Identification of Novel Biomarker Candidates for Hypertrophic Cardiomyopathy and Other Cardiovascular Diseases Leading to Heart Failure

H. REHULKOVA[1], P. REHULKA[1], A. MYSLIVCOVA FUCIKOVA[1], J. STULIK[1], R. PUDIL[2]

[1]Department of Molecular Pathology and Biology, Faculty of Military Health Sciences, University of Defence, Hradec Kralove, Czech Republic, [2]1st Department of Internal Medicine – Cardioangiology, Faculty Hospital, Hradec Kralove, Czech Republic

Summary

In-depth proteome discovery analysis represents new strategy in an effort to identify novel reliable specific protein markers for hypertrophic cardiomyopathy and other life threatening cardiovascular diseases. To systematically identify novel protein biomarkers of cardiovascular diseases with high mortality we employed an isobaric tag for relative and absolute quantitation (iTRAQ) proteome technology to make comparative analysis of plasma samples obtained from patients suffering from non-obstructive hypertrophic cardiomyopathy, stable dilated cardiomyopathy, aortic valve stenosis, chronic stable coronary artery disease and stable arterial hypertension. We found 128 plasma proteins whose abundances were uniquely regulated among the analyzed cardiovascular pathologies. 49 of them have not been described yet. Additionally, application of statistical exploratory analyses of the measured protein profiles indicated the relationship in pathophysiology of the examined cardiovascular pathologies.

Corresponding author

J. Stulik, Department of Molecular Pathology and Biology, Faculty of Military Health Sciences, University of Defence, Trebesska 1575, 500 01 Hradec Kralove, Czech Republic. E-mail: jiri.stulik@unob.cz

Introduction

Heart failure (HF) is not a uniform syndrome. It encompasses various causes as cardiomyopathy, congenital heart diseases, hypertension and ischemic heart disease. The molecular basis of this entity is not completely understood (Yang *et al.* 2015). Hypertrophic cardiomyopathy (HCM), one of the pathologies contributing to HF, has been defined morphologically by unexplained hypertrophy in the absence of haemodynamic stress, and at the histological level by myocyte disarray, fibrosis, and abnormalities of the intramyocardial small vessels. HCM is a monogenic cardiac disease with an autosomal dominant pattern of heritability and with prevalence in the general population of 1/500. More than 1,400 different mutations have been discovered so far, involving mutations in 11 genes of contractile sarcomeric proteins that have been shown to produce the disease (Hensley *et al.* 2015). Asymmetrical septal hypertrophy leads to a significant pressure gradient between the apical left ventricular (LV) chamber and the left ventricular outflow tract (LVOT), resulting in the so-called 'hypertrophic obstructive cardiomyopathy' (obstructive HCM) presentation (Hensley *et al.* 2015). The prevalence of LVOT obstruction in HCM is 20-30 % at rest and up to 70 % with provocation (Cooper *et al.* 2015). Majority of patients with obstructive HCM are asymptomatic but some can present with heart failure or arrhythmias. Just arrhythmias, premature sudden cardiac

deaths, heart failure and less frequently thromboembolic events are primary causes of mortality amongst HCM patients (Kumar *et al.* 2015). Ultimately, about 10 % of obstructive HCM patients progress to an end-stage dilated phase with LV wall thinning, cavity enlargement, and systolic dysfunction that resembles a dilated cardiomyopathy (DCM) (Xiao *et al.* 2015).

The classification of heart failure is traditionally based on the pathological cause of failure of the cardiac pump, the pathophysiological characteristics, and the acuity and severity of the heart failure. There is general assumption that identification of biomarker profiles of pathologies threatening by heart failure can provide additional criteria for risk stratification among HF patients (Braunwald 2008). From this point of view plasma is very useful clinical specimen for the identification of novel markers of cardiovascular diseases. Some plasma proteins like proteins of coagulation cascade, proteins of lipid transfer, proteins interacting with vessel walls or platelets and, finally, inflammatory proteins were found to be directly involved pathogenesis of the cardiovascular system (Anderson 2005). Plasma also represents the largest proteome of the human body exhibiting extremely dynamic ranges in protein concentration. This makes its quantitative analysis highly challenging. Nevertheless, current advanced mass spectrometry approaches preceded with different pre-fractionation techniques enable to quantitate several thousand plasma proteins in one run (Keshishian *et al.* 2015).

In our study we employed isobaric tag for relative and absolute quantitation (iTRAQ) technology labeling of peptides (Wiese *et al.* 2007) for comparison of plasma protein profiles of patients with HCM and other diseases of the cardiovascular system, which are accompanied with structural changes of the myocardium and can lead to its failure.

Methods

Study population

Study population consisted of five groups: patients with non-obstructive HCM, patients with aortic valve stenosis (AS), chronic stable coronary artery disease (CAD), stable arterial hypertension (AH), and stable DCM. The diagnosis of HCM was based on a history of illness, physical examination, echocardiography, and cardiac catheterization in accordance with European Society of Cardiology

recommendations (Lang *et al.* 2015). The term "stable" means that patient is clinically stable at least for one month before the enrollment to the study.

Patients were enrolled into the study consecutively. There was no selection bias according to age, gender and medication. However, the patients with significant concomitant disease, such as pulmonary disease, malignancy, autoimmune disorders, neurodegenerative disorders, thyroid disease, or concurrent viral disease were excluded. Baseline characteristics of the population of this study are shown in Table 1.

Results were compared to control group of healthy 46 blood donors aged 55 to 67 years (median age 60 years) with no evidence of cardiovascular disease or other severe disorders. Statistical analysis of the age differences showed no significant differences between all studied groups with one exception of aortic stenosis. There the median of age was significantly higher than in the other groups, because of the higher prevalence of degenerative etiology aortic stenosis.

The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by the ethical committee of our institution. Informed consent was obtained from each patient.

Plasma sample collection

Blood samples (before echocardiography and cardiac catheterization) were obtained from left forearm after an overnight fasting at the First Department of Internal Medicine of the University Hospital in Hradec Kralove. The control samples were obtained from healthy donors at the Transfusion Department of the same hospital. The BD P100v1.1 vacuum tubes containing protease inhibitors and EDTA anticoagulant were used (Becton Dickinson, Franklin Lakes, NJ). Within one hour, the blood samples were centrifuged for 20 min at 2500 g at 22 °C and the supernatant was removed, aliquoted and kept at -80 °C until a further analysis was performed.

Plasma immunoaffinity enrichment

An equal volume was taken from each patient sample to create pooled aliquots of each group, all in duplicate. These pooled aliquots were depleted from fourteen most abundant plasma proteins (albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, alpha-2-macroglobulin, alpha-1-acid glycoprotein, IgM, apolipoprotein A1, apolipoprotein A2, complement C3, and transthyretin) using HPLC immunoaffinity column

MARS Hu-14 (Multiple Affinity Removal System Agilent Technologies, Santa Clara, CA) according to manufacturer´s instructions. Obtained depleted fractions of plasma proteins were then concentrated and desalted using 3K Amicon Ultra centrifugal filter units (Millipore Corporation, Bedford, MA).

Table 1. Baseline characteristics of patients groups in the study.

	HCM (n=47)	DCM (n=21)	CAD (n=20)	AS (n=24)	AH (n=20)
Age (years)	58.4±12.4	57.6±12.0	62.8±9.2	70.6±11.6	54.7±15.8
Female gender, n (%)	12 (26)	4 (19)	3 (15)	6 (25)	7 (35)
NYHA functional class, n (%)					
I	21 (45)	4 (19)	13 (65)	3 (13)	20 (100)
II	18 (38)	14 (67)	5 (25)	15 (63)	0
III-IV	8 (17)	3 (14)	2 (10)	6 (25)	0
Atrial fibrillation, n (%)	14 (30)	6 (29)	3 (15)	6 (25)	2 (10)
Creatinine level (µmol.l^{-1})	94.0±29.8	88.3±18.0	89.7±29.0	109.9±44.7	91.7±18.9
Diabetes mellitus, n (%)	13 (28)	4 (19)	4 (20)	9 (38)	4 (20)
Hyperlipidemia, n (%)	23 (49)	9 (43)	19 (95)	15 (63)	13 (65)
Smoking, n (%)	20 (43)	12 (57)	15 (75)	9 (38)	9 (45)
Echocardiography					
LA (mm)	48.2±7.2	42.1±12.3	43.1±7.4	44.0±6.0	42.7±2.9
RV (mm)	26.1±3.8	26.5±7.7	26.8±2.8	26.7±3.4	27.2±2.4
IVST (mm)	19.4±4.4	10.4±2.6	11.3±1.9	13.9±2.1	12.1±1.3
LV ESD (mm)	31.9±7.1	50.3±15.8	45.5±11.0	38.1±7.1	35.7±3.8
LV EDD (mm)	47.1±7.3	58.2±16.0	57.7±8.8	51.9±4.7	52.3±3.3
PWT (mm)	13.2±2.7	10.2±0.9	10.9±1.0	12.6±1.5	11.1±1.3
LV EF (%)	67.1±9.9	29.2±13.7	45.1±15.8	55.5±14.0	64.2±6.6
FS	33.7±8.7	14.3±7.2	22.7±8.7	28.1±9.3	31.7±6.1
TAPSE (mm)	22.7±3.2	17.9±6.7	23.7±4.1	21.1±3.9	24.2±2.5
LVM (g)	344.8±129.9	279.6±70.8	266.1±50.9	291.0±62.7	240.9±34.7
LVMI (g.m^{-2})	171.4±60.2	141.2±33.0	127.8±24.9	150.6±31.4	115.7±15.5
Therapy **Medication, n (%)**					
Calcium channel blocker	22 (47)	1 (5)	2 (10)	0	11 (55)
B-blocker	30 (64)	19 (90)	20 (100)	22 (92)	10 (50)
Diuretic	20 (43)	16 (76)	7 (35)	21 (88)	9 (45)
Acetylosalicylic acid	7 (15)	5 (24)	20 (100)	12 (50)	3 (15)
ACE inhibitors/sartans	23 (49)	10 (48)	18 (90)	13 (54)	16 (80)
Implantable devices, n (%)					
DDD pacing	7 (15)	1 (5)	1 (5)	0	0
ICD/BiV	5 (11)	11 (52)	3 (15)	0	0

LA = left atrium, RV = right ventricle, IVST = interventricular septum thickness, LV ESD = left ventricular end-systolic diameter, LV EDD = left ventricle end-diastolic diameter, PWT = posterior wall thickness, FS = fractional shortening, TAPSE = tricuspid annulus plane systolic excursion, LVM = left ventricle mass, LVMI = left ventricle mass index, ACE = angiotensin converting enzyme, DDD = atrial and ventricular pacemaker, ICD/BiV = implantable cardioverter-defibrillator/biventricular pacing.

iTRAQ labeling of plasma samples

Aliquots of each sample with known protein concentration were reduced, cystein-blocked, digested with trypsin in the presence of urea, and resulted tryptic peptides labeled with iTRAQ tags according to the manufacturer´s instructions (AB Sciex, Foster City, USA). In our analysis setup two iTRAQ 8plex kits were used. HCM pooled sample aliquots were labeled with reagents 113 and 117 in both kits, control sample aliquots were labeled with reagent 114 and 118 in both kits as well. Furthermore, DCM and AS pooled sample aliquots were labeled each with two reagents in the first kit (115 and 119 for DCM; 116 and 121 for AS). Likewise, CAD and AH pooled sample aliquots were labeled with reagents in the second kit (115 and 119 for CAD; 116 and 121 for AH). Digested and labeled samples were combined together and peptides were fractionated into 22 fractions using reversed phase liquid chromatography separation with basic mobile phases at Waters 2695 Separations Module instrument (Waters, Milford, USA) using HPLC column Gemini C18 2.0 mm ID x 150 mm L, 3 µm (Phenomenex, Darmstadt, Germany) and cartridge Gemini C18 2.0 mm ID x 4.0 mm L (Phenomenex, Darmstadt, Germany). Mobile phase A consisted of deionized water (ELGA PURELAB Classic, Veolia Water Systems Ltd., Bucks, UK), mobile phase B was acetonitrile (gradient grade for LC, Merck, Darmstadt, Germany) and mobile phase C was represented by 200 mM ammonium formate (Sigma-Aldrich, Schnelldorf, Germany) buffered at pH 10.0 with concentrated formic acid (Sigma-Aldrich, Schnelldorf, Germany). The gradient was: 0 min, 85 % mobile phase A, 5 % mobile phase B; 2 min, 85 % A, 5 % B; 64 min, 35 % A, 55 % B; 65 min, 10 % A, 80 % B; 70 min, 10 % A, 80 % B; 71 min, 85 % A, 5 % B; 88 min, 85 % A, 5 % B; mobile phase C was kept constant at 10 % all the time. The flow was 160 µl min^{-1} and the two-minute fractions were collected from 8 min to 44 min. The collected fractions were acidified with 10 % formic acid and evaporated in a vacuum centrifuge (Concentrator 5301, Eppendorf, Germany).

LC-MS/MS analysis

The obtained fractions of iTRAQ labeled peptides were dissolved in 20 µl of 2 % acetonitrile (v/v)/0.05 % trifluoroacetic acid (v/v). The samples were loaded onto a Nano Trap precolumn (3 µm Acclaim PepMap100, C18, 75 µm x 20 mm, Dionex Corporation, Sunnyvale, USA) for preconcentration and then separated on the nano RP column (2 µm Acclaim PepMap RSLC, C18, 75 µm x 150 mm, Dionex Corporation, Sunnyvale, USA) connected in UltiMate 3000 RSLC-nano HPLC system (Dionex Corporation, Sunnyvale, USA). High purity reagents were used for nanoLC-MS/MS analysis: water (LC-MS Chromasolv, Fluka, Switzerland), acetonitrile (hypergrade for LC-MS, Merck, Darmstadt, Germany) and formic acid (additive eluent for LC-MS, Fluka, Germany). Mobile phase A was 0.1 % formic acid (v/v) and mobile phase B consisted of 80 % acetonitrile/0.1 % formic acid (v/v). The flow rate of the loading solvent was 200 µl min^{-1}; a constant flow rate of 300 nl min^{-1} was held during the separation on pre-concentration and separation column. The gradient was programmed as follows: 0 min, 4 % mobile phase B; 5.1 min, 4 % B; 53 min, 34 % B; 63 min, 55 % B; 64 min, 90 % B; 72 min, 90 % B; 73 min, 4 % B and 85 min, 4 % B. The eluent was directly sprayed using on-line coupled Nanospray Flex Ion source (Thermo Fisher Scientific, Waltham, USA) into Q Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, USA) which was run in a positive ion mode and set for data-dependent analysis with full MS scan (350-1,650 m/z) at resolution 70,000 and MS/MS resolution 17,500. The scanning run time was 0-67 min, isolation width of precursors 1.6 m/z and fixed first mass 100.0 m/z. The first 10 top peaks were fragmented with dynamic exclusion of 30 s and the intensity threshold was set to 50,000.

Protein identification

The raw MS/MS data were submitted to Proteome Discoverer 2.0.0.802 software (Thermo Fisher Scientific Inc.). The processing workflow consisted of extracting the quantification information from reporter ions, selection of MS/MS spectra (charge state – 2+ to 6+, precursor mass – 700 to 5,000 Da, total intensity threshold – 1,000, minimum peak count – 7), noise and top N (6 peaks per 100 Da) peak filtering, Mascot database searching and Percolator validation (strict target FDR – 0.01) of the peptide spectrum matches. The parameters for Mascot database searching were as follows: instrument – ESI-TRAP, protein database – UniProt human reference proteome UP000005640 (90434 entries, May 15, 2015), enzyme – trypsin, maximum missed cleavage sites – 1, taxonomy – all entries, precursor mass tolerance – 20 ppm, fragment mass tolerance – 20 mmu, modifications groups from quan method – iTRAQ 8plex, static modification –

methylthio (C), dynamic modifications – oxidation (M), acetyl (protein N-term), Gln→pyro-Glu (N-term Q), Glu→pyro-Glu (N-term E). The following consensus workflow consisted of peptide spectrum matches grouping, validation of peptides (strict target FDR – 0.01, relaxed target FDR – 0.05), peptide and protein filtering (peptide confidence – high, minimum peptide length – 6, minimum number of peptide sequences – 2), peptide and protein quantification (quantification values were rejected if not all quantification channels were present, experimental bias correction was done using median of master proteins) and result export into text files that were further processed with statistical software R. The MS proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD003597 (DOI: 10.6019/PXD003597).

Data analysis

GraphPad Prism software v6.05 (GraphPad Software, San Diego, CA) was used for statistical analyses of groups of patients. Continuous variables were tested for the normality of distribution (D´Agostino-Pearson test). Normally distributed variables are expressed as a mean ± standard deviation. Categorical variables are presented as frequencies (percentages). The statistical significance was designated at two-sided probability value of $p \leq 0.05$.

A statistical software R (v. 3.2.2; R Core Team 2015) with the user interface RStudio (v. 0.99.484) and installed packages extremevalues (v. 2.3.1), FactoMineR (v. 1.31.3), factoextra (v. 1.0.3) and scatterplot3d (v. 0.3-36) were used for processing of exported results from Proteome Discoverer software. The list of identified proteins with corresponding quantitation values was first filtered for representative master protein sequences having quantitation data available for all studied cohorts. The outlier values from the quantitative analysis of control experiment (ratio of two different iTRAQ labels of control sample) were removed (function "getOutliers" in the package "extremevalues") and the average and standard deviation of the lognormal ratio distribution were calculated. For further data processing, only the proteins exceeding the quantitative ratio of two standard deviations at least in one disease group were considered. The logarithm values of quantitative ratios were submitted to principal component analysis (PCA) in the R-package "FactoMineR" (function "PCA") with the graphical output created in the R-package "scatterplot3d"

(function "scatterplot3d") and the obtained data were submitted to hierarchical clustering analysis (HCA) in the R-package "FactoMineR" (function "HCPC" for data processing and function "plot.HCPC" for graphical output).

Results

Using iTRAQ peptide labeling strategy we carried out semiquantitative comparison of plasma proteome in samples collected from patients suffering from non-obstructive HCM, stable DCM, AS, stable AH, CAD and healthy donors. Before mass spectrometry analysis the pooled plasma samples of all clinical states were subjected to immunoaffinity depletion of the 14 most abundant proteins including albumin, immunoglobulin G, A and M, transferrin and others. After abundant protein depletion proteins were digested by trypsin and released peptides were labeled by iTRAQ reagent. The subsequent mass spectrometry analysis by nano-LC-ESI Orbitrap mass spectrometer generated quantitative dataset containing 476 proteins. Of them 293 proteins exhibited significant down- or up-regulation comparing the diseased specimens to healthy donors. When we focused only on proteins that showed disease-unique shift (i.e. the protein exhibits regulation shift in one direction for the particular disease and there is opposite or no change for other diseases) in plasma levels 128 proteins that can represent potential biomarker candidates were found. All of them are summarized in Table 2. Many of the proteins belonged to proteins of coagulation cascades and proteins involved in inflammatory and immune responses. Surprisingly there was great discrepancy in number of proteins with significantly altered plasma levels among individual cardiovascular diseases. The analysis of the pooled sera from patients with DCM proved the highest number of significant plasma protein changes, 67 proteins, followed by AS with 49 deregulated plasma proteins. Contrary, the least number of quantitative changes in protein spectrum, 8 proteins, were found in HCM plasma specimen. It was also interesting that some cardiovascular pathologies exerted significantly altered plasma levels of identical proteins but in opposite directions. The typical example is mannose-binding protein C whose level is up-regulated in HCM but down-regulated in DCM. Similarly the plasma level of serum paraoxonase/arylesterase 1 is increased in plasma sample from patients with AH and declined in patients suffering from AS. Besides proteins exhibiting

disease unique patterns there were additional proteins showing up- or down-regulation in two or more cardiovascular pathologies. In this case the highest number of proteins exerting the identical shift in their plasma levels was found in samples collected from patients with DCM and AS, 121 proteins, and in samples from patients with DCM, AS and CAD, 72 proteins (data not shown).

Table 2. A list of 128 proteins exerting significant alteration in plasma level comparing to healthy controls and down- or up-regulation uniquely associated with particular cardiovascular disease.

Accession	Description	HCM	DCM	CAD	AS	AH
Q14624-2	Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4	--	down	--	--	--
Q16853	Membrane primary amine oxidase	--	--	--	up	--
O95980	Reversion-inducing cysteine-rich protein with Kazal motifs	--	up	--	--	--
P01616	Ig kappa chain V-II region MIL	--	--	up	--	--
P55103	Inhibin beta C chain	--	down	--	--	--
B4DPQ0	Complement C1r subcomponent	--	down	--	--	--
P22792	Carboxypeptidase N subunit 2	--	down	--	--	--
P35590	Tyrosine-protein kinase receptor Tie-1	--	up	down	--	down
P27169	Serum paraoxonase/arylesterase 1	--	--	--	down	up
D3YTG3	Target of Nesh-SH3	--	--	--	--	up
E9PFZ2	Ceruloplasmin	--	--	down	up	up
P00390	Glutathione reductase, mitochondrial	--	--	--	up	--
Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1	--	down	--	--	--
P08581-2	Isoform 2 of Hepatocyte growth factor receptor	--	down	up	up	--
P55058	Phospholipid transfer protein	--	down	--	--	--
P18206	Vinculin	--	--	--	--	down
P04003	C4b-binding protein alpha chain	down	up	up	up	--
P04004	Vitronectin	--	--	--	down	--
Q02985	Complement factor H-related protein 3	--	--	--	up	--
P36222	Chitinase-3-like protein 1	--	--	down	--	--
P40189	Interleukin-6 receptor subunit beta	--	down	--	--	--
Q92823-5	Isoform 5 of Neuronal cell adhesion molecule	--	--	--	up	--
P09668	Pro-cathepsin H	up	--	down	--	down
P55056	Apolipoprotein C-IV	--	down	up	--	--
P10721	Mast/stem cell growth factor receptor Kit	--	down	--	--	--
P18065	Insulin-like growth factor-binding protein 2	--	--	--	up	--
Q15582	Transforming growth factor-beta-induced protein ig-h3	--	--	--	down	--
Q13201	Multimerin-1	--	down	--	--	--
P02743	Serum amyloid P-component	--	--	--	down	--
P80108	Phosphatidylinositol-glycan-specific phospholipase D	--	--	--	down	--
P24043	Laminin subunit alpha-2	up	--	--	--	--
Q14956	Transmembrane glycoprotein NMB	--	up	--	--	--
P15169	Carboxypeptidase N catalytic chain	--	down	--	--	--
P07333	Macrophage colony-stimulating factor 1 receptor	--	down	--	--	--
Q13508	Ecto-ADP-ribosyltransferase 3	--	--	down	up	--
O00187-2	Isoform 2 of Mannan-binding lectin serine protease 2	--	down	--	--	up
P24821	Tenascin	--	--	--	up	--
P18428	Lipopolysaccharide-binding protein	--	--	down	--	--
P51884	Lumican	--	down	--	--	--
P02750	Leucine-rich alpha-2-glycoprotein	--	--	--	up	--

Accession	Description	HCM	DCM	CAD	AS	AH
P40197	Platelet glycoprotein V	--	up	down	up	--
P35443	Thrombospondin-4	--	down	--	--	--
Q9H4A9	Dipeptidase 2	--	--	--	--	down
P11226	Mannose-binding protein C	up	down	--	--	--
P04070-2	Isoform 2 of Vitamin K-dependent protein C	--	down	--	--	--
O43157	Plexin-B1	--	--	down	--	--
Q15113	Procollagen C-endopeptidase enhancer 1	--	--	up	--	--
P17936-2	Isoform 2 of Insulin-like growth factor-binding protein 3	--	--	--	down	--
Q9NPY3	Complement component C1q receptor	--	down	--	--	--
P02753	Retinol-binding protein 4	--	--	--	down	--
P00740	Coagulation factor IX	--	down	--	--	--
P04217	Alpha-1B-glycoprotein	--	down	--	--	--
Q86UD1	Out at first protein homolog	--	up	--	--	--
A0A0A0MSD0	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	--	--	down	up	--
P25786-2	Isoform Long of Proteasome subunit alpha type-1	--	--	--	down	--
P09467	Fructose-1,6-bisphosphatase 1	--	--	--	down	--
Q6UXB8	Peptidase inhibitor 16	--	down	--	--	--
Q9HCL0	Protocadherin-18	--	--	up	--	--
P00451	Coagulation factor VIII	--	down	up	--	up
Q13093	Platelet-activating factor acetylhydrolase	--	--	--	--	down
O14791-2	Isoform 2 of Apolipoprotein L1	--	up	up	up	down
X6R8F3	Neutrophil gelatinase-associated lipocalin	--	--	--	up	--
P06396	Gelsolin	--	--	--	down	--
P13598	Intercellular adhesion molecule 2	--	up	--	--	--
Q9HBR0	Putative sodium-coupled neutral amino acid transporter 10	--	down	--	down	up
Q9NZK5	Adenosine deaminase CECR1	--	--	--	up	--
J3KQ66	Reelin	up	--	--	--	--
P55285	Cadherin-6	up	--	--	--	--
P04406	Glyceraldehyde-3-phosphate dehydrogenase	--	up	--	--	--
P33908	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	--	up	--	--	--
O14498	Immunoglobulin superfamily containing leucine-rich repeat protein	--	down	--	--	--
P02654	Apolipoprotein C-I	up	down	--	down	down
J3KNB4	Cathelicidin antimicrobial peptide	--	--	up	--	--
P23470	Receptor-type tyrosine-protein phosphatase gamma	--	--	--	up	--
P49747	Cartilage oligomeric matrix protein	--	down	--	--	--
Q13822-2	Isoform 2 of Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	--	up	--	--	--
P01042-2	Isoform LMW of Kininogen-1	--	down	--	--	--
Q99650	Oncostatin-M-specific receptor subunit beta	--	--	--	up	--
P62937	Peptidyl-prolyl cis-trans isomerase A	--	down	--	--	--
A6NC48	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2	--	down	--	--	up
P05154	Plasma serine protease inhibitor	--	--	--	down	--
R4GMU1	GDH/6PGL endoplasmic bifunctional protein	--	down	--	--	--
Q9Y4D7	Plexin-D1	--	down	--	--	--
P01042	Kininogen-1	--	--	--	down	--
P10909-2	Isoform 2 of Clusterin	--	--	--	down	--

Accession	Description	HCM	DCM	CAD	AS	AH
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1	--	up	--	--	--
P01034	Cystatin-C	--	--	--	up	--
P48637	Glutathione synthetase	--	--	down	--	--
P04278	Sex hormone-binding globulin	--	--	down	--	--
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	--	--	--	up	--
P0C0L4	Complement C4-A	--	down	--	--	--
Q99784	Noelin	--	down	--	--	--
P16035	Metalloproteinase inhibitor 2	--	up	--	--	--
P35555	Fibrillin-1	--	--	down	--	--
Q15166	Serum paraoxonase/lactonase 3	--	up	down	--	--
P15144	Aminopeptidase N	--	up	--	--	--
B1AHL2	Fibulin-1	--	up	--	--	--
P14543	Nidogen-1	--	--	--	--	down
P10599	Thioredoxin	--	down	up	--	up
Q15848	Adiponectin	--	--	--	up	--
Q86SQ4-3	Isoform 3 of G-protein coupled receptor 126	--	down	--	--	--
P07359	Platelet glycoprotein Ib alpha chain	--	down	--	--	--
P14618	Pyruvate kinase PKM	--	--	--	down	--
P02749	Beta-2-glycoprotein 1	--	down	--	--	--
P01344-3	Isoform 3 of Insulin-like growth factor II	--	--	--	down	--
P07738	Bisphosphoglycerate mutase	--	up	--	--	--
A0A0C4DH25	Protein IGKV3D-20 (Fragment)	--	up	up	up	down
Q9UBQ6	Exostosin-like 2	--	down	--	--	--
P16930	Fumarylacetoacetase	--	--	--	down	--
Q86UN3	Reticulon-4 receptor-like 2	--	down	--	down	up
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	--	up	--	--	--
P52209	6-phosphogluconate dehydrogenase, decarboxylating	--	up	--	--	--
P0C0L5	Complement C4-B	--	down	--	--	--
P02786	Transferrin receptor protein 1	--	down	--	up	up
P09960	Leukotriene A-4 hydrolase	--	--	--	up	--
O15031	Plexin-B2	--	--	--	up	--
O14818	Proteasome subunit alpha type-7	--	--	--	up	--
P00918	Carbonic anhydrase 2	--	up	--	--	--
O14786	Neuropilin-1	--	up	--	--	--
O75023-3	Isoform 3 of Leukocyte immunoglobulin-like receptor subfamily B member 5	--	--	--	--	up
P02144	Myoglobin	--	--	up	--	--
P12109	Collagen alpha-1(VI) chain	--	--	--	up	--
P13647	Keratin, type II cytoskeletal 5	up	up	--	down	up
P10586	Receptor-type tyrosine-protein phosphatase F	--	down	--	--	--
Q08380	Galectin-3-binding protein	--	--	--	up	--
P60174	Triosephosphate isomerase	--	--	up	--	--
Q96S96	Phosphatidylethanolamine-binding protein 4	--	down	--	--	--
Q13332	Receptor-type tyrosine-protein phosphatase S	--	--	--	up	down

a

b
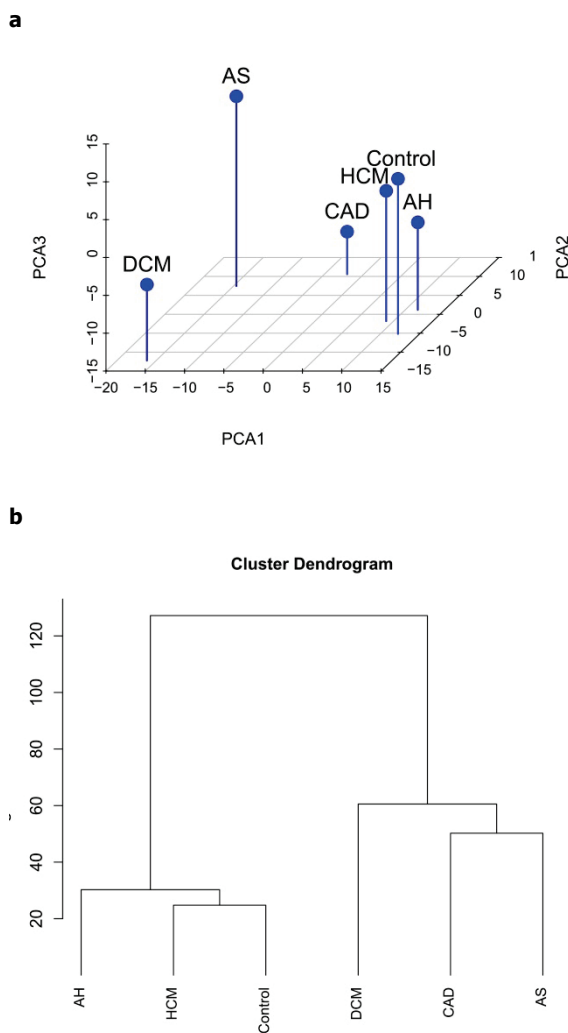
Fig. 1. Diagrams of disease groups and the healthy control group from (**a**) PCA and (**b**) HCA data processing of results obtained from iTRAQ analysis of plasma proteins. Only proteins with significantly different plasma concentration in at least one disease group (related to the control healthy group) were taken into account. PCA diagram shows a three-dimensional view of disease groups and control group projected in the three most significant PCA coordinates. HCA diagram then confirms the distinction of two separate categories consisting of control group, HCM and AH groups on one side and CAD, AS and DCM groups on the other side.

To further study the similarity or dissimilarity of plasma protein profiles of individual cardiovascular diseases we analyzed the acquired quantitative protein data by tools of exploratory data analyses. Globally, 293 proteins with differential abundances were submitted to PCA and HCA. Both procedures clearly divided the samples into two clusters that exhibited great discrepancy in plasma protein profiles. The first cluster was composed of the control, HCM and AH plasma specimens and the second one then involved DCM, CAD and AS plasma specimens. As for the former hierarchical clustering

showed that control and HCM samples exhibits the greatest similarity in protein profiles. Likewise, CAD and AS protein profiles exerted the close similarity in the second cluster (Fig. 1).

Discussion

Cardiovascular diseases form the group of pathologies which is highly perspective for searching for new biomarkers in plasma. There are multiple plasma proteins which are involved in cardiovascular pathogenesis like proteins of coagulation system, lipid transporters or inflammatory mediators. All these proteins are also on our list of identified potential biomarkers. The other perspective plasma proteins are released products of heart damage. Of them Troponin I is in combination with the cardiac isozyme of creatine kinase and myoglobin assumed as primary indicator of myocardial infarction (Jaffe 2001). From this point of view it is quite obvious that numerous proteomics studies of plasma samples have been already undertaken to discover candidates for new biomarkers. Nevertheless, the major progress in cardiovascular proteomics is associated with development of advanced liquid chromatography followed by tandem mass spectrometry approaches (Shen *et al.* 2014). Recently, the strengths of current achievements of proteomics in basic and clinical cardiovascular research were summarized in a scientific statement from the American Heart Association (Lindsey *et al.* 2015).

In our study we were primarily focused on identification of novel candidates for biomarkers of HCM. We employed the semiquantitative advanced LC-MS/MS approach based on peptide iTRAQ labeling and to get disease-specific data we also included into comparative analysis plasma samples from patients with other cardiovascular disorders like DCM, AH, CAD and AS. Comparing to healthy control we obtained 293 proteins with altered plasma concentrations. When we checked the list of potential biomarkers of cardiovascular disorders published in previous proteome studies (Anderson 2005, Keshishian *et al.* 2015, Prentice *et al.* 2010, Adonna *et al.* 2012, Rosello-Lleti *et al.* 2012), we found out that 49 newly described candidates occur in our study. The mutual comparison of protein profiles of five cardiovascular pathologies led to identification of the disease-unique biomarker candidates. As for HCM we found only 8 proteins with a significant shift in plasma concentration. Out of these proteins, 7 showed a unique shift in plasma concentrations in the case of HCM group. It was the least number of identified

potential biomarkers of all analyzed clinical states. This group of biomarker candidates involves one down-regulated protein, C4b-binding protein alpha chain and 6 up-regulated proteins covering well-known marker of cardiovascular diseases, apolipoprotein C-1, and additionally, mannose-binding protein C, pro-cathepsin H, laminin subunit alpha-2, reelin and cadherin 6. The quantitative changes of these proteins in plasma in patients with HCM have not been described, yet. C4b-binding protein is an important inhibitor of the classical and lectin pathway of complement activation and its binding to cardiomyocytes is an early marker of tissue damage (Trouw *et al.* 2008). In contrast to HCM its plasma level is up-regulated in DCM, CAD and AS plasma specimens. Mannose-binding protein C belongs to the collectin family and it activates the lectin pathway of complement. Recent publications document its role in myocardial ischaemia-reperfusion injury in rats and in vascular dysfunction and cardiomyopathy after acute hyperglycemia in mice (Pavlov *et al.* 2015, Pagowska-Klimek and Cedzynski 2014). The same protein, however, down-regulated is found in the DCM specimen. Fibrosis is a clinical feature of cardiac hypertrophy and therefore our proof of up-regulated laminin subunit alpha-2 makes sense. Additionally, perivascular fibrosis and myocardial hypertrophy of the left ventricle following isoproterenol administration in the rat is associated with positive laminin staining throughout whole ventricle (Grimm *et al.* 1998). Pro-cathepsin H belongs to lysosomal cysteine proteinases that exhibit strong collagenase and elastase activity. The increased activity and elevated expression of cathepsins was described in atherosclerosis and coronary heart diseases and hypertensive heart disease, as well (Hua and Nair 2015). Reelin is a large glycoprotein that was originally implicated in development of the nervous system, nevertheless, recent data also revealed its expression in the lymphatic vasculature (Lutter *et al.* 2012). Finally, cadherin 6 has a supportive role in platelet adhesion aggregation and thrombus formation (Dunne *et al.* 2012).

In contrast to HCM, DCM exerted the highest number of altered plasma proteins. Additionally, this disease shared a number of protein level changes with CAD and AS. Hence, we applied PCA and HCA tools to evaluate the relationships existing among the protein profiles of individual pooled disease samples. The analyses confirmed the great discrepancy between protein profiles of DCM, CAD and aortic valve stenosis on the one side and protein profiles of HCM and AH on the other side. This finding might indicate different pathophysiology of these two groups of cardiovascular diseases.

In conclusion immense technical progress in quantitative mass spectrometry represents great hope in uncovering the mechanisms of cardiovascular diseases together with the identification of novel biomarkers for early diagnosis. In our study we applied sophisticated iTRAQ approach for comparison of plasma protein profiles collected from patients with different cardiovascular pathologies. Despite of the study limitations (age, male/female ratio and medication differences), we found broad spectrum of proteins whose changes in plasma levels were unique for individual cardiovascular diseases. All these proteins have possible biomarker potential that will be subsequently validated in individual patient specimens using methods of targeted proteomics. Additionally, the examination of plasma protein profiles by methods of exploratory statistical analysis divided the analyzed cardiovascular pathologies into two clusters indicating their different pathophysiology.

Conflict of Interest
There is no conflict of interest.

Acknowledgements

References

ADDONA TA, SHI X, KESHISHIAN H, MANI DR, BURGESS M, GILLETTE MA, CLAUSER KR, SHEN D, LEWIS GD, FARRELL LA, FIFER MA, SABATINE MS, GERSZTEN RE, CARR SA: A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat Biotechnol* **29**: 635-643, 2011.
ANDERSON L: Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J Physiol* **563**: 23-60, 2005.

COOPER RM, SHAHZAD A, STABLES RH: Intervention in HCM: patient selection, procedural approach and emerging techniques in alcohol septal ablation. *Echo Res Pract* **2**: R25-R35, 2015.

DUNNE E, SPRING CM, REHEMAN A, JIN W, BERNDT MC, HEWMAN DK, NEWMAN PJ, NI H, DERMOT K: Cadherin 6 has a functional role in platelet aggregation and thrombus formation. *Arterioscler Thromb Vasc Biol* **32**: 1724-1731, 2012.

GRIMM D, ELSNER D, SCHUNKERT H, PFEIFER M, GRIESE D, BRUCKSCHLEGEL G, MUDERS F, RIEGGER GA, KROMER EP: Development of heart failure following isoproterenol administration in the rat: role of the renin-angiotensin system. *Cardiovasc Res* **37**: 91-100, 1998.

HENSLEY N, DIETRICH J, NYHAN D, MITTER N, YEE MS, BRADY M: Hypertrophic cardiomyopathy: a review. *Anesth Analg* **120**: 554-569, 2015.

HUA Y, NAIR S: Proteases in cardiometabolic diseases: Pathophysiology, molecular mechanisms and clinical applications. *Biochem Biophys Acta* **1852**: 195-208, 2015.

JAFFE AS: New standard for the diagnosis of acute myocardial infarction. *Cardiol Rev* **9**: 318-322, 2001.

KESHISHIAN H, BURGESS MW, GILLETTE MA, MERTINS P, CLAUSER KR, MANI DR, KUHN EW, FARRELL LA, GERSZTEN RE, CARR SA: Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol Cell Proteomics* **14**: 2375-2393, 2015.

KUMAR KR, MANDLEYWALA SN, LINK MS: Atrial and ventricular arrhythmias in hypertrophic cardiomyopathy. *Card Electrophysiol Clin* **7**: 173-186, 2015.

LANG RM, BADANO LP, MOR-AVI V, AFILALO J, ARMSTRONG A, ERNANDE L, FLACHSKAMPF FA, FOSTER E, GOLDSTEIN SA, KUZNETSOVA T, LANCELLOTTI P, MURARU D, PICARD MH, RIETZSCHEL ER, RUDSKI L, SPENCER KT, TSANG W, VOIGT JU: Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* **16**: 233-270, 2015.

LINDSEY ML, MAYR M, GOMES AV, DELLES C, ARRELL DK, MURPHY AM, LANGE RA, COSTELLO CE, JIN YF, LASKOWITZ DT, SAM F, TERZIC A, VAN EYK J, SRINIVAS PR: Transformative impact of proteomics on cardiovascular health and disease: A scientific statement from the American Heart Association. *Circulation* **132**: 852-872, 2015.

LUTTER S, XIE S, TATIN F, MAKINEN T: Smooth muscle-endothelial cell communication activates Reelin signaling and regulates lymphatic vessel formation. *J Cell Biol* **197**: 837-849, 2012.

PĄGOWSKA-KLIMEK I, CEDZYŃSKI M: Mannan-binding lectin in cardiovascular disease. *Biomed Res Int* **2014**: 616817, 2014.

PAVLOV VI, TAN YS, MCCLURE EE, LA BONTE LR, ZOU C, GORSUCH WB, STAHL GL: Human mannose-binding lectin inhibitor prevents myocardial injury and arterial thrombogenesis in a novel animal model. *Am J Pathol* **185**: 347-355, 2015.

PRENTICE RL, PACZESNY S, ARAGAKI A, AMON LM, CHEN L, PITTERI SJ, MCINTOSH M, WANG P, BUSON BUSALD T, HSIA J, JACKSON RD, ROSSOUW JE, MANSON JE, JOHNSON K, EATON C, HANASH SM: Novel proteins associated with risk for coronary heart disease or stroke among postmenopausal women identified by in-depth plasma proteome profiling. *Genome Med* **48**: 1-13, 2010.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, 2015.

ROSELLÓ-LLETÍ E, ALONSO J, CORTÉS R, ALMENAR L, MARTÍNEZ-DOLZ L, SÁNCHEZ-LÁZARO I, LAGO F, AZORÍN I, JUANATEY JR, PORTOLÉS M, RIVERA M: Cardiac protein changes in ischaemic and dilated cardiomyopathy: a proteomic study of human left ventricular tissue. *J Cell Mol Med* **16**: 2471-2486, 2012.

SHEN X, YOUNG R, CANTY JM, QU J: Quantitative proteomics in cardiovascular research: global and targeted strategies. *Proteomics Clin Appl* **8**: 488-505, 2014.

TROUW LA, OKROJ M, KUPREISHVILI K, LANDBERG G, JOHANSSON B, NIESSEN HWM, BLOM AM: C4b-binding protein is present in affected areas myocardial infarction during acute inflammatory phase and covers a larger area than C3. *PLoS ONE* **3**: 1-7, 2008.

WIESE S, REIDEGELD KA, MEYER HE, WARSCHEID B: Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics* **7**: 340-350, 2007.

XIAO Y, YANG KQ, JIANG Y, ZHOU XL: Recent progress in end-stage hypertrophic cardiomyopathy. *Am J Med Sci* **349**: 448-453, 2015.

YANG J, XU WW, HU SJ: Heart failure: advanced development in genetics and epigenetics. *Biomed Res Int* **2015**: 352734, 2015.