Expression profiling of human idiopathic dilated cardiomyopathy

Rafal Grzeskowiak\textsuperscript{a}, Henning Witt\textsuperscript{a}, Mario Drungowski\textsuperscript{a}, Rolf Thermann\textsuperscript{b}, Steffen Hennig\textsuperscript{a}, Andreas Perrot\textsuperscript{c}, Karl J. Osterziel\textsuperscript{c}, Dirk Klingbiel\textsuperscript{a}, Stefanie Scheid\textsuperscript{a}, Rainer Spang\textsuperscript{a},

Hans Lehrach\textsuperscript{a},* Patricia Ruiz\textsuperscript{a},*

\textsuperscript{a}Max-Planck Institute for Molecular Genetics, Ihnestr. 73, 14195 Berlin, Germany
\textsuperscript{b}Anadys Pharmaceuticals Europe GmbH, Meyerhofstr. 1, 69117 Heidelberg, Germany
\textsuperscript{c}Universitätsklinikum Charité/Kardiologie, Campus Buch & Virchow-Klinikum, Humboldt-Universität zu Berlin, Wiltbergstrasse 50, 13125 Berlin, Germany

Received 12 December 2002; accepted 7 April 2003

Abstract

Objective: To investigate the global changes accompanying human dilated cardiomyopathy (DCM) we performed a large-scale expression screen using myocardial biopsies from a group of DCM patients with moderate heart failure. By hierarchical clustering and functional annotation of the deregulated genes we examined extensive changes in the cellular and molecular processes associated to DCM.

Methods: The expression profiles were obtained using a whole genome covering library (UniGene RZPD1) comprising 30,336 cDNA clones and amplified RNA from myocardial biopsies from 10 DCM patients in comparison to tissue samples from four non-failing, healthy donors. Results: By setting stringent selection criteria 364 differentially expressed, sequence-verified non-redundant transcripts were identified with a false discovery rate of \(<0.001\). Numerous genes and ESTs were identified representing previously recognised, as well as novel DCM-associated transcripts. Many of them were found to be upregulated and involved in cardiomyocyte energetics, muscle contraction or signalling. Two hundred and twenty deregulated transcripts were functionally annotated and hierarchically clustered providing an insight into the pathophysiology of DCM. Data was validated using the MLP-deficient mouse, in which several differentially expressed transcripts identified in the human DCM biopsies could be confirmed. Conclusions: We report the first genome-wide expression profile analysis using cardiac biopsies from DCM patients at various stages of the disease. Although there is a diversity of links between the cytoskeleton and the initiation of DCM, we speculate that genes implicated in intracellular signalling and in muscle contraction are associated with early stages of the disease. Altogether this study represents the most comprehensive and inclusive molecular portrait of human cardiomyopathy to date.

© 2003 European Society of Cardiology. Published by Elsevier B.V. All rights reserved.

Keywords: UniGene; Cardiomyopathy; Cardiac biopsy; Expression profile; cDNA array

1. Introduction

Dilated cardiomyopathy (DCM), a major cause of human heart failure, is characterised by a progressive dilation primarily of the left ventricle [1]. In \(\sim30\%\) of DCM patients the disorder has been attributed to mutations in proteins of the sarcomere, the cytoskeleton or the cell/nuclear membrane (e.g., cardiac \(\beta\) myosin heavy chain, desmin or lamin A/C) [2,3]. Recently, the importance of cytoskeletal proteins (e.g., muscle-specific LIM-domain protein, desmin or dystrophin) has been emphasised in the pathogenesis of DCM [4,5]. Increased biomechanical stress on myocytes has been assumed to generate persistent signals in ventricular hypertrophy, long-term dilation and heart failure. Hypertrophic and apoptotic programs might be activated concomitantly, and the balance between these two may determine whether chamber dilation occurs or not (e.g., Gq, p38 MAP kinases or cardiotoxpin 1) [4,6]. Finally, the dysregulation of...
calcium cycling and calcium related signalling (e.g., related to phospholamban or calcineurin pathways) has been speculated as a further decisive factor underlying the pathogenesis of DCM [7].

During the development of heart failure cardiac myocytes undergo remodelling processes. While initially compensatory, they ultimately accelerate functional deterioration and the onset of cardiac failure. The analysis of such complex networks demands global gene expression approaches, e.g., subtractive hybridisation, large-scale sequencing, expressed-sequence-tag analysis, serial gene analysis or differential display. Genome-wide expression profiling has proven to be a very fruitful approach in deciphering complex changes in the transcriptome, translating these into alterations in cell structure, signalling and metabolism [8–10]. Although some expression profiling studies on heart failure have been reported these attempts provide rather a preliminary picture of the disease pathogenesis. They rely: (i) on samples from failing hearts from terminal patients, (ii) on restricted cDNA libraries with a limited number of clones and (iii) on a small number of patients [11–14]. The current study addresses these issues and expands previous findings by examining the gene expression profiles of myocardial biopsies from 10 clinically stable DCM patients using a whole-genome-covering cDNA library.

The UniGene RZPD1 (Resource Center Primary Database set 1) library represents one of the most comprehensive, non-redundant cDNA collections available to date [15]. Clones were selected and re-arrayed to produce a non-redundant clone set (for details see the Section 2). The RZPD1 library includes a substantial percentage of functionally unassigned cDNAs, offering a significant capability for gene discovery. The small myocardial biopsies used in this study were obtained during routine cardiac catheterisation from different DCM patients at various stages of the disease. This allowed a snapshot expression analysis of the pathophysiology of DCM providing a comprehensive and inclusive picture of heart failure pathogenesis.

Our results point to global changes in the cytoskeleton, in cellular energetics, and in calcium cycling and apoptotic pathways as being essential processes underlying dilative remodelling and cardiac dysfunction. We further demonstrate the applicability of gene expression profiling on cardiac biopsies, which may have important implications for future approaches in clinical diagnostics.

2. Methods

2.1. Human and mouse tissues samples

Myocardial biopsies (2–3 mm³) were obtained in the Charité/ Franz-Volhard-Klinik (Berlin, Germany) during routine heart catheterisation from clinically stable patients. Biopsies were taken from each patient only once at the timepoint of the first diagnosis of DCM. All patients gave written informed consent for diagnostic cardiac catheterization including endomyocardial biopsies. The indication for invasive examination was the clarification of the etiology of heart failure (e.g., exclusion of coronary artery disease). The most often suspected indication for taking biopsies was myocarditis or suspected storage diseases. In all patients coronary artery disease, myocarditis and storage diseases (including amyloidosis) were excluded. In addition, there were no positive family histories in these patients and no disease causing mutations were uncovered. Therefore, the diagnosis of idiopathic dilated cardiomyopathy (DCM) was made in all patients [1]. They received standard medication according to the ESC guidelines [16]. Patients in NYHA II–IV received angiotensin converting enzyme (ACE)-inhibitors, β-adrenergic blockers and diuretics. Patient A (NYHA I) received only ACE-inhibitors. Only endomyocardial samples, which were no longer needed for diagnostic purposes, were used for expression analysis. Biopsies were taken from each patient only once at the time-point of first diagnosis of DCM from the posterior-lateral wall of the left ventricle, flash frozen in liquid nitrogen and stored at −80°C. Gene expression and clinical data were pseudonymised. Control samples were obtained from four non-failing, healthy donor hearts considered unsuitable for transplantation because of technical reasons.

Total RNA was isolated from each biopsy in parallel using the standard Trizol method and subjected to two rounds of in vitro amplification using a modified version of the T7-based protocol [17] (MegaScript, Ambion). The concentration and integrity of the amplified RNA (aRNA) from each biopsy was measured using a Bioanalysier (Agilent Technologies) (Fig. 1 in the Methods Appendix, available on http://www.elsevier.com/homepage/sab/cardio/doc/Grzeskowiak_methods_appendix.pdf).

Whole heart samples from 12-month-old muscle LIM-domain protein (MLP) −/− mice and control littermates were kindly provided by Dr. E. Ehler (ETH, Zürich). As reported previously, 10-week-old MLP −/− mice show 100% prevalence of DCM as indicated by parameters such as end-diastolic diameter (4.97 mm), left ventricular posterior wall thickness (0.51 mm), and heart-to-body weight ratio (0.0097) [18]. Total RNA was isolated, and cDNA synthesised using a MMLV reverse transcriptase.

This investigation conforms with the Guide for the Care and Use of Laboratory Animals (NIH, 1996) and with the principles outlined in the Declaration of Helsinki (1997).

2.2. Evaluation of the RNA amplification protocol

Cells from two different human cell lines, dermal fibroblasts (F) and fetal kidney fibroblasts (N) were grown under standard cell culture conditions. Total RNA was isolated and 100 ng were used for either one or two rounds
of the T7-based amplification protocol [17]. The integrity, size range and concentration of the T-RNAs and the aRNAs were evaluated using the Agilent Bioanalyzer and agarose gel electrophoresis (Fig. 2 in the Methods Appendix). T-RNA (20 µg) and aRNA (100 ng) from both cell lines were independently radiolabeled and hybridised to nylon membranes carrying 21 888 cDNA clones (a subset of the UniGene RZPD1). After background subtraction, the intensities were normalised and differentially expressed transcripts were selected according to a $P$ value $<0.001$, a signal-to-background ratio $>3$ and a regulation ratio $>3$.

2.3. Production of cDNA arrays, hybridisation and sequence verification of cDNA clones

The human UniGene RZPD1 set containing 30 336 clones (http://www.rzpd.de, Resource Center and Primary Database homepage) was generated on the basis of the NCBI UniGene clustering. For each cluster a consensus sequence was computed and a representative clone selected. Selected clones were re-arrayed to produce a minimal clone set. Clones were selected by the RZPD on the basis of following criteria: (i) vitality; (ii) absence of contaminations with bacteriophages or other microorganisms; (iii) clone position close to the 3’ end; (iv) clone length between 500 bp and 1.2 kb; and (v) presence of 3’ and 5’ reads of the same clone within one cluster to ensure clone identity. cDNAs were PCR amplified using universal primers and the products were spotted onto 22×22-cm nylon membranes (Hybond N+, Amersham) using in-house robots. The quality of the PCR products was monitored by gel electrophoresis and that of the arrays by hybridisation of a PCR-primer-specific probe (see Methods Appendix). After background subtraction 94% of the cDNA clones showed a normal signal intensity distribution and produced mean signal-to-background ratios $>3$ (Fig. 3 in the Methods Appendix).

For hybridisations, 100 ng of aRNA was used per cDNA array. The aRNAs were primed with random hexamers and reverse transcribed in the presence of [α-32P]dCTP. The resulting radioactive cDNA was purified using Sephadex columns, denatured and added to denatured salmon sperm and human placenta DNA as blocking reagents. To aid grid positioning during image analysis 1 ng of radioactively labelled kanamycin cDNA was added to the complex probe. This mixture was added to 10 ml of hybridisation buffer (1 M NaCl, 1% SDS, 0.1×SSC) and incubated overnight at 65 °C. After three 20-min washes at 65 °C (0.1% SDS, 0.1×SSC), the filters were exposed for approximately 6 h on Fuji BAS screens and signal intensities read using a Fuji Bas-1800 reader (Fuji Photo Film). For expression profiling, radioactively labelled cDNA probes from each patient and pooled controls were hybridised, in parallel, onto arrays from the same PCR- and spotting-filter batch. Each hybridisation was performed twice using independently labelled cDNA targets (see experimental design in Fig. 1). For each patient the mean expression ratio for each gene was calculated.

The 655 differentially expressed cDNA clones were rearrayed. Bacterial clones were grown over night and plasmid DNAs used as sequencing templates together with vector-specific primers. Samples were precipitated and loaded onto ABI 3700 machines and the sequences obtained were analysed using BLASTN [19] against publicly available databases (GenEmbl, SwissProt, dbest, GeneOntology).

2.4. Data analysis, filter normalisation and statistical significance

Signal intensities for every cDNA clone and background spots were determined using the Visual Grid software (GPC-Biotech, Munich). Raw intensities were background corrected by subtracting the local background (estimated block-wise) and normalised. In brief, for each cDNA clone, a signal-to-noise ratio was calculated and signal intensities were log-transformed. Subsequently, an average intensity over all arrays was calculated and the clones were sorted accordingly. A moving average with a fixed length window (±50) was calculated for each filter to finally obtain an average curve. Then, the distortions were compensated by subtracting the filter-wise moving average curve from the intensities and adding the moving average of the means as previously described [20] (for details see the Methods Appendix).

To select differentially expressed transcripts we used the absolute difference between the groups measured in terms of the $P$ value of a two-sample t-test, the signal to background ratios, and the ratios of average expression levels between the DCM and the control group. The calculated $P$ values reflect the biological as well as the experimental variation. A transcript was determined as differentially expressed if both the corresponding signal-to-background ratio and the control-to-DCM ratio were $>3$, and if the corresponding $P$ value was less than 0.001, i.e., the transcript was deregulated in all patients. In addition, the false discovery rate (FDR) was calculated by random permutation of the class labels and applying the rejection criteria to the permuted data. The FDR estimates the proportion of genes falsely called differentially expressed among all genes called differentially expressed. The approach used was a minor modification of the widely applied SAM programme [20], accounting for the characteristics of an expression profiling approach using radio-labeled probes and nylon membranes (for details see the Methods Appendix).

2.5. Quantitative RT-PCR

Gene-specific primers for mouse and human genes were designed with GCG Prime software to generate 50–100-bp
Fig. 1. Experimental design. A set of two nylon membranes carrying 34,176 cDNA clones (UniGene RZPD1) was used to hybridise radiolabelled representations of RNAs from control and DCM samples, respectively. An overlap of 3,840 clones between both membranes provided internal controls. Each patient's expression profile ($n=10$) was compared to that of pooled non-failing heart, control samples ($n=4$) in two independent experiments. Differentially expressed genes in all patients were identified when a given transcript fulfilled the following criteria: regulation ratio $>3$, signal/background ratio $>3$, and $P$ value $<0.001$.

3. Results

3.1. Gene expression profiling of human idiopathic DCM

The clinical data obtained to characterise the patients are shown in Table 1. All DCM and control RNAs were amplified and those showing the required concentration, size-range and integrity after amplification were further processed (four control and 10 DCM samples, $n=14$). Following two rounds of amplification aRNAs ranged between 150 and 2500 bp in size and the efficiency of amplification obtained was ~1000-fold (Figs. 1 and 2 in the Methods Appendix, available on http://www.elsevier.com/homepage/sab/cardio/doc/Grzeskowiak_methods_appendix.pdf). To assess the linearity and reproducibility of the T7 amplification protocol we compared the gene expression profiles obtained using long amplicons (a complete list of primers is given in the Methods Appendix). All DCM and control aRNAs were pooled independently and 5 μg from each were used to generate cDNA templates (100 μl). The GeneAmp, 5700 Sequence Detection System (Applied Biosystems) was then used: 1 μl of cDNA from mouse and human DCM and control samples, respectively, was used in triplicate reactions (96-well format) with 1 μl of 10 pM gene specific primers and 25 μl SYBR GreenMix. After a first cycle of 10 min at 95 °C, the cycling conditions were 95 °C for 15 s and 60 °C for 1 min. The transcript levels were standardised to an appropriate internal control. Several routinely used house keeping genes (e.g., β-actin, hypoxanthine phosphoribosyltransferase, HPRT, glyceraldehyde-3-phosphate dehydrogenase, GAPDH) were tested, but excluded for the standardization experiments due to their deregulation in cardiac tissues. The gene encoding acidic ribosomal phosphoprotein PO (Arbp), however fulfilled the required criteria and was used as an internal control in each measurement. To control the product specificity, dissociation curves for each gene were evaluated. All experiments included no template, no RT and no primer controls to monitor reaction specificity and were repeated three times independently.
Table 1
Clinical data of the analysed patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age (years)</th>
<th>BMI (kg/m²)</th>
<th>EF (%)</th>
<th>FS (%)</th>
<th>LVEDD/BSA (mm/m²)</th>
<th>NYHA</th>
<th>Septum (mm)</th>
<th>A_mean (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>31</td>
<td>44</td>
<td>28</td>
<td>27</td>
<td>I</td>
<td>11</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>39</td>
<td>47</td>
<td>20</td>
<td>29</td>
<td>II</td>
<td>11</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>35</td>
<td>49</td>
<td>17</td>
<td>29</td>
<td>III</td>
<td>11</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>24</td>
<td>33</td>
<td>21</td>
<td>30</td>
<td>III</td>
<td>11</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>19</td>
<td>28</td>
<td>25</td>
<td>32</td>
<td>III</td>
<td>11</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>24</td>
<td>55</td>
<td>14</td>
<td>36</td>
<td>III</td>
<td>10</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>29</td>
<td>34</td>
<td>n</td>
<td>39</td>
<td>III</td>
<td>6</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>28</td>
<td>19</td>
<td>16</td>
<td>33</td>
<td>II</td>
<td>13</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>24</td>
<td>19</td>
<td>6</td>
<td>32</td>
<td>IV</td>
<td>10</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>26</td>
<td>20</td>
<td>14</td>
<td>42</td>
<td>III</td>
<td>12</td>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>

Ten DCM patients (A–J) are shown. The parameters measured were body mass index (BMI), ejection fraction (EF), fractional shortening (FS), left ventricular end-diastolic diameter/body surface area (LVEDD/BSA), New York Heart Association classification (NYHA; I–IV), intraventricular septum thickness (septum) and mean aortic pressure (A_mean); n, value not determined.

either total RNA (T-RNA) or amplified aRNA from two different human cell lines, dermal fibroblasts and fetal kidney fibroblasts. Expression profiling on 21,888 cDNA clones (a subset of the UniGene RZPD1) was performed. Four hundred and five and 395 clones were found to be differentially expressed between both cell lines when using T-RNA and aRNA, respectively. Notably, 305 (77%) of these clones were common to both approaches (Fig. 4 in the Methods Appendix). Therefore, we assumed adequate levels of reproducibility and linearity of the amplification protocol in our experiments, which agrees with previous reports using similar protocols [21].

Since our experimental design included an overlap of 3840 clones between the filters used (Fig. 1), internal controls for reproducibility and experimental variation were provided. When the regulation ratios, the signal-to-background ratios and the P values of this subset of clones were compared, we found that >90% of the clones produced comparable results. Thus indicating a high degree of reproducibility in our experiments (Fig. 5 in the Methods Appendix).

3.2. Selection of differentially expressed genes

The number of deregulated transcripts was calculated by comparing the distribution of the expression data to the distributions obtained by randomly permuting the class labels (DCM versus controls). We observed wide changes in gene expression levels between control and DCM samples. The selection criteria insured that the transcripts were indeed expressed (signal/background ratio >3), that there was a sufficiently high fold change in expression (>3) and a sufficiently high absolute change measured by the P value of a gene-wise t-test (P<0.001). Applied to the data, this procedure yielded 655 (~1.9%) differentially expressed clones in all patients, from which 459 were up- and 196 down-regulated. To verify that the criteria chosen were stringent enough, and that we could safely assume that the transcripts in this list were genuinely differentially expressed, we calculated the FDR of each transcript. After 2000 permutations we obtained a FDR of 0.0006, expecting 0.39=0.0006×655 to be false-positives.

The 655 differentially expressed transcripts were re-sequenced, and after excluding redundant clones representing the same transcript, we assembled a list of 364 non-redundant, differentially expressed genes. From these, 144 were ESTs and represented putative novel (i.e., so far not related to the disease) transcripts associated to DCM (the list of ESTs is given in Data Supplement). To annotate these ESTs, we blasted them against several public databases: human Ensembl, UniGene (human, mouse and rat), swplus (swissprot plus trembl), human dbest and Gene Ontology. One hundred and eighteen showed high BLAST scores (>10e-10) and 50% were found in the human Ensembl gene collection, providing additional information, e.g., locus, disease information and gene ontology (Data Supplement available on http://www.elsevier.com/homepage/sab/cardio/doc/Grzeskowiak_supplement.pdf).

To exclude that the changes in gene expression observed in the samples analysed are biased by the heterogenic composition of the biopsies we determined the variance of the expression levels between several cell-type specific genes. Signal intensities produced by cDNA clones representing neuronal-, endothelial-, fibroblast-, ECM- and muscle-specific genes were analysed. The normalised signal intensities of the selected genes were transformed to obtain a mean of zero and brought to a relative scale (in %). To visualise the variance and the distribution of the data, box plots were drawn. Each box represents 50% of the background corrected, normalised intensities of a given gene in all 10 DCM samples, and its boundary represents the inter-quartile range (IQR). The distribution of the signal intensities obtained for several genes such as occludin, Von Willebrand factor, vinculin, fibronectin, vimentin, cardiac troponin I (TNNI3), and cardiac myosin (heavy polypeptide 7, MYH7) were comparable in all biopsies (Fig. 2). The plots for most of the transcripts show a small variance, and even for those with a slightly
increased variability (e.g., occludin or vinculin) the intensities are distributed symmetrically to the median. In addition, cell-type-specific genes were grouped into four categories: cardiac myocyte, neuronal, endothelium and fibroblast, and their relative intensity percentage was calculated for all samples (Fig. 3). When the values were compared, we found that the most frequent cell type (>90%) in all the samples analysed were cardiac myocytes.

3.3. Validation of array data by quantitative PCR

Using quantitative PCR (Q-PCR) we determined the expression levels of 23 randomly selected genes in pooled DCM and control samples, respectively. The differential expression observed was confirmed for all genes tested. The transcripts analysed included candidate genes involved in essential cardiomyocyte cell processes (e.g., requiem, four-and-half LIM domain proteins 1 and 2) (Fig. 4A). Remarkably, not only genes with high expression levels were validated (e.g., skeletal and cardiac actin or myosin light chain 2), but also those producing low signal intensities on the arrays (e.g., requiem, GDP dissociation inhibitor 1, lipoprotein lipase or CD81) indicating an optimal sensitivity of our array approach. More importantly, the regulation ratios obtained with both techniques (arrays and Q-PCR) were comparable.

As a validation of the array data obtained and the statistical analysis applied, we determined the gene expression levels of several of the identified genes in a well-established model for DCM, the MLP sk mice [18]. The differential expression of eight randomly chosen genes (e.g., ZASP, calumenin, smpx and requiem) could also be observed in MLP knockout mice in comparison to control littermates (Fig. 4B).

3.4. Gene expression patterns associated to idiopathic DCM

From the differentially expressed and sequence verified transcripts, a functional annotation was found for 222 genes. The classification was based on the information available in Gene Cards (http://bioinfo.weizmann.ac.il/cards/index.html, GeneCards home page) and in the Celera Discovery System (http://cds.celera.com, Celera home page) databases. This allowed a global view on the pathophysiology of heart failure (Fig. 5). The largest group of genes consistently up-regulated in DCM patients were those involved in general cell growth and maintenance: energy pathways (e.g., TCA cycle, ATP synthesis), electron transport and lipid metabolism (e.g., lipid hydrolysis, fatty acid transport and oxidation). A similar degree of deregulation was also observed for genes contained in the ‘muscle contraction’ group (e.g., cytoskeletal components and regulators of calcium oscillations), as well as in the
Fig. 3. Myocytes are the most frequent cell type in DCM biopsies. Signal intensities from the corresponding cell type specific genes were grouped into four categories (cardiac myocyte, neurones, endothelium and fibroblast) and their relative abundance in each biopsy was calculated. Bars represent the percentage of cell-type specific gene expression of each group in each patients’ biopsy and in the pooled control samples. Note that in all biopsies the highest contribution obtained was of muscle-specific genes (i.e., cardiac troponin 1, myosin heavy chain 7, and myosin light chain 2) (>95%).

Fig. 4. Verification of array data by quantitative PCR (Q-PCR). The graph depicts the regulation ratios obtained using Q-PCR on human (DCM patients versus non-failing heart controls) and mouse (MLP −/− versus control littermates) samples. Mean values from three independent, normalised measurements are shown. (A) White and grey bars represent the regulation ratios obtained for 23 randomly selected genes by the array analysis and by the Q-PCR approach, respectively. (B) Bars represent the regulation ratios obtained for eight selected genes using cardiac samples from MLP −/− mice as compared to control littermates. (C) List of genes in alphabetical order. Note that the regulation ratios obtained with both approaches are very similar.
‘cell communication’ group (e.g., genes associated to ‘cell–cell/matrix adhesion’ and ‘intracellular signalling’).

To gain a comprehensive insight into the cellular processes deregulated in DCM, we grouped the expression profiles using a hierarchical clustering algorithm [22]. The log-transformed mean regulation ratios for each patient’s profile are displayed in the standard green/red colour coding (Fig. 6 and Data Supplement, available on http://www.elsevier.com/homepage/sub/cardio/doc/Grzeskowiak_supplement.pdf). Cellular energetics and muscle contraction appear to be the major processes activated in human DCM showing a strong upregulation in all patients. The largest number of up-regulated genes was found in the ‘energy pathways’ group: enzymes involved in glycolysis (e.g., P-fruktokinase, triose isomerase and P-glycerate kinase) and in the TCA cycle (e.g., succinate dehydrogenase units and succinate CoA ligase), as well as numerous proteins involved in oxidative phosphorylation. Likewise, overexpression of various mitochondrial carriers was also observed (e.g., SLC25A4, SLC25A1 and SLC25A11). The ‘lipid metabolism’ group contained genes involved in lipid degradation and fatty acid oxidation (e.g., lipoprotein lipase, fatty acid binding protein, acyl CoA isomerase and CoA reductase). An increase in expression levels of most of these transcripts was observed.

Cellular responses like activation of energy production and enhanced assembly of the contractile apparatus seem to be involved in the initial improvement of heart contractility and function. These compensatory responses are however followed by processes known to deteriorate cell and organ function. Accordingly, we observed the upregulation of several genes related to apoptosis (e.g., voltage-dependent anion channel 1 and 2, prostaglandin synthase D and kallikrein 11) and defence response (e.g., major histocompatibility complex I A, complement 1Q, peroxiredoxin 1, 3, 5 and annexin), with a simultaneous downregulation of several cell-cycle control genes (e.g., LPS-induced TNF factor α (PIG7), cyclin D binding Myb-like transcription factor 1 (DMTF1) and cyclin-dependent kinase 5, regulatory subunit 1 (CDK5R1)) (Fig. 6).

4. Discussion

4.1. A genome wide approach on cardiac biopsies from DCM patients

Despite recent advances in the understanding of heart failure in general and in dilated cardiomyopathy in particular, its complex pathophysiology, especially at the molecular level, remains poorly understood. Global analy-
sis of gene expression has proven to be a fruitful means for examining disease processes, and some initial studies have been attempted in human failing hearts [11–14]. The main limitations of these studies lie in the under-represented and biased cDNA array collections, the low number of patients analysed and the usage of cardiac samples from end-stage heart failure patients. We have addressed these issues by genome wide profiling cardiac biopsies from a group of DCM patients with moderate heart failure. To date, this is the largest and most unbiased gene set interrogated in an array-based cardiovascular study, offering a great potential for novel gene discovery.

The myocardial biopsies analysed were obtained from clinically well defined, stable patients. This allowed a more homogeneous transcriptome snapshot in contrast to the analysis of unstable patients with terminal heart failure. One of the limitations of employing cardiac biopsies, however, lies in their cellular heterogeneity, i.e., muscular and endothelial as well as fibrotic components can be obtained. We excluded that the changes in gene expression profiles were biased by the heterogenic composition of the biopsies, since the signal intensities obtained for cell-type specific genes (e.g., as cardiac troponin, fibronectin or Von Willebrand factor) were comparable in all biopsies (Fig. 6).
Furthermore, we found that the most frequent cell type in all samples were cardiac myocytes and that the contribution of fibrotic components in the biopsies was rather low (Fig. 3). Since cardiomyocytes are the most dominant cell type in the biopsies analysed, the upregulated transcription was mainly detected in this cell type. Altogether, we demonstrate the applicability and the reliability of the combination of an RNA amplification protocol and gene expression profiling on cardiac biopsies, which might have important implications for future approaches in clinical diagnostics.

4.2. Identification of putatively novel transcripts associated to human cardiomyopathies

The dissection of the gene expression profiles allowed an insight into several molecular events (e.g., an energy metabolism switch, apoptosis, and cessation of the cell cycle) that ultimately might directly contribute to the demise of heart function. The obtained data were validated using an independent model system, the MLP-deficient mouse, in which several differentially expressed transcripts identified in human DCM biopsies could be confirmed (Fig. 4B). This suggests that the MLP-deficient mice might serve as a valuable tool both for clinical evaluation of possible treatments, and for monitoring the preclinical onset of the disease.

Studies in human and mouse models of DCM, in which the heart chamber markedly enlarges, point to a critical role for muscle-cell Z-disc components in chamber dilatation, suggesting that the sensor might be located at the interface of the cytoskeleton and the sarcomere [3,18]. Recent evidence supports a selective role for the Z-disc protein MLP (muscle LIM protein) in mechanical stretch sensing [22]. The convergence of findings in humans and mouse models might suggest a single unifying molecular pathway. Indeed, in our analysis we have identified several potential candidates, e.g., enigma and myomesin 1 (skelemin, MYOM1). But there is a diversity of links between the cytoskeleton and the initiation of DCM. One of the members of the LIM domain protein family, enigma, was shown to interact with some sarcomeric components [23], as well as with tyrosine kinase receptors [24]. It is therefore tempting to speculate that enigma might be a convergence point between the sarcomere and signalling pathways, being involved in conveying mechanical stimuli into mechanisms responsible for ventricular dilation. Other LIM domain family members identified were four and a half LIM domain protein 1 [25], actin-associated LIM protein [26] and Z-band alternatively spliced PDZ motif (ZASP) [27], representing further putative DCM-associated candidate genes (Fig. 6). Intriguingly is also the upregulation of smpx, a novel stretch-responsive protein, which has been shown to regulate transcription and myocyte muscle hypertrophy [28].

We have observed extensive changes in cardiomyocyte energetics. The vast activation of oxidative ATP production was accompanied by an increase in lipid catabolism. A strong activation was observed for several enzymes involved in lipid breakdown and fatty acid oxidation, the main energy source for the cardiomyocyte. This may indicate a global switch in cardiomyocyte energetics towards a greater rate of lipid oxidation in DCM. Indeed, recent reports indicate that heart failure patients have a higher rate of lipid oxidation and decreased glucose uptake and carbohydrate oxidation [29]. Together with our findings this suggests that the chronic switch in the cardiomyocyte energetics may be an important factor in the progression of heart failure. Likewise, the overexpression of various nuclear-encoded mitochondrial carriers was also observed. Since one of them (adenine nucleotide translocator 4, SL25A4) has already been implicated in DCM pathology [30,31] these genes may indicate the mitochondrial transport system as crucial for heart failure.

Since cell death has been speculated to play a significant role in heart failure [4,32], the observed increase in expression levels of apoptosis-related genes may reinforce this hypothesis. Several profiles of pro-apoptotic genes correlate with disease (e.g., small edkr-rich factor 2, voltage-dependent anion channel 1 and 2, kallikrein 11 and prostaglandin D2 synthase). The initiation of apoptosis is associated with the release of cytochrome c from mitochondria to the cytoplasm and the processing of proteolytic caspases [33]. The dysregulation of VDACs represents an interesting finding since these proteins have been shown to play an essential role in the increase of mitochondrial membrane permeability, and to act as a convergence point for a variety of life-or-death signals [34]. Notably, cell cycle-controlling genes were down-regulated in DCM patients (e.g., anaphase-promoting complex subunit 7, cyclin-dependent kinase 5, cyclin D binding myb-like transcription factor 1 and fms-related tyrosine kinase 1) (Fig. 6). The calcineurin-mediated hypertrophic gene program and the dysregulation of calcium cycling have furthermore been implicated in the pathology of DCM [7,32]. Calsarclin has been shown to interact with the Z-disc protein α-actinin, linking the calcineurin-signalling pathway with the contractile apparatus [35,36]. Interestingly, several genes (e.g., calmodulin 1 and 2, myomesin, calumenin, protein kinase H11 and small muscle protein) show a progressive increase in expression levels, and may therefore be directly associated with the progression of DCM (Fig. 6).

Altogether, hypertrophic and apoptotic programs might be activated concomitantly, and the balance between these two may determine whether chamber dilation occurs or not. Which particular group of genes is critical for the onset and maintenance of DCM has still to be determined. From our data we may, however, speculate that genes with intracellular signalling function are associated with early stages of the disease, genes controlling muscle contraction with intermediate stages, and those associated with apop-
4.3. Outlook and future prospects

The transcripts involved in fundamental metabolic and signalling processes of the cardiomyocyte may indeed represent valuable candidates for diagnosis and eventually therapeutic approaches in human cardiomyopathies. On the basis of our findings, convenient custom-arrays could be developed as a diagnostic tool for a higher number of cardiac disease patients. Certainly a time study with repeated sampling of the same patients should be considered in future work. Furthermore, soluble and/or secreted proteins might be detected in blood samples of patients enabling the development of large, cost-effective and easily applicable screening approaches. Additional attempts will be necessary to further evaluate the significance and the function of the identified transcripts. These will hopefully facilitate a better understanding of the molecular mechanisms leading to DCM and perhaps contribute to its therapy.

Acknowledgements

We thank Drs M.V. Wiles, M.D. Schneider, B. Ivandic and E. Russo for critically reading the manuscript; B. Wulther, A. Sabah and U. Gruber for excellent technical assistance and J. Fontaneda for help with data and sequence analysis. We are also grateful to Dr. M.V. Wiles for help in establishing complex hybridisation protocols and to S. Kaiser (MetaGen) for help with real time quantitative PCR. This work was funded by the German Human Genome Project (BMBF 01KW9914/7) and Charité research grants.

References

against adenine nucleotide translocator in dilated cardiomyopathy and myocarditis—incidence and relation to cardiac function and morphology. Jpn Circ J 1993;57:1150–1158.