HHS Public Access

Author manuscript

Sci Transl Med. Author manuscript; available in PMC 2015 September 04.

Published in final edited form as:

Sci Transl Med. 2015 January 14; 7(270): 270ra6. doi:10.1126/scitranslmed.3010134.

Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease

Angharad M. Roberts[1,2,†], **James S. Ware**[2,3,4,†], **Daniel S. Herman**[4,5,6,†], **Sebastian Schafer**[7], **John Baksi**[2], **Alexander G. Bick**[4,5], **Rachel J. Buchan**[2], **Roddy Walsh**[2], **Shibu John**[2], **Samuel Wilkinson**[2], **Francesco Mazzarotto**[2], **Leanne E. Felkin**[2,3], **Sungsam Gong**[2], **Jacqueline A.L. MacArthur**[8], **Fiona Cunningham**[8], **Jason Flannick**[5,9], **Stacey B. Gabriel**[5], **David M. Altshuler**[5,9], **Peter S. Macdonald**[10], **Matthias Heinig**[7], **Anne M. Keogh**[10], **Christopher S. Hayward**[10], **Nicholas R. Banner**[3,11], **Dudley J. Pennell**[2,3], **Declan O'Regan**[1], **Tan Ru San**[12], **Antonio de Marvao**[1], **Timothy J. W. Dawes**[1], **Ankur Gulati**[2], **Emma J. Birks**[3,13], **Magdi H. Yacoub**[3], **Michael Radke**[14], **Michael Gotthardt**[14,15], **James G. Wilson**[16], **Christopher J. O'Donnell**[17], **Sanjay K. Prasad**[2], **Paul J.R. Barton**[2,3], **Diane Fatkin**[10], **Norbert Hubner**[7,15,18], **J. G. Seidman**[4], **Christine E. Seidman**[4,5,19,*,‡], and **Stuart A. Cook**[3,12,20,*,‡]

[1]Clinical Sciences Centre, Imperial College London, London, UK

[2]NIHR Cardiovascular Biomedical Research Unit at Royal Brompton & Harefield NHS Foundation Trust and Imperial College London, London, UK

[3]National Heart and Lung Institute, Imperial College London, London, UK

[4]Department of Genetics, Harvard Medical School, Boston, USA

[5]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[6]Department of Laboratory Medicine, University of Washington, Seattle, WA

[7]Cardiovascular and Metabolic Sciences, Max Delbrück Center for Molecular Medicine, Berlin, Germany

[8]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK

[9]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA

[*]Correspondence: stuart.cook@nhcs.com.sg or cseidman@genetics.med.harvard.edu.
[†]A.M.R., J.S.W. and D.S.H. contributed equally to this work.
[‡]C.E.S. and S.A.C are co-senior authors

Author contributions: *Data acquisition and primary analyses:* AMR, JSW, DSH, SS, JB, AJB, RJB, RW, SJ, SW, FM, LEF, SG, JALM, FC, JF, SBG, DMA, PSM, MH, AMK, CSH, NRB, DJP, DOR, TRS, ADM, TJWD, AG, EJB, MHY, MR, MG, JGW, CJOD, SKP, PJBR, DF, NH, JGS. *Study conception and design, data synthesis, statistical analyses & manuscript preparation:* AMR, JSW, DSH, PJRB, JGS, CES, SAC. All authors have seen and approved the final manuscript.

Competing Interests: The authors declare that they have no competing interests.

Data and materials availability: All genomic variants presented in the manuscript have been submitted to ClinVar (Accession SCV000189630-SCV000189803). RNAseq data is deposited at ArrayExpress (E-MTAB-2466). Data from population cohorts have previously been deposited into dbGaP.

[10]Cardiology Department, St Vincent's Hospital, Victor Chang Cardiac Research Institute, Darlinghurst NSW and Faculty of Medicine, University of New South Wales, Kensington NSW, Australia

[11]The Royal Brompton and Harefield NHS Foundation Trust, Harefield Hospital, Hill End Road, Harefield, Middlesex UK

[12]National Heart Centre Singapore, Singapore

[13]Department of Medicine, University of Louisville, Louisville, KY 40202

[14]Neuromuscular and Cardiovascular Cell Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany

[15]DZHK (German Centre for Cardiovascular Research), partner site Berlin, Germany

[16]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, Mississippi, USA 39216

[17]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA; Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, Maryland, USA

[18]Charité-Universitätsmedizin, Berlin, Germany

[19]Cardiovascular Division, Brigham and Women's Hospital and Howard Hughes Medical Institute

[20]Duke-National University of Singapore, Singapore

Abstract

The recent discovery of heterozygous human mutations that truncate full-length titin (TTN, an abundant structural, sensory, and signaling filament in muscle) as a common cause of end-stage dilated cardiomyopathy (DCM) provides new prospects for improving heart failure management. However, realization of this opportunity has been hindered by the burden of TTN truncating variants (TTNtv) in the general population and uncertainty about their consequences in health or disease. To elucidate the effects of TTNtv, we coupled *TTN* gene sequencing with cardiac phenotyping in 5,267 individuals across the spectrum of cardiac physiology, and integrated these data with RNA and protein analyses of human heart tissues. We report diversity of *TTN* isoform expression in the heart, define the relative inclusion of *TTN* exons in different isoforms, and demonstrate that these data, coupled with TTNtv position, provide a robust strategy to discriminate pathogenic from benign TTNtv. We show that TTNtv is the most common genetic cause for DCM in ambulant patients in the community, identify clinically important manifestations of TTNtv-positive DCM, and define the penetrance and outcomes of TTNtv in the general population. By integrating genetic, transcriptome, and protein analyses we provide evidence for a length-dependent, dominant negative mechanism of disease. These data inform diagnostic criteria and management strategies for TTNtv-positive DCM patients and for TTNtv that are identified as incidental findings.

Introduction

Non-ischemic dilated cardiomyopathy (DCM) has an estimated prevalence of 1:250, results in progressive cardiac failure, arrhythmia, and sudden death, and is the most frequent indication for cardiac transplantation (1, 2). Despite a strong genetic basis for DCM (2) and the recent advent of affordable and comprehensive exome and genome sequencing techniques that permit screening of all DCM genes (3–5), the application of clinical molecular diagnostics in DCM management remains limited (6), due to historically low mutational yield and a background of protein-altering variation of uncertain significance in the general population that make variant interpretation challenging (7–9).

TTN mutations can cause DCM (10, 11) and heterozygous mutations that truncate full-length titin (TTNtv, titin truncating variants) are the most common genetic cause for severe and familial DCM, accounting for approximately 25% of cases (12). TTNtv also occur in approximately 2% of individuals without overt cardiomyopathy (12–14), which exceeds the prevalence of non–ischemic DCM five-fold, and poses significant challenges for the interpretation of these variants in the era of accessible genome sequencing. Critical parameters that distinguish pathogenic TTNtv and their mechanisms of disease remain unknown.

Titin is a highly modular protein with ~90% of its mass composed of repeating immunoglobulin (Ig) and fibronectin-III (FN-III) modules that are interspersed with non-repetitive sequences with phosphorylation sites, PEVK motifs, and a terminal kinase (15). Two titin filaments with opposite polarity span each sarcomere, the contractile unit in striated muscle cells. The amino terminus of titin is embedded in the sarcomere Z-disk and participates in myofibril assembly, stabilization and maintenance (16). The elastic I-band behaves as a bidirectional spring, restoring sarcomeres to their resting length after systole and limiting their stretch in early diastole (17). The inextensible A-band binds myosin and myosin-binding protein and is thought to be critical for biomechanical sensing and signaling. The M-band contains a kinase (18) that may participate in strain-sensitive signaling and affect gene expression and cardiac remodeling in DCM (19, 20).

The *TTN* gene encodes 364 exons that undergo extensive alternative splicing to produce many isoforms ranging in size from 5,604 to 34,350 amino acids. In the adult myocardium two major full-length titin isoforms, N2BA and N2B, are robustly expressed in addition to low abundance short novex isoforms (Fig. 1). N2BA and N2B isoforms span the sarcomere Z-disk to M-band but differ primarily in the I-band. The longer N2BA isoform contains both the N2A and N2B segments while the N2B isoform lacks the unique N2A segment and contains fewer Ig domains and a smaller PEVK segment. The force required to stretch a titin molecule relates to its fractional extension (21), a parameter that shows non-linear dependence on the I-band composition. For a given sarcomere length the N2B isoform will have greater fractional extension and thus is stiffer than the longer N2BA isoform (20).

To explore further the spectrum of *TTN* genetic variation and transcript usage across the range of cardiac physiology, we studied five discovery cohorts, comprising healthy volunteers with full cardiovascular evaluations (n=308), community-based cohorts with

longitudinal clinical data (3,603 participants in the Framingham (22) and Jackson (23)) Heart Studies (FHS and JHS, respectively), prospectively-enrolled unselected ambulatory DCM patients (n=374) (Fig. S1), and end-stage DCM patients with left ventricular (LV) assist devices and/or considered for transplantation (n=155). Integrated analyses of sequencing and transcriptional data yielded strategies for considerably narrowing the subset of TTNtv that are likely pathogenic. We replicated these observations in two independent cohorts: patients with familial DCM (n=163), and ethnicity-matched population controls from the Women's Health Initiative (24) (WHI, n=667).

Results

Burden of TTNtv in health and DCM

In the discovery cohorts, we identified 56 TTNtv affecting *TTN* isoforms that span the sarcomere in 3,911 controls: 9 in healthy volunteers (2.9%), 16 in FHS (1.0%), and 31 in JHS participants (1.6%). Eighty-three DCM patients carried TTNtv (versus controls, OR = 13 (9–18) $P = 2.8\times10^{-43}$; Table 1): 49 unselected DCM patients (13%) and 34 end-stage DCM patients (22%). Comparing variants found in healthy individuals and DCM patients we found that nonsense, frameshift, and canonical splice site TTNtv were substantially enriched in DCM patients (OR = 17 (11–25); $P = 1.9\times10^{-45}$). Additional variants predicted to alter non-canonical splice signals were also enriched in DCM (OR = 4.2 (1.8–9.7); $P = 0.0017$), but not as strongly (comparison: $P = 0.0068$), and they often occurred in combination with another TTNtv (Tables S1–5).

Distribution of TTNtv in health and DCM

Our previous study of severe and familial DCM identified TTNtv that were located predominantly in the A-band (12). It is not known whether the A-band is more susceptible to truncating variation, whether TTNtv outside the A-band are excluded by alternate splicing, or whether the pathogenesis of TTN DCM is determined either by critical functional elements located in the A-band, or by the length of a truncated TTN protein product. To explore these concepts, we examined the distribution of TTNtv across the spectrum of health and disease.

We observed that TTNtv were non-uniformly distributed within and between study groups (Fig. 1). TTNtv were more commonly located in the A-band in DCM than controls (61/87 case variants located in A-band, versus 21/56 in controls, OR=3.9 (1.8–8.3), $p=1.4\times10^{-4}$), due both to an enrichment of A-band variants in DCM patients (compared with a uniform distribution; OR = 2.3 (1.4–3.6), $P = 3.4\times10^{-4}$), and an opposing trend, towards A-band sparing, in controls (35/56 variants outside A-band, OR = 0.58(0.34–1.0), p= 0.06). A-band enrichment was most pronounced in end-stage DCM patients (OR = 3.5(1.6–8.5), $P = 7.8\times10^{-4}$) with a concordant trend in less severe, unselected DCM (OR = 1.7 (0.97–3.1), $P = 0.07$) (Fig. 1 and Table 1). These distributions were not explained by DNA sequence susceptibility to truncating variation in the *TTN* gene or by differences in variant detection between cohorts (Figs. S2–S3).

The effect of TTNtv on different TTN isoforms

Distal exons, including those that encode the A-band of TTN, are constitutively expressed, whereas many proximal exons, particularly I-band exons, are variably spliced in different isoforms (Tables S6–S7). As recent studies suggest that variants affecting only a subset of gene transcripts are less likely to cause loss-of-function than variants impacting all isoforms (25), we compared TTNtv among different isoforms. TTNtv that altered both N2BA and N2B were strongly enriched in DCM patients versus controls (OR = 19(12–29), $P = 5.5\times10^{-46}$) and associated more strongly with DCM than TTNtv that affected only the N2BA isoform (OR=3.8(1.4–9.2), $P = 0.008$, Table 1).

By contrast, TTNtv found in controls were enriched in exons not incorporated into N2BA and N2B transcripts (such as novex exons and fetal isoforms; 13 variants in 7406 bp N2BA/N2B-excluded exons versus 49 variants in 103052 bp N2BA/N2B-included exons, OR=3.7(2.0–6.7), $P = 2.0\times10^{-4}$). Truncations in exons that are unique to the novex-3 isoform were not significantly different between cohorts (0.37% DCM versus 0.15% controls, OR=2.5(0.36–13), P =0.24) and the nominal two-fold excess in DCM was not robust to analyses that included only European subjects (0.26%, OR=1.4). In addition, the novex-3 isoform only spans the sarcomere Z-disk and proximal I-band (26) and LV expression levels in 105 samples from Genotype-Tissue Expression project (GTEx (27)) and in DCM patients (see below) were approximately 7.3% and 9.4% of N2BA and N2B isoform levels. Given these observations and the lack of evidence for pathogenicity of novex-3 truncations, mutations specific to this isoform were excluded from subsequent analyses.

Alternative splicing of TTN in the human heart

We generated RNA sequencing data from human LV samples (end-stage DCM hearts, n=84) and determined the mean usage of each *TTN* exon (Table S7), denoted as the proportion of transcripts that incorporate each exon (proportion spliced in, PSI; see Materials and Methods). Identical exons were alternatively spliced in LV tissues from DCM patients and GTEx donors (global PSI: R=0.98; Table S7 & Fig. S4). There were important differences between observed exon usage and conventional transcript definitions (see Table S7): 39/122 exons annotated as incorporated into the N2BA isoform were expressed in a small minority of transcripts (PSI < 0.15). Three exons annotated as constitutively expressed appeared to have variable usage (PSI 0.15–0.9), as did two exons absent from the conventional N2BA/N2B descriptions.

The *TTN* gene structure is organized to accommodate extensive splicing events. 85% of all *TTN* exons are symmetric, and consequently their exclusion would not alter the translation frame, whilst exome-wide only 68% of exons are symmetric ($P=1.2\times10^{-11}$). Exon symmetry was correlated with PSI. Only three exons (103, 104 and 106) among 175 exons with PSI < 0.99 are asymmetric, while 49 of 185 (27%) exons with PSI >0.99 are asymmetric (Table S7; $p=7\times10^{-13}$). Within the I-band, the domain with the overall lowest PSI, 93% of alternately spliced exons are symmetric. Moreover the cassette of I-band exons 103–106 is symmetric because these exons are always spliced together. As such, most I-

band exons can be excluded without resulting in a frameshift, including exons that might include TTNtv.

We used the mean PSI scores from the end-stage DCM patients to annotate each *TTN* exon's usage. The usage of the exons containing TTNtv differed between cohorts. Treating all cohorts as an ordered variable with four levels (healthy volunteers, general population, ambulatory unselected DCM, and end-stage DCM), we noted a strong relationship between cohort and mutant exon usage (Kruskal-Wallis chi-square, $P = 4.9\times10^{-3}$), with TTNtv-containing exons in controls having lower usage than TTNtv-containing exons in DCM patients ($P = 2.5\times10^{-4}$) (Fig. 2a). Based on this observation, we suggest that many TTNtv in controls may be tolerated because they fall in exons that are spliced out of the majority of expressed transcripts.

Protein coordinates of truncating variants

TTNtv in DCM cases could also be described as occurring more distally in the titin gene than TTNtv in controls. Treating all cohorts as an ordered variable (as above), we also observed a significant relationship between cohort phenotypes and TTNtv position (Kruskal-Wallis chi-square, $P =3.1\times10^{-3}$).

Diagnostic interpretation of TTNtv

As diagnostic sequencing of *TTN* will be most useful if causality can be confidently ascribed to individual variants, we estimated the probability of pathogenicity of TTNtv based on their relative frequency between cohorts. Applying this framework to our discovery cohort, we predicted that TTNtv produced by nonsense, frameshift, or canonical splice site mutations that affect highly expressed exons (PSI>0.9) had at least a 93% probability of pathogenicity (likelihood ratio (LR) = 14) when identified in an unselected patient with DCM, and an even higher probability of pathogenicity in end-stage disease (≥96%, LR= 24). When segregation data are available, we expect that probabilities will be even higher. These are conservative estimates, as we assumed an all or nothing model in which all TTNtv in controls were benign, giving an upper limit of the background noise.

Approximately 50% of the TTNtv identified in healthy volunteers and community-based cohorts occurred in low PSI exons, including novex-specific exons. Analyses of publicly available genomic datasets showed similar results. TTNtv occur in 1.1% alleles in the 1000 Genomes project (28, 29) and in 2.6% of alleles in the NHLBI GO Exome Sequencing Project (ESP, http://evs.gs.washington.edu/EVS/). Seven of 12 (58%) TTNtv in 1000 Genomes and 83 of 168 (49%) TTNtv in ESP are in novex or other low PSI exons (Tables S8 & S9).

To further explore the health effects associated with TTNtv in community-based cohorts we examined longitudinal follow up data. The majority of FHS and all JHS participants with TTNtv had normal cardiac parameters. However, among the small numbers of TTNtv-positive FHS subjects we observed a higher lifetime incidence of DCM morphology (dilated LV with impaired ejection fraction) in the absence of coronary artery disease (TTNtv-positive, 2 of 16 subjects; TTNtv-negative, 12 of 1,574 subjects; RR=16, $P = 0.008$; Tables

S10 & S11). While the two TTNtv (c.9727C>T, c.1245+3A>G) associated with DCM morphology are outside the A-band, both are in highly expressed exons (PSI = 1). The association between DCM morphology and TTNtv in highly expressed exons was even more marked (TTNtv-positive, 2 of 12 subjects; RR 22, $P = 0.005$).

TTNtv-positive FHS subjects had neither evidence of early heart failure nor early cardiovascular death (Table S12 and Fig. S5). While some of these TTNtv are unlikely to be pathogenic (e.g. TTNtv found in rare novex exons), others may cause DCM with reduced penetrance due to characteristics of the variant itself and/or the effects of additional protective and exacerbating genetic or environmental modifiers that are recognized in DCM (30, 31), factors that may also account for the mismatch between population prevalence of mutations in hypertrophic cardiomyopathy genes and overt disease (8).

Validation studies

The genetic and transcriptional analyses of the five discovery cohorts predicted that the pathogenicity of TTNtv was informed by isoform, exon usage, and variant position. To validate this hypothesis we identified TTNtv in an independent cohort of familial DCM patients (n=163) and 667 healthy participants in the WHI (24), a clinical trial that excluded individuals with chronic disease. In comparison to controls, TTNtv affecting both the N2BA and N2B isoforms were enriched in the DCM replication cohort (OR=78) 20–460), $P = 3.8\times10^{-21}$), and TTNtv in this cohort were located in more highly expressed exons ($P = 7.5\times10^{-4}$; Table 1, Figs. 1 & 2, Table S13). The predicted probability of pathogenicity of nonsense, frameshift, or canonical splice TTNtv in highly expressed exons in the DCM replication cohort was ≥98% (LR= 41).

Clinical stratification of DCM by TTN genotype

To better ascertain clinical phenotypes associated with TTNtv-positive DCM, we capitalized on quantitative cardiac MRI (CMR) (32, 33) in DCM patients. Among TTNtv-positive DCM patients we observed more severely impaired LV function, lower stroke volumes and thinner LV walls (Table 2) than in TTNtv-negative DCM patients. Multivariate regression confirmed that *TTN* genotype predicted phenotype severity after adjusting for important covariates (Tables S14 & S15). Mid-wall fibrosis, an important prognostic factor in DCM (34, 35), was similar in patients with and without TTNtv, but sustained ventricular tachycardia was more common in TTNtv-positive patients (OR = 6.7; $P = 0.001$) and robust to adjustment for LV ejection fraction. Consistent with these adverse intermediate phenotype associations, we observed a difference in the composite endpoint of LV assist device implantation, listing for cardiac transplantation, and all cause mortality. TTNtv-positive DCM patients reached this endpoint at earlier ages ($P = 0.015$, Fig. 3) and sooner after prospective enrolment (P= 0.05, Fig. 3).

Length-dependent association between TTNtv position and DCM

Motivated by the association between TTNtv location and disease status (Fig. 1), that may persist after controlling for PSI (Fig. 2b), we considered TTNtv in DCM patients as an allelic series to dissect positional effects and disease mechanism.

The distance from the N-terminus of the TTN protein to the TTNtv correlated with CMR indices (Figs. 4 & S6). Multivariate linear regression models showed that TTNtv location was significantly correlated with principle indices of heart function: ejection fraction and stroke volume ($P < 0.006$; Tables S14–15). Importantly, this positional effect on cardiac parameters was large, such that a C-terminal TTNtv would be associated with substantially reduced ejection fraction as compared with an N-terminal TTNtv (absolute reductions LV: −18±7%, $P = 0.006$; RV: −21±9%, $P = 7.3\times10^{-6}$) and indexed stroke volumes (absolute reductions: LV −22±8ml/m^2, $P = 0.0017$; RV −23±8ml/m^2, $P = 0.0013$). Among subjects with TTNtv, the variant position explained 19–23% of the observed variation (R^2) in phenotypic indices. Regression modelling of CMR data in FHS participants also suggested that the distance of the TTNtv from the N-terminus correlated with cardiac morphology; there was a consistent direction of effect across a range of phenotypic indices (Figs. S7 & S8 & Table S16–17).

We suggest that the phenotypic associations with exon usage and protein coordinate are potentially of substantial clinical importance for diagnostic variant interpretation. In addition, we deduced that a linear positional effect of TTNtv implied that mutant proteins produced dominant negative effects.

Molecular studies and mechanistic implications

Based on the observation that TTNtv exhibited length-dependent effects, we studied allele-specific *TTN* transcript expression and protein levels in human LV tissue to further explore whether TTNtv caused DCM through dominant negative effects or through haploinsufficiency. RNA sequencing showed comparable total *TTN* transcript levels in patients with or without TTNtv (Fig. 5a). Moreover, the relative expression of TTNtv and of other SNPs distributed throughout *TTN* transcripts showed robust expression of both alleles (Fig. 5b, and Table S18 and Fig. S9). We also observed no discernable difference in the abundance of N2BA and N2B protein isoforms in DCM patients either with or without TTNtv (Fig. 5c). Combined with the genetic data presented above, our analyses of *TTN* RNA and protein expression in LV tissues strongly suggest that TTNtv cause DCM by a dominant negative effect.

Discussion

The integrated analyses of *TTN* sequence, protein and transcriptional data, and quantitative phenotypic assessment in over 5200 healthy and DCM subjects define the spectrum of cardiac physiology associated with TTNtv. We demonstrate TTNtv occur in ~2% of the general population, in 13% of ambulatory unselected DCM patients and in 20% of end-stage DCM patients. We suggest that the clinical significance of TTNtv is largely predicated by exon usage and variant location (the distance of the TTNtv from the protein N-terminus). Incorporation of these data improved discrimination between pathogenic and benign variants in two independent study cohorts. That TTNtv exhibited length-dependent consequences and were highly expressed in human LV tissue indicate that these mutations cause DCM through a dominant negative mechanism.

Cardiomyopathy genes feature prominently in the American College of Medical Genetics and Genomics' list of genes in which mutations should be reported regardless of the primary indication for sequencing (36). Accurate interpretation of clinically actionable incidental findings in cardiomyopathy genes is both difficult and medically important because of the considerable population prevalence of protein-altering variants in cardiomyopathy genes (8, 36), a combined population prevalence of cardiomyopathies of approximately 0.7%, and the associated important medical consequences including heart failure and sudden death (37–39).

The true frequency of TTNtv across the general population has been unclear and the lack of penetrance of these variants is an issue of debate (13, 20, 40, 41). From analyses of over 4500 control subjects we identified TTNtv in 1.6% in individuals of African descent and 1.5% in individuals of European descent. Using transcript and mean *TTN* exon expression in human heart tissue, we provide novel insights into why some of these TTNtv are phenotypically silent. Truncations in the control subjects were more likely to affect minor *TTN* isoforms as compared with DCM cases, including novex-3, a low abundance isoform that does not span the cardiac sarcomere. Truncations that occurred specifically in novex-3 were not enriched in DCM patients. We also observed that while canonical splice variants were enriched in DCM cases, other variants predicted *in silico* to alter splicing showed more modest enrichment (0.4% vs. 1.7%, P= 0.0015; Table 1). While this observation may reflect limitations in current prediction algorithms, interpretation of non-canonical splice variants should be cautious unless informed by RNA evidence or robust segregation data. In addition, the impact of TTNtv in exons with intermediate expression levels (PSI 0.15–0.9), such as in the I-band, may be ameliorated because symmetrical exons may be excluded without deleterious consequences, or by an isoform switch from N2BA to N2B, or because short mutant proteins are less deleterious. Overall, TTNtv in low to intermediate expression exons (PSI<0.9), novex *TTN* isoforms, and at non-canonical splice sites accounted for 50% of all TTNtv identified in controls cohorts. These variants were associated with normal cardiac morphology and function and were not associated with DCM. We suggest that when TTNtv with these characteristics are incidentally identified in low risk individuals that clinical interpretation should not convey a high risk for DCM.

A small number of FHS participants (n=16) had TTNtv in highly expressed exons including 2 individuals with DCM morphology on cardiac imaging (RR=16). There was no increased risk for DCM in TTNtv-positive JHS participants, possibly due to phenotype ascertainment or ethnicity-specific genetics. Whether the differential risk associated with TTNtv in DCM and population cohorts reflects differences in phenotypic assessment, the very small numbers of cases, an aggregation of additional genetic factors in phenotypically ascertained DCM families, or other factors is unknown.

Patient stratification is a cornerstone of precision medicine (42). We propose that DCM due to TTNtv represents a specific patient subgroup who may benefit from more tailored clinical management. In DCM, sustained ventricular tachycardia, LV wall thickness (43) and LV ejection fraction predict outcome (44). TTNtv-positive patients reported here had poorer cardiac indices and earlier onset of heart failure or death. Despite small differences in the functional indices between groups, we showed that, as compared to TTNtv-negative

patients, TTNtv-positive DCM had substantially increased risk of sustained ventricular tachycardia (OR=6.8, P=0.001), perhaps related to increased wall stress (45). If these findings are replicated in prospective cohorts, TTNtv-positive DCM patients may benefit from a lower threshold for device therapy, as is practiced with *LMNA* DCM (46). Preliminary observations in five of six TTNtv-positive DCM patients who received mechanical unloading therapy support during this study had sustained recovery of cardiac function, raising the possibility that *TTN* DCM may prove amenable to targeted device therapy.

Mutations that lead to premature termination of the encoded protein often cause haploinsufficiency. By contrast, our analyses of allele-specific transcript expression and protein in human LV tissue indicate that *TTN* is highly and biallelically expressed: TTNtv were not subjected to substantial nonsense-mediated decay and levels of the major titin isoforms were not diminished. Recent studies of 2q31-q32 deletions encompassing the entire *TTN* locus that did not cause overt cardiac muscle disease (47), and immunohistochemical detection of truncated TTN in the sarcomeres of skeletal muscles from patients with recessive TTN mutations (48) supports this postulate. Rather, the correlations between TTNtv position and cardiac function reported here strongly suggest a dominant negative effect, as occurs with *MYBPC3* truncations that cause hypertrophic cardiomyopathy (49, 50), with increasing DCM severity associated with longer mutant proteins. Further studies are needed to determine if this length relationship results from an increased energetic cost associated with generation and turnover of longer mutant proteins, deleterious sequestration of non-sarcomeric intracellular factors that is proportional to protein length, or increased propensity for disruptive sarcomere protein interactions with longer mutant proteins.

We recognize several potential limitations despite these extensive analyses. There are many exons in which TTNtv were not observed, and we can only extrapolate our findings to these regions. Approximately 5% of the gene comprises repetitive exons with poor alignment: though coverage did not differ between cohorts, additional TTNtv in these regions may have been missed. Newer sequencing assays with longer read lengths are addressing this. The low burden of TTNtv in population cohorts yielded a small allelic series that combined with sensitivity of screening-quality phenotype ascertainment limited the power of genotype-phenotype analyses in these cohorts. Only a subset of the unselected DCM cohort had arrhythmia data, and on-going follow-up with replication is needed to further these assessments. Finally, molecular studies were limited by the scarcity of human cardiac tissue for study.

In conclusion, our data illuminate important characteristics that determine TTNtv pathogenicity and begin to dissect the molecular mechanisms by which these cause DCM. Nonsense, frameshift, and canonical splice site TTNtv, particularly those that truncate both principal isoforms of *TTN* and/or reside towards the carboxy-terminus, cause DCM with severely impaired LV function and life-threatening ventricular arrhythmias. In contrast, truncations that occur in novex-specific exons or other infrequently used *TTN* exons are less likely to be deleterious. An immediate clinical utility of our findings is improved variant interpretation that enables cascade screening of relatives, and gene and genotype-guided stratified management of DCM. Further elucidation of the myocyte processes altered by

large dominant negative mutant TTN proteins will be important to direct the development of therapies that prevent or attenuate the progression of TTNtv related DCM

Materials and Methods

STUDY DESIGN

We set out to compare the burden of rare TTN variants across five cohorts (detailed below) and to explore genotype-phenotype relationships within cohorts using standard cardiac investigations and techniques. There were no interventions. No genotype information was available at recruitment, so patient inclusion was blinded to genotype. Phenotype assessment was blinded to genotype but not to disease status. Study design and analyses underwent rigorous statistical review. All studies were carried using protocols that were reviewed and approved by institutional ethics committees and with informed consent from all participants. Tissue studies complied with U.K. Human Tissue Act guidelines.

COHORT DESCRIPTIONS AND SUBJECT SELECTION

DCM cohorts—374 unselected prospective patients of European ancestry who were referred to the cardiovascular magnetic resonance (CMR) unit of the Royal Brompton and Harefield Hospitals NHS Foundation Trust (RBHT) from July 2001 to August 2012 and diagnosed with idiopathic dilated cardiomyopathy (DCM) were studied. DCM was diagnosed by CMR findings of ejection fraction >2sd below and end diastolic volume >2sd above the mean normalized for age and sex (33, 51) by two independent Level 3 accredited CMR cardiologists (See Fig. S1). Patients with clinical symptoms or signs of active myocarditis or CMR evidence of infiltrative disease were excluded. CAD was assessed either by coronary angiography (249 patients), or non-invasive testing and clinical profiles (e.g. young relatives of an individual with idiopathic DCM, 101 patients). 24 patients had bystander CAD considered insufficient to produce CMR features of DCM (<2 myocardial segments with <25% late gadolinium enhancement).

155 randomly selected end-stage non-ischemic DCM patients who were listed for cardiac transplantation and/or left ventricular device implantation between 1993 and 2011 at RBHT and prospectively enrolled in a tissue bank were studied. Seventy-one of these patients were previously reported (12) Frozen LV samples from 84 patients were used for tissue studies.

163 DCM patients of European ancestry who were referred to the genetics research program at St Vincent's Hospital and Victor Chang Cardiovascular Institute were studied as a replication cohort. Patients had a positive family history (DCM or sudden explained cardiac death in ≥2 family members) and three had subclinical skeletal myopathy (elevated creatine kinase blood levels). DCM was diagnosed at a mean age of 42 years based on presenting clinical symptoms (exertional dyspnea, palpitations), and echocardiographic findings of LV dilation (LVDD>56 mm) with reduced systolic performance (EF<50%). DCM severity ranged from mild to severe: 32 patients (20%) required cardiac transplantation and 2 patients died from advanced heart failure prior to transplantation.

Healthy Volunteers and Population Cohorts—308 self-reported and clinically screened adult volunteers (age range 18–72 years, mean 40.3 years) of European ancestry

were prospectively recruited via advertisement at the MRC Clinical Sciences Centre, Imperial College London. Participants with previously documented cardiovascular disease, hypertension, diabetes or hypercholesterolemia were excluded.

Community-based cohorts—1,623 unrelated participants in the Framingham Heart Study (FHS) Offspring Cohort and 1,980 unrelated randomly selected participants in the Jackson Heart Study (JHS) were studied. The FHS is a multi-generation, prospective, population-based study aimed at identifying the causes of cardiovascular disease (22). In 1948, residents of Framingham, Massachusetts, of European Ancestry were enrolled. Between 1971 and 1975, the study enrolled a second generation, the Offspring Cohort, comprising 5,124 children of the original cohort and their spouses. The Offspring cohort has since been examined every three to eight years, with the last exam reported here, exam 8, occurring between 2005 and 2008. All FHS phenotypic data was retrieved from NCBI dbGaP (Accession: phs000007.v18.p7). *TTN* sequence data for FHS participants are available from NCBI dbGaP (Accession: phs000307.v3.p7). The JHS is an African American population-based, prospective study of cardiovascular disease (23). Between 2000 and 2003, the study enrolled 5,301 African-Americans aged 35 to 84 living in the Jackson, Mississippi metropolitan area. All JHS phenotypic data was retrieved from NCBI dbGaP (Accession: phs000286.v3.p1). *TTN* sequence data for JHS participants are available from NCBI dbGaP (Accession: phs000498.v1.p1).

Population control replication cohort—667 women of European ancestry from The Women's Health Initiative (WHI) participants with exome data (dbGaP Study Accession phs000200.v1.p1) were studied. These individuals are a subset of 161,808 postmenopausal WHI participants (aged 50 to 79 years) who were recruited and followed from 40 clinical centers across the United States between 1993–1998 (24). The WHI eligibility criteria included the ability to complete study visits with expected survival and local residency for at least 3 years. Subjects with medical conditions that would limit full participation in the study including individuals with advanced heart failure were excluded from enrolling.

Phenotype ascertainment details for all study cohorts are provided in Supplementary Materials.

TTN SEQUENCE DATA

DCM subjects and healthy volunteers—Targeted re-sequencing: *TTN* was sequenced in prospective DCM cases, healthy volunteers, and 103 end-stage DCM cases using a targeted approach. Custom hybridization capture probes were designed to target genes implicated in cardiovascular disease, including *TTN*. RNA baits were designed using Agilent's eArray platform. Baits targeted all exons of all Ensembl *TTN* transcripts (Ensemble version 54), including UTR, with a 100-bp extension into adjacent introns, and 1.25 kb of upstream sequence (Fig. S2, and Table S7). A total of 6340 unique 120-mer RNA baits were generated with increased bait tiling across the target (five-fold), covering a target region of 168369 bp, including 112916 protein-coding bases. DNA library preparation and target capture were performed according to the manufacturers' protocols before paired-end sequencing on the SOLiD 5500xl (Life Technologies). Reads were de-multiplexed and

aligned to human reference genome (hg19) in colour space using LifeScope v2.5.1 "targeted.reseq.pe" pipeline. SOLiD Accuracy Enhancement Tool (SAET) was used to improve colour call accuracy prior to mapping. All other LifeScope parameters were used as default. Duplicate reads and those mapping with a quality score <8 were removed. Variant calling was performed using diBayes (SNPs) and small.indel modules, as well as GATK v1.5–2.7 (55) and Samtools v0.1.18. Variants called by any of these methods were taken forward for Sanger validation. Alignment and coverage metrics were calculated using Picard v1.40, BedTools v2.12 and in-house Perl scripts. GATK CallableLoci Walker was used to identify target genomic regions covered sufficiently for variant calling (minimum depth >4 with base quality >20 and mapping quality >10).

In a subset of end stage patients (n=54) *TTN* was studied by whole genome sequencing (Complete Genomics, CA 94043, USA) and variants were called using Complete Genomic Analysis Tools (http://www.completegenomics.com/analysis-tools/cgatools/ version: 2.2.0.26). Variants were filtered out if they met any of the following conditions; i) low confidence or incomplete calls flagged by the caller, ii) read-depth less than 10, iii) allele frequency less than 0.15 for heterozygous calls. All remaining putative TTNtv were taken forward for Sanger validation.

No genotype information was available at recruitment, so patient inclusion was blinded to genotype. In addition to sequencing Titin we sequenced known DCM genes used in clinical practice and observed no difference in the prevalence of rare protein-altering variants in LMNA, MYH6, MYH7, TNNT2, SCN5A, nor in TTN missense variants between TTNtv +ve and TTNtv−ve cases in the prospectively recruited unselected DCM cohort.

Community-based cohorts—A custom set of hybridization capture probes were designed that targeted cardiovascular disease genes including *TTN*. Genomic DNA libraries were constructed for each sample and libraries were paired-end sequenced with an Illumina HiSeq2000 as previously described (8). Sequence reads were mapped to the hg19 human reference sequence with BWA (56). GATK v1.3 was used to recalibrate base quality scores, locally re-align reads, call single-nucleotide variants and small indels, and filter variant calls. Splice variants outside of the absolutely conserved intron bases were identified as described above for the DCM subjects. All *TTN* nonsense, frameshift, and splicing variants reported in FHS or JHS subjects were visually inspected using the Integrated Genomics Viewer, leading to the exclusion of 30 (FHS) and 4 (JHS) variants. FHS and JHS variants were not validated by an additional genotyping method.

Replication cohorts—The DCM replication cohort was studied by targeted sequencing approach using an Agilent custom capture that included *TTN*, followed by sequencing on the Illumina HiSeq 2000 platform. WHI exomes were captured and sequenced on the Illumina GenomeAnalyzerII and HiSeq platforms as previously described (57). Sequence reads for the replication cohorts were processed using an identical pipeline. Raw sequence reads for both WHI exomes and DCM replication cohorts were aligned to the human reference genome hg19 using Novoalign (http://novocraft.com). Duplicates were marked with Picard (http://picard.sourceforge.net). Indel realignment and Base Quality Score Recalibration was done with GATK v2.7. SNP and short insertions/deletions (indels) were called using a

pipeline derived from GATK v2.7 best practices. Variants were simultaneously called on all replication DCM cases and WHI controls using the UnifiedGenotyper joint variant calling module.

LV TISSUE STUDIES

Tissue studies were performed on LV tissue samples from the end-stage DCM cohort, snap frozen in liquid nitrogen at the time of acquisition. RNA sequencing was performed on all 84 samples, and protein studies in a subset of these. Control LV tissue used for protein studies was from unused donor hearts with no known cardiac disease, from the RBHT transplant program, stored and prepared as for the DCM samples.

Transcript studies—Total RNA was extracted from frozen LV samples from 84 end-stage DCM cases using Trizol (Life Technologies) following the manufacturer's protocol, and quantified by UV spectrophotometry. RNA quality was measured on the Agilent 2100 Bioanalyzer using Agilent's RNA 6000 reagents. RINs ranged between 6.3 and 8.7 with a mean of 7.6. 4μg of total RNA was used for library preparation with the TruSeq RNA Sample Preparation Kit (Illumina). Barcoded cDNA fragments of poly (A)+ RNA were then sequenced on a HiSeq 2000 (Illumina) using 2 × 100 bp PE chemistry. Pools of 6 samples were loaded on 3 lanes to avoid batch effects and obtain sufficient coverage for splicing analyses.

Reads were initially deconvoluted and aligned to the genome to detect and exclude multi-mapping sequences. The remaining sequences were then mapped against the GRCh37 reference genome and transcriptome using TopHat 1.4.1 (58) supplied with Ensembl (59) gene annotations. Splice junction detection was performed to allow split alignment across both known and novel splice sites.

RNAseq data was used to compare levels of N2BA, N2B, and novex-3 in DCM samples, and in 105 GTEx samples obtained from dbGaP (27), and processed using the same bioinformatic pipeline.

Reads from DCM samples were filtered stringently before calling point mutations in the RNAseq data. Only reads mapping to one unique position in either the genome or transcriptome with at most 2 mismatches in 100 basepairs were considered for further analyses. The SAMtools suite (60) was used to call TTNtv at all positions covered by a minimum of 10 reads, and putative variants present in at least 30% of reads covering the position were taken forward for Sanger validation.

Sequencing was also performed on paired DNA samples for all 84 individuals, as described above. Nineteen TTNtv were identified and confirmed in genomic DNA from 18 individuals (Supplementary Information). The allele balance of 12 TTNtv could be interrogated in both DNA & RNA (Table S18 and Fig. S9).

To assess allelic expression across the gene (Fig. 5b), base calls at SNP positions throughout *TTN* were quantified without applying any cut-off regarding the variant fraction. The read population supporting the SNP was then compared to the total number of reads to calculate

the allele balance. SNPs found in RNAseq were compared against SNPs found in the same sample by targeted NGS or whole genome sequencing and those SNPs found in both DNA and RNA were considered validated. There were no discordant zygosity calls.

To estimate *TTN* expression levels, all uniquely mapping reads that could be assigned to *TTN* unambiguously and did not intersect with any other known annotated transcripts were counted for each individual. These read numbers were then quantile-normalized to be compared across all samples.

To estimate exon usage, reads covering *TTN* exons (inclusion reads) and reads completely aligning before and after the exon but not within the exon borders (exclusion reads) were counted and normalized for exon length. The proportion of reads indicating incorporation of the exon compared to the number of reads deriving from isoforms excluding the particular exon indicates the proportion of transcripts that use the respective exon (proportion spliced in (PSI) score (61)). For each exon, the median value of PSI across 84 samples was taken as our estimate of exon usage.

Novex-3 expression was estimated by comparing the number of reads mapping to the last kilobase of the Novex-3 terminal exon (exon 48) and the last kilobase of the full-length terminal exon (364). The ratio was calculated for each sample from GTEx and DCM, and the mean taken as an estimate of Novex-3 abundance relative to other isoforms.

Estimated exon usage was derived from the RNA sequencing of 84 LV tissue samples; these values were then applied across all cohorts to give the estimated usage level of each exon containing a TTNtv (Fig. 2).

Protein studies—Analysis of titin isoform expression was by vertical 1% SDS agarose gel electrophoresis (VAGE). Protein samples from left ventricles were homogenized in sample buffer (8 M urea/2 M thiurea/0.05 M Tris pH 6.8/75 mM DTT/, 3% SDS/0.05% bromophenol blue) and titin isoforms were separated using an SDS/agarose gel electrophoresis system (62). Gels were Coomassie-stained to visualize the titin isoforms N2BA (several sizes, including both the N2A and N2B regions), N2B, and the proteolytic fragment T2.

VARIANT ANNOTATION

To facilitate standardized variant annotation in accordance with international guidelines we developed a Locus Reference Genomic sequence (LRG) (63) for *TTN* (http://www.lrg-sequence.org, LRG_391). Variants are described relative to an inferred complete meta-transcript (LRG_391_t1) manually curated by the HAVANA group that incorporates all *TTN* exons, with the exception of a single alternative terminal exon unique to the shorter novex-3 isoform (Figure 1 and Fig. S2 and Tables S6 & S7). Variants in the novex-3 terminal exon are reported relative to LRG_391_t2.

Variants were reported using Human Genome Variation Society nomenclature. The functional consequences of variants were predicted using the Ensembl perl API (64) Variant Effect Predictor (65). Variants were classified as truncating if their consequence included

one of following sequence ontology terms: “transcript_ablation”, “splice_donor_variant”, “splice_acceptor_variant”, “stop_gained”, “stop_lost” or “frameshift_variant”. To identify splice variants outside of the absolutely conserved 2 intron bases, Alamut (66) was used to calculate Maximum Entropy (67), Neural Network (NNSplice), Splice Site Finder (SSF), Human Splice Finder (HSF), and Gene Splicer (GS) scores for reference and alternate alleles. We used the FHS cohort variant frequencies to establish a threshold value for calling splice variant predictions: for each variant in the splicing region (donor: −3 to +6, acceptor: −20 to +3), the pairs of splicing scores were subtracted from one another and converted to percentiles. Variants scored >= 90^{th} percentile by at least 3 algorithms and >= 70^{th} percentile by all applied algorithms were considered conservative splicing variant predictions and the minimum absolute score change for each prediction algorithm was applied as threshold across all cohorts. These thresholds were selected to be more conservative than the previously applied Maximum Entropy Score difference <= −2 threshold (12) and to exclude variants found more frequently than in 1 in 1,000 individuals.

STATISTICAL ANALYSES

Statistical analyses were performed using R (http://www.R-project.org). Comparisons between groups were performed using Mann-Whitney or Fisher’s exact tests, as appropriate, except where indicated. Odds ratios are reported with 95% confidence intervals. Significance tests are two-tailed with alpha = 0.05 unless otherwise indicated. Standard linear regressions were used to evaluate the relationship between *TTN* genotype and cardiovascular phenotypes. Multivariate models were generated using known clinical covariates and optimised to minimise Bayesian information criterion. The relationships between morphologic parameters and *TTN* genotype were assessed by ANOVA between nested linear models.

Determining the likelihood that a TTNtv found in an individual with DCM is pathogenic—Excluding novex, low expression & predicted splice site variants:

$$\text{Total TTNtv frequency in controls} = 31/3911 = 0.79\%$$
$$\text{Total TTNtv frequency in unselected DCM} = 45/374 = 12\%$$

In order to estimate the proportion of variants in cases that are truly pathogenic, we take the frequency of TTNtv in controls without overt disease as an estimate of the burden of benign variation in both cohorts (0.79%). The burden of pathogenic TTNtv in unselected DCM is therefore conservatively estimated at 11.21%.

In an individual with DCM, the likelihood that a TTNtv is pathogenic = 11.21/0.79 = 14.2, and the probability of pathogenicity = 11.21/12 = 93.4%. The calculation for end-stage DCM is equivalent.

Stratification of DCM and linear modelling—For prospectively recruited DCM subjects, linear modelling was used to more fully assess the relationship between *TTN* genotype and cardiac phenotype. For each phenotype a model adjusting for age and sex was optimised using Bayesian information criteria, then compared with a similarly optimised

model that also included *TTN* genotype (defined by presence/absence of TTNtv and the distance of the TTNtv from the N-terminus) using ANOVA. *TTN* genotype was a significant predictor of five phenotypic indices (LV EF, RV EF, LV SVi, RV SVi and lateral WTi).

Stratification of community-based cohorts and linear modelling—FHS Offspring subjects were studied by echocardiography as part of exams 2 (N_total=3,787; N_sequenced=1,360), 4 (N_total=3,798; N_sequenced=1,517), 5 (N_total=3608; N_sequenced=1,514), 6 (N_total=3,335; N_sequenced=1,522), and 8 (N_total=2,898; N_sequenced=1,289). LV ejection fraction (LV_EF) was estimated as ($LVEDD^2$ – $LVESD^2$) / $LVEDD^2$.

To adjust each measure, linear regression models were built including as potential covariates age, sex, ln(weight), BSA, height, systolic blood pressure (SBP), diastolic blood pressure (DBP), hypertension status (HTN), diabetes status (DM), and HTN treatment status (HTN_tx), as well as interactions between age or sex and each other considered covariate. Diabetes status and hypertension treatment status were excluded from FHS exam 8 because these clinically assessed summary data were not available. Each model was built with all covariates and then step-wise optimised to minimise Bayesian information criterion. Final models are denoted as the first model listed for each cohort, exam, CMR/echocardiography combination in Supplementary Information.

Atop these baseline models, we added TTNtv status or TTNtv status plus TTNtv exon usage (PSI). The latter model is detailed in the Supplementary Information; in the tables, models including TTNtv status plus exon expression immediately follow each paired base model. While adding only TTNtv status to baseline models did not improve overall model prediction, adding TTNtv status plus TTNtv exon expression did appear to improve model performance for CMR data (Tables S16–17).

DCM outcome analysis—Kaplan-Meier and Fleming-Harrinton estimates were used to compare time to events between unselected DCM cohort subgroups (TTNtv+ve and TTNtv −ve). Models were right-censored at age 70, last contact time if aged <70, or the earliest adverse event time recorded for that patient. (Fig. S5)

FHS outcome analyses—Cox proportional hazard models were used to compare time to events between FHS cohort subgroups. Models were left-censored at exam 5 and right-censored at ‘cardiovascular disease time’ if there were no events recorded; ‘last contact time’, if available; or the latest event time recorded across all FHS Offspring subjects in the survival table. Multivariate models were generated and optimised, as described in stratification of community-based cohorts and linear modelling (above), initially considering potential covariates of age, sex, HDL cholesterol, total cholesterol, ln(triglycerides), BMI, systolic blood pressure, diastolic blood pressure, anti-hyperlipidemia treatment status, anti-hypertension treatment status, and all pairwise interactions between age or sex and the other considered covariates (Table S13 and Fig. S4).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all the patients, healthy volunteers and participants in the FHS, JHS and WHI for taking part in this research, and our team of research nurses across the hospital sites.

Funding:

The research was supported by the NIHR Biomedical Research Unit in Cardiovascular Disease at Royal Brompton & Harefield NHS Foundation Trust and Imperial College London, NIHR Imperial Biomedical Research Centre, British Heart Foundation (BHF) UK [SP/10/10/28431], European Molecular Biology Laboratory, Medical Research Council UK, Wellcome Trust UK [087183/Z/08/Z; 092854/Z/10/Z; WT095908], Fondation Leducq, Tanoto Foundation, Goh Foundation, Academy of Medical Sciences, Arthritis Research UK, Heart Research UK, CORDA, NMRC Singapore, Rosetrees Trust, European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement number 200754 – the GEN2PHEN project, National Human Genome Research Institute [U54 HG003067], National Institutes of Health (NIH) [HL080494; 5-T32-GM007748-33], Howard Hughes Medical Institute, and the Australian National Health and Medical Research Council. The Framingham Heart Study was supported by the National Heart, Lung and Blood Institute (NHLBI) [N01-HC-25195; 6R01-NS 17950] and genotyping services from Affymetrix, Inc. [N02-HL-6-4278]. The Jackson Heart Study is supported by NHLBI [N01-HC-95170; N01-HC-95171; N01-HC-95172], the National Institute for Minority Health and Health Disparities, and the National Institute of Biomedical Imaging and Bioengineering. The Women's Health Initiative Sequencing Project is supported by NHLBI [HL-102924], NIH, and U.S. Department of Health and Human Services through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221. This publication reflects only the author's views and the funders are not liable for any use that may be made of the information contained herein.

References and Notes

1. McNally EM, Golbus JR, Puckelwartz MJ. Genetic mutations and mechanisms in dilated cardiomyopathy. J Clin Invest. 2013; 123:19–26. [PubMed: 23281406]
2. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. Nature reviews Cardiology. 2013; 10:531–547. [PubMed: 23900355]
3. Ware JS, Roberts AM, Cook SA. Next generation sequencing for clinical diagnostics and personalised medicine: implications for the next generation cardiologist. Heart (British Cardiac Society). 2012; 98:276–281. [PubMed: 22128206]
4. Lakdawala NK, Funke BH, Baxter S, Cirino AL, Roberts AE, Judge DP, Johnson N, Mendelsohn NJ, Morel C, Care M, Chung WK, Jones C, Psychogios A, Duffy E, Rehm HL, White E, Seidman JG, Seidman CE, Ho CY. Genetic testing for dilated cardiomyopathy in clinical practice. Journal of cardiac failure. 2012; 18:296–303. [PubMed: 22464770]
5. Siegfried JD, Morales A, Kushner JD, Burkett E, Cowan J, Mauro AC, Huggins GS, Li D, Norton N, Hershberger RE. Return of genetic results in the familial dilated cardiomyopathy research project. Journal of genetic counseling. 2013; 22:164–174. [PubMed: 22886719]
6. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, Camm AJ, Ellinor PT, Gollob M, Hamilton R, Hershberger RE, Judge DP, Le Marec H, McKenna WJ, Schulze-Bahr E, Semsarian C, Towbin JA, Watkins H, Wilde A, Wolpert C, Zipes DP. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). Heart rhythm: the official journal of the Heart Rhythm Society. 2011; 8:1308–1339. [PubMed: 21787999]
7. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'Meara S, Latimer C, Dicks E, Menzies A, Stephens P, Blow M, Greenman C, Xue Y, Tyler-Smith C, Thompson D, Gray K, Andrews J, Barthorpe S, Buck G, Cole J, Dunmore R, Jones D, Maddison M, Mironenko T, Turner R, Turrell K, Varian J, West S, Widaa S, Wray P, Teague J, Butler A, Jenkinson A, Jia M, Richardson D, Shepherd R, Wooster R, Tejada MI, Martinez F, Carvill G, Goliath R, de Brouwer AP, van Bokhoven H, Van Esch H, Chelly J, Raynaud M, Ropers HH, Abidi FE, Srivastava AK,

Cox J, Luo Y, Mallya U, Moon J, Parnau J, Mohammed S, Tolmie JL, Shoubridge C, Corbett M, Gardner A, Haan E, Rujirabanjerd S, Shaw M, Vandeleur L, Fullston T, Easton DF, Boyle J, Partington M, Hackett A, Field M, Skinner C, Stevenson RE, Bobrow M, Turner G, Schwartz CE, Gecz J, Raymond FL, Futreal PA, Stratton MR. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. Nature genetics. 2009; 41:535–543. [PubMed: 19377476]

8. Bick AG, Flannick J, Ito K, Cheng S, Vasan RS, Parfenov MG, Herman DS, DePalma SR, Gupta N, Gabriel SB, Funke BH, Rehm HL, Benjamin EJ, Aragam J, Taylor HA Jr, Fox ER, Newton-Cheh C, Kathiresan S, O'Donnell CJ, Wilson JG, Altshuler DM, Hirschhorn JN, Seidman JG, Seidman C. Burden of rare sarcomere gene variants in the Framingham and Jackson Heart Study cohorts. American journal of human genetics. 2012; 91:513–519. [PubMed: 22958901]
9. Flannick J, Beer NL, Bick AG, Agarwala V, Molnes J, Gupta N, Burtt NP, Florez JC, Meigs JB, Taylor H, Lyssenko V, Irgens H, Fox E, Burslem F, Johansson S, Brosnan MJ, Trimmer JK, Newton-Cheh C, Tuomi T, Molven A, Wilson JG, O'Donnell CJ, Kathiresan S, Hirschhorn JN, Njolstad PR, Rolph T, Seidman JG, Gabriel S, Cox DR, Seidman CE, Groop L, Altshuler D. Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. Nature genetics. 2013; 45:1380–1385. [PubMed: 24097065]
10. Siu BL, Niimura H, Osborne JA, Fatkin D, MacRae C, Solomon S, Benson DW, Seidman JG, Seidman CE. Familial dilated cardiomyopathy locus maps to chromosome 2q31. Circulation. 1999; 99:1022–1026. [PubMed: 10051295]
11. Gerull B, Gramlich M, Atherton J, McNabb M, Trombitas K, Sasse-Klaassen S, Seidman JG, Seidman C, Granzier H, Labeit S, Frenneaux M, Thierfelder L. Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. Nat Genet. 2002; 30:201–204. [PubMed: 11788824]
12. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, Conner L, DePalma SR, McDonough B, Sparks E, Teodorescu DL, Cirino AL, Banner NR, Pennell DJ, Graw S, Merlo M, Di Lenarda A, Sinagra G, Bos JM, Ackerman MJ, Mitchell RN, Murry CE, Lakdawala NK, Ho CY, Barton PJ, Cook SA, Mestroni L, Seidman JG, Seidman CE. Truncations of titin causing dilated cardiomyopathy. The New England journal of medicine. 2012; 366:619–628. [PubMed: 22335739]
13. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. Circulation Cardiovascular genetics. 2012; 5:391–399. [PubMed: 22763267]
14. Pan S, Caleshu CA, Dunn KE, Foti MJ, Moran MK, Soyinka O, Ashley EA. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. Circ Cardiovasc Genet. 2012; 5:602–610. [PubMed: 23074333]
15. Kontrogianni-Konstantopoulos A, Ackermann MA, Bowman AL, Yap SV, Bloch RJ. Muscle giants: molecular scaffolds in sarcomerogenesis. Physiological reviews. 2009; 89:1217–1267. [PubMed: 19789381]
16. Gregorio CC, Trombitas K, Centner T, Kolmerer B, Stier G, Kunke K, Suzuki K, Obermayr F, Herrmann B, Granzier H, Sorimachi H, Labeit S. The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. The Journal of cell biology. 1998; 143:1013–1027. [PubMed: 9817758]
17. Linke WA, Granzier H. A spring tale: new facts on titin elasticity. Biophysical journal. 1998; 75:2613–2614. [PubMed: 9826585]
18. Leinwand LA, Tardiff JC, Gregorio CC. Mutations in the sensitive giant titin result in a broken heart. Circulation research. 2012; 111:158–161. [PubMed: 22773424]
19. Puchner EM, Alexandrovich A, Kho AL, Hensen U, Schafer LV, Brandmeier B, Grater F, Grubmuller H, Gaub HE, Gautel M. Mechanoenzymatics of titin kinase. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:13385–13390. [PubMed: 18765796]
20. LeWinter MM, Granzier HL. Titin is a major human disease gene. Circulation. 2013; 127:938–944. [PubMed: 23439446]
21. Kellermayer MS, Smith SB, Bustamante C, Granzier HL. Complete unfolding of the titin molecule under external force. Journal of structural biology. 1998; 122:197–205. [PubMed: 9724621]

22. Govindaraju DR, Cupples LA, Kannel WB, O'Donnell CJ, Atwood LD, D'Agostino RB Sr, Fox CS, Larson M, Levy D, Murabito J, Vasan RS, Splansky GL, Wolf PA, Benjamin EJ. Genetics of the Framingham Heart Study population. Advances in genetics. 2008; 62:33–65. [PubMed: 19010253]

23. Wyatt SB, Diekelmann N, Henderson F, Andrew ME, Billingsley G, Felder SH, Fuqua S, Jackson PB. A community-driven model of research participation: the Jackson Heart Study Participant Recruitment and Retention Study. Ethnicity & disease. 2003; 13:438–455. [PubMed: 14632263]

24. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. Controlled clinical trials. 1998; 19:61–109. [PubMed: 9492970]

25. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurles ME, Gerstein MB, Tyler-Smith C. A systematic survey of loss-of-function variants in human protein-coding genes. Science (New York, NY). 2012; 335:823–828.

26. Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. Circulation research. 2001; 89:1065–1072. [PubMed: 11717165]

27. Moore HM. Acquisition of normal tissues for the GTEx program. Biopreserv Biobank. 2013; 11:75–76. [PubMed: 24845427]

28. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]

29. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, Chinwalla A, Conrad DF, Fu Y, Grubert F, Hajirasouliha I, Hormozdiari F, Iakoucheva LM, Iqbal Z, Kang S, Kidd JM, Konkel MK, Korn J, Khurana E, Kural D, Lam HY, Leng J, Li R, Li Y, Lin CY, Luo R, Mu XJ, Nemesh J, Peckham HE, Rausch T, Scally A, Shi X, Stromberg MP, Stutz AM, Urban AE, Walker JA, Wu J, Zhang Y, Zhang ZD, Batzer MA, Ding L, Marth GT, McVean G, Sebat J, Snyder M, Wang J, Ye K, Eichler EE, Gerstein MB, Hurles ME, Lee C, McCarroll SA, Korbel JO. Mapping copy number variation by population-scale genome sequencing. Nature. 2011; 470:59–65. [PubMed: 21293372]

30. Villard E, Perret C, Gary Fo, Proust C, Dilanian G, Hengstenberg C, Ruppert V, Arbustini E, Wichter T, Germain M, Dubourg O, Tavazzi L, Aumont M-C, DeGroote P, Fauchier L, Trochu J-Nl, Gibelin P, Aupetit J-Fo, Stark K, Erdmann J, Hetzer R, Roberts AM, Barton PJR, Regitz-Zagrosek V, Aslam U, Duboscq-Bidot Lt, Meyborg M, Maisch B, Madeira H, WaldenstrÃ¶m A, Galve E, Cleland JG, Dorent R, Roizes G, Zeller T, Blankenberg S, Goodall AH, Cook S, Tregouet DA, Tiret L, Isnard R, Komajda M, Charron P, Cambien Fo. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. European Heart Journal. 2011; 32:1065–1076. [PubMed: 21459883]

31. Stark K, Esslinger UB, Reinhard W, Petrov G, Winkler T, Komajda M, Isnard R, Charron P, Villard E, Cambien F, Tiret L, Aumont MC, Dubourg O, Trochu JN, Fauchier L, Degroote P, Richter A, Maisch B, Wichter T, Zollbrecht C, Grassl M, Schunkert H, Linsel-Nitschke P, Erdmann J, Baumert J, Illig T, Klopp N, Wichmann HE, Meisinger C, Koenig W, Lichtner P, Meitinger T, Schillert A, Konig IR, Hetzer R, Heid IM, Regitz-Zagrosek V, Hengstenberg C. Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. PLoS genetics. 2010; 6:e1001167. [PubMed: 20975947]

32. Keenan NG, Pennell DJ. CMR of ventricular function. Echocardiography (Mount Kisco, NY). 2007; 24:185–193.

33. Parsai C, O'Hanlon R, Prasad SK, Mohiaddin RH. Diagnostic and prognostic value of cardiovascular magnetic resonance in non-ischaemic cardiomyopathies. J Cardiovasc Magn Reson. 2012; 14:54. [PubMed: 22857649]

34. Gulati A, Jabbour A, Ismail TF, Guha K, Khwaja J, Raza S, Morarji K, Brown TD, Ismail NA, Dweck MR, Di Pietro E, Roughton M, Wage R, Daryani Y, O'Hanlon R, Sheppard MN, Alpendurada F, Lyon AR, Cook SA, Cowie MR, Assomull RG, Pennell DJ, Prasad SK. Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy. JAMA: the journal of the American Medical Association. 2013; 309:896–908. [PubMed: 23462786]

35. Assomull RG, Prasad SK, Lyne J, Smith G, Burman ED, Khan M, Sheppard MN, Poole-Wilson PA, Pennell DJ. Cardiovascular magnetic resonance, fibrosis, and prognosis in dilated cardiomyopathy. Journal of the American College of Cardiology. 2006; 48:1977–1985. [PubMed: 17112987]

36. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genetics in medicine: official journal of the American College of Medical Genetics. 2013; 15:565–574. [PubMed: 23788249]

37. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MR, Towbin JA. Genetic evaluation of cardiomyopathy--a Heart Failure Society of America practice guideline. Journal of cardiac failure. 2009; 15:83–97. [PubMed: 19254666]

38. Kushner JD, Nauman D, Burgess D, Ludwigsen S, Parks SB, Pantely G, Burkett E, Hershberger RE. Clinical Characteristics of 304 Kindreds Evaluated for Familial Dilated Cardiomyopathy. Journal of cardiac failure. 2006; 12:422–429. [PubMed: 16911908]

39. Richardson P, McKenna W, Bristow M, Maisch B, Mautner B, O'Connell J, Olsen E, Thiene G, Goodwin J, Gyarfas I, Martin I, Nordet P. Report of the 1995 World Health Organization/ International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. Circulation. 1996; 93:841–842. [PubMed: 8598070]

40. McNally EM. Genetics: broken giant linked to heart failure. Nature. 2012; 483:281–282. [PubMed: 22422258]

41. Roncarati R, Viviani Anselmi C, Krawitz P, Lattanzi G, von Kodolitsch Y, Perrot A, di Pasquale E, Papa L, Portararo P, Columbaro M, Forni A, Faggian G, Condorelli G, Robinson PN. Doubly heterozygous LMNA and TTN mutations revealed by exome sequencing in a severe form of dilated cardiomyopathy. European journal of human genetics: EJHG. 2013

42. Robinson PN. Deep phenotyping for precision medicine. Human mutation. 2012; 33:777–780. [PubMed: 22504886]

43. Pahl E, Sleeper LA, Canter CE, Hsu DT, Lu M, Webber SA, Colan SD, Kantor PF, Everitt MD, Towbin JA, Jefferies JL, Kaufman BD, Wilkinson JD, Lipshultz SE. Incidence of and risk factors for sudden cardiac death in children with dilated cardiomyopathy: a report from the Pediatric Cardiomyopathy Registry. Journal of the American College of Cardiology. 2012; 59:607–615. [PubMed: 22300696]

44. Solomon SD, Anavekar N, Skali H, McMurray JJ, Swedberg K, Yusuf S, Granger CB, Michelson EL, Wang D, Pocock S, Pfeffer MA. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. Circulation. 2005; 112:3738–3744. [PubMed: 16330684]

45. Gjesdal O, Bluemke DA, Lima JA. Cardiac remodeling at the population level--risk factors, screening, and outcomes. Nature reviews Cardiology. 2011; 8:673–685. [PubMed: 22027657]

46. Pasotti M, Klersy C, Pilotto A, Marziliano N, Rapezzi C, Serio A, Mannarino S, Gambarin F, Favalli V, Grasso M, Agozzino M, Campana C, Gavazzi A, Febo O, Marini M, Landolina M, Mortara A, Piccolo G, Vigano M, Tavazzi L, Arbustini E. Long-term outcome and risk stratification in dilated cardiolaminopathies. Journal of the American College of Cardiology. 2008; 52:1250–1260. [PubMed: 18926329]

47. Prontera P, Bernardini L, Stangoni G, Capalbo A, Rogaia D, Ardisia C, Novelli A, Dallapiccola B, Donti E. 2q31.2q32.3 deletion syndrome: report of an adult patient. American journal of medical genetics Part A. 2009; 149A:706–712. [PubMed: 19248183]

48. Ceyhan-Birsoy O, Agrawal PB, Hidalgo C, Schmitz-Abe K, DeChene ET, Swanson LC, Soemedi R, Vasli N, Iannaccone ST, Shieh PB, Shur N, Dennison JM, Lawlor MW, Laporte J, Markianos

K, Fairbrother WG, Granzier H, Beggs AH. Recessive truncating titin gene, TTN, mutations presenting as centronuclear myopathy. Neurology. 2013; 81:1205–1214. [PubMed: 23975875]

49. Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R, Robbins J. A mouse model of myosin binding protein C human familial hypertrophic cardiomyopathy. The Journal of clinical investigation. 1998; 102:1292–1300. [PubMed: 9769321]
50. Sarikas A, Carrier L, Schenke C, Doll D, Flavigny J, Lindenberg KS, Eschenhagen T, Zolk O. Impairment of the ubiquitin-proteasome system by truncated cardiac myosin binding protein C mutants. Cardiovascular research. 2005; 66:33–44. [PubMed: 15769446]
51. Maceira AM, Prasad SK, Khan M, Pennell DJ. Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance. J Cardiovasc Magn Reson. 2006; 8:417–426. [PubMed: 16755827]
52. Grothues F, Smith GC, Moon JC, Bellenger NG, Collins P, Klein HU, Pennell DJ. Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. The American journal of cardiology. 2002; 90:29–34. [PubMed: 12088775]
53. Gulati A, Ismail TF, Jabbour A, Ismail NA, Morarji K, Ali A, Raza S, Khwaja J, Brown TD, Liodakis E, Baksi AJ, Shakur R, Guha K, Roughton M, Wage R, Cook SA, Alpendurada F, Assomull RG, Mohiaddin RH, Cowie MR, Pennell DJ, Prasad SK. Clinical utility and prognostic value of left atrial volume assessment by cardiovascular magnetic resonance in non-ischaemic dilated cardiomyopathy. European journal of heart failure. 2013; 15:660–670. [PubMed: 23475781]
54. Taylor MR, Carniel E, Mestroni L. Cardiomyopathy, familial dilated. Orphanet journal of rare diseases. 2006; 1:27. [PubMed: 16839424]
55. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010; 20:1297–1303. [PubMed: 20644199]
56. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England). 2009; 25:1754–1760.
57. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science (New York, NY). 2012; 337:64–69.
58. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics (Oxford, England). 2009; 25:1105–1111.
59. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Giron CG, Gordon L, Hourlier T, Hunt S, Johnson N, Juettemann T, Kahari AK, Keenan S, Kulesha E, Martin FJ, Maurel T, McLaren WM, Murphy DN, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ruffier M, Sheppard D, Taylor K, Thormann A, Trevanion SJ, Vullo A, Wilder SP, Wilson M, Zadissa A, Aken BL, Birney E, Cunningham F, Harrow J, Herrero J, Hubbard TJ, Kinsella R, Muffato M, Parker A, Spudich G, Yates A, Zerbino DR, Searle SM. Ensembl 2014. Nucleic acids research. 2014; 42:D749–755. [PubMed: 24316576]
60. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England). 2009; 25:2078–2079.
61. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008; 456:470–476. [PubMed: 18978772]
62. Warren CM, Krzesinski PR, Greaser ML. Vertical agarose gel electrophoresis and electroblotting of high-molecular-weight proteins. Electrophoresis. 2003; 24:1695–1702. [PubMed: 12783444]
63. MacArthur JA, Morales J, Tully RE, Astashyn A, Gil L, Bruford EA, Larsson P, Flicek P, Dalgleish R, Maglott DR, Cunningham F. Locus Reference Genomic: reference sequences for the reporting of clinically relevant sequence variants. Nucleic acids research. 2014; 42:D873–878. [PubMed: 24285302]

64. Rios D, McLaren WM, Chen Y, Birney E, Stabenau A, Flicek P, Cunningham F. A database and API for variation, dense genotyping and resequencing data. BMC bioinformatics. 2010; 11:238. [PubMed: 20459810]

65. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics (Oxford, England). 2010; 26:2069–2070.

66. Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Leone M, Mazoyer S, Muller D, Remenieras A, Revillion F, Rouleau E, Sokolowska J, Vert JP, Lidereau R, Soubrier F, Sobol H, Sevenet N, Bressac-de Paillerets B, Hardouin A, Tosi M, Sinilnikova OM, Stoppa-Lyonnet D. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. Human mutation. 2012; 33:1228–1238. [PubMed: 22505045]

67. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. Journal of computational biology: a journal of computational molecular cell biology. 2004; 11:377–394. [PubMed: 15285897]

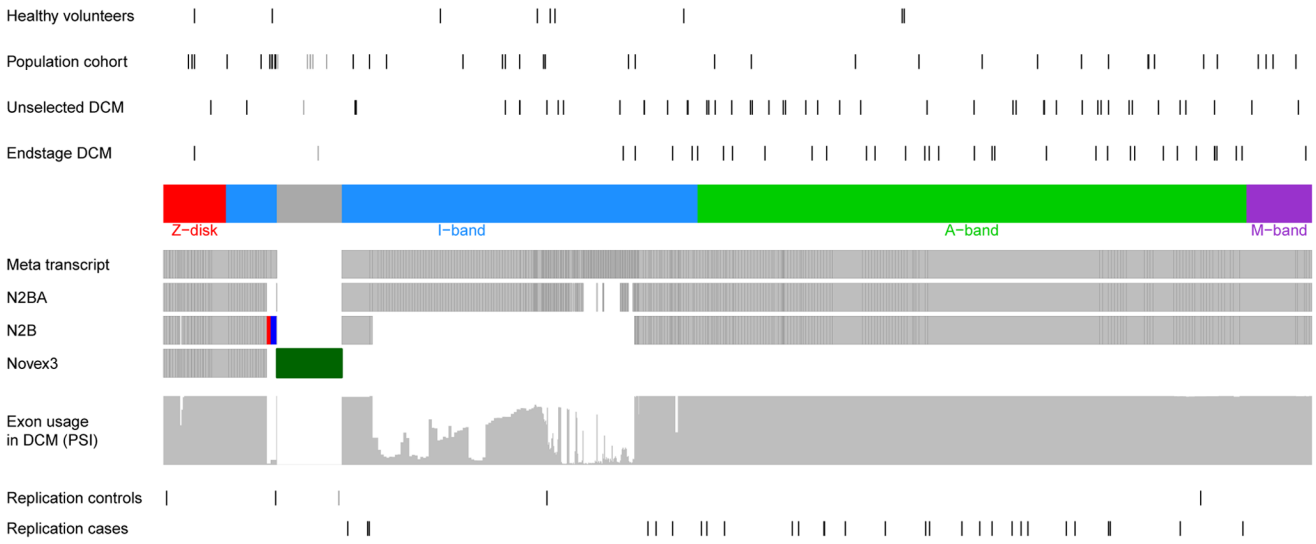

Fig. 1. Distribution of TTNtv in healthy individuals and DCM patients, and *TTN* exon usage in the heart

A schematic of the *TTN* meta-transcript is shown, with sarcomere regions demarcated. The meta-transcript (LRG_391_t1 / ENST00000589042) is a manually curated inferred complete transcript, incorporating all exons of all known *TTN* isoforms (including fetal and non-cardiac isoforms) with the exception of the large alternative terminal exon 48 (dark green) that is unique to the novex-3 transcript (LRG_391_t2 / ENST00000360870). Exon usage for the two principal adult cardiac isoforms, N2BA and N2B (ENST00000591111, ENST00000460472) is shown, though exon usage in vivo is variable (see below). Novex-1 and novex-2 are rare cardiac isoforms that differ from N2B by the inclusion of a single unique exon each (red and blue respectively within the N2B track). Exon usage in human LV is depicted as the “proportion spliced-in” (PSI, range 0–1; grey bars): the proportion of transcripts that include a given exon. TTNtv are located more distally in cases compared with controls, with A-band and distal I-band enrichment in end-stage (n=155) and unselected DCM patients (n=374) and corresponding depletion in the population (n=3603) and healthy volunteer (n=308) cohorts.

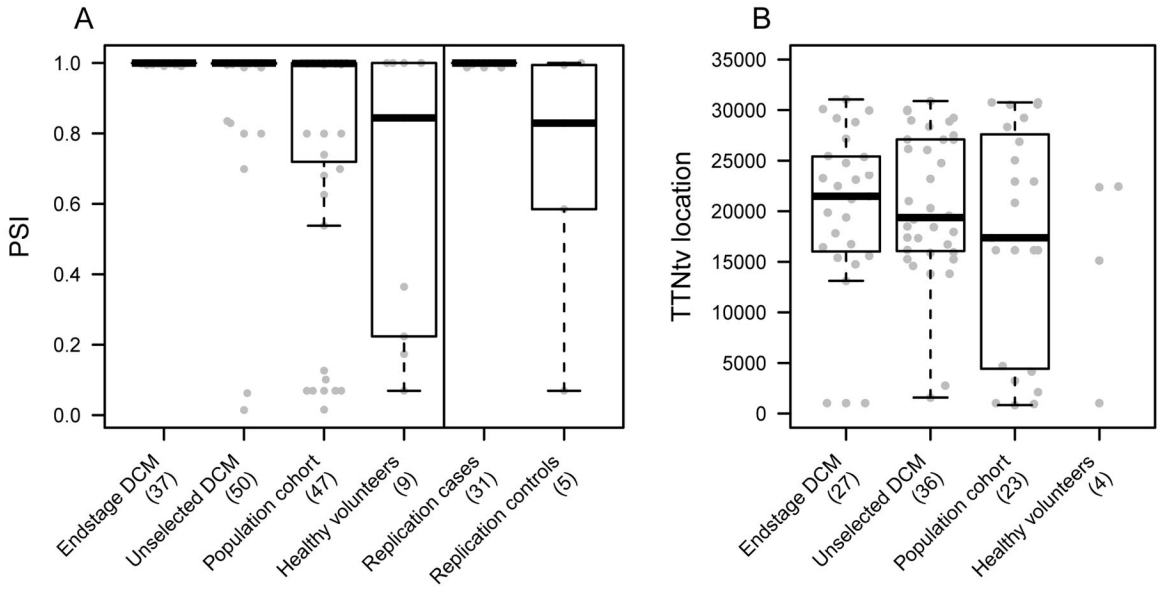

Fig. 2. Factors that discriminate TTNtv in health and disease

A: Usage of *TTN* exons containing TTNtv across all cohorts. Exon usage is represented as proportion spliced-in (PSI), which is an estimate of the proportion of transcripts that incorporate each exon. Each plotted data point represents the estimated PSI of an exon identified to have a TTNtv, separated by cohort. There was a strong relationship between the PSI of exons containing TTNtv and disease status ($P=4.9\times10^{-3}$, Kruskal-Wallis), with TTNtv in DCM cases found in more highly used exons ($P = 4.7\times10^{-4}$, Mann-Whitney). A similar difference was observed between the replication cohorts ($P = 7.5\times10^{-4}$). B: Relationships between TTNtv location, PSI and disease status. The positions of TTNtv are shown for constitutively expressed exons only (PSI=1).

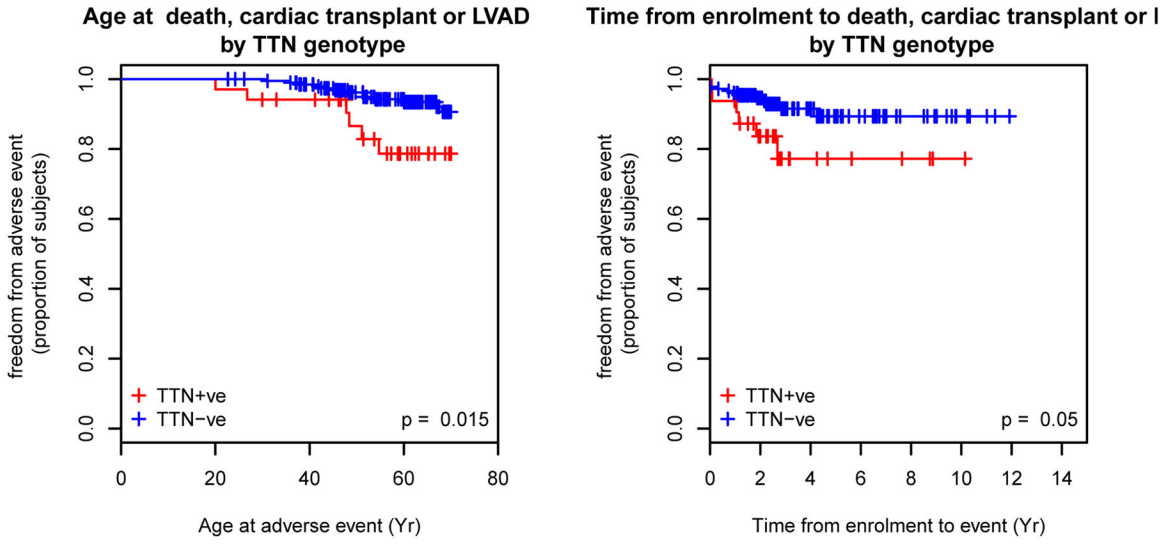

Fig. 3. TTNtv and survival in DCM

Outcomes in unselected DCM patients with (red) and without (blue) TTNtv. The left panel shows age censored at adverse event (death, cardiac transplant or left-ventricular assist device (LVAD) or at age 70 years. The right panel shows adverse events after enrollment, to control for ascertainment (interval censored from time of enrollment to age 70 years or adverse event). Event free survival is reduced in TTNtv-positive DCM ($P = 0.015$) due to altered disease progression both before and after presentation. A trend to younger presentation (Table 2), and worse outcomes after enrollment ($P = 0.05$) combine to give reduced survival overall.

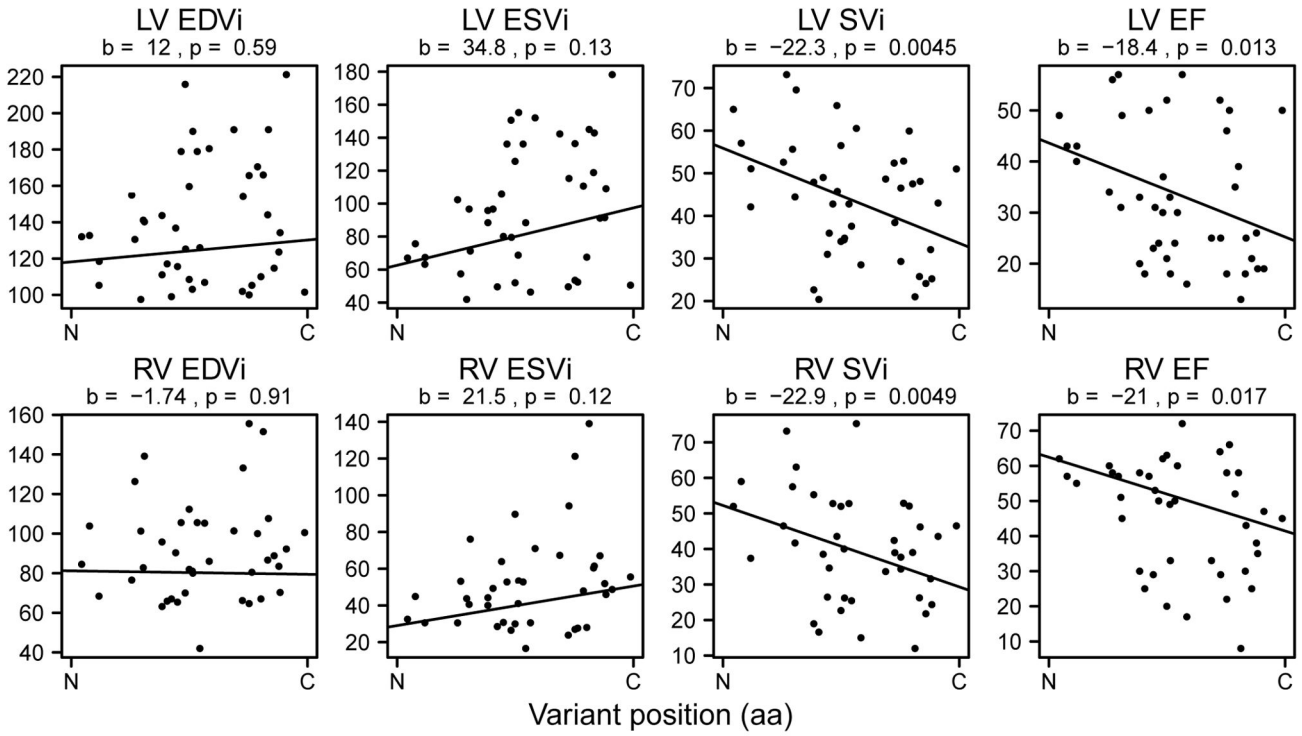

Fig. 4. Allelic dissection of the impact of TTNtv position on cardiac morphology and function

The relationships between TTNtv location and cardiac morphology and function assessed by CMR imaging in an allelic series of DCM cases. Genotype-phenotype relationships are shown for 43 TTNtv in unselected DCM patients. The TTNtv location (X axis) is plotted from the amino- (N) to carboxyl- (C) end of the protein. Distal (C-terminal) TTNtv were associated with worse cardiac contractile performance and associated with diminished indexed stroke volume (SVi) and ejection fraction (EF) of both left and right ventricles as compared to proximal truncations. A regression line is shown for each variable (Tables S14–S15). LV, left ventricle; RV, right ventricle; EDVi, indexed end diastolic stroke volume (ml/m2); ESVi, indexed end systolic volume (ml/m2), SVi, indexed stroke volume (ml/m2), EF ejection fraction (%).

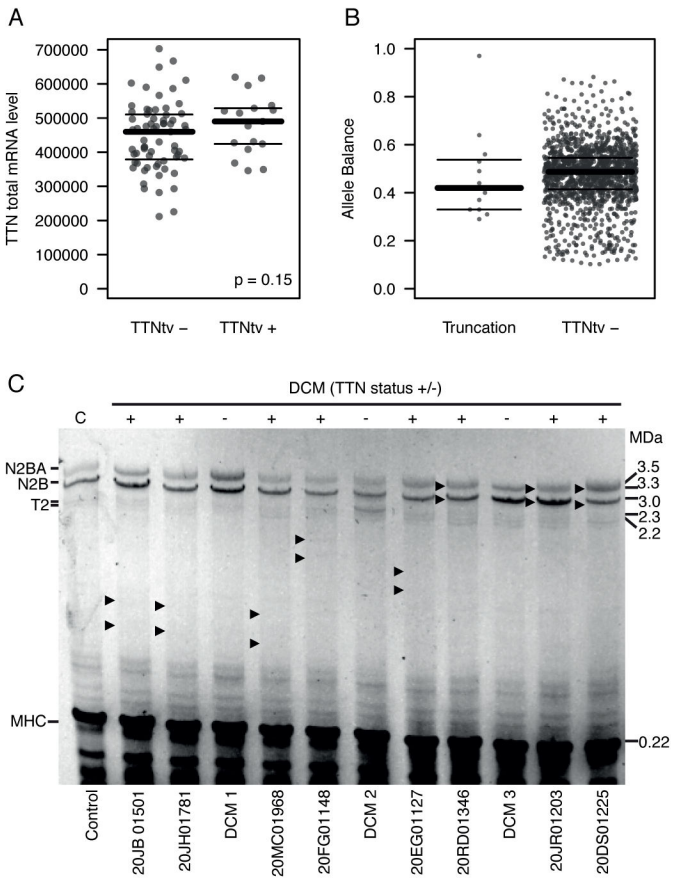

Fig. 5. *TTN* mRNA and protein expression in LV tissues from DCM patients with and without TTNtv

A. Levels of *TTN* mRNA in TTNtv-positive (n=18) and negative (n=66) patients (normalised read counts). B. Allelic balance of TTNtv compared to non-truncating *TTN* single nucleotide polymorphisms (SNPs), a surrogate for the proportion of transcripts with variant alleles (TTNtv or SNPs) among DCM patients with and without TTNtvs. The comparable allelic expression of TTNtv and SNPs does not support substantial nonsense-mediated decay (see also Fig. S9). Bars indicate median and quartiles. C. Protein electrophoresis from a healthy LV (C, lane 1) and LV from DCM patients (lanes 2–12: +, TTNtv-positive; −, TTNtv-negative). Sample IDs are shown for subjects with TTNtv: variant details are shown in Table S4. Truncated protein was not seen in TTNtv-positive samples. The arrows indicate approximate expected sizes of the truncated N2B and N2BA isoforms and T2 denotes TTN degradation product. Semi-quantitative analysis of TTN protein relative to myosin (MHC, myosin heavy chain) showed no reduction of TTN in TTNtv-positive samples.

Table 1

Burden of TTN truncating variants in DCM patients and controls

	Discovery Cohorts					P value (Discovery) DCM vs controls		Replication Cohorts		
	Healthy volunteers n=308	FHS n=1623	JHS n=1980	Unselected DCM n=371[a]	End-stage DCM n=155	Unselected DCM	All DCM	WHI n=667	Familial DCM n=163	P value DCM vs controls
Transcript affected by truncation										
N2BA & N2B	4	11	20	42	34	1.7×10^{-25}	5.5×10^{-46}	2	31	3.8×10^{-21}
N2BA only	4	2	8	7	0	0.0014	0.0083	2	0	1
Neither N2BA or N2B (Novex-3 terminal exon only[b])	1 (0)	8 (5)	4 (1)	1 (1)	1 (1)	1	0.7	5 (4)	0 (0)	1
Totals	9	21	32	50	35			9	31	
Sarcomere domain										
A-band (18235aa)	2	7	12	32	29	0.011	0.00014	1	25	0.015
Non A-band (17756aa)	7	9	19	18	8			4	6	
Totals	9	16	31	49 [f]	34 [f]			5	31	
Usage[c] of exon containing truncation										
Low (PSI <0.15)	1	4	4	2	0	0.25	0.38	1	0	1
Intermediate (PSI 0.15–0.9)	4	1	7	5	0	0.012	0.042	2	0	1
High (PSI>0.9)	4	11	20	42	34	1.7×10^{-25}	5.5×10^{-46}	2	31	3.8×10^{-21}
Totals	9	16	31	49[f]	34[f]			5	31	
Variant type										
Frameshift variant	2	3	7	22	15	4.9×10^{-16}	4.0×10^{-25}	1	15	1.9×10^{-10}
Stop gained	3	7	6	15	12	7.3×10^{-09}	2.3×10^{-15}	2	16	2.8×10^{-10}
Canonical splice sites[d]	2	3	7	9	5	2.9×10^{-05}	2.3×10^{-07}	1	0	1
Splice variant predictions[e]	2	3	11	4	5	0.089	0.0015	1	0	1
Totals	**9**	**16**	**31**	**49**[f]	**34**[f]	1.4×10^{-25}	2.8×10^{-43}	**5**	**31**	1.6×10^{-18}

Numbers of subjects with a TTNtv is shown for each group. TTNtv are classified by type, the affected transcript, and expression level of the variant-encoding exon.

Comparisons between groups were assessed by Fisher's exact test.

[a] 3/374 subjects were excluded from these analyses as they due to relatedness to other subjects.

[b] Variants that only impact the alternative terminal exon of novex-3 are excluded elsewhere.

[c] Exon usage levels are displayed categorically based on PSI

[d] Canonical splice sites refer to the 2 intronic base pairs at the 5′ and 3′ splice junctions

[e] Variants close to canonical splice sites that are predicted *in silico* to alter splicing

[f] total number of individuals with TTNtv; 4 individuals with DCM (1 unselected, 3 end-stage) carry a second TTNtv which is a splice variant prediction in all cases.

Cohort ethnicity: Caucasian: Healthy volunteer 75%, FHS 100%, JHS 0%, unselected DCM 88%, end-stage DCM 85%. African American: Healthy volunteer 2%, FHS 0%, JHS 100%, unselected DCM 4%, end-stage DCM 6%.

Table 2

Clinical characteristics of DCM patients with and without TTNtv

CMR and Clinical Data		TTNtv-negative (n=277)	TTNtv-positive (n=42)	P value
LV	EDVi	136 ± 38	137 ± 34.1	0.677
	ESVi	87.5 ± 39.1	93.7 ± 36.4	0.205
	SVi	48.2 ± 12.9	43.4 ± 14.1	**0.041**
	EF	37.5 ± 12.2	33.3 ± 13.3	**0.047**
RV	EDVi	89.4 ± 24.7	89.6 ± 26.1	0.972
	ESVi	45.2 ± 22.3	50.4 ± 25.9	0.234
	SVi	44.4 ± 12.6	39.2 ± 15.8	**0.031**
	EF	51.6 ± 14.1	45.2 ± 16	**0.036**
LVMi		95.4 ± 27.6	87.1 ± 18.3	0.106
Lateral WTi		3.13 ± 0.73	2.77 ± 0.713	**0.003**
Mid-wall fibrosis		94 / 270 (35%)	13 / 41 (32%)	0.869
Age at Diagnosis (years)		53.4 ± 13.4	49.3 ± 13.7	0.115
NYHA status 1/2/3/4		116/104/36/1	19/16/4/1	0.692
Sustained VT		20 / 97 (21%)	9 / 14 (64%)	**0.001**
Conduction Disease		82 / 227 (36%)	8 / 36 (22%)	0.130
Family history of DCM		24 / 218 (11%)	9 / 37 (24%)	**0.034**

Unselected DCM cohort. Values are means ± SD. Volume measurements are indexed to body surface area. LV, left ventricle; RV right ventricle; EDVi/ESVi end diastolic/systolic volume; SVi stoke volume; EF ejection fraction; LVMi LV mass; WTi wall thickness; VT ventricular tachycardia; NYHA New York Heart association functional class. P values, not corrected for multiple testing as not independent.