(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 August 2012 (16.08.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/107580 A1

(51) International Patent Classification:
G01N 33/68 (2006.01)

(21) International Application Number:
PCT/EP2012/052352

(22) International Filing Date:
10 February 2012 (10.02.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/441,325 10 February 2011 (10.02.2011) US
11154003.5 10 February 2011 (10.02.2011) EP

(71) Applicant *(for all designated States except US)*: **INSERM (INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE)** [FR/FR]; 101 rue de Tolbiac, F-75013 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHARRON, Philippe** [FR/FR]; Unité INSERM 956, Gh Pitie Salpetriere, 91 Boulevard de l'Hôpital, Faculté de Médecine, F-75634 Paris Cedex 13 (FR). **VILLARD, Eric** [FR/FR]; Unité INSERM 956, Gh Pitie Salpetriere, 91 Boulevard de l'Hôpital, Faculté de Médecine, F-75634 Paris Cedex 13 (FR). **CAMBIEN, François** [FR/FR]; Unité INSERM 956, Gh Pitie Salpetriere, 91 Boulevard de l'Hôpital, Faculté de Médecine, F-75634 Paris Cedex 13 (FR).

(74) Agent: **BOURGAREL, Denis**; Cabinet Plasseraud, Group of Mandatories No. 280, 235 cours Lafayette, F-69006 Lyon (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2012/107580 A1

(54) **Title**: IN VITRO DIAGNOSIS METHOD FOR PREDICTING A PREDISPOSITION TO CARDIOMYOPATHY

(57) **Abstract**: The invention relates to field of molecular diagnosis, in particular for the in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy, such as dilated cardiomyopathy in a subject. More specifically, the invention relates to an in vitro diagnosis method for detecting a genetic predisposition for cardiomyopathy in a human subject, said method comprising the steps of: a. providing a biological sample from said human subject; and, b. detecting the presence or absence of an HSPB7 disease associated variant and/or BAG3 disease associated variant in said biological sample, wherein the presence of said HSPB7 or BAG3 disease associated variant is indicative of a genetic predisposition to cardiomyopathy in said human subject.

IN VITRO DIAGNOSIS METHOD FOR PREDICTING A PREDISPOSITION TO CARDIOMYOPATHY

FIELD OF THE INVENTION:

The invention relates to the field of molecular diagnosis. The invention provides in vitro diagnosis methods for detecting a genetic predisposition to cardiomyopathy, such as dilated cardiomyopathy. More specifically, the invention relates to an in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a human subject, said method comprising the steps of:

a. providing a biological sample from a human subject; and,

b. detecting the presence or absence of an HSPB7 disease associated variant and/or BAG3 disease associated variant in said biological sample,

wherein the presence of said HSPB7 and/or BAG3 disease associated variant is indicative of a genetic predisposition for cardiomyopathy in said human subject.

BACKGROUND OF THE INVENTION:

Heart failure (HF) is a common chronic condition associated with a poor prognosis and an increasing economic burden[1,2]. The underlying causes of HF are heterogeneous. Idiopathic dilated cardiomyopathy (DCM) is a form of HF defined by the presence of left ventricular dilatation and left ventricular systolic dysfunction in the absence of an obvious etiology such as coronary artery disease (CAD), hypertension, valvular disease, or congenital defect[3,4,5]. DCM is a major cause of systolic heart failure and the leading indication for heart transplantation[6], it affects approximately 1/2500 adults and it is more common in men than in women[7].

The pathophysiology of DCM is poorly understood[3]. The disease is considered to be multifactorial with a possible implication of environmental factors and the existence of a strong genetic component attested by a high rate of familial aggregation[3]; 20% to 35% of DCM cases having an affected first-degree relative[3,4]. Genetic analyses of monogenic DCM have identified mutations in more than 30 genes, most of them encoding proteins of the cytoskeleton or the sarcomere[8]. These genes may carry mutations that are implicated in familial forms of the disease cases as well as common susceptibility alleles that are over-represented in sporadic cases[9]. However, the genetic basis of DCM, whether familial or

sporadic, is still largely unresolved and results of candidate gene association studies have been inconsistent[10].

Over the last few years, genome-wide association studies (GWASs) using high density genotyping arrays have discovered numerous new loci implicated in cardiovascular diseases[11,12,13]. Despite these successes no GWAS of DCM has been reported so far, probably as a consequence of the relatively low prevalence of the disease making the assembly of large clinically homogeneous cohorts of patients difficult.

Therefore, there is still need for providing simple and sensitive diagnostic method for predicting DCM predisposition.

In a first genome-wide association study of heart failure due to dilated cardiomyopathy (DCM) with discovery and replication cohorts comprising overall 2344 cases and 2410 controls, the inventors identified DCM biomarkers and in particular two DCM-associated SNPs, rs10927875 and rs2234962 with respective *P*-values of 9.5×10^{-10} and 4.0×10^{-12} in the combined data set. The first SNP is located at a locus on 1p36.13 which exhibit a yin/yang haplotype structure encompassing several genes including *HSPB7*. The second SNP on 10q26.11 is located within *BAG3* and is non-synonymous. Surprisingly, by sequencing of *BAG3* exons in patients with familial DCM, the inventors identified several damaging mutations which were absent in healthy individuals, suggesting that they are causal for DCM.

SUMMARY OF THE INVENTION:

A first object of the invention relates to an in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject, said method comprising the steps of:

a. providing a biological sample from said subject; and,

b. detecting the presence or absence of an HSPB7 disease associated variant and/or BAG3 disease associated variant in said biological sample,

wherein the presence of said HSPB7 and/or BAG3 disease associated variant is indicative of a predisposition to cardiomyopathy in said subject.

The method is particularly useful in predicting a predisposition to dilated cardiomyopathy, including familial or idiopathic dilated cardiomyopathy.

In one specific embodiment, said HSPB7 disease associated variant is a genetic mutation that decreases the expression level of HSPB7 gene without resulting in an alteration of the predicted amino acid sequence encoded by the HSPB7 gene.

Said genetic mutation may be for example, one or more single nucleotide polymorphism (SNP) mutations in HSPB7 gene locus. In a preferred embodiment, said genetic mutation is a SNP rs10927875 and/or rs945417.

In another specific embodiment, said BAG3 disease associated variant is a genetic mutation altering the predicted amino acid sequence encoded by the BAG3 gene. For example, said genetic mutation includes a non-synonymous single nucleotide polymorphism (SNP) in a BAG3 coding region. In one embodiment, said genetic mutation is a genetic mutation resulting in a deletion or a dysfunction of the BAG domain of the Bag3 protein. For example, in a related specific embodiment, said genetic mutation is a non-synonymous single nucleotide polymorphisme in the BAG domain of the Bag3 protein. In a preferred embodiment, said non-synonymous SNP in BAG3 coding region is rs2234962 or rs3858340.

The invention naturally further relates to a kit for carrying out the above-described method, said kit comprising:

a. means for detecting one or more SNPs in BAG3 and/or HSPB7 gene loci, and

b. optionally, instructions for use of the kit.

The invention also relates to a kit for carrying out the above-described method, said kit comprising:

a. means for detecting causal mutations in the full gene sequence of BAG3; and,

b. optionally, instructions for use of the kit.

DETAILED DESCRIPTION OF THE INVENTION:

The invention relates to an in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject, said method comprising the steps of:

a. providing a biological sample from said subject; and,

b. detecting the presence or absence of an HSPB7 disease associated variant and/or BAG3 disease associated variant in said biological sample,

wherein the presence of said HSPB7 and/or BAG3 disease associated variant is indicative of a genetic predisposition to cardiomyopathy in said subject.

As used herein, a subject has a genetic predisposition to a disease when this subject has a higher risk to develop such disease, compared to the average risk in a population to develop such disease. Of course, a predisposition does not mean that the subject will develop the disease. As used herein, "Detecting a predisposition to cardiomyopathy" therefore includes detecting a higher risk of developing the disease, or determining the susceptibility of that subject to developing the disease or to having a poor prognosis for the disease.

As used herein, the term "cardiomyopathy" refers to all myocardial disorder in which the heart muscle is structurally and functionally abnormal, in the absence of coronary artery disease, hypertension, valvular disease and congenital heart disease sufficient to cause the observed myocardial abnormality (Eur Heart J 2008;29:270-276). Cardiomyopathy includes without limitation, hypertrophic cardiomyopathy (HCM or HOCM), arrhythmogenic right ventricular cardiomyopathy (ARVC), isolated ventricular non-compaction, mitochondrial myopathy, dilated cardiomyopathy such as familial or idiopathic dilated cardiomyopathy (DCM), restrictive cardiomyopathy (RCM), Takotsubo cardiomyopathy and Loeffler endocarditis.

The diagnosis methods may also be applied to other cause of heart failure such as systolic or diastolic left ventricular dysfunction due to coronary artery disease, valvular diseases or arterial.

The invention more specifically relates to a method for detecting a genetic predisposition to dilated cardiomyopathy including familial or idiopathic dilated cardiomyopathy.

Providing a biological sample

The method of the invention can be carried out on any appropriate biological sample obtained from a subject.

As used herein, the term "biological sample" refers to a sample that contains either nucleic acid or protein materials reflecting the genomic information of cells, tissue or organs of the subject.

In one specific embodiment, said sample is obtained from a mammal, for example from rodents, cats, dogs, horses, primates or human. In one preferred embodiment, said sample is obtained from a human subject. For example, said biological sample may be obtained from urine, blood including without limitation peripheral blood or plasma, stool, sputum, bronchoalveolar fluid, endotracheal aspirates, wounds, cerebrospinal fluid, lymph node, exsudate and more generally any human biopsy tissue or body fluids, tissues or materials.

It is well known that genomic DNA of individuals can easily be purified from individual blood sample. Therefore, in a preferred embodiment, said biological sample is blood, more preferably human blood sample.

As used herein, a "disease associated variant" means any genotypic biomarker, such as a genetic mutation, that is associated with an increased or decreased risk of developing the disease.

In one embodiment, said disease associated variant is a genetic mutation in HPSB7 or BAG3 gene locus.

A "genetic mutation" refers to a nucleotide change (or nucleotide changes) in the wild type sequence of the corresponding gene. Said genetic mutation refers to a germline mutation that can be considered as a causal mutation (present only in patients, usually in a familial form of the disease, with a direct causal link with the disease). For example, mutations can occur within a gene or chromosome, including specific changes in non-coding regions of a chromosome, for instance changes in or near regulatory regions of genes. Types of mutations include, but are not limited to, base substitution point mutations (which are either transitions or transversions), deletions, and insertions. Missense mutations are those that introduce a different amino acid into the sequence of the encoded protein; nonsense mutations are those that introduce a new stop codon; and silent mutations are those that introduce the same amino acid often with a base change in the codon. In the case of insertions or deletions, mutations can be in-frame (not changing the frame of the overall sequence) or frame shift mutations, which may result in the misreading of a large number of codons (and often leads to abnormal termination of the encoded product due to the presence of a stop codon in the alternative frame).

In preferred embodiments of the diagnosis methods of the invention, a disease associated variant is a single nucleotide polymorphism (SNP) in HSPB7 or BAG3 gene locus.

As used herein, a "single nucleotide polymorphism (SNP)" is a single base (nucleotide) polymorphism in a DNA sequence among individuals in a population. Typically in the literature, a single nucleotide polymorphism (SNP) may fall within coding sequences of genes, non-coding regions of genes, or in the intergenic regions between genes. SNPs within a coding sequence will not necessarily change the amino acid sequence of the protein that is produced, due to degeneracy of the genetic code. A SNP in which both forms lead to the same polypeptide sequence is termed "synonymous" (sometimes called a silent mutation)-if a different polypeptide sequence is produced they are "nonsynonymous". A nonsynonymous change may either be missense or "nonsense", where a missense change results in a different amino acid, while a nonsense change results in a premature stop codon. The exact sequence of a SNP can be determined from the database of SNPs available at the NCBI website (Entrez SNP, dbSNP build 128, Jan. 28, 2009). The "position" of the nucleotide of interest gives the location in the genome of the SNP, referring to the nucleotide position from the p-terminus of the chromosome in the human genome, see the NCBI SNP website (dbSNP), available on the internet.

HSPB7 disease associated variant

The inventors have identified specific single nucleotide polymorphisms of HSPB7 that are associated to a higher risk of developing dilated cardiomyopathy. The inventors have further shown that such SNPs may be associated to a decreased level of expression of HSPB7 gene compared to expression level with a wild type sequence.

Therefore, in one embodiment, predisposition to cardiomyopathy may be diagnosed by detecting a genetic mutation in HSPB7 gene locus that decreases the expression level of HSPB7 gene in a biological sample obtained from a subject.

An example of human wild type HSPB7 gene is the gene comprising the nucleotide sequence as shown in Genbank accession number (NM_014424.4 GI:164519093 : with corresponding cDNA of SEQ ID NO:1), and coding for the polypeptide sequence of SEQ ID NO:2 (NP_055239.1 GI:7657202).

In one specific embodiment, said genetic mutation is a single nucleotide polymorphism (SNP) mutation in HSPB7 gene locus, which SNP in HSPB7 gene locus is associated with a decreased expression level of HSPB7 gene.

As used herein “HSPB7 gene locus” may comprise in the context of the present invention, the genetic sequence of HSPB7, the region of the promoter, the introns, the exons

including 3' and 5' untranscribed regions, and some intergenic regions involved in HSPB7 gene expression. The genetically associated region discovered by the inventors is covering 5 genes in the locus and extended from SPEN to CLCNKB (on chromosome 1p36.2-p36.1 interval 162670000-163700000; GRCh37/hg19 assembly) which comprise several other genes: SPEN (spen homolog, transcriptional regulator), HSPB7 (heat shock 27 kDa protein family, member 7), CLCNKA (chloride channel Ka) and CLCNKB (chloride channel kb).

In one preferred embodiment, said genetic mutation is the SNP rs10927875 or rs945417 as described for example in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi).

More specifically, rs10927875 is a SNP located in an intron of ZBTB17 (also referred as MIZ-1) (position 16299312; genomic release GRCh37). Rs945417 is a SNP located at position 16344625 (genomic release GRCh37) and located in the proximal promoter region of HSPB7 gene sequence.

Other genetic mutation or SNPs in HSPB7 gene locus may be detected in addition to rs10927875 and/or rs945417, in particular, other SNPs which are associated with a decreased expression of HSPB7. SNPs in HSPB7 have been described for example in Matkovitch et al.[21], to be associated to systolic heart failure.

BAG3 disease associated variant

The inventors have further identified specific SNPs of BAG3 gene that are associated to a higher risk of developing dilated cardiomyopathy. Such SNPs includes non-synonymous mutations in the coding regions and have been shown to be associated to an increased risk of developing cardiomyopathy. In some cases, the increased risk may therefore be linked to an altered Bag3 protein sequence. Sequencing of BAG3 exons in patients with familial dilated cardiomyopathy identified several damaging mutations absent in healthy individuals, suggesting a causal link of the mutations for dilated cardiomyopathy.

In one preferred embodiment, said genetic mutation is one or more single nucleotide polymorphism (SNP) mutations in BAG3 gene locus, for example, non synonymous single nucleotide polymorphism in BAG3 coding regions.

In one specific embodiment, genetic predisposition to cardiomyopathy may be diagnosed by detecting a genetic mutation that alters the predicted amino acid sequence encoded by the BAG3 gene in the biological sample.

As used herein "BAG3 gene locus" may comprise in the context of the present invention, the genetic sequence of BAG3, the region of the promoter, the introns, the exons including 3' and 5' untranscribed regions, and some intergenic regions involved in HSPB7 gene expression.. The BAG3 gene is located on chromosome 10q (position chr10:121400000-121440000; GRCh37/hg19 assembly).

An example of human wild type BAG3 gene is the gene comprising the nucleotide sequence as shown in Genbank accession number (NM_004281.3 GI:62530382; with corresponding cDNA of SEQ ID NO:3), and coding for the polypeptide sequence of SEQ ID NO:4 (NP_004272.2 GI:14043024).

In one specific embodiment, said genetic mutation is the single nucleotide polymorphism rs2234962 or rs3858340 as described for example in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi).

The SNP rs2234962 is a non-synonymous SNP (c. T757C, p. C151R) located within the coding sequence of BAG3 on chromosome 10q25.2-q26.2 (position 121429633 of GRCh37 Assembly).

The SNP rs3858340 is a non-synonymous SNP (c. C1526T, p. P407L) located within the coding sequence of BAG3 on chromosome 10q25.2-q26.2 (position 121436286 of GRCh37 Assembly).

Both SNPs were unambiguously associated with dilated cardiomyopathy as shown in the experimental part below.

Other genetic mutation or SNPs in BAG3 gene locus may be detected in addition to rs2234962 or rs3858340.

In other specific embodiments, said genetic mutation is one or more of the following SNPs set forth in Table 1, which have been identified in familial forms of dilated cardiomyopathy.

In particular, the following SNPs in BAG3 coding regions set forth in Table 1 may be detected in the diagnosis methods according to the invention, either as disease associated variants or in combination with detecting SNPs rs2234962 or rs3858340:

Table 1:

Exon	genomic position	variant name	protein effect	nature of variant	dbSNP reference
2	121429462	586A>T	I94F	missense	
2	121429633	757T>C	C151R	missense	rs2234962
3	121432011	1058delA	251RfsX56	frame shift	
4	121435991	1231C>T	R309X	nonsense	
4	121436219	1459_1466 delTCTTCCC	385QfsX56	frame shift	
4	121436246	1486_1487 delGA	395GfsX48	frame shift	
4	121436286	1526C>T	P407L	missense	rs3858340
4	121436429	1669G>A	E455K	missense	
4	121436468	1708G>A	V468M	missense	

In one specific embodiment, said BAG3 disease associated variant is a variant that is not associated to muscular dystrophy.

In another specific embodiment, said BAG3 disease associated variant is a Bag3 mutation resulting in a deletion or a dysfunction in the BAG domain of the Bag3 protein, for example one or more amino acid substitution or deletion in the BAG domain of Bag3 protein. As used herein, the term "BAG domain" may be defined by the following amino acid sequence of SEQ ID NO:75, corresponding to residues at amino acid position 421 to 498 of the Bag3 amino acid sequence of SEQ ID NO:4.

Means for detecting disease associated variants

The above-mentioned disease associated variants for HSPB7 or BAG3 may be detected in the biological sample either at the nucleic acid level or at the polypeptide level and includes methods to detect a genetic mutation in the genomic DNA or RNA transcripts and method to detect abnormal expression of the gene product or mutation in the gene product (RNA transcripts or polypeptide).

The presence of the genetic mutation may be detected on either one or both chromosomes, wherein the identification of the genetic mutation (e.g. one or more specific SNPs as described above) on at least one chromosome indicates that the subject is at risk for

developing cardiomyopathy, the subject has a cardiomyopathy, or has an early stage of cardiomyopathy.

A variety of techniques are known in the art for detecting a genetic mutation within a sample, including sequencing genotyping, microarrays, Restriction Fragment Length Polymorphism, Southern Blots, SSCP, dHPLC, single nucleotide primer extension, allele-specific hybridization, allele-specific primer extension, oligonucleotide ligation assay, and invasive signal amplification, MALDI-TOF mass spectrometry and fluorescence polarization (FP).

For example, specific genetic mutations of BAG3 or HSPB7 gene locus associated to dilated cardiomyopathy as described above may be readily detected on genomic DNA or transcript RNA or cDNA obtained from the biological sample.

For example, detecting a genetic mutation may be carried out by sequencing a nucleic acid comprising a fragment of BAG3 or HSPB7 gene locus and analysing the sequence for detecting the presence or the absence of said genetic mutation.

For sequencing the appropriate regions in BAG3 or HSPB7 gene locus, PCR amplification may be performed on genomic DNA from said biological sample allowing amplification of the fragments to be sequenced.

Thus, primers which span one or more fragments that comprise the putative location of a BAG3 or HSPB7 genetic mutation (e.g specific SNPs as described in the above paragraphs) may be used to detect, by sequencing, said BAG3 or HSPB7 genetic mutations.

Such primers which may be used in the diagnosis methods of the invention are short nucleic acid molecules, for instance DNA oligonucleotides of 10 nucleotides or more in length, which can be annealed to the complementary target nucleic acid molecule by nucleic acid hybridization to form a hybrid between the primer and the target nucleic acid strand. A primer can be extended along the target nucleic acid molecule by a polymerase enzyme. Therefore, primers can be used to amplify the target nucleic acid molecule, such as fragments of BAG3 coding regions or HSPB7 gene locus. The specificity of a primer increases with its length. Thus, for example, a primer that includes 30 consecutive nucleotides will anneal to a target sequence with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, to obtain greater specificity, probes and primers can be selected that include at least 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 or more

consecutive nucleotides. In particular examples, a primer is at least 15 nucleotides in length, such as at least 15 contiguous nucleotides complementary to a target nucleic acid molecule. Particular lengths of primers that can be used to practice the methods of the present disclosure include primers having at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 70, or more contiguous nucleotides complementary to the target nucleic acid molecule to be amplified, such as a primer of 15-70 nucleotides, 15-60 nucleotides, 15-50 nucleotides, or 15-30 nucleotides.

An "upstream" or "forward" primer is a primer 5' to a reference point on a nucleic acid sequence. A "downstream" or "reverse" primer is a primer 3' to a reference point on a nucleic acid sequence. In general, at least one forward and one reverse primer are included in an amplification reaction.

Nucleic acid probes and primers can be readily prepared based on the nucleic acid sequence of HSPB7 or BAG3 gene locus, for example SEQ ID NO:1 and SEQ ID NO:3.

PCR primer pairs can be derived from a known sequence by using computer programs intended for that purpose such as Primer 3 (v. 0.4.0 Whitehead Institute for Biomedical Research, Steve Rozen, and Helen Skaletsky).

Examples of primers that may be used for amplifying BAG3 and/or HSPB7 exons prior to sequencing are shown in the following table 2. Typical PCR amplification conditions that may be used are: 95°C 5 min followed by 35 cycles :[95°C-15sec / Tm (as shown in Table2) – 30sec / Elongation Temp – 30sec].

Table 2: Examples of PCR primers and conditions for sequence of BAG3 and HSPB7

BAG3 exon	primers (5'>3')	5' position (ref hg19)	PCR: size;Tm;MgCl2;DMSO
1	cgcgattatagccgatgact	121410858	562bp;62°C;2mM;10%
	gcctgccgtcgaggt	121411420	
2	aggagggttcacttcccagt	121429306	499bp;65°C;2mM;none
	atgccctgcatgtgaacag	121429805	
3	ggagtcatttgtggggtcat	121431722	497bp;55°C;2mM;none
	ccctggagacataccaccat	121432219	
4.1	caatttctgtgactttcagtcagtt	121435918	495bp;65°C;2mM;none

	tttgtcagtcttcttgccttca	121436413	
4.2	gccatcctggagaaggtaca	121436345	582bp;55°C;1mM;10%
	tttgcctccacccaagttac	121436927	
HSPB7 exon			
1	cagggacagtcggccttat	16344582	376bp;TD65-55°C;2mM;10%
	cagacgtcccctccctgt	16344207	
2	atgcaaggcctgctccat	16343748	271bp;TD65-55°C;2mM;10%
	gtctggggtcccaggatg	16343472	
3	ggggttagaatggggagaag	16342343	374bp;60°C;2mM;none
	ggctagaacctgggctgag	16341970	

In other embodiments, the genetic mutations are detected by specific hybridization of nucleic acid probes, such as oligonucleotide probes to genomic DNA or RNA transcripts or corresponding cDNA, containing the BAG3 or HSPB7 genetic mutations, for example containing the specific BAG3 or HSPB7 SNPs as described in the above paragraphs.

In a specific embodiment, Taqman 5' nuclease genotyping method is used as described in the Examples. In this method, oligonucleotide PCR primers are designed that flank the mutation in question and allow PCR amplification of the region. A third oligonucleotide probe is then designed to hybridize to the region containing the base subject to change between different alleles of the gene. This probe is labelled with fluorescent dyes at both the 5' and 3' ends. These dyes are chosen such that while in this proximity to each other the fluorescence of one of them is quenched by the other and cannot be detected. Extension by Taq DNA polymerase from the PCR primer positioned 5' on the template relative to the probe leads to the cleavage of the dye attached to the 5' end of the annealed probe through the 5' nuclease activity of the Taq DNA polymerase. This removes the quenching effect allowing detection of the fluorescence from the dye at the 3' end of the probe. The discrimination between different DNA sequences arises through the fact that if the hybridization of the probe to the template molecule is not complete (there is a mismatch of some form) the cleavage of the dye does not take place. Thus only if the nucleotide sequence of the oligonucleotide probe is completely complimentary to the template molecule to which it is bound will quenching be removed. A reaction mix can contain two different probe sequences each designed against different alleles that might be present thus allowing the detection of both alleles in one reaction.

Examples of primers and probes useful for detecting the SNPs of BAG and/or HSPB7 gene associated to dilated cardiomyopathy are shown in the following Table 3:

Table 3: Primers (a) and probes (b) and reaction conditions for Taqman genotyping

(a)

SNP	F	Forward Primer sequence (5'-3')	WT
	R	Reverse Primer sequence (5'-3')	M
rs16983785	F	TGA CCC CCA GGA TTT TTC TGT	T
	R	GTT TTT CAC TGC TAT TCT TGC ACA A	G
rs2234962	F	TCC CAG TCA CCT CTG CGG	T
rs2234962	R	CTC TCC TTA CCT CAG GTC CGT G	C
rs10927875	F	GTT TGG TTG GGA AAT TGT TGA TCT	C
rs10927875	R	TGG ACT GTG TAC TTC CAG AAT GTT AAG	T
rs945417	F	TGG AAT GTC AGG CTG TGA GCT	G
rs945417	R	ACT GTC CCT GGG CTT GGG	C
rs3858340	F	CAC TTT CAG CAC TCC TGG ATG TT	C
rs3858340	R	GGC TAC AGA AGA GAG GGC AGC	T

(b)

SNP	Probe sequence (5'Fam – 3' Tamra)	Annealing Temp
	Probe sequence (5'Tet – 3'Tamra)	
rs16983785	ACA AGT TTG AGC ATG TAA TTT GCA GTG CTT C	69 °C
rs16983785	CAA GTT TGA GCA TGT AAT TTG CAG GGC TT	
rs2234962	CCA GAA ACC ACT CAG CCA GAT AAA CAG TGT	67 °C
rs2234962	ACC ACT CAG CCA GAT AAA CAG CGT GG	
rs10927875	CAC TGG GCA ATG CTC ACG CCT T	67 °C
rs10927875	TGA TCA CTG GGC AAT GCT CAC ACC	
rs945417	TCT GGG CCA TGG ACA GGC CT	66 °C *
rs945417	CTC TGG GCC ATG GAG AGG CCT	
rs3858340	TCG GCT TCT CCT GGT TTT GGA GGT	67 °C
rs3858340	TCG GCT TCT CCT GGT TTT GGA AGT G	

In another embodiment, the method of detecting a genetic mutation in BAG3 or HSPB7 gene locus comprises use of a restriction enzyme that specifically recognizes sequence corresponding to wild type or genetic mutations of BAG3 or HSPB7 gene sequence.

Alternatively, said disease associated variant may be detected by detecting abnormal expression of gene product such as mRNA transcripts. Methods to quantify mRNA level from total mRNA cell extract are well known in the art and include quantitative PCR such as Real Time PCR amplification with labelled probes. Abnormal expression level may be determined by comparing the expression level to a control level. Significant difference in the expression level as determined by statistical analysis is indicative of abnormal expression level.

In another embodiment, a polypeptide disease associated variant is detected using a binding agent that specifically binds to a polypeptide disease associated variant. For example, an agent that specifically binds to a mutant variant of Bag3 corresponding to said disease associated variant but not to wild type Bag3 may be used. Said binding agent may be for example an antibody or antibody fragments comprising antigen-binding regions. In a preferred embodiment, a mutant variant of Bag3 is a mutant Bag3 polypeptide as encoded by a BAG3 coding sequence including one or more of the SNPs associated to cardiomyopathy as described above, such as rs2234962 and rs3858340 or one of the SNP described in Table 1.

Specific antibodies, or antibody fragments, reactive against particular disease associated variant, for example, a particular Bag3 mutant polypeptide may be selected by screening expression libraries encoding immunoglobulin genes using phage display technologies. In one embodiment, antibodies are used to detect mutant Bag3 protein, in particular the mutant Bag3 protein as encoded by a BAG3 coding sequence including one or more of the SNPs associated to cardiomyopathy as described above, such as rs2234962 and rs3858340 or one of the SNP described in Table 1.

Said antibodies or fragments thereof bind to mutant Bag3 protein but not to wild type Bag3 protein.

Alternatively, said antibodies or fragments thereof bind to wild type Bag3 protein but not to mutant Bag3 protein.

A person skilled in the art will understand that a number of methods can be used to detect and/or quantify specific polypeptide, including immunoassays such as Western Blots, ELISA, and immunoprecipitation followed by SDS-PAGE, as well as immunocytochemistry or immunohistochemistry.

Kits may be prepared for carrying out one of the above mentioned detection methods. Thus, the invention further relates to a kit for carrying out the above-described method, said kit comprising:

a. means for detecting a genetic mutation in BAG3 and/or HSPB7 gene locus, and,

b. optionally, instructions for use of the kit.

Said means for detecting a genetic mutation in BAG3 and/or HSPB7 gene locus may therefore comprises, specific primers or oligonucleotides probes as described above, or specific binding agents, such antibody or antibody fragments as described above.

The kits can include one or more isolated primers or primer pairs for amplifying a target nucleic acid in BAG3 or HSPB7 gene locus, such as a region comprising a SNP associated to cardiomyopathy as described above. For example, the kit can include primers for amplifying a haplotype including one, two, three, four, five SNPs in HSPB7 and/or BAG3, wherein the amplified sequence includes the SNP associated with cardiomyopathy.

The kit can further include one or more of a buffer solution, a conjugating solution for developing the signal of interest, or a detection reagent for detecting the signal of interest, each in separate packaging, such as a container.

In another example, the kit includes a plurality of size-associated marker target nucleic acid sequences for hybridization with a detection array. The target nucleic acid sequences can include oligonucleotides such as DNA, RNA, and peptide-nucleic acid, or can include PCR fragments.

In another example, the kit includes binding reagent to disease associated polypeptide variant, such antibodies or fragment thereof.

The kit can also include instructions in a tangible form, such as written instructions or in a computer-readable format.

In one preferred embodiment, said kit comprises

a. means for detecting one or more SNPs in BAG3 and/or HSPB7 gene locus selected from the group consisting rs10927875, rs945417, rs2234962 and rs3858340; and,

b. optionally, instructions for use of the kit.

In another specific embodiment, said kit comprise

a. means for detecting causal mutations in the full gene sequence of BAG3; and,

b. optionally, instructions for use of the kit.

The invention will now be further illustrated by the following figures and examples. However, these examples and figures should not be interpreted in any way as limiting the scope of the present invention.

Brief Description of the Figures

Figure 1: displays the association of these 14 SNPs with DCM in both the pooled and individual DNA analyses.

Figure 2: Ideogram from UCSC Genome Browser (assembly GRCh37/hg19) presenting the associated region on chromosome 1 and gene structure and 5'->3' orientation. Positions of each of the five SNPs in LD defining DCM associated haplotypes are indicated.

EXAMPLE:

1. Methods

Subjects. All participants included in the study were of European origin. Ethics committees approved the study protocols, and all participants gave written informed consent. All patients had a diagnosis of idiopathic dilated cardiomyopathy (DCM) according to conventional criteria[14], namely enlarged left ventricle diameter and low ejection fraction (≤45%) associated with the absence of causal factors, such as coronary artery disease. Only apparently sporadic cases without affected first degree relatives were included.

For the genome wide association study (GWAS), DCM cases were recruited through the *CARDIGENE* study (424 French patients)[42], the *EUROGENE* (EHF) study (463 DCM cases from Germany, Italy and France) and the *PHRC-DCM* study (292 French DCM cases). French controls were selected from the ECTIM Study (Etude Cas-Témoin sur l'Infarctus du Myocarde)[43] and the FITENAT Study[44]. Controls were also selected from healthy consultants or hospital professional workers in clinical centres in Italy (72 controls) and Germany (278 controls). Selection of controls was stratified according to age, gender

and geographic origin to match the distribution of DCM cases. Replications studies were conducted in two case-control studies, from Germany, 723 cases and 726 controls, and from United Kingdom, 442 cases and 576 controls.

Constitution of DNA pools. The DNA pools were stratified on population as well as gender and age when the numbers permitted. We required that at least 25 samples were mixed in a single pool. Each pool was constituted twice and each pool replicate was analysed on two arrays independently.

Quantification of allelic signals in DNA pools. The Human 610 quad *beadchip* was used to quantify allelic signals and the beadchips were scanned on an *i-scan* scanner (*Illumina*). The data delivered by the scanner were transferred to the *BeadStudio* software (*Illumina*). "Genotyping" of pools of DNA followed the same general protocol as individual genotyping, but after normalization in *BeadStudio* the genotype assignment usually performed by clustering was skipped. Allele frequencies within pools were estimated from the normalized quantitative signals (x and y) corresponding to the two alleles of each SNP: $p(x) = x/(x+y)$ and $p(y)=1-p(x)$. As four measurements (two replicates of pools and two arrays /replicate of pool) were done, four estimates of x and y were available for each SNP in each pool and were used for statistical modelling. Statistical analyses were performed using the R statistical software[45]. A mixed linear model (R: LME4) was used to analyse the pooled allelic signals of each SNP, in which p(x) was the dependent variable, disease status, age category, gender and study population were fixed effects and the pool identifier was a random effect. In addition, a weight was introduced in the model to account for the number of samples in each pool. To estimate allele frequencies from DNA pools It has been proposed to introduce a correction factor k which is specific of each SNP, to compensate for systematic differences in signal intensities ($p(x)=x/(x+ky)$)[46]. The coefficient k may be obtained from the ratio of allelic signals in heterozygotes by genotyping a sample of individuals using the same array as for the pools-GWAS. However, k may be difficult to estimate reliably[47]. In the present analysis, we were not interested in allele frequencies *per se,* but in association statistics (odds ratio), which are not affected by k, under the assumption that the difference in signal intensity is not related to disease status. As a consequence, despite some loss of precision due to a larger standard error of the association statistic, we did not include a k correction.

The association between DCM and individual genotypes was tested using a logistic regression model (R:GLM) assuming an additive allele effect and adjusted on age, gender and study population.

For in-depth investigation of the genomic context of DCM-associated SNPs, exploration and annotations of loci of interest were carried out using WGAviewer post-association genomic annotation tool[48]. To plot regional association results *LocusZoom*[49] was used.

Individual level validation of SNPs associated in the pools-GWAS. For each SNP needing validation at the individual level, 2 primers and 2 dual-labelled probes (5'FAM-3'TAMRA or 5'TET-3'TAMRA) were designed with the Primer Express software (Applied Biosystems). Twenty ng of each DNA were amplified on a 96-well *GeneAmp* PCR System 9700 (Applied Biosystems) in 7µl final volume with 900nM of each primer, 200nM of each probe, dNTP, 0.2 unit of BIOTAQ DNA Polymerase (Bioline) and a passive fluorescence reference Rox (Euromedex). Before and after the amplification, the fluorescence of the reaction was checked on an ABI Prism 7000 Real-Time PCR System (Applied Biosystem).

BAG3 exonic sequencing in patients with familial DCM. To search for mutations in *Bag3* in familial DCM, genomic DNA from 168 index patients was sequenced on an ABI 3100 capillary sequencing instrument (Applera). DNAs from other available family members were genotyped for the familial variant by PCR and sequencing. We also sequenced exons 2, 3 and 4 of *BAG3* in 364 individuals of European descent without known cardiac disease. Moreover, DNA from 95 controls of North African origin and 45 controls from Turkey were sequenced to check respectively for the presence of variants P115S and V468M in non-DCM populations.

2. Results

The results of a GWAS based on pools of DNA from patients with sporadic DCM and controls from several European populations and the replication of the main findings in two independent populations are reported hereafter.

The replicated loci suggest the implication of the genes encoding heat shock protein B7 (*HSPB7*) and B-cell lymphoma 2-associated athanogene 3 (*BAG3*) in DCM. The subsequent direct sequencing of all coding regions of the *BAG3* gene in patients with familial DCM identified several deleterious mutations.

The populations of patients and controls included in this study at the discovery and replication stages were all of European origin. We used strict criteria[14] for patient selection to exclude secondary or familial forms of cardiomyopathy. Overall 1179 sporadic DCM patients and 1108 controls were included in the discovery phase of the study while 1165

sporadic DCM patients and 1302 controls contributed to the replication phase. The discovery GWAS was performed on pools of DNA (pools-GWAS). Overall 26 DNA pools were constituted according to study population, disease status, gender and age. To improve the precision of effect estimates when comparing pools of DNA from patients and controls, each pool was duplicated and each duplicate was hybridized to two different genotyping arrays. The allele quantification in pools was performed using the *Illumina* 610 quad beadchip (*Illumina*)). After filtering out copy number variation (CNV) markers and single nucleotide polymorphisms (SNPs) with poor signal intensity, we retained 517,382 SNPs in the analysis.

Pools-GWAS, validation by individual DNA genotyping and replication

We used a mixed linear model to compare allelic signals between pools of patients and pools of controls. DCM status, age category, gender and study population were considered as fixed effects and pool replicates as random effects. At the discovery stage, a slightly liberal threshold of $P < 5 \times 10^{-7}$ was adopted for declaring a SNP significant, in an attempt to compensate for the random error added by the pooling. At this level of significance, 14 SNPs showed significant case-control differences in the pools-GWAS. To check the validity of these associations, these SNPs were genotyped on individual DNA in the same discovery samples (Figure 1).

In the individual-based analysis, two SNPs, rs2234962 on chromosome 10 (OR associated with minor allele: 0.52 (0.44-0.62), $P=1.13 \times 10^{-13}$) and rs10927875 on chromosome 1 (OR: 0.71 (0.62-0.81), $P=3.7 \times 10^{-7}$), remained significantly associated with DCM. A third SNP, rs16983785 on chromosome 21 (OR: 1.79 (1.42-2.26), $P=7.22 \times 10^{-7}$) was just above the pre-specified threshold. These three SNPs were tested for replication in 2 independent case-control studies of DCM from Germany and United Kingdom.

The pooled-GWAS associations were replicated for rs10927875 (OR: 0.82 (0.72-0.93), $P=0.0021$), and rs2234962 (OR: 0.82 (0.70-0.95), $P=0.0092$); while for rs16983785 replication was not achieved (OR: 1.25 (0.98-1.61), $P=0.076$). In the data set combining the discovery and replication samples, rs10927875 and rs2234962 were associated with respective P-values of 9.5×10^{-10} and 4.0×10^{-12}. The association of rs2234962 with DCM was replicated in men ($P=0.0016$) but not in women ($P=0.59$). This difference may largely be attributed to the low minor allele frequency (MAF) of this SNP in women included in the German replication control sample (0.138 compared to >0.188 in the other groups, data not shown). This low frequency is probably due to chance as the MAF of rs2234962 in German men and women was 0.214 and 0.218 respectively in a population-based sample of 6298 individuals from the Gutenberg Heart Study[15].

Further exploration of the rs10927875 association signal on chromosome 1

rs10927875 is not the best marker of DCM risk in the region. rs10927875 is located in an intron of *ZBTB17* (zinc finger and BTB domain containing 17 also frequently referred as *MIZ-1*) on chromosome 1p36.2-p36.1. The gene is located within a region exhibiting strong linkage disequilibrium (LD) and spanning several other genes: *SPEN* (spen homolog, transcriptional regulator), *HSPB7* (heat shock 27 kDa protein family, member 7), *CLCNKA* (chloride channel Ka) and *CLCNKB* (chloride channel Kb). Recently, using a custom genotyping array targeting putative cardiovascular candidate genes, Stark et al PLoS Genet, 2010 Oct 21;6(10):e1001167, identified and replicated an association between a SNP in HSPB7 (rs1739843) and DCM, in a study using the same group of patients as the German replication sample of DCM cases used in the present study. Stark et al. did not genotype rs10927875, however their result pointed to HSPB7 as being the gene responsible for the association with DCM. LD may explain the presence of several DCM-associated SNPs in the region (data not shown), especially in the sequences of HSPB7 (rs1763601, pools-GWAS: P<1.0 x 10-4, this SNP being a perfect proxy of rs1739843, the DCM-associated SNP identified by Stark et al. according to HAPMAP release 22) and CLCNKA (rs1805152, pools-GWAS: P=5.6 x 10-6). Because DNA-pooling introduces some error in the estimation of association statistics that may affect the observed relative effects of linked SNPs, we additionally genotyped rs1763601 and rs945417 (this proxy SNP was used in place of rs1805152 which we had difficulties to genotype by *Taqman* assay) on individual DNA in the discovery samples (data not shown). Two other tightly linked potentially functional SNPs were also genotyped, rs945425 which is located in the proximal promoter region of *CLCNKA* and may generate an alternative splice site, and rs1048261 which is located in the 3'UTR of *HSPB7*. The 4 additionally genotyped SNP are represented in their genomic context in Figure 2.

Three of the 4 additional SNPs exhibited stronger associations with DCM than the lead SNP identified in the pools-GWAS (rs1763601: P=3.4 x 10^{-9}; rs945417: 1.4 x 10^{-9}; rs945425: 4.2 x 10^{-7} and rs1048261: 7.9 x 10^{-9} as compared to P=1.35 x 10^{-7} for rs10927875) and for all these SNPs, the minor allele was associated with a reduced risk of DCM. In a stepwise logistic regression model including the 5 SNPs, only rs945417 remained significantly associated with DCM as a consequence of the LD in the region.

Yin/yang haplotypes. To get better insight into the haplotypic structure of the region and its impact on the association of the locus with DCM, haplotypic relative risks[16] were computed. The 5 genotyped SNPs determine 6 common haplotypes. The two major haplotypes CTGCT (0.508 and 0.603 in controls and DCM cases, respectively) and

TGCTA (0.301 and 0.230) differ at the five sites and are likely to represent ancestral alleles, the four less common haplotypes resulting from recombination between them. In comparison to CTGCT, the yin/yang haplotype TGCTA was associated with a reduced risk of DCM (OR: 0.64 [0.53-0.74], P=3.2 x 10^{-9}]. Yin/yang haplotypes are commonly observed in the human genome[17,18,19], and such configuration has already been reported for the *CLCNKA* and *CLCNKB* genes nearby *HSPB7*[20]. Recently, through a systematic sequencing of *HSPB7* in patients with sporadic DCM and controls, Matkovich et al.[21] identified 12 SNPs associated with systolic heart failure. Some of them are in strong LD with the SNPs investigated here or with tight proxies of them (according to HAPMAP[22]), suggesting that the variants reported in both studies correspond to the same yin/yang haplotypic structure.

***HSPB7* is the best candidate at the locus but does not exhibit any coding variant.** Among the genes located in the region, *HSPB7* that encodes the "cardiovascular heat shock protein"[23] and exhibit cardiac-specific expression, is an obvious candidate for an association with DCM. However none of the DCM-associated SNPs or tight proxies identified through systematic sequencing by Matkovich et al[21] or genotyped in the present study affects the coding sequence of *HSPB7*. Moreover, we sequenced the *HSPB7* exons in 168 independent index cases diagnosed with familial DCM but were unable to identify any coding variant affecting the sequence of the HSPB7 protein (data not shown).

***HSPB7* expression is strongly affected by *cis*-acting SNPs.** Given the lack of coding variant in the *HSPB7* sequence, we envisaged that sequence variations in the *HSPB7* region might be related to DCM risk through an effect on gene expression. To investigate this hypothesis we looked for eQTLs in the *HSPB7* region using two large eQTLs databases built from RNA expression data obtained from circulating monocytes (GHS_Express[15]) and from circulating monocytes and *in vitro*-derived macrophages (Cardiogenics_Express[24]). In both studies, *HSPB7* mRNA was not detected in monocytes. However, in the Cardiogenics resource, the gene was expressed in macrophages and several SNPs were associated in *cis* with its expression. The strongest associations were observed with rs945425 (P=5.8 x 10^{-57}), rs10927875 (P=5.3 x 10^{-42}), rs1763601 (P=1.9 x 10^{-36}) and rs1805152 (P=4.8 x 10^{-31}). These SNPs explained 34%, 26%, 23% and 20% of the inter-individual variability (R^2) of *HSPB7* expression respectively and their minor allele was associated with an increased *HSPB7* expression. As the minor alleles of rs945425 and rs1763601 were less common in DCM patients than in controls (see above), it might be hypothesized that an increased expression of *HSPB7* is associated with a reduced risk of DCM; however, the SNP most strongly correlating with *HSPB7* expression, rs945425, was not the one most differing between DCM patients and controls (rs945417).

This obviously does not preclude a possible involvement of *HSPB7* expression in the association, but as a consequence of the extensive LD and the yin/yang haplotypic structure in the region, an unexpected regulatory mechanisms at this locus may affect DCM risk. The other genes in the region were either not expressed in monocytes or macrophages (*CLCNKA*, *CLCNKB*) or their expression was not influenced by the DCM-associated SNPs (*ZBTB17*).

Further exploration of the rs2234962 association signal on chromosome 10

rs2234962 may directly affect DCM risk. rs2234962 is a non-synonymous SNP (c.T757C, p.C151R) located within the coding sequence of *BAG3* on chromosome 10q25.2-q26.2. Our results show that the minor allele of this variant is associated with a reduced risk of DCM. Although causality cannot be established from such data, several arguments support the possible direct involvement of this SNP in DCM susceptibility: 1. the cysteine at position 151 of the BAG protein is conserved across several species and the Polyphen tool[25] predicts that the substitution of an arginine at this position is probably damaging; 2. a plot of DCM-association *P*-values of SNPs located around rs2234962 shows that the other associated SNPs in the pools-GWAS exhibit rather modest *P*-values in comparison to rs2234962 (data not shown); 3. as rs2234962 could tag functional SNPs absent from the genotyping array used in our study, we examined the LD of this SNP with other variants in the region and observed that only 2 SNPs in HAPMAP release 22 were in tight LD ($r^2 > 0.80$) with the lead SNP and both were intronic with no evidence that they could be functional.

Several variants on *BAG3* may affect DCM risk. The HAPMAP database revealed that the *BAG3* sequence presents several non-synonymous SNPs. To complement our analysis we selected two of them, rs3858340 (c.C1526T, p.P407L) and rs35434411 (c.G518A, p.R71Q) on the basis of a MAF>5% in the HAPMAP CEU population, and individually genotyped them in our discovery cohorts. The rs35434411 was relatively uncommon in our data set and it was not associated with disease status (MAF 0.032 and 0.029 in patients and controls, respectively). Conversely, when tested alone, rs3858340 was significantly associated with DCM ($P=3.6 \times 10^{-5}$, MAF of 0.12 and 0.081 in patients and controls, respectively). A multiple logistic regression analysis suggested that both rs2234962 and rs3858340 independently contributed to DCM susceptibility ($P=3.1 \times 10^{-12}$ and $P=3.6 \times 10^{-3}$, respectively). In the discovery cohort, these two SNPs were in complete LD (D'=-1) defining 3 haplotypes, C757-C1526 (R-P), T757-C1526 (C-P) and T757-T1526 (C-L). We were therefore able to unambiguously determine the haplotypic pairs (i.e. diplotypes) carried by each individual and assessed their associations with DCM. The results suggest that both at the heterozygous and homozygous states, the combination of an arginine (R) at

position 151 and a proline (P) at position 407 of the BAG3 protein sequence protects against DCM.

Finally, the simultaneous association of the *HSPB7* and *BAG3* loci with DCM was tested by logistic regression in the discovery samples using individual genotypes. The results showed that their effect was additive (rs945417: $P<2.5 \times 10^{-8}$, rs2234962: $P<5.6 \times 10^{-13}$) with no evidence of interaction ($P=0.7$).

Rare mutations in BAG3 are present in patients with familial forms of DCM

Given the possible implication of non-synonymous variants in *BAG3* in sporadic DCM, we investigated whether mutations in *BAG3* might be implicated in familial forms of DCM. The four *BAG3* exons and intron-exon boundaries were sequenced in 168 independent index cases of mainly European DCM families. A total of 19 molecular variants were detected (see Table 4); among them, six were known SNPs referenced in dbSNP, four were present in a control group of 347 healthy individuals of European descent in which they were tested and three were unlikely to be deleterious. Among identified variants, 6 were known SNPs referenced in dbSNP and 4 were present in the control group of 347 healthy individuals of European descent in which they were tested. The 9 remaining variants were found once and none of them was present in the control group. All carriers of these mutations were heterozygous. The identified variants included three insertions/deletions resulting in a truncation of the encoded protein sequence (Q251RfsX56, R396GfsX48, S385QfsX56), a substitution creating a premature stop codon (p.R309X) and 5 missense mutations leading to single amino acid changes (I94F, P115S, P380S, E455K, V468M). Further analyses excluded 3 missense variants as possibly disease causing: I94F, despite perfect interspecies conservation, was not segregating in one affected relative of the family; P115S was not conserved and found in one control subject from North African origin out of 95 genotyped (the index case carrying this variant was of North African origin, explaining why a control group of similar origin was genotyped for this variant); P380S was not conserved. Software prediction of mutation consequences using Polyphen, SIFT and SNAP was in accordance with this conclusion as only E455K and V468M among the 5 missense mutants were predicted to impair the protein function.

Table 4: Identified SNPs and mutations in the coding sequences of BAG3 in 168 index cases with familial form of DCM

Exon	genomic position	variant name	protein effect	nature of variant	dbSNP reference	frequency in index cases, n/336 alleles (%)	frequency in controls, n/694 alleles (%)
2	121429394	212G>A	R71Q	missense	rs35434411	9 (0.026)	14 (0,020)
2	121429412	536C>T	P77L	missense		0	1 (0,001)
2	121429462	586A>T	I94F	missense		1 (0.003)	0
2	121429525	649C>T	P115S	missense		1 (0.003)	0
2	121429633	757T>C	C151R	missense	rs2234962	30 (0.089)	143 (0,206)
2	121429645	769G>A	A155T	missense	rs61756328	0	4 (0,006)
3	121431946	993G>A	Q229Q	synonymous		0	1 (0.001)
3	121432011	1058delA	251RfsX56	frame shift		1 (0.003)	0
3	121432036	1083C>G	P259P	synonymous		0	1 (0.001)
3	121432147	1194C>T	His296His	synonymous		0	1 (0.001)
4	121435991	1231C>T	R309X	nonsense		1 (0.003)	0
4	121436068	1308T>G	P334P	synonymous	rs3858339	43 (0.128)	64 (0.092)
4	121436204	1444C>T	P380S	missense		1 (0.003)	0
4	121436219	1459_1466 delTCTTCCC	385QfsX56	frame shift		1 (0.003)	0
4	121436246	1486_1487 delGA	395GfsX48	frame shift		1 (0.003)	0
4	121436286	1526C>T	P407L	missense	rs3858340	43 (0.128)	66 (0.095)
4	121436362	1602A>G	V432V	synonymous	rs196295	69 (0.205)	146 (0.210)
4	121436429	1669G>A	E455K	missense		1 (0.003)	0
4	121436468	1708G>A	V468M	missense		1 (0.003)	0

The six remaining variants, including four truncating and two missense mutations, affect the *BAG3* coding sequence and are likely to be disease causative. Two of the *BAG3* mutations were observed in relatively large families (6 and 5 proven mutation carriers, respectively) and the 4 others were observed in small families (1 to 2 mutation carriers in each of these families). The penetrance was high, with presence of DCM in 13/18 mutation carriers (72%) and possible DCM in 3 (LV dilatation and mild LVEF<60% in 3, congestive heart failure in 2). Two mutation carriers had a normal cardiac examination (a 7-year- old male and a 33-year-old female).

Phenotypically, the DCM patients carrying a mutation showed isolated DCM, without associated conduction defect or skeletal myopathy and a normal serum creatine kinase

level in all but one patient. Four patients had heart transplantation and five relatives, whose DNA was not available, died prematurely from cardiac cause with a previous diagnosis of DCM.

3. Discussion

We report the results of the first GWAS conducted in patients with HF due to DCM. For the discovery phase of this study, we used a DNA-pooling approach and identified 14 SNPs at different loci showing some evidence of association with DCM. Three of these associations were confirmed by individual genotyping and two of them were replicated in independent DCM patients and controls.

Even when a very strict methodology is applied, estimates of allelic effects based on DNA pools are affected by variation in the quantity and concentration of individual DNAs composing the pools and by quantification errors[26]. The reduced specificity of pools-GWAS may be accounted for by individual-based genotyping of the most significant hits in the same samples. We observed that associations observed in the pools-GWAS were more reproducible in individual genotyping when several SNPs in LD at the same locus were associated with the disease. The observation made at the *ZBTB17/HSPB7* locus is interesting in this regard, because the strongest hits were not the same in the pools-GWAS and in the individual genotyping. As a consequence of errors inherent to the DNA pooling approach, pools-GWAS also exhibit a reduced sensitivity, implying that our study may have missed some associations that would have been identified by an individual-based genotyping. Nevertheless, this approach is efficient[27,28,29] and the interest of the *Illumina* technologies for pools-GWAS has been emphasized[26].

The first DCM-associated SNP is located in a region exhibiting a yin/yang haplotype structure and encompassing the *HSPB7* gene. *HSPB7* is also called cardiovascular HSP as a consequence of its selective expression in cardiovascular tissues[30] and it belongs to the small HSP (sHSP) family[31], whose member's main function is to protect a variety of tissues by binding denatured proteins. The physiological response of muscle fibers to stress involves HSPs of high molecular mass, such as HSP70 and sHSPs. An example is provided by HSPB5 (αB-crystallin); a dominant mutation in this gene causes a severe form of desmin-related cardiomyopathy characterized by accumulation of misfolded proteins as a consequence of impaired autophagy[32]. In humans, genetic variants in *HSPB7* have recently been reported to be associated with advanced heart failure and systolic dysfunction of unspecific origin[21,33].

The second DCM-associated SNP affects the sequence of the protein encoded by *BAG3*. BAG3 is a member of a conserved family of cyto-protective co-chaperones proteins containing a conserved domain able to interact with HSC70/HSP70 and sHSPs proteins. BAG3 is involved in numerous activities including macro-autophagic protein degradation in aging cells[34]. BAG3 is mainly expressed in striated muscle and colocalizes with Z-disks (a sarcomeric protein assembly essential for actin anchoring in striated muscles). Following normal muscle development, *Bag*-3 deficient mice present a progressive myopathy with Z-disk disruption and develop a fulminant myopathy characterized by non-inflammatory myofibrillar degeneration with apoptotic features[35]. Autophagic degradation appears essential for maintaining Z-disk integrity and muscle contractility and Bag3 plays an essential role in this process which also involves small Hsps, such as HspB8[36].

In humans, a mutation in BAG3 (P209L) that affects a conserved motif of the BAG3 protein (I-P-V) known to play an important role in the interaction of BAG3 with sHSP[37] causes severe dominant childhood muscular dystrophy with cardiomyopathy[38]. In our study, we identified 6 mutations, all of them affecting the BAG domain of BAG3, either by substituting highly conserved amino acids or by deleting the whole domain. Surprisingly, none of the carriers of the mutations presented with myofibrillar disorders as observed for the P209L mutation, suggesting a specific role for the BAG domain deletion/dysfunction in the DCM phenotype. In *Bag3* knock-out mice fulminant myopathy and cardiomyopathy are observed only in homozygous -/- animals but not in heterozygous[35] while in drosophilia, the *Bag3* ortholog starvin deletion is associated with locomotion decline and myofiber sarcomeric disorganization even in heterozygous animals[36]. Further experiments are clearly needed to better understand if either poison peptide or haplo-insufficiency drives the pathophysiology of DCM related to *BAG3* mutations.

In conclusion, this GWAS identified two loci associated with sporadic DCM. The most likely candidate genes at these loci are *HSPB7* and *BAG3*. Beyond DCM, it will be of interest to investigate the implication of the two polymorphic loci in other forms of heart failure and muscle dysfunction, as well as in the age-related changes that affect the heart and muscle physiology. As recently discussed, drug targeting the "proteasis network"[36,41] may interfere with the progressive increase in muscle weakness seen in DCM patients and elderly people. An important aspect of this study is the discovery of *BAG3* sequence variants being involved in both sporadic and familial forms of DCM.

Table 5: Nucleotide sequences for practicing the method of the invention

NO :	Description	Sequence
1	cDNA HSB7	CTTTTACCTC TGACTGAAAT CCCACCTAAG GCGCTCGCAGGCATCACCTCTGACACAGAG ATGGAGCAGA CAGTTGCCGG GAACACAACC TTCTAGACCTTTAAGGAGAA AGCTGCCAAC TTTGTCCTCT TGGGGCTGCT GAATGTTCTGAGAGTCAGAG GGAGTGCAGG CACAGAATCC AGGTCCCCAT GGGGGCTGGCAGGGGACAGC TGCACATTTT ACACCCCCTG GCAGGGCTCA ACCTTTCCTTAGAGATGGGC TGTGAAGGTT GTCTGCCTTG AGGGAAGGAG CTCCTGACGTCACAGAGCGG TGGTGCTTCC ACCTACCCTG TCCCCAAGGA GCCTGCACACTGATGGAGGA GGTACAGCCG TGGAGGAGAA CCAGAGTCTG GAGGGGAGGCACAGTCCTCA GGAGCCCCCA GGCTGACGGG GGAAACAAAA CACACAATAAACCCCAAACT TGCAAGGTTC TGGGGCAGGC ACCGAGGTCC GAGTGGCAAGGAGACAGGAT CCAGCCCTGT GCAGTGGTGA GGAGGCCCCT GGGGCTGCTGTGATCAGCCT GGAATGTCAG GCTGTGAGCT CAGCCCTCTC GGCCTGCGTGTCGGGCTGTC AGCAGGGCCG GCGCACGGGA GAAGATTTCCAGCCCCCTCTGGGCCATGGA GAGGCCTAAA TTTAGCTAGT AGACAAGGGG CTGGCCCCAAGCCCAGGGAC AGTCGGCCTT ATAAAGCGGC CGCAGTCGGA GCCTGGCACGCTCGCCCAGA GGCCTGCGCC CACACCCTCT CCTGTCCAGC CCTCGCCCGCCTGGGCAGGG CCCGGCGCCG TCCGTGGATG AGCCACAGAA CCTCTTCCACCTTCCGAGCG GAGAGAAGTT TCCATTCCTC TTCCTCTTCC TCCTCCTCTTCCACCTCCTC CTCGGCCTCC CGTGCTCTCC CGGCCCAGGA CCCGCCCATGGAGAAGGCCC TGAGCATGTT TTCCGATGAC TTTGGCAGCT TCATGCGGCCCCACTCGGAG CCCCTGGCCT TCCCAGCCCG CCCCGGTGGG GCAGGCAACATCAAGACCCT AGGAGACGCC TATGAGTTTG CGGTGGACGT GAGAGACTTCTCACCTGAAG ACATCATTGT CACCACCTCC AACAACCACA TCGAGGTGCGGGCTGAGAAG CTGGCGGCTG ACGGCACTGT CATGAACACC TTCGCTCACAAGTGCCAGCT GCCGGAGGAC GTGGACCCGA CGTCGGTGAC CTCGGCTCTGCGGGAGGACG GCAGCCTCAC TATCCGGGCA CGGCGTCACC CGCATACAGAACACGTCCAG CAGACCTTCC GGACGGAGAT CAAAATCTGA GTGCCTCTCCCTTCCCTTTC CCTGTCCCCC CGCCCCACGC CTGCCAGCAA AGCCTCGCTAACCCCATTAC AACAGCTCCA GGACATCTCA GCCCAGGTTC TAGCCCCCACGCACCCCAGA CCCCAGGTGG ACCATCCTCC CAAACTAGGG CCCTCCACTCTATCCAGGGC AGGCCAGGGA CTCCCTGGCC TGACACATGA TGCCCAGATTTCAGATTTGG CCTCCGTCAC TTAATCCAGA GTACAGGGGC TGGGGTCAGGGAAGGAAGAT CTAAAGAACC CACTGTGGGT CAGGGGAATG GGACCAGCAGGACATATGGG CAAGCTCTGC AGGACAGACA GACAGACAAA CCCTCTGATCTATGAAGTCT CTGCAGGGCA AGGGGACCAG GGACCTGGAA CCCTCTTGGCCAAGGGGAGT GGGAGGGACA GAGGGAAGGT CACAGGCAAG GGTGCCTATCTAAGTGGAAC TAATTGCCCG AGGGCTCAGC AAGGCCAAGA GGAGACAGCCGTGACGGTAA ACTTCCCCTC TACCAGCCTC CAAGCCCCAC GCCAGCGAGC AGGCTGCCTG CCCACCCCGT GCCCCCAGCC AGCTGGCTGT GCCAGGGCAGAGCCATGCCA CATCTGTATA TAGATGGGGT TTTTCCAATA CAGCTGGTTCGTGATAAACT GCATGAAACT CCTGCCGTCC TGCGCCTGCT GGGGCCTCCAGGCAAGGCCA AGTGGGGTTG GGGGTGGGGC TGGTCCTTCT CCCTCCCACAGGCCTGTGTT CTTGGGGCTG CTCCCATGCA GACAGGATCA CCTAACAGAGATGGAAGCCA GGGCATGGAT GGGGCTTTGG GTCCTCGAGG TTGGACCCCAGCTTCTTGCC ACCTTCCCCT CCGGGCAGTC AGCTCTCCAT CCATCCCCCT CTTTAATCTA TGAATCTATA GGCTCGGTGT GTGTAACACA CACACCCCTA TCGTTGTCCT TCAAATACTC AGCATTACCA TTGGTTGAGG CCAAATTCAG AGCTTTCTCA AATCAGATTT ACAATCTCCA TTTTCATTAA CGGGGAAACA TCCCCGAGCC ACTGAGTGCT GTGCTTTGTC ACTGAAGGTT AGATCTGAAC CCAGGGTGTC AACTGCTGCT CTCAACTCCC CACCTCTGGG CACTGAGGAG TATTTCCCCT CATTCTACCT CTCTAAGGCT ATGCAACCCT CCCCACGTCT TCCAGCTGGG GGATGGGGGG GAGTCATAGG AAAAGCCCCC ATCTCCCATC TGGGATAGGG ACCTTCCATC AGCCTTAACC CTGGGAAATG CCTGCTGCCC CCAATGACTC TTGGTTTCGT CTCCCACATA CAGAAGCAGG GTGGAGGGGA AGGGTGGGTC TCAGTTAGCA GGGGTCCCCA GGGCAAGTCA GCCTCCTCCC TCCATGCCTC TCTGGTCAGT GTGCCTTAGG

		GTGGCCTCTC ACTCCCACCA CTCTGGGCCC CTTGGGGGAG GACTGGGGAG GGGGCCGTGG GAGAGCCCTG ACGCTGGAAC CTGTATACAC AATAAAGGAC AGTCTCACAG ACaaaaaaaa aaaaaaaa

2	HspB7 polypeptide	mshrtsstfr aersfhssss ssssstsssa sralpaqdpp mekalsmfsd dfgsfmrphs eplafparpg gagniktlgd ayefavdvrd fspediivtt snnhievrae klaadgtvmn tfahkcqlpe dvdptsvtsa lredgsltir arrhphtehv qqtfrteiki
3	cDNA BAG3	GCGGAGCTCC GCATCCAACC CCGGGCCGCG GCCAACTTCT CTGGACTGGA CCAGAAGTTT CTAGCCGGCC AGTTGCTACC TCCCTTTATC TCCTCCTTCC CCTCTGGCAG CGAGGAGGCT ATTTCCAGAC ACTTCCACCC CTCTCTGGCC ACGTCACCCC CGCCTTTAAT TCATAAAGGT GCCCGGCGCC GGCTTCCCGG ACACGTCGGC GGCGGAGAGG GGCCCACGGC GGCGGCCCGG CCAGAGACTC GGCGCCCGGA GCCAGCGCCC CGCACCCGCG CCCCAGCGGG CAGACCCCAA CCCAGCATGA GCGCCGCCAC CCACTCGCCC ATGATGCAGG TGGCGTCCGG CAACGGTGAC CGCGACCCTT TGCCCCCCGG ATGGGAGATC AAGATCGACC CGCAGACCGG CTGGCCCTTC TTCGTGGACC ACAACAGCCG CACCACTACG TGGAACGACC CGCGCGTGCC CTCTGAGGGC CCCAAGGAGA CTCCATCCTC TGCCAATGGC CCTTCCCGGG AGGGCTCTAG GCTGCCGCCT GCTAGGGAAG GCCACCCTGT GTACCCCCAG CTCCGACCAG GCTACATTCC CATTCCTGTG CTCCATGAAG GCGCTGAGAA CCGGCAGGTG CACCCTTTCC ATGTCTATCC CCAGCCTGGG ATGCAGCGAT TCCGAACTGA GGCGGCAGCA GCGGCTCCTC AGAGGTCCCA GTCACCTCTG CGGGGCATGC CAGAAACCAC TCAGCCAGAT AAACAGTGTG GACAGGTGGC AGCGGCGGCG GCAGCCCAGC CCCCAGCCTC CCACGGACCT GAGCGGTCCC AGTCTCCAGC TGCCTCTGAC TGCTCATCCT CATCCTCCTC GGCCAGCCTG CCTTCCTCCG GCAGGAGCAG CCTGGGCAGT CACCAGCTCC CGCGGGGGTA CATCTCCATT CCGGTGATAC ACGAGCAGAA CGTTACCCGG CCAGCAGCCC AGCCCTCCTT CCACCAAGCC CAGAAGACGC ACTACCCAGC GCAGCAGGGG GAGTACCAGA CCCACCAGCC TGTGTACCAC AAGATCCAGG GGGATGACTG GGAGCCCCGG CCCCTGCGGG CGGCATCCCC GTTCAGGTCA TCTGTCCAGG GTGCATCGAG CCGGGAGGGC TCACCAGCCA GGAGCAGCAC GCCACTCCAC TCCCCCTCGC CCATCCGTGT GCACACCGTG GTCGACAGGC CTCAGCAGCC CATGACCCAT CGAGAAACTG CACCTGTTTC CCAGCCTGAA AACAAACCAG AAAGTAAGCC AGGCCCAGTT GGACCAGAAC TCCCTCCTGG ACACATCCCA ATTCAAGTGA TCCGCAAAGA GGTGGATTCT AAACCTGTTT CCCAGAAGCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA AGTTCCCCCT GCTCCAGTTC CTTGTCCTCC TCCCAGCCCT GGCCCTTCTG CTGTCCCCTC TTCCCCCAAG AGTGTGGCTA CAGAAGAGAG GGCAGCCCCC AGCACTGCCC CTGCAGAAGC TACACCTCCA AAACCAGGAG AAGCCGAGGC TCCCCCAAAA CATCCAGGAG TGCTGAAAGT GGAAGCCATC CTGGAGAAGG TACAGGGGCT GGAGCAGGCT GTAGACAACT TTGAAGGCAA GAAGACTGAC AAAAAGTACC TGATGATCGA AGAGTATTTG ACCAAAGAGC TGCTGGCCCT GGATTCAGTG GACCCCGAGG GACGAGCCGA TGTGCGTCAG GCCAGGAGAG ACGGTGTCAG GAAGGTTCAG ACCATCTTGG AAAAACTTGA ACAGAAAGCC ATTGATGTCC CAGGTCAAGT CCAGGTCTAT GAACTCCAGC CCAGCAACCT TGAAGCAGAT CAGCCACTGC AGGCAATCAT GGAGATGGGT GCCGTGGCAG CAGACAAGGG CAAGAAAAAT GCTGGAAATG CAGAAGATCC CCACACAGAA ACCCAGCAGC CAGAAGCCAC AGCAGCAGCG ACTTCAAACC CCAGCAGCAT GACAGACACC CCTGGTAACC CAGCAGCACC GTAGCCTCTG CCCTGTAAAA ATCAGACTCG GAACCGATGT GTGCTTTAGG GAATTTTAAG TTGCATGCAT TTCAGAGACT TTAAGTCAGT TGGTTTTTAT TAGCTGCTTG GTATGCAGTA ACTTGGGTGG AGGCAAAACA CTAATAAAAG GGCTAAAAAG GAAAATGATG CTTTTCTTCT ATATTCTTAC TCTGTACAAA TAAAGAAGTT GCTTGTTGTT TGAGAAGTTT AACCCCGTTG CTTGTTGTTC TGCAGCCCTG TCTACTTGGG CACCCCCACC ACCTGTTAGC TGTGGTTGTG CACTGTCTTT TGTAGCTCTG GACTGGAGGG GTAGATGGGG AGTCAATTAC CCATCACATA AATATGAAAC ATTTATCAGA AATGTTGCCA TTTTAATGAG ATGATTTTCT TCATCTCATA ATTAAAATAC CTGACTTTAG AGAGAGTAAA ATGTGCCAGG AGCCATAGGA ATATCTGTAT GTTGGATGAC TTTAATGCTA CATTTTAAAA AAAGAAAATA AAGTAATAAT ATAACTCAAa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa

4	Bag3 polypeptide	msaathspmm qvasgngdrd plppgweiki dpqtgwpffv dhnsrtttwn dprvpsegpk etpssangps regsrlppar eghpvypqlr pgyipipvlh egaenrqvhp fhvypqpgmq rfrteaaaaa pqrsqsplrg mpettqpdkq cgqvaaaaaa qppashgper sqspaasdcs sssssaslps sgrsslgshq lprgyisipv iheqnvtrpa aqpsfhqaqk thypaqqgey qthqpvyhki qgddweprpl raaspfrssv qgassregsp arsstplhsp spirvhtvvd rpqqpmthre tapvsqpenk peskpgpvgp elppghipiq virkevdskp vsqkpppppse kvevkvppap vpcpppspgp savpsspksv ateeraapst apaeatppkp geaeappkhp gvlkveaile kvqgleqavd nfegkktdkk ylmieeyltk ellaldsvdp egradvrqar rdgvrkvqti lekleqkaid vpgqvqvyel qpsnleadqp lqaimemgav aadkgkknag naedphtetq qpeataaats npssmtdtpg npaap
5	BAG3 Exon 1	cgcgattatagccgatgact
6	BAG3 Exon 1	gcctgccgtcgaggt
7	BAG3 Exon 2	aggagggttcacttcccagt
8	BAG3 Exon 2	atgccctgcatgtgaacag
9	BAG3 Exon 3	ggagtcatttgtggggtcat
10	BAG3 Exon 3	ccctggagacataccaccat
11	BAG3 Exon 4.1	caatttctgtgactttcagtcagtt
12	BAG3 Exon 4.1	tttgtcagtcttcttgccttca
13	BAG3 Exon 4.2	gccatcctggagaaggtaca
14	BAG3 Exon 4.2	tttgcctccacccaagttac
15	HPSB7 Exon 1	cagggacagtcggccttat
16	HPSB7 Exon 1	cagacgtcccctccctgt
17	HPSB7 Exon 2	atgcaaggcctgctccat
18	HPSB7 Exon 2	gtctggggtcccaggatg
19	HPSB7 Exon 3	ggggttagaatggggagaag
20	HPSB7 Exon 3	ggctagaacctgggctgag
21	Primers F rs16983785	TGA CCC CCA GGA TTT TTC TGT
22	Primers R rs16983785	GTT TTT CAC TGC TAT TCT TGC ACA A
23	Primers F rs2234962	TCC CAG TCA CCT CTG CGG

24	Primers R rs2234962	CTC TCC TTA CCT CAG GTC CGT G
25	Primers F rs10927875	GTT TGG TTG GGA AAT TGT TGA TCT
26	Primers R rs10927875	TGG ACT GTG TAC TTC CAG AAT GTT AAG
27	Primers F rs945417	TGG AAT GTC AGG CTG TGA GCT
28	Primers R rs945417	ACT GTC CCT GGG CTT GGG
29	Primers F rs3858340	CAC TTT CAG CAC TCC TGG ATG TT
30	Primers R rs3858340	GGC TAC AGA AGA GAG GGC AGC
31	Probe rs16983785	ACA AGT TTG AGC ATG TAA TTT GCA GTG CTT C
32	Probe rs16983785	CAA GTT TGA GCA TGT AAT TTG CAG GGC TT
33	Probe rs2234962	CCA GAA ACC ACT CAG CCA GAT AAA CAG TGT
34	Probe rs2234962	ACC ACT CAG CCA GAT AAA CAG CGT GG
35	Probe rs10927875	CAC TGG GCA ATG CTC ACG CCT T
36	Probe rs10927875	TGA TCA CTG GGC AAT GCT CAC ACC
37	Probe rs945417	TCT GGG CCA TGG ACA GGC CT
38	Probe rs945417	CTC TGG GCC ATG GAG AGG CCT
39	Probe rs3858340	TCG GCT TCT CCT GGT TTT GGA GGT
40	Probe rs3858340	TCG GCT TCT CCT GGT TTT GGA AGT G
41	rs16983785	GACAAGTTTGAGCATGTAATTTGCAG**K**GCTTCTATTTTCACAGAAGCCAAAA

42	rs2234962	CAGAAACCACTCAGCCAGATAAACAG**Y**GTGGACAGGTGGCAGCGGCGGCGGC
43	rs10927875	cactgtAGATTTCACCTAACTAAAGG**Y**GTGAGCATTGCCCAGTGATCATCCA
44	rs945417	CTTGTCTACTAGCTAAATTTAGGCCT**S**TCCATGGCCCAGAGGGGGCTGGAAA
45	rs3858340	CAGCACTGCCCCTGCAGAAGCTACAC**Y**TCCAAAACCAGGAGAAGCCGAGGCT
46	rs1739843	CTGCCATCACCATCTCACACCTGTCA**Y**CTCCCACCCTCAAGCCCCACCCCCT
47	rs1763601	GGCAGGGCTGGAAACCCTGAGAGAAA**M**AGACACACCTTCTTTGAGATTTGGG
48	rs1805152	GCCTCTTGGGAGAGGCTCTTGCCGTC**R**CCTTCCCTGAGGGCATTGTGACTGG
49	rs945425	TTTGTCAAGGACGTTTTTAATCTCCA**Y**GGTAAACCCCGGCCGGCGGGGAAGG
50	rs1048261	CCCAGAGGTGGGGAGTTGAGAGCAGC**W**GTTGACACCCTGGGTTCAGATCTAA
51	rs35434411	TAGCAGGCGGCAGCCTAGAGCCCTCC**Y**GGGAAGGGCCATTGGCAGAGGATGG
52	rs61756328	AGCCAGATAAACAGTGTGGACAGGTG**R**CAGCGGCGGCGGCAGCCCAGCCCCC
53	rs3858339	GGCCCAGTTGGACCAGAACTCCCTCC**K**GGACACATCCCAATTCAAGTGATCC
54	rs196295	AAAGTGGAAGCCATCCTGGAGAAGGT**R**CAGGGGCTGGAGCAGGCTGTAGACA
55	rs7328410	aatatagacccagaaagtagattatt**K**tttgcttacaattgaggttgggggа
56	rs1991914	CTTAAAGTGTACAGTCTTCTCAAAGT**K**TGTGGGTGTAGAATTTTCTATTACA
57	rs13176432	CTTAGCAGAGGATAAGAGCAGAGACC**K**TCAGGCCTGGTCAAAGCATGTGCAG
58	rs10491858	CACCTGTGCTGGCATAGTGGAAAGGT**Y**CATCTGTTATCTAGGCCGAAGTGTT
59	rs5970164	TAGATACTGGCTTACATGCTACATGA**Y**GCCATTCCATTTCTTGCAAACACAT
60	rs856003	TGGGAAAAAGCTAACAAAATGGTAGA**Y**TTGAATCCAAATATATCAATAACAT

61	rs11543052	GAGCAGTGTGTGGATTACACTATCAC**Y**GGAAAAATACGAATTGAGAAGAAGG
62	rs1353456	TTACAACAGAGAAACGTCGCTATAGA**M**CTGCATTTTGTGAGCACACGGCCTT
63	rs2832070	TGCAAATTATTGGATTTGTCATTGCC**Y**TTTTGAAATACAGGCTACCAAATCA
64	rs1378796	AACTGTTCAAAGCCACTGGCTCATAG**M**CTGCTATCTCTATGAGGATGTTTAG
65	rs2290906	GAGGCACTGAGTGGCCAGGCACTCGA**R**GAAGGCAGCGTGGCATTCTGAGATG
66	rs2832227	ATTAACGAGATATCTTGATTACTAGG**R**TCTTTTGATTAGATCTCCTTTTGAT
67	rs17681175	CACAGGTACTCCTTGAACTATTGGCC**R**GGACATCGAATACTTGACAGCAGAT
68	rs7860026	ggctgttctcttttatgttcccatca**R**tggtatatgaggcactcactttctc
69	rs5970160	GCCCAGATTGGAAAAGTGCTCTGAGC**R**GTTCCTTTGTGACAATGGATGAACA
70	rs10886106	GCTTTATGAACACTCATTATTCTTAA**M**AATTAAACAAAAAGATACATTTATG
71	rs1805152	GCCTCTTGGGAGAGGCTCTTGCCGTC**R**CCTTCCCTGAGGGCATTGTGACTGG
72	rs5959428	ACATAAGACTGACTCCTCTTGATGCA**R**ACATACAGACATGGCAAAATAAGCT
73	rs2832057	TTCATTTTTGACATGGGTAATTTGTG**Y**TTTTGATCAAAGTCCAATCCTGCCA
74	rs12325933	GGGTAGCCCAGGGTGAGGCTCACGTC**Y**TCATCCTGCCCCTCGCCAGTGGGGA
75	Bag Domain	GVLKVEAILE KVQGLEQAVD NFEGKKTDKK YLMIEEYLTK ELLALDSVDP EGRADVRQAR RDGVRKVQTI LEKLEQKA

REFERENCES:

1. Lloyd-Jones, D. et al. Executive summary: heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation* **121**, 948-954 (2010).

2. Dickstein, K. et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *Eur. Heart J* **29**, 2388-2442 (2008).

3. Jefferies, J.L. & Towbin, J.A. Dilated cardiomyopathy. *Lancet* **375**, 752-762 (2010).

4. Hershberger, R., Kushner, J. & Parks, S. Dilated Cardiomyopathy Overview. *GeneReviews* (2007).

5. Elliott, P. 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* **29**, 270-276 (2008).

6. Everly, M.J. Cardiac transplantation in the United States: an analysis of the UNOS registry. *Clin Transpl* 35-43 (2008).

7. Codd, M.B., Sugrue, D.D., Gersh, B.J. & Melton, L.J. Epidemiology of idiopathic dilated and hypertrophic cardiomyopathy. A population-based study in Olmsted County, Minnesota, 1975-1984. *Circulation* **80**, 564-572 (1989).

8. Hershberger, R.E. et al. Coding sequence rare variants identified in MYBPC3, MYH6, TPM1, TNNC1, and TNNI3 from 312 patients with familial or idiopathic dilated cardiomyopathy. *Circ Cardiovasc Genet* **3**, 155-161 (2010).

9. Rampersaud, E. et al. Common susceptibility variants examined for association with dilated cardiomyopathy. *Ann. Hum. Genet* **74**, 110-116 (2010).

10. Tiret, L. et al. Lack of association between polymorphisms of eight candidate genes and idiopathic dilated cardiomyopathy: the CARDIGENE study. *J. Am. Coll. Cardiol* **35**, 29-35 (2000).

11. Samani, N.J. et al. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med* **357**, 443-453 (2007).

12. Yasuno, K. et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat. Genet* **42**, 420-425 (2010).

13. Smith, N.L. et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet* **3**, 256-266 (2010).

14. Mestroni, L. et al. Guidelines for the study of familial dilated cardiomyopathies. Collaborative Research Group of the European Human and Capital Mobility Project on Familial Dilated Cardiomyopathy. *Eur. Heart J* **20**, 93-102 (1999).

15. Zeller, T. et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5**, e10693 (2010).

16. Tregouet, D.A. & Garelle, V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. *Bioinformatics* **23**, 1038-1039 (2007).

17. Zhang, J., Rowe, W.L., Clark, A.G. & Buetow, K.H. Genomewide distribution of high-frequency, completely mismatching SNP haplotype pairs observed to be common across human populations. *Am. J. Hum. Genet* **73**, 1073-1081 (2003).

18. Curtis, D. & Vine, A.E. Yin yang haplotypes revisited - long, disparate haplotypes observed in European populations in regions of increased homozygosity. *Hum. Hered* **69**, 184-192 (2010).

19. Broadbent, H.M. et al. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet* **17**, 806-814 (2008).

20. Sile, S. et al. Haplotype diversity in four genes (CLCNKA, CLCNKB, BSND, NEDD4L) involved in renal salt reabsorption. *Hum. Hered* **65**, 33-46 (2008).

21. Matkovich, S.J. et al. Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease. *J. Clin. Invest* **120**, 280-289 (2010).

22. Thorisson, G.A., Smith, A.V., Krishnan, L. & Stein, L.D. The International HapMap Project Web site. *Genome Res* **15**, 1592-1593 (2005).

23. Krief, S. et al. Identification and characterization of cvHsp. A novel human small stress protein selectively expressed in cardiovascular and insulin-sensitive tissues. *J. Biol. Chem* **274**, 36592-36600 (1999).

24. Heinig, M. et al. A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* **467**, 460-464 (2010).

25. Ramensky, V., Bork, P. & Sunyaev, S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* **30**, 3894-3900 (2002).

26. Macgregor, S. et al. Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays. *Nucleic Acids Res* **36**, e35 (2008).

27. Huang, Y., Hinds, D.A., Qi, L. & Prentice, R.L. Pooled versus individual genotyping in a breast cancer genome-wide association study. *Genet Epidemiol* (2010).doi:10.1002/gepi.20517

28. Bossé, Y. et al. Identification of susceptibility genes for complex diseases using pooling-based genome-wide association scans. *Hum. Genet* **125**, 305-318 (2009).

29. Bostrom, M.A. et al. Candidate genes for non-diabetic ESRD in African Americans: a genome-wide association study using pooled DNA. *Hum. Genet* **128**, 195-204 (2010).

30. Krief, S. et al. Identification and characterization of cvHsp. A novel human small stress protein selectively expressed in cardiovascular and insulin-sensitive tissues. *J. Biol. Chem* **274**, 36592-36600 (1999).

31. Kappé, G. et al. The human genome encodes 10 alpha-crystallin-related small heat shock proteins: HspB1-10. *Cell Stress Chaperones* **8**, 53-61 (2003).

32. Tannous, P. et al. Autophagy is an adaptive response in desmin-related cardiomyopathy. *Proc. Natl. Acad. Sci. U.S.A* **105**, 9745-9750 (2008).

33. Cappola, T.P. et al. Common variants in HSPB7 and FRMD4B associated with advanced heart failure. *Circ Cardiovasc Genet* **3**, 147-154 (2010).

34. Gamerdinger, M. et al. Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J* **28**, 889-901 (2009).

35. Homma, S. et al. BAG3 deficiency results in fulminant myopathy and early lethality. *Am. J. Pathol* **169**, 761-773 (2006).

36. Arndt, V. et al. Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr. Biol* **20**, 143-148 (2010).

37. Fuchs, M. et al. Identification of the key structural motifs involved in HspB8/HspB6-Bag3 interaction. *Biochem. J* **425**, 245-255 (2010).

38. Selcen, D. et al. Mutation in BAG3 causes severe dominant childhood muscular dystrophy. *Ann. Neurol* **65**, 83-89 (2009).

39. McCollum, A.K., Casagrande, G. & Kohn, E.C. Caught in the middle: the role of Bag3 in disease. *Biochem. J* **425**, e1-3 (2010).

40. Nishino, I. et al. Primary LAMP-2 deficiency causes X-linked vacuolar cardiomyopathy and myopathy (Danon disease). *Nature* **406**, 906-910 (2000).

41. Balch, W.E., Morimoto, R.I., Dillin, A. & Kelly, J.W. Adapting proteostasis for disease intervention. *Science* **319**, 916-919 (2008).

42. Charron, P. et al. Identification of a genetic risk factor for idiopathic dilated cardiomyopathy. Involvement of a polymorphism in the endothelin receptor type A gene. CARDIGENE group. *Eur. Heart J* **20**, 1587-1591 (1999).

43. Parra, H.J. et al. A case-control study of lipoprotein particles in two populations at contrasting risk for coronary heart disease. The ECTIM Study. *Arterioscler. Thromb* **12**, 701-707 (1992).

44. Mazoyer, E. et al. Prevalence of factor V Leiden and prothrombin G20210A mutation in a large French population selected for nonthrombotic history: geographical and age distribution. *Blood Coagul. Fibrinolysis* **20**, 503-510 (2009).

45. R Development Core Team *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing: Vienna, Austria, 2008).à <http://www.R-project.org>

46. Sham, P., Bader, J.S., Craig, I., O'Donovan, M. & Owen, M. DNA Pooling: a tool for large-scale association studies. *Nat. Rev. Genet* **3**, 862-871 (2002).

47. Macgregor, S., Visscher, P.M. & Montgomery, G. Analysis of pooled DNA samples on high density arrays without prior knowledge of differential hybridization rates. *Nucleic Acids Res* **34**, e55 (2006).

48. Ge, D. et al. WGAViewer: software for genomic annotation of whole genome association studies. *Genome Res* **18**, 640-643 (2008).

49. Pruim, R.J. et al. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* (2010).doi:10.1093/bioinformatics/btq419

50. Fuchs, M. et al. Identification of the key structural motifs involved in HspB8/HspB6-Bag3 interaction. *Biochem. J* **425**, 245-255 (2010).

CLAIMS

1. An in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject, said method comprising the steps of:

   a. providing a biological sample from said subject; and,

   b. detecting the presence or absence of an HSPB7 disease associated variant and/or BAG3 disease associated variant in said biological sample,

   wherein the presence of said HSPB7 disease associated variant and/or BAG3 disease associated variant is indicative of a genetic predisposition to cardiomyopathy in said subject.

2. The diagnosis method according to Claim 1, wherein said cardiomyopathy is dilated cardiomyopathy.

3. The diagnosis method according to Claim 1 or 2, wherein said HSPB7 disease associated variant is a genetic mutation that decreases the expression level of HSPB7 gene without resulting in an alteration of the predicted amino acid sequence encoded by the HSPB7 gene.

4. The diagnosis method according to Claim 1 or 2, wherein said genetic mutation is one or more single nucleotide polymorphism (SNP) mutations in HSPB7 gene locus.

5. The diagnosis method according to Claim 4, wherein said SNP in HSPB7 gene locus is rs10927875 and/or rs945417.

6. The diagnosis method according to Claim 1 or 2, wherein said disease associated variant is a genetic mutation altering the predicted amino acid sequence encoded by the BAG3 gene.

7. The diagnosis method according to Claim 6, wherein said genetic mutation includes a non-synonymous single nucleotide polymorphism (SNP) in BAG3 coding regions.

8. The diagnosis method according to Claim 7, wherein said non-synonymous SNP in BAG3 coding region is rs2234962 or rs3858340.

9. The diagnosis method according to Claim 6, comprising detecting a variant polypeptide of Bag3 encoded by a genetic mutation as defined in any one of Claims 6-8 in said biological sample.

10. A kit for performing the diagnosis method of any one of Claims 1-9 comprising:

a. means for detecting one or more SNPs in BAG3 and/or HSPB7 gene loci or one or more causal mutation in the full gene sequence of BAG3;

b. optionally, instructions for use of the kit.

11. A kit for performing the diagnosis method of any Claim 10, comprising:

a. means for detecting said variant polypeptide of Bag3, such as antibodies that binds specifically to said variant polypeptide of Bag3 but not to corresponding wild type Bag3 protein; and,

b. optionally, instructions for use of the kit.

Retested SNP	Chr	Position	Locus	Pool-GWAS value [$]	*P*- Individual *P*-value [£]	MAF [µ] (cases/ controls)	Best proxy [&] of the retested SNP in the pool-GWAS and its association p-value with DCM
rs16983785	**21**	**29447289**	**TAK1L**	**3.8×10^{-11}**	**7.2×10^{-07}**	**0.102/0.059**	**rs2832227 (1/0.329) *P*=3.7×10^{-6}**
rs7328410	3	104824202	Intergenic	5.1×10^{-10}	0.035	0.056/ 0.045	NA
rs2234962	**0**	**121419623**	**BAG3**	**1.0×10^{-09}**	**1.1×10^{-13}**	**0.125/0.208**	**rs17681175 (0.712/0.156) *P*=2.1×10^{-5}**
rs1991914	4	4260014	OTOP1	1.6×10^{-08}	0.9	0.339/0.338	NA
rs13176432	5	114369723	intergenic	2.0×10^{-08}	0.0013	0.047/0.027	NA
rs10491858	9	1500495	intergenic	6.9×10^{-08}	0.0049	0.140/0.111	rs7860026 (1/1) *P*=4.9×10^{-5}
rs5970164	X	150849762	MAGEA4	1.2×10^{-07}	4.8×10^{-06}	0.103/0.167	rs5970160 (0.793/0.629) *P*=2.0×10^{-4}
rs856003	10	119393553	intergenic	1.2×10^{-07}	8.6×10^{-04}	0.195/0.157	rs10886106 (1.0/1.0) *P*=4.0×10^{-4}
rs10927875	**1**	**16171899**	**ZBTB17**	**1.3×10^{-07}**	**3.7×10^{-07}**	**0.269/0.341**	**rs1805152 (0.899/0.418) *P*=5.6×10^{-6}**
rs11543052	21	46880804	PRMT2	2.2×10^{-07}	0.0021	0.018/0.006	NA
rs1353456	X	78829377	intergenic	2.8×10^{-07}	1.0×10^{-4}	0.218/0.159	rs5959428 (1.0/0.8) *P*=7.8×10^{-6}
rs2832070	21	29059787	AF131217.1	3.1×10^{-07}	4.7×10^{-05}	0.155/0.111	rs2832057 (1.0/1.0) *P*=4.0×10^{-4}
rs1378796	3	158614482	VEPH1	4.4×10^{-07}	7.5×10^{-04}	0.099/0.074	NA
rs2290906	17	73605461	TNRC6C or TMC6	4.6×10^{-07}	5.8×10^{-05}	0.164/0.124	rs12325933 (1.0/1.0) *P*=9.0×10^{-4}

FIG 1

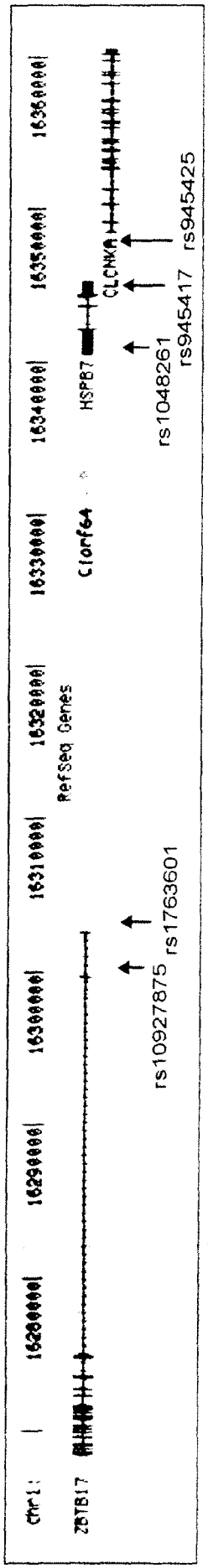
chr1:
16280000
16290000
16300000
16310000
16320000
16330000
16340000
16350000
16360000
ZBTB17
RefSeq Genes
C1orf64
HSPB7
CLCNKA
rs10927875
rs1763601
rs1048261
rs945417
rs945425

FIG 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2012/052352

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   [ ] on paper

   [X] in electronic form

   b. (time)

   [X] in the international application as filed

   [ ] together with the international application in electronic form

   [ ] subsequently to this Authority for the purpose of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/052352

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/68
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MATKOVICH SCOT J ET AL: "Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease", JOURNAL OF CLINICAL INVESTIGATION, vol. 120, no. 1, January 2010 (2010-01), pages 280-289, XP002644270, ISSN: 0021-9738 the whole document -----	1-4,10
A	WO 2010/001358 A2 (MOR RESEARCH APPLIC LTD [IL]; AMIR OFFER [IL]; LEWIS BASIL S [IL]) 7 January 2010 (2010-01-07) the whole document -----	1-5,10, 11

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 July 2012	16/07/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Hoff, Céline

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	VILLARD ERIC ET AL: "A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy.", EUROPEAN HEART JOURNAL MAY 2011 LNKD-PUBMED:21459883, vol. 32, no. 9, May 2011 (2011-05), pages 1065-1076, XP002644271, ISSN: 1522-9645 the whole document	1-5,10, 11
X	STARK KLAUS ET AL: "Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy.", PLOS GENETICS OCT 2010 LNKD-PUBMED:20975947, vol. 6, no. 10, October 2010 (2010-10), page E1001167, XP002644272, ISSN: 1553-7404 the whole document	1,2,4,10
X	CAPPOLA THOMAS P ET AL: "Common variants in HSPB7 and FRMD4B associated with advanced heart failure.", CIRCULATION. CARDIOVASCULAR GENETICS APR 2010 LNKD- PUBMED:20124441, vol. 3, no. 2, April 2010 (2010-04), pages 147-154, XP002644273, ISSN: 1942-3268 the whole document	1,2,4,10
X	DUYGU SELCEN ET AL: "Mutation in BAG3 Causes Severe Dominant Childhood Muscular Dystrophy", ANNALS OF NEUROLOGY, JOHN WILEY AND SONS, BOSTON, US, vol. 65, no. 1, 1 January 2009 (2009-01-01), pages 83-89, XP002661988, ISSN: 0364-5134, DOI: 10.1002/ANA.21553 the whole document	1,2,6,7, 9,10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2012/052352

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010001358 A2	07-01-2010	CA 2729640 A1	07-01-2010
		EP 2307572 A2	13-04-2011
		US 2011159493 A1	30-06-2011
		WO 2010001358 A2	07-01-2010

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2012/052352

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

1-11(partially)

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-5, 10, 11(all partially)

An in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject comprising the detection of the presence or absence of an HSPB7 disease associated variant.

---

2. claims: 1, 2, 6-11(all partially)

An in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject comprising the detection of the presence or absence of an BAG3 disease associated variant.

---

3. claims: 1-11(partially)

An in vitro diagnosis method for detecting a genetic predisposition to cardiomyopathy in a subject comprising the detection of the presence or absence of an HSPB7 and BAG3 disease associated variant.
