CURRENT CHALLENGES IN CARDIOVASCULAR MOLECULAR DIAGNOSTICS

EDITED BY : Matteo Vatta, Valeria Novelli, Luisa Mestroni,
Jeffrey A. Towbin, Carlo Napolitano and Guia Guffanti

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

CURRENT CHALLENGES IN CARDIOVASCULAR MOLECULAR DIAGNOSTICS

Topic Editors:

Matteo Vatta, Indiana University School of Medicine, United States

Valeria Novelli, Institute of Genomic Medicine, Fondazione Policlinico Universitario "A.Gemelli" and Centro Studi "Benito Stirpe" per la Prevenzione della Morte Improvvisa nel Giovane Atleta, Italy

Luisa Mestroni, University of Colorado Denver, United States

Jeffrey A. Towbin, St. Jude Children's Research Hospital, University of Tennessee Health Science Center, United States

Carlo Napolitano, Molecular Cardiology, Istituti Clinici Scientifici Maugeri, Italy

Guia Guffanti, McLean Hospital, Harvard Medical School, United States

Cover image: enzozo/Shutterstock.com, Khakimullin Aleksandr/Shutterstock.com and Sebastian Kaulitzki/Shutterstock.com

The field of cardiovascular genetics has tremendously benefited from the recent application of massive parallel sequencing technology also referred to as next generation sequencing (NGS). However, along with the discovery of additional genes associated with human cardiac diseases, the analysis of large dataset of genetic information uncovered a much more complex and variegated landscape, which often departs from the comfort zone of the monogenic Mendelian diseases image that clinical molecular geneticists have been well acquainted with for many decades. It is now clear that, in addition to highly penetrant genetic variants, which in isolation are able to recapitulate the full clinical presentation when expressed in animal models, we are now aware that a small but significant fraction of subjects presenting with cardiac muscle diseases such as cardiomyopathies or primary arrhythmias such as long QT syndrome (LQTS), may harbor at least two deleterious variants in the same gene (compound heterozygous) or in different gene

(double heterozygous). Although the clinical presentation in subjects with more than one deleterious variant appears to be more severe and with an earlier disease onset, it somehow changes the viewpoint of clinical molecular geneticists whose aim is to identify all possible genetic contributors to a human condition. In this light, the employment in clinical diagnostics of the NGS technology, allowing the simultaneous interrogation of a DNA target spanning from large panel of genes up to the entire genome, will definitely aid at uncovering all such contributors, which will have to be tested functionally to confirm their role in human cardiac conditions. The uncovering of all clinically relevant deleterious changes associated with a cardiovascular disease would probably increase our understanding of the clinical variability commonly occurring among affected family relatives, and potentially provide with unexpected therapeutic targets for the treatment of symptoms related to the presence of "accessory" deleterious genetic variants other than the key molecular culprit. The objective of this Research Topic is to explore the current challenges presenting to the cardiovascular genetics providers, such as clinical geneticists, genetic counselors, clinical molecular geneticists and molecular pathologists involved in the diagnosis, counseling, testing and interpretation of genetic tests results for the comprehensive management of patients affected by cardiovascular genetic disorders.

Table of Contents

Editorial: Current Challenges in Cardiovascular Molecular Diagnostics

Valeria Novelli[1,2]* and Matteo Vatta[3,4]*

[1] *Institute of Genomic Medicine, Fondazione Policlinico Universitario Agostino Gemelli, Rome, Italy,* [2] *Centro Studi "Benito Stirpe" per la prevenzione della morte improvvisa nel giovane atleta, Rome, Italy,* [3] *Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, United States,* [4] *Invitae, San Francisco, CA, United States*

Editorial on the Research Topic

Current Challenges in Cardiovascular Molecular Diagnostics

In the last 10 years, the development of massive parallel sequencing technology, commonly referred to as next-generation-sequencing (NGS) technologies, carried the genetic field in a new era, opening unexplored avenues in the research of inherited cardiovascular disease (1).

As any new technology, when based on solid experimental observations and on real innovation, primarily generates a lot of enthusiasm among researchers as well as among patients, especially when it reaches a broader audience through the media coverage. However, as for every technology, along with the promises, the technical limitations inevitably appear, prompting more questions and further development. The technical refinement and the need for more robust and reliable assay validation standards delayed the introduction of NGS in clinical practice and the development of consensus standard operating procedure for the incorporation of NGS in the laboratory guidelines as the new standard test for the molecular diagnosis of inherited cardiovascular disease.

Here, we provide a brief review of the potential new applications and current challenges associated with the widespread use of NGS and the strategies that still need to be implemented to consider NGS a critical and sustainable tool for the diagnosis of cardiovascular diseases, and the detection of all at-risk family members.

In this research topic, we tried to raise awareness about the complexity of the issues that cardiologists and genetic practitioners have face. In particular, here we discuss the challenges in cardiovascular molecular diagnostics by targeting four aspects spanning different clinical specialties and timeframes, from the diagnosis to the treatment of patients. The management of patients affected by a genetic cardiovascular disease has changed substantially over the last decade. As a result, it is of critical importance to developed common managing strategies among multi-disciplinary stakeholders, and being able to synthesize the large quantity of information generated by the healthcare procedures, including the genetic testing laboratory, in order to provide the best care options for the patients and their families (2).

PATIENTS SELECTION AND INDICATION TO GENETIC TESTING

In the last years, thanks to the development of the concept of precision genomic medicine, an unprecedented proliferation of genetics- and genomics-guided testing has been proposed.

Although a variety of testing guidelines have been indicated by the European and American Societies of Cardiology (AHA, ACC, ESC, and HRS), the "real world" scenario is more heterogeneous;

the optimal classification of individuals for molecular testing in inherited cardiovascular disease remains difficult (3). Genetic testing should be undertaken, indeed, only if considerable suspicion for an underlying genetic cardiovascular disease is present and always proposed after a comprehensive clinical evaluation, including, but not limited to, a detailed family history, cardiovascular work-up, and assessment for multisystem syndromes.

As we are all learning from NGS, the clinical utility of genetic testing is highly dependent on the pretest probability of each disease; common pitfalls associated with the inappropriate use of genetic testing, namely, poor phenotyping and inappropriate genetic test selection are very often able to hamper the diagnostic yield and the risk of encountering false-positive results.

NGS APPROACH TO ADOPT

Next-generation sequencing can be applied to panels of genes, to the exome, namely the targeting of all coding exons, or the whole genome in clinical settings exome sequencing (ES) and genome sequencing (GS) are mainly adopted for gene discovery and used as clinical testing only when a clinical diagnosis cannot unequivocally be established. Differently, gene panels represent a good compromise between testing just a few genes and obtaining information from the exome. This approach is usually employed when a clinical diagnosis has been reached and does not lend itself to the identification of novel genes (4).

In order to fulfill the diagnostic necessities and homogenize the diagnostic procedures for different cardiac conditions, the design of custom target sequencing panels requires an in-depth knowledge of the specific disease and accuracy in the selection of the genes to analyze, according to their level of evidence.

The numbers of genes included in each panel can differ between laboratories. Some laboratories apply the strategy to include only "major genes" for which substantial literature is reported. Other panels include a larger gene set that includes the aforementioned major genes and additional "minor genes," for which evidence is still accruing.

According to these considerations, it is important to highlight that when planning the development of a targeted gene panel, the main challenge is to define its main application and the targeted phenotypes for which the test is conceived.

VARIANT INTERPRETATION

The assessment of the pathogenicity of genetic variants is of critical importance. The high variability of the human genome calls to exercise extreme caution to avoid the misinterpretation of the identified genetic variants. Especially important for clinical genetic laboratories have been the development of large databases of control individuals, aiming at mimicking the genetic behavior of variants in a general population. When a variant is identified in a patient, we can now analyze its frequency in the largest open source database, namely the Genome Aggregation Database, which comprises of the data from sequencing 123,136 exomes and 15,496 genomes (http://gnomad.broadinstitute.org)

(5). In addition to the variant frequency, evidence such as family studies, functional analysis, and biocomputational assessments need to be considered.

Deciding how to categorize and weigh each type of evidence is really challenging, and it is, therefore, difficult to validate approaches to variant assessment, particularly for variants that have limited evidence, usually identified by GS or ES.

This issue actually needs the collective experience of experts in the community to begin to build commonly validated approach to variant classification. Starting from the collaboration of a group of experts ACMG and Association for Molecular Pathology in 2015 developed a framework for evidence evaluation. This framework is now in revision in order to be personalized according to the gene variation (6).

SUSTAINABILITY OF GENETIC TESTING IN THE "REAL WORLD"

When evaluating a genetic testing strategy, it is important to take into account the costs of that strategy and to determine if the increases in effectiveness are worth the additional costs that broader testing strategies incur. Genetic testing, although generally accepted by the medical community as an increasingly fundamental tool for patients' management, remains a relatively expensive test for which, identifying who should bear the economic burden, remains often challenging. In particular, the heterogeneity of political and socio-economical systems make genetic testing a very different experience for patients and their family in various countries across the globe. Too often, the economic burden of genetic testing is loaded onto the pockets of the patients and their families, while apart from the few countries with a national program to fund genetic testing as every other clinical test, other systems rely on third parties for the management of medical expenses (7). However, contrary to what occurred with genetic testing for cancer, the equal acknowledgment of the service provided by genetic laboratories to thousands of cardiovascular patients has yet to be achieved.

REMARK CONCLUSIVE

The last few decades have brought much technological advances, which have forever changed the landscape of clinical genetic testing. Despite the natural enthusiasm for those changes, much remains to be done for the optimal application of the massive parallel sequencing technology we commonly refer to as NGS. Although NGS represents a powerful testing tool, it can reach its full potential only if integrated with improved clinical diagnostics, refinement of the sequencing strategy, standardized variant interpretation, and economic sustainability, which requires genetic testing to be embraced as standard clinical practice by the healthcare community in its entirety.

AUTHOR CONTRIBUTIONS

VN and MV contributed extensively to the work presented in this Editorial.

REFERENCES

1. MacRae CA, Vasan RS. The future of genetics and genomics: closing the phenotype gap in precision medicine. *Circulation* (2016) 133(25):2634–9. doi:10.1161/CIRCULATIONAHA.116.022547

2. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* (2013) 14(4):295–300. doi:10.1038/nrg3463

3. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Europace* (2011) 13(8):1077–109. doi:10.1093/europace/eur245

4. Katsanis SH, Katsanis N. Molecular genetic testing and the future of clinical genomics. *Nat Rev Genet* (2013) 14(6):415–26. doi:10.1038/nrg3493

5. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* (2016) 536(7616):285–91. doi:10.1038/nature19057

6. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the association for molecular pathology. *Genet Med* (2015) 17(5):405–24. doi:10.1038/gim.2015.30

7. Shoenbill K, Fost N, Tachinardi U, Mendonca EA. Genetic data and electronic health records: a discussion of ethical, logistical and technological considerations. *J Am Med Inform Assoc* (2014) 21(1):171–80. doi:10.1136/amiajnl-2013-001694

The Brugada Syndrome: A Rare Arrhythmia Disorder with Complex Inheritance

Jean-Baptiste Gourraud[1,2,3,4], Julien Barc[2,3,4], Aurélie Thollet[1], Solena Le Scouarnec[2,3,4], Hervé Le Marec[1,2,3,4], Jean-Jacques Schott[1,2,3,4], Richard Redon[1,2,3,4] and Vincent Probst[1,2,3,4]*

[1] Service de Cardiologie, Centre Hospitalier Universitaire (CHU) de Nantes, l'institut du thorax, Nantes, France, [2] Institut National de la Santé et de la Recherche Médicale (INSERM) Unité Mixte de Recherche (UMR) 1087, l'institut du thorax, Nantes, France, [3] Centre National de la Recherche Scientifique (CNRS) UMR 6291, l'institut du thorax, Nantes, France, [4] l'institut du thorax, Université de Nantes, Nantes, France

For the last 10 years, applying new sequencing technologies to thousands of whole exomes has revealed the high variability of the human genome. Extreme caution should thus be taken to avoid misinterpretation when associating rare genetic variants to disease susceptibility. The Brugada syndrome (BrS) is a rare inherited arrhythmia disease associated with high risk of sudden cardiac death in the young adult. Familial inheritance has long been described as Mendelian, with autosomal dominant mode of transmission and incomplete penetrance. However, all except 1 of the 23 genes previously associated with the disease have been identified through a candidate gene approach. To date, only rare coding variants in the *SCN5A* gene have been significantly associated with the syndrome. However, the genotype/phenotype studies conducted in families with *SCN5A* mutations illustrate the complex mode of inheritance of BrS. This genetic complexity has recently been confirmed by the identification of common polymorphic alleles strongly associated with disease risk. The implication of both rare and common variants in BrS susceptibility implies that one should first define a proper genetic model for BrS predisposition prior to applying molecular diagnosis. Although long remains the way to personalized medicine against BrS, the high phenotype variability encountered in familial forms of the disease may partly find an explanation into this specific genetic architecture.

Keywords: Brugada syndrome, genetics, sudden death, cardiac arrhythmias, SCN5A

INTRODUCTION

The Brugada syndrome (BrS) is a rare inherited arrhythmia disease, first described in 1992, increasing the risk of ventricular fibrillation in apparently healthy young adults (1). It is suspected to be involved in 4–12% of cases of sudden cardiac death (SCD) in the general population and in at least 20% of SCD in patients with a structurally normal heart (1–3).

Clinical diagnosis is based on a specific electrocardiographic (ECG) pattern defined in three consecutive consensus conferences (4–6). This ECG pattern, previously known as "type 1" ECG pattern, is defined as a ST segment elevation with a coved-type morphology ≥0.2 mV in one lead among the right precordial leads V1 and V2, positioned in the second, third, or fourth intercostal space occurring either spontaneously or after provocative drug test with intravenous administration of Class I antiarrhythmic drugs (6) (**Figure 1**). The ECG pattern may be transient in affected

patients (7). To address this issue, unmasking drugs, such as ajmaline, flecainide, and procainamide, can be used to reveal this pattern (8), ajmaline showing higher sensitivity than flecainide and procainamide (4, 9, 10).

The high variability of the ECG pattern impairs proper assessment of its prevalence in the general population. Epidemiological studies have produced heterogeneous results regarding BrS incidence across the World. While estimated at 5 for 10,000 in western Europe and the USA, the prevalence of BrS seems higher in Southeast Asia, reaching 20 for 10,000 (11–13).

Aborted SCD is often the first symptom in BrS, with a mean age of 45 years at diagnosis and a four-time higher incidence in men than in women (14, 15). A third of the affected patients are identified after syncope, frequently preceded by vagal symptoms (14). The syncope could either be due to non-sustained VF or to a vaso-vagal episode without direct clinical relevance, rendering it hard for the practitioner to distinguish arrhythmic from non-arrhythmic etiology (16, 17). The majority of patients are asymptomatic at time of diagnosis. More than one-third of cases are identified during familial screening (14).

Implantation of a defibrillator is still the only efficient therapy in high-risk patients, with a 48% rate of appropriate device therapy at 10 years in patients with previous aborted sudden death. This rate falls to 12% among implanted asymptomatic patients, many affected patients remaining asymptomatic during all their life. Furthermore, device-related complications are frequent with a 30% risk at 10-year follow-up mainly due to lead dysfunction, inappropriate therapy, and infection (18, 19). These serious side effects in comparison to the very low arrhythmic risk for asymptomatic patients require accurate risk stratification and/ or efficient drug therapy.

Only few clinical parameters allow risk stratification in BrS. The effectiveness of ventricular stimulation is still a matter of debate, and symptoms and spontaneous ECG pattern are still the two major parameters enabling risk stratification for SCD (6, 14, 20–23).

There is still need for medical therapies that could reduce arrhythmia occurrence and prevent SCD. Because successful trials were reported in limited series of patients, quinidine has been expected to be "the drug" for BrS. However, several recent studies failed to demonstrate its beneficial effects (6, 24–27).
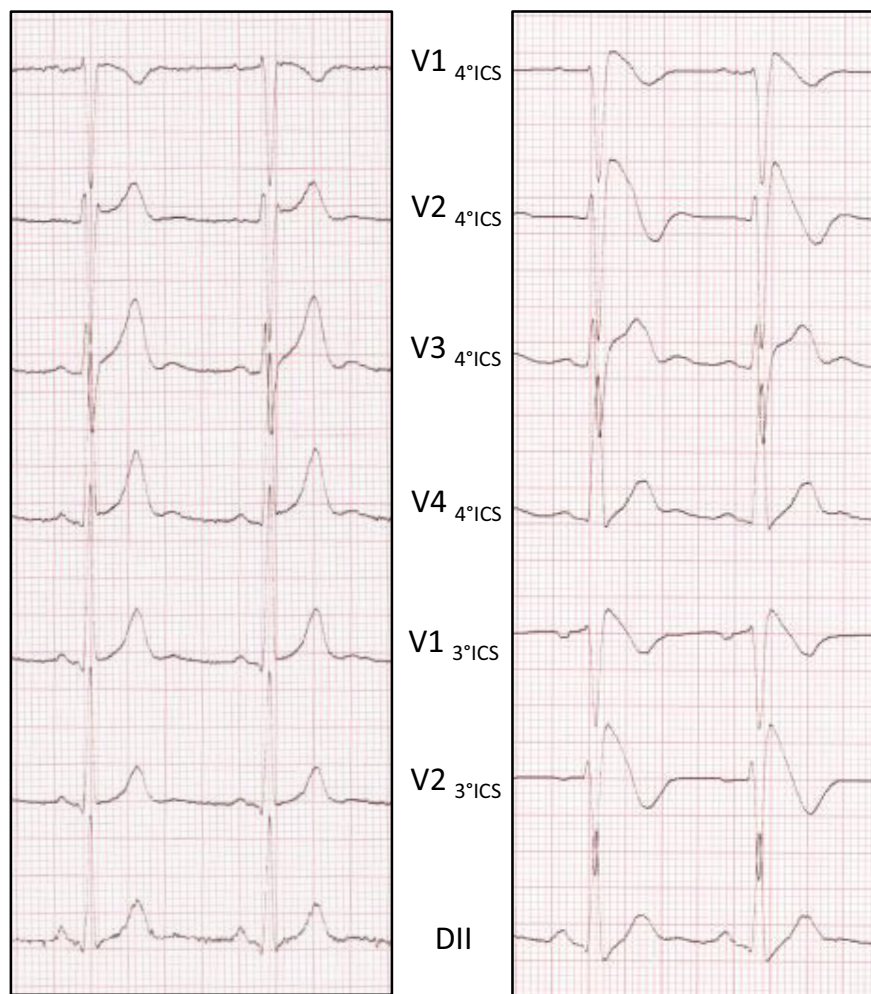
FIGURE 1 | Ajmaline testing reveals the Brugada ECG pattern. ECG pattern is recorded at 1 mm/10 mV and 25 mm/s. Baseline ECG without aspect of BrS (left side). Type 1 Brugada pattern on the right precordial leads at the end of the test (right side).

There is accumulating evidence that implantable defibrillator is an effective and accurate therapy for symptomatic patients (18). Many clinical parameters have also been proposed for asymptomatic patients, but risk prediction in the latter group of patients remains particularly challenging because of the lack of reproducible and reliable data (28).

TWO PATHOPHYSIOLOGICAL MODELS FOR BrS

Those unresolved questions concerning diagnosis and risk stratification for arrhythmia and therapy underlie the need for a better understanding of pathophysiological mechanisms in BrS. Two main pathophysiological hypotheses have been proposed to explain the ECG pattern.

Soon after the description of BrS, the first pathophysiological model was proposed, based on the existence of a transmural voltage gradient due to a repolarization heterogeneity across the ventricular wall (29, 30). According to this hypothesis, ST segment elevation could be due to either a loss of function of the sodium channel NaV1.5 responsible for the depolarization phase (phase 0 of the AP) favoring the expression of repolarization heterogeneity, an aggravation of this heterogeneity by a gain of function in one of the cardiac potassium channels responsible of

the repolarization phases (phases 1 and 3 of the AP), or a loss of function of the CaV1.2 calcium channel that participate to the phase 2 of the AP (29).

This hypothesis has been matter to debate since the second hypothesis, based on a conduction delay in the right ventricular outflow tract, emerged from clinical observations (31–35). This conduction delay could be responsible for voltage gradients between RV and RVOT during and explain the BrS ECG pattern.

Twenty years of genetic research based on both technological and methodological progresses have started to depict the complexity of BrS pathophysiology (36, 37). This review aims to provide an integrated synopsis of those two decades of research and to suggest future directions for further genetic investigations against BrS.

FROM A FAMILIAL DISEASE TO THE IDENTIFICATION OF RARE VARIANTS

With the initial report of two affected siblings, familial inheritance was suggested from the first description of the Brs in 1992 (1). Few years later, Kobayashi et al. described a two-generation family presenting with both SCD and persistent ST elevation in relatives (38), confirming the heritability of the disease. The genetic component of BrS was further demonstrated in several reports (39–41) (**Figure 2**). Today, familial history of SCD is
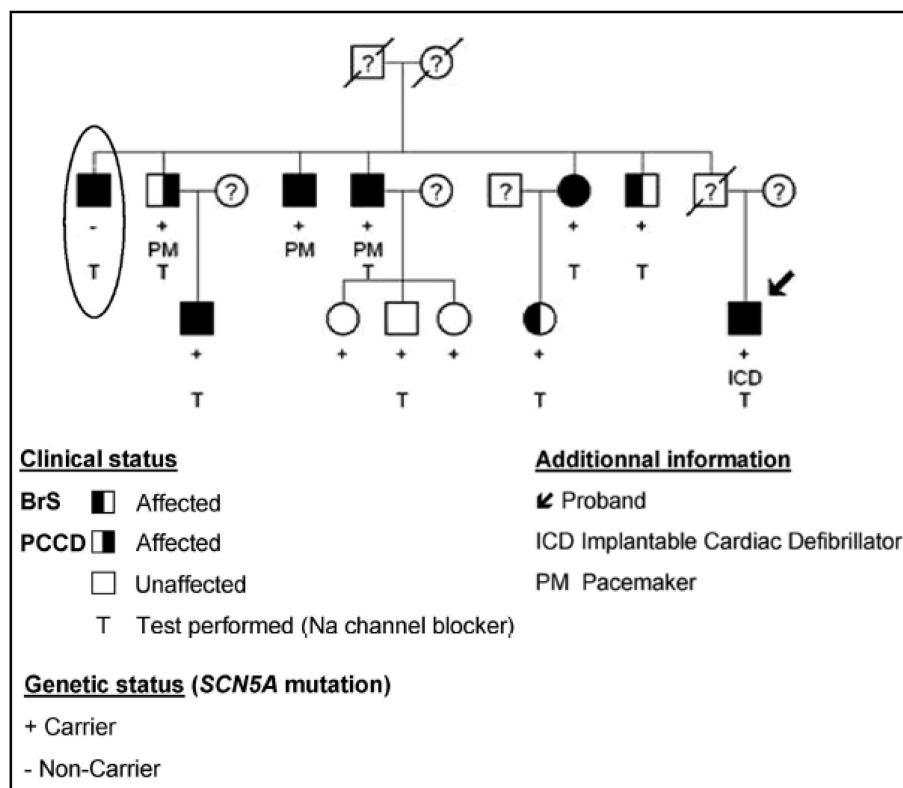
FIGURE 2 | The complex inheritance pattern of BrS. Modified from Ref. (41). Incomplete penetrance of the *SCN5A* mutation is illustrated by the presence of unaffected carriers of the mutation (+). The patient highlighted by an ellipse presents with a BrS ECG aspect, despite the absence of the familial mutation. Affected family members carrying the *SCN5A* mutation present with progressive cardiac conduction disease (PCCD) (right half-filled symbol), BrS (left half-filled symbol), or both diseases (full-filled symbol). PCCD consists of right bundle branch block with PR interval lengthening and led to complete AVB in three patients, in whom a pacemaker (PM) was implanted.

reported for about 26% of affected patients. Additionally, 36% of affected patients are identified during familial screening after SCD or identification of BrS in the proband (14).

Brugada syndrome has been consistently reported as a monogenic disease with autosomal dominant mode of inheritance, caused by rare genetic variants with large effect size (1, 38). Loss-of-function mutations in the SCN5A-encoded α-subunit of the cardiac sodium channel (Nav1.5) were first identified in 1998 (42). Mutations in *SCN5A* are detected in 20–25% of cases, *SCN5A* appearing as the major susceptibility gene for BrS (43). More than 300 rare variants in *SCN5A* have been reported, while the contribution of other genes remains extremely low (43, 44). In a pediatric population affected by BrS, the prevalence of *SCN5A* mutations seems to be even higher (45).

In this context, genetics was initially expected to help the clinical management of patients with BrS. Although some *SCN5A* mutations – particularly those leading to premature truncation of Nav1.5 – have been reported as associated with higher arrhythmic risk, no such result has been further confirmed in randomized studies (14, 46–48).

Despite evidence for strong familial inheritance, familial linkage analyses on BrS have been largely unsuccessful. Only one gene, *GPD1L,* has been identified as a BrS-susceptibility gene using this approach (49). The causing mutation in *GPD1L* has been shown to affect Na+ channel trafficking to the plasma membrane, by modifying its oxydation state (49, 50). Every other gene reported so far has been identified through a candidate approach based on direct sequencing of genes with a known (or suspected) role in cardiac electrical activity.

So far, 23 genes have been related to BrS (**Table 1**). Based on pathophysiological hypotheses, those genes can be divided according to whether they affect the sodium current I_{Na} (*SCN5A, SCN10A, GPD1L, SCN1B, SCN3B, RANGRF, SCN2B, PKP2, SLMAP,* and *FGF12*), the potassium current I_K (*KCNJ8, KCNH2, KCNE3, KCND3, KCNE5, KCND2, SEMA3A,* and *ABCC9*), or the calcium current I_{Ca} (*CACNA1C, CACNB2B,* and *CACNA2D1*).

LIMITS IN INTERPRETING RARE VARIANTS CARRIED BY PATIENTS WITH BrS

In the last decade, the emergence of massively parallel sequencing [or next-generation sequencing (NGS)] has considerably facilitated genetic screening and reduced its cost (72–76). Combined to the availability of the reference assembly of the human genome (77, 78), NGS-based approaches have revealed the high variability of the human genome, with at least 300–600 functional genetic variants detected in each exome (i.e., the whole coding portion of a single genome) (75) – and has retrospectively changed the interpretation of previous rare variants identified by candidate gene approach. The investigation of large number of exomes revealed the extraordinary prevalence of rare variants among each individual. As an illustration, the sequencing of 60,706 exomes identified about 7,500,000 variants from which 99% have a frequency of <1% (http://biorxiv.org/content/early/2015/10/30/030338).

Extreme caution should thus be taken when interpreting the rare genetic variants detected among patients with BrS, since the

TABLE 1 | The 23 reported susceptibility genes for BrS.

OMIM ranking	Gene	Protein	Prevalence in BrS cases	Functional effect of the mutation	Reference
BrS1	*SCN5A*	α subunit of the Nav1.5 sodium channel	20–25%	↘ I_{Na}	(42)
BrS2	*GPD1L*	Glycerol-3-phosphate dehydrogenase 1-like	Rare	↘ I_{Na}	(49)
BrS3	*CACNA1C*	α subunit α1C of the Cav1.2 calcium channel	1–2%	↘ I_{Ca-L}	(51)
BrS4	*CACNB2b*	β subunit Cavβ2b of calcium channel	1–2%	↘ I_{Ca-L}	(51)
BrS5	*SCN1b*	β subunit Navβ1 of sodium channel	Rare	↘ I_{Na}	(52)
BrS6	*KCNE3*	β subunit MiRP2 of potassium channel	Rare	↗ I_{to}	(53)
BrS7	*SCN3b*	β subunit Navβ3 of sodium channel	Rare	↘ I_{Na}	(54)
BrS8	*HCN4*	Hyperpolarization-activated cyclic nucleotide-gated channel 4	Rare	?	(55)
BrS9	*KCND3*	α subunit of the KV4.3 potassium channel	Rare	↗ I_{to}	(56)
BrS10	*KCNJ8*	α subunit of the KIR6.1 potassium channel	Rare	↗ I_{KATP}	(57)
BrS11	*CACNA2D1*	δ subunit Cavα2δ1 of calcium channel	Rare	↘ I_{Ca-L}	(58)
BrS12	*KCNE5*	β subunit of potassium channel	Rare	↗ I_{to}	(59)
BrS13	*RANGRF*	RAN guanine nucleotide release factor	Rare	↘ I_{Na}	(60)
BrS14	*KCND2*	α subunit of the KV4.2 potassium channel	Rare	↗ I_{to}	(61)
BrS15	*TRPM4*	Calcium-activated non-selective ion channel	Rare	?	(62)
BrS16	*SCN2B*	β subunit Navβ2 of sodium channel	Rare	↘ I_{Na}	(63)
BrS17	*PKP2*	Plakophilin 2	Rare	↘ I_{Na}	(64)
BrS18	*ABCC9*	ATP-sensitive potassium channels	Rare	↗ I_{KATP}	(65)
BrS19	*SLMAP*	Sarcolemma-associated protein	Rare	↘ I_{Na}	(66)
BrS20	*KCNH2*	α subunit of the HERG potassium channel	Rare	↗ I_{Kr}	(67)
BrS21	*SCN10A*	α subunit of the Nav1.8 sodium channel	<5%	↘ I_{Na}	(68, 69)
BrS22	*FGF12*	Fibroblast growth factor 12	Rare	↘ I_{Na}	(70)
BrS23	*SEMA3A*	Semaphorin family protein	Rare	↗ I_{to}	(71)

Functionnal effect on current are described with arrow, except for HCN4 mutation for which it remain unclear (?).

clinical implication of finding those variants remains doubtful in the absence of statistical association and/or of evidence supporting a functional effect in relation with cardiac electrical activity (79–82).

Furthermore, a recent study has illustrated the weakness of candidate approaches on small pedigrees, by highlighting the high frequency of some genetic variants previously associated with BrS among 6,500 individual exomes from the Exome Sequencing Project (83). One variant in particular, which was related to BrS based on functional evidence, showed a minor allele frequency of 4.4% among the 6,500 individuals. This result was confirmed in a healthy Danish control population, suggesting that a proportion of the genetic variants reported as causing BrS are actually not pathogenic. Interestingly, 93% of the *SCN5A* variants reported as causing BrS are not present among the control population, thus reinforcing the pivotal role of this gene.

By testing the burden of rare coding variants in 45 arrhythmia-susceptibility genes among 167 BrS cases versus 167 control individuals, we have also recently demonstrated the limitation of previous candidate approaches (44). Indeed, for every tested gene except *SCN5A*, rare variants were found in the same proportion in cases than in controls. **Figure 3** shows the distribution of rare variants among cases and controls for the protein products of four genes: *SCN5A*, *SCN10A*, *CACNA1C*, and *PKP2*. The distribution of rare variants across the functional domains of the *CACNA1C* product indicates that the C-terminal tail, which was previously considered as pathogenic in BrS, may in fact be highly polymorphic. On the opposite, most rare variants detected along the protein encoded by *PKP2* among BrS patients reside in a small interval coding for four amino acids. The *PKP2* gene has been previously associated with BrS by decreasing functional Na channel expression through modification of microtubule anchoring (64). The small *PKP2* interval emphasized in this study may be a preferential site of such interaction.

Rare genetic variants appear more evenly distributed across *SCN10A* and less predictive of any potential pathophysiological mechanism. In fact, the functional effects of these rare variants affecting *SCN10A* are largely debated. *SCN10A* gene, which encodes the sodium channel Na$_v$1.8, was initially described in neurons physiology (84, 85). Further investigations illustrated a potential role in cardiac electrophysiology, particulary as a modulator of cardiac conduction (86, 87). Recently, Hu et al. described rare variants in the *SCN10A* gene, in 16.7% of 150 patients affected with BrS (68). Furthermore, they demonstrated that the *SCN10A* variants R1268Q and R14L reduced cardiac sodium currents (68). However, although relevant biological effects are reported for some variants, most variants are also reported in control populations. Behr et al. have recently underlined this issue (69). Using an extended control population, they decreased the yield of such variants from 16.7% in the Hu et al.'s study to 5.1% in a different set of BrS probands (68). Additionally, only two over seven familial pedigrees available with such variants demonstrated segregation with the BrS.

Coding genetic variants in candidate genes are usually classified as likely pathogenic if they are extremely rare or absent from control populations. However, private genetic variants are found in control populations, and many rare variants predicted as damaging are carried by apparently healthy individuals (44, 83). As an example, in the *SCN5A* gene, rare functional variants can be found in about 2% of control patients and even in 5% in non-white population (88). Thus, considering *SCN5A*-mediated BrS account for about 20% of cases and that background noise of rare variant with minor allele frequency under 1/10,000 is approximately 2%, there is a 10/1 signal to noise ratio that means a 10% risk of false positive in possibly damaging rare *SCN5A* variants (82). As prevalence of asymptomatic BrS in the general population is unknown, this percentage may be over estimated. However, as BrS is a rare disease, the proportion of false positive
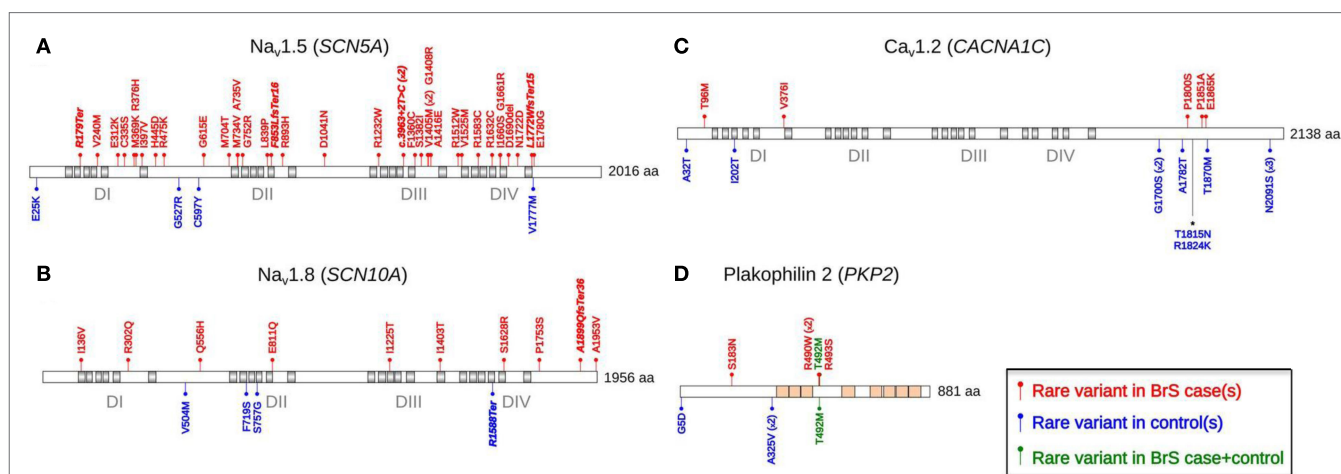
FIGURE 3 | The distribution of rare coding variants detected across four selected arrhythmia-susceptibility genes among 167 BrS cases and 167 healthy individuals. Modified from Ref. (44). *SCN5A* **(A)**, *SCN10A* **(B)**, *CACNA1C* **(C)**, and *PKP2* **(D)** are the four genes exhibiting the largest numbers of rare coding variants among BrS cases. Rare coding variants (minor allele frequency <0.1%) are represented in red (cases) and blue (controls). Green variants are detected in both cases and controls.

variants remains, in any case, too high to be confident with a direct translation of new rare variants in clinical practice.

On the opposite, some rare variants detected among BrS patients are reported as benign by prediction algorithms though they modify the function of the protein. As an example, while one *SCN3B* variant has been associated with BrS and reported as impacting the sodium current density, it is considered as benign by prediction algorithms such as SIFT and PolyPhen-2 (54, 89, 90). This demonstrates the strong limitations of such prediction algorithms and the need for functional studies and/or segregation analyses to better assess the causality of rare variants.

From that perspective, mutations in L-type calcium channels (*CACNA1C*, *CACNB2B*, and *CACNA2D1*) that were considered as associated with about 4% of BrS cases are of particular interest (43). The L-type calcium current I_{Ca-L} is a perfect candidate to explain BrS physiopathology, due to its central role in action potential dome (phases 2 and 3) and in the "depolarization" hypothesis (91). However, functional studies on mutations in L-type calcium channels are scarce in the literature. Moreover, mutations in *CACNA1C* among BrS cases and controls are mostly located within the C-terminal tail of $Ca_v1.2$, thus suggesting a high genetic variability of the domain (**Figure 3**). Although *CACNA1C* mutations seem to play lesser role than previously reported, this particular gene remains involved in a small subset of BrS cases, in particularly those with combined phenotypes of BrS and short QT syndrome (92).

These accumulated data demonstrate that in order to avoid misinterpretation of genetic variants: (1) functional prediction algorithms should be used cautiously and (2) ancestry-matched control populations should be systematically considered. Furthermore, familial segregation analysis and/or extended functional testing are mandatory before associating rare coding variants to disease susceptibility.

Following these guidelines, no previously reported susceptibility gene except *SCN5A* seems to contribute significantly to BrS pathophysiology. Although *SCN5A* remains the major gene involved in BrS with about 20% of carriers among probands (43, 44), a proportion of rare variants residing in this gene – particularly among those of uncertain functional effect – could play no role in relation with the disease (82).

THE COMPLEX INHERITANCE OF BrS: TOWARD A NEW GENETIC MODEL

Since the discovery of *SCN5A* as the first susceptibility gene for BrS, this disorder has been consistently reported as a monogenic disease with autosomal dominant mode of inheritance, caused by rare genetic variants with large effect size (1, 38, 42). *SCN5A* remains the only major susceptibility gene for BrS, with more than 300 coding variants described among more than 75% of the genetically diagnosed patients (43, 93). However, hardly any of the large family pedigrees with BrS provides evidence for Mendelian inheritance. Most familial forms indicate a genetic model with incomplete penetrance and remain genetically undiagnosed.

We have investigated the cosegregation of *SCN5A* mutations with BrS among large genotyped families (41). *SCN5A* mutations

exhibit low penetrance (61% after drug testing) in families, leading to poor genotype/phenotype correlations. More surprisingly, among five pedigrees, we could identify eight affected members who did not carry the familial *SCN5A* mutation (**Figure 2**). This lack of genotype/phenotype correlation is further emphasized in other families with variable cardiac phenotypes associated with a same *SCN5A* mutation. Indeed, although a Na current decrease could lead to cardiac conduction or sinus node dysfunction, the description of relatives sharing the same *SCN5A* mutation but presenting with either BrS or a progressive cardiac conduction disease question about the relevance of a monogenic model (94, 95). A similar issue involving *SCN5A* mutation has been described with BrS and long QT syndrome (96).

These observations have led us to seek for genetic factors modulating the risk of Brugada ECG phenotype. To explore the potential role of common genetic variants in susceptibility to Brs, we have recently coordinated an international genome-wide association study (GWAS) on BrS. By comparing allele frequencies of common haplotypes genome wide among 312 index cases versus 1,115 control individuals, we identified three loci associated with susceptibility to BrS (**Figure 4A**). The three hits were then replicated on independent case–control sets from Europe and Japan. We found that their cumulative effect on disease susceptibility was unexpectedly large, with an estimated odds ratio of 21.5 in the presence of more than four risk alleles versus less than two (**Figure 4B**). This study demonstrates that an aggregation of genetic polymorphisms can strongly influence the susceptibility to BrS and confirms that the mode of inheritance for this arrhythmia disorder is far more complex than previously described.

Two association signals reside at the *SCN5A-SCN10A* locus. Both common risk alleles have previously been associated with cardiac conduction traits in the general population (97). This finding demonstrates that genetic polymorphisms modulating cardiac conduction can also influence susceptibility to cardiac arrhythmia. One haplotype is located inside the *SCN10A* gene, of which involvement in the pathophysiology of BrS is still matter to debate. van den Boogaard et al. provided evidence that the *SCN10A* haplotype contain was an enhancer region for both *SCN10A* and *SCN5A* genes (98). They further demonstrated that a common variant (rs6801957) of this locus, associated with cardiac conduction trait and in high linkage disequilibrium with rs10428132, alters a transcription factor binding site for *TBX3*/*TBX5* and reduces the *SCN5A* expression (99). This may explain the high phenotype variability observed in BrS patients even within a same family.

The third association signal resides near the *Hey2* gene, which encodes a basic helix-loop-helix transcriptional repressor expressed in the cardiovascular system. The implication of this gene in susceptibility to BrS was previously unknown (100). Interestingly, Hey2 presents a gradient of expression across the ventricular wall in mirror image with *SCN5A* expression suggesting a possible (indirect) regulation mechanism. Despite no ECG changes, Hey2 heterozygous knockout mice (Hey2$^{+/-}$) present interesting findings for BrS pathophysiology. Conduction velocity seems specifically increase in the right outflow tract in which cellular action potential present both increase in AP upstroke
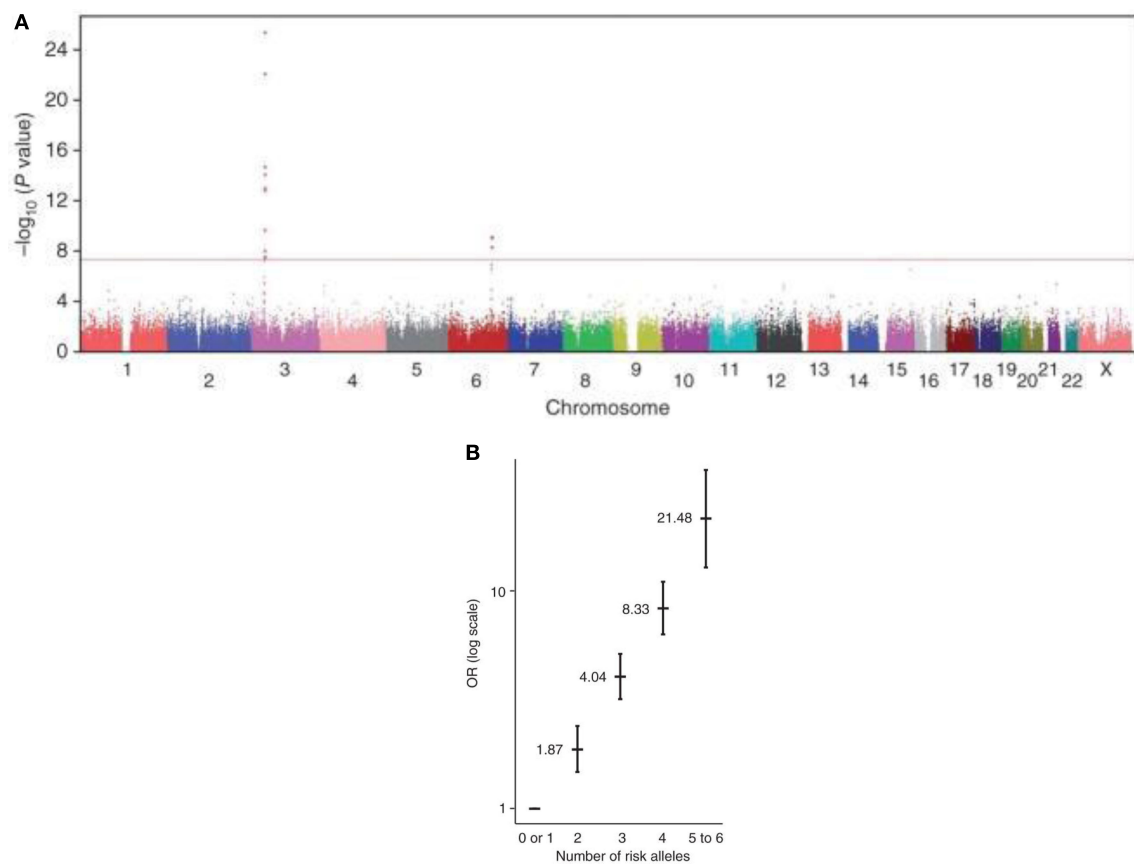
FIGURE 4 | A genome-wide association study on BrS. Modified from Ref. (101). **(A)** Manhattan plot revealing signal associations between two SNP (SCN10A and HEY2) and BrS. Statistical significance is represented with a red line ($P = 5 \times 10^{-8}$). A third haplotype of SCN5A reached genome-wide significance after replication analysis. **(B)** The cumulative effect of the six common risk alleles on susceptibility to BrS. Odds Ratio (OR) is plotted on the vertical axis and the cumulative number of alleles (from 0 to 6) in horizontal axis.

velocity and repolarization (101). These data uncovered the role of Hey2 in the cardiac electrical function and more specifically in the pathogenesis of BrS. Among its role on BrS phenotype, common variant in this gene could also presented with a protective role from ventricular fibrillation in BrS patients by regulating the repolarization current (102).

CONCLUSION

Almost two decades ago, the first description of a mutation in *SCN5A* gene has paved the way of genetics in BrS. As BrS was initially described as a Mendelian disease with low penetrance, many studies have been performed to track genetic variants in families affected by this syndrome. However, in most cases, studies were unable to show positive linkage. In a very large majority of cases, putative causing genes were identified through a "candidate gene approach" based on pathophysiological hypotheses. In these *a priori* approaches, the results were "validated" by the rarity of the genetic variants identified, while aberrant linkage results were "explained" by non-penetrance or phenocopies.

In the recent years, NGS technologies have dramatically expanded our capacity to sequence genomes. It has also revealed the high variability of the human genome, underlying the extreme caution that should be taken to avoid misinterpretation of the potential association of rare variants with BrS. Thus, recent burden tests have questioned the implication of several genes previously identified as there distribution was similar in the normal population and affected patients. For now, only rare variants in *SCN5A* gene seem to be significantly associated with the syndrome.

However, genotype/phenotype studies among BrS families with *SCN5A* mutation carriers have highlighted a complex mode of inheritance for this syndrome. In line with these reports, a GWAS has recently identified three common risk haplotypes for the Brugada ECG pattern.

It is now established that the molecular mechanisms leading to BrS involve both rare and common genetic variants, underlying the need for better understanding the genetic architecture of BrS prior to applying genetics as a diagnostic tool. For the next future, one of the challenges that could contribute to a more efficient

strategy for BrS would be to decipher the role of the combination of variants both for diagnosis and prognosis.

Another source of progress regarding risk stratification among BrS patients could go through the identification of specific ECG indices associated with higher risk of (fatal) arrhythmia. Genetic variants at the *SCN5A*, *HEY2*, and *SCN10A* loci have been associated with arrhythmia occurrence in independent studies (47, 102, 103). Integrating such effects toward establishing a global genetic model for BrS is the next step before including genetic testing into the clinical management of BrS.

Besides the direct benefit of this research on the BrS for itself, it appears increasingly that this primary electrical disorder affecting the young adult (with no identifiable structural abnormalities and presenting limited exposure to environment side effect) may represent a relevant model for the identification of markers and mechanism implied into broader common cardiac arrhythmias.

Retrospectively, *SCN10A* common variant identified in the BrS GWAS study have been also associated with the risk of VF in the context of myocardial infarction and with the pacemaker implantation rate (103, 104). Additionally, a protective role against developing AF has been suggested for both common variants previously identified as risk alleles for BrS at the SCN10A–SCN5A locus. This reinforces the interest of rare diseases to help identifying the pathophysiological bases of common pathologies. As they constitute homogenous groups of patients, rare arrhythmia disorders can provide new molecular insights that may be relevant to the broader health issue of SCD (105).

AUTHOR CONTRIBUTIONS

All authors authored sections of the manuscript, contributed to the figure design, and approved the final version.

REFERENCES

1. Brugada P, Brugada J. Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome: a multicenter report. *J Am Coll Cardiol* (1992) **20**:1391–6. doi:10.1016/0735-1097(92)90253-J

2. Juang J-M, Huang SKS. Brugada syndrome – an under-recognized electrical disease in patients with sudden cardiac death. *Cardiology* (2004) **101**:157–69. doi:10.1159/000076693

3. Papadakis M, Raju H, Behr ER, De Noronha SV, Spath N, Kouloubinis A, et al. Sudden cardiac death with autopsy findings of uncertain significance: potential for erroneous interpretation. *Circ Arrhythm Electrophysiol* (2013) **6**:588–96. doi:10.1161/CIRCEP.113.000111

4. Wilde AAM, Antzelevitch C, Borggrefe M, Brugada J, Brugada R, Brugada P, et al. Proposed diagnostic criteria for the Brugada syndrome: consensus report. *Circulation* (2002) **106**:2514–9. doi:10.1161/01.CIR.0000034169.45752.4A

5. Antzelevitch C, Brugada P, Borggrefe M, Brugada J, Brugada R, Corrado D, et al. Brugada syndrome: report of the second consensus conference endorsed by the Heart Rhythm Society and the European Heart Rhythm Association. *Circulation* (2005) **111**:659–70. doi:10.1161/01.CIR.0000152479.54298.51

6. Priori SG, Wilde AA, Horie M, Cho Y, Behr ER, Berul C, et al. HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of patients with inherited primary arrhythmia syndromes: document endorsed by HRS, EHRA, and APHRS in May 2013 and by ACCF, AHA, PACES, and AEPC in June 2013. *Heart Rhythm* (2013) **10**:1932–63. doi:10.1016/j.hrthm.2013.05.014

7. Miyazaki T, Mitamura H, Miyoshi S, Soejima K, Aizawa Y, Ogawa S. Autonomic and antiarrhythmic drug modulation of ST segment elevation in patients with Brugada syndrome. *J Am Coll Cardiol* (1996) **27**:1061–70. doi:10.1016/0735-1097(95)00613-3

8. Brugada R, Brugada J, Antzelevitch C, Kirsch GE, Potenza D, Towbin JA, et al. Sodium channel blockers identify risk for sudden death in patients with ST-segment elevation and right bundle branch block but structurally normal hearts. *Circulation* (2000) **101**:510–5. doi:10.1161/01.CIR.101.5.510

9. Wolpert C, Echternach C, Veltmann C, Antzelevitch C, Thomas GP, Spehl S, et al. Intravenous drug challenge using flecainide and ajmaline in patients with Brugada syndrome. *Heart Rhythm* (2005) **2**:254–60. doi:10.1016/j.hrthm.2004.11.025

10. Priori SG, Napolitano C, Gasparini M, Pappone C, Della Bella P, Giordano U, et al. Natural history of Brugada syndrome: insights for risk stratification and management. *Circulation* (2002) **105**:1342–7. doi:10.1161/hc1102.105288

11. Gallagher MM, Forleo GB, Behr ER, Magliano G, De Luca L, Morgia V, et al. Prevalence and significance of Brugada-type ECG in 12,012 apparently healthy European subjects. *Int J Cardiol* (2008) **130**:44–8. doi:10.1016/j.ijcard.2007.07.159

12. Miyasaka Y, Tsuji H, Yamada K, Tokunaga S, Saito D, Imuro Y, et al. Prevalence and mortality of the Brugada-type electrocardiogram in one city in Japan. *J Am Coll Cardiol* (2001) **38**:771–4. doi:10.1016/S0735-1097(01)01419-X

13. Nademanee K, Veerakul G, Nimmannit S, Chaowakul V, Bhuripanyo K, Likittanasombat K, et al. Arrhythmogenic marker for the sudden unexplained death syndrome in Thai men. *Circulation* (1997) **96**:2595–600. doi:10.1161/01.CIR.96.8.2595

14. Probst V, Veltmann C, Eckardt L, Meregalli PG, Gaita F, Tan HL, et al. Long-term prognosis of patients diagnosed with Brugada syndrome: results from the FINGER Brugada Syndrome Registry. *Circulation* (2010) **121**:635–43. doi:10.1161/CIRCULATIONAHA.109.887026

15. Benito B, Sarkozy A, Mont L, Henkens S, Berruezo A, Tamborero D, et al. Gender differences in clinical manifestations of Brugada syndrome. *J Am Coll Cardiol* (2008) **52**:1567–73. doi:10.1016/j.jacc.2008.07.052

16. Sacher F, Arsac F, Wilton SB, Derval N, Denis A, de Guillebon M, et al. Syncope in Brugada syndrome patients: prevalence, characteristics, and outcome. *Heart Rhythm* (2012) **9**:1272–9. doi:10.1016/j.hrthm.2012.04.013

17. Take Y, Morita H, Toh N, Nishii N, Nagase S, Nakamura K, et al. Identification of high-risk syncope related to ventricular fibrillation in patients with Brugada syndrome. *Heart Rhythm* (2012) **9**:752–9. doi:10.1016/j.hrthm.2011.11.045

18. Sacher F, Probst V, Maury P, Babuty D, Mansourati J, Komatsu Y, et al. Outcome after implantation of a cardioverter-defibrillator in patients with Brugada syndrome: a multicenter study-part 2. *Circulation* (2013) **128**:1739–47. doi:10.1161/CIRCULATIONAHA.113.001941

19. Sarkozy A, Boussy T, Kourgiannides G, Chierchia GB, Richter S, De Potter T, et al. Long-term follow-up of primary prophylactic implantable cardioverter-defibrillator therapy in Brugada syndrome. *Eur Heart J* (2007) **28**:334–44. doi:10.1093/eurheartj/ehm478

20. Priori SG, Gasparini M, Napolitano C, Della Bella P, Ottonelli AG, Sassone B, et al. Risk stratification in Brugada syndrome: results of the PRELUDE (PRogrammed ELectrical stimUlation preDictive valuE) registry. *J Am Coll Cardiol* (2012) **59**:37–45. doi:10.1016/j.jacc.2011.08.064

21. Conte G, Sieira J, Ciconte G, de Asmundis C, Chierchia GB, Baltogiannis G, et al. Implantable cardioverter-defibrillator therapy in Brugada syndrome: a 20-year single-center experience. *J Am Coll Cardiol* (2015) **65**:879–88. doi:10.1016/j.jacc.2014.12.031

22. Sroubek J, Probst V, Mazzanti A, Delise P, Hevia JC, Ohkubo K, et al. Programmed ventricular stimulation for risk stratification in the Brugada syndrome: a pooled analysis. *Circulation* (2016) **133**:622–30. doi:10.1161/CIRCULATIONAHA.115.017885

23. Probst V, Chatel S, Gourraud J-B, Marec HL. Risk stratification and therapeutic approach in Brugada syndrome. *Arrhythm Electrophysiol Rev* (2012) **1**:17–21. doi:10.15420/aer.2012.1.1.17

24. Belhassen B, Viskin S, Fish R, Glick A, Setbon I, Eldar M. Effects of electrophysiologic-guided therapy with Class IA antiarrhythmic drugs on the long-term outcome of patients with idiopathic ventricular fibrillation

with or without the Brugada syndrome. *J Cardiovasc Electrophysiol* (1999) **10**:1301–12. doi:10.1111/j.1540-8167.1999.tb00183.x

25. Hermida JS, Denjoy I, Clerc J, Extramiana F, Jarry G, Milliez P, et al. Hydroquinidine therapy in Brugada syndrome. *J Am Coll Cardiol* (2004) **43**:1853–60. doi:10.1016/j.jacc.2003.12.046

26. Belhassen B, Glick A, Viskin S. Efficacy of quinidine in high-risk patients with Brugada syndrome. *Circulation* (2004) **110**:1731–7. doi:10.1161/01.CIR.0000143159.30585.90

27. Probst V, Gourraud J-B. Quinidine in Brugada syndrome: still a long way to go…. *Circ Arrhythm Electrophysiol* (2015) **8**:1309–10. doi:10.1161/CIRCEP.115.003576

28. Adler A, Rosso R, Chorin E, Havakuk O, Antzelevitch C, Viskin S. Risk stratification in Brugada syndrome: clinical characteristics, electrocardiographic parameters, and auxiliary testing. *Heart Rhythm* (2016) **13**:299–310. doi:10.1016/j.hrthm.2015.08.038

29. Yan G-X, Antzelevitch C. Cellular basis for the electrocardiographic J wave. *Circulation* (1996) **93**:372–9. doi:10.1161/01.CIR.93.2.372

30. Diego JMD, Sun ZQ, Antzelevitch C. I(to) and action potential notch are smaller in left vs. right canine ventricular epicardium. *Am J Physiol* (1996) **271**:H548–61.

31. Fujiki A, Usui M, Nagasawa H, Mizumaki K, Hayashi H, Inoue H. ST segment elevation in the right precordial leads induced with Class IC antiarrhythmic drugs: insight into the mechanism of Brugada syndrome. *J Cardiovasc Electrophysiol* (1999) **10**:214–8. doi:10.1111/j.1540-8167.1999.tb00662.x

32. Nagase S, Kusano KF, Morita H, Fujimoto Y, Kakishita M, Nakamura K, et al. Epicardial electrogram of the right ventricular outflow tract in patients with the Brugada syndrome: using the epicardial lead. *J Am Coll Cardiol* (2002) **39**:1992–5. doi:10.1016/S0735-1097(02)01888-0

33. Tukkie R, Sogaard P, Vleugels J, de Groot IK, Wilde AA, Tan HL. Delay in right ventricular activation contributes to Brugada syndrome. *Circulation* (2004) **109**:1272–7. doi:10.1161/01.CIR.0000118467.53182.D1

34. Wilde AA, Postema PG, Di Diego JM, Viskin S, Morita H, Fish JM, et al. The pathophysiological mechanism underlying Brugada syndrome: depolarization versus repolarization. *J Mol Cell Cardiol* (2010) **49**:543–53. doi:10.1016/j.yjmcc.2010.07.012

35. Boukens BJ, Sylva M, de Gier-de Vries C, Remme CA, Bezzina CR, Christoffels VM, et al. Reduced sodium channel function unmasks residual embryonic slow conduction in the adult right ventricular outflow tract. *Circ Res* (2013) **113**:137–41. doi:10.1161/CIRCRESAHA.113.301565

36. Watanabe H, Minamino T. Genetics of Brugada syndrome. *J Hum Genet* (2016) **61**:57–60. doi:10.1038/jhg.2015.97

37. Spoonamore KG, Ware SM. Genetic testing and genetic counseling in patients with sudden death risk due to heritable arrhythmias. *Heart Rhythm* (2016) **13**:789–97. doi:10.1016/j.hrthm.2015.11.013

38. Kobayashi T, Shintani U, Yamamoto T, Shida S, Isshiki N, Tanaka T, et al. Familial occurrence of electrocardiographic abnormalities of the Brugada-type. *Intern Med* (1996) **35**:637–40. doi:10.2169/internalmedicine.35.637

39. Garg A, Finneran W, Feld GK. Familial sudden cardiac death associated with a terminal QRS abnormality on surface 12-lead electrocardiogram in the index case. *J Cardiovasc Electrophysiol* (1998) **9**:642–7. doi:10.1111/j.1540-8167.1998.tb00947.x

40. Brugada P, Brugada R, Brugada J. Sudden death in patients and relatives with the syndrome of right bundle branch block, ST segment elevation in the precordial leads V1to V3and sudden death. *Eur Heart J* (2000) **21**:321–6. doi:10.1053/euhj.1999.1751

41. Probst V, Wilde AA, Barc J, Sacher F, Babuty D, Mabo P, et al. SCN5A mutations and the role of genetic background in the pathophysiology of Brugada syndrome. *Circ Cardiovasc Genet* (2009) **2**:552–7. doi:10.1161/CIRCGENETICS.109.853374

42. Chen Q, Kirsch GE, Zhang D, Brugada R, Brugada J, Brugada P, et al. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* (1998) **392**:293–6. doi:10.1038/33305

43. Crotti L, Marcou CA, Tester DJ, Castelletti S, Giudicessi JR, Torchio M, et al. Spectrum and prevalence of mutations involving BrS1- through BrS12-susceptibility genes in a cohort of unrelated patients referred for Brugada syndrome genetic testing: implications for genetic testing. *J Am Coll Cardiol* (2012) **60**:1410–8. doi:10.1016/j.jacc.2012.04.037

44. Le Scouarnec S, Karakachoff M, Gourraud JB, Lindenbaum P, Bonnaud S, Portero V, et al. Testing the burden of rare variation in arrhythmia-susceptibility genes provides new insights into molecular diagnosis for Brugada syndrome. *Hum Mol Genet* (2015) **24**:2757–63. doi:10.1093/hmg/ddv036

45. Andorin A, et al. The impact of clinical and genetic findings on the management of young Brugada syndrome patients. *Heart Rhythm* (2016). doi:10.1016/j.hrthm.2016.02.013

46. Sommariva E, Pappone C, Martinelli Boneschi F, Di Resta C, Rosaria Carbone M, Salvi E, et al. Genetics can contribute to the prognosis of Brugada syndrome: a pilot model for risk stratification. *Eur J Hum Genet* (2013) **21**:911–7. doi:10.1038/ejhg.2012.289

47. Meregalli PG, Tan HL, Probst V, Koopmann TT, Tanck MW, Bhuiyan ZA, et al. Type of SCN5A mutation determines clinical severity and degree of conduction slowing in loss-of-function sodium channelopathies. *Heart Rhythm* (2009) **6**:341–8. doi:10.1016/j.hrthm.2008.11.009

48. Nishii N, Ogawa M, Morita H, Nakamura K, Banba K, Miura D, et al. SCN5A mutation is associated with early and frequent recurrence of ventricular fibrillation in patients with Brugada syndrome. *Circ J* (2010) **74**:2572–8. doi:10.1253/circj.CJ-10-044

49. London B, Michalec M, Mehdi H, Zhu X, Kerchner L, Sanyal S, et al. Mutation in glycerol-3-phosphate dehydrogenase 1 like gene (GPD1-L) decreases cardiac Na+ current and causes inherited arrhythmias. *Circulation* (2007) **116**:2260–8. doi:10.1161/CIRCULATIONAHA.107.703330

50. Fukuda K, Davies SS, Nakajima T, Ong BH, Kupershmidt S, Fessel J, et al. Oxidative mediated lipid peroxidation recapitulates proarrhythmic effects on cardiac sodium channels. *Circ Res* (2005) **97**:1262–9. doi:10.1161/01.RES.0000195844.31466.e9

51. Antzelevitch C, Pollevick GD, Cordeiro JM, Casis O, Sanguinetti MC, Aizawa Y, et al. Loss-of-function mutations in the cardiac calcium channel underlie a new clinical entity characterized by ST-segment elevation, short QT intervals, and sudden cardiac death. *Circulation* (2007) **115**:442–9. doi:10.1161/CIRCULATIONAHA.106.668392

52. Watanabe H, Koopmann TT, Le Scouarnec S, Yang T, Ingram CR, Schott JJ, et al. Sodium channel β1 subunit mutations associated with Brugada syndrome and cardiac conduction disease in humans. *J Clin Invest* (2008) **118**:2260–8. doi:10.1172/JCI33891

53. Delpón E, Cordeiro JM, Núñez L, Thomsen PE, Guerchicoff A, Pollevick GD, et al. Functional effects of KCNE3 mutation and its role in the development of Brugada syndrome. *Circ Arrhythm Electrophysiol* (2008) **1**:209–18. doi:10.1161/CIRCEP.107.748103

54. Hu D, Barajas-Martinez H, Burashnikov E, Springer M, Wu Y, Varro A, et al. A mutation in the β3 subunit of the cardiac sodium channel associated with Brugada ECG phenotype. *Circ Cardiovasc Genet* (2009) **2**:270–8. doi:10.1161/CIRCGENETICS.108.829192

55. Ueda K, Hirano Y, Higashiuesato Y, Aizawa Y, Hayashi T, Inagaki N, et al. Role of HCN4 channel in preventing ventricular arrhythmia. *J Hum Genet* (2009) **54**:115–21. doi:10.1038/jhg.2008.16

56. Giudicessi JR, Ye D, Tester DJ, Crotti L, Mugione A, Nesterenko VV, et al. Transient outward current (I(to)) gain-of-function mutations in the KCND3-encoded Kv4.3 potassium channel and Brugada syndrome. *Heart Rhythm* (2011) **8**:1024–32. doi:10.1016/j.hrthm.2011.02.021

57. Medeiros-Domingo A, Tan BH, Crotti L, Tester DJ, Eckhardt L, Cuoretti A, et al. Gain-of-function mutation S422L in the KCNJ8-encoded cardiac KATP channel Kir6.1 as a pathogenic substrate for J-wave syndromes. *Heart Rhythm* (2010) **7**:1466–71. doi:10.1016/j.hrthm.2010.06.016

58. Burashnikov E, Pfeiffer R, Barajas-Martinez H, Delpón E, Hu D, Desai M, et al. Mutations in the cardiac L-type calcium channel associated with inherited J-wave syndromes and sudden cardiac death. *Heart Rhythm* (2010) **7**:1872–82. doi:10.1016/j.hrthm.2010.08.026

59. Ohno S, Zankov DP, Ding WG, Itoh H, Makiyama T, Doi T, et al. KCNE5 (KCNE1L) variants are novel modulators of Brugada syndrome and idiopathic ventricular fibrillation. *Circ Arrhythm Electrophysiol* (2011) **4**:352–61. doi:10.1161/CIRCEP.110.959619

60. Olesen MS, Jensen NF, Holst AG, Nielsen JB, Tfelt-Hansen J, Jespersen T, et al. A novel nonsense variant in Nav1.5 cofactor MOG1 eliminates its sodium current increasing effect and may increase the risk of arrhythmias. *Can J Cardiol* (2011) **27**:e17–23. doi:10.1016/j.cjca.2011.01.003

61. Perrin MJ, Adler A, Green S, Al-Zoughool F, Doroshenko P, Orr N, et al. Evaluation of genes encoding for the transient outward current (Ito) identifies the KCND2 gene as a cause of J-wave syndrome associated with sudden cardiac death. *Circ Cardiovasc Genet* (2014) **7**:782–9. doi:10.1161/CIRCGENETICS.114.000623

62. Liu H, Chatel S, Simard C, Syam N, Salle L, Probst V, et al. Molecular genetics and functional anomalies in a series of 248 Brugada cases with 11 mutations in the TRPM4 channel. *PLoS One* (2013) **8**:e54131. doi:10.1371/journal.pone.0054131

63. Riuró H, Beltran-Alvarez P, Tarradas A, Selga E, Campuzano O, Vergés M, et al. A missense mutation in the sodium channel β2 subunit reveals SCN2B as a new candidate gene for Brugada syndrome. *Hum Mutat* (2013) **34**:961–6. doi:10.1002/humu.22328

64. Cerrone M, Lin X, Zhang M, Agullo-Pascual E, Pfenniger A, Chkourko Gusky H, et al. Missense mutations in plakophilin-2 cause sodium current deficit and associate with a Brugada syndrome phenotype. *Circulation* (2014) **129**:1092–103. doi:10.1161/CIRCULATIONAHA.113.003077

65. Hu D, Barajas-Martínez H, Terzic A, Park S, Pfeiffer R, Burashnikov E, et al. ABCC9 is a novel Brugada and early repolarization syndrome susceptibility gene. *Int J Cardiol* (2014) **171**:431–42. doi:10.1016/j.ijcard.2013.12.084

66. Ishikawa T, Sato A, Marcou CA, Tester DJ, Ackerman MJ, Crotti L, et al. A novel disease gene for Brugada syndrome: sarcolemmal membrane-associated protein gene mutations impair intracellular trafficking of hNav1.5. *Circ Arrhythm Electrophysiol* (2012) **5**:1098–107. doi:10.1161/CIRCEP.111.969972

67. Wang Q, Ohno S, Ding WG, Fukuyama M, Miyamoto A, Itoh H, et al. Gain-of-function KCNH2 mutations in patients with Brugada syndrome. *J Cardiovasc Electrophysiol* (2014) **25**:522–30. doi:10.1111/jce.12361

68. Hu D, Barajas-Martínez H, Pfeiffer R, Dezi F, Pfeiffer J, Buch T, et al. Mutations in SCN10A are responsible for a large fraction of cases of Brugada syndrome. *J Am Coll Cardiol* (2014) **64**:66–79. doi:10.1016/j.jacc.2014.04.032

69. Behr ER, Savio-Galimberti E, Barc J, Holst AG, Petropoulou E, Prins BP, et al. Role of common and rare variants in SCN10A: results from the Brugada syndrome QRS locus gene discovery collaborative study. *Cardiovasc Res* (2015) **106**:520–9. doi:10.1093/cvr/cvv042

70. Hennessey JA, Marcou CA, Wang C, Wei EQ, Wang C, Tester DJ, et al. FGF12 is a candidate Brugada syndrome locus. *Heart Rhythm* (2013) **10**:1886–94. doi:10.1016/j.hrthm.2013.09.064

71. Boczek NJ, Ye D, Johnson EK, Wang W, Crotti L, Tester DJ, et al. Characterization of SEMA3A-encoded semaphorin as a naturally occurring Kv4.3 protein inhibitor and its contribution to Brugada syndrome. *Circ Res* (2014) **115**:460–9. doi:10.1161/CIRCRESAHA.115.303657

72. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol* (2000) **18**:630–4. doi:10.1038/76469

73. Porreca GJ, Zhang K, Li JB, Xie B, Austin D, Vassallo SL, et al. Multiplex amplification of large sets of human exons. *Nat Methods* (2007) **4**:931–6. doi:10.1038/nmeth1110

74. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* (2005) **437**:376–80. doi:10.1038/nature03959

75. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* (2012) **337**:64–9. doi:10.1126/science.1219240

76. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* (2013) **493**:216–20. doi:10.1038/nature11690

77. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* (2001) **291**:1304–51. doi:10.1126/science.1058040

78. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* (2001) **409**:860–921. doi:10.1038/35057062

79. Tabor HK, Auer PL, Jamal SM, Chong JX, Yu JH, Gordon AS, et al. Pathogenic variants for Mendelian and complex traits in exomes of 6,517 European and African Americans: implications for the return of incidental results. *Am J Hum Genet* (2014) **95**:183–93. doi:10.1016/j.ajhg.2014.07.006

80. Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, Shirts BH, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res* (2015) **25**:305–15. doi:10.1101/gr.183483.114

81. Van Driest SL, Wells QS, Stallings S, Bush WS, Gordon A, Nickerson DA, et al. Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records. *JAMA* (2016) **315**:47–57. doi:10.1001/jama.2015.17701

82. Ackerman MJ. Genetic purgatory and the cardiac channelopathies: exposing the variants of uncertain/unknown significance issue. *Heart Rhythm* (2015) **12**:2325–31. doi:10.1016/j.hrthm.2015.07.002

83. Risgaard B, Jabbari R, Refsgaard L, Holst AG, Haunsø S, Sadjadieh A, et al. High prevalence of genetic variants previously associated with Brugada syndrome in new exome data. *Clin Genet* (2013) **84**:489–95. doi:10.1111/cge.12126

84. Akopian AN, Sivilotti L, Wood JN. A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons. *Nature* (1996) **379**:257–62. doi:10.1038/379257a0

85. Renganathan M, Cummins TR, Waxman SG. Contribution of Na(v)1.8 sodium channels to action potential electrogenesis in DRG neurons. *J Neurophysiol* (2001) **86**:629–40.

86. Pallante BA, Giovannone S, Fang-Yu L, Zhang J, Liu N, Kang G, et al. Contactin-2 expression in the cardiac Purkinje fiber network. *Circ Arrhythm Electrophysiol* (2010) **3**:186–94. doi:10.1161/CIRCEP.109.928820

87. Yang T, Atack TC, Stroud DM, Zhang W, Hall L, Roden DM. Blocking Scn10a channels in heart reduces late sodium current and is antiarrhythmic. *Circ Res* (2012) **111**:322–32. doi:10.1161/CIRCRESAHA.112.265173

88. Kapa S, Tester DJ, Salisbury BA, Harris-Kerr C, Pungliya MS, Alders M, et al. Genetic testing for long-QT syndrome distinguishing pathogenic mutations from benign variants. *Circulation* (2009) **120**:1752–60. doi:10.1161/CIRCULATIONAHA.109.863076

89. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* (2009) **4**:1073–81. doi:10.1038/nprot.2009.86

90. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* (2010) **7**:248–9. doi:10.1038/nmeth0410-248

91. Antzelevitch C. J wave syndromes: molecular and cellular mechanisms. *J Electrocardiol* (2013) **46**:510–8. doi:10.1016/j.jelectrocard.2013.08.006

92. Watanabe H, Minamino T. Role of mutations in L-type calcium channel genes in Brugada syndrome, early repolarization syndrome, and idiopathic ventricular fibrillation associated with right bundle branch block. *Circ J* (2013) **77**:1689–90. doi:10.1253/circj.CJ-13-0641

93. Kapplinger JD, Tester DJ, Alders M, Benito B, Berthet M, Brugada J, et al. An international compendium of mutations in the SCN5A-encoded cardiac sodium channel in patients referred for Brugada syndrome genetic testing. *Heart Rhythm* (2010) **7**:33–46. doi:10.1016/j.hrthm.2009.09.069

94. Schott JJ, Alshinawi C, Kyndt F, Probst V, Hoorntje TM, Hulsbeek M, et al. Cardiac conduction defects associate with mutations in SCN5A. *Nat Genet* (1999) **23**:20–1. doi:10.1038/12618

95. Kyndt F, Probst V, Potet F, Demolombe S, Chevallier JC, Baro I, et al. Novel SCN5A mutation leading either to isolated cardiac conduction defect or Brugada syndrome in a large French family. *Circulation* (2001) **104**:3081–6. doi:10.1161/hc5001.100834

96. Bezzina C, Veldkamp MW, van Den Berg MP, Postma AV, Rook MB, Viersma JW, et al. A single Na+ channel mutation causing both long-QT and Brugada syndromes. *Circ Res* (1999) **85**:1206–13. doi:10.1161/01.RES.85.12.1206

97. Pfeufer A, van Noord C, Marciante KD, Arking DE, Larson MG, Smith AV, et al. Genome-wide association study of PR interval. *Nat Genet* (2010) **42**:153–9. doi:10.1038/ng.517

98. van den Boogaard M, Wong LY, Tessadori F, Bakker ML, Dreizehnter LK, Wakker V, et al. Genetic variation in T-box binding element functionally affects SCN5A/SCN10A enhancer. *J Clin Invest* (2012) **122**:2519–30. doi:10.1172/JCI62613

99. van den Boogaard M, Smemo S, Burnicka-Turek O, Arnolds DE, van de Werken HJ, Klous P, et al. A common genetic variant within SCN10A modulates cardiac SCN5A expression. *J Clin Invest* (2014) **124**:1844–52. doi:10.1172/JCI73140

100. Reamon-Buettner SM, Borlak J. HEY2 mutations in malformed hearts. *Hum Mutat* (2006) **27**:118. doi:10.1002/humu.9390

101. Bezzina CR, Barc J, Mizusawa Y, Remme CA, Gourraud JB, Simonet F, et al. Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. *Nat Genet* (2013) **45**:1044–9. doi:10.1038/ng.2712

102. Nakano Y, Ochi H, Onohara Y, Toshishige M, Tokuyama T, Matsumura H, et al. Common variant near HEY2 has a protective effect on ventricular fibrillation occurrence in Brugada syndrome by regulating the repolarization current. *Circ Arrhythm Electrophysiol* (2016) **9**. doi:10.1161/CIRCEP.115.003436

103. Bezzina CR, Pazoki R, Bardai A, Marsman RF, de Jong JS, Blom MT, et al. Genome-wide association study identifies a susceptibility locus at 21q21 for ventricular fibrillation in acute myocardial infarction. *Nat Genet* (2010) **42**:688–91. doi:10.1038/ng.623

104. Holm H, Gudbjartsson DF, Arnar DO, Thorleifsson G, Thorgeirsson G, Stefansdottir H, et al. Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* (2010) **42**:117–22. doi:10.1038/ng.511

105. Andreasen L, Nielsen JB, Darkner S, Christophersen IE, Jabbari J, Refsgaard L, et al. Brugada syndrome risk loci seem protective against atrial fibrillation. *Eur J Hum Genet* (2014) **22**:1357–61. doi:10.1038/ejhg.2014.46

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Mitochondrial Cardiomyopathies

Ayman W. El-Hattab[1] and Fernando Scaglia[2]*

[1]*Division of Clinical Genetics and Metabolic Disorders, Department of Pediatrics, Tawam Hospital, Al-Ain, United Arab Emirates,* [2] *Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA*

Mitochondria are found in all nucleated human cells and perform various essential functions, including the generation of cellular energy. Mitochondria are under dual genome control. Only a small fraction of their proteins are encoded by mitochondrial DNA (mtDNA), whereas more than 99% of them are encoded by nuclear DNA (nDNA). Mutations in mtDNA or mitochondria-related nDNA genes result in mitochondrial dysfunction leading to insufficient energy production required to meet the needs for various organs, particularly those with high energy requirements, including the central nervous system, skeletal and cardiac muscles, kidneys, liver, and endocrine system. Because cardiac muscles are one of the high energy demanding tissues, cardiac involvement occurs in mitochondrial diseases with cardiomyopathies being one of the most frequent cardiac manifestations found in these disorders. Cardiomyopathy is estimated to occur in 20–40% of children with mitochondrial diseases. Mitochondrial cardiomyopathies can vary in severity from asymptomatic status to severe manifestations including heart failure, arrhythmias, and sudden cardiac death. Hypertrophic cardiomyopathy is the most common type; however, mitochondrial cardiomyopathies might also present as dilated, restrictive, left ventricular non-compaction, and histiocytoid cardiomyopathies. Cardiomyopathies are frequent manifestations of mitochondrial diseases associated with defects in electron transport chain complexes subunits and their assembly factors, mitochondrial transfer RNAs, ribosomal RNAs, ribosomal proteins, translation factors, mtDNA maintenance, and coenzyme Q_{10} synthesis. Other mitochondrial diseases with cardiomyopathies include Barth syndrome, Sengers syndrome, *TMEM70*-related mitochondrial complex V deficiency, and Friedreich ataxia.

Keywords: hypertrophic cardiomyopathy, dilated cardiomyopathy, restrictive cardiomyopathy, non-compaction cardiomyopathy, histiocytoid cardiomyopathies, Barth syndrome, Friedreich ataxia

INTRODUCTION

Metabolic disorders account for a minority of causes of cardiomyopathies. However, diagnosing a metabolic disease as a cause for cardiomyopathy can have prognostic and therapeutic implications. Major groups of metabolic disorders associated with cardiomyopathy include organic acidemias (e.g., propionic acidemia), fatty acid oxidation defects (e.g., very long chain acyl CoA dehydrogenases deficiency), lysosomal storage diseases (e.g., Fabry disease), glycogen storage diseases (e.g., Pompe disease), congenital disorders of glycosylation, and mitochondrial disorders (1).

Mitochondrial diseases are a clinically and genetically heterogeneous group of disorders that result from dysfunction of the mitochondrial respiratory chain, which is responsible for the generation of most cellular energy (2, 3). Because cardiac muscles are one of the high energy demanding

tissues, cardiac involvement occurs in large number of mito-chondrial diseases. The most frequent cardiac manifestations of mitochondrial diseases are cardiomyopathies. Arrhythmias and conduction defects, pulmonary hypertension, pericardial effu-sion, dilated aortic root, and coronary heart disease can also be seen in mitochondrial diseases (4, 5).

In this article, we review normal mitochondrial structure and function, pathogenesis of mitochondrial diseases, clinical aspects of mitochondrial cardiomyopathies, mitochondrial diseases frequently associated with cardiomyopathies, and diagnosis and management of mitochondrial cardiomyopathies.

NORMAL MITOCHONDRIAL STRUCTURE AND FUNCTION

Mitochondria are found in all nucleated human cells each of which typically contains in its cytoplasm several hundred mitochondria depending on the energy needs for the tissue. Mitochondria are composed of two bilayer membranes that create two distinct compartments: an intermembrane space and a matrix space within the inner membrane. The mitochondrial outer membrane is smooth, whereas the inner mitochondrial membrane is highly folded, forming structures called cristae. The large surface area of the inner mitochondrial membrane accommodates energy-generating multipolypeptide enzyme complexes called respira-tory chain or electron transport chain (ETC) complexes (2).
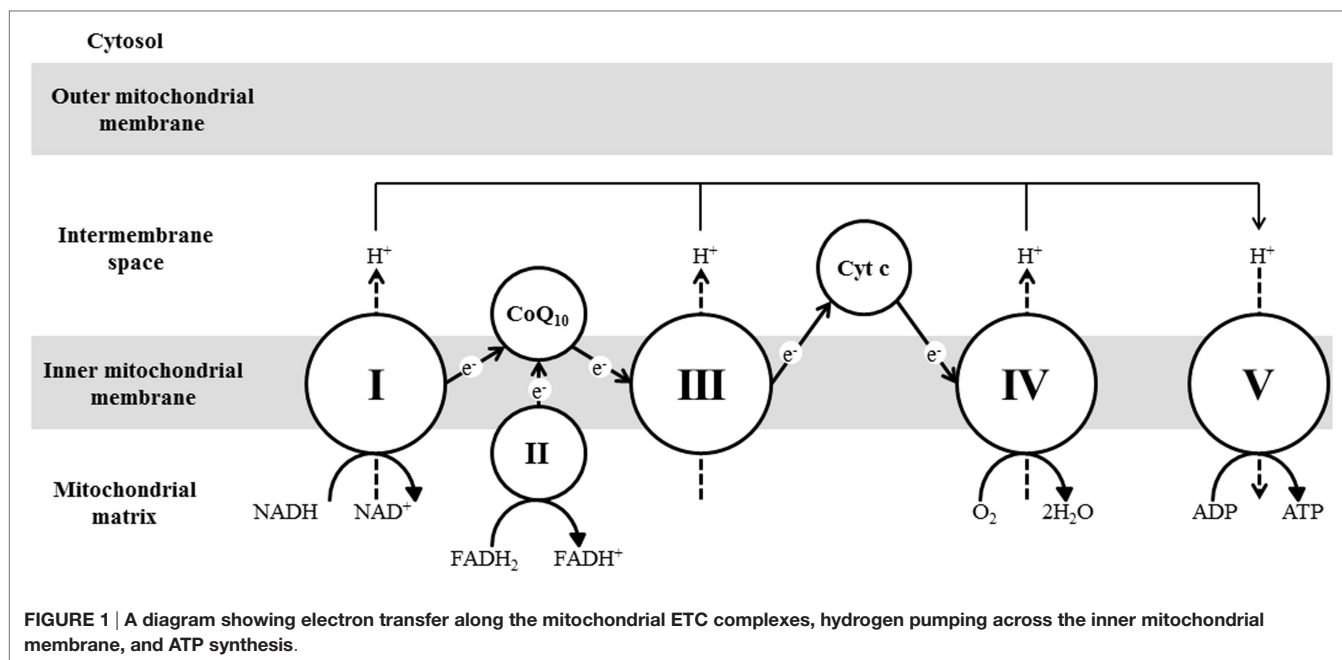
Approximately 1,500 proteins are involved in maintain-ing mitochondrial structure and function; however, <1% are encoded by mitochondrial DNA (mtDNA), while more than 99% of mitochondrial proteins are encoded by nuclear DNA (nDNA). Therefore, mitochondria are under dual genome control. Each mitochondrion contains mtDNA in the form of a multicopy, 16.6 kb circular double-stranded DNA. The mtDNA encodes 13 essential polypeptides for the ETC complexes and 24 different RNAs, including 2 ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs) (3, 6). The remaining ETC complexes subunits, as well as proteins needed to assemble the ETC complexes (assembly factors), maintain mtDNA, and transport molecules across the mitochondrial membranes, are encoded by nDNA, synthesized on cytoplasmic ribosomes, and imported into mitochondria (7). Unlike nDNA, which replicates with each cell division, mtDNA replicates continuously and independently of cell division. Two nDNA-encoded enzymes play major roles in mtDNA replication: DNA polymerase gamma that functions in replication and repair of mtDNA, and the twinkle protein that serves the function of a DNA helicase that is required for mtDNA replication (8). Transcription of mtDNA produces a polycistronic precursor RNA that is then processed to produce individual mRNA, tRNA, and rRNA molecules. The nDNA-encoded mitochondrial RNA polymerase and mitochondrial transcriptions factors are needed for the mitochondrial transcription process (9). The mRNAs for the 13 mtDNA-encoded proteins are translated on mitochondrial ribosomes. Mitochondrial tRNAs and rRNAs are required for this process in addition to several nDNA-encoded proteins, including mitochondrial ribosomal proteins and mitochondrial translation factors (9). The nDNA-encoded mitochondrial polypeptides are synthesized on cytosolic ribosomes and transported into the mitochondria *via* mitochondrial protein import systems, includ-ing the translocase of the outer membrane (TOM) and translocase of the inner membrane (TIM) complexes (7, 10).

Mitochondria perform various essential functions, including the generation of most of the energy needed by cells in the form of ATP in a process called oxidative phosphorylation (OXPHOS) carried out by the ETC complexes in the inner mitochondrial membrane. Complexes I, II, III, and IV make up the ETC, whereas complex V is the ATP synthase. Hydrogen atoms generated from different catabolic pathways bind to nicotinamide adenine dinu-cleotide (NAD^+) and flavin adenine dinucleotide (FAD) to yield NADH and $FADH_2$, respectively. NADH is oxidized by complex I (NADH dehydrogenase), and the electrons are transported through flavin mononucleotide (FMN) and multiple iron–sulfur (Fe–S) centers in complex I until they are transferred to coenzyme Q_{10} (CoQ_{10}). CoQ_{10} also accepts hydrogen atoms from $FADH_2$ generated by β-oxidation and the TCA enzyme succinate dehy-drogenase (complex II). Electrons are subsequently transferred from CoQ_{10} to complex III ($bc1$ complex) within which the elec-trons move through cytochrome b, cytochrome $c1$, and the Fe–S components. The electrons are then transferred from complex III to cytochrome c, which transfers the electrons to complex IV (cytochrome c oxidase). Within this complex, the electrons are transferred through copper centers and cytochromes a and $a3$ and ultimately combine with O_2 to generate H_2O. The energy that is released during electron transfer is used to pump protons from inside the mitochondrial matrix across the inner mitochondrial membrane into the intermembrane space through complexes I, III, and IV. The resulting electrochemical gradient forces protons to move back through a proton channel in complex V (ATP syn-thase), which utilizes this energy in synthesizing ATP. The ETC complexes are multipolypeptides encoded by both mtDNA and nDNA except for complex II, which is encoded entirely by nDNA (11) (**Figure 1**).

MITOCHONDRIAL DYSFUNCTION AND DISEASES

Mutations in mtDNA or mitochondria-related nDNA genes result in mitochondrial dysfunction leading to mitochondrial diseases (12). Defects in mtDNA can be either point muta-tions or rearrangements. Point mutations in mtDNA can affect protein-encoding genes or genes encoding tRNA or rRNA. These mutations are maternally inherited and typically associ-ated with very variable phenotypes. Rearrangements of mtDNA include deletions and duplications that differ in size and position but typically encompass several genes. These rearrangements are usually sporadic arising *de novo* but can be maternally inherited (13). Mutations in nDNA genes are inherited in an autosomal recessive, autosomal dominant, or X-linked manner. Mitochondrial dysfunction can result from mutations in nDNA genes encoding ETC complexes subunits or their assembly fac-tors (11), mitochondrial import complexes (10), mitochondrial ribosomal proteins and translational factors (14), and CoQ_{10} biosynthesis enzymes (15). The mtDNA is maintained by a group

FIGURE 1 | **A diagram showing electron transfer along the mitochondrial ETC complexes, hydrogen pumping across the inner mitochondrial membrane, and ATP synthesis**.

of nDNA-encoded proteins that function either in mitochondrial deoxyribonucleoside triphosphate (dNTP) synthesis or mtDNA replication. Mutations in any of these genes result in depletion of the mitochondrial dNTP pool or impaired mtDNA replication, leading to severe reduction in mtDNA content (mtDNA depletion). Inadequate amount of mtDNA results in impaired synthesis of key subunits of ETC complexes (16). Finally, Fe–S clusters are ubiquitous cofactors that are composed of iron and inorganic sulfur. These clusters are important prosthetic groups that are required for the function of proteins involved in various activities, including electron transport in ETC complexes. Defects in the process of Fe–S clusters can result in impaired ETC activity and mitochondrial dysfunction (17).

Defects in mtDNA- or nDNA-encoded mitochondrial proteins result in mitochondrial respiratory chain dysfunction leading to impaired OXPHOS and inability to generate sufficient energy to meet the needs for various organs, particularly those with high energy demand, including the central nervous system, skeletal and cardiac muscles, kidneys, liver, and endocrine system (2, 3). Additionally, due to the impaired OXPHOS, NADH cannot be utilized and the NADH:NAD ratio increases, which results in the inhibition of the TCA cycle. Pyruvate, produced through glycolysis, is increased due to the TCA cycle inhibition. Both elevated pyruvate and NADH:NAD ratio result in shifting the equilibrium of lactate dehydrogenase toward the production of lactate from pyruvate. Lactate can accumulate, causing systemic acidosis. Lactic acidosis is among one of the common features of mitochondrial disorders (6).

In addition to ATP deficiency, consequences of mitochondrial dysfunction include aberrant calcium handling, excessive reactive oxygen species (ROS) production, apoptosis dysregulation, and nitric oxide (NO) deficiency all of which contribute to the pathogenesis of mitochondrial diseases (18). During OXPHOS, a small part of oxygen is partially reduced and converted to

ROS (superoxide and hydrogen peroxide). Under normal conditions, ROS can be scavenged by various enzymes, including the mitochondrial superoxide dismutase and glutathione peroxidase (19, 20). ROS, whose generation is enhanced as a result of OXPHOS blockade, can irreversibly modify many cellular macromolecules leading to cellular toxicity. Increased ROS production in mitochondrial diseases can result in protein, lipid, and DNA damage, which can potentially lead to further cellular damage and dysfunction (19, 20). One of the mitochondrial functions is calcium buffering. In addition, mitochondrial ATP production is needed to fuel calcium pumps in the plasma membrane and endoplasmic reticulum. Therefore, mitochondrial dysfunction can result in aberrant calcium handling. This model could contribute to the frequent involvement of muscle and nerve tissues in mitochondrial diseases, since these cells rely heavily on ATP and on fluctuating levels of intracellular calcium (21, 22). Mitochondria are also major regulators of apoptosis. In response to several intracellular stress conditions, supermolecular channels called mitochondrial permeability transition pores open resulting in increased mitochondrial inner membrane permeability. Apoptosis is initiated when the inner mitochondrial membrane becomes permeable leading to the release of several toxic mitochondrial proteins into the cytosol, including cytochrome *c*. These proteins activate latent forms of caspases, resulting in the execution of apoptosis. Therefore, excessive cell loss can contribute to the pathology in mitochondrial diseases (23). Finally, there is growing evidence that NO deficiency occurs in mitochondrial diseases and can play a major role in the pathogenesis of several complications observed in these diseases, including stroke-like episodes, myopathy, diabetes, and lactic acidosis. NO deficiency in mitochondrial disorders is multifactorial in origin, including impaired NO production and postproduction sequestration (24).

The mtDNA in cells can be identical (homoplasmy) or a mixture of two or more types (heteroplasmy). Some mtDNA

mutations affect all copies of the mtDNA (homoplasmic mutations), while most of the mutations are present in only some copies of mtDNA and cells harbor a mixture of mutant and normal mtDNA (heteroplasmic mutations) (25). When cell divides, the mitochondria are distributed in a stochastic process in daughter cells. Therefore, when a cell harboring a heteroplasmic mtDNA mutation divides, it is a matter of chance whether the mutant mtDNAs will be partitioned into one daughter cell or another. Therefore, over time, the percentage of mutant mtDNAs can differ in different tissues and organs. This process, which is called replicative segregation, explains why the heteroplasmy percentage of mutant mtDNA may vary among organs and tissues within the same individual. The different tissues and organs rely on mitochondrial energy to different extents. As the percentage of mutated mtDNA increases, energy production declines. When the proportion of mutant mtDNA crosses a critical threshold level, the impaired energy production will result in organ dysfunction and clinical manifestations. The threshold level varies among different organs and tissues depending on their energy requirement (26). The replicative segregation and different organ threshold levels can explain in part the varied clinical phenotypes observed in individuals with mtDNA mutations. On the other hand, the clinical phenotypes of nDNA-related mitochondrial diseases are typically more homogenous than the mtDNA-related disease, as all the mitochondria are similarly affected (3).

Mitochondrial disorders are not uncommon with a minimum prevalence of 1 in 5,000 (12). Mitochondria are essential components of all nucleated cells. Therefore, mitochondrial dysfunction affects many organs, particularly those with high energy requirements. Insufficient energy for various organs results in multi-organ dysfunction and the variable manifestations observed in mitochondrial diseases, including epilepsy, intellectual disability, skeletal and cardiac myopathies, hepatopathies, endocrinopathies, and nephropathies (2, 3, 6). Although the vast majority of mitochondrial diseases involve multiple organ systems, some mitochondrial diseases may affect a single organ (e.g., Leber hereditary optic neuropathy, and non-syndromic sensorineural hearing loss) (2). Mitochondrial diseases can begin at any age. Many patients with mitochondrial diseases display a cluster of clinical features that fall into a discrete clinical syndrome such as Kearns–Sayre syndrome, mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), myoclonic epilepsy with ragged-red fibers (MERRF), neurogenic weakness with ataxia and retinitis pigmentosa (NARP), mitochondrial neurogastrointestinal encephalopathy (MNGIE), and Alpers syndrome. However, there is often considerable clinical variability, and many affected individuals do not fit into one particular syndrome (2, 3, 6).

CLINICAL ASPECTS OF MITOCHONDRIAL CARDIOMYOPATHIES

Mitochondrial cardiomyopathy can be described as a myocardial disorder characterized by abnormal myocardial structure and/or function secondary to genetic defects resulting in the impairment of the mitochondrial respiratory chain, in the absence of concomitant coronary artery disease, hypertension, valvular disease,

and congenital heart disease (27). Cardiomyopathy is estimated to occur in 20–40% of children with mitochondrial diseases (5, 28). Therefore, screening for cardiomyopathy is a standard part of the management of children and adults with known or suspected mitochondrial disease (29).

Mitochondrial cardiomyopathies can vary in severity from asymptomatic status to severe manifestations, including heart failure, arrhythmias, and sudden cardiac death. Cardiac manifestations can be precipitated or worsen during metabolic decompensation episodes that are often caused by stressors, such as febrile illnesses or surgery, and can be accompanied by acute heart failure (27). It has been reported that mortality in children with mitochondrial diseases is significantly higher in those with cardiomyopathy than in those without (28). The clinical manifestations of mitochondrial cardiomyopathies are often accompanied by other manifestations of the multi-organ involvement of mitochondrial diseases. On the other hand, mitochondrial cardiomyopathy can occur in the absence of known mitochondrial disease, of which it may be the first or the sole clinical manifestation (29).

Hypertrophic cardiomyopathy is the most common form; however, mitochondrial cardiomyopathies might also present as dilated, restrictive, left ventricular non-compaction, and histiocytoid cardiomyopathies (4). Hypertrophic cardiomyopathy is the most frequent cardiac manifestation in mitochondrial diseases and can occur in more than 50% of individuals with mitochondrial cardiomyopathies (5). It can be detected as early as the antenatal period and may be the only manifestation of a mitochondrial disease or a part of a multi-organ disease. Obstructive hypertrophic cardiomyopathy rarely occurs, but hypertrophic cardiomyopathy frequently develops into systolic dysfunction followed by decompensation and dilatation of the left ventricle (30). Dilated cardiomyopathy, which can be primary or secondary following hypertrophic cardiomyopathy, occurs less frequently than hypertrophic cardiomyopathy, whereas restrictive cardiomyopathy is a rare manifestation of mitochondrial diseases (31). Although left ventricular non-compaction cardiomyopathy is also a rare finding in mitochondrial diseases, among individuals with non-compaction, mitochondrial diseases are highly prevalent. Left ventricular non-compaction cardiomyopathy is generally more frequent in males and tends to develop during pregnancy in females. Occasionally, it may disappear during the disease course in some individuals with mitochondrial diseases (32). Histiocytoid cardiomyopathy (Purkinje fiber dysplasia) is histologically characterized by morphological and functional abnormalities of cardiomyocytes and Purkinje cells with a cytoplasm like in histiocyte foam cells, which contain glycogen and lipids. It has been reported exclusively in individuals with mitochondrial diseases (33).

MITOCHONDRIAL DISEASES FREQUENTLY ASSOCIATED WITH CARDIOMYOPATHIES

Cardiomyopathies are frequent manifestations of mitochondrial diseases associated with defects in ETC complexes subunits and

their assembly factors, mitochondrial tRNAs, rRNAs, ribosomal proteins, translation factors, mtDNA maintenance, and CoQ$_{10}$ synthesis. Other mitochondrial diseases with cardiomyopathies include Barth syndrome and other 3-methylglutaconic aciduria disorders, and Friedreich ataxia (**Table 1**).

Complex I deficiency, which is clinically and genetically heterogeneous, can present with hypertrophic cardiomyopathy that might be isolated or associated with multi-organ disease. Cardiomyopathy has been reported with mutations in

TABLE 1 | Mitochondrial diseases frequently associated with cardiomyopathies.

Mitochondrial diseases	Genes	Clinical manifestations
ETC complexes deficiencies		
Complex I deficiency	Subunit mtDNA genes: *MTND1*, *MTND2*, *MTND4*, *MTND5*, and *MTND6* Subunit nDNA genes: *NDUFV1*, *NDUFV2*, *NDUFS1*, *NDUFS2*, *NDUFS3*, *NDUFS4*, *NDUFS6*, *NDUFS7*, *NDUFS8*, *NDUFA2*, *NDUFA11*, *NDUFAF3*, *NDUFA10*, *NDUFB3*, *NDUFB9*, and *NDUFA1* Assembly genes: *NDUFAF2*, *NDUFAF4*, *NDUFAF5*, *NUBPL*, *NDUFAF1*, *FOXRED1*, and *ACAD9*	Hypertrophic cardiomyopathy Growth failure Developmental delay Epilepsy Ataxia Weakness Spasticity Leukoencephalopathy Macrocephaly Sensorineural deafness Hepatic dysfunction Lactic acidosis Hypoglycemia
Complex II deficiency	Subunit genes: *SDHA* and *SDHD*	Hypertrophic, dilated, and non-compaction cardiomyopathies Growth failure Developmental delay Weakness Spasticity Ataxia Epilepsy Leukodystrophy Contractures Ophthalmoplegia Pigmentary retinopathy Optic atrophy Lactic acidosis
Complex III deficiency	*MTCYB*	Hypertrophic, dilated, and histiocytoid cardiomyopathies Growth failure Exercise intolerance Optic atrophy Stroke-like episodes Epilepsy Lactic acidosis Hypoglycemia
Complex IV deficiency	Subunit mtDNA genes: *MTCO1*, *MTCO2*, and *MTCO3* Subunit nDNA genes: *COX6B1* Assembly factors: *COX10*, *COX14*, *COX15*, *COX20*, *SCO1*, *SCO2*, *COA3*, and *COA5*	Dilated, hypertrophic, and histiocytoid cardiomyopathies Growth failure Developmental delay Ataxia Epilepsy Hypotonia

(Continued)

TABLE 1 | Continued

Mitochondrial diseases	Genes	Clinical manifestations
		Sensorineural hearing loss Optic atrophy Pigmentary retinopathy Liver dysfunction Renal tubulopathy Lactic acidosis
Mitochondrial tRNA genes		
MERRF (myoclonic epilepsy with ragged-red fibers) syndrome	*MTTK*	Dilated and histiocytoid cardiomyopathy Epilepsy Ataxia Weakness Sensorineural hearing loss Short stature Lactic acidosis
MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) syndrome	*MTTL1*	Hypertrophic cardiomyopathy Muscle weakness Stroke-like episodes Dementia Epielpsy Sensorineural hearing loss Lactic acidosis Diabetes Short stature
Mitochondrial DNA depletion		
Mitochondrial neurogastrointestinal encephalopathy syndrome (MNGIE)	*TYMP*	Hypertrophic cardiomyopathy Gastrointestinal dysmotility Cachexia Ptosis Ophthalmoplegia Hearing loss Peripheral neuropathy Leukoencephalopathy
CoQ10 deficiency		
Coenzyme Q10 deficiency	*COQ2*, *COQ4*, *COQ6*, *COQ7*, *COQ9*, *ADCK3*, *PDSS1*, and *PDSS2*	Hypertrophic cardiomyopathy Growth failure Developmental delay Weakness Epilepsy Ataxia Pigmentary retinopathy Sensorineural hearing loss Liver dysfunction Renal impairment Pancytopenia Lactic acidosis

(Continued)

TABLE 1 | Continued

Mitochondrial diseases	Genes	Clinical manifestations
3-Methylglutaconic acidurias		
3-Methylglutaconic aciduria type II (Barth syndrome)	TAZ	Non-compaction, dilated, and hypertrophic cardiomyopathies Growth failure Weakness Arrhythmias Neutropenia
3-Methylglutaconic aciduria, type V (dilated cardiomyopathy and ataxia syndrome)	DNAJC19	Dilated and non-compaction cardiomyopathies Growth failure Ataxia Testicular dysgenesis Anemia
Mitochondrial complex V deficiency	TMEM70	Hypertrophic cardiomyopathy Growth failure Developmental delay Hypotonia Ataxia Epilepsy Leukodystrophy Distinctive facial features Lactic acidosis Hyperammonemia
Sengers syndrome	AGK	Hypertrophic cardiomyopathy Growth failure Cataracts Hypotonia Weakness Lactic acidosis
Defects in iron–sulfur cluster		
Friedreich ataxia	FXN	Hypertrophic cardiomyopathy Ataxia Dysarthria Peripheral sensory neuropathy Diabetes mellitus

mitochondrial (e.g., *MTND1* and *MTND5*) and nuclear (e.g., *NDUFS2*, *NDUFV2*, and *NDUFA2*) genes encoding complex I subunits, and nuclear genes that encode complex I assembly factors (e.g., *ACAD9* and *NDUFAF1*) (34, 35). Complex II is entirely encoded by nDNA, and its deficiency has been reported in individuals with hypertrophic, dilated, and non-compaction cardiomyopathies who carried mutations in complex II subunits genes (*SDHA* and *SDHD*) (36, 37). Complex III deficiency can also cause cardiomyopathy that can either be isolated or accompanied with multi-organ involvement. Hypertrophic, dilated, and histiocytoid cardiomyopathies were reported in individuals with complex III deficiency and mutations in the *MTCYB* gene encoding cytochrome *b* (38–40). Dilated, hypertrophic, and histiocytoid cardiomyopathies have been reported in complex IV deficiencies associated with mutations in complex IV subunit genes (*COX6B1*,

MTCO2, and *MTCO3*) and complex IV assembly factors genes (*SURF1* and *SCO2*) (41, 42).

Mutations in several mitochondrial tRNA genes (e.g., *MTTK* causing MERRF syndrome and *MTTL1* causing MELAS syndrome) have been reported with multi-organ mitochondrial diseases or isolated cardiomyopathies. Cardiomyopathies associated with pathogenic variants in genes encoding mitochondrial tRNAs are usually hypertrophic, but can also be dilated or histiocytoid cardiomyopathy (29, 43). Hypertrophic cardiomyopathy has been reported with mutations in the mitochondrial 16S rRNA gene (*MTRNR2*) and restrictive cardiomyopathy with the m.1555A>G mutation in the mitochondrial 12S rRNA gene (*MTRNR1*) that is typically associated with aminoglycoside-induced hearing loss (44, 45). Mutations in genes coding mitochondrial ribosomal proteins (e.g., *MRPL3* and *MRPL44*) can cause hypertrophic cardiomyopathy accompanied by multi-organ disease (46, 47). Mutations in *TSFM*, encoding a mitochondrial translation elongation factor, can be associated with hypertrophic or dilated cardiomyopathy associated with multi-organ disease (48).

Mitochondrial neurogastrointestinal encephalopathy (MNGIE) syndrome is an mtDNA depletion syndrome caused by deficiency of thymidine phosphorylase, resulting in imbalances in mitochondrial nucleotide pools. Clinical features of MNGIE include progressive gastrointestinal dysmotility and cachexia, ptosis, ophthalmoplegia, hearing loss, demyelinating peripheral neuropathy, and leukoencephalopathy. Cardiac manifestations are usually asymptomatic ventricular hypertrophy and bundle branch block (49, 50).

Defects in CoQ_{10} biosynthesis result in primary CoQ_{10} deficiency which is a phenotypically and genetically heterogeneous condition with various clinical presentations, including encephalomyopathy, isolated myopathy, cerebellar ataxia, and nephrotic syndrome. Hypertrophic cardiomyopathy has been reported with mutations in genes involved in CoQ_{10} biosynthesis (*COQ2*, *COQ4*, and *COQ9*) (51, 52).

Barth syndrome is an X-linked disorder characterized by cardiomyopathy, skeletal myopathy, growth retardation, neutropenia, and increased urinary levels of 3-methylglutaconic acid. It is caused by mutations in the *TAZ* gene that codes for tafazzin, a phospholipid transacylase located in the inner mitochondrial membrane and plays an important role in the remodeling of cardiolipin. Cardiomyopathies are commonly left ventricular non-compaction and dilated cardiomyopathies, whereas hypertrophic cardiomyopathy appears to be less common. Other cardiac manifestations of Barth syndrome are arrhythmia (including supraventricular and ventricular tachycardia) and sudden death (53, 54).

Barth syndrome is one of a small group of disorders characterized by 3-methylglutaconic aciduria as a discriminative feature, where excretion of 3-methylglutaconic acid is significant and consistent. Other disorders in this group that might be associated with cardiomyopathy are caused by mutations in *DNAJC19*, *TMEM70*, and *AGK* (55). 3-Methylglutaconic aciduria associated with *DNAJC19* mutations (dilated cardiomyopathy and ataxia syndrome), results from deficient mitochondrial protein import and is characterized by dilated cardiomyopathy or left ventricular non-compaction, non-progressive cerebellar ataxia,

testicular dysgenesis, and growth failure (56). Mutations in *TMEM70* (mitochondrial complex V deficiency), encoding a protein involved in the insertion of ATP synthase (complex V) into the mitochondrial membrane, result in multi-organ mitochondrial disease with hypertrophic cardiomyopathy (57). Sengers syndrome, caused by mutations in *AGK*, might also be accompanied by 3-methylglutaconic aciduria and is characterized by hypertrophic cardiomyopathy, cataracts, myopathy, exercise intolerance, and lactic acidosis. The *AGK* gene product is an acylglycerol kinase and is involved in the assembly of ANT1, a mitochondrial adenine nucleotide transporter (58).

Friedreich ataxia is an autosomal recessive neurodegenerative disorder caused by mutations of *FXN*, which encodes frataxin, a mitochondrial iron-binding protein involved in the synthesis of the Fe–S clusters required by the ETC complexes. The clinical presentation includes progressive ataxia after the teenage years, dysarthria, loss of lower limb reflexes, peripheral sensory neuropathy, and diabetes mellitus. The cardiac manifestations include hypertrophic cardiomyopathy (59, 60).

DIAGNOSIS AND MANAGEMENT OF MITOCHONDRIAL CARDIOMYOPATHIES

The diagnosis of mitochondrial diseases is based on clinical recognition, biochemical screening, histopathological studies, functional assays, and molecular genetic testing. Due to the multi-organ involvement in the majority of mitochondrial diseases, evaluation of these diseases should include a systematic screening for all the targeted organs, e.g., neuroimaging, hearing assessment, ophthalmologic examination, liver function test, and serum creatinine phosphokinase (2). Biochemical screening tests for mitochondrial disorders include the determination of plasma lactate, blood glucose, urine organic acids, and plasma amino acids. Although lactic acidemia is a common biochemical feature of many mitochondrial disorders, it is neither specific nor sensitive (61). Hypoglycemia can be seen in children with mitochondrial diseases and urine organic acid analysis can show non-specific findings, including elevated lactate, ketone bodies, and TCA intermediates. A plasma amino acid profile may show elevation in plasma alanine level which reflects lactic acidemia and branched-chain amino acids which are catabolized in mitochondria (18).

Analysis of a fresh skeletal muscle biopsy is considered the gold standard in the diagnosis of mitochondrial disorders. The histology of affected muscles typically shows ragged-red fibers, which can be demonstrated using the modified Gomori trichrome stains, and contains peripheral and intermyofibrillar accumulation of abnormal mitochondria. Examining the muscle under an electron microscopy can demonstrate mitochondrial proliferation and abnormal mitochondrial morphology in mitochondrial myopathies. Histochemical staining for different ETC complexes can be used to estimate the severity and heterogeneity of ETC complexes deficiencies in the muscle tissue (3). Mitochondrial function can be assessed by measuring the enzymatic activity of different ETC complexes using a spectrophotometric methodology that utilizes specific electron acceptors and donors.

This assessment is usually carried out on skeletal muscle, skin fibroblast, or liver tissue (11). Mitochondrial function can also be assessed using the extracellular flux analyzer, Seahorse instrument, which can simultaneously measure mitochondrial respiration and glycolysis (62). Cardiac muscle biopsy is more invasive and can be performed in a patient with rapid disease progression or when biochemical testing in fibroblasts and skeletal muscle and molecular testing have not led to a conclusive diagnosis (29).

Molecular testing includes assessment of mtDNA content and DNA sequencing. Increased mtDNA content suggests a compensatory mechanism due to deficient mitochondrial function, whereas reduced mtDNA content implies defects in mtDNA biosynthesis, leading to mtDNA depletion. Measurement of mtDNA copy number is performed by real-time quantitative polymerase chain reaction using a mtDNA probe and a unique nuclear gene reference (63). Variable DNA sequencing options are available. If the clinical features of a mitochondrial disease are consistent with a recognizable syndrome, the mtDNA or nDNA gene known to be responsible for that syndrome can be tested to confirm the diagnosis. If a maternally inherited mitochondrial disease is suspected, the whole mtDNA can be sequenced. When genetically heterogeneous nDNA gene-related mitochondrial disease (e.g., mtDNA depletion syndromes) panel tests that include the known genes associated with such disease can be helpful. Next-generation massively parallel sequencing, which allows simultaneous sequencing of multiple genes at high coverage and low cost, has been widely used method for these gene panels. When the clinical picture is not consistent with a disease related to a specific gene or group of genes, a more extensive panel that includes all the known nDNA-related mitochondrial genes or whole exome or genome sequencing methodology can be considered (64, 65).

Currently, there are no satisfactory therapies available for mitochondrial disorders. Treatment remains largely symptomatic and does not significantly alter the course of the disease. Several cofactor supplementations have been tried with limited data supporting their benefits for most of them (6). So far, the only mitochondrial cardiomyopathies with an effective and specific metabolic treatment are those caused by CoQ_{10} deficiency. CoQ_{10} (ubiquinone) supplementation for patients with CoQ_{10} deficiency results in restoring the electron flow and a dramatic improvement in clinical manifestations associated with CoQ_{10} deficiency (66).

Heart transplantation was reported to be performed in 14% of patients with Barth syndrome (53). With respect to other mitochondrial diseases, although multi-organ diseases are considered a relative contraindication for solid organ transplantation, heart transplantation might be successful when clinical expression is limited to the myocardium or manifestations outside the heart are mild and appear non-progressive (29).

Ongoing clinical trials for potential treatment of mitochondrial diseases include the use of Bendavia, a mitochondrial permeability transition pore inhibitor, RTA 408, a potent activator of Nrf2 which is a regulator of cellular resistance to oxidants, and cysteamine bitartrate, an antioxidant (67) (http://Clinicaltrials.gov).

CONCLUSION

Hypertrophic, dilated, non-compaction, and histiocytoid cardiomyopathies can be the only feature or part of multi-organ mitochondrial diseases. Cardiomyopathies occur in approximately one-third of children with mitochondrial diseases and increase the mortality in these children. Therefore, screening for cardiomyopathy is a standard part of the management of individuals with known or suspected mitochondrial disease. Diagnosing mitochondrial diseases remains challenging in many cases and treatment remains largely symptomatic, as there are no satisfactory therapies available that significantly alter the course of the disease. Therefore, a lot of work is still need to be done to facilitate early diagnosis through discovering new disease biomarkers and novels genes involved in mitochondrial function and to find new treatment strategies that can restore the mitochondrial function.

AUTHOR CONTRIBUTIONS

Dr. AE-H has written the initial draft. Dr. FS has reviewed and modified the draft.

REFERENCES

1. Gilbert-Barness E. Review: metabolic cardiomyopathy and conduction system defects in children. *Ann Clin Lab Sci* (2004) 34:15–34.
2. Chinnery PF. Mitochondrial disorders overview. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJ, et al., editors. *GeneReviews(®)*. Seattle, WA: University of Washington Seattle (2014).
3. Munnich A, Rötig A, Rio M. Defects of the respiratory chain. *Inborn Metabolic Diseases Diagnosis and Treatment*. Berlin: Springer (2012). p. 223–38.
4. Finsterer J, Kothari S. Cardiac manifestations of primary mitochondrial disorders. *Int J Cardiol* (2014) 177:754–63. doi:10.1016/j.ijcard.2014.11.014
5. Scaglia F, Towbin JA, Craigen WJ, Belmont JW, Smith EO, Neish SR, et al. Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics* (2004) 114:925–31. doi:10.1542/peds.2004-0718
6. El-Hattab AW, Scaglia F. Mitochondrial disorders. In: Lee B, Scaglia F, editors. *Inborn Errors of Metabolism: From Neonatal Screening to Metabolic Pathways*. New York, NY: Oxford University Press (2015). p. 180–202.
7. Ferramosca A, Zara V. Biogenesis of mitochondrial carrier proteins: molecular mechanisms of import into mitochondria. *Biochim Biophys Acta* (2013) 1833:494–502. doi:10.1016/j.bbamcr.2012.11.014
8. Young MJ, Copeland WC. Human mitochondrial DNA replication machinery and disease. *Curr Opin Genet Dev* (2016) 38:52–62. doi:10.1016/j.gde.2016.03.005
9. Hällberg BM, Larsson N-G. Making proteins in the powerhouse. *Cell Metab* (2014) 20:226–40. doi:10.1016/j.cmet.2014.07.001
10. MacKenzie JA, Payne RM. Mitochondrial protein import and human health and disease. *Biochim Biophys Acta* (2007) 1772:509–23. doi:10.1016/j.bbadis.2006.12.002
11. Thorburn DR, Sugiana C, Salemi R, Kirby DM, Worgan L, Ohtake A, et al. Biochemical and molecular diagnosis of mitochondrial respiratory chain disorders. *Biochim Biophys Acta* (2004) 1659:121–8. doi:10.1016/j.bbabio.2004.08.006
12. Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders – past, present and future. *Biochim Biophys Acta* (2004) 1659:115–20. doi:10.1016/j.bbabio.2004.09.005
13. Chinnery PF, Hudson G. Mitochondrial genetics. *Br Med Bull* (2013) 106:135–59. doi:10.1093/bmb/ldt017
14. Coenen MJH, Antonicka H, Ugalde C, Sasarman F, Rossi R, Heister JGAMA, et al. Mutant mitochondrial elongation factor G1 and combined oxidative phosphorylation deficiency. *N Engl J Med* (2004) 351:2080–6. doi:10.1056/NEJMoa041878
15. Potgieter M, Pretorius E, Pepper MS. Primary and secondary coenzyme Q10 deficiency: the role of therapeutic supplementation. *Nutr Rev* (2013) 71:180–8. doi:10.1111/nure.12011
16. El-Hattab AW, Scaglia F. Mitochondrial DNA depletion syndromes: review and updates of genetic basis, manifestations, and therapeutic options. *Neurotherapeutics* (2013) 10:186–98. doi:10.1007/s13311-013-0177-6
17. Rouault TA. Biogenesis of iron-sulfur clusters in mammalian cells: new insights and relevance to human disease. *Dis Model Mech* (2012) 5:155–64. doi:10.1242/dmm.009019
18. El-Hattab AW, Scaglia F. Mitochondrial cytopathies. *Cell Calcium* (2016). doi:10.1016/j.ceca.2016.03.003
19. Balaban RS, Nemoto S, Finkel T. Mitochondria, oxidants, and aging. *Cell* (2005) 120:483–95. doi:10.1016/j.cell.2005.02.001
20. Hayashi G, Cortopassi G. Oxidative stress in inherited mitochondrial diseases. *Free Radic Biol Med* (2015) 88:10–7. doi:10.1016/j.freeradbiomed.2015.05.039
21. Giorgi C, Agnoletto C, Bononi A, Bonora M, De Marchi E, Marchi S, et al. Mitochondrial calcium homeostasis as potential target for mitochondrial medicine. *Mitochondrion* (2012) 12:77–85. doi:10.1016/j.mito.2011.07.004
22. Willems PH, Valsecchi F, Distelmaier F, Verkaart S, Visch H-J, Smeitink JAM, et al. Mitochondrial Ca2+ homeostasis in human NADH:ubiquinone oxidoreductase deficiency. *Cell Calcium* (2008) 44:123–33. doi:10.1016/j.ceca.2008.01.002
23. Galluzzi L, Kepp O, Trojel-Hansen C, Kroemer G. Mitochondrial control of cellular life, stress, and death. *Circ Res* (2012) 111:1198–207. doi:10.1161/CIRCRESAHA.112.268946
24. El-Hattab AW, Emrick LT, Craigen WJ, Scaglia F. Citrulline and arginine utility in treating nitric oxide deficiency in mitochondrial disorders. *Mol Genet Metab* (2012) 107:247–52. doi:10.1016/j.ymgme.2012.06.018
25. Ballana E, Govea N, de Cid R, Garcia C, Arribas C, Rosell J, et al. Detection of unrecognized low-level mtDNA heteroplasmy may explain the variable phenotypic expressivity of apparently homoplasmic mtDNA mutations. *Hum Mutat* (2008) 29:248–57. doi:10.1002/humu.20639
26. Saneto RP, Sedensky MM. Mitochondrial disease in childhood: mtDNA encoded. *Neurotherapeutics* (2013) 10:199–211. doi:10.1007/s13311-012-0167-0
27. Meyers DE, Basha HI, Koenig MK. Mitochondrial cardiomyopathy: pathophysiology, diagnosis, and management. *Tex Heart Inst J* (2013) 40:385–94.
28. Holmgren D, Wåhlander H, Eriksson BO, Oldfors A, Holme E, Tulinius M. Cardiomyopathy in children with mitochondrial disease; clinical course and cardiological findings. *Eur Heart J* (2003) 24:280–8. doi:10.1016/S0195-668X(02)00387-1
29. Brunel-Guitton C, Levtova A, Sasarman F. Mitochondrial diseases and cardiomyopathies. *Can J Cardiol* (2015) 31:1360–76. doi:10.1016/j.cjca.2015.08.017
30. Okajima Y, Tanabe Y, Takayanagi M, Aotsuka H. A follow up study of myocardial involvement in patients with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS). *Heart* (1998) 80:292–5.
31. Bates MGD, Nesbitt V, Kirk R, He L, Blakely EL, Alston CL, et al. Mitochondrial respiratory chain disease in children undergoing cardiac transplantation: a prospective study. *Int J Cardiol* (2012) 155:305–6. doi:10.1016/j.ijcard.2011.11.063
32. Finsterer J. Cardiogenetics, neurogenetics, and pathogenetics of left ventricular hypertrabeculation/noncompaction. *Pediatr Cardiol* (2009) 30:659–81. doi:10.1007/s00246-008-9359-0
33. Finsterer J. Histiocytoid cardiomyopathy: a mitochondrial disorder. *Clin Cardiol* (2008) 31:225–7. doi:10.1002/clc.20224
34. Brecht M, Richardson M, Taranath A, Grist S, Thorburn D, Bratkovic D. Leigh syndrome caused by the MT-ND5 m.13513G>A mutation: a case presenting with WPW-like conduction defect, cardiomyopathy, hypertension and hyponatraemia. *JIMD Rep* (2015) 19:95–100. doi:10.1007/8904_2014_375
35. Fassone E, Rahman S. Complex I deficiency: clinical features, biochemistry and molecular genetics. *J Med Genet* (2012) 49:578–90. doi:10.1136/jmedgenet-2012-101159

36. Alston CL, Ceccatelli Berti C, Blakely EL, Oláhová M, He L, McMahon CJ, et al. A recessive homozygous p.Asp92Gly SDHD mutation causes prenatal cardiomyopathy and a severe mitochondrial complex II deficiency. *Hum Genet* (2015) 134:869–79. doi:10.1007/s00439-015-1568-z

37. Jain-Ghai S, Cameron JM, Al Maawali A, Blaser S, MacKay N, Robinson B, et al. Complex II deficiency – a case report and review of the literature. *Am J Med Genet A* (2013) 161A:285–94. doi:10.1002/ajmg.a.35714

38. Andreu AL, Checcarelli N, Iwata S, Shanske S, DiMauro S. A missense mutation in the mitochondrial cytochrome b gene in a revisited case with histiocytoid cardiomyopathy. *Pediatr Res* (2000) 48:311–4. doi:10.1203/00006450-200009000-00008

39. Carossa V, Ghelli A, Tropeano CV, Valentino ML, Iommarini L, Maresca A, et al. A novel in-frame 18-bp microdeletion in MT-CYB causes a multisystem disorder with prominent exercise intolerance. *Hum Mutat* (2014) 35:954–8. doi:10.1002/humu.22596

40. Hagen CM, Aidt FH, Havndrup O, Hedley PL, Jespersgaard C, Jensen M, et al. MT-CYB mutations in hypertrophic cardiomyopathy. *Mol Genet Genomic Med* (2013) 1:54–65. doi:10.1002/mgg3.5

41. Abdulhag UN, Soiferman D, Schueler-Furman O, Miller C, Shaag A, Elpeleg O, et al. Mitochondrial complex IV deficiency, caused by mutated COX6B1, is associated with encephalomyopathy, hydrocephalus and cardiomyopathy. *Eur J Hum Genet* (2015) 23:159–64. doi:10.1038/ejhg.2014.85

42. Marin-Garcia J, Goldenthal MJ, Ananthakrishnan R, Pierpont ME. The complete sequence of mtDNA genes in idiopathic dilated cardiomyopathy shows novel missense and tRNA mutations. *J Card Fail* (2000) 6:321–9. doi:10.1054/jcaf.2000.19232

43. Bannwarth S, Procaccio V, Lebre AS, Jardel C, Chaussenot A, Hoarau C, et al. Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *J Med Genet* (2013) 50:704–14. doi:10.1136/jmedgenet-2013-101604

44. Liu Z, Song Y, Li D, He X, Li S, Wu B, et al. The novel mitochondrial 16S rRNA 2336T>C mutation is associated with hypertrophic cardiomyopathy. *J Med Genet* (2014) 51:176–84. doi:10.1136/jmedgenet-2013-101818

45. Santorelli FM, Tanji K, Manta P, Casali C, Krishna S, Hays AP, et al. Maternally inherited cardiomyopathy: an atypical presentation of the mtDNA 12S rRNA gene A1555G mutation. *Am J Hum Genet* (1999) 64:295–300. doi:10.1086/302188

46. Distelmaier F, Haack TB, Catarino CB, Gallenmüller C, Rodenburg RJ, Strom TM, et al. MRPL44 mutations cause a slowly progressive multisystem disease with childhood-onset hypertrophic cardiomyopathy. *Neurogenetics* (2015) 16:319–23. doi:10.1007/s10048-015-0444-2

47. Galmiche L, Serre V, Beinat M, Assouline Z, Lebre A-S, Chretien D, et al. Exome sequencing identifies MRPL3 mutation in mitochondrial cardiomyopathy. *Hum Mutat* (2011) 32:1225–31. doi:10.1002/humu.21562

48. Ahola S, Isohanni P, Euro L, Brilhante V, Palotie A, Pihko H, et al. Mitochondrial EFTs defects in juvenile-onset Leigh disease, ataxia, neuropathy, and optic atrophy. *Neurology* (2014) 83:743–51. doi:10.1212/WNL.0000000000000716

49. Hirano M. Mitochondrial neurogastrointestinal encephalopathy disease. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJ, et al., editors. *GeneReviews(®)*. Seattle, WA: University of Washington Seattle (2016).

50. Nishino I, Spinazzola A, Papadimitriou A, Hammans S, Steiner I, Hahn CD, et al. Mitochondrial neurogastrointestinal encephalomyopathy: an autosomal recessive disorder due to thymidine phosphorylase mutations. *Ann Neurol* (2000) 47:792–800. doi:10.1002/1531-8249(200006)47:6<792::AID-ANA12>3.3.CO;2-P

51. Desbats MA, Lunardi G, Doimo M, Trevisson E, Salviati L. Genetic bases and clinical manifestations of coenzyme Q10 (CoQ 10) deficiency. *J Inherit Metab Dis* (2015) 38:145–56. doi:10.1007/s10545-014-9749-9

52. Doimo M, Desbats MA, Cerqua C, Cassina M, Trevisson E, Salviati L. Genetics of coenzyme q10 deficiency. *Mol Syndromol* (2014) 5:156–62. doi:10.1159/000362826

53. Clarke SLN, Bowron A, Gonzalez IL, Groves SJ, Newbury-Ecob R, Clayton N, et al. Barth syndrome. *Orphanet J Rare Dis* (2013) 8:23. doi:10.1186/1750-1172-8-23

54. Jefferies JL. Barth syndrome. *Am J Med Genet C Semin Med Genet* (2013) 163C:198–205. doi:10.1002/ajmg.c.31372

55. Wortmann SB, Duran M, Anikster Y, Barth PG, Sperl W, Zschocke J, et al. Inborn errors of metabolism with 3-methylglutaconic aciduria as discriminative feature: proper classification and nomenclature. *J Inherit Metab Dis* (2013) 36:923–8. doi:10.1007/s10545-012-9580-0

56. Ojala T, Polinati P, Manninen T, Hiippala A, Rajantie J, Karikoski R, et al. New mutation of mitochondrial DNAJC19 causing dilated and noncompaction cardiomyopathy, anemia, ataxia, and male genital anomalies. *Pediatr Res* (2012) 72:432–7. doi:10.1038/pr.2012.92

57. Cízková A, Stránecký V, Mayr JA, Tesarová M, Havlícková V, Paul J, et al. TMEM70 mutations cause isolated ATP synthase deficiency and neonatal mitochondrial encephalocardiomyopathy. *Nat Genet* (2008) 40:1288–90. doi:10.1038/ng.246

58. Haghighi A, Haack TB, Atiq M, Mottaghi H, Haghighi-Kakhki H, Bashir RA, et al. Sengers syndrome: six novel AGK mutations in seven new families and review of the phenotypic and mutational spectrum of 29 patients. *Orphanet J Rare Dis* (2014) 9:119. doi:10.1186/s13023-014-0119-3

59. Bidichandani SI, Delatycki MB. Friedreich ataxia. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJ, et al., editors. *GeneReviews(®)*. Seattle, WA: University of Washington Seattle (2014).

60. Weidemann F, Störk S, Liu D, Hu K, Herrmann S, Ertl G, et al. Cardiomyopathy of Friedreich ataxia. *J Neurochem* (2013) 126(Suppl 1):88–93. doi:10.1111/jnc.12217

61. Mitochondrial Medicine Society's Committee on Diagnosis, Haas RH, Parikh S, Falk MJ, Saneto RP, Wolf NI, et al. The in-depth evaluation of suspected mitochondrial disease. *Mol Genet Metab* (2008) 94:16–37. doi:10.1016/j.ymgme.2007.11.018

62. Bonnen PE, Yarham JW, Besse A, Wu P, Faqeih EA, Al-Asmari AM, et al. Mutations in FBXL4 cause mitochondrial encephalopathy and a disorder of mitochondrial DNA maintenance. *Am J Hum Genet* (2013) 93:471–81. doi:10.1016/j.ajhg.2013.07.017

63. Bai R-K, Wong L-JC. Simultaneous detection and quantification of mitochondrial DNA deletion(s), depletion, and over-replication in patients with mitochondrial disease. *J Mol Diagn* (2005) 7:613–22. doi:10.1016/S1525-1578(10)60595-8

64. Legati A, Reyes A, Nasca A, Invernizzi F, Lamantea E, Tiranti V, et al. New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. *Biochim Biophys Acta* (2016) 1857:1326–35. doi:10.1016/j.bbabio.2016.02.022

65. Metzker ML. Sequencing technologies – the next generation. *Nat Rev Genet* (2010) 11:31–46. doi:10.1038/nrg2626

66. Di Giovanni S, Mirabella M, Spinazzola A, Crociani P, Silvestri G, Broccolini A, et al. Coenzyme Q10 reverses pathological phenotype and reduces apoptosis in familial CoQ10 deficiency. *Neurology* (2001) 57:515–8. doi:10.1212/WNL.57.3.515

67. Viscomi C, Bottani E, Zeviani M. Emerging concepts in the therapy of mitochondrial disease. *Biochim Biophys Acta* (2015) 1847:544–57. doi:10.1016/j.bbabio.2015.03.001

Molecular and Genetic Insights into Thoracic Aortic Dilation in Conotruncal Heart Defects

W. Aaron Kay*

Department of Medicine, Krannert Institute of Cardiology, Indiana University School of Medicine, Indianapolis, IN, USA

Thoracic aortic dilation (AD) has commonly been described in conotruncal defects (CTDs), such as tetralogy of Fallot, double outlet right ventricle and transposition of the great arteries, and truncus arteriosus. Several theories for this have been devised, but fairly recent data indicate that there is likely an underlying histologic abnormality, similar to that seen in Marfan and other connective tissue disease. The majority of aortic dissection in the general population occurs after the age of 45 years, and there have been very few case reports of aortic dissection in CTD. Given advances in cardiac surgery and increasing survival over the past several decades, there has been rising concern that, as patients who have survived surgical correction of these defects age, there may be increased morbidity and mortality due to aortic dissection and aortic regurgitation. This review discusses the most recent developments in research into AD in CTD, including associated genetic mutations.

Keywords: conotruncal, tetralogy of Fallot, transposition of great arteries, aortic aneurysm, aortic dissection, truncus arteriosus

INTRODUCTION

Conotruncal cardiac defects (CTDs) include a variety of congenital heart defects, such as tetralogy of Fallot (TOF), truncus arteriosus (TA), double outlet right ventricle (DORV), and transposition of the great arteries (TGA). These defects represent 5–10% of congenital heart disease and, generally, lead to severe cyanosis, necessitating repair in the newborn period or early in infancy. A common observation in CTD is thoracic aortic dilation (AD). It has been known for over half a century that AD is common in TOF (1, 2). Landmark work by Niwa demonstrated that the incidence of AD in adults with repaired tetralogy approaches 15% (3). Progressive dilation of the neo-aortic root is out of proportion to somatic growth in TGA after arterial switch surgery (4), and AD is found in the majority of patients with TA who survive initial repair (5).

Aortic dilation can become more clinically relevant if it leads to significant aortic valve regurgitation, aortic dissection, or worsened LV–aorta interaction (6). AD can lead to significant morbidity and mortality, with the chief worry being ascending aortic or aortic root dissection, which is often fatal without emergency surgery, and as a result, clinicians seek to evaluate patients to avoid this complication, generally by assessing aortic size on non-invasive imaging, and intervening with elective surgery in a more controlled setting. Specific guidelines exist for elective surgery to prevent dissection in Marfan syndrome and certain other connective tissue disease, but specific guidelines

Abbreviations: AD, aortic dilation; CTD, conotruncal defects; DORV, double outlet right ventricle; FBN1, fibrillin-1 gene; TGA, transposition of the great arteries; TOF, tetralogy of Fallot.

for elective surgery for AD in CTD have not been established (7). One interesting article over several years evaluated the outcomes of 81 adults with CTD and AD who had surgery over several decades and came to the conclusion that elective surgery in CTD should be delayed unless the maximum aortic dimension is at least 5.5 cm, unless there is documented rapid growth of the ascending aorta, or a worrisome family history of aortic dissection or aneurysm (8).

In the past, it was thought that the risk of aortic dissection in these patients was low, perhaps due to the low incidence of systemic hypertension, smoking, and other traditional vascular risk factors in this population (9). However, there have now been four case reports of dissection in TOF (10–13). All these cases except the most recent one dissected at diameters >6 cm. Perhaps, aortic cannulation during the initial repair could be blamed for a dissection late after repair, but this seems unlikely given that the dissections occurred anywhere from 9 years to more than 30 years after the initial repair (10–13). It would stand to reason that, if direct injury to the aorta during initial cannulation was the culprit, there would be a significant number of dissections reported in childhood in the literature, when, in fact, the youngest patient identified in the literature was already 18 years of age, more than 17 years after initial compete TOF repair (12).

Regarding TGA, fortunately, no case reports of aortic dissection in TGA after arterial switch have been published, but perhaps none have occurred due to the relative young age of this population, given that most survivors of the arterial switch procedure are <30 years old. To date, there has only been one case report of an aortic dissection in TGA after Mustard repair (14). In this case, the patient had been lost to cardiology follow-up for over two decades, had several pregnancies, and smoked cigarettes, and it was unknown what size the aorta was prior to dissection. Isolated cases of dissection in CTD have also been found in review of administrative databases (15). There has been only one case report of elective aortic root replacement in a TGA patient after Mustard procedure – which was performed in a 30-year-old man, for an aortic aneurysm, measuring 4.5 cm (16). Still, the majority of patients with dissection in CTD have been older than 45 years of age. However, as the population of CTD survivors' ages, more patients might be at risk for aortic dissection or other complication.

It is important to note that aortic dissection is not the only concern or cause of morbidity due to AD in patients with CTD. The dilation itself can lead to significant aortic valve regurgitation, which can lead to an increased pulsatile load on the left ventricle (or systemic right ventricle), leading directly to decreased cardiac output, or indirectly *via* decreasing coronary arterial blood flow (17, 18). Aortic regurgitation may additionally worsen not just due to dilation but also due to stiffness of the aortic root (19). The need for aortic valve replacement is fortunately fairly uncommon in CTD, although the presence of aortic regurgitation is fairly common in TGA after arterial switch (3, 20). After arterial switch surgery, Losay and others noted that freedom from aortic regurgitation was 78% at 10 years and 69% at 15 years; however, freedom from aortic valve replacement was 98% at 10 years and 97% at 15 years (21). In a study by Marino and others, severe neo-aortic valve regurgitation was present in 3.7% and trivial to mild regurgitation in 81% of patients at mid-term follow-up (22).

POSSIBLE MECHANISMS FOR AORTIC ROOT DILATION IN CONOTRUNCAL DEFECTS

There are a few hypotheses for why AD in CTD occurs, independent of standard risk factors, such as hypertension, aging, pregnancy, and smoking. The first is that AD occurs due to hemodynamic stress on the aorta from a chronic right-left-shunt. Evidence to support this hypothesis include data showing that AD is worse with worsened degrees of right ventricular outflow tract stenosis and is worse in patients with pulmonary atresia than in patients with pulmonary stenosis (23). A second hypothesis is that volume loading of the aorta, *via* a surgical systemic-to-pulmonary shunt, will increase flow through the aortic valve, thus leading to dilation of the proximal aorta *via* increased wall stress (24). A longer duration between shunting and complete repair has been found to correlate with AD in repaired TOF (23, 25). One study showed a 12% increase in mean aortic diameter after surgical aortopulmonary shunting (25). Other observations that have been associated with larger aortic dimensions in TOF have included a right rather than left aortic arch and male sex.

Newer data suggest that CTD may be associated with a primary problem with aortic histology, i.e., a true aortopathy. Evidence of aortopathy in TOF has been found early in life, on fetal echocardiography (26), and also on histologic studies. Even in infants, higher histologic grading scores in TOF patients have been seen, thus making it likely that there is an intrinsic abnormality of the aortic tissue leading to dilation, rather than long-term hemodynamic stress (20). Histologic abnormalities were reported in a cohort of 15 repaired TOF patients with ascending aortic aneurysms, and this was further supported in another study demonstrating elastin fragmentation in the ascending aorta in 74.5% of 98 consecutive patients undergoing complete TOF repair (27, 28).

There are many variables that affect the size of the aortic root and ascending aorta in general. Although there have been similarities of aortic root histology seen between Marfan syndrome and some CTDs, it is notable that the risk of dissection in Marfan is significantly higher, which begs the question why the Marfan phenotype is so much more dangerous. It is possible that the histologic abnormality in the aorta found in TOF may be less severe than the abnormality found in Marfan syndrome (27).

Most research regarding AD in CTD predominantly focused on TOF with fewer papers focused on TGA or truncus. DORV is rarely considered by itself in the literature, as it is a diagnosis encompassing a broad spectrum of pathophysiology, depending on where the ventricular septal defect is in an individual patient and also on the relationship of the great arteries to one another. It most commonly presents with TOF-like physiology, wherein the patient has a subaortic ventricular septal defect with pulmonary stenosis. In most of the literature, DORV with this physiology is included as a TOF variant.

Conclusions have been difficult to draw, given differing definitions of what constitutes an aortic aneurysm. Most pediatric centers have used Z-scores, whereas most adult studies have looked at either absolute dimensions or dimensions indexed to either body surface area or height. A common definition for "aneurysm" is an observed to expected ratio of >1.5 of the normal population at a specific aortic segment. A large study using cardiac MRI evaluated normal values in a control population and can be used as a helpful baseline (29). Very few studies have performed longitudinal measurements to demonstrate if the AD found in CTD is likely to progress over time. In one study of children with both TOF and TGA, independent predictors of a longitudinal increase in Z-scores of the ascending aorta included male sex and presence of aortic regurgitation (30). A study of adults with repaired tetralogy, utilizing MRI as the imaging technique, showed minimal growth in TOF over the course of 3 years, with a very small number of aortas increasing in size from below to above a threshold value of 5 cm (31).

As survival of CTD patients has improved, numerous patients with CTD have been able to have pregnancies of their own. Pregnancy is known to be an independent variable for the structural change of the aortic media, but it is unknown how likely these changes are to regress after delivery and whether the changes that occur with gestation are additive to the normal aging process in this population (27).

GENETIC AND MOLECULAR FINDINGS

The investigation of AD in CTD from a molecular and genetic standpoint, compared to the robust database found in Marfan syndrome, is still in its infancy. There have been very few articles investigating the genetic or histologic associations of AD and CTD, to date. Marfan syndrome, with a prevalence of 1:5000, is in most cases due to a mutation in the fibrillin-1 gene (*FBN1*) (32). Fibrillin-1, together with other proteins of the extracellular matrix, forms thread-like microfibrils, which create structural support and elasticity to tissues. Mutations in *FBN1* lead to breakdown of microfibril architecture, which can lead to aortic aneurysms and other complications. There are numerous mutations in the *FBN1* mutation database, with nearly 3000 reported mutations to date, varying from point mutations to large rearrangements.

Research demonstrating that the histology of the aorta in patients with congenital heart disease is similar to that of Marfan patients (27) has led to small studies investigating the role of fibrillin in CTD. Given the much lower incidence of dissection in TOF, it is possible that the histologic abnormality in the aorta found in TOF may be less severe than the abnormality found in Marfan syndrome (27). In a study of 74 consecutive patients undergoing intracardiac repair or TOF, full-thickness aortic wall biopsies were performed, and there was a 50.9% prevalence of *FBN1* gene polymorphisms or mutations in those with a dilated aorta (28). Additionally, the risk of AD was found to be eight times higher in patients with these variants. Abnormal histology, defined as a lamellar count <60, was associated with a risk of AD 15.97 times higher than in normal controls.

The DiGeorge or velocardial facial syndrome, due to a 22q12 deletion, is commonly associated with CTD, and 22q11.2 mutation

has been found to be associated with larger aortic root size in TOF (33). Patients with 22q11.2 mutation are more likely to have right aortic arch and pulmonary atresia than non-syndromic patients, so it is unclear, in TOF, if the larger aortic size is due directly to the genetic mutation or rather due to a change in hemodynamics. However, one paper noted that the 22q11.2 deletion itself, even in the absence of CTD, appears to be associated with AD, where AD was noted to have an incidence of 10.8% (34).

Linkage analysis has been used to find novel gene mutations that appear to correlate with TOF and other CTD, but, to date, no studies have been performed to evaluate for novel mutations that explicitly explain AD in CTD.

It is possible that some of the patients in the literature who had CTD and aortic dissection may have had undiagnosed Marfan syndrome or other known connective tissue disease, as the lack of diagnosis may have been retrospectively made on phenotypic, rather than genotypic, grounds. There are numerous variables, including the underlying tissue strength, and varying changes in physiology that frustrate the ability to tease out.

FUTURE DIRECTIONS

Further research into AD in CTD will be much more likely to proceed if more patients are found to suffer aortic complications. Next generation sequencing, such as whole exome sequencing, may be very helpful at identifying novel gene mutations that could be responsible for AD in CTD. Genome-wide linkage analysis and exome sequencing together, recently, led to the discovery of a novel *TGFB3* mutation as a cause of syndromic aortic aneurysm and aortic dissection in series of 470 index cases with thoracic aortic aneurysms who had been screened for all known gene mutations associated with thoracic aortic aneurysms (35). Current standard genetic panels to test for aortopathy genes only include 20–25 gene mutations, but these panels will expand greatly as new candidate gene mutations are discovered. *In silico* analysis and more advanced informatics technology will greatly facilitate the ability to translate these findings to clinical practice.

SUMMARY

The American Heart Association (AHA) and American College of Cardiology (ACC) provide guidelines on the management of thoracic aortic diseases (7), but the most recent guidelines do not provide a clear management decision for how to manage AD in CTD patients. It is exceedingly rare for an aortic dissection to occur in childhood, other than in infancy, due to very severe genetic problems or iatrogenic causes; thus, aortic dissection is largely considered an adult-onset problem. The reader is directed to the current adult congenital heart disease clinical management guidelines (36), which are due for an update in the near future.

The decision of when to intervene for AD in CTD must be weighed on a number of factors, including the number of prior cardiac interventions (and thus, likelihood of morbidity of an elective procedure), the rate of growth of the aorta over time, other lesions that require operative management, and perhaps

most importantly, a genetic risk profile. For patients with a strong family history of thoracic aortic aneurysm, aortic dissection, or a known genetic mutation likely to lead to aortic dissection, perhaps, a lower threshold for intervention should be used.

Over time, we may discover advantages to elective aortic root replacement to improve LV–aorta coupling, even in patients not thought to be at acute risk of aortic dissection, as surgical techniques improve. Perhaps, newer discoveries will lead to new therapies that prevent, or even reverse, aortopathies.

Ultimately, only time will tell what the true risk for aortic complications in CTD is, and if the incidence grows over time as this population ages, more research will help determine who is at most risk. It is difficult to determine a risk profile when so few patients have had aortic complications. Ideally, in this era of

personalized medicine and high throughput genetic sequencing, every patient will have a unique genetic signature that can be used to tailor his or her unique risk.

AUTHOR CONTRIBUTIONS

WK wrote this entire article himself, including reviewing appropriate background literature.

FUNDING

REFERENCES

1. Bahnson H, Spencer F, Landtman B, Wolf M, Neill C, Taussig H. Surgical treatment and follow-up of 147 cases of tetralogy of Fallot treated by correction. *J Thorac Cardiovasc Surg* (1962) 1962(44):419–32.

2. Rieker R, Berman M, Stansel H. Postoperative studies in patients with tetralogy of Fallot. *Ann Thorac Surg* (1975) 19(1):17–26. doi:10.1016/S0003-4975(10)65727-X

3. Niwa K. Aortic root dilatation in tetralogy of Fallot long-term after repair – histology of the aorta in tetralogy of Fallot: evidence of intrinsic aortopathy. *Int J Cardiol* (2005) 103(2):117–9. doi:10.1016/j.ijcard.2004.07.002

4. Schwartz M, Gauvreau K, del Nido P, Mayer JE, Colan SD. Long-term predictors of aortic root dilation and aortic regurgitation after arterial switch operation. *ACC Curr J Rev* (2004) 13(12):54–5. doi:10.1016/j.accreview.2004.11.021

5. Carlo W, McKenzie E, Slesnick T. Root dilation in patients with truncus arteriosus. *Congenit Heart Dis* (2011) 6(3):228–33. doi:10.1111/j.1747-0803.2011.00520.x

6. Niwa K. Aortopathy in congenital heart disease in adults: aortic dilatation with decreased aortic elasticity that impacts negatively on left ventricular function. *Korean Circ J* (2013) 43(4):215. doi:10.4070/kcj.2013.43.4.215

7. Hiratzka L, Bakris G, Beckman J, Bersin R, Carr V, Casey D, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with thoracic aortic disease. *J Am Coll Cardiol* (2010) 55(14):e27–129. doi:10.1016/j.jacc.2010.02.010

8. Stulak J, Dearani J, Burkhart H, Sundt T, Connolly H, Schaff H. Does the dilated ascending aorta in an adult with congenital heart disease require intervention? *J Thorac Cardiovasc Surg* (2010) 140(6):S52–7. doi:10.1016/j.jtcvs.2010.08.052

9. Dearani J, Burkhart H, Stulak J, Sundt T, Schaff H. Management of the aortic root in adult patients with conotruncal anomalies. *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu* (2009) 12(1):122–9. doi:10.1053/j.pcsu.2009.01.013

10. Rathi V, Doyle M, Williams R, Yamrozik J, Shannon R, Biederman R. Massive aortic aneurysm and dissection in repaired tetralogy of Fallot; diagnosis by cardiovascular magnetic resonance imaging. *Int J Cardiol* (2005) 101(1):169–70. doi:10.1016/j.ijcard.2004.05.037

11. Kim W, Seo J, Kim S, Song J, Lee J, Na C. Aortic dissection late after repair of tetralogy of Fallot. *Int J Cardiol* (2005) 101(3):515–6. doi:10.1016/j.ijcard.2004.03.026

12. Konstantinov I, Fricke T, d'Udekem Y, Robertson T. Aortic dissection and rupture in adolescents after tetralogy of Fallot repair. *J Thorac Cardiovasc Surg* (2010) 140(5):e71–3. doi:10.1016/j.jtcvs.2010.06.045

13. Wijesekera V, Kiess M, Grewal J, Chow R, Raju R, Leipsic J, et al. Aortic dissection in a patient with a dilated aortic root following tetralogy of Fallot repair. *Int J Cardiol* (2014) 174(3):833–4. doi:10.1016/j.ijcard.2014.04.167

14. Nowitz A. Acute ascending aortic dissection 41 years after mustard procedure. *J Cardiothorac Vasc Anesth* (2013) 27(4):735–9. doi:10.1053/j.jvca.2012.04.006

15. Frischhertz B, Shamszad P, Pedroza C, Milewicz D, Morris S. Thoracic aortic dissection and rupture in conotruncal cardiac defects: a population-based study. *Int J Cardiol* (2015) 184:521–7. doi:10.1016/j.ijcard.2015.03.061

16. Ono M, Goerler H, Boethig D, Westhoff-Bleck M, Hartung D, Breymann T. Valve-sparing operation for aortic root aneurysm late after mustard procedure. *Ann Thorac Surg* (2007) 83(6):2224–6. doi:10.1016/j.athoracsur.2006.12.043

17. Schaefer B, Lewin M, Stout K, Gill E, Prueitt A, Byers P, et al. The bicuspid aortic valve: an integrated phenotypic classification of leaflet morphology and aortic root shape. *Heart* (2008) 94(12):1634–8. doi:10.1136/hrt.2007.132092

18. Chowdhury U, Mishra A, Ray R, Kalaivani M, Reddy S, Venugopal P. Histopathologic changes in ascending aorta and risk factors related to histopathologic conditions and aortic dilatation in patients with tetralogy of Fallot. *J Thorac Cardiovasc Surg* (2008) 135(1):69–77.e11. doi:10.1016/j.jtcvs.2007.06.011

19. Murakami T, Nakazawa M, Momma K, Imai Y. Impaired distensibility of neo-aorta after arterial switch procedure. *Ann Thorac Surg* (2000) 70(6):1907–10. doi:10.1016/S0003-4975(00)01837-3

20. Tan J, Davlouros PA, McCarthy KP, Gatzoulis MA, Ho SY. Intrinsic histological abnormalities of aortic root and ascending aorta in tetralogy of Fallot: evidence of causative mechanism for aortic dilatation and aortopathy. *Circulation* (2005) 112(7):961–8. doi:10.1161/CIRCULATIONAHA.105.537928

21. Losay J, Touchot A, Capderou A, Piot J, Belli E, Planché C, et al. Aortic valve regurgitation after arterial switch operation for transposition of the great arteries. *J Am Coll Cardiol* (2006) 47(10):2057–62. doi:10.1016/j.jacc.2005.12.061

22. Marino B, Wernovsky G, McElhinney D, Jawad A, Kreb D, Mantel S, et al. Neo-aortic valvar function after the arterial switch. *Cardiol Young* (2006) 16(05):481. doi:10.1017/S1047951106000953

23. Niwa K, Siu SC, Webb GD, Gatzoulis MA. Progressive aortic root dilatation in adults late after repair of tetralogy of Fallot. *Circulation* (2002) 106(11):1374–8. doi:10.1161/01.CIR.0000028462.88597.AD

24. Capelli H, Ross D, Somerville J. Aortic regurgitation in tetrad of Fallot and pulmonary atresia. *Am J Cardiol* (1982) 49(8):1979–83. doi:10.1016/0002-9149(82)90218-1

25. Bhat A, Smith C, Hawker R. Late aortic root dilatation in tetralogy of Fallot may be prevented by early repair in infancy. *Pediatr Cardiol* (2004) 25(6):654–9. doi:10.1007/s00246-003-0665-2

26. DeVore G, Siassi B, Platt L. Fetal echocardiography. VIII. Aortic root dilatation: a marker for tetralogy of Fallot. *Am J Obstet Gynecol* (1988) 159:129–36. doi:10.1016/0002-9378(88)90508-X

27. Niwa K, Perloff J, Bhuta S, Laks H, Drinkwater D, Child J, et al. Structural abnormalities of great arterial walls in congenital heart disease: light and electron microscopic analyses. *Circulation* (2001) 103(3):393–400. doi:10.1161/01.CIR.103.3.393

28. Chowdhury U, Mishra A, Balakrishnan P, Sharma S, Kabra M, Ray R, et al. Role of fibrillin-1 genetic mutations and polymorphism in aortic dilatation in patients undergoing intracardiac repair of tetralogy of Fallot. *J Thorac Cardiovasc Surg* (2008) 136(3):757–66.e10. doi:10.1016/j.jtcvs.2007.12.044

29. Burman E, Keegan J, Kilner P. Aortic root measurement by cardiovascular magnetic resonance: specification of planes and lines of measurement and corresponding normal values. *Circ Cardiovasc Imaging* (2008) 1(2):104–13. doi:10.1161/CIRCIMAGING.108.768911

30. Trippel A, Pallivathukal S, Pfammatter J, Hutter D, Kadner A, Pavlovic M. Dimensions of the ascending aorta in conotruncal heart defects. *Pediatr Cardiol* (2014) 35(5):831–7. doi:10.1007/s00246-014-0862-1

31. Kay W, Cook S, Daniels C. Evaluation by MRA of aortic dilation late after repair of tetralogy of Fallot. *Int J Cardiol* (2013) 167(6):2922–7. doi:10.1016/j.ijcard.2012.07.015

32. Ramachandra C, Mehta A, Guo K, Wong P, Tan J, Shim W. Molecular pathogenesis of Marfan syndrome. *Int J Cardiol* (2015) 187:585–91. doi:10.1016/j.ijcard.2015.03.423

33. John A, Rychik J, Khan M, Yang W, Goldmuntz E. 22q11.2 deletion syndrome as a risk factor for aortic root dilation in tetralogy of Fallot. *Cardiol Young* (2013) 24(02):303–10. doi:10.1017/S1047951113000309

34. John A, McDonald-McGinn D, Zackai E, Goldmuntz E. Aortic root dilation in patients with 22q11.2 deletion syndrome. *Am J Med Genet* (2009) 149A(5):939–42. doi:10.1002/ajmg.a.32770

35. Bertoli-Avella A, Gillis E, Morisaki H, Verhagen J, de Graaf B, van de Beek G, et al. Mutations in a TGF-β ligand, TGFB3, cause syndromic aortic aneurysms and dissections. *J Am Coll Cardiol* (2015) 65(13):1324–36. doi:10.1016/j.jacc.2015.01.040

36. Warnes C, Williams R, Bashore T, Child J, Connolly H, Dearani J, et al. ACC/AHA 2008 guidelines for the management of adults with congenital heart disease. *J Am Coll Cardiol* (2008) 52(23):e143–263. doi:10.1016/j.jacc.2008.10.001

frontiers
in Cardiovascular Medicine

Challenges in Molecular Diagnostics of Channelopathies in the Next-Generation Sequencing Era: Less Is More?

Valeria Novelli[1,2], Patrick Gambelli[3], Mirella Memmi[3] and Carlo Napolitano[3]*

[1] Medical Genetics Unit, Fondazione Policlinico Universitario A. Gemelli, Rome, Italy, [2] Centro Studi "Benito Stirpe" per la prevenzione della morte improvvisa nel giovane atleta, Rome, Italy, [3] Molecular Cardiology, IRCCS Fondazione Salvatore Maugeri, Pavia, Italy

INTRODUCTION

Inherited arrhythmogenic diseases (IADs – also called cardiac channelopathies) are defined as a group of genetic diseases characterized by electrically unstable substrate in a structurally normal heart (1). Genetic testing in cardiac channelopathies has completed its transition from a research-based activity to that of a clinical genetic service. In parallel, the advancements of the sequencing technologies are providing ways to sequence several genes at a relatively low cost. This is progressively changing the approach to the genetic diagnosis of IADs. Indeed, while Sanger-based genetic testing was traditionally limited to the well characterized, most prevalent genes, next-generation sequencing (NGS) allows screening even the "minor" disease genes with very short turnaround time. This approach to genetic testing is highly efficient, but it is also generating remarkable interpretative problems mostly related to the high prevalence of variants of unknown significance (VUS), i.e., not clearly related to disease pathophysiology. This issue is relevant to several genetic diseases of the heart but is particularly evident in IADs, where the familial co-segregation analysis is hampered by the incomplete penetrance and the variable expressivity. How to design the most appropriate screening approach is challenging and it requires in-depth knowledge of the specific diseases of interest. In this Opinion article, we will review the available NGS approaches and try to outline the available strategies to optimize the performance of this genetic testing methodology.

GENETIC TESTING OF INHERITED ARRHYTHMIAS IN THE ERA OF NGS

Long QT syndrome (LQTS), Brugada syndrome (BrS), and catecholaminergic polymorphic ventricular tachycardia (CPVT) are the main channelopathies that can cause for sudden cardiac death (SCD) in children or young adults. In the last few years, the number of genes and genetic variants associated with these diseases has increased. For example, there are 15 known LQTS[1] genes and at least 16 BrS genes.[2] Importantly, however, in each disease, there are few major genes and a larger number of genes accounting for few cases each. The "minor" genes are usually poorly characterized in terms of function and pathophysiological role. As a consequence, the identification of mutations in these genes often leads to results that are difficult to interpret. Therefore, the HRS expert consensus

---

[1] http://www.ncbi.nlm.nih.gov/books/NBK1129/
[2] http://www.ncbi.nlm.nih.gov/books/NBK1517/

statement on the diagnosis and management of patients with inherited arrhythmias syndromes (2) has outlined the indications for genetic testing on the basis of the epidemiological relevance of the genes and the clinical implications of genetic testing for each disease (i.e., how much the identification of the mutation can impact the clinical management).

The NGS, a massive parallel sequencing technology that revolutionized the genetic diagnostics, allows large-scale and rapid assessment of the entire human genome (3). In principle, there are three approaches that can be used: (1) whole genome sequencing (WGS), applied to sequence the entire genome, coding, and non-coding regions; (2) whole exome sequencing (WES) used to analyze only the "exome," which represents 1% of the whole genome; (3) target resequencing panel (TRS) of genes, adopted to sequence selected gene sets/panels (4).

The first two approaches, WGS and WES, are mainly applied for research purposes, for discovery of new disease genes, while TRS is commonly used for the diagnosis in the clinical setting (5).

Recently, Pua et al. reported a comparison study applying different approaches of sequencing, such as TRS, WES, and WGS (6). Analyzing a custom panel, including 174 genes involved in inherited cardiac disease, they investigated the performances of the approaches across this set of genes. Results showed that TRS approach achieved a higher coverage (>99.8% at ≥20× read depth) compared with the other approaches (88.1 and 99.3%; WES and WGS, respectively, at ≥20× read depth). Furthermore, this approach has been reported to be faster and more affordable.

APPROACH TO NGS IN INHERITED ARRHYTHMIAS

In the pre-NGS era, the analysis of yield of genetic testing provided a clear evidence of the tight link between the severity of the clinical phenotype and rate of identified mutations. Bai et al. (7) showed a high yield of screening (64, 51, and 13% for LQTS, CPVT, and BrS, respectively) in patients with a conclusive diagnosis compared with the borderline cases (14, 13 and 2%) (7). A similar concept also applies in the NGS era. Clinicians are tempted to use the fast and efficient NGS technology as a diagnostic tool when clinical examinations are inconclusive. This can lead to the identification of a high rate of VUS, especially on minor genes. Thus, the selection of genes to be included in TRS is crucial. In general, there are three NGS strategies available:

(1) Comprehensive cardio panel: 60–180 genes covering all known genes (channelopathies and/or cardiomyopathies).
(2) Comprehensive arrhythmias panel: 20–60 genes restricted to arrhythmogenic conditions (**Table 1**).
(3) "Key gene" panels: few genes (3–6 genes), with high evidence, related to a specific phenotype.

In patients with conclusive diagnosis, use of TRS panels with a limited set of well-characterized genes should be considered the first step to reduce the number of tests with uncertain findings (first tier) (**Figure 1**).

The optimal strategy in subjects, who turn out negative in the first step, is much less defined. After the exclusion of the

TABLE 1 | Genes included in comprehensive arrhythmias panels.

Genes	Location
ABCC9	12p12.1
ACTN2	1q43
AKAP9	7q21.2
ANK2	4q25–q26
ANKRD1	10q23.31
ANKX2.5	5q35.1
CACNA1C	12p13.33
CACNA2D1	7q21.11
CACNB2	10p12.33–p12.31
CALM1	14q32.11
CALM2	2p21
CALM3	19q13.32
CASQ2	1p13.1
CAV3	3p25.3
CTNNA3	10q21.3
DES	2q35
DSC2	18q12.1
DSG2	18q12.1
DSP	6p24.3
EMD	Xq28
GPD1L	3p22.3
HCN4	15q24.1
JUP	17q21.2
KCND3	1p13.2
KCNE1	21q22.12
KCNE2	21q22.11
KCNE3	11q13.4
KCNE5	Xq23
KCNH2	7q36.1
KCNJ2	17q24.3
KCNJ5	11q24.3
KCNJ8	12p12.1
KCNQ1	11p15.5–p15.4
LDB3	10q23.2
LMN	1q22
PDLIM3	4q35.1
PKP2	12p11.21
PLN	6q22.31
PRKAG2	7q36.1
RANGRF	17p13.1
RBM20	10q25.2
RYR2	1q43
SCN10A	3p22.2
SCN1B	19q13.11
SCN2B	11q23.3
SCN3B	11q24.1
SCN4B	11q23.3
SCN5A	3p22.2
SLMAP	3p14.3
SNTA1	20q11.21
TBX5	12q24.21
TGFB3	14q24.3
TMEM43	3p25.1
TNNI3	19q13.42
TNNT2	1q32.1
TRDN	6q22.31
TRPM4	19q13.33
TTN	2q31.2

FIGURE 1 | Sequencing strategy for genetic testing.
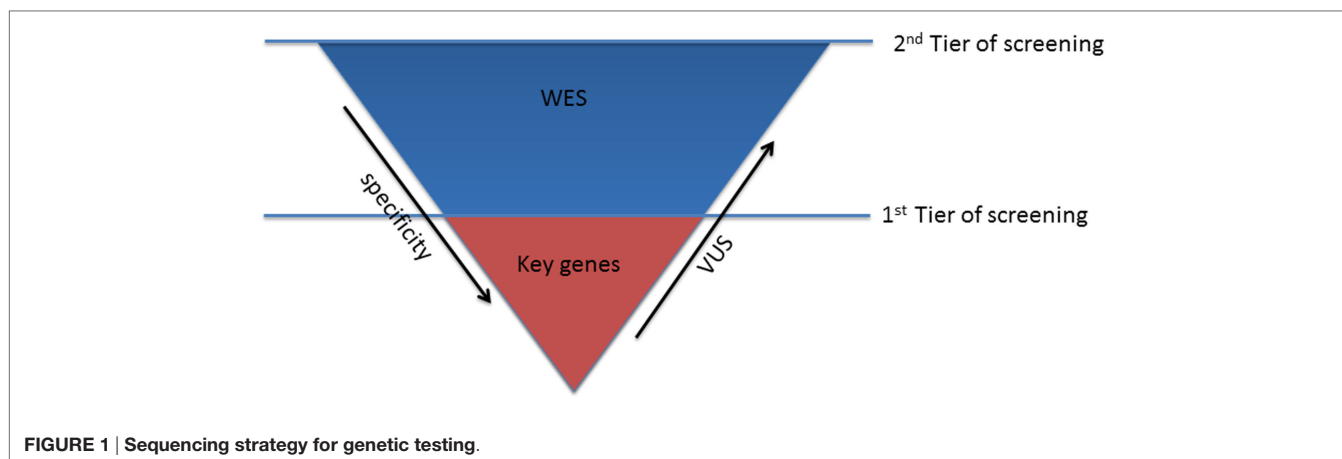
keys genes, the second tier of screening, using WES approach, can be considered (**Figure 1**). This choice might be preferable over the use of comprehensive cardio panels, due to the limited evidence of the minor genes associated with channelopathies, which cannot justify the investment required for the design and production of the second larger disease-specific gene panel. Therefore, WES will guarantee the consideration of all the mutations in the minor genes that have not been unraveled yet and consider also rare genetic variants in novel genes, still unrelated with the phenotype.

Nevertheless, it is clear that there may be hurdles also in the interpretations of WES data. Independently from the screening approach, it should be considered that for diagnostic purposes the presence of (1) a clear pathophysiological link between genetic variant and the phenotype and (2) the co-segregation within families, still remain crucial for the interpretation (8).

An example of the first tier strategy was recently reported by Millat et al. (9). Analyzing a cohort of 15 LQTS with a key panel, including only the main five genes associated with LQTS (*KCNQ1*, *KCNH2*, *SCN5A*, *KCNE1,* and *KCNE2*), they compared TRS and Sanger sequencing. The results showed that Sanger efficiently sequenced all the 69 exons compared with the TRS that sequenced 55/69 exons (86% of the targeted regions). NGS–TRS showed cost and turnaround time advantages over Sanger method. The study by Millat et al. highlights a very relevant problem, which is common to all NGS platforms: lack of coverage of specific regions of genes. In some cases, the problem can be particularly relevant. For example, several exons of *KCNH2*, a highly prevalent LQTS gene (~35% of patients), are not completely sequenced due to their high CG rich sequence. Thus, integration with Sanger sequencing of uncovered regions is often required with a consequence impact on costs and turnaround time.

Another interesting study on the evaluation of the first tier approach has been reported by Steffensen et al. in a cohort of 39 patients analyzed for the main genes associated with LQTS (*KCNQ1*, *KCNH2*, *SCN5A*, and *KCNE1*) (10). Results showed a high percentage of patients (17 patients; 44%) carrying variants classified as pathogenic compared with patients carrying VUS or VUS likely pathogenic (11 patients; 28%) and no alterations (13 patients; 34%).

Enlarging the screening to other seven minor genes (*ANK2*, *KCNJ2*, *CACNA1C*, *CAV3*, *SCN4B*, *AKAP9*, and *SNTA1*) associated with LQTS, the authors identified only three more variants, two classified as VUS and one as likely benign, demonstrating a very limited contribution, when including minor genes in the screening but a significant increase in the cost of the genetic tests.

The use of large panels, inclusive of all the minor genes may have additional limitations, as reported by Alfares et al. in patients with hypertrophic cardiomyopathy (HCM) (11). They tested over 9 years, 2,912 probands referred for clinical HCM genetic testing with different approaches: 11-gene panel, 18-gene panel, and a 50-gene pan cardiomyopathy panel. Results showed that the majority of positive tests were due to pathogenic or likely pathogenic variants in the *MYBPC3* and *MYH7* genes (83%), the two well-characterized genes routinely screened even with Sanger sequencing. Furthermore, analyzing a subset of 202 HCM patients with 18-gene panel and the pan cardiomyopathy panel, none of the probands had a causative variant outside the 18 "classic" HCM genes, suggested that use of the extended cardiomyopathy gene panel is useless for patients with HCM and should be reserved for patients with atypical clinical phenotypes (12).

SUMMARY AND FUTURE DEVELOPMENTS

Next-generation sequencing technology has improved significantly over the past few years. However, there are still some limitations that need to be considered in terms of sensitivity of the uncovered regions (lower than Sanger sequencing). Moreover, the high-throughput capability is revealing itself as a double edge sword: on the one hand, it allows amazingly short turnaround time and reduced costs, but on the other hand, it reveals an increased rate of VUS and tests that are not conclusive (and therefore clinically irrelevant). A way to overcome this problem can be the implementation of a shared knowledge based on VUS. International collaborative efforts for the annotation of genetic variants are currently being explored as a mean to improve the interpretation capabilities for NGS results (13). The ClinVar

database[3] is a publicly available tool for deposition and retrieval of variant data and annotations (14). This effort is expected to support the decision on the pathogenicity of identified variants and, most importantly, to resolve the classification of VUS. Meanwhile, the most appropriate use of NGS is that of a phenotype-driven approach with sequencing panels with a limited number of well-known genes and used in patients with clear clinical indications

for genetic testing. Therefore, if causative mutations are not identified on the "key" disease IAD genes, the analysis should take a "research track" with the use of WES. However, patients should be counseled accordingly.

---

[3]https://www.ncbi.nlm.nih.gov/clinvar/

AUTHOR CONTRIBUTIONS

All authors (VN, MM, PG, and CN) contributed extensively to the work presented in this opinion article.

REFERENCES

1. Bezzina CR, Lahrouchi N, Priori SG. Genetics of sudden cardiac death. *Circ Res* (2015) 116(12):1919–36. doi:10.1161/CIRCRESAHA.116.304030
2. Priori SG, Wilde AA, Horie M, Cho Y, Behr ER, Berul C, et al. HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of patients with inherited primary arrhythmia syndromes: document endorsed by HRS, EHRA, and APHRS in May 2013 and by ACCF, AHA, PACES, and AEPC in June 2013. *Heart Rhythm* (2013) 10(12):1932–63. doi:10.1016/j.hrthm.2013.05.014
3. Lubitz SA, Ellinor PT. Next-generation sequencing for the diagnosis of cardiac arrhythmia syndromes. *Heart Rhythm* (2015) 12(5):1062–70. doi:10.1016/j.hrthm.2015.01.011
4. de Koning TJ, Jongbloed JD, Sikkema-Raddatz B, Sinke RJ. Targeted next-generation sequencing panels for monogenetic disorders in clinical diagnostics: the opportunities and challenges. *Expert Rev Mol Diagn* (2015) 15(1):61–70. doi:10.1586/14737159.2015.976555
5. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* (2013) 14(4):295–300. doi:10.1038/nrg3463
6. Pua CJ, Pua CJ, Bhalshankar J, Miao K, Walsh R, John S, et al. Development of a comprehensive sequencing assay for inherited cardiac condition genes. *J Cardiovasc Transl Res* (2016) 9(1):3–11. doi:10.1007/s12265-016-9673-5
7. Bai R, Napolitano C, Bloise R, Monteforte N, Priori SG. Yield of genetic screening in inherited cardiac channelopathies: how to prioritize access to genetic testing. *Circ Arrhythm Electrophysiol* (2009) 2(1):6–15. doi:10.1161/CIRCEP.108.782888
8. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet* (2014) 59(1):5–15. doi:10.1038/jhg.2013.114
9. Millat G, Chanavat V, Rousson R. Evaluation of a new high-throughput next-generation sequencing method based on a custom AmpliSeq library

and ion torrent PGM sequencing for the rapid detection of genetic variations in long QT syndrome. *Mol Diagn Ther* (2014) 18(5):533–9. doi:10.1007/s40291-014-0099-y
10. Steffensen AB, Refaat MM, David JP, Mujezinovic A, Calloe K, Wojciak J, et al. High incidence of functional ion-channel abnormalities in a consecutive Long QT cohort with novel missense genetic variants of unknown significance. *Sci Rep* (2015) 5:10009. doi:10.1038/srep10009
11. Alfares AA, Kelly MA, McDermott G, Funke BH, Lebo MS, Baxter SB, et al. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genet Med* (2015) 17(11):880–8. doi:10.1038/gim.2014.205
12. Ackerman MJ. Genetic purgatory and the cardiac channelopathies: exposing the variants of uncertain/unknown significance issue. *Heart Rhythm* (2015) 12(11):2325–31. doi:10.1016/j.hrthm.2015.07.002
13. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen – the clinical genome resource. *N Engl J Med* (2015) 372(23):2235–42. doi:10.1056/NEJMsr1406261
14. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* (2016) 44(D1):D862–8. doi:10.1093/nar/gkv1222

Next-Generation Sequencing in Post-mortem Genetic Testing of Young Sudden Cardiac Death Cases

Najim Lahrouchi[1], Elijah R. Behr[2] and Connie R. Bezzina[1]*

[1] *Department of Clinical and Experimental Cardiology, Heart Center, AMC, Amsterdam, Netherlands,* [2] *Cardiology Clinical Academic Group, St George's University of London, London, UK*

Sudden cardiac death (SCD) in the young (<40 years) occurs in the setting of a variety of rare inherited cardiac disorders and is a disastrous event for family members. Establishing the cause of SCD is important as it permits the pre-symptomatic identification of relatives at risk of SCD. Sudden arrhythmic death syndrome (SADS) is defined as SCD in the setting of negative autopsy findings and toxicological analysis. In such cases, reaching a diagnosis is even more challenging and post-mortem genetic testing can crucially contribute to the identification of the underlying cause of death. In this review, we will discuss the current achievements of "the molecular autopsy" in young SADS cases and provide an overview of key challenges in assessing pathogenicity (i.e., causality) of genetic variants identified through next-generation sequencing.

Keywords: sudden cardiac death, post-mortem genetic testing, molecular autopsy, next-generation sequencing, channelopathy, cardiomyopathy

INTRODUCTION

Each year, thousands of individuals die suddenly before the age of 35. Sudden cardiac death (SCD) in this age category has an estimated incidence of 0.005–0.2 per 1000 individuals per year, which is lower than in the general adult population (1). The causes of SCD in the young can be grouped into (1) structural heart disease, where the heart is structurally abnormal and (2) the channelopathies in which the heart is structurally normal (**Figure 1**) (2). Post-mortem analysis of young SCD cases uncovers a structural cardiac pathology in the majority of cases. However, a subset of around 30% remains unexplained (3). Sudden arrhythmic death syndrome (SADS) is defined as SCD in the setting of a negative autopsy and toxicological analysis (4, 5). In these cases, reaching a diagnosis is challenging and post-mortem genetic testing, the so-called *molecular autopsy*, can crucially contribute to the identification of the underlying (genetic) cause of death (6). This is important for clinical and genetic evaluation of surviving family members that are potentially at risk of SCD (7). The recent advances in sequencing technologies (next-generation sequencing) have made it possible to screen in detail large proportions of the human genome at relatively low cost. However, despite these significant developments, distinguishing true disease-causing genetic variants from the bulk of genetic variation that is not directly associated with the SCD phenotype is of major importance (8). In this review, we will discuss the current achievements of the molecular autopsy in young SADS cases and provide an overview of key challenges in assessing pathogenicity (i.e., causality) of genetic variants identified through next-generation sequencing (NGS).

FIGURE 1 | Overview of causes of sudden cardiac death in the young based on post-mortem studies. HCM, hypertrophic cardiomyopathy; ARVC, arrhythmogenic right ventricular cardiomyopathy. LQTS, long QT syndrome; CPVT, catecholaminergic polymorphic ventricular tachycardia; SADS, sudden arrhythmic death syndrome. Reprinted from Semsarian et al. (3) with permission of the publisher.
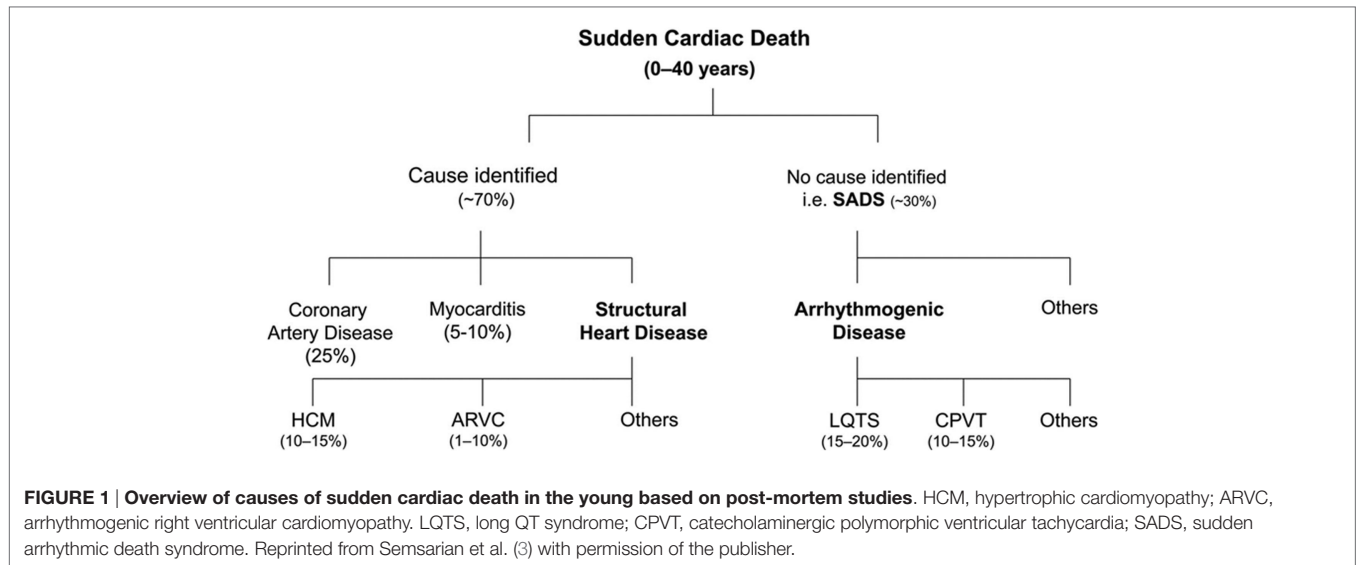
THE CARDIAC CHANNELOPATHIES

The cardiac channelopathies form a group of inherited disorders associated with the occurrence of arrhythmia and SCD in the presence of a structurally normal heart. These diseases are caused by mutations in genes that encode cardiac ion channel subunits or proteins that regulate and interact with ion channels. The underlying genetic defect leads to cardiac electrical disturbances that have the potential to initiate lethal cardiac arrhythmia (2). The cardiac channelopathies include, among others, the Long QT syndrome (LQTS), the Short QT syndrome (SQTS), Brugada syndrome (BrS), and catecholaminergic polymorphic ventricular tachycardia (CPVT) (5).

Long QT Syndrome

The LQTS is characterized by prolongation of the QT-interval on the surface electrocardiogram (ECG) associated with syncope and SCD as a result of *torsades des pointes* (TdP) ventricular tachycardia (VT) (9). The disease is genetically heterogeneous and has an estimated prevalence of 1:2000 (10). The inheritance pattern is generally autosomal dominant and mutations in 16 different genes have been associated with the disorder (11). Together, mutations in three major LQTS-causing genes account for ~90% of genotype-positive LQTS patients (7, 12). These genes include *KCNQ1* encoding for the Kv7.1 potassium channel (LQT1, 40–55%), *KCNH2* (LQT2, 30–45%) encoding for the Kv11.1 potassium channel, and *SCN5A* (LQT3, 5–10%) that encodes for the Nav1.5 sodium channel. Genotype–phenotype studies have uncovered genotype-specific clinical presentations that can contribute to the diagnosis of SADS cases based on the circumstances of the SCD (13). In LQT1, cardiac events occur typically during exercise and more specifically during swimming and diving, whereas in LQT2 symptoms are often triggered by sudden auditory stimuli. Patients with LQT3 usually present with symptoms during rest or sleep. The 13 minor LQTS-associated genes have been linked to LQTS in small studies with varying

evidence of disease association (2). LQTS can also present with extra-cardiac features. The Jervell and Lange-Nielsen (JLN) syndrome is characterized by significant QTc-interval prolongation accompanied by severe arrhythmias and sensorineural deafness. JLNS is caused by homozygous or compound heterozygous mutations in *KCNQ1* (14) or *KCNE1* (15). The Andersen–Tawil syndrome (LQT7) presents with QTc-interval prolongation, hypokalemic periodic paralysis and facial dysmorphism. The disease is caused by mutations in KCNJ2 (16). Timothy syndrome (LQT8) presents with severe QTc-prolongation, cardiac arrhythmia, syndactyly, autism, and malignant hypoglycemia. The most common associated mutation is the heterozygous G406R mutation in *CACNA1C* (17). The presence of extra-cardiac features has the potential to contribute to the unequivocal identification of the underlying genetic defect and identify an overlooked clinical diagnosis.

Catecholaminergic Polymorphic Ventricular Tachycardia

Catecholaminergic polymorphic ventricular tachycardia is an inherited arrhythmia syndrome characterized by the onset of life-threatening arrhythmia during exercise or acute emotional stress (18). These patients have a normal resting ECG and the disease can be diagnosed using exercise-stress testing or Holter recording, revealing typical bidirectional or polymorphic VT (5). When left untreated SCD occurs in up to 30% of cases before the age of 40 (19, 20). The autosomal-dominant form of CPVT is caused by mutations in *RYR2* (21) encoding for the ryanodine receptor, whereas an autosomal recessive and more rare form is caused by biallelic mutations in *CASQ2* (22) that encodes for the calsequestrin-2 protein. In addition, mutations in *TRDN*, *CALM1*, *KCNJ2*, and *ANKB* have also been identified in a small set of CPVT patients (2). Mutations in *RYR2* can be identified in ~60% of CPVT cases that have a classical phenotype and these mutations are mainly located in clusters within the gene (21, 23, 24). Genotype-phenotype studies have been conducted

and these data suggest a higher arrhythmia risk associated with mutations in the C-terminal portion of the protein (25).

Brugada Syndrome

Brugada syndrome can present with syncope due to polymorphic VT and SCD as a result of ventricular fibrillation. SCD most commonly occurs during rest or sleep and it typically occurs in males in the fourth decade of life (5, 26). Recent guidelines state that BrS is diagnosed when a coved ST-segment elevation of ≥0.2 mV is present in at least one precordial lead, either occurring spontaneously or after administration of a sodium channel-blocking agent (5). The typical ECG pattern can be concealed and may be intermittently present. In addition to sodium channel blockers, the typical BrS ECG pattern can also be induced by pyrexia (26). Loss-of-function mutations in *SCN5A*, encoding for the Nav1.5 sodium channel, are identified in ~16% of BrS cases (27). In addition to *SCN5A*, multiple other genes have been associated with this disorder (2). Even though the yield of genetic testing is low, genetic testing of *SCN5A* can identify a pathogenic mutation that could contribute to further genetic risk stratification in the family (5, 7). The observation that within some families the *SCN5A* mutation does not segregate with the disease suggests a potential modifying or more complex role for other genetic factors (28). Furthermore, a recent study suggested a more complex form of inheritance for the BrS with an important role for common genetic variation in disease susceptibility (29).

Short QT Syndrome

The SQTS presents with a short QT-interval on the surface ECG (<350 ms) predisposing to supraventricular and ventricular arrhythmia and is associated with a high risk of SCD (30, 31). The disorder is genetically heterogeneous and inherited in an autosomal-dominant mode. SQTS has been associated with pathogenic variants in genes that encode potassium channels (*KCNQ1*, *KCNH2*, and *KCNJ2*), which are also implicated in LQTS (32–34). Importantly, SQTS-causing variants in these genes lead to a gain-of-function on the affected channel, whereas the LQTS-causing variants lead to a loss-of-function. In addition, Cav1.2 L-type calcium channel subunits (*CACNA1C*, *CACNB2*) have been associated with SQTS (35). Even though in half of SQTS cases familial disease is present, the yield of genetic testing is around 14% (36).

THE CARDIOMYOPATHIES

The inherited cardiomyopathies include hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and arrhythmogenic cardiomyopathy (ACM) (37). The hallmark of HCM is unexplained ventricular hypertrophy, and myocyte disarray and fibrosis during histological analysis (38). The disease has an autosomal-dominant mode of inheritance in the majority of cases, with mutations predominantly located in genes encoding sarcomeric proteins. Most mutations are found in *MYBPC3* and *MYH7* (39, 40). SCD occurs in only a small subset of HCM cases (38). DCM can present with heart failure due to dilatation of the left ventricle and systolic dysfunction (41). In approximately one-third of patients with idiopathic DCM, a positive family

history for DCM can be identified (42). The inheritance pattern varies and is most commonly autosomal dominant or autosomal recessive, whereas X-linked inheritance or mitochondrial inheritance is less common (43, 44). The disease is genetically heterogeneous and more than 30 genes have been associated with DCM, although the evidence of disease association is highly variable. The most common genetic causes of DCM are found in *TTN*, *MYH7*, *LMNA*, and *TNNT2* (43). Importantly, mutations in *LMNA* have been associated with a form of DCM with significant cardiac conduction abnormalities and the occurrence of cardiac arrhythmia. Therefore, the identification of a mutation in *LMNA* during molecular autopsy has the potential to offer pre-symptomatic intervention (e.g., implantable defibrillator, pacemaker) to surviving family members carrying the familial *LMNA* mutation (45). ACM is characterized by fibrofatty infiltration of the myocardium and a high susceptibly to ventricular arrhythmia and SCD at young age (46). The disease is most commonly inherited in an autosomal-dominant fashion and gene mutations are mostly found in the following desmosomal genes: *PKP2*, *JUP*, *DSP*, *DSC2*, and *DSG2* (47). ACM has a variable disease expressivity and reduced penetrance among mutation carriers (48). It may affect the right ventricle predominantly (arrhythmogenic right ventricular cardiomyopathy – ARVC), the left ventricle, or both. Genetic testing in ACM can be helpful to identify family members at risk (7).

THE MOLECULAR AUTOPSY

Post-mortem genetic testing, using DNA extracted from blood or other tissue after death, has an important role in the identification of the underlying genetic cause in SADS cases (i.e., SCD cases with negative toxicology and pathology analysis). This process has been termed the "molecular autopsy" (**Figure 2**). Recent guidelines recommend the use of post-mortem genetic testing in cases where clinical evidence suggests a diagnosis of the LQTS or CPVT (5, 7).

In 1999, the identification of LQT1 as the underlying cause of death in a 19-year-old female was reported by Ackerman and colleagues (49). Several years after this report, Chugh and colleagues analyzed 5 LQTS-associated genes (*KCNQ1*, *KCNH2*, *SCN5A*, *KCNE1*, and *KCNE2*) in 12 sudden unexplained death cases in whom no diagnosis could be established after thorough post-mortem analysis of 270 adult SCD cases. Through this analysis, the authors identified the same *KCNH2* missense mutation in 2 out of 12 cases (yield of genetic testing: 17%) (50). Shortly afterwards, another study reported the post-mortem genetic analysis in 10 cases of juvenile (13–29 years) sudden unexplained death cases and identified LQTS-associated mutations in two patients (51). Subsequently, multiple similar post-mortem genetic studies have been conducted by several groups (52–58). In one study, 33 young cases were examined for LQT1–6 genes, and a putative pathogenic mutation was identified in 15% of patients (59). Tester and colleagues conducted a post-mortem analysis in 49 cases screening 18 exons of the CPVT-associated gene *RYR2* (60). In a subsequent study in the same cohort, these authors analyzed the three major genes associated with LQTS (*KCNQ1*, *KCNH2*, and *SCN5A*) (61). The genetic yield of CPVT and that

FIGURE 2 | Overview of the molecular autopsy process. Adjusted from Semsarian et al. (3) with permission of the publisher.
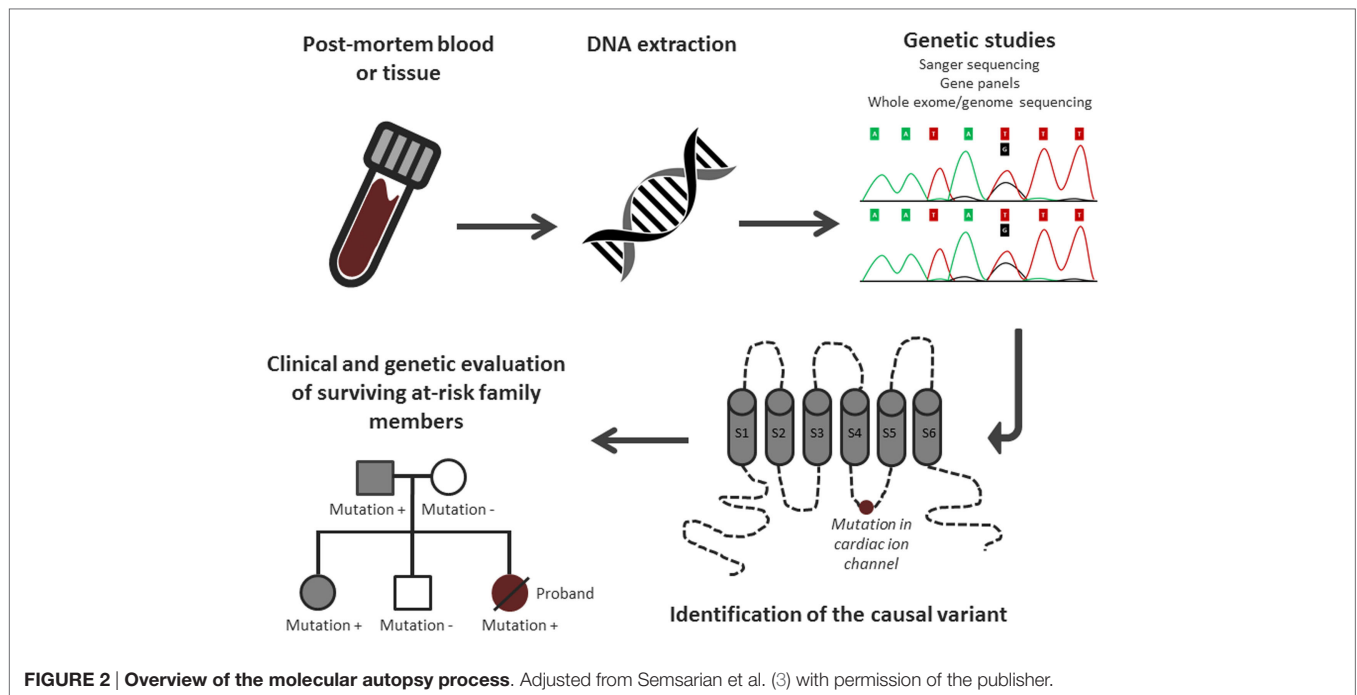
of LQTS genetic testing were, respectively, 14 and 20%, with an overall genetic yield reaching 35%. In an extended cohort of 173 autopsy-negative sudden unexplained death cases from the same group, five genes associated with LQTS and *RYR2* were screened (62). In this expanded analysis, 25% out of the 173 cases carried a potentially pathogenic variant in a LQTS-associated gene (14.5%) and *RYR2* (12%). Even though SCD was the presenting event in the majority of these patients, nearly 60% of the mutation positive cases had a family history of cardiac events. These studies showed that a significant proportion of unexplained death in the young is caused by cardiac channelopathies.

NEXT-GENERATION SEQUENCING MOLECULAR AUTOPSY STUDIES

The above-mentioned molecular autopsy studies have investigated a small number of channelopathy-associated genes. Recent advances in sequencing technologies (next-generation sequencing) have now made it possible to screen in detail an increasing number of genes in cardiac gene panels (i.e., >100 genes) at relatively low cost and using a limited amount of DNA. In addition, whole-exome sequencing (WES), where the coding regions of all ~22,000 genes is sequenced, has been introduced in post-mortem genetic testing as well. It is important to note that these NGS-based studies did not only consider more genes, but also extended to the inclusion of genes involved in the inherited cardiomyopathies (in addition to the channelopathy genes). The role of the cardiomyopathy-associated genes in normal-cardiac autopsy SCD cases remains largely unexplored. In evaluating these NGS-based studies, one should keep in mind that they not only screened varying numbers of genes but also employed different methods of variant prioritization (based on minor

allele frequency (MAF) in the general population as cut-off, *in silico* prediction tools for variant pathogenicity). Therefore, the genetic yield of these studies should be interpreted in relation to the varying variant curation and categorization. Bagnall and colleagues, conducted a post-mortem WES study in 28 sudden unexplained death cases and identified three rare variants in the major LQTS-associated genes when they focused their analysis on only a small panel of four genes (*KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2*) (63). In subsequent analyses, more than 70 arrhythmia and cardiomyopathy-associated genes were included and this led to the identification of an additional variant in *CACNAC1* that had been previously reported in a LQTS family. Of note, this additional analysis (using a MAF cut-off of <0.1% in 7500 publically available exomes) identified a large number of variants of unknown significance (VUSs), attesting to the complexity of analyzing such data. In a more recent study, WES followed by the analysis of 135 genes associated with cardiac channelopathies and cardiomyopathies was performed in 59 SADS victims (age range: 1–51 years) (64). Of these, 20 cases had subtle post-mortem cardiac structural abnormalities not reaching the diagnostic criteria for one of the cardiomyopathies. A primary analysis using a filtering MAF ≤0.02% based on the NHLBI exome sequencing project identified rare variants in seven probands. Three of these variants were located in ion channel genes of which two were known LQTS-associated *de novo* variants in *SCN5A* and one known CPVT-associated variant in *RYR2.* The other four rare variants were found in cardiomyopathy-associated genes. In a secondary analysis, using a MAF cut-off of 0.02–0.5%, previously reported variants were identified in an additional 10 probands. However, the clinical significance of these variants has yet to be determined.

Recently, Hertz and colleagues screened 52 SCD cases with non-diagnostic structural cardiac abnormalities during autopsy using a gene panel consisting of 100 genes previously associated

with cardiac channelopathies and cardiomyopathies (65). Genetic variants were prioritized using MAF in control populations (<1%), measures of evolutionary sequence conservation, prediction of deleteriousness, and prior disease association of the variant in the Human Genome Mutation Database (HGMD). Variants were subsequently classified as (a) likely, (b) unknown, or (c) unlikely to have functional effects by two physicians. Fifteen individuals (29%) were identified as carriers of variants with "likely functional effects" according to their classification system. In another study, Ackerman and colleagues performed WES and gene-specific analysis of 117 sudden death-susceptibility genes in 14 cases of sudden unexplained death in the young (66). In their analyses, eight rare variants in six genes were identified in seven cases. More recently, the same authors performed WES in 21 cases in whom no mutation was found during the screening of *KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2* (67). Interestingly, three variants (*CALM2*-F90L, *CALM2*-N98S and *PKP2*-N634fs) were classified as pathogenic according to the quideline recommendations of the American College of Medical Genetics (ACMG) (68). Of the 18 remaining cases, 7 carried at least 1 VUS in 1 of the 100 genes associated with SCD.

Thus far, several comparable post-mortem genetic studies using NGS have been conducted recently by several groups (69–72). Collectively, from these studies, it is clear that expanding the number of tested genes from small channelopathy panels to large panels containing a broader set of channelopathy genes, and even the cardiomyopathy-associated genes, increases the yield of likely causal variants only slightly as opposed to the large number of VUSs that are uncovered. The interpretation of these variants is challenging and their clinical utility is currently minimal. In addition, the large majority of SADS cases remain unexplained despite NGS screening of large gene panels.

IMPLICATING GENETIC VARIANTS IDENTIFIED THROUGH NGS IN THE MOLECULAR AUTOPSY

Post-mortem genetic testing using NGS is plagued by the same issues as genetic testing in patients with aborted SCD (and many other disorders) with the added complication that one cannot undertake further clinical tests in the deceased patient. The incorporation of NGS in post-mortem genetic testing requires the capability of assessing the genetic variants identified. False assignment of causality can have significant consequences for patients and their families (73). Even though assessing pathogenicity (i.e., causality) of genetic variants is complex, there are several steps to aid in this process (74, 68). It is important to note that each of these steps contributes to rather than determines the classification of a given variant.

Gene-Level Implication
Unlike the major channelopathy or cardiomyopathy-associated genes, some of the minor associated genes have been implicated in disease in small studies and evidence of disease association has not always been robust (absence of linkage data or absence of recurrent implication of the gene in independent families).

Including these minor genes in NGS panels often leads to the identification of a plethora of VUSs. Their clinical utility in establishing the diagnosis in a SADS case and for genetic risk stratification of family members is, therefore, likely to be small. Therefore, the evaluation of an identified variant should start with the assessment of the published data linking that gene to a specific form of disease. In addition, these data should also be taken into consideration during the design of clinical channelopathy and cardiomyopathy gene panels.

Variant-Level Implication
The assessment of a genetic variant has to take into account the large background of genetic variation in the human genome. Healthy individuals carry multiple rare protein-altering variants and this has been described as "genetic background noise." Consequently, one of the first important steps in variant prioritization is filtering using the variant MAF in the general population using large ancestry-matched publically available reference databases, such as the Exome Aggregation Consortium (WES data from >60,000 individuals) (75). However, rarity of a variant does not, by definition, implicate disease causality.

After the identification of a genetic variant in a SADS case, co-segregation with disease status should whenever possible be performed in surviving family members. *De novo* inheritance of rare genetic variation in an SCD-associated gene in a SADS case, with unaffected parents, provides strong evidence for disease association. Of importance, parental mosaicism, as opposed to *de novo* inheritance should be taken into account during genetic counseling as this could lead to the false assumption that siblings are genetically unaffected. Parental mosaicism has been described previously in Timothy syndrome (LQT8) (76).

The previous identification of the genetic variant in an independent proband displaying the same or similar phenotype is also highly valuable. Such previous associations can be found by scanning the literature and by using in-house or public databases of disease variants. Of importance, these previous published studies should be assessed carefully (i.e., study design, co-segregation in the family, functional data) to assess the strength of disease association. In this regard, some of the previously published "pathogenic" variants in the literature have later been shown to be at such a high MAF in the general population that their role in disease is questioned (77–79). The assessment of a variant's pathogenicity would benefit from centralized depositories that include curated evidence for previously identified disease-associated variants.

Computational prediction tools, such as sorting intolerant from tolerant (SIFT) and PolyPhen2, can be helpful in the process but should be handled with caution. Measures of evolutionary sequence conservation among species (orthologs) and among proteins derived from same ancestral gene (paralogs) can have value in the assessment of variants. Paralog annotation tools have been applied to the cardiac channelopathies and are freely available online (80). The Grantham score is a measure of the difference in the physicochemical properties of the amino acid substitutions and a higher score indicates larger differences between amino acids (81). Combining these *in silico* prediction tools has been performed for *KCNQ1*, *KCNH2*, and *SCN5A* and

has shown a synergistic utility during the assessment of genetic variation within these genes (82, 83). Most of the recently conducted NGS-based port-mortem genetic testing studies have also incorporated *in silico* prediction tools in variant prioritization (65, 69, 70). Despite these developments, prediction algorithms should not be regarded as stand-alone evidence of pathogenicity. Although certain classes of genetic variation, such as splice-site or truncating variants, are much more likely to affect the protein, their role should be assessed in the specific gene context and if loss of function is a known mechanism of disease. Functional studies can contribute to the understanding of a variant's biological consequences. However, these studies are labor-intensive and require specialized research centers.

CONCLUSION AND FUTURE DIRECTIONS

Next-generation sequencing (NGS) has made it possible to screen large gene panels, spanning not only the channelopathy genes but also the cardiomyopathy genes, in search for the cause of SCD. While these panels have made it possible to broadly screen for genetic variation, it comes with the challenge of interpreting any identified VUS. As seen for the cardiac channelopathies and cardiomyopathies, the genetic architecture of SADS is characterized by large genetic and allelic heterogeneity, which adds to the difficulty of genetic screening in these patients. Even though the majority of SADS cases remain elusive after NGS screening, the generated data make it possible to combine similar datasets

through future international collaboration. This has the huge potential to demonstrate statistically an excess of rare genetic variation in known SCD genes (or more interestingly in new genes) in comparison to controls through burden testing (74). Even though presumed to be monogenic, the genetic architecture of SADS is largely unknown in the majority of cases and such case-control studies could point toward a genetic model in which an accumulation of rare genetic variation is required to develop symptoms. However, implementation of the oligogenic model in the segregation within families will be challenging and may require different approaches.

AUTHOR CONTRIBUTIONS

All authors researched data for the article, discussed its content, and wrote, edited, and reviewed the manuscript.

FUNDING

REFERENCES

1. Meyer L, Stubbs B, Fahrenbruch C, Maeda C, Harmon K, Eisenberg M, et al. Incidence, causes, and survival trends from cardiovascular-related sudden cardiac arrest in children and young adults 0 to 35 years of age: a 30-year review. *Circulation* (2012) **126**:1363–72. doi:10.1161/CIRCULATIONAHA.111.076810

2. Bezzina CR, Lahrouchi N, Priori SG. Genetics of sudden cardiac death. *Circ Res* (2015) **116**:1919–36. doi:10.1161/CIRCRESAHA.116.304030

3. Semsarian C, Ingles J, Wilde AAM. Sudden cardiac death in the young: the molecular autopsy and a practical approach to surviving relatives. *Eur Heart J* (2015) **36**:1290–6. doi:10.1093/eurheartj/ehv063

4. Behr ER, Casey A, Sheppard M, Wright M, Bowker TJ, Davies MJ, et al. Sudden arrhythmic death syndrome: a national survey of sudden unexplained cardiac death. *Heart* (2007) **93**:601–5. doi:10.1136/hrt.2006.099598

5. Priori SG, Wilde AA, Horie M, Cho Y, Behr ER, Berul C, et al. HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of patients with inherited primary arrhythmia syndromes: document endorsed by HRS, EHRA, and APHRS in May 2013 and by ACCF, AHA, PACES, and AEPC in June 2013. *Heart Rhythm* (2013) **10**:1932–63. doi:10.1016/j.hrthm.2013.05.014

6. Miles CJ, Behr ER. The role of genetic testing in unexplained sudden death. *Transl Res* (2016) **168**:59–73. doi:10.1016/j.trsl.2015.06.007

7. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm* (2011) **8**:1308–39. doi:10.1016/j.hrthm.2011.05.020

8. Lubitz SA, Ellinor PT. Next generation sequencing for the diagnosis of cardiac arrhythmia syndromes. *Heart Rhythm* (2015) **12**(5):1062–70. doi:10.1016/j.hrthm.2015.01.011

9. Schwartz PJ, Moss AJ, Vincent GM, Crampton RS. Diagnostic criteria for the long QT syndrome. An update. *Circulation* (1993) **88**:782–4. doi:10.1161/01.CIR.88.2.782

10. Schwartz PJ, Stramba-Badiale M, Crotti L, Pedrazzini M, Besana A, Bosi G, et al. Prevalence of the congenital long-QT syndrome. *Circulation* (2009) **120**:1761–7. doi:10.1161/CIRCULATIONAHA.109.863209

11. Mizusawa Y, Horie M, Wilde AA. Genetic and clinical advances in congenital long QT syndrome. *Circ J* (2014) **78**(12):2827–33. doi:10.1253/circj.CJ-14-0905

12. Hedley PL, Jørgensen P, Schlamowitz S, Wangari R, Moolman-Smook J, Brink PA, et al. The genetic basis of long QT and short QT syndromes: a mutation update. *Hum Mutat* (2009) **30**:1486–511. doi:10.1002/humu.21106

13. Schwartz PJ, Priori SG, Spazzolini C, Moss AJ, Vincent GM, Napolitano C, et al. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. *Circulation* (2001) **103**:89–95. doi:10.1161/01.CIR.103.1.89

14. Neyroud N, Tesson F, Denjoy I, Leibovici M, Donger C, Barhanin J, et al. A novel mutation in the potassium channel gene KVLQT1 causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nat Genet* (1997) **15**:186–9. doi:10.1038/ng0297-186

15. Schulze-Bahr E, Wang Q, Wedekind H, Haverkamp W, Chen Q, Sun Y, et al. KCNE1 mutations cause Jervell and Lange-Nielsen syndrome. *Nat Genet* (1997) **17**:267–8. doi:10.1038/ng1197-267

16. Plaster NM, Tawil R, Tristani-Firouzi M, Canún S, Bendahhou S, Tsunoda A, et al. Mutations in Kir2.1 cause the developmental and episodic electrical phenotypes of Andersen's syndrome. *Cell* (2001) **105**:511–9. doi:10.1016/S0092-8674(01)00342-7

17. Splawski I, Timothy KW, Sharpe LM, Decher N, Kumar P, Bloise R, et al. Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* (2004) **119**:19–31. doi:10.1016/j.cell.2004.09.011

18. Leenhardt A, Lucet V, Denjoy I, Grau F, Ngoc DD, Coumel P. Catecholaminergic polymorphic ventricular tachycardia in children. A 7-year follow-up of 21 patients. *Circulation* (1995) **91**:1512–9. doi:10.1161/01.CIR.91.5.1512

19. Priori SG, Napolitano C, Memmi M, Colombi B, Drago F, Gasparini M, et al. Clinical and molecular characterization of patients with catecholaminergic polymorphic ventricular tachycardia. *Circulation* (2002) **106**:69–74. doi:10.1161/01.CIR.0000020013.73106.D8

20. Hayashi M, Denjoy I, Extramiana F, Maltret A, Buisson NR, Lupoglazoff JM, et al. Incidence and risk factors of arrhythmic events in catecholaminergic polymorphic ventricular tachycardia. *Circulation* (2009) **119**:2426–34. doi:10.1161/CIRCULATIONAHA.108.829267

21. Priori SG, Napolitano C, Tiso N, Memmi M, Vignati G, Bloise R, et al. Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation* (2001) **103**:196–200. doi:10.1161/01.CIR.103.2.196

22. Lahat H, Pras E, Olender T, Avidan N, Ben-Asher E, Man O, et al. A missense mutation in a highly conserved region of CASQ2 is associated with autosomal recessive catecholamine-induced polymorphic ventricular tachycardia in Bedouin families from Israel. *Am J Hum Genet* (2001) **69**:1378–84. doi:10.1086/324565

23. Laitinen PJ, Brown KM, Piippo K, Swan H, Devaney JM, Brahmbhatt B, et al. Mutations of the cardiac ryanodine receptor (RyR2) gene in familial polymorphic ventricular tachycardia. *Circulation* (2001) **103**:485–90. doi:10.1161/01.CIR.103.4.485

24. Medeiros-Domingo A, Bhuiyan ZA, Tester DJ, Hofman N, Bikker H, van Tintelen JP, et al. The RYR2-encoded ryanodine receptor/calcium release channel in patients diagnosed previously with either catecholaminergic polymorphic ventricular tachycardia or genotype negative, exercise-induced long QT syndrome: a comprehensive open reading frame mutational analysis. *J Am Coll Cardiol* (2009) **54**:2065–74. doi:10.1016/j.jacc.2009.08.022

25. van der Werf C, Nederend I, Hofman N, van Geloven N, Ebink C, Frohn-Mulder IM, et al. Familial evaluation in catecholaminergic polymorphic ventricular tachycardia: disease penetrance and expression in cardiac ryanodine receptor mutation-carrying relatives. *Circ Arrhythm Electrophysiol* (2012) **5**:748–56. doi:10.1161/CIRCEP.112.970517

26. Antzelevitch C, Brugada P, Borggrefe M, Brugada J, Brugada R, Corrado D, et al. Brugada syndrome: report of the second consensus conference. *Heart Rhythm* (2005) **2**:429–40. doi:10.1016/j.hrthm.2005.01.005

27. Crotti L, Marcou CA, Tester DJ, Castelletti S, Giudicessi JR, Torchio M, et al. Spectrum and prevalence of mutations involving BrS1- through BrS12-susceptibility genes in a cohort of unrelated patients referred for Brugada syndrome genetic testing: implications for genetic testing. *J Am Coll Cardiol* (2012) **60**:1410–8. doi:10.1016/j.jacc.2012.04.037

28. Probst V, Wilde AA, Barc J, Sacher F, Babuty D, Mabo P, et al. SCN5A mutations and the role of genetic background in the pathophysiology of Brugada syndrome. *Circ Cardiovasc Genet* (2009) **2**:552–7. doi:10.1161/CIRCGENETICS.109.853374

29. Bezzina CR, Barc J, Mizusawa Y, Remme CA, Gourraud JB, Simonet F, et al. Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. *Nat Genet* (2013) **45**:1044–9. doi:10.1038/ng.2712

30. Gussak I, Brugada P, Brugada J, Wright RS, Kopecky SL, Chaitman BR, et al. Idiopathic short QT interval: a new clinical syndrome? *Cardiology* (2000) **94**:99–102. doi:10.1159/000047299

31. Gaita F, Giustetto C, Bianchi F, Wolpert C, Schimpf R, Riccardi R, et al. Short QT syndrome: a familial cause of sudden death. *Circulation* (2003) **108**:965–70. doi:10.1161/01.CIR.0000085071.28695.C4

32. Brugada R, Hong K, Dumaine R, Cordeiro J, Gaita F, Borggrefe M, et al. Sudden death associated with short-QT syndrome linked to mutations in HERG. *Circulation* (2004) **109**:30–5. doi:10.1161/01.CIR.0000109482.92774.3A

33. Bellocq C, van Ginneken AC, Bezzina CR, Alders M, Escande D, Mannens MM, et al. Mutation in the KCNQ1 gene leading to the short QT-interval syndrome. *Circulation* (2004) **109**:2394–7. doi:10.1161/01.CIR.0000130409.72142.FE

34. Priori SG, Pandit SV, Rivolta I, Berenfeld O, Ronchetti E, Dhamoon A, et al. A novel form of short QT syndrome (SQT3) is caused by a mutation in the KCNJ2 gene. *Circ Res* (2005) **96**:800–7. doi:10.1161/01.RES.0000162101.76263.8c

35. Antzelevitch C, Pollevick GD, Cordeiro JM, Casis O, Sanguinetti MC, Aizawa Y, et al. Loss-of-function mutations in the cardiac calcium channel underlie a new clinical entity characterized by ST-segment elevation, short QT intervals, and sudden cardiac death. *Circulation* (2007) **115**:442–9. doi:10.1161/CIRCULATIONAHA.106.668392

36. Mazzanti A, Kanthan A, Monteforte N, Memmi M, Bloise R, Novelli V, et al. Novel insight into the natural history of short QT syndrome. *J Am Coll Cardiol* (2014) **63**:1300–8. doi:10.1016/j.jacc.2013.09.078

37. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, et al. Classification of the cardiomyopathies: a position statement from the European Society Of Cardiology Working Group on myocardial and pericardial diseases. *Eur Heart J* (2008) **29**:270–6. doi:10.1093/eurheartj/ehm342

38. Maron BJ, Maron MS. Hypertrophic cardiomyopathy. *Lancet* (2013) **381**:242–55. doi:10.1016/S0140-6736(12)60397-3

39. Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, et al. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation* (2003) **107**:2227–32. doi:10.1161/01.CIR.0000066323.15244.54

40. Li Q, Gruner C, Chan RH, Care M, Siminovitch K, Williams L, et al. Genotype-positive status in patients with hypertrophic cardiomyopathy is associated with higher rates of heart failure events. *Circ Cardiovasc Genet* (2014) **7**:416–22. doi:10.1161/CIRCGENETICS.113.000331

41. Jefferies JL, Towbin JA. Dilated cardiomyopathy. *Lancet* (2010) **375**:752–62. doi:10.1016/S0140-6736(09)62023-7

42. Petretta M, Pirozzi F, Sasso L, Paglia A, Bonaduce D. Review and metaanalysis of the frequency of familial dilated cardiomyopathy. *Am J Cardiol* (2011) **108**:1171–6. doi:10.1016/j.amjcard.2011.06.022

43. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol* (2013) **10**:531–47. doi:10.1038/nrcardio.2013.105

44. Towbin JA. Inherited cardiomyopathies. *Circ J* (2014) **78**:2347–56. doi:10.1253/circj.CJ-14-0893

45. van Rijsingen IA, Arbustini E, Elliott PM, Mogensen J, Hermans-van Ast JF, van der Kooi AJ, et al. Risk factors for malignant ventricular arrhythmias in lamin a/c mutation carriers a European cohort study. *J Am Coll Cardiol* (2012) **59**:493–500. doi:10.1016/j.jacc.2011.08.078

46. Basso C, Bauce B, Corrado D, Thiene G. Pathophysiology of arrhythmogenic cardiomyopathy. *Nat Rev Cardiol* (2012) **9**:223–33. doi:10.1038/nrcardio.2011.173

47. Marcus FI, Edson S, Towbin JA. Genetics of arrhythmogenic right ventricular cardiomyopathy: a practical guide for physicians. *J Am Coll Cardiol* (2013) **61**:1945–8. doi:10.1016/j.jacc.2013.01.073

48. Lodder EM, Bezzina CR. Arrhythmogenic right ventricular cardiomyopathy: growing evidence for complex inheritance. *Circ Cardiovasc Genet* (2013) **6**:525–7. doi:10.1161/CIRCGENETICS.113.000367

49. Ackerman MJ, Tester DJ, Porter CJ, Edwards WD. Molecular diagnosis of the inherited long-QT syndrome in a woman who died after near-drowning. *N Engl J Med* (1999) **341**:1121–5. doi:10.1056/NEJM199910073411504

50. Chugh SS, Senashova O, Watts A, Tran PT, Zhou Z, Gong Q, et al. Postmortem molecular screening in unexplained sudden death. *J Am Coll Cardiol* (2004) **43**:1625–9. doi:10.1016/j.jacc.2003.11.052

51. Di Paolo M, Luchini D, Bloise R, Priori SG. Postmortem molecular analysis in victims of sudden unexplained death. *Am J Forensic Med Pathol* (2004) **25**:182–4. doi:10.1097/01.paf.0000127406.20447.8a

52. Creighton W, Virmani R, Kutys R, Burke A. Identification of novel missense mutations of cardiac ryanodine receptor gene in exercise-induced sudden death at autopsy. *J Mol Diagn* (2006) **8**:62–7. doi:10.2353/jmoldx.2006.050081

53. Nishio H, Iwata M, Suzuki K. Postmortem molecular screening for cardiac ryanodine receptor type 2 mutations in sudden unexplained death: R420W mutated case with characteristics of status thymico-lymphaticus. *Circ J* (2006) **70**:1402–6. doi:10.1253/circj.70.1402

54. Gladding PA, Evans CA, Crawford J, Chung SK, Vaughan A, Webster D, et al. Posthumous diagnosis of long QT syndrome from neonatal screening cards. *Heart Rhythm* (2010) **7**:481–6. doi:10.1016/j.hrthm.2009.12.023

55. Winkel BG, Larsen MK, Berge KE, Leren TP, Nissen PH, Olesen MS, et al. The prevalence of mutations in KCNQ1, KCNH2, and SCN5A in an unselected national cohort of young sudden unexplained death cases. *J Cardiovasc Electrophysiol* (2012) **23**:1092–8. doi:10.1111/j.1540-8167.2012.02371.x

56. Larsen MK, Berge KE, Leren TP, Nissen PH, Hansen J, Kristensen IB, et al. Postmortem genetic testing of the ryanodine receptor 2 (RYR2) gene in a

cohort of sudden unexplained death cases. *Int J Legal Med* (2013) **127**:139–44. doi:10.1007/s00414-011-0658-2

57. Dean J, Cann F, Corbett M, O'Sullivan D, Tennant S, Hailey H, et al. Molecular autopsy for sudden cardiac death – outcome of a practical approach. *Heart Lung Circ* (2014) **23**:e3. doi:10.1016/j.hlc.2014.07.007

58. Stattin EL, Westin IM, Cederquist K, Jonasson J, Jonsson BA, Mörner S, et al. Genetic screening in sudden cardiac death in the young can save future lives. *Int J Legal Med* (2016) **130**:59–66. doi:10.1007/s00414-015-1237-8

59. Skinner JR, Crawford J, Smith W, Aitken A, Heaven D, Evans CA, et al. Prospective, population-based long QT molecular autopsy study of postmortem sudden death in 1 to 40 year olds. *Heart Rhythm* (2011) **8**:412–9. doi:10.1016/j.hrthm.2010.11.016

60. Tester DJ, Spoon DB, Valdivia HH, Makielski JC, Ackerman MJ. Targeted mutational analysis of the RyR2-encoded cardiac ryanodine receptor in sudden unexplained death: a molecular autopsy of 49 medical examiner/coroner's cases. *Mayo Clin Proc* (2004) **79**:1380–4. doi:10.4065/79.11.1380

61. Tester DJ, Ackerman MJ. Postmortem long QT syndrome genetic testing for sudden unexplained death in the young. *J Am Coll Cardiol* (2007) **49**:240–6. doi:10.1016/j.jacc.2006.10.010

62. Tester DJ, Medeiros-Domingo A, Will ML, Haglund CM, Ackerman MJ. Cardiac channel molecular autopsy: insights from 173 consecutive cases of autopsy-negative sudden unexplained death referred for postmortem genetic testing. *Mayo Clin Proc* (2012) **87**:524–39. doi:10.1016/j.mayocp.2012.02.017

63. Bagnall RD, Das KJ, Duflou J, Semsarian C. Exome analysis-based molecular autopsy in cases of sudden unexplained death in the young. *Heart Rhythm* (2014) **11**:655–62. doi:10.1016/j.hrthm.2014.01.017

64. Nunn LM, Lopes LR, Syrris P, Murphy C, Plagnol V, Firman E, et al. Diagnostic yield of molecular autopsy in patients with sudden arrhythmic death syndrome using targeted exome sequencing. *Europace* (2015). doi:10.1093/europace/euv285

65. Hertz CL, Christiansen SL, Ferrero-Miliani L, Dahl M, Weeke PE, LuCamp, et al. Next-generation sequencing of 100 candidate genes in young victims of suspected sudden cardiac death with structural abnormalities of the heart. *Int J Legal Med* (2016) **130**:91–102. doi:10.1007/s00414-015-1261-8

66. Narula N, Tester DJ, Paulmichl A, Maleszewski JJ, Ackerman MJ. Post-mortem whole exome sequencing with gene-specific analysis for autopsy-negative sudden unexplained death in the young: a case series. *Pediatr Cardiol* (2015) **36**:768–78. doi:10.1007/s00246-014-1082-4

67. Anderson JH, Tester DJ, Will ML, Ackerman MJ. Whole exome molecular autopsy following exertion-related sudden unexplained death in the young. *Circ Cardiovasc Genet* (2016). doi:10.1161/CIRCGENETICS.115.001370

68. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) **17**:405–24. doi:10.1038/gim.2015.30

69. Farrugia A, Keyser C, Hollard C, Raul JS, Muller J, Ludes B. Targeted next generation sequencing application in cardiac channelopathies: analysis of a cohort of autopsy-negative sudden unexplained deaths. *Forensic Sci Int* (2015) **254**:5–11. doi:10.1016/j.forsciint.2015.06.023

70. Hertz CL, Christiansen SL, Ferrero-Miliani L, Fordyce SL, Dahl M, Holst AG, et al. Next-generation sequencing of 34 genes in sudden unexplained death victims in forensics and in patients with channelopathic cardiac diseases. *Int J Legal Med* (2015) **129**:793–800. doi:10.1007/s00414-014-1105-y

71. Neubauer J, Haas C, Bartsch C, Domingo-Medeiros A, Berger W. Post-mortem whole-exome sequencing (WES) with a focus on cardiac disease-associated genes in five young sudden unexplained death (SUD) cases. *Int J Legal Med* (2016). doi:10.1007/s00414-016-1317-4

72. Buscemi L, Alessandrini F, Perna G, Tagliabracci A. Next-generation sequencing of 68 genes in sudden unexplained death of young individuals in forensics. *Forensic Sci Int Genet Suppl Ser* (2015) **5**:e138–40. doi:10.1016/j.fsigss.2015.09.056

73. Ackerman MJ. Genetic purgatory and the cardiac channelopathies: exposing the variants of uncertain/unknown significance issue. *Heart Rhythm* (2015) **12**:2325–31. doi:10.1016/j.hrthm.2015.07.002

74. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature* (2014) **508**:469–76. doi:10.1038/nature13127

75. *Exome Aggregation Consortium (ExAC)*. Cambridge, MA (2016). Available from: http://exac.broadinstitute.org; http://exac.broadinstitute.org/about

76. Etheridge SP, Bowles NE, Arrington CB, Pilcher T, Rope A, Wilde AA, et al. Somatic mosaicism contributes to phenotypic variation in Timothy syndrome. *Am J Med Genet A* (2011) **155A**:2578–83. doi:10.1002/ajmg.a.34223

77. Refsgaard L, Holst AG, Sadjadieh G, Haunsø S, Nielsen JB, Olesen MS. High prevalence of genetic variants previously associated with LQT syndrome in new exome data. *Eur J Hum Genet* (2012) **20**:905–8. doi:10.1038/ejhg.2012.23

78. Risgaard B, Jabbari R, Refsgaard L, Holst AG, Haunsø S, Sadjadieh A, et al. High prevalence of genetic variants previously associated with Brugada syndrome in new exome data. *Clin Genet* (2013) **84**:489–95. doi:10.1111/cge.12126

79. Jabbari J, Jabbari R, Nielsen MW, Holst AG, Nielsen JB, Haunsø S, et al. New exome data question the pathogenicity of genetic variants previously associated with catecholaminergic polymorphic ventricular tachycardia. *Circ Cardiovasc Genet* (2013) **6**:481–9. doi:10.1161/CIRCGENETICS.113.000118

80. Ware JS, Walsh R, Cunningham F, Birney E, Cook SA. Paralogous annotation of disease-causing variants in long QT syndrome genes. *Hum Mutat* (2012) **33**:1188–91. doi:10.1002/humu.22114

81. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* (1974) **185**:862–4. doi:10.1126/science.185.4154.862

82. Giudicessi JR, Kapplinger JD, Tester DJ, Alders M, Salisbury BA, Wilde AA, et al. Phylogenetic and physicochemical analyses enhance the classification of rare nonsynonymous single nucleotide variants in type 1 and 2 long-QT syndrome. *Circ Cardiovasc Genet* (2012) **5**:519–28. doi:10.1161/CIRCGENETICS.112.963785

83. Kapplinger JD, Giudicessi JR, Ye D, Tester DJ, Callis TE, Valdivia CR, et al. Enhanced classification of Brugada syndrome-associated and long-QT syndrome-associated genetic variants in the SCN5A-encoded Na(v)1.5 cardiac sodium channel. *Circ Cardiovasc Genet* (2015) **8**:582–95. doi:10.1161/CIRCGENETICS.114.000831

The Role of Genetic Testing in the Identification of Young Athletes with Inherited Primitive Cardiac Disorders at Risk of Exercise Sudden Death

Francesco Danilo Tiziano[1], Vincenzo Palmieri[2], Maurizio Genuardi[1]* and Paolo Zeppilli[2]

[1] Istituto di Medicina Genomica, Università Cattolica del Sacro Cuore, Roma, Italy, [2] Unità di Medicina dello Sport, Fondazione Policlinico "A. Gemelli", Università Cattolica del Sacro Cuore, Roma, Italy

Although relatively rare, inherited primitive cardiac disorders (IPCDs) in athletes have a deep social impact since they often present as sudden cardiac death (SCD) of young and otherwise healthy persons. The diagnosis of these conditions is likely underestimated due to the lack of shared clinical criteria and to the existence of several borderline clinical pictures. We will focus on the clinical and molecular diagnosis of the most common IPCDs, namely hypertrophic cardiomyopathies, long QT syndrome, arrhythmogenic right ventricular cardiomyopathy, and left ventricular non-compaction. Collectively, these conditions account for the majority of SCD episodes and/or cardiologic clinical problems in athletes. In addition to the clinical and instrumental tools for the diagnosis of IPCD, the viral technological advances in genetic testing have facilitated the molecular confirmation of these conditions. However, genetic testing presents several issues: the limited sensitivity (globally, around 50%), the low prognostic predictive value, the probability to find pathogenic variants in different genes in the same patient, and the risk of non-interpretable results. In this review, we will analyze the pros and cons of the different clinical approaches for the presymptomatic identification, the diagnosis and management of IPCD athletes, and we will discuss the indications to the genetic testing for patients and their relatives, particularly focusing on the most complex scenarios, such as presymptomatic tests, uncertain results, and unexpected findings.

Keywords: athletes, sudden cardiac death, genetics, medical, hypertrophic cardiomyopathy, long QT syndrome, arrhythmogenic right ventricular displasia, isolated non-compact myocardium

INTRODUCTION

Inherited primitive cardiac disorders (IPCDs) comprise a wide and heterogeneous group of conditions. Two major subcategories of IPCDs are universally recognized: primitive cardiomyopathies and primitive electric disorders of the heart. Primitive cardiomyopathies can be defined as disorders characterized by morphologically and functionally abnormal myocardium, in the absence of other diseases that can cause the observed cardiac phenotype (1). This definition is aimed at distinguishing primitive cardiomyopathies from conditions in which the cardiac involvement is secondary to a systemic disorder. Primitive electric disorders are conditions characterized by the presence heart electric conduction disturbances with a morphologically normal myocardium (2). However, while

the pathophysiological mechanisms are different, the two groups of IPCDs encompass a continuous spectrum of diseases: rhythm disturbances occur in cardiomyopathies secondary to myocardial disarray and are the leading cause of death. Most IPCDs are Mendelian conditions, most commonly transmitted as autosomal dominant traits with incomplete penetrance, show familial recurrence and have a high degree of genetic heterogeneity (many genes causing the same or similar phenotypes).

First level diagnostic tools are standard cardiologic investigations, such as electrocardiography (ECG) and/or echocardiography (EchoCG), which can be supplemented with cardiac magnetic resonance and/or Holter ECG when needed (3). These will not be discussed in this review, since papers more focused on clinical aspects have been recently published (3).

The prevalence of IPCDs in general is likely underestimated, due to the reduced penetrance (see below), and to the existence of a wide gray zone between normal and definitely pathological instrumental findings.

In the setting of sports medicine, the identification and clinical management of IPCD becomes even more complex: the younger age of the population at risk and the presence of common features between IPCD and athlete heart may delay or prevent a timely diagnosis of these conditions; additionally, the extreme exertion may exacerbate underlying cardiac defects. The onset of IPCD symptoms in athletes is often dramatic, since these are the leading cause of sudden cardiac death (SCD). Thus, presymptomatic identification of affected individuals is of paramount relevance. Following an IPCD diagnosis, it is crucial to evaluate the eligibility for competitive or recreational sports (that implies a lesser cardiac impact), to establish the prognosis, to prevent the occurrence of fatal events, and, last but not least, to evaluate the risk for the athlete's offspring and sibship.

The main relevance of genetic testing for IPCDs is in identifying at risk subjects or to solve diagnostic uncertainties. The most commonly altered genes involved in IPCDs are listed in **Table 1**. The introduction of next-generation sequencing (NGS) platforms in diagnostic laboratories and the consequential reduction of costs of molecular analysis per patient have improved the diagnostic yield and reduced the time interval between sampling and final reports. However, the data accrued have revealed the complexity of the genetics of IPCD (4–10). With the traditional Sanger sequencing diagnostic approach, single genes were investigated sequentially, one after another, and often testing was interrupted when a pathogenic or likely pathogenic variant was found. NGS-based approaches allow to test several genes simultaneously. As a consequence, the finding of individuals carrying ≥2 rare pathogenic or potentially pathogenic variants in the same or different genes has become not uncommon. At the same time, there has been a surge in the numbers of variants of uncertain significance (VUS) detected. Moreover, it has become clear that allelic variants in the same gene can be associated with different phenotypes, increasing the difficulties inherent to the interpretation of genetic test results. These findings suggest that the variable phenotypic spectrum of IPCDs cannot be only accounted for by classical Mendelian mechanisms, and point toward the involvement of an oligogenic model with strong environmental influences.

TABLE 1 | The genes most commonly altered in cardiomyopathies.

Gene name	Gene symbol	Frequency (%)
Hypertrophic cardiomyopathy		
Beta-myosin heavy chain	*MYH7*	15–25
Cardiac myosin-binding protein C	*MYBPC3*	15–25
Cardiac troponin T	*TNNT2*	<5
Cardiac troponin I	*TNNI3*	<5
Alpha-tropomyosin	*TPM1*	<5
LIM-binding domain 3	*LBD3*	1–5
Ventricular regulatory myosin light chain	*MYL2*	<2
Myosin light chain 3	*MYL3*	<1
Cardiac muscle alpha actin	*ACTC1*	<1
Long QT syndrome		
Potassium channel, voltage gated, member 1 (Kv7.1; LQT1)	*KCNQ1*	40–55
Potassium channel, voltage gated, member 2 (Kv11.1; LQT2)	*KCNH2*	30–45
Sodium channel, voltage gated, type V, α subunit (Nav1.5; LQT3)	*SCN5A*	5–10
Arrhythmogenic right ventricular cardiomyopathy		
Plakophilin 2	*PKP2*	10–45
Desmoplakin	*DSP*	10–15
Desmoglein	*DSG2*	7–10
Desmocollin 2	*DSC-2*	2
Junction plakoglobin	*JUP*	<1

Genes that are rarely mutated have not been included [adapted from Bos et al. (11), Mizusawa et al. (12), and Iyer and Chin (13)].

In this review, we will focus on the clinical and molecular diagnosis of the most common IPCDs in athletes, namely hypertrophic cardiomyopathies (HCMs), long QT syndrome (LQTS), arrhythmogenic right ventricular cardiomyopathy (ARVC), and left ventricular non-compaction (LVNC). We will also discuss about the *TTN* gene, one of the largest genes in our genome encoding for the giant protein Titin, which is often altered in different clinical conditions. We will finally discuss the prognostic utility of genetic testing, and the counseling approaches to IPCD patients and their families.

HYPERTROPHIC CARDIOMYOPATHIES

Hypertrophic cardiomyopathies belong to the wider spectrum of cardiomyopathies that include also the dilative and restrictive phenotypes. HCM is diagnosed on the basis of left ventricular hypertrophy, in the absence of abnormal loading conditions. The estimated prevalence in the young adults is about 0.1–0.2%, which does not likely reflect the prevalence in the general population that is expected to be higher (14): available data have been obtained through clinical studies, and thus do not take into account the ascertainment biases related to later onset of symptoms and to the presence of borderline patients (14). According to studies performed in the US, HCM is the most common cause of SCD in young athletes (15–17); it is noteworthy that in countries where the preparticipation screening by ECG is mandatory by law, the incidence of HCM as cause of SCD among athletes is dramatically lower (18).

About half of HCM cases are familial with an autosomal dominant pattern of inheritance. More than 20 genes have been

related to HCM: b-myosin heavy chain (*MYH7*) and cardiac myosin-binding protein C (*MYBPC3*) account for about 50% of the cases; the other genes are rarely affected, with some involved in a single family so far. Incomplete penetrance is an important issue for HCM management. Environmental factors (such as intense training) and/or modifier genes may increase the risk of clinical manifestations especially during exercise or sport. If this was the case, one should expect to find more frequently clinical/instrumental signs of HCM among athletes compared with the general population. However, to the best of our knowledge, there are no available data on the true prevalence of HCM among professional athletes. Since the main complication of this condition is sudden death, the development of primary prevention programs aimed at identifying at risk subjects is very important, despite the relatively low frequency of HCM (16–18). There are not yet enough sensitive clinical markers that may help to identify HCM patients at risk of sudden death. The most reliable predictors are family history of sudden death related to HCM, syncope or presyncope events, ventricular tachyarrhythmia, marked hypotension during training, extreme left ventricle hypertrophy, and extended late enhancement at cardiac MRI (19–22), but their performance is far from satisfactory. A quantitative approach for the assessment of the risk of sudden death in HCM has been reported by O'Mahony et al. (23), the so-called HCM risk SCD. In this case, the risk is estimated on the basis of data collected in a retrospective longitudinal study by taking into account different variables.

Rather than for patients with clear HCM phenotypes, genetic testing may be useful for the proper interpretation of borderline patients falling in the gray zone. However, reduced penetrance, genetic heterogeneity, and high VUS frequency make the interpretation of the clinical significance of genetic variants challenging. Furthermore, a preliminary NGS-based study reported the occurrence of double heterozygosity in a high proportion of HCM patients, 2 of the 11 patients with pathogenic variants. These patients were reported as having a more severe phenotype compared with patients with a single disease causing variant (24).

With regard to practical implications for molecular diagnosis, according to the guidelines of the European Society of Cardiology (ESC), genetic testing could be offered to all patients fulfilling the HCM diagnostic criteria. Irrespective of the sequencing methodology employed, genetic analysis should include the most commonly implicated sarcomere protein genes (1). Following the identification of a definite pathogenic variant in the proband, genetic testing can be offered to all relatives on a voluntary basis. If no causative variants are found in the proband, relatives should be advised to undergo clinical reassessment should symptoms of HCM manifest.

LONG QT SYNDROME

Long QT syndrome is defined by the finding of a prolonged QT interval in standard ECG recording. It is generally accepted that the normal duration of the QT interval is 0.37–0.44 s. Based on this criterion, the diagnosis of LQTS is apparently easy, but 15% of subjects in the general population have a QT interval >0.44 s

(0.44–0.47 s) and 25–35% of individuals with a pathogenic variant in one of the LQTS genes has a normal QT interval (25, 26). This latter observation deserves some additional comments: at this stage, it is very difficult to establish whether the finding of a variant considered disease causing in asymptomatic patients is due to reduced penetrance or if, in the light of more recent concepts of molecular genetics, it is the consequence of a wrong interpretation, and the observed DNA change in a VUS, or even a rare benign, not clinically relevant, variation.

Since the diagnostic value of QT interval measurement on its own is not sufficient, a scoring method based on multiple parameters is currently used [**Table 2**; (27)]. LQTS belongs to the wider nosologic group of the channelopathies, and its cumulative prevalence is about 1/2,500: as in the case of HCM, this is likely an underestimate, due to the wide phenotypic heterogeneity. The majority of cases are familial (about 90%). As in the case of HCM, few genes account for the vast majority of cases: specifically, defects in *KCNQ1*, *KCNH2*, and *SCN5A* are found in about 80% of patients. Double heterozygotes are not uncommon: two pathogenic or likely pathogenic variants in different genes are observed in about 10% of patients, and these often display more severe phenotypes (28).

The diagnosis of LQTS in athletes is complicated by the correlation between duration of the QTc interval and exercise, and the extreme variation of heart rate reached by athletes. In two studies, the prevalence of LQT was 0.6 and 0.4% in an Italian and a British athlete population, respectively (29, 30), that is about 10- to 15-fold higher than in the general population. In the British

TABLE 2 | LQTS diagnostic criteria [from Schwartz and Crotti (25)].

	Points
Electrocardiographic findings[a]	
A. QTc[b]	
≥480 ms	3
460–479 ms	2
450–459 ms (in males)	1
B. QTc[b] fourth minute of recovery from exercise stress	
Test ≥480 ms	1
C. Torsade de pointes[c]	2
D. T wave alternans	1
E. Notched T wave in 3 leads	1
F. Low heart rate for age[d]	0.5
Clinical history	
A. Syncope[c]	
With stress	2
Without stress	1
B. Congenital deafness	0.5
Family history	
A. Family members with definite LQTS[e]	1
B. Unexplained sudden cardiac death below 30 years of age among immediate family members[e]	0.5

[a]*In the absence of medications or disorders known to affect these electrocardiographic features.*
[b]*QTc calculated by Bazett's formula where QTc = QT/√RR.*
[c]*Mutually exclusive.*
[d]*Resting heart rate below the second percentile for age.*
[e]*The same family member cannot be counted in A and B.*
Score: ≤1 point: low probability of LQTS; 1.5–3 points: intermediate probability of LQTS; ≥3.5 points high probability.

study, molecular screening of *KCNQ1*, *KCNH2*, and *SCN5A* was performed in five of the seven patients with LQT, three of whom had a QT interval >0.50 s and additional signs reinforcing the suspicion of LQTS; a pathogenic variant was found only in one patient. However, the apparently low yield of genetic testing in this cohort of patients could be related to technical limitations. A proper diagnosis of LQTS in athletes is of particular relevance: besides the obvious implications for the patient and the family, it entails also important career implications, since it is suggested that it may represent a contraindication to competitive sport disciplines involving moderate- and high-intensity strenuous exertion (31–33).

ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY

Arrhythmogenic right ventricular cardiomyopathy is a cardiac muscle disease characterized by life-threatening ventricular arrhythmias. The estimated prevalence is about 1:2,500–5,000. ARVC is considered one of the major causes of sudden death in young individuals and in athletes (18). ARVC is generally associated with ECG alterations, including negative T wave in right precordial leads, ventricular arrhythmias with a left bundle branch block morphology, epsilon waves, and others. However, some of these abnormalities are not specific and may be found in other pathological conditions with a different prognosis, such as myocarditis (34). The extensive use of ECG screening may help the sports physician to suspect the diagnosis, while genetic testing may be very useful for the differential diagnosis with more benign conditions.

Cardiac pathology shows dystrophy of the right ventricular myocardium with fibrofatty replacement. The clinical picture may include a subclinical asymptomatic phase; ventricular fibrillation, or an electrical disorder with palpitations and syncope, due to tachyarrhythmias of right ventricular origin, may be the first presentation. Most ARVC genes encode for proteins of mechanical cell junctions (*DSC2*, *DSG2*, *DSP*, *PKP2*, *JUP*, and *DES*), while others encode for structural proteins of the nuclear membrane (*LMNA* and *TMEM43*) or membrane channels (*RYR2*). ARVC has an autosomal dominant pattern of inheritance with incomplete penetrance. Pathogenic variations in the nine ARVC genes identified so far account for about 50% of cases. Double heterozygotes have been reported also in this condition. Clinical diagnosis may be achieved by demonstrating functional and structural alterations of the right ventricle, depolarization and repolarization abnormalities, arrhythmias with left bundle branch block morphology, and fibrofatty replacement upon endomyocardial biopsy [see Basso et al. (35) for a review]. Albeit rare, the diagnosis of ARVC is of crucial importance for athletes, due to the risk of sudden death: the condition was originally described as the most common cause of death in sportsmen. However, it is now evident that the condition has a wide phenotypic variability, including very mild asymptomatic cases: sport activity may increase the risk of ventricular arrhythmias in asymptomatic subjects with pathogenic variants in desmosomal genes (36).

LEFT VENTRICULAR NON-COMPACTION

Left ventricular non-compaction is due to the precocious arrest of myocardial compaction during the first weeks of the embryonic development. This causes persistence of prominent trabeculae in the ventricular cavity. The disease spectrum is very wide: the first reported cases were of patients with a marked dilation of the left ventricle and high risk of death (37–40), but asymptomatic and barely progressive segmental forms have also been reported, in which the lack of compaction involves only part of the left ventricle. LVNC is a rare disorder with prevalence <0.1%, although it has been increasingly diagnosed over the last few years. Similar to many other rare disorders, with increasing knowledge its diagnosis has become more common, and among newly identified cases, there is an increasing proportion of asymptomatic subjects, including athletes, with mild phenotypic expression. Indeed, heart hypertrabeculation has been observed in up to 18.3% of athletes (41), about 8% of which fulfill the diagnostic criteria of LVNC. In our experience, a multiparametric evaluation, based on morphological and functional parameters, such as the thickness of the residual compact layer and the presence of major conduction defects and arrhythmias, may help to discriminate between true cardiomyopathies and "benign" forms of LVNC (42). In the latter group of patients, the risk of sudden death, heart failure, and life-threatening arrhythmias is likely low. In any case, close follow-up is still recommended.

Genetic testing is not particularly useful for the molecular confirmation of LVNC for at least two reasons: the detection rate of pathogenic variants is relatively low, about 40%, and the genes involved in LVNC are also responsible also for other cardiomyopathies, complicating the interpretation of positive test results (43). Based on these findings, it has been proposed that LVNC may be a phenotypic variant of other cardiomyopathies, characterized by impaired general development of the sarcomeric proteins: this pathogenic model is supported by the cooccurrence in the same family of LVNC and different cardiomyopathies. It is conceivable that genetic background and environmental factors may play a relevant role in the onset of LVNC (44).

Regarding the relationship with sport activity, the number of incidental diagnoses has increased over time, often in asymptomatic athletes. Of note, hypertrabeculation of the left ventricle may physiologically occur in athletes, particularly in elite and black sportsmen. Thus, it is crucial to distinguish between the true cardiomyopathy and the benign segmental LVNC for the assessment of the risk of serious life-threatening events (45).

TITIN: A TITAN OR A GIANT WITH CLAY FEET?

Although if it is only one of the genes involved in cardiomyopathies, the titin gene (*TTN*) deserves a separate discussion due to its peculiarities. *TTN* is one of the largest genes in our genome and encodes for the largest human protein. The titin protein has several functions in both cardiac and skeletal muscle. Due to the size, prior to the advent of NGS, the mutational analysis of *TTN* was limited to few exons. The exact number of isoforms

is unknown, although it has been estimated that at least one-third of *TTN* exons may give raise to alternative splicing events (46, 47). *TTN* has been associated with both dominant and recessive disorders and is currently considered one of the most commonly altered genes in human disease (48), causing at least 10 different conditions, involving skeletal muscle, heart, or both. However, accruing data on genomic variations in the general population have shown that rare *TTN* variants overall are common, with at least 2–3% of healthy individuals bearing monoallelic truncating mutations. Rare and private missense variants are extremely common as well (47, 49). We should then expect a prevalence of recessive pathogenic variants of at least 1/4,000–10,000, much higher than the cumulative prevalence of titinopathies. Thus, it seems that at least a part, if not the majority, of truncating *TTN* variants is benign and does not cause pathological phenotypes on their own. The lack of pathogenicity of *TTN* alleles potentially causing complete loss of function could be explained by alternative splicing events rescuing the gene function.

Based on these observations, the pathogenic role of *TTN* variants should be assessed cautiously, especially for the potential application to presymptomatic-predictive testing in healthy relatives of patients in whom *TTN* alterations have been detected. It is likely that in the few next years, with accruing genomic data in the general population and the spreading of NGS platforms for the diagnosis of cardiomyopathies, the pathogenicity of *TTN* variants will be largely elucidated.

DISCUSSION

In this review, we have highlighted critical aspects associated with the clinical and genetic diagnosis of the IPCDs. The most recent findings on the variability of the human genome are quickly changing the approach to DNA variant interpretation. Indeed, a systematic assessment of variants, including those previously interpreted as pathogenic, is ongoing for several genes associated with inherited conditions. Overall, the following issues are associated with all types of IPCD and complicate their diagnosis and management: (1) wide genetic heterogeneity, (2) incomplete penetrance, (3) relatively high frequency of double heterozygotes, and (4) effect of environmental factors (largely unknown as well, besides sport activity). In the light of these characteristics, IPCDs could be considered as complex traits determined by the predominant effect of single gene variants, rather than as monogenic disorders. Considering IPCDs as oligogenic multifactorial disorders has three main implications: (1) genotype–phenotype correlations are unclear, (2) difficulty in establishing prognosis and risks for patients, and consequently, (3) genetic testing has a limited predictive power both in affected patients and in asymptomatic relatives at risk. Therefore, in our opinion, the use of the risk figures estimated according to Mendelian inheritance is not fully appropriate for predictive purposes. It could be useful to develop a risk assessment model similar to those applied for the familial predisposition to breast cancer, taking into account the presence of multiple factors (e.g., family history, level of exposure to physical activity, presence of multiple gene variants) (50). Indeed, with few exceptions (51–53),

due to poor genotype–phenotype correlations, results of clinical investigations provide better prognostic information than those of genetic testing.

These problems become even more complicated in athletes, who can display some features resembling those of IPCD as a consequence of physiological rearrangements of the myocardium with training. These subjects, who are mostly young, present some additional issues: the ascertainment of a variant known to be the cause of an IPCD phenotype may have strong implications for the prosecution of their sport career, for reproductive choices, and for their families. The offer of a predictive test for relatives should be considered with caution, and only when the pathogenicity of the variant detected in the proband has been clearly established according to consensus criteria (54). The opportunity of testing underage relatives of athletes should also be carefully scrutinized, especially when presymptomatic diagnosis may be beneficial, such as in the case of LQTS, which can manifest as infant sudden death.

Variant interpretation is the main issue in molecular diagnosis of IPCDs, and it is further exacerbated by the small size of many pedigrees. This hampers analysis of variant segregation with respect to the phenotype, one of the most useful points of evidence for clinical interpretation of genetic variants, as well the estimation of penetrance values. Ideally, novel or unclassified variants should be validated by functional studies; however, with few exceptions, these are not performed in clinical diagnostic laboratories and are associated with several issues, such as the choice of the cellular model and feasibility, since most sarcomeric mRNAs are quite large and not easily manageable. In conclusion, unless validated functional tests have been performed, the main hints suggesting pathogenicity of a variant are its identification in different patients or *de novo* occurrence.

Another open issue is the significance of double heterozygosity and related counseling. So far, it is common opinion that these subjects may in general display a more severe phenotype but the cohorts published so far are too small to draw definite conclusions.

TTN deserves separate considerations. In particular, given the difficulties inherent with *TTN* molecular testing and interpretation, one might wonder whether it is appropriate to include this gene in diagnostic panels and in genetic reports for patients, or rather, whether it should still be investigated in research settings for epidemiological purposes. These might shed light on the pathogenicity of *TTN* truncating and missense variants, as well as and on their clinical relevance; indeed, it remains to be established if *TTN* variants act as main phenotypic drivers and/or as a risk factor for the appearance of the clinical manifestation of some IPCDs, insufficient alone to determine a phenotype.

Similar to *TTN*, also in HCM the large amount of whole exome data that are accumulating in the different databases is disclosing that presumptive pathogenic variants can be found in "controls" at a higher rate than expected for the prevalence of the condition (55–57). On the one hand, this excess of pathogenic variants could be accounted for by reduced penetrance: carriers identified in the general population may or may not develop signs of HCM over time, but should be considered as asymptomatic

subjects. On the other hand, these findings may indicate that the effects of these gene variants are too weak to cause appearance of the phenotype on their own. In conclusion, the refinement of clinical diagnosis of IPCD, coupled with the new technological tools available in molecular genetics, has opened the Pandora box of cardiac primitive defects. Now, the pieces of this puzzle need to be reconstructed in order to provide patients and athletes with more accurate information and best care. But, there is still a long way to go.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

1. Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). *Eur Heart J* (2014) 35:2733–79. doi:10.1093/eurheartj/ehu284

2. Bezzina CR, Lahrouchi N, Priori SG. Genetics of sudden cardiac death. *Circ Res* (2015) 116:1919–36. doi:10.1161/CIRCRESAHA.116.304030

3. Myerson M, Sanchez-Ross M, Sherrid MV. Preparticipation athletic screening for genetic heart disease. *Prog Cardiovasc Dis* (2012) 54:543–52. doi:10.1016/j.pcad.2012.03.003

4. Bortot B, Athanasakis E, Brun F, Rizzotti D, Mestroni L, Sinagra G, et al. High-throughput genotyping robot-assisted method for mutation detection in patients with hypertrophic cardiomyopathy. *Diagn Mol Pathol* (2011) 20:175–9. doi:10.1097/PDM.0b013e31820b34fb

5. Fokstuen S, Munoz A, Melacini P, Iliceto S, Perrot A, Ozcelik C, et al. Rapid detection of genetic variants in hypertrophic cardiomyopathy by custom DNA resequencing array in clinical practice. *J Med Genet* (2011) 48:572–6. doi:10.1136/jmg.2010.083345

6. Meder B, Haas J, Keller A, Heid C, Just S, Borries A, et al. Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circ Cardiovasc Genet* (2011) 4:110–22. doi:10.1161/CIRCGENETICS.110.958322

7. Faita F, Vecoli C, Foffa I, Andreassi MG. Next generation sequencing in cardiovascular diseases. *World J Cardiol* (2012) 4:288–95. doi:10.4330/wjc.v4.i10.288

8. Lopes LR, Rahman MS, Elliott PM. A systematic review and meta-analysis of genotype-phenotype associations in patients with hypertrophic cardiomyopathy caused by sarcomeric protein mutations. *Heart* (2013) 99:1800–11. doi:10.1136/heartjnl-2013-303939

9. Loar RW, Bos JM, Will ML, Ommen SR, Ackerman MJ. Genotype-phenotype correlations of hypertrophic cardiomyopathy when diagnosed in children, adolescents, and young adults. *Congenit Heart Dis* (2015) 10:529–36. doi:10.1111/chd.12280

10. Lopes LR, Zekavati A, Syrris P, Hubank M, Giambartolomei C, Dalageorgou C, et al. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. *J Med Genet* (2013) 50:228–39. doi:10.1136/jmedgenet-2012-101270

11. Bos JM, Towbin JA, Ackerman MJ. Diagnostic, prognostic, and therapeutic implications of genetic testing for hypertrophic cardiomyopathy. *J Am Coll Cardiol* (2009) 54(3):201–11. doi:10.1016/j.jacc.2009.02.075

12. Mizusawa Y, Horie M, Wilde AA. Genetic and clinical advances in congenital long QT syndrome. *Circ J* (2014) 78:2827–33. doi:10.1253/circj.CJ-14-0905

13. Iyer VR, Chin AJ. Arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D). *Am J Med Genet C Semin Med Genet* (2013) 163C:185–97. doi:10.1002/ajmg.c.31368

14. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. *Circulation* (1995) 92:785–9. doi:10.1161/01.CIR.92.4.785

15. Maron BJ, Haas TS, Doerer JJ, Thompson PD, Hodges JS. Comparison of U.S. and Italian experiences with sudden cardiac deaths in young competitive athletes and implications for preparticipation screening strategies. *Am J Cardiol* (2009) 104:276–80. doi:10.1016/j.amjcard.2009.03.037

16. Maron BJ. Distinguishing hypertrophic cardiomyopathy from athlete's heart physiological remodelling: clinical significance, diagnostic strategies and implications for preparticipation screening. *Br J Sports Med* (2009) 43:649–56. doi:10.1136/bjsm.2008.054726

17. Maron BJ. Sudden death in hypertrophic cardiomyopathy. *J Cardiovasc Transl Res* (2009) 2:368–80. doi:10.1007/s12265-009-9147-0

18. Corrado D, Basso C, Rizzoli G, Schiavon M, Thiene G. Does sports activity enhance the risk of sudden death in adolescents and young adults? *J Am Coll Cardiol* (2003) 42:1959–63. doi:10.1016/j.jacc.2003.03.002

19. Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N Engl J Med* (1989) 321:1372–8. doi:10.1056/NEJM198911163212005

20. Elliott PM, Gimeno JR, Thaman R, Shah J, Ward D, Dickie S, et al. Historical trends in reported survival rates in patients with hypertrophic cardiomyopathy. *Heart* (2006) 92:785–91. doi:10.1136/hrt.2005.068577

21. Spirito P, Bellone P, Harris KM, Bernabo P, Bruzzi P, Maron BJ. Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy. *N Engl J Med* (2000) 342:1778–85. doi:10.1056/NEJM200006153422403

22. Delcrè SD, Di Donna P, Leuzzi S, Miceli S, Bisi M, Scaglione M, et al. Relationship of ECG findings to phenotypic expression in patients with hypertrophic cardiomyopathy: a cardiac magnetic resonance study. *Int J Cardiol* (2013) 167:1038–45. doi:10.1016/j.ijcard.2012.03.074

23. O'Mahony C, Jichi F, Pavlou M, Monserrat L, Anastasakis A, Rapezzi C, et al. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM risk-SCD). *Eur Heart J* (2014) 35:2010–20. doi:10.1093/eurheartj/eht439

24. Golbus JR, Puckelwartz MJ, Dellefave-Castillo L, Fahrenbach JP, Nelakuditi V, Pesce LL, et al. Targeted analysis of whole genome sequence data to diagnose genetic cardiomyopathy. *Circ Cardiovasc Genet* (2014) 7:751–9. doi:10.1161/CIRCGENETICS.113.000578

25. Vyas H, Hejlik J, Ackerman MJ. Epinephrine QT stress testing in the evaluation of congenital long-QT syndrome: diagnostic accuracy of the paradoxical QT response. *Circulation* (2006) 113:1385–92. doi:10.1161/CIRCULATIONAHA.105.600445

26. Priori SG, Napolitano C, Schwartz PJ. Low penetrance in the long-QT syndrome: clinical impact. *Circulation* (1999) 99:529–33. doi:10.1161/01.CIR.99.4.529

27. Schwartz PJ, Crotti L. QTc behavior during exercise and genetic testing for the long-QT syndrome. *Circulation* (2011) 124:2181–4. doi:10.1161/CIRCULATIONAHA.111.062182

28. Tester DJ, Will ML, Haglund CM, Ackerman MJ. Compendium of cardiac channel mutations in 541 consecutive unrelated patients referred for long QT syndrome genetic testing. *Heart Rhythm* (2005) 2:507–17. doi:10.1016/j.hrthm.2005.07.012

29. Corrado D, Basso C, Pavei A, Michieli P, Schiavon M, Thiene G. Trends in sudden cardiovascular death in young competitive athletes after implementation of a preparticipation screening program. *JAMA* (2006) 296:1593–601. doi:10.1001/jama.296.13.1593

30. Basavarajaiah S, Wilson M, Whyte G, Shah A, Behr E, Sharma S. Prevalence and significance of an isolated long QT interval in elite athletes. *Eur Heart J* (2007) 28:2944–9. doi:10.1093/eurheartj/ehm404

31. Zipes DP, Ackerman MJ, Estes NA III, Grant AO, Myerburg RJ, Van Hare G. Task force 7: arrhythmias. *J Am Coll Cardiol* (2005) 45:1354–63. doi:10.1016/j.jacc.2005.02.014

32. Pelliccia A, Corrado D, Bjørnstad HH, Panhuyzen-Goedkoop N, Urhausen A, Carre F, et al. Recommendations for participation in competitive sport and leisure-time physical activity in individuals with cardiomyopathies,

myocarditis and pericarditis. *Eur J Cardiovasc Prev Rehabil* (2006) 13:876–85. doi:10.1097/01.hjr.0000238393.96975.32

33. Biffi A, Delise P, Zeppilli P, Giada F, Pelliccia A, Penco M, et al. Italian Society of Sports Cardiology and Italian Sports Medicine Federation. Italian cardiological guidelines for sports eligibility in athletes with heart disease: part 1. *J Cardiovasc Med (Hagerstown)* (2013) 14:477–99. doi:10.2459/JCM.0b013e32835f6a21

34. Dello Russo A, Pieroni M, Santangeli P, Bartoletti S, Casella M, Pelargonio G, et al. Concealed cardiomyopathies in competitive athletes with ventricular arrhythmias and an apparently normal heart: role of cardiac electroanatomical mapping and biopsy. *Heart Rhythm* (2011) 8:1915–22. doi:10.1016/j.hrthm.2011.07.021

35. Basso C, Corrado D, Marcus FI, Nava A, Thiene G. Arrhythmogenic right ventricular cardiomyopathy. *Lancet* (2009) 373:1289–300. doi:10.1016/S0140-6736(09)60256-7

36. James CA, Bhonsale A, Tichnell C, Murray B, Russell SD, Tandri H, et al. Exercise increases age-related penetrance and arrhythmic risk in arrhythmogenic right ventricular dysplasia/cardiomyopathy-associated desmosomal mutation carriers. *J Am Coll Cardiol* (2013) 62:1290–7. doi:10.1016/j.jacc.2013.06.033

37. Kohli SK, Pantazis AA, Shah JS, Adeyemi B, Jackson G, McKenna WJ, et al. Diagnosis of left-ventricular non-compaction in patients with left-ventricular systolic dysfunction: time for a reappraisal of diagnostic criteria? *Eur Heart J* (2008) 29:89–95. doi:10.1093/eurheartj/ehm481

38. Bellavia D, Michelena HI, Martinez M, Pellikka PA, Bruce CJ, Connolly HM, et al. Speckle myocardial imaging modalities for early detection of myocardial impairment in isolated left ventricular non-compaction. *Heart* (2010) 96:440–7. doi:10.1136/hrt.2009.182170

39. Nucifora G, Aquaro GD, Pingitore A, Masci PG, Lombardi M. Myocardial fibrosis in isolated left ventricular non-compaction and its relation to disease severity. *Eur J Heart Fail* (2011) 13:170–6. doi:10.1093/eurjhf/hfq222

40. Melendez-Ramirez G, Castillo-Castellon F, Espinola-Zavaleta N, Meave A, Kimura-Hayama ET. Left ventricular noncompaction: a proposal of new diagnostic criteria by multidetector computed tomography. *J Cardiovasc Comput Tomogr* (2012) 6:346–54. doi:10.1016/j.jcct.2012.07.001

41. Gati S, Chandra N, Bennett RL, Reed M, Kervio G, Panoulas VF, et al. Increased left ventricular trabeculation in highly trained athletes: do we need more stringent criteria for the diagnosis of left ventricular non-compaction in athletes? *Heart* (2013) 99:401–8. doi:10.1136/heartjnl-2012-303418

42. Poscolieri B, Bianco M, Vessella T, Gervasi S, Palmieri V, Zeppilli P. Identification of benign form of ventricular non-compaction in competitive athletes by multiparametric evaluation. *Int J Cardiol* (2014) 176:1134–6. doi:10.1016/j.ijcard.2014.07.288

43. Hoedemaekers YM, Caliskan K, Michels M, Frohn-Mulder I, van der Smagt JJ, Phefferkorn JE, et al. The importance of genetic counseling, DNA diagnostics, and cardiologic family screening in left ventricular noncompaction cardiomyopathy. *Circ Cardiovasc Genet* (2010) 3:232–9. doi:10.1161/CIRCGENETICS.109.903898

44. McNally EM, Barefield DY, Puckelwartz MJ. The genetic landscape of cardiomyopathy and its role in heart failure. *Cell Metab* (2015) 21:174–82. doi:10.1016/j.cmet.2015.01.013

45. Caselli S, Attenhofer Jost CH, Jenni R, Pelliccia A. Left ventricular noncompaction diagnosis and management relevant to pre-participation

screening of athletes. *Am J Cardiol* (2015) 116:801–8. doi:10.1016/j.amjcard.2015.05.055

46. Neiva-Sousa M, Almeida-Coelho J, Falcão-Pires I, Leite-Moreira AF. Titin mutations: the fall of Goliath. *Heart Fail Rev* (2015) 20:579–88. doi:10.1007/s10741-015-9495-6

47. Gerull B. The rapidly evolving role of titin in cardiac physiology and cardiomyopathy. *Can J Cardiol* (2015) 31:1351–9. doi:10.1016/j.cjca.2015.08.016

48. Chauveau C, Rowell J, Ferreiro A. A rising titan: TTN review and mutation update. *Hum Mutat* (2014) 35:1046–59. doi:10.1002/humu.22611

49. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* (2012) 366:619–28. doi:10.1056/NEJMoa1110186

50. Mavaddat N, Peock S, Frost D, Ellis S, Platte R, Fineberg E, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. *J Natl Cancer Inst* (2013) 105:812–22. doi:10.1093/jnci/djt095

51. Semsarian C, Healey MJ, Fatkin D, Giewat M, Duffy C, Seidman CE, et al. A polymorphic modifier gene alters the hypertrophic response in a murine model of familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol* (2001) 33:2055–60. doi:10.1006/jmcc.2001.1466

52. Suzuki M, Carlson KM, Marchuk DA, Rockman HA. Genetic modifier loci affecting survival and cardiac function in murine dilated cardiomyopathy. *Circulation* (2002) 105:1824–9. doi:10.1161/01.CIR.0000014926.32463.89

53. Wheeler M, Pavlovic A, DeGoma E, Salisbury H, Brown C, Ashley EA. A new era in clinical genetic testing for hypertrophic cardiomyopathy. *J Cardiovasc Transl Res* (2009) 2:381–91. doi:10.1007/s12265-009-9139-0

54. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17:405–24. doi:10.1038/gim.2015.30

55. Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, et al. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. *Eur J Hum Genet* (2013) 21:918–28. doi:10.1038/ejhg.2012.283

56. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet* (2012) 5:391–9. doi:10.1161/CIRCGENETICS.112.962928

57. Pan S, Caleshu CA, Dunn KE, Foti MJ, Moran MK, Soyinka O, et al. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. *Circ Cardiovasc Genet* (2012) 5:602–10. doi:10.1161/CIRCGENETICS.112.963421

The Current Landscape of Genetic Testing in Cardiovascular Malformations: Opportunities and Challenges

Benjamin J. Landis[1,2]* and Stephanie M. Ware[1,2]

[1] Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, USA, [2] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Human cardiovascular malformations (CVMs) frequently have a genetic contribution. Through the application of novel technologies, such as next-generation sequencing, DNA sequence variants associated with CVMs are being identified at a rapid pace. While clinicians are now able to offer testing with NGS gene panels or whole exome sequencing to any patient with a CVM, the interpretation of genetic variation remains problematic. Variable phenotypic expression, reduced penetrance, inconsistent phenotyping methods, and the lack of high-throughput functional testing of variants contribute to these challenges. This article elaborates critical issues that impact the decision to broadly implement clinical molecular genetic testing in CVMs. Major benefits of testing include establishing a genetic diagnosis, facilitating cost-effective screening of family members who may have subclinical disease, predicting recurrence risk in offsprings, enabling early diagnosis and anticipatory management of CV and non-CV disease phenotypes, predicting long-term outcomes, and facilitating the development of novel therapies aimed at disease improvement or prevention. Limitations include financial cost, psychosocial cost, and ambiguity of interpretation of results. Multiplex families and patients with syndromic features are two groups where disease causation could potentially be firmly established. However, these account for the minority of the overall CVM population, and there is increasing recognition that genotypes previously associated with syndromes also exist in patients who lack non-CV findings. In all circumstances, ongoing dialog between cardiologists and clinical geneticists will be needed to accurately interpret genetic testing and improve these patients' health. This may be most effectively implemented by the creation and support of CV genetics services at centers committed to pursuing testing for patients.

Keywords: genetics, congenital heart disease, phenotyping, next-generation sequencing, phenomics, genomics, mutation

Abbreviations: CMA, chromosomal microarray analysis; CNV, copy number variant; CVM, cardiovascular malformation; HPO, Human Phenotype Ontology; IPCCC, International Pediatric and Congenital Cardiac Code; NGS, next-generation sequencing; VUS, variant of uncertain significance; WES, whole exome sequencing.

INTRODUCTION

Cardiovascular malformations (CVMs) are the most common birth defects with an incidence estimated at approximately 8/1000 live births (1). Taking into account very high rates of CVMs in spontaneous abortuses, common malformations, such as BAV [present in 1.3% of the population (2)], and latent cardiac diseases, such as aortic dilation, which are not included in the birth incidence of CVMs, genetically mediated CVMs are likely much more common than previously thought. When considering the etiology of CVMs, as opposed to the proportion of CVM cases that manifest as disease at birth, the incidence increases to approximately 5% (1). The common nature of these birth defects, combined with their heterogeneous etiologies, makes genetic evaluation both important and complex.

The underlying causes of CVMs are varied and can include cytogenetic abnormalities, single gene disorders, epigenetic alterations, environmental etiologies, or most commonly, multifactorial etiologies. Chromosomal abnormalities account for 12–14% of all live-born cases and 20–33% of fetal cases (1, 3–5). CVMs can occur as isolated findings, as part of a well-defined syndrome, or in conjunction with additional extracardiac anomalies not formally recognized as a syndrome (6). The American Heart Association has summarized reasons for establishing a genetic diagnosis for cardiac conditions (7). The benefits of a genetic diagnosis include improved longitudinal and acute medical management (8). In addition, a genetic diagnosis allows for the provision of anticipatory guidance, risk stratification for family members, and recurrence risk information (7). Despite an increasing awareness of the genetic basis of CVMs and the clinical importance of making an accurate diagnosis, there remain many questions about the best approach to clinical application of molecular or cytogenetic testing in individuals with a CVM.

Recently, we summarized the overall progress in the molecular genetic analyses of CVMs and current recommendations for clinical application of genetic testing (9). In particular, we reviewed the utility and limitations of chromosomal microarray analysis (CMA) and the emerging clinical roles for whole exome sequencing (WES) and other NGS technologies for CVMs. Here, we focus on the opportunities and challenges of clinical NGS testing and highlight the importance of phenotyping to improve clinical genetic testing interpretation and to drive etiology-centered research. NGS technologies generate abundant amounts of precise human genetic data, but imprecise phenotype data limit the power to determine genotype–phenotype correlation (10). We propose that deep phenotyping of CVMs and existing phenomic analysis methods provide major opportunities for progress analogous to the recently realized efforts in genomics and developmental biology. The integration of genetic findings with deep phenotyping will improve our understanding of disease etiology and advance medical care.

EPIDEMIOLOGY OF CVMs

Cardiovascular malformations represent the single largest cause of infant mortality resulting from birth defects (4). Approximately 25% of infants with CVMs are thought to have syndromic conditions based on the findings of multiple congenital anomalies or neurodevelopmental delays (11). The distinction between syndromic and non-syndromic, or isolated, CVMs can be subtle, and criteria to differentiate these categories are inconsistent between studies. In addition, as genetic diagnostic modalities have become more sophisticated, the spectrum of genetic syndromic conditions has expanded, and therefore earlier assessment of syndromic cases may represent an underestimate.

The high heritability of CVMs provides evidence for an important genetic role in these birth defects. Specific CVMs show strong familial clustering in first-degree relatives, ranging from 3- to 80-fold compared to the prevalence in the population (12). Heritability for some types of CVMs is as high as 70–90%, indicating the strong genetic contribution (13–15). Not all families show evidence of similar types of CVMs, and familial clustering of discordant CVMs has also been documented (16). Because CVMs are so common, the majority of cases occur in individuals without a family history of CVMs despite a high heritability. The prevalence of familial CVM will likely increase as more patients with CVMs survive into adulthood. Epidemiologic studies may underestimate the number of familial cases due to the high rate of miscarriages of fetuses with CVMs and reproductive decisions to limit future pregnancies in families with a child with a CVM.

The sibling or offspring recurrence risk across all types of CVMs is estimated at 1–4%. This empiric recurrence risk suggests that the majority of CVMs have a multifactorial etiology (17, 18). These estimates represent an average of different risks across the population and include individuals with higher recurrence risks due to Mendelian inheritance as well as individuals with lower risks due to a *de novo* event in the affected individual or a teratogenic etiology. Empiric recurrence risks for specific types of CVMs, such as left ventricular outflow tract obstructive defects, are higher. While the incidence of CVMs appear to be similar in most populations, there are some specific types of CVM that show important differences (14, 19, 20). In addition, there is an increased rate of CVMs in populations with increased consanguinity, often attributed to autosomal recessive mutations in disease genes (21–25). Family history of CVMs is one of the most consistently identified risk factors for identifying a CVM prenatally.

THE GENETIC BASIS OF CVMs

Cardiovascular malformations can be subdivided into syndromic and non-syndromic cases. Aneuploidies (disorders of chromosome number) are frequent causes of syndromic CVMs. As genetic testing technologies have evolved, CMA has emerged as a test with higher resolution and increased sensitivity over routine chromosome analysis (i.e., karyotype) for detecting abnormalities. Submicroscopic chromosome deletions and duplications [also known as copy number variants (CNVs)] underlie many genetic syndromes, and the term genomic disorder is used to refer to these conditions. Gene dosage is an important concept underlying CVMs. For many genes, a missing (deletion) or extra (duplication) copy of that gene results in no phenotypic

consequences. In contrast, dosage-sensitive genes produce abnormal phenotypes in the absence of two functional genes. 22q11.2 deletion syndrome and Williams–Beuren syndrome are two examples of genomic disorders that are commonly associated with CVMs related to dosage-sensitive genes (*TBX1* and *ELN*, respectively). Variants within *TBX1* and *ELN* are associated with CVMs in non-syndromic patients (26, 27). This fact illustrates an important principle: understanding the genetic basis for syndromic CVMs can identify genes responsible for non-syndromic isolated CVMs.

Because of the increased yield with CMA, it should be the first-line test for genetic analysis in infants with CVMs except in cases that are classic aneuploidies (9). CMA is considered standard of care testing for individuals with developmental disability or multiple congenital anomalies, and it has been shown to be cost-effective (28, 29). Importantly, CNVs have emerged as important causes of both syndromic and non-syndromic CVMs, occurring in approximately 3–25% of syndromic cases and 3–10% of non-syndromic cases (6, 30).

In addition to aneuploidies, chromosome rearrangements, and CNVs as causes of CVMs, mutations at the nucleotide level are also important genetic causes. These mutations are often inherited in a Mendelian fashion, and autosomal dominant, autosomal recessive, and *X*-linked inheritance patterns have been documented for both syndromic and non-syndromic CVMs (31–33). For dominantly inherited conditions, such as Noonan or Holt–Oram syndromes, individual recurrence risks for offspring with the syndrome is 50%. Importantly, not all patients with a particular syndrome have associated heart defects, and the proportion can vary by syndrome. Furthermore, the presence or severity of a CVM in the parent does not predict the severity in the child.

The genetic architecture of CVMs suggests that a majority of non-syndromic cases result from multifactorial causes and behave as a complex trait. Similar to other conditions inherited as a complex trait, isolated CVMs may show familial clustering with reduced penetrance. Nevertheless, Mendelian inheritance does occur, albeit less frequently, and *de novo* mutations are another important cause (34, 35). The distinction between monogenic and complex traits can be overly simplistic, as is drawing a distinct boundary between syndromic and non-syndromic causes. Indeed, variants in genes known to cause syndromic forms of CVMs are now identified in non-syndromic cases. In addition, traits that appear to be monogenic can be influenced by variation in multiple genes, termed modifier genes. The reverse is also true: complex traits can be predominantly influenced by variation in a single gene. These findings likely explain the decreased penetrance and variable expressivity that are so common among both syndromic and non-syndromic CVMs. Currently, there remain many unknowns about the contribution of common variants, rare variants, CNVs, *de novo* mutations, epigenetics, and environmental exposures to the development of CVMs. For these reasons, recurrence risks for apparently isolated CVMs can be difficult to assign, and even in cases of Mendelian inheritance, decreased penetrance and variable expressivity present dilemmas to predicting genetic effect on phenotype. There is need for a systematic approach to accurate and detailed phenotyping in order to begin characterizing these complexities. In addition, these factors are important considerations when contemplating molecular genetic testing in the CVM population.

In an effort to better understand genetic causes of CVMs, systems biology approaches have been used to assess functional convergence of causative CVM genes, effectively combining knowledge of genetics and developmental biology. Interestingly, these approaches have suggested that different CVM risk factors are more likely to act on distinct components of a common functional network than to directly converge on a single genetic or molecular target (36, 37). Developmental pathways acting independently or coordinately contribute to heart development and have been the subject of recent reviews (38, 39). These pathways often exhibit extensive cross talk, and a particular signal can be antithetically regulated at different developmental time points. Systems biology suggests a highly complex milieu in which individual or multiple genetic variants could potentially act to disrupt normal heart morphogenesis. The web of interactions of signaling and transcriptional networks highlighted by these approaches hint at the possibility that some CVMs may result from additive effects of multiple low-effect susceptibility alleles. The integration of genetic analysis with developmental biology knowledge provides a powerful platform for variant interpretation and candidate gene identification, but expanded databases and prediction methods are needed. Improving the assimilation of this information with careful cardiac phenotyping from human studies represents an opportunity to advance our understanding of the etiology of specific CVMs.

SEQUENCED-BASED APPROACHES TO THE GENETICS OF CVMs

The importance of CNV analysis in both syndromic and non-syndromic CVMs has been documented (6, 9, 40–44). Genetic testing in infants with CVMs is frequently underutilized but indicated in all infants with complex CVMs, except in cases warranting syndrome-specific testing (9, 45, 46). Decisions about additional genetic testing after CMA are less straightforward. The increased sophistication of genetic testing technology provides the ability to interrogate an ever increasing array of genes to identify the molecular basis of disease. Distinguishing testing that has clinical utility is necessary, but few evidence-based guidelines exist, in part, because of difficulties with phenotyping. As a result, clinical experience is the primary criterion utilized in deciding on genetic testing, and substantial practice variation exists for CVMs.

With the development of NGS, large gene sequencing panels have become both technically feasible and cost-effective. As a result, NGS panels for CVMs are developing rapidly. For example, genetic testing for Noonan syndrome has been available for several years, with additional genes being added to NGS panels as they are identified. The current yield of testing using NGS Noonan syndrome panels in suspected cases is approximately 70–85%. As another example, testing for heterotaxy syndrome, situs inversus, and primary ciliary dyskinesia are combined into one NGS panel available from several commercial laboratories.

Several studies have also documented the utility of NGS panels in diagnostic evaluation of CVMs in non-syndromic multiplex families. Blue et al. used a custom NGS panel consisting of 57 genes known to cause CVMs to sequence 16 probands from multiplex families (47). After identifying potential disease-causing variants with the panel in probands, affected family members were tested to confirm segregation with disease. Five variants in 4 genes, *TBX5*, *TFAB2B*, *ELN*, and *NOTCH1*, were concluded to be likely disease-causing among the 16 families, giving a diagnostic yield of 31%. A similar study by Jia et al. utilized a slightly different 57 gene panel in 13 multiplex non-syndromic families (48). Altogether, 44 rare variants were identified. After bioinformatics predictions and testing for segregation in other family members, a likely disease-causing variant was established in 6 of 13 families, giving a diagnostic yield of 46%. The causative genes identified in this study (*NOTCH1*, *TBX5*, and *MYH6*) partially overlapped those of Blue et al. Finally, in a recent study using a panel of 97 genes in 78 unrelated probands with bicuspic aortic valve, 33 potential disease-causing rare variants were identified (49). However, these variants were identified in only 16 of the subjects, indicating that many carried more than one potential disease-causing variant. Because all but two variants were inherited from an unaffected family member, the clinical interpretation of the pathogenicity is difficult. Together, these cases highlight benefits and limitations of NGS panels in non-syndromic patients. First, a substantial number of rare variants will be identified even with relatively small panels. Second, diagnostic yield is high in multiplex families, especially when family members are available for follow-up testing of variant segregation with disease. However, in isolated cases, our current approaches for variant classification and functional prediction make clinical interpretation difficult. Third, careful phenotyping is critical, and distinction of syndromic versus non-syndromic isolated disease is often difficult even in multiplex families. For example, mutations in *TBX5* causes Holt–Oram syndrome, which is characterized by upper limb defects that are highly variable but thought to be completely penetrant with careful examination. In the study by Blue et al. (47), the authors note that subtle hand anomalies may have been missed because radiologic examination was not performed in either family. Finally, while segregation with disease provides strong evidence for pathogenicity of variants, the reduced penetrance of many CVMs suggests that a variant inherited from an unaffected parent does not necessarily rule out disease causation or susceptibility.

Large gene panels have the advantage of increasing the sensitivity of the test, but they also increase the likelihood of identifying variants of uncertain significance (VUS). These increase in direct proportion to the number of genes tested, increasing the complexity of the interpretation and genetic counseling. Importantly, the strength of evidence for disease causality for genes on current panels differs. Some well-established disease-causing genes have a wealth of information about variants, but genes more recently implicated in disease may have much less information available. The latter situation increases the likelihood of finding a VUS. In all cases, it is important for patients to understand that a negative genetic test result does not rule out a genetic cause. The composition of gene panels varies by testing lab. It is critical that the ordering physician understands these factors to order the most appropriate test.

Whole exome sequencing interrogates the coding regions of every gene using an NGS approach. First offered as a clinical genetic test in 2011, the clinical scenarios in which WES is utilized continue to expand. For less than twice the cost of most large targeted gene panels, WES provides sequence data for all known genes, making it comparatively cost-effective. It can be superior to targeted panels for rare syndromes with CVMs in which a genetic cause is suspected but the differential diagnosis is challenging. WES has also been shown to be effective in multiplex families with CVMs. Large, multiplex families with concordant CVMs are good candidates for identifying monogenic disease variants. In addition, recently, a large multiplex family with discordant CVMs across four generations was studied by WES followed by targeted sequencing of candidates (50). A missense variant in *MYH6* was identified in 10 of 11 affected family members and absent in 10 unaffected family members. An additional four unaffected family members also carried the variant. This study not only illustrates the utility of WES for large families but also highlights the complexity of analysis and the challenges that variable expressivity and non-penetrance pose for conclusive interpretation of causality when variants are identified.

Interpretation of causality of a rare variant in a candidate gene is theoretically simplified when the variant occurs *de novo* in the proband. In these cases, the variant is frequently interpreted as likely disease-causing. Therefore, in clinical WES, parental samples are typically requested, if available, in order to aid interpretation. The multisite research study by the Pediatric Cardiac Genomics Consortium provides insight into the frequency of *de novo* variants that are likely disease-causing in a large CVMs cohort (34). Using a trio design to study 362 non-syndromic probands with CVMs, including conotruncal defects, left ventricular outflow defects, and heterotaxy, 249 protein-altering *de novo* variants were identified. Compared with control trios, CVM probands had more *de novo* variants in genes highly expressed during cardiac development and more *de novo* variants with likely damaging effects. The variants were enriched for methylation pathways and were thought to explain approximately 10% of CVMs in the cohort. In a follow-up study of this cohort in which 1213 trios were studied, more *de novo* variants were identified in cases as compared to controls (35). Interestingly, many of these variants were identified in genes known to be important for heart development, and approximately one-third were in genes known to cause syndromic CVMs. Furthermore, there was a striking overlap of variants in genes previously associated with neurodevelopmental delay. These findings may have important clinical impact not only for guiding genetic testing but also for identifying individuals with CVMs who are at increased risk for neurodevelopmental disability and for implementing early intervention.

Limitations to WES in clinical practice include the high likelihood of identifying VUSs, the decreased depth of sequencing as compared to targeted panels, and the increased likelihood of identifying a mutation for a disease unrelated to the clinical presentation or reason for performing the genetic testing. The latter situation mandates pretest genetic counseling to discuss the possibility of secondary or incidental findings. At this time,

WES may be the test of choice for syndromic CVMs in which the syndrome is not recognized. It should be considered for both syndromic and apparently non-syndromic CVMs that are inherited in a Mendelian fashion, particularly if the differential is broad or would require multiple targeted panels to test. WES in cases of isolated, non-syndromic CVMs is more controversial due to interpretation ambiguity and financial cost of testing. However, recent data indicate that the incidence of disease-causing *de novo* mutations is high and should prompt consideration of WES especially when parents are available for testing (34, 35).

THE IMPORTANCE OF PHENOTYPING

As high-throughput technologies, such as NGS, have developed and spread, the volume of genetic data available in clinical and research databases has amassed very quickly. These molecular data are mostly considered to be highly accurate. Accordingly, it is critical that equally accurate phenotype information be used for interpretation of genetic variants. However, the progress in molecular and bioinformatics techniques has vastly outpaced methods to collect and organize detailed and accurate phenotype data across the spectrum of human health and disease (51). Unfortunately, phenotype information associated with genetic diagnoses has not historically been collected and/or reported in a consistent manner. Thus, there is now a pressing need to improve phenotyping practices. The field of phenomics has emerged to address this need, consisting of (1) detailed and accurate phenotype data collection, termed deep phenotyping and (2) computational phenomic analysis (10, 52, 53). With the proper motivation and resources, there is a tortuous but passable route to implement a deep phenotyping approach for clinical testing and etiologic research of CVMs. The following sections describe the current status of phenotype data collection and analysis across the spectrum of human disease, review the current phenotype classification systems commonly used in the clinical care and research of CVMs, highlight the current phenotyping challenges in clinical CVM genetic testing, and emphasize the critical need to harmonize existing phenotype data to advance the field.

DATABASE APPROACHES TO DEEP PHENOTYPING

Deep phenotyping has been defined as "the precise and comprehensive analysis of phenotypic abnormalities in which the individual components of the phenotype are observed and described" (10). Because there are virtually unlimited ways to describe the phenotype of a patient in the clinical setting, there needs to be a constrained language or set of phenotype definitions to apply systematically in order to analyze differences and similarities between patients. An example of this problem of phenotype unboundedness exists in the Online Mendelian Inheritance in Man database (http://omim.org/), where manually curated phenotype data are highly detailed but unconstrained (54). An ontology is one approach used to organize phenotype data into a structure that is robust for computational analysis. An ontology consists of a set of definitions (or terms) that are assembled as

a directed acyclic graph. A number of biomedical ontologies have been developed, including the Gene Ontology, Disease Ontology, Mammalian Phenotype Ontology, and the Human Phenotype Ontology (HPO) (55–58). The HPO is a manually curated ontology that was first developed in 2007 and has since grown to include more than 10,000 terms (each term represents a phenotype definition) (58). The HPO is hierarchically ordered so that the terms at the highest level of the graph consist of the broadest phenotypes. Each term is subdivided into more specific subclass phenotypes until reaching the lowest tier consisting of the most detailed and specific phenotypes. In the HPO, a phenotype term "points" (as a unidirectional edge) to each of its phenotype superclass terms.

In recent years, the HPO has become a heavily used system for phenotyping in the field of human genetics. For instance, the International Standards for Cytogenomic Arrays Consortium was among the first large-scale genotype–phenotype initiatives to adopt the HPO system and demonstrate effectiveness (59). This consortium subsequently became the basis for the Clinical Genome Resource (ClinGen), sponsored by the National Institutes of Health. ClinGen aims to facilitate and establish standards for large collaborative efforts to make genotype–phenotype discoveries and implement these discoveries clinically (60). ClinGen utilizes a public database, ClinVar, as the primary repository of variant and phenotype data. The data are compiled from diverse sources, including domain-specific databases, clinical and research molecular laboratories, clinical providers, and others (61). Similar to the International Standards for Cytogenomic Arrays Consortium, ClinVar utilizes the HPO to define phenotypes and structure data. A number of other databases containing genotype–phenotype data, such as DECIPHER and PhenomeCentral, also utilize the HPO (62, 63).

Whereas the usage of the HPO has increased among genetics providers and investigators, there are many alternative phenotype classification systems in practice. Most of these systems, such as the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10), are not designed for the purpose of genetic discovery. Thus, in order to explore genotype–phenotype relationships leveraging separate data sets that potentially contain valuable phenotype information, it is necessary to cross-link systems by mapping phenotypes. These mappings have been created for a number of data sets, but harmonizing databases with different language definitions and structures presents significant challenges and limitations (54). At the very least, the HPO illustrates the motivation and value of establishing a standardized language for deep phenotyping. Importantly, there are user-friendly software applications, such as Phenotips, that enable HPO format data entry (64). Thus, the HPO represents a promising system for phenotyping CVMs in clinical and research settings that would align with many major genotype–phenotype efforts across human disease.

CVM NOMENCLATURE AND CLASSIFICATION METHODS

Abbott published the first classification method of CVM phenotypes in the *Atlas of Congenital Cardiac Disease* in 1936

(65). Since then, the precision and accuracy of diagnostic modalities, such as echocardiography and cardiac magnetic resonance imaging, as well as understanding of the morphological and molecular aspects of CV development, have advanced significantly. This has naturally led to the adoption of newer CVM phenotype nomenclatures and classification systems over time. In addition, different practical needs, such as health-care billing, clinical outcomes research, and epidemiology, have given rise to a heterogeneous set of CVM classification systems. One example is the International Pediatric and Congenital Cardiac Code (IPCCC), which was created in 2005 by the International Nomenclature Committee for Pediatric and Congenital Heart Disease. This group includes experts in the fields of pediatric cardiology and cardiothoracic surgery (66). A stated goal of the IPCCC is to facilitate clinical outcomes research across medical centers. Another important CVM phenotyping system, The Fyler Classification System, was created at Boston Children's Hospital to facilitate local outcomes research and enhance inter-provider clinical communication (67). Since its creation roughly five decades ago, the "Fyler codes" have been expanded and are mapped to the IPCCC. Another frequently used classification system for outcomes research is implemented in the Society of Thoracic Surgeons' National Congenital Heart Surgery Database (68). Finally, the Botto classification system was developed and tested using data from the National Birth Defects Prevention Study with the principal aim to investigate the etiology of CVMs using epidemiological data (69). Unique among other classification systems, the Botto system emphasized morphological and developmental concepts by grouping individual CVMs into three hierarchical levels. With the recognition that different hypotheses and statistical approaches may require minimum group sizes to achieve adequate power, groupings were also partly based on the known frequency and distribution of individual CVM phenotypes (70). Taken together, there is a long precedent for organizing CVM phenotypes but lack of a consensus nomenclature system. Clinically, this lack of consensus creates barriers in communication with non-cardiac specialists and hinders the attempt to establish genotype–phenotype correlations. In research studies, this lack of consensus creates difficulties comparing results between studies utilizing different classification schemes.

OBSTACLES TO DEEP PHENOTYPING IN CVMs

The lack of a consensus nomenclature system for classifying CVMs can lead to significant confusion and obstacles to the goal of identifying genotype–phenotype relationships. Not only is there a lack of consensus on methods to group CVMs but also many CVMs have synonymous definitions that vary in use across, and in some cases within, clinical programs. For instance, there are at least six synonyms for the diagnosis of perimembranous ventricular septal defect, including infracristal, conoventricular, membranous, paramembranous, and type 1 (71, 72). Furthermore, not all CVMs encountered in patients will fit cleanly into commonly used CVM definitions. For example,

a ventricular septal defect is typically defined by the anatomic location of the defect within the ventricular septum, but a defect extending across these anatomic boundaries is not uncommon (e.g., a perimembranous ventricular septal defect that extends into the muscular or inlet portions of the septum). These variations may be developmentally significant. Furthermore, a number of CVMs have morphological subtypes that are classified much differently between systems, such as the Collett and Edwards versus Van Praagh systems, to define subtypes of persistent truncus arteriosus (73, 74). Distinguishing CVM subtypes may improve detection of single gene defects with NGS panels or WES. Taken together, the heterogeneity in routine clinical definition of CVMs is a major impetus for genotype–phenotype databases to utilize a controlled vocabulary structured to manage these intricacies.

There are additional obstacles to consider when organizing data or reports from different clinical programs. For instance, the standard operating procedures in pediatric cardiology imaging laboratories are not uniform across programs despite established recommendations (75). The level of detail provided in clinical reports, such as echocardiography findings, can be variable, and nomenclatures are variable between report-generating software (66). Many aspects of cardiac imaging interpretation depend upon qualitative judgment and experience of the reader, and diagnoses may change or resolve as the patient ages or follow-up studies are performed. Even in circumstances where quantification is feasible, technically standardized, and clinically useful, such as measuring anatomic dimensions, there may be a lack of consensus normative reference databases of healthy children (76). For example, there at least five published normal data sets for calculating z-scores of aortic diameters in children (77). Our experience is that calculated z-scores range widely depending on the normal data set selected.

With the goal of deep phenotyping in mind, a complete study that includes documentation of negative findings is key to fully defining the patient's phenotype. However, this may require multiple studies or even different imaging modalities. For instance, in some cases, it can be difficult to absolutely rule in or rule out the presence of extracardiac vascular anomalies, such as abnormal aortic arch sidedness or persistent left superior vena cava with transthoracic echocardiography. Subtle anomalies of coronary artery branching are very difficult to characterize with echocardiography and may not be rigorously interrogated if considered clinically insignificant. Whether these types of subclinical data would advance the understanding of genetics and developmental mechanisms is not known but is quite possible. Additionally, the patient's age at the time of study may impact not only the technical quality but also the actual diagnoses. Cardiovascular hemodynamics begin to change from the time of birth, cardiac morphology may change as the child grows, and a complete diagnosis may not be reached until normal physiological events, such as closure of the ductus arteriosus. In spite of all of these potential confounders and challenges, the fact that the clinical care of patients is absolutely dependent on accurately characterizing the patient's phenotype promises to facilitate the implementation of deep phenotyping of CVMs.

MAXIMIZING THE OPPORTUNITIES FOR GENOTYPE–PHENOTYPE CORRELATIONS

In the field of genetics, there has been important progress in the analysis of phenotype data using computational techniques, sometimes referred to as phenomic analysis. Most phenomic analysis to date has consisted of algorithms used to prioritize lists of candidate disease-causing genes based on phenotype data. Gene prioritization algorithms are useful for interpreting variants identified with NGS techniques, such as clinical WES. The premise for these phenotype-based algorithms is to utilize "semantic similarity," or the mathematical similarity between a given individual's phenotype and the phenotypes of reference disease populations, such as those with established genetic disorders. This similarity measure can then be used as the score for prioritizing which variants are most likely to contribute to the individual's phenotype. Some prediction techniques exclusively utilize phenotype similarity algorithms (78, 79). Alternatively, phenotype-based scores are one component of multidimensional variant prioritization applications that combine algorithms using multiple features, such as the predicted effect of a variant on protein function (80). Variant prioritization applications that incorporate human phenotype data in this manner include Phevor, Phen-Gen, and Exomiser (81–83). There is evidence that incorporation of structured human phenotype data does improve performance (80). Importantly, computational algorithms based on semantic similarity to compare phenotypes across species have also been implemented in applications, such as Exomiser. There is ongoing work to advance phenotype-based computational methods. The accuracy of these methods is likely to improve as more deep phenotyping data are generated and shared.

With the goal of discovering genotype–phenotype relationships for CVMs, the National Heart, Lung, and Blood Institute's Bench to Bassinet program has generated an unprecedented volume of exome data for patients with CVMs, which have led to major advances toward defining the genetic basis of CVMs (34, 35, 84, 85). This study used a phenotype nomenclature system based on the IPCCC (85). Meanwhile, a large-scale forward genetic screening approach using chemical mutagenesis in mice recently led to novel insights to the mechanisms driving abnormal cardiovascular development (86). Critically, this study undertook a detailed phenotyping approach using fetal echocardiography, postmortem 3D imaging, and histopathological evaluation of unprecedented scale. To illustrate the study's scope, over 80,000 mouse fetuses were scanned with fetal echocardiography, and over 200 mutant lines with CVMs were identified. The CVMs were classified according to the Mammalian Phenotype Ontology system but were also mapped to human phenotypes using the Fyler codes. The genetic and phenotype data generated from these two large-scale studies present seemingly unbounded opportunities for computational analyses. These include the opportunity to integrate cross-species phenotype data, which will have a key role in advancing understanding of disease pathogenesis (87). These data sets potentially represent the foundation onto which clinical genetic testing data and data from other research enterprises can be added using a uniform phenotyping language. There is the opportunity for the field of CV genetics to harmonize phenotype

data with emerging standards used by large genotype–phenotype data sets within the broader field of genomics by mapping to the HPO. Given strong evidence that the genetic basis of non-syndromic CVMs overlaps with neurodevelopmental and other non-cardiac anomalies (35), the integration with other domain-specific genotype–phenotype data sets are likely to produce significant results.

At present, there are clear challenges to implementing the practices of phenomics into routine clinical interpretation of variants and genotype–phenotype research. Some of these challenges are ubiquitous, but others are unique to CVM phenotyping. Most are practical challenges that can be overcome through the efforts of highly motivated clinical and research programs. There is a clear need to adopt a standardized domain-specific CVM nomenclature where individual phenotypes are defined for every patient. Until a uniform nomenclature is adopted, phenotypes will have to be mapped between databases, which pose the risk for error and misclassification (88). On a clinical basis, the established variant databases, such as ClinVar, represent a great opportunity to begin to systematically adopt the reporting of deep phenotyping data. Of equal importance, molecular laboratories should start to require that detailed CVM phenotype data accompany genetic testing requests, which will help force improved clinical practices. These processes will be facilitated if caregivers treating patients with CVMs standardize clinical reporting practices in a manner that is both clinically practical and robust for data analysis. Harmonizing phenotype data across species will facilitate new discoveries. The development of high-throughput, quantitative methods for CVM phenotyping, such as automated digital analysis of imaging data, akin to facial image analysis, may speed discovery by breaking the bottleneck created by the highly specialized, labor-intensive nature of clinical CVM phenotyping (52, 89). While the resources required to advance CVM phenotyping are significant, these will be well worth the added investment to maximize the utility of currently funded genotyping projects. Of equal importance, the clinical interpretation of genetic testing will be improved with deep CVM phenotyping.

INTERPRETATION OF GENETIC TESTING

The tremendous effort in genomic and phenomic research has a direct effect on clinical testing. Clinical genetic testing moves rapidly to incorporate the most recent research results that have clinical utility and aid patient diagnosis or management. However, because this is an area of rapid accumulation of new data, clinical genetic testing results are not always straightforward since they represent a probability of causing or contributing to disease (90). There are two stages of interpretation of clinical genetic testing results. The clinical laboratory performs the first stage. Variants are classified, compared with ethnic and race-specific information in databases, analyzed using bioinformatic prediction programs, and classified into one of five categories: (1) benign, (2) likely benign, (3) VUS, (4) likely pathogenic, or (5) pathogenic (91). New guidelines have standardized and increased the stringency of interpretation, with more clear criteria for strength of evidence required for interpretation (91). Nevertheless, the interpretations provided for a given variant may differ between clinical genetic

testing laboratories. In addition, updates and revisions of the laboratory interpretation may occur as more information is obtained from larger cohorts. For this reason, families should also maintain a relationship with the CV genetics providers, as VUSs often get reclassified over time. A second stage is the interpretation provided by the clinician. Molecular testing results should be one piece of evidence in a diagnostic evaluation. These results need to be interpreted in the context of the patient's medical history, physical exam findings, disease course, and family history to arrive at a diagnosis. Family history information and the segregation of a potential disease-causing variant within the family may be important information to guide the clinical interpretation of the genetic testing results, especially in cases where novel genetic variants are identified. For CVMs, in which Mendelian inheritance may not be seen or decreased penetrance may make segregation with disease difficult to establish, there are increased challenges to the interpretation of genetic testing results.

A CVM genetic testing workflow begins with the ascertainment of high-quality deep phenotype data (**Figure 1**). The genetic testing laboratory can improve their interpretation of genetic data when provided with clear phenotype information. The diagnostic interpretation of the clinical care team, longitudinal

follow-up and outcome, and family-based clinical information and genetic testing results are all used by the testing laboratory to refine interpretation. Communicating the patient's phenotype to the testing laboratory or clinical databases, such as ClinVar, is a critical step that is highly susceptible to errors, such as misclassifications or omissions. How can the genetics provider who orders genetic testing communicate the CVM phenotype accurately? The accuracy and completeness of the diagnosis may depend on the sources of the information, which include clinical notes, imaging study reports, procedure notes, or administrative diagnostic codes. The optimal source of this information likely depends on factors specific to patient and medical system. In order to minimize the errors, ideally the genetics provider must have access to all pertinent information (e.g., echocardiography reports, operative reports, cardiac catheterization reports), have sufficient background understanding and experience in CVM diagnoses to accurately define the patient's CVM phenotype, and have a cardiologist readily available when clarifications are needed. While this process may be effectively conducted by a team of investigators devoted to a specific research project, undoubtedly in most pediatric cardiology centers, there are immense practical challenges to clinically implementing the above
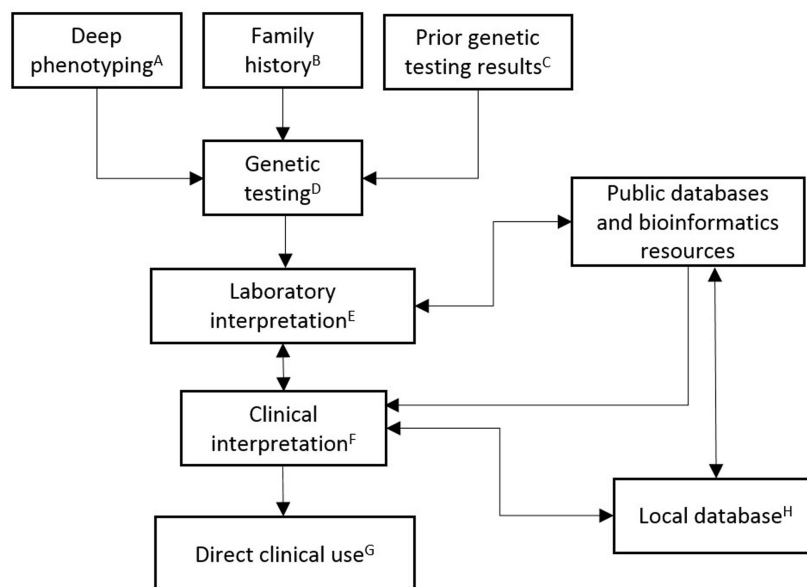
FIGURE 1 | **Schematic outlining the clinical implementation of CVM genetic testing**. **(A)** Deep phenotyping data include complete CVM diagnoses, congenital non-CV malformations, dysmorphic exam findings, neurodevelopmental abnormalities, and other pertinent medical history. Specification of relevant negative findings, including radiographic studies (e.g., head/renal ultrasound), neurodevelopmental evaluation, and specific cardiac evaluations, is important for robust datasets. Age that diagnoses were established or ruled out should be included. Phenotype data should be collected in a structured format (e.g., HPO). **(B)** Family history data are input as a three-generation pedigree, including documentation of relatives with negative cardiac screening. **(C)** Prior genetic testing results data include dates and testing laboratory. **(D)** Genetic testing decisions are patient, family, and disease differential specific. Current clinically available testing options include single gene (e.g., sequencing or deletion/duplication testing), multiple gene (e.g., NGS panels), or genome-wide (e.g., chromosomal analysis, CMA, or whole exome sequencing) testing. **(E)** Laboratory interpretation of genetic testing is based upon ACMG guidelines. High-quality patient data should be provided with the orders for genetic testing. **(F)** Clinical interpretation of genetic testing combines multidisciplinary CV genetics knowledge/expertise with the laboratory interpretation. **(G)** Direct clinical use includes providing results and counseling to family, reporting to health-care providers, recommending treatment, making appropriate subspecialty referrals, making appropriate plan for longitudinal monitoring, and instituting cascade genetic testing and/or family-based cardiac imaging as indicated. **(H)** Local database compiles high-quality phenotype and genotype data for multiple uses, including longitudinal follow up (e.g., completion of cardiac screening in family members or reassessment of variant interpretation), documentation of clinical practices and outcomes, and periodic data harvests for dissemination to public databases and peer-reviewed publication.

scenario for every patient undergoing genetic testing. However, a multidisciplinary CV genetics program consisting of geneticists, cardiologists, genetic counselors, and molecular biologists, which fosters cross-disciplinary education and communication, is actually well suited to meet these needs. These collaborative groups of professionals improve the accuracy of the probabilistic genetic testing information and provide more expertise to the diagnosis and management of the patient.

There remain great opportunities for improving our ability to interpret the results of genetic variation and predicting impact. These are important priorities in all clinical fields that incorporate genetic testing into the diagnosis and management of patients. In the future, identification of genetic modifiers that contribute to phenotypic presentation and explain a portion of the variability and reduced penetrance in these disorders is necessary. This focus will need to include an improvement in our understanding of the impact of rare genetic variation in the population as well as the functional significance of common polymorphisms.

SUMMARY

In conclusion, there is strong evidence to support CMA testing as a first-line genetic test for infants with clinically significant CVMs. Molecular genetic testing with NGS panels is useful for the evaluation of CVM patients in whom a specific genetic syndrome is suspected. In cases where genetic conditions are highly suspected but a specific syndrome is not recognized, WES may be indicated. NGS panels or WES may be diagnostic in multiplex families with CVMs. Data supporting the potential utility of expanded NGS CVM-gene panels or WES in isolated non-syndromic CVM patients are accumulating, but clinical sensitivity is currently unknown and conclusive variant interpretation remains

problematic. Systems biology provides evidence that many CVM genes functionally converge on signaling and transcriptional pathways. Given these considerations, WES or whole genome sequencing will likely ultimately replace NGS panels. However, broader testing will result in ambiguous variant interpretation in CVM patients due in part to variable and expression and reduced penetrance. Incomplete phenotype information and lack of standardized methods for phenotyping also remain significant obstacles. Collaboration between genetics and cardiac care providers and molecular testing laboratories is needed to optimize variant interpretation. There are currently major opportunities to integrate and analyze molecular and phenotype data from human and animal research projects to advance our understanding of the cause of CVMs.

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

1. Hoffman JI, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol* (2002) 39(12):1890–900. doi:10.1016/S0735-1097(02)01886-7
2. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, et al. Heart disease and stroke statistics – 2011 update: a report from the American Heart Association. *Circulation* (2011) 123(4):e18–209. doi:10.1161/CIR.0b013e3182009701
3. Centers for Disease Control and Prevention (CDC). Hospital stays, hospital charges, and in-hospital deaths among infants with selected birth defects – United States, 2003. *MMWR Morb Mortal Wkly Rep* (2007) 56(2):25–9.
4. Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Erickson JD. Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979-1997. *Circulation* (2001) 103(19):2376–81. doi:10.1161/01.CIR.103.19.2376
5. Hoffman JI. Incidence of congenital heart disease: II. Prenatal incidence. *Pediatr Cardiol* (1995) 16(4):155–65. doi:10.1007/BF00801907
6. Lander J, Ware SM. Copy number variation in congenital heart defects. *Curr Genet Med Rep* (2014) 2(3):168–78. doi:10.1007/s40142-014-0049-3
7. Pierpont ME, Basson CT, Benson DW Jr, Gelb BD, Giglia TM, Goldmuntz E, et al. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* (2007) 115(23):3015–38. doi:10.1161/CIRCULATIONAHA.106.183056
8. Landis BJ, Cooper DS, Hinton RB. CHD associated with syndromic diagnoses: peri-operative risk factors and early outcomes. *Cardiol Young* (2016) 26(1):30–52. doi:10.1017/S1047951115001389
9. Cowan JR, Ware SM. Genetics and genetic testing in congenital heart disease. *Clin Perinatol* (2015) 42(2):373–93,ix. doi:10.1016/j.clp.2015.02.009
10. Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat* (2012) 33(5):777–80. doi:10.1002/humu.22080
11. Ferencz C, Loffredo CA, Corea-Villasenor A, Wilson PD. *Genetic and Environmental Risk Factors for Major Cardiovascular Malformations: The Baltimore-Washington Infant Study 1981-1989.* Armonk, NY: Futura Publishing Co., Inc (1997). 463 p.
12. Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. Recurrence of congenital heart defects in families. *Circulation* (2009) 120(4):295–301. doi:10.1161/CIRCULATIONAHA.109.857987
13. Cripe L, Andelfinger G, Martin LJ, Shooner K, Benson DW. Bicuspid aortic valve is heritable. *J Am Coll Cardiol* (2004) 44(1):138–43. doi:10.1016/j.jacc.2004.03.050
14. McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, et al. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: segregation, multiplex relative risk, and heritability. *Am J Med Genet A* (2005) 134a(2):180–6. doi:10.1002/ajmg.a.30602
15. Hinton RB Jr, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, Benson DW. Hypoplastic left heart syndrome is heritable. *J Am Coll Cardiol* (2007) 50(16):1590–5. doi:10.1016/j.jacc.2007.07.021
16. Oyen N, Poulsen G, Wohlfahrt J, Boyd HA, Jensen PK, Melbye M. Recurrence of discordant congenital heart defects in families. *Circ Cardiovasc Genet* (2010) 3(2):122–8. doi:10.1161/CIRCGENETICS.109.890103
17. Burn J, Brennan P, Little J, Holloway S, Coffey R, Somerville J, et al. Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* (1998) 351(9099):311–6. doi:10.1016/S0140-6736(97)06486-6

18. Gill HK, Splitt M, Sharland GK, Simpson JM. Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J Am Coll Cardiol* (2003) 42(5):923–9. doi:10.1016/S0735-1097(03)00853-2

19. Canfield MA, Honein MA, Yuskiv N, Xing J, Mai CT, Collins JS, et al. National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001. *Birth Defects Res A Clin Mol Teratol* (2006) 76(11):747–56. doi:10.1002/bdra.20294

20. Pradat P, Francannet C, Harris JA, Robert E. The epidemiology of cardiovascular defects, part I: a study based on data from three large registries of congenital malformations. *Pediatr Cardiol* (2003) 24(3):195–221. doi:10.1007/s00246-002-9401-6

21. Becker S, Al Halees Z. First-cousin matings and congenital heart disease in Saudi Arabia. *Community Genet* (1999) 2(2–3):69–73.

22. Becker SM, Al Halees Z, Molina C, Paterson RM. Consanguinity and congenital heart disease in Saudi Arabia. *Am J Med Genet* (2001) 99(1):8–13. doi:10.1002/1096-8628(20010215)99:1<8::AID-AJMG1116>3.0.CO;2-U

23. Chehab G, Chedid P, Saliba Z, Bouvagnet P. Congenital cardiac disease and inbreeding: specific defects escape higher risk due to parental consanguinity. *Cardiol Young* (2007) 17(4):414–22. doi:10.1017/S1047951107000704

24. Nabulsi MM, Tamim H, Sabbagh M, Obeid MY, Yunis KA, Bitar FF. Parental consanguinity and congenital heart malformations in a developing country. *Am J Med Genet A* (2003) 116a(4):342–7. doi:10.1002/ajmg.a.10020

25. Ramegowda S, Ramachandra NB. Parental consanguinity increases congenital heart diseases in South India. *Ann Hum Biol* (2006) 33(5–6):519–28. doi:10.1080/03014460600909349

26. Griffin HR, Topf A, Glen E, Zweier C, Stuart AG, Parsons J, et al. Systematic survey of variants in TBX1 in non-syndromic tetralogy of Fallot identifies a novel 57 base pair deletion that reduces transcriptional activity but finds no evidence for association with common variants. *Heart* (2010) 96(20):1651–5. doi:10.1136/hrt.2010.200121

27. Merla G, Brunetti-Pierri N, Piccolo P, Micale L, Loviglio MN. Supravalvular aortic stenosis: elastin arteriopathy. *Circ Cardiovasc Genet* (2012) 5(6):692–6. doi:10.1161/CIRCGENETICS.112.962860

28. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* (2010) 86(5):749–64. doi:10.1016/j.ajhg.2010.04.006

29. Regier DA, Friedman JM, Marra CA. Value for money? Array genomic hybridization for diagnostic testing for genetic causes of intellectual disability. *Am J Hum Genet* (2010) 86(5):765–72. doi:10.1016/j.ajhg.2010.03.009

30. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* (2009) 41(8):931–5. doi:10.1038/ng.415

31. Ferencz C, Boughman JA, Neill CA, Brenner JI, Perry LW. Congenital cardiovascular malformations: questions on inheritance. Baltimore-Washington Infant Study Group. *J Am Coll Cardiol* (1989) 14(3):756–63. doi:10.1016/0735-1097(89)90122-8

32. Grech V, Gatt M. Syndromes and malformations associated with congenital heart disease in a population-based study. *Int J Cardiol* (1999) 68(2):151–6. doi:10.1016/S0167-5273(98)00354-4

33. Meberg A, Hals J, Thaulow E. Congenital heart defects – chromosomal anomalies, syndromes and extracardiac malformations. *Acta Paediatr* (2007) 96(8):1142–5. doi:10.1111/j.1651-2227.2007.00381.x

34. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* (2013) 498(7453):220–3. doi:10.1038/nature12141

35. Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* (2015) 350(6265):1262–6. doi:10.1126/science.aac9396

36. Lage K, Mollgard K, Greenway S, Wakimoto H, Gorham JM, Workman CT, et al. Dissecting spatio-temporal protein networks driving human heart development and related disorders. *Mol Syst Biol* (2010) 6:381. doi:10.1038/msb.2010.36

37. Lage K, Greenway SC, Rosenfeld JA, Wakimoto H, Gorham JM, Segre AV, et al. Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. *Proc Natl Acad Sci U S A* (2012) 109(35):14035–40. doi:10.1073/pnas.1210730109

38. Rana MS, Christoffels VM, Moorman AF. A molecular and genetic outline of cardiac morphogenesis. *Acta Physiol (Oxf)* (2013) 207(4):588–615. doi:10.1111/apha.12061

39. Kathiriya IS, Nora EP, Bruneau BG. Investigating the transcriptional control of cardiovascular development. *Circ Res* (2015) 116(4):700–14. doi:10.1161/CIRCRESAHA.116.302832

40. Breckpot J, Thienpont B, Peeters H, de Ravel T, Singer A, Rayyan M, et al. Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. *J Pediatr* (2010) 156(5):810–7, 7.e1–7.e4. doi:10.1016/j.jpeds.2009.11.049

41. Breckpot J, Thienpont B, Arens Y, Tranchevent LC, Vermeesch JR, Moreau Y, et al. Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenet Genome Res* (2011) 135(3–4):251–9. doi:10.1159/000331272

42. Thienpont B, Mertens L, de Ravel T, Eyskens B, Boshoff D, Maas N, et al. Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. *Eur Heart J* (2007) 28(22):2778–84. doi:10.1093/eurheartj/ehl560

43. Richards AA, Santos LJ, Nichols HA, Crider BP, Elder FF, Hauser NS, et al. Cryptic chromosomal abnormalities identified in children with congenital heart disease. *Pediatr Res* (2008) 64(4):358–63. doi:10.1203/PDR.0b013e31818095d0

44. Lalani SR, Shaw C, Wang X, Patel A, Patterson LW, Kolodziejska K, et al. Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. *Eur J Hum Genet* (2013) 21(2):173–81. doi:10.1038/ejhg.2012.155

45. Baker K, Sanchez-de-Toledo J, Munoz R, Orr R, Kiray S, Shiderly D, et al. Critical congenital heart disease – utility of routine screening for chromosomal and other extracardiac malformations. *Congenit Heart Dis* (2012) 7(2):145–50. doi:10.1111/j.1747-0803.2011.00585.x

46. Connor JA, Hinton RB, Miller EM, Sund KL, Ruschman JG, Ware SM. Genetic testing practices in infants with congenital heart disease. *Congenit Heart Dis* (2014) 9(2):158–67. doi:10.1111/chd.12112

47. Blue GM, Kirk EP, Giannoulatou E, Dunwoodie SL, Ho JW, Hilton DC, et al. Targeted next-generation sequencing identifies pathogenic variants in familial congenital heart disease. *J Am Coll Cardiol* (2014) 64(23):2498–506. doi:10.1016/j.jacc.2014.09.048

48. Jia Y, Louw JJ, Breckpot J, Callewaert B, Barrea C, Sznajer Y, et al. The diagnostic value of next generation sequencing in familial nonsyndromic congenital heart defects. *Am J Med Genet A* (2015) 167(8):1822–9. doi:10.1002/ajmg.a.37108

49. Bonachea EM, Zender G, White P, Corsmeier D, Newsom D, Fitzgerald-Butt S, et al. Use of a targeted, combinatorial next-generation sequencing approach for the study of bicuspid aortic valve. *BMC Med Genomics* (2014) 7:56. doi:10.1186/1755-8794-7-56

50. Arrington CB, Bleyl SB, Matsunami N, Bonnell GD, Otterud BE, Nielsen DC, et al. Exome analysis of a family with pleiotropic congenital heart disease. *Circ Cardiovasc Genet* (2012) 5(2):175–82. doi:10.1161/CIRCGENETICS.111.961797

51. Brookes AJ, Robinson PN. Human genotype-phenotype databases: aims, challenges and opportunities. *Nat Rev Genet* (2015) 16(12):702–15. doi:10.1038/nrg3932

52. Freimer N, Sabatti C. The human phenome project. *Nat Genet* (2003) 34(1):15–21. doi:10.1038/ng0503-15

53. Lanktree MB, Hassell RG, Lahiry P, Hegele RA. Phenomics: expanding the role of clinical evaluation in genomic studies. *J Investig Med* (2010) 58(5):700–6. doi:10.231/JIM.0b013e3181d844f7

54. Robinson PN, Kohler S, Bauer S, Seelow D, Horn D, Mundlos S. The human phenotype ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet* (2008) 83(5):610–5. doi:10.1016/j.ajhg.2008.09.017

55. Gene Ontology Consortium. Creating the gene ontology resource: design and implementation. *Genome Res* (2001) 11(8):1425–33. doi:10.1101/gr.180801

56. Schriml LM, Arze C, Nadendla S, Chang YW, Mazaitis M, Felix V, et al. Disease ontology: a backbone for disease semantic integration. *Nucleic Acids Res* (2012) 40(Database issue):D940–6. doi:10.1093/nar/gkr972

57. Smith CL, Eppig JT. The Mammalian Phenotype Ontology as a unifying standard for experimental and high-throughput phenotyping data. *Mamm Genome* (2012) 23(9–10):653–68. doi:10.1007/s00335-012-9421-3

58. Kohler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, et al. The Human Phenotype Ontology project: linking molecular biology

and disease through phenotype data. *Nucleic Acids Res* (2014) 42(Database issue):D966–74. doi:10.1093/nar/gkt1026

59. Riggs ER, Jackson L, Miller DT, Van Vooren S. Phenotypic information in genomic variant databases enhances clinical care and research: the International Standards for Cytogenomic Arrays Consortium experience. *Hum Mutat* (2012) 33(5):787–96. doi:10.1002/humu.22052

60. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen – the Clinical Genome Resource. *N Engl J Med* (2015) 372(23):2235–42. doi:10.1056/NEJMsr1406261

61. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* (2014) 42(Database issue):D980–5. doi:10.1093/nar/gkt1113

62. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am J Hum Genet* (2009) 84(4):524–33. doi:10.1016/j.ajhg.2009.03.010

63. Buske OJ, Girdea M, Dumitriu S, Gallinger B, Hartley T, Trang H, et al. PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. *Hum Mutat* (2015) 36(10):931–40. doi:10.1002/humu.22851

64. Girdea M, Dumitriu S, Fiume M, Bowdin S, Boycott KM, Chenier S, et al. PhenoTips: patient phenotyping software for clinical and research use. *Hum Mutat* (2013) 34(8):1057–65. doi:10.1002/humu.22347

65. Abbott ME. *Atlas of Congenital Cardiac Disesae.* New York: The American Heart Association (1936).

66. Franklin RC, Jacobs JP, Krogmann ON, Beland MJ, Aiello VD, Colan SD, et al. Nomenclature for congenital and paediatric cardiac disease: historical perspectives and The International Pediatric and Congenital Cardiac Code. *Cardiol Young* (2008) 18(Suppl 2):70–80. doi:10.1017/S1047951108002795

67. Newburger JW, Fyler J, Colan SD. Donald Charles Fyler: 1924-2011. *Cardiol Young* (2011) 21(5):485–9. doi:10.1017/S1047951111001089

68. Mavroudis C, Bove EL, Cameron DE, Drinkwater DC Jr, Edwards FH, Hammon JW Jr, et al. The Society of Thoracic Surgeons' National Congenital Heart Surgery Database. *Ann Thorac Surg* (1995) 59(2):554–6. doi:10.1016/0003-4975(94)00947-6

69. Botto LD, Lin AE, Riehle-Colarusso T, Malik S, Correa A; National Birth Defects Prevention Study. Seeking causes: classifying and evaluating congenital heart defects in etiologic studies. *Birth Defects Res A Clin Mol Teratol* (2007) 79(10):714–27. doi:10.1002/bdra.20403

70. Khoury MJ, Moore CA, James LM, Cordero JF. The interaction between dysmorphology and epidemiology: methodologic issues of lumping and splitting. *Teratology* (1992) 45(2):133–8. doi:10.1002/tera.1420450206

71. Spicer DE, Hsu HH, Co-Vu J, Anderson RH, Fricker FJ. Ventricular septal defect. *Orphanet J Rare Dis* (2014) 9:144. doi:10.1186/s13023-014-0144-2

72. Gelehrter S, Ensing G. Ventricular septal defects. In: Eidem BW, editor. *Echocardiography in Pediatric and Adult Congenital Heart Disease.* Philadelphia, PA: Lippincott Williams & Wilkins (2010). p. 158–75.

73. Collett RW, Edwards JE. Persistent truncus arteriosus; a classification according to anatomic types. *Surg Clin North Am* (1949) 29(4):1245–70.

74. Van Praagh R. Truncus arteriosus: what is it really and how should it be classified? *Eur J Cardiothorac Surg* (1987) 1(2):65–70. doi:10.1016/1010-7940(87)90014-5

75. Lai WW, Geva T, Shirali GS, Frommelt PC, Humes RA, Brook MM, et al. Guidelines and standards for performance of a pediatric echocardiogram: a report from the Task Force of the Pediatric Council of the American Society of Echocardiography. *J Am Soc Echocardiogr* (2006) 19(12):1413–30. doi:10.1016/j.echo.2006.09.001

76. Lopez L, Colan SD, Frommelt PC, Ensing GJ, Kendall K, Younoszai AK, et al. Recommendations for quantification methods during the performance of a pediatric echocardiogram: a report from the Pediatric Measurements Writing Group of the American Society of Echocardiography Pediatric and Congenital Heart Disease Council. *J Am Soc Echocardiogr* (2010) 23(5):465–95; quiz 576–7. doi:10.1016/j.echo.2010.03.019

77. Cantinotti M, Scalese M, Molinaro S, Murzi B, Passino C. Limitations of current echocardiographic nomograms for left ventricular, valvular, and arterial dimensions in children: a critical review. *J Am Soc Echocardiogr* (2012) 25(2):142–52. doi:10.1016/j.echo.2011.10.016

78. Zemojtel T, Kohler S, Mackenroth L, Jager M, Hecht J, Krawitz P, et al. Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Sci Transl Med* (2014) 6(252):252ra123. doi:10.1126/scitranslmed.3009262

79. Masino AJ, Dechene ET, Dulik MC, Wilkens A, Spinner NB, Krantz ID, et al. Clinical phenotype-based gene prioritization: an initial study using semantic similarity and the human phenotype ontology. *BMC Bioinformatics* (2014) 15:248. doi:10.1186/1471-2105-15-248

80. Smedley D, Robinson PN. Phenotype-driven strategies for exome prioritization of human Mendelian disease genes. *Genome Med* (2015) 7(1):81. doi:10.1186/s13073-015-0199-2

81. Singleton MV, Guthery SL, Voelkerding KV, Chen K, Kennedy B, Margraf RL, et al. Phevor combines multiple biomedical ontologies for accurate identification of disease-causing alleles in single individuals and small nuclear families. *Am J Hum Genet* (2014) 94(4):599–610. doi:10.1016/j.ajhg.2014.03.010

82. Javed A, Agrawal S, Ng PC. Phen-Gen: combining phenotype and genotype to analyze rare disorders. *Nat Methods* (2014) 11(9):935–7. doi:10.1038/nmeth.3046

83. Robinson PN, Kohler S, Oellrich A, Wang K, Mungall CJ, Lewis SE, et al. Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res* (2014) 24(2):340–8. doi:10.1101/gr.160325.113

84. Kaltman JR, Schramm C, Pearson GD. The National Heart, Lung, and Blood Institute Bench to Bassinet Program: a new paradigm for translational research. *J Am Coll Cardiol* (2010) 55(12):1262–5. doi:10.1016/j.jacc.2009.11.055

85. Gelb B, Brueckner M, Chung W, Goldmuntz E, Kaltman J, Kaski JP, et al. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. *Circ Res* (2013) 112(4):698–706. doi:10.1161/CIRCRESAHA.111.300297

86. Li Y, Klena NT, Gabriel GC, Liu X, Kim AJ, Lemke K, et al. Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature* (2015) 521(7553):520–4. doi:10.1038/nature14269

87. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: the glass half empty. *Circ Res* (2013) 112(4):707–20. doi:10.1161/CIRCRESAHA.112.300853

88. Strickland MJ, Riehle-Colarusso TJ, Jacobs JP, Reller MD, Mahle WT, Botto LD, et al. The importance of nomenclature for congenital cardiac disease: implications for research and evaluation. *Cardiol Young* (2008) 18(Suppl 2): 92–100. doi:10.1017/S1047951108002515

89. Chen W, Xia X, Huang Y, Chen X, Han JJ. Bioimaging for quantitative phenotype analysis. *Methods* (2016) 102:20–25. doi:10.1016/j.ymeth.2016.01.017

90. Ingles J, Semsarian C. Conveying a probabilistic genetic test result to families with an inherited heart disease. *Heart Rhythm* (2014) 11(6):1073–8. doi:10.1016/j.hrthm.2014.03.017

91. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17(5):405–24. doi:10.1038/gim.2015.30

Clinical Genetic Testing for the Cardiomyopathies and Arrhythmias: A Systematic Framework for Establishing Clinical Validity and Addressing Genotypic and Phenotypic Heterogeneity

John Garcia, Jackie Tahiliani, Nicole Marie Johnson, Sienna Aguilar, Daniel Beltran, Amy Daly, Emily Decker, Eden Haverfield, Blanca Herrera, Laura Murillo, Keith Nykamp and Scott Topper*

Invitae Corporation, San Francisco, CA, USA

Advances in DNA sequencing have made large, diagnostic gene panels affordable and efficient. Broad adoption of such panels has begun to deliver on the promises of personalized medicine, but has also brought new challenges such as the presence of unexpected results, or results of uncertain clinical significance. Genetic analysis of inherited cardiac conditions is particularly challenging due to the extensive genetic heterogeneity underlying cardiac phenotypes, and the overlapping, variable, and incompletely penetrant nature of their clinical presentations. The design of effective diagnostic tests and the effective use of the results depend on a clear understanding of the relationship between each gene and each considered condition. To address these issues, we developed simple, systematic approaches to three fundamental challenges: (1) evaluating the strength of the evidence suggesting that a particular condition is caused by pathogenic variants in a particular gene, (2) evaluating whether unusual genotype/phenotype observations represent a plausible expansion of clinical phenotype associated with a gene, and (3) establishing a molecular diagnostic strategy to capture overlapping clinical presentations. These approaches focus on the systematic evaluation of the pathogenicity of variants identified in clinically affected individuals, and the natural history of disease in those individuals. Here, we applied these approaches to the evaluation of more than 100 genes reported to be associated with inherited cardiomyopathies and arrhythmias including hypertrophic cardiomyopathy, dilated cardiomyopathy, arrhythmogenic right ventricular dysplasia or cardiomyopathy, long QT syndrome, short QT syndrome, Brugada, and catecholaminergic polymorphic ventricular tachycardia, and to a set of related syndromes such as Noonan Syndrome and Fabry disease. These approaches provide a framework for delivering meaningful and accurate genetic test results to individuals with hereditary cardiac conditions.

Keywords: genetic testing, cardiomyopathies, arrhythmias, ARVD/C, curation

INTRODUCTION

The dramatic reduction in the cost of DNA sequencing, combined with advances in the understanding of the genetic and phenotypic heterogeneity of cardiac conditions, has led to the adoption of large panels of genes as a cost-effective clinical tool to establish a molecular diagnosis in affected individuals. But, although the economics and the potential yield of such tests have improved, the fundamental principles of diagnostic testing – analytic validity, clinical validity, and clinical utility – have not.

Genetic diagnosis of inherited cardiac conditions is especially challenging due to the extensive genetic heterogeneity and the overlapping, variable, and incompletely penetrant nature of the clinical presentations. The design of effective diagnostic tests and the effective use of the results of such tests depend on a clear understanding of the relationship between each gene and each considered condition, and a clear understanding of the extent of the overlap in clinical presentation.

As part of an effort to develop a scalable framework for developing, launching, and supporting diagnostic genetic tests across a broad range of clinical areas, we set out to create general methods for establishing the *clinical validity* of a gene, and for designing focused and clinically useful panel tests.

Establishing the clinical validity of a multi-gene panel depends on an accurate and detailed understanding of the clinical validity of each included gene. While clinical laboratories have been implicitly making assessments regarding clinical validity for years, there has been a lack of clarity about evidentiary requirements for establishing clinical validity and, especially for rare, multigenic conditions, accessible data and methods to be applied in reaching this conclusion.

Within clinical genetic diagnostics, clinical validity "measures the accuracy with which a test identifies a person with the clinical condition in question" (1–3) and depends on such quantitative measures as sensitivity, specificity, positive predictive value and negative predictive value. While there are disagreements about quantitative thresholds for these measures, this is clear: a test that cannot return a positive result has no sensitivity and no positive predictive value, and cannot be considered clinically valid. Conceptually, clinical validity can be understood as a proven, causal connection between a gene and a human disease.

For diagnostic testing, the question of clinical validity should be thought of as only the first requirement. Once answered positively, it opens onto a series of more detailed considerations that also need evaluation. If it has been proven that the gene causes human disease (i.e., if clinical validity has been established) then one can reasonably ask: which disease(s)? How certain are we that we understand the boundaries of the phenotypic heterogeneity that can derive from this gene? What is the yield in different populations? Which specific variants are causal? How certain are we that we completely understand the molecular mechanisms? If, however, it has not yet been proven that a gene causes disease, then questions of expressivity and mechanism of disease are clouded by this more fundamental uncertainty.

This paper aims to (1) propose a method for establishing clinical validity of a gene, (2) propose a method for grouping genes together into meaningful panel tests, and (3) apply those methods to evaluate a set of cardiac genes and conditions. We also describe the kinds of specific results that may be expected from clinical testing of different classes of genes, and the appropriate use of those results in clinical care.

MATERIALS AND METHODS

When we first began curating gene–condition relationships, we established a working group to develop a framework for evaluating and documenting relevant evidence and our conclusions about that evidence. The working group consisted of lab directors, genetic counselors, and scientists with experience from a diversity of diagnostic and research labs. We first discussed and compared methods used in those environments, and quickly came to the conclusion that, while different approaches generally considered the same *types* of evidence (linkage studies, animal, cellular, and molecular models, observations of variants in affected individuals and pedigrees), there was little consensus about how to rigorously and reproducibly synthesize that evidence. In fact, in many cases that synthesis was not supported by an established method, but rather left to the professional judgment of a single individual.

As a starting point, the working group established a simple point-based framework that focused on a comprehensive cataloging of the relevant experimental and observational evidence, judging the strength of that evidence, and requiring that multiple pieces of evidence were present. We then applied that preliminary framework to the curation of a set of genes putatively involved in an array of hereditary cancers. While the framework was generally effective, we found that progress was slow, that much of the research we were doing had little long-term utility in the context of a diagnostic lab, and that there were regularly cases that led to a kind of logical conflict: genes supported by extensive and generally convincing research, where we were still unable to deliver positive results without additional clinical observations or experimental data.

We therefore re-convened the working group to reconsider the framework, with the following goals: (1) clearly define the purpose of the curation research effort, (2) identify the specific information that directly supported those aims, and (3) develop a reproducible and auditable approach to meet those aims. The group identified and discussed particular cases that led to disagreements or inconsistency when using the previous approach, and worked through a series of thought experiments to elucidate edge cases. Through this process, the logical necessity of harmonizing variant evaluation and gene clinical validity evaluation emerged. As described below, the method that was developed simply defers the general question of the clinical validity of a gene to the specific question of the pathogenicity of clinically observed variants.

With this framework in place, we then applied the method to the set of genes suggested to cause hereditary cardiac conditions. For each considered gene, we used the Human Gene Mutation Database (HGMD) to provide a list of published, clinically observed variants. The presence of a variant in the HGMD database does not necessarily indicate pathogenicity, and variants were independently researched and interpreted using an implementation of the ACMG variant interpretation guidelines. In the case where none of the variants listed in HGMD were

determined to be convincingly pathogenic, a literature search for more recent case reports of other clinically observed variants was performed.

RESULTS

Establishing the Clinical Validity of Gene–Condition Relationships
Method for Establishing the Clinical Validity of a Particular Gene

The approach described here addresses the question of distinguishing between genes that have been proven to cause human disease and genes that currently only have preliminary evidence suggesting an association. It depends on a simple insight: an accurate methodology for evaluating *variant pathogenicity* must provide results that are consistent with an accurate methodology for evaluating the *clinical validity* of the gene. If the methodologies provide inconsistent results, one of the methodologies must be delivering an incorrect conclusion. This dependency suggests that the approaches can and should be harmonized.

This dependency can be demonstrated with a logical argument:

- Clinical validity of a gene is established when we know that the gene causes human disease.
- A gene causes human disease because at least one variant in the gene causes human disease.
- Therefore, clinical validity of a gene is established when we know of at least one variant in the gene that causes human disease.
- If the variant classification schema is robust, we know that a variant causes disease only when there is sufficient evidence to classify the variant as pathogenic.
- Therefore:
  - If there is sufficient evidence to classify a variant as pathogenic, then we know that the variant causes disease, we know that the gene causes disease, and we know that a clinical test of the gene can be clinically valid.
- Conversely:
  - If there is NOT sufficient evidence to classify any clinically observed variants as pathogenic (i.e., if ALL clinically observed variants must be classified as VUS), then we do not know of a variant in the gene that has been proven to cause disease in humans, we can not be certain that the gene causes disease, and we do not know that a test of the gene would be clinically valid.

This argument reduces the question of evaluating the strength of the evidence supporting a causal *gene–condition* relationship to the more tractable question of the formal classification of clinically observed *variants*.

This approach depends on a robust and rigorous variant classification method, such as a careful implementation of the American College of Medical Genetics (ACMG) variant classification guidelines. These guidelines classify observations about the consequence and context of a variant into one of a series of "evidence types." Each evidence type contributes a predetermined amount to the argument that a variant is benign

or pathogenic, and thresholds are suggested for the amount of evidence required to reach a certain classification. Admissible evidence may include family segregation data, observations in multiple, unrelated, clinically affected individuals, absence in healthy controls, animal models demonstrating recapitulation of a human disease phenotype, and functional data demonstrating an aberrant effect on the protein or transcript. The conclusion that a variant is pathogenic generally requires more than one type of strong evidence consistent with pathogenicity, and generally requires the observation of the variant in multiple, unrelated, similarly affected individuals.

The number of different pathogenic variants a gene harbors is not relevant to the general question of clinical validity. In fact there are many cases, especially in genes with gain-of-function disease mechanism, where all known incidences of the disease are caused by the same, specific pathogenic variant. In these cases, the clinical validity of the gene test is still established by the determination that that single variant is pathogenic.

One possible objection is that the assessment of a variant's pathogenicity depends on already knowing the strength of the gene–condition relationship. We would argue that a careful application of the variant classification framework already takes the relevant uncertainties into account. The ACMG framework provides a series of cautions about using information of tangential relevance: for example, case reports from individuals with possibly unrelated presentations, or assumptions about molecular mechanism. The power of case reports should be modulated based on the relevance and specificity of the presenting phenotype. Before the gene has been well established as a cause of disease, any case report should be treated with this caution. The significance of the effect of the change on the RNA or protein should be modulated based on the understanding of the molecular mechanism of disease. Before the gene has been well established as a cause of disease, all sequence observations should be treated with this caution. If these cautions are respected, it will be impossible to conclude that a variant is pathogenic without substantial evidence of multiple types supporting the conclusion.

If there is a clinically observed variant that can be classified as pathogenic, then it has been proven that the gene causes human disease.

Gene–Condition Strength Terminology

We describe the strength of the evidence supporting a possible relationship between a gene and a particular condition as either "strong," "suggested," or "emerging."

- "Strong" is used in cases where there exists at least one clinically observed variant supported by sufficient evidence to classify that variant as pathogenic. "Strong" indicates that the relationship has been proven.
- "Suggested" is used in cases where some preliminary evidence exists suggesting a causal relationship, but the relationship has not yet been formally proven.
- "Emerging" is used to describe a growing suspicion that a *specific* condition is caused by a gene that has already been proven to cause disease.

If a gene has at least one "strong" relationship, then clinical validity for that gene has been established. If a gene has no "strong" relationships, then clinical validity has not been established, and the gene remains a "Gene of Uncertain Significance."

Establishing Specificity in the Associated Clinical Phenotype

It is important to be as specific as possible with regard to the condition(s) associated with a gene. However in some cases we recognize that, while it is clear that the gene causes disease, there are too few case reports to derive any confidence about *which* disease, or if that disease matches neatly with any known and established clinical entities.

We address this question by a two-step process. We first establish that the gene causes disease by considering its relationship to a generic entity referred to as "GENE-related conditions." We then consider if that generic entity can be refined into one or more specific conditions.

Our approach to this question relies on a heuristic. In order to consider the relationship between a gene and a specific condition as "strong," we require the observation of a pathogenic variant in three unrelated individuals who manifest the specific condition. If only one or two case reports describing individuals with pathogenic variants and manifesting a specific condition are available, we consider the specific gene–condition relationship to be "emerging." The decision to require three individuals is not a statistical assessment, but is meant as a simple hedge against coincidences of individual expressivity or complex individual genotype; a non-classical clinical presentation in an affected individual may simply reflect an expansion of disease presentation or a modification of the presentation due to other genetic or external modifiers.

A gene can be well established as the cause of one specific condition, and also purported to cause an additional condition. In some cases, the conditions are distinct enough to be thought of as separate entities, and in some cases, the second condition should be thought of as a phenotypic expansion of the gene-specific, clinical manifestation. This distinction can be somewhat subjective. In general, if individuals with pathogenic variants and both phenotypes are reported, we consider this to be evidence in support of the idea that the phenotype associated with that gene is complex, rather than the idea that the gene causes two distinct conditions.

This is of special importance in cardiac genetics where some distinct conditions exist, where common mechanisms may cause one clinical condition to progress to presenting features of a second, and where there is extensive clinical overlap between some closely related conditions. Furthermore, because some cardiac conditions display reduced penetrance and/or later onset, the simple fact of observing a pathogenic variant in a single individual with an unexpected phenotype is not sufficient to establish a relationship between this variant and the carrier's condition.

Evaluation of Cardiac Genes: Examples and Summary

For each gene purported to be associated with a cardiac condition, we evaluated the evidence supporting the pathogenicity of the published, clinically observed variants. The full conclusions of our assessments of the genes associated with arrhythmias, cardiomyopathies, and the related syndromes are presented in **Tables 1** and **2**, and detailed examples of Strong, Suggested, and Emerging relationships are described below.

Establishing a Single Condition as "Strong"

The *MYH7* gene has long been understood to be a cause of Hypertrophic Cardiomyopathy (HCM), a relationship which is clearly illustrated by the well-known, pathogenic p.Arg403Gln variant. This variant is absent from control populations, but has been shown to strongly segregate with HCM in four families with an overall LOD score of 3.4 (4–7). In addition, experimental studies have demonstrated that this change leads to defective ATPase activity and significantly alters actin motility (8–13). Furthermore, this variant has been shown to cause HCM in both transgenic mouse and rabbit models (14, 15). The clinical, functional, population, and animal data clearly establish this variant as pathogenic, and an abundance of individuals with this variant and a classic HCM phenotype have been reported. This pathogenic variant in *MYH7* causes HCM, and the link between *MYH7* and HCM is therefore established.

Multiple "Strong" Conditions Caused by the Same Gene

KCNQ1, the potassium voltage-gated channel, is an example of a gene that causes two clinically distinct conditions: Jervell and Lange-Nielsen Syndrome (JNLS), and long QT syndrome (LQTS).

Jervell and Lange-Nielsen Syndrome is an autosomal recessive, multisystem disorder characterized by congenital profound bilateral sensorineural hearing loss and prolonged QT interval at a young age. Onset of cardiac symptoms typically occurs in childhood, and arrhythmia due to JLNS may result in recurrent syncope, seizure-like activity, or sudden cardiac arrest/death. In addition to congenital hearing loss and cardiac symptoms, some individuals with JLNS have also been found to have anemia and elevated levels of the hormone gastrin.

Although a range of variants can be pathogenic, a common pathogenic JLNS variant is p.Arg518*, a founder mutation in the Swedish population. It has been observed in the homozygous state in a number of JLNS patients, and as a compound heterozygote with other truncating variants (16–18).

In addition, heterozygous carriers of pathogenic variants are affected by LQTS of varying severity. LQTS is characterized by a prolonged QTc interval on electrocardiogram (ECG/EKG) and cardiac arrhythmia, such as torsade de pointes, that may result in recurrent syncope, seizure-like activity, and sudden cardiac arrest/death (19, 20). Although mild hearing loss can sometimes be an associated symptom of LQTS, it is a recognizably distinct clinical entity from JLNS. In one study, 12 heterozygous carriers of the pathogenic p.Arg518* variant were demonstrated to have prolonged QT segments and normal hearing (18).

This series of case studies of individuals with variants that are known to be pathogenic, and who are affected by LQTS and *not* JLNS, establishes the relationship between *KCNQ1* and LQTS.

TABLE 1 | Gene–condition strengths for selected cardiomyopathies.

	HCM	DCM	ARVD/C	LVNC	Overlapping cardiomyopathy syndromes
MYBPC3	Strong	Strong		Suggested	
MYH7	Strong	Strong		Strong	Strong: Laing distal myopathy
PLN	Strong	Strong	Strong	Emerging	
TNNC1	Strong	Strong			
TNNT2	Strong	Strong		Strong	
TPM1	Strong	Strong		Suggested	
TNNI3	Strong	Strong		Suggested	
TTN		Strong			Strong: titinopathies
LMNA		Strong	Suggested	Suggested	Strong: laminopathies
DSP		Strong	Strong		Strong: Carvajal syndrome
RBM20		Strong	Suggested		
VCL	Emerging	Strong			
TAZ		Strong		Strong	Strong: Barth syndrome
DSG2		Strong	Strong		
DES		Strong	Strong		Strong: desminopathy
BAG3		Strong			Strong: myofibrillar myopathy
SCN5A		Strong	Strong	Suggested	
DMD		Strong			Strong: Duchenne muscular dystrophy
RAF1		Strong			
ACTC1	Strong	Emerging		Strong	
ACTN2	Strong	Emerging	Emerging		
CSRP3	Strong	Suggested			
FHL1	Strong				Strong: Emery–Dreifuss muscular dystrophy
GLA					Strong: Fabry disease
MYL2	Strong				
MYL3	Strong				
PRKAG2	Strong				Strong: glycogen storage disease
HCN4				Strong	
RYR2		Emerging	Strong	Strong	
LAMP2		Emerging			Strong: Danon disease
TTR					Strong: transthyretin amyloidosis
ELAC2					Strong: combined oxidative phosphorylation deficiency
MTO1					Strong: combined oxidative phosphorylation deficiency
CAV3	Suggested				Strong: caveolinopathies
ALMS1					Strong: Alstrom syndrome
EMD		Emerging			Strong: Emery–Dreifuss muscular dystrophy
FKRP					Strong: muscular dystrophy
FKTN					Strong: muscular dystrophy
JUP		Suggested	Strong		Strong: Naxos disease
SDHA					Strong: mitochondrial complex II deficiency
TMEM43			Strong		Suggested: Emery–Dreifuss muscular dystrophy
DSC2		Suggested	Strong		Strong: RVC with palmoplantar keratoderma and wooly hair
PKP2		Suggested	Strong		
TCAP	Suggested	Suggested			Strong: limb girdle muscular dystrophy
ABCC9		Suggested			Strong: Cantu syndrome
SGCD		Suggested			Strong: limb girdle muscular dystrophy

This table presents assessments of the strength of each gene–condition associations across the cardiomyopathies. Specific references for each cell in this table are available in the Supplementary Material.

The following genes have only "suggested" relationships to cardiac conditions, and are therefore classified as preliminary evidence genes: LDB3, ANKRD1, PDLIM3, MYPN, NEXN, CALR3, JPH2, MYLK2, MYOM1, MYOZ2, PRDM16, CRYAB, CTF1, FHL2, GATA6, GATAD1, ILK, LAMA4, NEBL, NPPA, TMPO, TXNRD2, DTNA, CTNNA3.

TABLE 2 | Gene–condition strengths for selected arrhythmias.

	Long QT syndrome	Short QT syndrome	Brugada syndrome	CPVT	Overlapping arrhythmia syndrome
KCNH2	Strong	Strong	Suggested		
KCNQ1	Strong	Emerging			Strong: Jervell Lange-Nielsen syndrome
SCN5A	Strong		Strong		
CASQ2				Strong	
RYR2				Strong	
CAV3	Strong				
KCNE1	Strong				Strong: Jervell Lange-Nielsen syndrome
KCNE2	Strong				
CALM1	Strong			Strong	
CALM2	Strong			Strong	
CALM3	Strong			Strong	
TRDN	Strong			Strong	
KCNJ2	Emerging	Strong		Emerging	Strong: Anderson-Tawil
CACNB2		Strong	Strong		
GPD1L			Strong		
CACNA1C	Strong	Suggested	Suggested		Strong: Timothy syndrome

This table presents assessments of the strength of each gene–condition associations across the arrhythmias. Specific references for each cell in this table are available in the supplemental materials.
The following genes have only "suggested" relationships to cardiac conditions, and are therefore classified as preliminary evidence genes: SCN4B, SNTA1, TRPM4, KCNE3, KCNE5, RANGRF, SLMAP, KCNJ8, SCN3B, SCN2B, and SCN10A.
CPVT, catecholaminergic polymorphic ventricular tachycardia.

"Strong" First Condition and "Emerging" Second Condition

In addition to LQTS and JLNS, there is some emerging evidence that certain variants in *KCNQ1* also cause Short QT Syndrome (SQTS). Short QT syndrome is characterized by a shortened QTc interval on electrocardiogram and cardiac arrhythmias that may result in syncope, seizure-like activity, and/or sudden cardiac arrest/death (21, 22). To date, there are two relevant case reports: (1) a 70-year-old male has been observed with SQT and a p.Val307Leu variant. While there is functional evidence that p.Val307Leu could contribute to a gain-of-function phenotype (23), this variant would still be formally classified as a VUS until additional case reports or segregation data became available. (2) An infant with severe fetal bradycardia, irregular rhythm, and short QT who has a *de novo* p.Val141Met variant. In this case also, there is some functional evidence that this mutation has a gain-of-function effect (24). The *de novo* observation contributes strongly to this variant's pathogenic classification.

At this time, there is one report of an individual with a pathogenic variant in *KCNQ1* and a severe short QT/arrhythmogenic phenotype. It is quite likely that certain gain-of-function mutations in *KCNQ1* cause short QT; however, it is also possible that

these variants are coincidental observations in individuals whose true causative variant remains undiscovered. Until additional case reports come to light, we classify the relationship between *KCNQ1* and short QT syndrome as "emerging."

Single Condition Example: "Suggested"

SCN10A may be a gene that causes Brugada syndrome, although this has not yet been proven. While *SCN5A* is the primary cause of Brugada syndrome, rare *SCN10A* variants have been found in about 16% of *SCN5A*-negative Brugada patients. All told, there have been around 30 missense changes observed in Brugada patients (25, 26); however, a detailed evaluation of the underlying evidence regarding each of these variants leads us to conclude that every one of these variants should be classified as VUS. Most of these variants are supported by little evidence beyond an observation in an individual, absence in the general population, and computational predictors. There are two reported variants that have been explored more thoroughly, p.Arg14Leu and p.Arg1268Gln. These variants have each been observed in four individuals with Brugada signs, and experimental evidence seems to demonstrate that in an *in vitro* co-expression model, introducing these variants into *SCN10A* leads to a significant reduction in *SCN5A* current (26). However, these variants are also relatively common in the general population, with hundreds of observations in ExAC. Taken together, and in the absence of compelling segregation data, even these two variants must remain classified as VUSes.

At this point, there exist no clinically observed variants in *SCN10A* that can be classified as pathogenic, and therefore we cannot be certain that pathogenic variants in *SCN10A* cause Brugada syndrome. For this reason, we classify the relationship between *SCN10A* and Brugada as "suggested."

Syndromic Genes and Isolated Phenotypes

Pathogenic variants in genes that are primarily associated with syndromes can sometimes manifest clinically as isolated cardiac conditions. This can be because other symptoms have not yet developed, because other symptoms are subtly present and have escaped notice, or because the gene truly also causes the isolated phenotype. How should we think about the range of conditions caused by mutations in these genes?

For example, do some pathogenic variants in *DMD* cause isolated dilated cardiomyopathy (DCM)? The evidence that they may include: (1) in one large family, isolated DCM in the absence of Duchenne or Becker muscular dystrophy mapped to the *DMD* gene (27), (2) a collection of case studies identified classically pathogenic *DMD* variants (exonic deletions or splice variants) in individuals with DCM and without additional features associated with muscular dystrophy (28–30), (3) a study which evaluated the *DMD* gene (called DYS in this paper) in a series of 436 male patients diagnosed with isolated DCM, the authors identified pathogenic deletions or splice variants in 34 patients (31). Upon closer inspection, many of these individuals with "isolated" DCM had elevated serum creatine phosphokinase and/or mild skeletal myopathy. However, there were six individuals with classically pathogenic *DMD* variants who did not have any signs of latent or undiagnosed muscular dystrophy.

There are, therefore, a collection of individuals with isolated DCM whose phenotypes can clearly be explained by identified, pathogenic variants in *DMD*. We would argue that this series of patients establishes the relationship between *DMD* and isolated DCM. We would also suggest, however, that this is largely a semantic distinction. One can say, "Pathogenic variants in *DMD* can, in some cases, lead to isolated DCM" or one can say "Pathogenic variants in *DMD* cause Becker syndrome, which in some cases can present as (and not progress beyond) isolated DCM" and these amount to much the same thing in practice. A patient who has what is apparently isolated DCM should be evaluated for potential pathogenic variants in the *DMD* gene, as such a variant may be the cause of this patient's condition. Likewise, a patient with a pathogenic variant in *DMD* should be carefully examined and monitored for other symptoms of Becker muscular dystrophy as additional symptoms may be subtle or may appear with a later onset.

FHL1 presents another example of this logic. *FHL1* is a well established, "strong" cause of Emery Dreifuss muscular dystrophy (EDMD) and may also cause isolated HCM. Some EDMD patients develop HCM (32), and there are many reports of patients with pathogenic variants in *FHL1* who present with isolated HCM and who have no other symptoms of EDMD. These include: (1) a small pedigree of three individuals with isolated HCM that segregates with an *FHL1* truncation variant, (2) an unrelated individual with isolated HCM and an apparently *de novo* frameshift variant (33), and (3) a three generation pedigree manifesting HCM that segregates with different truncating *FHL1* variants (34). This series of individuals with pathogenic variants in *FHL1*, with a clinical diagnosis of HCM but with no evidence of EDMD, establishes the relationship between *FHL1* and isolated HCM. However, there is no clear genotype/phenotype correlation distinguishing variants that cause EDMD from variants that cause isolated HCM, and the differing manifestations may be due to other genetic or environmental factors specific to these individuals or families. Patients with isolated HCM should be evaluated for variants in the *FHL1* gene, and patients with pathogenic *FHL1* variants should be carefully examined for features of EDMD.

Panel Design and Clinical Overlap among Cardiac Conditions

Conventional cardiac evaluations may not accurately determine an individual's true, underlying diagnosis. For example, left ventricular hypertrophy observed on an echocardiographic evaluation is typically associated with isolated HCM but may also be the primary presenting feature of an unrecognized syndromic condition such as Noonan syndrome or Fabry disease (35). "End-stage" HCM is characterized by left ventricular dilation, and isolated echocardiogram results can easily lead to a misdiagnosis of primary DCM (36). Ventricular arrhythmia, conduction disease, cardiac arrest or unexplained syncope, in the absence of secondary causes, could either represent a primary inherited arrhythmia syndrome or the early clinical presentation of an arrhythmogenic cardiomyopathy (37–41). Clinicians and professional organizations have recognized the importance of comprehensive genetic testing to aid in the diagnosis of cardiac conditions (42), and panel design should address these issues of overlapping and misleading clinical presentation.

We propose that a comprehensive panel test designed for the molecular diagnosis of a particular condition should include the following classes of genes:

- Genes that have been conclusively proven to cause the condition in question.
- Genes suspected but not yet proven to cause the condition in question.
- Genes that have been conclusively proven to cause a condition within the clinical differential. This category should include genes that cause a condition that can progress into the condition in question, genes that cause a condition that can be confused with the condition in question, and genes that cause a syndrome that include the condition in question as a primary feature.

We suggest that the clinical validity of a panel is established when that panel includes a set of genes that account for a substantial proportion of the genetic causes of the disease in question. Conversely, a panel is NOT valid if it omits certain genes that account for a substantial proportion of the known genetic risk. A clinically valid panel may also include genes for which some preliminary evidence of clinical validity exists ("preliminary evidence genes").

A panel test for HCM should include, therefore, genes proven or suspected to cause isolated HCM, and genes *proven* to cause conditions that can present with HCM as a primary feature, such as Fabry disease.

Likewise, a panel for DCM should include genes proven or suspected to cause isolated DCM, genes proven to cause HCM (because HCM can progress to, and be observed as, DCM), and genes proven to cause arrhythmogenic right ventricular dysplasia or cardiomyopathy (ARVD/C, because ARVD/C can present as DCM).

A selected mapping of clinically presenting features to their potential underlying clinical conditions is presented in **Table 3**. While this mapping is not meant to be comprehensive, it is intended to illustrate some of the common discrepancies and overlaps.

DISCUSSION

The three pillars of effective diagnostic medicine are analytic validity, clinical validity, and clinical utility. Establishing the clinical validity of a multi-gene panel depends on an accurate and detailed understanding of the strength of the evidence that establishes a causal relationship between the included genes and human disease. We have established methods to establish gene-level clinical validity, to construct meaningful panel tests, and have applied these methods to a set of cardiac gene–condition pairs.

Before the advent of exome sequencing, gene–conditions associations were traditionally established through the use of gene-mapping techniques. This approach required the ascertainment of multiple, large affected families to provide

TABLE 3 | Clinical overlap of selected inherited arrhythmias and cardiomyopathies.

		HCM	DCM	ARVD/C	Long QT	Short QT	Brugada	CPVT	Unspecified arrhythmia	Noonan syndrome/ RASopathy	Fabry disease[a]	DMD/ Becker[a]
							...may actually represent a true case of:					
A clinical presentation of…	HCM	✓								✓	✓	
	DCM	✓	✓	✓								✓
	Unspecified cardiomyopathy	✓	✓	✓						✓	✓	✓
	ARVD/C		✓	✓			✓	✓	✓			
	Unspecified arrhythmia	✓	✓	✓	✓	✓	✓	✓	✓			

Specific cardiac presentations can reflect a wide range of underlying conditions. Understanding the spectrum of conditions that can present with particular features is an important first consideration in the planning of diagnostic panels. This table presents some of these relationships.
[a]Noonan, Fabry and DMD/Becker are included as representative examples only. Other overlapping syndromic conditions exist but are not represented in this table [adapted from Pagon et al. (44)].

sufficient power to establish linkage to relatively small genomic regions. Genes within the identified regions were then analyzed further for possibly causal variants, or for functional or biological relevance. This was an effective strategy when the cost of expansively sequencing an individual was prohibitive, but was limited in that it relied on the availability of large pedigrees or multiple pedigrees that shared the same underlying genetic cause. Sufficient families are generally only available in the cases where a single gene explains a substantial number of cases of a particular condition.

As the cost of sequencing has come down, it has become feasible to bypass the process of narrowing the genomic search space, and to move directly to the search for causal variants. This has allowed the clinical research community to take greater advantage of isolated unrelated individuals and small pedigrees to generate meaningful genetic hypotheses. The increased accessibility of exome sequencing, for example, has led to an explosion of hypotheses about gene–condition relationships. The consequence is that we have a greater appreciation of the specific genetic diversity underlying many conditions, but also that the amount of data available to support a particular hypothesis is often substantially limited. As diagnostic testing moves to include these genes in routinely available tests, there is a need for an efficient and reproducible method for evaluating the strength of the evidence suggesting a causal relationship. Meaningful panel design, and the appropriate understanding and use of results derived from the included genes, depend on this fundamental understanding.

Clinical Validity of a Panel Test

This paper aims to provide a method for establishing clinical validity of individual genes that is consistent with the general ACCE framework. We also note that, although clinical practice is quickly moving to embrace panel testing, a clear framework does not exist for establishing clinical validity on a panel level. Some entities suggest that clinical validity of a panel is established when each and every included gene has established validity; however, this assertion is refuted by the rapid adoption of exome testing as a viable, clinically valid option.

For many conditions, the bulk of the diagnostic yield of a panel test is accounted for by pathogenic variants in a small number of genes, and is supplemented by a "long tail" of genes that account for rare cases. In addition, there can be real benefit to patients to testing genes in advance of their clinical validity being conclusively established.

We therefore propose that the clinical validity of the panel test is largely established by the inclusion of genes that account for the bulk of the diagnostic yield for that condition. Conversely, a panel test should be considered to be out-of-date and no longer clinically valid if it fails to include such genes. For example, a clinically valid test for HCM must include *MYBPC3*, as pathogenic variants in this gene account for a substantial portion of HCM cases, and an HCM panel that fails to include this gene should not be. However, a panel should not be bounded by the current state of information, for the reasons described below. An effective HCM panel may also include a series of preliminary evidence genes that may turn out to contribute additional clinical sensitivity, but that cannot do so at this point in time.

Utility of Findings in the Three Classes of Genes

Broadly speaking, genes are included in a panel for one of the three reasons listed below. The utility of findings in the gene depends on the categorization of the gene and the reason for inclusion.

Genes That Definitively Cause a Condition within the Patient's Differential

When pathogenic variants are identified in these genes, these variants likely represent a causal explanation for the individual's condition. Pathogenic variants in these genes can inform the prognosis, management, and treatment of the affected individual. Pathogenic variants are also material to the health and clinical management of the proband's family members. Asymptomatic relatives who carry the variant may be candidates for more aggressive screening, monitoring, or prophylactic interventions. Asymptomatic relatives who do not carry the pathogenic familial variants may be returned to standard monitoring protocols for their demographic.

When variants of uncertain significance (VUSes) are identified in these genes, testing of similarly affected family members may be useful in understanding the clinical significance of the variant. Segregation of the variant with disease can inform the relevance of the variant to the particular family, and may inform the formal classification of the variant.

Most of the diagnostic yield from a panel test is derived from these genes, and testing broadly beyond this class of genes does not substantially increase that yield (35).

Genes That Definitively Cause a Related or Similar Condition, but That Have Not Been Definitively Proven to Cause the Proband's Condition

The clinical overlap between many cardiac conditions is extensive. Furthermore, we know that our understanding of the full phenotypic heterogeneity of many of these genes may be limited. It will come as no surprise when evidence emerges demonstrating that pathogenic variants in any one gene can lead to a larger range of phenotypes than we currently appreciate. Because of this, more clinicians are opting to test genes that have been definitively proven to cause a related disease in their diagnostic testing regimens. However, the utility of findings in such genes is different than that described above.

When pathogenic variants are identified in these genes, they can mean one of a few things. The variant may represent the true cause of the patient's condition, and may indicate that the patient represents an expansion of the clinical phenotype previously associated with the gene. However, due to the prevalence of some cardiac conditions, families affected by more than one condition are not uncommon. The pathogenic variant may, therefore, be an incidental finding, and may indicate that the individual is also at risk for a second condition. The pathogenic finding is still relevant to asymptomatic family members; however, caution should be applied as the observation in the first proband may suggest that the variant is incompletely penetrant in this family. Discovery of pathogenic or uncertain variants in these genes in a patient should stimulate a thorough review of the clinical presentation through the lens of the new hypothesis. The patient may have subtle features of the associated clinical condition that were not initially appreciated.

When a VUS is identified in one of these genes, segregation data can be difficult to parse. Segregation analysis depends on the co-occurrence of the variant with an associated condition. But if it is not yet certain that the gene causes the condition, individuals with that condition are not necessarily informative. If the variant does not segregate with the condition in the family, then it is certainly not likely to be the cause of disease. In this situation, clinicians and patients must wait for additional information to emerge regarding the spectrum of clinical presentations associated with clearly pathogenic variants in the gene.

Genes That Have Not yet Been Proven to Cause Any Condition: Only "Suggested" Gene–Condition Relationships

Genetic testing panels routinely include "candidate" or "preliminary evidence" genes (genes with no more than "suggested" relationships to any clinical condition), and for good reason. The cost of generating and holding additional patient data has become marginal, we expect our understanding of genetics to improve rapidly over the next years, and an appropriate use of this information does not increase downstream clinical cost or burden.

It is, by definition, not possible to identify pathogenic variants in genes which have not been proven to cause any condition. Variants that are identified in these genes are not used to guide monitoring or treatment decisions. They are also not used to inform risk in family members. However, variants identified in these genes are held in the patient record or by the lab so that, if and when new information becomes available, that information can become useful to the patient without having to endure the cost and time of a second genetic test. Additionally, testing these genes can help identify patients and families who may be referred to research studies to help support an expansion of our understanding of the condition and the genetics.

Other Necessary Information for the Accurate Interpretation of Variants

This paper focuses on establishing the clinical validity of particular genes. It should be clear, however, that although clinical validity is a primary consideration, it does not encapsulate all of the relevant details one would need to accurately interpret novel sequence variants. Such details include molecular mechanism of disease, inheritance patterns associated with disease, penetrance, age-of-onset, severity of disease, the consequence of homozygous variants, relevant protein domains, the frequency of *de novo* variants, etc. The framework for rigorously establishing these qualities is beyond the scope of this paper. If it has been proven that the gene causes disease, then questions of "how?" and "by what mechanism?" become relevant.

Other Gene–Condition Classification Efforts

Besides the approach outlined in this paper, there exist other, promising efforts to tackle this essential question of the evaluation of the strength of the purported gene–condition relationship. Among the most promising is being developed by the ClinGen Gene Curation Working Group, a part of the broader ClinGen effort (43). This working group is developing a framework for evaluating the strength of the evidence that supports a gene–condition relationship that is similarly based on the structured evaluation of underlying evidence. We expect that group's efforts to ultimately become the authoritative source for this sort of information. However in order for that to happen, that approach must be finalized and then broadly accepted and adopted with the support of the larger clinical genomics community. Equally importantly, it must be supported and maintained by an extensive community-based curation effort that will march through the Mendeliome and the array of possibly associated conditions. For labs and clinicians working in clinical genetics now, it is simply not possible to defer patient care until these community-based resources can mature.

CONCLUSION

The design of effective diagnostic tests, the clinical validity of those tests, and the effective use of the results of such tests,

depends on a clear understanding of the relationship between each gene and each considered condition. This paper clearly describes a general methodology establishing clinical validity of a gene that can easily be applied across clinical areas. For an active clinical lab, the benefits of the variant-centric approach to the question of clinical validity should be evident: it allows the lab to maintain one consistent lens for assessing clinical molecular genetics, capitalizes on the variant classification method and the infrastructure to support that method, and reduces logical inconsistencies that arise from using different schema to evaluate the relationship between genetic changes and human disease.

AUTHOR CONTRIBUTIONS

The conceptual framework was developed by ST, KN, and JG, with input from the individuals listed in the acknowledgments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fcvm.2016.00020

REFERENCES

1. Burke W. Clinical validity and clinical utility of genetic tests. *Curr Protoc Hum Genet* (2009) Chapter 9:Unit9.15. doi:10.1002/0471142905.hg0915s60
2. Botkin JR, Teutsch SM, Kaye CI, Hayes M, Haddow JE, Bradley LA, et al. EGAPP Working Group. Outcomes of interest in evidence-based evaluations of genetic tests. *Genet Med* (2010) 12(4):228–35. doi:10.1097/GIM.0b013e3181cdde04
3. Burke W. Genetic tests: clinical validity and clinical utility. *Curr Protoc Hum Genet* (2014) 81:9.15.1–8. doi:10.1002/0471142905.hg0915s81
4. Abraham MR, Bottomley PA, Dimaano VL, Pinheiro A, Steinberg A, Traill TA, et al. Creatine kinase adenosine triphosphate and phosphocreatine energy supply in a single kindred of patients with hypertrophic cardiomyopathy. *Am J Cardiol* (2013) 112(6):861–6. doi:10.1016/j.amjcard.2013.05.017
5. Epstein ND, Cohn GM, Cyran F, Fananapazir L. Differences in clinical expression of hypertrophic cardiomyopathy associated with two distinct mutations in the beta-myosin heavy chain gene. A 908Leu – Val mutation and a 403Arg – Gln mutation. *Circulation* (1992) 86(2):345–52. doi:10.1161/01.CIR.86.2.345
6. Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, et al. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell* (1990) 62(5):999–1006. doi:10.1016/0092-8674(90)90274-I
7. Woo A, Rakowski H, Liew JC, Zhao MS, Liew CC, Parker TG, et al. Mutations of the beta myosin heavy chain gene in hypertrophic cardiomyopathy: critical functional sites determine prognosis. *Heart* (2003) 89(10):1179–85. doi:10.1136/heart.89.10.1179
8. Belus A, Piroddi N, Scellini B, Tesi C, D'Amati G, Girolami F, et al. The familial hypertrophic cardiomyopathy-associated myosin mutation R403Q accelerates tension generation and relaxation of human cardiac myofibrils. *J Physiol* (2008) 586(15):3639–44. doi:10.1113/jphysiol.2008.155952
9. Cuda G, Fananapazir L, Epstein ND, Sellers JR. The in vitro motility activity of beta-cardiac myosin depends on the nature of the beta-myosin heavy chain gene mutation in hypertrophic cardiomyopathy. *J Muscle Res Cell Motil* (1997) 18(3):275–83. doi:10.1023/A:1018613907574
10. Roopnarine O, Leinwand LA. Functional analysis of myosin mutations that cause familial hypertrophic cardiomyopathy. *Biophys J* (1998) 75(6):3023–30. doi:10.1016/S0006-3495(98)77743-4
11. Volkmann N, Lui H, Hazelwood L, Trybus KM, Lowey S, Hanein D. The R403Q myosin mutation implicated in familial hypertrophic cardiomyopathy causes disorder at the actomyosin interface. *PLoS One* (2007) 2(11):e1123. doi:10.1371/journal.pone.0001123
12. Palmiter KA, Tyska MJ, Haeberle JR, Alpert NR, Fananapazir L, Warshaw DM. R403Q and L908V mutant beta-cardiac myosin from patients with familial hypertrophic cardiomyopathy exhibit enhanced mechanical performance

at the single molecule level. *J Muscle Res Cell Motil* (2000) 21(7):609–20. doi:10.1023/a:1005678905119
13. Yamashita H, Tyska MJ, Warshaw DM, Lowey S, Trybus KM. Functional consequences of mutations in the smooth muscle myosin heavy chain at sites implicated in familial hypertrophic cardiomyopathy. *J Biol Chem* (2000) 275(36):28045–52. doi:10.1074/jbc.M005485200
14. Geisterfer-Lowrance AA, Christe M, Conner DA, Ingwall JS, Schoen FJ, Seidman CE, et al. A mouse model of familial hypertrophic cardiomyopathy. *Science* (1996) 272(5262):731–4. doi:10.1126/science.272.5262.731
15. Marian AJ, Wu Y, Lim DS, McCluggage M, Youker K, Yu QT, et al. A transgenic rabbit model for human hypertrophic cardiomyopathy. *J Clin Invest* (1999) 104(12):1683–92. doi:10.1172/JCI7956
16. Hedley PL, Jørgensen P, Schlamowitz S, Wangari R, Moolman-Smook J, Brink PA, et al. The genetic basis of long QT and short QT syndromes: a mutation update. *Hum Mutat* (2009) 30(11):1486–511. doi:10.1002/humu.21106
17. Tranebjaerg L, Bathen J, Tyson J, Bitner-Glindzicz M. Jervell and Lange-Nielsen syndrome: a Norwegian perspective. *Am J Med Genet* (1999) 89(3):137–46. doi:10.1002/(SICI)1096-8628(19990924)89:3<137::AID-AJMG4>3.0.CO;2-C
18. Winbo A, Stattin EL, Diamant UB, Persson J, Jensen SM, Rydberg A. Prevalence, mutation spectrum, and cardiac phenotype of the Jervell and Lange-Nielsen syndrome in Sweden. *Europace* (2012) 14(12):1799–806. doi:10.1093/europace/eus111
19. Schwartz PJ, Moss AJ, Vincent GM, Crampton RS. Diagnostic criteria for the long QT syndrome. An update. *Circulation* (1993) 88(2):782–4. doi:10.1161/01.CIR.88.2.782
20. Giudicessi JR, Ackerman MJ. Genotype- and phenotype-guided management of congenital long QT syndrome. *Curr Probl Cardiol* (2013) 38(10):417–55. doi:10.1016/j.cpcardiol.2013.08.001
21. Mazzanti A, Kanthan A, Monteforte N, Memmi M, Bloise R, Novelli V, et al. Novel insight into the natural history of short QT syndrome. *J Am Coll Cardiol* (2014) 63:1300–8. doi:10.1016/j.jacc.2013.09.078
22. Giustetto C, Schimpf R, Mazzanti A, Scrocco C, Maury P, Anttonen O, et al. Long-term follow-up of patients with short QT syndrome. *J Am Coll Cardiol* (2011) 58:587–95. doi:10.1016/j.jacc.2011.03.038
23. Bellocq C, van Ginneken AC, Bezzina CR, Alders M, Escande D, Mannens MM, et al. Mutation in the *KCNQ1* gene leading to the short QT-interval syndrome. *Circulation* (2004) 109(20):2394–7. doi:10.1161/01.CIR.0000130409.72142.FE
24. Hong K, Piper DR, Diaz-Valdecantos A, Brugada J, Oliva A, Burashnikov E, et al. De novo *KCNQ1* mutation responsible for atrial fibrillation and short QT syndrome in utero. *Cardiovasc Res* (2005) 68(3):433–40. doi:10.1016/j.cardiores.2005.06.023
25. Behr ER, Savio-Galimberti E, Barc J, Holst AG, Petropoulou E, Prins BP, et al. Role of common and rare variants in *SCN10A*: results from the Brugada

syndrome QRS locus gene discovery collaborative study. *Cardiovasc Res* (2015) 106(3):520–9. doi:10.1093/cvr/cvv042

26. Hu D, Barajas-Martínez H, Pfeiffer R, Dezi F, Pfeiffer J, Buch T, et al. Mutations in *SCN10A* are responsible for a large fraction of cases of Brugada syndrome. *J Am Coll Cardiol* (2014) 64(1):66–79. doi:10.1016/j.jacc.2014.04.032

27. Towbin JA, Hejtmancik JF, Brink P, Gelb B, Zhu XM, Chamberlain JS, et al. X-linked dilated cardiomyopathy. Molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus. *Circulation* (1993) 87:1854–65. doi:10.1161/01.CIR.87.6.1854

28. Milasin J, Muntoni F, Severini GM, Bartoloni L, Vatta M, Krajinovic M, et al. A point mutation in the 5' splice site of the dystrophin gene first intron responsible for X-linked dilated cardiomyopathy. *Hum Mol Genet* (1996) 5:73–9. doi:10.1093/hmg/5.1.73

29. Bies RD, Maeda M, Roberds SL, Holder E, Bohlmeyer T, Young JB, et al. A 5' dystrophin duplication mutation causes membrane deficiency of alpha-dystroglycan in a family with X-linked cardiomyopathy. *J Mol Cell Cardiol* (1997) 29:3175–88. doi:10.1006/jmcc.1997.0568

30. Kimura S, Ikezawa M, Ozasa S, Ito K, Ueno H, Yoshioka K, et al. Novel mutation in splicing donor of dystrophin gene first exon in a patient with dilated cardiomyopathy but no clinical signs of skeletal myopathy. *J Child Neurol* (2007) 22:901–6. doi:10.1177/0883073807304705

31. Diegoli M, Grasso M, Favalli V, Serio A, Gambarin FI, Klersy C, et al. Diagnostic work-up and risk stratification in X-linked dilated cardiomyopathies caused by dystrophin defects. *J Am Coll Cardiol* (2011) 58(9):925–34. doi:10.1016/j.jacc.2011.01.072

32. Gueneau L, Bertrand AT, Jais JP, Salih MA, Stojkovic T, Wehnert M, et al. Mutations of the *FHL1* gene cause Emery-Dreifuss muscular dystrophy. *Am J Hum Genet* (2009) 85(3):338–53. doi:10.1016/j.ajhg.2009.07.015

33. Friedrich FW, Wilding BR, Reischmann S, Crocini C, Lang P, Charron P, et al. Evidence for *FHL1* as a novel disease gene for isolated hypertrophic cardiomyopathy. *Hum Mol Genet* (2012) 21(14):3237–54. doi:10.1093/hmg/dds157

34. Hartmannova H, Kubanek M, Sramko M, Piherova L, Noskova L, Hodanova K, et al. Isolated X-linked hypertrophic cardiomyopathy caused by a novel mutation of the four-and-a-half LIM domain 1 gene. *Circ Cardiovasc Genet* (2013) 6(6):543–51. doi:10.1161/CIRCGENETICS.113.000245

35. Alfares AA, Kelly MA, McDermott G, Funke BH, Lebo MS, Baxter SB, et al. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genet Med* (2015) 17(11):880–8. doi:10.1038/gim.2014.205

36. Biagini E, Coccolo F, Ferlito M, Perugini E, Rocchi G, Bacchi-Reggiani L, et al. Dilated-hypokinetic evolution of hypertrophic cardiomyopathy: prevalence, incidence, risk factors, and prognostic implications in pediatric and adult patients. *J Am Coll Cardiol* (2005) 46(8):1543–50. doi:10.1016/j.jacc.2005.04.062

37. Forleo C, Carmosino M, Resta N, Rampazzo A, Valecce R, Sorrentino S, et al. Clinical and functional characterization of a novel mutation in lamin a/c gene in a multigenerational family with arrhythmogenic cardiac laminopathy. *PLoS One* (2015) 10(4):e0121723. doi:10.1371/journal.pone.0121723

38. Hoogendijk MG. Diagnostic dilemmas: overlapping features of Brugada syndrome and arrhythmogenic right ventricular cardiomyopathy. *Front Physiol* (2012) 3:144. doi:10.3389/fphys.2012.00144

39. Li D, Morales A, Gonzalez-Quintana J, Norton N, Siegfried JD, Hofmeyer M, et al. Identification of novel mutations in *RBM20* in patients with dilated cardiomyopathy. *Clin Transl Sci* (2010) 3(3):90–7. doi:10.1111/j.1752-8062.2010.00198.x

40. McGregor SM, Husain AN. A brief review and update of the clinicopathologic diagnosis of arrhythmogenic cardiomyopathy. *Arch Pathol Lab Med* (2015) 139(9):1181–6. doi:10.5858/arpa.2014-0114-RS

41. McNair WP, Ku L, Taylor MR, Fain PR, Dao D, Wolfel E, et al. *SCN5A* mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* (2004) 110(15):2163–7. doi:10.1161/01.CIR.0000144458.58660.BB

42. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm* (2011) 8(8):1308–39. doi:10.1016/j.hrthm.2011.05.020

43. ClinGen. *Clinical Validity Classifications as of October 2015*. (2015). Available from: https://www.clinicalgenome.org/site/assets/files/2657/current_clinical_validity_classifications.pdf

44. Pagon RA, Adam MP, Ardinger HH, Wallace SE, Lora A, Bean LJH, et al., editors. *GeneReviews*. Seattle: University of Washington (1993-2016). Available from: http://www.ncbi.nlm.nih.gov/books/NBK1116/

CrossMark

Genetic Evaluation and Use of Chromosome Microarray in Patients with Isolated Heart Defects: Benefits and Challenges of a New Model in Cardiovascular Care

Benjamin M. Helm[1]* and Samantha L. Freeze[2]

[1] *Department of Medical and Molecular Genetics, Indiana University School of Medicine, IU Health, Indianapolis, IN, USA,*
[2] *Department of Pediatrics, Indiana University School of Medicine, IU Health, Indianapolis, IN, USA*

Congenital heart defects (CHDs) are common birth defects and result in significant morbidity and global economic impact. Genetic factors play a role in most CHDs; however, identification of these factors has been historically slow due to technological limitations and incomplete understanding of the impact of human genomic variation on normal and abnormal cardiovascular development. The advent of chromosome microarray (CMA) brought tremendous gains in identifying chromosome abnormalities in a variety of human disorders and is now considered part of a standard evaluation for individuals with multiple congenital anomalies and/or neurodevelopmental disorders. Several studies investigating use of CMA found that this technology can identify pathogenic copy-number variations (CNVs) in up to 15–20% of patients with CHDs with other congenital anomalies. However, there have been fewer studies exploring the use of CMA for patients with isolated CHDs. Recent studies have shown that the diagnostic yield of CMA in individuals with seemingly isolated CHD is lower than in individuals with CHDs and additional anomalies. Nevertheless, positive CMA testing in this group supports chromosome variation as one mechanism underlying the development of isolated, non-syndromic CHD – either as a causative or risk-influencing genetic factor. CMA has also identified novel genomic variation in CHDs, shedding light on candidate genes and pathways involved in cardiac development and malformations. Additional studies are needed to further address this issue. Early genetic diagnosis can enhance the medical management of patients and potentially provide crucial information about recurrence. This information is critical for genetic counseling of patients and family members. In this review, we review CMA for the non-genetics cardiology provider, offer a summary of CNV in isolated CHDs, and advocate for the use of CMA as part of the cardiovascular genetics evaluation of patients with isolated CHDs. We also provide perspective regarding the benefits and challenges that lie ahead for this model in the clinical setting.

Keywords: chromosome microarray, congenital heart defects, copy-number variation, genetic counseling, clinical genetics

INTRODUCTION

Congenital heart defects (CHDs) are a common group of human malformations with significant morbidity and economic impact (1–3). The prevalence of CHDs in the general population has increased with ongoing advancements in medical and surgical care so that survival to adulthood is relatively common (4). In fact, the population of adults with CHDs is now larger than the number of children with CHDs. Despite the birth incidence of CHDs remaining relatively stable over the last half-century, the true global prevalence of CHDs is likely underestimated (2, 5). As more individuals with CHDs survive and reach reproductive age, questions regarding heritability, etiology, and recurrence risks will be common.

The majority of all CHDs are isolated or non-syndromic, but about 20–30% of infants with CHDs have extracardiac malformations (6). These cases often constitute well-known chromosomal and single-gene syndromes (e.g., trisomy 21, trisomy 18, and Noonan syndrome). However, complex rare diseases with CHDs and multiple congenital anomalies may remain undiagnosed, despite expert evaluation and/or the use of genetic testing. The underlying causes for the vast majority of CHDs remain unknown, especially in the case of apparently isolated or non-syndromic CHDs.

Approximately 20–30% of CHDs can be attributed to a single identifiable genetic or environmental cause (6–8), while the remaining cases are thought to be multifactorial. Examples of environmental risk factors include maternal disease (like maternal hyperphenylalaninemia, rubella, diabetes) and fetal teratogens (like alcohol, retinoic acid, and lithium) (9–11). CHDs are genetically heterogeneous, with numerous confirmed or proposed genetic risk factors, including single-gene variation, aneuploidy, chromosome rearrangements, and chromosome deletions/duplications. There are at least 55 human genes implicated in CHDs, but over 500 have been identified in mouse models (12). It is likely that the same magnitude will be eventually identified in humans. However, it is estimated that about 70–80% of CHDs have an unknown or multifactorial basis (13, 14). The complexity of genetic contributions probably reflects the complexity of cardiac development, and it is accepted that CHD development is influenced by multiple genetic (and environmental) factors. A multifactorial etiology emphasizing genetic contributions has been proposed for CHDs based on recurrence risk data (~1–4% across all lesions) and that the fact that family history is a consistent risk factor for CHDs (15–21). These recurrence risks generally increase as the number of affected first-degree relatives increases. Heritability estimates have been relatively high for specific classes of CHDs, namely, the left ventricular outflow tract obstructions (LVOTO) (22–24). The available evidence suggests that most CHDs have some genetic basis, but this is complicated further by variable expressivity and incomplete penetrance, even in families with an identified gene mutation or chromosome abnormality predisposing to the development of CHDs. Additionally, variants in the same genes can result in a spectrum of cardiac phenotypes.

As more individuals with CHDs survive and reach reproductive age, questions regarding inheritance and recurrence risk become increasingly important for reproductive planning and counseling. The recent advent of genomic technologies like chromosome microarray (CMA) and next-generation sequencing are providing additional diagnostic ability and refining recurrence information. As knowledge of the genetic bases of CHDs increases, genetic evaluation, testing, and counseling will continue to be important parts of the management of patients with CHDs. Current understanding of the multifactorial basis of CHDs is growing but far from complete, and cytogenetic analysis remains a valuable tool in the evaluation of patients with CHDs.

CYTOGENETICS AND CHROMOSOME MICROARRAY FOR THE CARDIOLOGIST: A REVIEW

Chromosome analysis has been a standard for investigating causes for developmental delay/intellectual disability, autism spectrum disorder, and congenital anomalies (25, 26). However, standard chromosome analysis (i.e., karyotype) has an estimated 3% detection rate for pathogenic chromosome abnormalities. Conventional chromosome analysis detects well-known chromosome aneuploidies (like trisomies 13, 18, and 21 or Turner syndrome) in about 10% of cases of CHDs (27). The innovation of CMA technology has increased the detection of chromosome abnormalities thought to be causative in individuals with developmental delay and congenital anomalies from 3% to about 15–20% (25). Karyotype has a genomic resolution of ~5–10 million base-pairs (megabases, or Mb); chromosome anomalies smaller than this are not consistently or reliably detected. Current CMA platforms generally have a genomic resolution of $\geq$250 thousand base-pairs (kilobases, or kb), though some platforms may have a resolution down to individual genes (1 kb).

One evident challenge of this increased genomic resolution is that smaller chromosome variations that have unknown clinical significance can be identified (28). This contrasts conventional chromosome testing (karyotype) in which large imbalances that are detected are all likely pathogenic, and it is uncommon to identify variants of unknown significance. Due to the increased diagnostic ability of CMA, Miller et al. (25) suggested that they be used as a first-line test over standard karyotype – though there are certain scenarios in which karyotype may be an ideal test (balanced chromosome rearrangements, family histories with multiple miscarriages, and/or reduced fertility).

Chromosome microarray is ideal for detecting chromosomal imbalances and copy-number variations (CNVs) in patients with birth defects and early developmental impairments (25, 29, 30). CNVs are generally defined as chromosomal deletions or duplications that cannot be detected using traditional chromosome analysis, generally sized 1 kb or greater. These CNVs are also referred to as "microdeletions" and "microduplications." Additionally, other chromosomal imbalances can be detected like gross aneuploidy and higher-order amplifications like triplications. Interpretation of the clinical significance of CNVs is typically based on the overall size, gene content, location of breakpoints, and deletion vs. duplication of a chromosome region. Because the clinical significance of many CNVs may be uncertain, the American College of Medical Genetics and Genomics published guidelines to assist

in predicting the pathogenicity of CNVs (31). Importantly, when a dose-sensitive gene is involved in the CNV region, deletion or duplication may have profound effects on the function of the gene and its protein products and potentially affect other downstream gene functions.

Chromosome microarray is performed by two strategies: array-based comparative genomic hybridization (aCGH) platforms or by single-nucleotide polymorphism (SNP) platforms. Array-based CGH utilizes short DNA sequence oligonucleotide probes, whereas SNP-based arrays use SNPs as probes. SNP microarrays also provide genotype information by detecting allelic copies of single base-pairs throughout the genome. Loss/gain of oligonucleotide probes on the aCGH platform and loss/gain of SNPs on the SNP-based platform, both indicate deletions and duplications, respectively. Current CMA platforms may merge these two strategies in the form of "oligo-SNP" microarrays. It is imperative that ordering providers understand the benefits and limitations of CMA platforms and be able to interpret and communicate results to patients/families.

COPY-NUMBER VARIATION AND CHD: A REVIEW OF THE LITERATURE

While many CNVs are associated with well-described genetic syndromes, the role of CNVs in the development of all CHDs is not entirely known at this time. A few examples of well-characterized syndromes with CHDs caused by CNVs include Williams syndrome (7q11.23 deletion), DiGeorge syndrome (22q11.2 deletion syndrome), and Smith–Magenis syndrome (17p11.2 deletion). It should be noted that these conditions typically include other congenital anomalies, dysmorphic features, and neurodevelopmental disorders. Assessment by a clinical geneticist should be standard in these and similar cases with CHDs due to the presence of congenital anomalies and/or developmental delay.

It is estimated that pathogenic CNVs are present in 15–20% of patients with CHDs and extracardiac features (32–35). The submicroscopic deletions and duplications associated with these syndromes generally are not detected by routine chromosome analysis; therefore, emphasizing the importance of CMA as a part of the diagnostic workup in patients with CHDs. Although CNVs play an important role in the development of genetic syndromes with CHDs, most CHDs do not occur in the context of a genetic syndrome. While the exact contribution of CNVs to isolated CHDs is unclear, studies show that ~4–14% of individuals with isolated CHDs have pathogenic or suspected pathogenic CNVs (36, 37), though others have suggested 3–10% (13, 38). Geng et al. (39) retrospectively reviewed 514 CHD cases that had CMA testing, contrasting the yields between isolated and syndromic cases. They found pathogenic or likely pathogenic results for 4.3–9.3% of isolated CHD cases. The yield was higher for syndromic cases when excluding aneuploidies. Additional large-scale studies are necessary to further specify and support these estimates.

The few studies that have investigated the contribution of CNV to the development of isolated CHDs are providing insight into additional genes and pathways involved in cardiovascular development and malformation. These studies can also provide additional understanding about heritability, recurrence, variable expression, and incomplete penetrance of CHDs in families. In the studies summarized in **Table 1**, there are examples of apparently isolated CHDs that were found to have CNVs overlapping known syndromic regions [e.g., 22q11.2 deletion and duplication, 16p11.2 duplication; see Silversides et al. (40)]. It is unclear if those patients had been evaluated for and/or diagnosed clinically with a genetic syndrome, or they had been unrecognized or only presented with mild features. These studies have not only provided additional information about candidate genes and pathways associated with CHDs or risk of CHDs but also highlight that even apparently isolated CHDs may actually be syndromic. This information can inform patient evaluation and may lead to early diagnosis, which can have positive impact on management and genetic counseling. Utility of genetic testing depends largely on accurate phenotyping of the CHD lesion and the presence of extra-cardiac features. Further studies with meticulous phenotyping and goals to assess broad classes of CHDs lesions should be undertaken to further refine this estimate. This also highlights the critical importance of involvement of clinical geneticists in the evaluation of seemingly isolated CHDs.

CMA FOR THE CHD POPULATION: INTERPRETATION OF RESULTS

Chromosome microarray is recommended as a first-tier clinical genetic test in cases of isolated CHDs due to the relatively high rate of detection of pathogenic CNVs. Positive results in the patient with isolated CHDs can provide important information for practitioners and family members when making decisions regarding ongoing care and family planning. The negative CMA result can also be critical in guiding next steps for care and in limiting the differential for any given patient. Many well-described chromosomal conditions can be eliminated as diagnoses by a normal CMA result. This elimination can guide further genetic testing decisions and options for additional clinical testing to hone in on the exact diagnosis for the patient.

While implementing the use of CMA in the diagnostic evaluation of patients with CHDs has uncovered previously unknown pathogenic chromosome variation, it has also presented the unique challenges of interpreting variants of uncertain significance (VUS). Designation of VUS is typically reserved for deletions or duplications that have not been previously described, have not been seen in studied control populations, and for which there are incomplete data regarding genes in the affected region (45). Adding to this difficulty are the concepts of incomplete penetrance and variable expressivity. Some VUS results have been reported in the literature with highly variable phenotypic features due to variable expressivity, adding further complexity to the interpretation of the contribution of any given VUS to the phenotype of the patient. Incomplete penetrance of CNVs also complicates the recommendations and counseling provided to families, as accurate risk prediction for certain health concerns cannot be provided. Of particular concern is the patient with significant morbidity who inherits a CNV of unknown significance

TABLE 1 | A summary of CNVs identified by CMA in non-syndromic CHDs reported in the literature.

Study	CNVs (limited list)	Confirmed or putative candidate genes for CHDs noted by authors (limited list)	CHD types in study	Other notes
Thienpont et al. (41)	4q34 deletion 5q35.1 deletion 9q34.3 deletion 22q11.2 duplication	NKX2-5 NOTCH1	AS, TOF, CoA, VSD, truncus arteriosus, PS	
Erdogan et al. (36)	1q21.1 deletion 2p22.3 duplication 17p11.2 deletion 22q11.2 duplication	GJA5 LTBP1 (See note in last column) TBX1, CRKL	Any CHD: majority were VSD; TOF, PS, CoA, ASD, AS, HLHS, and AVSD	The 17p11.2 deletion is causative for Smith–Magenis syndrome. Features of this disorder were not appreciated until after the test result
Greenway et al. (42)	1q21.2 deletion and duplication 2p23.3 duplication 3p25.1 duplication 9q34.3 deletion 20p12.2 deletion[a] 22q11.2 deletion[a]	GJA5, PRKAB2, CHD1L, BCL9 ASXL2, KIF3C, RAB10 RAF1 NOTCH1 JAG1 TBX1, CRKL	TOF	Study involved only subjects with TOF
Silversides et al. (40)	1q21.1 duplication 1q32.2 deletion 3p25.1 deletion 7q21.11 deletion 7p15.3 deletion 7p22.2 deletion 8p23.1 deletion[a] 8p23.3 duplication	GJA5 PLXNA2 RAF1 SEMA3E, SEMA3D DNAH11 SNX8 GATA4, ANGPT2 ARHGEF10	TOF	Study involved only subjects with TOF
Soemedi et al. (37)	1q21.1 duplication 4q34 deletion 5q14.1q14.3 duplication 5q35.3 duplication 8p23.1	GJA5 HAND2 SSBP2, TMEM167A, VCAN, EDIL3 CNOT6 GATA4	TOF, ASD, VSD, CoA, complex left-sided defect, TAPVR	Other rare CNVs identified with unconfirmed candidate genes associated with cardiac development; TOF overly represented
Fakhro et al. (43)	1q32.3 duplication 2p25.1 duplication 3p24.1-p23 deletion 3p24.1 duplication 7q36.1 deletion 8p23.1 deletion 9q34.11 duplication	NEK2 ROCK2 TGFBR2, RBMS3, GADL1 TGFBR2, GADL1 GALNT11 GATA4 NUP188	Heterotaxy with: D-TGA; dextrocardia; VSD, ASD, PAPVR; malposed great arteries, CoA	Study involved CHDs with heterotaxy
Zhao et al. (44)	3q21.3 duplication 16q23.1 duplication 18q23 duplication 22q11.2 deletion[a]	PLXNA1 WWOX NFATC1 TBX1, CRKL	ASD, VSD, PDA, TOF, Ebstein anomaly, tricuspid incompetence	Study involved 100 Han Chinese subjects

AS, aortic stenosis; ASD, atrial septal defect; AVSD, atrioventricular septal defect; CHD, congenital heart defect; CoA, coarctation of the aorta; D-TGA, dextro-transposition of the great arteries; HLHS, hypoplastic left heart syndrome; PAPVR, partial anomalous pulmonary venous return; PDA, patent ductus arteriosus; PS, pulmonary stenosis; TAPVR, total anomalous pulmonary venous return; TOF, tetralogy of Fallot; VSD, ventricular septal defect.

[a]It was unclear if these reports included patients with clinical diagnoses of syndromic disorders (i.e., DiGeorge syndrome for 22q11/2 deletion or Alagille syndrome for the 20p12.2 deletion). It could be that these reports were either unrecognized syndromes or individuals who were mildly affected.

from a typical-appearing parent. This situation requires discernment from both the calling laboratory and the health care team in order to provide an accurate risk assessment to the "unaffected" parent as well as the affected child and the significance of the familial CNV. It may also be difficult to provide accurate recurrence risk information for reproductive decision-making if the contribution of the CNV to the affected patient's phenotype is unclear. Parental studies should be offered in the event that a VUS is found in a child in order to aid in interpretation and significance of the CNV regardless of whether the parents have similar or dissimilar phenotypes. However, insurance coverage

and justification of how this information will impact the parental medical management may prove to be difficult and require the adamant support from the health care team in order to secure insurance coverage. The likelihood of discovering a VUS should be outlined to the family as part of the pretest informed consent process.

Another challenge when using CMA in the CHD population is the "one-hit fallacy" or the notion that any specific CHD is caused by one particular genetic variation alone. CHD is a multifactorial disease caused by both environmental and genetic factors. The contribution of any one CNV to the overall risk for

CHDs is difficult to assess. While ~20% of CHDs can be attributed to a known cause (syndromic, teratogenic, etc.), the vast majority of CHDs is non-syndromic, isolated defects exhibiting a multifactorial inheritance pattern. In any one case of isolated CHD, there may be multiple genes involved, each providing a minimal contribution to the patient's risk, interacting with various environmental factors to form a complex model of CHD development.

One area in which health care providers can aid in the interpretation of a CNV is to provide accurate and thorough phenotyping prior to the completion of genetic testing. By performing CMA for a patient, a genome-wide net is cast in the hopes of finding an explanation for the patient's particular phenotype. By casting such a wide net, results can often be complicated by overlapping clinical diagnoses and lack of genotype/phenotype correlations. CNVs must also be considered in the context of size, location, and gene involvement. Understanding of the clinical significance of a CNV involving genes that have yet to be well described or that have yet to be implicated in a particular phenotype can prove to be difficult. One example of distinguishing cardiac phenotypes is the presence of an atrial septal defect (ASD) and the classification of primum vs. secundum ASD. Primum ASDs are within the spectrum of atrioventricular canal defects, whereas secundum ASDs are a malformation of the atrial septum (46). This classification distinguishes the CHDs from a developmental perspective and can aid in the interpretation by narrowing focus on genes associated with the responsible developmental process. Accurate and specific phenotyping will require coordinated efforts from cardiologists and clinical geneticists.

The process of interpretation of CNVs is ongoing and constantly evolving. As CMA continues to be performed as a first-line test for patients with CHDs, CNVs classified as VUS will continue to present challenging clinical scenarios for health care providers. While it is important that both the laboratory and the health care team work together to interpret CNVs and assign appropriate labels of pathogenicity, it is also important to acknowledge current limitations in our understanding of the human genome and the contribution of variation to cardiovascular disease phenotypes. As our knowledge continues to increase, the opportunity to further refine and identify novel phenotypes presents an exciting challenge for the cardiovascular genetics community.

THE IMPORTANCE OF GENETICS CARE PROVIDER INVOLVEMENT WITH CHD

Our understanding of the association between CNVs and syndromic genetic diagnoses is increasing. There are many examples of newly described microdeletion and microduplication conditions with CHDs (47). Variable expressivity of these conditions and the generally small number of patients described in the literature can make it difficult to recognize associated features and make an accurate diagnosis. Even more well-described syndromes, such as DiGeorge syndrome, 1p36 deletion syndrome, and Williams syndrome, can go undetected for many years in patients with mild or variable presentations. Early involvement of the clinical genetics team provides the opportunity for earlier recognition of syndromic conditions, which can result in more

comprehensive medical interventions and therapies as well as improved prognosis, compared to those patients who receive a syndromic diagnosis later in life. There is also increasing recognition that many delineated syndromes have broader phenotypic variability than previously thought (34, 35, 48). Many syndromes may not be recognized earlier in life due to absence of the "classic" defining features (49). CHD, which is present at birth, can provide a framework for the genetics provider to begin the process of creating a differential for the patient due to the higher prevalence of certain types of CHD lesions in certain genetic conditions (50, 51). CMA, as a first-line genetic test, can detect causative CNVs for many syndromic conditions with a CHD component well before other hallmark features of the diagnosis can be recognized. When CMA is negative, additional genetic testing, including sequencing of genes associated with known conditions and/or whole-exome sequencing, may be warranted for patients with multi-system involvement or features suggestive of a particular genetic condition.

Clinical geneticists and genetic counselors serve as valuable resources to family members of patients with CHD. Early syndromic recognition by the geneticist physician and continued follow-up by a genetic counselor can provide valuable information to the family regarding the anticipation of developmental delays and disabilities, available therapies, and social services that might benefit their child. Early diagnosis can also refine recurrence risk estimates and allow families to make informed reproductive planning decisions. Genetic evaluation and risk assessment can prove to be a powerful tool for empowering families to use genetic information to make informed health decisions. A unique role of the genetics team is the ability to clearly communicate familial risk for CHDs and recommendation of family screening protocols. First-degree family members with certain types of CHDs are at an increased risk to also have undetected CHDs. For example, LVOTO heart defects are understood to be a heritable class of defects, and family members have an increased risk of also having a CHD (22). Studies show that in up to 20% of cases, there is at least one other affected relative in the family with variability in type of LVOTO present. Therefore, screening by echocardiogram is recommended for all first-degree family members of someone affected with an LVOTO class of heart defect (23, 52).

An emphasis must be placed on the coordinated efforts of the cardiologist, clinical geneticist, and genetic counselor in the evaluation, management, and follow-up with patients with CHDs and their family members. This same approach should be used when considering CMA testing and interpretation in this population. A multidisciplinary approach provides a comprehensive care model for patients and families. Genetic testing through the use of CMA, even in patients with apparently isolated CHDs, can aid in delineation of diagnosis, accurate risk assessment for family members, and refinement of recurrence risk estimates for reproductive decision-making. Accurate phenotyping and diagnosis can improve patient outcomes and access to necessary evaluations, therapies, and social services. Genetic counseling and education can empower patients with CHDs and their relatives to use their understanding of the genetic basis of cardiovascular disease to in turn choose effective strategies for health maintenance and appropriate psychosocial coping mechanisms.

AUTHOR CONTRIBUTIONS

BH conceived this review article and provided 50% of the work presented in the manuscript. SF completed the remaining 50% of the manuscript. Both authors contributed significant effort in the writing process.

REFERENCES

1. Hoffman JI, Kaplan S, Liberthson RR. Prevalence of congenital heart disease. *Am Heart J* (2004) 147(3):425–39. doi:10.1016/j.ahj.2003.05.003
2. Hoffman JI, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol* (2002) 39:1890–900. doi:10.1016/S0735-1097(02)01886-7
3. Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Rickson JD. Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979-1997. *Circulation* (2001) 103(19):2376–81. doi:10.1161/01.CIR.103.19.2376
4. Brickner ME, Hillis LD, Lange RA. Congenital heart disease in adults: first of two parts. *N Engl J Med* (2000) 342:256–63. doi:10.1056/NEJM200002033420507
5. Marelli AJ, Ionescu-Ittu R, Mackie AS, Guo L, Dendukuri N, Kaoucahe M. Lifetime prevalence of congenital heart disease in the general population from 2000 to 2010. *Circulation* (2014) 130:749–56. doi:10.1161/CIRCULATIONAHA.113.008396
6. Ferencz C, Boughman JA, Neill CA, Brenner JI, Perry LW. Congenital cardiovascular malformations: questions on inheritance. Baltimore-Washington infant study group. *J Am Coll Cardiol* (1989) 14:756–63. doi:10.1016/0735-1097(89)90122-8
7. Meberg A, Hals J, Thaulow E. Congenital heart defects – chromosomal anomalies, syndromes and extracardiac malformations. *Acta Paediatr* (2007) 96:1142–5. doi:10.1111/j.1651-2227.2007.00381.x
8. Grech V, Gatt M. Syndromes and malformations associated with congenital heart disease in a population-based study. *Int J Cardiol* (1999) 68:151–6. doi:10.1016/S0167-5273(98)00354-4
9. Seagraves NJ, McBride KL. Cardiac teratogenicity in mouse maternal phenylketonuria: defining phenotype parameters and genetic background influences. *Mol Genet Metab* (2012) 107(4):650–8. doi:10.1016/j.ymgme.2012.08.001
10. Corrigan N, Brazil DP, McAuliffe F. Fetal cardiac effects of maternal hyperglycemia during pregnancy. *Birth Defects Res A Clin Mol Teratol* (2009) 85(6):523–30. doi:10.1002/bdra.20567
11. Botto LD, Correa A. Decreasing the burden of congenital heart disease: an epidemiologic evaluation of risk factors and survival. *Prog Pediatr Cardiol* (2003) 18:111–21. doi:10.1016/S1058-9813(03)00084-5
12. Andersen TA, Troelsen KDL, Larsen LA. Of mice and men: molecular genetics of congenital heart disease. *Cell Mol Life Sci* (2014) 71:1327–52. doi:10.1007/s00018-013-1430-1
13. Cowan JR, Ware SM. Genetics and genetic testing in congenital heart disease. *Clin Perinatol* (2015) 42(2):373–93. doi:10.1016/j.clp.2015.02.009
14. Blue GM, Kirk EP, Sholler GF, Harvey RP, Winlaw DS. Congenital heart disease: current knowledge about causes and inheritance. *Med J Aust* (2012) 197:155–9. doi:10.5694/mja12.10811
15. Lalani SR, Belmont JW. Genetic basis of congenital cardiovascular malformations. *Eur J Med Genet* (2014) 57:402–13. doi:10.1016/j.ejmg.2014.04.010
16. Gill HK, Splitt M, Sharland GK, Simpson JM. Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J Am Coll Cardiol* (2003) 42:923–9. doi:10.1016/S0735-1097(03)00853-2
17. Loffredo CA, Hirata J, Wilson PD, Ferencz C, Lurie IW. Atrioventricular septal defects: possible etiologic differences between complete and partial defects. *Teratology* (2001) 63:87–93. doi:10.1002/1096-9926(200102)63:2<87::AID-TERA1014>3.0.CO;2-5
18. Oyen N, Poulsen G, Wohlfahrt J, Boyd HA, Jensen PK, Melbye M. Recurrence of discordant congenital heart defects in families. *Circ Cardiovasc Genet* (2010) 3(2):122–8. doi:10.1161/CIRCGENETICS.109.890103
19. Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. Recurrence of congenital heart defects in families. *Circulation* (2009) 120:295–301. doi:10.1161/CIRCULATIONAHA.109.857987
20. Burn J, Brennan P, Little J, Holloway S, Coffey R, Somerville J, et al. Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* (1998) 351:311–6. doi:10.1016/S0140-6736(97)06486-6
21. Zavala C, Jimenez D, Rubio R, Castillo-Sosa ML, Diaz-Arauzo A, Salamanca F. Isolated congenital heart defects in first degree relatives of 185 affected children. Prospective study in Mexico City. *Arch Med Res* (1992) 23:177–82.
22. Hinton RB, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, Benson WB. Hypoplastic left heart syndrome is heritable. *J Am Coll Cardiol* (2007) 50:1590–5. doi:10.1016/j.jacc.2007.07.021
23. McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, et al. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: segregation, multiplex relative risk, and heritability. *Am J Med Genet A* (2005) 134(2):180–6. doi:10.1002/ajmg.a.30602
24. Cripe L, Andelfinger G, Martin LJ, Shooner K, Benson DW. Bicuspid aortic valve is heritable. *J Am Coll Cardiol* (2004) 44:138–43. doi:10.1016/j.jacc.2004.03.050
25. Miller DR, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* (2010) 86(5):749–64. doi:10.1016/j.ajhg.2010.04.006
26. Rauch A, Hoyer J, Zweier C, Kraus C, Becker C, Zenker M, et al. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet* (2006) 140(19):2063–74. doi:10.1002/ajmg.a.31416
27. Wimalasundera RC, Gardiner HM. Congenital heart disease and aneuploidy. *Prenat Diag* (2004) 24:1136–42. doi:10.1002/pd.1068
28. Wincent J, Anderlid BM, Lagerberg M, Nordenskjold M, Schoumans J. High-resolution molecular karyotyping in patients with developmental delay and/or multiple congenital anomalies in a clinical setting. *Clin Genet* (2011) 79(2):147–57. doi:10.1111/j.1399-0004.2010.01442.x
29. Kharbanda M, Tolmie J, Joss S. How to use… microarray comparative genomic hybridisation to investigate developmental disorders. *Arch Dis Child Educ Pract Ed* (2015) 100(1):24–9. doi:10.1136/archdischild-2014-306022
30. Moeschler JB, Shevell M; The Committee on Genetics. Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics* (2014) 134(3):903–18. doi:10.1542/peds.2014-1839
31. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST; Working Group of the American College of Medical Genetics Laboratory Quality Assurance Committee. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med* (2011) 13(7):680–5. doi:10.1097/GIM.0b013e3182217a3a
32. Breckpot J, Thienpont B, Peeters H, de Ravel T, Singer A, Rayyan M, et al. Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. *J Pediatr* (2010) 156(5):e811–7. doi:10.1016/j.jpeds.2009.11.049
33. Goldmuntz E, Paluru P, Glessner J, Hakonarson H, Biegel JA, White PS, et al. Microdeletions and microduplications in patients with congenital heart disease and multiple congenital anomalies. *Congenit Heart Dis* (2011) 6(6):592–602. doi:10.1111/j.1747-0803.2011.00582.x
34. Lalani SR, Shaw C, Wang X, Patel A, Patterson LW, Kolodziejska K, et al. Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. *Eur J Hum Genet* (2013) 21(2):173–81. doi:10.1038/ejhg.2012.155
35. Lalani SR, Ware SM, Wang X, Zapata G, Tian Q, Franco LM, et al. MCTP2 is a dosage-sensitive gene required for cardiac outflow tract development. *Hum Mol Genet* (2013) 22(21):4339–48. doi:10.1093/hmg/ddt283
36. Erdogan F, Larsen LA, Zhang L, Tumer Z, Tommerup N, Chen W, et al. High frequency of submicroscopic genomic aberrations detected by tiling path

array comparative genome hybridisation in patients with isolated congenital heart disease. *J Med Genet* (2008) 45(11):704–9. doi:10.1136/jmg.2008.058776

37. Soemedi R, Wilson IJ, Bentham J, Darlay R, Topf A, Zelenika D, et al. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am J Hum Genet* (2012) 91(3):489–501. doi:10.1016/j.ajhg.2012.08.003

38. Lander J, Ware SM. Copy number variation in congenital heart defects. *Curr Genet Med Rep* (2014) 2:168–78. doi:10.1007/s40142-014-0049-3

39. Geng J, Picker J, Zheng Z, Zhang X, Wang J, Hisama F, et al. Chromosome microarray testing for patients with congenital heart defects reveals novel disease causing loci and high diagnostic yield. *BMC Genomics* (2014) 15:1127. doi:10.1186/1471-2164-15-1127

40. Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, et al. Rare copy number variations in adult with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet* (2012) 8(8):e1002843. doi:10.1371/journal.pgen.1002843

41. Thienpont B, Mertens L, de Ravel T, Eyskens B, Boshoff D, Maas N, et al. Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. *Eur Heart J* (2007) 28:2778–84. doi:10.1093/eurheartj/ehl560

42. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* (2009) 41(8):931–5. doi:10.1038/ng.415

43. Fakhro KA, Choi M, Ware SM, Belmont JW, Towbin JA, Lifton RP, et al. Rare copy number variants in congenital heart disease patients identify unique genes in left-right patterning. *Proc Nat Acad Sci U S A* (2011) 108(7):2915–20. doi:10.1073/pnas.1019645108

44. Zhao W, Niu G, Shen B, Zheng Y, Gong F, Wang X, et al. High-resolution analysis of copy number variants in adults with simple-to-moderate congenital heart disease. *Am J Med Genet A* (2013) 161A:3087–94. doi:10.1002/ajmg.a.36177

45. Reiff M, Bernhardt BA, Mulchandani S, Soucier D, Cornell D, Pyeritz RE, et al. "What does it mean?" uncertainties in understanding results of chromosomal microarray testing. *Genet Med* (2012) 14(2):250–8. doi:10.1038/gim.2011.52

46. Webb G, Gatzoulis MA. Congenital heart disease for the adult cardiologist recent progress and overview. *Circulation* (2006) 114:1645–53. doi:10.1161/CIRCULATIONAHA.105.592055

47. Nevado J, Mergener R, Palomares-Bralo M, Souza KR, Vallespin E, Mena R, et al. New microdeletion and microduplication syndromes: a comprehensive review. *Genet Mol Biol* (2014) 37(1 Suppl):210–9. doi:10.1590/S1415-47572014000200007

48. Verhagen JM, de Leeuw N, Papatsonis DN, Grijseels EW, de Krijger RR, Wesssels MW. Phenotypic variability associated with a large recurrent 1q21.1 microduplication in a three-generation family. *Mol Syndromol* (2015) 6(2):71–6. doi:10.1159/000431274

49. Krepischi-Santos ACV, Vianna-Morgante AM, Jehee FS, Passos-Bueno MR, Knijnenburg J, Szuhai K, et al. Whole-genome array-CGH screening in undiagnosed syndromic patients: old syndromes revisited and new alterations. *Cytogenetic Genome Res* (2006) 115:254–61. doi:10.1159/000095922

50. Burch M, Sharland M, Shinebourne E, Smith G, Patton M, McKenna W. Cardiologic abnormalities in Noonan syndrome: phenotypic diagnosis and echocardiographic assessment of 118 patients. *J Am Coll Cardiol* (1993) 22(4):1189–92. doi:10.1016/0735-1097(93)90436-5

51. Goldmuntz E, Clark BJ, Mitchell LE, Jawad AF, Cuneo BF, Reed L, et al. Frequency of 22q11 deletions in patients with conotruncal defects. *J Am Coll Cardiol* (1998) 32(2):492–8. doi:10.1016/S0735-1097(98)00259-9

52. Kerstjens-Frederikse WS, Du Marchie Sarvaas GJ, Ruiter JS, Van Den Akker PC, Temmerman AM, Van Melle JP, et al. Left ventricular outflow tract obstruction: should cardiac screening be offered to first-degree relatives? *Heart* (2011) 97(15):1228–32. doi:10.1136/hrt.2010.211433

frontiers
in Cardiovascular Medicine

A Review of the Giant Protein Titin in Clinical Molecular Diagnostics of Cardiomyopathies

Marta Gigli[1,2†], Rene L. Begay[1†], Gaetano Morea[1,2], Sharon L. Graw[1], Gianfranco Sinagra[2], Matthew R. G. Taylor[1], Henk Granzier[3] and Luisa Mestroni[1]*

[1] Adult Medical Genetics Program, Cardiovascular Institute, University of Colorado Denver, Aurora, CO, USA, [2] Department of Cardiology, Hospital and University of Trieste, Trieste, Italy, [3] Molecular Cardiovascular Research Program, University of Arizona, Tucson, AZ, USA

Titin (*TTN*) is known as the largest sarcomeric protein that resides within the heart muscle. Due to alternative splicing of *TTN*, the heart expresses two major isoforms (N2B and N2BA) that incorporate four distinct regions termed the Z-line, I-band, A-band, and M-line. Next-generation sequencing allows a large number of genes to be sequenced simultaneously and provides the opportunity to easily analyze giant genes such as *TTN*. Mutations in the *TTN* gene can cause cardiomyopathies, in particular dilated cardiomyopathy (DCM). DCM is the most common form of cardiomyopathy, and it is characterized by systolic dysfunction and dilation of the left ventricle. *TTN* truncating variants have been described as the most common cause of DCM, while the real impact of *TTN* missense variants in the pathogenesis of DCM is still unclear. In a recent population screening study, rare missense variants potentially pathogenic based on bioinformatic filtering represented only 12.6% of the several hundred rare *TTN* missense variants found, suggesting that missense variants are very common in *TTN* and are frequently benign. The aim of this review is to understand the clinical role of *TTN* mutations in DCM and in other cardiomyopathies. Whereas *TTN* truncations are common in DCM, there is evidence that *TTN* truncations are rare in the hypertrophic cardiomyopathy (HCM) phenotype. Furthermore, *TTN* mutations can also cause arrhythmogenic right ventricular cardiomyopathy (ARVC) with distinct clinical features and outcomes. Finally, the identification of a rare *TTN* missense variant cosegregating with the restrictive cardiomyopathy (RCM) phenotype suggests that *TTN* is a novel disease-causing gene in this disease. Clinical diagnostic testing is currently able to analyze over 100 cardiomyopathy genes, including *TTN*; however, the size and presence of extensive genetic variation in *TTN* presents clinical challenges in determining significant disease-causing mutations. This review discusses the current knowledge of *TTN* genetic variations in cardiomyopathies and the impact of the diagnosis of *TTN* pathogenic mutations in the clinical setting.

Keywords: titin, *TTN*, familial cardiomyopathy, cardiovascular genetics, clinical genetics, heart failure, clinical diagnosis

INTRODUCTION

Dilated cardiomyopathy (DCM) is defined by the presence of left ventricular (LV) or biventricular dilatation and systolic dysfunction in the absence of hypertension, valvular disease, or coronary artery disease sufficient to cause global systolic impairment (1). The prevalence of the disease is about 1:2,500, and DCM explains about half of the heart failure cases in the United States. About 35–40% of DCM cases are classified as "idiopathic" or "familial/genetic" cardiomyopathy (2). Other causes of the DCM phenotype are ischemic, congenital, valvular, inflammatory, or cardiotoxic heart disease. Finally, other rare cardiomyopathies, such as hypertrophic cardiomyopathy (HCM), arrhythmogenic right ventricular cardiomyopathy (ARVC), and restrictive cardiomyopathy (RCM), have genetic causes.

In this setting, genetics can justify a significant proportion of DCM cases (up to 25%), so the disease can be classified into genetic and non-genetic forms (3). In DCM, the most common form of cardiomyopathy, more than 50 genes have been associated with the phenotype, usually with incomplete penetrance and variable expressivity, and frequently with familial transmission (2–4). Evidence suggests that familial DCM is inherited in an autosomal dominant pattern in about 90% of cases, but few cases follow an autosomal recessive, x-linked, or mitochondrial pattern of inheritance (5–7). Genes most frequently involved in the disease are encoding structural proteins of the sarcomere (titin and myosin heavy chain), cytoskeleton (desmin), nuclear membrane (lamin A/C), membrane proteins and ion channels (phospholamban and presenilin), protein of the dystrophin-glycoprotein complex (dystrophin and sarcoglycan), desmosomes (desmoplakin and desmoglein), mitochondrial proteins (frataxin), and extracellular matrix proteins (alpha-laminin) (8).

Titin (TTN) encodes the largest human protein, whose name stems from the word Titans, giants of Greek mythology. Among the genes involved in cardiomyopathies, TTN plays a central role because of its frequency and the key structural, mechanical, and regulatory role within the sarcomere in the striated muscle (9). The TTN gene consists of 364 exons, located on chromosome 2q31, that produces maximally a 4,200-kDa protein which is composed of ~38,000 amino acid residues. The size and complex structure of the TTN protein provides architectural support, maintaining the sarcomeric organization during contraction, and developing passive tension during muscle stretching. It also has a sensory and signaling role through the multiple TTN-binding proteins that are organized in signaling hot spots (10–12). The protein is organized in four structural and functional regions: the N-terminal Z-line (anchor to the sarcomeric Z-disk), the I-band (responsible for elastic properties), A-band regions (with a stabilizer role of the thick filament), and the C-terminal M-line extremity (overlap in antiparallel orientation with another C-terminal TTN molecule; modulation of TTN expression and turnover with the tyrosine kinase domain) (10).

Truncation mutations of TTN are the most frequent in DCM where 25% of cases are familial forms and 18% are sporadic forms of DCM (13). However, it remains to be confirmed that TTN truncating mutations are always pathogenic (3, 14). Interestingly, truncations in the A-band region of TTN accounts for up to 25% of DCM cases (15). Furthermore, TTN is involved in the pathogenesis of other cardiomyopathies such as HCM and ARVC that is considered to be a genetic disease (30–50% of cases are familial), and RCM.

After the introduction of next-generation sequencing (NGS), the study of TTN gene mutations, previously difficult to analyze due to its size and complexity, has now allowed the identification of more than 60,000 TTN missense variants (reported in the 1000 Genomes Project) (16, 17). The aim of this review is to discuss the challenges in diagnosing the correlation between TTN mutations and the different types of cardiomyopathy in the clinical setting.
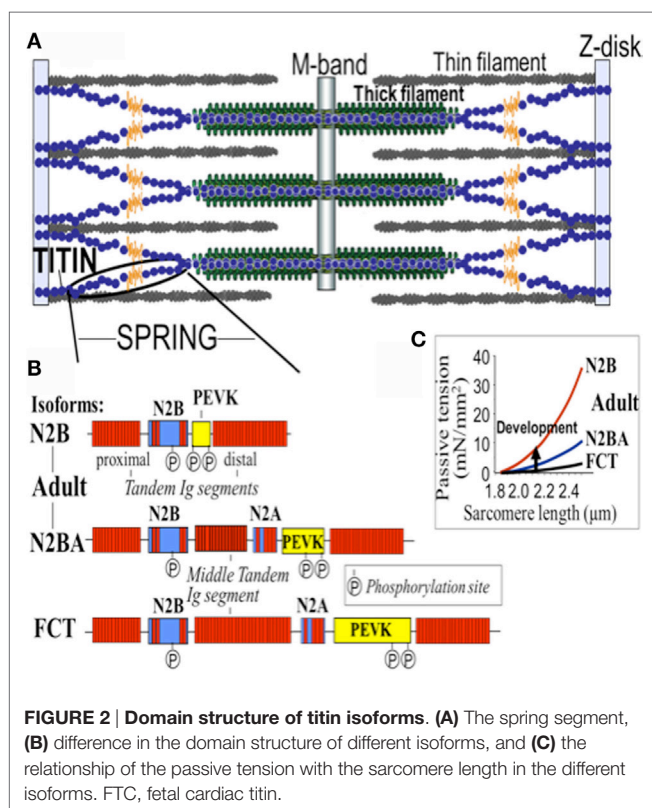
MECHANISTIC STUDIES OF TTN

Titin is the largest human protein. Two TTN filaments with opposite polarity span each sarcomere, namely, the contractile unit in striated muscle cells. TTN is responsible for sarcomere passive stiffness generation (18). TTN is composed of a Z-disk at its N-terminus, whereas the remaining part of the molecule is composed of the elastic I-band region (consisting of tandem Ig segments of serially-linked Ig-like domains), the spring-like PEVK region (is composed of proline (P), glutamate (E), valine (V), and lysine (K)), three unique sequences of Novex1, 2, and 3, cardiac-specific N2B and N2A domains, a thick A-band region, and a M-band region embedding the C-terminus (**Figures 1 and 2**) (19–21). The extensible I-band region gradually lengthens and develops passive tension when the sarcomere is stretched during diastole (15). The inextensible A-band binds myosin and myosin-binding protein C (MyBP-C), whereas the M-band contains a kinase that affects gene expression and cardiac remodeling (22).

The 364 exons of TTN undergo extensive alternative splicing to encode different isoforms. In cardiomyocytes, three different isoforms of titin are expressed: adult N2BA, adult N2B, and the fetal cardiac titin (FCT) isoforms. The I-band sequence defines the different proprieties of each isoform, whereas the Z-disk, A-band, and M-line regions are extremely conserved (22). The isoforms, N2BA and N2B are expressed 30–40 and 60–70% respectively, within the TTN protein in healthy adult human heart. The ratio between these two isoforms is a major determinant of the cardiomyocyte stiffness (18). Due to the longer extensible I-band region, the N2BA titin isoform is more compliant than N2B titin (23–25). The compliant N2BA contains additional spring elements in the PEVK and tandem Ig regions and is therefore associated with low cardiomyocyte passive tension (25). The TTN-based passive tension is established by the TTN expression ratio in the human heart. There is a strong relationship between the TTN-based passive tension and the size of the I-band region: the larger the elastic I-band region and the lower the passive tension (22). Variable isoform expression and TTN splicing have become of great importance in different cardiac diseases, including DCM, whereby the compliant N2BA isoform is upregulated and is associated with decreasing passive stiffness and increasing chamber compliance (23, 24, 26, 27).

A recent study by Roberts et al. suggested that the clinical significance of TTN truncating variants is largely predicated by
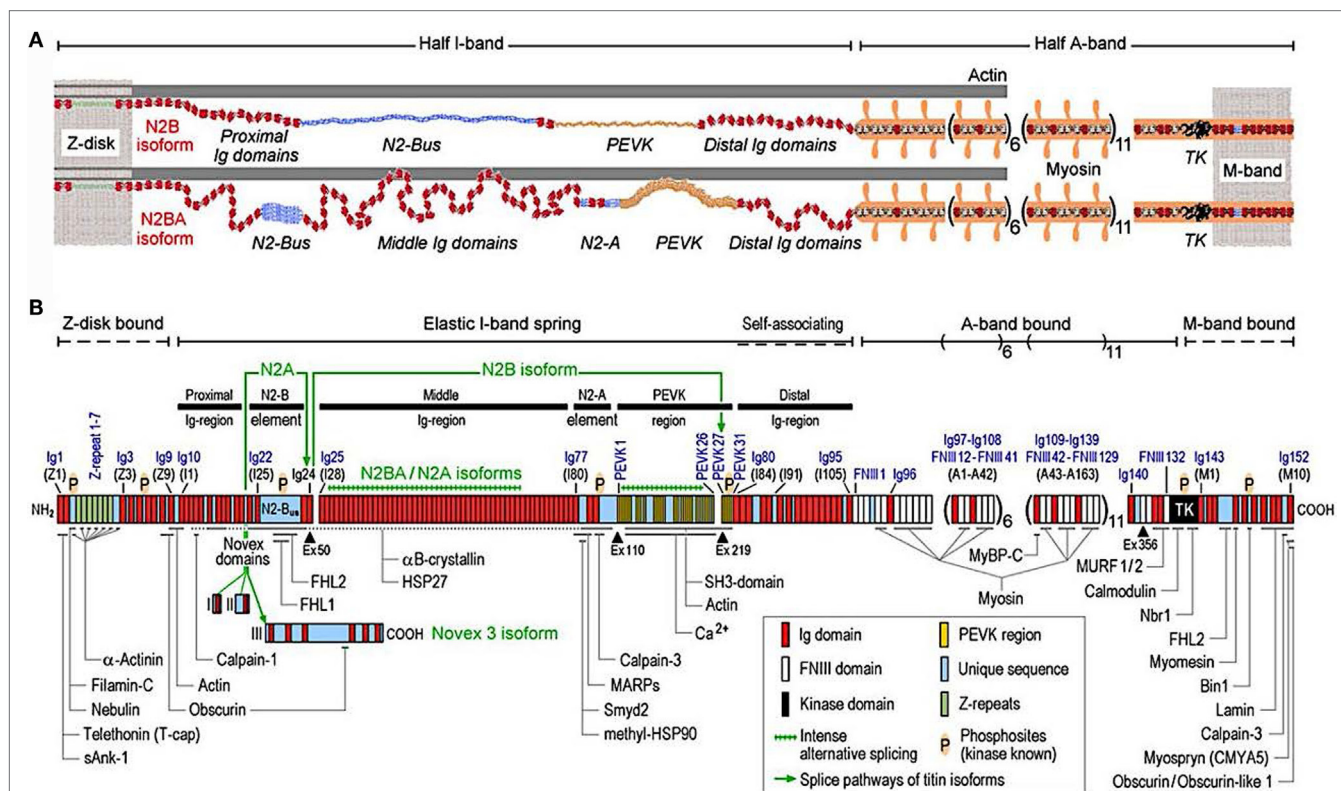
FIGURE 1 | Domain structure of titin isoforms and binding sites of titin ligands. (A) N2B and N2BA titin isoforms represented in the cardiac half-sarcomere, **(B)** Domain structure of titin sequence, Q8WZ42-1, with ligand binding sites represented (from Linke and Hamdani, with permission) (60).

FIGURE 2 | Domain structure of titin isoforms. (A) The spring segment, **(B)** difference in the domain structure of different isoforms, and **(C)** the relationship of the passive tension with the sarcomere length in the different isoforms. FTC, fetal cardiac titin.

the exon usage and variant location (the distance of the truncating variant from the protein N-terminus) (28). Furthermore, the authors compared *TTN* truncating variants among different isoforms and found *TTN* truncating variants altering both N2BA and N2B were overrepresented in DCM patients versus controls and more strongly associated with DCM as compared with the *TTN* truncations involving the N2BA isoform only. Conversely, the *TTN* truncations of the controls were composed of exons not incorporated into N2BA and N2B transcripts (28).

The *TTN* gene structure is organized to accommodate extensive splicing events. Roberts et al. defined a percentage spliced in (PSI) score based on RNA sequencing data from end-stage DCM and donor heart in order to find the mean usage of each *TTN* exon (28). The PSI estimates the proportion of transcripts that incorporate a given exon. A high PSI was given to an exon constitutively expressed and present in all *TTN* isoforms, while a low PSI was usually present only in one isoform and had a lower expression. Moreover, exon symmetry was related to PSI: only 3 exons among the 175 with PSI < 0.99 were asymmetric versus 27% of those with PSI > 0.99. Interestingly, the authors found that more than 80% of all *TTN* exons were symmetric and that their exclusion would not alter the translational reading frame. For instance, in the I-band, the region with the lower PSI, 93% of alternately spliced exons were symmetric: a truncating variant in that region will fall in exons spliced out or not expressed in the majority of the transcripts and should not have such a

deleterious effect. While the stiffness of TTN is defined primarily by the I-band segment sequence of each isoform, it is well known that the cardiac passive tension can be affected by multiple post-translational modifications of contractile and regulatory proteins (29). Few studies have discovered that protein kinase phosphorylation significantly alters the stiffness of N2B and PEVK spring elements (30, 31). The N2B spring element is phosphorylated by PKA and PKG with a reduction in passive tension (29, 32).

The mechanisms responsible for the changes in TTN isoform expression are still not completely understood; however, it has been shown that RNA-Binding Motif Protein 20 (RBM20), a RNA splicing factor, plays an important role in this process and a reduced expression of RBM20 can alter TTN splicing and isoform expression in human (33) and mice (34), leading to DCM.

Therefore, the TTN-based myocardial stiffness is determined by the TTN isoform composition and the phosphorylation state of TTN's elastic I-band. Different kinases can modify the TTN elasticity in different ways; indeed, it is known that changes in post-translational modification (in particular hypophosphorylation) plays a role in the pathophysiology of heart disease (13).

TITIN IN THE PATHOGENESIS OF DILATED CARDIOMYOPATHY

Dilated cardiomyopathy is a primary myocardial disease with variable natural history and clinical presentation affecting young individuals with a potential long life expectancy. A genetic etiology is demonstrated in ~30% of cases (35), and the giant muscle TTN protein has been recognized as the major human disease-causing gene for DCM (9). The advances in contemporary DNA sequencing and the introduction of NGS have allowed the screening of TTN in large cohorts of patients with DCM and in the past few years have been prolific in the description of new DCM-related TTN mutations. A comprehensive cohort study by Herman et al. (16) on 312 DCM patients reported TTN truncating mutations to be the cause of DCM in 25 and 18% of familial DCM and sporadic cases, respectively. TTN truncating mutations found in subjects with DCM were overrepresented in the A-band region and were absent from the Z-disk and M-band regions. Interestingly, TTN truncation variants were also present in up to 2% of the control population, but the control subjects were less enriched for the A-band region of TTN including the Z-band variants. A recent study by Pugh et al. (36) confirmed the presence of truncating variants in the general population (1.65%) and demonstrated that truncating variants located in the A-band are more common in patients with DCM compared with controls. The rate of TTN truncating variants found by Pugh et al., in the DCM cohort was ~14%. In addition, a reduced frequency of variants in the I-band was identified in probands compared with controls, whereas no differences were detected in the Z and M bands.

The TTN gene has also been evaluated in the European Atlas study of 639 patients with sporadic or familial DCM by NGS. Mutations in TTN were identified in 19% of familial and 11% of sporadic cases (37). Noteworthy, 44% of patients with a truncating TTN variant also presented an additional known disease-causing variant in at least one other gene involved in the pathogenesis of DCM; thus in these cases, the TTN variant

may not be the only contributor leading to the pathogenesis of DCM (37).

A large study recently compared the burden of rare TTN variants across five cohorts of healthy volunteers, participants in the Framingham Heart Study, participants in the Jackson Heart Study, cohort of unselected ambulatory patients with DCM, and end-stage DCM cases. The authors confirmed that TTN truncations were not uniformly distributed within and between study groups, being more common in patients with DCM (22%), but with a rate in the healthy volunteers ranging between 1 and 2.9% (28). The TTN truncation variants in the DCM cohort were located predominantly in the A-band, as already described in previous studies mentioned above (16, 36).

The role of TTN truncation mutations in the pathogenesis of DCM has been largely recognized. However, the high prevalence of missense variants and the potential modifier effects make it difficult to elucidate the effective role of TTN missense variants in DCM. Some of these variants are proposed to be pathogenic, but other variants are of unknown significance (VUS). In order to address this challenge, a recent multicenter study sequenced the TTN gene in a cohort of 147 DCM patients (38). In this cohort, 13 TTN truncating variants had previously been reported (16), and 348 missense variants were filtered by bioinformatic algorithms resulting in 44 out of 348 (involving 37 probands) classified as "severe" or likely pathogenic. Among the nine families with TTN variants classified as "severe," five were considered false positives due to discordant cosegregation analysis among affected relatives, whereas four families had "severe" TTN variants that cosegregated with the DCM phenotype. The remaining 28 probands harbored "severe" variants that could not be assessed by cosegregation (possibly pathogenic). Furthermore, the outcome of TTN missense variants carriers did not differ significantly from the other DCM patients (**Figure 3**). Interestingly, the distribution of the likely and possibly TTN severe missense variants across TTN domains was again non-random and was overrepresented in the A-band region of TTN. Specifically, variants were overrepresented in the C-zone of the A-band, which consists of a super repeat of 11 immunoglobulin-like domains and Fn-III domains shown to bind to MyBP-C and subfragment myosin-1, and is essential for the length dependency of force development and calcium sensitivity (39). Therefore, although the real impact of TTN missense variants in the pathogenesis of DCM is still unclear, the clustering of variants in the A-band in DCM may suggest that some A-band missense variants may have a functional detrimental effect on contractility and should be further investigated.

TITIN IN OTHER FORMS OF CARDIOMYOPATHY

Hypertrophic Cardiomyopathy

Hypertrophic cardiomyopathy is a common and inherited cardiomyopathy with a prevalence of 1 in 500 (40). HCM presents as an unexplained LV hypertrophy, myocardial disarray, and fibrosis that translate in increased risk of life-threatening ventricular arrhythmias, sudden cardiac death, and an increased life-long risk of heart failure (41, 42). In the majority of cases, HCM has
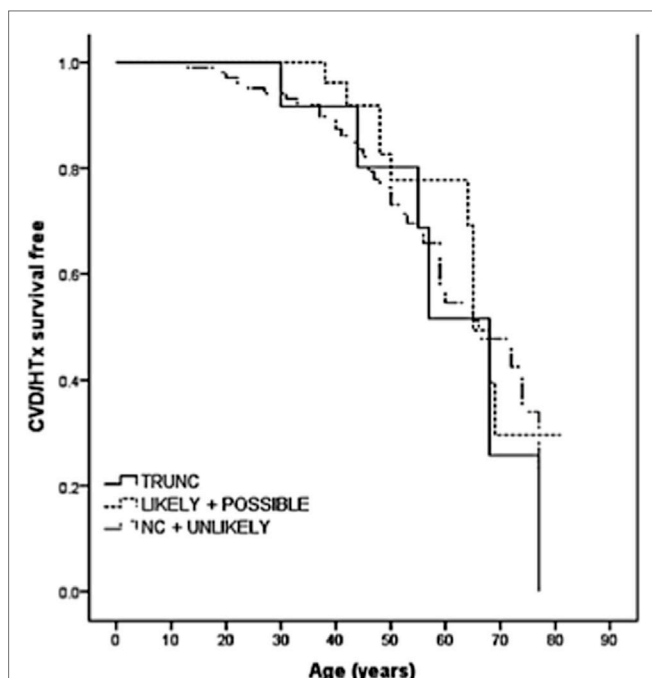
FIGURE 3 | Long-term survival curves in *TTN* variant carriers. Kaplan–Meier event-free survival for cardiovascular death (CVD) or heart transplantation (HTx) based on *TTN* variant categories: truncations (TRUNC); "likely" and "possibly" missense variants; non-carriers (NC), and "unlikely" with lack of cosegregation. *TTN* indicates titin gene (from Begay et al., with permission) (38).

an autosomal dominant trait and mutations in at least 11 different genes. These genes encode for sarcomeric proteins that are responsible for 50–65% of familial cases (9). While *TTN* truncation mutations are common in DCM, there is evidence that *TTN* truncations are rare in the HCM phenotype, with a frequency similar to control populations (16). Using high-throughput sequencing in 142 HCM probands, Lopes et al. found 219 *TTN* rare variants with 209 being novel missense variants (43). However, this cohort of individuals potentially had a sarcomeric gene mutation that likely caused HCM, and the actual pathogenic role of these *TTN* variants in unknown.

Restrictive Cardiomyopathy

Restrictive cardiomyopathy is a very rare form of cardiomyopathy, characterized by preserved biventricular systolic function and a restrictive physiology determining an impaired LV filling despite normal cavity size and frequently normal wall thickness. RCM can be secondary to idiopathic or system disease. It is believed that a significant proportion of RCM cases are genetically determined (42). The pattern of inheritance can be autosomal dominant, autosomal recessive, or x-linked (44). The overall prognosis of RCM is poor, usually resulting in progressive biventricular heart failure with a high mortality rate in the absence of heart transplantation. Interestingly, RCM overlaps in clinical features with HCM (42). Recently, a study using linkage analysis that reported a *TTN* missense variant (*TTN*: c.22862A>G) cosegregating with RCM in six affected individuals of a family. The most common

genes were excluded due to lack of complete cosegregation. Interestingly, some healthy individuals also harbored the *TTN* missense variant resulting in an incomplete penetrance (44). The identification of a rare missense variant in *TTN* cosegregating with the RCM disease phenotype suggests that *TTN* is a novel disease-causing gene for RCM.

Arrhythmogenic Right Ventricular Cardiomyopathy

Arrhythmogenic right ventricular cardiomyopathy is considered to be a genetic disease (30–50% of cases) mainly with autosomal dominant pattern of inheritance (45). ARVC is characterized by fibrofatty replacement of the myocardium, predominantly of the right ventricle, although the left ventricle can also be involved. Typical symptoms include palpitations, cardiac syncope, and cardiac arrest due to ventricular arrhythmias. Heart failure may develop later in life as a result of this disease (46). One study by Taylor et al. in which the investigators analyzed by direct sequencing of 312 exons of *TTN* (311 expressing TTN protein) found *TTN* mutations to be associated with the ARVC phenotype (47). Among seven different probands with an ARVC phenotype, eight *TTN* rare variants (two *TTN* variants present in one proband) were identified (47). In addition to this study, another investigation by Brun et al. compared the clinical outcomes of ARVC patients with *TTN* mutations, desmosomal mutations, and patients with no identifiable mutation (non-carriers) (45). In this study, 13% of *TTN* rare variants were accounted for in their population of subjects. Among the 67 ARVC affected patients (39 ARVC families), 11 harbored rare *TTN* variants and 8 desmosomal genes variants. The *TTN* carriers had increased supraventricular arrhythmias, and conduction disease compared with non-carriers (45), while desmosomal gene variant carriers had the worse prognosis. In conclusion, these studies suggest that *TTN* mutations can cause ARVC and *TTN* mutation carriers have distinct clinical features and outcomes.

TITIN AS A GENE MODIFIER

TTN variants are very frequent; of them, pathogenic mutations are relatively rare and most variants are probably benign. However, a portion of these variants could have a *modifier* gene effect. For instance, *TTN* has been proposed as a modifier gene in combination with the Lamin A/C (*LMNA*) gene (48, 49). A *modifier* gene is not the causal gene, but it may affect the phenotypic expression (50). In a study by Roncarati et al., the authors reported a *TTN* missense mutation modifying the DCM phenotype primarily caused by a *LMNA* mutation. The authors analyzed 41 Italian patients using whole exome sequencing (WES). Fourteen individuals harbored *LMNA*: c.656A>C mutation, and of those five also carried a novel *TTN* missense mutation (*TTN*: c.14563C>T) as well (48). *LMNA* gene mutations are known to be causative of a specific phenotype expression of DCM (51). According to Taylor et al., patients carrying a *LMNA* mutation show a poor prognosis and experience high event-rates compared with non-carriers of a *LMNA* mutation (52). Upholding the structure of the nucleus, chromatin arrangement, and gene expression is encoded by the

LMNA gene for the A-type lamins (53). In a study by Roncarati et al., the presence of the *TTN* variant and the *LMNA* mutation carriers modified DCM patients' clinical course and disease severity, with double heterozygotes requiring earlier heart transplantation (four individuals) compared with those only harboring the *LMNA* mutation alone. Furthermore, histological studies showed more evidence that double heterozygote individuals had worse outcomes on a cellular level (48). In conclusion, this study suggests a modifier role of *TTN* variants that contribute to the complexity of the DCM phenotype.

CLINICAL ASSESSMENT OF TITIN VARIANTS

Titin has been known to be cause a DCM phenotype for many years; however, the systematic analysis and the complete meaning of its contribution to DCM have been precluded by its giant size and sequencing technical limitations (54). As discussed earlier, using NGS, Herman et al. found that heterozygous mutations truncating the full-length TTN are the most common causes of DCM; occurring in ~25% of familial cases of DCM and 18% of sporadic cases. However, *TTN* truncating variants were also found in ~2% of healthy controls (16, 55), raising concern about the correct clinical interpretations of *TTN* variants. The finding of a *TTN* truncating variant in a patient before the onset of clinical manifestation of disease thus requires further in-depth analysis to support pathogenicity (9). Additional factors, such as band location and PSI score, might help to differentiate pathogenic truncation mutations from benign variants (28, 36). This is of particular importance considering that most DCM patients present late in the course of the disease (advanced disease presenting with heart failure or sudden cardiac death), while the early detection of asymptomatic DCM might be critical to enable early intervention that may prevent the progression to advanced disease (56). Moreover, *TTN* truncation variants may be found in association with other disease-related genes, increasing the concerns about the actual role of some *TTN* mutations (37).

Analysis of a large number of genes has led to the identification of sequence VUS. These VUS are one of the main challenges of NGS, because cardiologists and clinical geneticists are faced with uncertainty of the clinical meaning of VUS findings (57).

To date indeed, a large number of identified *TTN* truncating variants are still classified as VUS, and the high prevalence of missense variants in *TTN*, and their potential modifier roles make interpretation difficult in both research and the clinical settings. The location of *TTN* truncating variants can contribute to a better definition of genetic findings, because as already mentioned, *TTN* truncating variants associated with DCM are located predominantly in the A-band (16, 38). The availability of multiple family members to test for cosegregation with disease, the absence in population databases (ClinVar, ExAC, 1000 Genomes Project, and NHLBI Exome Sequencing Project), prediction software (PolyPhen, SIFT, GERP), and functional data also add in the understanding of classifying the pathogenicity of *TTN* variants (9). Most importantly, NGS has to be considered

a diagnostic test in development and testing results need to be interpreted cautiously in close collaboration between bioformaticians, cardiologists, molecular biologists, and clinical geneticists preferably in expert centers. Many novel variants identified by NGS and classified as VUS present as an inconclusive test result, pending further evidence (57).

Once the pathogenic effect has been defined, another concern is the variability in phenotype expression based on the presence and type of *TTN* mutation variants. Roberts et al. found more severe impaired LV function, lower stroke volume, and thinner LV walls in *TTN* truncating than in *TTN* truncating negative DCM patients (**Figure 4**) (28). In this cohort, the *TTN* genotype independently predicted phenotype severity. Furthermore, TTN truncating positive patients more frequently suffer from sustained ventricular tachycardia (28). In the future, if larger prospective studies confirm these findings, *TTN* mutations might influence the decision-making process for the selection of candidates to an implantable cardioverter defibrillator (ICD) implantation, such as in other cardiomyopathies at high risk of life-threatening arrhythmias (58).

Mutations in *TTN* and other proteins affecting *TTN* splicing are associated with the development of DCM, but these mechanisms are still not completely understood (59). Variable isoform expression and *TTN* splicing have become of great importance in DCM, and are associated with decreasing passive stiffness and increasing chamber compliance (26). Both mechanisms might be important in the process of DCM in connection to *TTN* mutations. By genetic approaches or by splicing or posttranslational modifications *TTN* appears to be a target for future therapeutic interventions (9).

Regarding the universe of *TTN* missense variants, the situation is even more challenging because *TTN* missense variants are very common and their real meaning is still unknown. A recent study demonstrated that missense variants did not correlate with the clinical measures of disease severity or progression and indicated that the DCM phenotype caused by *TTN* missense variants are not distinguishable from other types of DCM (**Figure 3**). According to the authors, *TTN* rare missense mutations should not be currently interpreted as disease-causing in most situations (38). Nevertheless, there is some interesting evidence that *TTN* missense mutations may have a modifier role leading to a greater severity of cardiomyopathy (17, 48). In the future, a better understanding of the *TTN* missense variants in DCM will be elucidated with large-scale *TTN* sequencing and functional investigations on *TTN* variant domains.

Finally, despite the recent advances in genetic studies and in the understanding of the different effects of specific gene mutations in the pathogenesis of DCM, the clinical approach to diagnosing cardiomyopathy affected families remains largely based on the general recommendations for heart failure management, familiar screening programs, and systematic follow-up. The continuous improvement in technologies, such as the increasing evidence concerning the clinical expression of different gene variants might lead in the future to an individualized clinical approach to identifying carriers of different mutations.
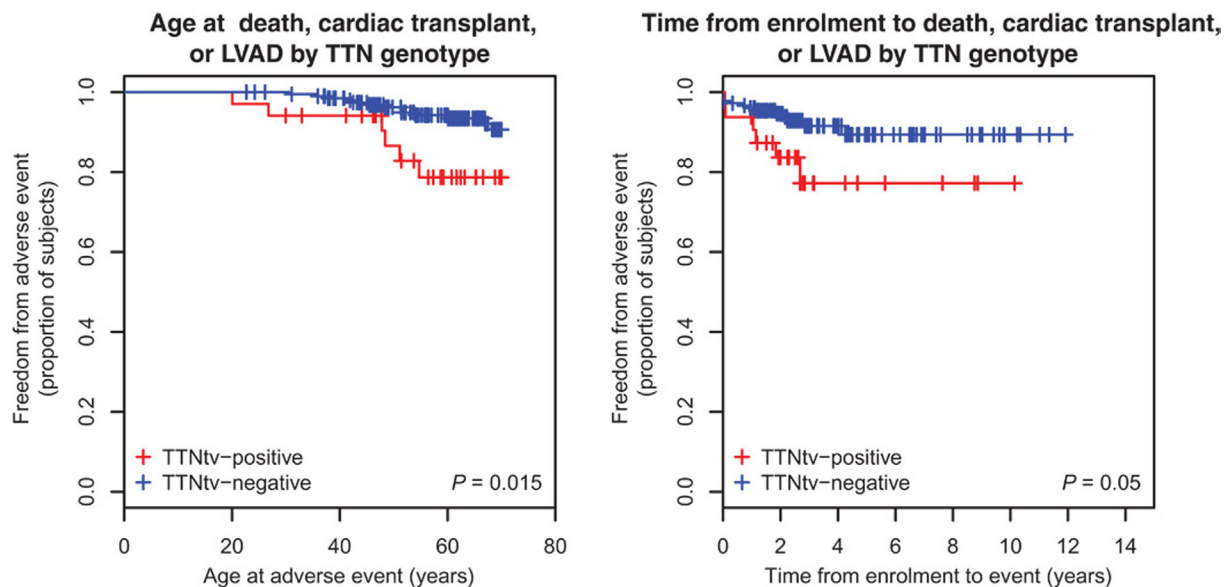
FIGURE 4 | Survival of *TTN* truncation carriers. Patients carriers of a *TTN* truncation variant (TTNtv) had a worse clinical outcome when considering the age of adverse event (death, cardiac transplant or left ventricular assisted device) (*P* = 0.015). They also had a worse clinical outcome when considering the time of event from enrollment (from Roberts et al., with permission) (28).

CONCLUSION

Titin is the largest protein in striated muscle. *TTN* variants have been shown to cause the following cardiac diseases: DCM, RCM, HCM, and ARVC. The advancement of NGS has allowed researchers to analyze the whole *TTN* gene, which has revealed the leading role of this gene in DCM. Challenges are the high genetic variability of the gene, the large number of missense and truncation variants found in control populations, and the criteria for clinical diagnosis of many variants demand individualized clinical diagnosis platforms for *TTN* carriers. Future studies will clarify whether the early identification of *TTN*-related cardiomyopathies might positively influence the natural history of disease by the early initiation of therapeutic management.

AUTHOR CONTRIBUTIONS

All authors have contributed significantly, read, and approved the manuscript. In particular, MG, RB, and GM: drafting of the manuscript; GS, MT, HG, and LM: revising critically the manuscript for important intellectual content.

FUNDING

REFERENCES

1. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, et al. Classification of the cardiomyopathies: a position statement from the European Society Of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* (2008) 29(2):270–6. doi:10.1093/eurheartj/ehm342

2. Hershberger RE, Morales A, Siegfried JD. Clinical and genetic issues in dilated cardiomyopathy: a review for genetics professionals. *Genet Med* (2010) 12(11):655–67. doi:10.1097/GIM.0b013e3181f2481f

3. Pinto YM, Elliott PM, Arbustini E, Adler Y, Anastasakis A, Böhm M, et al. Proposal for a revised definition of dilated cardiomyopathy, hypokinetic non-dilated cardiomyopathy, and its implications for clinical practice: a position statement of the ESC working group on myocardial and pericardial diseases. *Eur Heart J* (2016) 37(23):1850–8. doi:10.1093/eurheartj/ehv727

4. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol* (2013) 10(9):531–47. doi:10.1038/nrcardio.2013.105

5. Burkett EL, Hershberger RE. Clinical and genetic issues in familial dilated cardiomyopathy. *J Am Coll Cardiol* (2005) 45(7):969–81. doi:10.1016/j.jacc.2004.11.066

6. Hershberger RE, Cowan J, Morales A, Siegfried JD. Progress with genetic cardiomyopathies: screening, counseling, and testing in dilated, hypertrophic, and arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Heart Fail* (2009) 2(3):253–61. doi:10.1161/CIRCHEARTFAILURE.108.817346

7. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MR, Towbin JA. Genetic evaluation of cardiomyopathy – a Heart Failure Society of America practice guideline. *J Card Fail* (2009) 15(2):83–97. doi:10.1016/j.cardfail.2009.01.006

8. Sinagra G, Mestroni L, Camerini F. *Genetic Cardiomyopathies. A Clinical Approach*. London: Springer (2012).

9. Gerull B. The rapidly evolving role of titin in cardiac physiology and cardiomyopathy. *Can J Cardiol* (2015) 31(11):1351–9. doi:10.1016/j.cjca.2015.08.016

10. Chauveau C, Rowell J, Ferreiro A. A rising titan: TTN review and mutation update. *Hum Mutat* (2014) 35(9):1046–59. doi:10.1002/humu.22611

11. Witt CC, Ono Y, Puschmann E, McNabb M, Wu Y, Gotthardt M, et al. Induction and myofibrillar targeting of CARP, and suppression of the Nkx2.5 pathway in the MDM mouse with impaired titin-based signaling. *J Mol Biol* (2004) 336(1):145–54. doi:10.1016/j.jmb.2003.12.021

12. Granzier H, Wu Y, Siegfried L, LeWinter M. Titin: physiological function and role in cardiomyopathy and failure. *Heart Fail Rev* (2005) 10(3):211–23. doi:10.1007/s10741-005-5251-7

13. LeWinter MM, Granzier HL. Titin is a major human disease gene. *Circulation* (2013) 127(8):938–44. doi:10.1161/CIRCULATIONAHA.112.139717

14. Akinrinade O, Ollila L, Vattulainen S, Tallila J, Gentile M, Salmenperä P, et al. Genetics and genotype-phenotype correlations in Finnish patients with dilated cardiomyopathy. *Eur Heart J* (2015) 36(34):2327–37. doi:10.1093/eurheartj/ehv253

15. Granzier HL, Irving TC. Passive tension in cardiac muscle: contribution of collagen, titin, microtubules, and intermediate filaments. *Biophys J* (1995) 68(3):1027–44. doi:10.1016/S0006-3495(95)80278-X

16. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* (2012) 366(7):619–28. doi:10.1056/NEJMoa1110186

17. Merlo M, Sinagra G, Carniel E, Slavov D, Zhu X, Barbati G, et al. Poor prognosis of rare sarcomeric gene variants in patients with dilated cardiomyopathy. *Clin Transl Sci* (2013) 6(6):424–8. doi:10.1111/cts.12116

18. Hidalgo C, Granzier H. Tuning the molecular giant titin through phosphorylation: role in health and disease. *Trends Cardiovasc Med* (2013) 23(5):165–71. doi:10.1016/j.tcm.2012.10.005

19. Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, et al. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ Res* (2001) 89(11):1065–72. doi:10.1161/hh2301.100981

20. Labeit S, Barlow DP, Gautel M, Gibson T, Holt J, Hsieh CL, et al. A regular pattern of two types of 100-residue motif in the sequence of titin. *Nature* (1990) 345(6272):273–6. doi:10.1038/345273a0

21. Lange S, Xiang F, Yakovenko A, Vihola A, Hackman P, Rostkova E, et al. The kinase domain of titin controls muscle gene expression and protein turnover. *Science* (2005) 308(5728):1599–603. doi:10.1126/science.1110463

22. LeWinter MM, Granzier H. Cardiac titin: a multifunctional giant. *Circulation* (2010) 121(19):2137–45. doi:10.1161/CIRCULATIONAHA.109.860171

23. Neagoe C, Opitz CA, Makarenko I, Linke WA. Gigantic variety: expression patterns of titin isoforms in striated muscles and consequences for myofibrillar passive stiffness. *J Muscle Res Cell Motil* (2003) 24(2–3):175–89. doi:10.1023/A:1026053530766

24. Cazorla O, Freiburg A, Helmes M, Centner T, McNabb M, Wu Y, et al. Differential expression of cardiac titin isoforms and modulation of cellular stiffness. *Circ Res* (2000) 86(1):59–67. doi:10.1161/01.RES.86.1.59

25. Lahmers S, Wu Y, Call DR, Labeit S, Granzier H. Developmental control of titin isoform expression and passive stiffness in fetal and neonatal myocardium. *Circ Res* (2004) 94(4):505–13. doi:10.1161/01.RES.0000115522.52554.86

26. Nagueh SF, Shah G, Wu Y, Torre-Amione G, King NM, Lahmers S, et al. Altered titin expression, myocardial stiffness, and left ventricular function in patients with dilated cardiomyopathy. *Circulation* (2004) 110(2):155–62. doi:10.1161/01.CIR.0000135591.37759.AF

27. Opitz CA, Leake MC, Makarenko I, Benes V, Linke WA. Developmentally regulated switching of titin size alters myofibrillar stiffness in the perinatal heart. *Circ Res* (2004) 94(7):967–75. doi:10.1161/01.RES.0000124301.48193.E1

28. Roberts AM, Ware JS, Herman DS, Schafer S, Baksi J, Bick AG, et al. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci Transl Med* (2015) 7(270):270ra276. doi:10.1126/scitranslmed.3010134

29. Yamasaki R, Wu Y, McNabb M, Greaser M, Labeit S, Granzier H. Protein kinase A phosphorylates titin's cardiac-specific N2B domain and reduces passive tension in rat cardiac myocytes. *Circ Res* (2002) 90(11):1181–8. doi:10.1161/01.RES.0000021115.24712.99

30. Belin RJ, Sumandea MP, Allen EJ, Schoenfelt K, Wang H, Solaro RJ, et al. Augmented protein kinase C-alpha-induced myofilament protein phosphorylation contributes to myofilament dysfunction in experimental congestive heart failure. *Circ Res* (2007) 101(2):195–204. doi:10.1161/CIRCRESAHA.107.148288

31. Hidalgo C, Hudson B, Bogomolovas J, Zhu Y, Anderson B, Greaser M, et al. PKC phosphorylation of titin's PEVK element: a novel and conserved pathway for modulating myocardial stiffness. *Circ Res* (2009) 105(7):631–638, 617 p following 638. doi:10.1161/CIRCRESAHA.109.198465

32. Krüger M, Kötter S, Grützner A, Lang P, Andresen C, Redfield MM, et al. Protein kinase G modulates human myocardial passive stiffness by phosphorylation of the titin springs. *Circ Res* (2009) 104(1):87–94. doi:10.1161/CIRCRESAHA.108.184408

33. Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med* (2012) 18(5):766–73. doi:10.1038/nm.2693

34. Methawasin M, Hutchinson KR, Lee EJ, Smith JE III, Saripalli C, Hidalgo CG, et al. Experimentally increasing titin compliance in a novel mouse model attenuates the Frank-Starling mechanism but has a beneficial effect on diastole. *Circulation* (2014) 129(19):1924–36. doi:10.1161/CIRCULATIONAHA.113.005610

35. Morita H, Seidman J, Seidman CE. Genetic causes of human heart failure. *J Clin Invest* (2005) 115(3):518–26. doi:10.1172/JCI24351

36. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, et al. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med* (2014) 16(8):601–8. doi:10.1038/gim.2013.204

37. Haas J, Frese KS, Peil B, Kloos W, Keller A, Nietsch R, et al. Atlas of the clinical genetics of human dilated cardiomyopathy. *Eur Heart J* (2015) 36(18):1123a–35a. doi:10.1093/eurheartj/ehu301

38. Begay RL, Graw S, Sinagra G, Merlo M, Slavov D, Gowan K, et al. Role of titin missense variants in dilated cardiomyopathy. *J Am Heart Assoc* (2015) 4(11):e002645. doi:10.1161/JAHA.115.002645

39. Muhle-Goll C, Habeck M, Cazorla O, Nilges M, Labeit S, Granzier H. Structural and functional studies of titin's fn3 modules reveal conserved surface patterns and binding to myosin S1 – a possible role in the Frank-Starling mechanism of the heart. *J Mol Biol* (2001) 313(2):431–47. doi:10.1006/jmbi.2001.5017

40. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. *Circulation* (1995) 92(4):785–9. doi:10.1161/01.CIR.92.4.785

41. Lopes LR, Rahman MS, Elliott PM. A systematic review and meta-analysis of genotype-phenotype associations in patients with hypertrophic cardiomyopathy caused by sarcomeric protein mutations. *Heart* (2013) 99(24):1800–11. doi:10.1136/heartjnl-2013-303939

42. Cahill TJ, Ashrafian H, Watkins H. Genetic cardiomyopathies causing heart failure. *Circ Res* (2013) 113(6):660–75. doi:10.1161/CIRCRESAHA.113.300282

43. Lopes LR, Zekavati A, Syrris P, Hubank M, Giambartolomei C, Dalageorgou C, et al. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. *J Med Genet* (2013) 50(4):228–39. doi:10.1136/jmedgenet-2012-101270

44. Peled Y, Gramlich M, Yoskovitz G, Feinberg MS, Afek A, Polak-Charcon S, et al. Titin mutation in familial restrictive cardiomyopathy. *Int J Cardiol* (2014) 171(1):24–30. doi:10.1016/j.ijcard.2013.11.037

45. Brun F, Barnes CV, Sinagra G, Slavov D, Barbati G, Zhu X, et al. Titin and desmosomal genes in the natural history of arrhythmogenic right ventricular cardiomyopathy. *J Med Genet* (2014) 51(10):669–76. doi:10.1136/jmedgenet-2014-102591

46. Basso C, Corrado D, Marcus FI, Nava A, Thiene G. Arrhythmogenic right ventricular cardiomyopathy. *Lancet* (2009) 373(9671):1289–300. doi:10.1016/S0140-6736(09)60256-7

47. Taylor M, Graw S, Sinagra G, Barnes C, Slavov D, Brun F, et al. Genetic variation in titin in arrhythmogenic right ventricular cardiomyopathy-overlap syndromes. *Circulation* (2011) 124(8):876–85. doi:10.1161/CIRCULATIONAHA.110.005405

48. Roncarati R, Viviani Anselmi C, Krawitz P, Lattanzi G, von Kodolitsch Y, Perrot A, et al. Doubly heterozygous LMNA and TTN mutations revealed by exome sequencing in a severe form of dilated cardiomyopathy. *Eur J Hum Genet* (2013) 21(10):1105–11. doi:10.1038/ejhg.2013.16

49. Lin F, Worman HJ. Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J Biol Chem* (1993) 268(22):16321–6.

50. Marian AJ. Modifier genes for hypertrophic cardiomyopathy. *Curr Opin Cardiol* (2002) 17(3):242–52. doi:10.1097/00001573-200205000-00006

51. Tesson F, Saj M, Uvaize MM, Nicolas H, Ploski R, Bilinska Z. Lamin A/C mutations in dilated cardiomyopathy. *Cardiol J* (2014) 21(4):331–42. doi:10.5603/CJ.a2014.0037

52. Taylor MR, Fain PR, Sinagra G, Robinson ML, Robertson AD, Carniel E, et al. Natural history of dilated cardiomyopathy due to lamin A/C gene mutations. *J Am Coll Cardiol* (2003) 41(5):771–80. doi:10.1016/S0735-1097(02)02954-6

53. Broers JL, Ramaekers FC, Bonne G, Yaou RB, Hutchison CJ. Nuclear lamins: laminopathies and their role in premature ageing. *Physiol Rev* (2006) 86(3):967–1008. doi:10.1152/physrev.00047.2005

54. Ho CY, Charron P, Richard P, Girolami F, Van Spaendonck-Zwarts KY, Pinto Y. Genetic advances in sarcomeric cardiomyopathies: state of the art. *Cardiovasc Res* (2015) 105(4):397–408. doi:10.1093/cvr/cvv025

55. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet* (2012) 5(4):391–9. doi:10.1161/CIRCGENETICS.112.962928

56. Hershberger RE. Cardiovascular genetic medicine: evolving concepts, rationale, and implementation. *J Cardiovasc Transl Res* (2008) 1(2):137–43. doi:10.1007/s12265-008-9031-3

57. Mogensen J, van Tintelen JP, Fokstuen S, Elliott P, van Langen IM, Meder B, et al. The current role of next-generation DNA sequencing in routine care of patients with hereditary cardiovascular conditions: a viewpoint paper of the European Society of Cardiology working group on myocardial and pericardial diseases and members of the European Society of Human Genetics. *Eur Heart J* (2015) 36(22):1367–70. doi:10.1093/eurheartj/ehv122

58. Pasotti M, Klersy C, Pilotto A, Marziliano N, Rapezzi C, Serio A, et al. Long-term outcome and risk stratification in dilated cardiolaminopathies. *J Am Coll Cardiol* (2008) 52(15):1250–60. doi:10.1016/j.jacc.2008.06.044

59. LeWinter MM, Granzier HL. Cardiac titin and heart disease. *J Cardiovasc Pharmacol* (2014) 63(3):207–12. doi:10.1097/FJC.0000000000000007

60. Linke WA, Hamdani N. Gigantic business: titin properties and function through thick and thin. *Circ Res* (2014) 114(6):1052–68. doi:10.1161/CIRCRESAHA.114.301286

Understanding the Causes and Implications of Endothelial Metabolic Variation in Cardiovascular Disease through Genome-Scale Metabolic Modeling

Sarah McGarrity[1], Haraldur Halldórsson[2], Sirus Palsson[1,3], Pär I. Johansson[4] and Óttar Rolfsson[1,5]*

[1] Center for Systems Biology, University of Iceland, Reykjavik, Iceland, [2] Department of Pharmacology and Toxicology, School of Health Sciences, University of Iceland, Reykjavik, Iceland, [3] Sinopia Biosciences Inc., San Diego, CA, USA, [4] Section for Transfusion Medicine, Capital Region Blood Bank, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, [5] Department of Biochemistry and Molecular Biology, School of Health Sciences, University of Iceland, Reykjavik, Iceland

High-throughput biochemical profiling has led to a requirement for advanced data inter-pretation techniques capable of integrating the analysis of gene, protein, and metabolic profiles to shed light on genotype–phenotype relationships. Herein, we consider the current state of knowledge of endothelial cell (EC) metabolism and its connections to cardiovascular disease (CVD) and explore the use of genome-scale metabolic models (GEMs) for integrating metabolic and genomic data. GEMs combine gene expression and metabolic data acting as frameworks for their analysis and, ultimately, afford mecha-nistic understanding of how genetic variation impacts metabolism. We demonstrate how GEMs can be used to investigate CVD-related genetic variation, drug resistance mech-anisms, and novel metabolic pathways in ECs. The application of GEMs in personalized medicine is also highlighted. Particularly, we focus on the potential of GEMs to identify metabolic biomarkers of endothelial dysfunction and to discover methods of stratifying treatments for CVDs based on individual genetic markers. Recent advances in systems biology methodology, and how these methodologies can be applied to understand EC metabolism in both health and disease, are thus highlighted.

Keywords: endothelium, metabolism, personalized/precision medicine, metabolomics, metabolic modeling, genetics

INTRODUCTION

Cardiovascular disease (CVD) includes acute and chronic conditions, such as stroke and coronary heart disease (1). CVD results in a shortened life span and is the biggest cause of death worldwide (1–3). The endothelium is the single cell layer that lines blood vessels and lymphatic system and its dysfunction contributes to the development of CVD (4, 5). Endothelial cells (ECs) play an important role in controlling vascular tone and by secreting or expressing surface molecules, they ensure appro-priate regulation of blood flow, counteracting intravascular activation of platelets, and coagulation (6, 7). Moreover, cardiac ECs have been shown to affect the ventricular myocardium. Thus, the

force-frequency response of cardiac muscle in the presence of increased cardiac workload is blunted after damage to the cardiac endothelium (8).

A vascular surface that normally is thromboresistant, anti-inflammatory, vasodilatory, and antiproliferative can turn into a surface that is thrombogenic, proinflammatory, vasoconstrictive, and stimulatory of smooth muscle cell proliferation. Often this change is reactive and transient restoring vascular homeostasis. However, in diseases such as atherosclerosis, hypertension, and diabetes mellitus (DM) such changes, known as endothelial dysfunction, may be prolonged and critical for disease progression. The extent of pathological metabolic perturbation is determined by an interaction of lifestyle factors, such as diet and exercise with underlying genetic factors (9–12). Consequently, health-care interventions may be more effective if adapted to an individual.

Metabolic modeling offers insights into cellular metabolism (13). Below, we consider endothelial metabolic alterations, their contribution to endothelial dysfunction, and integrated analysis of this information with genome-scale metabolic models (GEMs) to advance personalized health care.
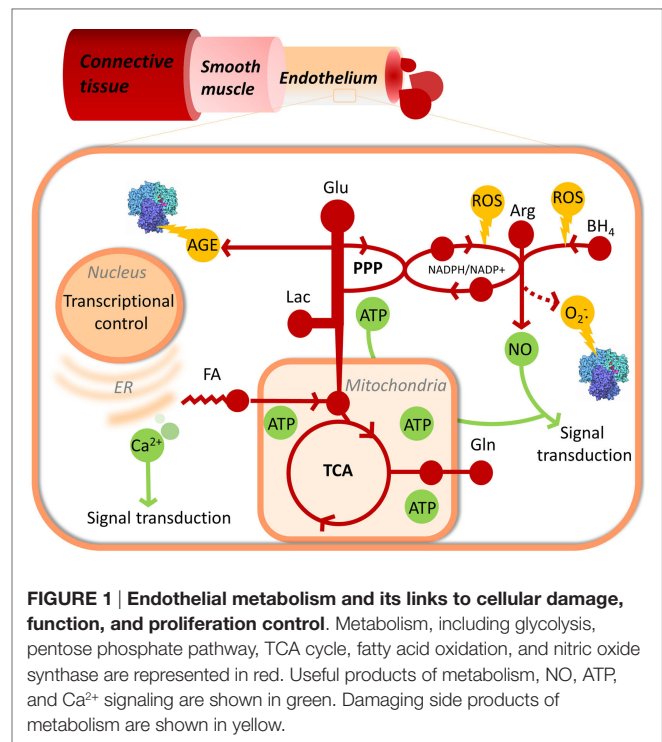
ENDOTHELIAL METABOLISM

Endothelial cell metabolism has been investigated in multiple contexts including angiogenesis, hypoxia, shear stress, glycemia, and response to perturbations with mediators of vascular health including thrombin, sphingosine-1-phosphate, and more (14–19). The endothelium operates with variable nutrient availability and oxygen partial pressures in a manner that is EC subtype specific (20) and results in altered synergy in the oxidation of its core nutrients glucose, fatty acids, and amino acids (17, 21–23) that are reviewed specifically elsewhere (24, 25) but considered collectively here and illustrated in **Figure 1**.

Glycolysis Affects Endothelial Proliferation and Angiogenesis

Endothelial cells oxidize glucose largely by glycolysis, allowing maximal availability of oxygen for transendothelial transport to perivascular cells (26–29). Carbons from glucose are primarily excreted as lactate with only 1 in 200 pyruvate equivalents contributing to oxidative phosphorylation (26). Laminar shear stress, the frictional force created by blood flow, promotes anti-inflammatory, anti-thrombotic, and anti-oxidative properties in ECs and helps to maintain quiescence largely *via* the transcription factor Kruppel-like factor 2 (30) that acts to repress phosphofructokinase-2/fructose-2,6-bisphosphatase-3 (PFKFB3) thereby promoting a quiescent phenotype (16).

In response to angiogenic factors induced by injury or in pathological conditions such as hypoxia, nutrient deprivation, or tissue damage, ECs quickly form new vasculature by sprouting. During vessel sprouting, glycolysis is increased further, mediated by increased activity of PFKFB3, the loss of which impairs vessel formation (26). Increased glycolysis without oxidation of pyruvate relies on lactate dehydrogenase to supply NAD+, and the activity of PFKFB3 is reflected in both intracellular and secreted lactate of ECs (31). Furthermore, lactate is involved in

FIGURE 1 | **Endothelial metabolism and its links to cellular damage, function, and proliferation control**. Metabolism, including glycolysis, pentose phosphate pathway, TCA cycle, fatty acid oxidation, and nitric oxide synthase are represented in red. Useful products of metabolism, NO, ATP, and Ca²⁺ signaling are shown in green. Damaging side products of metabolism are shown in yellow.

PFKFB3-mediated endothelial proliferation, tube formation, and Akt activation providing a plausible explanation for PFKFB3-mediated angiogenesis (31). Lactate dehydrogenase activity also increases with EC subtype proliferation rate. In pulmonary microvascular ECs, rapid angiogenesis is dependent on lactate dehydrogenase A expression (14).

Endothelial-dependent vascular function correlates with blood glucose levels (32–37). In hyperglycemia, glyceraldehyde-3-phosphate dehydrogenase is inactivated, impeding glycolysis (38). A build-up of fructose-6-phosphate, a glycolytic intermediate, impacts hexosamine biosynthesis generating N-acetylglucosamine that glycosylates and modifies angiogenic proteins including Notch and vascular endothelial growth factor receptor 2 (39–45) and, inhibits eNOS (46). Excess glucose also enters the polyol pathway, producing excess advanced glycation end products (AGEs) (47, 48). AGEs alter the binding of erythrocytes and platelets to the endothelium (49, 50), and clinical arterial responsiveness correlates negatively with the ratio of AGEs to soluble receptor of AGEs (51).

Fatty Acid and Amino Acids Metabolism

Fatty acid-binding protein 4 (FABP4) is an intracellular fatty acid chaperone protein that impacts the peroxisome proliferator-activated receptor transcription pathway (52). Circulating levels of FABP4 are associated with endothelial dysfunction in DM patients (53) and increased risk of atherosclerosis and cerebrovascular malformations (54, 55).

Fatty acid oxidation (FAO) accounts for roughly 14% of ATP production in cultured EC (22). Carnitine palmitoyl transferase (CPT1A), a long-chain fatty acid shuttle protein regulated by

AMP-activated protein kinase, is a key point of FAO regulation (22, 56, 57). Palmitate has been shown to contribute carbons to nucleotide formation *via* the tricarboxylic acid (TCA) cycle. When CPT1A was knocked down *in vitro*, vessel sprouting was impaired due to low levels of deoxy ribonucleotides. CPT1A knockdown in mice produced impaired retinal vessel formation (58).

In addition to glucose and fatty acids, amino acids contribute to EC metabolism and function (59). Specifically glutamine fuels anaplerotic reactions *via* the TCA cycle (23, 29, 60). Internalization of glutamine occurs *via* solute carrier family 1 member 5 (23, 29), and inhibition of glutaminase causes premature senescence and reduced proliferation in ECs (61). The most intensely investigated amino acid with respect to endothelial dysfunction is, however, arginine in the context of its conversion to the vasorelaxant nitric oxide (NO) by endothelial nitric oxide synthase (eNOS).

Endothelial Nitric Oxide Is Important to Vascular Function and Its Production Is Affected by Genetic and Metabolic Factors

In addition to causing vasorelaxation, NO affects smooth muscle cell proliferation, aggregation and adhesion of platelets and leukocytes, important processes to atherosclerosis and other CVD (62, 63). When eNOS has insufficient arginine, a result of competition with arginase, and/or lacks the cofactor tetrahydrobiopterin, it produces reactive oxygen species (ROS) instead of the products NO and citrulline – in a pathological state known as uncoupling (64–71). Furthermore, the pressure of O_2^- causes rapid inactivation of endothelium-derived NO (72). Indeed, arginase and eNOS activities and genotypes in addition to tetrahydrobiopterin levels have all been linked to endothelial function (73–76).

Altered NOS activity due to inhibition by asymmetric dimethylarginine (ADMA) encourages NOS uncoupling leading to endothelial dysfunction. ADMA levels, and the ratio of ADMA to arginine, have been connected to several aspects of CVD risk (77–80).

Genetic variation in eNOS affects some measures of recovery of blood flow control in acute myocardial infarction (73). Inhibiting arginase activity, which reduces eNOS uncoupling, is helpful in restoring endothelial function in both coronary artery disease and after ischemia–reperfusion injury (64, 65). Genetic variation in NOS1 has also been linked with CVD in various studies (75, 76). Furthermore, the ROS scavenger methionine sulfoxide reductase A, important to reducing the effect of uncoupled NOS and other ROS, is affected by genetic variation relevant to coronary artery disease risk (81, 82).

Interestingly, the extracellular presence of certain amino acids – ornithine, ʟ-lysine, ʟ-homoarginine, ʟ-glutamine, ʟ-leucine, or ʟ-serine – decreases NO and increases endothelium-dependent vascular resistance. This effect is reversible by adding arginine to the medium and was shown to be dependent on y⁺L and y⁺ family amino acid transporters (83).

DECODING ENDOTHELIAL METABOLISM AND FUNCTION THROUGH COMPUTATIONAL MODELING

The previous section highlights the complexity of the contribution of metabolism to endothelial dysfunction. Importantly, some of the most common human metabolic gene alterations impact enzymes that are of importance to endothelial metabolic phenotypes. These include pyruvate kinase and (84) glucose-6-phosphate dehydrogenase, which alters CVD risk (85), in addition to those already mentioned above. The variability of the effect of these mutations on cardiovascular phenotypes highlights the problem of untangling complex genetic diseases (12). This complexity is aggravated by lifestyle choices that impact the expression and activity of these genes (9, 86–89). How altered gene expression and the environment combine to advance CVD can, however, be explored on the metabolic level, through metabolic systems analysis using genome-scale models of endothelial metabolism. For CVD research, genome-scale modeling promises to contribute to the definition of endothelial metabolism under different physiological conditions, allow the differentiation of individual endothelial metabolic phenotypes that can be related to CVD states and ultimately contribute to individualized therapy. In the following sections, we explain the concept of GEMs, their current and potential applications toward increasing the understanding of endothelial metabolism, and how this could lead to novel discoveries to combat CVD on the individual level.

GEMs Provide Snapshots of Metabolism

Genome-scale metabolic models are computational models that can be used to describe and investigate the metabolic flux phenotype of a cell based on disparate biochemical information. GEMs are built from biochemical component knowledge-bases, also termed biochemical network reconstructions (90). Reconstructions are organism specific and account for genetic, and biochemical components, and their interactions, based on annotated biological information sourced from literature. All metabolic reactions and metabolites contained within a reconstruction can be represented as a numerical matrix, which is comprised of the stoichiometric factors of reactants and products of each metabolic reaction. In this format, the metabolome is subject to computational research allowing metabolic reaction flux at steady state through metabolic pathways to be computed (91).

Genome-scale metabolic reconstructions aim to account for as many as possible biochemical interactions that have been described in an organism (e.g., a human). While reconstructions afford a mechanistic description of genotype–phenotype relationships, they are not context specific. However, when constrained with cell or context-specific data, for example gene expression information of ECs, reconstructions afford GEMs that are descriptive of the biological event and cell of interest. Gene expression data of a HUVEC cells at normoxia vs. hypoxia would for instance generate two GEMs based on the same reconstruction thereby providing two snapshots descriptive of metabolic flux through reactions as defined by the two expression datasets. Essentially, reconstructions define the biochemical components

of an organism, while context-specific polyomic data are required to generate a GEM of a particular cell or cellular event. Genomic, proteomic, and/or metabolomic fingerprints can thus be analyzed and compared within the context of GEMs (92).

The methodology of building, curating, and analyzing reconstructions and GEMs is commonly referred to as constraint-based analysis. Various software has been developed to facilitate constraint-based analysis including the COBRA and RAVEN toolbox's for Matlab, Merlin and CORDA (93–96). Detailed protocols describing the necessary stages of building and curation are established (90, 97–99). Ultimately, constraint-based analysis of GEMs allows holistic exploration of metabolic phenotypes *in silico* and affords realistic hypotheses of biochemical mechanisms (92). In the past 5 years, multiple applications of GEMs descriptive of human metabolism have materialized that may contribute to the understanding of how genetic and environmental factors collectively contribute to CVD disease phenotypes when applied to endothelial metabolic research.

GEMs Differentiate between Metabolic Phenotypes

In the context of CVD, GEMs that are descriptive of healthy and CVD endothelial metabolism can be produced. As recently reviewed in Väremo et al. (100), GEMs of various tissues have been built and applied to the investigation of CVD-related disorders, including DM and metabolic syndrome, although not yet endothelium (101–104). Transcriptional changes in cardiomyocytes of DM patients have been analyzed using the myocyte-specific GEM, iMyocyte2419, revealing deregulation of metabolic pathways ultimately linked to dihydro-lipoamide dehydrogenase, a unique characteristic of myocyte response in DM (101).

Genome-scale metabolic models serve as a biomarker discovery tool, and a tool to discover potentially "druggable" metabolic (105). Computational techniques exist that predict the pathways likely to be responsible for differences between two metabolic states, identifying these differences allows reactions, linked to genes in a GEM, to be selected as drug targets, for example in hepatocellular carcinoma and Alzheimer's, or metabolites to be identified as potential biomarkers for example for drug resistance in ovarian cancer (100, 106–108). Changes due in FAO in ECs leading to alterations in EC permeability – clinically important to sepsis – have been detected using a GEM. Altering FAO using drugs was shown to alter permeability, which may be clinically useful (109), future discoveries of this type may be linked to NO synthesis or clotting factor production useful for modulating CVD risk factors.

GEMs Can Define Endothelial Metabolism

Genome-scale metabolic models that are descriptive of core endothelial metabolism have already been produced. Patella et al recently used endothelial proteomic data to constrain the human reconstruction, Recon 1 (110), to generate a GEM that describes EC cell core metabolism during tube formation in matrigel (109). FAO was identified as an area of metabolism that is altered during tube formation. CPT1A inhibition affects ATP production *via* the TCA cycle and oxidative phosphorylation.
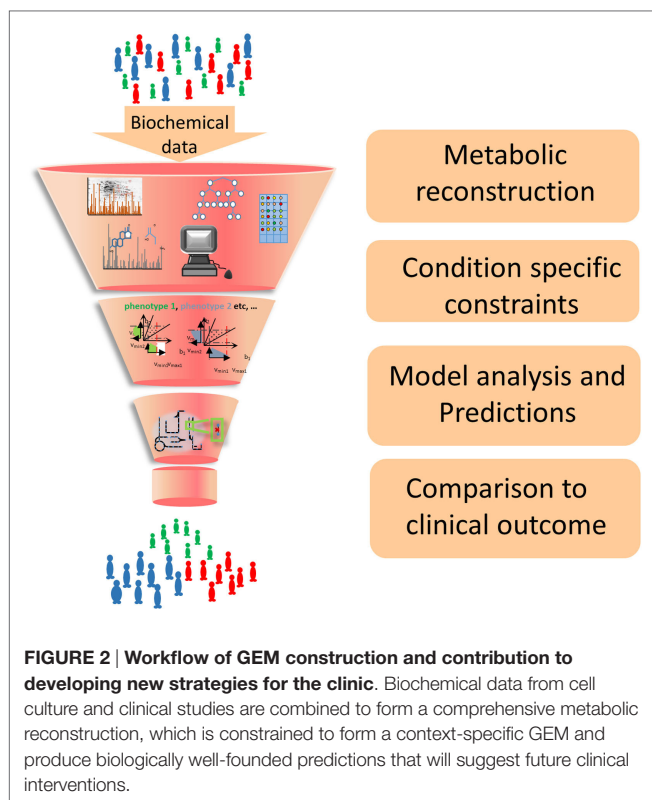
Downstream, this alters Ca^{2+} signaling and junctional proteins *via* phospho-signaling to alter endothelial permeability, which were partially reversed by pyruvate supplementation (109). Automated GEMs have also been generated for colon and cerebral cortex ECs (111), though these models were not applied to CVD research.

Although automatically generated GEMs of EC metabolism have been used to reveal basal endothelial metabolic pathway usage, further curation and validation of EC GEMs would be beneficial. Investigations of vascular endothelial metabolism in different conditions and with different genetic backgrounds could be achieved, allowing genetic variation outside the context of core energy metabolism to be queried. For example, due to the inherent connectivity of metabolic reactions within GEMs, alterations in the release of sphingosine-1-P (a sphingolipid involved in vascular and immune signaling pathways) from ECs could be hypothesized and related to alterations in core energy metabolism induced by global metabolic expression profiles. The release of sphingosine-1-P from ECs and its contribution to individual vascular health could thus be proposed on biochemical alterations on the systems level as opposed to mutations in sphingosine kinase alone.

GEMs Can Be Personalized to Account for Individual Genetic Variation

Computational modeling can contribute to decisions regarding the suitability of a treatment for individual patients. GEMs could be produced for individuals based on genomics and subsequently used to stratify patients and personalize medical interventions for CVD. GEMs maybe based on generalized transcriptomic data from a pool of samples from a cell type (112) or a set of models may be created from individual samples and comparing the metabolic phenotypes predicted by each, allowing links between metabolism and broader phenotype, such as drug resistance in cancer cells, to be explored and may lead to insights about predictive biomarkers and druggable targets (108, 113, 114). Various algorithms for selecting active reactions for context-specific models based on transcriptomic and proteomic data are available including INIT and iMAT. These approaches have differing strengths and weaknesses that have been described and compared elsewhere (98).

Individualized hepatocellular carcinoma models have been used to predict patient outcomes based on the predicted production of acetate, identified as a key metabolic pathway for survival (114). Twenty-four individualized GEMs of erythrocytes were created based on genetic and metabolic data. These captured altered dynamics of erythrocyte metabolism and allowed the identification of individuals at risk to drug-induced anemia based upon their genomic sequence (115). These examples highlight a potential workflow, exemplified in **Figure 2**, to contribute to the personalization and stratification of medical treatments in the clinic. In the future, it is envisioned that an EC GEM could be used in a similar fashion by comparing GEMs CVD patients and healthy individuals to identify key metabolic changes to CVD for example those that increase production of atherosclerotic plaques.

FIGURE 2 | Workflow of GEM construction and contribution to developing new strategies for the clinic. Biochemical data from cell culture and clinical studies are combined to form a comprehensive metabolic reconstruction, which is constrained to form a context-specific GEM and produce biologically well-founded predictions that will suggest future clinical interventions.

computational models of various aspects of vascular biology have also been made. These address some important biophysical parameters that are currently outside the scope of GEM modeling. This includes assessing the effects of shear stress on blood vessel reactivity and growth as well as the effects on blood cell/endothelium interactions of flow (116–122). Models describing the effects of circulation on endothelial metabolites have also been built (123). Endothelial NO interactions (124–126), Ca^{2+} signaling (127) along with protein (128) and mechanical (119) signaling have also been addressed with computational modeling. Models have been individualized using patient data and have explored the effects of stenting on blood flow (129–132).

Integrating biophysical and signaling parameters with GEMs would generate a more complete understanding of the role of endothelial metabolism for CVD. In addition, these future GEMs would allow retrospective analysis of biophysical and genomic data that have been generated in the last few decades from population studies (86, 133), whose analysis is currently confined to multivariate statistical and comparative analysis techniques for the identification of CVD risk factors. Such an effort could allow, for example, *in silico* querying of the effect of LDL deposition on global endothelial metabolism. Indeed, computational analysis of LDL metabolism has already proposed novel approaches to combat CVD (134–136).

Realistic computational predictions of the effects of genetic and environmental perturbations on endothelial metabolism are possible and beneficial. There has been some exploration of CVD with GEMs and analysis of EC metabolism with GEMs; however, the full potential of this technique is only just beginning to be explored. Existing and future models will allow clinicians and researchers to investigate variable endothelial function *in silico* in a data-driven manner, to optimize future clinical interventions.

CONCLUSION

Developing personalized CVD therapeutic interventions relies on the ability to account for genotypic and phenotypic variation. Variability in disease phenotypes can be captured and understood in the context of GEM's to facilitate this process.

Genome-scale metabolic models provide an integrated approach in studying EC metabolism. They allow analysis of the multiple factors affecting ECs in the body, facilitating the exploration of the relationship of genotype to metabolic phenotype. This offers the possibility of producing personalized predictions of CVD risk and treatment, that account for both genetic and lifestyle factors. Currently, GEMs are the only biochemical model type that can account for both of these factors within a predictive modeling framework (92).

Genome-scale metabolic models are only one type of model used to account for EC function. Focused and mechanistic

AUTHOR CONTRIBUTIONS

SM and OR wrote the manuscript and conceived the ideology. HH, SP, and PJ conceived the ideology and contributed to writing of the manuscript.

ACKNOWLEDGMENTS

REFERENCES

1. WHO. *Cardiovascular Diseases (CVDs)*. World Health Organization (2016). Available from: http://www.who.int/mediacentre/factsheets/fs317/en/
2. Field JM, Hazinski MF, Sayre MR, Chameides L, Schexnayder SM, Hemphill R, et al. Part 1: executive summary: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation* (2010) **122**(Suppl 3):640–57. doi:10.1161/CIRCULATIONAHA.110.970889
3. Townsend N, Nichols M, Scarborough P, Rayner M. Cardiovascular disease in Europe – epidemiological update 2015. *Eur Heart J* (2015) **36**(40):ehv428. doi:10.1093/eurheartj/ehv428
4. Landmesser U, Drexler H. The clinical significance of endothelial dysfunction. *Curr Opin Cardiol* (2005) **20**(6):547–51. doi:10.1097/01.hco.0000179821.11071.79
5. Vanhoutte PM, Shimokawa H, Feletou M, Tang EHC. Endothelial dysfunction and vascular disease – a thirthieth anniversary update. *Acta Physiol (Oxf)* (2015). doi:10.1111/apha.12646
6. Kazmi R, Boyce S, Lwaleed B. Homeostasis of hemostasis: the role of endothelium. *Semin Thromb Hemost* (2015) **41**(6):549–55. doi:10.1055/s-0035-1556586
7. Santillo M, Colantuoni A, Mondola P, Guida B, Damiano S. NOX signaling in molecular cardiovascular mechanisms involved in the blood pressure homeostasis. *Front Physiol* (2015) **6**:194. doi:10.3389/fphys.2015.00194

8. Shen X, Tan Z, Zhong X, Tian Y, Wang X, Yu B, et al. Endocardial endothelium is a key determinant of force-frequency relationship in rat ventricular myocardium. *J Appl Physiol* (2013) **115**(3):383–93. doi:10.1152/japplphysiol.01415.2012

9. Choi B, Steiss D, Garcia-Rivas J, Kojaku S, Schnall P, Dobson M, et al. Comparison of body mass index with waist circumference and skinfold-based percent body fat in firefighters: adiposity classification and associations with cardiovascular disease risk factors. *Int Arch Occup Environ Health* (2016) **89**(3):435–48. doi:10.1007/s00420-015-1082-6

10. Hwang M-H, Kim S. Type 2 diabetes: endothelial dysfunction and exercise. *J Exerc Nutrition Biochem* (2014) **18**(3):239–47. doi:10.5717/jenb.2014.18.3.239

11. Kwaśniewska M, Kozińska J, Dziankowska-Zaborszczyk E, Kostka T, Jegier A, Rębowska E, et al. The impact of long-term changes in metabolic status on cardiovascular biomarkers and microvascular endothelial function in middle-aged men: a 25-year prospective study. *Diabetol Metab Syndr* (2015) **7**:81. doi:10.1186/s13098-015-0074-8

12. Hamrefors V. Common genetic risk factors for coronary artery disease: new opportunities for prevention? *Clin Physiol Funct Imaging* (2015). doi:10.1111/cpf.12289

13. Rolfsson O, Palsson B, Thiele I. The human metabolic reconstruction Recon 1 directs hypotheses of novel human metabolic functions. *BMC Syst Biol* (2011) **5**:155. doi:10.1186/1752-0509-5-155

14. Parra-Bonilla G, Alvarez DF, Alexeyev M, Vasauskas A, Stevens T. Lactate dehydrogenase a expression is necessary to sustain rapid angiogenesis of pulmonary microvascular endothelium. *PLoS One* (2013) **8**(9):e75984. doi:10.1371/journal.pone.0075984

15. Tuleta I, França CN, Wenzel D, Fleischmann B, Nickenig G, Werner N, et al. Intermittent hypoxia impairs endothelial function in early preatherosclerosis. *Adv Exp Med Biol* (2015) **858**:1–7. doi:10.1007/5584_2015_114

16. Doddaballapur A, Michalik KM, Manavski Y, Lucas T, Houtkooper RH, You X, et al. Laminar shear stress inhibits endothelial cell metabolism via KLF2-mediated repression of PFKFB3. *Arterioscler Thromb Vasc Biol* (2015) **35**(1):137–45. doi:10.1161/ATVBAHA.114.304277

17. Zhang Z, Apse K, Pang J, Stanton RC. High glucose inhibits glucose-6-phosphate dehydrogenase via cAMP in aortic endothelial cells. *J Biol Chem* (2000) **275**(51):40042–7. doi:10.1074/jbc.M007505200

18. Halldórsson H, Thors B, Thorgeirsson G. Thrombin or Ca(++)-ionophore-mediated fall in endothelial ATP levels independent of poly(ADP-Ribose) polymerase activity and NAD levels – comparison with the effects of hydrogen peroxide. *Nucleosides Nucleotides Nucleic Acids* (2015) **34**(4):246–57. doi:10.1080/15257770.2014.984072

19. Mahajan-Thakur S, Böhm A, Jedlitschky G, Schrör K, Rauch BH. Sphingosine-1-phosphate and its receptors: a mutual link between blood coagulation and inflammation. *Mediators Inflamm* (2015) **2015**:831059. doi:10.1155/2015/831059

20. Viator RJ, Khader H, Hingorani N, Long S, Solodushko V, Fouty B. Hypoxia-induced increases in glucose uptake do not cause oxidative injury or advanced glycation end-product (AGE) formation in vascular endothelial cells. *Physiol Rep* (2015) **3**(7):e12460. doi:10.14814/phy2.12460

21. Koziel A, Woyda-Ploszczyca A, Kicinska A, Jarmuszkiewicz W. The influence of high glucose on the aerobic metabolism of endothelial EA.hy926 cells. *Pflugers Arch* (2012) **464**(6):657–69. doi:10.1007/s00424-012-1156-1

22. Dagher Z, Ruderman N, Tornheim K, Ido Y. Acute regulation of fatty acid oxidation and AMP-activated protein kinase in human umbilical vein endothelial cells. *Circ Res* (2001) **88**(12):1276–82. doi:10.1161/hh1201.092998

23. Lohmann R, Souba WW, Bode BP. Rat liver endothelial cell glutamine transporter and glutaminase expression contrast with parenchymal cells. *Am J Physiol* (1999) **276**(3 Pt 1):G743–50.

24. Eelen G, de Zeeuw P, Simons M, Carmeliet P. Endothelial cell metabolism in normal and diseased vasculature. *Circ Res* (2015) **116**(7):1231–44. doi:10.1161/CIRCRESAHA.116.302855

25. Tabas I, García-Cardeña G, Owens G. Recent insights into the cellular biology of atherosclerosis. *J Cell Biol* (2015) **209**(1):13–22. doi:10.1083/jcb.201412052

26. De Bock K, Georgiadou M, Schoors S, Kuchnio A, Wong B, Cantelmo A, et al. Role of PFKFB3-driven glycolysis in vessel sprouting. *Cell* (2013) **154**(3):651–63. doi:10.1016/j.cell.2013.06.037

27. Dobrina A, Rossi F. Metabolic properties of freshly isolated bovine endothelial cells. *Biochim Biophys Acta* (1983) **762**(2):295–301. doi:10.1016/0167-4889(83)90084-8

28. Krützfeldt A, Spahr R, Mertens S, Siegmund B, Piper HM. Metabolism of exogenous substrates by coronary endothelial cells in culture. *J Mol Cell Cardiol* (1990) **22**(12):1393–404. doi:10.1016/0022-2828(90)90984-A

29. Leighton B, Curi R, Hussein A, Newsholme EA. Maximum activities of some key enzymes of glycolysis, glutaminolysis, Krebs cycle and fatty acid utilization in bovine pulmonary endothelial cells. *FEBS Lett* (1987) **225**(1–2):93–6. doi:10.1016/0014-5793(87)81137-7

30. Parmar KM, Larman HB, Dai G, Zhang Y, Wang ET, Moorthy SN, et al. Integration of flow-dependent endothelial phenotypes by Kruppel-like factor 2. *J Clin Invest* (2006) **116**(1):49–58. doi:10.1172/JCI24787

31. Xu Y, An X, Guo X, Habtetsion TG, Wang Y, Xu X, et al. Endothelial PFKFB3 plays a critical role in angiogenesis. *Arterioscler Thromb Vasc Biol* (2014) **34**(6):1231–9. doi:10.1161/ATVBAHA.113.303041

32. Zhang X-G, Zhang Y-Q, Zhao D-K, Wu J-X, Zhao J, Jiao X-M, et al. Relationship between blood glucose fluctuation and macrovascular endothelial dysfunction in type 2 diabetic patients with coronary heart disease. *Eur Rev Med Pharmacol Sci* (2014) **18**(23):3593–600.

33. Temelkova-Kurktschiev TS, Koehler C, Henkel E, Leonhardt W, Fuecker K, Hanefeld M. Postchallenge plasma glucose and glycemic spikes are more strongly associated with atherosclerosis than fasting glucose or HbA1c level. *Diabetes Care* (2000) **23**(12):1830–4. doi:10.2337/diacare.23.12.1830

34. Azuma K, Kawamori R, Toyofuku Y, Kitahara Y, Sato F, Shimizu T, et al. Repetitive fluctuations in blood glucose enhance monocyte adhesion to the endothelium of rat thoracic aorta. *Arterioscler Thromb Vasc Biol* (2006) **26**(10):2275–80. doi:10.1161/01.ATV.0000239488.05069.03

35. Torimoto K, Okada Y, Mori H, Tanaka Y. Relationship between fluctuations in glucose levels measured by continuous glucose monitoring and vascular endothelial dysfunction in type 2 diabetes mellitus. *Cardiovasc Diabetol* (2013) **12**:1. doi:10.1186/1475-2840-12-1

36. Jiao X-M, Zhang X-G, Xu XU-P, Yi C, Bin C, Cheng Q-P, et al. Blood glucose fluctuation aggravates lower extremity vascular disease in type 2 diabetes. *Eur Rev Med Pharmacol Sci* (2014) **18**(14):2025–30.

37. Service FJ, Molnar GD, Rosevear JW, Ackerman E, Gatewood LC, Taylor WF. Mean amplitude of glycemic excursions, a measure of diabetic instability. *Diabetes* (1970) **19**(9):644–55. doi:10.2337/diab.19.9.644

38. Du X, Matsumura T, Edelstein D, Rossetti L, Zsengellér Z, Szabó C, et al. Inhibition of GAPDH activity by poly(ADP-ribose) polymerase activates three major pathways of hyperglycemic damage in endothelial cells. *J Clin Invest* (2003) **112**(7):1049–57. doi:10.1172/JCI18127

39. Kim DH, Seok YM, Kim IK, Lee I-K, Jeong SY, Jeoung NH. Glucosamine increases vascular contraction through activation of RhoA/Rho kinase pathway in isolated rat aorta. *BMB Rep* (2011) **44**(6):415–20. doi:10.5483/BMBRep.2011.44.6.415

40. Rajapakse AG, Ming X-F, Carvas JM, Yang Z. O-linked beta-N-acetylglucosamine during hyperglycemia exerts both anti-inflammatory and pro-oxidative properties in the endothelial system. *Oxid Med Cell Longev* (2009) **2**(3):172–5. doi:10.4161/oxim.2.3.8482

41. Rajapakse AG, Ming X-F, Carvas JM, Yang Z. The hexosamine biosynthesis inhibitor azaserine prevents endothelial inflammation and dysfunction under hyperglycemic condition through antioxidant effects. *Am J Physiol Heart Circ Physiol* (2009) **296**(3):H815–22. doi:10.1152/ajpheart.00756.2008

42. Wu G, Haynes TE, Yan W, Meininger CJ. Presence of glutamine:fructose-6-phosphate amidotransferase for glucosamine-6-phosphate synthesis in endothelial cells: effects of hyperglycaemia and glutamine. *Diabetologia* (2001) **44**(2):196–202. doi:10.1007/s001250051599

43. Slawson C, Copeland RJ, Hart GW. O-GlcNAc signaling: a metabolic link between diabetes and cancer? *Trends Biochem Sci* (2010) **35**(10):547–55. doi:10.1016/j.tibs.2010.04.005

44. Benedito R, Roca C, Sörensen I, Adams S, Gossler A, Fruttiger M, et al. The notch ligands Dll4 and Jagged1 have opposing effects on angiogenesis. *Cell* (2009) **137**(6):1124–35. doi:10.1016/j.cell.2009.03.025

45. Vaisman N, Gospodarowicz D, Neufeld G. Characterization of the receptors for vascular endothelial growth factor. *J Biol Chem* (1990) **265**(32):19461–6.

46. Federici M, Menghini R, Mauriello A, Hribal ML, Ferrelli F, Lauro D, et al. Insulin-dependent activation of endothelial nitric oxide synthase is impaired

by O-linked glycosylation modification of signaling proteins in human coronary endothelial cells. *Circulation* (2002) **106**(4):466–72. doi:10.1161/01. CIR.0000023043.02648.51

47. Du XL, Edelstein D, Dimmeler S, Ju Q, Sui C, Brownlee M. Hyperglycemia inhibits endothelial nitric oxide synthase activity by posttranslational modification at the Akt site. *J Clin Invest* (2001) **108**(9):1341–8. doi:10.1172/JCI11235

48. Wautier J-L, Schmidt AM. Protein glycation: a firm link to endothelial cell dysfunction. *Circ Res* (2004) **95**(3):233–8. doi:10.1161/01. RES.0000137876.28454.64

49. Wautier J-L, Wautier M-P. Erythrocytes and platelet adhesion to endothelium are mediated by specialized molecules. *Clin Hemorheol Microcirc* (2004) **30**(3–4):181–4.

50. Vlassara H, Fuh H, Donnelly T, Cybulsky M. Advanced glycation endproducts promote adhesion molecule (VCAM-1, ICAM-1) expression and atheroma formation in normal rabbits. *Mol Med* (1995) **1**(4):447–56.

51. Kajikawa M, Nakashima A, Fujimura N, Maruhashi T, Iwamoto Y, Iwamoto A, et al. Ratio of serum levels of AGEs to soluble form of RAGE is a predictor of endothelial function. *Diabetes Care* (2015) **38**(1):119–25. doi:10.2337/dc14-1435

52. Makowski L, Brittingham KC, Reynolds JM, Suttles J, Hotamisligil GS. The fatty acid-binding protein, aP2, coordinates macrophage cholesterol trafficking and inflammatory activity. Macrophage expression of aP2 impacts peroxisome proliferator-activated receptor gamma and IkappaB kinase activities. *J Biol Chem* (2005) **280**(13):12888–95. doi:10.1074/jbc.M413788200

53. Aragonès G, Ferré R, Lázaro I, Cabré A, Plana N, Merino J, et al. Fatty acid-binding protein 4 is associated with endothelial dysfunction in patients with type 2 diabetes. *Atherosclerosis* (2010) **213**(1):329–31. doi:10.1016/j.atherosclerosis.2010.07.026

54. Yeung DCY, Xu A, Cheung CWS, Wat NMS, Yau MH, Fong CHY, et al. Serum adipocyte fatty acid-binding protein levels were independently associated with carotid atherosclerosis. *Arterioscler Thromb Vasc Biol* (2007) **27**(8):1796–802. doi:10.1161/ATVBAHA.107.146274

55. Cataltepe S, Arikan MC, Liang X, Smith TW, Cataltepe O. Fatty acid binding protein 4 expression in cerebral vascular malformations: implications for vascular remodelling. *Neuropathol Appl Neurobiol* (2015) **41**(5):646–56. doi:10.1111/nan.12159

56. Dagher Z, Ruderman N, Tornheim K, Ido Y. The effect of AMP-activated protein kinase and its activator AICAR on the metabolism of human umbilical vein endothelial cells. *Biochem Biophys Res Commun* (1999) **265**(1):112–5. doi:10.1006/bbrc.1999.1635

57. Zabielska MA, Borkowski T, Slominska EM, Smolenski RT. Inhibition of AMP deaminase as therapeutic target in cardiovascular pathology. *Pharmacol Rep* (2015) **67**(4):682–8. doi:10.1016/j.pharep.2015.04.007

58. Schoors S, Bruning U, Missiaen R, Queiroz KCS, Borgers G, Elia I, et al. Fatty acid carbon is essential for dNTP synthesis in endothelial cells. *Nature* (2015) **520**(7546):192–7. doi:10.1038/nature14362

59. Pan M, Fischer CP, Wasa M, Lukaszewicz G, Stevens BR, Bode BP, et al. Amino acid metabolism and the vascular endothelium: regulation and disease implications. *Shock* (1995) **4**(2):79–88. doi:10.1097/00024382-199508000-00001

60. Wu G, Haynes TE, Li H, Meininger CJ. Glutamine metabolism in endothelial cells: ornithine synthesis from glutamine via pyrroline-5-carboxylate synthase. *Comp Biochem Physiol Part A Mol Integr Physiol* (2000) **126**(1):115–23. doi:10.1016/S1095-6433(00)00196-3

61. Unterluggauer H, Mazurek S, Lener B, Hütter E, Eigenbrodt E, Zwerschke W, et al. Premature senescence of human endothelial cells induced by inhibition of glutaminase. *Biogerontology* (2008) **9**(4):247–59. doi:10.1007/s10522-008-9134-x

62. Kawashima S, Yokoyama M. Dysfunction of endothelial nitric oxide synthase and atherosclerosis. *Arterioscler Thromb Vasc Biol* (2004) **24**(6):998–1005. doi:10.1161/01.ATV.0000125114.88079.96

63. Yang Y, Wu Z, Meininger CJ, Wu G. ʟ-Leucine and NO-mediated cardiovascular function. *Amino Acids* (2015) **47**(3):435–47. doi:10.1007/s00726-014-1904-y

64. Kovamees O, Shemyakin A, Eriksson M, Angelin B, Pernow J. Arginase inhibition improves endothelial function in patients with familial hypercholesterolaemia irrespective of their cholesterol levels. *J Intern Med* (2015). doi:10.1111/joim.12461

65. Kövamees O, Shemyakin A, Pernow J. Effect of arginase inhibition on ischemia-reperfusion injury in patients with coronary artery disease with and without diabetes mellitus. *PLoS One* (2014) **9**(7):e103260. doi:10.1371/journal.pone.0103260

66. Daff S. NO synthase: structures and mechanisms. *Nitric Oxide* (2010) **23**(1):1–11. doi:10.1016/j.niox.2010.03.001

67. Vásquez-Vivar J, Kalyanaraman B, Martásek P, Hogg N, Masters BS, Karoui H, et al. Superoxide generation by endothelial nitric oxide synthase: the influence of cofactors. *Proc Natl Acad Sci U S A* (1998) **95**(16):9220–5. doi:10.1073/pnas.95.16.9220

68. Vásquez-Vivar J, Martásek P, Whitsett J, Joseph J, Kalyanaraman B. The ratio between tetrahydrobiopterin and oxidized tetrahydrobiopterin analogues controls superoxide release from endothelial nitric oxide synthase: an EPR spin trapping study. *Biochem J* (2002) **362**(Pt 3):733–9. doi:10.1042/bj3620733

69. Ali ZA, Rinze R, Douglas G, Hu Y, Xiao Q, Qi W, et al. Tetrahydrobiopterin determines vascular remodeling through enhanced endothelial cell survival and regeneration. *Circulation* (2013) **128**(11 Suppl 1):S50–8. doi:10.1161/CIRCULATIONAHA.112.000249

70. Xia Y, Tsai AL, Berka V, Zweier JL. Superoxide generation from endothelial nitric-oxide synthase. A Ca2+/calmodulin-dependent and tetrahydrobiopterin regulatory process. *J Biol Chem* (1998) **273**(40):25804–8. doi:10.1074/jbc.273.40.25804

71. Landmesser U, Dikalov S, Price SR, McCann L, Fukai T, Holland SM, et al. Oxidation of tetrahydrobiopterin leads to uncoupling of endothelial cell nitric oxide synthase in hypertension. *J Clin Invest* (2003) **111**(8):1201–9. doi:10.1172/JCI200314172

72. Paolocci N, Biondi R, Bettini M, Lee C-I, Berlowitz CO, Rossi R, et al. Oxygen radical-mediated reduction in basal and agonist-evoked NO release in isolated rat heart. *J Mol Cell Cardiol* (2001) **33**(4):671–9. doi:10.1006/jmcc.2000.1334

73. Machado-Silva W, Alfinito-Kreis R, Carvalho LSF, Quinaglia-E-Silva JC, Almeida OLR, Brito CJ, et al. Endothelial nitric oxide synthase genotypes modulate peripheral vasodilatory properties after myocardial infarction. *Gene* (2015) **568**(2):165–9. doi:10.1016/j.gene.2015.05.042

74. Umman B, Cakmakoglu B, Cincin ZB, Kocaaga M, Emet S, Tamer S, et al. Identification of gene variants related to the nitric oxide pathway in patients with acute coronary syndrome. *Gene* (2015) **574**(1):76–81. doi:10.1016/j.gene.2015.07.081

75. Levinsson A, Olin A-C, Björck L, Rosengren A, Nyberg F. Nitric oxide synthase (NOS) single nucleotide polymorphisms are associated with coronary heart disease and hypertension in the INTERGENE study. *Nitric Oxide* (2014) **39**:1–7. doi:10.1016/j.niox.2014.03.164

76. Manso H, Krug T, Sobral J, Albergaria I, Gaspar G, Ferro JM, et al. Variants within the nitric oxide synthase 1 gene are associated with stroke susceptibility. *Atherosclerosis* (2012) **220**(2):443–8. doi:10.1016/j.atherosclerosis.2011.11.011

77. Stamboul K, Lorin J, Lorgis L, Guenancia C, Beer J-C, Touzery C, et al. Atrial fibrillation is associated with a marker of endothelial function and oxidative stress in patients with acute myocardial infarction. *PLoS One* (2015) **10**(7):e0131439. doi:10.1371/journal.pone.0131439

78. Xuan C, Tian Q-W, Li H, Zhang B-B, He G-W, Lun L-M. Levels of asymmetric dimethylarginine (ADMA), an endogenous nitric oxide synthase inhibitor, and risk of coronary artery disease: a meta-analysis based on 4713 participants. *Eur J Prev Cardiol* (2016) **23**(5):502–10. doi:10.1177/2047487315586094

79. Shivkar RR, Abhang SA. Ratio of serum asymmetric dimethyl arginine (ADMA)/nitric oxide in coronary artery disease patients. *J Clin Diagn Res* (2014) **8**(8):CC04–6. doi:10.7860/JCDR/2014/7849.4665

80. Masaki N, Hakuno D, Toya T, Shiraishi Y, Kujiraoka T, Namba T, et al. Association between brachial-ankle pulse wave velocity and the ratio of ʟ-arginine to asymmetric dimethylarginine in patients undergoing coronary angiography. *J Cardiol* (2015) **65**(4):311–7. doi:10.1016/j.jjcc.2014.06.005

81. Tabas I, Garcia-Cardena G, Owens GK. The cell biology of disease: recent insights into the cellular biology of atherosclerosis. *J Cell Biol* (2015) **209**(1):13–22. doi:10.1083/jcb.201412052

82. García-Bermúdez M, López-Mejías R, González-Juanatey C, Castañeda S, Miranda-Filloy JA, Blanco R, et al. Association of the methionine sulfoxide reductase A rs10903323 gene polymorphism with cardiovascular disease in

patients with rheumatoid arthritis. *Scand J Rheumatol* (2012) **41**(5):350–3. doi:10.3109/03009742.2012.677063

83. Kakoki M, Kim H-S, Edgell C-JS, Maeda N, Smithies O, Mattson DL. Amino acids as modulators of endothelium-derived nitric oxide. *Am J Physiol Renal Physiol* (2006) **291**(2):F297–304. doi:10.1152/ajprenal.00417.2005

84. Carey PJ, Chandler J, Hendrick A, Reid MM, Saunders PW, Tinegate H, et al. Prevalence of pyruvate kinase deficiency in northern European population in the north of England. Northern Region Haematologists Group. *Blood* (2000) **96**(12):4005–6.

85. Meloni L, Manca MR, Loddo I, Cioglia G, Cocco P, Schwartz A, et al. Glucose-6-phosphate dehydrogenase deficiency protects against coronary heart disease. *J Inherit Metab Dis* (2008) **31**(3):412–7. doi:10.1007/s10545-008-0704-5

86. Tsao CW, Vasan RS. Cohort profile: the Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol* (2015) **44**(6):1800–13. doi:10.1093/ije/dyv337

87. Jankovic N, Geelen A, Streppel MT, de Groot LC, Kiefte-de Jong JC, Orfanos P, et al. WHO guidelines for a healthy diet and mortality from cardiovascular disease in European and American elderly: the CHANCES project. *Am J Clin Nutr* (2015) **102**(4):745–56. doi:10.3945/ajcn.114.095117

88. Miao C, Bao M, Xing A, Chen S, Wu Y, Cai J, et al. Cardiovascular health score and the risk of cardiovascular diseases. *PLoS One* (2015) **10**(7):e0131537. doi:10.1371/journal.pone.0131537

89. Eckel N, Meidtner K, Kalle-Uhlmann T, Stefan N, Schulze MB. Metabolically healthy obesity and cardiovascular events: a systematic review and meta-analysis. *Eur J Prev Cardiol* (2015) 1–11. doi:10.1177/2047487315623884

90. Thiele I, Palsson B. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* (2010) **5**(1):93–121. doi:10.1038/nprot.2009.203

91. Rolfsson Ó, Palsson BO. Decoding the jargon of bottom-up metabolic systems biology. *Bioessays* (2015) **37**(6):588–91. doi:10.1002/bies.201400187

92. Bordbar A, Monk J, King Z, Palsson B. Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet* (2014) **15**(2):107–20. doi:10.1038/nrg3643

93. Becker S, Feist A, Mo M, Hannum G, Palsson B, Herrgard M. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox. *Nat Protoc* (2007) **2**(3):727–38. doi:10.1038/nprot.2007.99

94. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. *PLoS Comput Biol* (2013) **9**(3):e1002980. doi:10.1371/journal.pcbi.1002980

95. Dias O, Rocha M, Ferreira EC, Rocha I. Reconstructing genome-scale metabolic models with merlin. *Nucleic Acids Res* (2015) **43**(8):3899–910. doi:10.1093/nar/gkv294

96. Schultz A, Qutub AA. Reconstruction of tissue-specific metabolic networks using CORDA. *PLoS Comput Biol* (2016) **12**(3):e1004808. doi:10.1371/journal.pcbi.1004808

97. Pfau T, Pacheco MP, Sauter T. Towards improved genome-scale metabolic network reconstructions: unification, transcript specificity and beyond. *Brief Bioinform* (2015):bbv100. doi:10.1093/bib/bbv100

98. Machado D, Herrgård M. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput Biol* (2014) **10**(4):e1003580. doi:10.1371/journal.pcbi.1003580

99. Mardinoglu A, Gatto F, Nielsen J. Genome-scale modeling of human metabolism – a systems biology approach. *Biotechnol J* (2013) **8**(9):985–96. doi:10.1002/biot.201200275

100. Väremo L, Nookaew I, Nielsen J. Novel insights into obesity and diabetes through genome-scale metabolic modeling. *Front Physiol* (2013) **4**:92. doi:10.3389/fphys.2013.00092

101. Väremo L, Scheele C, Broholm C, Mardinoglu A, Kampf C, Asplund A, et al. Proteome- and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes. *Cell Rep* (2015) **11**(6):921–33. doi:10.1016/j.celrep.2015.04.010

102. Kumar A, Harrelson T, Lewis NE, Gallagher EJ, LeRoith D, Shiloach J, et al. Multi-tissue computational modeling analyzes pathophysiology of type 2 diabetes in MKR mice. *PLoS One* (2014) **9**(7):e102319. doi:10.1371/journal.pone.0102319

103. Mardinoglu A, Kampf C, Asplund A, Fagerberg L, Hallström BM, Edlund K, et al. Defining the human adipose tissue proteome to reveal metabolic alterations in obesity. *J Proteome Res* (2014) **13**(11):5106–19. doi:10.1021/pr500586e

104. Edwards LM, Sigurdsson MI, Robbins PA, Weale ME, Cavalleri GL, Montgomery HE, et al. Genome-scale methods converge on key mitochondrial genes for the survival of human cardiomyocytes in hypoxia. *Circ Cardiovasc Genet* (2014) **7**(4):407–15. doi:10.1161/CIRCGENETICS.113.000269

105. Yizhak K, Gabay O, Cohen H, Ruppin E. Model-based identification of drug targets that revert disrupted metabolism and its application to ageing. *Nat Commun* (2013) **4**:2632. doi:10.1038/ncomms3632

106. Stempler S, Yizhak K, Ruppin E. Integrating transcriptomics with metabolic modeling predicts biomarkers and drug targets for Alzheimer's disease. *PLoS One* (2014) **9**(8):e105383. doi:10.1371/journal.pone.0105383

107. Shlomi T, Cabili MN, Ruppin E. Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol Syst Biol* (2009) **5**(1):263. doi:10.1038/msb.2009.22

108. Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol Syst Biol* (2014) **10**(3):721. doi:10.1002/msb.145122

109. Patella F, Schug Z, Persi E, Neilson L, Erami Z, Avanzato D, et al. Proteomics-based metabolic modeling reveals that fatty acid oxidation (FAO) controls endothelial cell (EC) permeability. *Mol Cell Proteomics* (2015) **14**(3):621–34. doi:10.1074/mcp.M114.045575

110. Duarte N, Becker S, Jamshidi N, Thiele I, Mo M, Vo T, et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci U S A* (2007) **104**(6):1777–82. doi:10.1073/pnas.0610772104

111. Agren R, Bordel S, Mardinoglu A, Pornputtapong N, Nookaew I, Nielsen J. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput Biol* (2012) **8**(5):e1002518. doi:10.1371/journal.pcbi.1002518

112. Fouladiha H, Marashi S-A, Shokrgozar M. Reconstruction and validation of a constraint-based metabolic network model for bone marrow-derived mesenchymal stem cells. *Cell Prolif* (2015) **48**(4):475–85. doi:10.1111/cpr.12197

113. Motamedian E, Ghavami G, Sardari S. Investigation on metabolism of cisplatin resistant ovarian cancer using a genome scale metabolic model and microarray data. *Iran J Basic Med Sci* (2015) **18**(3):267–76.

114. Björnson E, Mukhopadhyay B, Asplund A, Pristovsek N, Cinar R, Romeo S, et al. Stratification of hepatocellular carcinoma patients based on acetate utilization. *Cell Rep* (2015) **13**(9):2014–26. doi:10.1016/j.celrep.2015.10.045

115. Bordbar A, McCloskey D, Zielinski DC, Sonnenschein N, Jamshidi N, Palsson BO. Personalized whole-cell kinetic models of metabolism for discovery in genomics and pharmacodynamics. *Cell Syst* (2015) **1**(4):283–92. doi:10.1016/j.cels.2015.10.003

116. Kirby P, Buerk DG, Parikh J, Barbee KA, Jaron D. Mathematical model for shear stress dependent NO and adenine nucleotide production from endothelial cells. *Nitric Oxide* (2015) **52**:1–15. doi:10.1016/j.niox.2015.10.004

117. Koo A, Nordsletten D, Umeton R, Yankama B, Ayyadurai S, García-Cardeña G, et al. In silico modeling of shear-stress-induced nitric oxide production in endothelial cells through systems biology. *Biophys J* (2013) **104**(10):2295–306. doi:10.1016/j.bpj.2013.03.052

118. Kang HG, Shim EB, Chang K-S. A new multiphysics model for the physiological responses of vascular endothelial cells to fluid shear stress. *J Physiol Sci* (2007) **57**(5):299–309. doi:10.2170/physiolsci.RP005407

119. Lim YC, Cooling MT, Long DS. Computational models of the primary cilium and endothelial mechanotransmission. *Biomech Model Mechanobiol* (2015) **14**(3):665–78. doi:10.1007/s10237-014-0629-x

120. Nordgaard H, Swillens A, Nordhaug D, Kirkeby-Garstad I, Van Loo D, Vitale N, et al. Impact of competitive flow on wall shear stress in coronary surgery: computational fluid dynamics of a LIMA-LAD model. *Cardiovasc Res* (2010) **88**(3):512–9. doi:10.1093/cvr/cvq210

121. McDougall SR, Anderson ARA, Chaplain MAJ. Mathematical modelling of dynamic adaptive tumour-induced angiogenesis: clinical implications and therapeutic targeting strategies. *J Theor Biol* (2006) **241**(3):564–89. doi:10.1016/j.jtbi.2005.12.022

122. Pontrelli G, Halliday I, Spencer TJ, König CS, Collins MW. Modelling the glycocalyx-endothelium-erythrocyte interaction in the microcirculation: a computational study. *Comput Methods Biomech Biomed Engin* (2015) **18**(4):351–61. doi:10.1080/10255842.2013.799146

123. Comerford A, David T, Plank M. Effects of arterial bifurcation geometry on nucleotide concentration at the endothelium. *Ann Biomed Eng* (2006) **34**(4):605–17. doi:10.1007/s10439-005-9046-8

124. Kar S, Kavdia M. Modeling of biopterin-dependent pathways of eNOS for nitric oxide and superoxide production. *Free Radic Biol Med* (2011) **51**(7):1411–27. doi:10.1016/j.freeradbiomed.2011.06.009

125. Kar S, Kavdia M. Endothelial NO and O_2^- production rates differentially regulate oxidative, nitroxidative, and nitrosative stress in the microcirculation. *Free Radic Biol Med* (2013) **63**:161–74. doi:10.1016/j.freeradbiomed.2013.04.024

126. Ghonaim NW, Fraser GM, Ellis CG, Yang J, Goldman D. Modeling steady state SO_2-dependent changes in capillary ATP concentration using novel O_2 micro-delivery methods. *Front Physiol* (2013) **4**:260. doi:10.3389/fphys.2013.00260

127. Munaron L, Scianna M. Multilevel complexity of calcium signaling: modeling angiogenesis. *World J Biol Chem* (2012) **3**(6):121–6. doi:10.4331/wjbc.v3.i6.121

128. Wei XN, Han BC, Zhang JX, Liu XH, Tan CY, Jiang YY, et al. An integrated mathematical model of thrombin-, histamine-and VEGF-mediated signalling in endothelial permeability. *BMC Syst Biol* (2011) **5**(1):112. doi:10.1186/1752-0509-5-112

129. Sakellarios AI, Papafaklis MI, Siogkas P, Athanasiou LS, Exarchos TP, Stefanou K, et al. Patient-specific computational modeling of subendothelial LDL accumulation in a stenosed right coronary artery: effect of hemodynamic and biological factors. *Am J Physiol Heart Circ Physiol* (2013) **304**(11):H1455–70. doi:10.1152/ajpheart.00539.2012

130. Nicolás M, Peña E, Malvè M, Martínez MA. Mathematical modeling of the fibrosis process in the implantation of inferior vena cava filters. *J Theor Biol* (2015) **387**:228–40. doi:10.1016/j.jtbi.2015.09.028

131. Kagadis GC, Skouras ED, Bourantas GC, Paraskeva CA, Katsanos K, Karnabatidis D, et al. Computational representation and hemodynamic characterization of in vivo acquired severe stenotic renal artery geometries using turbulence modeling. *Med Eng Phys* (2008) **30**(5):647–60. doi:10.1016/j.medengphy.2007.07.005

132. Cunnane EM, Mulvihill JJE, Barrett HE, Walsh MT. Simulation of human atherosclerotic femoral plaque tissue: the influence of plaque material model on numerical results. *Biomed Eng Online* (2015) **14**(Suppl 1):S7. doi:10.1186/1475-925X-14-S1-S7

133. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* (2007) **165**(9):1076–87. doi:10.1093/aje/kwk115

134. Holzinger ER, Dudek SM, Frase AT, Krauss RM, Medina MW, Ritchie MD. ATHENA: a tool for meta-dimensional analysis applied to genotypes and gene expression data to predict HDL cholesterol levels. *Pac Symp Biocomput* (2013):385–96. doi:10.1142/9789814447973_0038

135. Metts B, Thatcher S, Lewis E, Karounos M, Cassis L, Smith R, et al. DDDAS design of drug interventions for the treatment of dyslipidemia in ApoE(-/-) mice. *J Dev drugs* (2013) **2**(2):107. doi:10.4172/2329-6631.1000107

136. van Schalkwijk DB, van Ommen B, Freidig AP, van der Greef J, de Graaf AA. Diagnostic markers based on a computational model of lipoprotein metabolism. *J Clin Bioinforma* (2011) **1**(1):29. doi:10.1186/2043-9113-1-29

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Validation and Utilization of a Clinical Next-Generation Sequencing Panel for Selected Cardiovascular Disorders

Patrícia B. S. Celestino-Soper[1], Hongyu Gao[1], Ty C. Lynnes[1], Hai Lin[1,2], Yunlong Liu[1,2], Katherine G. Spoonamore[3], Peng-Sheng Chen[3] and Matteo Vatta[1,3]*

[1] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA, [2] Center for Computational Biology and Bioinformatics, Indiana University Purdue University Indianapolis, Indianapolis, IN, USA, [3] Division of Cardiology, Department of Medicine, Krannert Institute of Cardiology, Indiana University School of Medicine, Indianapolis, IN, USA

The development of high-throughput technologies such as next-generation sequencing (NGS) has allowed for thousands of DNA loci to be interrogated simultaneously in a fast and economical method for the detection of clinically deleterious variants. Whenever a clinical diagnosis is known, a targeted NGS approach involving the use of disease-specific gene panels can be employed. This approach is often valuable as it allows for a more specific and clinically relevant interpretation of results. Here, we describe the customization, validation, and utilization of a commercially available targeted enrichment platform for the scalability of clinical diagnostic cardiovascular genetic tests, including the design of the gene panels, the technical parameters for the quality assurance and quality control, the customization of the bioinformatics pipeline, and the post-bioinformatics analysis procedures. Regions of poor base coverage were detected and targeted by Sanger sequencing as needed. All panels were successfully validated using genotype-known DNA samples either commercially available or from research subjects previously tested in outside clinical laboratories. In our experience, utilizing several of the sub-panels in a clinical setting with 33 real-life cardiovascular patients, we found that 20% of tests requested were reported to have at least one pathogenic or likely pathogenic variant that could explain the patient phenotype. For each of these patients, the positive results may aid the clinical team and the patients in best developing a disease management plan and in identifying relatives at risk.

Keywords: next-generation sequencing, sequencing panels, cardiovascular, panel validation, clinical sequencing

INTRODUCTION

In the clinical genetics setting, most deleterious DNA variants can be detected by DNA sequencing. The development of high-throughput technologies such as next-generation sequencing (NGS) has allowed for thousands of target regions to be interrogated in a fast and economical approach, when compared to the more traditional technique of Sanger sequencing. Different NGS approaches such as whole-genome sequencing (WGS) and whole-exome sequencing (WES) have been employed especially for gene discovery. In particular, WGS and WES can be used as clinical testing modalities when a clinical diagnosis cannot be unequivocally established or for genetic disorders for which no other established clinical testing is available. However, when a clinical diagnosis has been

reached, a more targeted NGS approach involving the use of comprehensive disease-specific gene panels can be employed. In the clinical setting for example, gene panels may be designed to target genes associated with a disease or a group of related diseases depending on the level of complexity of clinical and phenotypic overlap. This approach is often valuable as it allows for a more specific and clinically relevant interpretation of results with variants in genomic loci *a priori* selected for their disease association. Additionally, when compared to WGS and WES, gene panels have the practical benefit of having more robust sequence coverage of target loci, lower cost, and faster turnaround time (1, 2).

Here, we describe the customization, clinical validation, and utilization of a commercial NGS panel, the TruSight One (TSO) panel developed by Illumina, Inc., which targets the coding regions of 4,813 genes associated with human disease, enriching for over 62,000 exons and their splice sites. Although NGS is currently considered to be a well-established technique, the clinical validation of recently available commercial kits still remains a constant challenge and a necessary step to ensure the high quality of clinical practice. Here, we show that the method of choice was technically reliable for sequencing and base calling, and that the annotation and filtering methods selected from the literature successfully detected variants in the targeted regions. Target regions were enriched and captured using the Illumina Nextera TSO Enrichment Kit and sequenced using solid-state sequencing-by-synthesis technology employing the Illumina MiSeq desktop sequencer system. The sequencing data were processed using an in-house custom bioinformatics pipeline with variant calls generated using the Burrows–Wheeler Aligner (BWA) followed by GATK analysis, which generated a variant call format (.vcf) file to be used for final interpretation. We subdivided the TSO panel into six smaller panels for testing of the exonic and splicing regions of genes associated with cardiovascular diseases according to disease phenotype, including arrhythmogenic right ventricular cardiomyopathy (ARVC), dilated cardiomyopathy/left ventricular non-compaction (DCM/LVNC), hypertrophic cardiomyopathy (HCM), Marfan syndrome/Loeys–Dietz syndrome (MFS/LDS), thoracic aortic aneurysms and dissections (TAAD), and a comprehensive cardiomyopathy (CMP) panel. Splitting into sub-panels allows for the proper test requisition by the physician while minimizing the risk of incidental findings and the presence of confounding variants. Several sub-panels were designed to have overlapping genes. In addition, the large CMP panel allows physicians to request the sequencing of genes in a more a comprehensive approach, usually when the clinical presentation is not very predictive of a particular type of CMP. The performance for each sub-panel was established after bioinformatics analyses which detected regions of poor coverage. These regions were targeted by Sanger sequencing as needed. Overall, all panels were successfully validated using a series of available genotype-known samples. We also describe our experience utilizing several of the sub-panels in a clinical setting with a group of 33 real-life cardiovascular patients (35 NGS tests requested). In conclusion, the utilization of the validated TSO sub-panels has provided us with a method to efficiently and economically search for thousands of clinically significant variants in one single experiment. Given the success of this project, we aim to continue the validation of additional sub-panels for other human disorders.

MATERIALS AND METHODS

NGS Cardiovascular Panels

The commercial TSO panel consisted of all 4,813 genes. We subdivided the gene content of the TSO panel into six clinical NGS panels which were further validated. The six clinical NGS panels made available for ordering of clinical testing each comprised of the sequencing of all coding regions and the immediate flanking regions of each exon of a specific group of genes. The CMP and the TAAD panels were also made available as reflex options (per physician request) when negative results were reported. **Table 1** describes each validated panel and the genes they cover. Table S1

TABLE 1 | Cardiovascular genetics next-generation sequencing (NGS) panel gene content.

NGS sub-panel	Number of genes	Genes covered
Comprehensive cardiomyopathy (CMP) and reflex CMP	61	*ABCC9, ACTC1, ACTN2, ANKRD1, BAG3, CASQ2, CAV3, CRYAB, CSRP3, CTNNA3, DES, DMD, DSC2, DSG2, DSP, DTNA, EMD, EYA4, FHL1, FHL2, FKTN, GATAD1, GLA, JPH2, JUP, KLF10, LAMA4, LAMP2, LDB3, LMNA, MYBPC3, MYH6, MYH7, MYL2, MYL3, MYLK2, MYOZ2, NEBL, NEXN, PDLIM3, PKP2, PLN, PRKAG2, PSEN1, PSEN2, RAF1, RBM20, RYR2, SCN5A, SGCD, TAZ, TCAP, TMEM43, TMPO, TNNC1, TNNI3, TNNT2, TPM1, TTN, TTR, VCL*
Hypertrophic cardiomyopathy	18	*ACTC1, ACTN2, CSRP3, GLA, LAMP2, MYBPC3, MYH7, MYL2, MYL3, MYOZ2, NEXN, PLN, PRKAG2, TNNC1, TNNI3, TNNT2, TPM1, TTR*
Dilated cardiomyopathy	33	*ABCC9, ACTC1, ACTN2, ANKRD1, BAG3, CRYAB, CSRP3, DES, DMD, DTNA, EMD, EYA4, GATAD1, LAMP2, LDB3, LMNA, MYBPC3, MYH7, NEXN, PLN, RAF1, RBM20, SCN5A, SGCD, TAZ, TCAP, TNNC1, TNNI3, TNNT2, TPM1, TTN, TTR, VCL*
Arrythmogenic right ventricular cardiomyopathy	8	*CTNNA3, DSC2, DSG2, DSP, JUP, LMNA, PKP2, TMEM43*
Marfan syndrome and Loeys–Dietz syndrome	3	*FBN1, TGFBR1, TGFBR2*
Thoracic aortic aneurysms and dissections (TAAD) and reflex TAAD	18	*ACTA2, CBS, COL3A1, COL5A1, COL5A2, ELN, FBN1, FBN2, MED12, MYH11, MYLK, PLOD1, SLC2A10, SMAD3, SMAD4, TGFB2, TGFBR1, TGFBR2*

in Supplementary Material contains the gene symbol, gene name, genomic coordinates, and main gene transcript for each gene that was sequenced using one of the NGS panels.

Validation Samples

The validation samples consisted of 23 genotype-known and 3 genotype-unknown samples (Table S2 in Supplementary Material). These samples were tested for their ability to result in successful libraries and sequence runs, as well as for evaluation and validation of the current NGS panel. DNA extraction methodology for each sample is listed in Table S2 in Supplementary Material. Validation samples were anonymous Coriell gDNA specimens or clinical samples that have been recruited through and Indiana University IRB-approved Genetic Registry Specimen Repository with minimal description in the report and no listed identifiable information.

Patient Samples

Patient samples consisted of clinical cases sent to the Indiana University School of Medicine Molecular Genetics Diagnostic Laboratory of the Department of Medical and Molecular Genetics, Indiana University School of Medicine, in the period of January 2015 to December 2015, with requisition for testing with one the following cardiovascular NGS panels: ARVC, DCM/LVNC, HCM, MFS/LDS, TAAD, or CMP. Patient samples listed in this manuscript have been described without identifiable information. All patient DNAs used for testing were extracted from whole blood collected in purple-top tube using the Qiagen's Gentra Puregene Blood Kit (Qiagen, Germantown, MD, USA) following the manufacturer's instructions. Samples were de-identified and only information about the NGS test requested, and variants identified were retained for the purpose of analysis of results for this manuscript.

Validation Runs

Table S3 in Supplementary Material summarizes the scheme used to prepare libraries using Illumina's TSO panel as specified by the standard and operating procedure (SOP). Each kit had reagents available for three runs. A repeat of NA12878 within a run was used for intra-run variability studies. A repeat of NA12878 between runs was used for inter-run variability studies. Additionally, different operators performed the tests according to the established technical protocols and SOP guidelines to allow for the evaluation of runs from libraries prepared by different operators (A or B). Furthermore, inter-lot, inter-day, and inter-run variability were assessed from runs of NA11931 between TSO_009 and TSO_013, and runs of NA11829 between TSO_010 and TSO_014. Single individuals (1-plex) or pools of three individuals (3-plex) were sequenced per flow-cell.

Library Preparation and Sequence Data Generation

The Illumina TSO NGS panel was developed by Illumina, Inc. (San Diego, CA, USA) and includes over 125,395 80-mer probes that were designed against the human NCBI GRCh37/hg19 reference genome assembly. Information about the

expected performance, targeted regions, content design, and other information can be found in the TSO full gene list, TSO Data Sheet, and TSO Technical Note in the manufacturer's website.

Optimization and validation experiments were set up manually following the manufacturer's instructions. Experiments were performed loading either a single sample (1-plex) or three samples (3-plex) per MiSeq run per flow-cell. After quantitation using Qubit (Thermo Fisher Scientific, Waltham, MA, USA), genomic DNA underwent Nextera tagmentation, which converts input genomic DNA into adapter-tagged libraries. Next, libraries were denatured and biotin-labeled probes specific to the targeted region were used for hybridization. The pool was then enriched for the targeted regions by adding streptavidin beads that bind to the biotinylated probes. Biotinylated DNA fragments bound to the streptavidin-coated beads were magnetically pulled down from the solution. The enriched DNA fragments were then eluted from the beads and hybridized for a second capture. Library preparation underwent quality control (QC) using an Agilent TapeStation, which was employed before library preparation and Qubit quantitation after library preparation. These steps provided the necessary metrics to assess the efficiency of fragmentation within the desired size range and the successful adapters/barcoding addition to each sample's DNA fragment. Prepared libraries were then loaded on to a flow-cell for sequencing with the Illumina MiSeq desktop sequencer system, which acquired sequencing data points and generated a bam and a fastq file for sequence reads. The resulting sequence data were submitted to analysis if the data passed the acceptance and rejection criteria for analytic runs according to manufacturer's instruction.

Bioinformatics Pipeline

To analyze and characterize data generated from targeted resequencing, the following softwares were implemented in our bioinformatics pipeline: Trim Galore (version 0.3.2) to remove adaptor sequences and low quality reads; BWA (version 0.7.5a) (3) to align reads to human reference genome UCSC GRCh37/hg19; GenomeAnalysisTK-2.8-1 (4, 5) for local realignment, base quality recalibration, and variants identification; SAMtools (version 0.1.19) (6) and picard-tools-1.105 (http://picard.sourceforge.net) to manipulate alignment files; VCFtools (version 0.1.10) (7) and BEDTools (8) (version 2.17.0) to further process resulted variant VCF files; ANNOVAR (revision 529) (9) for annotating variants; and Human Gene Mutation Database (HGMD) (10–12) professional used for further characterizing variants. R (13) and PERL were used for additional data analysis and characterization. All data processing steps were compiled into an all-steps-in-one bash script. Running scripts and parameters applied are available at http://compbio.iupui.edu/group/6/pages/clinicalsequencing. Additionally, our target file, the TruSight One Sequencing Panel Manifest downloaded from Illumina, can be found at: support.illumina.com/content/dam/illumina-support/documents/downloads/productfiles/trusight/trusight-one-manifest-may-2014.zip.

Following these procedures generated a final report of variants from targeted gene regions. The report consisted of

variant mapping information, gene annotation, amino acid change annotation (synonymous or non-synonymous variants), variant functional annotation [SIFT, PolyPhen, LRTs, and MutationTaster (14–20)], variant evolutionary conservation annotation [PhyloP and GERP++ (21–23)], variant presence and allele frequencies in currently publicly sequenced populations (dbSNP identifiers, 1000 Genomes Project allele, NHLBI-ESP 6500 exome project), and known disease-related functional annotation from the HGMD database. Quality parameters such as variant quality, read depth, mapping quality, and fisher strand bias were included in the final variants report as well (see Table S4 in Supplementary Material). Although the TSO panel includes 4,813 genes, in the clinical setting, only the genes included in the panel requested by the referring physician, genetic counselor, or other appropriate health care provider were analyzed and only variants for the requested panel were available for post-bioinformatics analyses of variants. Further information regarding the bioinformatics pipeline, can be found in the Supplementary Material.

Post-Bioinformatics Analyses

The TSO sequencing panel was first tested in 3-plex experiments (three individuals pooled per flow-cell run), as specified by the manufacturer. Validation of coverage and SNP performance was completed using 3-plex and 1-plex runs. Variants found in the validation samples were compared to secondary data as specified in Table S2 in Supplementary Material for concordance and evaluation of several metrics including false positive (FP) and false negative (FN) rates, analytic sensitivity, analytic specificity, overall genotype concordance (OGC), non-reference sensitivity (NRS), non-reference discrepancy (NRD), non-reference genotype concordance (NRGC), and precision.

Samples received for clinical testing were run as 1-plex or 3-plex NGS panel experiments (panel selection as requested for each patient). Sanger (BigDye) sequencing was used to provide data for bases with insufficient coverage in exonic and splicing (±2 nucleotides from the exon) regions of genes of interest in the NGS panel run (<15× or <10× sequence depth, as needed). Several regions were recurrently found to have lower than 15× sequence depth in 1-plex validation runs (Table S5 in Supplementary Material) and were included in the default Sanger sequencing for clinical testing for each selected panel. "Products were sequenced using an Applied Biosystems 3500 xl Genetic Analyzer in conjunction with the ABI BigDye® Terminator v3.1 Cycle Sequencing kit chemistry and protocol (ABI, Foster City, CA). Sequences where aligned to each gene and analyzed using Mutation Surveyor software V4.0.7 (SoftGenetics, State College, PA)." The limitations of the Sanger sequencing method are that the presence of DNA structural rearrangements (such as the deletion of an exon or multiple exons) may not be detected by sequence analysis. Additional tests analyzing DNA structural rearrangements should be recommended to those patients who are negative for sequencing analysis. Additionally, variants that may be found within known segmental duplication (SegDup) regions listed in Table S6 in Supplementary Material cannot be amplified and

sequenced unambiguously by PCR and BigDye sequence, and therefore cannot be reported. Variants found outside those loci listed in Table S6 in Supplementary Material were attempted to be confirmed unambiguously by PCR and BigDye sequencing if they were classified as pathogenic/likely pathogenic or variant of uncertain clinical significance (VUS).

Variants found in clinical test samples were evaluated for their clinical effect as being pathogenic, likely pathogenic, gene modifier, VUS, likely benign, or benign as explained in the Supplementary Material (special cases may differ from the classification procedures). The first step included separating the variants based on their presence or absence in the HGMD database, in order to facilitate the review process, since the HGMD database provides some curation for variants with known disease association. Variants deemed to be pathogenic/likely pathogenic or VUS were confirmed by Sanger sequencing as deemed necessary by the laboratory director on a case-by-case basis. In our post-bioinformatics analysis, we have mostly adhered to the current American College of Medical Genetics and Genomics (ACMG) guidelines for the standard interpretation of genetic variants (24). However, some parameters such as frequency have been adapted to reflect the current knowledge about the increased complex inheritance pattern in several cardiac syndromes, previously regarded as pure monogenic Mendelian diseases such as dilated cardiomyopathy (DCM), HCM, and ARVC, in which 5–10% of cases can present with two or more deleterious variants (25).

RESULTS

Panel Validation

Six NGS panels comprising of a select group of genes from the Illumina, Inc. TSO panel was optimized and validated using our in-house bioinformatics approach as described in Section "Materials and Methods." **Table 2** summarizes the metrics of the validation studies. Overall, the validation data for the NGS TSO and the cardiovascular sub-panels gave consistent and accurate genotype calls. A more detailed explanation of the validation results found in **Table 2** is presented below.

Coverage

The overall coverage (sequence depth) of target bases for the TSO panel (see **Figure 1**) was dependent on the concentration of final library used in the sequencing run (compare TSO_002_NA12878 15 versus 18pM runs) as well as on the final number of samples pooled per sequencing run (compare for example, NA11829 in 3-plex run TSO_010 and in 1-plex run TSO_014). The same was true for all cardiovascular sub-panels as summarized on **Table 2**. Overall, better coverage was obtained from higher concentration of library used in the run and in 1-plex experiments (as was found later for patient runs). Regions of systematic low coverage were often found to fall within the first exon of the targeted genes, likely due to high GC content in these regions, which may affect probe binding. Other factors, such as an inherent suboptimal performance of certain capture probes, may also play a role in the decreased coverage of some regions; however, since the TSO panel is a commercial

TABLE 2 | Summary of validation for next-generation sequencing TruSight One (TSO) and cardiovascular sub-panels.

1-Plex experiment (average of experiments ± SD)

Metrics[b]	TSO panel (full)	Comprehensive cardiomyopathy (CMP) and reflex	Hypertrophic cardiomyopathy (HCM)	Dilated cardiomyopathy/left ventricular non-compaction (DCM/LVNC)	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	Marfan syndrome/ Loeys–Dietz syndrome (MFS/ LDS)	Thoracic aortic aneurysms and dissections (TAAD) and reflex
Target region mean depth (x)	301.25 ± 35.56	309 ± 37.7	277 ± 30.38	322 ± 39.71	319.75 ± 43.22	240.25 ± 28.55	292.75 ± 32.84
Fraction of regions target depth ≥15×	0.97 ± 0.00	0.99 ± 0.00	1.00 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00
Accuracy	100%	100%	100%	100%	100%	100%	100%
Analytical sensitivity	0.96 ± 0.01	0.96 ± 0.02	1.00 ± 0.00	0.95 ± 0.03	0.98 ± 0.04	1.00 ± 0.00	0.96 ± 0.03
Analytical specificity	0.89 ± 0.03	0.93 ± 0.03	0.92 ± 0.04	0.93 ± 0.03	0.91 ± 0.05	1.00 ± 0.00	0.99 ± 0.02
False negative (FN) SNP rate	0.04 ± 0.01	0.04 ± 0.02	0.00 ± 0.00	0.05 ± 0.03	0.02 ± 0.04	0.00 ± 0.00	0.04 ± 0.03
False positive (FP) SNP rate	0.11 ± 0.03	0.07 ± 0.03	0.08 ± 0.04	0.07 ± 0.03	0.09 ± 0.05	0.00 ± 0.00	0.01 ± 0.02
Overall genotype concordance (OGC) range	0.81–0.94	0.82–0.94	0.83–1.00	0.81–0.97	0.81–1.00	N/A[a]	0.87–1.00
Non-reference sensitivity (NRS) range	0.91–0.98	0.87–1.00	0.94–1.00	0.86–1.00	0.81–1.00	N/A[a]	0.89–1.00
Non-reference discrepancy (NRD) range	0.07–0.20	0.06–0.19	0.00–0.17	0.04–0.20	0.00–0.19	N/A[a]	0.00–0.13
Non-reference genotype concordance (NRGC) range	0.90–0.97	0.87–1.00	0.94–1.00	0.86–1.00	0.81–1.00	N/A[a]	0.89–1.00
Precision range	0.85–0.97	0.86–0.98	0.85–1.00	0.87–1.00	0.81–1.00	N/A[a]	0.95–1.00
Number of loci with >15× coverage (length in basepairs)[c]	Not assessed	34 (3,329 bp)	3 (221 bp)	21 (1,561 bp)	6 (423 bp)	1 (100 bp)	6 (787 bp)
Number of SegDups loci (length in basepairs)[d]	Not assessed	5 (1,914 bp) 2[e] (197 bp)	0 (0 bp)	2 (208 bp) 0[e] (0 bp)	2 (197 bp)	0 (0 bp)	6 (1,917 bp) 1[e] (173 bp)

3-Plex experiment (average of experiments ± SD)

| Metrics[b]	TSO panel	CMP and CMP reflex	HCM	DCM/LVNC	ARVC	MFS/LDS	TAAD and TAAD reflex
Target region mean depth (x)	117.48 ± 11.92	118.03 ± 10.42	108.41 ± 12.01	122.66 ± 10.57	124.07 ± 11.58	94.48 ± 9.75	114.76 ± 11.74
Fraction of regions target depth ≥15×	0.95 ± 0.01	0.98 ± 0.00	0.98 ± 0.01	0.98 ± 0.00	0.97 ± 0.01	0.98 ± 0.01	0.98 ± 0.00
Accuracy (%)	100	100	100	100	100	100	100
Analytical sensitivity	0.96 ± 0.01	0.96 ± 0.03	0.99 ± 0.02	0.94 ± 0.03	0.98 ± 0.04	0.97 ± 0.09	0.97 ± 0.04
Analytical specificity	0.89 ± 0.03	0.93 ± 0.03	0.88 ± 0.03	0.94 ± 0.03	0.91 ± 0.06	0.98 ± 0.06	0.96 ± 0.04
FN SNP rate	0.04 ± 0.01	0.04 ± 0.03	0.01 ± 0.02	0.06 ± 0.03	0.02 ± 0.04	0.03 ± 0.09	0.03 ± 0.04
FP SNP rate	0.11 ± 0.03	0.07 ± 0.03	0.12 ± 0.08	0.06 ± 0.03	0.09 ± 0.06	0.02 ± 0.06	0.04 ± 0.04
OGC range	0.81–0.94	0.82–0.94	0.83–1.00	0.81–0.97	0.81–1.00	N/A[a]	0.87–1.00
NRS range	0.91–0.98	0.87–1.00	0.94–1.00	0.86–1.00	0.81–1.00	N/A[a]	0.89–1.00
NRD range	0.07–0.20	0.06–0.19	0.00–0.17	0.04–0.20	0.00–0.19	N/A[a]	0.00–0.13
NRGC range	0.90–0.97	0.87–1.00	0.94–1.00	0.86–1.00	0.81–1.00	N/A[a]	0.89–1.00
Precision range	0.85–0.97	0.86–0.98	0.85–1.00	0.87–1.00	0.81–1.00	N/A[a]	0.95–1.00

[a]Quantitative measurements OGC, NRS, NRD, NRGC, and precision were not computed for the MFS/LDS panel since only three variants were available for these analyses, and there was not enough power for the performance of the quantitative measurements above. Qualitative measurements of quality control for library preparation, sequencing performance, coverage, and SNP performance indicated that the variability among the runs listed above were within acceptable range.

[b]Samples MotherLP and ProbandJP were not used in the calculations of FN SNP rate, FP SNP rate, analytic sensitivity, analytic specificity, OGC, NRS, NRD, NRGC, and precision range. Only GenReg samples were used in accuracy calculations.

[c]Number of loci (with corresponding length in basepairs) that presented with lower than 15× coverage after 1-plex runs.

[d]Number of loci (with length in basepairs) that mapped to a region of known segmental duplications (SegDups).

[e]Number of loci (with corresponding number of basepairs in parenthesis) with known SegDup that were successfully validated to be unambiguously amplified and sequenced out of the total regions and corresponding basepairs listed.
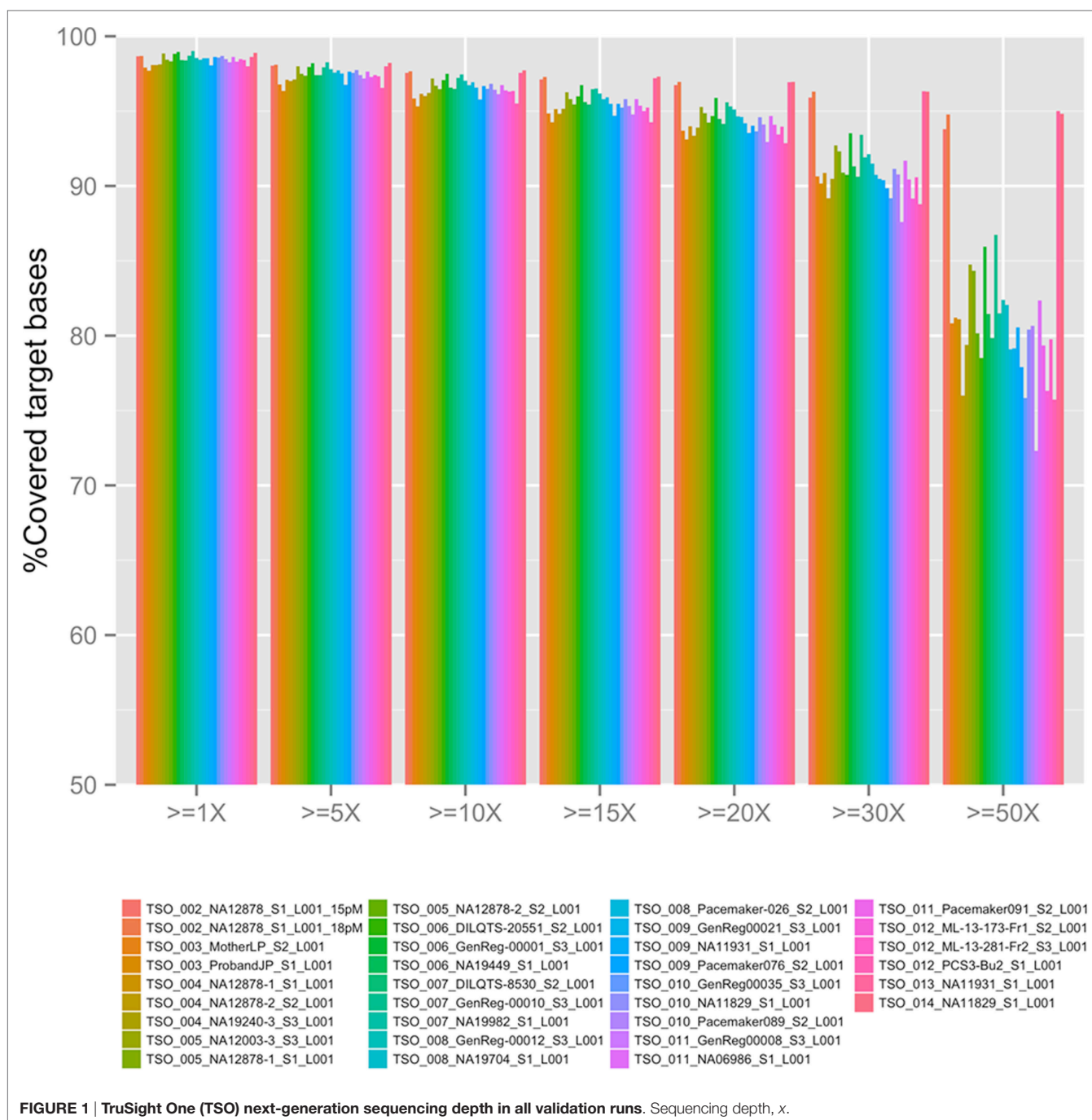
FIGURE 1 | TruSight One (TSO) next-generation sequencing depth in all validation runs. Sequencing depth, *x*.

off-the-shelf product, we were not given the choice to add custom optimized probes to mitigate this problem. Regions of less than 15× depth of coverage from 1-plex experiments were selected for BigDye (Sanger) sequencing validation for each panel (number and length of loci are listed in **Table 2**). Loci pertaining to regions of known SegDups for each panel were also attempted to be validated. It is possible that additional loci may need to be Sanger sequenced after NGS testing of a given patient (for example, for confirmation or for testing of additional regions with low coverage or within SegDups).

Accuracy

De-identified DNA samples with various genotypes previously tested at an independent clinical laboratory were assayed. The assay showed complete concordance with expected results for all panels, following a blinded analysis. These results show validation of the TSO panel (and its sub-panels), of the bioinformatics pipeline, and of the post-bioinformatics filtering of variants. In addition, using a 1-plex run with NA12878 (TSO_002_18pM run), we determined the maximum length of indels properly detected by the TSO panel to be of 22 nucleotides in a homozygous

state, and 30 nucleotides in a heterozygous state (both cases with satisfactory quality and sequence depth; see indel information in Materials and Methods in Supplementary Material).

Analytical Sensitivity, Analytical Specificity, FN Rates, and FP Rates

Table 2 lists the analytical sensitivity, analytical specificity, FN rates, and FP rates obtained in 1-plex and 3-plex NGS validation experiments when the data obtained (from a run, prior to BigDye confirmation) were compared to data from an outside source. Samples MotherLP and ProbandJP were not used in these calculations. Overall, very similar analytical sensitivity, analytical specificity, FN rates, and FP rates were obtained between 1-plex and 3-plex experiments.

BigDye confirmation was performed to test the FP and FN variants found in 1-plex experiments. Following BigDye confirmation, the results for 1-plex experiments shown in **Table 2** were corrected to reflect the final analytical sensitivity, analytical specificity, and FN and FP rates for the exonic and splicing targeted regions that obtained sequence depth of ≥15× of genes in the six NGS panels (**Table 3**). The values in **Table 3** represent the true expected reportable performance of the six NGS panels for 1-plex runs (since exonic and splicing regions of genes with <15× or <10× sequence depth were covered by BigDye sequencing, as deemed necessary). With an average of 1, 0.996, 0, and 0.004 for the sensitivity, specificity, FN rate, and FP rate, respectively, our panels demonstrated an excellent performance for the clinical application.

Assay Precision

The overall precision of each panel was calculated by running three different samples various times. Runs were compared to secondary data available (Illumina Platinum Genomes and 1000 G project) and also to a series of repeated runs in our laboratory. The repeatability was tested by the intra-run variability (two libraries of the same starting genomic DNA sample run twice in the same sequencing experiment), while the reproducibility was tested by the inter-run variability (two libraries of the same starting genomic DNA sample run twice in the separate sequencing experiments). Additionally, inter-operator variability was tested by allowing operator A to perform experiment TSO_004 while operator B performed experiment TSO_005. Furthermore, inter-lot, inter-day, and inter-run variability were assessed from runs of NA11931 between TSO_009 and TSO_013, and runs of NA11829

between TSO_010 and TSO_014. Several measurements were used to assess variability. The OGC, NRS, NRD, and NRGC were computed as previously published (26, 27). OGC, NRS, NRD, and NRGC were calculated treating each replicate alternatively as comparison set and evaluation set. Precision was calculated as True Positive SNPs divided by SNPs obtained from the Miseq run. **Table 2** describes the range of each measurement for both MiSeq runs compared to secondary data and to MiSeq repeated runs. Overall, values obtained from comparing our TSO NGS experiments to other methods used by secondary testing sites showed more variability than when comparing to our repeated runs (Table S7 in Supplementary Material).

Assay Robustness

DNAs obtained from different sources (whole blood, cell lines, buccal swabs, and frozen post-mortem blood) passed all QC steps from DNA extraction to library preparation, to required MiSeq sequencing metrics. Only one MiSeq instrument is available in our molecular genetics laboratory. Table S2 in Supplementary Material summarizes the various runs performed and **Table 2** summarizes the precision obtained under various conditions. It is evident that this assay is sufficiently robust to accommodate variations among consumables, technologists, and origin of DNA. However, assessment of the impact of contaminants on the performance of the test has not been systematically performed.
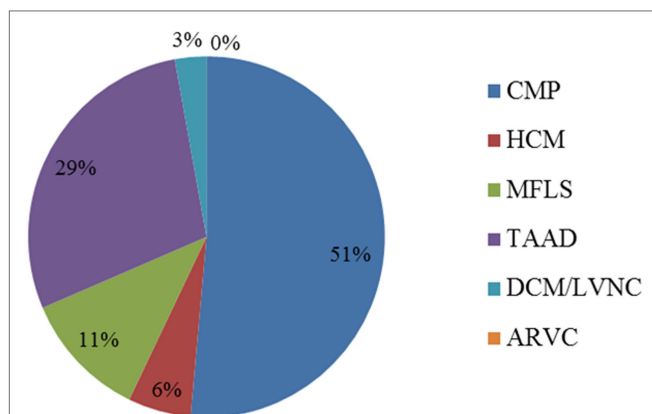
Patient Testing

Results from de-identified patient samples received by our clinical laboratory in the period of January 2015 to December 2015 with requisition for testing using one the cardiovascular NGS panels were collected. The patient reported results and variants found per panel requested are detailed in Table S8 in Supplementary Material. Overall, we were requested to perform NGS tests for 33 patients in the period selected for the writing of this manuscript, with two patients (IDs 24 and 32) having sequencing reported for two panels (CMP as a reflex). The distribution of the type of panel requested reflects the specific patient population of the requesting health professional, for which they deem to have the necessity to order a clinical genetic testing. Out of the 35 panels requested for testing, about half (18) were CMP panels. The second most ordered NGS test was the TAAD panel (10, or 29%). There were no requests for the ARVC panel during the period selected (**Figure 2**). Overall, 20% of all tests

TABLE 3 | Corrected SNP performance validation in 1-plex comprehensive comprehensive cardiomyopathy (CMP) panel validation after BigDye sequencing.

Metrics	TruSight One panel	CMP and reflex	Hypertrophic cardiomyopathy	Dilated cardiomyopathy/ left ventricular non-compaction	Arrhythmogenic right ventricular cardiomyopathy	Marfan syndrome/ Loeys–Dietz syndrome	Thoracic aortic aneurysms and dissections and reflex
Analytical sensitivity	N/A[a]	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	N/A[a]	1.00 ± 0.00
Analytical specificity	N/A[a]	0.99 ± 0.00	1.00 ± 0.00	0.99 ± 0.00	1.00 ± 0.00	N/A[a]	1.00 ± 0.00
False negative (FN) SNP rate	N/A[a]	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	N/A[a]	0.00 ± 0.00
False positive (FP) SNP rate	N/A[a]	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	N/A[a]	0.00 ± 0.00

[a]Measurements not calculated for the designated panels (no BigDye sequencing performed).

FIGURE 2 | Distribution of panels ordered for patient testing (%). The ARVC panel was not ordered for patient testing during the time-frame selected. Two patients (IDs 24 and 32) had sequencing reported for two panels (CMP as a reflex). ARVC, arrhythmogenic right ventricular cardiomyopathy panel; CMP, comprehensive cardiomyopathy panel; DCM/LVNC, dilated cardiomyopathy/left ventricular non-compaction panel; HCM, hypertrophic cardiomyopathy panel; MFLS, Marfan syndrome/Loeys–Dietz syndrome panel; TAAD, thoracic aortic aneurysms and dissections panel.

requested resulted in a positive result, meaning that at least one pathogenic or likely pathogenic variant was found in the patient tested. The CMP panel had the highest positive result rate with 28% of patients tested being reported to have at least one variant that could explain their phenotype. Additionally, about 43% of all NGS tests were reported to have no pathogenic or likely pathogenic variants, but to have at least one VUS, and about 37% of panels tested had a negative result (no pathogenic, likely pathogenic, or VUS was found). All HCM panels were reported as negative; however, only two HCM panels were requested in the period analyzed (**Figure 3**). **Figure 3** shows a graphic representation of our clinical sample population pick-up rates. Overall, the positive rate of approximately 28% obtained for the CMP panel (our largest sample size) is consistent with the previously published expected positive rate of 30% for patients diagnosed with DCM—we compare the published DCM rate to our CMP rate since in most cases, when the ordering physicians for our patients suspected a diagnosis of DCM they tended to order the larger, more comprehensive CMP panel (28). For other panels, the positive rates were heavily dependent on the diagnostic criteria and interpretation of clinical presentation used by the clinician prior to patient genetic testing. Additionally, in many cases, the panel requested by the physician may have been used as a differential diagnosis to exclude a specific disease. Therefore, the rate we obtained for the HCM, MFLS, TAAD, and DCM/LVNC panels may not be an accurate representation of expected pick-up rates for a given definitively diagnosed patient population.
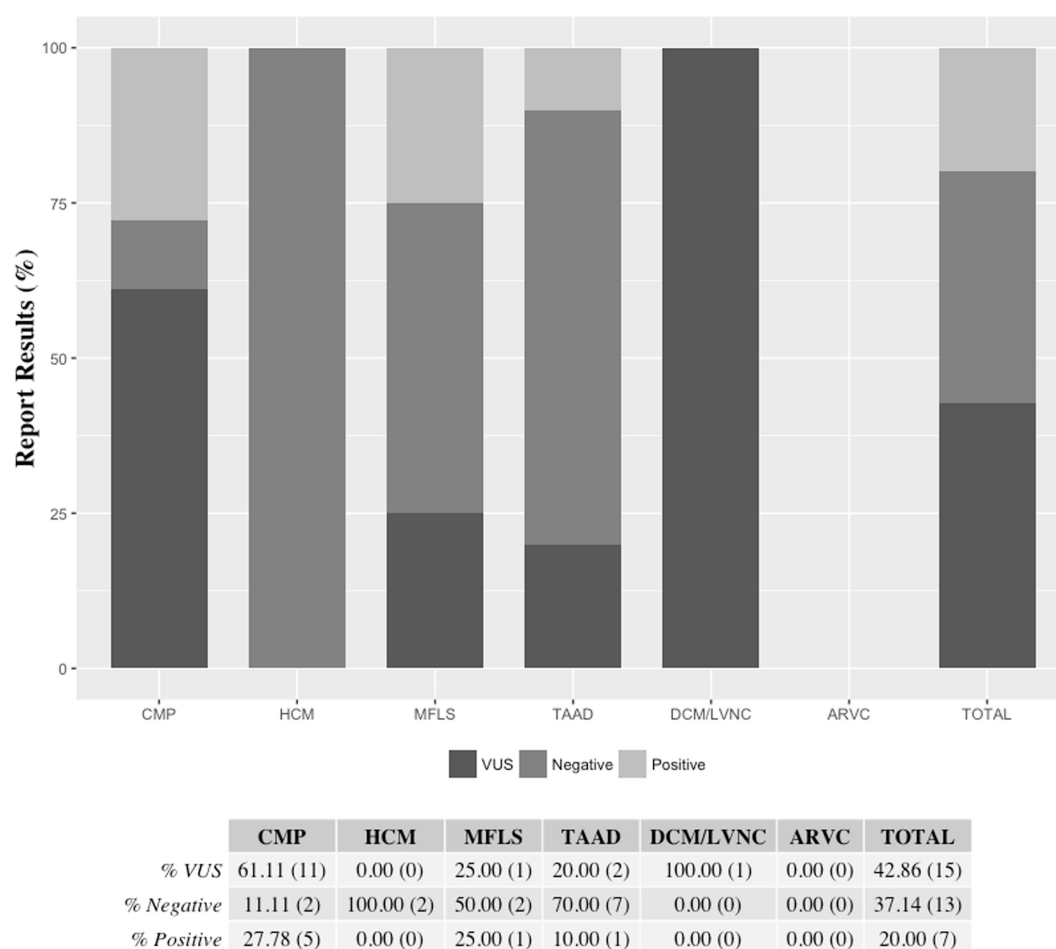
DISCUSSION

Sequencing information may be used as an aid to clinicians in determining disease diagnosis, follow-up procedures, genetic

counseling, therapeutic strategy, and treatment of disorders based on variants found in the gene(s) analyzed. Laboratory-developed NGS panel tests can identify an individual's genotype from genomic DNA with focus on specific disorders, groups of genes, phenotypes, and other variables in an efficient and cost-effective way. In this study, we present our results from the optimization and validation of the Illumina TSO NGS panel utilizing the Illumina MiSeq and an in-house bioinformatics pipeline for clinical testing of cardiovascular disorders (including HCM, DCM/LVNC, ARVC, MFS/LDS, TAAD, and comprehensive CMP). Our validation demonstrated that our procedures fulfilled the requirements of a clinical assay for detection of nucleotide base alterations, and small deletions and insertions with a desirable clinical test level of quality to detect constitutive genomic variants. Compared to the use of Sanger sequencing, at the current pricing and established turnaround time for clinical samples at our laboratory, one would save approximately 9.5 times the cost, and 10.2 times the time when using our NGS approach for an average gene, such as *LMNA*. Compared to other NGS targeting technologies, the hybridization capture-based approach that we used (as opposed to amplicon-based approach) allowed us to obtain a high quality NGS panel with clinically acceptable sensitivity, specificity, accuracy, precision, and coverage. Previous studies have shown that amplicon methods tend to be suboptimal and may generate higher FP and FN rates as well as lower coverage and uniformity (29). Finally, with regards to our optimized bioinformatics pipeline, we employed the most widely used tools to identify variants following the best practices of the GATK. Although, to our knowledge, there is no single state-of-the-art pipeline that is currently available for clinical NGS panel studies, our validation studies of our in-house developed bioinformatics pipeline have also shown clinically acceptable high quality results. A limitation of our study is that it is based on a small sample size, which may render it to be of insufficient power to address the genotypic variability of future samples and the true analytic sensitivity and specificity. Future studies are necessary to increase the power of our current assessment.

Our sub-panel approach included the selection of genes associated with cardiovascular diseases according to disease phenotype. Among the genes selected for each panel, several belong to a list of known pathogenic (KP) and/or expected pathogenic (EP) actionable variants, according to the ACMG recommendations on incidental findings: 18 genes with actionable KP/EP variants out of 61 CMP panel genes (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *TPM1*, *MYL3*, *ACTC1*, *PRKAG2*, *GLA*, *MYL2*, *LMNA*, *RYR2*, *PKP2*, *DSP*, *DSC2*, *TMEM43*, *DSG2*, and *SCN5A*), 10/18 HCM panel genes (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *TPM1*, *MYL3*, *ACTC1*, *PRKAG2*, *GLA*, and *MYL2*), 8/33 DCM/LVNC panel genes (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *TPM1*, *ACTC1*, *LMNA*, and *SCN5A*), 6/8 ARVC panel genes (*LMNA*, *PKP2*, *DSP*, *DSC2*, *TMEM43*, and *DSG2*), 3/3 MFS/LDS panel genes (*FBN1*, *TGFBR1*, and *TGFBR2*), and 8/18 TAAD panel genes (*COL3A1*, *FBN1*, *TGFBR1*, *TGFBR2*, *SMAD3*, *ACTA2*, *MYLK*, and *MYH11*) (30).

In our experience with 33 patients referred for clinical genetic testing using the given NGS panels, we found a positive result

	CMP	**HCM**	**MFLS**	**TAAD**	**DCM/LVNC**	**ARVC**	**TOTAL**
% VUS	61.11 (11)	0.00 (0)	25.00 (1)	20.00 (2)	100.00 (1)	0.00 (0)	42.86 (15)
% Negative	11.11 (2)	100.00 (2)	50.00 (2)	70.00 (7)	0.00 (0)	0.00 (0)	37.14 (13)
% Positive	27.78 (5)	0.00 (0)	25.00 (1)	10.00 (1)	0.00 (0)	0.00 (0)	20.00 (7)

FIGURE 3 | Reported results per panel requested. Final patient panel results were reported as either being positive (at least one pathogenic or likely pathogenic variant was found), VUS (at least one VUS but no pathogenic or likely pathogenic variant was found), or negative (no pathogenic, likely pathogenic, or VUS was found). The ARVC panel was not ordered for patient testing during the time-frame selected. Two patients (IDs 24 and 32) had sequencing reported for two panels (CMP as a reflex) Percentages are shown followed by the actual number of reports of each category in parenthesis. ARVC, arrhythmogenic right ventricular cardiomyopathy panel; CMP, comprehensive cardiomyopathy panel; DCM/LVNC, dilated cardiomyopathy/left ventricular non-compaction panel; HCM, hypertrophic cardiomyopathy panel; MFLS, Marfan syndrome/Loeys–Dietz syndrome panel; TAAD, thoracic aortic aneurysms and dissections panel; VUS, variant of uncertain clinical significance.

(pathogenic or likely pathogenic variant) for 20% of the panels tested. The highest positive rate resulted from CMP panels (28%), which was also the NGS panel that was the most requested in the period analyzed (51% of all panels requested). Patients with a positive test result may have a more appropriate management of their clinical phenotype and they were, as well as their relatives, recommended to receive continued clinical evaluation, follow-up, and genetic counseling. Our laboratory offers targeted testing for the specific variant(s) detected in the proband to at-risk relatives using Sanger sequencing technology, and many of the families took advantage of this service.

From the 33 patients referred for clinical genetic testing using the given NGS panels, about 43% of all tests were reported to have at least one VUS, but not a definitive pathogenic or likely pathogenic variant. The functional significance of these variants is not known at present and their contribution to the patient's disease

phenotype could not be determined at the time of reporting. However, these VUSs are good candidates for functional studies, and the analysis of other affected relatives of the patient tested may help support a potential pathogenic role of these variants if they co-segregate with the disease phenotype in the families studied.

From the 33 patients referred for clinical genetic testing using the given NGS panels, about 37% of all tests were reported to be negative. Many reasons may be related to a negative result. For example, a complicated clinical phenotype, or confounding factors such as environmental causes may result in a challenging choice for the most appropriate test to run in order to achieve the diagnosis of the proband. On the laboratory side, there are several technical limitations that could be associated with negative results. For example, the enrichment design employed in the commercial kit used for our NGS assays targets and detects variants in the

coding sequence and adjacent splicing and intronic sequences of the desired genes, while variants in deep intronic, non-coding, and regulatory regions that could affect gene expression were not targeted by our NGS assays. In addition, our clinical NGS tests were not designed for the purpose of detecting copy number variants due to large deletions and duplications encompassing all or a large portion of a gene (the maximum length of indels we detected was of 23 nucleotides in a homozygous state and 31 nucleotides in a heterozygous state). Moreover, our NGS methodology and depth of coverage were designed for constitutional genetics and may not detect low level mosaicism. Likewise, there could be some coding and splice site regions of genes that may present with an intrinsic sequence characteristics leading to suboptimal data. Finally, although our panels have been designed to include the great majority of genes known to be involved in each of the cardiovascular disorders listed here, every day, scientific progress reveals new genes that may be causing or be associated with these diseases. A benefit of our sub-panel design approach, in which a large panel was subdivided into smaller panels, is the fact that new genes and new sub-panels may be quickly validated from the list of 4,813 genes in the TSO panel as new literature points to new genes being involved in cardiovascular diseases. This validation would only consist of developing Sanger sequencing for regions of systematic low coverage and regions of known SegDups for the new genes and the calculation of parameters, such as FN, FP, accuracy, sensitivity, and specificity. For example, we are currently working on the validation and clinical implementation of NGS sub-panels for testing in Noonan spectrum disorders, long QT syndrome, hypertension, lipid disorders, and comprehensive arrhythmias. Additionally, our TSO panel and validation strategy may be used in the future for an array of non-cardiovascular diseases, including neurological, metabolic, skeletal disorders, and cancer, to name a few.

WEBSITES

1000 genomes browser (URL: http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/).

1000 G project release (genotypic information for NA12003, NA19449, NA19982, NA19704, NA11931, NA11829, and NA06986) (URL: ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/release/20110521/).

Bioinformatics pipeline scripts (http://compbio.iupui.edu/group/6/pages/clinicalsequencing).

Coriell Cell Repositories, Camden, NJ, USA (URL: http://ccr.coriell.org/).

dbSNP (URL: http://www.ncbi.nlm.nih.gov/SNP/).

GATK UnifiedGenotyper (URL: https://www.broadinstitute.org/gatk/gatkdocs/org_broadinstitute_gatk_tools_walkers_genotyper_UnifiedGenotyper.php#–heterozygosity).

Illumina (URL: http://www.illumina.com/).

NA12878 genotype information (URL: https://www.illumina.com/platinumgenomes.html).

NA19240 genotype information (URL: ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/pilot_data/release/2010_07/trio/snps/).

Picard-tools-1.105 (URL: http://picard.sourceforge.net).

TSO Data Sheet (URL: https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_one_panel.pdf).

TSO Full Gene List (URL: https://www.illumina.com/products/by-type/clinical-research-products/trusight-one.html).

TSO Technical Note (URL: https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/technote_trusight_one_panel.pdf).

UCSC browser (URL: http://genome.ucsc.edu/).

ETHICS STATEMENT

The study is exempt from the Helsinki declaration for studies involving human subjects because it either employed commercial human cell lines (from Coriell Institute) or de-identified and anonymized human subjects specimens.

AUTHOR CONTRIBUTIONS

PC-S, HG, YL, and MV were responsible for conception and design of the experiments and manuscript drafting. PC-S, HG, TL, and HL performed the experiments. PC-S, HG, TL, HL, YL, and MV were responsible for data generation, analysis, and interpretation. PC-S, HG, TL, HL, YL, PS-C, and MV were responsible for revising the content critically for intellectual content and for final approval of the manuscript.

ACKNOWLEDGMENTS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fcvm.2017.00011/full#supplementary-material.

REFERENCES

1. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* (2013) 14(4):295–300. doi:10.1038/nrg3463

2. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* (2013) 15(9):733–47. doi:10.1038/gim.2013.92

3. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* (2009) 25(14):1754–60. doi:10.1093/bioinformatics/btp324

4. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* (2010) 20(9):1297–303. doi:10.1101/gr.107524.110

5. Hoffman RM, Smith AY. What we have learned from randomized trials of prostate cancer screening. *Asian J Androl* (2011) 13(3):369–73. doi:10.1038/aja.2010.181

6. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics* (2009) 25(16):2078–9. doi:10.1093/bioinformatics/btp352

7. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics* (2011) 27(15):2156–8. doi:10.1093/bioinformatics/btr330

8. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010) 26(6):841–2. doi:10.1093/bioinformatics/btq033

9. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* (2010) 38(16):e164. doi:10.1093/nar/gkq603

10. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, et al. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat* (2003) 21(6):577–81. doi:10.1002/humu.10212

11. Cooper DN, Stenson PD, Chuzhanova NA. The Human Gene Mutation Database (HGMD) and its exploitation in the study of mutational mechanisms. *Curr Protoc Bioinformatics* (2006) Chapter 1:Unit1.13. doi:10.1002/0471250953.bi0113s12

12. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protoc Bioinformatics* (2012) Chapter 1:Unit1.13. doi:10.1002/0471250953.bi0113s39

13. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (2008).

14. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic acids research* (2003) 31(13):3812–4. doi:10.1093/nar/gkg509

15. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* (2009) 4(7):1073–81. doi:10.1038/nprot.2009.86

16. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* (2002) 30(17):3894–900. doi:10.1093/nar/gkf493

17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* (2010) 7(4):248–9. doi:10.1038/nmeth0410-248

18. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. *Genome Res* (2009) 19(9):1553–61. doi:10.1101/gr.092619.109

19. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* (2010) 7(8):575–6. doi:10.1038/nmeth0810-575

20. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* (2014) 11(4):361–2. doi:10.1038/nmeth.2890

21. Siepel A, Pollard KS, Haussler D. New methods for detecting lineage-specific selection. *Proceedings of the 10th International Conference on Research in Computational Molecular Biology (RECOMB '06)*. Venice, Italy: (2006). p. 190–205.

22. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res* (2010) 20(1):110–21. doi:10.1101/gr.097857.109

23. Cooper GM, Goode DL, Ng SB, Sidow A, Bamshad MJ, Shendure J, et al. Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nat Methods* (2010) 7(4):250–1. doi:10.1038/nmeth0410-250

24. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17(5):405–24. doi:10.1038/gim.2015.30

25. Vatta M, Spoonamore KG. Use of genetic testing to identify sudden cardiac death syndromes. *Trends Cardiovasc Med* (2015) 25(8):738–48. doi:10.1016/j.tcm.2015.03.007

26. Linderman MD, Brandt T, Edelmann L, Jabado O, Kasai Y, Kornreich R, et al. Analytical validation of whole exome and whole genome sequencing for clinical applications. *BMC Med Genomics* (2014) 7:20. doi:10.1186/1755-8794-7-20

27. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* (2011) 43(5):491–8. doi:10.1038/ng.806

28. Hershberger RE, Morales A, Siegfried JD. Clinical and genetic issues in dilated cardiomyopathy: a review for genetics professionals. *Genet Med* (2010) 12(11):655–67. doi:10.1097/GIM.0b013e3181f2481f

29. Samorodnitsky E, Jewell BM, Hagopian R, Miya J, Wing MR, Lyon E, et al. Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. *Hum Mutat* (2015) 36(9):903–14. doi:10.1002/humu.22825

30. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* (2013) 15(7):565–74. doi:10.1038/gim.2013.73

31. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* (2009) 42(2):377–81. doi:10.1016/j.jbi.2008.08.010

Multivariate Methods for Genetic Variants Selection and Risk Prediction in Cardiovascular Diseases

Alberto Malovini[1]*, Riccardo Bellazzi[1,2], Carlo Napolitano[3] and Guia Guffanti[4]

[1] Laboratory of Informatics and Systems Engineering for Clinical Research, IRCCS Fondazione Salvatore Maugeri, Pavia, Italy, [2] Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy, [3] Molecular Cardiology Laboratories, IRCCS Fondazione Salvatore Maugeri, Pavia, Italy, [4] Department of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA, USA

Over the last decade, high-throughput genotyping and sequencing technologies have contributed to major advancements in genetics research, as these technologies now facilitate affordable mapping of the entire genome for large sets of individuals. Given this, genome-wide association studies are proving to be powerful tools in identifying genetic variants that have the capacity to modify the probability of developing a disease or trait of interest. However, when the study's goal is to evaluate the effect of the presence of genetic variants mapping to specific chromosomes regions on a specific phenotype, the candidate loci approach is still preferred. Regardless of which approach is taken, such a large data set calls for the establishment and development of appropriate analytical methods in order to translate such knowledge into biological or clinical findings. Standard univariate tests often fail to identify informative genetic variants, especially when dealing with complex traits, which are more likely to result from a combination of rare and common variants and non-genetic determinants. These limitations can partially be overcome by multivariate methods, which allow for the identification of informative combinations of genetic variants and non-genetic features. Furthermore, such methods can help to generate additive genetic scores and risk stratification algorithms that, once extensively validated in independent cohorts, could serve as useful tools to assist clinicians in decision-making. This review aims to provide readers with an overview of the main multivariate methods for genetic data analysis that could be applied to the analysis of cardiovascular traits.

Keywords: SNPs, multivariate methods, risk scores, risk stratification, cardiovascular diseases

INTRODUCTION

The interaction of several genetic and environmental factors modulates the clinical expression of common cardiovascular diseases (CVDs), such as coronary artery disease (CAD), cerebrovascular disease, peripheral arterial disease, and stroke. Poor diet, physical inactivity, smoking, and harmful use of alcohol have all been established as key risk factors that can affect the clinical expression

of many CVDs (1). While predisposition to CVD as indicated by the presence of family history suggests that genetic factors play a role in the expression of the trait, the characteristics of inheritance often do not follow Mendelian patterns. For multifactorial diseases, this atypical pattern of inheritance impairs the elucidation of the genetic underpinnings. Indeed, multiple genetic factors with variable effects and effect size have to be identified to account for such a complex "polygenic" inheritance. On the other hand, the variable expressivity commonly found in monogenic cardiac diseases, even among subjects with the same genetic defect, represents a major limitation for the definition of genotype-based risk stratification algorithms (2).

Over the last decade, genome-wide association studies (GWASs) successfully identified more than 1,100 associations of genetic markers with cardiovascular traits, such as stroke, CAD, peripheral arterial disease, variability of the human electrocardiogram, and monogenic cardiac diseases (3). Although providing strong evidence of statistical association with these traits (p-value $<1 \times 10^{-8}$), single genetic variants identified by GWASs only explain a small proportion of the disease risk or phenotype variability (4–6). As an example, the recently identified CAD-associated variants reviewed in Ref. (4) induce each an average increase in terms of disease risk of ~18% [odds ratio (OR) = 1.18] (5). Further refining in genetic risk prediction and resuming multi-markers information in CVD will require alternative analytical strategies.

In the following sections, this review will address the main multivariate approaches to perform genetic variants selection from GWAS or candidate region studies, how the deriving findings could be modeled to define specific risk profiles and risk stratification algorithms and how to evaluate the prediction accuracy of the defined models.

IDENTIFICATION OF INFORMATIVE GENETIC VARIANTS

Identifying informative genetic markers among millions of candidates generated by microarrays or next generation sequencing (NGS) platforms has historically been a process of ranking variants according to their level of statistical association with a specific trait. This is first estimated by one-SNP-at-a-time testing approaches, and then a subset of these associated variants is selected based on a defined significance threshold (7). More recently, methods have emerged that are better suited for large cohorts of individuals deeply characterized by phenotypic measurements. Multivariate machine learning methods can be applied to identify informative subsets of genetic variants and non-genetic factors that jointly contribute to the overall phenotype expression (8). Annotating the identified markers could then be performed by accessing resources providing information on genomic variants previously associated with a trait of interest (3, 9, 10) and functional annotation tools (11–15). Once validated on independent cohorts of individuals, functional studies will allow researchers to translate evidence of statistical association and informative predictive models into biologically relevant findings (16).

Multivariate Methods for Common Genetic Variants Selection

Multivariate approaches of feature selection allow researchers to identify a subset or a combination of informative common genetic variants and non-genetic covariates that underlies the risk of developing a trait (17). These approaches offer a method that can overcome the limitations of the one-variant-at-a-time testing strategy characterizing univariate tests, which are incapable of capturing the multifactorial characteristics of many cardiovascular traits (e.g., additive effects of multiple variants, interactions between genetic and non-genetic factors) (18). In general, these approaches select informative variables not based on the strength of their statistical association with the trait, but rather on the basis of their capability to correctly predict the trait value in independent data.

A distinction has then to be made between multivariate methods for the analysis of binary traits (i.e., when the dependent variable indicates the presence or absence of a specific condition) and methods for quantitative traits analysis (i.e., when the dependent variable is characterized by a continuous distribution).

Binary Traits Analysis

The analysis of binary traits offers several alternatives that draw from both frequentist and Bayesian methods (**Table 1**). In order to identify informative sets of genetic and non-genetic variables expected to jointly affect a disease phenotype, stepwise logistic regression is one of the most consolidated approaches. The first step of this approach consists in testing simultaneously an initial set of SNPs in a logistic regression model as predictors of disease status which is represented by the binary-dependent variable. Then, different models are subsequently compared with the initial model to estimate whether a different set of predictors improved the fit, which is measured by goodness of fit metrics such as deviance or log-likelihood (19). Identifying the optimal model can be performed by a forward search strategy (the selection starts with the intercept of the regression, and then sequentially adds into the model the predictor that most improves the fit), a backward search strategy (it starts by including all variables, and sequentially deletes the predictor that has the lowest impact on the fit), or a combination of both (19). However, it is important to consider that this approach may prove computationally intensive when large sets of variables need to be analyzed, making the task of feature selection difficult.

The Least Absolute Shrinkage and Selection Operator (LASSO) (22) is a shrinkage method that represents a sound alternative to stepwise regression for the identification of informative genetic variants. The LASSO approach silences non-informative variables by setting their regression coefficient to 0 through a penalty parameter called lambda (λ). The optimal value to be assigned to λ can be learned by a resampling strategy performed on the data: the value guaranteeing the lowest average classification error on the test sets will be applied to the regression model. Vaarhorst and colleagues (34) used LASSO to identify predictors of coronary heart disease (CHD), starting from a set of candidate variants, whereas Hughes and colleagues (35) applied the algorithm to the identification of genetic variants to define a risk score for

TABLE 1 | Summary of the main multivariate methods for common variants analysis.

Phenotype	Method	Main software packages	Analysis of entire GWAS datasets	Advantages	Disadvantages
Binary traits					
	Stepwise logistic regression (19)	Orange (20), WEKA (21), stats[a], MASS[a]	Limited to candidate variants	Results can be easily interpreted	Results could be negatively influenced by collinearity; computationally intensive; R implementations[a] require advanced computer skills
	LASSO (22)	Orange (20), PLINK (23), HyperLASSO (24), glmnet[a], lars[a], penalized[a], ldlasso[a], scikit-learn[b]	Yes (HyperLASSO), otherwise the analysis is limited to candidate variants	Fast computation; internal CV to learn the optimal λ parameter	Does not necessarily yield good results in presence of high collinearity and when the number of variants exceeds the number of examples; R[a], Python[b], and PLINK implementations require advanced computer skills
	Elastic net (25)	elasticnet[a], glmnet[a], scikit-learn[b]	Limited to candidate variants	Combines strengths of LASSO and Ridge regression (26), overcoming issues due to collinearity, and unbalanced variants/samples ratio	Requires advanced computer skills
	BOSS (27)	BOSS	Limited to candidate variants	Works properly also when the number of features exceeds the number of samples	Computationally intensive; requires advanced computer skills
	BoNB (28)	BoNB	Yes	Fast computation; robust to LD between variants	Requires advanced computer skills
	Classification trees (29)	Orange (20), WEKA (21), rpart[a], tree[a], scikit-learn[b]	Limited to candidate variants	Fast computation; easy to interpret	May not perform well in the presence of complex interactions, overfitting may lead to instability; R[a] and Python[b] implementations require advanced computer skills
	Random forest (30)	Orange (20), WEKA (21), randomForest[a], randomForestSRC[a], scikit-learn[b], RFF (31)	Yes (RFF) otherwise the analysis is limited to candidate variants	Robust to noise; fast computation	Results are difficult to interpret; R[a], Python[b] and RFF implementations require advanced computer skills
	ABACUS (32)	ABACUS[a]	Candidate regions mapping to specific pathways	Able to simultaneously consider common and rare variants and different directions of genotype effect	Requires advanced computer skills
Time to event					
	Stepwise Cox proportional hazard model	Survival[a], MASS[a]	Limited to candidate variants	Results can be easily interpreted	Results could be negatively influenced by collinearity; computationally intensive; requires advanced computer skills
	LASSO (22)	glmnet[a], penalized[a] coxnet[a]	Limited to candidate variants	Fast computation; internal CV to learn the optimal λ parameter	Does not necessarily yield good results in presence of high collinearity and when the number of variants exceeds the number of examples; requires advanced computer skills
	Elastic net (25)	coxnet[a]	Limited to candidate variants	Combines strengths of LASSO and Ridge regression (26), overcoming issues due to collinearity, and unbalanced variants/samples ratio	Requires advanced computer skills
	Classification (survival) trees (29)	rpart[a]	Limited to candidate variants	Fast computation; easy to interpret	May not perform well in the presence of complex interactions, overfitting may lead to instability; requires advanced computer skills
	Random forest (30)	randomForestSRC[a]	Limited to candidate variants	Robust to noise; fast computation	Results are difficult to interpret; requires advanced computer skills
Quantitative traits					
	Stepwise linear regression	stats[a], MASS[a]	Limited to candidate variants	Results can be easily interpreted	Results could be negatively influenced by collinearity; computationally intensive; requires advanced computer skills
	LASSO (22)	Orange (20), PLINK (23), HyperLASSO (24), glmnet[a], lars[a], penalized[a], ldlasso[a], scikit-learn[b]	Yes (HyperLASSO), otherwise the analysis is limited to candidate variants	Fast computation; internal CV to learn the optimal λ parameter	Does not necessarily yield good results in presence of high collinearity and when the number of variants exceeds the number of examples; R[a], Python[b], and PLINK implementations require advanced computer skills

(Continued)

TABLE 1 | Continued

Phenotype	Method	Main software packages	Analysis of entire GWAS datasets	Advantages	Disadvantages
	Elastic net (25)	Elasticnet[a], glmnet[a], scikit-learn[b]	Limited to candidate variants	Combines strengths of LASSO and Ridge regression (26), overcoming issues due to collinearity, and unbalanced variants/samples ratio	Requires advanced computer skills
	GUESS (33)	GUESS/R2GUESS[a]	Yes	Fast parallel computation	Requires advanced computer skills
	Regression trees (29)	Orange (20), rpart[a], tree[a], scikit-learn[b]	Limited to candidate variants	Fast computation; easy to interpret	May not perform well in the presence of complex interactions, overfitting may lead to instability; R[a] and Python[b] implementations require advanced computer skills
	Random forest (30)	Orange (20), randomForest[a], randomForestSRC[a], scikit-learn[b], RFF (31)	Yes (RFF) otherwise the analysis is limited to candidate variants	Robust to noise; fast computation	Results are difficult to interpret; R[a], Python[b], and RFF implementations require advanced computer skills

Phenotype, dependent variable's distribution; method, algorithm or method; main software packages, main softwares, packages, or functions implementing the described method; analysis of entire GWAS datasets, indicates whether the method can be applied to whole GWAS data; advantages, advantages of the method; disadvantages, disadvantages of the method.

[a]*R package.*
[b]*Python package.*

coronary risk prediction. The elastic net (25) is an extension of the LASSO that is robust to extreme correlations among predictors, which also provides a more efficient, effective system for handling the analysis of unbalanced datasets.

Bayesian methods, such as the binary outcome stochastic search (BOSS) (27) and bags of naive Bayes (BoNB) (28) algorithms, also provide alternative approaches. BOSS is a feature selection approach deriving from the method described in Ref. (36) based on a latent variable model that links the observed outcome to the underlying genetic variants mapping to candidate regions of interest. A Markov Chain Monte Carlo approach is used for model search and to evaluate the posterior probability of each predictor in determining the latent variable profile (27). A latent variable profile is defined as a stochastic vector of same size of the number of SNPs; the vector may assume 0/1 values, thus expressing the fact that a marker is considered (value equal to 1) or not (value equal to 0) as a predictor of the outcome. The model estimates the posterior probability of such latent variable; as a consequence, the most likely latent variable will determine the set of SNPs with the highest risk prediction potential for developing a disease. BoNB (28) is an algorithm for genetic biomarkers selection from the simultaneous analysis of genome-wide SNP data based on the naive Bayes (NB) (37) classification framework. The predictive value (marginal utility) of each genetic variant is assessed by a resampling strategy. By randomly shuffling the genotypes of an informative variant, an overall decrease in terms of classification accuracy will be observed, and if an uninformative variant is permuted, no substantial loss will be observed. This strategy, coupled with appropriate statistical tests, allows BoNB to identify informative sets of SNPs. These methods have been tested on real datasets on type 1 (28, 38) and type 2 diabetes (27), respectively.

Classification and regression trees (RTs) methods (29) fall under the category of decision tree learning. In these tree structures, leaves represent the predicted phenotypic outcome, whereas nodes and branches represent the set of genetic variants and clinical covariates that predict the phenotypic outcome. These methods recursively partition data into subsets according to the variables' values: each partition corresponds with a "split" based on the set of variables being considered, defining a tree-like structure (19). Classification trees (CTs) are designed to analyze categorical traits and facilitate the identification of informative interactions between variables and stratifications in the data starting from a limited numbers of predictors.

Random forests (RFs) (30) are based on CTs, as they aggregate a large collection of de-correlated trees, and then average them (19). RFs generate a multivariate ranking of the analyzed variables according to their predictive importance with respect to the outcome. Even more, they can be easily applied to analyze unbalanced datasets, and they are able to account for correlation and informative interactions among features. Such characteristics make this approach particularly appealing for high-dimensional genomic data analysis (39). RFs have been applied to identify genetic variants influencing coronary artery calcification in hypertensive subjects (40), bicuspid aortic valve condition (41), and high-density lipoprotein (HDL) cholesterol level (42). Maenner and colleagues (43) applied RFs to identify SNPs involved in gene-by-smoking interactions related to the early-onset of CHD using the Framingham Heart Study data.

ABACUS is an Algorithm based on a BivAriate CUmulative Statistic, which allows identifying combinations of common and rare genetic variants associated with a disease by focusing on predefined SNPs-sets (e.g., belonging to specific pathways) (32). ABACUS calculates a statistic for each pair of SNPs within each SNPs set and generates an aggregated score measuring the cumulative evidence of association of the SNPs annotated in the SNP set. This method has been tested on GWAS on type 1 and type 2 diabetes (32).

Specific implementations of LASSO, elastic net, CTs, RFs, and stepwise Cox proportional hazard regression (44) have been also

proposed for the identification of SNPs associated with time to event outcomes (**Table 1**).

Quantitative Traits Analysis

Many of the feature selection methods for binary traits derive from algorithms originally established for quantitative traits analyses (**Table 1**). Linear regression (45) coupled with stepwise feature selection is probably one of the most commonly applied approaches when dealing with the task of identifying informative predictors with respect to continuous traits starting from a limited set of variables.

The LASSO and the elastic net shrinkage algorithms for regression problems work similarly for classification. Warren and colleagues (46) used LASSO and HyperLASSO (24) to predict low-density lipoprotein (LDL) and HDL cholesterol, two lipid traits of clinical relevance. Bottolo and colleagues (33) published the results from the validation and implementation of a method called Graphical Unit Evolutionary Stochastic Search (GUESS), a Bayesian variable selection approach able to analyze single and multiple responses, searching for the best combinations of SNPs to predict the traits. The authors applied the method to study genetic regulation of lipid metabolism in the Gutenberg Health Study (GHS), confirming the association of previously identified loci for blood lipid phenotypes.

Though largely similar to CTs, RTs differ from CTs in that the dependent variable is continuous, and a regression model is fitted to each node to perform the task of prediction. Additionally, RFs for regression problems are also widely employed and implemented in specific analytical packages.

MULTIVARIATE MODELS FOR DECISION SUPPORT

Demographic, clinical, and genetic risk factors identified by the previously described methods or selected based on prior knowledge can be combined in order to define specific predictive models, which could assist clinicians during the decision make steps of the clinical practice (47–49). Such models can be defined by making use of the above mentioned methods. For example, multilocus genetic risk profiles can be defined by weighting genetic variants by the corresponding regression coefficients (50, 51). Similarly, tree-based approaches or regression methods can be applied to define risk stratification algorithms combining genetic and non-genetic information (49, 51).

Multilocus Genetic Risk Profiles

The theory of multifactorial, polygenic liability relies on the combined effect of multiple common genetic variants, each explaining a small amount of phenotypic variance and possibly interacting with environmental factors, all contributing to the overall risk (52, 53). Polygenic risk score (PRS) approaches were introduced to examine the load of genetic risk associated with a given disease by simultaneously testing a broad set of common variants (54). Essentially, the PRS approach capitalizes on the identification of genetic risk variants derived from large,

mega-, or meta-analyses for specific disorders and generates an index of genetic vulnerability associated with the disease (54). Affected subjects present higher values of the PRS than not affected subjects. The advantage of polygenic modeling is that the genetic vulnerability is represented by a larger set of gene-mapping variants contributing to the risk of the disease, rather than a single genetic variant. There are several different ways to implement polygenic modeling approaches (55). All methods rely on selecting variants on a training set using univariate or multivariate approaches or focusing on candidate loci identified by previous studies. The risk alleles of the identified sets of genetic variants are then used to generate a PRS either by summing the number of risk alleles ("un-weighted" approach) or by weighting the number of risk alleles by the effect size of the association deriving from regression models ("weighted" approach) (50). Either way, the PRS is tested for association in a replication sample *via* traditional regression-based statistics and standard metrics are used to estimate its predictive power (56).

Polygenic risk score usually explain 1–5% of the variation in complex traits, which is already an improvement compared with GWAS single genetic variants, which typically yield relatively small increment of risk with ORs <1.5-fold, with the exception of traits such as height, for which a GWAS identified a SNP explaining almost 5% of the phenotypic variance (53, 57). PRS have been applied to several CVD studies and are found to be a significant predictor of CAD (58, 59), incident cardiovascular (60), CHD (61), atrial fibrillation, and stroke (62). Furthermore, Pfeufer and colleagues (63) assessed the cumulative effect of SNPs modulating the QT interval in the general population. For a more comprehensive review of PRS findings in CVD, we encourage readers to consider the report by Abraham and Inouye (51).

Risk Stratification Algorithms

Risk stratification algorithms are designed to be intuitive tools that can assist clinicians in identifying patients at high risk of adverse events, thus informing decision-making by following a defined set of logical steps (64–66). These algorithms can be derived by the integration of genetic information (e.g., single SNPs, mutations on causative loci, PRSs) with known clinical and behavioral risk factors by appropriate multivariate methods. When defined by regression methods, they can be interrogated by nomograms, graphical tools that allow interpreting the risk of developing a certain trait based on an individual's characteristics (67).

Priori et al. (47) proposed a risk stratification algorithm to identify long QT syndrome (LQTS) patients at high risk of adverse cardiac events (defined as occurrence of syncope, cardiac arrest, or sudden death before the age of 40 years and in absence of therapies). LQTS is a genetic disorder caused by mutations that affect ion-channel encoding genes or other genes that indirectly modulate the function of ion channels. The algorithm was based on the combination of information about the presence of genetic variants on one of the three main LQTS genes (*KCNQ1*, *KCNH2*, and *SCN5A* defining LQT1, LQT2, and LQT3), gender, and QT interval duration (≥500 or <500 ms), which are known

independent risk predictors in LQTS. Three risk groups were identified based on the observed probability of an adverse cardiac event: low risk (probability <30%), intermediate risk (30–49%), and high risk (≥50%). Based on the published risk stratification algorithms for LQT1, LQT2, and LQT3 patients (47, 68), Tomás and colleagues (48) investigated whether common variants on *NOS1AP* locus can add additional insights for risk stratification in this group of patients. The authors demonstrated that the presence of the *NOS1AP* rs10494366 variant improved event risk stratification for previously identified LQT1, LQT2, and LQT3 patients. The presence of the GG or GT genotype of *NOS1AP* rs10494366 increased the risk of cardiac events compared with homozygotes for the T allele in all the subgroups of LQTS patients defined by different combinations of gender and genetic locus (**Figure 1**).

Talmud et al. (69) evaluated whether the inclusion of information regarding the genotype of rs10757274 on 9p21.3 locus to the risk factors defining the Framingham risk score (FRS) allowed increasing the accuracy in identifying patients at risk of CHD in a prospective study. Results showed that, although rs10757274 did not add substantially to the usefulness of the FRS for predicting future events, it did improve reclassification of CHD risk, and thus may have clinical utility.

Ripatti et al. (58) tested 13 SNPs – associated with myocardial infarction or CAD by previous GWASs – in a case–control design including 3,829 CHD cases and 48,897 control participants and a prospective cohort design including 30,725 individuals free of CVD. In prospective cohort analyses, the weighted PRS defined using the set of selected SNPs was significantly associated with a first CHD event. Furthermore, when compared with the bottom quintile of the PRS distribution, individuals in the top quintile shared a 1.66-fold increased covariates-adjusted risk of CHD. When focusing on its risk prediction capability, the PRS did not improve the C index over clinical risk factors but increased

QTc ≥500 ms rs366 minor	4.08	6.24	7.18	10.97
QTc ≥500 ms rs366 common	2.78	4.25	4.89	7.47
QTc <500 ms rs366 minor	1.47	2.24	2.58	3.95
QTc <500 ms rs366 common	1	1.53	1.76	2.69
	LQT1 Males	LQT1 Females	LQT2/3 Males	LQT2/3 Females

☐ HR <2   ☐ HR < 2 < 4   ☐ HR 4-5   ☐ HR > 5

FIGURE 1 | rs10494366 common variant on *NOS1AP* modulates risk of events in LQTS (48). The schema reports the combined hazard ratios (HRs) from Cox regression by risk categories. The risk stratification schema includes the common variant rs10494366 on *NOS1AP* gene and known risk predictors in LQTS, represented by: QTc ≥ 500 ms, gender, and LQTS subgroup. Each box shows the combined HR for patients sharing clinical and genetic characteristics. The reference category (HR = 1) is represented by individuals LQT1, males, QT < 500 ms and homozygote for the common allele of *NOS1AP* rs10494366. Reprinted from the manuscript by Tomás and colleagues (48) with permission from Elsevier.

slightly the integrated discrimination index (*p*-value <0.001). Similar results were obtained from the case–control analyses.

MODEL ASSESSMENT STRATEGIES

Once multivariate sets of SNPs, PRSs or risk stratification algorithms are defined on an initial cohort (training set), their accuracy in predicting the condition of new examples must be assessed on independent populations (test set). In the absence of independent cohorts, it is possible to rely on resampling strategies like K-Fold Cross Validation (K-Fold CV) (19), holdout (70), and bootstrap (71). Several metrics are available to evaluate and compare the discriminative power of predictive models on the test set, based on the trait's distribution (72, 73).

CONCLUSION

The goal of this review is to provide readers with an overview about the main multivariate methods that can be applied to the identification of informative genetic variants and to the definition of risk prediction tools in the context of CVDs. It is important to note that some methods described have been applied to intermediate phenotypes that could be considered precursors to their manifestation as cardiovascular traits, but these methods have not yet been applied to the analysis of cardiovascular traits. Their application to large CVD cohorts could lead to interesting findings.

Multivariate methods allow the identification of complex additive effects due to the presence of multiple genetic variants on specific loci or complex interactions among genetic and non-genetic risk factors able to modulate the probability of developing a specific disease or its severity.

Still, the task of identifying informative combinations of genetic variants by multivariate search strategies can be extremely computationally intensive due to the high number of models to be explored and, in many cases, to the impossibility of parallelizing the analyses. Missing values represent a common limitation to these approaches, although it could be partially solved by resorting to multivariate imputation methods. Furthermore, large sets of samples thoroughly characterized in terms of phenotype characteristics are needed in order to avoid overfitting issues and to increase the probability of defining models whose predictive performances can be confirmed in independent cohorts.

AUTHOR CONTRIBUTIONS

AM and GG conceived the study and drafted the manuscript. RB and CN conceived the study and revised the manuscript critically for important intellectual content. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS
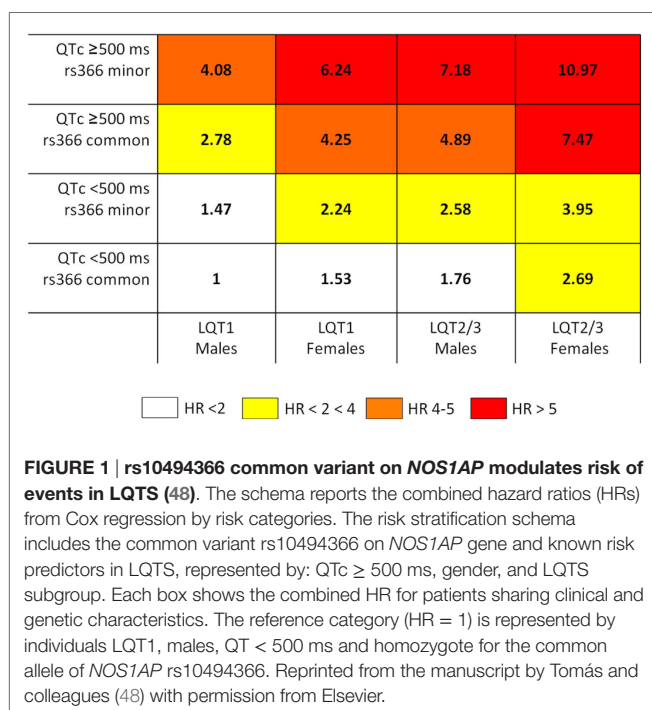
REFERENCES

1. WHO. *Global Status Report on Noncommunicable Diseases 2014*. Geneva: WHO (2014).

2. Priori SG. Inherited arrhythmogenic diseases: the complexity beyond monogenic disorders. *Circ Res* (2004) 94:140–5. doi:10.1161/01.RES.0000115750.12807.7E

3. Burdett T, Hall PN, Hasting E, Hindorff LA, Junkins HA, Klemm AK, et al. *The NHGRI-EBI Catalog of Published Genome-Wide Association Studies [Online]*. (2016). Available from: www.ebi.ac.uk/gwas

4. Roberts R. Genetics of coronary artery disease. *Circ Res* (2014) 114:1890–903. doi:10.1161/CIRCRESAHA.114.302692

5. Bjorkegren JL, Kovacic JC, Dudley JT, Schadt EE. Genome-wide significant loci: how important are they? Systems genetics to understand heritability of coronary artery disease and other common complex disorders. *J Am Coll Cardiol* (2015) 65:830–45. doi:10.1016/j.jacc.2014.12.033

6. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* (2015) 47:1121–30. doi:10.1038/ng.3396

7. Panagiotou OA, Ioannidis JP; Genome-Wide Significance Project. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int J Epidemiol* (2012) 41:273–86. doi:10.1093/ije/dyr178

8. Okser S, Pahikkala T, Airola A, Salakoski T, Ripatti S, Aittokallio T. Regularized machine learning in the genetic prediction of complex traits. *PLoS Genet* (2014) 10:e1004754. doi:10.1371/journal.pgen.1004754

9. Stenson PD, Mort M, Ball EV, Howells K, Phillips AD, Thomas NS, et al. The human gene mutation database: 2008 update. *Genome Med* (2009) 1:13. doi:10.1186/gm13

10. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* (2014) 42:D980–5. doi:10.1093/nar/gkt1113

11. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* (2009) 4:1073–81. doi:10.1038/nprot.2009.86

12. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* (2010) 7:248–9. doi:10.1038/nmeth0410-248

13. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* (2011) 39:e118. doi:10.1093/nar/gkr407

14. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* (2012) 7:e46688. doi:10.1371/journal.pone.0046688

15. Limongelli I, Marini S, Bellazzi R. PaPI: pseudo amino acid composition to score human protein-coding variants. *BMC Bioinformatics* (2015) 16:123. doi:10.1186/s12859-015-0554-8

16. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet* (2013) 93:779–97. doi:10.1016/j.ajhg.2013.10.012

17. Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet* (2015) 16:85–97. doi:10.1038/nrg3868

18. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* (2007) 23:2507–17. doi:10.1093/bioinformatics/btm344

19. Hastie T, Tibshirani R, Friedman R. *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. New York: Springer (2009).

20. Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, et al. Orange: data mining toolbox in Python. *J Mach Learn Res* (2013) 14:2349–53.

21. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. *SIGKDD Explor* (2009) 11:10–8. doi:10.1145/1656274.1656278

22. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* (1996) 58:267–8.

23. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* (2015) 4:7. doi:10.1186/s13742-015-0047-8

24. Hoggart CJ, Whittaker JC, De Iorio M, Balding DJ. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS Genet* (2008) 4:e1000130. doi:10.1371/journal.pgen.1000130

25. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Series B Stat Methodol* (2005) 67:301–20. doi:10.1111/j.1467-9868.2005.00503.x

26. Jain RK. Ridge regression and its application to medical data. *Comput Biomed Res* (1985) 18:363–8. doi:10.1016/0010-4809(85)90014-X

27. Russu A, Malovini A, Puca AA, Bellazzi R. Stochastic model search with binary outcomes for genome-wide association studies. *J Am Med Inform Assoc* (2012) 19:e13–20. doi:10.1136/amiajnl-2011-000741

28. Sambo F, Trifoglio E, Di Camillo B, Toffolo GM, Cobelli C. Bag of naive Bayes: biomarker selection and classification from genome-wide SNP data. *BMC Bioinformatics* (2012) 13(Suppl 14):S2. doi:10.1186/1471-2105-13-S14-S2

29. Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees*. Belmont, CA: Wadsworth Publishing Company (1984).

30. Breiman L. Random forests. *Mach Learn* (2001) 45:5–32. doi:10.1023/A:1010933404324

31. Yang W, Charles Gu C. Random forest fishing: a novel approach to identifying organic group of risk factors in genome-wide association studies. *Eur J Hum Genet* (2014) 22:254–9. doi:10.1038/ejhg.2013.109

32. Di Camillo B, Sambo F, Toffolo G, Cobelli C. ABACUS: an entropy-based cumulative bivariate statistic robust to rare variants and different direction of genotype effect. *Bioinformatics* (2014) 30:384–91. doi:10.1093/bioinformatics/btt697

33. Bottolo L, Chadeau-Hyam M, Hastie DI, Zeller T, Liquet B, Newcombe P, et al. GUESS-ing polygenic associations with multiple phenotypes using a GPU-based evolutionary stochastic search algorithm. *PLoS Genet* (2013) 9:e1003657. doi:10.1371/journal.pgen.1003657

34. Vaarhorst AA, Lu Y, Heijmans BT, Dolle ME, Bohringer S, Putter H, et al. Literature-based genetic risk scores for coronary heart disease: the cardiovascular registry Maastricht (CAREMA) prospective cohort study. *Circ Cardiovasc Genet* (2012) 5:202–9. doi:10.1161/CIRCGENETICS.111.960708

35. Hughes MF, Saarela O, Stritzke J, Kee F, Silander K, Klopp N, et al. Genetic markers enhance coronary risk prediction in men: the MORGAM prospective cohorts. *PLoS One* (2012) 7:e40922. doi:10.1371/journal.pone.0040922

36. Bottolo L, Richardson S. Evolutionary stochastic search for Bayesian model exploration. *Bayesian Anal* (2010) 5:35. doi:10.1214/10-BA523

37. Russell S, Norvig P. *Artificial Intelligence: A Modern Approach*. 2nd ed. Upper Saddle River, NJ: Prentice Hall (2003).

38. Sambo F, Malovini A, Sandholm N, Stavarachi M, Forsblom C, Makinen VP, et al. Novel genetic susceptibility loci for diabetic end-stage renal disease identified through robust naive Bayes classification. *Diabetologia* (2014) 57:1611–22. doi:10.1007/s00125-014-3256-2

39. Chen X, Ishwaran H. Random forests for genomic data analysis. *Genomics* (2012) 99:323–9. doi:10.1016/j.ygeno.2012.04.003

40. Sun YV, Bielak LF, Peyser PA, Turner ST, Sheedy PF II, Boerwinkle E, et al. Application of machine learning algorithms to predict coronary artery calcification with a sibship-based design. *Genet Epidemiol* (2008) 32:350–60. doi:10.1002/gepi.20309

41. Wooten EC, Iyer LK, Montefusco MC, Hedgepeth AK, Payne DD, Kapur NK, et al. Application of gene network analysis techniques identifies AXIN1/PDIA2 and endoglin haplotypes associated with bicuspid aortic valve. *PLoS One* (2010) 5:e8830. doi:10.1371/journal.pone.0008830

42. Heidema AG, Feskens EJ, Doevendans PA, Ruven HJ, van Houwelingen HC, Mariman EC, et al. Analysis of multiple SNPs in genetic association studies: comparison of three multi-locus methods to prioritize and select SNPs. *Genet Epidemiol* (2007) 31:910–21. doi:10.1002/gepi.20251

43. Maenner MJ, Denlinger LC, Langton A, Meyers KJ, Engelman CD, Skinner HG. Detecting gene-by-smoking interactions in a genome-wide association study of early-onset coronary heart disease using random forests. *BMC Proc* (2009) 3(Suppl 7):S88. doi:10.1186/1753-6561-3-s7-s88

44. Cox DR. Regression models and life-tables. *J R Stat Soc Series B Stat Methodol* (1972) 34:187–220.

45. Hocking RR. The analysis and selection of variables in linear regression. *Biometrics* (1976) 32:1–49. doi:10.2307/2529336

46. Warren H, Casas JP, Hingorani A, Dudbridge F, Whittaker J. Genetic prediction of quantitative lipid traits: comparing shrinkage models to gene scores. *Genet Epidemiol* (2014) 38:72–83. doi:10.1002/gepi.21777

47. Priori SG, Schwartz PJ, Napolitano C, Bloise R, Ronchetti E, Grillo M, et al. Risk stratification in the long-QT syndrome. *N Engl J Med* (2003) 348:1866–74. doi:10.1056/NEJMoa022147

48. Tomás M, Napolitano C, De Giuli L, Bloise R, Subirana I, Malovini A, et al. Polymorphisms in the NOS1AP gene modulate QT interval duration and risk of arrhythmias in the long QT syndrome. *J Am Coll Cardiol* (2010) 55:2745–52. doi:10.1016/j.jacc.2009.12.065

49. Wasan PS, Uttamchandani M, Moochhala S, Yap VB, Yap PH. Application of statistics and machine learning for risk stratification of heritable cardiac arrhythmias. *Expert Syst Appl* (2013) 40(7):2476–86. doi:10.1016/j.eswa.2012.10.054

50. Sebastiani P, Solovieff N, Sun JX. Naive Bayesian classifier and genetic risk score for genetic risk prediction of a categorical trait: not so different after all! *Front Genet* (2012) 3:26. doi:10.3389/fgene.2012.00026

51. Abraham G, Inouye M. Genomic risk prediction of complex human disease and its clinical application. *Curr Opin Genet Dev* (2015) 33:10–6. doi:10.1016/j.gde.2015.06.005

52. Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. *Nat Rev Genet* (2004) 5:89–100. doi:10.1038/nrg1270

53. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* (2009) 461:747–53. doi:10.1038/nature08494

54. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* (2009) 460:748–52. doi:10.1038/nature08185

55. Simonson MA, Wills AG, Keller MC, McQueen MB. Recent methods for polygenic analysis of genome-wide data implicate an important effect of common variants on cardiovascular disease risk. *BMC Med Genet* (2011) 12:146. doi:10.1186/1471-2350-12-146

56. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* (2013) 9:e1003348. doi:10.1371/journal.pgen.1003348

57. Visscher PM. Sizing up human height variation. *Nat Genet* (2008) 40:489–90. doi:10.1038/ng0508-489

58. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet* (2010) 376:1393–400. doi:10.1016/S0140-6736(10)61267-6

59. Tikkanen E, Havulinna AS, Palotie A, Salomaa V, Ripatti S. Genetic risk prediction and a 2-stage risk screening strategy for coronary heart disease. *Arterioscler Thromb Vasc Biol* (2013) 33:2261–6. doi:10.1161/ATVBAHA.112.301120

60. Havulinna AS, Kettunen J, Ukkola O, Osmond C, Eriksson JG, Kesaniemi YA, et al. A blood pressure genetic risk score is a significant predictor of incident cardiovascular events in 32,669 individuals. *Hypertension* (2013) 61:987–94. doi:10.1161/HYPERTENSIONAHA.111.00649

61. Ganna A, Magnusson PK, Pedersen NL, de Faire U, Reilly M, Arnlov J, et al. Multilocus genetic risk scores for coronary heart disease prediction. *Arterioscler Thromb Vasc Biol* (2013) 33:2267–72. doi:10.1161/ATVBAHA.113.301218

62. Tada H, Shiffman D, Smith JG, Sjogren M, Lubitz SA, Ellinor PT, et al. Twelve-single nucleotide polymorphism genetic risk score identifies individuals at increased risk for future atrial fibrillation and stroke. *Stroke* (2014) 45:2856–62. doi:10.1161/STROKEAHA.114.006072

63. Pfeufer A, Sanna S, Arking DE, Muller M, Gateva V, Fuchsberger C, et al. Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat Genet* (2009) 41:407–14. doi:10.1038/ng.362

64. Cousins MS, Shickle LM, Bander JA. An introduction to predictive modeling for disease management risk stratification. *Dis Manage* (2002) 5:157–67. doi:10.1089/109350702760301448

65. Thanassoulis G, Vasan RS. Genetic cardiovascular risk prediction: will we get there? *Circulation* (2010) 122:2323–34. doi:10.1161/CIRCULATIONAHA.109.909309

66. Hosein FS, Bobrovitz N, Berthelot S, Zygun D, Ghali WA, Stelfox HT. A systematic review of tools for predicting severe adverse events following patient discharge from intensive care units. *Crit Care* (2013) 17:R102. doi:10.1186/cc12747

67. Shaw LJ, Min JK, Hachamovitch R, Hendel RC, Borges-Neto S, Berman DS. Nomograms for estimating coronary artery disease prognosis with gated stress myocardial perfusion SPECT. *J Nucl Cardiol* (2012) 19:43–52. doi:10.1007/s12350-011-9468-7

68. European Heart Rhythm Association, Heart Rhythm Society, Zipes DP, Camm AJ, Borggrefe M, Buxton AE, et al. ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force and the European Society of Cardiology Committee for Practice Guidelines (writing committee to develop guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death). *J Am Coll Cardiol* (2006) 48:e247–346. doi:10.1016/j.jacc.2006.07.010

69. Talmud PJ, Cooper JA, Palmen J, Lovering R, Drenos F, Hingorani AD, et al. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. *Clin Chem* (2008) 54:467–74. doi:10.1373/clinchem.2007.095489

70. Schorfheide F, Wolpin KI. On the use of holdout samples for model selection. *Am Econ Rev* (2012) 102:477–81. doi:10.1257/aer.102.3.477

71. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York & London: Taylor & Francis (1994).

72. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast* (2006) 22:679–88. doi:10.1016/j.ijforecast.2006.03.001

73. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *J Mach Learn Technol* (2011) 2:37–63.

frontiers
in Cardiovascular Medicine

CrossMark

Cardiovascular Cascade Genetic Testing: Exploring the Role of Direct Contact and Technology

Amy C. Sturm[1,2]*

[1] Department of Internal Medicine, Division of Human Genetics, Ohio State University Wexner Medical Center, Columbus, OH, USA, [2] Ohio State University Wexner Medical Center, Dorothy M. Davis Heart and Lung Research Institute, Columbus, OH, USA

> Cascade screening is one of the more forceful demonstrations that molecular biology and genetics are not just a tool for researchers, but represent an important and by now essential component of good medical care.
>
> – Peter J. Schwartz (1)

INTRODUCTION

There is much attention and excitement in the current health care environment on the potential of precision medicine based on a patient's genomic data. Today, what arguably remains as one of the most valuable and informative genetic tests is that of predictive testing for a known familial pathogenic variant. Predictive genetic testing determines whether the pathogenic variant previously identified in an affected family member(s) is present or not in relatives at risk. Previous research has documented that affected individuals undergoing genetic testing cite obtaining genetic information for others as being the most important, if not the only, motivation for undergoing genetic testing (2). Predictive, cascade testing is able to separate at-risk relatives who require vigilant serial screening from those who do not. For those with the predisposition, clinical screening allows for early identification of the family's phenotype, which when present, may require lifelong medical therapy, implantation of devices, and/or other types of medical management. Relatives who test negative for the familial variant can typically be released from lifelong screening. In addition, it is also then known that their children are not at increased risk for the family's disease. This approach can save the health care system, and the family itself, thousands of dollars. Cascade screening is imperative with "high-stakes" cardiovascular conditions, such as familial hypercholesterolemia (FH), long QT syndrome (LQTS), and other inherited arrhythmias, as well as other heritable cardiovascular phenotypes, including cardiomyopathies and aneurysms, where there is an increased risk for sudden cardiac death and severe morbidities such as heart failure.

The value of cascade screening for highly penetrant cardiovascular (and cancer) phenotypes has been acknowledged by public health officials. The United States Centers for Disease Control Office of Public Health Genomics classifies cascade screening of at-risk relatives for certain conditions, FH being one, as a Tier 1 genomic application, meaning it meets the criteria for analytic and clinical validity and utility and therefore has evidence supporting its implementation into practice (3). Cascade screening may include targeted genetic testing as well as clinical screening (e.g., lipid panel) of at-risk relatives.

This opinion provides a brief summary of research in this area and poses questions to facilitate future discussion regarding the potential for direct contact of at-risk relatives. As a practicing genetic

counselor in clinical genetic medicine for 14 years who has provided genetic counseling and testing to thousands of families with heritable cardiovascular and cancer conditions, it is my opinion that more could be done to provide assistance to probands for at-risk relative notification and that genetic counselors are in the ideal position to facilitate cascade testing and lead forward-thinking research in this area (4).

CASCADE SCREENING: WHERE ARE WE NOW?

Cascade screening is a mechanism for identifying people at risk for a genetic condition by a process of systematic family tracing. It should begin with first-degree relatives (parents, siblings, and children) and then extend to second- and third-degree relatives in a stepwise, cascade fashion, moving through the pedigree in sequential steps as additional family members are diagnosed until all at-risk relatives have been screened (5). Cascade screening for FH is a cost-effective method for identifying new cases of FH (6–8). Cascade screening in families with inherited arrhythmia syndromes has been shown to lead to immediate prophylactic treatment, including drug treatment or implantation of pacemakers or cardioverter defibrillators (9). However, cascade screening is not effective unless at-risk relatives are first notified of their risk, the health implications of the inherited condition in their family, the availability of testing, with subsequent uptake. However, uptake of genetic counseling and predictive genetic testing has been shown to be inadequate (10). While there is support from payers, public health, and health care providers (HCPs) regarding the importance of cascade testing, how best to inform relatives of their risk and systematically implement cascade testing has yet to be determined.

   Psychological, educational, geographical, and other barriers exist to family communication of genetic risk information. Ethical factors and family dynamics, including maintenance of confidentiality and privacy, potential for psychological harm and genetic discrimination (i.e., life insurance), balancing the right "not to know" with "duty to warn," among others, must be considered (11). The currently recommended approach for FH, made by the International FH Foundation, includes the following: (1) the proband's HCP should construct a pedigree that facilitates identification of at-risk relatives who should be offered testing; (2) the HCP should discuss risk notification with the proband; and (3) the proband should be provided with written information that includes general information about the family's condition, the benefit, and availability of preventive therapies, emphasizes health consequences without testing, and be encouraged to share this with relatives (12). This approach should be taken with other highly penetrant autosomal dominant conditions. In one study specific to inherited arrhythmias and cardiomyopathies, probands were asked to distribute "family letters" containing information on risks, genetic and other screening tests, and preventive options to relatives at risk. In this study, 57% of informed relatives underwent screening (80% in arrhythmia families; 45% in cardiomyopathy families), and this was statistically significant

when compared to the group where no family letter was provided (35%). While such "family letters" increased the number of relatives who presented for evaluation, over 50% in cardiomyopathy families and 43% overall of at-risk relatives had no documentation that they underwent cascade evaluation (13).

   It has been suggested that it is not outrageous to expect that clinicians, once they have diagnosed a patient with a genetic arrhythmia, "track down" all at-risk family members and determine their genetic status (1). However, realistically, implementation of this approach is problematic since many health care systems do not support this type of family-centric care model. Specifically, a recent review presents health policy-related limitations faced in the United States to effective implementation of cascade screening and includes (1) a low rate of reimbursement for comprehensive genetic counseling services; (2) an individual, versus family-centric, approach to prevention and insurance coverage; (3) insufficient genetic risk assessment and knowledge by a majority of HCPs without genetics credentials; and (4) a shortage of genetics specialists (in rural areas especially) (14). In order to begin addressing and overcoming these challenges, research should be conducted demonstrating effectiveness of novel methods and tools that have the capacity to efficiently notify relatives of risk. These tools should provide education, offer support, and provide attainable next steps with calls to action so that probands can be assisted, and their relatives can understand their own risk and be supported to act on it.

DIRECT CONTACT IN CASCADE SCREENING: SHOULD WE TAKE A MORE ACTIVE APPROACH?

Different methods of informing relatives of risk exist including (1) proband, or family-mediated, contact; (2) proband, or family-mediated, contact *with assistance* (provision of materials, such as a family letter or other written information aids, by the HCP to the proband); and (3) direct contact of at-risk relatives by the clinical service itself.

   Research suggests that clinical providers may take an active approach and directly contact relatives to notify them of their risk without compromising privacy or autonomy, with significantly higher numbers of relatives whose genetic status is clarified for greater efficiency, and with high levels of acceptability (15–18). A thematic analysis of FH proband interviews found that probands believed they had insufficient authority or control to persuade family members to attend screening and that they welcomed greater assistance from the clinic for contact with relatives (19). Also in support of direct contact is increased accuracy, as errors may occur in proband-mediated transmission of genetic testing result information through families (20). However, a prior study found that FH patients who expressed a preference regarding cascading method favored indirect contact because they considered it less threatening to family members (21). A genetic counseling intervention study that offered direct contact to the index patient as a last option for assistance in informing at-risk relatives reported no uptake; only eight index patients

were offered this service, however, and none of the patients in this study had cardiovascular phenotypes (22). A recent literature review concluded that most studies support direct contact of relatives *via* letter mailed from the provider and that provider-initiated communication more often resulted in relatives being tested compared to other methods of communication (16).

Regarding additional Tier 1 conditions, a prospective study of families with *BRCA* mutations associated with Hereditary Breast Ovarian Cancer syndrome compared proband-mediated contact to a direct contact intervention protocol that included a letter and subsequent phone call to at-risk relatives (17). This study concluded that the direct contact protocol nearly doubled the number of relatives tested and was also found to be psychologically safe. A direct contact study in families with Lynch syndrome, or hereditary non-polyposis colorectal cancer, demonstrated high approval in those who consented to participate, with a third of newly diagnosed mutation carriers having cancer identified in their first post-test colonoscopy. This type of data demonstrates acceptability of direct contact risk notification programs, as well as efficacy, feasibility, and also ethical responsibility.

From the perspective of those potentially at risk, a study conducted in Australia assessing community members' viewpoints showed that over 90% of respondents indicated their desire to be informed about a familial risk of FH and to be offered screening, with evidence of strong community support for direct contact by an FH clinic (23). The "right *to* know" must also be considered.

FUTURE DIRECTIONS

Research evaluating genetic counseling interventions focused on strengthening family communication, the number of relatives informed of risk, and the impact on uptake of genetics services is ongoing and will help inform future efforts (22, 24). A randomized controlled trial studying whether a specifically designed genetic counseling intervention that included telephone support up to three times post new genetic diagnosis showed no overall significant difference for the level of family communication between the intervention and control groups (25). In this study, the level of family communication was the highest for conditions with appropriate treatments or active surveillance, such as LQTS and hypertrophic cardiomyopathy. While promising, the level still only reached ~30%. These data again beg the question regarding the potential role for direct contact, especially in "high stakes" conditions.

Most, if not all, of the research conducted to date specific to direct contact has been done outside the United States. Therefore, there is a real need for research to determine whether direct contact methods would be acceptable to probands, at-risk relatives, and HCPs within the United States. How many probands might indeed welcome and appreciate this assistance and support and opt in to programs that work *with* and/or *for* them to assist in disclosure of risk information to relatives? This opinion piece does not propose that we break probands' confidentiality and throw privacy to the wind. Instead, it hopes to promote additional

conversation and brainstorming that may lead to the development and testing of innovative models of care for probands with highly penetrant, yet manageable conditions. The ultimate goal is that we will have greater impact in our work with these families where there are clear risk-reducing interventions. Probands and family members should be engaged in shaping these models and the research testing them, starting now!

The next question becomes, what is feasible now in the landscape of our current health care system? Can we systematize the collection of informed consent from probands to directly share their protected health information with relatives for which they provide the clinic contact information? Can we offer probands active assistance in family communication of genetic risk information? In the pediatric setting, is there a role for standardized direct contact of HCPs caring for the at-risk children in our pedigrees with FH, other Tier One conditions, and beyond? This may be a service welcomed by the affected parent proband, who may appreciate greater assistance in coordination of care for their at-risk children and other pediatric members of their family.

Advances in web-based technologies and novel models for the delivery of genetic counseling may be able to bring cascade testing more effectively and efficiently to larger numbers of at-risk relatives. For example, home-based online genetic counseling sessions for cardiovascular genetic cascade screening can be effective (26), allowing at-risk individuals to access their genetic risk information at the time of their choosing and without having to travel to a hospital or clinic, a barrier mentioned previously. In addition, interactive e-learning and decisional support e-tools available *via* informative websites and mobile applications have been used in pre-test genetic counseling with high knowledge and satisfaction, leading toward the "e-informed" patient (27). Mobile health applications have been shown to result in more "activated" patients – defined as individuals who believe their roles are important, that they have the confidence and knowledge needed to take action, and that they can engage in health-promoting behaviors (28), such as predictive genetic testing. Probands with higher activation may lead toward more at-risk relatives notified of their risk. In turn, e-learning information, such as an informational video about the family's inherited cardiovascular disease, could then be delivered to relatives, who may then become activated themselves to pursue cascade testing.

The power of preventive genetic and genomic information is real – that is not the question. How to ensure this information gets into the hands of all that need it, including children, however, needs more active attention.

In conclusion, a powerful quote from Newson and Humphries (11): "Our biology does not stop: the risk of developing coronary heart disease as a consequence of FH will still be present, even if relatives live in ignorance."

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

1. Schwartz PJ. Cascades or waterfalls, the cataracts of genetic screening are being opened on clinical cardiology. *J Am Coll Cardiol* (2010) 55(23):2577–9. doi:10.1016/j.jacc.2009.12.064

2. Hallowell N, Foster C, Eeles R, Ardern-Jones A, Murday V, Watson M. Balancing autonomy and responsibility: the ethics of generating and disclosing genetic information. *J Med Ethics* (2003) 29(2):74–9. doi:10.1136/jme.29.2.74

3. Centers for Disease Control and Prevention. *Genomic Tests and Family Healthy History by Levels of Evidence*. (2016). Available from: http://www.cdc.gov/genomics/gtesting/tier.htm

4. Sturm AC. The role of genetic counselors for patients with familial hypercholesterolemia. *Curr Genet Med Rep* (2014) 2(2):68–74. doi:10.1007/s40142-014-0036-8

5. Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *Eur Heart J* (2013) 34(45):3478a–90a. doi:10.1093/eurheartj/eht273

6. Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HA. Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Health Technol Assess* (2000) 4(29):1–123. doi:10.3310/hta4290

7. Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HA. Cost effectiveness analysis of different approaches of screening for familial hypercholesterolaemia. *BMJ* (2002) 324(7349):1303. doi:10.1136/bmj.324.7349.1303

8. Nherera L, Marks D, Minhas R, Thorogood M, Humphries SE. Probabilistic cost-effectiveness analysis of cascade screening for familial hypercholesterolaemia using alternative diagnostic and identification strategies. *Heart* (2011) 97(14):1175–81. doi:10.1136/hrt.2010.213975

9. Hofman N, Tan HL, Alders M, van Langen IM, Wilde AA. Active cascade screening in primary inherited arrhythmia syndromes: does it lead to prophylactic treatment? *J Am Coll Cardiol* (2010) 55(23):2570–6. doi:10.1016/j.jacc.2009.12.063

10. Christiaans I, Birnie E, Bonsel GJ, Wilde AA, van Langen IM. Uptake of genetic counselling and predictive DNA testing in hypertrophic cardiomyopathy. *Eur J Hum Genet* (2008) 16(10):1201–7. doi:10.1038/ejhg.2008.92

11. Newson AJ, Humphries SE. Cascade testing in familial hypercholesterolaemia: how should family members be contacted? *Eur J Hum Genet* (2005) 13(4):401–8. doi:10.1038/sj.ejhg.5201360

12. Watts GF, Gidding S, Wierzbicki AS, Toth PP, Alonso R, Brown WV, et al. Integrated guidance on the care of familial hypercholesterolaemia from the International FH Foundation. *Int J Cardiol* (2014) 171(3):309–25. doi:10.1016/j.ijcard.2013.11.025

13. van der Roest WP, Pennings JM, Bakker M, van den Berg MP, van Tintelen JP. Family letters are an effective way to inform relatives about inherited cardiac disease. *Am J Med Genet A* (2009) 149A(3):357–63. doi:10.1002/ajmg.a.32672

14. George R, Kovak K, Cox SL. Aligning policy to promote cascade genetic screening for prevention and early diagnosis of heritable diseases. *J Genet Couns* (2015) 24(3):388–99. doi:10.1007/s10897-014-9805-5

15. Suthers GK, Armstrong J, McCormack J, Trott D. Letting the family know: balancing ethics and effectiveness when notifying relatives about genetic testing for a familial disorder. *J Med Genet* (2006) 43(8):665–70. doi:10.1136/jmg.2005.039172

16. Allison M. Communicating risk with relatives in a familial hypercholesterolemia cascade screening program: a summary of the evidence. *J Cardiovasc Nurs* (2015) 30(4):E1–12. doi:10.1097/JCN.0000000000000153

17. Sermijn E, Delesie L, Deschepper E, Pauwels I, Bonduelle M, Teugels E, et al. The impact of an interventional counselling procedure in families with a BRCA1/2 gene mutation: efficacy and safety. *Fam Cancer* (2016) 15(2):155–62. doi:10.1007/s10689-015-9854-4

18. van Maarle MC, Stouthard MEA, Marang-van de Mheen PJ, Klazinga NS, Bonsel GJ. How disturbing is it to be approached for a genetic cascade screening programme for familial hypercholesterolaemia? Psychological impact and screenees' views. *Community Genet* (2001) 4(4):244–52. doi:10.1159/000064200

19. Hardcastle SJ, Legge E, Laundy CS, Egan SJ, French R, Watts GF, et al. Patients' perceptions and experiences of familial hypercholesterolemia, cascade genetic screening and treatment. *Int J Behav Med* (2015) 22(1):92–100. doi:10.1007/s12529-014-9402-x

20. Vos J, Menko F, Jansen AM, van Asperen CJ, Stiggelbout AM, Tibben A. A whisper-game perspective on the family communication of DNA-test results: a retrospective study on the communication process of BRCA1/2-test results between proband and relatives. *Fam Cancer* (2011) 10(1):87–96. doi:10.1007/s10689-010-9385-y

21. Hallowell N, Jenkins N, Douglas M, Walker S, Finnie R, Porteous M, et al. Patients' experiences and views of cascade screening for familial hypercholesterolemia (FH): a qualitative study. *J Community Genet* (2011) 2(4):249–57. doi:10.1007/s12687-011-0064-y

22. Forrest LE, Burke J, Bacic S, Amor DJ. Increased genetic counseling support improves communication of genetic information in families. *Genet Med* (2008) 10(3):167–72. doi:10.1097/GIM.0b013e318164540b

23. Maxwell SJ, Molster CM, Poke SJ, O'Leary P. Communicating familial hypercholesterolemia genetic information within families. *Genet Test Mol Biomarkers* (2009) 13(3):301–6. doi:10.1089/gtmb.2008.0138

24. Gaff C, Hodgson J. A genetic counseling intervention to facilitate family communication about inherited conditions. *J Genet Couns* (2014) 23(5):814–23. doi:10.1007/s10897-014-9696-5

25. Hodgson J, Metcalfe S, Gaff C, Donath S, Delatycki MB, Winship I, et al. Outcomes of a randomised controlled trial of a complex genetic counselling intervention to improve family communication. *Eur J Hum Genet* (2016) 24(3):356–60. doi:10.1038/ejhg.2015.122

26. Otten E, Birnie E, Ranchor AV, van Langen IM. Online genetic counseling from the providers' perspective: counselors' evaluations and a time and cost analysis. *Eur J Hum Genet* (2016). doi:10.1038/ejhg.2015.283

27. Birch PH. Interactive e-counselling for genetics pre-test decisions: where are we now? *Clin Genet* (2015) 87(3):209–17. doi:10.1111/cge.12430

28. Ledford CJ, Ledford CJ, Canzona MR, Cafferty LA, Hodge JA. Mobile application as a prenatal education and engagement tool: a randomized controlled pilot. *Patient Educ Couns* (2015) 99(4):578–82. doi:10.1016/j.pec.2015.11.006

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Who Pays? Coverage Challenges for Cardiovascular Genetic Testing in U.S. Patients

Katherine G. Spoonamore[1]* and Nicole M. Johnson[2]

[1] Department of Medicine, Krannert Institute of Cardiology, Indiana University School of Medicine, Indianapolis, IN, USA,
[2] Invitae Corporation, San Francisco, CA, USA

Inherited cardiovascular (CV) conditions are common, and comprehensive care of affected families often involves genetic testing. When the clinical presentations of these conditions overlap, genetic testing may clarify diagnoses, etiologies, and treatments in symptomatic individuals and facilitate the identification of asymptomatic, at-risk relatives, allowing for often life-saving preventative care. Although some professional society guidelines on inherited cardiac conditions include genetic testing recommendations, they quickly become outdated owing to the rapid expansion and use of such testing. Currently, these guidelines primarily discuss the benefits of targeted genetic testing for identifying at-risk relatives. Although most insurance policies acknowledge the benefit and the necessity of this testing, many exclude coverage for testing altogether or are vague about coverage for testing in probands, which is imperative if clinicians are to have the best chance of accurately identifying pathogenic variant(s) in a family. In response to uncertainties about coverage, many commercial CV genetic testing laboratories have shouldered the burden of working directly with commercial payers and protecting patients/institutions from out-of-pocket costs. As a result, many clinicians are unaware that payer coverage policies may not match professional recommendations for CV genetic testing. This conundrum has left patients, clinicians, payers, and laboratories at an impasse when determining the best path forward for meaningful and sustainable testing. Herein, we discuss the need for all involved parties to recognize their common goals in this process, which should motivate collaboration in changing existing frameworks and creating more sustainable access to genetic information for families with inherited CV conditions.

Keywords: genetic testing, insurance coverage, cardiovascular genetics, preventative care, access barriers, cascade testing

INTRODUCTION

Inherited cardiovascular (CV) conditions include arrhythmias, cardiomyopathies, aortopathies, and dyslipidemias. These conditions affect more than 1 in 200 individuals, and several of them have considerable phenotypic overlap. Therefore, comprehensive care of affected patients and families often involves multi-gene panel-based genetic testing, which can clarify diagnoses and etiologies. Since becoming available in the early to mid-2000s, panel-based CV genetic tests have seen widening clinical adoption. The range of conditions covered by current commercial CV genetic testing and the number of genes included in analyses have also expanded exponentially.

Advances in CV diagnostics have spurred changes for insurance companies, genetic testing laboratories, and professional cardiology societies, which have created coverage policies, developed expanded panel-based tests, and formulated care guidelines, respectively. However, the question of who pays for CV genetic testing is ongoing and relates to both the applicability of genetic testing in probands and at-risk relatives and the need for sustainability in laboratory services and payer policies.

Herein, we review the current status of genetic testing guidelines for inherited CV conditions and the roles of payers and genetic testing laboratories in providing access to testing for affected families. We also discuss the inconsistencies among clinical approaches, professional society guidelines, payer policies, and laboratory practices that have influenced this access. Finally, we highlight the shared goals of all stakeholders and discuss how these overlapping interests are a starting point on the path to sustainable, accessible genetic testing for patients with inherited CV conditions.

HISTORY OF CV GENETIC TESTING

Although genotype–phenotype correlations remain in their infancy in CV genetics, genetic test results can have key impacts on patient care and management by clarifying clinical presentation and etiology, aiding decision-making about surgical procedures, and guiding medication selection and surveillance strategies. For example, long QT syndrome subtyping (for types 1, 2, and 3) provides some indication of both responsiveness to certain treatments and the presence of higher-risk situations that may trigger cardiac events. In cardiac hypertrophy, genetic testing can identify underlying causal conditions (e.g., Fabry disease, Danon disease, Pompe disease, transthyretin amyloidosis). Alfares et al. (1) found that 3% of hypertrophic cardiomyopathy (HCM) patients who underwent genetic testing had an undetected syndromic disease that presented an opportunity for more effective treatment (e.g., enzyme replacement therapy in Fabry disease). Earlier clarification of cardiomyopathy etiology through genetic testing in children, in which metabolic causes are much more frequent, can also improve treatment outcomes. When cardiomyopathies are associated with conduction disease or higher arrhythmogenic potential, the increased likelihood of changes in certain genes (e.g., *SCN5A*, *LMNA*) warrant closer surveillance and specific intervention from a cardiac electrophysiologist and a heart failure/cardiomyopathy specialist (2). Moreover, in the case of overlapping aortopathies (Marfan syndrome versus Loeys–Dietz syndrome), genetic test results can guide the timing of surgical intervention, which differs based on the etiology of aortic disease (3). The risks for recurring aortic aneurysms/dissection, the most vulnerable portions of the aorta, and the involvement of additional vasculature also vary according to etiology; therefore, genetic testing can guide the choice of imaging method and frequency of ongoing surveillance.

Until recently, long turnaround times – typically 8–12 weeks – could be expected for CV genetic tests. Therefore, the results have generally been less routinely useful for planning immediate patient care. Furthermore, CV genetic test results may not have direct management implications for the individual tested. Often, the primary benefit of CV genetic testing comes in uncovering pathogenic variants in probands that can then be used for targeted genetic testing to identify at-risk family members (cascade screening) and plan their surveillance and, equally important, reduce risk in family members to the population baseline when they test negative for variants. Early diagnosis of hereditary CV conditions improves outcomes; therefore, early identification of at-risk family members improves outcomes as well.

Historically, CV genetic testing has been covered only sporadically by insurance and has been cost prohibitive for patients. The ability to direct family cardiac screening is valuable for both patients and payers, but this reason alone is not always a convincing argument for why payers should cover testing for probands. For example, Medicare specifically prohibits the genetic testing of both affected patients and asymptomatic at-risk family members, if the test will benefit individuals other than Medicare patients themselves.

PROFESSIONAL GUIDELINES FOR CV GENETIC TESTING

Practice guidelines drafted by professional cardiology and genetics societies aim to provide patient care recommendations (evidence-based, when possible) that lead to the best clinical outcomes. The guidelines for the inherited arrhythmias and cardiomyopathies currently address genetic testing most thoroughly. These guidelines were published in 2009 by the Heart Failure Society of America (4) and in 2011 by the American College of Cardiology Foundation/American Heart Association (ACCF/AHA) (5) and the Heart Rhythm Society/European Heart Rhythm Association (HRS/EHRA) (2).

Genetic testing is a class I recommendation ("is recommended") in probands for just 5 of the 13 conditions covered in the HRS/EHRA document, including individuals with strong clinical suspicion of long QT syndrome, catecholaminergic polymorphic ventricular tachycardia, HCM, and dilated cardiomyopathy in the presence of conduction disease or a family history of premature unexpected death, as well as individuals who have survived an out-of-hospital cardiac arrest when a specific channelopathy or cardiomyopathy is suspected. However, cascade testing for a pathogenic variant, previously identified in a family proband, is a class I recommendation for all but 1 of the 13 conditions included.

The available professional guidelines are sometimes inconsistent. For example, unlike the HRS/EHRA statement described above, the ACCF/AHA guidelines for HCM recommend genetic testing only for probands with atypical presentations that raise suspicion of an underlying syndromic etiology. For all other individuals with HCM, the guidelines classify genetic testing as a class II recommendation ("is reasonable"), specifically to facilitate the identification of at-risk family members (5).

Unlike the National Comprehensive Cancer Network guidelines for oncology, which are updated annually to provide recommendations for genetic testing, the guidelines provided by professional societies in CV medicine are updated too infrequently to serve as comprehensive recommendations for patient management. In addition, the National Comprehensive Cancer Network guidelines are highly specific about the genes to test and how the results of testing will influence management and surveillance of the proband undergoing testing. A lack of available

clinical data on the use of genetic testing to improve long-term outcomes in patients with inherited CV conditions means that much of the professional guidance for cardiac genetic testing is based on expert opinion and experience rather than accumulated evidence. As such, the NCCN guidelines are closely followed by many major payers, unlike the current cardiology recommendations, which are rarely consistent with insurance coverage policies and may be considered insufficient by payers for determining which tests and which individuals to cover in affected families.

The American College of Medical Genetics and Genomics has published a "must-report" guideline related to clinically useful pathogenic test results of whole-exome or whole-genome analysis (6). The guideline specifies 56 genes for which findings are important for all patients to know and should be conveyed by clinicians even if they are secondary to a patient's original indication for genetic analysis. This guideline underscores the value of genetic information for clinicians engaged in patient care and their desire to use genetic test results to guide the care they provide. Thirty-one of these genes are related to cardiac conditions; however, commercial payers do not cover genetic testing for some of these genes, even in probands suspected of having the condition.

PAYER POLICIES

Payer policies are driven by the goal of providing quality health care to all clients in a sustainable, cost-effective way. The downstream cost savings of initiating appropriate genetic testing in a family proband with cardiomyopathy followed by targeted genetic testing in related family members are considerable (1, 7, 8). These savings occur when relatives who did not inherit a known familial pathogenic variant can be released from further cardiac surveillance, and the testing and intervention recommendations for at-risk family members can be refined and optimized. However, evidence of these cost savings has not yet translated to wider payer coverage for genetic testing in probands.

Coverage policies for CV genetic testing are inconsistent among payers. When policies do include panel testing, different payers sometimes cover testing for different genes for the same condition. Because professional guidelines do not offer up-to-date, gene-specific, evidence-based guidance, it is unclear who is selecting the genes to be covered and what information is guiding or informing the selections.

For clinicians and families, these inconsistencies impede efficient decision-making and delivery of care. Written policies on medical necessity for specific CV conditions are often unavailable, which means that clinicians and patients have no assurance that genetic testing will be authorized or covered. This lack of specific documentation exists even for Medicare/Medicaid policies, in which coverage details for non-oncology genetic testing rarely exist and, when present, are tied to medical necessity. Medical necessity often remains undetermined by payers until a claim is submitted, and the expectations of patients and clinicians about the medical necessity of testing often differ significantly from the definitions adhered to by payers.

Even when CV genetic testing is covered, clinical decisions are further complicated by the extent of coverage provided. In many cases, payers cover only testing deemed medically necessary for the individual covered. Familial probands must undergo genetic testing to determine the underlying genetic cause of an inherited CV condition before cascade testing can begin. Until then, at-risk family members whose genetic testing is medically necessary and covered by insurance cannot obtain authorization for testing. Coverage denials based on medical necessity in probands create obstacles for both determining the underlying genetic cause of a familial CV condition and allowing at-risk family members to take advantage of genetic testing that is covered under their policies. Coverage/no-coverage combinations within affected families can become ongoing catch-22s in the management of life-threatening health conditions.

To the best of our knowledge, only one policy, from the commercial payer Aetna (9), successfully navigates the murky waters of proband/at-risk relative coverage for genetic testing. This policy states that the payer will cover *oncologic* genetic testing for a non-member familial proband whose own insurance has denied coverage, if the results are needed to pursue the medically necessary targeted testing of a covered at-risk family member. How often this coverage clause is used or honored, and what mechanism the payer has established to extend such coverage are unknown, but this example may be a model for consideration in CV genetics, in particular, because the primary benefit of most CV genetic tests is the identification of at-risk relatives.
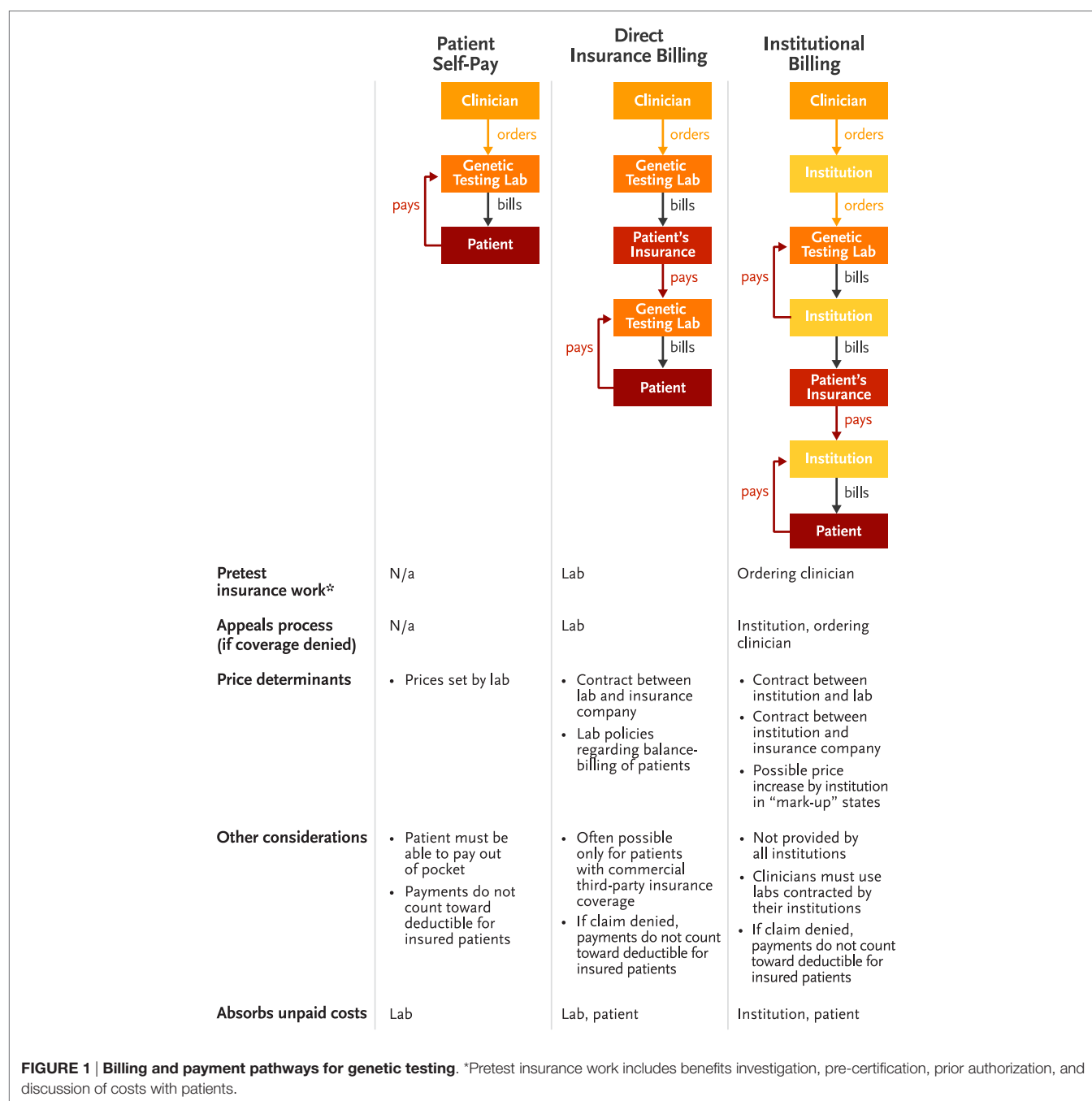
COMMERCIAL LABORATORY BILLING POLICIES

Genetic testing laboratories aim to provide quality, maximally accessible genetic testing to patients and clinicians. Costs and payment processes for genetic testing are generally dictated by the method in which tests are ordered and billed. Institutional billing, in which an institution (hospital) pays the laboratory performing the test and then bills and collects the payment from both the patient's insurance and the patient, is the only option for some clinicians. Other institutions do not allow clinicians to use institutional billing, and instead, require them to work with laboratories that can bill payers directly.

Laboratories that cannot bill insurers directly may require that all orders be handled by an institutional billing process at the clinician's facility. To cover the costs associated with billing and collecting from both patients and payers, institutions in so-called mark-up states may charge more for testing. **Figure 1** demonstrates the complexity of current billing processes for genetic testing.

Many commercial laboratories that bill patient insurance companies directly also devote customer service resources to payment planning or cost reductions for qualifying patients who sometimes bear significant out-of-pocket costs despite having insurance coverage. However, these additional services may be unavailable to patients with non-commercial insurance and are advantageous only to clinicians who can send samples directly to testing laboratories.

Cost often emerges as a key factor for patients in deciding whether to pursue testing recommended by their clinicians. Having laboratories take responsibility for billing and coverage/cost determination has increased both patient access and clinician

FIGURE 1 | **Billing and payment pathways for genetic testing**. *Pretest insurance work includes benefits investigation, pre-certification, prior authorization, and discussion of costs with patients.

utilization of CV genetic testing but has not necessarily improved insurance coverage for these tests. Furthermore, taking these processes out of the hands of clinicians has created a situation in which genetic testing stakeholders do not always realize that treatment decisions, professional guideline recommendations, and payer coverage policies are misaligned.

CHALLENGES IN THE CV CLINIC

The disconnect between practice guidelines and coverage policies presents barriers to the timely and effective provision of care

to patients with inherited CV disorders. Patient confusion can arise when testing that clinicians call "recommended" is considered "experimental" or "investigational" by payers. With vastly different payer policies or no clear policy to rely on, clinicians have difficulty determining whether patients can proceed with genetic testing and, if so, when testing can take place and what out-of-pocket expenses will be incurred (e.g., which tests are covered and which are not and which billing process – the institution's or the laboratory's – will yield the lowest out-of-pocket expense). The time required to find answers to these questions could be better spent counseling patients, and the delays in testing that occur while insurance policies are being clarified

can create added concern for patients facing potentially serious diagnoses.

Furthermore, clinicians are inadequately trained to advise patients about the implications of various billing policies, and their lack of expertise may introduce legal liabilities. Few clinicians can differentiate among a pre-verification, a pre-determination, and a pre-authorization, for example, and even if they can, obtaining these clearances from payers often does not guarantee coverage. Discussions about expense are appropriate and necessary in decisions about patient care and management. However, compared with other routinely ordered medical tests (e.g., echocardiogram, electrocardiogram, magnetic resonance imaging) in cardiology clinics, orders for genetic testing frequently require clinicians to take a more prominent financial/insurance counseling role because uncertainties about coverage put cost at the center of diagnosis and treatment decisions.
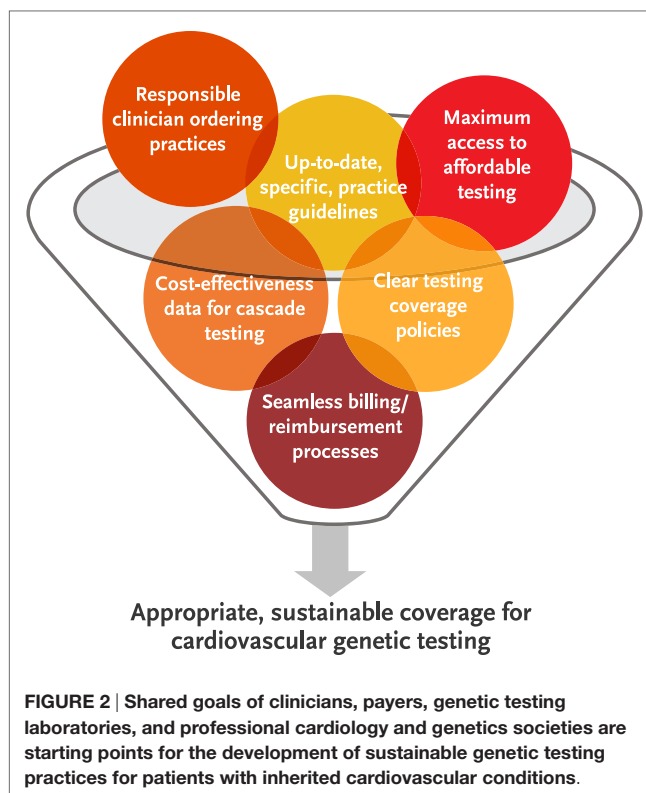
In some cases, clinicians may alter the genetic testing strategy in a family based on the type of insurance coverage available for the required test – for example, selecting a different relative with a better insurance situation for testing. Gathering and assessing all of the necessary documentation to make decisions about testing logistics has the potential to be a complicated process that prevents some patients from receiving recommended and appropriate CV genetic testing in a timely manner.

PATH FORWARD

Genetic testing in the management of inherited CV conditions is here to stay. Its utility in the care of families with inherited CV conditions has been established, and genotype–phenotype correlations will likely become more refined as sequencing technologies advance. Therefore, the establishment of clear professional guidelines and consistent payer policies is crucial if affected families are to benefit from the availability of accurate and effective testing. To make recommendations and coverage work hand in hand, payers, laboratories, clinicians, and professional CV and genetics societies must collaborate and recognize their shared goals in caring for these patients (**Figure 2**).

Because all stakeholders agree that cascade testing can improve outcomes through early identification of individuals at risk for inherited CV conditions, the key issues are primarily those about coverage and billing. Should a family member's insurance policy pay for CV genetic testing in a proband in some scenarios, as exemplified by the Aetna policy described above? If so, insurers must collectively determine what policy clauses or mechanisms are required to ensure that patients can benefit regardless of the coverage combinations within their families. Input from professional organizations – perhaps through more frequent updating of published CV genetic testing practice guidelines – is likely to help in standardizing the types of tests clinicians order and payers cover.

In the meantime, if billing continues to be handled by laboratories, clinicians should collect data about which tests are being covered and denied to improve responsible ordering practices. Many clinicians prefer laboratory billing because it saves time, and owing to laboratory-based customer service resources, this path often provides assurance that patients will not see unexpected bills. However, if this billing arrangement is not ensuring coverage and

FIGURE 2 | **Shared goals of clinicians, payers, genetic testing laboratories, and professional cardiology and genetics societies are starting points for the development of sustainable genetic testing practices for patients with inherited cardiovascular conditions**.

hides gaps in coverage, it will neither provide sustainable patient access to testing nor help clinicians advocate for changes in payer policies. Clinicians must recognize opportunities for engaging with the billing process and educating payers and guideline writers on the need for and applications of CV genetic testing.

From a health economics standpoint, available data suggest that genetic testing can provide cost savings (1, 7, 8). However, additional data are needed to demonstrate its cost-effectiveness. Also needed are specific, up-to-date practice guidelines backed by appropriate cardiology and genetics societies (e.g., HRS, ACCF, AHA, National Society of Genetic Counselors, American College of Medical Genetics and Genomics) to encourage appropriate guideline implementation and reduce misdirected use of genetic testing, which drives up health-care costs for payers without benefiting patients and families. These could be frequently re-evaluated, making any necessary updates or changes, akin to the process followed by the National Comprehensive Cancer Network for the Clinical Practice Guidelines in Oncology regarding Genetic/Familial High-Risk Assessment. Discussions between clinicians and payers about coverage will be critical as the costs of genetic testing decrease.

The current landscape of CV genetic testing is complex and involves stakeholders with different purposes, constraints, and scopes of care. No single entity can resolve the current challenges alone, and all parties must understand each other's points of view and recognize opportunities for clearing the path toward more effective and accessible genetic testing coverage. For example, clinicians are well positioned to partner with payers to help conduct necessary research about the clinical utility of currently available CV genetic testing and cost-effectiveness of cascade screening,

and professional societies are uniquely positioned to update and maintain consensus testing guidelines that can inform both clinicians and payers. Only through such collective understanding and discussion will processes and policies emerge that both safeguard clinician and patient access to testing and guarantee sustainability for laboratories and payers.

AUTHOR CONTRIBUTIONS

KS and NJ contributed equally to the conception of this work, the drafting and critical revision of the content, and approval of the final version to be published. They agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy and integrity of any part of the work are appropriately investigated and resolved.

REFERENCES

1. Alfares AA, Kelly MA, McDermott G, Funke BH, Lebo MS, Baxter SB, et al. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genet Med* (2015) 17:880–8. doi:10.1038/gim.2014.205

2. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Europace* (2011) 13:1077–109. doi:10.1093/europace/eur245

3. Hiratzka LF, Bakris GL, Beckman JA, Bersin RM, Carr VF, Casey DE Jr, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with thoracic aortic disease. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine. *J Am Coll Cardiol* (2010) 55:e27–129. doi:10.1016/j.jacc.2010.02.015

4. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MR, Towbin JA, et al. Genetic evaluation of cardiomyopathy – a Heart Failure Society of America practice guideline. *J Card Fail* (2009) 1:83–97. doi:10.1016/j.cardfail.2009.01.006

5. Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, et al. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation* (2011) 124:2761–96. doi:10.1161/CIR.0b013e318223e230

6. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* (2013) 15:565–74. doi:10.1038/gim.2013.73

7. Ingles J, McGaughran J, Scuffham PA, Atherton J, Semsarian C. A cost-effectiveness model of genetic testing for the evaluation of families with hypertrophic cardiomyopathy. *Heart* (2012) 98:625–30. doi:10.1136/heartjnl-2011-300368

8. Wordsworth S, Leal J, Blair E, Legood R, Thomson K, Seller A, et al. DNA testing for hypertrophic cardiomyopathy: a cost-effectiveness model. *Eur Heart J* (2010) 31:926–35. doi:10.1093/eurheartj/ehq067

9. Aetna Clinical Policy Bulletin Number 0140. *Genetic Testing*. (2015). Available from: http://www.aetna.com/cpb/medical/data/100_199/0140.html

Advantages of publishing in Frontiers

OPEN ACCESS

Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW

Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION

Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS

No limit to article distribution and re-use

TRANSPARENT

Editors and reviewers acknowledged by name on published articles

SUPPORT

By our Swiss-based editorial team

IMPACT METRICS

Advanced metrics track your article's impact

GLOBAL SPREAD

5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK

Our network increases readership for your article

Find us on