European Heart Journal (2015) **36**, 1123–1135
doi:10.1093/eurheartj/ehu301

BASIC SCIENCE

Atlas of the clinical genetics of human dilated cardiomyopathy

Jan Haas[1,2], Karen S. Frese[1,2], Barbara Peil[3], Wanda Kloos[1], Andreas Keller[4], Rouven Nietsch[1,2], Zhu Feng[1], Sabine Müller[4], Elham Kayvanpour[1,2], Britta Vogel[1], Farbod Sedaghat-Hamedani[1,2], Wei-Keat Lim[6], Xiaohong Zhao[6], Dmitriy Fradkin[6], Doreen Köhler[1], Simon Fischer[1], Jennifer Franke[1], Sabine Marquart[1,2], Ioana Barb[1,2], Daniel Tian Li[1,2], Ali Amr[1,2], Philipp Ehlermann[1], Derliz Mereles[1,2], Tanja Weis[1,2], Sarah Hassel[1,2], Andreas Kremer[7], Vanessa King[6], Emil Wirsz[6,5], Richard Isnard[11], Michel Komajda[11], Alessandra Serio[8], Maurizia Grasso[8], Petros Syrris[9], Eleanor Wicks[9], Vincent Plagnol[9], Luis Lopes[9], Tenna Gadgaard[13], Hans Eiskjær[13], Mads Jørgensen[19], Diego Garcia-Giustiniani[16], Martin Ortiz-Genga[16], Maria G. Crespo-Leiro[17], Rondal H. Lekanne Dit Deprez[10], Imke Christiaans[10], Ingrid A. van Rijsingen[10], Arthur A. Wilde[10], Anders Waldenstrom[18], Martino Bolognesi[15], Riccardo Bellazzi[14], Stellan Mörner[18], Justo Lorenzo Bermejo[3], Lorenzo Monserrat[16,17], Eric Villard[11], Jens Mogensen[12], Yigal M. Pinto[10], Philippe Charron[11], Perry Elliott[9], Eloisa Arbustini[8], Hugo A. Katus[1,2,20], and Benjamin Meder[1,2,20]*

[1]Department of Internal Medicine III, University of Heidelberg, Germany; [2]DZHK (German Centre for Cardiovascular Research), Germany; [3]Institute of Medical Biometry and Informatics (IMBI), University Hospital Heidelberg, Germany; [4]Department of Human Genetics, Saarland University, Germany; [5]Siemens AG, Erlangen, Germany; [6]Siemens Corporate Technology, Princeton, USA; [7]Siemens AG, Österreich; [8]Fondazione IRCCS Policlinico San Matteo, Pavia, Italy; [9]University College London, Great Britain; [10]Academisch Medisch Centrum (AMC), Amsterdam, Netherlands; [11]INSERM UMRS-956, UPMC Univ Paris 6, AP-HP, Hôpital Pitié-Salpêtrière, Paris, France; [12]Department of Cardiology, Odense University Hospital and Institute of Clinical Research, University of Southern Denmark; [13]Department of Cardiology, Aarhus University Hospital, Denmark; [14]Biomedical Informatics Laboratory, University of Pavia, Italy; [15]Department of BioSciences, University of Milano, Italy; [16]Health-in-code, La Coruna, Spain; [17]Cardiology Department, Biomedical Research Institute INIBIC, A Coruña, Spain; [18]Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; [19]Department of Clinical Genetics, Vejle Hospital, 7100 Vejle, Denmark; and [20]Klaus Tschira Institute for Computational Cardiology, University of Heidelberg, Germany

Online publish-ahead-of-print 27 August 2014

See page 1074 for the editorial comment on this article (doi:10.1093/eurheartj/ehu402)

Aim Numerous genes are known to cause dilated cardiomyopathy (DCM). However, until now technological limitations have hindered elucidation of the contribution of all clinically relevant disease genes to DCM phenotypes in larger cohorts. We now utilized next-generation sequencing to overcome these limitations and screened all DCM disease genes in a large cohort.

Methods and results In this multi-centre, multi-national study, we have enrolled 639 patients with sporadic or familial DCM. To all samples, we applied a standardized protocol for ultra-high coverage next-generation sequencing of 84 genes, leading to 99.1% coverage of the target region with at least 50-fold and a mean read depth of 2415. In this well characterized cohort, we find the highest number of known cardiomyopathy mutations in plakophilin-2, myosin-binding protein C-3, and desmoplakin. When we include yet unknown but predicted disease variants, we find titin, plakophilin-2, myosin-binding protein-C 3, desmoplakin, ryanodine receptor 2, desmocollin-2, desmoglein-2, and SCN5A variants among the most commonly mutated genes. The overlap between DCM, hypertrophic cardiomyopathy (HCM), and channelopathy causing mutations is considerably high. Of note, we find that >38% of patients have compound or combined mutations and 12.8% have three or even more mutations. When comparing patients recruited in the eight participating European countries we find remarkably little differences in mutation frequencies and affected genes.

* Corresponding author: Benjamin Meder, University Hospital of Heidelberg, Department of Internal Medicine III, Im Neuenheimer Feld 669, 69120 Heidelberg, Germany, Tel.: +496221564835, Fax: +496221564645, Email: benjamin.meder@med.uni-heidelberg.de

Conclusion This is to our knowledge, the first study that comprehensively investigated the genetics of DCM in a large-scale cohort and across a broad gene panel of the known DCM genes. Our results underline the high analytical quality and feasibility of Next-Generation Sequencing in clinical genetic diagnostics and provide a sound database of the genetic causes of DCM.

Keywords Cardiomyopathy • Genetics • Patients • Diagnosis

Translational Perspective

We were able to show that targeted Next-Generation Sequencing is well suited to be applied in clinical routine diagnostics, substantiating the ongoing paradigm shift from low- to high-throughput genomics in medicine. By means of our atlas of the genetics of human DCM, we aspire to soon be able to apply our findings to the individual patient with cardiomyopathy in daily clinical practice.

Introduction

Dilated cardiomyopathy (DCM) accounts for 30–40% of all heart failure cases in large clinical trials and is the leading cause of heart transplantation. There is ample data on the familial aggregation of DCM and in recent registers familial forms of DCM account for 30–50% of all DCM cases. With an autosomal-dominant inheritance being the predominant pattern of transmission, some familial cases also present by an autosomal recessive or X-linked recessive trait. Particularly in western countries, the small size of contemporary families may obscure the genetic nature of the disease and it is important to consider that also sporadic DCM cases can be due to genetic mutations.[1]

Tremendous advances have been made in understanding the genetic basis of DCM. Linkage analyses in families and candidate gene sequencing as well as genome-wide association studies (GWAS) in large cohorts[2] have contributed to the identification of risk variants and disease causing mutations in >30 disease genes, many of which encode for structural components of the heart muscle, such as the sarcomere or the cardiac z-disc. Recently, we and others have used Next-Generation Sequencing (NGS) approaches to dissect the genetic causes of DCM and established a comprehensive methodology for the clinical genetic testing of all currently known disease genes.[3] However, the existing studies are either limited by the small number of investigated patients or the restriction to only a subset of disease genes, prohibiting a more detailed dissection of the role of DNA-alterations in DCM.

We are here presenting the results of the gene sequencing study of the European INHERITANCE project including 639 patients with sporadic or proven familial DCM enrolled in eight different clinical centres (Denmark, Sweden, France, Italy, Germany, UK, Netherlands, and Spain). We aimed to systematically investigate not only the clinically relevant DCM genes but also genes causative for other inherited cardiomyopathies and to systematically benchmark the analytical performance of NGS as a novel technology being introduced into broad clinical application.

Methods

Patients and study design

This multi-centre study was conducted in accordance with the principles of the Declaration of Helsinki. All participants from all centres have given written informed consent and the study was approved by the ethic committees of the participating study centres.

Dilated cardiomyopathy was diagnosed according to the WHO/International Society and Federation of Cardiology Task Force clinical criteria,[4] defined as a myocardial disorder characterized by the presence of left ventricular dilatation and systolic impairment, in the absence of abnormal loading conditions (e.g. hypertension, valve disease) or coronary artery disease (CAD) sufficient to cause global systolic dysfunction.[5] Specific criteria were the presence of left ventricular dilatation (>117% of the predicted value corrected for age and body surface area using the Henry equation) and left ventricular systolic dysfunction (LVEF <45%) in the absence of abnormal loading conditions (e.g. hypertensive heart disease, primary valve disease) or CAD with stenosis >50% of at least one main vessel in coronary angiography. CAD was ruled out by coronary angiography in 53% of patients ($n = 338$). Patients having clinic suspicion or evidence for myocarditis [late-Gadolinium enhancement typical for myocarditis in cardiac MRI (69 patients (11%) of all patients undergone MRI), or evidence from myocardial biopsies (121 patients (19%) of all patients undergone biopsy)] or history of cardio-toxic therapies were excluded. Familial DCM was defined according to ESC definitions[5] and evaluated according to the position statement of the European Society of Cardiology, Working Group on Myocardial and Pericardial Diseases.[6] In detail, all familial DCM cases had at least one additional affected family member with DCM or one case of sudden cardiac death earlier 35 years of age in the pedigree. All criteria had to be met at an initial diagnosis of DCM in probands or relatives. Clinical data for association analysis were assessed during the index visit, where it was required that the patient was clinically stable for at least 4 weeks. The enrolment was at the index visit of patients between 2009 and 2011, NGS was performed between March 2011 and August 2013 after DNA was extracted at the recruiting centre and shipped pseudonymized to the INHERITANCE sequencing core centre (Department of Cardiology, University of Heidelberg, Germany).

Target enrichment, next-generation sequencing, and data analysis

SureSelect^{XT} Target Enrichment System (Agilent; Waldbronn, Germany) was used for capturing of the desired regions. Design of the capture baits was done using eArray (Agilent Technologies, Santa Clara, CA, USA). Raw data analysis was performed with an in-house pipeline based on variant calling of the Genome Analysis Toolkit (GATK). Annotation of the variants was mainly done using ANNOVAR. Genotype–phenotype association tests were done with the SAS software. For a detailed description of the data analysis, please refer to the Supplemental Material online.

Variant classification

In this study, we relied on distinct, well-defined categories for the classification of variants. Variants were classified as benign when present in

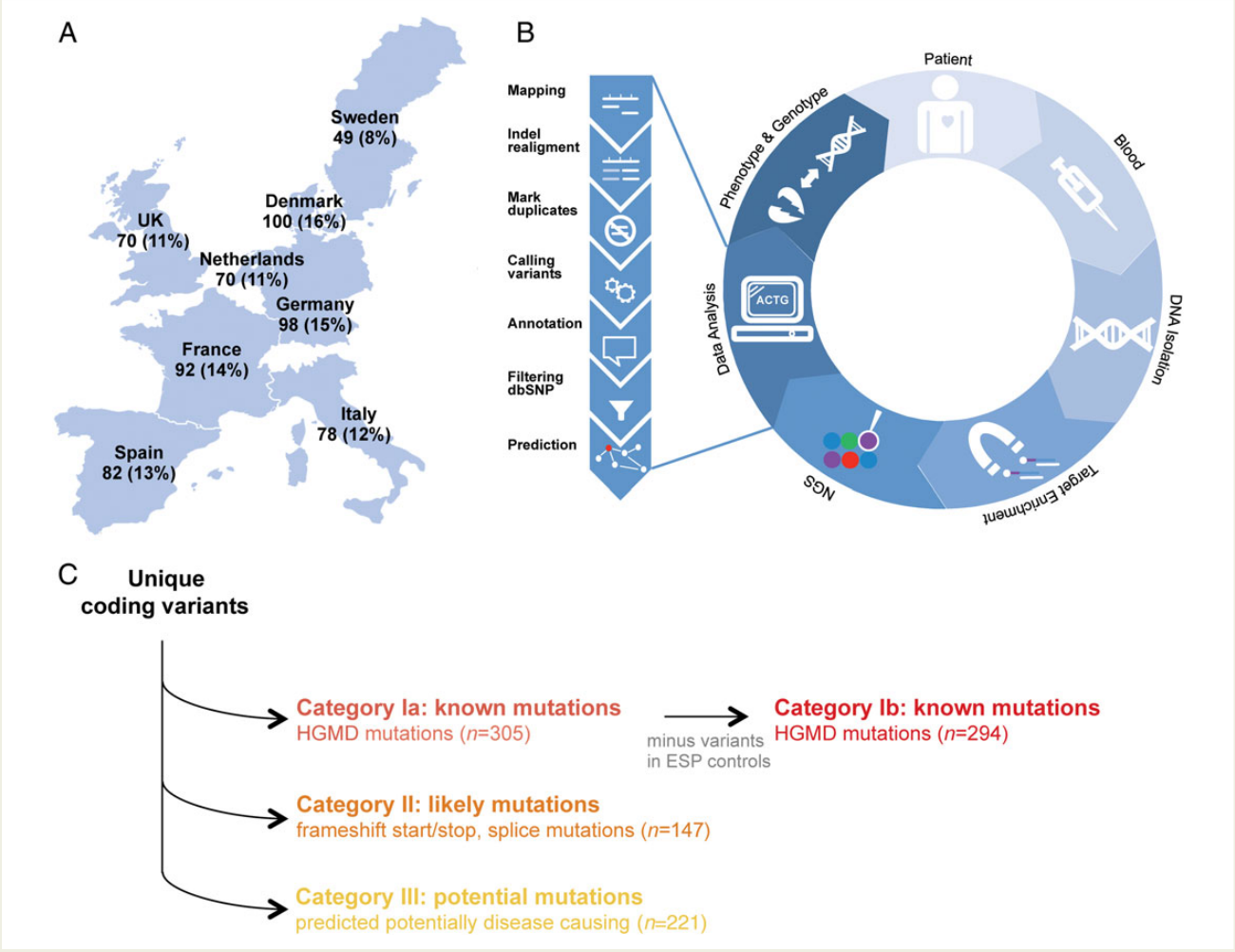

Figure 1 Study design. (*A*) The map is showing the participating European countries and the number of patients included (absolute and in brackets relative). Map (EU27-European_Union_map.svg) by Kolja21, used under CC BY. (*B*) Applied workflow, consisting of target-enrichment of cardiomyopathy related genes, next-generation sequencing and data analysis. (*C*) Overview of variant classification. Total number of patients in the whole cohort is given in brackets.

'dbSNP137common' (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/snp137Common.sql) and flagged as validated-by-frequency, which means that those variants have been found with an allele-frequency of ≥1% in populations. For further determination of the likelihood to being disease relevant mutations, we defined distinct categories (see also *Figure 1C*): category Ia consists of coding human genome mutation database (HGMD) disease mutations (heart muscle diseases and channelopathies) and either are non-synonymous, frameshift insertions or deletions, splice or start/stop mutations. The same definition was applied for category Ib, where we additionally removed variants present in the 4300 individuals of the European-Americans cohort of the NHLBI GO Exome Sequencing Project (ESP) database (http://evs.gs.washington.edu/EVS/). As category II, we defined all not common, truncating variants that are either frameshift insertions/deletions, splice, or start/stop variants. Finally, all not common non-synonymous variants with prediction 'disease' were classified as category III. The predictions were based on SNPs&GO (February/March 2013; http://snps.biofold.org/snps-and-go//snps-and-go.html).[7]

Genotype–phenotype association analysis

Association analyses were carried out using the SAS software version 9.2 (SAS Institute, Inc., Cary, NC, USA). Patients were characterized by gender, country, family history of DCM, LVEF, LVEDD, NYHA classification, age at diagnosis, heart transplantation, and implantation of ICD according to available information. Based on the above-described variant classification, the number of variants, the number of patients carrying at least one variant, the number of patients carrying at least two variants in a single investigated gene, and the number of patients carrying at least two variants in any investigated gene were calculated for each group. Probability values and effect sizes with 95% confidence intervals were estimated based on logistic and Poisson regression assuming a dominant penetrance model. To identify the most relevant DCM variants according to their function, a stepwise forward model selection was carried out. The list of explanatory variables included the variant carrier status in genes grouped by function (ion flux, nucleus, cell membrane, sarcomeric, cystoskeleton, and intercalated disc). Significant explanatory variables at the 5% level (score test) entered the models and they were not removed if they remained significant at the 10% level. Because of the exploratory

Table 1 Patient characteristics

Characteristics	DCM patients (n = 639)
Gender (%)	
Woman	212 (34)
Man	405 (66)
Family history familial DCM (%)	
Sporadic	271 (51)
Familial	265 (49)
LVEF (%)	31.2 ± 12.1
LVEDD (mm)	64.4 ± 11.2
NYHA functional class (%)	
I	145 (28)
II	164 (32)
III	165 (32)
IV	39 (8)
Received heart transplantation (%)	
No	465 (80)
Yes	113 (20)
Received ICD implantation (%)	
No	344 (73)
Yes	130 (27)

nature of the genotype–phenotype association analysis, no correction for multiple testing has been applied.

Results

Ultra-high coverage next-generation sequencing enables comprehensive diagnostics of dilated cardiomyopathy

For the purpose of this study, we designed and optimized a custom target-enrichment assay based on in-solution hybridization, targeting relevant genes involved in human DCM being summarized in Supplementary material online, *Table S1*. The custom target region encompasses all known coding exons of each gene. In total, we analysed 639 patients with known DCM diagnosed according to current ESC guidelines. Importantly, all 639 samples from the eight countries analysed in this study were processed according to the same standard operating procedures (SOPs) and quality control measures, allowing detailed inter-sample comparisons. *Figure 1* details the origin of patients and the workflow applied. As shown in *Table 1*, the proportional number of patients with familial DCM was 49%. Left ventricular ejection fraction, indicating disease severity, was 31.2% (± 12.1), while the NYHA functional class, being a measure of the individual clinical status, was mainly in I–III. Heart transplantation was performed in 113 patients and another 130 received ICD implantation. Gender distribution was as expected with more male patients being affected. For detailed patient characteristics, please refer to *Table 1*.

As a pre-requisite for clinically applicable tests, we reached a very high 50-fold target coverage of 99.1% over all genes by iteratively improving the enrichment assay in a preceding establishment phase, being relevant to conclude positive as well as negative test results in individual patients. The mean coverage over all DCM genes was as high as 2526-fold (*Figure 2A*). To determine the accuracy, specificity, and sensitivity of the proposed diagnostic tool, we followed-up 25 randomly selected amplicons containing at least one variant by Sanger sequencing. From 5909 readable bases in the Sanger sequencing, we observed 5879 true-negative calls (TN) and 26 true-positive (TP) calls. Only three false-positive (FP) calls and one false-negative (FN) call were found, together resulting in a sensitivity [TP/(TP + FN)] of 96.3% and specificity [TN/(TN + FP)] of 100% and accuracy [(TP + TN)/(TP + FP + FN + TN)] of 99.9%. To further increase the depth of variants for benchmarking, we enriched and sequenced the well-genotyped HapMap sample NA12878 using the same methodology. When testing the standard GATK cut-off values, we find a sensitivity of 100% and an accuracy of 85.4%. Using our filter off-sets described in materials and methods, we achieved a much higher accuracy of 91.3% by maintaining a sensitivity of 99.0%. After manual inspection of the false-positives that could not be validated by Sanger, we postulated that at least some of these must be true variants. Hence, we exemplarily subcloned genomic DNA of patients with a variant call in the NGS data set and negative Sanger sequencing and sequenced individual clones. As shown in Supplementary material online, *Figure S3*, wild-type as well as mutant clones from one individual can be found, which is not obvious in the Sanger sequence, indicating that the actual accuracy of NGS is even higher than the estimated and that Sanger sequencing may miss at least some of the variants.

Distribution of mutations in dilated cardiomyopathy patients

In total, we identified 8269 unique genetic variants, adding up to 359 669 variants in the 639 patients across the investigated target region. On average, each patient carries 563 variants in this region. To gain information on the relevance of each variant, we performed a stepped filtering approach, first by eliminating known common variants. We thereafter annotated the remaining variants using ANNOVAR, snpEff, Genometrax (Biobase), and SNPs&GO.

We then applied the classification presented in *Figure 1C*. A known cardiomyopathy mutation (category Ia; reported in HGMD as cardiomyopathy or channelopathy variant) was found in 305 patients (48% of all patients). When we additionally excluded variants observed in a large non-DCM control cohort (ESP whole exome sequencing project[8]) (=category Ib), we still find in 294 patients (46%) a known mutation previously reported as disease causing. *Figure 2B* gives an overview over the distribution of mutations across the screened DCM genes. When considering only DCM-causing mutations by excluding mutations of other cardiomyopathies, a known mutation is found in 101 patients (16% of all patients).

Since many cardiomyopathy cases will be caused by rare or private mutations, which are not yet annotated in databases such as HGMD, we next searched in all patients for 'likely' disease mutations (category II). The 'likely' mutations include frameshift insertions/deletions, stop-gain/-loss variants, and splice-site mutations within the target genes. In addition to the category Ia variants, our analysis yielded insertions with resulting frameshifts in three different genes across 13 patients and

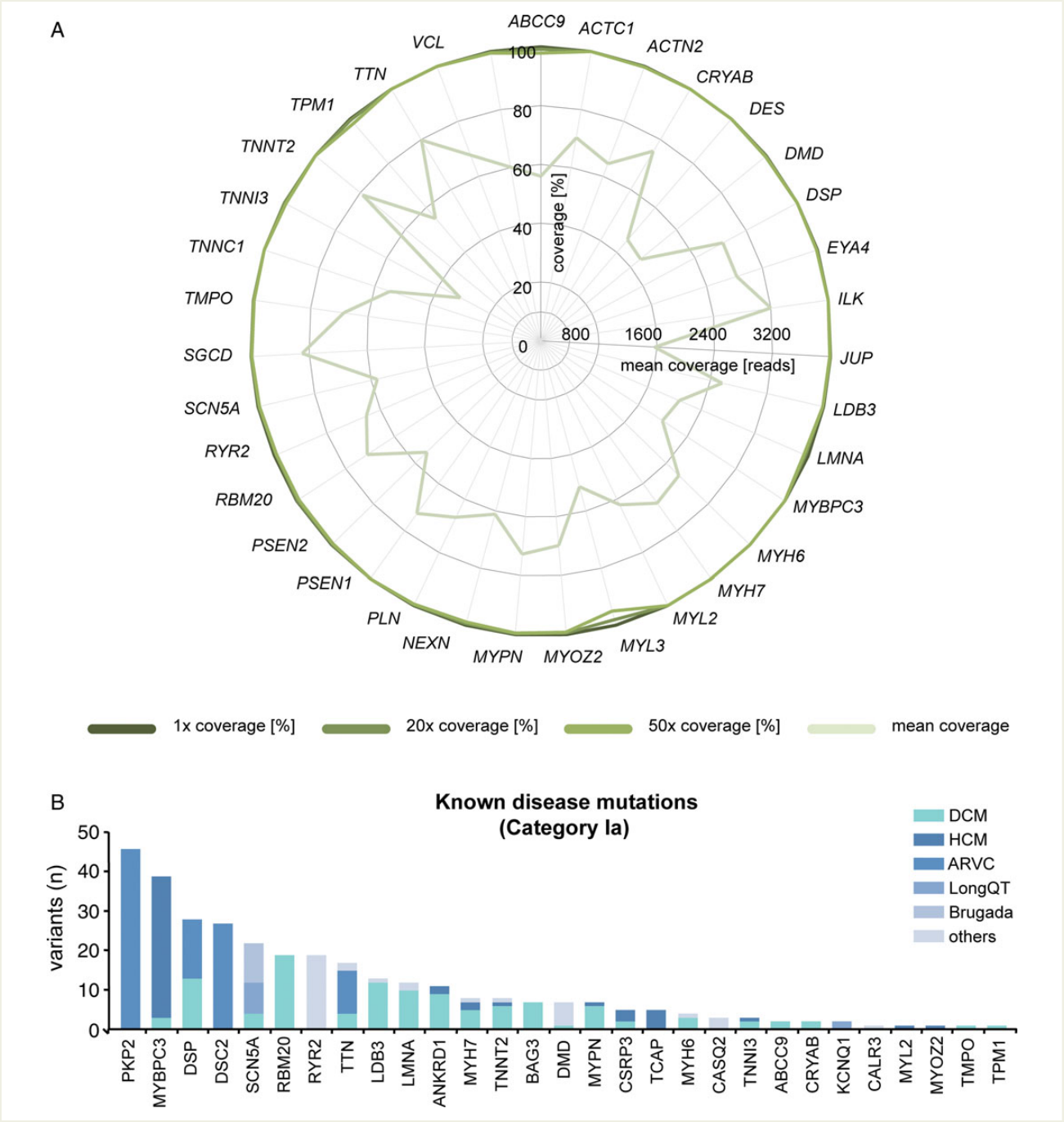

Figure 2 Sequence coverage and variant distribution per Gene. (*A*) Graph representing the sequence coverage of dilated cardiomyopathy genes. Vertical axis is showing the target coverage for every gene, colour codes give the read depth of coverage. The horizontal axis is giving the mean coverage in reads per gene (pale green line). Over all dilated cardiomyopathy genes a mean coverage of 2526 could be found. (*B*) Bar graph demonstrating the distribution of identified known disease causing mutations (category Ia) in the whole patient cohort over all tested dilated cardiomyopathy genes. Colour code is giving information on the different CMP subtypes as annotated in HGMD. Dilated cardiomyopathy mutations are painted in petrol, hypertrophic cardiomyopathy, ARVC, LongQT, and Brudgada mutations are coloured in blue, ranging from dark to pale.

frameshift deletions in 10 DCM genes covering a total of 37 patients. We also identified 11 individual splicing variants in 8 DCM genes in a total of 37 patients and 60 stop-gain/-loss variants in 17 genes in another 67 patients. Altogether, we find 117 previously not annotated highly 'likely' pathogenic variants in 26 genes for 147 patients (23%).

To search for 'potential' disease mutations (category III), we selected all non-common, non-synonymous variants, and applied bioinformatics methods to predict a detrimental effect of each variant on the protein function. By using SNPs&GO, we classified as many as 939 variants as neutral and 141 unique variants as potentially disease causing. These

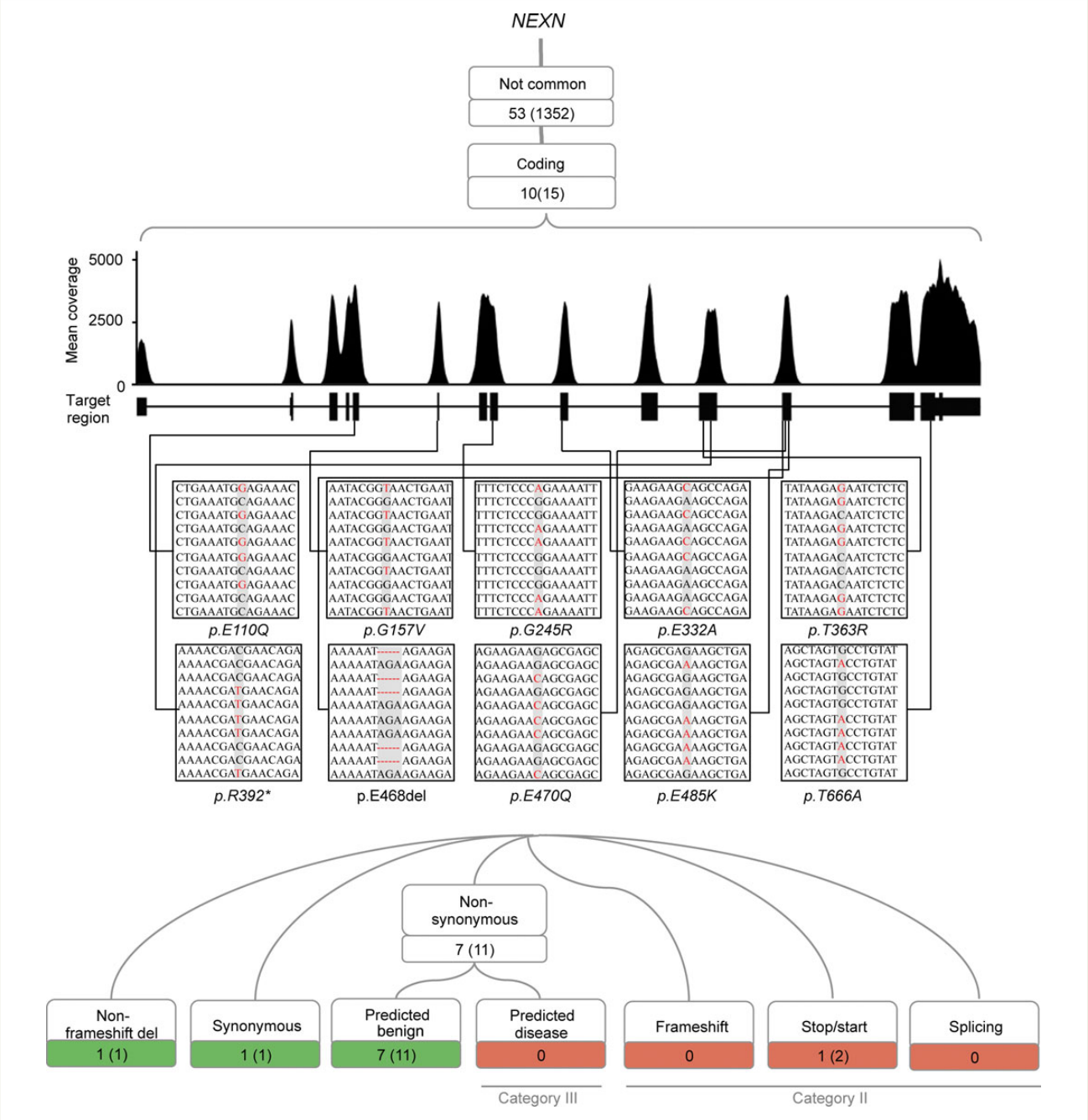

Figure 3 Variant distribution in *NEXN*. The high coverage of each gene, exemplified for the *nexilin* gene, enables a reliable read mapping, variant calling and subsequent filtering of genetic variants. The black boxes below the coverage plot are representing the targeted region of the *NEXN* gene. Sequence alignments show the variants and their chromosomal location. The scheme at the bottom of the figure gives an overview on the variant classification. In total, seven different non-synonymous variants and one stop mutation in two patients and one non-frameshift variant in one patient were found in *NEXN*. Numbers in brackets are the total sum of variants found in the cohort. Predicted benign = green; deleterious = red.

141 variants were detected in 221 patients and are rare judged by their frequency. *Figures 3* and *4* exemplify the distribution of variants for representative disease genes (*Nexilin, Titin, Lamin A/C*). See also Supplementary material online, *Figure S1* for the whole list of genes. To investigate whether the variants from category II–III indeed represent rare mutations, we calculated the percentage of Singletons, which means variants found uniquely in only one patient. Here, the rate was 21.8% in the categories II–III, compared with 1.6% in the removed variants, underlining the stringency of the classification approach.

Looking at the total number of variants within the DCM genes, the majority of variants (13%) can be found in the *titin* gene (*TTN*) (*Figure 5A*). This is not surprising since *TTN* is the largest human gene and accounts for >20% of the total target region. When normalizing, the number of variants to the size of each gene

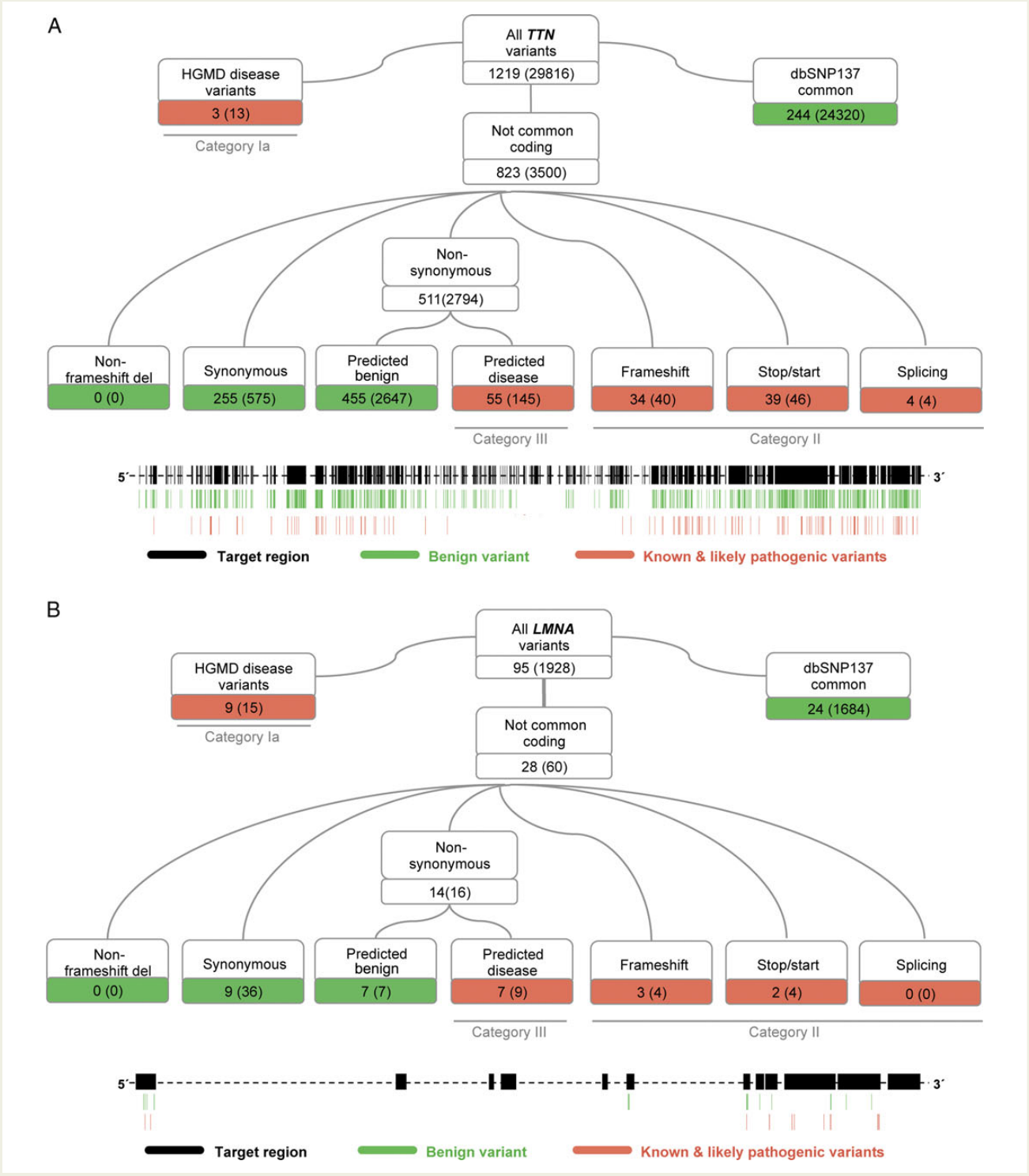

Figure 4 Variant classification in *TTN* and *LMNA*. Diagrams are showing the number of individual variants of different types for (*A*) *titin* and (*B*) *lamin*. Numbers in brackets are the total sum of variants found in the cohort. HGMD disease variants (category Ia) were annotated using the Biobase Human Genome Mutation Database, regardless of their appearance in dbSNP137common. All subcategories (non-synonymous, frameshift, stop/start, predicted disease, predicted benign, splicing, synonymous) were annotated after removing those common variants. For prediction of non-synonymous variants, SNPs&Go was used. Colours of the boxes indicate a potential benign (green) or deleterious (red) effect. Variant categorization is indicated below the deleterious variant types. Sketch below the diagram is showing the target region of the gene (black) and the distribution of common (green) and known and likely/potential pathogenic variants (red).

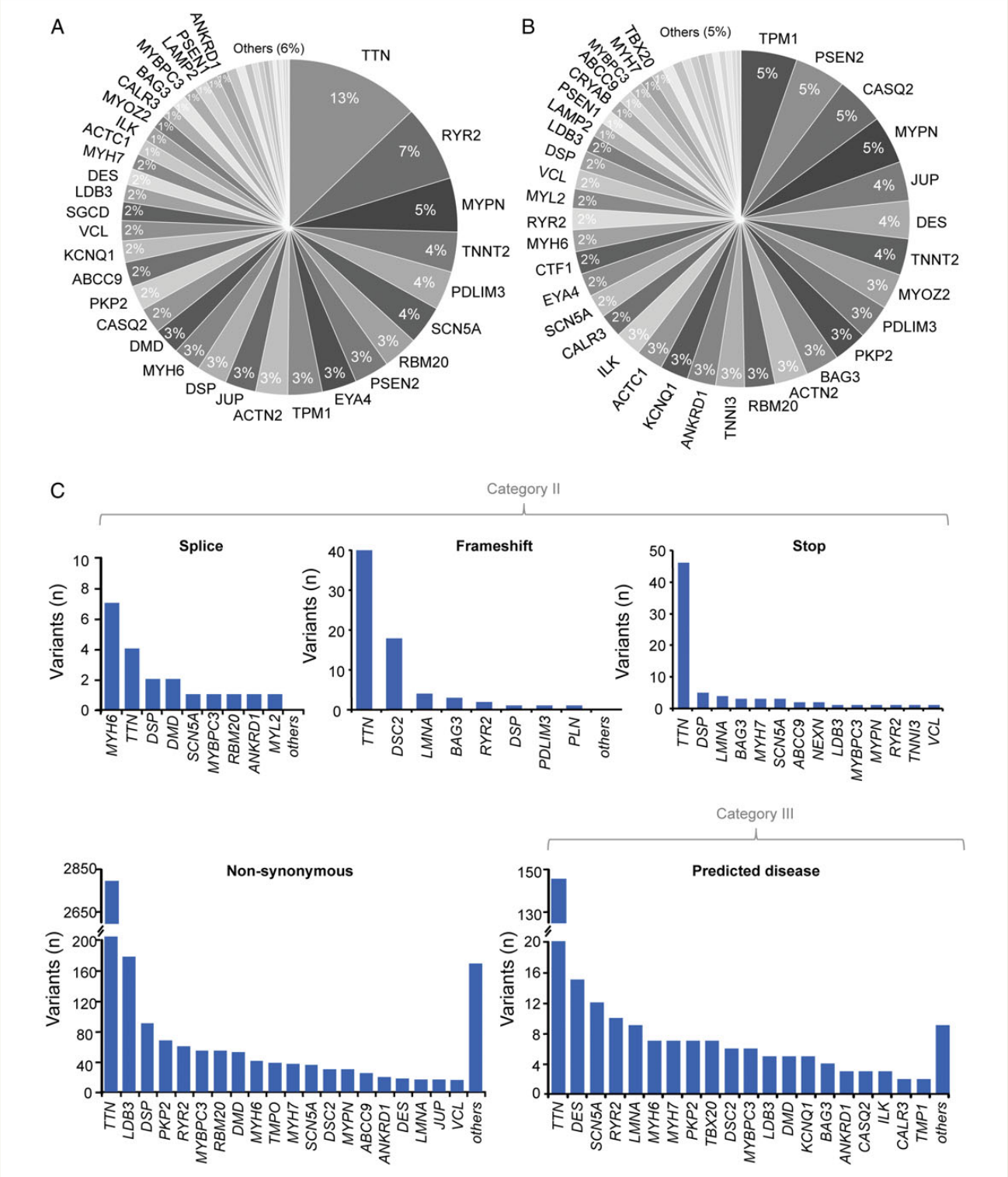

Figure 5 Variant distribution in dilated cardiomyopathy genes. Pie charts representing the total number of variants per dilated cardiomyopathy genes (*A*) and after normalization to the gene size (*B*). (*C*) Blue bar graphs are showing the total number of variants over all patients by predicted effect type for each gene (splice, frameshift, stop, non-synonymous, and predicted disease). Curly bracket is grouping the likely truncating variants (category II) and indicates the potential disease causing variants (category III).

(*Figure 5B*), a rather even distribution can be found, disproving the existence of instability hotspots in DCM genes. A detailed view on the distribution of all likely and potentially pathogenic variants is given in *Figure 5C* showing the number of splice, frameshift, stop, non-synonymous, and predicted disease causing (SNPs&GO) non-synonymous variants.

Distribution of the functional effects of dilated cardiomyopathy mutations

Numerous studies suggest different phenotypic manifestations or severities depending on the gene affected, type and number of mutations. Hence, we tested if DCM patients in our cohort might carry multiple disease mutations, e.g. compound mutations (category Ib–III). Strikingly, such compound heterozygous states were found in 49 (7%) patients and combined heterozygous mutations were found in 243 patients (38%). Remarkably, we detected in 82 patients (12.8%) at least 3 mutations (*Table 2*). Considering only the very stringent category Ib variants of annotated disease mutations after exclusion of variants detected in additional control cohorts, still 82 patients (12.8%) carry at least 2 known disease mutations. As expected, using logistic regression, we find a significant association of patients having a disease mutation and familial DCM ($P = 0.03$, category Ib–III) (*Table 3*). To test if those results are driven by an effect of the large *ttn* gene, we repeated the analysis after exclusion of any *ttn* variant and still find 79 patients (12.4%) with at least two category Ib variants (Supplementary material online, *Table S5*).

When looking more closely at the variants annotated using the HGMD database (CatIb), a large portion of disease causing mutations are known to cause arrhythmogenic right ventricular cardiomyopathy (ARVC) (31%), HCM (16%), or channelopathies (6%) (*Figure 6A*), indicating a marked overlap not only related to disease genes, but also to specific mutations in cardiomyopathies. Hence, based on current literature, all genes investigated in this study were classified according to the different cell components or functions they contribute to (Supplementary material online, *Table S1*).[9]

Table 2 Multiple mutations affecting single patients

Number of mutations	HGMD[a] variant pos patients (%)	Category Ib–III[b] variant pos patients (%)
0	345 (54.0)	171 (26.7)
≥1	294 (46.0)	468 (73.2)
≥2	82 (12.8)	243 (38.0)
≥3	14 (2.2)	82 (12.8)
≥4	2 (0.3)	16 (2.5)

[a]Category Ib.
[b]Either category Ib or category II or category III.

Next, we summed up the number of patients carrying a category Ib mutation in the different groups. *Figure 6B* details the groups and the identified number of patients. Based on this classification, the sarcomere group shows the highest number of patients having a mutation (14%), followed by ion flux (13%), z-disc/cytoskeleton (12%), and intercalated disc (11%).

We next asked whether we might identify specific genotype–phenotype associations in this large cohort. This would have direct clinical implications, since a genotype-guided risk assessment could improve patient selection for intensified monitoring or directed therapies. First, we investigated the gene groups introduced above and the available phenotypes. In an exploratory association analysis by using a stepwise forward selection, we could identify a logistic regression model for the group 'nucleus' being a significant predictor for ICD-carrier status in DCM (unadjusted $P = 0.02$). The odds ratio (OR) for patients carrying such variants was 2.44 [95% confidence interval (95% CI): 1.13–5.28]. This association was mainly driven by the nuclear gene *RBM20*, having an OR of 5.65 (1.89–16.86; $P = 0.002$). For the age at diagnosis, which might be relevant for establishing genetic testing in relatives, we find a significant association between *MYH6* and *ADRB3* mutations (*Tables 4* and *5*). Other clinically relevant associations were found for *SMYD1*, which we here suggest as novel disease gene for DCM, as well as for *alpha-crystalin B* (*CRYAB*) mutations, both were associated with LVEF. Alterations in left-ventricular diameter (LVEDD) could be seen in association with *TBX20* (OR: 0.45, 0.22–0.94; $P = 0.03$). Since a significant number of patients also received heart transplantation (HTX) due to end-stage heart failure (20% of the cohort), we investigated associations with HTX. Here, we found a significant association with *MYPN* having an OR of 4.23 (1.04–17.18).

Differences of dilated cardiomyopathy mutations across Europe

This study includes a total of 639 patients from eight countries, enabling us to investigate the geographical distribution of cardiomyopathy relevant variants. When considering genetic variants from category Ib–III (Supplementary material online, *Table S6*), we observe small yet statistically significant differences between countries regarding the rate of mutations (variants per patient), with Germany showing the lowest rate of 0.98 and Great Britain showing the highest rate of 1.51 (global $P = 0.04$, Poisson regression). The rate of mutation positive patients for single genes and across countries is depicted in *Figure 7*. For example, the rate of patients carrying a TTN variant ranged from 0.56 in Dutch patients to 0.20 in German patients

Table 3 Association of familial and sporadic dilated cardiomyopathy with number of mutation positive results

	No. of patients	No. of patients having category Ib–III[a] variant	*P*-value from logistic regression	OR (95% CI)
Sporadic DCM	271	185	0.03	Reference
Familial DCM	265	203		1.52 (1.04–2.23)

[a]Either category Ib or category II or category III.
OR, odds ratio; CI, confidence interval.

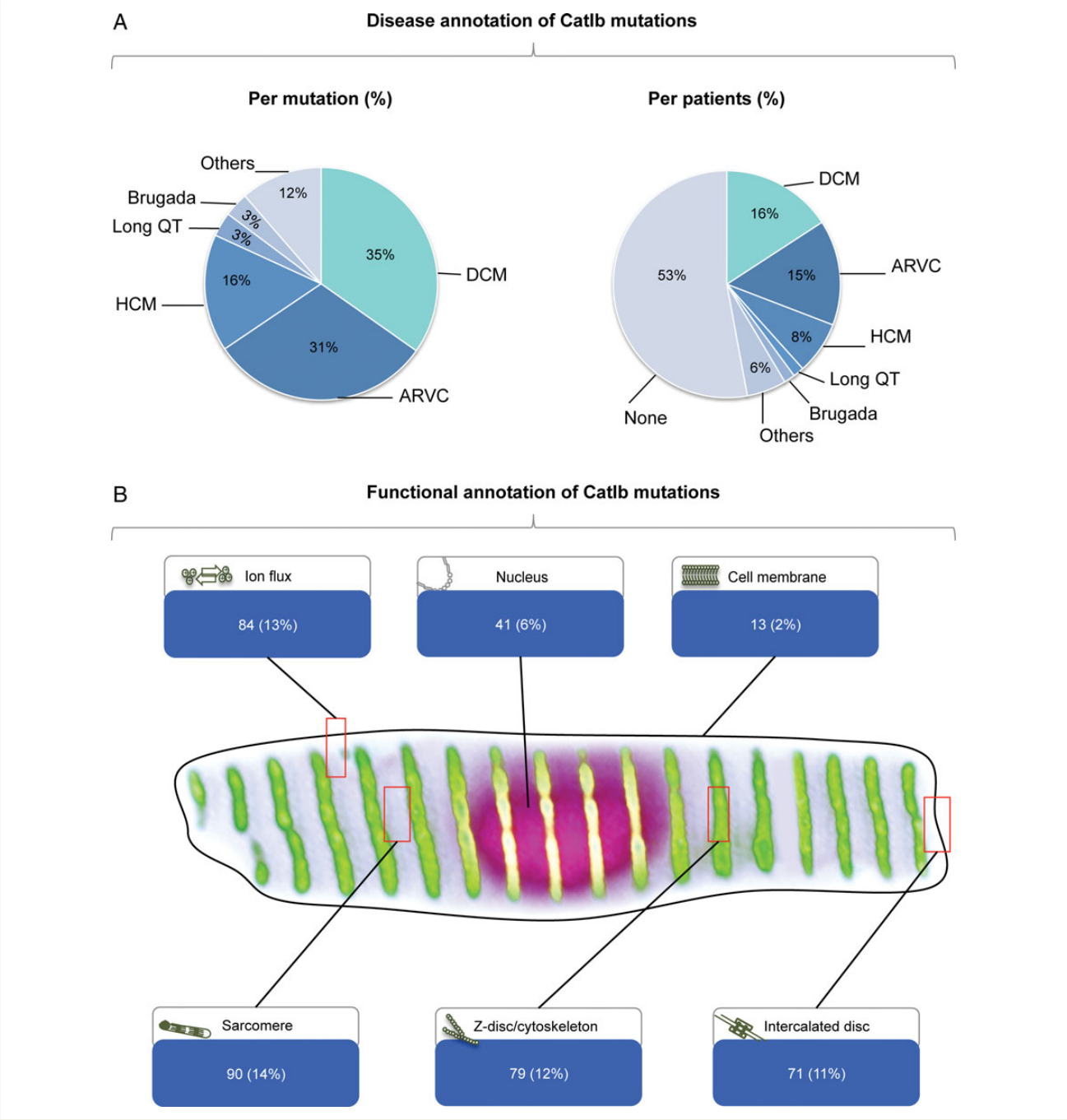

Figure 6 Variant distribution and classification by cell components. (*A*) Variants were annotated using HGMD and filtered by Exome Sequencing Project variants (category Ib). Within all annotated mutations, we find the highest number described to cause dilated cardiomyopathy (35%), ARVC (31%) and hypertrophic cardiomyopathy (16%). Also, a substantial number is found for Long QT and Brugada syndrome (left pie chart). The right pie chart is giving percentages of affected patients in relation to dilated cardiomyopathy-cohort size ($n = 639$). B) Genes were grouped based on their contribution to different cellular compartments or functions in six groups (ion flux, nucleus, cell membrane, sarcomere, z-disc/cytoskeleton or intercalated disc). Given are the number (percentage) of patients carrying a category Ib disease mutation. Patients having a variant in more than one group were counted multiple times.

(global $P < 0.001$, logistic regression). However, mutation frequencies of DCM genes are clearly more homogeneous than previously reported in smaller studies, suggesting that genetic testing for DCM can be applied in a uniform setting across Europe.

Discussion

To our knowledge, this is the most comprehensive study on the contribution of DCM-causing genes to date. The data reported shed light

Table 4 Association between genotypes and dilated cardiomyopathy

Gene	Phenotype	No. of patients	Patients with category Ib-III[a] variants	*P*-value from logistic regression	OR (95% CI)
CRYAB	LVEF	582	2	0.04	0.05 (0.00–0.81)
MYPN	Received HTX				
	No	465	4		Reference
	Yes	113	4	0.04	4.23 (1.04–17.18)
RBM20	Received ICD				
	No	344	5	0.002	Reference
	Yes	130	10		5.65 (1.89–16.86)
SMYD1	LVEF	582	2	0.03	4.42 (1.14–17.10)
TBX20	LVEDD	538	7	0.03	0.45 (0.22–0.94)
MYH6	Age at diagnosis	439	8	0.03	0.63 (0.41–0.96)
ADRB3			2	0.04	0.36 (0.14–0.94)

Only significant associations are shown (unadjusted $\alpha = 5\%$, two-sided).
[a]Either category Ib or category II or category III.
OR, odds ratio; CI, confidence interval. For LVEF OR are calculated per 10% step. For LVEDD OR are calculated per 10 mm step. For age at diagnosis, OR are calculated per 10 years step.

Table 5 Mutation counts by country

Country	No. of patients	No. of category Ib-III[a] variants	*P*-value from poisson regression	RR (95% CI)
Denmark	100	130	0.04	Reference
England	70	106		1.16 (0.90–1.51)
France	92	111		0.93 (0.72–1.20)
Germany	98	93		0.73 (0.56–0.95)
Italy	78	97		0.96 (0.74–1.24)
Netherlands	70	99		1.09 (0.84–1.41)
Spain	82	107		1.00 (0.78–1.30)
Sweden	49	71		1.11 (0.83–1.49)

[a]Either category Ib or category II or category III.
RR, relative risk; CI, confidence interval.

on the distribution of genes, the number of mutations and mutational burden of patients with DCM.

Next-generation sequencing technologies (NGS) have emerged as a fast alternative to Sanger-sequencing, providing the analytical characteristics for the comprehensive exploration of genetic mechanisms.[10] However, NGS retains some weaknesses, such as the incomplete representation and coverage of exons, bearing the risk of limited sensitivity and detection failure of clinically relevant mutations. The use of target enrichment followed by NGS, for high-throughput genetic testing of disease genes for DCM, HCM and other cardiomyopathies has now become feasible and technically validated as shown by the nearly complete coverage and high accuracy of the approach described here.

To reduce the number of likely and potentially pathogenic variants, bioinformatics analyses must include a filter step, e.g. excluding common variants present in databases like dbSNP. However, it is known that dbSNP is 'contaminated' with a small but yet substantial number of pathogenic alleles.[11] On the other hand, disease mutation databases contain potentially benign variants, previously classified as disease causing. Therefore, to substantiate the disease causing nature of a variant, further investigations must include a screening of well-phenotyped control cohorts from diverse populations and a detailed follow-up on the clinical circumstances in each patient and family, for example, by co-segregation analyses. But since families are often small and many patients are classified as sporadic, the ultimate solution of this intricate problem seems illusive. Mestroni and Taylor[12] recently reviewed how the progress in genetic research already changed the view on the genetic basis of DCM. It is hence obvious that bioinformatics strategies must grow together with the enormous amount of data from adequately sized and high-quality NGS studies. With regard to a functional prediction, Thusberg *et al.*[13] observed significant differences between algorithms, but found that algorithms such as the here applied SNPs&GO perform already surprisingly well. However, at least by now, proving new variants or genes as disease causing requires further efforts to validate them through functional and familial studies and large-scale population-based control cohorts sequenced by NGS.

Previous publications have given rise to the notion that disease-causing gene mutations may underlie geographical disparities.[14] In our study, we enrolled patients from eight countries to create a European map of cardiomyopathy relevant variants and find only modest differences across the countries. We observed the lowest number of variants in the German cohort when compared with the other countries. This effect may be due to a higher number of patients in Germany bearing sporadic forms of DCM. However, differences caught our attention for Sweden, where we detected the most category Ib–III variants in *DSP* and *MYBPC3* compared with the other countries, as well as for the Netherlands, where we observed the highest number in *TTN* variants. With respect to *PLN*, we only uncovered one single mutation in our study population, suggesting that the

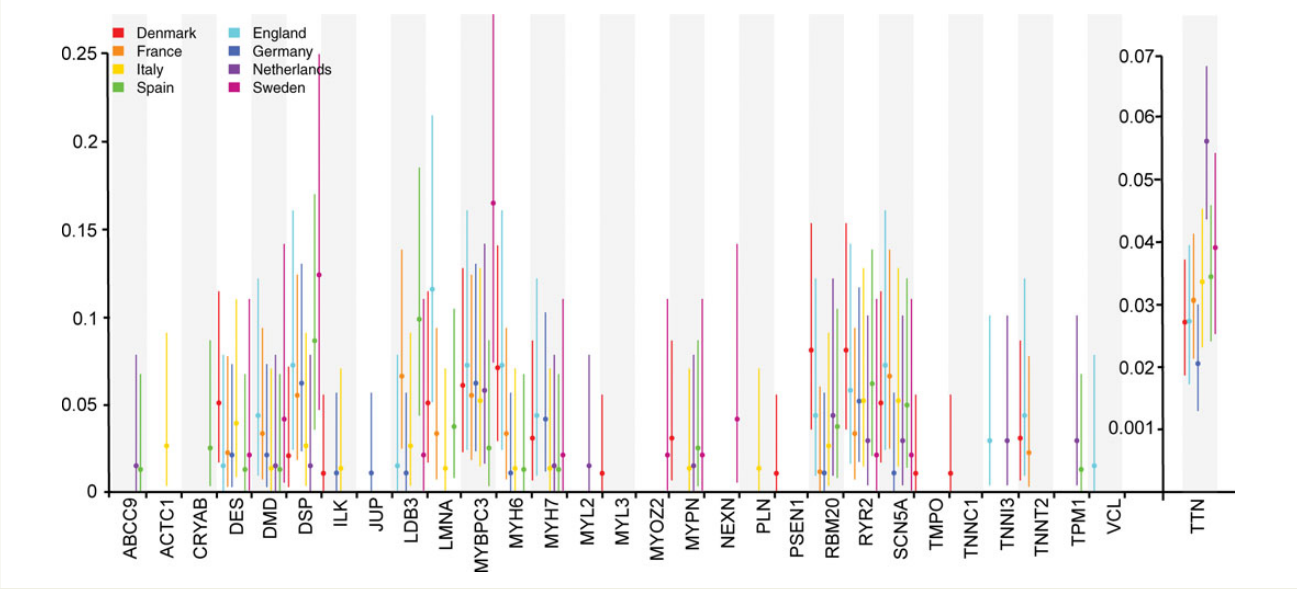

Figure 7 Country-specific rates of mutation positive patients by dilated cardiomyopathy genes. The rates of patients carrying at least one category Ib–III variant per gene are represented by coloured dots, 95% Clopper–Pearson confidence interval are indicated by coloured lines. Countries without mutations in a certain gene are not plotted.

previously mentioned mutation frequencies might express a very local phenomenon rather than a common cause of DCM. The overall distribution of mutations in DCM disease genes all over the participating countries appeared to be more homogeneous than expected.

In clinical routine, only detailed workup of cases and their families allows to uncover familial aggregation.[15] Often this is impeded by small family structures, unavailable family members or incomplete penetrance, which classifies many cases as sporadic or idiopathic. However, this does not exclude a genetic cause of DCM in this individual patient. By introducing familial and sporadic DCM cases in our study, we were able to compare both groups in a comprehensive and well-controlled manner. By doing so, we observed significantly higher mutation rates in familial cases than in the sporadic ones (OR: 1.52). However, as shown, we find also in many cases of idiopathic DCM a known and well-characterized disease mutation. Hence, the current work underlines that even if a definite familial DCM may not be proven, a genetic aetiology cannot be ruled out. The decision to genetically test sporadic cases should be, however, taken carefully to avoid unnecessary costs and inconclusive results. By limiting testing to familial DCM and risk groups of idiopathic DCM, e.g. with documented arrhythmias, would be an apparent conclusion.

Genotype–phenotype correlations will be of increasing importance to predict the clinical manifestations of genetically diagnosed patients. However, there are few existing studies on genotype–phenotype correlations in selected cohorts. Convincing data exist for the clinical impact of *LMNA*. Here, several studies have repeatedly shown a poor prognosis for *LMNA* mutation carriers due to the occurrence of ventricular arrhythmias and sudden cardiac death.[16] Similarly, it was shown that *RBM20* mutation carriers with DCM present with a fast progression of heart failure and high risk for arrhythmias.[17] In our study, the often-maligned stop or frameshift mutations in *LMNA* ranked within the top three genes of our tested gene panel and we find *LMNA* to having together with *RBM20* the most category Ib–III mutations for genes with a comparable size. While we observed a significant association of ICD-carriers with the functional gene group 'nucleus' to which *LMNA* belongs, a statistically significant effect could only be seen for *RBM20*, where we find an association with the ICD-carrier status. It should be kept in mind that present results rely on a small number of observations for some strata and that probability values are not corrected for multiplicity. When using an exact test, the probability value *MYPN* and HTX ($P = 0.04$) decreases to $P = 0.0051$, while the probability value for *CRYAB* ($P = 0.04$) decreases to $P = 0.02$ and *RBM20* ($P = 0.002$) remains unchanged. While for *RBM20* our results may be seen as validation of previous findings and hence underline the importance for testing this gene now routinely, the newly found associations require additional replication in independent cohorts of DCM patients.

Remarkably, we find in our DCM cohort a high percentage of mutations previously described for HCM, ARVC, or channelopathies, which questions the hypothesis of the allelic nature of cardiomyopathies. For instance, we found *plakophilin-2* to be the most frequently affected gene when considering only known pathogenic mutations. *Plakophilin-2* mutations represent a frequent cause of ARVC, which is characterized by the degeneration of cardiomyocytes and resulting arrhythmias.[18] A similar finding was recently provided by Pugh *et al.*,[19] underlining that our findings are not spurious. While they hypothesized that misdiagnosis in their broad referral population of an ordering provider could be one potential explanation, we nearly can exclude this due to the controlled setting of our study. To avoid phenotypic misclassification, we relied on experienced clinicians for recruiting to our

best knowledge only DCM patients and performed additional phenotyping by coronary angiography, cMRI, or myocardial biopsy where appropriate. We rather hypothesize a pathophysiological link between conductance defects and resulting cardiac mechanical disparities. However, even then a certain overlap of phenotypes will explain at least a part of these findings.

We were to our knowledge first to introduce genetic testing of cardiomyopathy patients for the large *TTN* gene using NGS.[3] In accordance with recent findings by Herman *et al.*[20] reporting *TTN* truncating mutations to be the cause of DCM in ~25% of familial and in 18% of sporadic cases, we were able to identify such mutations in 19% of familial and 11% of sporadic cases. Interestingly, 44% of all patients with a truncating *TTN* variant also had additional known disease-causing variants in at least one other gene, suggesting that in these cases, the *TTN* variant may not be the sole cause of DCM, underlining the importance to investigate a broad gene panel rather than selected genes.[19] In this regard, this study sheds for the first time light on the role of *TTN* and other mutations to DCM in a considerable high number of patients. This is also underlined by the high clinical sensitivity of our approach of 46% (considering only known mutations, cat Ib) up to 73.2% (considering cat Ib–III), which is already in the range of the sensitivity reported for HCM.[21] This is in contrast to previous studies that reported lower sensitivity values, maybe due to more inhomogeneous cohorts or the use of NGS in subgroups of patients.[19] This maybe also the reason why *MYBPC3 m*utations, which we surprisingly find very frequently, could not be associated with DCM in some previous studies and hence was long debated if it contributes to DCM at all. For the group of patients without Ib–III variants (26.7%), family-based whole exome sequencing or alternative approaches should be applied to identify possible genetic defects in so far unknown disease genes, with *BAG3* being one recent example.[22,23]

Even if genetic investigations do not resolve a familial disease, other mechanisms such as epigenetic modifications (microRNAs, Histon-modifications, DNA methylation) should be taken into account.[10,24] Only very recently, first studies underlined the role of DNA methylation on heart failure and DCM.[25,26] While the studies did not provide evidence for the causality or inheritance of the observed changes, it will be interesting to investigate the shown proof-of-principle in pedigrees with DCM. Furthermore, the increasing numbers of genetic susceptibility loci harbouring common polymorphisms add to the concept that DCM is a complex disorder, with rare and common genetic variants and environmental factors driving disease onset and outcome.

With improved accuracy, efficiency, and decreasing cost, NGS is becoming the sole standard for gene sequencing. It must be noted that high analytical standards must be the basis for introducing such methodology for the diagnostics of DCM. We were able to show high-analytical fidelity and consistency in a large patient cohort, demonstrating that targeted NGS is already well suited to be applied in clinical routine diagnostics, substantiating the ongoing paradigm shift from low- to high-throughput genomics in medicine. However, many third party providers do not publish quality indices that should be prerequisite for introducing their technologies into the clinics. Our study hopefully stimulates not only the single physician to demand for transparency in regard to quality but also helps to impact on future recommendations for the next-generation of genetic testing. By means of our atlas of the genetics of human DCM, we aspire cardiologist and human geneticists to soon be able to apply our findings to the individual patient with cardiomyopathy in daily clinical practice.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

This work was supported by grants from the European Union (FP7 INHERITANCE and BestAgeing), the 'Bundesministerium für Bildung und Forschung' (BMBF): German Center for Cardiovascular Research (DZHK), NGFN II, NGFN-plus and NGFN-transfer, the medical faculty of the University of Heidelberg (B.M.), and INSIGHT DCM (BMBF and Agence nationale pour la recherche).

Funding

This work was also supported by grants from Assistance Publique - Hôpitaux de Paris (PHRC AOM04141), the 'Fondation LEDUCQ' (Eurogene Heart Failure network), the Société Française de Cardiologie/ Fédération Française de cardiologie, the CONNY-MAEVA charitable foundation.

Conflict of interest: A.K., W.-K.L., X.Z., D.F., A.K., V.K., E.W. are employees of Siemens AG. Siemens AG is an international provider of diagnostic assays and hold several patients on diagnostic assays. B.M. and H.A.K. conduct research together with Siemens AG, Erlangen.

References

1. McNally EM, Golbus JR, Puckelwartz MJ. Genetic mutations and mechanisms in dilated cardiomyopathy. *J Clin Invest* 2013;**123**:19–26.
2. Meder B, Ruhle F, Weis T, Homuth G, Keller A, Franke J, Peil B, Lorenzo Bermejo J, Frese K, Huge A, Witten A, Vogel B, Haas J, Volker U, Ernst F, Teumer A, Ehlermann P, Zugck C, Friedrichs F, Kroemer H, Dorr M, Hoffmann W, Maisch B, Pankuweit S, Ruppert V, Scheffold T, Kuhl U, Schultheiss HP, Kreutz R, Ertl G, Angermann C, Charron P, Villard E, Gary F, Isnard R, Komajda M, Lutz M, Meitinger T, Sinner MF, Wichmann HE, Krawczak M, Ivandic B, Weichenhan D, Gelbrich G, El-Mokhtari NE, Schreiber S, Felix SB, Hasenfuss G, Pfeufer A, Hubner N, Kaab S, Arbustini E, Rottbauer W, Frey N, Stoll M, Katus HA. A genome-wide association study identifies 6p21 as novel risk locus for dilated cardiomyopathy. *Eur Heart J* 2014;**35**:1069–1077.
3. Meder B, Haas J, Keller A, Heid C, Just S, Borries A, Boisguerin V, Scharfenberger-Schmeer M, Stahler P, Beier M, Weichenhan D, Strom TM, Pfeufer A, Korn B, Katus HA, Rottbauer W. Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circ Cardiovasc Genet* 2011;**4**:110–122.
4. 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.
5. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, Dubourg O, Kuhl U, Maisch B, McKenna WJ, Monserrat L, Pankuweit S, Rapezzi C, Seferovic P, Tavazzi L, Keren A. Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 2008;**29**:270–276.
6. Charron P, Arad M, Arbustini E, Basso C, Bilinska Z, Elliott P, Helio T, Keren A, McKenna WJ, Monserrat L, Pankuweit S, Perrot A, Rapezzi C, Ristic A, Seggewiss H, van Langen I, Tavazzi L. European Society of Cardiology Working Group on M, Pericardial D. Genetic counselling and testing in cardiomyopathies: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 2010;**31**:2715–2726.
7. Capriotti E, Altman RB. Improving the prediction of disease-related variants using protein three-dimensional structure. *BMC Bioinformatics* 2011;**12**(Suppl. 4):S3.
8. Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, Sajadieh A, Haunso S, Svendsen JH, Olesen MS. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. *Eur J Hum Genet* 2013;**21**:918–928.

9. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. *N Engl J Med* 2011;**364**:1643–1656.
10. Frese KS, Katus HA, Meder B. Next-generation sequencing: from understanding biology to personalized medicine. *Biology (Basel)* 2013;**2**:378–398.
11. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;**12**:745–755.
12. Mestroni L, Taylor MR. Genetics and genetic testing of dilated cardiomyopathy: a new perspective. *Discov Med* 2013;**15**:43–49.
13. Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat* 2011;**32**:358–368.
14. van der Zwaag PA, van Rijsingen IA, Asimaki A, Jongbloed JD, van Veldhuisen DJ, Wiesfeld AC, Cox MG, van Lochem LT, de Boer RA, Hofstra RM, Christiaans I, van Spaendonck-Zwarts KY, Lekanne Dit Deprez RH, Judge DP, Calkins H, Suurmeijer AJ, Hauer RN, Saffitz JE, Wilde AA, van den Berg MP, van Tintelen JP. Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy. *Eur J Heart Fail* 2012;**14**:1199–1207.
15. Meder B, Katus HA. Clinical and genetic aspects of hypertrophic and dilated cardiomyopathy. *Internist (Berl)* 2012;**53**:408–414, 417–418.
16. van Rijsingen IA, Nannenberg EA, Arbustini E, Elliott PM, Mogensen J, Hermans-van Ast JF, van der Kooi AJ, van Tintelen JP, van den Berg MP, Grasso M, Serio A, Jenkins S, Rowland C, Richard P, Wilde AA, Perrot A, Pankuweit S, Zwinderman AH, Charron P, Christiaans I, Pinto YM. Gender-specific differences in major cardiac events and mortality in lamin A/C mutation carriers. *Eur J Heart Fail* 2013;**15**:376–384.
17. Refaat MM, Lubitz SA, Makino S, Islam Z, Frangiskakis JM, Mehdi H, Gutmann R, Zhang ML, Bloom HL, MacRae CA, Dudley SC, Shalaby AA, Weiss R, McNamara DM, London B, Ellinor PT. Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. *Heart Rhythm* 2012;**9**:390–396. S1547-5271(11)01218-5 (pii).
18. van der Smagt JJ, van der Zwaag PA, van Tintelen JP, Cox MG, Wilde AA, van Langen IM, Ummels A, Hennekam FA, Dooijes D, Gerbens F, Bikker H, Hauer RN, Doevendans PA. Clinical and genetic characterization of patients with arrhythmogenic right ventricular dysplasia/cardiomyopathy caused by a plakophilin-2 splice mutation. *Cardiology* 2012;**123**:181–189.
19. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, Bowser M, Harrison B, Aaron D, Mahanta LM, Lakdawala NK, McDermott G, White ET, Rehm HL, Lebo M, Funke BH. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med* 2014;**16**:601–608.
20. 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. *N Engl J Med* 2012;**366**:619–628. 10.1056/NEJMoa1110186.
21. Hershberger RE, Cowan J, Morales A, Siegfried JD. Progress with genetic cardiomyopathies: screening, counseling, and testing in dilated, hypertrophic, and arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Heart Fail* 2009;**2**:253–261.
22. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, Ruppert V, Arbustini E, Wichter T, Germain M, Dubourg O, Tavazzi L, Aumont MC, Degroote P, Fauchier L, Trochu JN, Gibelin P, Aupetit JF, Stark K, Erdmann J, Hetzer R, Roberts AM, Barton PJ, Regitz-Zagrosek V, Aslam U, Duboscq-Bidot L, Meyborg M, Maisch B, Madeira H, Waldenstrom 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 F. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J* 2011;**32**:1065–1076.
23. Norton N, Li D, Rieder MJ, Siegfried JD, Rampersaud E, Zuchner S, Mangos S, Gonzalez-Quintana J, Wang L, McGee S, Reiser J, Martin E, Nickerson DA, Hershberger RE. 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.
24. Vogel B, Keller A, Frese KS, Leidinger P, Sedaghat-Hamedani F, Kayvanpour E, Kloos W, Backe C, Thanaraj A, Brefort T, Beier M, Hardt S, Meese E, Katus HA, Meder B. Multivariate miRNA signatures as biomarkers for non-ischaemic systolic heart failure. *Eur Heart J* 2013;**34**:2812–2822.
25. Haas J, Frese KS, Park YJ, Keller A, Vogel B, Lindroth AM, Weichenhan D, Franke J, Fischer S, Bauer A, Marquart S, Sedaghat-Hamedani F, Kayvanpour E, Kohler D, Wolf NM, Hassel S, Nietsch R, Wieland T, Ehlermann P, Schultz JH, Dosch A, Mereles D, Hardt S, Backs J, Hoheisel JD, Plass C, Katus HA, Meder B. Alterations in cardiac DNA methylation in human dilated cardiomyopathy. *EMBO Mol Med* 2013;**5**:413–429.
26. Movassagh M, Vujic A, Foo R. Genome-wide DNA methylation in human heart failure. *Epigenomics* 2011;**3**:103–109.