Cardiovascular Research Advance Access published March 2, 2012

Cardiovascular Research
doi:10.1093/cvr/cvs080

Whole-genome analysis of gene expression associates the ubiquitin-proteasome system with the cardiomyopathy phenotype in disease-sensitized congenic mouse strains

Boris T. Ivandic[1]*†, Sergey E. Mastitsky[2†], Frank Schönsiegel[1†], Raffi Bekeredjian[1], Roland Eils[2,3], Norbert Frey[4], Hugo A. Katus[1], and Benedikt Brors[2]

[1]Division of Cardiology, Department of Internal Medicine III, University Hospital Heidelberg, Im Neuenheimer Feld 410, Heidelberg 69120, Germany; [2]Division of Theoretical Bioinformatics, German Cancer Research Center, Im Neuenheimer Feld 580, Heidelberg 69120, Germany; [3]Department of Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology, and Bioquant Center, University of Heidelberg, Im Neuenheimer Feld 267, Heidelberg 69120, Germany; and [4]Division of Cardiology, Department of Medicine III, University Hospital Kiel, Arnold-Heller-Str. 3, Kiel 24105, Germany

Received 21 September 2011; revised 12 January 2012; accepted 30 January 2012

Time for primary review: 21 days

Aims Penetrance and phenotypic expressivity of cardiomyopathies are modulated by modifier genes both in model systems and patients. We aimed to dissect the disease-modifying mechanisms by examining genome-wide gene expression in a new set of mouse (*Mus musculus*) congenic strains.

Methods and results Mutant alleles of the genes calsarcin-1 (Myoz2), sarcoglycan-delta (Sgcd), and muscle LIM protein (Csrp2) were each transferred onto inbred strain backgrounds C57BL/6, C3H/He, 129S1/Sv, and FVB/N, respectively. At 9–10 weeks of age, left ventricular pump function (fractional shortening, FS) was determined by echocardiography in non-sedated congenic animals. Gene expression was then analysed in myocardial tissue using the Affymetrix Mouse 430.2 microarray platform. Variance stabilization, linear mixed-effects modelling, correlations, gene functional classification, and pathway analysis were conducted using the standard software. Strain background FVB/N appeared to protect against the consequences of gene inactivation. Sgcd-deficient congenics showed normal FS, which was consistent with their hypertrophic cardiomyopathy phenotype. Animals with other allele/background combinations developed an impaired ventricular pump function (FS $<$65%). Gender did not influence FS significantly, yet it determined the sets of genes that were differentially expressed in mice with low FS. In particular, genes encoding the elements of the ubiquitin-proteasome system (UPS) were strongly correlated with the cardiac impairment (absolute Spearman $r \geq 0.7$) in both males and females.

Conclusion Gene expression profiling in a novel set of congenic strains revealed an association between the UPS and myocardial contractile function, indicating that the UPS may be an important modifier of phenotypic variability in cardiomyopathies.

Keywords Cardiomyopathy • Congenic strain • Mouse • Gene expression • Proteasome

1. Introduction

Human cardiomyopathy is considered to be a genetic disorder, and over 40 disease-causing mutations have been identified in patients and model systems. In clinical practice, cardiomyopathies are classified according to their characteristic phenotypes such as hypertrophic cardiomyopathy (HCM), dilative cardiomyopathy (DCM), and restrictive cardiomyopathy, although identical disease genes may cause such diverse phenotypes.[1,2] Furthermore, in families with inherited cardiomyopathy, the disease phenotype and its penetrance are highly

* Corresponding author. MVZ synlab Leverkusen GmbH, Paracelsusstr. 13, D-51375 Leverkusen Germany. Email: boris.ivandic@synlab.com

† These three authors contributed equally.

variable even in family members carrying the same disease-causing mutation. These observations suggest a critical role of disease modifiers for phenotypic expressivity and the clinical progression of the disorder.

Genetically engineered mice are valuable model systems that are useful in dissecting the molecular pathogenesis of cardiomyopathies. However, similar to clinical observations, the penetrance and phenotypic expressivity of cardiomyopathies may vary considerably depending on the genetic background of the genetically engineered mouse model.[3,4] While patients may display a broad range of cardiomyopathy phenotypes as a consequence of genetic heterogeneity, inbred laboratory mouse strains with a given cardiomyopathy gene mutation exhibit a narrow range of phenotypes, like a subset of patients or a family with inherited cardiomyopathy. It is possible to study and compare their phenotypic characteristics in an unlimited number of their progeny. Rockman and co-workers have done pioneering work to identify genetic modifiers of DCM in transgenic mouse models.[5] We employed a similar approach and created a set of 10 disease-sensitized congenic inbred strains, including four closely related *Mus musculus domesticus* backgrounds. The aim was to use this novel resource to elucidate the molecular mechanisms of cardiomyopathy variability, not only for understanding the disease mechanisms, but also for improved risk prediction and therapeutic strategies.

Here, we report the identification of major genetic modifiers in cardiomyopathy mouse models obtained by analysing global gene expression patterns in the heart tissue. Starting with three established knock-out models on a mixed genetic background, we transferred by breeding their defective alleles (in the genes encoding for muscle LIM protein (Csrp2), myozenin 2 (Myoz2), and δ-sarcoglycan (Sgcd)) into four common inbred mouse strain backgrounds (C57BL/6, 129S1/Sv, FVB/N, and C3H/He). These new disease-sensitized congenic mouse models exhibited considerable variation in ventricular pump function, which was the central target for the identification of modifier genes by correlation analysis of cardiac gene expression patterns.

2. Methods

2.1 Animals

The inbred strains C57BL/6NCrl (C57BL/6), C3H/HeNCrl (C3H/He), FVB/NCrl (FVB/N), and 129S1/SvImJ@CRL (129S1/Sv) were purchased from Charles River Laboratories (Sulzfeld, Germany). Strain B6.129−Sgcd^{tm1Mcn}/J (B6.Sgcd) was obtained from The Jackson Laboratory (Bar Habor, MA, U.S.A.).[6] Strains B6.129−Myoz2^{tm1} (B6.Myoz2) and B6.129−Csrp2^{tm1} (B6.Csrp2) were received on mixed genetic backgrounds from Norbert Frey.[7,8] Both strains were backcrossed to C57BL/6NCrl for 11 generations to obtain homogeneously inbred C57BL/6 backgrounds: B6.Sgcd, B6.Csrp2, and B6.Myoz2, which exhibited targeted mutations in the genes *Sgcd*, *Csrp2* and *Myoz2*, respectively. B6.Myoz2, B6.Sgcd, and B6.Csrp2 were then bred with each inbred strain C3H/He, FVB/N, and 129S1/Sv to transfer the defective allele onto a new genetic background. In each generation, positive and negative selection based on densely genome-wide-typed SNP markers was used to identify the optimal backcross partners, which carried the largest proportion of the new genetic background while retaining the defective allele.[9] This optimized breeding strategy was applied for another 11 generations (F11) followed by brother-sister mating to finally obtain congenic strains B6.Myoz2, B6.Sgcd, B6.Csrp2, C3H.Myoz2, C3H.Sgcd, FVB.Myoz2, FVB.Sgcd, 129.Myoz2, 129.Sgcd, and 129.Csrp2, with homogenous backgrounds and homozygous defective alleles (*Table 1*). All animals were maintained in facilities with 15 air changes per hour and a 12/12-h light/dark cycle. Room temperature was 22 ± 2°C; relative humidity was 55 ± 10%. Up to five mice were maintained in Macrolon type II cages on hardwood bedding. The animals had free access to tap water and standard rodent chow (ssniff® M-Z Ereich low-phytoestrogen from Ssniff GmbH, Germany). This chow contained 48% of calories from carbohydrates, 37% from protein, and 15% from fat. This study conformed to the Directive 2010/63/EU of the European Parliament and the German Animal Welfare Act and was approved by the Institutional Animal Care of the University of Heidelberg and the Use Committee of the Regierungspraesidium Baden-Wuerttemberg.

Table 1 Composition and size of the experimental groups

Strain	Allele			Total
	Csrp2	Myoz2	Sgcd	
129S1/Sv	6M[a]/6F	6M/6F	6M/6F	18M/18F
C57BL/6	4M/5F	6M/6F	6M/4F	16M/15F
C3H/He	n.a.[b]	6M/6F	5M/4F	11M/10F
FVB/N	n.a.	5M/5F	6M/6F	11M/11F
Total:	10M/11F	23M/23F	23M/20F	56M/54F

[a]The number denotes sample size, and the letter stands for sex (M, male; F, female).
[b]Not available.

2.2 Echocardiography

A Sonos 5500 (Philips, Eindhoven, Netherlands) with a S12 transducer (12 MHz) was used to determine left ventricular function by transthoracic echocardiography. The echocardiographer was blinded for the mice genotype. All examinations were performed on conscious animals to prevent anaesthesia-related impairment of cardiac function.[10] Left ventricular parasternal short-axis M-mode views were obtained at the papillary muscle level. M-mode images were digitally stored for subsequent analysis. The left ventricular end-diastolic (LVEDD) and end-systolic internal diameters (LVESD) were determined and averaged over three consecutive beats using the American Society of Echocardiography leading edge method.[11] Fractional shortening (FS) was calculated as FS (%) = [(LVEDD − LVESD)/LVEDD] × 100. In these non-sedated mice an FS <65% was considered indicative of the impaired left ventricular pump function.[10] FS measurements in the animals of wild-type strains are presented in Supplementary material online, *Figure S1*.

2.3 Gene expression profiling using microarrays

At the age of 9−10 weeks, sedated mice were sacrificed by cervical dislocation. Hearts were rapidly removed, rinsed with isotonic phosphate-buffered saline (pH 7.4), blotted dry using filter paper, and dissected. Ascending aorta, left and right atrial tissues were removed. The remaining ventricles were shock frozen using liquid nitrogen and stored at −80°C pending isolation of total RNA. Hearts were thawed in Trizol (Invitrogen, Karlsruhe, Germany) and homogenized using a MM301 mixer mill (Retsch, Haan, Germany). After electrophoretic quality control and quantification, total RNA was sent to a commercial service company (MFT Services, Tuebingen, Germany) that performed genome-wide gene expression profiling using the Affymetrix Mouse 430.2 microarray platform according to their standard protocols.

Preprocessing of the microarray data was conducted in the R v2.11 computing environment[12] using the affy v1.28.0[13] and vsn v3.18.0[14] packages, which are part of the Bioconductor project.[15] Variance stabilizing normalization[14] and robust multi-array average-based summarization[16]

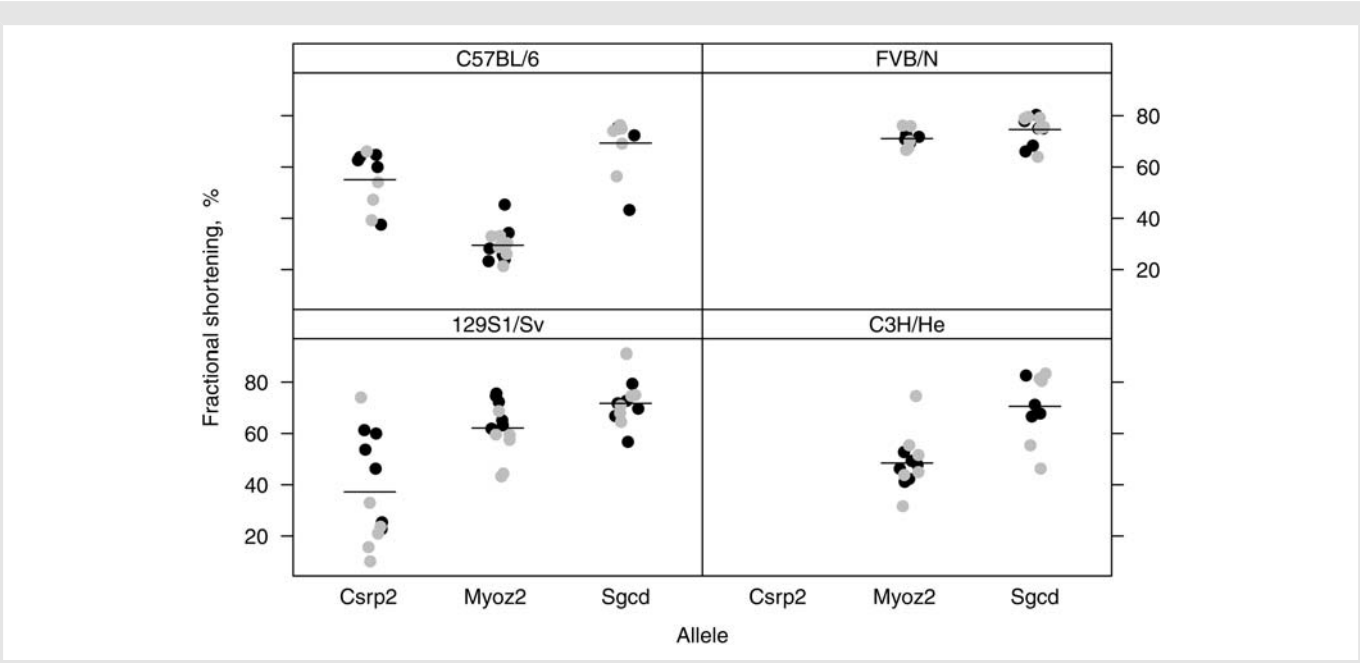

Figure 1 Fractional shortening in the examined cardiomyopathy models. Black and grey markers denote female and male animals, respectively. Short horizontal lines denote the median levels.

of the expression measurements were carried out without adjustment to the perfect match values, i.e. ignoring the values from mismatch probes.[17] The resulting expression values are presented on the $\log_2$ scale. To reduce multiple testing issues in the subsequent statistical analysis, the data were non-specifically filtered to discard the 10% lowest variability probe sets (based on the gene-wise standard deviation).

2.4 Statistical analysis

Statistical analysis was conducted in the R v2.11 computing environment[12] and included three steps.

2.4.1 Step 1

A linear mixed-effects modelling approach was applied to test the statistical significance of the association between FS and factors such as 'Strain', 'Allele' (i.e. the mutant cardiomyopathy-causing genes), 'Sex', 'Age', and the interaction between 'Strain' and 'Sex'. As indicated in *Table 1*, two of the strain/allele combinations were not available for the study, making its design unbalanced by allele. To address this issue, the effect of 'Allele' was modelled as a random intercept in the following model:

$$\mathrm{FS} = \beta_0 + \beta_1 \mathrm{Strain} + \beta_2 \mathrm{Sex} + \beta_3 \mathrm{Age} + \beta_4 \mathrm{Strain} \times \mathrm{Sex} + a + \varepsilon, \quad (1)$$

where β_0 is the average FS over all observations; β_1, β_2, and β_3 are the coefficients reflecting the effects of 'Strain', 'Sex' and 'Age', respectively; β_4 is the effect of the interaction between 'Strain' and Sex'; a is the random intercept allowed to vary with the levels of 'Allele' (it is assumed to be normally distributed with mean 0 and variance σ_a^2); ε are the model residuals, assumed to be normally distributed with mean 0. An exploratory data analysis revealed high heterogeneity of FS measurements in strain/allele combinations (*Figure 1*). The residual variance was, therefore, allowed to vary with each combination of strain and allele (see the work by Pinheiro and Bates[18] for detail on this type of parameterization).

Estimation of the coefficients of Model (1) was conducted with the nlme v3.1–97 package for R using the restricted maximum likelihood method.[18] Insignificant terms were stepwise backward eliminated from this saturated model. Selection of the optimal model was supported by the log-likelihood ratio test (L) and examination of the Akaike's Information Criterion AIC (models with smaller AIC values are considered optimal).[18] Pairwise comparisons of the mean FS levels in the animal groups of interest were performed with the Tukey test.[19]

2.4.2 Step 2

Spearman correlation coefficients (r_s) were calculated between the FS measurements and the normalized expression values of each transcript represented on the microarrays. Because we expected gender-specific gene expression patterns,[20] correlation analysis was carried out separately for male and female animals. P-values of the resulting coefficients were adjusted to account for the false discovery rate (FDR) following the Benjamini–Hochberg procedure.[21] Only the genes with $r_s \geq |0.70|$ and the FDR-adjusted $P < 0.01$ were considered to be strongly associated with FS, and were then passed on to subsequent gene functional classification and pathway analyses.

2.4.3 Step 3

Two types of functional classification were available in the DAVID knowledgebase (Database for Annotation, Visualization and Integrative Discovery).[22] They were applied to the lists of FS-correlated genes in order to identify biological mechanisms modulating cardiomyopathy. DAVID allows one to organize a list of genes into functionally related clusters. Using default settings, this analysis was run for gene lists obtained from separate analyses of male and female animals. However, we set the classification stringency parameter to 'lowest' when examining the relatively short list of FS-correlated genes shared by both males and females. The biological significance of the discovered functional groups was ranked based on the EASE score, an enrichment score that is the log-transformed geometric mean of the cluster members' P-values obtained from a modified Fisher's exact test. High EASE scores suggested a close association between the clustered genes in terms of their biological functions.[22] The main biological functions of a cluster were derived by visual examination of a 2D gene annotation heatmap offered by the tool.

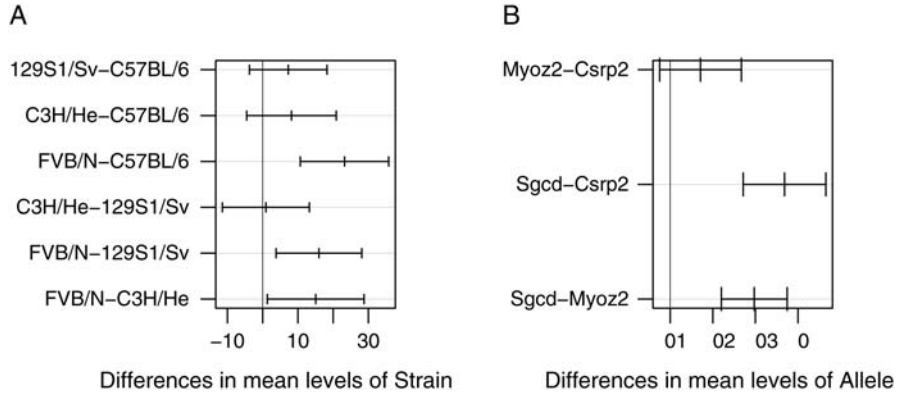

Figure 2 95% confidence intervals visualizing the results of pairwise comparisons of the mean FS levels grouped by strain (*A*) and mutant allele (*B*). The intervals that do not cross the vertical zero line correspond to statistically significant differences at $P < 0.05$.

Pathway analysis was conducted by mapping the lists of FS-correlated genes to the annotation terms from the KEGG database (Kyoto Encyclopedia of Genes and Genomes). This was performed using default settings in DAVID (in particular, the cut-off *P*-value 0.1 was used in the modified Fisher's exact test for enrichment). Following the recommendations of the developers of DAVID,[22] we treated the resulting *P*-values of enrichment tests as scores to rank the importance of a pathway, rather than using them as a statistical decision-making tool. For a more comprehensive ranking, we additionally report fold enrichment for each of the revealed pathways relative to the population background. Fold enrichment values of 1.5 and above are considered biologically important and worth the researcher's attention.[22]

3. Results

3.1 Fractional shortening depends on strain background and mutant allele

A mixed-effects statistical modelling approach was applied to determine the factors associated with FS in the examined cardiomyopathy models. FS substantially varied among the experimental groups (*Figure 1*), ranging from an average of 29.0 ± 6.5% in B6.Myoz2 to 75.0 ± 5.5% in FVB.Sgcd (mean ± SD here and elsewhere). In half of the groups, FS was >65%, indicating that these animals maintained a good ventricular function despite their carrier status.

Sex, age, and their combined effect appeared to be statistically insignificant ($P < 0.05$, *F*-tests). They were stepwise-backward eliminated from the initial Model (1), resulting in its reduction to

$$\mathrm{FS} = \beta_0 + \beta_1 \mathrm{Strain} + a + \varepsilon. \quad (2)$$

All parametric terms of Model (2), i.e. the intercept β_0 and the strain effect β_1, were highly significant ($P < 0.001$ in both cases, *F*-tests). In addition, the contribution of mutant alleles to the variance of FS in Model (2) was statistically different from 0 at $P < 0.05$ ($\sigma_a^2 = 54.0$, approximate lower 95% confidence limit = 4.5, upper 95% confidence limit = 653.8). The random effect of alleles was kept in the Model (2) because its removal would have caused a considerable drop in the explanatory power of this model ($L = 11.7$, $P < 0.001$; AIC increase from 883.5 to 893.2). Visual examination of the distribution of residuals produced by Model (2) showed no considerable deviation from normality and violation of the homogeneity of variance, suggesting that the model was valid (see Supplementary material online, *Figure S2* and *S3*).

Mean FS in congenic background FVB/N (73.0 ± 4.8%) differed significantly from FS of any other background (49.8 ± 19.6% in C57BL/6, 57.0 ± 20.3% in 129S1/Sv, and 57.9 ± 15.9% in C3H/He). No differences were found in any of the pairwise comparisons between the 129S1/Sv, C57BL/6 and C3H/He strains (*Figure 2A*). By allele, mean FS in Sgcd-deficient animals (71.1 ± 9.4%) significantly differed from that in Myoz2−(52.0 ± 17.6%) and in Csrp2-deficient mice (44.9 ± 19.3%). No difference was observed between Myoz2 and Csrp2 mutants (*Figure 2B*).

In summary, the observed FS values were strongly associated with strain background and mutant allele.

3.2 Correlation analysis between gene expression and ventricular function

To test the association between gene expression and ventricular (dys)function, we calculated Spearman correlation coefficients between the FS measurements and the $\log_2$-expression values of each microarray gene transcript. The correlation analysis was carried out separately for male and female mice because sex is known to be an important confounder at the level of gene expression.[20]

In male mice, we identified 181 transcripts (corresponding to 159 genes) that were highly significantly correlated with FS ($r_s \geq |0.70|$, FDR-adjusted $P < 0.001$; see Supplementary material online, *Table S1*). Most of these transcripts (74%) were negatively correlated with FS, i.e. their expression increased with ventricular dysfunction (*Figure 3*). Based on the expression levels of these transcripts, males could be divided into two distinctive equally sized groups according to their degree of ventricular impairment (see Supplementary material online, *Figure S4*). One of the groups had a mean FS of 72.4 ± 9.2%, and was composed of all the Sgcd mutants and the FVB.Myoz2 mice. The other group had a mean FS of 42.6 ± 17.8% and included all other animals.

In female mice, we found 184 transcripts (168 genes) highly correlated with the cardiac impairment ($r_s \geq |0.70|$, FDR-adjusted $P <$ 0.001). As in male mice, most of the transcripts (66%) were negatively correlated with FS (see Supplementary material online, *Table S1*).

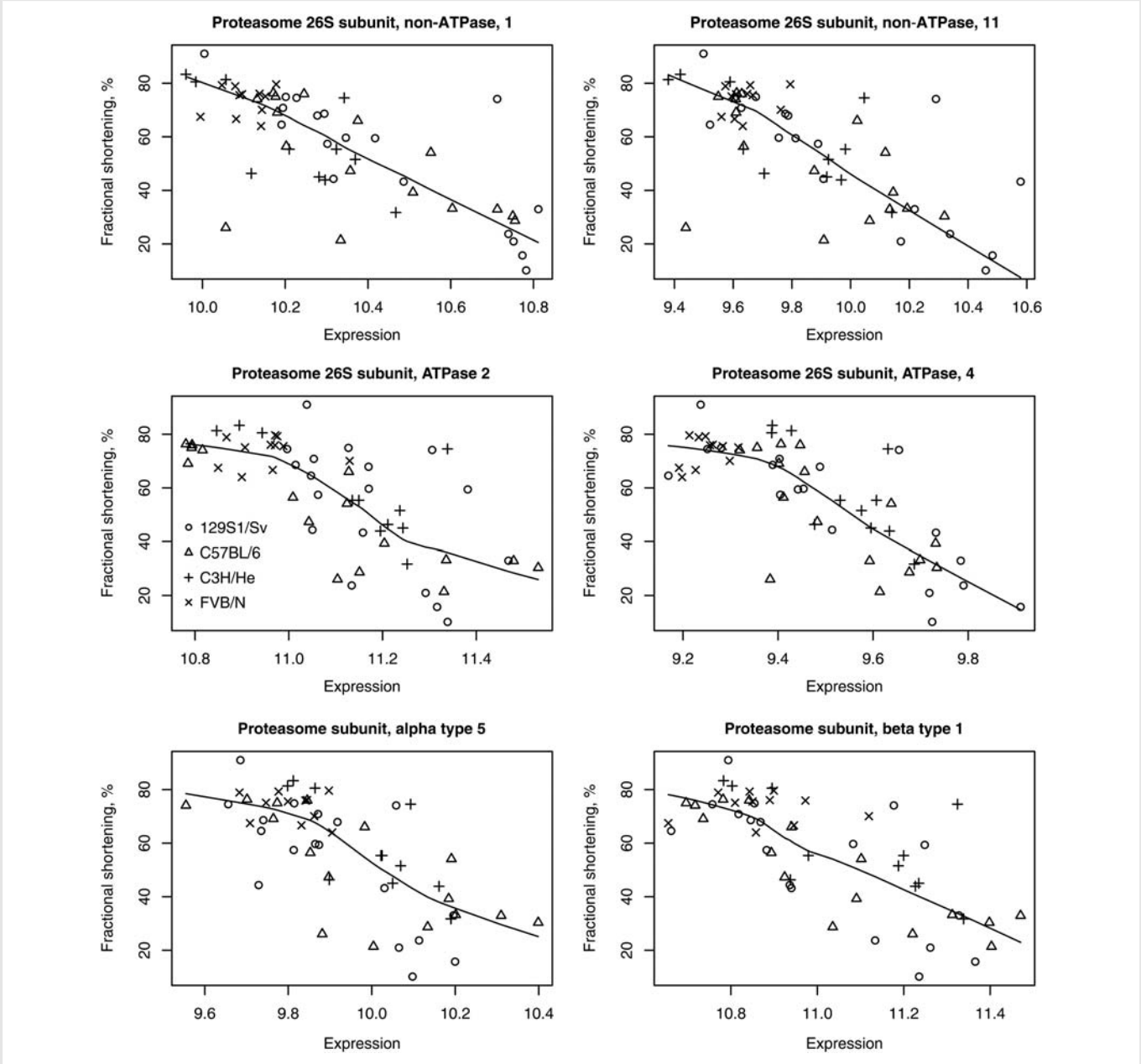

Figure 3 An example of six genes encoding the members of the proteasome complex in male animals. Their expression levels were significantly negatively correlated with FS in the examined mouse strains. Smoothing lines were added to each panel to highlight the patterns.

Based on the expression values of these transcripts, all female mice could similarly be divided into two groups according to their degree of ventricular dysfunction (see Supplementary material online, *Figure S5*). One of these groups had a mean FS of 71.9 $\pm$ 5.5% and was composed of FVB.Myoz2 animals and all Sgcd mutants, except for one B6.Sgcd animal that belonged to the second group showing an unusually low FS (43.3%). The second group had a mean FS of 49.2 $\pm$ 16.2% and included all other animals.

In total, 280 genes were highly correlated with FS in the examined cardiomyopathy models (see Supplementary material online, *Table S1*). Of these, only 48 genes (17%) were shared by males and females (see Supplementary material online, *Table S1*). All shared transcripts had the same sign of correlation in males and females, with the majority (90%) being negatively correlated with FS. Thus, most of these genes were overexpressed in animals with severe heart dysfunction.

3.3 Functional classification of genes correlated with FS

We used the gene classification tool in DAVID to organize the lists of genes correlated with FS into functionally related groups and to identify their major biological functions.

Two functionally related clusters were discovered among the 159 genes that were highly correlated with FS in male mice (see Supplementary material online, *Table S2*). The first cluster was composed of five members involved in metal ion binding. The second cluster listed 16 genes encoding transmembrane proteins and glycoproteins.

Table 2 Pathways found to be represented by the FS-correlated genes

Pathway (enrichment *P*-value; fold enrichment)	Gene names	Correlation with FS
Males		
Proteasome ($P < 0.001$; 16.4)	Proteasome 26S subunit, non-ATPase, 11	−0.708
	Proteasome 26S subunit, ATPase 2	−0.725
	Proteasome 26S subunit, ATPase, 4	−0.735
	Proteasome 26S subunit, non-ATPase, 1	−0.745
	Proteasome subunit, alpha type 5	−0.714
	Proteasome subunit, beta type 1	−0.728
	Proteasome subunit, beta type 4	−0.760
Hypertrophic cardiomyopathy ($P = 0.038$; 5.3); dilated cardiomyopathy ($P = 0.048$; 4.8)	Calcium channel, voltage-dependent, Alpha2/delta subunit 1	−0.768
	Dystrophin, muscular dystrophy	−0.761
	Myosin, heavy polypeptide 7, cardiac muscle, beta	−0.738
	Transforming growth factor, beta 2	−0.738, −0.728[a]
ABC transporters ($P = 0.06$; 1.9)	ATP-binding cassette, sub-family B, member 4	−0.730
	ATP-binding cassette, sub-family C, member 1	−0.756
	ATP-binding cassette, sub-family G, member 2	−0.726
Focal adhesion ($P = 0.098$; 3.1)	RIKEN cDNA 2900073G15 gene	−0.703
	Cyclin D2	−0.726, −0.702, −0.711
	Laminin, gamma 2	−0.719
	Platelet-derived growth factor, C polypeptide	−0.769, −0.768
	B-cell leukaemia/lymphoma 2	−0.738
Females		
Tyrosine metabolism ($P = 0.047$; 8.5)	Dopamine beta hydroxylase	0.725
	Monoamine oxidase B	0.700
	Phenylethanolamine-*N*-Methyltransferase	0.747
Proteasome ($P = 0.065$; 7.0)	Proteasome 26S subunit, ATPase, 4	−0.726
	Proteasome 26S subunit, non-ATPase, 1	−0.727
	Proteasome 26S subunit, non-ATPase, 13	−0.744
Amyotrophic lateral sclerosis ($P = 0.091$; 5.8)	BCL2-associated agonist of cell death	−0.766
	Glutathione peroxidase 1	−0.799
	Mitogen-activated protein kinase 12	−0.754
Both genders		
Proteasome ($P = 0.064$; 27.1)	Proteasome 26S subunit, ATPase, 4	See above
	Proteasome 26S subunit, non-ATPase, 1	See above

[a]Multiple correlation coefficients are given for genes that were represented by more than one probe on the microarray chip.

In female mice, 26 of the 168 FS-correlated genes were allocated to 4 functional clusters (see Supplementary material online, *Table S2*) involved in metal ion binding (in particular, potassium and calcium channels), transmembrane transport, proteinase activity, and ribonucleotide binding. These functional clusters were composed of 4–14 members. In both male and female mice, the highest enrichment scores were observed in the gene groups encoding for metal ion-binding proteins (e.g. zinc finger proteins) and transmembrane transporters related to electrolyte homeostasis and signal transduction (e.g. solute carriers) (see Supplementary material online, *Table S2*).

Among the genes that were correlated with FS and were shared by both genders, three functionally related clusters were identified, with three to five members each. These genes encoded for leucin-rich repeats, hormone-like proteins, and transmembrane proteins (see Supplementary material online, *Table S2*).

3.4 Pathway analysis of genes correlated with FS

To determine whether FS-correlated genes were enriched with representatives of specific metabolic pathways, we attributed KEGG annotation terms to these genes using the functional annotation tool of DAVID.

In male mice, members of the following five pathways were found to be enriched: proteasome, HCM, DCM, ABC transporters, and focal adhesion proteins. Expression levels of all genes mapped on these pathways were negatively correlated with FS (*Table 2*, male mice). The proteasome pathway exhibited the most significant enrichment *P*-value (<0.001) and the highest fold enrichment (16.4), suggesting a strong association between this pathway and cardiomyopathy-related impairment of ventricular function. The detection of the HCM and DCM pathways, which were represented

by the same four genes (*Table 2*, male mice), was not surprising and reflected the fact that the experimental animals suffered from genetic cardiomyopathy. ABC transporters and focal adhesion pathways showed less significant enrichment *P*-values but relatively high (>1.5) fold enrichment values, suggesting that their activation does play a role in cardiomyopathy.

In female mice, FS-correlated genes were enriched with representatives of three pathways, namely tyrosine metabolism, proteasome, and amyotrophic lateral sclerosis (*Table 2*, female mice). All genes of the proteasome and the ALS pathways were negatively correlated with FS. At the same time, the tyrosine metabolism genes were positively related to FS, suggesting a deactivation of this pathway in animals with contractile dysfunction. The tyrosine metabolism had the most significant enrichment *P*-value (0.047) and the highest fold enrichment (8.5), suggesting a specific, yet unknown role for female cardiomyopathy phenotypes.

Only the proteasome pathway was enriched in the list of the FS-correlated genes shared by males and females (*Table 2*, both genders). As the proteasome metabolism was also found to be overrepresented in separate analyses of the male and female mice, the activation of this pathway seems to play a particularly important role in the response to genetic cardiomyopathy, irrespective of gender.

4. Discussion

Many authors study gene expression patterns in a murine knock-out model in comparison with appropriate wild-type litter mate controls with the aim of understanding the mechanisms of genetic cardiomyopathies. However, not uncommonly, these mouse models may exhibit an attenuated or no disease phenotype at all.[4] This may be explained by a very weak pathogenic effect of the gene under study and/or attenuation of penetrance caused by modifier genes in the genetic background of the strain in which the gene under study was inactivated. Although this critical issue can only be resolved by studying the effects of the defective alleles in several strain backgrounds, this is rarely done because it is very cumbersome to produce such disease-sensitized congenic strains. Our study aimed at identifying general genetic modifiers of cardiomyopathy rather than individual cardiomyopathy gene effects. Therefore, we transferred well-known cardiomyopathy genes into four common strain backgrounds to produce the largest group of disease-sensitized congenic cardiomyopathy models published thus far.

At the level of gene expression, one may expect a very complex picture of specific, disease-causing gene fingerprints and compensatory gene expression in response to genetic inactivation of a cardiomyopathy gene. Attempting to identify major genetic modifiers without prior assumptions on candidate genes, our approach was rather 'statistical': by comparing many cardiomyopathy models, each on several common inbred backgrounds, we aimed to identify common gene expression patterns shared by the majority of the diverse models. We hypothesized that these common patterns would point to non-specific genetic modifiers of the cardiomyopathy phenotype, whereas gene expression patterns unique to one model would more likely translate the inactivation of a specific cardiomyopathy gene.

Not unexpectedly, mean levels of FS significantly varied among the examined cardiomyopathy models. Use of the linear mixed-effects modelling allowed us to identify strain background as the major factor that determined these differences. For example, mutants on an FVB/N background exhibited normal ventricular pump function with little variance, suggesting that strong compensatory mechanisms existed in this strain background. Notably, FVB/N is a frequently used strain for transgenic gene addition experiments. Also, FS was significantly associated with the type of mutant alleles. Mean FS in Sgcd-deficient mice was consistent with a normal ventricular pump function and exhibited only little variance. This may be related to the finding that the cardiomyopathy showed rather a hypertrophic than a dilative phenotype in these mice.[20]

We identified a total of 280 genes whose expression was highly significantly associated with the level of ventricular dysfunction in the animal models examined. In line with previous findings,[20] we also observed gender-specific gene activation patterns. Most of the identified significant genes were overexpressed in animals with low FS, suggesting that compensatory pathways were activated. In male mice, these genes were enriched with members of five different pathways, with the proteasome pathway showing the highest enrichment. In female mice, three pathways were found, with the tyrosine pathway being most enriched. The administration of anti-oestrogens to the mice may help us to confirm that the observed gender-specific gene enrichments were indeed part of a sexual dimorphism. In contrast, only the proteasome pathway was activated in both males and females.

The ubiquitin-proteasome system (UPS) has already received attention as a potentially important determinant of cardiac dysfunction.[23–25] An increased expression of the proteasome subunits was previously reported in animal models of hypertrophic and dilated cardiomyopathies[26,27] and even in cardiomyopathy patients.[28,29] To explain these findings, the authors hypothesized that impaired pump function produced overload-related cardiac wall stress that may have stimulated the synthesis of excessive amounts of damaged myocardial proteins. These proteins had to be rapidly degraded with the help of the UPS to avoid the activation of pro-apoptotic signals.[24,26] Preferential degradation of damaged proteins may have, in turn, competitively inhibited UPS proteolytic capacity for other substrates, with the consequence being hypertrophy.[30] Indeed, impaired proteolytic capacity of the UPS, resulting in accumulation of large amounts of ubiquitinated but not degraded proteins, was repeatedly observed in patients and animals with end-stage heart failure.[28,29,31] In addition to the hypothesis of competitive inhibition of UPS activity, Predmore *et al.*[25] provided evidence that the UPS may become dysfunctional due to oxidative post-translational modifications of the proteasome itself. Interestingly, another line of evidence showed that pharmacological inhibition of the proteasome prevented and even reversed (mal)adaptive myocardial hypertrophy and remodelling.[24] In summary, accumulating evidence suggests a certain role of the UPS in cardiomyopathy.[32] However, it remains controversial whether activation or inhibition is favourable for myocardial function.

Additional functionally related genes were correlated with FS. Their biological functions were associated mainly with metal ion binding, transmembrane proteins, proteinase activity, ribonucleotide binding, and signal transduction. These findings emphasize the complexity of the development and progression of cardiomyopathy and are in agreement with a number of studies. For example, a group of four protein kinases (see Supplementary material online, *Table S2*) correlated with FS in female mice. This group included the mitogen-activated protein kinases Map3k7 and Mapk12, known transducers of extracellular stress signals. Map3k7 mediates signal transduction induced by the transforming growth factor-beta, which has been

implicated in cardiac fibrosis in mice with mutant sarcomere proteins and HCM.[33] As a result, we observed an increased expression of both Map3k7 and the transforming growth factor beta-regulated gene 1 (Tbrg1) in females with severe ventricular dysfunction (see Supplementary material online, *Table S1*). Another example is the atrial natriuretic peptide, a circulating peptide hormone of cardiac origin that is known to compensate for heart failure by its diuretic, natriuretic, and vasodilatory actions and inhibitory effects on renin and aldosterone secretion.[34] Schoensiegel *et al.*[20] have previously reported increased concentrations of the atrial natriuretic peptide (Nt-proANP) in a subset of congenic strains with heart failure. Our study confirmed the high expression levels of the natriuretic peptide precursor type A gene (Nppa) in both males and females with ventricular dysfunction (see Supplementary material online, *Table S1*).

In conclusion, we examined the largest existing collection of disease-sensitized congenic mouse models for cardiomyopathy, allowing the comparison of mutant alleles on several strain backgrounds. We contributed additional evidence implicating a potential role for UPS as a major modifier of the cardiomyopathy phenotype. Elements of the UPS may be pharmacologically targeted in the future to clarify whether their activation or inhibition can improve cardiac function in genetic cardiomyopathy. In contrast to previous reports, our findings resulted from an unbiased, genome-wide analysis of gene expression.

The aim of our work was to generate hypotheses rather than studying candidate gene pathways or confirming previous reports of other groups. Therefore, our study had the limitations of any report on whole-genome analysis of gene expression: a focus on bioinformatics rather than wet lab evidence of molecular mechanisms. Our conclusion is based on a statistical association analysis that cannot differentiate whether myocardial overexpression of the UPS genes is the cause or effect of the impaired pump function in genetic cardiomyopathy. Investigations at repeated time points during the development of cardiomyopathy may help answer this question. Future independent studies should address the UPS activity and protein levels as well as the structure of the 26S proteasome. The determination of oxidized and/or ubiquitinated myocardial proteins should be included in such studies as well as a sophisticated characterization of the new congenic models, e.g. *in vivo* pressure/volume loops and perfusion of isolated hearts.

Supplementary material

Supplementary material is available at *Cardiovascular Research* online.

Acknowledgement

We thank Mrs. Sybille Menges-Wirth (Division of Cardiology, University Hospital Heidelberg) for her excellent technical assistance.

Conflict of interest: none declared.

Funding

This work was mainly supported by the National Genome Research Network (NGFN) funded by the Bundesministerium für Bildung und Forschung (BMBF) [grant numbers 01GS0420 (project NHK-S19T13) to B.I. and 01GR0438 (project PMM-S19T18) to B.I.].

References

1. Fatkin D, Otway R, Richmond Z. Genetics of dilated cardiomyopathy. *Heart Fail Clin* 2010;**6**:129–140.
2. Maron BJ, McKenna WJ, Danielson GK, Kappenberger LJ, Kuhn HJ, Seidman CE *et al.* American College of Cardiology; Committee for Practice Guidelines. American College of Cardiology/European Society of Cardiology clinical expert consensus document on hypertrophic cardiomyopathy. A report of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents and the European Society of Cardiology Committee for Practice Guidelines. European Society of Cardiology. *J Am Coll Cardiol* 2003;**42**:1687–1713.
3. Doetschman T. Influence of genetic background on genetically engineered mouse phenotypes. *Methods Mol Biol* 2009;**530**:423–433.
4. Montagutelli X. Effect of the genetic background on the phenotype of mouse mutations. *J Am Soc Nephrol* 2000;**11**:S101–S105.
5. Suzuki M, Carlson KM, Marchuk DA, Rockman HA. Genetic modifier loci affecting survival and cardiac function in murine dilated cardiomyopathy. *Circulation* 2002; **105**:1824–1829.
6. Hack AA, Lam MY, Cordier L, Shoturma DI, Ly CT, Hadhazy MA *et al.* Differential requirement for individual sarcoglycans and dystrophin in the assembly and function of the dystrophin-glycoprotein complex. *J Cell Sci* 2000;**113**:2535–2544.
7. Arber S, Hunter JJ, Ross J Jr, Hongo M, Sansig G, Borg J *et al.* MLP-deficient mice exhibit a disruption of cardiac cytoarchitectural organization, dilated cardiomyopathy, and heart failure. *Cell* 1997;**88**:393–403.
8. Frey N, Barrientos T, Shelton JM, Frank D, Rütten H, Gehring D *et al.* Mice lacking calsarcin-1 are sensitized to calcineurin signaling and show accelerated cardiomyopathy in response to pathological biomechanical stress. *Nat Med* 2004;**10**:1336–1343.
9. Wong GT. Speed congenics: applications for transgenic and knock-out mouse strains. *Neuropeptides* 2002;**36**:230–236.
10. Schoensiegel F, Ivandic B, Geis NA, Schrewe A, Katus HA, Bekeredjian R. High throughput echocardiography in conscious mice: training and primary screens. *Ultraschall Med* 2011;**32**:S124–S129.
11. Sahn DJ, DeMaria A, Kisslo J, Weyman A. Recommendations regarding quantitation in M-mode echocardiography: results of a survey of echocardiographic measurements. *Circulation* 1978;**58**:1072–1083.
12. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2010. Available at: http://www.R-project.org
13. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy–analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;**20**:S307–S315.
14. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to quantification of differential expression. *Bioinformatics* 2002;**18**:S96–S104.
15. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Gen Biol* 2004;**5**:R80.
16. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Bioinformatics* 2003;**4**:S249–S264.
17. Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S eds. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York: Springer, 2005.
18. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer-Verlag, 2000.
19. Zar HJ. *Biostatistical Analysis*. 3rd ed. New Jersey: Prentice Hall, 1996.
20. Schoensiegel F, Bekeredjian R, Schrewe A, Weichenhan D, Frey N, Katus HA *et al.* Atrial natriuretic peptide and osteopontin are useful markers of cardiac disorders in mice. *Comparative Med* 2007;**57**:546–553.
21. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 1995;**57**:289–300.
22. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**:44–57.
23. Willis MS, Patterson C. Into the heart: the emerging role of the ubiquitin-proteasome system. *J Mol Cell Cardiol* 2006;**41**:567–579.
24. Hedhli N, Depre C. Proteasome inhibitors and cardiac cell growth. *Cardiovasc Res* 2010;**85**:321–329.
25. Predmore JM, Wang P, Davis F, Bartlone S, Westfall MV, Dyke DB *et al.* Ubiquitin proteasome dysfunction in human hypertrophic and dilated cardiomyopathies. *Circulation* 2010;**121**:997–1004.
26. Depre C, Wang Q, Yan L, Hedhli N, Peter P, Chen L *et al.* Activation of the cardiac proteasome during pressure overload promotes ventricular hypetrophy. *Circulation* 2006;**114**:1821–1828.
27. Hedhli N, Wang L, Wang Q, Rashed E, Tian Y, Sui X *et al.* Proteasome activation during cardiac hypertrophy by the chaperone H11 Kinase/Hsp22. *Cardiovasc Res* 2008;**77**:497–505.
28. Hein S, Arnon E, Kostin S, Schönburg M, Elsässer A, Polyakova V *et al.* Progression from compensated hypertrophy to failure in the pressure-overloaded human heart: structural deterioration and compensatory mechanisms. *Circulation* 2003;**107**: 984–991.
29. Weeks J, Morrison K, Mullen A, Wait R, Barton P, Dunn MJ. Hyperubiquitination of proteins in dilated cardiomyopathy. *Proteomics* 2003;**3**:208–216.

30. Sarikas A, Carrier L, Schenke C, Doll D, Flavigny J, Lindenberg KS *et al*. Impairment of the ubiquitin-proteasome system by truncated cardiac myosin binding protein. *Cardiovasc Res* 2005;**66**:33–44.
31. Tsukamoto O, Minamino T, Okada K, Shintani Y, Takashima S, Kato H *et al*. Depression of proteasome activities during the progression of cardiac dysfunction in pressure-overloaded heart of mice. *Biochem Biophys Res Commun* 2006;**340**: 1125–1133.
32. Schlossarek S, Carrier L. The ubiquitin-proteasome system in cardiomyopathies. *Curr Opin Cardiol* 2011;**26**:190–195.
33. Teekakirikul P, Eminaga S, Toka O, Alcalai R, Wang L, Wakimoto H *et al*. Cardiac fibrosis in mice with hypertrophic cardiomyopathy is mediated by non-myocyte proliferation and requires Tgf-β. *J Clin Invest* 2010;**120**:3520–3529.
34. Nishikimi T, Maeda N, Matsuoka H. The role of natriuretic peptides in cardioprotection. *Cardiovasc Res* 2006;**69**:318–328.