and the challenge of whole-genome sequencing. *Nat Rev Genet* **9**, 152-156 (2008).

46. G. E. Wright, P. G. Koornhof, A. A. Adeyemo, N. Tiffin, Ethical and legal implications of whole genome and whole exome sequencing in african populations. *BMC Med Ethics* **14**, 21 (2013).

47. B. M. Knoppers, Y. Joly, J. Simard, F. Durocher, The emergence of an ethical duty to disclose genetic research results: International perspectives. *Eur J Hum Genet* **14**, 1170-1178 (2006).

48. M. E. Grove, M. N. Wolpert, M. K. Cho, S. S. Lee, K. E. Ormond, Views of genetics health professionals on the return of genomic results. *J Genet Couns* **23**, 531-538 (2014).

49. P. Borry, M. Shabani, H. C. Howard, Is there a right time to know? The right not to know and genetic testing in children. *J Law Med Ethics* **42**, 19-27 (2014).

50. R. C. Green, J. S. Berg, W. W. Grody, S. S. Kalia, B. R. Korf *et al.*, Acmg recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* **15**, 565-574 (2013).

51. J. S. Berg, M. J. Khoury, J. P. Evans, Deploying whole genome sequencing in clinical practice and public health: Meeting the challenge one bin at a time. *Genetics in Medicine* **13**, 499-504 (2011).

52. R. R. Fabsitz, A. McGuire, R. R. Sharp, M. Puggal, L. M. Beskow *et al.*, Ethical and practical guidelines for reporting genetic research results to study participants: Updated guidelines from a national heart, lung, and blood institute working group. *Circ Cardiovasc Genet* **3**, 574-580 (2010).

53. E. B. Bookman, A. A. Langehorne, J. H. Eckfeldt, K. C. Glass, G. P. Jarvik *et al.*, Reporting genetic results in research studies: Summary and recommendations of an nhlbi working group. *Am J Med Genet A* **140**, 1033-1040 (2006).

54. C. V. Fernandez, E. Kodish, C. Weijer, Informing study participants of research results: An ethical imperative. *IRB* **25**, 12-19 (2003).

55. A. L. McGuire, J. A. Hamilton, R. Lunstroth, L. B. McCullough, A. Goldman, DNA data sharing: Research participants' perspectives. *Genet Med* **10**, 46-53 (2008).

56. H. K. Tabor, B. E. Berkman, S. C. Hull, M. J. Bamshad, Genomics really gets personal: How exome and whole genome sequencing challenge the ethical framework of human genetics research. *Am J Med Genet A* **155A**, 2916-2924 (2011).

57. M. Gymrek, A. L. McGuire, D. Golan, E. Halperin, Y. Erlich, Identifying personal genomes by surname inference. *Science* **339**, 321-324 (2013).

58. T. Caulfield, A. L. McGuire, M. Cho, J. A. Buchanan, M. M. Burgess *et al.*, Research ethics recommendations for whole-genome research: Consensus statement. *PLoS Biol* **6**, e73 (2008).

59. P. C. f. t. S. o. B. Issues, Privacy and progress in whole genome sequencing. *Report to the President*, (2012).

60. B. Alsoufi, B. Schlosser, M. Mori, C. McCracken, T. Slesnick *et al.*, Influence of morphology and initial surgical strategy on survival of infants with tricuspid atresia. *Ann Thorac Surg* **100**, 1403-1409; discussion 1409-1410 (2015).

61. R. Sittiwangkul, A. Azakie, G. S. Van Arsdell, W. G. Williams, B. W. McCrindle, Outcomes of tricuspid atresia in the fontan era. *The Annals of Thoracic Surgery* **77**, 889-894 (2004).

62. Z. Lin, H. Ge, J. Xue, G. Wu, J. Du *et al.*, Comparison of extracardiac conduit and lateral tunnel for functional single-ventricle patients: A meta-analysis. *Congenit Heart Dis* **12**, 711-720 (2017).

63. A. Child, S. R. Kondapally Seshasai, R. Bastiaenen, J. A. Aragon-Martin, S. Gati *et al.*, Cardiovascular management of adults with marfan syndrome. *European Cardiology Review* **11**, 102 (2016).

64. S. Chakraborty, M. D. Combs, K. E. Yutzey, Transcriptional regulation of heart valve progenitor cells. *Pediatr Cardiol* **31**, 414-421 (2010).

65. M. Azhar, J. Schultz Jel, I. Grupp, G. W. Dorn, 2nd, P. Meneton *et al.*, Transforming growth factor beta in cardiovascular development and function. *Cytokine Growth Factor Rev* **14**, 391-407 (2003).

66. D. G. Molin, U. Bartram, K. Van der Heiden, L. Van Iperen, C. P. Speer *et al.*, Expression patterns of tgfbeta1-3 associate with myocardialisation of the outflow tract and the development of the epicardium and the fibrous heart skeleton. *Dev Dyn* **227**, 431-444 (2003).

67. P. Vrljicak, A. C. Chang, O. Morozova, E. D. Wederell, K. Niessen *et al.*, Genomic analysis distinguishes phases of early development of the mouse atrio-ventricular canal. *Physiol Genomics* **40**, 150-157 (2010).

68. M. Azhar, K. Brown, C. Gard, H. Chen, S. Rajan *et al.*, Transforming growth factor beta2 is required for valve remodeling during heart development. *Dev Dyn* **240**, 2127-2141 (2011).

69. A. H. Li, N. A. Hanchard, D. Furthner, S. Fernbach, M. Azamian *et al.*, Whole exome sequencing in 342 congenital cardiac left sided lesion cases reveals extensive genetic heterogeneity and complex inheritance patterns. *Genome Med* **9**, 95 (2017).

70. C. Ward, Clinical significance of the bicuspid aortic valve. *Heart* **83**, 81-85 (2000).

71. J. M. Schwarz, D. N. Cooper, M. Schuelke, D. Seelow, Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nat Methods* **11**, 361-362 (2014).

72. G. Takaesu, S. Kishida, A. Hiyama, K. Yamaguchi, H. Shibuya *et al.*, Tab2, a novel adaptor protein, mediates activation of tak1 mapkkk by linking tak1 to traf6 in the il-1 signal transduction pathway. *Mol Cell* **5**, 649-658 (2000).

73. A. Besse, B. Lamothe, A. D. Campos, W. K. Webster, U. Maddineni *et al.*, Tak1-dependent signaling requires functional interaction with tab2/tab3. *J Biol Chem* **282**, 3918-3928 (2007).

74. B. Thienpont, L. Zhang, A. V. Postma, J. Breckpot, L. C. Tranchevent *et al.*, Haploinsufficiency of tab2 causes congenital heart defects in humans. *Am J Hum Genet* **86**, 839-849 (2010).

75. S. Morioka, M. Inagaki, Y. Komatsu, Y. Mishina, K. Matsumoto *et al.*, Tak1 kinase signaling regulates embryonic angiogenesis by modulating endothelial cell survival and migration. *Blood* **120**, 3846-3857 (2012).

76. H. Sanjo, K. Takeda, T. Tsujimura, J. Ninomiya-Tsuji, K. Matsumoto *et al.*, Tab2 is essential for prevention of apoptosis in fetal liver but not for interleukin-1 signaling. *Molecular and Cellular Biology* **23**, 1231-1238 (2003).

77. Y. Yang, B. Seed, Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes. *Nat Biotechnol* **21**, 447-451 (2003).

78. K. Weiss, C. Applegate, T. Wang, D. A. Batista, Familial tab2 microdeletion and congenital heart defects including unusual valve dysplasia and tetralogy of fallot. *Am J Med Genet A* **167A**, 2702-2706 (2015).

79. M. A. Prabhu, A. Vupputuri, S. Shekar, M. S. Harikrishnan, P. G. Pai *et al.*, An unusual type of accessory pathway in tricuspid atresia. *J Cardiol Cases* **14**, 181-184 (2016).

80. T. Koide, T. Hayata, K. W. Cho, Negative regulation of hedgehog signaling by the cholesterogenic enzyme 7-dehydrocholesterol reductase. *Development* **133**, 2395-2405 (2006).

81. W. L. Xiao, D. Z. Zhang, H. Xu, C. Z. Zhuang, Dhcr7 regulates palatal shelf fusion through regulation of shh and bmp2 expression. *Biomed Res Int* **2016**, 7532714 (2016).

82. Y. Peng, R. Myers, W. Zhang, E. Alexov, Computational investigation of the missense mutations in dhcr7 gene associated with smith-lemli-opitz syndrome. *Int J Mol Sci* **19**, (2018).

83. C. A. Wassif, P. Zhu, L. Kratz, P. A. Krakowiak, K. P. Battaile *et al.*, Biochemical, phenotypic and neurophysiological characterization of a genetic mouse model of rsh/smith--lemli--opitz syndrome. *Hum Mol Genet* **10**, 555-564 (2001).

84. H. Yu, S. B. Patel, Recent insights into the smith-lemli-opitz syndrome. *Clin Genet* **68**, 383-391 (2005).

85. F. Moccia, S. Baruffi, S. Spaggiari, D. Coltrini, R. Berra-Romani *et al.*, P2y1 and p2y2 receptor-operated ca2+ signals in primary cultures of cardiac microvascular endothelial cells. *Microvasc Res* **61**, 240-252 (2001).

86. E. Hochhauser, R. Cohen, M. Waldman, A. Maksin, A. Isak *et al.*, P2y2 receptor agonist with enhanced stability protects the heart from ischemic damage in vitro and in vivo. *Purinergic Signal* **9**, 633-642 (2013).

87. D. De Oliveira Moreira, H. Santo Neto, M. J. Marques, P2y2purinergic receptors are highly expressed in cardiac and diaphragm muscles ofmdxmice, and their expression is decreased by suramin. *Muscle & Nerve* **55**, 116-121 (2017).

88. F. G. Khalafalla, S. Greene, H. Khan, K. Ilves, M. M. Monsanto *et al.*, P2y2 nucleotide receptor prompts human cardiac progenitor cell activation by modulating hippo signaling. *Circ Res* **121**, 1224-1236 (2017).

89. Y. Tian, Y. Liu, T. Wang, N. Zhou, J. Kong *et al.*, A microrna-hippo pathway that promotes cardiomyocyte proliferation and cardiac regeneration in mice. *Science Translational Medicine* **7**, (2015).

90. A. Nishimura, C. Sunggip, S. Oda, T. Numaga-Tomita, M. Tsuda *et al.*, Purinergic p2y receptors: Molecular diversity and implications for treatment of cardiovascular diseases. *Pharmacol Ther* **180**, 113-128 (2017).

91. A. Bondue, S. Tannler, G. Chiapparo, S. Chabab, M. Ramialison *et al.*, Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. *J Cell Biol* **192**, 751-765 (2011).

92. L. Zhou, J. Liu, M. Xiang, P. Olson, A. Guzzetta *et al.*, Gata4 potentiates second heart field proliferation and hedgehog signaling for cardiac septation. *Proc Natl Acad Sci U S A* **114**, E1422-E1431 (2017).

93. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang *et al.*, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).

94. G. Euler, Good and bad sides of TGFbeta-signaling in myocardial infarction. *Front Physiol* **6**, 66 (2015).

95. S. I. Kim, M. E. Choi, TGF-beta-activated kinase-1: New insights into the mechanism of tgf-beta signaling and kidney disease. *Kidney Res Clin Pract* **31**, 94-105 (2012).

96. Q. Liu, Y. Chen, M. Auger-Messier, J. D. Molkentin, Interaction between NFkappaB and nfat coordinates cardiac hypertrophy and pathological remodeling. *Circ Res* **110**, 1077-1086 (2012).

97. J. H. Kim, N. Kim, Regulation of nfatc1 in osteoclast differentiation. *J Bone Metab* **21**, 233-241 (2014).

98. X. Fang, J. Robinson, J. Wang-Hu, L. Jiang, D. A. Freeman *et al.*, Camp induces hypertrophy and alters DNA methylation in hl-1 cardiomyocytes. *Am J Physiol Cell Physiol* **309**, C425-436 (2015).

99. Y. Chen, X. Cao, Nfat directly regulates nkx2-5 transcription during cardiac cell differentiation. *Biol Cell* **101**, 335-349 (2009).

100. B. Zhou, B. Wu, K. L. Tompkins, K. L. Boyer, J. C. Grindley *et al.*, Characterization of nfatc1 regulation identifies an enhancer required for gene expression that is specific to pro-valve endocardial cells in the developing heart. *Development* **132**, 1137-1146 (2005).

101. H. Hara, N. Takeda, T. Fujiwara, H. Yagi, S. Maemura *et al.*, Activation of tgf-beta signaling in an aortic aneurysm in a patient with loeys-dietz syndrome caused by a novel loss-of-function variant of TGFBR1. *Hum Genome Var* **6**, 6 (2019).

102. M. M. Winslow, M. Pan, M. Starbuck, E. M. Gallo, L. Deng *et al.*, Calcineurin/nfat signaling in osteoblasts regulates bone mass. *Developmental Cell* **10**, 771-782 (2006).

103. I. Lobo, Same genetic mutation, different genetic disease phenotype. *Nature* **1**, 64 (2008).

104. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott *et al.*, Genome engineering using the crispr-cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).

105. K. J. Carroll, C. A. Makarewich, J. McAnally, D. M. Anderson, L. Zentilin *et al.*, A mouse model for adult cardiac-specific gene deletion with crispr/cas9. *Proc Natl Acad Sci U S A* **113**, 338-343 (2016).