Genetics in Medicine

A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy

--Manuscript Draft--

Manuscript Number:	GIM-D-17-00727R2
Article Type:	Article
Section/Category:	Molecular Genetics and Pharmacogenetics
Keywords:	Dilated cardiomyopathy; mutation; genetic testing
Corresponding Author:	Diane Fatkin, MD Victor Chang Cardiac Research Institute Darlinghurst, NSW AUSTRALIA
First Author:	Diane Fatkin, MD
Order of Authors:	Diane Fatkin, MD
	Claire Horvat
	Renee Johnson
	Lien Lam
	Jacob Munro
	Francesco Mazzarotto
	Angharad Roberts
	Daniel Herman
	Michael Parfenov
	Alireza Haghighi
	Barbara McDonough
	Stephen DePalma
	Anne Keogh
	Peter Macdonald
	Christopher Hayward
	Amy Roberts
	Paul Barton
	Leanne Felkin
	Eleni Giannoulatou
	Stuart Cook
	JG Seidman
	Christine Seidman
Manuscript Region of Origin:	AUSTRALIA
Abstract:	Purpose: We evaluated strategies for identifying disease-causing variants in genetic testing for dilated cardiomyopathy (DCM). Methods: Cardiomyopathy gene panel testing was performed in 532 DCM patients and 527 healthy control subjects. Rare variants in 41 genes were stratified using variant-level and gene-level characteristics. Results: A majority of DCM cases and controls carried rare protein-altering cardiomyopathy gene variants. Variant-level characteristics alone had limited discriminative value. Differentiation between groups was substantially improved by

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	addition of gene-level information that incorporated ranking of genes based on literature evidence for disease association. The odds of DCM were increased to nearly 9-fold for truncating variants or high-impact missense variants in the subset of 14 genes that had the strongest biological links to DCM ($P < 0.0001$). For some of these genes, DCM-associated variants appeared to be clustered in key protein functional domains. Multiple rare variants were present in many family probands, however, there was generally only one "driver" mutation that co-segregated with disease. Conclusions - Rare variants in cardiomyopathy genes can be effectively stratified by combining variant-level and gene-level information. Prioritization of genes based on their a priori likelihood of disease causation is a key factor in identifying clinically-actionable variants in cardiac genetic testing.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

March 20, 2018.

James P. Evans, MD, PhD
Editor-in-Chief
Genetics in Medicine
University of North Carolina
Department of Genetics
5095 Genetic Medicine Building
Campus Box 7264
Chapel Hill, NC 275899-7264, USA

Re: GIM-D-17-00727R1
"A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy"

Dear Dr Evans

Thankyou for the opportunity to submit a further revised version of the above manuscript.

We are very pleased that the concerns raised by Reviewers #2 and #3 were addressed during our first revision.

Reviewer #1 had only one minor comment regarding the clinical screening of the control cohort. These details were actually provided in our first revision in an expanded Supplementary Methods. However, to clarify this point for the Reviewer, we have now added the following sentences to the Methods and Discussion in the main text:

Revised MS #2, page 7, para 1, Methods: "The absence of cardiovascular disease in control subjects was ascertained by medical history, and was confirmed by cardiac magnetic resonance imaging in 319 individuals (61%).

Revised MS #2, page 20, para 2, Discussion: "The lack of cardiac investigations in ~40% control subjects is a limitation of this study".

We hope that this work will now be suitable for publication.

Yours Sincerely,

Diane Fatkin, MD (corresponding author).
Molecular Cardiology Division
Victor Chang Cardiac Research Institute
405 Liverpool St,
PO Box 699, Darlinghurst NSW 2010, Australia
PH: 61 2 9295 8618; FAX: 61 2 9295 8601; email: d.fatkin@victorchang.edu.au

GIM-D-17-00727R1
"A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy"

RESPONSES TO REVIEWERS

We are very pleased that the concerns raised by Reviewers #2 and #3 were addressed during our first revision. Reviewer #1 had only one minor comment that we have responded to below.

Reviewer #1

R1-1: ***"****In the response to reviewers, they state "some" control subjects had a cardiac evaluation and in the discussion, they state "The lack of cardiac investigation in all control subjects is a limitation." From this description, it is not evident if 10% or 90% of the cohort was screened for cardiac disease. Furthermore, the cardiac screening of the control cohorts was not described. Is it possible to estimate the percent of the control population that had a cardiac evaluation and describe whether this evaluation was based on cardiac imaging, questionnaires or medical history? This would greatly help the readers can have a better sense of the control group".*

AU: In our first revision of the Supplementary Methods, we did provide information about the clinical evaluation of control subjects (see page 2 in the revised Supplemental Material) as follows:

Boston control subjects: "208 control subjects (aged 57 ± 15 years, 24% males) were recruited from non-cardiac outpatient clinics at Brigham and Women's Hospital as part of a broad volunteer biobank study ("Our Genes"). These individuals had no history of cardiovascular disease, but did not undergo ECG or echocardiographic screening".

UK control subjects: "…319 adult healthy volunteers (42 ± 14 years, 46% males) who had been recruited prospectively via advertisement at the London Institute of Medical Sciences, Imperial College, London. The volunteers all underwent CMR scanning, had no history of medical illness, were not taking regular medication, and did not have evidence of cardiac structural or functional impairment when scanned".

Due to space limitations, the detailed information about the study cohorts was not able to be included in the main manuscript text. We have now added the following text:
Revised MS #2, page 7, para 1, Methods: "The absence of cardiovascular disease in control subjects was ascertained by medical history, and was confirmed by cardiac magnetic resonance imaging in 319 individuals (61%).

Revised MS #2, page 20, para 2, Discussion: "The lack of cardiac investigations in ~40% control subjects is a limitation of this study".

Revised text

A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy

Short title: Genetic Testing in Dilated Cardiomyopathy

Claire Horvat, PhD[1], Renee Johnson, PhD[1], Lien Lam, PhD[2,8], Jacob Munro, BSc, MIT[3], Francesco Mazzarotto, PhD[4], Angharad M. Roberts, MRCP[4], Daniel S. Herman, MD, PhD[2], Michael Parfenov, MD, PhD[2], Alireza Haghighi MD, PhD[2,8,9], Barbara McDonough, RN[2], Steven R. DePalma, PhD[2], Anne M. Keogh, MD,[5,6,7] Peter S. Macdonald, MB BS, PhD[5,6,7], Christopher S. Hayward, MD[5,6,7], Amy Roberts, MD[10], Paul J.R. Barton, PhD[4], Leanne E. Felkin, PhD[4], Eleni Giannoulatou, DPhil[3], Stuart A. Cook, PhD, MRCP[4,11], J.G. Seidman, PhD[2,8], Christine E. Seidman, MD[2,9], Diane Fatkin, MD[1,6,7].

[1] Molecular Cardiology Division, [3] Molecular, Structural and Computational Biology Division and [5] Cardiac Physiology and Transplantation Division, Victor Chang Cardiac Research Institute, Sydney, New South Wales, Australia.

[2] Department of Genetics, Harvard Medical School, Boston MA, USA.

[4] NIHR Cardiovascular Biomedical Research Unit, Royal Brompton & Harefield NHS Foundation Trust, and Imperial College London, London, England.

[6] Cardiology Department, St Vincent's Hospital, Sydney, New South Wales, Australia.

[7] Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia.

[8] Howard Hughes Medical Institute, Boston MA, USA.

[9] Cardiovascular Division, Brigham and Women's Hospital, Boston MA, USA.

[10] Boston Children's Hospital, Boston MA, USA.

[11] Duke-National University of Singapore, and National Heart Center, Singapore

Correspondence:

Diane Fatkin, Victor Chang Cardiac Research Institute, 405 Liverpool St, Darlinghurst, New South Wales 2010, Australia.

Ph: +61 2 9295 8618, Fax: +61 2 9295 8770

Email: d.fatkin@victorchang.edu.au.

DISCLOSURE

The authors declare no conflict of interest.

FUNDING SOURCES

This study was supported by grants from the National Health and Medical Research Council of Australia (D.F., P.S.M.), Harvard Club of Australia Foundation (D.F.), Estate of the Late RT Hall (D.F.), Simon Lee Foundation (D.F.), British Heart Foundation (S.A.C.), National Institute for Health Research Cardiovascular BRU at the Brompton and Harefield & Harefield NHS Foundation Trust and Imperial College London (S.A.C., P.J.B.), Howard Hughes Medical Institute (C.E.S.), National Institutes of Health (D.S.H., J.G.S.), and the Leducq Foundation (S.A.C., J.G.S., C.E.S.).

Abstract

Purpose: We evaluated strategies for identifying disease-causing variants in genetic testing for dilated cardiomyopathy (DCM).

Methods: Cardiomyopathy gene panel testing was performed in 532 DCM patients and 527 healthy control subjects. Rare variants in 41 genes were stratified using variant-level and gene-level characteristics.

Results: A majority of DCM cases and controls carried rare protein-altering cardiomyopathy gene variants. Variant-level characteristics alone had limited discriminative value. Differentiation between groups was substantially improved by addition of gene-level information that incorporated ranking of genes based on literature evidence for disease association. The odds of DCM were increased to nearly 9-fold for truncating variants or high-impact missense variants in the subset of 14 genes that had the strongest biological links to DCM (P <0.0001). For some of these genes, DCM-associated variants appeared to be clustered in key protein functional domains. Multiple rare variants were present in many family probands, however, there was generally only one "driver" mutation that co-segregated with disease.

Conclusions **-** Rare variants in cardiomyopathy genes can be effectively stratified by combining variant-level and gene-level information. Prioritization of genes based on their *a priori* likelihood of disease causation is a key factor in identifying clinically-actionable variants in cardiac genetic testing.

Key Words

Dilated cardiomyopathy, mutation, genetic testing, next-generation sequencing

INTRODUCTION

Determining variant pathogenicity is *the* major challenge in cardiomyopathy genetic testing, spawning debate in clinics worldwide. Sequence data interpretation has been biased historically by a focus on affected patients and a full appreciation of the normal spectrum of variation in cardiomyopathy "disease genes" has been lacking. In recent years, it has become apparent that many rare protein-altering variants reported to be pathogenic are also seen in the general population.[1-4] These disturbing findings raise doubts about the discriminative efficacy of variant annotation pipelines used in literature reports and the reliability of genetics results provided to patients. Given the increasing role of personal genome sequencing in clinical practice, there is a critical need for an improved strategy for identifying the subset of rare variants that are truly disease-causing.

Dilated cardiomyopathy (DCM) is the most common cardiomyopathy and frequently has a genetic etiology. More than 100 genes have been implicated to date. [5-8] Genetic testing panels generally contain sets of putative DCM-associated genes and have expanded over time to include genes with direct and indirect links to other cardiac and skeletal myopathies. This inclusive approach potentially reduces negative screening results but magnifies the problem of rare variant interpretation.

The aim of this study was to investigate parameters that might improve discrimination between rare cardiomyopathy gene variants in patients with DCM and healthy control subjects. Our analysis included evaluation of variant-level and gene-level characteristics, as well as assessment of the total numbers of rare variants in each individual. This type of information about personal burden of rare variants is unable to be determined from studies of single genes or small gene panels in DCM patients, and is generally not available in population databases where single variants in de-identified subjects are listed separately. Our

findings provide a new framework for variant prioritization and highlight the key role of a subset of genes in DCM pathogenesis.

MATERIALS AND METHODS

Study subjects

Participants provided informed written consent, and protocols were approved by the institutional human ethics committees. Patients with familial or sporadic idiopathic DCM (n=532, aged 47 ± 19 years, 69% males) and healthy control subjects (n=527, aged 49 ± 15 years, 38% males) were recruited from Brigham and Women's Hospital, Boston Children's Hospital, Victor Chang Cardiac Research Institute, Royal Brompton & Harefield NHS Foundation Trust, and the London Institute of Medical Sciences, Imperial College, London (Supplementary Methods, Supplementary Figure S1). The absence of cardiovascular disease in control subjects was ascertained by medical history, and was confirmed by cardiac magnetic resonance imaging in 319 individuals (61%). All subjects had self-reported European ancestry and population stratification was evaluated using principal component analysis (Supplementary Methods, Supplementary Table S1).

A DCM replication cohort was comprised of Australian familial DCM probands (n=101, aged 45 ± 16 years, 67% males). Replication control data were obtained for subjects aged >70 years drawn from the Alzheimer's Disease Sequencing Project (n=2971) and the Medical Genome Reference Bank (n=1144). These replication control cohorts were selected specifically to account for the age-related penetrance of DCM and because per person data were available.

Gene sequencing

Genomic DNA libraries were constructed using standard library preparation protocols. Fragments were ligated to adaptors, amplified, purified, then hybridized to custom arrays enriched for coding regions of genes associated with DCM and other inherited cardiac disorders. A 69-gene panel was used by the Boston laboratory (n=203 DCM cases, 208

controls), and a 64-gene panel was used by the UK laboratory (n=320 DCM cases, 319 controls) (Supplementary Figure S1, Supplementary Table S2). Sequencing was performed using the Illumina HiSeq 2000 or SOLiD 5500xl platforms. Selected variants were evaluated in probands and family members using Sanger sequencing.

Sequence data analysis

Sequence data were processed and aligned to the hg19 (GRCh37) human genome reference using Novoalign and the Genome Analysis Toolkit, or Lifescope v2.5.1. Data for the 41 genes that were represented on both the 69-gene and 64-gene panels, and for which variants were detected in cases, were included in this study (Supplementary Methods, Supplementary Table S2). Coding-sequence variants that were truncating (stop gain, splice donor or acceptor site gain or loss, frame-shift indels) or missense, and that met quality metrics for mapping, read depth, and allelic balance were evaluated. Population data for variant minor allele frequency (MAF) were obtained from the Exome Sequencing Project (ESP) and Exome Aggregation Consortium (ExAC) v1 databases. *In silico* pathogenicity predictions for missense variants were made using PolyPhen2, SIFT, PROVEAN, and MetaSVM.[9-11]

Gene group allocation

A PubMed literature search was undertaken to assess available evidence supporting roles in DCM pathogenesis for each of the 41 genes evaluated. A semiquantitative scoring system was devised to rank genes using grades of A (strong) to D (weak) based on genetic, *in vitro* and animal model data (Supplementary Methods).

Statistical analysis

Statistical comparisons between groups were made using Fisher's exact or chi-squared (2 x 2, 2 x 4) tests. To identify the cells that contributed most to 2 x 4 chi-squared tests, standardized residuals were calculated and deviations of absolute value >2 were considered significant. To evaluate variant distribution in annotated domains of group A genes, missense variants identified in DCM cases were pooled with those of Walsh et al.[12] and compared with ExAC missense variants (MAF 0.1 %) within the same protein domains using Fisher's exact test. $P < 0.05$ was used to denote statistical significance.

RESULTS

Rare variants in cardiomyopathy gene are common in cases and controls

Genetic screening was performed in 532 DCM patients and 527 healthy control subjects. Study participants were of European ancestry, and no significant population stratification effects were detected on a principal component analysis (Supplementary Figure S2). Variants that were protein-altering (truncating, missense) and rare (defined here by MAF <0.1% in the European [EA] subgroup in the ESP database) in a set of 41 cardiomyopathy-associated genes were evaluated. Variants that met these criteria were found in 407 (77%) DCM cases and in 348 (66%) control subjects ($P = 0.0002$), with the number of rare variants per person ranging from 0 to 13 (mean 1.63) in DCM cases and from 0 to 8 (mean 1.24) in controls ($P < 0.0001$). To explore strategies for identifying disease-associated variants, we compared 770 different variants found in DCM cases and 589 variants in control subjects (Supplementary Table S3).

TTN variants

There was an excess of truncating *TTN* variants (*TTN*tv) in DCM cases (83 [11%] variants vs controls, 6 [1%] variants; $P < 0.0001$; Supplementary Table S4). *TTN* missense variants comprised ~ 40% of all variants found in cases and controls. Unlike missense variants in other genes (see below), *TTN* missense variants were not differentiated between groups by assessment of novelty or *in silico* functional predictions (Supplementary Table S5), and these were excluded from subsequent variant-level, gene-level, and per person analyses.

Limited discriminative value of variant characteristics

There were relatively few truncating variants in cardiomyopathy genes other than *TTN*, and these were equally seen in DCM cases (25 [3%] variants) and controls (17 [3%] variants; *P*

=0.75). To start to stratify the larger numbers of missense variants found in both groups, we looked at novelty and *in silico* functional predictions.

Variant novelty. The absence of a variant in healthy subjects has been used as a criterion of pathogenicity in many DCM mutation reports.[13-15] However, assessment of novelty varies with the size of the reference cohort, and differences between groups were only found by using the ExAC database (>60,000 subjects), with DCM cases having more novel variants than controls (P <0.0001; Supplementary Table S5). Surprisingly, more than half of the missense variants in our dataset have been identified previously in genetic testing studies and are listed in ClinVar (Supplementary Table S5).

Variant functional predictions. To assess variant function, we evaluated subsets of variants that were (i) uniformly predicted to be deleterious in PolyPhen2, SIFT, and PROVEAN,[9,10] or (ii) predicted deleterious by MetaSVM, which provides an ensemble score derived from 10 separate algorithms.[11] Both of these approaches yielded significant differences between DCM cases and controls, with modestly better results for MetaSVM's ensemble score than for the three-program consensus score (Supplementary Table S5). When novelty and functional predictions were combined, discrimination between groups was increased: DCM cases were twice as likely as controls to have variants that were absent from ExAC *or* deleterious (MetaSVM), and nearly five-fold more likely to have variants that were absent from ExAC *and* deleterious (MetaSVM) (Figure 1A).

Importance of gene-based parameters

Since variant level parameters incompletely differentiated DCM cases from controls, we next sought to determine whether there might be a hierarchy for pathogenicity between genes.

ExAC metrics of expected genetic variation. ExAC provides two gene constraint metrics for expected in the general population.[16] The "probability of being loss-of-function

intolerant" (pLI) scores for our 41 genes ranged from 0 to 1,[16] with only 10 genes assessed as extremely intolerant (pLI >0.9; Table 1). DCM-associated truncating variants were enriched in genes with high pLI scores (11 of 23 [48%] variants vs controls, 1 of 17 [6%] variants, P = 0.005), but >50% variants were in genes with intermediate or low pLI scores. Twenty-six genes had expected intolerance to missense variants, indicated by positive Z-scores (Table 1).[16] There was a statistically-significant but relatively modest excess of missense variants in these genes in DCM cases (182 [52%] variants vs controls, 129 [43%] variants; P = 0.018).

Residual Variation Intolerance Score (RVIS). RVIS is another metric of expected genetic variation.[17] In this scoring system, genes that cause Mendelian disorders should have little population variability (25th percentile RVIS), while genes associated with complex disorders are likely to be highly variable (100th percentile RVIS). In keeping with the Mendelian model, we found that 18 of our genes were in the 25th percentile bin (Table 1). However, while DCM-associated truncating variants were preferentially distributed in 25th percentile genes (P = 0.01, 2 x 4 chi-squared test; Supplementary Table S6), this was not the case for missense variants (P = 0.67).

Variant burden per gene. There were only three genes, *RBM20*, *MYH7*, and *LMNA*, in which there were significantly more rare variants in DCM cases than in controls (Supplementary Table S3). These findings strongly suggest that these genes harbor pathogenic rare variants, but do not directly inform the interpretation of any single variant. For most genes, the relatively small numbers of variants limited statistical comparisons.

DCM gene ranking. We used a semiquantitative method to rank genes according to literature evidence for roles in DCM pathogenesis (Table 1, Supplementary Table S7). Only 14 of 41 genes had robust evidence for disease association and were classified as group A. There was a prominent peak of DCM-associated truncating variants in these genes, with control-associated truncating variants mostly found in gene groups C and D (P = 0.01, 2 x 4

chi-squared test; Figure 1B, Supplementary Table S8). DCM-associated missense variants also showed a peak in group A genes ($P = 0.018$), with marked enrichment of variants that were novel (ExAC) + deleterious (MetaSVM) ($P = 0.001$, Figure 1C). Adding the "group A" gene parameter to these variant parameters increased the odds ratio for DCM from approximately 5-fold to nearly 9-fold (Figure 1A).

A similar pattern was seen for the subset of ClinVar-listed truncating and missense variants, with a majority of pathogenic/likely-pathogenic variants occurring in DCM cases and in group A genes (P=0.0007, Supplementary Figure S3A). Variants annotated specifically as pathogenic/likely-pathogenic for DCM were exclusive to group A genes, comprising 17 of the 94 (18%) group A variants in DCM cases and 1 of 65 (1.5%) group A variants in controls. Of the 30 total pathogenic/likely pathogenic group A variants in DCM cases, 24 (80%) were also identified as "damaging" by our criteria (Supplementary Figure S3B). The 6 ClinVar variants that were not captured by our method included 5 very rare variants (1 to 5 alleles in ExAC), 3 of which were annotated pathogenic/likely-pathogenic for DCM, and one variant (40 alleles in ExAC) that was likely-pathogenic for HCM.

Yield of variants per person

We next looked at the yield per person of prioritized variants (Table 2). For this analysis, truncating variants and "novel (ExAC) + deleterious (MetaSVM)" missense variants are referred to as "damaging" group A gene variants (Supplementary Table S9). Damaging variants were present in 65 (12.2%) DCM cases and in 8 (1.5%) control subjects (P <0.0001). Similar patterns for the prevalence of damaging group A variants were seen in an independent cohort of familial DCM cases (17 probands [16.8%]), and in two replication control cohorts from the Alzheimer's Disease Sequencing project (54 subjects [1.8%]) and the Medical Genome Reference Bank (33 subjects [2.9%]).

When *TTN*tv were also considered, 151 (28.4%) of our DCM cases were positive for damaging variants and/or *TTN*tv compared to 14 (2.7%) controls (P <0.0001). Damaging variants and/or *TTN*tv were present in 32 (31.7%) probands in the familial DCM replication cohort, and in 61 (2.1%) and 46 (4.0%) subjects, respectively in the two control replication cohorts. It has been questioned whether *TTN*tv are sufficient alone to cause DCM or require "second hit" genetic and/or acquired factors for DCM manifestation. We found no differences in the background burden of variants in *TTN*tv+ DCM cases (1.58 ± 1.74 variants, range 0-9) when compared with *TTN*tv-/damaging+ DCM cases (1.17 ± 1.22 variants, range 0-5), *TTN*tv-/damaging- DCM cases (1.30 ± 1.55 variants, range 0-13), or *TTN*tv-/damaging- control subjects (1.20 ± 1.29, range 0-8; P = 0.3819, Kruskal-Wallis test). These data suggest that *TTN*tv carriers generally do not show an excess of second damaging variants.

Family segregation

To further test our criteria for variant prioritization, we performed co-segregation analysis in 28 DCM families in which DNA samples from ≥3 informative individuals were available. All rare variants that were identified in each family proband were evaluated in the respective family members.

Truncating group A gene variants. Group A gene truncating variants were evaluated in 3 families (Supplementary Figure S4). In Family BY, the proband carried a *BAG3* stop codon with segregation analysis consistent with linkage (odds ratio 1/18, LOD score 1.3). The truncated protein would lack the signature BAG domain that binds to HSP70 and is required for chaperone activity.[18] Numerous *BAG3* truncating mutations have been reported in DCM patients, and reduced levels of BAG3 protein have been seen in ventricular tissues from patients with heart failure.[19-21]

The Family DF proband and her affected brother carried a *MYH7* splice acceptor site variant. This variant was absent from 3 unaffected individuals aged >40 years, but the family size was insufficient to show statistically-significant linkage (odds ratio 1/6, LOD score 0.8). *MYH7* truncating variants were found in 2 additional sporadic DCM cases in our cohort (Supplementary Table S9) but have rarely been reported in DCM cases,[12,22] and there is uncertainty about their clinical significance.

The Family C proband carried a *SCN5A* stop codon that would truncate the C-terminus with loss of a PY-motif (binding site for Nedd4 ubiquitin ligase) and a PDZ-domain binding motif (interacts with the cytoskeletal adapter protein, syntrophin), potentially giving rise to protein degradation or trafficking defects.[23,24] The cardiac phenotype associated with *SCN5A* mutations often includes arrhythmias and conduction-system abnormalities as well as DCM,[25,26] and atrial fibrillation was present in all affected individuals in this kindred. The apparent non-segregation of this variant with DCM (LOD score -2.5) may be confounded by two possible phenocopies who had other plausible acquired causes of disease.

Missense variants. Segregation of 30 missense variants was evaluated in 14 kindreds (Supplementary Table S10). Only 4 of the 30 variants had LOD scores >1 (suggestive of linkage relative to family size) and all of these were damaging group A gene variants (Supplementary Figure S4). Affected individuals in Family CZ carried a p.Arg369Gln *MYH7* variant and had both DCM and left ventricular non-compaction. This variant, in the myosin motor domain, has been associated with both phenotypes.[6,27] The p.Arg634Trp *RBM20* variant, present in all affected individuals in Family AB, lies within the arginine-serine-rich (RS) domain which is a putative DCM mutation hotspot.[28,29] There were also two co-segregating *DES* variants, p.Leu398Pro (Family FK) and p.Lys449Thr (Family FG), the latter associated with myofibrillar myopathy.[30] Four additional kindreds (Families BG, BK, FR, KS) had damaging group A gene variants with LOD scores <1 (Supplementary Figure S5).

The damaging variants were present in 15 of 16 affected individuals in these families, with genotype-phenotype discordance driven mainly by non-penetrant variant carriers. In all families tested, there was no evidence for co-segregation of any variants that did not meet the criteria for being "damaging" or that were in genes other than group A.

Multiple variants. With expanded genetic testing panels, it is not uncommon to find several rare variants in DCM cases, prompting hypotheses of multiple mutations.[7,31] In an extension of the current analysis, we reviewed data for individuals who had undergone testing with the 69-gene testing panel and looked at all 69 genes as well as increasing the threshold level of MAF to <1%. Family co-segregation analysis was then performed in 5 kindreds in which the proband was found to carry ≥5 variants. In 4 of these families (Families BY, FK, AB, BA), only one of the proband's multiple variants segregated with DCM (Supplementary Figure S6). In the remaining family (Family GX), there were no co-segregating variants. Four of the proband's 5 rare variants were inherited from her unaffected father, and none were damaging group A variants. These studies demonstrate the value of family analysis and support the expectation that a single "driver" variant will be present when DCM appears as a Mendelian trait.

Location of variants in group A genes

To explore potential effects of variant location within group A genes, we compared the distribution of missense variants identified in DCM cases (derived from our own cohorts and two clinical laboratories[12]) with rare (MAF <0.1%) missense variants in ExAC (Figure 2, Supplementary Table S11). In *MYH7*, DCM variants occurred in all protein domains but were preferentially located in the myosin motor (Figure 2A). DCM-associated variants in *RBM20* were significantly enriched in the RS domain that is associated with altered titin splicing and myocardial compliance (Figure 2B).[28,29] More than half of the damaging *RBM20* variants

resided in the RS domain or in a glutamate-rich region that is also associated with titin splicing defects.[32] DCM-associated *LMNA* variants were mostly located in the coiled-coil rod domain, particularly in coil 2, which is critical for dimer formation (Figure 2C),[33] and there was a cluster of DCM-associated variants in the S4 voltage-sensor, repeat I, of *SCN5A* (Figure 2E). Variants in the S4 transmembrane segments cause arrhythmic forms of DCM and have been associated with gain-of-function effects and gating pore currents[25,26,34-36] Clustering of DCM variants was also apparent in the α-tropomyosin binding domain of *TNNT2,* and several damaging variants in *TPMI* resided in the cardiac troponin T binding domain (Figure 2G and 2H).

DISCUSSION

Our data show that rare protein-altering cardiomyopathy gene variants are not unique to DCM patients but also occur commonly in apparently healthy individuals. Because of this, effective discriminitive strategies are clearly needed in order to derive meaningful results from genetic testing. Here we provide a new approach for assessment of rare variants that combines both variant-level and gene-level information. We propose that ranking genes based on their *a priori* likelihood of disease causation is a key factor in identifying clinically-actionable variants.

Only 14 of 41 genes achieved "group A" status. Our literature-based gene grading method was subsequently validated by the predilection for damaging variants in DCM cases to occur in group A genes and by the similar clustering of pathogenic/likely-pathogenic ClinVar variants. Gene group is a dynamic parameter, and is highly dependent on the contemporary knowledge base. Assessment of genetic evidence is intrinsically biased by the size of the kindreds evaluated, the frequency in which genes have been screened, and published reports, and definitive conclusions cannot be drawn if there is insufficient information. Genes in which variants co-segregated with DCM in multiple large kindreds ($\geq$5 affected) scored highly but for many genes, only small families or single cases have been studied, raising the possibility that variants, even if function-altering, might co-segregate by chance or be incidental findings. Similarly, *in vitro* functional data are helpful if these are available but are often lacking from the literature. Group A genes typically had animal data showing spontaneous development of DCM in models expressing heterozygous loss-of-function alleles or human missense mutations. This level of evidence was missing from many studies in which only homozygous loss-of-function animals have been studied. While such models can implicate genes in normal cardiac function, it cannot be assumed that there is a direct correlation with gene "dose" and that heterozygous counterparts will have a similar,

albeit less severe, phenotype. In homozygous loss-of-function animals, there may also be profound effects on cardiac development that independently predispose to contractile impairment.

The group A genes *MYH7*, *RBM20*, and *LMNA*, had significantly more rare variants in DCM cases than controls, and frequently harbored damaging variants. *MYH7* and *LMNA* have consistently appeared on lists of the "most frequently mutated" genes in DCM genetics studies, and were the top 2 genes (after *TTN*) showing an excess of rare variants (MAF <0.0001) in a recent analysis of 1315 DCM patients referred to diagnostic genetic testing laboratories.[6,7,12,22] *RBM20* is a more recent addition to the disease gene list and has been less frequently screened. Despite their overall enrichment in DCM cases, not all rare variants in *MYH7*, *RBM20*, and *LMNA* are necessarily deleterious and additional information is required for clinical interpretation of any specific variant. Interestingly, variants in these genes had a non-random distribution in DCM cases: *MYH7* variants clustered in the myosin motor, *RBM20* variants in the RS domain, and *LMNA* variants in the coiled-coil rod. As the numbers of reported DCM-associated variants expands, additional patterns of variant distribution patterns may become apparent.

Our method for stratifying rare variants provides a framework for future studies and further refinements are anticipated. We expect that over time, more genes will achieve group A status, and ongoing rigorous gene ranking is warranted. For missense variants, the criterion of absence from ExAC may underestimate the yield of deleterious variants and varying MAF threshold levels need to be evaluated. It is notable, for example, that the few ClinVar-annotated pathogenic/likely-pathogenic group A gene variants that were not captured by our method were mostly extremely rare in ExAC. Many of the Clinvar-listed variants are annotated as pathogenic/likely-pathogenic for disorders other than DCM. This phenotypic heterogeneity suggests that at least some of these variants might be modifiers rather than

directly causative of disease. Another explanation is that the ultimate clinical manifestations of any given variant may be flavored by the context of each person's genetic and lifestyle factors. The detection of pathogenic/likely-pathogenic variants for diverse disorders suggests that our prioritization method could be applied to a range of genetic testing indications, with the caveat that disease-specific gene lists would be needed.

The lack of cardiac investigations in ~40% control subjects is a limitation of this study. We focussed on subjects of European ancestry to avoid confounding effects of ethnicity-related background genetic variation, and the applicability of our findings to other populations has yet to be determined. In assessing individual burden of genetic variation, only rare cardiomyopathy gene variants were considered. The extent to which rare, low-frequency and common variants in a broad range of cardiac genes might cumulatively contribute to a myopathic substrate is unknown.

Ongoing efforts by the American College of Medical Genetics to standardize variant annotation methods are an important step towards establishing international guidelines for genetic testing in DCM and other inherited human disorders.[37] The cost-efficacy of first-line screening of a core set of key DCM disease genes (e.g. group A genes) vs extended panels or genome-wide testing warrants further analysis. As genome sequencing is poised to become part of mainstream healthcare, there is a pressing need for curation of clinical and genetic information in databases such as ClinGen,[38] functional evaluation of variants, and clinical trials in genotyped patients. Advancements in these areas should expedite implementation of personalized medicine.

In summary, our data suggest that interpretation of genetic testing results cannot be made on the basis of variant characteristics alone, and that the effect size of protein-altering variants is primarily determined by the biological importance of the gene involved and relevance to DCM pathogenesis.

ACKNOWLEDGMENTS

We thank Erica Mazaika, Monique Ohanian, Santiago Pineda, Celine Santiago. Magdalena Soka, and Gunjan Trivedi for assistance with data collection, collaborating physicians who referred patients to the study and participating families. See Supplementary Acknowledgments for the Medical Genome Research Bank.

REFERENCES

1. Pan S, Caleshu CA, Dunn KE, et al. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. *Circ Cardiovasc Genet* 2012;5:602-610.
2. Norton N, Robertson PD, Rieder MK, et al. Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. *Circ Cardiovasc Genet* 2012;5:167-174.
3. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet* 2012;5:391-399.
4. Haggerty CM, James CA, Calkins H, et al. Electronic health record phenotype in subjects with genetic variants associated with arrhythmogenic right ventricular cardiomyopathy: a study of 30,716 subjects with exome sequencing. *Genet Med* 2017;19:1245-1252.
5. Fatkin D, Seidman CE, Seidman JG. Genetics and disease of ventricular muscle. *Cold Spring Harb Perspect Med* 2014;4;a021063.
6. Pugh TJ, Kelly MA, Gowrisankar S, et al. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med* 2014:16:601-608.
7. Haas J, Frese KS, Peil B, et al. Atlas of the clinical genetics of human dilated cardiomyopathy. *Eur Heart J* 2015;36:1123-1135.
8. McNally EM, Mestroni L. Dilated cardiomyopathy: genetic determinants and mechanisms. *Circ Res* 2017;121:731-748.
9. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-249.
10. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acid Res* 2012;40:W452-W457.

11. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 2015;24:2125-2137.

12. Walsh R, Thomson KL, Ware JS, et al. Exome Aggregation Consortium, MacArthur DG, Farrall M, Cook SA, Watkins H. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet Med* 2017;19:192-203.

13. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* 1998;280:750-752.

14. Fatkin D, MacRae C, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999;341:1715-1724.

15. Kamisago M, Sharma SD, dePalma SR, et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* 2000;343:1688-1696.

16. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;285-291.

17. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 2013;9:e1003709.

18. Hishiya A, Kitazawa T, Takayama S. BAG3 and Hsc70 interact with actin capping protein CapZ to maintain myofibrillar integrity under mechanical stress. *Circ Res* 2010;107:1220-1231.

19. Villard E, Perret C, Gary F, et al. A genome-wide association study identified two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J* 2011;32:1065-1076.

20. Norton N, Li D, Rieder MJ, et al. Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. *Am J Hum Genet* 2011;88:273-282.

21. Feldman AM, Begay RL, Knezevic T, et al. Decreased levels of BAG3 in a family with a rare variant and in idiopathic dilated cardiomyopathy. *J Cell Physiol* 2014;229:1697-1702.

22. Kayvanpour E, Sedaghat-Hamedani F, Amr A, et al. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin Res Cardiol* 2017;106:127-139.

23. Van Bemmelen MX, Rougier JS, Gavillet B, et al. Cardiac voltage-gated sodium channel Nav1.5 is regulated by Nedd4-2 mediated ubiquitination. *Circ Res* 2004;95:284-291.

24. Gavillet B, Rougier JS, Domenighetti AA, et al. Cardiac sodium channel Nav1.5 is regulated by a multiprotein complex composed of syntrophins and dystrophin. *Circ Res* 2006;99:407-414.

25. McNair WP, Ku L, Taylor MR, et al. *SCN5A* mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* 2004;110:2163-2167.

26. Olson TM, Michels VV, Ballew JD, et al. Sodium channel mutations and susceptibility to heart failure and atrial fibrillation. *JAMA* 2005;293:447-454.

27. Dellefave LM, Pytel P, Mewborn S, et al. Sarcomere mutations in cardiomyopathy with left ventricular hypertrabeculation. *Circ Cardiovasc Genet* 2009;2:442-449.

28. Brauch KM, Karst ML, Herron KJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol* 2009;54:930-941.

29. Guo W, Schafer S, Greaser ML, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med* 2012;18:766-773.

30. Selcen D, Ohno K, Engel AG. Myofibrillar myopathy: clinical, morphological and genetic studies in 63 patients. *Brain* 2004;127:439-451.

31. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*. 2013;10:531-547.

32. Beqqali A, Bollen IA, Rasmussen TB, et al. A mutation in the glutamate-rich region of RNA-binding motif protein 20 causes dilated cardiomyopathy through missplicing of titin and impaired Frank-Starling mechanism. *Cardiovasc Res* 2016;112:452-463.

33. Gruenbaum Y, Foisner R. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. Annu Rev Biochem 2015;84:131-164.

34. McNair WP, Sinagra G, Taylor MR, et al. *SCN5A* mutations associate with arrhythmic dilated cardiomyopathy and commonly localize to the voltage-sensing mechanism. *J Am Coll Cardiol* 2011;57:2160-8.

35. Mann SA, Castro ML, Ohanian M, et al. R222Q *SCN5A* mutation is associated with reversible ventricular ectopy and dilated cardiomyopathy. *J Am Coll Cardiol* 2012;60:1566-1573.

36. Moreau A, Gosselin-Badaroudine P, Boutjdir M, Chahine M. Mutations in the voltage sensors of domain I and II of Nav1.5 that are associated with arrhythmias and dilated cardiomyopathy generate gating pore currents. *Front Pharmacol* 2015;6:301.

37. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-424.

38. Rehm HL, Berg JS, Brooks LD, et al. ClinGen – the clinical genome resource. *N Engl J Med* 2015;372:2235-2242.

Figure legends

Figure 1. Criteria for variant prioritization. (**A**) Missense rare variants in cardiomyopathy genes in DCM cases and controls were compared using a number of criteria including novelty with respect to the European subgroup of the Exome Sequencing Project (EA-ESP) and non-Finnish European (NFE) or all subjects in the Exome Aggregation Consortium (ExAC), *in silico* functional predictions (PolyPhen2, SIFT, PROVEAN, MetaSVM), RVIS score and gene group. The odds ratios (OR) for DCM are displayed in a forest plot. (**B**) Distribution of truncating variants across gene groups in controls (white bars) and DCM cases (blue bars). (**C**) Distribution of missense variants across gene groups in controls (white bars) and DCM cases (blue bars). The subset of "novel (ExAC) + deleterious (MetaSVM)" missense variants are denoted in the solid sections of bars. There were significant effects of clinical status and gene group for truncating variants (**B,** $P = 0.01$; 2 x 4 chi-squared test) and missense variants (**C;** $P = 0.018$); see Supplementary Table S8 for statistical analysis. For all panels, *TTN* truncating and missense variants were excluded.

Figure 2. Rare variant distribution in DCM cases and control subjects. The location of variants in different protein domains of group A genes is shown in panels A to L. In each panel, variants identified in DCM patients in the primary cohort (top row) and in the familial DCM (FDCM) replication cohort (second row) are shown: all variants (pink, "DCM all"), damaging variants (red, "DCM HP"). Variant types are indicated by: circle =missense, square = splice site change, diamond = frameshift, star = stop codon. The third row shows DCM variants identified by two clinical diagnostic laboratories ("DCM Walsh")[12]. Missense variants identified in control subjects are shown below the protein schematic: control subjects in this study (Controls), Exome Aggregation Consortium (ExAC) database. See Supplementary Table S11 for protein domain coordinates and statistical analysis.

Revised text (tracked)

A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy

Short title: Genetic Testing in Dilated Cardiomyopathy

Claire Horvat, PhD[1], Renee Johnson, PhD[1], Lien Lam, PhD[2,8], Jacob Munro, BSc, MIT[3], Francesco Mazzarotto, PhD[4], Angharad M. Roberts, MRCP[4], Daniel S. Herman, MD, PhD[2], Michael Parfenov, MD, PhD[2], Alireza Haghighi MD, PhD[2,8,9], Barbara McDonough, RN[2], Steven R. DePalma, PhD[2], Anne M. Keogh, MD,[5,6,7] Peter S. Macdonald, MB BS, PhD[5,6,7], Christopher S. Hayward, MD[5,6,7], Amy Roberts, MD[10], Paul J.R. Barton, PhD[4], Leanne E. Felkin, PhD[4], Eleni Giannoulatou, DPhil[3], Stuart A. Cook, PhD, MRCP[4,11], J.G. Seidman, PhD[2,8], Christine E. Seidman, MD[2,9], Diane Fatkin, MD[1,6,7].

[1] Molecular Cardiology Division, [3] Molecular, Structural and Computational Biology Division and [5] Cardiac Physiology and Transplantation Division, Victor Chang Cardiac Research Institute, Sydney, New South Wales, Australia.

[2] Department of Genetics, Harvard Medical School, Boston MA, USA.

[4] NIHR Cardiovascular Biomedical Research Unit, Royal Brompton & Harefield NHS Foundation Trust, and Imperial College London, London, England.

[6] Cardiology Department, St Vincent's Hospital, Sydney, New South Wales, Australia.

[7] Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia.

[8] Howard Hughes Medical Institute, Boston MA, USA.

[9] Cardiovascular Division, Brigham and Women's Hospital, Boston MA, USA.

[10] Boston Children's Hospital, Boston MA, USA.

[11] Duke-National University of Singapore, and National Heart Center, Singapore

Correspondence:

Diane Fatkin, Victor Chang Cardiac Research Institute, 405 Liverpool St, Darlinghurst, New South Wales 2010, Australia.

Ph: +61 2 9295 8618, Fax: +61 2 9295 8770

Email: d.fatkin@victorchang.edu.au.

DISCLOSURE

The authors declare no conflict of interest.

FUNDING SOURCES

This study was supported by grants from the National Health and Medical Research Council of Australia (D.F., P.S.M.), Harvard Club of Australia Foundation (D.F.), Estate of the Late RT Hall (D.F.), Simon Lee Foundation (D.F.), British Heart Foundation (S.A.C.), National Institute for Health Research Cardiovascular BRU at the Brompton and Harefield & Harefield NHS Foundation Trust and Imperial College London (S.A.C., P.J.B.), Howard Hughes Medical Institute (C.E.S.), National Institutes of Health (D.S.H., J.G.S.), and the Leducq Foundation (S.A.C., J.G.S., C.E.S.).

Abstract

Purpose: We evaluated strategies for identifying disease-causing variants in genetic testing for dilated cardiomyopathy (DCM).

Methods: Cardiomyopathy gene panel testing was performed in 532 DCM patients and 527 healthy control subjects. Rare variants in 41 genes were stratified using variant-level and gene-level characteristics.

Results: A majority of DCM cases and controls carried rare protein-altering cardiomyopathy gene variants. Variant-level characteristics alone had limited discriminative value. Differentiation between groups was substantially improved by addition of gene-level information that incorporated ranking of genes based on literature evidence for disease association. The odds of DCM were increased to nearly 9-fold for truncating variants or high-impact missense variants in the subset of 14 genes that had the strongest biological links to DCM (P <0.0001). For some of these genes, DCM-associated variants appeared to be clustered in key protein functional domains. Multiple rare variants were present in many family probands, however, there was generally only one "driver" mutation that co-segregated with disease.

Conclusions **-** Rare variants in cardiomyopathy genes can be effectively stratified by combining variant-level and gene-level information. Prioritization of genes based on their *a priori* likelihood of disease causation is a key factor in identifying clinically-actionable variants in cardiac genetic testing.

Key Words

Dilated cardiomyopathy, mutation, genetic testing, next-generation sequencing

INTRODUCTION

Determining variant pathogenicity is *the* major challenge in cardiomyopathy genetic testing, spawning debate in clinics worldwide. Sequence data interpretation has been biased historically by a focus on affected patients and a full appreciation of the normal spectrum of variation in cardiomyopathy "disease genes" has been lacking. In recent years, it has become apparent that many rare protein-altering variants reported to be pathogenic are also seen in the general population.[1-4] These disturbing findings raise doubts about the discriminative efficacy of variant annotation pipelines used in literature reports and the reliability of genetics results provided to patients. Given the increasing role of personal genome sequencing in clinical practice, there is a critical need for an improved strategy for identifying the subset of rare variants that are truly disease-causing.

Dilated cardiomyopathy (DCM) is the most common cardiomyopathy and frequently has a genetic etiology. More than 100 genes have been implicated to date.[5-8] Genetic testing panels generally contain sets of putative DCM-associated genes and have expanded over time to include genes with direct and indirect links to other cardiac and skeletal myopathies. This inclusive approach potentially reduces negative screening results but magnifies the problem of rare variant interpretation.

The aim of this study was to investigate parameters that might improve discrimination between rare cardiomyopathy gene variants in patients with DCM and healthy control subjects. Our analysis included evaluation of variant-level and gene-level characteristics, as well as assessment of the total numbers of rare variants in each individual. This type of information about personal burden of rare variants is unable to be determined from studies of single genes or small gene panels in DCM patients, and is generally not available in population databases where single variants in de-identified subjects are listed separately. Our

findings provide a new framework for variant prioritization and highlight the key role of a subset of genes in DCM pathogenesis.

MATERIALS AND METHODS

Study subjects

Participants provided informed written consent, and protocols were approved by the institutional human ethics committees. Patients with familial or sporadic idiopathic DCM (n=532, aged 47 ± 19 years, 69% males) and healthy control subjects ~~without cardiovascular disease~~ (n=527, aged 49 ± 15 years, 38% males) were recruited from Brigham and Women's Hospital, Boston Children's Hospital, Victor Chang Cardiac Research Institute, Royal Brompton & Harefield NHS Foundation Trust, and the London Institute of Medical Sciences, Imperial College, London (Supplementary Methods, Supplementary Figure S1). The absence of cardiovascular disease in control subjects was ascertained by medical history, and was confirmed by cardiac magnetic resonance imaging in 319 individuals (61%). All subjects had self-reported European ancestry and population stratification was evaluated using principal component analysis (Supplementary Methods, Supplementary Table S1).

A DCM replication cohort was comprised of Australian familial DCM probands (n=101, aged 45 ± 16 years, 67% males). Replication control data were obtained for subjects aged >70 years drawn from the Alzheimer's Disease Sequencing Project (n=2971) and the Medical Genome Reference Bank (n=1144). These replication control cohorts were selected specifically to account for the age-related penetrance of DCM and because per person data were available.

Gene sequencing

Genomic DNA libraries were constructed using standard library preparation protocols. Fragments were ligated to adaptors, amplified, purified, then hybridized to custom arrays enriched for coding regions of genes associated with DCM and other inherited cardiac disorders. A 69-gene panel was used by the Boston laboratory (n=203 DCM cases, 208

Formatted: Justified

controls), and a 64-gene panel was used by the UK laboratory (n=320 DCM cases, 319 controls) (Supplementary Figure S1, Supplementary Table S2). Sequencing was performed using the Illumina HiSeq 2000 or SOLiD 5500xl platforms. Selected variants were evaluated in probands and family members using Sanger sequencing.

Sequence data analysis

Sequence data were processed and aligned to the hg19 (GRCh37) human genome reference using Novoalign and the Genome Analysis Toolkit, or Lifescope v2.5.1. Data for the 41 genes that were represented on both the 69-gene and 64-gene panels, and for which variants were detected in cases, were included in this study (Supplementary Methods, Supplementary Table S2). Coding-sequence variants that were truncating (stop gain, splice donor or acceptor site gain or loss, frame-shift indels) or missense, and that met quality metrics for mapping, read depth, and allelic balance were evaluated. Population data for variant minor allele frequency (MAF) were obtained from the Exome Sequencing Project (ESP) and Exome Aggregation Consortium (ExAC) v1 databases. *In silico* pathogenicity predictions for missense variants were made using PolyPhen2, SIFT, PROVEAN, and MetaSVM.[9-11]

Gene group allocation

A PubMed literature search was undertaken to assess available evidence supporting roles in DCM pathogenesis for each of the 41 genes evaluated. A semiquantitative scoring system was devised to rank genes using grades of A (strong) to D (weak) based on genetic, *in vitro* and animal model data (Supplementary Methods).

Statistical analysis

Statistical comparisons between groups were made using Fisher's exact or chi-squared (2 x 2, 2 x 4) tests. To identify the cells that contributed most to 2 x 4 chi-squared tests, standardized residuals were calculated and deviations of absolute value >2 were considered significant. To evaluate variant distribution in annotated domains of group A genes, missense variants identified in DCM cases were pooled with those of Walsh et al.[12] and compared with ExAC missense variants (MAF 0.1 %) within the same protein domains using Fisher's exact test. P <0.05 was used to denote statistical significance.

RESULTS

Rare variants in cardiomyopathy gene are common in cases and controls

Genetic screening was performed in 532 DCM patients and 527 healthy control subjects. Study participants were of European ancestry, and no significant population stratification effects were detected on a principal component analysis (Supplementary Figure S2). Variants that were protein-altering (truncating, missense) and rare (defined here by MAF <0.1% in the European [EA] subgroup in the ESP database) in a set of 41 cardiomyopathy-associated genes were evaluated. Variants that met these criteria were found in 407 (77%) DCM cases and in 348 (66%) control subjects ($P = 0.0002$), with the number of rare variants per person ranging from 0 to 13 (mean 1.63) in DCM cases and from 0 to 8 (mean 1.24) in controls ($P < 0.0001$). To explore strategies for identifying disease-associated variants, we compared 770 different variants found in DCM cases and 589 variants in control subjects (Supplementary Table S3).

TTN variants

There was an excess of truncating *TTN* variants (*TTN*tv) in DCM cases (83 [11%] variants vs controls, 6 [1%] variants; $P < 0.0001$; Supplementary Table S4). *TTN* missense variants comprised ~ 40% of all variants found in cases and controls. Unlike missense variants in other genes (see below), *TTN* missense variants were not differentiated between groups by assessment of novelty or *in silico* functional predictions (Supplementary Table S5), and these were excluded from subsequent variant-level, gene-level, and per person analyses.

Limited discriminative value of variant characteristics

There were relatively few truncating variants in cardiomyopathy genes other than *TTN*, and these were equally seen in DCM cases (25 [3%] variants) and controls (17 [3%] variants; *P*

=0.75). To start to stratify the larger numbers of missense variants found in both groups, we looked at novelty and *in silico* functional predictions.

Variant novelty. The absence of a variant in healthy subjects has been used as a criterion of pathogenicity in many DCM mutation reports.[13-15] However, assessment of novelty varies with the size of the reference cohort, and differences between groups were only found by using the ExAC database (>60,000 subjects), with DCM cases having more novel variants than controls (P <0.0001; Supplementary Table S5). Surprisingly, more than half of the missense variants in our dataset have been identified previously in genetic testing studies and are listed in ClinVar (Supplementary Table S5).

Variant functional predictions. To assess variant function, we evaluated subsets of variants that were (i) uniformly predicted to be deleterious in PolyPhen2, SIFT, and PROVEAN,[9,10] or (ii) predicted deleterious by MetaSVM, which provides an ensemble score derived from 10 separate algorithms.[11] Both of these approaches yielded significant differences between DCM cases and controls, with modestly better results for MetaSVM's ensemble score than for the three-program consensus score (Supplementary Table S5). When novelty and functional predictions were combined, discrimination between groups was increased: DCM cases were twice as likely as controls to have variants that were absent from ExAC *or* deleterious (MetaSVM), and nearly five-fold more likely to have variants that were absent from ExAC *and* deleterious (MetaSVM) (Figure 1A).

Importance of gene-based parameters

Since variant level parameters incompletely differentiated DCM cases from controls, we next sought to determine whether there might be a hierarchy for pathogenicity between genes.

ExAC metrics of expected genetic variation. ExAC provides two gene constraint metrics for expected in the general population.[16] The "probability of being loss-of-function

intolerant" (pLI) scores for our 41 genes ranged from 0 to 1,[16] with only 10 genes assessed as extremely intolerant (pLI >0.9; Table 1). DCM-associated truncating variants were enriched in genes with high pLI scores (11 of 23 [48%] variants vs controls, 1 of 17 [6%] variants, *P* = 0.005), but >50% variants were in genes with intermediate or low pLI scores. Twenty-six genes had ~~had~~ expected intolerance to missense variants, indicated by positive Z-scores (Table 1).[16] There was a statistically-significant but relatively modest excess of missense variants in these genes in DCM cases (182 [52%] variants vs controls, 129 [43%] variants; *P* = 0.018).

Residual Variation Intolerance Score (RVIS). RVIS is another metric of expected genetic variation.[17] In this scoring system, genes that cause Mendelian disorders should have little population variability (25th percentile RVIS), while genes associated with complex disorders are likely to be highly variable (100th percentile RVIS). In keeping with the Mendelian model, we found that 18 of our genes were in the 25th percentile bin (Table 1). However, while DCM-associated truncating variants were preferentially distributed in 25th percentile genes (*P* = 0.01, 2 x 4 chi-squared test; Supplementary Table S6), this was not the case for missense variants (*P* = 0.67).

Variant burden per gene. There were only three genes, *RBM20*, *MYH7*, and *LMNA*, in which there were significantly more rare variants in DCM cases than in controls (Supplementary Table S3). These findings strongly suggest that these genes harbor pathogenic rare variants, but do not directly inform the interpretation of any single variant. For most genes, the relatively small numbers of variants limited statistical comparisons.

DCM gene ranking. We used a semiquantitative method to rank genes according to literature evidence for roles in DCM pathogenesis (Table 1, Supplementary Table S7). Only 14 of 41 genes had robust evidence for disease association and were classified as group A. There was a prominent peak of DCM-associated truncating variants in these genes, with

control-associated truncating variants mostly found in gene groups C and D (P = 0.01, 2 x 4 chi-squared test; Figure 1B, Supplementary Table S8). DCM-associated missense variants also showed a peak in group A genes (P = 0.018), with marked enrichment of variants that were novel (ExAC) + deleterious (MetaSVM) (P = 0.001, Figure 1C). Adding the "group A" gene parameter to these variant parameters increased the odds ratio for DCM from approximately 5-fold to nearly 9-fold (Figure 1A).

A similar pattern was seen for the subset of ClinVar-listed truncating and missense variants, with a majority of pathogenic/likely-pathogenic variants occurring in DCM cases and in group A genes (P=0.0007, Supplementary Figure S3A). Variants annotated specifically as pathogenic/likely-pathogenic for DCM were exclusive to group A genes, comprising 17 of the 94 (18%) group A variants in DCM cases and 1 of 65 (1.5%) group A variants in controls~~.~~. Of the 30 total pathogenic/likely pathogenic group A variants in DCM cases, 24 (80%) were also identified as "damaging" by our criteria (Supplementary Figure S3B). The 6 ClinVar variants that were not captured by our method included 5 very rare variants (1 to 5 alleles in ExAC), 3 of which were annotated pathogenic/likely-pathogenic for DCM, and one variant (40 alleles in ExAC) that was likely-pathogenic for HCM.

Yield of variants per person

We next looked at the yield per person of prioritized variants (Table 2). For this analysis, truncating variants and "novel (ExAC) + deleterious (MetaSVM)" missense variants are referred to as "damaging" group A gene variants (Supplementary Table S9). Damaging variants were present in 65 (12.2%) DCM cases and in 8 (1.5%) control subjects (P <0.0001). Similar patterns for the prevalence of damaging group A variants were seen in an independent cohort of familial DCM cases (17 probands [16.8%]), and in two replication

control cohorts from the Alzheimer's Disease Sequencing project (54 subjects [1.8%]) and the Medical Genome Reference Bank (33 subjects [2.9%]).

When *TTN*tv were also considered, 151 (28.4%) of our DCM cases were positive for damaging variants and/or *TTN*tv compared to 14 (2.7%) controls (P <0.0001). Damaging variants and/or *TTN*tv were present in 32 (31.7%) probands in the familial DCM replication cohort, and in 61 (2.1%) and 46 (4.0%) subjects, respectively in the two control replication cohorts. It has been questioned whether *TTN*tv are sufficient alone to cause DCM or require "second hit" genetic and/or acquired factors for DCM manifestation. We found no differences in the background burden of variants in *TTN*tv+ DCM cases (1.58 ± 1.74 variants, range 0-9) when compared with *TTN*tv-/damaging+ DCM cases (1.17 ± 1.22 variants, range 0-5), *TTN*tv-/damaging- DCM cases (1.30 ± 1.55 variants, range 0-13), or *TTN*tv-/damaging- control subjects (1.20 ± 1.29, range 0-8; P = 0.3819, Kruskal-Wallis test). These data suggest that *TTN*tv carriers generally do not show an excess of second damaging variants.

Family segregation

To further test our criteria for variant prioritization, we performed co-segregation analysis in 28 DCM families in which DNA samples from ≥3 informative individuals were available. All rare variants that were identified in each family proband were evaluated in the respective family members.

Truncating group A gene variants. Group A gene truncating variants were evaluated in 3 families (Supplementary Figure S4). In Family BY, the proband carried a *BAG3* stop codon with segregation analysis consistent with linkage (odds ratio 1/18, LOD score 1.3). The truncated protein would lack the signature BAG domain that binds to HSP70 and is required for chaperone activity.[18] Numerous *BAG3* truncating mutations have been reported in DCM

patients, and reduced levels of BAG3 protein have been seen in ventricular tissues from patients with heart failure.[19-21]

The Family DF proband and her affected brother carried a *MYH7* splice acceptor site variant. This variant was absent from 3 unaffected individuals aged >40 years, but the family size was insufficient to show statistically-significant linkage (odds ratio 1/6, LOD score 0.8). *MYH7* truncating variants were found in 2 additional sporadic DCM cases in our cohort (Supplementary Table S9) but have rarely been reported in DCM cases,[12,22] and there is uncertainty about their clinical significance.

The Family C proband carried a *SCN5A* stop codon that would truncate the C-terminus with loss of a PY-motif (binding site for Nedd4 ubiquitin ligase) and a PDZ-domain binding motif (interacts with the cytoskeletal adapter protein, syntrophin), potentially giving rise to protein degradation or trafficking defects.[23,24] The cardiac phenotype associated with *SCN5A* mutations often includes arrhythmias and conduction-system abnormalities as well as DCM,[25,26] and atrial fibrillation was present in all affected individuals in this kindred. The apparent non-segregation of this variant with DCM (LOD score -2.5) may be confounded by two possible phenocopies who had other plausible acquired causes of disease.

Missense variants. Segregation of 30 missense variants was evaluated in 14 kindreds (Supplementary Table S10). Only 4 of the 30 variants had LOD scores >1 (suggestive of linkage relative to family size) and all of these were damaging group A gene variants (Supplementary Figure S4). Affected individuals in Family CZ carried a p.Arg369Gln *MYH7* variant and had both DCM and left ventricular non-compaction. This variant, in the myosin motor domain, has been associated with both phenotypes.[6,27] The p.Arg634Trp *RBM20* variant, present in all affected individuals in Family AB, lies within the arginine-serine-rich (RS) domain which is a putative DCM mutation hotspot.[28,29] There were also two co-segregating *DES* variants, p.Leu398Pro (Family FK) and p.Lys449Thr (Family FG), the latter

associated with myofibrillar myopathy.[30] Four additional kindreds (Families BG, BK, FR, KS) had damaging group A gene variants with LOD scores <1 (Supplementary Figure S5). The damaging variants were present in 15 of 16 affected individuals in these families, with genotype-phenotype discordance driven mainly by non-penetrant variant carriers. In all families tested, there was no evidence for co-segregation of any variants that did not meet the criteria for being "damaging" or that were in genes other than group A.

Multiple variants. With expanded genetic testing panels, it is not uncommon to find several rare variants in DCM cases, prompting hypotheses of multiple mutations.[7,31] In an extension of the current analysis, we reviewed data for individuals who had undergone testing with the 69-gene testing panel and looked at all 69 genes as well as increasing the threshold level of MAF to <1%. Family co-segregation analysis was then performed in 5 kindreds in which the proband was found to carry ≥5 variants. In 4 of these families (Families BY, FK, AB, BA), only one of the proband's multiple variants segregated with DCM (Supplementary Figure S6). In the remaining family (Family GX), there were no co-segregating variants. Four of the proband's 5 rare variants were inherited from her unaffected father, and none were damaging group A variants. These studies demonstrate the value of family analysis and support the expectation that a single "driver" variant will be present when DCM appears as a Mendelian trait.

Location of variants in group A genes

To explore potential effects of variant location within group A genes, we compared the distribution of missense variants identified in DCM cases (derived from our own cohorts and two clinical laboratories[12]) with rare (MAF <0.1%) missense variants in ExAC (Figure 2, Supplementary Table S11). In *MYH7*, DCM variants occurred in all protein domains but were preferentially located in the myosin motor (Figure 2A). DCM-associated variants in *RBM20*

were significantly enriched in the RS domain that is associated with altered titin splicing and myocardial compliance (Figure 2B).[28,29] More than half of the damaging *RBM20* variants resided in the RS domain or in a glutamate-rich region that is also associated with titin splicing defects.[32] DCM-associated *LMNA* variants were mostly located in the coiled-coil rod domain, particularly in coil 2, which is critical for dimer formation (Figure 2C),[33] and there was a cluster of DCM-associated variants in the S4 voltage-sensor, repeat I, of *SCN5A* (Figure 2E). Variants in the S4 transmembrane segments cause arrhythmic forms of DCM and have been associated with gain-of-function effects and gating pore currents[25,26,34-36] Clustering of DCM variants was also apparent in the α-tropomyosin binding domain of *TNNT2*, and several damaging variants in *TPMI* resided in the cardiac troponin T binding domain (Figure 2G and 2H).

DISCUSSION

Our data show that rare protein-altering cardiomyopathy gene variants are not unique to DCM patients but also occur commonly in apparently healthy individuals. Because of this, effective discriminitive strategies are clearly needed in order to derive meaningful results from genetic testing. Here we provide a new approach for assessment of rare variants that combines both variant-level and gene-level information. We propose that ranking genes based on their *a priori* likelihood of disease causation is a key factor in identifying clinically-actionable variants.

Only 14 of 41 genes achieved "group A" status. Our literature-based gene grading method was subsequently validated by the predilection for damaging variants in DCM cases to occur in group A genes and by the similar clustering of pathogenic/likely-pathogenic ClinVar variants. Gene group is a dynamic parameter, and is highly dependent on the contemporary knowledge base. Assessment of genetic evidence is intrinsically biased by the size of the kindreds evaluated, the frequency in which genes have been screened, and published reports, and definitive conclusions cannot be drawn if there is insufficient information. Genes in which variants co-segregated with DCM in multiple large kindreds ($\geq$5 affected) scored highly but for many genes, only small families or single cases have been studied, raising the possibility that variants, even if function-altering, might co-segregate by chance or be incidental findings. Similarly, *in vitro* functional data are helpful if these are available but are often lacking from the literature. Group A genes typically had animal data showing spontaneous development of DCM in models expressing heterozygous loss-of-function alleles or human missense mutations. This level of evidence was missing from many studies in which only homozygous loss-of-function animals have been studied. While such models can implicate genes in normal cardiac function, it cannot be assumed that there is a direct correlation with gene "dose" and that heterozygous counterparts will have a similar,

albeit less severe, phenotype. In homozygous loss-of-function animals, there may also be profound effects on cardiac development that independently predispose to contractile impairment.

The group A genes *MYH7*, *RBM20*, and *LMNA*, had significantly more rare variants in DCM cases than controls, and frequently harbored damaging variants. *MYH7* and *LMNA* have consistently appeared on lists of the "most frequently mutated" genes in DCM genetics studies, and were the top 2 genes (after *TTN*) showing an excess of rare variants (MAF <0.0001) in a recent analysis of 1315 DCM patients referred to diagnostic genetic testing laboratories.[6,7,12,22] *RBM20* is a more recent addition to the disease gene list and has been less frequently screened. Despite their overall enrichment in DCM cases, not all rare variants in *MYH7*, *RBM20*, and *LMNA* are necessarily deleterious and additional information is required for clinical interpretation of any specific variant. Interestingly, variants in these genes had a non-random distribution in DCM cases: *MYH7* variants clustered in the myosin motor, *RBM20* variants in the RS domain, and *LMNA* variants in the coiled-coil rod. As the numbers of reported DCM-associated variants expands, additional patterns of variant distribution patterns may become apparent.

Our method for stratifying rare variants provides a framework for future studies and further refinements are anticipated. We expect that over time, more genes will achieve group A status, and ongoing rigorous gene ranking is warranted. For missense variants, the criterion of absence from ExAC may underestimate the yield of deleterious variants and varying MAF threshold levels need to be evaluated. It is notable, for example, that the few ClinVar-annotated pathogenic/likely-pathogenic group A gene variants that were not captured by our method were mostly extremely rare in ExAC. Many of the Clinvar-listed variants are annotated as pathogenic/likely-pathogenic for disorders other than DCM. This phenotypic heterogeneity suggests that at least some of these variants might be modifiers rather than

directly causative of disease. Another explanation is that the ultimate clinical manifestations of any given variant may be flavored by the context of each person's genetic and lifestyle factors. The detection of pathogenic/likely-pathogenic variants for diverse disorders suggests that our prioritization method could be applied to a range of genetic testing indications, with the caveat that disease-specific gene lists would be needed.

The lack of cardiac investigations in ~40% ~~all~~ control subjects is a limitation of this study. We focussed on subjects of European ancestry to avoid confounding effects of ethnicity-related background genetic variation, and the applicability of our findings to other populations has yet to be determined. In assessing individual burden of genetic variation, only rare cardiomyopathy gene variants were considered. The extent to which rare, low-frequency and common variants in a broad range of cardiac genes might cumulatively contribute to a myopathic substrate is unknown.

Ongoing efforts by the American College of Medical Genetics to standardize variant annotation methods are an important step towards establishing international guidelines for genetic testing in DCM and other inherited human disorders.[37] The cost-efficacy of first-line screening of a core set of key DCM disease genes (e.g. group A genes) vs extended panels or genome-wide testing warrants further analysis. As genome sequencing is poised to become part of mainstream healthcare, there is a pressing need for curation of clinical and genetic information in databases such as ClinGen,[38] functional evaluation of variants, and clinical trials in genotyped patients. Advancements in these areas should expedite implementation of personalized medicine.

In summary, our data suggest that interpretation of genetic testing results cannot be made on the basis of variant characteristics alone, and that the effect size of protein-altering variants is primarily determined by the biological importance of the gene involved and relevance to DCM pathogenesis.

ACKNOWLEDGMENTS

We thank Erica Mazaika, Monique Ohanian, Santiago Pineda, Celine Santiago. Magdalena Soka, and Gunjan Trivedi for assistance with data collection, collaborating physicians who referred patients to the study and participating families. See Supplementary Acknowledgments for the Medical Genome Research Bank.

REFERENCES

1. Pan S, Caleshu CA, Dunn KE, et al. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. *Circ Cardiovasc Genet* 2012;5:602-610.
2. Norton N, Robertson PD, Rieder MK, et al. Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. *Circ Cardiovasc Genet* 2012;5:167-174.
3. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet* 2012;5:391-399.
4. Haggerty CM, James CA, Calkins H, et al. Electronic health record phenotype in subjects with genetic variants associated with arrhythmogenic right ventricular cardiomyopathy: a study of 30,716 subjects with exome sequencing. *Genet Med* 2017;19:1245-1252.
5. Fatkin D, Seidman CE, Seidman JG. Genetics and disease of ventricular muscle. *Cold Spring Harb Perspect Med* 2014;4;a021063.
6. Pugh TJ, Kelly MA, Gowrisankar S, et al. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med* 2014:16:601-608.
7. Haas J, Frese KS, Peil B, et al. Atlas of the clinical genetics of human dilated cardiomyopathy. *Eur Heart J* 2015;36:1123-1135.
8. McNally EM, Mestroni L. Dilated cardiomyopathy: genetic determinants and mechanisms. *Circ Res* 2017;121:731-748.
9. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-249.
10. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acid Res* 2012;40:W452-W457.

11. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 2015;24:2125-2137.
12. Walsh R, Thomson KL, Ware JS, et al. Exome Aggregation Consortium, MacArthur DG, Farrall M, Cook SA, Watkins H. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet Med* 2017;19:192-203.
13. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* 1998;280:750-752.
14. Fatkin D, MacRae C, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999;341:1715-1724.
15. Kamisago M, Sharma SD, dePalma SR, et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* 2000;343:1688-1696.
16. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;285-291.
17. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 2013;9:e1003709.
18. Hishiya A, Kitazawa T, Takayama S. BAG3 and Hsc70 interact with actin capping protein CapZ to maintain myofibrillar integrity under mechanical stress. *Circ Res* 2010;107:1220-1231.
19. Villard E, Perret C, Gary F, et al. A genome-wide association study identified two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J* 2011;32:1065-1076.

20. Norton N, Li D, Rieder MJ, et al. Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. *Am J Hum Genet* 2011;88:273-282.

21. Feldman AM, Begay RL, Knezevic T, et al. Decreased levels of BAG3 in a family with a rare variant and in idiopathic dilated cardiomyopathy. *J Cell Physiol* 2014;229:1697-1702.

22. Kayvanpour E, Sedaghat-Hamedani F, Amr A, et al. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin Res Cardiol* 2017;106:127-139.

23. Van Bemmelen MX, Rougier JS, Gavillet B, et al. Cardiac voltage-gated sodium channel Nav1.5 is regulated by Nedd4-2 mediated ubiquitination. *Circ Res* 2004;95:284-291.

24. Gavillet B, Rougier JS, Domenighetti AA, et al. Cardiac sodium channel Nav1.5 is regulated by a multiprotein complex composed of syntrophins and dystrophin. *Circ Res* 2006;99:407-414.

25. McNair WP, Ku L, Taylor MR, et al. *SCN5A* mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* 2004;110:2163-2167.

26. Olson TM, Michels VV, Ballew JD, et al. Sodium channel mutations and susceptibility to heart failure and atrial fibrillation. *JAMA* 2005;293:447-454.

27. Dellefave LM, Pytel P, Mewborn S, et al. Sarcomere mutations in cardiomyopathy with left ventricular hypertrabeculation. *Circ Cardiovasc Genet* 2009;2:442-449.

28. Brauch KM, Karst ML, Herron KJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol* 2009;54:930-941.

29. Guo W, Schafer S, Greaser ML, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med* 2012;18:766-773.

30. Selcen D, Ohno K, Engel AG. Myofibrillar myopathy: clinical, morphological and genetic studies in 63 patients. *Brain* 2004;127:439-451.

31. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*. 2013;10:531-547.

32. Beqqali A, Bollen IA, Rasmussen TB, et al. A mutation in the glutamate-rich region of RNA-binding motif protein 20 causes dilated cardiomyopathy through missplicing of titin and impaired Frank-Starling mechanism. *Cardiovasc Res* 2016;112:452-463.

33. Gruenbaum Y, Foisner R. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. Annu Rev Biochem 2015;84:131-164.

34. McNair WP, Sinagra G, Taylor MR, et al. *SCN5A* mutations associate with arrhythmic dilated cardiomyopathy and commonly localize to the voltage-sensing mechanism. *J Am Coll Cardiol* 2011;57:2160-8.

35. Mann SA, Castro ML, Ohanian M, et al. R222Q *SCN5A* mutation is associated with reversible ventricular ectopy and dilated cardiomyopathy. *J Am Coll Cardiol* 2012;60:1566-1573.

36. Moreau A, Gosselin-Badaroudine P, Boutjdir M, Chahine M. Mutations in the voltage sensors of domain I and II of Nav1.5 that are associated with arrhythmias and dilated cardiomyopathy generate gating pore currents. *Front Pharmacol* 2015;6:301.

37. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-424.

38. Rehm HL, Berg JS, Brooks LD, et al. ClinGen – the clinical genome resource. *N Engl J Med* 2015;372:2235-2242.

Figure legends

Figure 1. Criteria for variant prioritization. (**A**) Missense rare variants in cardiomyopathy genes in DCM cases and controls were compared using a number of criteria including novelty with respect to the European subgroup of the Exome Sequencing Project (EA-ESP) and non-Finnish European (NFE) or all subjects in the Exome Aggregation Consortium (ExAC), *in silico* functional predictions (PolyPhen2, SIFT, PROVEAN, MetaSVM), RVIS score and gene group. The odds ratios (OR) for DCM are displayed in a forest plot. (**B**) Distribution of truncating variants across gene groups in controls (white bars) and DCM cases (blue bars). (**C**) Distribution of missense variants across gene groups in controls (white bars) and DCM cases (blue bars). The subset of "novel (ExAC) + deleterious (MetaSVM)" missense variants are denoted in the solid sections of bars. There were significant effects of clinical status and gene group for truncating variants (**B,** $P = 0.01$; 2 x 4 chi-squared test) and missense variants (**C;** $P = 0.018$); see Supplementary Table S8 for statistical analysis. For all panels, *TTN* truncating and missense variants were excluded.

Figure 2. Rare variant distribution in DCM cases and control subjects. The location of variants in different protein domains of group A genes is shown in panels A to L. In each panel, variants identified in DCM patients in the primary cohort (top row) and in the familial DCM (FDCM) replication cohort (second row) are shown: all variants (pink, "DCM all"), damaging variants (red, "DCM HP"). Variant types are indicated by: circle =missense, square = splice site change, diamond = frameshift, star = stop codon. The third row shows DCM variants identified by two clinical diagnostic laboratories ("DCM Walsh")[12]. Missense variants identified in control subjects are shown below the protein schematic: control subjects in this study (Controls), Exome Aggregation Consortium (ExAC) database. See Supplementary Table S11 for protein domain coordinates and statistical analysis.

A Gene-Centric Strategy for Identifying Disease-Causing

Rare Variants in Dilated Cardiomyopathy

DISCLOSURE

The authors declare no conflict of interest.

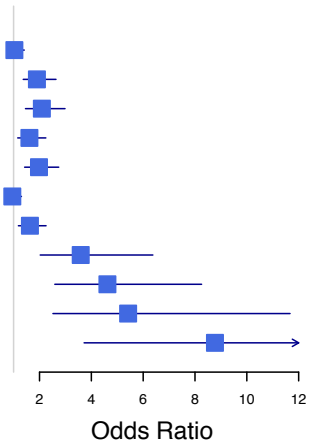
FUNDING SOURCES

This study was supported by grants from the National Health and Medical Research Council of Australia (D.F., P.S.M.), Harvard Club of Australia Foundation (D.F.), Estate of the Late RT Hall (D.F.), Simon Lee Foundation (D.F.), British Heart Foundation (S.A.C.), National Institute for Health Research Cardiovascular BRU at the Brompton and Harefield & Harefield NHS Foundation Trust and Imperial College London (S.A.C., P.J.B.), Howard Hughes Medical Institute (C.E.S.), National Institutes of Health (D.S.H., J.G.S.), and the Leducq Foundation (S.A.C., J.G.S., C.E.S.).

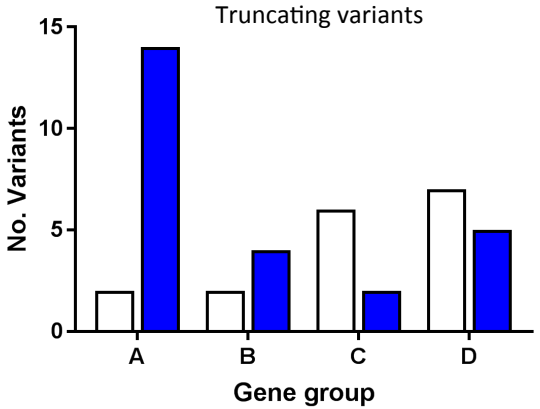
Figure

A

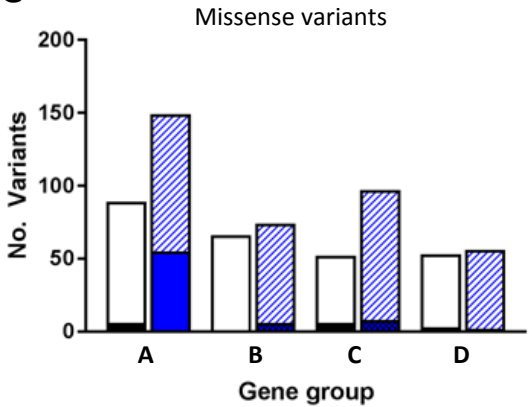
Criteria	OR	p–value
Novel (EA–ESP)	1.03	0.8278
Novel (ExAC–NFE)	1.90	0.0001
Novel (ExAC)	2.09	0.0001
Damaging (PolyPhen2/SIFT/PROVEAN)	1.61	0.0046
Damaging (MetaSVM)	1.98	<0.0001
RVIS 25th percentile	0.95	0.7716
Group A gene	1.63	0.0028
Novel (ExAc) + Damaging (PolyPhen2/SIFT/PROVEAN)	3.59	<0.0001
Novel (ExAC) + Damaging (MetaSVM)	4.62	<0.0001
Novel (ExAC) + Damaging (MetaSVM) + RVIS 25th percentile	5.42	<0.0001
Novel (ExAC) + Damaging (MetaSVM) + Group A gene	8.77	<0.0001



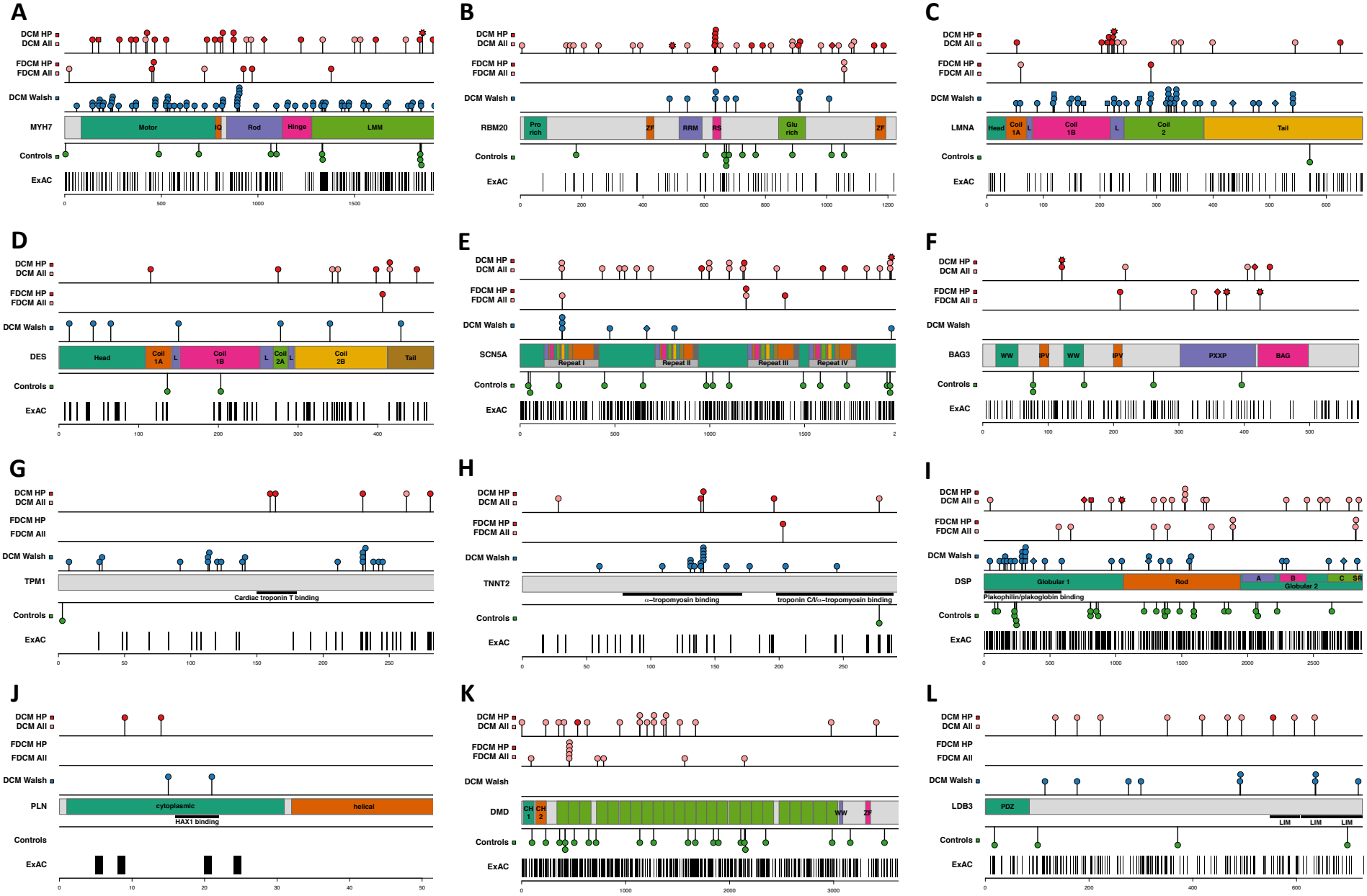
B



C



Figure


A
DCM HP
DCM All
FDCM HP
FDCM All
DCM Walsh
MYH7
Motor
IQ
Rod
Hinge
LMM
Controls
ExAC
B
RBM20
Pro rich
ZF
RRM
RS
Glu rich
C
LMNA
Head
Coil 1A
L
Coil 1B
Coil 2
Tail
D
DES
Coil 2A
Coil 2B
E
SCN5A
Repeat I
Repeat II
Repeat III
Repeat IV
F
BAG3
WW
IPV
PXXP
BAG
G
TPM1
Cardiac troponin T binding
H
TNNT2
α–tropomyosin binding
troponin C/I/α–tropomyosin binding
I
DSP
Globular 1
Rod
A
B
C
SR
Globular 2
Plakophilin/plakoglobin binding
J
PLN
cytoplasmic
helical
HAX1 binding
K
DMD
CH 1
CH 2
WW
ZF
L
LDB3
PDZ
LIM
LIM
LIM

Table

Table 1. Gene characteristics and evidence for role in DCM pathogenesis

Gene	Protein	Gene group	RVIS[a]	pLI score[b]	Z Score[b]	Genetic data[c]	*In vitro* defects	Animal model[d]	Other phenotypes
MYH7	Myosin heavy chain 7	A	0.28	0.00	6.54	+++	Yes	+++[e]	HCM, RCM, LVNC, CHD, DM, MSM, SPM
DSP	Desmoplakin	A	1.32	1.00	0.91	+++	Yes	+++	ARVC, LVNC, palmoplantar keratoderma, skin-fragility woolly hair syndrome
SCN5A	Cardiac sodium channel	A	1.99	1.00	2.53	+++	Yes	+++	CD, AF, LQTS, BS, VF, LVNC
LMNA	Lamin A/C	A	9.96	0.99	3.37	+++	Yes	+++	CD, EDMD, >10 additional phenotypes/overlap syndromes.
LDB3	LIM domain binding protein 3	A	10.16	0.00	0.32	+++	Yes	+++	HCM, LVNC, ARVC, MM
DMD	Dystrophin	A	11.28	1.00	-4.82	+++	Yes	+++	DMD, BMD
DES	Desmin	A	21.56	0.00	2.34	+++	Yes	+++	RCM, ARVC, HCM, CD, MM, LGMD-2R, NSS
TPM1	Tropomyosin α-1 chain	A	25.15	0.80	3.42	+++	Yes	+++	HCM, RCM, LVNC
TNNC1	Troponin C	A	39.68	0.51	2.22	+++	Yes	+++	HCM
TNNT2	Troponin T	A	41.64	0.01	1.54	+++	Yes	+++	HCM, RCM, LVNC
BAG3	BCL-2 associated athanogene	A	62.14	0.53	-1.01	+++	Yes	+++	MM
PLN	Phospholamban	A	62.38	0.11	0.57	+++	Yes	+++	HCM, ARVC
RBM20	RNA-binding protein 20	A	90.98	NA	NA	+++	Yes	+	
TTN	Titin	A	98.04	0.00	-5.48	+++	Yes	++	HCM, ARVC, RCM, LVNC, CM, CNM, HMERF, LGMD-2J, TMD
VCL	Vinculin	B	2.23	0.99	3.11	++	Yes	++	HCM
LAMA4	Laminin subunit α-4	B	4.68	0.00	-1.14	+	Yes	++	
ILK	Integrin-linked protein kinase	B	14.97	0.04	0.96	+	Yes	++	
MYBPC3	Cardiac myosin binding	B	20.54	0.00	0.69	++	Yes	++	HCM, LVNC

	protein 3								
CSRP3	Muscle LIM protein	B	28.93	0.00	-0.66	++	Yes	++	HCM
CRYAB	αβ-crystallin	B	35.42	0.01	0.38	++	Yes	+++	MM, DM, cataracts
ACTC1	Cardiac actin	B	38.28	0.95	5.25	++	Yes	+++	HCM, RCM, LVNC, CHD
NEXN	Nexilin	B	50.34	0.00	-1.32	++	NA	++	HCM, CHD
TAZ	Tafazzin	B	59.76	0.97	2.42	+++	Yes	++	Barth syndrome (may present as infantile X-linked DCM), LVNC, EFE
LAMP2	Lysosome-associated membrane glycoprotein 2	B	61.73	0.95	0.41	++	Yes	+	Danon disease (may present as HCM or DCM)
TNNI3	Troponin I	B	64.11	0.17	1.88	++	Yes	+	HCM, RCM, AF
ANKRD1	Cardiac ankyrin repeat protein	B	65.76	0.00	-0.01	++	Yes	No	HCM, CHD
DSG2	Desmoglein-2	B	98.32	0.00	-1.20	++	NA	+++	ARVC
MYH6	Myosin heavy chain 6	C	0.66	0.00	2.87	++	NA	++[e]	HCM, CHD, SSS
ABCC9	K_{ATP} channel SUR2A subunit	C	1.82	0.00	4.89	+	Yes	++	AF, BS, Cantu syndrome, coronary vasospasm
ACTN2	α-actinin	C	2.47	1.00	1.76	+	Yes	+	HCM
PDLIM3	PDZ and LIM domain protein 3	C	20.70	0.00	-0.04	+	NA	++	
SGCD	δ - sarcoglycan	C	34.32	0.00	-0.23	++	Yes	++	LGMD-2F
SYNE1	Nesprin-1	C	41.65	0.00	-0.95	++	Yes	+	EDMD, cerebellar ataxia, autism, arthrogryposis
FXN	Frataxin	C	76.67	0.82	0.47	+	NA	++	Friedreich's ataxia (may develop LVH + late DCM)
CAV3	Caveolin-3	C	77.70	0.34	1.19	+	Yes	++	HCM, LQTS, SIDS, LGMD-1C, RMD-2, DM, HCK
TCAP	Telethonin	C	78.28	0.08	0.45	+	Yes	+	HCM, LGMD-2G
FKTN	Fukutin	C	82.25	0.00	-0.64	++	Yes	NA	LGMD-2M (may present as DCM)
SDHA	Succinate dehydrogenase	D	10.95	0.00	2.32	++	Yes	NA	Leigh syndrome (rarely associated

	flavoprotein subunit								with pediatric recessive DCM)
DTNA	α-dystrobrevin	D	21.65	0.92	1.17	+	NA	No	LVNC, CHD
LAMA2	Laminin subunit α-2	D	79.64	0.00	-1.58	+	NA	NA	Merosin-deficient congenital MD/LGMD (rarely includes DCM)
SYNM	Synemin	D	NA	0.00	-0.15	NA	NA	NA	Accumulates in skeletal myopathy

[a] Residual Variation Intolerance Score (RVIS) – see text and reference 17.
[b] Probability of being loss-of-function intolerant (pLI) and missense Z-scores - see text and reference 16.
[c] Genetic evidence: +++ co-segregation of variants in multiple (3+) families, with at least one family showing variant co-segregation in 5 or more affected individuals; ++ co-segregation of variant in 1 or more small families (<5 affected); + variant demonstrated in single cases (with/without positive family history), or SNPs associated with DCM susceptibility.
[d] Animal models: +++ Spontaneous development of DCM in animal model with a heterozygous gene loss-of-function or expressing a human disease-associated non-synonymous variant; ++ spontaneous DCM present in homozygous gene loss-of-function model, or stress-induced DCM in model of a human non-synonymous variant; + No DCM but other relevant phenotypes, eg. isolated ventricular dilation or contractile dysfunction, evidence of reduced contractile reserve, skeletal myopathy. See Supplementary Methods and Supplementary Table S7 for more details of gene group allocation and animal data.
[e] α-MHC is major myosin isoform in murine ventricle: human *MYH7* mutations modelled in murine *myh6* gene.
AF, atrial fibrillation; ARVC, arrhythmogenic right ventricular cardiomyopathy; BMD, Becker muscular dystrophy; BS, Brugada syndrome; CD, conduction-system defects; CHD, congenital heart disease; CM, congenital myopathy; CNM, centronuclear myopathy; DCM, dilated cardiomyopathy; DM, distal myopathy; DMD, Duchenne muscular dystrophy; EDMD, Emery-Dreifuss muscular dystrophy; EFE, endomyocardial fibroelastosis; HCK, high creatine kinase; HCM, hypertrophic cardiomyopathy; HMERF, hereditary myopathy with early respiratory failure; LGMD, limb girdle muscular dystrophy; LQTS, long QT syndrome; LVH, left ventricular hypertrophy; LVNC, left ventricular non-compaction; MD, muscular dystrophy; MM, myofibrillar myopathy; MSM, myosin storage myopathy; NA, not available; NSS, neurogenic scapuloperoneal syndrome; RCM, restrictive cardiomyopathy; RMD, rippling muscle disease; SIDS, sudden infant death syndrome; SPM, scapuloperoneal myopathy; SSS, sick sinus syndrome; TMD, tibial muscular dystrophy; VF, ventricular fibrillation.

Table

Table 2. Yield per person of *TTN*tv and damaging[a] group A gene variants in the primary cohort and in replication cohorts.

Variants	DCM cases			Control subjects		
	Primary cohort (n=532 subjects)	FDCM (n=101 subjects)		Primary cohort (n=527 subjects)	Alzheimer's (n=2971 subjects)	MGRB (n=1144 subjects)
Single *TTN*tv	85 (16.0%)	15 (14.9%)		6 (1.1%)	6 (0.2%)	13 (1.1%)
Two *TTN*tv	1 (0.2%)	0		0	0	0
*TTN*tv + group A trunc	1 (0.2%)	0		0	0	0
*TTN*tv + group A missense	5 (0.9%)	1 (1.0%)		0	0	1 (0.1%)
Group A trunc only	13 (2.4%)	6 (5.9%)		2 (0.4%)	9 (0.3%)	3 (0.3%)
Group A trunc + missense	0	1 (1.0%)		0	0	0
Group A missense only	46 (8.6%)†	9 (8.9%)		6 (1.1%)	46 (1.5%)	29 (2.5%)
None	381 (71.6%)	69 (68.3%)		513 (97.3%)	2910 (97.9%)	1098 (96.0%)

[a] Damaging variants included truncating (trunc) or "novel (ExAC) + deleterious (MetaSVM)" missense variants in group A genes (*TTN* missense variants were excluded).† Three cases had 2 variants. ADSP, Alzheimer's Disease Sequencing Project; FDCM, familial dilated cardiomyopathy; MGRB, Medical Genome Reference Bank.

A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy

SUPPLEMENTAL MATERIAL

CONTENTS

Supplementary Methods	2
Supplementary Figure Legends	7
Supplementary Figures	9
Supplementary Tables	15
Supplementary Acknowledgments	43
Supplementary References	45

Supplementary Methods

Study subjects and sequencing

1. Primary cohort

The primary cohort was comprised of 532 patients with familial or sporadic idiopathic DCM (aged 47 ± 19 years, 69% males) and 527 healthy subjects without cardiovascular disease (aged 49 ± 15 years, 38% males). These study subjects were recruited from Brigham and Women's Hospital, Boston Children's Hospital, Victor Chang Cardiac Research Institute, Royal Brompton & Harefield NHS Foundation Trust, and the London Institute of Medical Sciences, Imperial College, London. Additional detail about the composition of the study groups and the overall study design are provided below, and in Supplementary Figure S1.

Boston subjects: 143 patients (aged 31 ± 18 years, 63% males) were recruited from among DCM patients clinically evaluated at Brigham and Women's Hospital or Boston Children's Hospital, or by referral for genetics research studies. All patients had familial or sporadic idiopathic DCM and were evaluated by medical history, physical examination, ECG and transthoracic echocardiography. 208 control subjects (aged 57 ± 15 years, 24% males) were recruited from non-cardiac outpatient clinics at Brigham and Women's Hospital as part of a broad volunteer biobank study ("Our Genes"). These individuals had no history of cardiovascular disease, but did not undergo ECG or echocardiographic screening. Patient and control DNA samples were sequenced at Harvard Medical School (the "Boston laboratory") using a 69-gene panel and the Illumina HiSeq 2000 sequencing platform (Supplementary Figure S1).

Australian subjects: 60 Australian patients (aged 45 ± 12 years, 50% males) with familial DCM were recruited for genetics research studies at the Victor Chang Cardiac Research Institute. Familial DCM was defined by the presence of DCM and/or early (<35 years) sudden unexplained death in two or more individuals in the absence of another heritable cardiac or systemic cause. All subjects and participating relatives were evaluated by history and physical examination, ECG and transthoracic echocardiography. Proband DNA samples were sent to the Boston laboratory where

they were sequenced using a 69-gene panel and the Illumina HiSeq 2000 sequencing platform (Supplementary Figure S1).

UK subjects: The UK study subjects included 329 patients with DCM (50 ± 16 years, 74% males) who were referred to the Royal Brompton and Harefield NHS Foundation Trust for cardiac magnetic resonance imaging (CMR), or with severe, end-stage DCM requiring heart transplantation or left-ventricular assist device, and 319 adult healthy volunteers (42 ± 14 years, 46% males) who had been recruited prospectively via advertisement at the London Institute of Medical Sciences, Imperial College, London. The volunteers all underwent CMR scanning, had no history of medical illness, were not taking regular medication, and did not have evidence of cardiac structural or functional impairment when scanned. CMR scans were performed using a 1.5 T Philips Achieva (Best, Netherlands). Cine images were acquired using a balanced-steady-state free-precession sequence in standard 2-, 3-and 4-chamber long-axis views (TE/TR 1.5/3.0 ms, flip angle 60°), followed by sequential 8 mm short-axis slices (2 mm gap) from the atrio-ventricular ring to the apex. Late gadolinium enhancement was not assessed. Ventricular volumes were measured and indexed using standard techniques. DNA samples from the UK-based subjects were sequenced at the NIHR Royal Brompton and Harefield Cardiovascular Biomedical Research Unit (the "UK laboratory" using a 64-gene panel and the SOLiD 5500xl sequencing platform. (Supplementary Figure S1).

2. Replication cohorts:

DCM replication cohort: An independent cohort of 101 Australian familial DCM probands (aged 45 ± 16 years, 67% males) with self-reported European ancestry was recruited as described above for the Australian subjects in the primary cohort, and was evaluated in the Boston laboratory using a 69-gene panel and a Illumina HiSeq 2000 sequencing platform (Supplementary Figure S1).

Control replication cohorts: De-identified data from two control groups were obtained (Supplementary Figure S1). These groups were selected specifically because (i) subjects were aged

over 70 years and hence were considered unlikely to carry significant cardiovascular disease-associated genetic variants, and (ii) information was able to be obtained about the numbers of variants carried by each individual.

The Alzheimer's Disease Sequencing Project (www.niagads.org/adsp/) is a case-control cohort of 10,909 individuals, mainly of European American ancestry, of which 4,767 were unaffected. We analyzed data from 2,971 individuals, who were confirmed to be of European ancestry using principal component analysis. This cohort was primarily ascertained by the presence or absence of Alzheimer's Disease and no specific cardiac investigations were performed. Whole-exome sequencing libraries were prepared using Agilent's SureSelect Human All Exon kits V3 and V5 or Roche VCRome. Sequencing was performed on a HiSeq2000 with paired ends reads, with a mean depth of coverage of 50x to 150x.

Data were also obtained for 1144 subjects in the Medical Genome Reference Bank (MGRB; https://sgc.garvan.org.au/initiatives/mgrb). The MGRB is comprised of whole-genome sequencing data and phenotypic information from healthy Australians over 70 years of age. MGRB participants, consented through contributing studies, 45 and UP (Sax Institute, Sydney) and the ASPirin in Reducing Events in the Elderly (ASPREE) clinical trial (Monash University, Melbourne) had no self-reported history of cardiovascular disease, degenerative neurological disorders or cancer at the time of consent into the study. Whole-genome sequencing was performed at the Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, using HiSeq X sequencers. On the basis of single nucleotide variant patterns, the MGRB cohort is estimated to have 97% non-Finnish European ancestry.

Principal components analysis

In order to avoid any possible population stratification within the primary cohort of DCM patients and control subjects, we performed a Principal Components Analysis (PCA) using SMARTPCA software.[1] This analysis was performed using 127 single nucleotide variants (listed in

Supplementary Table S1) that were shared in sequence data generated by the Boston and UK laboratories and that had a genotyping rate >88% in each cohort. Since no large-scale underlying differences between the Boston and UK cases and controls were detected (Supplementary Figure 2), we proceeded with using the combined primary cohort for the downstream analysis without excluding any individuals.

Genes sets evaluated

Genetic studies were performed using re-sequencing of custom panels of 69 genes (Boston laboratory) or 64 genes (UK laboratory). Forty-four genes were represented on both of these panels, but 3 of these genes *(CTF1, DNAJC19,* and *EMD,)* were not able to be evaluated further for roles in DCM pathogenesis as there were no rare variants detected in any of the DCM cases. The remaining set of 41 genes was included in our analyses. These gene lists are provided in Supplementary Table S2.

Gene group allocation

Although there is a long list of putative DCM "disease genes", the amount of supporting data for roles in DCM pathogenesis for each of these genes varies widely. For some genes, there are multiple mutations, co-segregation of variants in large families and animal models that recapitulate the human phenotype. For other genes, there are only predicted-pathogenic variants in single cases and the argument for disease causation is substantially weaker. In order to better rank genes, we reviewed the available literature (PubMed) and devised a semiquantitative scoring method that utilized grades of A (strong) to D (weak) for the strength of genetic, *in vitro,* and *in vivo* evidence of disease association. Gene grade needs to be considered as a dynamic parameter and is limited by the current knowledge base. For many genes, it can be expected that the allocated grade will change over time as more genetic and functional information is accrued.

Genetic evidence was obtained from gene screening studies in patients with familial and sporadic DCM. These studies had been primarily undertaken to look for rare disease-causing variants. Genetic data were scored as: +++ co-segregation of variants in multiple (3 or more) families, with at least one family showing variant co-segregation in 5 or more affected individuals; ++ co-segregation of variant in 1 or more small families (<5 affected); + variant demonstrated in single cases (with/without positive family history), or SNVs associated with DCM susceptibility. *In vitro* functional data were scored as: Yes or No, depending on whether there was experimental evidence to demonstrate that a specific variant had function-altering effects. Animal model data were scored as: +++ spontaneous development of DCM in animals that carry a heterozygous loss-of-function variant or expressing a human disease-associated non-synonymous variant; ++ spontaneous DCM present in homozygous gene loss-of-function model, or stress-induced DCM in model of a human non-synonymous variant; + no DCM but other relevant phenotypes, eg. isolated ventricular dilation or contractile dysfunction, evidence of reduced contractile reserve, skeletal myopathy.

The overall grade was based on the combined level of genetic and functional scores:

A – strong: +++ genetic evidence, *in vitro* evidence present, +++ animal evidence.

B – moderate: genetic and animal evidence present but less than +++ for both categories ± *in vitro* evidence.

C – some: ++/+ genetic evidence, ± *in vitro* evidence, ++/+/absent animal evidence.

D – weak: ++/+/absent genetic evidence (association of gene with related pathologies but rarely associated with DCM, reported in single cases only, or no variants reported), no animal evidence.

Supplementary Figure Legends

Figure S1. Flow diagram outlining the composition of the primary cohort and replication cohorts, and the sequencing performed in each group.

Figure S2. Principal components analysis (PCA) of study subjects in the primary cohort. In this plot, DCM cases are shown in black and controls in red. The DCM cases were recruited from the USA (n=143, black triangle), Australia (n=60, black square), and the UK (black circle). Control subjects from the USA (n=208, red triangle) and UK (n=319, red circle) are shown.

Figure S3. Distribution of ClinVar variants**.** (**A**) Total number of truncating and missense variants present in ClinVar for each gene group are shown for controls (white bars) and DCM cases (blue striped bars), with further breakdown into variants considered to be likely pathogenic/pathogenic (LP/P) by general consensus of submissions to ClinVar, for any disease (light solid section of bars) or DCM specifically (dark solid section of bars). There is significant enrichment of LP/P variants for group A genes ($P = 0.0007$). (**B**) Venn diagram showing overlap of damaging variants (classification criteria used in this paper) and variants annotated as LP/P in ClinVar. All variants identified in group A genes for DCM cases (light blue outer circle) are indicated. The subsets of variants assessed as damaging (novel (ExAC) + deleterious (MetaSVM)) (dark blue) or likely pathogenic/pathogenic (white) are shown. The overlap of circles shows variants that are considered disease-causing across both methods. *TTN* truncating and missense variants were excluded from analyses in (A) and (B).

Figure S4. Family co-segregation analysis. Pedigrees showing genotype-phenotype correlations for selected rare truncating and missense variants identified in family probands. The presence (+) or absence (-) of variants in relatives is shown. DCM status is denoted as affected (solid symbols), unaffected (open symbols), or unknown (gray symbols); probands are indicated by arrows.

Figure S5. Genotype-phenotype correlations for selected missense variants. All rare variants (MAF < 0.1%) in the 41-gene data-set that were present in the family proband were evaluated in family members. The presence (+) or absence (-) of variants in relatives is shown. DCM status is denoted as affected (solid symbols), unaffected (open symbols), or unknown (gray symbols); probands are indicated by arrows.

Figure S6. Pedigrees for 5 kindreds in which multiple (5+) rare variants were identified in the family proband. For this analysis, all variants with a MAF < 1.0% in any of the genes in the 69-gene targeted re-sequencing panel were included. The presence (+) or absence (-) of variants in relatives is shown, with co-segregating variants in bold type. DCM status is denoted as affected (solid symbols), unaffected (open symbols), or unknown (gray symbols); probands are indicated by arrows.

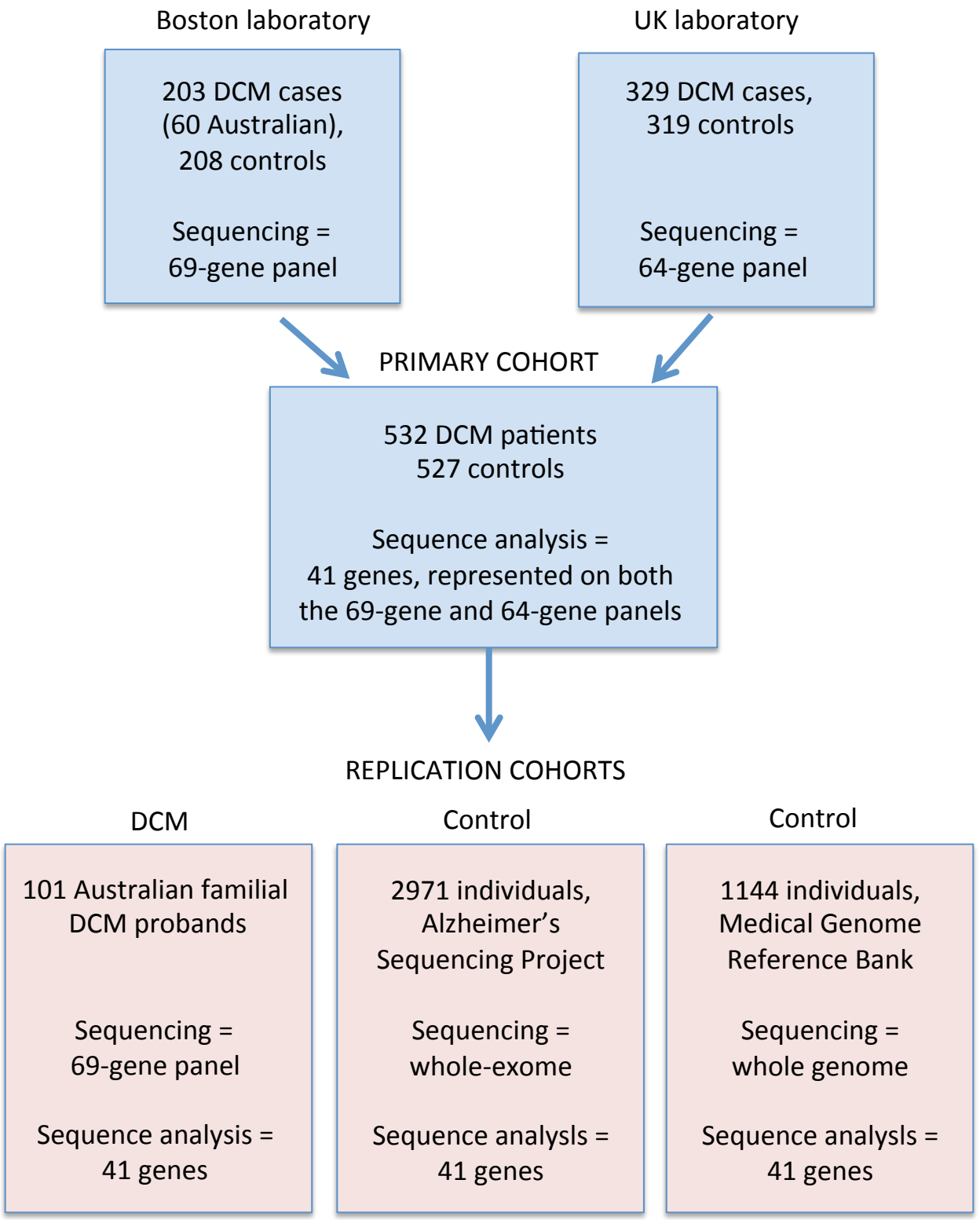

Boston laboratory
203 DCM cases
(60 Australian),
208 controls
Sequencing =
69-gene panel
UK laboratory
329 DCM cases,
319 controls
Sequencing =
64-gene panel
PRIMARY COHORT
532 DCM patients
527 controls
Sequence analysis =
41 genes, represented on both
the 69-gene and 64-gene panels
REPLICATION COHORTS
DCM
101 Australian familial
DCM probands
Sequencing =
69-gene panel
Sequence analysis =
41 genes
Control
2971 individuals,
Alzheimer's
Sequencing Project
Sequencing =
whole-exome
Sequence analysls =
41 genes
Control
1144 individuals,
Medical Genome
Reference Bank
Sequencing =
whole genome
Sequence analysls =
41 genes

Figure S1

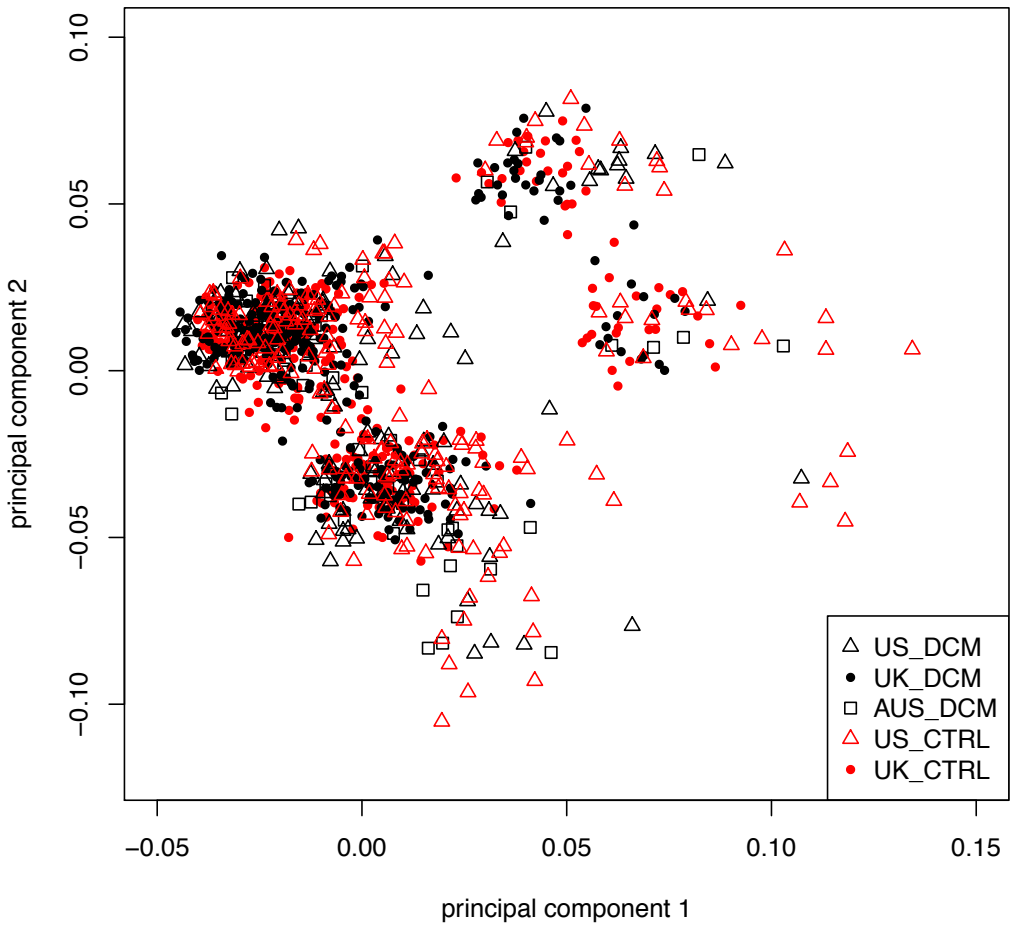
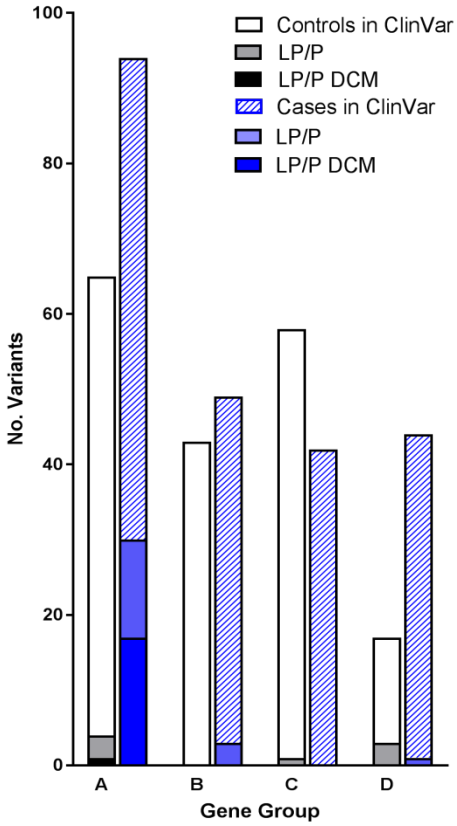

principal component 2
0.10
0.05
0.00
−0.05
−0.10
−0.05
0.00
0.05
0.10
0.15
principal component 1
US_DCM
UK_DCM
AUS_DCM
US_CTRL
UK_CTRL

Figure S2

A


Controls in ClinVar
LP/P
LP/P DCM
Cases in ClinVar
LP/P
LP/P DCM
No. Variants
100
80
60
40
20
0
A
B
C
D
Gene Group

B

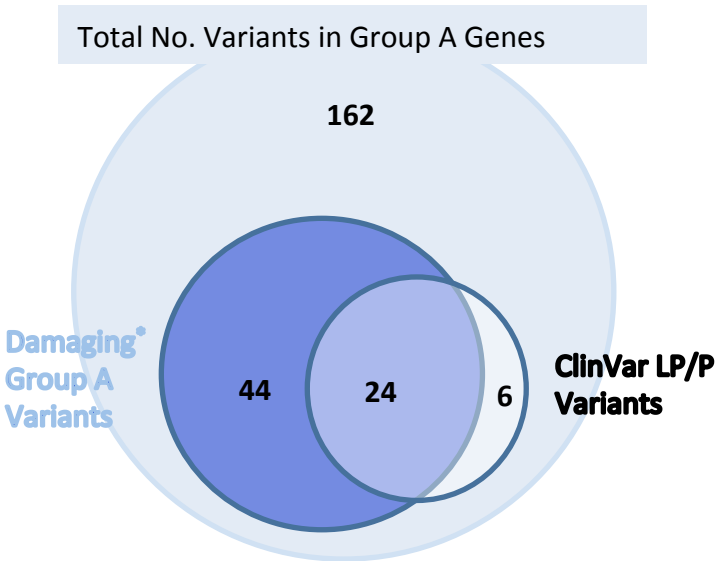

Total No. Variants in Group A Genes
162
Damaging* Group A Variants
44
24
6
ClinVar LP/P Variants

Figure S3

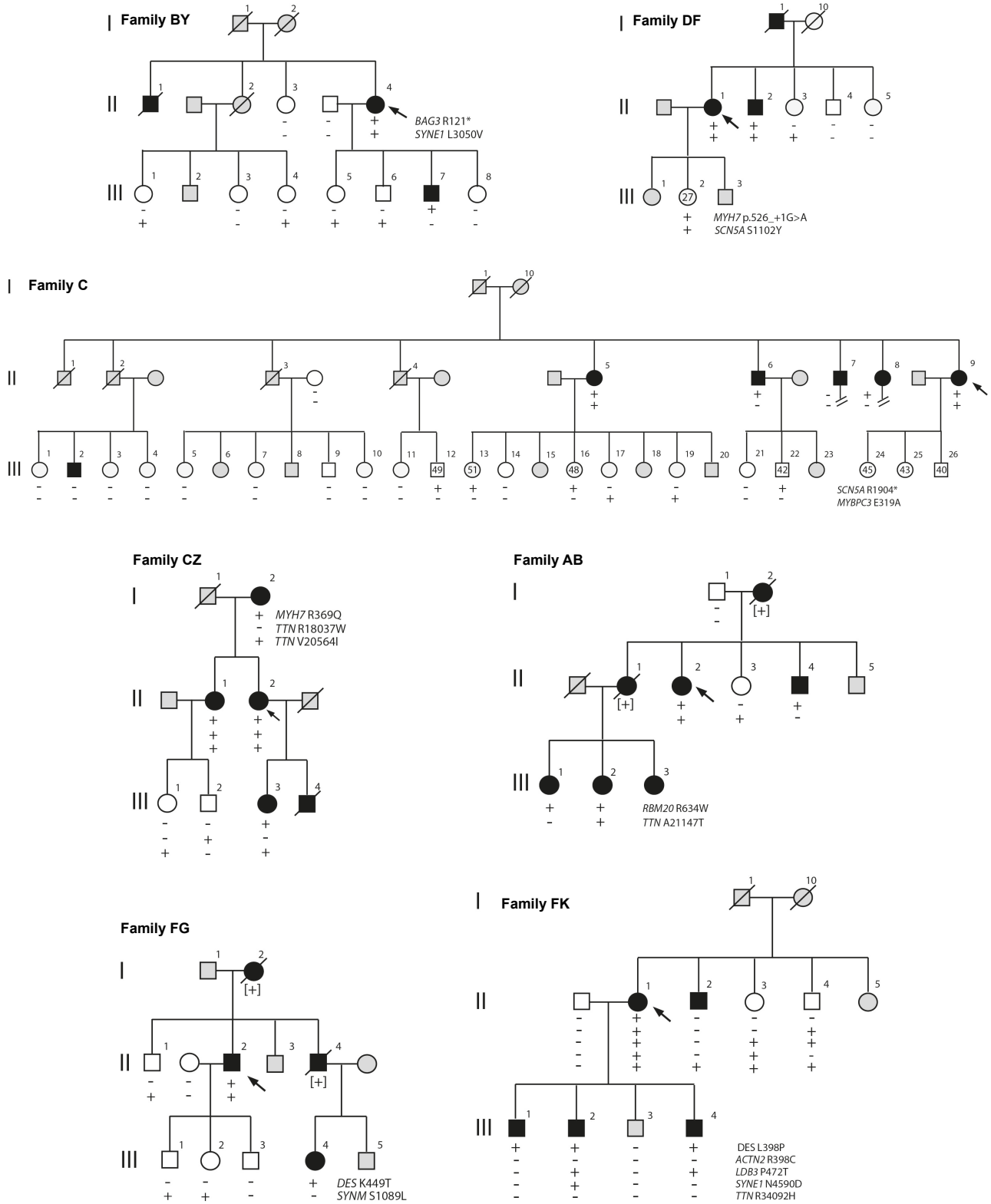

Family BY
BAG3 R121*
SYNE1 L3050V
Family DF
MYH7 p.526_+1G>A
SCN5A S1102Y
Family C
SCN5A R1904*
MYBPC3 E319A
Family CZ
MYH7 R369Q
TTN R18037W
TTN V20564I
Family AB
RBM20 R634W
TTN A21147T
Family FG
DES K449T
SYNM S1089L
Family FK
DES L398P
ACTN2 R398C
LDB3 P472T
SYNE1 N4590D
TTN R34092H

Figure S4

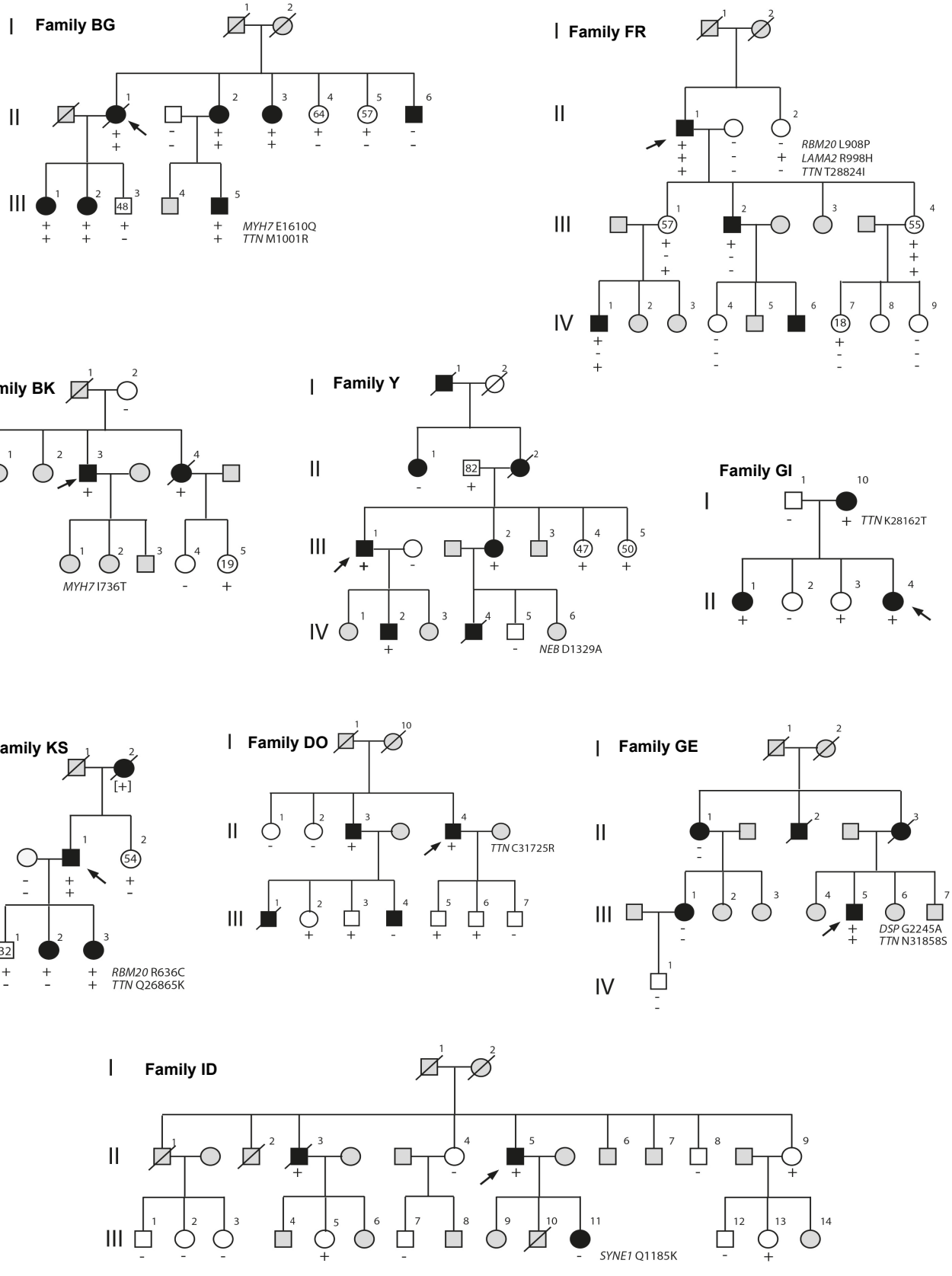

Family BG
MYH7 E1610Q
TTN M1001R
Family FR
RBM20 L908P
LAMA2 R998H
TTN T28824I
Family BK
MYH7 I736T
Family Y
NEB D1329A
Family GI
TTN K28162T
Family KS
RBM20 R636C
TTN Q26865K
Family DO
TTN C31725R
Family GE
DSP G2245A
TTN N31858S
Family ID
SYNE1 Q1185K

Figure S5

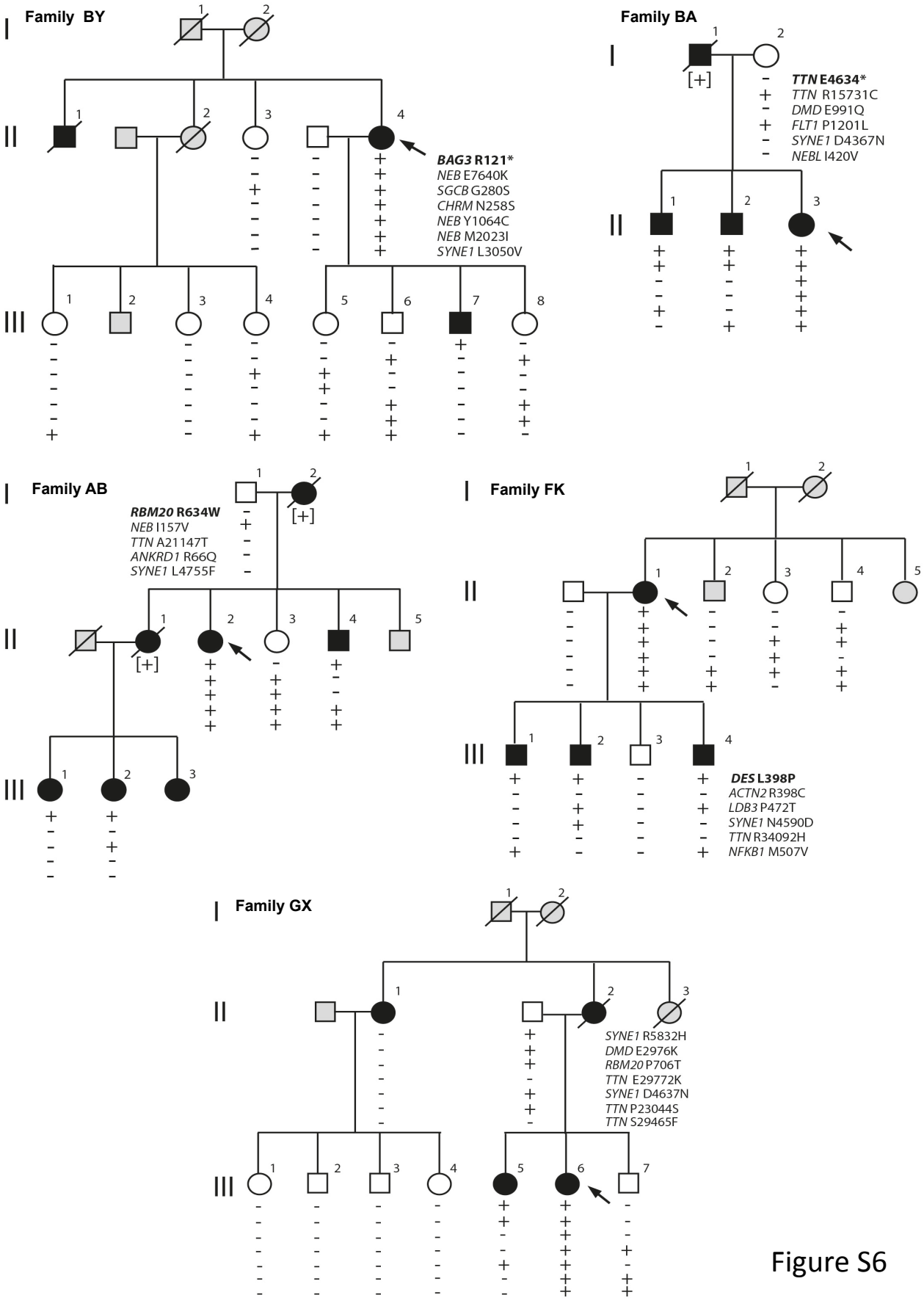

Family BY
BAG3 R121*
NEB E7640K
SGCB G280S
CHRM N258S
NEB Y1064C
NEB M2023I
SYNE1 L3050V
Family BA
TTN E4634*
TTN R15731C
DMD E991Q
FLT1 P1201L
SYNE1 D4367N
NEBL I420V
Family AB
RBM20 R634W
NEB I157V
TTN A21147T
ANKRD1 R66Q
SYNE1 L4755F
Family FK
DES L398P
ACTN2 R398C
LDB3 P472T
SYNE1 N4590D
TTN R34092H
NFKB1 M507V
Family GX
SYNE1 R5832H
DMD E2976K
RBM20 P706T
TTN E29772K
SYNE1 D4637N
TTN P23044S
TTN S29465F

Figure S6

Table S1. Single nucleotide variants used in the Principal Components Analysis.

Chromosome	SNV	Position	Allele 1	Allele 2
1	rs1048302	16340879	T	G
1	rs1048261	16340951	A	T
1	rs1749913	78383467	G	A
1	rs4483454	78399207	C	G
1	rs3729547	201334382	G	A
1	rs1573230	201335899	C	T
1	rs10800776	201338553	T	C
1	rs868407	201341341	C	T
1	rs11405	227069677	T	C
1	rs1295644	227071377	T	C
1	rs2236910	227073410	G	C
1	rs2802267	227078955	T	C
1	rs10753428	227081622	A	G
1	rs1341865	236882110	G	A
1	rs1341864	236882303	T	C
1	rs1341863	236883421	C	T
1	rs1341862	236883564	G	A
1	rs2288601	236902594	C	G
1	rs2282366	236924506	A	G
1	rs1803032	236926445	T	C
2	rs7260	26413917	C	T
2	rs1865324	26417533	C	A
2	rs13387855	26420708	T	G
2	rs7593175	26454996	A	G
2	rs2037814	73675669	T	G
2	rs4145333	179444768	C	G
2	rs2562831	179585266	C	T
2	rs2742348	179600563	G	A
2	rs10803917	179614952	A	G
2	rs922984	179615887	T	C
2	rs922985	179615931	C	G
2	rs922986	179615994	T	C
2	rs7585334	179620951	C	T
2	rs6433728	179621477	C	T
2	rs2291310	179623758	C	T
2	rs6705594	179623939	T	C
2	rs2291311	179629461	C	T
2	rs2291313	179632710	T	C
2	rs3816781	179639297	C	T
2	rs2291302	179643886	A	G
2	rs2291301	179643934	A	G
2	rs1552280	179644035	G	A
2	rs10497520	179644855	T	C
3	rs7430407	38622467	T	C
3	rs7645173	38627629	A	G
3	rs6599222	38648062	C	T

3	rs6599230	38674712	T	C
3	rs1035002	52488202	C	T
4	rs1609993	103514658	T	C
4	rs4635850	186423637	G	A
4	rs10866276	186423677	G	A
4	rs2306705	186427841	T	C
4	rs4862545	186456617	A	G
5	rs2288461	230980	A	G
5	rs1126417	231111	T	C
5	rs4434364	155771773	A	G
6	rs2806234	7563983	T	G
6	rs2076296	7567970	C	T
6	rs2076304	7572262	A	G
6	rs1016835	7576527	G	A
6	rs2806229	7578819	T	G
6	rs2842691	7578823	G	A
6	rs4945897	112443134	G	A
6	rs2032567	112457390	C	T
6	rs2032568	112457471	G	A
6	rs9387061	112508769	T	G
6	rs9400522	112508770	G	T
6	rs4404787	129381026	C	A
6	rs2229848	129807629	C	T
6	rs2229849	129807699	G	C
6	rs2297740	129835814	G	A
6	rs6929656	133827148	T	C
6	rs2813495	152476952	G	C
6	rs998147	152489294	T	C
6	rs684807	152534946	T	A
6	rs2147377	152540278	A	C
6	rs1890100	152638209	A	G
6	rs2306916	152647681	A	T
6	rs6911096	152652034	A	T
6	rs9479297	152658142	T	C
6	rs9479298	152658191	T	C
6	rs7775637	152688567	C	T
6	rs4523096	152793412	C	A
9	rs4146164	118418	T	C
9	rs6560541	71690684	G	T
9	rs7866579	71693128	G	C
10	rs2270552	75863750	C	T
10	rs2131956	75871735	C	G
10	rs703258	75879511	T	A
10	rs3740345	88445385	G	C
10	rs6585012	112543217	G	A
10	rs1417635	112572458	G	C
10	rs7070640	112590810	G	C
10	rs942077	112595719	G	C

10	rs196294	121435955	A	C
10	rs196295	121436362	A	G
11	rs896818	47364762	A	G
11	rs2070665	116707684	A	G
11	rs5070	116708020	A	G
12	rs704183	22005510	A	G
12	rs697250	22017422	A	G
12	rs697251	22017486	C	G
12	rs1356368	22035873	C	T
12	rs704216	22047151	G	T
12	rs10770865	22063115	A	G
12	rs4762719	22063737	G	T
12	rs4762720	22063749	T	C
13	rs9510700	23898454	C	T
13	rs1800354	23898664	A	G
13	rs9579177	28882948	G	A
13	rs7993418	28883061	G	A
13	rs4771238	28885672	C	T
13	rs7328692	28886113	A	G
13	rs2296188	28893484	T	C
13	rs7990486	28970917	C	T
14	rs434273	23872666	T	C
15	rs589759	35081574	C	T
15	rs3743248	99670989	T	C
15	rs260087	99671760	T	C
15	rs3210942	99675028	A	G
17	rs1053651	37822311	A	C
17	rs2696297	48245180	C	T
17	rs2696288	48252804	T	C
19	rs7252610	55665584	A	C
23	rs2293668	31224684	A	G
23	rs2293667	31224881	G	A
23	rs1800280	31496350	C	T

Table S2. Genes represented on the 69-gene (Boston) and 64-gene (UK) sequencing panels.

Gene	Boston panel	UK panel	Genes analyzed in this study[a]
ABCC9	X	X	X
ACTC1	X	X	X
ACTN2	X	X	X
ALMS1	X	Not tested	
ANKRD1	X	X	X
APOA1	X	Not tested	
BAG3	X	X	X
CALR3	Not tested	X	
CASQ2	Not tested	X	
CAV3	X	X	X
CHRM2	X	Not tested	
CRYAB	X	X	X
CSRP3	X	X	X
CTF1	X	X	No variants
DES	X	X	X
DMD	X	X	X
DNAJC19	X	X	No variants
DSC2	Not tested	X	
DSG2	X	X	X
DSP	X	X	X
DTNA	X	X	X
EMD	X	X	No variants
EYA4	X	Not tested	
FHL2	X	Not tested	
FKRP	X	Not tested	
FKTN	X	X	X
FLT1	X	Not tested	
FOXD4	X	Not tested	
FXN	X	X	X
GATA4	Not tested	X	
GLA	Not tested	X	
HADHA	X	Not tested	
HFE	X	Not tested	
HOPX	X	Not tested	
HSPB7	X	Not tested	
ILK	X	X	X
JPH2	Not tested	X	
JUP	Not tested	X	
LAMA2	X	X	X
LAMA4	X	X	X
LAMP2	X	X	X
LDB3	X	X	X
LMNA	X	X	X
MYBPC3	X	X	X
MYH6	X	X	X

MYH7	X	X	X
MYL2	Not tested	X	
MYL3	Not tested	X	
MYOZ2	Not tested	X	
MYPN	Not tested	X	
NEB	X	Not tested	
NEBL	X	Not tested	
NEXN	X	X	X
NFKB1	X	Not tested	
PDLIM3	X	X	X
PKP2	Not tested	X	
PLEC	X	Not tested	
PLN	X	X	X
PRKAG2	Not tested	X	
PSEN1	X	Not tested	
PSEN2	X	Not tested	
PTPN11	Not tested	X	
RAF1	Not tested	X	
RBM20	X	X	X
RYR2	Not tested	X	
SCN5A	X	X	X
SDHA	X	X	X
SGCA	X	Not tested	
SGCB	X	Not tested	
SGCD	X	X	X
SGCG	X	Not tested	
SOD2	X	Not tested	
SOS1	Not tested	X	
SYNE1	X	X	X
SYNM	X	X	X
TAF10	X	Not tested	
TAZ	X	X	X
TBX20	X	Not tested	
TCAP	X	X	X
TMEM43	Not tested	X	
TMPO	X	Not tested	
TNNC1	X	X	X
TNNI3	X	X	X
TNNT2	X	X	X
TP63	Not tested	X	
TPM1	X	X	X
TRIM63	Not tested	X	
TTN	X	X	X
VCL	X	X	X

[a] Genes were excluded from this study if they were (i) not represented on both the Boston and UK panels and (ii) no rare variants were found in DCM cases.

Table S3. Yield of rare (MAF< 0.1%) variants per gene in DCM patients and control subjects.

Gene	Variants (n=770) in DCM cases			Variants (n=589) in control subjects		
	Truncating	**Missense**	**All**	**Truncating**	**Missense**	**All**
TTN	83	287	370	6	254	260
SYNE1	0	54	54	3	61	64
LAMA2	2	29	31	4	28	32
RBM20	2	27	29[a]	0	10	10
MYH7	3	24	27[a]	0	9	9
DSP	3	16	19	0	23	23
MYBPC3	0	19	19	1	14	15
SYNM	1	18	19	2	18	20
SCN5A	1	17	18	1	14	15
DMD	0	16	16	0	20	20
ACTN2	1	13	14	0	11	11
LMNA	3	11	14[a]	0	1	1
LAMA4	0	13	13	0	14	14
MYH6	0	12	12	0	16	16
DSG2	0	10	10	0	10	10
LDB3	0	10	10	0	4	4
DES	0	8	8	0	2	2
NEXN	0	8	8	0	4	4
DTNA	1	6	7	0	2	2
VCL	1	6	7	0	9	9
BAG3	2	4	6	0	4	4
ILK	1	5	6	0	6	6
FKTN	0	5	5	1	1	2
TNNT2	0	5	5	0	1	1
TPM1	0	5	5	1	1	2
SDHA	1	3	4	1	5	6
TCAP	0	4	4	0	2	2
ABCC9	1	2	3	0	4	4
ACTC1	0	3	3	0	0	0
ANKRD1	0	3	3	0	6	6
LAMP2	1	2	3	0	1	1
PDLIM3	0	3	3	0	4	4
TNNC1	0	3	3	0	0	0
CRYAB	0	2	2	0	1	1
PLN	0	2	2	0	0	0
SGCD	0	2	2	0	3	3
TNNI3	1	1	2	1	1	2
CAV3	0	1	1	0	2	2
CSRP3	0	1	1	0	0	0
FXN	0	1	1	2	0	2
TAZ	0	1	1	0	0	0

[a] $P < 0.05$ compared with controls (2x2 chi-squared test; all *TTN* variants excluded).

Table S4. Characteristics of *TTN*tv identified in DCM cases.

Genomic position (Hg19)	Transcript effect	Protein effect	Protein domain	Exon; PSI	Sample ID	Ref.	MAF (ExAC)
179639647	c.6790+1G>T		I-band	29;1.00	10JL01453		Absent
179623709	c.10303+2T>C		I band	44;1.00	10CP01917		0.000008253
179606048	c.11912G>A	p.Trp3971*	I-band, N2B	49;0.95	LW-II-3	R[a]	Absent
179605315	c.12643_12644delCA	p.Gln4215fs	I-band, N2B	49;0.95	10DW00512	R[a]	Absent
179605203	c.12757C>T	p.Gln4253*	I-band, N2B	49;0.95	14MW01583	R[a]	Absent
179604264	c.13696C>T	p.Gln4566*	I-band, N2B	49;0.95	HFA-42	H[a]	0.000008294
179604060	c.13900G>T	p.Glu4634*	I-band, N2B	49;0.95	BA-III-3	R[a] W	Absent
179603994	c.13966C>T	p.Gln 4656*	I-band, N2B	49;0.95	DCM-22	F[a]	Absent
179591957	c.20134delG	p.Ser6712fs	I-band	70;0.69	HFA-9	H[a]	Absent
179571660	c.29062delG	p.Ala9688fs	I-band, N2A	102;0.88	12SK00375	R[a]	Absent
179558736	c.31427-1G>A		I-band	119;0.79	D13KD	H[a]	Absent
179554624	c.31763-1G>A		I-band, PEVK	123;0.88	12PB00376	R[a]	0.00027334
179547630	c.32888-1delG		I-band	136;0.86	10SB00367	R[a]	Absent
179506963	c.40558+1G>A		I-band, PEVK	220;0.96	HFA-26, HFA-41	H[a]	0.0001195
179498590	c.42636del	p.Ala14213fs	I-band	232;0.98	FW-II-1	R[a]	Absent
179498055	c.42947-2A>G		I-band	234;0.97	MIY-11	H[a]	Absent
179497076	c.43544_43545insA	p.Ala14514fs	I-band	237;0.98	MDT-11	H[a]	0.00003414
179495982	c.43792delG	p.Val14598*	I-band	238;0.96	10JM01592	R[a]	Absent
179494967	c.44281+1G>A		I-band	240;0.96	20MC01968	R[a]	0.00000831
179494168	c.44284C>T	p.Arg14762*	I-band	241;0.96	KI-III-2	R[a]	Absent
179487495	c.44816-1G>A		I-band	244;0.96	10BM02061	C	0.00001637
179486244	c.45307C>T	p.Arg15103*	I-band	246;0.97	12RH00086	W	Absent
179485580	c.45756dupA	p.Tyr15253fs	I-band	248;0.97	10BC00144		Absent
179485525	c.45812T>G	p.Leu15271*	I-band	248;0.97	10WS00448	R[a]	Absent
179483495	c.46782C>A	p.Tyr15594*	I-band	252;0.89	20JH01781	R[a] W	Absent
179482120	c.47692C>T	p.Arg15898*	I-band	255;0.86	AP-III-4	R[a]	0.00001504
179480145	c.48527G>A	p.Trp16176*	A-band	260;0.96	10CH00929	R[a]	Absent
179478864	c.49259delA	p.Glu16420fs	A-band	263;0.96	HFA-68	H[a]	Absent
179478665	c.49346-1G>A		A-band	264;0.95	20JB01501	R[a]	Absent

179478552	c.49458G>A	p.Trp16486*	A-band	264;0.95	MO-II-2	R^a	Absent
179477885	c.49648+2delT		A-band	265;0.96	DM-III-2, MDD-22, MIV-14	H^a	Absent
179477082	c.50170C>T	p.Arg16724*	A-band	267;0.93	10RN00513	R^a	Absent
179472127	c.53287+1G>T		A-band	278,0.94	HFA-71	H^a	Absent
179466192	c.55525_55531delGACAGGA	p.Asp18509fs	A-band	288;0.91	10JF01881	R^a	Absent
179464422	c.56206delA	p.Thr18736fs	A-band	290;0.96	FQ-III-11	R^a	Absent
179463603	c.56834delG	p.Gly18945fs	A-band	292;0.94	BM-III-25	R^a	Absent
179458947	c.58172delA	p.Asp19391fs	A-band	298;0.90	20PD01505	R^a	Absent
179458293	c.58732+2T>C		A-band	299;0.89	10AL00811	R^a	Absent
179458075	c.58858delC	p.Glu19619fs	A-band	300;0.90	HFA-63	H^a	Absent
179457380	c.59352delT	p.Glu19785fs	A-band	302;0.90	AM-III-7	R^a	Absent
179457005	c.40558G>C		A-band	303;0.91	DCMUKB12	H^a	Absent
179455162	c.61290T>A	p.Cys20430*	A-band	305;0.97	MEH-11	H^a	Absent
179454957	c.61495C>T	p.Arg20499*	A-band	305;0.97	MDJ-21	H^a W	Absent
179454576	c.61876C>T	p.Arg20626*	A-band	305;0.97	MEW-11	H^a	0.00006641
179453427	c.63025C>T	p.Arg21009*	A-band	305;0.97	10TM00933	R^a W	Absent
179452435	c.63601C>T	p.Arg21201*	A-band	307;0.91	MAO-92, 20EG01127	H^a	0.000008362
179450018	c.64453G>A	p.Arg21485*	A-band	310;0.91	DCMUKB4	W	0.00002508
179441649	c.69412+1G>A		A-band	325;0.93	DCMUKB2	H^a	Absent
179441479	c.69491_69492_delGT	p.Val23164fs	A-band	326;0.94	CS-III-12	R^a	Absent
179441341	c.69630C>A	p.Tyr23210*	A-band	326;0.94	10CS01472	R^a	Absent
179438189	c.72668delT	p.Pro24223fs	A-band	327;0.95	MEK-111	H^a	Absent
179435975	c.74880_74883insACTT	p.Pro24962fs	A-band	327;0.95	DI-II-6	R^a	Absent
179434743	c.76115insA	p.Asn25372fs	A-band	327;0.95	CT-II-2, R-IV-2, 20SF01123	R^a	Absent
179433758	c.77101_77102insT	p.Gln25700fs	A-band	327;0.95	HFA-57	H^a	Absent
179432761	c.78095_78098delAAAG	p.Arg26032fs	A-band	327;0.95	AV-IV-2	R^a F^a	Absent
179432675	c.78184G>T	p.Glu26062*	A-band	327;0.95	10BK01803	R^a	Absent
179432351	c.78507delT	p.Gly26170fs	A-band	327;0.95	10JC00262	R^a	Absent
179431868	c.78991C>T	p.Arg26331*	A-band	327;0.95	BR-IV-1, EA-II-7	R^a F^a	Absent
179429589	c.81262_81269delCAGATGCT	p.Gln27088fs	A-band	327;0.95	10PP00413, 10DM00077	R^a	Absent
179429538	c.81321C>G	p.Tyr27107*	A-band	327;0.95	12NP00167	R^a	Absent
179429340	c.81518delC	p.Pro27173fs	A-band	327;0.95	20DH01701	R^a	Absent

179428345	c.82513delA	p.Ile27505fs	A-band	327;0.95	10CP00605	R[a]	Absent
179426483	c.84376C>T	p.Gln28126*	A-band	327;0.95	LT-II-1	R[a]	Absent
179425091	c.85768C>T	p.Arg28590*	A-band	327;0.95	MGW-11	H[a]	Absent
179424496	c.86363G>A	p.Trp28788*	A-band	327;0.95	MIP-13	H[a]	Absent
179424036	c.86821+2T>A		A-band	327;0.95	MIA-1	H[a] WN	0.000008913
179423219	c.86967G>A	p.Trp28989*	A-band	328;0.96	12JL00046	R[a]	Absent
179422457	c.87624C>A	p.Tyr29208*	A-band	329;0.95	DCMUKB6	H[a]	Absent
179413670	c.92683C>T	p.Arg30895*	A-band	340;0.98	10KF00073	R[a]	Absent
179412902	c.93451G>T	p.Glu31151*	A-band	340;0.98	MBG_121	H[a]	Absent
179412874	c.93479G>A	p.Trp31160*	A-band	340;0.98	10AH00506		Absent
179412245	c.94103_94107delTTAAA	p.Thr31366fs	A-band	340;0.98	HFA-46	H[a]	0.000008475
179411592	c.94562dupC	p.Thr31522fs	A-band	342;0.98	20JM01785	R[a]	Absent
179410799	c.95164C>T	p.Gln31722*	A-band	344;0.99	MHQ-12	H[a]	Absent
179410544	c.95415_95416+2delCAGT		A-band	344;0.99	10KW01906	R[a]	Absent
179406990	c.97492+1G>C		A-band	350;1.00	DCM-30, HFA-66, HFA-83	H[a] F[a]	Absent
179404523	c.98265_98268dupAACA	p.His32757fs	A-band	353;1.00	10CS01784	R[a] W	Absent
179403522	c.99034A>T	p.Lys33012*	A-band	355;1.00	MHX-11	H[a]	Absent
179401029	c.100445C>A	p.Ser33482*	A-band	358;0.99	20JR01203	H[a] R[a]	Absent
179400319	c.101021_101022delGA	p.Arg33674fs	A-band	359;1.00	20RD01346	R[a]	Absent
179399346	c.101996G>A	p.Trp33999*	A-band	359;1.00	14CT01557	R[a]	Absent
179393848	c.106629delA	p.Ala35544fs	A-band	361;0.99	12PF00041	R[a]	Absent
179393000	c.107377+1G>A		A-band	362;1.00	20JM01785	R[a]	0.00001661

[a] This patient reported previously; H = Herman[2], R = Roberts[3], F = Fatkin[4], W = Walsh[5], C = Ceyhan-Birsoy[6].
MAF (ExAC), minor allele frequency in the Exome Aggregation Consortium database; PSI, proportion spliced-in.

Table S5. Characteristics of rare missense variants in *TTN* and 40 other cardiomyopathy genes.

Parameter	*TTN* gene			Other cardiomyopathy genes		
	DCM (n=287 variants)	**Control (n=254 variants)**	***P* value**	**DCM (n=375 variants)**	**Control (n=312 variants)**	***P* value**
Present in ClinVar	171 (60%)	154 (61%)	0.86	208 (55%)	179 (57%)	0.64
Novel (EA-ESP)	163 (57%)	141(56%)	0.79	235 (63%)	193 (62%)	0.87
Novel (NFE-ExAC)	96 (33%)	71 (28%)	0.19	150 (40%)	81 (26%)	0.0001
Novel (ExAC)	76 (27%)	50 (20%)	0.07	121 (32%)	58 (19%)	<0.0001
Deleterious (PolyPhen2/SIFT/PROVEAN)	100 (35%)	95 (37%)	0.59	134 (36%)	80 (26%)	0.005
Deleterious (MetaSVM)	49 (17%)	42 (17%)	0.91	152 (41%)	80 (26%)	<0.0001
Novel (ExAC) + deleterious (PolyPhen2/ SIFT/PROVEAN)	26 (9%)	18 (7%)	0.43	61 (16%)	16 (5%)	<0.0001
Novel (ExAC) + deleterious (MetaSVM)	16 (6%)	8 (3%)	0.21	71 (19%)	15 (5%)	<0.0001

EA-ESP, European American subset; ESP, Exome Sequencing Project; ExAC, Exome Aggregation Consortium; NFE, non-Finish European subset; PolyPhen2, SIFT, PROVEAN, MetaSVM are *in silico* variant prediction algorithms (see Methods).

Table S6. Distribution across RVIS percentiles of truncating and missense variants in 40 cardiomyopathy genes (*TTN* excluded).

	RVIS percentile				P value	Residuals				Standardized residuals			
	25th	50th	75th	**100th**	**(2x4)[a]**	**25th**	**50th**	**75th**	**100th**	**25th**	**50th**	**75th**	**100th**
All variants													
DCM	205	78	17	81	0.75	0.11	-0.60	0.10	0.39	0.24	-1.01	0.16	0.66
Control	163	73	13	60		-0.12	0.67	-0.11	-0.44	-0.24	1.01	-0.16	-0.66
Truncating variants													
DCM	16	0	4	4	0.01	1.05	-1.36	0.53	-1.06	2.43	-2.28	0.91	-2.03
Control	4	3	1	7		-1.33	1.72	-0.67	1.35	-2.43	2.28	-0.91	2.03
Missense variants													
DCM	189	78	13	77	0.67	-0.15	-0.36	-0.20	0.72	-0.33	-0.61	-0.30	1.19
Control	159	70	12	53		0.17	0.40	0.22	-0.79	0.33	0.61	0.30	-1.19
Missense variants, novel (ExAC)													
DCM	69	22	5	25	0.49	0.16	-0.65	-0.20	0.51	0.43	-1.31	-0.36	1.01
Control	30	15	3	8		-0.24	0.96	0.29	-0.76	-0.43	1.31	0.36	-1.01
Missense variants, deleterious (MetaSVM)													
DCM	94	19	7	28	0.29	-0.16	-0.03	1.11	-0.12	-0.47	-0.05	1.93	-0.23
Control	51	10	0	16		0.23	0.04	-1.54	0.17	0.47	0.05	-1.93	0.23
Missense variants, novel + deleterious													
DCM	47	8	4	12	0.15	0.24	-0.83	0.38	0.13	0.94	-2.17	0.94	0.34
Control	8	5	0	2		-0.51	1.81	-0.84	-0.28	-0.94	2.17	-0.94	-0.34

[a] 2 x 4 chi-square test was used to assess differences between cases and controls across different RVIS percentiles. To identify the cells that contribute most to the statistical significant results, raw residuals were obtained by subtracting expected from observed values and standardized residuals were calculated by dividing the raw residuals by the square root of the residual cell variance. The standardized residuals follow a standard normal distribution so any deviations from +2 or –2 corresponds to statistical significance and lack of fit of null hypothesis in that cell.

Table S7. Animal model evidence for roles of genes in DCM pathogenesis and list of key references supporting gene group allocations.

Gene	Protein	DCM in mouse model		Other species	Gene group	References
		Loss-of-function[a]	**Human NS mutation**			
MYH7	Myosin heavy chain 7	See *MYH6*[b]; no murine *myh7* KO	Adult KI[b] mice	Zebrafish embryo morphants Adult fly morphants	A	7,8[c],9-12
DSP	Desmoplakin	KO mice: EL (-/-), adult biV ARVC (+/-); cKO mice: adult biV ARVC (-/-)	Adult Tg mice (ARVC mutation)		A	13-19
SCN5A	Cardiac sodium channel	KO mice: EL (-/-); normal FS (+/-); adult Tg mice (80-90% ↓protein): DCM	Adult KI mice	Zebrafish embryo MO: developmental defects	A	20-29
LMNA	Lamin A/C	KO mice: NN DCM/death (-/-); adult DCM (+/-)	Adult KI mice	Zebrafish embryo MO	A	8[c],30-38
LDB3	LIM domain binding protein 3	KO mice: NN death (-/-); adult cKO mice (-/-)	Adult Tg mice	Zebrafish embryo MO: thin ventricular walls	A	39-44
DMD	Dystrophin	Adult mdx mice	NA	KO flies	A	45-53
DES	Desmin	Adult KO mice: DCM (-/-)	Adult KI mice	Zebrafish embryo MO: DCM + skeletal myopathy	A	38,54-63
TPM1	Tropomyosin α-1 chain	KO mice: EL (-/-), normal systolic function (+/-).	Adult Tg mice		A	64-71
TNNC1	Troponin C	KI mice (D73N, Ca^{2+} de-sensitizer): EL (-/-), adult DCM (+/-)	NA	Zebrafish embryo morphant: developmental defects, ↓contraction	A	66, 72-76
TNNT2	Troponin T	KO mice: DCM, NN death (-/-), normal FS (+/-).	Adult Tg/KI mice	Zebrafish embryo morphants/MO	A	7,8[c],38,72,77-83
BAG3	BCL-2 associated athanogene	KO mice: NN death (-/-)	Adult Tg mice	Zebrafish embryo MO: ↓contraction, PE	A	84-90

PLN	Phospholamban	Adult Tg mice (over-expression): DCM; KO mice: hypercontractile (-/-)	Adult Tg mice		A	8[c],91-101
RBM20	RNA-binding protein 20	NA	NA	Adult KO rats: LV dilation, normal FS	A	8[c],102-108
TTN	Titin	*TTN*tv Tg mice: EL (-/-), adult stress-induced DCM (+/-)	NA	Zebrafish embryo morphants: ↓contraction, PE (-/-), normal contraction (+/-); Adult *TTN*tv rats: stress-induced DCM	A	109-118
VCL	Vinculin	KO mice: EL (-/-); adult stress-induced DCM (+/-); cKO mice: adult DCM (-/-)	NA	Zebrafish embryo MO: DCM + skeletal myopathy	B	38,119-123
LAMA4	Laminin subunit α-4	Adult KO mice: LVH, late DCM (-/-)	NA	Zebrafish embryo MO: PE	B	124,125
ILK	Integrin-linked protein kinase	Adult KO mice (-/-); adult cKO mice (-/-)	NA	Zebrafish embryo MO: cardiac developmental defects; zebrafish embryo morphant: ↓contraction, PE	B	38,124,126-128
MYBPC3	Cardiac myosin binding protein 3	KO/KI mice: NN/adult DCM, LVH (-/-), normal adult FS, ± LVH (+/-)	NA	Zebrafish embryo MO: ventricular hypertrophy, normal contraction, PE	B	8[c],129-134
CSRP3	Muscle LIM protein	KO mice: LVH, NN death or adult DCM (-/-); no early phenotype (+/-)	Adult KI mice: LVH, ↓contractile reserve	Zebrafish embryo MO	B	38,135-138
CRYAB	αβ-crystallin	NA	Adult Tg mice: LVH, late DCM	Zebrafish embryo MO: ↓contraction, PE; NS mutant flies	B	139-145
ACTC1	Cardiac actin	KO mice: EL/NN death (-/-)	Adult Tg mice	Zebrafish embryo morphants: cardiac dilation/development	B	146-151

				defects, ↑contraction, PE		
NEXN	Nexilin	KO mice: DCM, EFE, NN death (-/-), normal adult FS (+/-)	NA	Zebrafish embryo MO/injected morphants	B	152,153
TAZ	Tafazzin	Adult iKO mice (-/-)	NA	Zebrafish embryo MO: developmental defects, PE	B	154-158
LAMP2	Lysosome-associated membrane glycoprotein 2	Adult KO mice: ↑ HW/BW, ↓myofibril force (-/-)	NA		B	159-162
TNNI3	Troponin I	KO mice: NN death, ↑ myocyte resting tension, ↓ Ca^{2+} sensitivity (-/-)	NA	Morphant flies	B	8[c],163-167
ANKRD1	Cardiac ankyrin repeat protein	Adult KO mice: normal FS (-/-)	NA		B	168-170
DSG2	Desmoglein-2	KO mice: EL (-/-), viable, no overt changes (+/-); cKO mice: adult RV dilation, histopathology (-/-)	Adult Tg mice (ARVC mutation)		B	14,171-175
MYH6	Myosin heavy chain 6	KO mice[b]: EL (-/-), adult ↓contraction (+/-)	NA	KO xenopus: developmental defects, ↓contraction	C	9,66,129,176-179
ABCC9	K_{ATP} channel SUR2A subunit	KO mice: DCM, NN death (-/-)	NA	KO flies: pacing-induced heart failure	C	180-182
ACTN2	α-actinin	NA	NA	Zebrafish embryo MO: developmental defects, ↓contraction	C	136,183-185
PDLIM3	PDZ and LIM domain protein 3	KO mice: adult DCM (RV>LV) (-/-)	NA		C	186,187
SGCD	δ - sarcoglycan	KO mice: skeletal myopathy, late (>50 weeks) DCM (-/-)	Adult KI mice: no DCM	KO Syrian hamster: skeletal myopathy + DCM/HCM; zebrafish embryo MO: skeletal myopathy + DCM	C	38,188-197

SYNE1	Nesprin-1	Adult KO mice: normal FS (-/-)	NA	Zebrafish MO: variable developmental defects	C	198-202
FXN	Frataxin	KO mice: EL (-/-); cKO mice: adult LVH/DCM (-/-)	NA	KO flies	C	203-209
CAV3	Caveolin-3	KO mice: no overt cardiac changes (-/-); Tg overexpression: adult DCM	NA		C	210-213
TCAP	Telethonin	Adult KO mice: normal baseline, post-TAC LVH, DCM (-/-), no change (+/-)	NA	Zebrafish embryo MO	C	38,135,214-216
FKTN	Fukutin	KO mice: EL (-/-)	NA		C	217-221
SDHA	Succinate dehydrogenase flavoprotein subunit	NA	NA		D	222,223
DTNA	α-dystrobrevin	Adult KO mice: no DCM, isoproterenol-induced ↑mortality (-/-)	NA		D	224-226
LAMA2	Laminin subunit α-2	KO mice: SKM (-/-)	NA		D	227-233
SYNM	Synemin	KO mice: SKM (-/-)	NA		D	234-237

[a] Most studies have evaluated homozygous gene loss-of-function models; data from heterozygous mice are shown when available.
[b] α-MHC is major myosin isoform in murine ventricle: human *MYH7* mutations modelled in murine *myh6* gene.
[c] Review article with reference list for genetic studies.
ARVC, arrhythmogenic right ventricular cardiomyopathy; biV, biventricular; cKO, cardiac-specific knockout; DCM, dilated cardiomyopathy; EFE, endocardial fibroelastosis; EL, embryonic lethal; FS, fractional shortening; HW/BW, heart weight: body weight ratio; iKO, inducible knockout; KASH, Klarsicht Anc-1 Syne Homology domain; KI, knock-in; KO, knockout; LV, left ventricle; LVH, left ventricular hypertrophy; MO, morpholino; NA, not available; NN, neonatal; NS, non-synonymous variant; PE, pericardial effusion; RV, right ventricle; SKM, skeletal myopathy; TAC, transverse aortic constriction; Tg, transgenic; *TTN*tv, truncating variants in the *TTN* (titin) gene.

Table S8. Distribution across gene groups of truncating and missense variants in 40 cardiomyopathy genes (*TTN* excluded).

	Gene group				P value	Residuals				Standardized residuals			
	A	B	C	D	(2x4)[a]	A	B	C	D	A	B	C	D
All variants													
DCM	162	78	99	61	0.002	1.99	-0.25	-1.48	-0.67	3.67	-0.41	-2.60	-1.09
Control	91	68	110	60		-2.19	0.27	1.63	0.74	-3.67	0.41	2.60	1.09
Truncating variants													
DCM	14	4	2	5	0.01	1.45	0.23	-1.27	-0.80	2.90	0.39	-2.21	-1.49
Control	2	2	6	7		-1.76	-0.28	1.53	0.97	-2.90	-0.39	2.21	1.49
All missense variants													
DCM	148	74	97	56	0.018	1.66	-0.29	-1.23	-0.46	3.05	-0.48	-2.16	-0.75
Control	89	66	104	53		-1.82	0.32	1.35	0.51	-3.05	0.48	2.16	0.75
Missense variants, novel (ExAC)													
DCM	76	14	25	6	0.0009	1.56	-1.13	-0.87	-0.94	3.98	-2.17	-1.76	-1.72
Control	18	14	19	7		-2.26	1.64	1.26	1.36	-3.98	2.17	1.76	1.72
Missense variants, deleterious (MetaSVM)													
DCM	85	24	27	16	0.0005	1.45	0.50	-1.61	-1.09	3.44	0.92	-3.16	-2.01
Control	26	9	29	16		-2.01	-0.69	2.23	1.51	-3.44	-0.92	3.16	2.01
Missense variants, novel + deleterious													
DCM	55	6	8	2	0.001	0.65	0.47	-1.05	-1.05	2.90	1.17	-2.74	-2.58
Control	6	0	6	3		-1.42	-1.02	2.28	2.28	-2.90	-1.17	2.74	2.58

[a] A 2 x 4 chi-square test was used to assess differences between cases and controls across different gene groups. To identify the cells that contribute most to the statistical significant results, raw residuals were obtained by subtracting expected from observed values and standardized residuals were calculated by dividing the raw residuals by the square root of the residual cell variance. The standardized residuals follow a standard normal distribution so any deviations from +2 or –2 corresponds to statistical significance and lack of fit of null hypothesis in that cell.

Table S9. Damaging[a] variants in group A genes identified in DCM cases.

Gene	Genomic position (Hg19)	Transcript effect	Protein effect	Protein domain	Variant carriers	Other studies (no. probands)	ClinVar[b]	ACMG[c]
MYH7	14:23901919	c.431G>T	p.Gly144Val	Motor	10EC00826	DCM (3)	Absent	LP
MYH7	14:23900157	c.848A>G	p.Tyr283Cys	Motor	DCM-24	HCM (1)	LP (NP)	LP
MYH7	14:23899092	c.1030G>A	p.Glu344Lys	Motor	10EH00627		Absent	VUS
MYH7	14:23899016	c.1106G>A	p.Arg369Gln	Motor	CZ-III-4	DCM (4)	LP (DCM)	P
MYH7	14:23898294	c.1277C>T	p.Ala426Val	Motor	DCM-2		Absent	VUS
MYH7	14:23898175	c.1396G>C	p.Glu466Gln	Motor switch II	DCM-23		LP (NP)	VUS
MYH7	14:23897714	c.1573G>A	p.Glu525Lys	Motor	DCMUKB5	DCM (1)	P (DCM)	P
MYH7	14:23897708	c.526_+1G>A		Motor	DF-III-1		Absent	VUS
MYH7	14:23894983	c.2207T>C	p.Ile736Thr	Motor, converter	BK-III-3	DCM (1), HCM (>10)	P (HCM)	LP
MYH7	14:23894584	c.2330G>A	p.Arg777Lys	Motor, converter	20JO01157		Absent	VUS
MYH7	14:23894500	c.2414T>C	p.Leu805Pro	Neck, IQ domain	20DH01138		Absent	VUS
MYH7	14:23894204	c.2453T>A	p.Ile818Asn	Neck	20RW01823		Absent	VUS
MYH7	14:23894037	c.2620G>C	p.Glu874Gln	Rod/hinge	DD-III-4		Absent	VUS
MYH7	14:23894036	c.2621A>T	p.Glu874Val	Rod/hinge	10LL01020		Absent	VUS
MYH7	14:23892760	c.3094delG	p.Asp1032fs	Rod/hinge	12RC00100		Absent	VUS
MYH7	14:23889113	c.3667G>A	p.Glu1223Lys	Hinge	MDU-111		VUS (DCM+)	VUS
MYH7	14:23885338	c.4828G>C	p.Glu1610Gln	Rod, LMM	BG-III-1		LP (NP)	VUS
MYH7	14:23884251	c.5512A>G	p.Lys1838Glu	Rod, LMM	MDJ_21		Absent	VUS
MYH7	14:23884207	c.5556C>G	p.Tyr1852*	Rod, LMM	12DB00017		VUS (NP)	VUS
MYH7	14:23883041	c.5717C>G	p.Ala1906Gly	Rod, LMM	DCM-26		LP (DCM)	VUS
RBM20	10:112544606	c.1486C>T	p.Gln496*		12PF00041		Absent	VUS
RBM20	10:112572055	c.1900C>T	p.Arg634Trp	RS domain	AB-III-2	DCM (1)	LP (NP)	LP
RBM20	10:112572061	c.1906C>T	p.Arg636Cys	RS domain	KS-II-1, MCH-123	DCM (1)	P (DCM)	P
RBM20	10:112572062	c.1907G>A	p.Arg636His	RS domain	MAK-13	DCM (4)	P (DCM)	P

RBM20	10:112572068	c.1913C>T	p.Pro638Leu	RS domain	MJ-11	DCM (4)	P (DCM)	P
RBM20	10:112572419	c.2264G>A	p.Arg755His		10IP00668		Absent	VUS
RBM20	10:112572526	c.2371C>T	p.Arg791Trp		HFA-71		VUS (DCM)	VUS
RBM20	10:112581100	c.2723T>C	p.Leu908Pro	Glutamate-rich	FR-I-1		VUS (NP)	VUS
RBM20	10:112581114	c.2737G>A	p.Glu913Lys	Glutamate-rich	20RJ01217	DCM (2)	P (DCM)	LP
RBM20	10:112581427	c.3051_3054del CTCC	p.Ser1018fs		10LS01497		Absent	VUS
RBM20	10:112590831	c.3464T>C	p.Val1155Ala		DCMUKB8		VUS (NP)	VUS
RBM20	10:112590924	c.3557A>G	p.His1186Arg	ZF, C2H2-type	10MG01841		Absent	VUS
LMNA	1:156084868	c.159A>C	p.Glu53Asp	Coil 1A	20JT01288		Absent	VUS
LMNA	1:156104287	c.607G>A	p.Glu203Lys	Coil 1B	12PH00095	DCM (2)	P (DCM)	P
LMNA	1:156104594	c.640-2A>G		Coil 1B	20BB01870		Absent	LP
LMNA	1:156104603	c.647G>A	p.Arg216His	Coil 1B	LZ-III-4		VUS (CMT)	VUS
LMNA	1:156104617	c.661C>T	p.Arg221Cys	Linker 2	10LW02013		VUS (DCM)	VUS
LMNA	1:156104622	c.667_668delGA	p.Glu223fs	Linker 2	20PF01735		Absent	LP
LMNA	1:156104629	c.673C>T	p.Arg225*	Linker 2	MIG111	DCM±CD(6)	P (DCM)	P
LMNA	1:156108455	c.1875C>A	p.Ser625Arg	Tail	14SD01620		Absent	VUS
DES	2:220283527	c.343C>A	p.Leu115Ile	Coil 1A	DCMUKC1		Absent	VUS
DES	2:220285304	c.823A>G	p.Arg275Gly	Coil 2A	14AM01556	DCM (1)	Absent	VUS
DES	2:220286231	c.1193T>C	p.Leu398Pro	Coil 2B	FK-II-1		Absent	VUS
DES	2:220286282	c.1244G>A	p.Arg415Gln	Tail	12JW00204		Absent	VUS
DES	2:220290442	c.1346A>C	p.Lys449Thr	Tail	FG-II-2	MFM (1)	P (MFM)	VUS
SCN5A	3:38622781	c.2869A>G	p.Asn957Asp	Cytoplasmic	20RJ01812		Absent	VUS
SCN5A	3:38616909	c.3545A>G	p.Asp1182Gly	Cytoplasmic	20CP01865		Absent	VUS
SCN5A	3:38595793	c.4790T>C	p.Val1597Ala	Transmembrane, S3, repeat IV	14AM01556		Absent	VUS
SCN5A	3:38592729	c.5134G>A	p.Gly1712Ser	Extracellular	20AB01158	BS (1)	NP (BS)	VUS
SCN5A	3:38591991	c.872C>T	p.Arg1958*	Cytoplasmic, C-terminus	C-II-9	DCM (1)	LP (NP)	VUS
BAG3	10:121429543	c.361C>T	p.Arg121*		BY-III-4		Absent	P

BAG3	10:121429544	c.362G>A	p.Arg121Gln		20JO01157		VUS (NP)	VUS
BAG3	10:121436309	c.1243_1244insCT	p.Pro416fs	PXXP domain	20PW01714		Absent	LP
BAG3	10:121436382	c.1316T>A	p.Val439Glu	BAG domain	MBG-121		Absent	VUS
TPM1	15:63351866	c.479G>A	p.Arg160His	cTNT binding	DCM-21		LP (LVNC)	VUS
TPM1	15:63351878	c.491A>G	p.Glu164Gly	cTNT binding	MDK-11		Absent	VUS
TPM1	15:63354462	c.688G>A	p.Asp230Asn		DCM-20	DCM (3)	P (DCM)	P
TPM1	15:63356332	c.842T>C	p.Met281Thr		SS200	HCM (5)	P (HCM)	VUS
TNNT2	1:201333470	c.415C>T	p.Arg139Cys	α-TM binding site	10AB00132	DCM (1)	LP (NP)	LP
TNNT2	1:201333464	c.421C>T	p.Arg141Trp	α-TM binding site	MGY-123	DCM (10), HCM (1)	P (LNVC, DCM) LP (HCM)	P
TNNT2	1:201331144	c.586C>T	p.Arg196Trp		10MP01703		Absent	VUS
DSP	6:7574464	c.2276_2277insA	p.Thr760fs	Globular, N-terminal	10AS00411	DCM (1)	Absent	LP
DSP	6:7575030	c.2436+2T>C		Globular, N-terminal	20GF01948		LP (NP)	LP
DSP	6:7579556	c.3133C>T	p.Arg1045*	Globular, N-terminal	10KS01707		P (ARVC)	P
PLN	6:118880109	c.25C>T	p.Arg9Cys	Cytoplasmic	20CP01865	DCM (2)	P (DCM)	P
PLN	6:118880125	c.41G>C	p.Arg14Thr	PKA motif	DCMUKB11		Absent	VUS
DMD	X: 32591954	c.1612G>T	p.Asp538Tyr	Spectrin 2 repeat	10RN00513		Absent	VUS
LDB3	10:88476516	c.1664A>G	p.Asn555Ser	LIM zinc-binding 1	10TM00933		Absent	VUS

[a] Damaging variants included truncating (trunc) or novel (ExAC) + deleterious (MetaSVM) missense variants in group A genes (*TTN* excluded).
[b] Variants present in the ClinVar database are classified as pathogenic (P) or likely-pathogenic (LP) in the context of specific disorders (shown in brackets). For some variants, the associated disorders were not provided (NP).
[c] Classification of variants according to the American College of Medical Genetics criteria:[238] pathogenic, likely-pathogenic, or variant of unknown significance (VUS).

α-TM, α-tropomyosin; ARVC, arrhytmogenic right ventricular cardiomyopathy; BS, Brugada syndrome; CD, conduction-system abnormalities; CMT, Charcot-Marie-Tooth disease; cTNT, cardiac troponin T; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; LMM, light meromyosin region; LVNC, left ventricular non-compaction; MFM, myofibrillar myopathy; PKA, protein kinase A; ZF, zinc finger.

Table S10. Co-segregation analysis of rare (MAF < 0.1%) missense variants in the 41-gene dataset in 14 DCM families.

Family	Gene	Variant	Gene group	MetaSVM	MAF (ExAc)	No. affected/ carriers[a]	LOD score[b]
AB	*RBM20*	p.Arg634Trp	A	Deleterious	Absent	6/6	1.46
	TTN	p.Ala21147Thr	A	Deleterious	0.0007573	2/3	-3.82
BG	*MYH7*	p.Glu1610Gln	A	Deleterious	Absent	6/9	-1.47
	TTN	p.Met1001Arg	A	Tolerated	0.0001072	6/6	-3.41
BK	*MYH7*	p.Ile736Thr	A	Deleterious	Absent	2/3	-0.19
CZ	*MYH7*	p.Arg369Gln	A	Deleterious	Absent	4/4	1.11
	TTN	p.Arg18037Trp	A	Tolerated	0.00004172	2/3	0.12
	TTN	p.Val205641Ile	A	Tolerated	0.000008327	4/5	-0.18
DO	*TTN*	p.Cys31725Arg	A	Deleterious	Absent	2/6	0.28
FG	*DES*	p.Lys449Thr	A	Deleterious	Absent	4/4	1.33
	SYNM	p.Ser1089Leu	D	Tolerated	0.0003569	1/4	-0.39
FK	*DES*	p.Leu398Pro	A	Deleterious	Absent	4/4	1.38
	ACTN2	p.Arg398Cys	C	Tolerated	0.00003295	1/2	0.68
	LDB3	p.Pro472Thr	A	Tolerated	Absent	3/5	-3.47
	SYNE1	p.Asn4590Asp	C	Tolerated	0.00009071	2/3	-2.94
	TTN	p.Arg34092His	A	Tolerated	0.00004973	1/3	0.6
FR	*RBM20*	p.Leu908Pro	A	Deleterious	Absent	3/6	0.40
	LAMA2	p.Arg998His	D	Tolerated	0.0001426	1/3	0.58
	TTN	p.Thr28241Ile	A	Tolerated	0.00008418	2/4	-1.11
GE	*DSP*	p.Gly2245Ala	A	Tolerated	0.000008364	1/1	-0.17
	TTN	p.Asn31858Ser	A	Tolerated	0.00003312	1/1	-0.17
GI	*TTN*	p.Lys28162Thr	A	Tolerated	Absent	3/4	-0.08
GX	*DMD*	p.Glu2976Lys	A	Tolerated	0.00002711	2/3	0
	RBM20	p.Pro706Thr	A	Tolerated	Absent	1/2	0
	SYNE1	p.Arg5832His	C	Tolerated	0.00001647	2/3	0
	TTN	p.Glu29772Lys	A	Tolerated	0.0002655	2/2	0
ID	*SYNE1*	p.Gln1185Lys	C	Tolerated	0.00002471	2/5	0.17

KS	*RBM20*	p.Arg636Cys	A	Deleterious	Absent	3/5	-1.11
	TTN	p.Gln28506Lys	A	Tolerated	0.0003342	2/2	-∞
Y	*NEB*	p.Asp1329Ala	D	Deleterious	0.0008056	3/6	-2.64

[a] Includes all affected individuals and unaffected individuals aged >40 years.
[b] LOD scores for each variant (theta = 0, penetrance 0.7) were based on all genotyped family members; pedigrees are shown in Supplementary Figures 4 and 5.

Table S11. Distribution of rare variants in DCM cases and controls[a] across different protein functional domains.

Protein domain	Protein location (aa)	No. variants		Odd ratio	Confidence interval	*P* value
		DCM cases	Controls			
MYH7 (β-myosin heavy chain)						
Myosin motor	85 - 778	42	108	3.35	2.24 - ∞	**2.35E-07**
Neck (IQ)	781 - 810	2	15	0.81	0.13 - ∞	0.72
Tail - Rod/hinge	838 - 1126	15	130	0.65	0.37 - ∞	0.95
Tail - Hinge	1127 - 1279	5	12	2.61	0.87- ∞	0.08
Tail – Light meromyosin (LMM)	1279 - 1935	30	303	0.43	0.29 - ∞	1.00
All domains	1-1935	100	609	NA	NA	NA
RBM20 (RNA-binding protein 20)						
Proline-rich	13 - 86	0	1	0	0 - ∞	1.00
Zinc finger	412 - 436	0	0	0	0 - ∞	1.00
RNA-recognition motif (RRM)	518 - 593	2	4	0.94	0.12 - ∞	0.68
Arg-Ser-rich (RS)	628 - 655	9	4	5.05	1.56 - ∞	**0.008**
Glutamate-rich	843 - 931	6	6	2.02	0.61 - ∞	0.20
Zinc finger	1159 - 1193	1	1	1.89	0.05 - ∞	0.58
All domains	1 - 1227	41	77	NA	NA	NA
LMNA (lamin A/C)						
Head	1 - 33	0	24	0	0 - ∞	1.00
Rod - coil 1A	34 - 70	4	0	∞	4.12 - ∞	**0.001**
Rod - linker 1	71 - 80	0	2	0	0 - ∞	1.00
Rod - coil 1B	81 - 218	9	22	1.98	0.86 - ∞	0.09
Rod- linker 2	219 - 242	6	10	2.86	0.98 - ∞	**0.05**
Rod - coil 2	243 - 383	20	40	3.04	1.64 - ∞	**0.001**
Tail	384 - 646	9	113	0.20	0.09 - ∞	1.00
All domains	1 - 646	48	211	NA	NA	NA

DES (desmin)						
Head	1 - 108	3	30	1.28	0.29 - ∞	0.47
Rod - coil 1A	109 - 141	1	11	1.12	0.05 - ∞	0.62
Rod - linker 1	142 - 151	1	0	∞	0.64 - ∞	0.08
Rod - coil 1B	152 - 252	0	61	0	0 - ∞	1.00
Rod - linker 12	253 - 268	0	0	0	0 - ∞	1.00
Rod - coil 2A	269 - 287	2	2	13.44	1.37 - ∞	**0.03**
Rod - linker 2	288 - 295	0	1	0	0 - ∞	1
Rod - coil 2B	296 - 412	5	57	1.11	0.36 - ∞	0.53
Tail	413 - 470	4	34	1.58	0.45 - ∞	0.32
All domains	1 - 470	16	196	NA	NA	NA
SCN5A (cardiac sodium channel protein 5)						
Cytoplasmic, N-terminus	1 - 126	0	60	0	0 - ∞	1.00
Repeat I	127 - 415	6	110	1.78	0.70 - ∞	0.16
S1, repeat I	127 - 150	0	1	0	0 - ∞	1.00
Extracellular	151 - 158	0	1	0	0 - ∞	1.00
S2, repeat I	159 - 178	0	25	0	0 - ∞	1.00
Cytoplasmic	179 - 191	0	28	0	0 - ∞	1.00
S3, repeat I	192 - 210	0	1	0	0 - ∞	1.00
Extracellular	211- 216	0	17	0	0 - ∞	1.00
S4, repeat I, voltage sensor	217 - 236	6	5	43.35	12.64 - ∞	**2.89E-07**
Cytoplasmic	237 - 252	0	3	0	0 - ∞	1.00
S5, repeat I	253 - 276	0	1	0	0 - ∞	1.00
Extracellular	277 - 389	0	28	0	0 - ∞	1.00
S6, repeat I	390 - 415	0	0	0	0 - ∞	1.00
Cytoplasmic	416 - 711	6	213	0.80	0.32 - ∞	0.75
Repeat II	712 - 939	1	36	0.82	0.04 - ∞	0.71
S1, repeat II	712 - 736	0	2	0	0 - ∞	1.00
Extracellular	737 - 747	0	2	0	0 - ∞	1.00

S2, repeat II	748 - 771	0	4	0	0 - ∞	1.00
Cytoplasmic	772 - 779	0	1	0	0 - ∞	1.00
S3, repeat II	780 - 799	0	9	0	0 - ∞	1.00
Extracellular	800 - 805	0	5	0	0 - ∞	1.00
S4, repeat II, voltage sensor	806 - 825	1	3	10.18	0.38 - ∞	0.12
Cytoplasmic	826 - 841	0	4	0	0 - ∞	1.00
S5, repeat II	842 - 862	0	0	0	0 - ∞	1.00
Extracellular	863 - 913	0	5	0	0 - ∞	1.00
S6, repeat II	914 - 939	0	1	0	0 - ∞	1.00
Cytoplasmic	940 - 1200	9	137	2.35	1.08 - ∞	**0.035**
Repeat III	1201 - 1470	2	148	0.36	0.06 - ∞	0.97
S1, repeat III	1201 - 1224	0	4	0	0 - ∞	1.00
Extracellular	1225 - 1237	0	0	0	0 - ∞	1.00
S2, repeat III	1238 - 1263	0	31	0	0 - ∞	1.00
Cytoplasmic	1264 - 1269	0	1	0	0 - ∞	1.00
S3, repeat III	1270 - 1291	0	18	0	0 - ∞	1.00
Extracellular	1292 - 1295	0	43	0	0 - ∞	1.00
S4, repeat III, voltage sensor	1296 - 1317	0	25	0	0 - ∞	1.00
Cytoplasmic	1318 - 1336	0	3	0	0 - ∞	1.00
S5, repeat III	1337 - 1359	1	11	2.77	0.13 - ∞	0.33
Extracellular	1360 - 1443	1	9	3.39	0.15 - ∞	0.28
S6, repeat III	1444 - 1470	0	3	0	0 - ∞	1.00
Cytoplasmic	1471 - 1523	0	17	0	0 - ∞	1.00
Repeat IV	1524 - 1772	2	55	1.09	0.18 - ∞	0.56
S1, repeat IV	1524 - 1547	0	3	0	0 - ∞	1.00
Extracellular	1548 - 1558	0	1	0	0 - ∞	1.00
S2, repeat IV	1559 - 1582	0	3	0	0 - ∞	1.00
Cytoplasmic	1583 - 1588	0	1	0	0 - ∞	1.00
S3, repeat IV	1589 - 1612	1	16	1.89	0.09 - ∞	0.43
Extracellular	1613 - 1622	0	1	0	0 - ∞	1.00
S4, repeat IV, voltage sensor	1623 - 1644	0	10	0	0 - ∞	1.00

Cytoplasmic	1645 - 1659	0	2	0	0 - ∞	1.00
S5, repeat IV	1660 - 1682	0	9	0	0 - ∞	1.00
Extracellular	1683 - 1747	1	7	4.36	0.19 - ∞	0.23
S6, repeat IV	1748 - 1772	0	2	0	0 - ∞	1.00
Cytoplasmic	1773 - 2017	5	150	0.99	0.36 - ∞	0.58
All domains	1 - 2017	31	926	NA	NA	NA
BAG3 (BAG3)						
WW1	20 - 54	0	9	0	0 - ∞	1.00
IPV	87 -101	0	3	0	0 - ∞	1.00
WW2	124 - 154	0	15	0	0 - ∞	1.00
IPV	200 - 213	1	5	14.57	0.54 - ∞	0.09
PXXP	302 - 417	2	80	1.87	0.25 - ∞	0.38
BAG	421 - 498	1	14	5.17	0.21- ∞	0.21
All domains	1 - 575	6	379	NA	NA	NA
TPM1 (α-tropomyosin)						
Cardiac troponin T binding	150 - 180	2	1	5.42	0.41 - ∞	0.19
All domains	1 - 284	24	62	NA	NA	NA
TNNT2 (cardiac troponin T)						
α-tropomyosin binding	78 - 171	15	33	8.96	3.76 - ∞	**2.85E-06**
Troponin C/I/α-tropomyosin binding	198 - 289	4	93	0.23	0.07 - ∞	1.00
All domains	2 - 298	23	193	NA	NA	NA
DSP (desmoplakin)						
Globular 1	1 - 1056	20	536	1.05	0.62 - ∞	0.48
Plakophilin/plakoglobin binding	1 - 584	15	269	1.78	0.99 - ∞	**0.05**
Rod	1057 - 1945	18	503	0.99	0.57 - ∞	0.57
All Globular 2	1946 - 2871	12	345	0.95	0.50 - ∞	0.61

Globular 2 - domain A	1960 - 2208	0	96	0	0 - ∞	1.00
Globular 2 – domain B	2244 - 2446	3	89	0.93	0.24 - ∞	0.63
Globular 2 – domain C	2609 - 2822	3	73	1.15	0.30 - ∞	0.50
Globular 2 - SR	2824 - 2847	2	14	4.07	0.64 - ∞	0.10
All domains	1 - 2871	50	1384	NA	NA	NA
PLN (phospholamban)						
Cytoplasmic	1 - 31	4	6	0	0 - ∞	1.00
HAX binding	16 - 22	1	3	0.37	0.01 - ∞	0.93
Helical	32 - 52	0	0	0	0 - ∞	1.00
All domains	1 - 52	4	6	NA	NA	NA
DMD (dystrophin)						
Actin binding - CH1	15 - 119	1	16	2.53	0.12 - ∞	0.35
Actin binding - CH2	134 - 237	1	41	0.97	0.05 - ∞	0.65
Spectrin 1	339 - 447	2	88	0.90	0.15 - ∞	0.66
Spectrin 2	448 - 556	6	23	12.82	4.69 - ∞	**4.04E-05**
Spectrin 3	559 - 667	1	58	0.67	0.03 - ∞	0.78
Spectrin 4	719 - 828	2	47	1.74	0.29 - ∞	0.34
Spectrin 5	830 - 934	0	36	0	0 - ∞	1.00
Spectrin 6	943 - 1045	1	45	0.88	0.04 - ∞	0.69
Spectrin 7	1048 - 1154	2	38	2.17	0.36 - ∞	0.26
Spectrin 8	1157 - 1263	1	9	4.53	0.20 - ∞	0.22
Spectrin 9	1266 - 1367	3	42	3.05	0.75 - ∞	0.09
Spectrin 10	1368 - 1463	1	39	1.02	0.05 - ∞	0.64
Spectrin 11	1468 - 1568	2	18	4.66	0.74 - ∞	0.08
Spectrin 12	1571 - 1676	1	60	0.65	0.03 - ∞	0.79
Spectrin 13	1679 - 1778	0	21	0	0 - ∞	1.00
Spectrin 14	1779 - 1874	0	37	0	0 - ∞	1.00
Spectrin 15	1877 - 1979	0	48	0	0 - ∞	1.00
Spectrin 16	1992 - 2101	0	34	0	0 - ∞	1.00

Spectrin 17	2104 - 2208	1	61	0.64	0.03 - ∞	0.80
Spectrin 18	2211 - 2318	0	17	0	0- ∞	1.00
Spectrin 19	2319 - 2423	0	29	0	0 - ∞	1.00
Spectrin 20	2475 - 2577	0	20	0	0 - ∞	1.00
Spectrin 21	2580 - 2686	0	27	0	0 - ∞	1.00
Spectrin 22	2689 - 2802	0	36	0	0 - ∞	1.00
Spectrin 23	2808 - 2930	0	16	0	0 - ∞	1.00
Spectrin 24	2935 - 3040	1	46	0.86	0.04 - ∞	0.70
WW	3055 - 3088	0	1	0	0 - ∞	1.00
Zinc finger	3307 - 3354	0	1	0	0 - ∞	1.00
All domains	1 - 3685	28	1112	NA	NA	NA
LDB3 (LIM domain-binding protein 3)						
PDZ	1-85	0	63	0	0 - ∞	1
LIM zinc-binding 1	549-607	2	40	2.57	0.41 - ∞	0.21
LIM zinc-binding 2	608-667	3	59	2.71	0.66 - ∞	0.13
LIM zinc-binding 3	668-727	1	48	1.00	0.05 - ∞	0.64
All domains	1 - 727	19	914	NA	NA	NA

[a] DCM cases include this study (n=532), familial DCM replication cohort (n=101) and patients evaluated in two clinical diagnostic laboratories (reported by Walsh et al.[5]), population control variants were derived from the ExAC database.

Supplementary Acknowledgments

We thank the Medical Genome Reference Bank (MGRB) for providing healthy elderly control sequencing data. The MGRB is comprised of whole-genome sequencing data and phenotypic information from up to 4000 healthy Australians over 70 years of age. MGRB participants, consented through contributing studies, 45 and UP (Sax Institute, Sydney) and the ASPirin in Reducing Events in the Elderly (ASPREE) clinical trial (Monash University, Melbourne) are free from cardiovascular disease, degenerative neurological disorders and of a history of cancer at the time of consent into the study.

ASPREE Acknowledgements

Authors would like to acknowledge the ASPREE Healthy Ageing Biobank, ASPREE Investigator Group and ASPREE Collaborating Practitioners listed on www.aspree.org, as well as ASPREE representatives John McNeil, Robyn Woods and Paul Lacaze for their contributions to the MGRB project. ASPREE was funded by the National Institute on Aging and the National Cancer Institute at the National Institutes of Health (grant number U01AG029824); the National Health and Medical Research Council of Australia (grant numbers 334047, 1127060); Monash University (Australia). The ASPREE Healthy Ageing Biobank was supported by the Commonwealth Scientific and Industrial Research Organisation (Australia), the Victorian Cancer Agency (Australia) and Monash University (Australia). Authors acknowledge the dedicated and skilled staff in Australia and the U.S. for the conduct of the ASPREE trial, and the ASPREE participants who willingly volunteered.

45 and Up Acknowledgements

This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au). The 45 and Up Study is managed by the Sax Institute in

collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; NSW Government Family & Community Services – Ageing, Carers and the Disability Council NSW; and the Australian Red Cross Blood Service. We thank the many thousands of people participating in the 45 and Up Study.

Supplementary References

1. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904-909.
2. Herman DS, Lam L, Taylor MR, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012;366:619-628.
3. Roberts AM, Ware JS, Herman DS, et al. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci Transl Med* 2015;7:270ra6.
4. Fatkin D, Lam L, Herman DS, et al. Titin truncating mutations: a rare cause of dilated cardiomyopathy in the young. *Prog Ped Cardiol* 2016;40:41-45.
5. Walsh R, Thomson KL, Ware JS, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet Med* 2017;19:192-203.
6. Ceyhan-Birsoy O, Agrawal PB, Hidalgo C, et al. Recessive truncating titin gene, *TTN*, mutations presenting as centronuclear myopathy. *Neurology* 2013;81:1205-1214.
7. Kamisago M, Sharma SD, DePalma SR, et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* 2000;343:1686-1696.
8. Kayvanpour E, Sedaghat-Hamedani F, Amr A, et al. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin Res Cardiol* 2017;106:127-139.
9. Jones KW, Grupp IL, Doetschman T, et al. Ablation of the murine alpha myosin heavy chain gene leads to dosage effects and functional deficits in the heart. *J Clin Invest* 1996;98:1906-17.

10. Schmidt JP, Debold EP, Ahmad F, et al. Cardiac missense mutations cause dilated cardiomyopathy in mouse models and depress motor function. *Proc Natl Acad Sci USA* 2006;103:14525-14530.

11. Auman HJ, Coleman H, Riley HE, Olale F, Rsai HJ, Yelon D. Functional modulation of cardiac form through regionally confined cell shape changes. *PLoS Biol* 2007;5:e53.

12. Cammarato A, Dambacher CM, Knowles AF, et al. Myosin transducer mutations differentially affect motor function, myofibril structure, and the performance of skeletal and cardiac muscles. *Mol Biol Cell* 2008;19:553-562.

13. Elliott P, O'Mahony C, Syrris P, et al. Prevalence of desmosomal protein gene mutations in patients with dilated cardiomyopathy. *Circ Cardiovasc Genet* 2010;4:314-322.

14. Garcia-Pavia P, Syrris P, Salas C, et al. Desmosomal protein gene mutations in patients with idiopathic dilated cardiomyopathy undergoing cardiac transplantation: a clinicopathological study. *Heart* 2011;97:1744-1752.

15. Lopez-Avala JM, Gomez-Milanes I, Sanchez Munoz JJ, et al. Desmoplakin truncations and arrhythmogenic left ventricular cardiomyopathy: characterizing a phenotype. *Europace* 2014;16:1838-1846.

16. Gallicano GI, Bauer C, Fuchs E. Rescuing desmoplakin function in an extra-embryonic ectoderm reveals the importance of this protein in embryonic heart, neuroepithelium, skin, and vasculature. *Development* 2001;128:929-941.

17. Garcia-Gras E, Lombardi R, Giocondo MJ, et al. Suppression of canonical Wnt/beta-catenin signalling by nuclear plakoglobin recapitulates phenotype of arrhythmogenic right ventricular cardiomyopathy. *J Clin Invest* 2006;116:2012-2021.

18. Lyon RC, Mezzano V, Wright AT, et al. Connexin defects underlie arrhythmogenic right ventricular cardiomyopathy in a novel mouse model. *Hum Mol Genet* 2014;23:1134-1150.

19. Yang Z, Bowles NE, Scherer SE, et al. Desmosomal dysfunction due to mutations in desmoplakin causes arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Res* 2006;99:646-655.

20. McNair WP, Ku L, Taylor MR, et al. *SCN5A* mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* 2004;110:2163-2167.

21. Olson TM, Michels VV, Ballew JD, et al. Sodium channel mutations and susceptibility to heart failure and atrial fibrillation. *JAMA* 2005;293:447-454.

22. Ge J, Sun A, Paajanen V, et al. Molecular and clinical characterization of a novel *SCN5A* mutation associated with atrioventricular block and dilated cardiomyopathy. *Circ Arrhythm Electrophysiol* 2008;1:83-92.

23. McNair WP, Sinagra G, Taylor MR, et al. *SCN5A* mutations associate with arrhythmic dilated cardiomyopathy and commonly localize to the voltage-sensing mechanism. *J Am Coll Cardiol* 2011;57:2160-2168.

24. Mann SA, Castro ML, Ohanian M, et al. R222A *SCN5A* mutation is associated with reversible ventricular ectopy and dilated cardiomyopathy. *J Am Coll Cardiol* 2012;60:1566-1573.

25. Papadatos GA, Wallerstein PM, Head CE, et al. Slowed conduction and ventricular tachycardia after targeted disruption of the cardiac sodium channel gene Scn5a. *Proc Natl Acad Sci* USA 2002;99:6210-6215.

26. Royer A, van Veen TA, Le Bouter S, et al. Mouse model of *SCN5A*-linked hereditary Lenegre's disease: age-related conduction slowing and myocardial fibrosis. *Circulation* 2005;111:1738-1746.

27. Hesse M, Kondo CS, Clark RB, et al. Dilated cardiomyopathy is associated with reduced expression of the cardiac sodium channel Scn5a. *Cardiovasc Res* 2007;75:498-509.

28. Watanabe H, Yang T, Stroud DM, et al. Striking in vivo phenotype of a disease-associated human *SCN5A* mutation producing minimal changes in vitro. *Circulation* 2011;124:1001-1011.

29. Bennett JS, Stroud DM, Becker JR, Roden DM. Proliferation of embryonic cardiomyocytes in zebrafish requires the sodium channel scn5Lab. *Genesis* 2013;51:562-574.

30. Fatkin D, MacRae CA, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999;341:1715-1724.

31. Muchir A, Medioni J, Laluc M, et al. Nuclear envelope alterations in fibroblasts from patients with muscular dystrophy, cardiomyopathy, and partial lipodystrophy carrying lamin A/C mutations. *Muscle Nerve* 2005;30:444-450.

32. Raharjo WH, Enarson P, Sullivan T, Stewart CL, Burke B. Nuclear envelope defects associated with *LMNA* mutations cause dilated cardiomyopathy and Emery-Dreifuss muscular dystrophy. *J Cell Sci* 2001;114:4447-4457.

33. Cowan J, Li D, Gonzalez-Quintana J, Morales A, Hershberger RE. Morphological analysis of 13 *LMNA* variants identified in a cohort of 324 unrelated patients with idiopathic or familial dilated cardiomyopathy. *Circ Cardiovasc Genet* 2010;3:6-14.

34. Nikolova V, Leimena C, McMahon AC, et al. Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J Clin Invest* 2004;113:357-369.

35. Chandar S, Yeo LS, Leimena C, et al. Effects of mechanical stress and carvedilol in lamin A/C-deficient dilated cardiomyopathy. *Circ Res* 2010;106:573-582.

36. Mounkes LC, Kozlov SV, Rottman JN, Stewart CL. Expression of an LMNA-N195K variant of A-type lamins results in cardiac conduction defects and death in mice. *Hum Mol Genet* 2005;14:2167-2180.

37. Arimura T, Helbling-Leclerc A, Massart C, et al. Mouse model carrying H222P-Lmna mutation develops muscular dystrophy and dilated cardiomyopathy similar to human striated muscle laminopathies. *Hum Mol Genet* 2005;14:155-169.

38. Vogel B, Meder B, Just S, et al. *In-vivo* characterization of human dilated cardiomyopathy genes in zebrafish. *Biochem Biophys Res Commun* 2009;390:516-522.

39. Vatta M, Mohapatra B, Jimenez S, et al. Mutations in Cypher/ZASP in patients with dilated cardiomyopathy and left ventricular non-compaction. *J Am Coll Cardiol* 2003;42:2014-2027.

40. Arimura T, Hayashi T, Terada H, et al. A Cypher/ZASP mutation associated with dilated cardiomyopathy alters the binding affinity to protein kinase C. *J Biol Chem* 2004;279:6746-6752.

41. Zhou Q, Chu PH, Huang C, et al. Ablation of Cypher, a PDZ-LIM domain Z-line protein, causes a severe form of congenital myopathy. *J Cell Biol* 2001;155:605-612.

42. Zheng M, Cheng H, Li X, et al. Cardiac-specific ablation of Cypher leads to a severe form of dilated cardiomyopathy with premature death. *Hum Mol Genet* 2009;18:701-713.

43. Li Z, Ai T, Samani K, et al. A ZASP missense mutation, S196L, leads to cytoskeletal and electrical abnormalities in a mouse model of cardiomyopathy. *Circ Arrhythm Electrophysiol* 2010;3:646-656.

44. Van der Meer DL, Marques IJ, Leito JT, et al. Zebrafish cypher is important for somite formation and heart development. *Dev Biol* 2006;299:356-372.

45. Berko BA, Swift M. X-linked dilated cardiomyopathy. *N Engl J Med* 1987;316:1186-1191.

46. Muntoni F, Cau M, Ganau A, et al. Deletion of the dystrophin muscle-promoter region associated with X-linked dilated cardiomyopathy. *N Engl J Med* 1993;329:921-925.

47. Feng J, Yan J, Buzin CH, Sommer SS, Towbin JA. Comprehensive mutation scanning of the dystrophin gene in patients with non-syndromic X-linked dilated cardiomyopathy. *J Am Coll Cardiol* 2002;40:1120-1124.

48. Feng J, Yan J, Buzin CH, Towbin JA, Sommer SS. Mutations in the dystrophin gene are associated with sporadic dilated cardiomyopathy. Mol Genet Metabol 2002;77:119-26.

49. Diegoli M, Grasso M, Favalli V, et al. Diagnostic work-up and risk stratification in X-linked dilated cardiomyopathies caused by dystrophin defects. *J Am Coll Cardiol* 2011;58:925-934.

50. Henderson DM, Lee A, Ervasti JM. Disease-causing missense mutations in actin binding domain 1 of dystrophin induce thermodynamic instability and protein aggregation. *Proc Natl Acad Sci USA* 2010;107:9632-9637.

51. Singh SM, Bandi S, Shah DD, Armstrong G, Mallela KM. Missense mutation Lys18Asn in dystrophin that triggers X-linked dilated cardiomyopathy decreases protein stability, increases protein unfolding, and perturbs protein structure, but does not affect protein function. *PLoS One* 2014;9:e110439.

52. Bostick B, Yue Y, Long C, Duan D. Prevention of dystrophin-deficient cardiomyopathy in twenty-one month-old carrier mice by mosaic dystrophin expression or complementary dystrophin/utrophin expression. *Circ Res* 2008;102:121-130.

53. Taghli-Lamallem O, Akasaka T, Hogg G, et al. Dystrophin deficiency in Drosophila reduces lifespan and causes a dilated cardiomyopathy phenotype. *Aging Cell* 2008;7:237-249.

54. Li D, Tapscoft T, Gonzalez O, et al. Desmin mutation responsible for idiopathic dilated cardiomyopathy. *Circulation* 1999;100:461-464.

55. Miyamoto Y, Akita H, Shiga N, et al. Frequency and clinical characteristics of dilated cardiomyopathy caused by desmin gene mutations. *Eur Heart J* 2001;22:2284-2289.

56. Mavroidis M, Panagopoulou P, Kostavasili I, Weisleder N, Capetanaki Y. A missense mutation in desmin tail domain linked to human dilated cardiomyopathy promotes cleavage of the head domain and abolishes its Z-disc location. *FASEB J* 2008;22:3318-3327.

57. Taylor MR, Slavov D, Ku L, et al. Prevalence of desmin mutations in dilated cardiomyopathy. *Circulation* 2007;115:1244-1251.

58. Bergman JE, Veenstra-Knol HE, van Essen AJ, et al. Two related Dutch families with a clinically variable presentation of the cardioskeletal myopathy caused by a novel S13F mutation in the desmin gene. *Eur J Med Genet* 2007;50:355-366.

59. Van Spaendonck-Zwarts KY, van der Kooi AJ, van den Berg MP, et al. Recurrent and founder mutations in the Netherlands: the cardiac phenotype of *DES* founder mutations p.S13F and p.N342D. *Neth Heart J* 2012;20:219-228.

60. Tse HF, Ho JC, Choi SW, et al. Patient-specific induced-pluripotent stem cells-derived cardiomyocytes recapitulate the pathogenic phenotypes of dilated cardiomyopathy due

to a novel *DES* mutation identified by whole exome sequencing. *Hum Mol Genet* 2013;22:1395-1403.

61. Brodehl A, Dieding M, Biere N, et al. Functional characterization of the novel *DES* mutation p.L136P associated with dilated cardiomyopathy reveals a dominant filament assembly defect. *J Mol Cell Cardiol* 2016;91:207-214.
62. Milner DJ, Taffet GE, Wang X, et al. The absence of desmin leads to cardiomyocyte hypertrophy and cardiac dilation with compromised systolic function. *J Mol Cell Cardiol* 1999;31:2063-2076.
63. Clemen CS, Stockigt F, Strucksberg KH, et al. The toxic effect of R350P mutant desmin in striated muscle of man and mouse. *Acta Neuropathol* 2015;129:297-315.
64. Olson TM, Kishimoto NY, Whitby FG, Michels VV. Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *J Mol Cell Cardiol* 2001;33:723-732.
65. Lakdawala NK, Dellefave L, Redwood CS, et al. Familial dilated cardiomyopathy caused by an alpha-tropomyosin mutation: the distinctive natural history of sarcomeric dilated cardiomyopathy. *J Am Coll Cardiol* 2010;55:320-329.
66. Hershberger RE, Norton N, Morales A, Li D, Siegfried JD, Gonzalez-Quintana J. Coding sequence rare variants identified in *MYBPC3*, *MYH6*, *TPM1*, *TNNC1*, and *TNNI3* from 312 patients with familial or idiopathic dilated cardiomyopathy. *Circ Cardiovasc Genet* 2010;3:155-161.
67. Van de Meerakker JB, Christiaans I, Barnett P, et al. A novel alpha-tropomyosin mutation associates with dilated and non-compaction cardiomyopathy and diminishes actin binding. *Biochim Biophys Acta* 2013;1833:833-839.

68. Borovikov YS, Karpicheva OE, Avrova SV, Robinson P, Redwood CS. The effect of the dilated cardiomyopathy-causing mutation Glu54Lys of alpha-tropomyosin in actin-myosin interactions during the ATPase cycle. *Arch Biochem Biophys* 2009;489:20-24.

69. Chang AN, Greenfield NJ, Singh A, Potter JD, Pinto JR. Structural and protein interaction effects of hypertrophic and dilated cardiomyopathy mutations in alpha-tropomyosin. *Front Physiol* 2014;5:460.

70. Rethinasamy P, Muthuchamy M, Hewett T, et al. Molecular and physiological effects of alpha-tropomyosin ablation in the mouse. *Circ Res* 1998;82:116-123.

71. Rajan S, Ahmed RP, Jagatheesan G, et al. Dilated cardiomyopathy mutant tropomyosin mice develop cardiac dysfunction with significantly decreased fractional shortening and myofilament calcium sensitivity. *Circ Res* 2007;101:205-214.

72. Mogensen J, Murphy RT, Shaw T, et al. Severe disease expression of cardiac troponin C and T mutations in patients with idiopathic dilated cardiomyopathy. *J Am Coll Cardiol* 2004;44:2033-2040.

73. Van Spaendonck-Zwarts KY, van Tintelen JP, van Veldhuisen DJ, et al. Peripartum cardiomyopathy as part of familial dilated cardiomyopathy. *Circulation* 2010;121:2169-2175.

74. Pinto JR, Siegfried JD, Parvatiyar MS, et al. Functional characterization of *TNNC1* rare variants identified in dilated cardiomyopathy. *J Biol Chem* 2011;286:34404-34412.

75. McConnell BK, Singh S, Fan Q, et al. Knock-in mice harboring a Ca^{2+} desensitizing mutation in cardiac troponin C develop early onset dilated cardiomyopathy. *Front Physiol* 2015;6:242.

76. Sogah VM, Serluca FC, Fishman MC, Yelon DL, MacRae CA, Mably JD. Distinct troponin C isoform requirements in cardiac and skeletal muscle *Dev Dyn* 2010;239:3115-3123.

77. Campbell N, Sinagra G, Jones KL, et al. Whole exome sequencing identifies a troponin T mutation hot spot in familial dilated cardiomyopathy. *PLoS One* 2013;8:e78104.

78. Li D, Czernuszewicz GZ, Gonzalez O, et al. Novel cardiac troponin T mutation as a cause of familial dilated cardiomyopathy. *Circulation* 2001;104:2188-2193.

79. Lombardi R, Bell A, Senthil V, et al. Differential interactions of thin filament proteins in two cardiac troponin T mouse models of hypertrophic and dilated cardiomyopathies. *Cardiovasc Res* 2008;79:109-117.

80. Ahmad F, Banerjee SK, Lage ML, et al. The role of cardiac troponin T quantity and function in cardiac development and dilated cardiomyopathy. *PLoS One* 2008;3:e2642.

81. Nishii K, Morimoto S, Minakami R, et al. Targeted disruption of the cardiac troponin T gene causes sarcomere disassembly and defects in heartbeat within the early mouse embryo. *Dev Biol* 2008;322:65-73.

82. Ramratnam M, Salama G, Sharma RK, et al. Gene-targeted mice with the human troponin T R141W mutation develop dilated cardiomyopathy with calcium desensitization. *PLoS One* 2016;11:e016781.

83. Sehnert AJ, Huq A, Weinstein BM, Walker C, Fishman M, Stainier DY. Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nat Genet* 2002;31:106-110.

84. Norton N, Li D, Rieder MJ, et al. Genome-wide studies of copy number variation and exome-sequencing identify rare variants in *BAG3* as a cause of dilated cardiomyopathy. *Am J Hum Genet* 2011;88:273-282.

85. Villard E, Perret C, Gary F, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J* 2011;32:1065-1076.

86. Franaszczyk M, Bilinska ZT, Sobieszczanska-Malek M, et al. The *BAG3* gene variants in Polish patients with dilated cardiomyopathy: four novel mutations and a genotype-phenotype correlation. *J Transl Med* 2014;12:192.

87. Feldman AM, Begay RL, Knezevic T, et al. Decreased levels of BAG3 in a family with a rare variant and in idiopathic dilated cardiomyopathy. *J Cell Physiol* 2014;229:1697-1702.

88. Arimura T, Ishikawa T, Nunoda S, Kawai S, Kimura A. Dilated cardiomyopathy-associated *BAG3* mutations impair Z-disc assembly and enhance sensitivity to apoptosis in cardiomyocytes. *Hum Mutat* 2011;32:1481-1491.

89. Homma S, Iwasaki M, Shelton GD, Engvall E, Reed JC, Takayama S. BAG3 deficiency results in fulminant myopathy and early lethality. *Am J Pathol* 2006;169:761-773.

90. Quintana MT, Parry TL, He J, et al. Cardiomyocyte-specific human Bcl2-associated athanogene 3 P209L expression induces mitochondrial fragmentation, Bcl2-associated athanogene 3 haploinsufficiency and actives p38 signaling. *Am J Pathol* 2016;186:1989-2007.

91. Schmitt JP, Kamisago M, Asahi M, et al. Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* 2003;299:1410-1413.

92. Haghighi K, Kolokathis F, Pater L, et al. Human phospholamban null results in dilated cardiomyopathy revealing a critical difference between mouse and human. *J Clin Invest* 2003;111:869-876.

93. Haghighi K, Kolokathis F, Gramolini AO, et al. A mutation in the human phospholamban gene, deleting arginine 14, results in lethal, hereditary cardiomyopathy. *Proc Natl Acad Sci USA* 2006;103:1388-1393.

94. DeWitt MM, MacLeod HM, Soliven B, McNally EM. Phospholamban R14 deletion results in late-onset, mild, hereditary dilated cardiomyopathy. *J Am Coll Cardiol* 2006;48:1396-1398.

95. Posch MG, Perrot A, Geier C, et al. Genetic deletion of arginine 14 in phospholamban causes dilated cardiomyopathy with attenuated electrocardiographic R amplitudes. *Heart Rhythm* 2009;6:480-486.

96. Medeiros A, Biagi DG, Sobreira TJ, et al. Mutations in the human phospholamban gene in patients with heart failure. *Am Heart J* 2011;162:1088-1095.

97. Van der Zwaag PA, van Rijsingen IA, de Ruiter R, et al. Recurrent and founder mutations in the Netherlands – Phospholamban p.Arg14del mutation causes arrhythmogenic cardiomyopathy. *Neth Heart J* 2013;21:286-293.

98. Haghighi K, Pritchard T, Bossuyt J, et al. The human phospholamban Arg14-deletion mutant localizes to plasma membrane and interacts with the Na/K-ATPase. *J Mol Cell Cardiol* 2012;52:773-782.

99. Ceholski DK, Trieber CA, Young HS. Hydrophobic imbalance in the cytoplasmic domain of phospholamban is a determinant for lethal dilated cardiomyopathy. *J Biol Chem* 2012;287:16521-16529.

100. Luo W, Grupp IL, Harrer J, et al. Targeted ablation of the phospholamban gene is associated with markedly enhanced myocardial contractility and loss of beta-agonist stimulation. *Circ Res* 1994;75:401-409.

101. Kadambi VJ, Ponniah S, Harrer JM, et al. Cardiac-specific overexpression of phospholamban alters calcium kinetics and resultant cardiomyocyte mechanics in transgenic mice. *J Clin Invest* 1996;97:533-539.

102. Brauch KM, Karst ML, Herron KJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol* 2009;54:930-941.

103. Li D, Morales A, Gonzalez-Quintana J, et al. Identification of novel mutations in *RBM20* in patients with dilated cardiomyopathy. *Clin Transl Sci* 2010;3:90-97.

104. Wells QS, Becker JR, Su YR, et al. Whole exome sequencing identifies a causal *RBM20* mutation in a large pedigree with familial dilated cardiomyopathy. *Circ Cardiovasc Genet* 2013;6:317-326.

105. Beqquali A, Bollen IA, Rasmussen TB, et al. A mutation in the glutamate-rich region of *RBM20* causes dilated cardiomyopathy through missplicing of titin and improved Frank-Starling mechanism. *Cardiovasc Res* 2016;112:452-463.

106. Guo W, Schafer S, Greaser ML, et al. *RBM20*, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med* 2012;18:766-773.

107. Beraldi R, Li X, Martinez Fernandez A, et al. Rbm20-deficient cardiogenesis reveals early disruption of RNA processing and sarcomere remodelling establishing a developmental etiology for dilated cardiomyopathy. *Hum Mol Genet* 2014;23:3779-3791.

108. Wyles SP, Li X, Hrstka SC, et al. Modeling structural and functional deficiencies of *RBM20* familial dilated cardiomyopathy using human induced pluripotent stem cells. *Hum Mol Genet* 2016;25:254-265.

109. Gerull B, Gramlich M, Atherton J, et al. Mutations of *TTN*, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat Genet* 2002;30:201-204.

110. Gerull B, Atherton J, Geupel A, et al. Identification of a novel frameshift mutation in the giant muscle filament titin in a large Australian family with dilated cardiomyopathy. *J Mol Med (Berl)* 2006;84:478-483.

111. Herman DS, Lam L, Taylor MR, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012;366:619-628.

112. Gramlich M, Michely B, Krohne C, et al. Stress-induced dilated cardiomyopathy in a knock-in mouse model mimicking human titin-based disease. *J Mol Cell Cardiol* 2009;47:352-358.

113. Zhou Q, Kesteven S, Wu J, et al. Pressure overload by transverse aortic constriction induces maladaptive hypertrophy in a titin-truncated mouse model. *Biomed Res Int* 2015;2015:163564.

114. Hinson JT, Chopra A, Nafissi N, et al. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy. *Science* 2015;349:982-986.

115. Schafer S, de Marvao A, Adami E, et al. Titin truncating variants affect heart function in disease cohorts and the general population. *Nat Genet* 2017;49:46-53.

116. Xu X, Meiler SE, Zhong TP, et al. Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin. *Nat Genet* 2002;30:205-209.

117. Seeley M, Huang W, Chen Z, Wolff WO, Lin X, Xu X. Depletion of zebrafish titin reduces cardiac contractility by disrupting the assembly of Z-discs and A-bands. *Circ Res* 2007;100:238-245.

118. Myhre JL, Hills JA, Prill K, Wohlegemuth SL, Pilgrim DB. The titin A-band rod domain is dispensible for initial thick filament assembly in zebrafish. *Dev Biol.* 2014;387:93-108.

119. Olson TM, Illenberger S, Kishimoto NY, Huttelmaier S, Keating MT, Jockusch BM. Metavinculin mutations alter actin interaction in dilated cardiomyopathy. *Circulation* 2002;105:431-437.

120. Wells QS, Ausborn NL, Funke BH, et al. Familial dilated cardiomyopathy associated with congenital defects in the setting of a novel *VCL* mutation (Lys815Arg) in conjunction with a known *MYBPC3* variant. *Cardiogenetics* 2011;1:e10.

121. Xu W, Baribault H, Adamson ED. Vinculin knockout results in heart and brain defects during embryonic development. *Development* 1998;125:327-337.

122. Zemljic-Harpf AE, Ponrartana S, Avalos RT, et al. Heterozygous inactivation of the vinculin gene predisposes to stress-induced cardiomyopathy. *Am J Pathol* 2004;165:1033-1044.

123. Zemljic-Harpf AE, Miller JC, Henderson SA, et al. Cardiac-myocyte-specific excision of the vinculin gene disrupts cellular junctions, causing sudden death or dilated cardiomyopathy. *Mol Cell Biol* 2007;27:7522-7537.

124. Knoll R, Postel R, Wang J, et al. Laminin-alpha-4 and integrin-linked kinase mutations cause human cardiomyopathy via simultaneous defects in cardiomyocytes and endothelial cells. *Circulation* 2007;116:515-525.

125. Wang J, Hoshijima M, Lam J, et al. Cardiomyopathy associated with microcirculation dysfunction in laminin alpha4 chain-deficient mice. *J Biol Chem* 2006;281:213-220.

126. White DE, Coutu P, Shi YF, et al. Targeted ablation of *ILK* from the murine heart results in dilated cardiomyopathy and spontaneous heart failure. *Genes Dev* 2006;20:2355-2360.

127. Quang KL, Maguy A, Qi XY, et al. Loss of cardiomyocyte integrin-linked kinase produces an arrhythmogenic cardiomyopathy in mice. *Circ Arrhythm Electrophysiol* 2015;8:921-932.

128. Bendig G, Grimmler M, Huttner IG, et al. Integrin-linked kinase, a novel component of the cardiac mechanical stretch sensor, controls contractility in the zebrafish heart. *Genes Dev* 2006;20:2361-2372.

129. Morales A, Painter T, Li R, et al. Rare variant mutations in pregnancy-associated or peripartum cardiomyopathy. *Circulation* 2010;121:2176-2182.

130. McConnell BK, Jones KA, Fatkin D, et al. Dilated cardiomyopathy in homozygous myosin-binding protein-C mutant mice. *J Clin Invest* 1999;104:1235-1244.

131. McConnell BK, Fatkin D, Semsarian C, et al. Comparison of two murine models of familial hypertrophic cardiomyopathy. *Circ Res* 2001;88:383-389.

132. Harris SP, Bartley CR, Hacker TA, et al. Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. *Circ Res* 2002;90:594-601.

133. Carrier L, Knoll R, Vignier N, et al. Asymmetric septal hypertrophy in heterozygous cMyBP-C null mice. *Cardiovasc Res* 2004;63:293-304.

134. Chen YH, Pai CW, Huang SW, et al. Inactivation of myosin binding protein C homolog in zebrafish as a model for human cardiac hypertrophy and diastolic dysfunction. *J Am Heart Assoc* 2013;2:e000231.

135. Knoll R, Hoshijima M, Hoffmann HM, et al. The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy. *Cell* 2002;111:943-955.

136. Mohapatra B, Jimenez S, Lin JH, et al. Mutations in the muscle LIM protein and alpha-actinin-2 genes in dilated cardiomyopathy and endocardial fibroelastosis. *Mol Genet Metabol* 2003;80:207-215.

137. Arber S, Hunter JJ, Ross J, et al. MLP-deficient mice exhibit a disruption of cardiac cytoarchitecture organization, dilated cardiomyopathy, and heart failure. *Cell* 1997;88:393-403.

138. Knoll R, Kostin S, Klede S, et al. A common MLP (muscle LIM protein) variant is associated with cardiomyopathy. *Circ Res* 2010;106:695-704.

139. Inagaki N, Hayashi T, Arimura T, et al. Alpha-B-crystallin mutation in dilated cardiomyopathy. *Biochem Biophys Res Commun* 2006;342:379-386.

140. Pilotto A, Marziliano N, Pasotti M, Grasso M, Costante AM, Arbustini E. αβ-crystallin mutation in dilated cardiomyopathies: low prevalence in a consecutive series of 200 unrelated probands. *Biochem Biophys Res Commun* 2006;346:1115-1117.

141. van der Smagt JJ, Vink A, Kirkels JH, et al. Congenital posterior pole cataract and adult onset dilating cardiomyopathy: expanding the phenotype of αβ-crystallinopathies. *Clin Genet* 2014;85:381-385.

142. Zhu Y, Bogomolovas J, Labeit S, Granzier H. Single molecule force spectroscopy of the cardiac titin N2B element: effects of the molecular chaperone αβ-crystallin with disease-causing mutations. *J Biol Chem* 2009;284:13914-13923.

143. Wang X, Osinska H, Klevitsky R, et al. Expression of R120G-αβ-crystallin causes aberrant desmin and αβ-crystallin aggregation and cardiomyopathy in mice. *Circ Res* 2001;89:84-91.

144. Buhrdel JB, Hirth S, Kessler M, et al. In vivo characterization of human myofibrillar myopathy genes in zebrafish. *Biochem Biophys Res Commun* 2015;461:217-223.

145. Xie HB, Cammarato A, Rajasekaran NS, et al. The NADPH metabolic network regulates human αβ-crystallin cardiomyopathy and reductive stress in Drosophila melanogaster. *PLoS Genet* 2013;9:e1003544.

146. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* 1998;280:750-752.

147. Debold P, Saber W, Cheema Y, et al. Human actin mutations associated with hypertrophic and dilated cardiomyopathies demonstrate distinct thin filament regulatory properties in vitro. *J Mol Cell Cardiol* 2010;48:286-292.

148. Kumar A, Crawford K, Close L, et al. Rescue of cardiac α-actin-deficient mice by enteric smooth muscle γ-actin. *Proc Natl Acad Sci USA* 1997;94:4406-4411.

149. Toko H, Takahashi H, Kayama Y, et al. Ca^{2+}/calmodulin-dependent kinase IIδ causes heart failure by accumulation of p53 in dilated cardiomyopathy. *Circulation* 2010;122:891-899.

150. Song W, Dyer E, Stuckey D, et al. Investigation of a transgenic mouse model of familial dilated cardiomyopathy. *J Mol Cell Cardiol* 2010;49:380-389.

151. Glenn NO, McKane M, Kohli V, et al. The W-loop of alpha-cardiac actin is critical for heart function and endocardial cushion morphogenesis in zebrafish. *Mol Cell Biol* 2012;32:3527-3540.

152. Hassel D, Dahme T, Erdmann J, et al. Nexilin mutations destabilize cardiac Z-disks and lead to dilated cardiomyopathy. *Nat Med* 2009;15:1281-1288.

153. Aherrahrou Z, Schlossararek S, Stoelting S, et al. Knock-out of nexillin in mice leads to dilated cardiomyopathy and endomyocardial fibroelastosis. *Basic Res Cardiol* 2016;111:6.

154. D'Adamo P, Fassone L, Gedeon A, et al. The X-linked gene G4.5 is responsible for different infantile dilated cardiomyopathies. *Am J Hum Genet* 1997;61:862-867.

155. Bissler JJ, Tsoras M, Goring HH, et al. Infantile dilated X-linked cardiomyopathy, G4.5 mutations, altered lipids, and ultrastructural malformations of mitochondria in heart, liver, and skeletal muscle. *Lab Invest* 2002;82:335-344.

156. Man E, Lafferty KA, Funke BH, et al. NGS identifies *TAZ* mutation in a family with X-linked dilated cardiomyopathy. *BMJ Case Rep* 2013; pii: bcr2012007529.

157. Acehan D, Vaz F, Houtkooper RH, et al. Cardiac and skeletal muscle defects in a mouse model of human Barth syndrome. *J Biol Chem* 2011;286:899-908.

158. Khuchua Z, Yue Z, Batts L, Strauss AW. A zebrafish model of human Barth syndrome reveals the essential role of tafazzin in cardiac development and function. *Circ Res* 2006;99:201-208.

159. Taylor MR, Ku L, Slavov D, et al. Danon disease presenting with dilated cardiomyopathy and a complex phenotype. *J Hum Genet* 2007;52:830-835.

160. Echaniz-Laguna A, Mohr M, Epailly E, et al. Novel Lamp-2 gene mutation and successful treatment with heart transplantation in a large family with Danon disease. *Muscle Nerve* 2006;33:393-397.

161. Balmer C, Ballhausen D, Bosshard NU, et al. Familial X-linked cardiomyopathy (Danon disease): diagnostic confirmation by mutation analysis of the *LAMP2* gene. *Eur J Pediatr* 2005 164:509-514.

162. Tanaka Y, Guhde G, Suter A, et al. Accumulation of autophagic vacuoles and cardiomyopathy in LAMP-2-deficient mice. *Nature* 2000;406:902-906.

163. Carballo S, Robinson P, Otway R, et al. Identification and functional characterization of cardiac troponin I as a novel disease gene in autosomal dominant dilated cardiomyopathy. *Circ Res* 2009;105:375-382.

164. Murphy RT, Mogensen J, Shaw A, Kubo T, Hughes S, McKenna WJ. Novel mutation in cardiac troponin I in recessive idiopathic dilated cardiomyopathy. *Lancet* 2004;363:371-372.

165. Memo M, Leung MC, Ward DG, et al. Familial dilated cardiomyopathy mutations uncouple troponin I phosphorylation from changes in myofibrillar Ca^{2+} sensitivity. *Cardiovasc Res* 2013;99:65-73.

166. Huang XP, Pi YQ, Lee KJ, et al. Cardiac troponin I gene knockout. A mouse model of myocardial troponin I deficiency. *Circ Res* 1999;84:1-8.

167. Wolf MJ, Amrein H, Izatt JA, Choma MA, Reedy MC, Rockman HA. Drosophila as a model for the identification of genes causing adult human heart disease. *Proc Natl Acad Sci USA* 2006;103:1394-1399.

168. Moulik M, Vatta M, Witt SH, et al. *ANKRD1*, the gene encoding cardiac ankyrin repeat protein, is a novel dilated cardiomyopathy gene. *J Am Coll Cardiol* 2009;54:325-333.

169. Duboscq-Bidot L, Charron P, Ruppert V, et al. Mutations in the *ANKRD1* gene encoding CARP are responsible for human dilated cardiomyopathy. *Eur Heart J* 2009;30:2128-2136.

170. Zhong L, Chiusa M, Cadar AG, et al. Targeted inhibition of ANKRD1 disrupts sarcomeric ERK-GATA4 signal transduction and abrogates phenylephrine-induced cardiomyocyte hypertrophy. *Cardiovasc Res* 2015;106:261-271.

171. Posch MG, Posch MJ, Geier C, et al. A missense variant in desmoglein-2 predisposes to dilated cardiomyopathy. *Mol Genet Metab* 2008;95:74-80.

172. Syrris P, Ward D, Asimaki A, et al. Desmoglein-2 mutations in arrhythmogenic right ventricular cardiomyopathy: a genotype-phenotype characterization of familial disease. *Eur Heart J* 2007;28:581-588.

173. Eshkind L, Tian Q, Schmidt A, Franke WW, Windoffer R, Leube RE. Loss of desmoglein 2 suggests essential functions for early embryonic development and proliferation of embryonal stem cells. *Eur J Cell Biol* 2002;81:592-598.

174. Kant S, Holthofer B, Magin TM, Krusche CA, Leube RE. Desmoglein 2-dependent arrhythmogenic cardiomyopathy is caused by a loss of adhesive function. *Circ Cardiovasc Genet* 2015;8:553-563.

175. Pilichou K, Remme CA, Basso C, et al. Myocyte necrosis underlies progressive myocardial dystrophy in mouse *dsg2*-related arrhythmogenic right ventricular cardiomyopathy. *J Exp Med* 2009;206:1787-1802.

176. Carniel E, Taylor MR, Sinagra G, et al. Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. *Circulation* 2005;112:54-59.

177. Merlo M, Sinagra G, Carniel E, et al. Poor prognosis of rare sarcomeric gene variants in patients with dilated cardiomyopathy. *Clin Transl Sci* 2013;6:424-428.

178. Theis JL, Zimmermann MT, Evans JM, et al. Recessive *MYH6* mutations in hypoplastic left heart with reduced ejection fraction. *Circ Cardiovasc Genet* 2015;8:564-571.

179. Abu-Daya A, Sater AK, Wells DE, Mohun TJ, Zimmerman LB. Absence of heartbeat in the *Xenopus tropicalis* mutation *muzak* is caused by a nonsense mutation in cardiac myosin *myh6*. *Dev Biol* 2009;336:20-29.

180. Bienengraeber M, Olston TM, Selivanov VA, et al. *ABCC9* mutations identified in human dilated cardiomyopathy disrupt catalytic KATP channel gating. *Nat Genet* 2004;36:382-387.

181. Fahrenbach JP, Stoller D, Kim G, et al. Abcc9 is required for the transition to oxidative metabolism in the newborn heart. *FASEB J* 2014;28:2804-2815.

182. Akasaka T, Klinedinst S, Ocorr K, Bustamante EL, Kim SK, Bodmer R. The ATP-sensitive potassium (KATP) channel-encoded dSUR gene is required for Drosophila heart function and is regulated by tinman. *Proc Natl Acad Sci USA* 2006;103:11999-12004.

183. Bagnall RD, Molloy LK, Kalman JM, Semsarian C. Exome sequencing identifies a mutation in the *ACTN2* gene in a family with idiopathic ventricular fibrillation, left ventricular noncompaction, and sudden death. *BMC Med Genet* 2014;15:99.

184. Gupta V, Discenza M, Guyon JR, Kunkel LM, Beggs AH. α-actinin-2 deficiency results in sarcomeric defects in zebrafish that cannot be rescued by α-actinin-3 revealing functional differences between sarcomeric isoforms. *FASEB J* 2012;26:1892-1908.

185. Yang J, Xu X. α-actinin-2 is required for the lateral alignment of Z discs and ventricular chamber enlargement during zebrafish cardiogenesis. *FASEB J* 2012;26:4230-4242.

186. Arola AM, Sanchez X, Murphy RT, et al. Mutations in *PDLIM3* and *MYOZ1* encoding myocyte Z line protein are infrequently found in idiopathic dilated cardiomyopathy. *Mol Genet Metabol* 2007;90:435-440.

187. Pashmforoush M, Pomies P, Peterson KL, et al. Adult mice deficient in actinin-associated LIM-domain protein reveal; a developmental pathway for right ventricular cardiomyopathy. *Nat Med* 2001;7:591-597.

188. Tsubata S, Bowles KR, Vatta M, et al. Mutations in the human δ-sarcoglycan gene in familial and sporadic dilated cardiomyopathy. *J Clin Invest* 2000;106:655-662.

189. Bauer R, Hudson J, Muller HD, et al. Does delta-sarcoglycan-associated autosomal-dominant cardiomyopathy exist? *Eur J Hum Genet* 2009;17:1148-1153.

190. Karkkainen S, Miettinen R, Tuomainen P, et al. A novel mutation, Arg71Thr, in the δ-sarcoglycan gene is associated with dilated cardiomyopathy. *J Mol Med* 2003;81:795-800.

191. Sylvius N, Duboscq-Bidot L, Bouchier C, et al. Mutational analysis of the β- and δ-sarcoglycan genes in a large number of patients with familial and sporadic dilated cardiomyopathy. *Am J Med Genet* 2003;120A:8-12.

192. Politano L, Nigro V, Passamano L, et al. Evaluation of cardiac and respiratory involvement in sarcoglycanopathies. *Neuromusc Dis* 2001;11:178-185.

193. Campbell MD, Witcher M, Gopal A, Michele DE. Dilated cardiomyopathy mutations in δ-sarcoglycan exert a dominant-negative effect on cardiac myocyte mechanical stability. *Am J Physiol Heart Circ Physiol* 2016;310:H1140-1150.

194. Coral-Vazquez R, Cohn RD, Moore SA, et al. Disruption of the sarcoglycan-sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell* 1999;98:465-474.

195. Sabharwal R, Weiss RM, Zimmerman K, Domeniq O, Cicha MZ, Chapleau MW. Angiotensin-dependent autonomic dysregulation precedes dilated cardiomyopathy in a mouse model of muscular dystrophy. *Exp Physiol* 2015;100:776-795.

196. Rutschow D, Bauer R, Gohringer C, et al. S151A δ-sarcoglycan mutation causes a mild phenotype of cardiomyopathy in mice. *Eur J Hum Genet* 2014;22:119-125.

197. Sakamoto A, Ono K, Abe M, et al. Both hypertrophic and dilated cardiomyopathies are caused by mutation of the same gene, δ-sarcoglycan, in hamster: an animal model of disrupted dystrophin-associated glycoprotein complex. *Proc Natl Acad Sci USA* 1997;94:13873-13878.

198. Zhang Q, Bethmann C, Worth NF, et al. Nesprin-1 and -2 are involved in the pathogenesis of Emery-Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Hum Mol Genet* 2007;16:2816-2833.

199. Puckelwartz MJ, Kessler EJ, Kim G, et al. Nesprin-1 mutations in human and murine cardiomyopathy. *J Mol Cell Cardiol* 2009;48:600-608.

200. Zhou C, Li C, Zhou B, et al. Novel nesprin-1 mutations associated with dilated cardiomyopathy cause nuclear envelope disruption and defects in myogenesis. *Hum Mol Genet* 2017;26:2258-2276.

201. Haskell GT, Jensen BC, Samsa LA, et al. Whole exome sequencing identifies truncating variants in nuclear envelope genes in patients with cardiovascular disease. *Circ Cardiovasc Genet* 2017;10:e001443.

202. Zhang J, Felder A, Liu Y, et al. Nesprin 1 is critical for nuclear positioning and anchorage. *Hum Mol Genet* 2010;19:329-341.

203. Child JS, Perloff JK, Bach PM, Wolfe AD, Perlman S, Kark RA. Cardiac involvement in Friedreich's ataxia: a clinical study of 75 patients. *J Am Coll Cardiol* 1986;7:1370-1378.

204. Alboliras ET, Shub C, Gomez MR, et al. Spectrum of cardiac involvement in Friedreich's ataxia: clinical, electrocardiographic and echocardiographic observations. *Am J Cardiol* 1986;58:518-524.

205. Poulsen SH, Sogaard P, Nielsen-Kudsk JE, Egeblad H. Dilated cardiomyopathy in Friedreich's ataxia: 2D echo and tissue-Doppler analysis of left ventricular and atrial function. *Eur J Echocardiog* 2003;4:331-333.

206. Kipps A, Alexander M, Colan SD, et al. The longitudinal course of cardiomyopathy in Friedreich's ataxia during childhood. *Pediatr Cardiol* 2009;30:306-310.

207. Cossee M, Puccio H, Gansmuller A, et al. Inactivation of the Friedreich's ataxia mouse gene leads to early embryonic lethality without iron accumulation. *Hum Mol Genet* 2000;9:1219-1226.

208. Seznec H, Simon D, Monassier L, et al. Idebenone delays the onset of cardiac functional alteration without correction of Fe-S enzymes deficit in a mouse model for Friedreich ataxia. *Hum Mol Genet* 2004;13:1017-1024.

209. Tricoire H, Palandri A, Bourdais A, Camadro JM, Monnier V. Methylene blue rescues heart defects in a Drosophila model of Friedreich's ataxia. *Hum Mol Genet* 2014;23:968-979.

210. Traverso M, Gazzerro E, Assereto S, et al. Caveolin-3 T78M and T78K missense mutations lead to different phenotypes in vivo and in vitro. *Lab Invest* 2008;88:275-283.

211. Catteruccia M, Sanna T, Santorelli FM, et al. Rippling muscle disease and cardiomyopathy associated with a mutation in the *CAV3* gene. *Neuromuscul Disord* 2009;10:779-783.

212. Galbiati F, Engelman JA, Volonte D, et al. Caveolin-3 null mice show a loss of caveolae, changes in the microdomain distribution of the dystrophin-glycoprotein complex, and t-tubule abnormalities. *J Biol Chem* 2001;276:21425-21433.

213. Aravamudan B, Volonte D, Ramani R, et al. Transgenic overexpression of caveolin-3 in the heart induces a cardiomyopathic phenotype. *Hum Mol Genet* 2003;12:2777-2788.

214. Hayashi T, Arimura T, Itoh-Satoh M, et al. Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. *J Am Coll Cardiol* 2004;44:2192-2201.

215. Hershberger RE, Parks SB, Kushner JD, et al. Coding sequence mutations identified in *MYH7*, *TNNT2*, *SCN5A*, *CSRP3*, *LDB3*, and *TCAP* from 313 patients with familial or idiopathic dilated cardiomyopathy. *Clin Transl Sci* 2008;1:21-26.

216. Knoll R, Linke WA, Zou P, et al. Telethonin deficiency is associated with maladaptation to biomechanical stress in the mammalian heart. *Circ Res* 2011;109:758-769.

217. Murakami T, Hayashi YK, Noguchi S, et al. Fukutin gene mutations cause dilated cardiomyopathy with minimal muscle weakness. *Ann Neurol* 2006;60:597-602.

218. Arimura T, Hayashi YK, Murakami T, et al. Mutational analysis of fukutin gene in dilated cardiomyopathy and hypertrophic cardiomyopathy. *Circ J* 2009;73:158-161.

219. Riisager M, Duno M, Hansen FJ, Krag TO, Vissing CR, Vissing J. A new mutation of the fukutin gene causing late-onset limb girdle muscular dystrophy. *Neuromusc Disord* 2013;23:562-567.

220. Tachikawa M, Kanagawa M, Yu CC, Kobayashi K, Toda T. Mislocalization of fukutin protein by disease-causing missense mutations can be rescued with treatments directed at folding amelioration. *J Biol Chem* 2012;287:8398-8406.

221. Kurahashi H, Taniguchi M, Meno C, et al. Basement membrane fragility underlies embryonic lethality in fukutin-null mice. *Neurobiol Dis* 2005;19:208-217.

222. Levitas A, Muhammad E, Harel G, et al. Familial neonatal isolated cardiomyopathy caused by mutation in the flavoprotein subunit of succinate dehydrogenase. *Eur J Hum Genet* 2010;18;1160-1165.

223. Jain-Ghai S, Cameron JM, Al Maawali A, et al. Complex II deficiency: a case report and review of the literature. *Am J Med Genet* 2013;161A:285-294.

224. Ichida F, Tsubata S, Bowles KR, et al. Novel gene mutations in patients with left ventricular noncompaction or Barth syndrome. *Circulation* 2001;103:1256-1263.

225. Grady RM, Grange RW, Lau KS, et al. Role for α-dystrobrevin in the pathogenesis of dystrophin-dependent muscular dystrophies. *Nat Cell Biol* 1999;1:215-220.

226. Strakova J, Dean JD, Sharpe KM, Mayers TA, Odom GL, Townsend D. Dystrobrevin increases dystrophin's binding to the dystrophin-glycoprotein complex and provides protection during cardiac stress. *J Mol Cell Cardiol* 2014;76:106-115.

227. Gavassini BF, Carboni N, Nielsen JE, et al. Clinical and molecular characterization of limb-girdle muscular dystrophy due to *LAMA2* mutations. *Muscle Nerve* 2011;44:703-709.

228. Carboni N, Marrosu G, Porcu M, et al. Dilated cardiomyopathy with conduction defects in a patient with partial merosin deficiency due to mutations in the laminin-α2-chain gene: a chance association or a novel phenotype. *Muscle Nerve* 2011;44:826-828.

229. Marques J, Duarte ST, Costa S, et al. Atypical phenotype in two patients with *LAMA2* mutations. *Neuromusc Disord* 2014;24:419-424.

230. Harris E, McEntagart M, Topf A, et al. Clinical and neuroimaging findings in two brothers with limb girdle muscular dystrophy due to *LAMA2* mutations. *Neuromusc Disord* 2017;27:170-174.

231. Miyagoe Y, Hanaoka K, Nonaka I, et al. Laminin alpha2 chain-null mutant mice by targeted disruption of the Lama2 gene: a new model of merosin (laminin2)-deficient congenital muscular dystrophy. *FEBS Lett* 1997;415:33-39.

232. Kuang W, Xu H, Vilquin JT, Engvall E. Activation of the lama2 gene in muscle regeneration: abortive regeneration in laminin alpha-2 deficiency. *Lab Invest* 1999;79:1601-1613.

233. Guo LT, Zhang XU, Kuang W, et al. Laminin α2 deficiency and muscular dystrophy: genotype-phenotype correlation in mutant mice. *Neuromusc Disord* 2003;13:207-215.

234. Mizuno Y, Thompson TG, Guyon JR, et al. Desmuslin, an intermediate filament protein that interacts with alpha-dystrobrevin and desmin. *Proc Natl Acad Sci USA* 2001;98:6156-6161.

235. Russell MA, Lund LM, Haber R, McKeegan K, Cianciola N, Bond M. The intermediate filament protein, synemin, is an AKAP in the heart. *Arch Biochem Biophys* 2006;456:204-215.

236. Lund LM, Kerr JP, Lupinetti J, et al. Synemin isoforms differentially organize cell junctions and desmin filaments in neonatal cardiomyocytes. *FASEB J* 2012;26:137-148.

237. Garcia-Pelagio KP, Muriel J, O'Neill A, et al. Myopathic change in murine skeletal muscle lacking synemin. *Am J Physiol Cell Physiol* 2015;308:C448-C462.

238. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-424.

LICENSE TO PUBLISH

ACMG
American College of Medical Genetics and Genomics
Translating Genes into Health®

Manuscript Number: GIM-D-17-00727

Journal Name: Genetics in Medicine (the "Journal")

Proposed Title of the Contribution: A gene-centric strategy for identifying disease-causing rare variants in dilated cardiomyopathy (the "Contribution")

Author(s): Claire Hervel, PhD1, Renee Johnson, PhD1, Lien Lam, PhD2,8, Jacob Munro, BSc, MIT3, Francesco Mazzarotto, PhD4, Angharad M. Roberts, MRCP4, Daniel S. Herman, MD, PhD2, Michael Parfenov, MD, PhD2, Alireza Haghighi MD, PhD2,8,9, Barbara McDonough, RN2, Steven R. DePalma, PhD2, Anne M. Keogh, MD,5,6,7 Peter S. Macdonald, MB BS, PhD5,6,7, Christopher S. Hayward, MD5,6,7, Amy Roberts, MD10, Paul J.R. Barton, PhD4, Leanne E. Felkin, Ph (the "Author(s)")

To: *American College of Medical Genetics and Genomics* (the "Society")

1. In consideration of the Society evaluating the Contribution for publication (and publishing the Contribution if the Society so decides) the Author(s) grant to the Society for the full term of copyright and any extensions thereto, subject to clause 2 below, the exclusive right and irrevocable license:
   (a) to edit, adapt, publish, reproduce, distribute, display and store the Contribution in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world;
   (b) to translate the Contribution into other languages, create adaptations, summaries or extracts of the Contribution or other derivative works based on the Contribution and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts and derivative works;
   (c) to license others to do any or all of the above; and
   (d) to re-license article metadata without restriction (including but not limited to author name, title, abstract, citation, references, keywords and any additional information, as determined by the Society).
2. Ownership of copyright remains with the Author(s), and provided that, when reproducing the Contribution or extracts from it or the Supplementary Information (defined below), the Author(s) acknowledge first and reference publication in the Journal, the Author(s) retain only the following non-exclusive rights:
   (a) to reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the Author(s);
   (b) they and any academic institution where they work may reproduce the Contribution for the purpose of course teaching;
   (c) to post a copy of the Contribution as accepted for publication after peer review (in a locked word processing file, or a PDF version thereof) on the Author(s)' own web sites, or institutional repositories, or the Author(s)' funding body(s)' archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the Contribution on the Society's web site; and
   (d) to reuse figures or tables created by the Author(s) and contained in the Contribution in oral presentations and other works created by them.
3. The Author(s) grant to the Society for the full term of copyright and any extensions thereto the same rights that have been granted in respect of the Contribution as set out in clause 1 above, in and to all supplementary material in any form (including without limitation images, videos, tables and/or graphs) submitted by the Author(s) to the Society with or in connection with the Contribution ("Supplementary Information") but on a non-exclusive basis.
4. The Society acknowledges that an earlier version of the Contribution and/or Supplementary Information may have been submitted to a pre-print service (in accordance with that service's standard license terms).
5. The Author(s) jointly and severally warrant and represent to the Society and Nature America, Inc. trading as Nature Publishing Group ("NPG") that:
   (a) the Author(s) are the sole Author(s) of and sole owners of the copyright in the Contribution and the Supplementary Information and the Contribution and the Supplementary Information is the original work of the Author(s) and not copied (in whole or part) from another work. If however the Contribution or the Supplementary Information includes materials from other sources, the Author(s) warrant they have obtained the permission of the owners of the copyright in all such materials to enable them to grant the rights contained herein. Copies of all such permissions are attached to this license;
   (b) all of the facts contained in the Contribution and the Supplementary Information are true and accurate;
   (c) the signatory (the Author or the employer) who has signed this Agreement below has full right, power and authority to enter into this Agreement and grant the rights herein on behalf of all of the Authors;
   (d) nothing in the Contribution or the Supplementary Information is obscene, defamatory, libelous, violates any right of privacy or publicity, infringes any intellectual property rights (including without limitation copyright, patent, database or trademark rights) or any other human, personal or other rights of any person or entity or is otherwise unlawful; and
   (e) nothing in the Contribution or the Supplementary Information infringes any duty of confidentiality which any of the Author(s) may owe to anyone else or violates any contract, express or implied, of any of the Author(s), and all of the institutions in which work recorded in the Contribution or the Supplementary Information was created or carried out, have authorized such publication.
6. The Author(s) authorize the Society to take such steps as it considers necessary at its own expense in the Author(s) name and on their behalf if the Society believes that a third party is infringing or is likely to infringe copyright in the Contribution and/or Supplementary Information including but not limited to initiating legal proceedings.
7. The Author(s) hereby waive or agree not to assert (where such waiver is not possible at law) any and all moral rights they may now or in the future hold in connection with the Contribution and the Supplementary Information.
8. The Author(s) shall cooperate fully with the Society in relation to any legal action that might arise from the publication of the Contribution and/or Supplementary Information and the Author(s) shall give the Society access at reasonable times to any relevant accounts, documents and records within the power or control of the Author(s).
9. The Author(s) agree that NPG is intended to have the benefit of and shall have the right to enforce the terms of this Agreement.
10. If the Contribution is rejected by the Society and not published, all rights under this license shall revert to the Author(s).
11. This Agreement shall be governed by the laws of the State of New York without regards to any conflict of law provisions. The parties submit to the exclusive jurisdiction of the courts located in New York, New York

Signed for and on behalf of the Author(s): Diane Fatkin

Print name: Diane Fatkin

Date: February 15, 2018

Address: Victor Chang Cardiac Research Institute, 405 Liverpool St, Darlinghurst, New South Wales 2010, Australia.

Corresponding Author Name: Fatkin
Manuscript Number: GIM-D-17-00727

Reporting Checklist

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. **Please respond completely to all questions relevant to your manuscript. For sections that are not applicable please fill in NA.** For more information, please read the journal's Guide to Authors.

⬓ Check here to confirm that the following information is available in the Material & Methods section:

- the **exact sample size (*n*)** for each experimental group/condition, given as a number, not a range;
- a **description of the sample collection** allowing the reader to understand whether the samples represent **technical or biological replicates** (including how many animals, litters, culture, etc.);
- a **statement of how many times the experiment shown was replicated in the laboratory**;
- **definitions of statistical methods and measures**: (For small sample sizes (n<5) descriptive statistics are not appropriate, instead plot individual data points)
  - very common tests, such as *t*-test, simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - **statistical test results**, e.g., ***P* values**;
  - definition of **'center values'** as **median or mean**;
  - definition of **error bars as s.d. or s.e.m. or c.i.**

Please ensure that the answers to the following questions are reported both **in the manuscript itself and in the space below.** We encourage you to include a specific subsection in the methods section each for statistics, reagents and animal models. Below, provide the text as it appears in the manuscript as well as the page number.

Statistics and general methods	**Text AND page number from manuscript**
1. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size? (Give text and page #)	NA
For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established? (Give text and page #)	Supplementary methods, page 2
3. If a method of randomization was used to determine how samples/animals were allocated to experimental groups and processed, describe it. (Give text and page #)	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA

4. If the investigator was blinded to the group allocation during the experiment and/or when assessing the outcome, state the extent of blinding. (Give text and page #)	NA
For animal studies, include a statement about blinding even if no blinding was done.	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)?	Yes
Is there an estimate of variation within each group of data?	NA
Is the variance similar between the groups that are being statistically compared? (Give text and page #)	NA

Reagents	**Text AND page number from manuscript**
6. Report the source of antibodies (vendor and catalog number)	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination	NA

Animal Models	**Text AND page number from manuscript**
8. Report species, strain, sex and age of animals	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA

10. We recommend consulting the ARRIVE guidelines (*PLoS Biol.* **8**(6), e1000412,2010) to ensure that other relevant aspects of animal studies are adequately reported.

Human subjects	**Text AND page number from manuscript**
11. Identify the committee(s) approving the study protocol.	page 7, para 1
12. Include a statement confirming that informed consent was obtained from all subjects.	Yes, page 7, para 1
13. For publication of patient photos, include a statement confirming that consent to publish was obtained. For more information, please see http://www.icmje.org/recommendations/browse/roles-and-responsibilities/protection-of-research-participants.html.	NA
14. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent).	NA

15. For phase II and III randomized controlled trials, please refer to the CONSORT statement and submit the CONSORT checklist with your submission.

16. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines.

Data deposition	**Text AND page number from manuscript**
17. Provide accession codes for deposited data. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Microarray data	NA

Deposition is strongly recommended for many other datasets for which structured public repositories exist; more details on our data policy are available in the Guide to Authors. We encourage the provision of other source data in supplementary information or in unstructured repositories such as Figshare and Dryad. We encourage publication of Data Descriptors (see Scientific Data) to maximize data reuse.

18. If computer code was used to generate results that are central to the paper's conclusions, include a statement in the Methods section under "**Code availability"** to indicate whether and how the code can be accessed. Include version information as necessary and any restrictions on availability.	NA

COLOR ARTWORK FORM

This form must be completed for all papers. We will be unable to process your paper through to production until we receive instructions concerning color files. Please upload this form with your revised files.

JOURNAL: ***Genetics in Medicine***

ARTICLE TITLE A Gene-Centric Strategy for Identifying Disease-Causing Rare Variants in Dilated Cardiomyopathy

MANUSCRIPT NUMBER GIM-D-17-00727 / Fatkin

CORRESPONDING AUTHOR NAME

Genetics in Medicine has charges for figures printed in color: $500 for the first color figure and $250 for each subsequent figure. Current ACMG members (excluding Student Members) who are first or senior/corresponding authors are exempt,* as are authors who have opted for Open Access.*

Please note that figures can appear online in color in the HTML version of your manuscript, and in black and white in the PDF/print version of the manuscript. Color figures will also be at the discretion of the editorial office.

Please check:

☑ Yes, my manuscript contains material that must be printed in color. I agree to pay the color charges in full and hereby authorize Nature Publishing Group to invoice me for the cost of reproducing color artwork in print.

☐ Yes, my manuscript contains material that should be in color in the online HTML version, but in black and white in the PDF/print version of the manuscript. No charges incurred.

☐ Yes, my manuscript contains material that must be printed in color, but I have chosen Open Access for my article or am an ACMG member, so am exempt.

☐ No, my manuscript does not contain material that must be printed in color.

Which figures should be printed in color? (e.g. Figures 1a, 2, 3b)

Figure 1 and Figure 2

College Member's Name: _____

Membership will be verified by the ACMG, and false claims will be liable for the full color charge rate. Signed completion of this form constitutes a full and total acceptance of the terms listed. College membership has no bearing on the review process at *Genetics in Medicine*, and papers from members and nonmembers alike will be given equal consideration.

Color figures will be set close to the citation and in the best possible position.

20% VAT will be added to the total charge amount upon invoicing. This applies to all EU authors who do not provide a valid VAT number upon returning this form. Customers outside of the EU will not be charged VAT but local taxes will be added where applicable.

Signature: _____ VAT no. (if applicable): _____

Print Name: DIANE FATKIN Date: February 15, 2018

Updated 8/16/16

nature publishing group npg