European Journal of Human Genetics (2013) 21, 918–928

www.nature.com/ejhg

EJHG Open

ARTICLE

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

Charlotte Andreasen[1,2,5], Jonas B Nielsen[1,2,5], Lena Refsgaard[1,2], Anders G Holst[1,2], Alex H Christensen[1,2], Laura Andreasen[1,2], Ahmad Sajadieh[3], Stig Haunsø[1,2,4], Jesper H Svendsen[1,2,4] 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 (2203 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

[1]The Danish National Research Foundation Centre for Cardiac Arrhythmia, Copenhagen, Denmark; [2]Laboratory of Molecular Cardiology, The Heart Centre, University Hospital of Copenhagen, Rigshospitalet, Denmark; [3]Department of Cardiology, Copenhagen University Hospital of Bispebjerg, Bispebjerg, Denmark; [4]Department of Medicine and Surgery, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark
[5]These authors contributed equally to this work.
*Correspondence: MS Olesen, Department of Cardiology, Rigshospitalet, Laboratory for Molecular Cardiology, 9312, Juliane Mariesvej 20, Copenhagen O, OE 2100, Denmark. Tel: +45 3545 6506; Fax: +45 3545 6500; E-mail: morten.salling.olesen@rh.regionh.dk
Received 25 April 2012; revised 4 September 2012; accepted 23 November 2012; published online 9 January 2013

cohorts 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)[11] and *The Human Gene Mutation Database* (updated June 2012)[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 HGMD 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*[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 '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 $\sim$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 *in vivo* or *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[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,[15] 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.[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[17] and with regard to the Cornell[18] and the Sokolow–Lyon criteria for ventricular hypertrophy.[19]

RESULTS

Hypertrophic cardiomyopathy

In the ESP population, we identified 94 out of 687 variants previously associated with HCM (14%). Ninety-tree 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 $\geq$13 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).

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.

Table 1 Variants associated with hypertrophic cardiomyopathy present in the ESP population

			European Americans—genotype			*African Americans—genotype*			*All—genotype*				
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
ACTC1	—	—	—	—	—	—	—	—	—	—	—	—	—
ACTN2	c.1484C>T	T495M	0	1	4299	0	0	2203	0	1	6502	Yes/NA	Benign
AGK	—	—	—	—	—	—	—	—	—	—	—	—	—
ANKRD1	c.368C>T	T123M	0	2	4298	0	1	2202	0	3	6500	No/Yes	Benign
	c.838A>G	I280V	0	3	4297	0	0	2203	0	3	6500	No/Yes	Benign
CALM3	—	—	—	—	—	—	—	—	—	—	—	—	—
CALR3	c.218G>A	R73Q	0	0	4300	0	1	2202	0	1	6502	No/No	Probably damaging
	c.245A>G	K82R	0	13	4287	0	0	2203	0	13	6490	No/No	Benign
CASQ2	—	—	—	—	—	—	—	—	—	—	—	—	—
CAV3	—	—	—	—	—	—	—	—	—	—	—	—	—
COA5	—	—	—	—	—	—	—	—	—	—	—	—	—
COX15	c.649C>T	R217W	0	0	4300	0	1	2202	0	1	6502	NA/Yes	Probably damaging
CSRP3	c.136A>C	S46R	0	3	4290	0	0	2199	0	3	6489	Yes/NA	Probably damaging
	c.190C>T	R64C	0	2	4291	0	0	2199	0	2	6490	No/NA	Possibly damaging
	c.299G>A	R100H	0	7	4286	0	0	2199	0	7	6485	NA/NA	Possibly damaging
DES	—	—	—	—	—	—	—	—	—	—	—	—	—
FXN	—	—	—	—	—	—	—	—	—	—	—	—	—
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
	c.674G>A	S225N	0	1	4299	0	0	2203	0	1	6502	NA/Yes	Benign
MRPL3	—	—	—	—	—	—	—	—	—	—	—	—	—
MYH6	c.166G>A[b]	G56R	33	693	3574	14	256	1933	47	949	5507	No/NA	Possibly damaging
MYH7	c.77C>T	A26V	0	1	4299	0	0	2203	0	2	6502	NA/NA	Benign
	c.115G>A	V39M	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.958G>A	V320M	0	2	4298	0	0	2203	0	2	6501	No/NA	Probably damaging
	c.976G>C	A326P	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
	c.1988G>A	R663H	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
	c.2183C>T	A728V	0	2	4298	0	1	2202	0	3	6500	Yes/NA	Benign
	c.2359C>T	R787C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Possibly damaging
	c.2360G>A	R787H	0	0	4300	0	1	2202	0	1	6502	NA/NA	Possibly damaging
	c.2389G>A	A797T	0	0	4300	0	1	2202	0	1	6502	Yes/Yes	Benign
	c.2585C>T[b]	A862V	0	0	4300	0	1	2202	0	1	6502	NA/NA	Benign
	c.2608C>T	R870C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.2945T>C	M982T	0	19	4281	0	3	2200	0	22	6481	NA/NA	Possibly damaging
	c.3301G>A	G1101S	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
	c.3981C>A	N1327K	0	1	4298	0	0	2202	0	1	6500	Yes/NA	Possibly damaging
	c.4052C>T	T1351M	0	2	4298	0	0	2203	0	2	6501	NA/NA	Benign
	c.4258C>T	R1420W	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.4377G>T	K1459N	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.4423C>T[b]	R1475C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.4472C>G[b]	S1491C	0	99	4201	0	13	2190	0	112	6391	NA/NA	Benign
	c.4909G>A	A1637T	0	1	4299	0	5	2198	0	6	6497	NA/NA	Benign
	c.5305C>A	L1769M	0	2	4295	0	0	2203	0	2	6498	NA/NA	Probably damaging
	c.5536C>T	R1846C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.5561C>T	T1854M	0	1	4299	0	0	2203	0	1	6502	NA/NA	Possibly damaging
MYBPC3	c.13G>C	G5R	0	7	4152	0	3	1964	0	10	6116	No/NA	Probably damaging
	c.184A>C	T62P	0	3	4286	0	0	2193	0	3	6479	NA/NA	Benign
	c.223G>A	D75N	0	1	4272	0	0	2191	0	1	6463	NA/NA	Probably damaging
	c.461T>C	I154T	0	0	4149	0	2	2020	0	2	6169	NA/NA	Benign
	c.478C>T	R160W	0	1	4150	0	1	2025	0	2	6175	NA/NA	Probably damaging
	c.529C>T	R177C	0	1	4198	0	0	2065	0	1	6263	NA/NA	Probably damaging
	c.624G>C	Q208H	0	5	4214	0	0	2112	0	5	6326	NA/NA	Probably damaging
	c.646G>A	A216T	0	1	4203	0	6	2068	0	7	6271	NA/NA	Benign
	c.649A>G[b]	S217G	0	10	4149	0	2	2073	0	12	6267	NA/NA	Possibly damaging
	c.682G>A	D228N	0	1	4193	0	0	2093	0	1	6286	NA/NA	Benign

Table 1 (Continued)

Gene	Variant	Amino acid	European Americans—genotype			African Americans—genotype			All—genotype			Family co-segregation/ functional effect[a]	Polyphen-2 prediction
			Minor/ minor	Minor/ major	Major/ major	Minor/ minor	Minor/ major	Major/ major	Minor/ minor	Minor/ major	Major/ major		
	c.706A>G[b]	S236G	69	922	3206	8	223	1851	77	1145	5057	No/NA	Possibly damaging
	c.787G>A	G263R	0	1	4073	0	0	1894	0	1	5967	NA/NA	Possibly damaging
	c.818G>A	R273H	0	1	4061	0	0	1903	0	1	5964	Yes/NA	Probably damaging
	c.833G>A	G278E	0	1	4109	0	62	1909	0	63	6018	NA/NA	Benign
	c.836G>C	G279A	0	0	4114	0	1	1972	0	1	6086	NA/NA	Benign
	c.977G>A[b]	R326Q	0	32	4233	0	0	2161	0	32	6394	Yes/NA	Possibly damaging
	c.1246G>A	G416S	0	0	4224	0	3	2111	0	3	6335	NA/NA	Probably damaging
	c.1286C>T	A429V	0	0	4250	0	5	2124	0	5	6374	NA/NA	Possibly damaging
	c.1321G>A	E441K	0	2	4241	0	3	2135	0	5	6376	NA/NA	Possibly damaging
	c.1370C>T	T457M	0	1	4169	0	0	2047	0	1	6216	NA/NA	Probably damaging
	c.1373G>A	R458H	0	1	4160	0	1	2037	0	2	6197	NA/NA	Benign
	c.1468G>A	G490R	0	3	4235	0	1	2125	0	4	6360	Yes/NA	Probably damaging
	c.1504C>T	R502W	0	1	4259	0	0	2159	0	1	6418	Yes/NA	Probably damaging
	c.1519G>A	G507R	0	3	4265	0	14	2155	0	17	6420	NA/NA	Probably damaging
	c.1564G>A	A522T	0	7	4257	0	15	2146	0	22	6403	No/No	Benign
	c.1624G>C	E542Q	0	0	4208	0	1	2098	0	1	6306	Yes/NA	Probably damaging
	c.1813G>A[b]	D605N	0	3	4169	0	1	2006	0	4	6175	NA/NA	Possibly damaging
	c.1855G>A	E619K	0	8	4161	0	0	2027	0	8	6188	Yes/NA	Possibly damaging
	c.2035C>T	P679S	0	0	4151	0	1	1985	0	1	6136	NA/NA	Probably damaging
	c.2269G>A	V757M	0	1	4211	0	0	2098	0	1	6309	NA/NA	Possibly damaging
	c.2311G>A	V771M	0	1	3954	0	1	1946	0	2	5900	Yes/NA	Probably damaging
	c.2320G>A	A774T	0	1	4009	0	0	1981	0	1	5990	NA/NA	Benign
	c.2374T>C	W792R	0	1	4157	0	0	2038	0	1	6195	NA/NA	Probably damaging
	c.2429G>A	R810H	0	1	4243	0	0	2141	0	1	6384	NA/NA	Probably damaging
	c.2497G>A	A833T	0	15	4250	0	0	2165	0	15	6415	Yes/NA	Possibly damaging
	c.2498C>T[b]	A833V	0	2	4263	1	95	2069	1	97	6332	NA/NA	Probably damaging
	c.2618C>A	P873H	0	1	4173	0	0	2007	0	1	6180	NA/NA	Probably damaging
	c.2686G>A	V896M	0	40	4111	0	6	1979	0	46	6090	NA/NA	Benign
	c.2870C>G	T957S	0	11	4131	0	0	1947	0	11	6078	NA/NA	Benign
	c.2873C>T	T958I	0	2	4140	0	0	1951	0	2	6091	NA/NA	Benign
	c.2882C>T	P961L	0	0	4151	0	1	1967	0	1	6118	No/NA	Benign
	c.3049G>A	E1017K	0	1	4245	0	0	2144	0	1	6389	NA/NA	Benign
	c.3137C>T	T1046M	0	1	4265	0	0	2187	0	1	6452	NA/NA	Benign
	c.3392T>C[b]	I1131T	0	4	4149	0	3	2013	0	7	6162	NA/NA	Benign
	c.3641G>A	W1214X	0	1	4121	0	0	1881	0	1	6002	NA/NA	Unknown
	c.3742G>A	G1248R	0	3	4141	0	0	1973	0	3	6114	NA/NA	Probably damaging
MYL2	c.37G>A	A13T	0	6	4294	0	0	2203	0	6	6497	NA/Yes	Possibly damaging
	c.141C>A	N47K	0	5	4295	0	0	2203	0	5	6498	NA/Yes	Benign
	c.401A>C	E134A	0	5	4295	0	0	2203	0	5	6498	NA/NA	Probably damaging
MYL3	—	—	—	—	—	—	—	—	—	—	—	—	—
MYLK2	c.260C>T	A87V	0	1	4278	0	0	2186	0	1	6464	Yes/Yes	Probably damaging
	c.284C>A	A95E	0	1	4288	1	27	2170	1	27	2170	Yes/Yes	Probably damaging
MYO6	—	—	—	—	—	—	—	—	—	—	—	—	—
MYOM1	—	—	—	—	—	—	—	—	—	—	—	—	—
MYOZ2	c.738A>G	I246M	0	2	4298	0	1	2202	0	3	6500	NA/NA	Benign
MYPN	c.59A>G	Y20C	0	10	4290	0	2	2201	0	12	6491	NA/Yes	Probably damaging
	c.3481C>A	L1161I	0	3	4297	0	0	2203	0	3	6500	NA/NA	Probably damaging
NDUFAF1	—	—	—	—	—	—	—	—	—	—	—	—	—
NDUFV2	—	—	—	—	—	—	—	—	—	—	—	—	—
NEXN	—	—	—	—	—	—	—	—	—	—	—	—	—
OBSCN	c.13031G>A	R4344Q	0	14	4172	14	302	1754	14	316	5926	NA/Yes	Possibly damaging
PRKAG2	—	—	—	—	—	—	—	—	—	—	—	—	—
RAF1	—	—	—	—	—	—	—	—	—	—	—	—	—
SLC25A3	—	—	—	—	—	—	—	—	—	—	—	—	—
SLC25A4	—	—	—	—	—	—	—	—	—	—	—	—	—

Table 1 (Continued)

			European Americans—genotype			African Americans—genotype			All—genotype				
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
SRI	c.334T>C	F112L	0	1	4299	0	31	2172	0	32	6471	NA/NA	Benign
TAZ	—	—	—	—	—	—	—	—	—	—	—	—	—
TCAP	c.316C>T	R106C	0	36	4263	0	11	2191	0	47	6454	NA/NA	Probably damaging
	c.458G>A	R153H	0	2	4294	0	0	2198	0	2	6492	Yes/Yes	Benign
TNNC1	—	—	—	—	—	—	—	—	—	—	—	—	—
TNNI3	c.244C>T	P82S	0	1	3801	0	70	1652	0	71	5453	NA/NA	Benign
	c.484C>T	R162W	0	1	4156	0	0	1953	0	1	6109	NA/Yes	Probably damaging
	c.586G>A	D196N	0	1	4157	0	0	1974	0	1	6131	NA/NA	Probably damaging
TNNT2	c.83C>T[b]	A28V	0	4	4296	0	0	2203	0	4	6499	NA/NA	Benign
	c.230C>T	P77L	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
	c.732G>T	E244D	0	0	4300	0	1	2202	0	1	6502	Yes/Yes	Benign
	c.740A>G	K247R	3	121	4176	58	540	1605	61	661	5781	NA/Yes	Benign
	c.832C>T	R278C	0	5	4294	0	1	2202	0	6	6496	Yes/Yes	Probably damaging
	c.857G>A	R286H	0	1	4298	0	0	2201	0	1	6499	NA/NA	Probably damaging
TPM1	c.515T>C	I172T	0	0	4300	0	1	2202	0	1	6502	NA/NA	Benign
TTN	—	—	—	—	—	—	—	—	—	—	—	—	—
VCL	c.829C>A	L277M	0	1	4299	0	0	2203	0	1	6502	NA/Yes	Benign

[a]NA indicates no data available.
[b]Likely disease-causing mutation, but with questionable pathogenicity.

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%) were 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%) were 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 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 (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). ECG's from geno-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

			European Americans—genotype			*African Americans—genotype*			*All—genotype*				
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
ABCC9	—	—	—	—	—	—	—	—	—	—	—	—	—
ACTC1	—	—	—	—	—	—	—	—	—	—	—	—	—
ACTN2	c.26A>G	Q9R	0	7	4293	0	0	2203	0	7	6496	NA/Yes	Benign
	c.2323C>T	H775Y	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
ANKRD1	c.197G>A	R66Q	0	9	4291	0	0	2203	0	9	6494	No/No	Benign
	c.313C>T	P105S	0	4	4295	0	0	2203	0	4	6498	No/Yes	Probably damaging
	c.319G>T	V107L	0	0	4300	0	39	2164	0	39	6464	No/Yes	Benign
	c.827C>T	A276V	0	50	4250	0	4	2199	0	54	6449	No/Yes	Benign
BAG3	—	—	—	—	—	—	—	—	—	—	—	—	—
CHRM2	—	—	—	—	—	—	—	—	—	—	—	—	—
CRYAB	460G>A	G154S	0	9	4288	0	1	2200	0	10	6488	NA/NA	Benign
	c.470G>A	R157H	0	1	4296	0	1	2200	0	2	6496	NA/Yes	Probably damaging
CSRP3	c.10T>C	W4R	0	46	4247	0	2	2197	0	48	6444	NA/Yes	Possibly damaging
	c.148G>A	A50T	0	2	4291	0	1	2198	0	3	6489	NA/NA	Possibly damaging
	c.206A>G	K69R	0	1	4292	0	0	2199	0	1	6491	NA/Yes	Possibly damaging
	c.214G>A[b]	G72R	0	2	4291	0	0	2199	0	2	6490	NA/NA	Probably damaging
CTF1	—	—	—	—	—	—	—	—	—	—	—	—	—
DES	c.893C>T	S298L	0	2	4298	0	1	2202	0	3	6500	NA/Yes	Probably damaging
	c.934G>A	D312N	0	0	4300	0	7	2196	0	7	6496	NA/Yes	Probably damaging
	c.1048C>T	R350W	0	1	4299	0	0	2203	0	1	6502	NA/Yes	Probably damaging
	c.1375G>A	V459I	0	1	4299	2	153	2048	2	154	6347	NA/Yes	Benign
DMD	c.5016T>A	N1672K	0	5	4295	7	244	1951	7	249	6246	NA/NA	Possibly damaging
	c.9682T>C	F3228L	0	0	4300	0	1	2201	0	1	6501	NA/NA	Probably damaging
DNAJC19	—	—	—	—	—	—	—	—	—	—	—	—	—
DOLK	—	—	—	—	—	—	—	—	—	—	—	—	—
DSC2	c.907G>A[b]	V303M	0	2	4298	0	0	2203	0	2	6501	No/NA	Benign
DSG2	c.1003A>G	T335A	0	5	4094	0	0	1863	0	5	5957	No/NA	Probably damaging
DSP	c.5498A>T[b]	E1833V	0	112	4188	0	9	2194	0	121	6382	Yes/NA	Probably damaging
	c.5513G>A	R1838H	0	3	4297	0	0	2203	0	3	6500	No/NA	Probably damaging
	c.6881C>G	A2294G	0	3	4297	0	0	2203	0	3	6500	No/NA	Probably damaging
	c.7123G>C	G2375R	0	1	4299	0	0	2203	0	1	6502	No/NA	Probably damaging
EYA4	—	—	—	—	—	—	—	—	—	—	—	—	—
FKTN	—	—	—	—	—	—	—	—	—	—	—	—	—
FLT1	c.162G>C	R54S	0	5	4295	0	1	2202	0	6	6497	No/NA	Probably damaging
FOXD4	—	—	—	—	—	—	—	—	—	—	—	—	—
GATAD1	—	—	—	—	—	—	—	—	—	—	—	—	—
HSPB7	—	—	—	—	—	—	—	—	—	—	—	—	—
ILK	—	—	—	—	—	—	—	—	—	—	—	—	—
LAMA2	—	—	—	—	—	—	—	—	—	—	—	—	—
LAMA4	c.3217C>T	p.R1073X	0	1	4299	0	0	2203	0	1	6502	No/Yes	Unknown
LDB3	c.349G>A	D117N	0	33	4122	0	51	1994	0	84	6116	NA/Yes	Benign
	c.566C>T	S189L	0	2	4298	0	0	2203	0	2	6501	Yes/Yes	Benign
	c.1049C>T	T350I	0	1	4299	0	0	2203	0	1	6502	NA/NA	Possibly damaging
	c.1051A>G	T351A	0	5	4295	0	0	2203	0	5	6498	NA/NA	Benign
	c.2092G>A[b]	A698T	0	6	4294	0	0	2203	0	6	6497	NA/NA	Probably damaging
LMNA	c.565C>T	R189W	0	1	4299	0	0	2203	0	1	6502	No/NA	Possibly damaging
	c.1303C>T	R435C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
MURC	c.384C>G	N128K	0	0	4299	0	5	2198	0	5	6497	Yes/Yes	Probably damaging
	c.458T>C	L153P	0	0	4300	0	1	2202	0	1	6502	No/Yes	Probably damaging
	c.971C>T[b]	P324L	0	1	4299	0	5	2198	0	6	6497	NA/Yes	Benign
MYBPC3	c.961G>A	V321M	0	4	4246	0	1	2157	0	5	6403	NA/NA	Probably damaging
	c.977G>A[b]	R326Q	0	32	4233	0	0	2161	0	32	6394	NA/NA	Possibly damaging
	c.1814A>G[b]	D605G	0	1	4171	0	0	2008	0	1	6179	NA/NA	Possibly damaging

Table 2 (Continued)

			European Americans—genotype			*African Americans—genotype*			*All—genotype*				
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
MYH6	c.824T>A[b]	I275N	0	3	4297	0	0	2203	0	3	6500	NA/NA	Benign
	c.1702C>T[b]	R568C	0	1	4299	0	1	2202	0	2	6501	NA/NA	Probably damaging
	c.3010G>T	A1004S	0	11	4289	0	1	2202	0	12	6491	No/NA	Benign
	c.4318G>C	A1440P	0	1	4299	0	0	2203	0	1	6502	NA/NA	Possibly damaging
	c.4505G>A[b]	R1502Q	0	3	4297	0	0	2203	0	3	6500	NA/NA	Probably damaging
MYH7	c.2890G>C[b]	V964L	0	6	4294	0	0	2203	0	6	6497	NA/NA	Probably damaging
	c.3286G>T[b]	D1096Y	0	2	4298	0	0	2203	0	2	6501	NA/NA	Probably damaging
	c.4985G>A	R1662H	0	0	4300	0	1	2202	0	1	6502	NA/NA	Benign
	c.5494C>T	R1832C	0	1	4299	0	0	2203	0	1	6502	No/NA	Probably damaging
	c.5588G>A[b]	R1863Q	0	1	4299	0	0	2203	0	1	6502	NA/NA	Possibly damaging
MYPN	c.59A>G	Y20C	0	10	4290	0	2	2201	0	12	6491	NA/Yes	Probably damaging
	c.3335C>T	P1112L	0	21	4279	0	6	2197	0	27	6476	No/Yes	Probably damaging
	c.3583G>A	V1195M	0	0	4300	0	4	2199	0	4	6499	No/Yes	Probably damaging
NEBL	c.180G>C	K60N	0	61	4236	0	6	2196	0	67	6432	NA/Yes	Probably damaging
	c.604G>A	G202R	0	19	4281	0	4	2199	0	23	6480	NA/Yes	Benign
	c.1775C>A	A592E	0	1	4298	0	16	2187	0	17	6485	NA/Yes	Benign
NEXN	c.1955A>G	Y652C	0	1	4086	0	0	1843	0	1	5929	NA/Yes	Probably damaging
PKP2	c.184C>A	Q62K	0	3	4188	0	0	2110	0	3	6298	NA/NA	Benign
	c.419C>T	S140F	0	25	4275	0	1	2202	0	26	6477	NA/NA	Benign
	c.505A>G	S169G	0	16	4281	0	5	2198	0	21	6479	NA/NA	Benign
	c.1759G>A	V587I	0	40	4260	0	2	2201	0	42	6461	No/NA	Possibly damaging
PLN	—	—	—	—	—	—	—	—	—	—	—	—	—
POLG	c.2207A>G	N736S	0	2	4297	0	5	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	1	1590	0	0	692	0	1	2282	Yes/NA	Possibly damaging
SCN5A	c.647C>T[b]	S216L	0	11	4205	0	1	2044	0	12	6249	NA/Yes	Probably damaging
	c.1336G>A	E446K	0	1	4188	0	0	2050	0	1	6238	No/NA	Probably damaging
	c.3835G>A	V1279I	0	1	4299	0	0	2203	0	1	6502	No/NA	Possibly damaging
	c.6013C>G[b]	P2005A	0	13	4169	0	1	2005	0	14	6174	NA/NA	Benign
SGCD	—	—	—	—	—	—	—	—	—	—	—	—	—
SYNE1	—	—	—	—	—	—	—	—	—	—	—	—	—
SYNM	—	—	—	—	—	—	—	—	—	—	—	—	—
TCAP	c.260G>A	R87Q	0	1	4290	0	1	2194	0	2	6484	NA/Yes	Probably damaging
TCF21	c.65G>T[b]	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/Yes	Probably damaging
TNNC1	—	—	—	—	—	—	—	—	—	—	—	—	—
TNNI3	—	—	—	—	—	—	—	—	—	—	—	—	—
TTN	c.96206G>A	R32069Q	0	0	4115	0	3	1920	0	3	6035	NA/NA	Probably damaging
TNNT2	c.83C>T[b]	A28V	0	4	4296	0	0	2203	0	4	6499	NA/NA	Benign
TPM1	—	—	—	—	—	—	—	—	—	—	—	—	—
TXNRD2	—	—	—	—	—	—	—	—	—	—	—	—	—
VCL	c.2923C>T	R975W	0	1	4299	0	0	2203	0	1	6502	Yes/Yes	Probably damaging
	c.3373C>T	R1125C	0	1	4299	0	0	2203	0	1	6502	NA/NA	Benign
VPS13A	c.9403C>T	R3135X	0	0	4293	0	1	2197	0	1	6490	NA/NA	Unknown

[a]NA indicates no data available.
[b]Likely disease-causing mutation, but with questionable pathogenicity.

Table 3 Variants associated with arrhythmogenic right ventricular cardiomyopathy present in the ESP population

			European Americans—genotype			*African Americans—genotype*			*All—genotype*				
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
DES	—	—	—	—	—	—	—	—	—	—	—	—	—
DSC2	c.304G>A	E102K	0	7	4291	0	3	2200	0	10	6491	No/Yes	Benign
	c.327A>G[b]	I109M	0	1	4295	0	1	2201	0	2	6496	Yes/NA	Benign
	c.1018A>G	T340A	0	1	4299	0	0	2202	0	1	6501	NA/NA	Probably damaging
	c.1721G>A[b]	S574N	0	4	4296	0	0	2203	0	4	6499	NA/NA	Benign
	c.2194T>G	L732V	0	11	4289	0	3	2200	0	14	6489	NA/NA	Benign
	c.2471C>T	S824L	0	1	4299	0	2	2201	0	3	6500	No/NA	Probably damaging
	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	5958	NA/NA	Probably damaging
	c.473T>G	V158G	0	65	4059	0	9	1864	0	74	5923	Yes/NA	Benign
	c.581C>T[b]	S194L	0	1	4098	0	0	1858	0	1	5956	NA/NA	Probably damaging
	c.716T>C[b]	V239A	0	1	4140	0	2	1908	0	3	6048	NA/NA	Probably damaging
	c.961T>A	F321I	0	2	4105	0	0	1872	0	2	5977	NA/NA	Probably damaging
	c.1003A>G	T335A	0	5	4094	0	0	1863	0	5	5957	No/NA	Probably damaging
	c.1051A>G[b]	S351G	0	0	4090	0	24	1807	0	24	5897	NA/NA	benign
	c.1174G>A	V392I	0	17	4093	0	2	1873	0	19	5966	No/NA	Benign
	c.1478A>G[b]	N493S	0	1	4150	0	1	1936	0	2	6086	NA/NA	Probably damaging
	c.1550C>T[b]	A517V	0	0	4174	0	21	1940	0	21	6114	NA/No	Probably damaging
	c.1912G>A	G638R	0	2	4132	0	0	1914	0	2	6046	NA/NA	Probably damaging
	c.2137G>A	E713K	16	620	6558	0	72	1929	16	692	6587	NA/NA	Benign
	c.2434G>T	G812C	0	1	4084	0	0	1845	0	1	5929	NA/NA	Probably damaging
	c.2759T>G	V920G	0	41	4094	0	6	1942	0	47	6036	NA/NA	Benign
DSP	c.88G>A	V30M	0	16	4253	0	0	2169	0	16	6422	NA/Yes	Benign
	c.688G>A	D230N	1	3	4296	0	0	2203	1	3	6499	NA/NA	Probably damaging
	c.1696G>A	A566T	0	5	4295	0	0	2203	0	5	6498	NA/NA	Probably damaging
	c.2422C>T	R808C	0	2	4298	0	0	2203	0	2	6501	NA/Yes	Probably damaging
	c.2423G>A	R808H	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
	c.2723G>A[b]	R908H	0	13	4287	0	1	2202	0	14	6489	No/NA	Possibly damaging
	c.2815G>A	G939S	0	1	4298	3	166	2034	3	167	6332	NA/NA	Benign
	c.4372C>G	R1458G	0	18	4282	0	4	2199	0	22	6481	NA/NA	Benign
	c.5324G>T	R1775I	0	1	4299	0	0	2203	0	1	6502	No/NA	Benign
	c.7123G>C	G2375R	0	1	4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
JUP	c.1219G>A	V407I	0	1	4299	0	0	2203	0	1	6502	No/NA	Benign
	c.427G>A[b]	A143T	0	1	4298	0	0	2203	0	1	6501	No/NA	Possibly damaging
PKP2	c.76G>A	D26N	0	56	4044	0	4	2068	0	60	6112	NA/NA	Possibly damaging
	c.184C>A	Q62K	0	3	4188	0	0	2110	0	3	6298	NA/NA	Benign
	c.419C>T	S140F	0	25	4275	0	1	2202	0	26	6477	NA/NA	Benign
	c.505A>G	S169G	0	16	4281	0	5	2198	0	21	6479	NA/NA	Benign
	c.1114G>C[b]	A372P	0	6	4294	0	0	2203	0	6	6497	NA/NA	benign
	c.1237C>T	R413X	0	0	4300	0	1	2202	0	1	6502	No/NA	Unknown
	c.1465G>A	G489R	0	1	4295	0	0	2202	0	1	6497	NA/NA	Benign
	c.1577C>T	T526M	0	20	4280	0	28	2175	0	48	6455	NA/NA	Possibly damaging

Table 3 (Continued)

			European Americans—genotype		African Americans—genotype		All—genotype			
Gene	*Variant*	*Amino acid*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Minor/ major*	*Major/ major*	*Minor/ minor*	*Major/ major*	*Family co-segregation/ functional effect*[a]	*Polyphen-2 prediction*
	c.1592T>G[b]	I531S	0	42	4258	0	0	2203	0	42	6461	NA/No	Benign
	c.1759G>A	V587I	0	40	4260	0	2	2201	0	42	6461	No/NA	Possibly damaging
	c.2062T>C	S688P	0	0	4300	0	2	2201	0	2	6501	NA/NA	Probably damaging
	c.2431C>A	R811S	0	5	4295	0	0	2203	0	5	6498	NA/NA	Benign
	c.2434G>A[b]	D812N	0	0	4300	0	1	2201	0	1	6502	NA/NA	Benign
	c.2615C>T[b]	T872I	0		4299	0	0	2203	0	1	6502	NA/NA	Probably damaging
RYR2	—	—	—	—	—	—	—	—	—	—	—	—	—
TGFβ3	—	—	—	—	—	—	—	—	—	—	—	—	—
TMEM43	c.718C>T[b]	R240C	0	0	4300	0	1	2202	0	1	6502	NA/NA	Probably damaging
TTN	c.26542C>T	H8848Y	0	34	4146	0	3	2011	0	34	6157	NA/NA	Benign
	c.50846T>C	I16949T	0	1	4091	0	0	1839	0	1	5930	NA/NA	Benign
	c.55735G>A	A18579T	0	1	4116	0	0	1917	0	1	6033	NA/NA	Probably damaging
	c.99872T>C	M33291T	0	13	4130	0	0	1929	0	13	6059	NA/NA	Benign

[a]NA indicates no data available.
[b]Likely disease-causing mutation, but with questionable pathogenicity.

Table 4 Cardiomyopathy-associated variants in ESP and in control population

				ESP population		Control population
				Genotype frequency (%)		Genotype frequency (%)
Gene	*Amino acid*	*rs#*	*Disease association*	*African Americans*	*European Americans*	*Northern European*
CSRP3	W4R	rs45550635	DCM	0.09	1.07	1.12
DSG2	V158G	—	ARVC	0.48	1.58	1.69
DSP	V30M	rs121912998	ARVC	0.00	0.37	0.94
MYBPC3	V896M	rs35078470	HCM	0.30	0.96	0.94
MYH6	A1004S	rs143978652	DCM	0.05	0.26	0.37
MYH7	M982T	rs145532615	HCM	0.14	0.44	0.56
PKP2	D26N	rs143004808	ARVC	0.19	1.37	0.94

Abbreviations: ARVC, arrhythmogenic right ventricular cardiomyopathy; DCM, dilated cardiomyopathy; ESP, Exome Sequencing Project; HCM, hypertrophic cardiomyopathy.

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,[20] 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,[21–23] and also 75–95% of ARVC and HCM patients display ECG abnormalities.[24,25] 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 in to 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.[26–28] 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,[29,30] but lack of family co-segregation has also been reported.[31]

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-pathogenic.

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, Fondsbørsvekselerer Henry Hansen og Hustru Karla Hansen, født Westergaards Legat, The Research Foundation at Rigshospitalet, Direktør Ib Henriksens 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**: 81–90.
3 Raju H, Alberg C, Sagoo GS, Burton H, Behr ER: Inherited cardiomyopathies. *BMJ* 2011; **343**: d6966–d6966.
4 Maron BJ, Towbin JA, Thiene G *et al*: Contemporary definitions and classification of the cardiomyopathies. *Circulation* 2006; **113**: 1807–1816.
5 Wexler 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 Marian 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 Kapplinger JD, Landstrom AP, 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://evs.gs.washington.edu/EVS/.
11 A genetic variants database for arrhythmogenic right ventricular dysplasia/cardiomyopathy—van der Zwaag -2009- Human Mutation—Wiley Online 2009.
12 HGMD—Human Gene Mutation Database for Human Genetics Research—BIOBASE Biological Databases. Available at: http://www.biobase-international.com/product/hgmd.
13 Herman DS, Lam L, Taylor MRG *et al*: Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012; **366**: 619–628.
14 Adzhubei IA, Schmidt S, Peshkin L *et al*: A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–249.
15 Olesen MS, Holst AG, Jabbari J *et al*: Genetic loci on chromosomes 4q25, 7p31, and 12p12 are associated with onset of lone atrial fibrillation before the age of 40 years. *Can J Cardiol* 2012; **28**: 191–195.
16 Sajadieh 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.
17 Marcus FI, McKenna WJ, Sherrill D *et al*: Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. *Eur Heart J* 2010; **31**: 806–814.
18 Casale PN, Devereux 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.
19 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.
20 Refsgaard L, Holst AG, Sadjadieh G *et al*: High prevalence of genetic variants previously associated with LQT syndrome in new exome data. *Eur J Hum Genet* 2012; **20**: 905–908.
21 Luk A, Ahn E, Soor GS, Butany J: Dilated cardiomyopathy: a review. *J Clin Pathol* 2009; **62**: 219–225.
22 Basso C, Corrado D, Marcus FI, Nava A, Thiene G: Arrhythmogenic right ventricular cardiomyopathy. *Lancet* 2009; **373**: 1289–1300.
23 Christiaans I, Carrier L: Hypertrophic cardiomyopathy. I: Baars HF, Doevendans PAFM, van der Smagt JJ, red. *Clinical Cardiogenetics*. London: Springer, 2011; 47–61.
24 Maron BJ: Hypertrophic cardiomyopathy: a systematic review. *JAMA* 2002; **287**: 1308–1320.
25 Azaouagh A, Churzidse S, Konorza T, Erbel R: Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a review and update. *Clin Res Cardiol* 2011; **100**: 383–394.
26 Rampersaud E, Kinnamon 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.