Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/euprot

Investigation of heart proteome of different consomic mouse strains. Testing the effect of polymorphisms on the proteome-wide trans-variation of proteins

Stefanie Forler[a], *Oliver Klein*[b,c], *Sebastian Köhler*[a], *Peter N. Robinson*[a], *Henning Witt*[d], *Marc Sultan*[d], *Murat Eravci*[e], *Vera Regitz-Zagrosek*[f], *Hans Lehrach*[d], *Joachim Klose*[a,b,c,*]

[a] *Institute for Medical Genetics and Human Genetics, Charité – University Medicine, Berlin, Germany*
[b] *Berlin-Brandenburg Center for Regenerative Therapies, Charité – University Medicine, Berlin, Germany*
[c] *Core Unit Proteomics, Berlin-Brandenburg Center for Regenerative Therapies, Charité – University Medicine, Berlin, Germany*
[d] *Max Planck Institute for Molecular Genetics, Berlin, Germany*
[e] *Max-Delbrück-Centrum for Molecular Medicine, Berlin, Germany*
[f] *Institute of Gender in Medicine, Center for Cardiovascular Research, Charité – University Medicine, Berlin, Germany*

ARTICLE INFO

Article history:
Received 24 November 2014
Received in revised form
10 March 2015
Accepted 12 March 2015
Available online 27 March 2015

Keywords:
2-D electrophoresis
Heart hypertrophy
Heart proteome
Mass spectrometry
Mouse chromosomal substitution strains

ABSTRACT

We investigated to which extent polymorphisms of an individual affect the proteomic network. Consomic mouse strains (CS) were used to study the trans-effect of the cis-variant (polymorphic) proteins of the strain PWD/Ph on the proteins of the host strain C57BL/6J. The cardiac proteome of ten CSs was analyzed by 2-DE and MS. Cis-variant PWD proteins altered a high number of C57BL/6J proteins, but the number of trans-variant proteins differed considerably between different CSs. Cardiac hypertrophy was induced in CSs. We found that high variability of the proteome, as induced by polymorphisms in CS14, acts protective against the complex disease.

Abbreviations: SNP, single nucleotide polymorphisms; CS, consomic strain; Chr, chromosome; GWAS, genome-wide association studies; QTL, quantitative trait loci; MGI, mouse genome informatics; TAC, transversal aortic constriction; IVS, intraventricular septum; LVID, left ventricular end-diastolic diameter; LV PW, left ventricular posterior wall; LVM, left ventricular mass; BW, body weight; EF, ejection fraction; FS, fractional shortening; 2-DE, two dimensional gel electrophoresis.

* *Corresponding author at*: Institute for Medical Genetics and Human Genetics, Charité – University Medicine, Berlin, Germany. Tel.: +49 30 450566133; fax: +49 30 450566904.
E-mail address: joachim.klose@charite.de (J. Klose).

http://dx.doi.org/10.1016/j.euprot.2015.03.002

1. Introduction

Common multifactorial diseases are caused by genetic and/or environmental perturbations. Moreover, disease associated modifier genes are involved which influence the individual severity of the disease. For example, familial hypertrophic cardiomyopathy (FHC) is caused by a mutation in a gene encoding the sarcomeric β-myosin binding protein [1]. In order to investigate the modifier effect, 26 patients, all carriers of the same mutation in the sarcomeric gene MyBP-C, were investigated with regard to particular polymorphisms [2,3]. In spite of the presence of the same mutation, the severity of the disease differed among the 26 family members considerably, and this was found to be due to distinct polymorphisms, apparently acting as disease modifiers.

Recent developments in DNA analysis have led to comprehensive genome-wide association studies (GWAS) aimed at detecting genes/single nucleotide polymorphisms (SNPs) that modify the expression of a disease caused by genetic or environmental factors. In these studies a general observation is that many SNPs associated with a disease can be detected, but the vast majority of effect sizes contributing to the disease phenotype are small [4]. For example, 47 distinct genetic variants associated with blood pressure and hypertension were found to explain collectively only a few percent of the heritability of blood pressure [5]. It has been concluded that the complete genetic architecture of blood pressure involves possibly hundreds of genetic variants [6], but that current GWAS results have identified only a subset of this architecture [5].

Chromosomal substitution strains (consomic strains, CS) of the rat have been used as an animal model to screen the chromosomes separately for disease modifier loci acting as quantitative trait loci (QTL). Two strains (BN and SS) of the rat differing in sensitivity against hypertension have been studied in this way [7–9]. The different chromosomes were screened for QTLs that show association with high blood pressure. QTLs related to blood pressure were identified, but surprisingly on almost every chromosome [8]. The results suggested that global responses to an initial stimulus of blood flow, such as salt intake, change the expression of genes throughout the genome [10].

The results considered show that the phenotypic expression of a disease in response to a genetic variant or external stimulus involves many genes, but with different consequences for each individual due to the polymorphisms [4]. This suggests behind the action of a particular disease gene a global acting disease modifier effect [11]. The global modifier would result from the cellular proteomic network and the connected regulatory system. The efficiency of the functional network in compensating the deleterious effect of the stimulus would depend on the number, type and composition of the interacting individual variants that form the genetic background of each individual.

It is a particular feature of the genetic system, that the genome of each individual is composed of two sets of alleles, that from the mother and that from the father, which become combined under heterozygote conditions arbitrarily. Recently the genome of one person was analyzed by next generation sequencing and the maternal genome was compared with the paternal genome with respect to polymorphisms [12]. The two genomes differed by about 11,000 SNPs resulting in non-synonymous amino acid exchanges in coding genes. In addition, 178,616 SNPs were found in non-coding regions up to 10 kb upstream of genes and 6599 in transcription factor binding sites. The arbitrary combination of the alleles of so many genes in an individual may alter fundamental characteristics of biological systems such as the robustness against perturbations. This was defined as a property that allows a system to maintain its function even under internal (genetic) or external (environment) perturbations [13,14]. In a cell the proteomic network of interaction and expression is highly coordinated and tightly controlled [15]. Therefore, this system must be maintained by adapting the effect of the different polymorphisms (and other mutations) of each individual [16]. This is a prerequisite for normal development and differentiation of cells and tissues. But adapting the polymorphisms to the proteomic system, i.e. to the network of proteins including expression, regulation, interaction and distribution of proteins, may lead to conditions not necessarily "perfect" [17] in each case. In consequence, the proteomic system may differ between individuals in robustness against perturbations [18]. Robustness against perturbations and its limitation has been studied in various molecular biological systems [13,14,19–21]. Moreover, mathematical models have been proposed to define robustness [14,16,17,21].

The aim of our investigation was to define in the proteome-wide expression of proteins a criterion that may indicate the capacity of an individual proteome to react robustly against perturbations. As an animal model CSs of the mouse were used, and the proteome of the heart was investigated by large-gel 2-dimensional electrophoresis (2-DE) [22] and mass spectrometry (MS). The consomic mouse model used consists of 20 C57BL/6J (B6) strains, each strain differing from each other by one chromosome (the two homozygotes) that has been replaced by the corresponding chromosome from another strain, the PWD/Ph (PWD) mouse strain [23]. Therefore, the large number of polymorphisms which exist between B6 and PWD was reduced in each strain to the polymorphisms (alleles) of one single chromosome, the PWD-Chr. This allows us to test the extent of the trans-effect of the different sets of variants encoded by the PWD-Chr of each CS on the variability of the B6 proteome (Fig. 1). According to our hypothesis high variability in a CS (compared to B6 to detect the effect of the PWD-Chr) may reflect high elasticity of the proteomic network, i.e. high buffer capacity of proteins in up- and down-regulation and high flexibility in using alternative pathways in protein–protein interaction. High variability might be an indication of high robustness of the proteomic network of this CS against perturbations. In another CS, however, the particular composition of the trans-acting polymorphisms may confine the robustness of the proteome. The frequency of trans-variant proteins of a CS induced by the cis-variant proteins (the polymorphic proteins of the consomic chromosome PWD) by protein-protein interaction was taken as a parameter to compare the robustness of the proteomic network between the CSs. If the trans-effect induced in the proteomic network by the cis-variant genes/proteins is high, this would indicate a high potential of this CS to establish homeostasis in the proteomic network even in case of perturbation (high robustness).

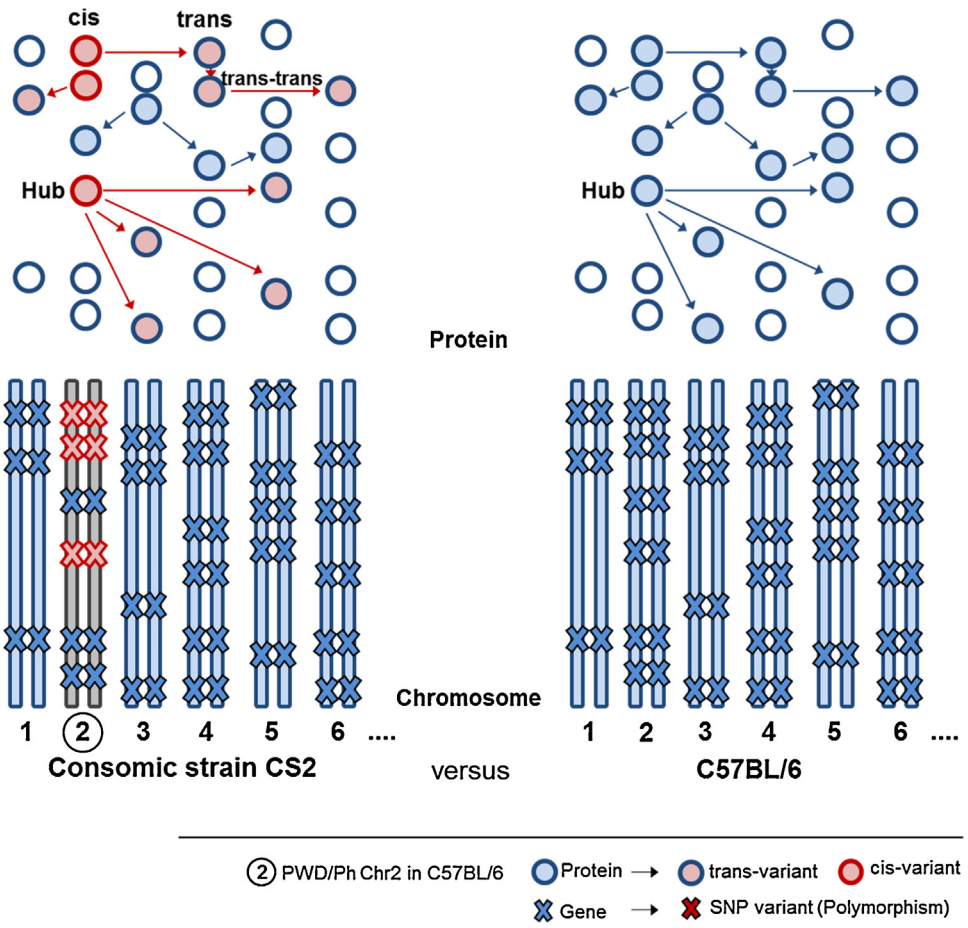

Fig. 1 – Schematic presentation of cis- and trans-acting proteins in a consomic mouse strain (CS2) versus strain C57BL/6 [18].

Under converse conditions homeostasis may be only "partial" [17].

We investigated the heart proteome of ten CSs in males and in females. Three CSs found to differ particularly in the frequency of trans-variant proteins in the heart proteome were tested in their response to induced cardiac pressure overload. Furthermore, we investigated the frequency of variant proteins in the F1(B6xPWD) generation, the situation in that all the 20 chromosomes from the mother (B6) come together with all the 20 chromosomes from the father (PWD). In parallel, the variation in RNA expression was analyzed in 7 CSs and compared with the variation of proteins.

2. Experimental procedures

2.1. *Experimental animals*

Consomic strains (CSs) (chromosome substitution strains) of the mouse based on the inbred C57BL/6J (B6) mouse carrying a single chromosome from the mouse of the PWD/Ph inbred strain were used, and included CSs with Chr^{PWD} No 1, 5, 9, 12, 14, 15, 16, 17, 19, or Y; males and females (Acad. Science, Inst. Molec. Genet., Prague; available from the Jackson Lab., USA). B6 mice were used for comparison. Each experiment was repeated six times with six single animals. At 9 weeks the animals were sacrificed by cervical dislocation, the heart was rapidly removed, cut, rinsed with physiological saline solution, and then flash-frozen in liquid nitrogen and stored at −80 °C until use. All animal procedures were performed in accordance with the guidelines of the Charité Medical University Berlin, were approved by the *Landesamt für Gesundheit und Soziales* (LaGeSo, Berlin, Germany) for the use of laboratory animals (G0206/09), and followed the "Principles of Laboratory Animal Care" [24], as well as the current version of the German Law on the Protection of Animals.

2.2. *Transversal aortic constriction (TAC)*

The mice, 9 weeks of age, were anesthetized with ketamine hydrochloride (80 mg/ml)/xylazine hydrochloride (12 mg/ml) solution administered by intraperitoneal injection at a dose of 1 mg/kg. Animals were placed in the supine position under a dissecting stereoscope. Briefly, after induction of anesthesia, the mice were intubated and artificially ventilated (respirator: Hugo Basile model 7025; FMI). After sternotomy, the aorta and the carotid arteries were exposed, and the transverse aorta was ligated by tying a 6-0 silk suture (FST) against a 26-gauge needle. The needle was then removed, leaving a narrowing of 0.4 mm in diameter. Sham animals underwent an identical surgical procedure without placement of the suture. Animals recovered from anesthesia under warming conditions and

normal ventilation. Mice were killed 63 days after TAC or sham surgery.

2.3. Echocardiography

Two-dimensional short- and long-axis views of the left ventricle (LV) were obtained by transthoracic echocardiography with the Vevo 770 Imaging System (Visual Sonics) as described by Barrick et al. [25]. M-mode tracings were recorded and used to determine intraventricular septum (IVS), LV end-diastolic diameter (LVID), and posterior wall thickness (LVPW) over three cardiac cycles in systole and diastole. 2× LVPWd = thickness/LVIDd was calculated and is an index for concentric myocardial hypertrophy. LV ejection fraction (EF) was evaluated with the formula $\%EF = (LV_{Vold} - LV_{Vols}) \times 100$.

Echocardiography was performed after 2, 3, 4, 6, and 9 weeks in all operated mice after TAC or sham surgery. Detailed protocols are described in the supplemental data for this article available online at the *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology* website.

2.4. Histology

As the most relevant parameter for cardiac remodeling we measured in selected consomic strains the total content of fibrosis in the heart. Paraffin-embedded tissues, from the animals killed at 9 weeks after TAC or sham surgery, were cut in 5-μm sections. Picrosirius Red (SR) staining was used for fibrosis calculation. The fibrosis score was calculated as following: (fibrotic area/total picture area) × 100. Stained sections were quantified using AxioVision Rel.4.6 (Zeiss) and ImageJ 1.38 x software.

2.5. Extraction of total heart proteins

Total protein extracts were prepared from the entire heart. The extraction procedure has been published previously and validated [26,27]. 120 mg frozen tissue, 1.6 parts (v/w) of buffer P (50 mM TRIZMA® Base (Sigma–Aldrich, Steinheim, Germany), 50 mM KCl and 20% (w/v) glycerol at pH 7.5) supplemented with a final CHAPS concentration of 4% (w/v), 0.08 parts of protease inhibitor solution 1 (1 Complete™ tablet (Roche Applied Science, Mannheim, Germany) dissolved in 2 ml of buffer P) and 0.02 parts of protease inhibitor solution 2 (1.4 μM pepstatin A and 1 mM phenylmethylsulfonyl fluoride in ethanol) were grounded to fine powder in a mortar pre-cooled in liquid nitrogen. The tissue powder was transferred into a 2 ml tube (Eppendorf, Hamburg, Germany), and supplied with glass beads (0.034 units of glass beads per combined weight of tissue, buffers and inhibitors in mg; glass beads, 2.5 mm ± 0.05 mm diameter, Worf Glaskugeln GmbH, Mainz, Germany). Each sample was sonicated 12 times in an ice-cold water bath for 20 s each, with cooling intervals of 1 min 40 s in between. The homogenate was stirred for 30 min at 4 °C in the presence of 0.025 parts (v/w) of benzonase (Merck, Darmstadt, Germany) and a final concentration of 5 mM magnesium chloride. Subsequently, solid urea and thiourea to give a final concentration of 6.5 and 2 M, respectively were added, and stirring was continued for 30 min at room temperature until urea and thiourea were completely dissolved. The protein extract was supplied with 70 mM dithiothreitol (Biorad, Munich, Germany), 2% (v/w) of ampholyte mixture Servalyte pH 2-4 (Serva, Heidelberg, Germany), corrected by the amount of urea added (correction factor = sample weight prior to addition of urea/sample weight after addition of urea), and stored at −80 °C. Protein concentrations were determined in sample aliquots without urea using Biorad DC Protein Assay according to the protocol supplied by the manufacturer.

2.6. Two-dimensional gel electrophoresis (2-DE)

Protein samples of the heart from single mice ($n = 6$), 9 weeks of age, of the different CSs were separated by 2-DE. To detect proteins genetically variant between the CS and the parental strain B6, we used the high-resolution large-gel 2-dimensional electrophoresis technique (2-DE), developed in our laboratory and able to resolve more than 10,000 protein spots from complex protein samples [22]. Accordingly, isoelectric focusing (IEF) was performed in capillary tubes (0.75 mm inner diameter) using carrier ampholytes (pH 3–10). The gel format in the second dimension was 40 cm (IEF) × 30 cm (SDS-PAGE). 2-D gels were run in batches of two. Sample pairs consisting of one CS and the control B6 were run always in parallel in both dimensions of 2-DE, IEF and SDS-PAGE. Extracts of 30 μg protein were applied to the anodic end of the IEF gel. Proteins were visualized in the SDS-PAGE polyacrylamide gel by highly sensitive silver staining [22]. 2-DE gels were dried and scanned at 300 dpi and 16 bits gray scale using the Microtek Scan Maker 9800XL scanner (Evestar GmbH, Willich, Germany).

2-DE is at present the only method which allows large scale analysis of genetic variation of proteins with the inclusion of quantitative and structural genetic variation, as well as variation in protein isoforms and posttranslational modifications [28].

2.7. Quantitative analysis of proteins revealed by 2-DE

Protein spot patterns ($n = 6$) were evaluated by Delta2D imaging software version 4.0 (DECODON, Greifswald, Germany). 2-DE spot patterns of CS and control (B6) were matched using the Delta2D "exact" mode matching protocol. First, sample pairs (CS/B6) were matched separately. Subsequently, all gels were matched to create a "match link" between all 2-DE spot patterns using match vectors [29]. Using "union" mode, a fusion image was generated, including the visible spots for each 2-DE gel from each group (males and females, B6 and CS). Only the fusion image was used for spot detection. Spots were not edited manually after spot detection. A set of 2.345 spots of the fusion image was transferred to all other 2-DE gel images for each strain. This ensures that for each stage investigated the ID for each spot on a gel is identical. Relative spot volume intensities (fractions of 100%) were used for quantitative protein analysis. After background subtraction, normalized spot intensity values were copied into Excel spreadsheets for statistical analysis. Data sets were analyzed applying a Students t test to determine statistical significance of alterations (significance threshold $p < 0.05$). The Kolmogorov–Smirnov Z test provided by the statistical analysis software SPSS Statistics 21 (SPSS Inc., Chicago, USA) was used to investigate whether the data was normally distributed. Only fold changes

exceeding 20% were considered [30]. In case of CS versus control (B6) comparisons a t test was used to compare sample pairs which run side by side during both IEF and 2D-PAGE [26].

The rate of false positive protein changes in our 2-DE patterns was evaluated as described elsewhere [31]. Moreover, each "real" protein spot differing significantly was analyzed visually with the Delta2D software tool "Gel Image Region" devised for the evaluation of separation of protein spots and distinction of intensities.

2.8. Protein identification

For protein identification by MS, 0.4 mg of the respective protein extracts were separated on 1.5 mm (inner diameter) IEF and 1.0 mm-thick SDS-PAGE Gels and stained using a MS-compatible silver staining protocol [32]. Protein spots of interest were excised from 2-DE gels and subjected to in-gel tryptic digestion [32]. Trypsin was provided by Promega (Madison, WI, USA). Tryptic fragments were analyzed by nanoflow high-performance liquid chromatography (nanoHPLC; Proxeon Easy-nLC, Odense, Denmark)/electrospray ionization (ESI)-MS and -MS/MS on a LCQ Deca XP ion trap instrument (Thermo Finnigan, Waltham, MA, USA). Nano-HPLC was directly coupled to ESI-MS analysis. Protein band eluates of 18 μl were loaded onto a SSPE traps C18 pre-column (5 μm, 120 Å, 100 μm I.D. × 20 mm; NanoSeparations) using 0.1% (v/v) tri-fluoroacetic acid (Merck) at a flow rate of 20 μl/min. Peptides were separated in an analytical C18 column (5 μm, 120 Å, 75 μm I.D. × 10 cm). The elution gradient was created by mixing 0.1% (v/v) formic acid in water (solvent A, Merck) and 0.1% (v/v) formic acid in acetonitrile (solvent B, Merck) and run at a flow rate of 200 nl/min. The gradient was started at 5% (v/v) solvent B and increased linearly up to 50% (v/v) solvent B after 40 min. ESI-MS data acquisition was performed throughout the LC run. Three scan events: (i) full scan, (ii) zoom scan of most intense ion in full scan and (iii) MS/MS scan of the most intense ion in full scan were applied sequentially. No MS/MS scan on single charged ions was performed. Raw data were extracted by TurboSEQUEST algorithm, and trypsin autolytic fragments and known keratin peptides were subsequently filtered. All DTA files generated by BioWorks version 3.2 (Thermo Scientific, Waltham, MA, USA) were merged and converted to MASCOT generic format files (MGF).

Mass spectra were analyzed using our in-house MASCOT software package license version 2.2 automatically searching the database Uniprot_sprot Version 290709 (531,473 sequences; 188,463,640 residues). MS/MS ion search was performed with following set of parameters: (i) taxonomy: Mus musculus (16,441 sequences), (ii) proteolytic enzyme: trypsin, (iii) maximum of accepted missed cleavages: 1, (iv) mass value: monoisotopic, (v) peptide mass tolerance: 0.8 Da, (vi) fragment mass tolerance: 0.8 Da and (vii) variable modifications: oxidation of methionine and acrylamide adducts (propionamide) on cysteine. No fixed modifications were considered. Only proteins with scores corresponding to $p < 0.05$, with at least two independent peptides identified were considered. The cut off score for individual peptides using ESI-MS identification was equivalent to $p < 0.05$ for each peptide.

Additionally matrix assisted laser desorption ionization (MALDI)-time of flight (TOF) mass spectrometry MS/MS (MALDI TOF/TOF) analyses were used to identify proteins, which could not be identified by ESI-MS. 1.5 μl of peptide extract were spotted onto to Anchor Chip (Bruker, Bremen, Germany), dried and mixed with 1.5 μl of matrix solution containing 0.5 mg/ml α-cyano-4-hydroxycinammic acid in 70% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid. Protein identification was performed by MALDI TOF/TOF (Autoflex III) mass spectrometer (Bruker Daltonik) in positive reflector ion mode. A peptide calibration standard (Bruker Daltonik) was used for external calibration. The MS measurement settings in flexControl were as follows: detection range of *m*/*z* 800–3500, 400 laser shots per spot. MS spectra were annotated and processed in flexAnalysis 3.0 (BrukerDaltonik) with SNAP algorithm, S/N 1, quality factor threshold 50, and "TopHat" baseline subtraction. Parent mass values from generated mass list were analyzed in LIFT mode (flexControl 3.0). The fragmentation mode MS/MS was set as following: detector gain boost 110%, laser power boost 130%, and analog offset 0.5% and parameters, precursor ion selection of 0.65%. MS/MS spectra were annotated and processed in flexAnalysis 3.0 (BrukerDaltonik) with SNAP algorithm, S/N 1, Quality Factor Threshold 50, Smooth and Baseline Subtraction.

MS/MS ion search was performed in Biotools 2.2 (BrukerDaltonics, Bremen, Germany) and MASCOT software version 2.2 (in house license, matrixsince, London, UK) with the following set of parameters: (i) database: Uniprot_sprot Version 290709 (531,473 sequences; 188,463,640 residues), (ii) taxonomy: Mus musculus (16,441 sequences), (iii) proteolytic enzyme: trypsin, (iv) maximum of accepted missed cleavages: 1, (v) mass value: monoisotopic, (vi) peptide mass tolerance: 70 ppm, (vii) fragment mass tolerance: 0.7 Da and (viii) variable modifications: oxidation of methionine. No fixed modifications were considered.

2.9. Label-free quantification by mass spectrometry

For online LC–MS/MS analysis 10 μl of the mouse heart total protein extracts (approx. 133 μg) was diluted with 50 mM ammonium bicarbonate digestion buffer to a concentration of 1 μg/μl. Ten micrograms of these proteins were then reduced with dithiothreitol (DTT) and alkylated with iodoacetamide (IAA) prior to in solution protein digestion with trypsin. After desalting of the peptide solution with zip-tips, peptide mixtures (WT BL-6, CS12 and CS14; 5 biological replicates each) were separated by reversed phase chromatography using the Proxeon Easy nLC II system (Thermo Fisher Scientific) on in-house manufactured 25-cm fritless silica microcolumns with an inner diameter of 75 μm. Columns were packed with ReproSil-Pur C18-AQ 3-μm resin (Dr. Maisch GmbH). Separation was performed using a 10–60% ACN gradient (240 min) with 0.1% formic acid at a flow rate of 200 nl/min. Eluting peptides were directly ionized by electrospray ionization and transferred into the orifice of a linear trap quadrupole Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific). Mass spectrometry was performed in the data-dependent mode with one full scan in the Orbitrap ($m/z = 300–1700$; resolution = 60,000; target value = 1×10^6). The five most intense ions with a charge state greater than 1 were selected (target value, 5000; monoisotopic precursor selection enabled) and fragmented in the linear trap quadrupole using CID

(35% normalized collision energy and wideband activation enabled). Dynamic exclusion for selected precursor ions was 60 s. Protein identification and label free quantification were performed with the MaxQuant software package (version 1.2.2.5) [33]. Statistical analysis of LFQ-data was performed with Welch's *t*-test [33–35]. The correction of potentially false positives in the LFQ-data was performed by calculations with a permutation based false discovery rate [36]. The MS/MS spectra were matched against International Protein Index (IPI) mouse database (version 3.84) and against 248 frequently observed laboratory contaminants as provided by the MaxQuant software package, and calculated using the MaxQuant search engine Andromeda [37]. Search parameters were as follows: Precursor mass tolerance was set to 20 ppm, the mass tolerance for fragment ions was set to 0.5 Da, tryptic specificity allowing for maximum 1 missed cleavage sites and K/R cleavages (accounts for in-source fragmentation of tryptic peptides followed by a proline residue), fixed modification of carbamidomethylation of cysteine residues, and variable modification of oxidation of methionine residues. "Requantify" and "match between runs" were enabled. The target-decoy-based false discovery rate (FDR) for peptide and protein identification was set to 1% for peptides and proteins and the minimum peptide length was set to 6 amino acids.

2.10. *Pathway enrichment analysis in the protein data set*

Official SwissProt-ID numbers (Uniprot, http://www.uniprot.org/) were used to investigate differences in cell function between the variant proteins of different CSs. In order to detect enrichments of specific pathways in the dataset of altered protein expression, a "WEB-based Gene SeT AnaLysis Toolkit" (WEBGESTALT) tool supplied by the Vanderbilt University at http://bioinfo.vanderbilt.edu/webgestalt/was used. We used the Enrichment analysis tool for KEGG (Kyoto Encyclopedia of Genes and Genomes) [38]. The following parameters were used to create the KEGG tables: 1. Reference set: species "Mus musculus", significance level: $p < 0.05$, minimum number of genes: 2. Statistical methods available were "hypergeometric test" and "Fisher's exact test". For our data the results were the same with either test.

2.11. *Proteome-network analysis*

The protein–protein interactions were analyzed for the proteins variant in CS12 (group c12), variant in CS14 (c14) and variant in both CS12 and Cs14 (c12/14), males. We used the STRING protein association network (http://string-db.org, version 9.1), but used only Mus musculus proteins and interactions that had an overall score of at least 700 (high confidence interactions). The complete network consists of 12,993 proteins. For the resulting network we could easily calculate the degree (number of neighbors) and clustering coefficient for each protein. We also calculated the shortest-path distance for all pairs of nodes. For a particular group of proteins (c12, c14, c12/c14) we calculated the average degree and clustering coefficient. We also calculated the shortest-path average, by taking each member and determining the average shortest-path to all other members of the group. Note that proteins that were not reachable by any path were excluded from the shortest-path analyses.

Random sampling: In order to measure the statistical significance of the average degree/clustering coefficient/shortest-path, we randomly constructed a group of proteins, with the same number of proteins and evaluated this group using the same measure. We performed this procedure 10,000 times. The *p*-value for the real average is then the fraction of randomizations that achieved a higher/lower average value. In order to correct for the group specific degree distribution we changed the randomization procedure, such that for each node (with degree x) of the original group we drew a protein from the set X. This set X consists of all proteins that have the same degree x, or have a degree of $x+1$ or $x-1$.

2.12. *RNA analysis*

Total RNA was prepared from frozen heart samples (males and females) using the RNeasy Mini Kit (Qiagen GmbH, Hilden, Germany) and according to manufacturer's instructions. We have profiled the heart transcriptome of the consomic mouse strains, including two biological replicates per strain. Each paired-end RNA-seq library was prepared from 10 μg of total RNA using a strand-specific strategy and following the protocol described in [39]. Sequencing was carried out on the HiSeq 2000 platform (Illumina) by running 2×51 cycles according to the manufacturer instructions. Sequencing reads were aligned to the mm9 assembly of the mouse reference genome using BWA [40]. Gene levels were then quantified in reads per kilobase of exon model per million mapped reads (RPKM) [41] and using the Ensembl v.53 (Mus musculus) annotation. Differential expression was inferred using the EdgeR package from Bioconductor [42].

3. Results

3.1. *Variation of proteins due to polymorphisms*

From the 20 consomic mouse strains B6-Chr^{PWD} available we selected ten strains arbitrarily for investigation in males and females: CS1, 5, 9, 12, 14, 15, 16, 17, 19, Y (CS No of PWD chromosome). Total heart protein extracts from these animals were separated by large-gel 2-DE and the proteins revealed by using highly sensitive silver staining [22]. The 2-DE protein patterns show about 4500 protein spots. Out of these 2345 spots, well separated and reproducible ($n = 6$) in position and quantity, were evaluated in each strain by comparing CS versus B6 (Fig. 2, Suppl. Tab. IA and IB). In males/females 364/375 variant protein spots were detected within the ten CSs; 151/175 were identified by mass spectrometry (MS) (Suppl. Tab. II). Protein spots variant in the F1(B6xPWD) animals (versus B6) were also identified (Suppl. Tab. II). The number of variant protein spots varied considerably between the ten CS. In females between 12 (CS1), 23 (CS12), on one side and 80 (CS14) on the other side, in males between 26 (CS12) and 69 (CS14) (besides of 13 in CSY) (Tab. I). In order to confirm our results CS12 and CS14 males were taken to test the same heart extracts investigated by 2-DE-MS by using label-free LC–MS.

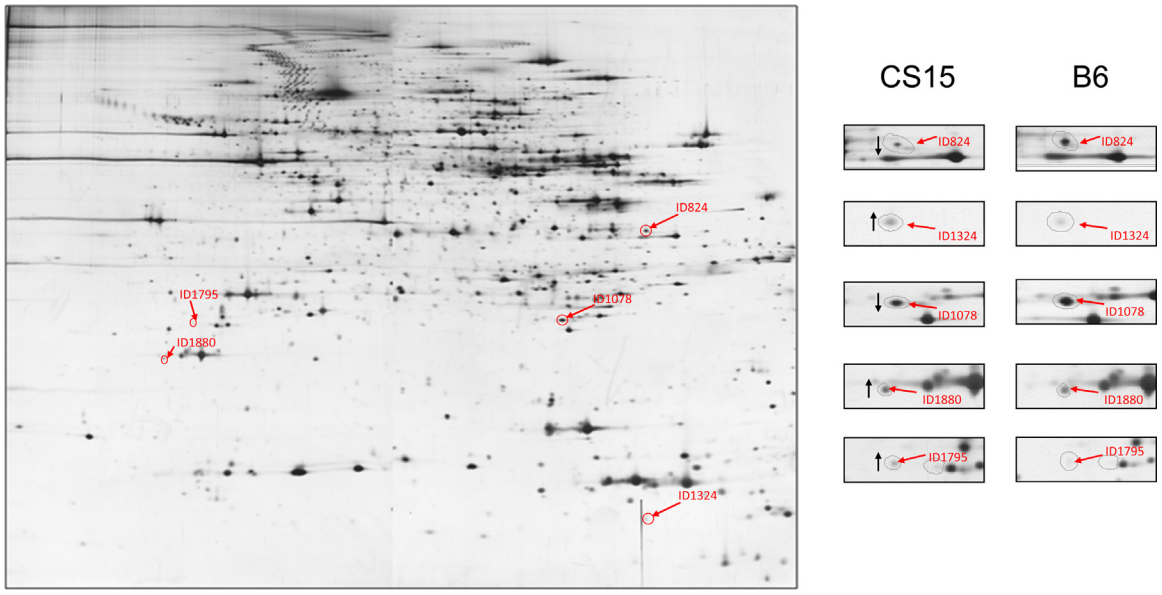

Fig. 2 – The 2-DE protein pattern of the heart of CS15 is shown. As an example, five protein spots variant between CS15 and B6 (males) are indicated. On the right site the five spots are shown in comparison to B6.

The proteins identified in CS12 as well as in CS14 and B6 amounted to 719. In comparison, CS12 versus B6 showed 39 variant proteins, CS14 versus B6 revealed 119 variants (Suppl. Tab. III). Whereas by this approach a higher number of variants were detected than by 2-DE-MS, the frequency of variants in CS14 compared to CS12 (factor 3.05) was quite similar to that found by 2-DE-MS (factor 2.65). Even after correction of the alpha error by calculating the permutation based false discovery rate (FDR) (Suppl. Tab. III), the ratio of the remaining significant values of CS12 to CS14 (factor 2.43) are quite similar to the relation that was observed with 2-DE-MS. Moreover, the low number of cis-variants (6%, see below) detected by 2-DE-MS was confirmed by LC–MS showing 5% cis-variants among the 119 variant proteins of CS14.

Supplementary tables related to this article can be found, in the online version, at doi:10.1016/j.euprot.2015.03.002.

Remark: In our 2-DE pattern, in the high molecular range, a very tight group of a very large number (by far more than one hundred) of tiny spots occur, identified as far as possible and found to be Myh6 (myosin heavy chain). These spots may represent the myosin chain complex occurring under electrophoretic conditions as tiny spots, fused in the upper part of the gel. The myosin complex could not be reproducible identified and quantified for different CSs. But quantitative variation compared to B6 was found for many spots in CS14 (males) and to some extent for males and females of CS9 and CS17.

The frequency of variant proteins detected in each CS showed a highly significant agreement between males and females (Fig. 3) and demonstrates in this way the reproducibility of the 2-DE protein patterns. As an interesting finding, these variants were to a large extent trans-variant proteins apparently resulting via the protein-protein interaction network from a small number of cis-variant proteins (6% found in males, 9% in females out of the total number of variants), the variant proteins which mapped to the consomic PWD-Chr.

In order to compare the network effect (trans-effect) of the cis-variant PWD proteins between the different CSs we used the frequency of single nucleotide polymorphisms (SNP, Mouse genome Informatics (MGI), http://www.informatics.jax.org/) of the genes of the corresponding PWD-Chr as a reference value. This was tested for coding-non-synonymous SNPs, non-coding SNPs (locus-region, mRNA-UTR), and total SNPs (including introns). We found that, within the ten CSs tested the number of variant proteins (almost all trans-variants, see above) did not correlate with the number of variant genes/SNPs (Tab. IA males, IB females). Apparently, the difference in the frequency of trans-variant proteins in the different CSs is a protein-network effect. As considered in the following, “core-proteins” may play a central role in transmitting genetic variation of protein expression.

We identified so far in our 2-DE pattern of the mouse heart 491 proteins. We compared these proteins with the list of proteins (1124 proteins) of the “Central Proteome” published by Burkard et al. [43], combined with the list of proteins (1581 proteins) of the “Eukaryotic Core Proteome” published by Weiss et al. [44]. This resulted in a list of 2074 proteins, called here core-proteins. The central proteome resulted from analyzing seven different human cell types, and the eukaryotic core proteome from five different organisms. The two lists were combined here in terms of mouse genes/proteins. Core proteins are responsible for the common fundamental cell functions, expressed as high-abundant proteins, and they are highly connected in the proteomic network [43,44]. Core-proteins are known to be hub-proteins, i.e. proteins which interact directly with many other proteins [45] and hold the whole network together [11]. Comparison with the heart proteins as revealed in our 2-DE pattern showed that 49% of these were core proteins. Genes encoding core proteins were found to correspond in human tissues to 6.4% [46]. Based on the Burkard/Weiss list we calculated about 9%.

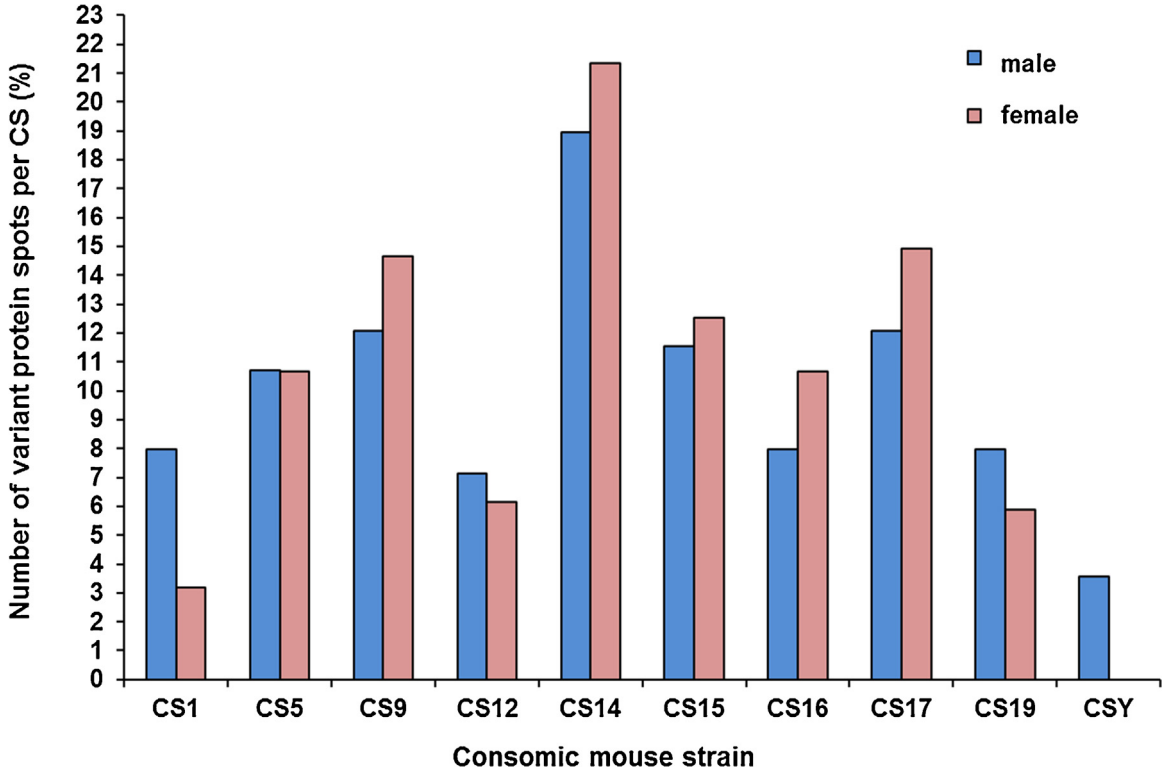

Fig. 3 – Frequency of variant proteins in the heart proteome of consomic mouse strains (CS) in males and females. The frequency of variant protein spots in males and females is highly correlated: $r = 0.920$ (Pearson correlation), $p = 0.0001$ (Students *t* test).

Apparently, core-proteins were drastically enriched in our 2-DE protein patterns. This is certainly due to the fact that 2-DE detects preferentially high-abundant proteins. Among the variant proteins detected in our CSs about 50% (males 48%, females 54%) were core-proteins. We found among the proteins in CSs that 38% (females 29%) were variant in several strains among the ten stains investigated indicating multiply connected proteins. Core proteins contribute most extensively to protein-protein interaction [43] in contrast to low-abundant tissue-specific proteins. In this context it is an interesting finding that CS14 with the high number of variant proteins showed 45% core proteins among these variants, whereas CS12 with low variability in the proteome revealed only 26% core proteins among the variants. In this calculation the representative number of identified proteins obtained by label free MS/MS analysis (Suppl. Tab. III, IV, V) was used and suggest that the differences between CSs in the frequency of trans-variant proteins reflect differences in the frequency of protein-protein interaction (flexibility) based on the core proteins.

Supplementary table related to this article can be found, in the online version, at doi:10.1016/j.euprot.2015.03.002.

3.2. *F1 generation versus CSs*

One would expect that F1(B6xPWD) animals show in the heart protein pattern a similar number of variant protein spots as the 20 different CSs in total, if the overlapping variants within the 20 strains would be ignored. But our analysis revealed, surprisingly, that the number of variant proteins in F1 animals was drastically reduced. We counted the different variant proteins of the ten strains investigated and estimated from this result the total number of variants which may result from all 20 strains (Fig. 4). We found that up to 700–800 variant protein spots have to be expected in a corresponding protein pattern of F1 animals. However, only 86 variants in males and 70 in females were detected. Under the heterozygote conditions in F1 some variants may disappear in case of recessivity, but additional variants would also occur due to positional variant protein spots which generate two variants in F1. Therefore, heterozygosity cannot explain the high discrepancy in the number of variant proteins between the CSs and F1. Obviously, each PWD chromosome has a high potential to create variation on the protein level based on polymorphisms, but in presence of all 20 chromosomes in F1 the variability of proteins is drastically reduced.

3.3. *Males and females*

Out of the 104 identified proteins variant in male CSs and 118 variant in female CSs only 50 were identical (and not necessarily in the same strain). Moreover, there is a tendency in females for down-regulation of proteins, in males for up-regulation; only CS9 and CS19 showed the opposite relationship, and this in a pronounced extent (Fig. 5). These findings suggest that the genetic regulation of the proteomic network differs between males and females proteome-wide. This is certainly due to the sex-specific hormones, but also to hormone-independent sex-specific regulation of transcription [47]. The effect of

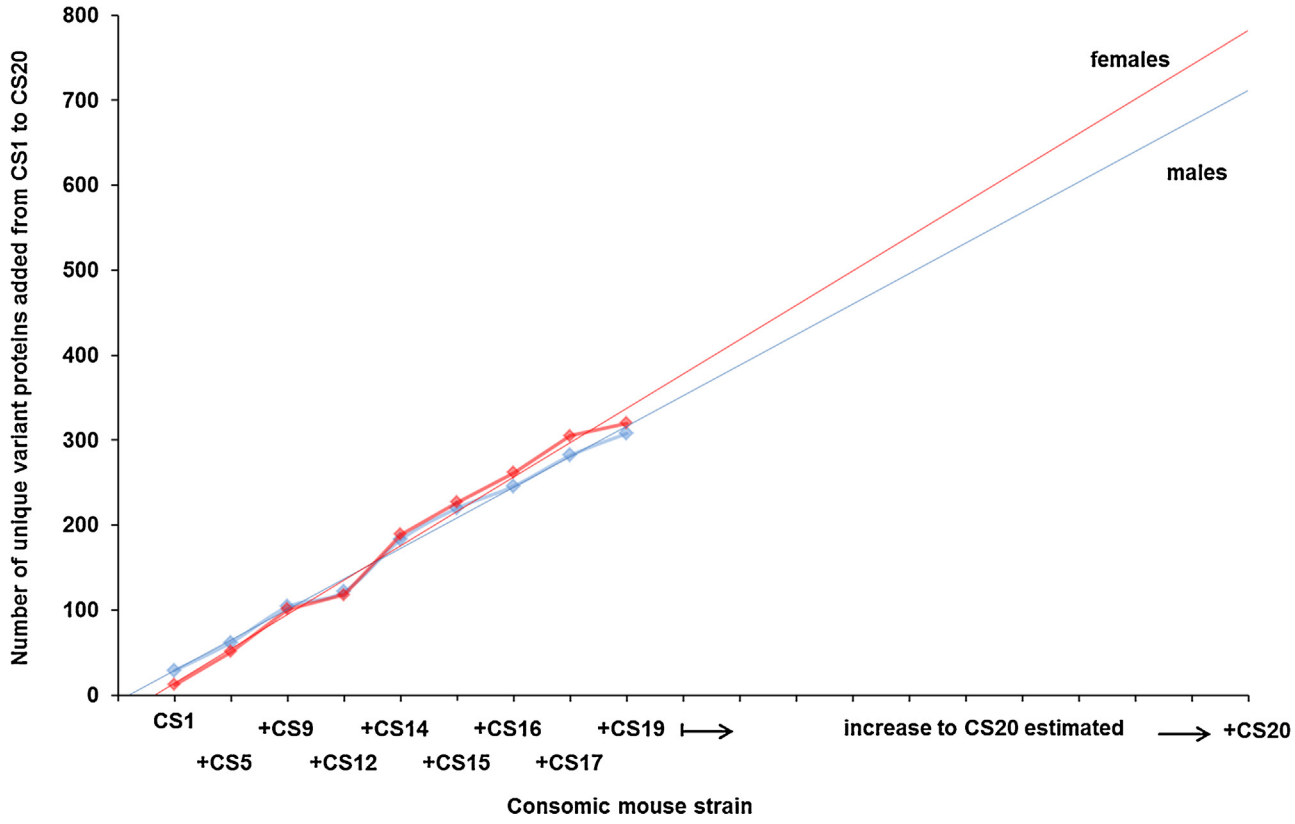

Fig. 4 – Total number of variant proteins (CS versus B6) of the heart in 20 consomic strains (CS). The number of variant protein spots detected in 9 CS strains by 2-D-electrophoresis were added from strain to strain without overlap and supplemented by the remaining 11 strains using a linear trend line.

sex-specificity became more obvious, when the protein patterns from males and females were compared directly within a CS. Males compared with females for instance in CS1 and in CS14 showed sex-specific differences in a representative number of proteins, but in contrast to the pronounced differences in the genetic variability found between these two strains (Tab. I), the number of sex-specifically variant proteins was similar: 72 in CS1, 82 in CS14. At present we investigate females after ovarectomy in different CSs. The results will show to which extent genetic variation in the proteome involves regulatory effects of hormones [45,47].

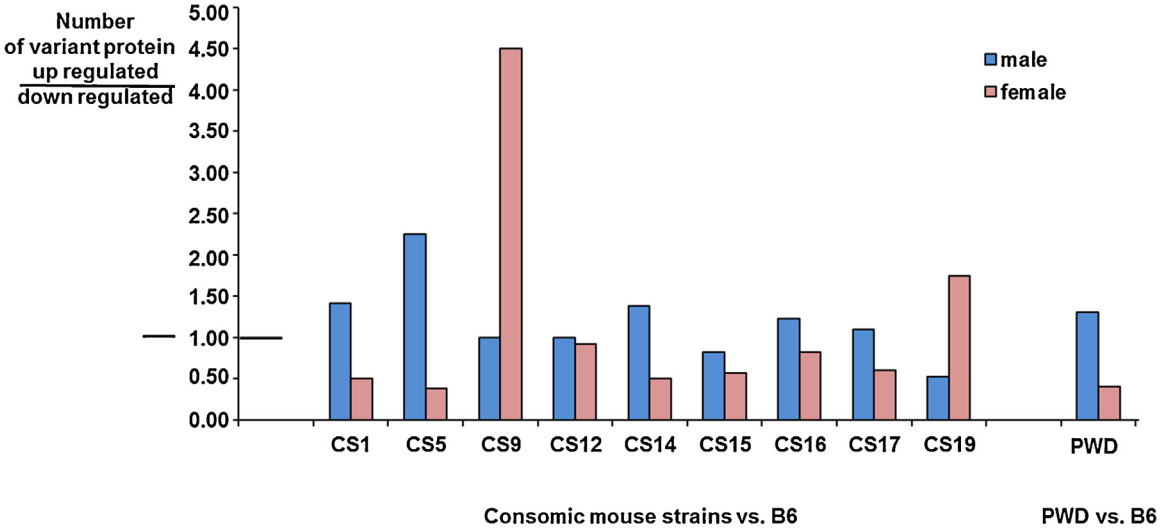

Fig. 5 – Quantitative regulation of variant proteins in the heart of consomic strains (CS). Ratio of up- versus down-regulation of proteins in males and females in different CSs and in the parental strains.

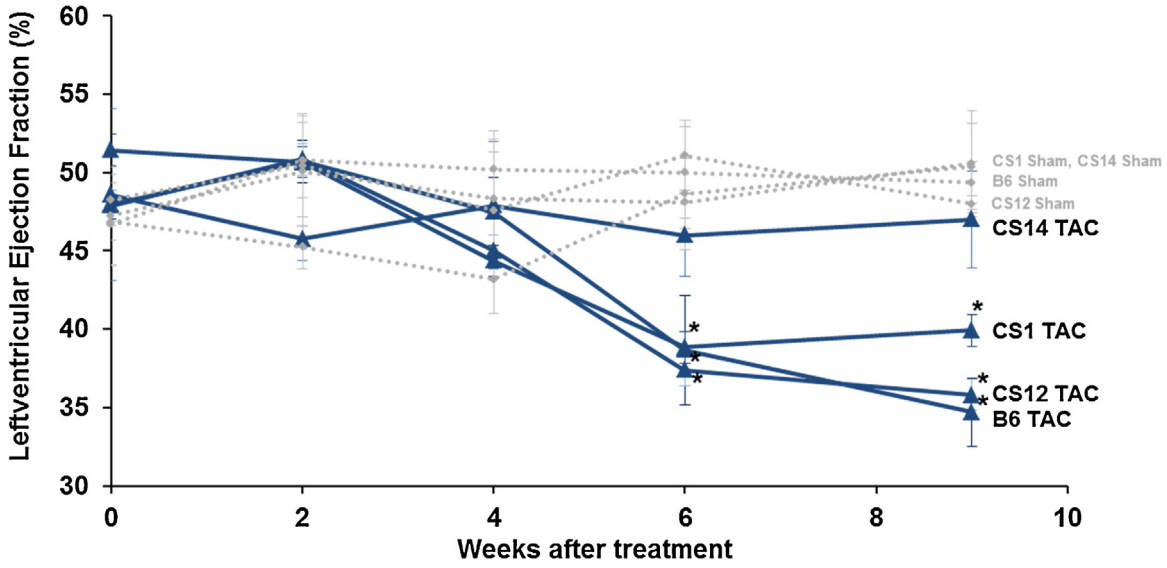

Fig. 6 – Heart ejection fraction of consomic mouse strains (males) CS1, CS12, and CS14 and parental strain B6 in the course of 9 weeks after transversal aortic constriction (TAC). CSs TAC treated blue, Sham gray; *$p < 0.05$ TAC versus Sham.

3.4. *mRNA expression in CSs*

In seven CSs (CS5, 12, 14, 15, 16, 17, 19, males)[1] the mRNA expression in the heart tissue was investigated. In total 1196 variant (versus B6) mRNAs, out of 13,198 mRNAs evaluated, were detected in the seven strains (Tab. II and Suppl. Tab. VI) including 249 (26%) cis-variants and 947 trans-variants. The distribution of the cis-variants over the CSs correlated with the number of genes of the corresponding PWD CS-Chr, but this was not observed for the trans-variant mRNAs (Tab. II). The trans-variants varied among the seven CSs between 42 (CS17) and 247 (CS15). No correlation was found between variant mRNAs and variant proteins, evaluated on cis + trans variants ($p = 0.372$, $r = 0.401$; data not shown). Since mRNA uncovered a much broader abundance range than the proteins under our conditions, only a small number of variant mRNAs overlapped with the variant proteins. Furthermore, among the 612 different mRNAs detected in the seven CSs only 5% corresponded to core-proteins according to Burkard/Weiss (see above).

Supplementary table related to this article can be found, in the online version, at doi:10.1016/j.euprot.2015.03.002.

3.5. *Heart phenotypes of consomic mouse strains under normal conditions and after induced heart hypertrophy*

To study the consequences of proteome-wide effects of polymorphisms on the heart phenotype of different CSs, we investigated in males and females morphological parameters such as septum thickness (IVS), left ventricular internal diameter (LVID), left ventricular posterior wall (LVPW), left ventricular mass (LVM), and body weight (BW) in three representative consomic strains, CS1, CS12, CS14, and in the parental strain B6 by echocardiographic measurements (Suppl. Tab. VII). CS1 represents the situation low variation in proteins versus high variation in genes of the consomic PWD-Chr and CS14 the opposite constellation, whereas CS12 also showed low variation in proteins, but in context with relatively low variation in genes (Tab. I). We also evaluated two physiological properties of the heart, the left ventricular ejection fraction (EF) and the left ventricular fractional shortening (FS) (Suppl. Tab. VII). We found that BW and LVM were increased in males in all 3 consomic strains in comparison with B6. In CS1 the increase in LVM was due to ventricular dilatation and in CS14 to increased LVPW thickness; in CS12 males the LVIDd was increased but at the borderline of significance. Heart function measured as EF and FS was maintained in the consomic males. In contrast, in consomic females BW was not increased in comparison with B6. LVM/BW was decreased in the consomic strains CS1, CS12 and CS14. This was due to lower wall thickness in all three strains as well as to cavity dilatation in CS1 and CS12. There was a trend toward decreased function (EF, FS) in CS1 females. Thus the consomic strains deviated in cardiac growth parameters under resting conditions from B6 and this was different in males and females.

Supplementary table related to this article can be found, in the online version, at doi:10.1016/j.euprot.2015.03.002.

In the following we used our consomic mouse model as an approach to test the effect of the genetic background (given by the PWD-Chr.) on the robustness of the proteomic system against perturbations. We tested the response of the heart to cardiac pressure overload under the different conditions of the proteomic network observed. The extent of trans-variant proteins induced by the polymorphisms was taken as a parameter. According to our hypotheses, high variability in the proteome may reflect high flexibility of the protein network to overcome restrictions in functional pathways.

We selected out of the 10 CSs investigated so far the most contrasting strains – CS1 and CS12 with low frequency of

[1] CS1 and CS9 were no longer available from the Max-Planck Institute.

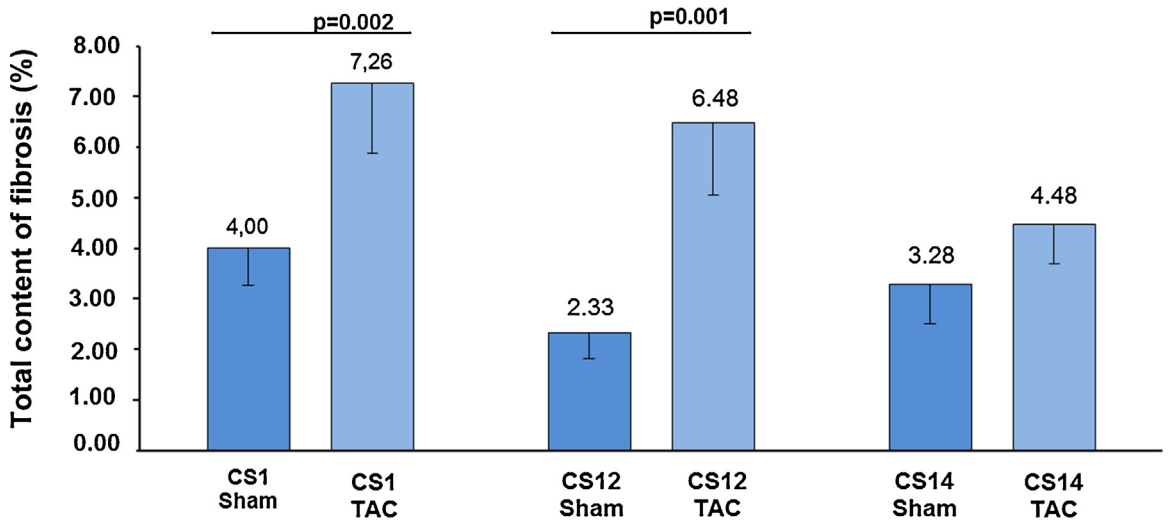

Fig. 7 – Heart hypertrophy induced in consomic mouse strains (males). Heart hypertrophy was induced in consomic mouse strains CS1, CS12, and CS14 by transversal aortic constriction (TAC). The total content of fibrosis of the heart was determined.

protein changes versus CS14 with high frequency of protein variation. In these three strains (males) transversal aortic constriction (TAC) was induced at 9 weeks of age. The control groups were sham operated. Testing different echocardiographic parameters under TAC conditions showed an increase in left ventricular mass (LVM) and a decrease in ejection fraction (EF) in CS1, CS12 and in B6 but not in CS14 (Suppl. Tab. VIII, and Fig. 6). Parameters of systolic function decreased in CS1, CS12, and in B6, but not in CS14 (Suppl. Tab. VIII). Moreover, under TAC conditions the total fibrosis content of the heart was increased (CS1: $p = 0.002$; CS12: $p = 0.001$) in CS1 and CS12 mice (CS1: Sham: 4.00%; TAC: 7.26%; CS12: Sham: 2.33%; TAC: 6.48%) but was in the normal range in CS14 (Sham: 3.28%, TAC: 4.48%) (Fig. 7). Thus, adaptation to pressure overload was best achieved in CS14 whereas worst conditions were found in CS1 and CS12.

Supplementary table related to this article can be found, in the online version, at doi:10.1016/j.euprot.2015.03.002.

3.6. *Proteome-network analysis*

We used the representative number of variant proteins identified in CS12 and CS14 by label-free LC–MS/MS to compare the structure of the protein network of these two strains. CS12 was the strain with the lowest number of trans-variant proteins found among the 10 CSs (males) investigated whereas in CS14 the highest variability of proteins was found (Tab. I, Fig. 3). The STRING protein association data were used to analyze three groups of proteins: Proteins variant only in CS12 (group c12), only in CS14 (c14) and variant in both CS12 and CS14 (c12/14). We calculated the average clustering coefficient (cliquishness), the number of neighbors (degree) and the shortest-path distance (proximity) in these groups, and compared the data obtained with the result obtained for randomly drawn groups with the same number of proteins (see Methods).

We found for c12, including 19 proteins (7 proteins not found), a significantly decreased shortest-path distance ($p = 0.022$), i.e. the pairs of proteins of that group were closer in the network than expected by chance. The degree and clustering coefficient did not demonstrate a statistically significant difference. For c14, including 93 proteins (13 not found), we also found a decreased shortest-path distance ($p = 0.0001$). Additionally, we found an increase in the clustering coefficient ($p = 0.042$) and degree ($p = 0.027$). The c12/14 group (12 proteins, 1 protein not found) also showed a decreased shortest-path distance ($p = 0.001$), but the degree and clustering coefficient did not show any statistical difference.

Comparing the results between the three groups, the c14 group of proteins was the only one showing statistical significant deviations in all three measurements and, more importantly, this group had by far the shortest-path distance. We determined the group-specific degree distributions randomizations (see Methods) and found that the only measure that remains highly significant is the shortest-path distance of the c14 ($p = 0.0073$).

3.7. *Enrichment analysis of functional protein classes of the protein network*

After enrichment analysis of the variant proteins with the KEGG database, functional protein classes were found in CS14, variant versus B6, which differed from those found in CS12, variant versus B6. In CS14 proteins closely connected with the variants were related in particular to the circulatory system, such as proteins involved in cardiac muscle contraction, and proteins of the energy metabolism: oxidative phosphorylation, citrate cycle. In contrast, in CS12 the interacting proteins found were not related to heart contraction: complement and coagulation cascades (immune system), ECM-receptor interaction (environmental information processing), and regulation of actin cytoskeleton (cell mobility) (Fig. 8).

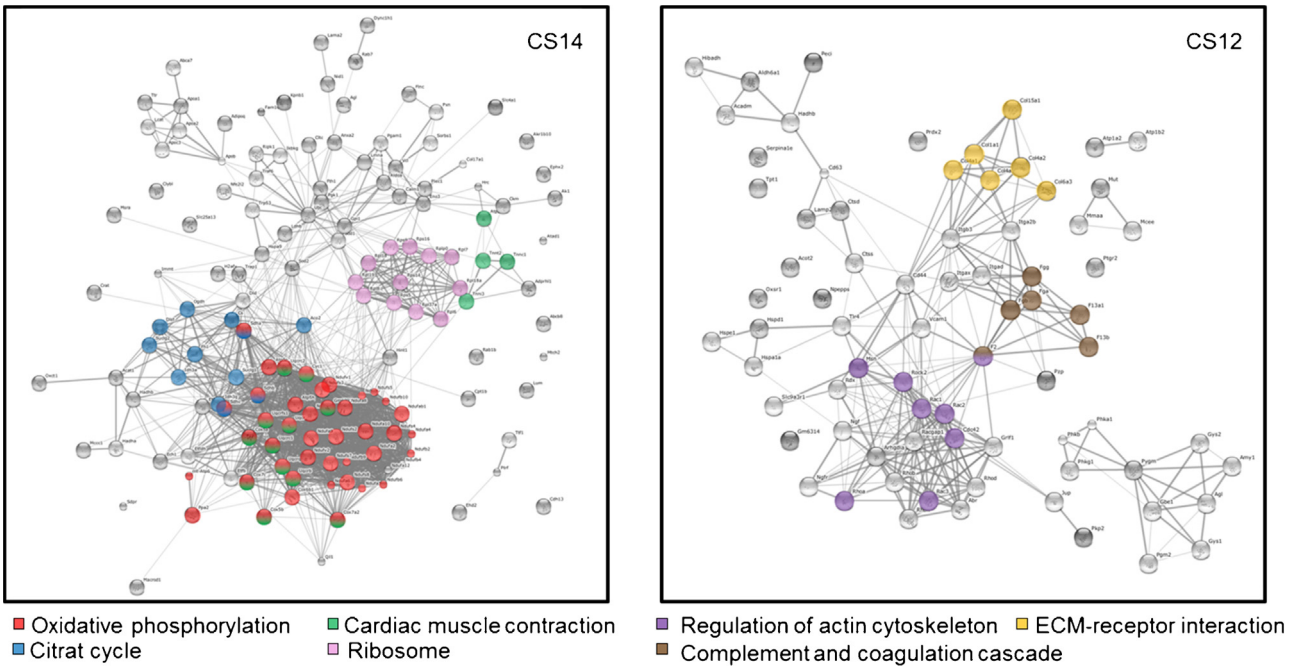

Fig. 8 – Protein interaction network generated in the CS14 and CS12 proteome (males) by the cis-trans and trans-trans variant proteins found. Protein classes dominating the network are indicated.

4. Discussion

In our study of the heart proteome of ten consomic mouse strains we consider each strain as an "individual" because each strain possesses its own set of polymorphisms contributed by the consomic PWD chromosome. Our question was to which extent different sets of polymorphisms modify the proteomic network of interaction toward conditions which differ in robustness against perturbations, in particular against complex diseases such as heart hypertrophy. In other words, could the particular set of SNPs inherent in the genome of each CS lead to strain specific networks of protein interactions which act as global disease modifiers with different capacities in protecting against the disease?

Consomic mouse strains offer an optimal model system to investigate the key question in this problem: to which extent does a polymorphic protein, i.e. a cis-variant protein coded by the PWD chromosome of the CS, affect the proteins from the B6 chromosomes by trans-variation. Our observation was that only a few cis-variants occurred in each CS, but able to alter quantitatively a several-fold greater number of B6 proteins by trans-variation. However, the response of the different CSs to the polymorphisms was of quite different extent (e.g. CS1 – 12 variants versus CS14 – 80 variants, females) and did not correlate with the frequency of the variant genes of the PWD chromosomes (Table 1). Apparently, the genetic effect of the CSs led to a proteome network effect and this to an extent characteristic for each CS. According to our results this may be explained on the basis of core-proteins, i.e. by the frequency and types of core-proteins that were affected by trans-variation in each CS. In consequence, the conditions in the proteomic network of each strain may be quite different, leading to differences in the buffer capacity of proteins, in the potential of the proteomic network for alternative pathways, in the kind of modification of proteins and in other conditions. As a consequence, homeostasis in the proteomic system may not necessarily be maintained in each CS and, therefore, proteomic conditions may result, which are restricted in robustness against perturbations.

Robustness can be defined as a property that allows a system to maintain its function despite external or internal perturbations and has been observed from the level of gene transcription to the level of systemic homeostasis [13]. Robustness has been explained by using the bow-tie structure as a model [13]. The bow-tie structure is represented by a highly conserved and highly connected core cluster interfaced with less connected in- and out-clusters. The robustness against input processes such as disease conditions depends on the flexibility of the core cluster, i.e. on the ability of the core processes to offer alternative pathways or to compensate over-expression or down regulation of gene products due to mutations. The output is then the maintenance of specific functionalities and homeostasis in the system [13]. In terms of our investigation performed on the proteome of the mouse heart the core cluster would be given by the core proteins, as published by Burkard et al. [43] and Weiss et al. [44] and detected in our protein patterns to a high extent. The "input" in our proteome system would be the different sets of polymorphisms given by the different PWD chromosomes of the different CSs. The "output" is then given by the cis- and trans-variant proteins which occur in each CS. The robustness of the proteome then depends on the flexibility of the core proteome. The high abundant and highly connected core proteins have the capacity to compensate proteome-wide variation in up- and down-regulation and maintain in this way homeostasis in

Table 1 – Correlation of polymorphisms and variant proteins. Frequency of variant proteins detected in the heart proteome of the consomic strains (CS) versus B6 compared with the frequency of SNPs identified in genes of the corresponding CS-chromosomes PWD.

CS-strains	SNPs in genes of CS-chromosomes[a]						Variant proteins in CSs	
	Coding-non-synonymous SNPs		Non-coding SNPs: locus-region + mRNA-UTR		Total SNPs including introns			
A. Males								
CS 1	171	22.41%	1224	23.64%	14,554	21.61%	29	7.97%
CS 5	106	13.89%	622	12.01%	9140	13.57%	39	10.71%
CS 9	83	10.88%	612	11.82%	6796	10.09%	44	12.09%
CS12	67	8.78%	372	7.19%	5999	8.91%	26	7.14%
CS14	87	11.40%	659	12.73%	9216	13.68%	69	18.96%
CS15	60	7.86%	526	10.16%	7548	11.21%	42	11.54%
CS16	47	6.16%	349	6.74%	4359	6.47%	29	7.97%
CS17	78	10.22%	422	8.15%	5040	7.48%	44	12.09%
CS19	64	8.39%	391	7.55%	4696	6.97%	29	7.97%
CSY	0	0.00%	0	0.00%	0	0.00%	13	3.57%
Correlation variant proteins – SNPs		$r = 0.319$ $p = 0.369$		$r = 0.366$ $p = 0.298$		$r = 0.422$ $p = 0.224$		
B. Females								
CS 1	171	22.41%	1224	23.64%	14,554	21.61%	12	3.20%
CS 5	106	13.89%	622	12.01%	9140	13.57%	40	10.67%
CS 9	83	10.88%	612	11.82%	6796	10.09%	55	14.67%
CS12	67	8.78%	372	7.19%	5999	8.91%	23	6.13%
CS14	87	11.40%	659	12.73%	9216	13.68%	80	21.33%
CS15	60	7.86%	526	10.16%	7548	11.21%	47	12.53%
CS16	47	6.16%	349	6.74%	4359	6.47%	40	10.67%
CS17	78	10.22%	422	8.15%	5040	7.48%	56	14.93%
CS19	64	8.39%	391	7.55%	4696	6.97%	22	5.87%
Correlation variant proteins – SNPs		$r = -0.320$ $p = 0.402$		$r = -0.228$ $p = 0.556$		$r = -0.191$ $p = 0.622$		

[a] Mouse Genomics Informatics (MGI) Database.

the proteome. However, whereas homeostasis is a stabilizing factor necessary to keep protein expression and interaction in balance, enough flexibility in the system must be maintained to allow the proteomic network to react against perturbation [48]. The scope of variability in the proteome above homeostasis would then determine the capacity of robustness inherent in a proteome (Table 2).

We found in our consomic mouse model relatively low variability in the heart proteome of CS1 and CS12, but high variability in CS14. Treatment of CS12 and CS14 by TAC, that induced cardiac pressure overload, showed that CS12 was strongly affected, but in CS14 cardiac pressure was within the normal range. In addition, an interesting observation was the high variability of Myh6 in the myosin spot complex in

Table 2 – Variant mRNAs in consomic strains. Frequency of variant mRNAs in the heart transcriptome of consomic mouse strains (CS) compared to the B6 strain and related to the number of genes per CS-chromosome PWD; males.

CS-strains	Genes/Chr[a]		Variant mRNAs in CSs			
			Cis-variant		Trans-variant	
CS 5	1300	19.90%	70	28.11%	132	13.94%
CS12	838	12.83%	33	13.25%	81	8.55%
CS14	1101	16.86%	29	11.65%	190	20.06%
CS15	793	12.14%	28	11.24%	247	26.08%
CS16	671	10.27%	25	10.04%	109	11.51%
CS17	1102	16.87%	37	14.86%	42	4.44%
CS19	727	11.13%	27	10.84%	146	15.42%
Correlation variant mRNAs – genes				$r = 0.796$ $p = 0.032$		$r = -0.259$ $p = 0.575$

[a] Mouse Genomics Informatics (MGI) Database.

the 2-DE protein pattern of CS14. Myh6 is a protein tightly involved in regulation of heart contraction and function [49]. Myh6 maps on the consomic Chr 14 and may contribute to the robustness of the CS14 heart.

According to our hypothesis the high variability of proteins found in CS14 reflects a high flexibility of the proteomic network based on core proteins and may explain the robustness of CS14 against the induced cardiac pressure overload. We studied this effect with tools of bioinformatics. The comprehensive data obtained by LC–MS/MS were analyzed. The results showed that the variant proteins found in CS14 were connected with many other proteins resulting in a complex network (compared to random networks). Interestingly, 45% of the primarily variant proteins were core proteins, which are known to be highly connected (hubs) with other proteins [43]. Moreover, in CS14 to a high extent, the proteins were connected by "shortest-path" distances. These characteristics of the CS14 proteome contrast with those found for the CS12 proteome (see Results) and suggest that the genetic background (the polymorphisms introduced) in CS14 offered the conditions which allow to establish homeostasis in the proteomic network, but also flexibility to maintain robustness in the system.

When we looked for the functional protein classes characterizing the sub-nets, resulting from the interaction of the trans-variant proteins with related proteins, a quite different situation was found in CS14 versus CS12 (Fig. 8). In CS14 the proteins connected were closely related to cardiac contraction. Taking this observation into account in addition to the close connection of proteins in CS14 by "shortest-path", that may allow fast reaction to perturbation, another feature of CS14 may explain the robustness of this strain against cardiac pressure overload. However, additional consomic strains and sub-strains (congenic strains) have to be investigated to draw a general conclusion from these experiments.

Whereas the polymorphisms in the "genetic architecture" (defined by Stranger et al. [4]) of a genome may lead to individual differences in the degree of robustness of the proteomic system, mechanisms are involved to reduce genetic variation in order to keep the system stable. According to Jarosz et al. [19] these mechanisms include: (1) functional redundancy by duplication of chromosome regions or single genes, (2) functional redundant pathways and epistasis, and (3) the manner in which the genome is organized into networks of interacting proteins and functional modules. Picotti et al. [50] have shown in a comprehensive proteome analysis of genetically different strains of yeast that epistatic interactions between loci affecting protein levels are a common phenomenon and may partly explain "missing heritability" observed in traditional association studies [51]. Our results obtained from the proteome analysis of CSs and from the F1 (PWDxB6) generation show an impressive reduction of genetic variation of proteins. Although F1 has the polymorphisms of all the 20 PWD chromosomes (B6/PWD heterozygotes) that should result in a large number (700–800) of variant proteins as calculated from the ten consomic strains investigated (Fig. 4), the number of variants in F1 was found to be below 100. Epistasis defined as, "the situation in which the alleles at one gene cover up or alter the expression of alleles at another gene" probably compensates to a large extent over-expression and down regulation of proteins in order to keep the proteomic system in homeostasis. Shao et al. [52] investigated phenotypes of blood, bone and metabolic traits on consomic mouse strains B6-Chr$^{A/J}$. The results showed that the overall phenotypic difference between the parental strains was much less than the sum of the phenotypic differences attributable to individual chromosome substitutions. This has been explained as an effect of epistasis on quantitative trait loci. These results support our finding at the proteome level and demonstrate the consequences of the global proteomic arrangement for the phenotypic traits. It would be interesting to find out whether particular classes of proteins are preferentially restricted in variation. It is unlikely that this process acts randomly. If the expression of polymorphic genes is, to some extent, selectively normalized, this would also contribute to the individualization of the genetic architecture of each proteome.

The results from our mouse model lead to the conclusion that protein expression and interaction in a wild population differs proteome-wide between individuals due to the individual genetically variant proteins (polymorphisms) which alter many other proteins by trans-variation. However, as a consequence, homeostasis may not become established perfectly in the proteome of each individual and this may restrict the robustness of the system against perturbations. Testing different conditions of the heart proteome in response to cardiac pressure overload suggests that the proteomic network acts as a global disease modifier. High variability in the proteomic network may indicate high buffer capacity in case of perturbation.

Conflict of interest

The authors declare no conflict of interest.

Transparency document

The Transparency document associated with this article can be found in the online version.

Acknowledgements

This work was supported by Research Grant Nos. KL237/11-2 and AOBJ:587765 (Forschergruppe) from the Deutsche Forschungsgemeinschaft (DFG). This work was also supported by grants from the BCRT obtained from the German Federal Ministry of Education and Research (BMBF). The authors thank Dr. Jiri Forejt, Academy of Sciences of the Czech Rpublic, Prague for supporting us with consomic animals, and we thank Dr. Marie-Laure Yaspo, Max Planck Institute for Molecular Genetics, Berlin and Prof. Dr. Matthias Selbach, Max-Delbrück-Centrum for Molecular Medicine, Berlin for cooperation.

REFERENCES

[1] Chen H, Hwang H, McKee LA, Perez JN, Regan JA, Constantopoulos E, et al. Temporal and morphological

impact of pressure overload in transgenic FHC mice. Front Physiol 2013;4:205.
[2] Ortlepp JR, Vosberg HP, Reith S, Ohme F, Mahon NG, Schroder D, et al. Genetic polymorphisms in the renin–angiotensin–aldosterone system associated with expression of left ventricular hypertrophy in hypertrophic cardiomyopathy: a study of five polymorphic genes in a family with a disease causing mutation in the myosin binding protein C gene. Heart 2002;87:270–5.
[3] Griendling KK, Murphy TJ, Alexander RW. Molecular biology of the renin–angiotensin system. Circulation 1993;87:1816–28.
[4] Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics 2011;187:367–83.
[5] Johnson T, Gaunt TR, Newhouse SJ, Padmanabhan S, Tomaszewski M, Kumari M, et al. Blood pressure loci identified with a gene-centric array. Am J Hum Genet 2011;89:688–700.
[6] Munroe P, Johnson T, Caulfield MJ. The genetic architecture of blood pressure variation. Curr Cardiovasc Risk Rep 2009;3:418–25.
[7] Mao L. Genetic background specific hypoxia resistance in rat is correlated with balanced activation of a cross-chromosomal genetic network centering on physiological homeostasis. Front Genet 2012;3:208.
[8] Stoll M, Cowley Jr AW, Tonellato PJ, Greene AS, Kaldunski ML, Roman RJ, et al. A genomic-systems biology map for cardiovascular function. Science 2001;294:1723–6.
[9] Tian Z, Greene AS, Usa K, Matus IR, Bauwens J, Pietrusz JL, et al. Renal regional proteomes in young Dahl salt-sensitive rats. Hypertension 2008;51:899–904.
[10] Cowley Jr AW. Genomics and homeostasis. Am J Physiol Regul Integr Comp Physiol 2003;284:R611–27.
[11] Vidal M, Cusick ME, Barabasi AL. Interactome networks and human disease. Cell 2011;144:986–98.
[12] Suk EK, McEwen GK, Duitama J, Nowick K, Schulz S, Palczewski S, et al. A comprehensively molecular haplotype-resolved genome of a European individual. Genome Res 2011;21:1672–85.
[13] Kitano H. Biological robustness. Nat Rev Genet 2004;5:826–37.
[14] Kitano H. Towards a theory of biological robustness. Mol Syst Biol 2007;3:137.
[15] Wu L, Candille SI, Choi Y, Xie D, Jiang L, Li-Pook-Than J, et al. Variation and genetic control of protein abundance in humans. Nature 2013;499:79–82.
[16] Ma W, Trusina A, El-Samad H, Lim WA, Tang C. Defining network topologies that can achieve biochemical adaptation. Cell 2009;138:760–73.
[17] Waldherr S, Streif S, Allgower F. Design of biomolecular network modifications to achieve adaptation. IET Syst Biol 2012;6:223–31.
[18] Forler S, Klein O, Klose J. Individualized proteomics. J Proteomics 2014:56–61.
[19] Jarosz DF, Taipale M, Lindquist S. Protein homeostasis and the phenotypic manifestation of genetic diversity: principles and mechanisms. Annu Rev Genet 2010;44:189–216.
[20] Kitano H. Violations of robustness trade-offs. Mol Syst Biol 2010;6:384.
[21] Barkai N, Leibler S. Robustness in simple biochemical networks. Nature 1997;387:913–7.
[22] Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis 1995;16:1034–59.
[23] Gregorova S, Divina P, Storchova R, Trachtulec Z, Fotopulosova V, Svenson KL, et al. Mouse consomic strains: exploiting genetic divergence between Mus m. musculus and Mus m. domesticus subspecies. Genome Res 2008;18:509–15.
[24] U.S. Office of Science and Technology Policy. Laboratory animal welfare; U.S. government principles for the utilization and care of vertebrate animals used in testing, research and training; notice. Fed Regist 1985;50:20864–5.
[25] Barrick CJ, Rojas M, Schoonhoven R, Smyth SS, Threadgill DW. Cardiac response to pressure overload in 129S1/SvImJ and C57BL/6J mice: temporal- and background-dependent development of concentric left ventricular hypertrophy. Am J Physiol Heart Circ Physiol 2007;292:H2119–30.
[26] Hartl D, Irmler M, Romer I, Mader MT, Mao L, Zabel C, et al. Transcriptome and proteome analysis of early embryonic mouse brain development. Proteomics 2008;8:1257–65.
[27] Zabel C, Sagi D, Kaindl AM, Steireif N, Klare Y, Mao L, et al. Comparative proteomics in neurodegenerative and non-neurodegenerative diseases suggest nodal point proteins in regulatory networking. J Proteome Res 2006;5:1948–58.
[28] Klose J, Nock C, Herrmann M, Stuhler K, Marcus K, Bluggel M, et al. Genetic analysis of the mouse brain proteome. Nat Genet 2002;30:385–93.
[29] Berth M, Moser FM, Kolbe M, Bernhardt J. The state of the art in the analysis of two-dimensional gel electrophoresis images. Appl Microbiol Biotechnol 2007;76:1223–43.
[30] Klein O, Rohwer N, de Molina KF, Mergler S, Wessendorf P, Herrmann M, et al. Application of two-dimensional gel-based mass spectrometry to functionally dissect resistance to targeted cancer therapy. Proteomics Clin Appl 2013;7:813–24.
[31] Zabel C, Mao L, Woodman B, Rohe M, Wacker MA, Klare Y, et al. A large number of protein expression changes occur early in life and precede phenotype onset in a mouse model for Huntington disease. Mol Cell Proteomics 2009;8:720–34.
[32] Nebrich G, Herrmann M, Sagi D, Klose J, Giavalisco P. High MS-compatibility of silver nitrate-stained protein spots from 2-DE gels using ZipPlates and AnchorChips for successful protein identification. Electrophoresis 2007;28:1607–14.
[33] Luber CA, Cox J, Lauterbach H, Fancke B, Selbach M, Tschopp J, et al. Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity 2010;32:279–89.
[34] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 2008;26:1367–72.
[35] Cox J, Matic I, Hilger M, Nagaraj N, Selbach M, Olsen JV, et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc 2009;4:698–705.
[36] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116–21.
[37] Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res 2011;10:1794–805.
[38] Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 2014;42:D199–205.
[39] Borodina T, Adjaye J, Sultan M. A strand-specific library preparation protocol for RNA sequencing. Methods Enzymol 2011;500:79–98.
[40] Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 2010;26:589–95.
[41] Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621–8.

[42] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.
[43] Burkard TR, Planyavsky M, Kaupe I, Breitwieser FP, Burckstummer T, Bennett KL, et al. Initial characterization of the human central proteome. BMC Syst Biol 2011;5:17.
[44] Weiss M, Schrimpf S, Hengartner MO, Lercher MJ, von Mering C. Shotgun proteomics data from multiple organisms reveals remarkable quantitative conservation of the eukaryotic core proteome. Proteomics 2010;10:1297–306.
[45] Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. Proc Natl Acad Sci U S A 2007;104:8685–90.
[46] Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, et al. A compendium of gene expression in normal human tissues. Physiol Genomics 2001;7:97–104.
[47] Xu X, Coats JK, Yang CF, Wang A, Ahmed OM, Alvarado M, et al. Modular genetic control of sexually dimorphic behaviors. Cell 2012;148:596–607.
[48] Gaspar ME, Csermely P. Rigidity and flexibility of biological networks. Brief Funct Genomics 2012;11:443–56.
[49] Citro L, Naidu S, Hassan F, Kuppusamy ML, Kuppusamy P, Angelos MG, et al. Comparison of human induced pluripotent stem-cell derived cardiomyocytes with human mesenchymal stem cells following acute myocardial infarction. PLoS ONE 2014;9:e116281.
[50] Picotti P, Clement-Ziza M, Lam H, Campbell DS, Schmidt A, Deutsch EW, et al. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature 2013;494:266–70.
[51] Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci U S A 2012;109:1193–8.
[52] Shao H, Burrage LC, Sinasac DS, Hill AE, Ernest SR, O'Brien W, et al. Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. Proc Natl Acad Sci U S A 2008;105:19910–4.