New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants

Charlotte Andreasen1,2,5, Jonas B Nielsen1,2,5, Lena Refsgaard1,2, Anders G Holst1,2, Alex H Christensen1,2, Laura Andreasen1,2, Ahmad Sajadieh3, Stig Haunso1,2,4, Jesper H Svendsen1,2,4 and Morten S Olesen*1,2

Cardiomyopathies are a heterogeneous group of diseases with various etiologies. We focused on three genetically determined cardiomyopathies: hypertrophic (HCM), dilated (DCM), and arrhythmogenic right ventricular cardiomyopathy (ARVC). Eighty-four genes have so far been associated with these cardiomyopathies, but the disease-causing effect of reported variants is often dubious. In order to identify possible false-positive variants, we investigated the prevalence of previously reported cardiomyopathy-associated variants in recently published exome data. We searched for reported missense and nonsense variants in the NHLBI-Go Exome Sequencing Project (ESP) containing exome data from 6500 individuals. In ESP, we identified 94 variants out of 687 (14%) variants previously associated with HCM, 58 out of 337 (17%) variants associated with DCM, and 38 variants out of 209 (18%) associated with ARVC. These findings correspond to a genotype prevalence of 1:4 for HCM, 1:6 for DCM, and 1:5 for ARVC. PolyPhen-2 predictions were conducted on all previously published cardiomyopathy-associated missense variants. We found significant overrepresentation of variants predicted as being benign among those present in ESP compared with the ones not present. In order to validate our findings, seven variants associated with cardiomyopathy were genotyped in a control population and this revealed frequencies comparable with the ones found in ESP. In conclusion, we identified genotype prevalences up to more than one thousand times higher than expected from the phenotype prevalences in the general population (HCM 1:500, DCM 1:2500, and ARVC 1:5000) and our data suggest that a high number of these variants are not monogenic causes of cardiomyopathy.

European Journal of Human Genetics (2013) 21, 918–928; doi:10.1038/ejhg.2012.283; published online 9 January 2013

Keywords: cardiomyopathy; exome; next-generation sequencing; HCM; DCM; ARVC

INTRODUCTION
Cardiomyopathy is a diverse group of cardiac disorders characterized by mechanical and/or electrical dysfunction of the cardiac muscle. The diseases are associated with significant morbidity and mortality and are a known risk factor for sudden cardiac death.1–3 Over time, several classification systems have evolved based on etiology, anatomy, physiology, or histopathological expression.4 In newer classification systems, major types include hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and arrhythmogenic right ventricular cardiomyopathy (ARVC).

Hypertrophic cardiomyopathy is characterized by a non-dilated, hypertrophic left ventricle with variable degrees of diastolic dysfunction, whereas DCM is characterized by dilated ventricular cavities and systolic dysfunction.4–7 In ARVC, progressive fibrofatty replacement of the normal cardiac tissue predisposes to ventricular tachycardia and sudden death.8,9 The prevalence of these three cardiomyopathies in the general population has been estimated to be 1:500, 1:2500, and 1:5000, respectively.3

Inherited cardiomyopathy has traditionally been considered a monogenic disorder and to date hundreds of variants in 84 genes have been associated with these syndromes. However, some associations are based on weak family phenotype–genotype co-segregation and/or the absence of the variant in a limited number of controls.

Until recently, there has only been limited knowledge regarding the genetic variation in the general population, especially with regard to low-frequency variants. This was changed in June 2011 when whole exome data from the NHLBI GO Exome Sequencing Project (ESP) was published (latest update June 2012).10 In order to identify possible false-positive cardiomyopathy variants reported in the literature, we aimed to investigate the prevalence of previously cardiomyopathy-associated variants in the new ESP exome data and compare the prevalence of these variants with the expected prevalences of monogenic cardiomyopathies in the same population.

METHODS
In ESP, next-generation sequencing of all protein coding regions in 6500 individuals, including both European Americans (4300 individuals) and African Americans (2200 individuals), from different population studies were carried out.10 No clinical data were available on the ESP population, nor at request. By literature search, we found inclusion and exclusion criteria on 9/12
cohort used in ESP. None of these has specifically included persons with cardiomyopathies or other heart diseases and at least two cohorts have excluded such patients.

The databases ARVD/C Genetic Variants database (last update April 2012)\textsuperscript{11} and The Human Gene Mutation Database (updated June 2012)\textsuperscript{12} were searched for missense and nonsense cardiomyopathy-associated variants involving the three major types; HCM, DCM, and ARVC. All genes in ARVD/C Genetic Variants database were evaluated and in HGM the search term ‘Cardiomyopathy’ was used. In total, 84 genes associated with cardiomyopathy were identified. Genes were then evaluated one by one and the ones associated with any of the above-mentioned cardiomyopathies were selected. Additionally, we included the recently reported DCM-associated TTN nonsense variants published by Herman et al\textsuperscript{13} in order to include all genes so far associated with DCM. All identified variants were then systematically searched for in ESP. Only variants classified by one of the databases as being pathogenic/disease causing were included in the analyses. Variants of unknown pathogenicity or variants classified as ‘disease-causing mutation’ are marked with \textsuperscript{b} in Tables 1–3, but in order to make a conservative approach, these variants were excluded from our calculations. Due to lack of data regarding variants positioned in promoters, introns and UTRs regions in ESP, these could not be included.

In addition to taking all identified variants associated with HCM, DCM, and ARVC, into account for the calculation of genotype prevalences, we also did a more conservative approach. Based on the frequencies of HCM, DCM, and ARVC in the general population (1:500, 1:2500, and 1:5000, respectively), the estimated number of individuals in the ESP data that can be expected to be affected by HCM, DCM, and ARVC are \(\approx 13\), 3, and 2, respectively. These values roughly represent the number of times a given variant with complete penetrance can be present in the exome database and still theoretically be the cause of monogenic forms of the respective cardiomyopathies.

The literature was searched for functional data and family co-segregation of all the cardiomyopathy-associated variants identified in the ESP population. Positive functional data were defined as any \textit{in vivo} or \textit{in vitro} model, demonstrating results differing from the wild-type model. Co-segregation was defined as at least two family members in two generations both having the phenotype and the genotype.

Additionally, we conducted a PolyPhen-2 prediction\textsuperscript{14} on all previously reported missense variants. Variants were, by PolyPhen-2, predicted to be ‘benign,’ ‘possible damaging,’ or ‘probably damaging.’ As nonsense variants cannot be evaluated by PolyPhen-2, we classified these as of ‘unknown pathogenicity.’ In an analysis, we evaluated differences in distributions of the four categories of pathogenicity between the variants identified in ESP vs variants not identified in ESP with the use of Fisher’s exact test. A \(P\)-value \(<0.05\) was considered as statistical significant. In case of a statistical significant difference, we also evaluated the difference in proportions of variants being predicted as benign for variants identified in ESP vs variants not identified in ESP, also with the use of Fisher’s exact test.

Using a Taqman assay as previously described,\textsuperscript{13} we genotyped seven variants with a pathogenic association and a prevalence in the proportion of ESP with European American ancestry high enough (10:6500) to have a modest chance of being detected in our own control population (\(N=534\)). The control population of Northern European ancestry consisted of men and women between the age of 55–75 years with no history of arrhythmias or other cardiac diseases and with available ECGs as previously described.\textsuperscript{16} The ECGs from geno-positive controls were evaluated by two independent experienced ECG readers with regard to the 2010 task force ECG criteria for ARVC\textsuperscript{17} and with regard to the Cornell\textsuperscript{18} and the Sokolow–Lyon criteria for ventricular hypertrophy.\textsuperscript{19}

\textbf{RESULTS}

\textbf{Hypertrophic cardiomyopathy}

In the ESP population, we identified 94 out of 687 variants previously associated with HCM (14%). Ninety-three missense and one nonsense variants were identified, affecting 1672 individuals in total (homozygote = 76, heterozygote = 1596). Eighteen variants with family co-segregation analyses and 16 variants with functional characterization different form wild-type were identified in ESP. On average, the genes investigated were sequenced in 6286 individuals, corresponding to a genotype prevalence of 1:4 (1672:6286). PolyPhen-2 analysis of the 94 HCM-associated variants present in ESP predicted 39 (41%) to be probably damaging, 14 (15%) to be possibly damaging, and 40 (43%) to be benign. Only one nonsense variant was found in ESP and classified as being of unknown pathogenicity (Table 1). Of the remaining 593 HCM-associated variants not present in ESP, 324 (55%) were predicted to be probably damaging, 108 (18%) possibly damaging and 107 (18%) were predicted to be benign. Fifty-four nonsense variants were classified as being of unknown pathogenicity. This difference in the distribution of the four categories of pathogenicity was statistical significant both for the overall comparison (\(P<0.0001\)) and when comparing the proportion of variants predicted to be benign for variants identified in ESP vs variants not identified in ESP (43% vs 18%, respectively, \(P<0.0001\)).

Fourteen of the 94 variants were identified in \(\geq13\) individuals, though above our conservative cutoff value. These variants affected a total of 1474 individuals, which is equivalent to a HCM genotype prevalence of 1:4 (1474:5810). If variants predicted to be benign by Polyphen-2 (43%) were additionally excluded, the genotype prevalence was 1:7. The cardiomyopathy-associated variants identified in the ESP population are listed in Table 1.

Two variants (\(MYBPC3\) p.V896M and \(MYH7\) p.M982T) were, based on our criteria, selected for genotyping in our control population. Five individuals were heterozygous carriers of the \(MYBPC3\) p.V896M variant and three carried the \(MYH7\) p.M982T variant. This corresponds to genotype prevalences of 0.94 and 0.56%, respectively, which are comparable to those found in ESP (0.96 and 0.44%, respectively; Table 4).

\textbf{Dilated cardiomyopathy}

In DCM, we found 58 out of 337 variants previously associated with DCM (17%). Two out of the 58 variants were nonsense variants. Both nonsense variants (\(LAMA4\) p.R1073X and \(VSP13A\) p.R3135X) were only found in a single individual and both were heterozygous for the variant. A total of 1043 individuals were affected. On average, the genes investigated have been screened in 6314 individuals, and this results in a DCM genotype prevalence of 1:6 (1043:6314). Four variants with convincing segregation analyses were identified and 26 variants were found to have functional effects. PolyPhen-2 analysis of the 58 DCM-associated variants predicted 26 (45%) to be probably damaging, 11 (19%) possibly damaging, and 19 (33%) variants were predicted to be benign whereas two nonsense variants were classified as being of unknown pathogenicity (Table 2). Of the remaining 279 DCM-associated variants, not present in ESP, 134 (48%) were predicted to be probably damaging, 43 (15%) possibly damaging, and 56 (20%) were predicted to be benign. Forty-six nonsense variants were classified as being of unknown pathogenicity. This difference in the distribution of the four categories of pathogenicity was statistical significant both for the overall comparison (\(P=0.013\)) and when comparing the proportion of variants predicted to be benign for variants identified in ESP vs variants not identified in ESP (33 vs 20%, respectively, \(P=0.039\)).

Thirty-five out of the 58 variants were identified in three or more individuals, though above our conservative cutoff value. These variants affected a total of 963 individuals giving a genotype prevalence of 1:7 (963:6334). If variants predicted to be benign by Polyphen-2 (33%) were additionally excluded, the genotype prevalence was 1:10. The DCM-associated variants identified in the ESP population are listed in Table 2.

\textsuperscript{919}
### Table 1: Variants associated with hypertrophic cardiomyopathy present in the ESP population

| Gene   | Variant | Amino acid | European Americans—genotype | African Americans—genotype | All—genotype | Family co-segregation/ functional effect | Polyphen-2 prediction |
|--------|---------|------------|-----------------------------|-----------------------------|--------------|----------------------------------------|----------------------|
|        |         |            | Minor/ | Minor/ | Major/ | Minor/ | Minor/ | Major/ | Minor/ | Major/ | Minor/ | Major/ |                           |                      |
|        |         |            | minor  | major  |        | minor  | major  |        | minor  | major  |        | minor  | major  |                           |                      |
| ACTC1  | c.1484C>T | T495M      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | Yes/NA | Benign                   |                      |
| ACTN2  | c.1814G>A | T123M      | 0     | 2      | 4298   | 0     | 1      | 2202   | 0     | 3      | 6500   | No/Yes | Benign                   |                      |
| ANKR    | c.838A>G  | I280V      | 0     | 3      | 4297   | 0     | 0      | 2203   | 0     | 3      | 6500   | No/Yes | Benign                   |                      |
| CALM3  | c.218G>A  | R73Q       | 0     | 0      | 4300   | 0     | 1      | 2202   | 0     | 1      | 6502   | No/No  | Possibly damaging         |                      |
| CALR3  | c.214G>A  | R82R       | 0     | 13     | 4287   | 0     | 0      | 2203   | 0     | 13     | 6490   | No/No  | Benign                   |                      |
| CASQ2  | c.80G>A   | A216T      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| CASQ3  | c.1468G>A | G526R      | 0     | 1      | 4299   | 0     | 1      | 2202   | 0     | 2      | 6502   | NA/Yes | Benign                   |                      |
| CAV3   | c.27G>A   | A82V       | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| COX5   | c.20G>A   | A67V       | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| COX7   | c.1493G>A | R498C      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| DES    | c.26G>A   | A82V       | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| JPH2   | c.1513G>A | G505S      | 0     | 19     | 2356   | 0     | 28     | 930    | 0     | 47     | 3286   | Yes/No | Benign                   |                      |
| KLF10  | c.610G>A  | A204T      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 2      | 6502   | NA/Yes | Benign                   |                      |
| MRPL3  | c.1469G>A | G526R      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/Yes | Benign                   |                      |
| MYH6   | c.166G>A  | G56R       | 33    | 693    | 3574   | 14    | 256    | 1933   | 47    | 949    | 5507    | No/NA  | Possibly damaging         |                      |
| MYH7   | c.777G>T  | A26V       | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/NA  | Benign                   |                      |
|        | c.322C>T  | A107V      | 0     | 1      | 4299   | 0     | 0      | 2203   | 0     | 1      | 6502   | NA/NA  | Benign                   |                      |
| MYBPC3 | c.13G>C   | G5R        | 0     | 7      | 4152   | 0     | 3      | 1964   | 10    | 6116   | No/NA  | Probable damaging         |                      |
| Gene | Variant | Amino acid change | African Americans—genotype | European Americans—genotype | All—genotype | Major/Minor | European Americans—genealogy | Family co—aggregation/functionaleffect | Polyphen—2 prediction |
|------|---------|------------------|-----------------------------|-----------------------------|-------------|-------------|-----------------------------|-------------------------------|-------------------------------|
| MYL2 | c.763C>A | c.763G>G | 473 | 473 < 473 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| MYLK2 | c.3481C>T | c.3481G>G | I246M | I246M < I246M | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| MYOZ2 | c.59A>G | c.59G>G | Y20C | Y20C < Y20C | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| MYPN | c.3481C>T | c.3481G>G | I246M | I246M < I246M | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| NEXN | c.13031G>A | c.13031G>G | A87V | A87V < A87V | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| PRKAG2 | c.2870C>T | c.2870G>G | T958I | T958I < T958I | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| RAF1 | c.13303G>A | c.13303G>G | R434Q | R434Q < R434Q | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SLC25A4 | c.3394G>A | c.3394G>G | N47K | N47K < N47K | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Questioning cardiomyopathy-associated genetic variants

C. Andreasen et al.

European Journal of Human Genetics
Two variants (CSRP3 p.W4R and MYH6 p.A1004S) were selected for genotyping in our control population. Six individuals were heterozygote carriers of the CSRP3 p.W4R variant and two carried the MYH6 p.A1004S variant. The prevalences were thus comparable to the ones in ESP (1.12 vs 1.07% and 0.37 vs 0.26%, respectively; Table 4). One individual carrying the CSRP3 p.W4R variant fulfilled the Cornell ECG criteria for ventricular hypertrophy; however, this individual died at the age of 73 and was never diagnosed with cardiomyopathy. The ECGs from the rest of the genotype-positive individuals were normal and without signs of ventricular hypertrophy.

### Arrhythmogenic right ventricular cardiomyopathy

Thirty-eight out of 209 variants associated with ARVC (18%) were found in the ESP population. One nonsense and 37 missense variants were identified, affecting a total of 1404 individuals. Only one variant with convincing family co-segregation and three variants with functional characterization different from wild-type were identified in ESP. Twenty-eight of the 38 variants were identified in two or more individuals. On average, the genes investigated in ARVC have been sequenced in 6354 individuals thus corresponding to an ARVC genotype prevalence of 1:5 (1407:6354). PolyPhen-2 analysis of the 38 ARVC-associated variants predicted 14 (37%) to be probably damaging, 3 (8%) to be possibly damaging, and 20 (53%) to be predicted to be benign whereas one nonsense variant was classified as being of unknown pathogenicity (Table 3). Of the remaining 171 ARVC-associated variants, not present in ESP, 77 (45%) were predicted to be probably damaging, 14 (8%) possibly damaging, and 21 (12%) predicted to be benign. Fifty-nine nonsense variants were classified as being of unknown pathogenicity. This difference in the distribution of the four categories of pathogenicity was statistically significant both for the overall comparison ($P < 0.0001$) and when comparing the proportion of variants predicted to be benign for variants identified in ESP vs variants not identified in ESP (53 vs 12%, respectively, $P < 0.001$).

Twenty-eight variants were present in two or more individuals, though above our conservative cutoff value, and this still corresponded to an ARVC genotype prevalence of 1:5 (1393:6359). If variants predicted to be benign by Polyphen-2 (53% were additionally excluded, the genotype prevalence was 1:11. The ARVC-associated variants identified in the ESP population are listed in Table 3.

Three variants (PKP2 p.D26N; DSG2 p.V158G; and DSP p.V30M) were genotyped in the control population and five individuals were heterozygote carriers of the PKP2 p.D26N variant, nine of DSG2 p.V158G, and five of DSP p.V30M. One individual was carrier of both the DSG2 p.V158G and the DSP p.V30M variant. The variant frequencies were comparable to those found in ESP (0.94 vs 1.37%; 1.69 vs 1.58%; and 0.94 vs 0.37%, respectively; Table 4). ECGs from genotype-positive individuals were normal and without signs of ARVC or ventricular hypertrophy.

### DISCUSSION

The present study identified a high prevalence of cardiomyopathy-associated genetic variants in recently published population-based exome data. Fourteen percent of all previously HCM-associated variants and 18% of all DCM- and ARVC-associated variants were identified in ESP. Thus, a much higher prevalence of cardiomyopathy-associated genetic variants were identified in ESP than expected from the phenotype prevalences in the general population.

In order to validate the marked overrepresentation of variants associated with HCM, DCM, and ARVC in ESP, we genotyped seven variants in seven different genes associated with cardiomyopathy in a
### Table 2: Variants associated with dilated cardiomyopathy present in the ESP population

| Gene     | Variant          | Amino acid | European Americans—genotype | African Americans—genotype | All—genotype | Family co-segregation/functional effect | Polyphen-2 prediction |
|----------|------------------|------------|------------------------------|---------------------------|--------------|----------------------------------------|-----------------------|
| ABC9     |                  |            |                              |                           |              |                                        |                       |
| ACTC1    |                  |            |                              |                           |              |                                        |                       |
| ACTN2    | c.26A>G          | Q9R        | 0                            | 7                         | 4293         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| ANKRD1   | c.197G>A         | R66Q       | 0                            | 9                         | 4291         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| CRYAB    | c.5016T>A        | A1095N     | 0                            | 2                         | 4293         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| CSRP3    | c.1010C>T        | R337P      | 0                            | 2                         | 4299         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| CTBF1    | c.893C>T         | S298L      | 0                            | 2                         | 4293         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| DMD      | c.5016T>A        | A1095N     | 0                            | 2                         | 4293         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| DNAJC19  |                  |            |                              |                           |              |                                        |                       |
| DOLK     |                  |            |                              |                           |              |                                        |                       |
| DSC2     | c.907G>A         | V303M      | 0                            | 2                         | 4293         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| DSP      | c.1003A>G        | T335A      | 0                            | 5                         | 4094         | 0                                      | 1863                  |
|          |                  |            |                              |                           |              |                                        |                       |
| EYA4     |                  |            |                              |                           |              |                                        |                       |
| FKB1     |                  |            |                              |                           |              |                                        |                       |
| FOXD4    |                  |            |                              |                           |              |                                        |                       |
| GATA1    |                  |            |                              |                           |              |                                        |                       |
| HSPB7    |                  |            |                              |                           |              |                                        |                       |
| ILK      |                  |            |                              |                           |              |                                        |                       |
| LAMA2    |                  |            |                              |                           |              |                                        |                       |
| LAMC3    | c.3217C>T        | p.R1073X   | 0                            | 1                         | 4299         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| LDB3     | c.349G>A         | D117N      | 0                            | 33                        | 4122         | 0                                      | 1994                  |
|          |                  |            |                              |                           |              |                                        |                       |
| LMX1     | c.1042C>T        | T351I      | 0                            | 1                         | 4299         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| LMNA     | c.2092G>A        | A698T      | 0                            | 5                         | 4094         | 0                                      | 1863                  |
|          |                  |            |                              |                           |              |                                        |                       |
| MURC     | c.384C>A         | G128E      | 0                            | 5                         | 4295         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| MYBPC3   | c.961G>A         | V321M      | 0                            | 4                         | 4295         | 0                                      | 2203                  |
|          |                  |            |                              |                           |              |                                        |                       |
| NCKX1     | c.1814A>G        | G605V      | 0                            | 1                         | 4171         | 0                                      | 2008                  |

*Note:* The table continues with additional columns and variants for each gene, including amino acid changes, genotype information, and functional effects. The last column indicates whether the variant is considered benign, possibly damaging, probably damaging, or unknown based on Polyphen-2 prediction.
Table 2 (Continued)

| Gene | Variant | Amino acid | European Americans—genotype | African Americans—genotype | All—genotype | Family co-segregation/ functional effect | Polyphen-2 prediction |
|------|---------|------------|-----------------------------|-----------------------------|-------------|------------------------------------------|----------------------|
|      |         |            | Minor/ major                 | Minor/ major                 | Minor/ major | Minor/ major |                                |                      |
| MYH6 | c.824T>A | I275N      | 0  3  4297                   | 0  0  2203                   | 0  3  6500   | NA/NA        | Benign                           |                      |
|      | c.1702C>T | R568C      | 0  1  4299                   | 0  1  2202                   | 0  2  6501   | NA/NA        | Benign                           |                      |
|      | c.3010G>C | A1004S     | 0  11 4289                  | 0  1  2202                   | 0  12 6491   | No/NA        | Benign                           |                      |
|      | c.4318G>C | A1440P     | 0  1  4297                   | 0  0  2203                   | 0  3  6500   | NA/NA        | Probably damaging                 |                      |
| MYH7 | c.2890G>C | V646L      | 0  6  4294                   | 0  0  2203                   | 0  6  6497   | NA/NA        | Probably damaging                 |                      |
|      | c.3286G>T | D1096Y     | 0  2  4298                   | 0  0  2203                   | 0  2  6501   | NA/NA        | Probably damaging                 |                      |
|      | c.4505G>A | R1502Q     | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | No/NA        | Probably damaging                 |                      |
|      | c.5494C>T | R1863C     | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | No/NA        | Probably damaging                 |                      |
| MYPN | c.59A>G  | Y20C       | 0  1  4290                   | 0  0  2201                   | 0  12 6491   | NA/Yes       | Probably damaging                 |                      |
|      | c.333C>T | P1112L     | 0  21 4279                  | 0  6  2197                   | 0  27 6476   | No/Yes       | Probably damaging                 |                      |
|      | c.3583G>A | V1195M     | 0  0  4300                   | 0  16 2187                  | 0  17 6485   | NA/Yes       | Benign                           |                      |
| NEBL | c.180G>C | K60N       | 0  61 4236                  | 0  6  2196                   | 0  67 6432   | No/Yes       | Probably damaging                 |                      |
|      | c.604A>G | G202R      | 0  19 4281                  | 0  4  2199                   | 0  23 6480   | No/Yes       | Benign                           |                      |
|      | c.1776C>A | A592E      | 0  1  4298                   | 0  16 2187                  | 0  17 6485   | NA/Yes       | Benign                           |                      |
| NEXN | c.1955A>G | Y652C      | 0  1  4086                   | 0  0  843                  | 0  17 6499   | NA/Yes       | Probable damaging                |                      |
| PKP2 | c.184C>T | W62K       | 0  3  4188                   | 0  0  2110                   | 0  3  6298   | NA/Yes       | Benign                           |                      |
|      | c.419C>T | S140F      | 0  25 4275                  | 0  1  2202                   | 0  26 6477   | NA/Yes       | Benign                           |                      |
|      | c.505A>G | S169G      | 0  16 4281                  | 0  5  2198                   | 0  21 6479   | NA/Yes       | Benign                           |                      |
|      | c.1759G>A | V587I      | 0  40 4260                  | 0  2  2201                   | 0  42 6461   | NA/Yes       | Possibly damaging                |                      |
| PLN  | c.2207A>G | N736S      | 0  2  4297                   | 0  0  2195                   | 0  7  6492   | NA/NA        | Possibly damaging                |                      |
|      | PSEN1   |            |                             |                             |             |             |                                    |                      |
| RBM20| c.1364C>T | S455L      | 0  33 1558                  | 0  0  692                   | 0  33 2250   | NA/NA        | Benign                           |                      |
|      | c.2147G>A | R716Q      | 0  1500                    | 0  0  692                   | 0  1  2282   | Yes/NA       | Probably damaging                |                      |
| SCN5A| c.647C>T | S216L      | 0  11 4205                  | 0  1  2044                  | 0  12 6249   | NA/Yes       | Probable damaging                |                      |
|      | c.1336G>A | E446K      | 0  1  4188                   | 0  0  2050                   | 0  1  6238   | No/NA        | Possibly damaging                |                      |
|      | c.3835G>A | V1279I     | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | No/NA        | Possibly damaging                |                      |
|      | c.6013G>A | P2005A     | 0  13 4169                  | 0  1  2005                   | 0  14 6174   | NA/NA        | Benign                           |                      |
| SGCD |         |            |                             |                             |             |             |                                    |                      |
| SYNE1|         |            |                             |                             |             |             |                                    |                      |
| SYNM |         |            |                             |                             |             |             |                                    |                      |
| TCF21| c.250G>A | R87Q       | 0  1  4290                   | 0  1  2194                   | 0  2  6484   | NA/Yes       | Probably damaging                |                      |
|      | c.65G>T  | G22V       | 0  17 4283                  | 84 598 1521                 | 84 615 5804  | NA/NA        | Benign                           |                      |
| TMPO | c.2068C>T | R690C      | 0  4  4295                  | 0  4  2196                  | 0  8  6491   | No/NA        | Probably damaging                |                      |
| TNN1 |         |            |                             |                             |             |             |                                    |                      |
| TNN3 |         |            |                             |                             |             |             |                                    |                      |
| TTN  | c.96206G>A | R32069Q   | 0  0  4115                   | 0  3  1920                   | 0  3  6035   | NA/NA        | Probably damaging                |                      |
|      | c.83C>T  | A28V       | 0  4  4296                  | 0  0  2203                   | 0  4  6499   | NA/NA        | Benign                           |                      |
| TTN2 |         |            |                             |                             |             |             |                                    |                      |
| TNNT2| c.306G>T | S1025G     | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | Yes/Yes      | Probably damaging                |                      |
| TPM1 |         |            |                             |                             |             |             |                                    |                      |
| TXNRD2|          |            |                             |                             |             |             |                                    |                      |
| VCL  | c.2923C>T | R975W      | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | NA/NA        | Benign                           |                      |
|      | c.3373C>T | R1125C     | 0  1  4299                   | 0  0  2203                   | 0  1  6502   | NA/NA        | Benign                           |                      |
| VPS13A|            |            |                             |                             |             |             |                                    |                      |

aNA indicates no data available.
bLikely disease-causing mutation, but with questionable pathogenicity.
| Gene | Variant | Amino acid | European Americans—genotype | African Americans—genotype | All—genotype | Family co-segregation/functional effect | Polyphen-2 prediction |
|------|---------|------------|-----------------------------|---------------------------|-------------|----------------------------------------|-------------------------|
| DES  | c.304G>A | E102K      | 0 7 4291                    | 0 3 2200                  | 0 10 6491   | No/Yes                                | Benign                  |
| DES  | c.372A>G  | I109M      | 0 1 4295                    | 0 1 2201                  | 0 2 6496    | Yes/NA                                | Benign                  |
| DES  | c.1018A>G | T340A      | 0 1 4299                    | 0 0 2202                  | 0 1 6501    | NA/NA                                | Probably damaging       |
| DES  | c.1721G>A | S574N      | 0 4 4296                    | 0 0 2203                  | 0 4 6499    | NA/NA                                | Benign                  |
| DES  | c.2194T>G | L732V      | 0 11 4289                   | 0 3 2200                  | 0 14 6489   | NA/NA                                | Benign                  |
| DSC2 | c.2471C>T | S824L      | 0 1 4299                    | 0 2 2201                  | 0 3 6500    | No/NA                                | Probably damaging       |
| DSC2 | c.2587G>A | G863R      | 0 5 4295                    | 0 1 2202                  | 0 6 6497    | NA/NA                                | Probably damaging       |
| DSG2 | c.166G>A  | V56M       | 0 29 4082                   | 0 2 1876                  | 0 31 5998   | NA/NA                                | Probably damaging       |
| DSG2 | c.473T>G  | V158G      | 0 65 4099                   | 0 9 1864                  | 0 74 5923   | Yes/NA                                | Benign                  |
| DSG2 | c.581C>T  | S194L      | 0 1 4098                    | 0 0 1858                  | 0 1 5956    | NA/NA                                | Probably damaging       |
| DSG2 | c.716T>G  | V239A      | 0 1 4140                    | 0 2 1908                  | 0 3 6048    | NA/NA                                | Probably damaging       |
| DSP  | c.1478A>G | N493S      | 0 1 4150                    | 0 1 1936                  | 0 2 6086    | NA/NA                                | Probably damaging       |
| DSP  | c.1550C>T | A517V      | 0 0 4174                    | 0 21 1940                 | 0 21 6114   | NA/NA                                | Probably damaging       |
| DSP  | c.1912A>G | G638R      | 0 2 4132                    | 0 0 1914                  | 0 2 6046    | NA/NA                                | Probably damaging       |
| DSP  | c.2137G>A | E713K      | 16 620 6558                 | 0 72 1929                 | 16 692 6587 | NA/NA                                | Benign                  |
| DSP  | c.2434G>T | G812C      | 0 1 4084                    | 0 0 1845                  | 0 1 5929    | NA/NA                                | Probably damaging       |
| DSP  | c.2795T>G | V920G      | 0 41 4094                   | 0 6 1942                  | 0 47 6036   | NA/NA                                | Probably damaging       |
| JUP  | c.88G>A   | V30M       | 0 16 4253                   | 0 0 2169                  | 0 16 6422   | NA/Yes                                | Benign                  |
| JUP  | c.688G>A  | D230N      | 1 3 4296                    | 0 0 2203                  | 1 3 6499    | NA/NA                                | Probably damaging       |
| JUP  | c.1696G>A | A566T      | 0 1 4295                    | 0 0 2203                  | 0 5 6498    | NA/NA                                | Probably damaging       |
| JUP  | c.2422C>T | R808C      | 0 2 4298                    | 0 0 2203                  | 0 2 6501    | NA/NA                                | Probably damaging       |
| JUP  | c.1219G>A | V407I      | 1 1 4299                    | 0 0 2203                  | 0 1 6502    | NA/NA                                | Probably damaging       |
| JUP  | c.2723G>A | R908H      | 0 13 4287                   | 0 1 2202                  | 0 14 6489   | NA/NA                                | Possibly damaging       |
| PKP2 | c.766>G>A | D26N       | 0 56 4044                   | 0 4 2068                  | 0 60 6112   | NA/NA                                | Possibly damaging       |
| PKP2 | c.184G>C  | G62K       | 0 3 4188                    | 0 0 2110                  | 0 3 6298    | NA/NA                                | Benign                  |
| PKP2 | c.419C>T  | S140F      | 0 25 4275                   | 0 1 2202                  | 0 26 6477   | NA/NA                                | Benign                  |
| PKP2 | c.505A>G  | S169G      | 0 16 4281                   | 0 5 2198                  | 0 21 6479   | NA/NA                                | Benign                  |
| PKP2 | c.1114G>C  | A372P      | 0 6 4294                    | 0 0 2203                  | 0 6 6497    | NA/NA                                | Benign                  |
| PKP2 | c.1237C>T | R413X      | 0 0 4300                    | 0 1 2202                  | 0 1 6502    | NA/NA                                | Benign                  |
| PKP2 | c.1465G>A | G489R      | 0 1 4295                    | 0 0 2202                  | 0 1 6497    | NA/NA                                | Benign                  |
| PKP2 | c.1577C>T | T526M      | 0 20 4280                   | 0 28 2175                 | 0 48 6455   | NA/NA                                | Possibly damaging       |
second population with clinical data available and no history of arrhythmias or other cardiac diseases. Thirty-four out of the 534 control subjects carried at least one of the variants. The seven genotyped variants were present with frequencies comparable with those found in ESP (Table 4), and all geno-positive controls, except from one individual, had ECGs without any signs of cardiomyopathy (eg, no hypertrophy or signs of ARVC). Thus, overrepresentation of cardiomyopathy-associated variants in ESP does not seem to be a major problem. In a recent paper, we also established that prevalences of four other variants genotyped in a control population were comparable to those of ESP. These results thus indicate that ESP consists of individuals representative of the general population.

A control population with available echocardiograms would have been preferable, but such a control population was not available. However, symptoms and signs of cardiomyopathy do not usually appear beyond the age of 50–60 years in these diseases, and also 75–95% of ARVC and HCM patients display ECG abnormalities. This indicates that our control population is well suited since it consists of 534 people all above the age of 55 with no reported signs of cardiovascular diseases. It is of course possible that a small fraction of the control population might develop cardiomyopathy in a very late age and that variant carriers are displaying reduced penetrance. However, this is not very likely, since we found a high number of carriers of the seven genotyped variants, and the fact that all geno-positive individuals except one had ECGs without any signs of cardiomyopathy and no history of cardiac diseases.

A genotype prevalence of 1:4 for HCM, 1:6 for DCM, and 1:5 for ARVC is unlikely to be caused by reduced or age-related penetrance. Even when taking into consideration a penetrance as low as 20% (reported for some genotypes), it would still result in genotype prevalences being massively overrepresented.

PolyPhen-2 predicted a statistically significant higher proportion of the variants present in ESP to be benign compared with the variants not present in ESP (43 vs 18% for HCM, 33 vs 20% for DCM, and 53 vs 12% for ARVC). This analysis further questions the pathogenic role of at least some of the variants present in ESP.
In the lack of phenotypic data available on the ESP population, we defined a cutoff value based on the expected prevalences of the respective cardiomyopathies in the same population. In this definition, variants with prevalence above this cutoff were assumed not to be monogenic causes of cardiomyopathy. However, taking this conservative cutoff into account revealed genotype prevalences similar to the ones obtained when including all cardiomyopathy-associated variants. Such a cutoff is, of course, somewhat arbitrary because of uncertainty regarding true prevalences of the cardiomyopathies in the general population (ESP) and because variants with reduced penetrance or recessive inheritance are not taken into account. However, most variants listed in the ARVC database and in HGMD are reported as monogenic, autosomal dominant causes of the cardiomyopathies.

Interpretation of the significance of the cardiomyopathy-associated variants with prevalences below our cutoff and thus present in a very low frequency in the ESP data is much less straightforward. These rare variants may be monogenic causes of cardiomyopathy, disease-modifiers, or benign. A small number of studies have associated genetic variation with increased susceptibility for cardiomyopathy in a non-monogenic manner. For this reason, we can only exclude high-prevalent variants as monogenic causes of cardiomyopathy, but we cannot make a conclusion about possible disease-modifying effects.

It is noteworthy that four genes associated with HCM in the HGMD database (COX15, OBSCN, SRI, and VCL) only had one variant that has been associated with cardiomyopathy (COX15 p.R217W, OBSCN p.4344Q, SRI p.F112L, and VCL p.L277M). These four variants were also present in ESP and both OBSCN p.4344Q and SRI p.F112L had prevalences above our defined cutoff values. Similarly, in five genes associated with DCM only one variant was identified in each gene (DSG2 p.T335A, FLT1 p.R54S, POLG p.N736S, TMPO p.R690C, and VPS13A p.R3135X) and all of these were also present in ESP. Only DSG2 p.T335A and VPS13A p.R3135X were below our cutoff value. Our data suggest that the genes OBSCN, SRI, FLT1, POLG, and TMPO require a revaluation regarding their disease causation with HCM and DCM.

A number of variants with functional effects or family co-segregation were identified in ESP. Functional characterization and family co-segregation analyses within families are valuable tools in determining the pathogenicity of identified sequence variants. However, small family sizes and reduced penetrance often hampers segregation analyses. In addition, functional characterization in model systems may not be representative of in vivo human physiology and an observed difference in a model system may not be of clinical importance. As an example, the CSRP3 (alternative symbol MLP) p.W4R variant has been associated with cardiomyopathy in functional systems, but lack of family co-segregation has also been reported.

Genetic screening is gaining ground in the identification of patients and family members at an increased risk of cardiomyopathies. Identification of a misclassified genetic variant in cardiomyopathy patients might lead to erroneous risk stratification, misdiagnosis of family members and this could have potentially devastating clinical consequences. It is therefore important that variants being reported as causative of cardiomyopathies are truly disease causing.

In conclusion, we identified a massive overrepresentation of previously cardiomyopathy-associated genetic variants in new population-based exome data. With genotype prevalences up to one thousand times higher than expected from the phenotype prevalence in the general population, we suspect a high number of these genetic variants to be only modest disease-modifiers or even non- pathetic.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This work was supported by The Danish National Research Foundation Centre for Cardiac Arrhythmia, The John and Birthe Meyer Foundation, The Research Foundation of the Heart Centre Rigshospitalet, The Arvid Nilsson Foundation, Fondenbørsværselener Henry Hansen og Hustru Karla Hansen, fonde Westergaards Legat, The Research Foundation at Rigshospitalet, Direktør Ib Henriksen’s fond, and Villadsen Family Foundation.

1. Kimura A: Molecular etiology and pathogenesis of hereditary cardiomyopathy. Circ J 2008; 72: Suppl A: A38–A48.
2. Kimura A: Molecular basis of hereditary cardiomyopathy: abnormalities in calcium sensitivity, stretch response, stress response and beyond. J Hum Genet 2010; 55: 81–91.
3. Reij H, Alberc G, Sago BS, Burton H, Behr ER: Inherited cardiomyopathies. BMJ 2011; 343: d6966–d6966.
4. Maron BJ, Toebin JA, Thiene G et al: Contemporary definitions and classification of the cardiomyopathies. Circulation 2006; 113: 1807–1816.
5. Wesler R, Elton T, Pleister A, Feldman D: Cardiomyopathy: an overview. Am Fam Physician 2009; 79: 778–784.
6. Elliott P, Andersson B, Arbustini E et al: Classification of the cardiomyopathies: a position statement from the European society of cardiology working group on myocardial and pericardial diseases. Eur Heart J 2008; 29: 270–276.
7. Marhan AJ, Salek L, Lutucuta S: Molecular genetics and pathogenesis of hypertrophic cardiomyopathy. Minerva Med 2001; 92: 435–451.
8. Saffitz JE: Arrhythmogenic cardiomyopathy. Circulation 2011; 124: e390–e392.
9. Kaplingier JD, Landstrom KE, Salisbury BA et al: Distinguishing arrhythmogenic right ventricular cardiomyopathy/dysplasia-associated mutations from background genetic noise. J Am Coll Cardiol 2011; 57: 2317–2327.
10. Exome Variant Server: Available at: http://exac.broadinstitute.org/Exomesequence.
11. A genetic variants database for arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J 2005; 26: 1402–1409.
12. Marcus FI, McKenna WJ, Sherrick D et al: Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J 2010; 31: 806–814.
13. Levy H, Raffel HB, Tanguay JF et al: Hypertrophic cardiomyopathy: a systematic review. JAMA 2004; 291: 2457–2464.
14. Reesink HP, Dumont K, Eggink G: The pathophysiology of hypertrophic cardiomyopathy. J Am Coll Cardiol 2006; 47: 1580–1588.
15. Sadosh A, Nielsen OW, Rasmussen V, Hein HO, Hansen JF: Prevalence and prognostic significance of daily-life silent myocardial ischaemia in middle-aged and elderly subjects with no apparent heart disease. Eur Heart J 2005; 26: 1402–1409.
16. Marcus FI, McKenna WJ, Sherrick D et al: Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J 2010; 31: 806–814.
17. Cano JF, Neves RB, Alonso DR, Campo E, Kligfield P: Improved sex-specific criteria of left ventricular hypertrophy for clinical and computer interpretation of electrocardiograms: validation with autopsy findings. Circulation 1987; 75: 565–572.
18. Sokolow M, Lyon TP: The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. Am Heart J 1949; 37: 161–186.
19. Reisgaard L, Holst A, Sajdajieh G et al: High prevalence of genetic variants previously associated with LQT syndrome in new exome data. Eur J Hum Genet 2012; 20: 505–908.
20. Taitt S, Ahl E, Sorensen J, Borresen A: Dilated cardiomyopathy: a review. J Clin Pathol 2009; 62: 219–225.
21. Basso C, Corrado D, Marcus F, Nava A, Thiene G: Arrhythmogenic right ventricular cardiomyopathy. Lancet 2009; 373: 1289–1300.
22. Christians I, Carrier L: Hypertrophic cardiomyopathy. I: Baars HF, Doevendans PAFM, van der Smagt JJ, ed. Clinical Cardiogenetics. London: Springer, 2001; 27–61.
23. Maron BJ: Hypertrophic cardiomyopathy: a systematic review. JAMA 2002; 287: 1308–1320.
24. Azouaou A, Churzidse S, Korora T, Erdel R: Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a review and update. Clin Res Cardiol 2011; 100: 383–394.
25. Rampersaud E, Kinnaman DD, Hamilton K et al: Common susceptibility variants examined for association with dilated cardiomyopathy. Ann Hum Genet 2010; 74: 110–116.
27 Stark K, Esslinger UB, Reinhard W et al: Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. PLoS Genet 2010; 6: e1001167.
28 Lahtinen AM, Lehtonen E, Marjamaa A et al: Population-prevalent desmosomal mutations predisposing to arrhythmogenic right ventricular cardiomyopathy. Heart Rhythm 2011; 8: 1214–1221.
29 Knöll R, Hoshijima M, Hoffman HM et al: The cardiac mechanical stretch sensor machinery involves a Z Disc Complex that is Defective in a Subset of Human Dilated Cardiomyopathy. Cell 2002; 111: 943–955.
30 Knöll R, Kostin S, Klede S et al: A common MLP (Muscle LIM Protein) variant is associated with cardiomyopathy. Circ Res 2010; 106: 695–704.

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