Targeted Next-Generation Sequencing for the Molecular Genetic Diagnostics of Cardiomyopathies

Benjamin Meder, MD*; Jan Haas, MS*; Andreas Keller, PhD; Christiane Heid, MS; Steffen Just, PhD; Anne Borries, MS; Valesca Boisguerin, PhD; Maren Scharfenberger-Schmeer, PhD; Peer Stähler; Markus Beier, PhD; Dieter Weichenhan, PhD; Tim M. Strom, MD; Arne Pfeuffer, MD, PhD; Bernhard Korn, PhD; Hugo A. Katus, MD; Wolfgang Rottbauer, MD

**Background**—Today, mutations in more than 30 different genes have been found to cause inherited cardiomyopathies, some associated with very poor prognosis. However, because of the genetic heterogeneity and limitations in throughput and scalability of current diagnostic tools up until now, it is hardly possible to genetically characterize patients with cardiomyopathy in a fast, comprehensive, and cost-efficient manner.

**Methods and Results**—We established an array-based subgenomic enrichment followed by next-generation sequencing to detect mutations in patients with hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM). With this approach, we show that the genomic region of interest can be enriched by a mean factor of 2169 compared with the coverage of the whole genome, resulting in high sequence coverage of selected disease genes and allowing us to define the genetic pathogenesis of cardiomyopathies in a single sequencing run. In 6 patients, we detected disease-causing mutations, 2 microdeletions, and 4 point mutations. Furthermore, we identified several novel nonsynonymous variants, which are predicted to be harmful, and hence, might be potential disease mutations or modifiers for DCM or HCM.

**Conclusions**—The approach presented here allows for the first time a comprehensive genetic screening in patients with hereditary DCM or HCM in a fast and cost-efficient manner. (Circ Cardiovasc Genet. 2011;4:110-122.)

**Key Words:** genetic testing ■ genetics ■ heart diseases ■ cardiomyopathy dilated ■ cardiomyopathy hypertrophic

Because of steadily increasing efforts in genomic research over the past decade, the genetic pathogenesis of many heritable diseases could be resolved. For most of these inherited disorders, more than 1 disease gene could be identified. For instance, inherited forms of heart muscle diseases, so-called cardiomyopathies, can be caused by mutations in at least 30 different genes.1–10 Accordingly, in-depth genetic testing of patients with cardiomyopathy often is not possible with reasonable efforts and costs11; however, this is of immense clinical importance, because some genetic forms of heart muscle diseases are associated with disease manifestation at an early age, an overall poor prognosis, or a high incidence of sudden cardiac death.1,12,13

**Clinical Perspective on p 122**

Introduction and steady improvement of next-generation sequencing (NGS) technologies recently have led to a dramatic increase in parallelism, and hence, sequencing throughput, overcoming limitations of traditional capillary sequencing.14,15 Until now, however, cost-efficient NGS for the diagnosis of inherited diseases has not entered clinical practice, especially because of a lack of efficient reduction of genomic complexity and established protocols. We and others have shown recently that microarray-based sequence enrichment is feasible and highly efficient before sequencing.16–20 Here, we established microarray-based target enrichment followed by SOLiD NGS for the comprehensive and cost-efficient genetic diagnosis of cardiomyopathies.

**Methods**

**Patients and DNA Sample Preparation**

All study participants gave written informed consent, and the study was approved by the Ethic Committee of the University of Heidelberg (Heidelberg, Germany). From each patient, 5 μg of genomic DNA was dissolved in 50 μL of water and fragmented with the

Received July 15, 2010; accepted December 16, 2010.

From the Department of Internal Medicine III (B.M., J.H., C.H., S.J., H.A.K., W.R.), University of Heidelberg; Biomarker Discovery Center (A.K.); febit biomed gmbh (A.B., V.B., P.S., M.B.); and German Cancer Research Center (DKFZ) (M.S.-S., D.W., B.K.), Heidelberg, Germany; Institute of Human Genetics (T.M.S.), Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; Institute of Human Genetics (A.P.), Klinikum rechts der Isar, Technische Universität München, Munich, Germany; and Department of Internal Medicine II (W.R.), University of Ulm, Ulm, Germany.

*Drs Meder and Haas contributed equally to this work.

The online-only Data Supplement is available at http://circgenetics.ahajournals.org/cgi/content/full/CIRCGENETICS.110.958322/DC1.

Correspondence to Wolfgang Rottbauer, MD, Department of Medicine II, University of Ulm, Albert-Einstein-Allee 23, D-89081 Ulm, Germany. E-mail wolfgang.rottbauer@uniklinik-ulm.de

© 2011 American Heart Association, Inc.

Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org DOI: 10.1161/CIRCGENETICS.110.958322

110
Covaris AFA system (Covaris Inc; Woburn, MA) to a size of 100 to 300 bp, as judged by agarose gel electrophoresis. Preparation of the adaptor-ligated DNA library was performed according to the manufacturer’s instructions (ABI SOLiD 3G/H11001 platform; Applied Biosystems; Carlsbad, CA). The adaptor-ligated DNA library was loaded onto an agarose gel, and the size fractions of 200 to 400 bp were excised. Sample concentration and quality was assessed by spectrometry (Nanodrop 1000; Thermo Scientific; Waltham, MA).

Targeted Sequence Enrichment

For the sequence capture array design, current literature was accessed and gene sequences of interest retrieved from reference databases (http://www.ensembl.org and http://genome.ucsc.edu). In addition to known cardiomyopathy disease genes (Table 1), we included 18 novel potential cardiomyopathy genes. The final microarray design covered the coding regions of 47 genes, in total, 0.273 Mb.

Light-activated in situ oligonucleotide synthesis was performed essentially as described, using a digital micromirror device (Texas Instruments; Dallas, TX) for light-directed activation on an activated microfluidic array consisting of a glass/silicon-glass sandwich (febit biomed gmbh; Heidelberg, Germany). Each chip consisted of 8 physically separated arrays with 15 000 individual DNA oligonucleotide features, resulting in a total content of about 120 000 features. For enrichment of the 47 selected genes, 2 arrays were used as an enrichment matrix for each patient sample. Fifty oligonucleotide-long capture probes were distributed over the coding sequences, with an average 9-bp tiling density to target both sense and antisense strands in an alternating manner.

The adaptor-ligated genomic DNA library, which has been amplified by 2 to 12 polymerase chain reaction (PCR) cycles and 500 μg RNA per sample (transferred from baker’s yeast; buffered aqueous solution, 10.5 mg/mL, 5 × 1 mL; Sigma-Aldrich; Bayern, Germany) were dried in a Speed-Vac and dissolved in Hybmix-4 (febit biomed). Hybridization was performed overnight for 16 hours at 50°C, with agitation of the sample. The probes on each microarray were denatured with water at 80°C, and 2 μL of the sample mixture was injected automatically onto the biochip in the HybSelector (febit biomed). Hybridization was performed overnight for 16 hours at 50°C, with agitation of the sample. After hybridization, each array was washed twice with 6 × SSPE at room temperature and 0.5 × SSPE at 45°C. Each array was subsequently washed with 2 mL of SSPE at room temperature. Hybrid-

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Exons</th>
<th>Coding</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTC1</td>
<td>15q14</td>
<td>6</td>
<td>3693</td>
<td>DCM, HCM, ASD, LVNC</td>
</tr>
<tr>
<td>ACTN2</td>
<td>1q43</td>
<td>21</td>
<td>4528</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>CSRP3</td>
<td>11p15.1</td>
<td>5</td>
<td>1341</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>DES</td>
<td>2q35</td>
<td>9</td>
<td>2248</td>
<td>DCM, myopathy</td>
</tr>
<tr>
<td>DMD</td>
<td>Xp21.2</td>
<td>79</td>
<td>14 069</td>
<td>DCM, myopathy</td>
</tr>
<tr>
<td>DSC2</td>
<td>18q12.1</td>
<td>16</td>
<td>5245</td>
<td>ARVC, DCM</td>
</tr>
<tr>
<td>DSG2</td>
<td>18q12.1</td>
<td>15</td>
<td>5653</td>
<td>ARVC, DCM</td>
</tr>
<tr>
<td>DSP</td>
<td>6p24.3</td>
<td>24</td>
<td>9730</td>
<td>DCM, striate palmoplantar keratoderma, woolly hair, ARVC</td>
</tr>
<tr>
<td>ILK</td>
<td>11p15.4</td>
<td>12</td>
<td>1797</td>
<td>DCM</td>
</tr>
<tr>
<td>LDB3</td>
<td>10q23.2</td>
<td>13</td>
<td>5433</td>
<td>HCM, DCM, LVNC, myopathy</td>
</tr>
<tr>
<td>LMNA</td>
<td>1q22</td>
<td>12</td>
<td>3135</td>
<td>DCM, myopathy, CSD, SCID, SVT, VT, AF</td>
</tr>
<tr>
<td>MYBPC3</td>
<td>11p11.2</td>
<td>32</td>
<td>4217</td>
<td>DCM, HCM, LVNC, SCID</td>
</tr>
<tr>
<td>MYH6</td>
<td>14q11.2</td>
<td>37</td>
<td>5937</td>
<td>HCM</td>
</tr>
<tr>
<td>MYH7</td>
<td>14q11.2</td>
<td>38</td>
<td>6030</td>
<td>DCM, HCM</td>
</tr>
<tr>
<td>MYL2</td>
<td>12q24.11</td>
<td>7</td>
<td>829</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>MYL3</td>
<td>3p21.31</td>
<td>6</td>
<td>929</td>
<td>HCM</td>
</tr>
<tr>
<td>MYPN</td>
<td>10q21.3</td>
<td>19</td>
<td>5756</td>
<td>DCM</td>
</tr>
<tr>
<td>NEXN</td>
<td>1p31.1</td>
<td>12</td>
<td>3389</td>
<td>DCM</td>
</tr>
<tr>
<td>PLN</td>
<td>6q22.31</td>
<td>1</td>
<td>1714</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>TA2</td>
<td>Xq28</td>
<td>10</td>
<td>1811</td>
<td>DCM, Barth syndrome, LVNC</td>
</tr>
<tr>
<td>TBX20</td>
<td>7p14.3</td>
<td>8</td>
<td>1823</td>
<td>DCM</td>
</tr>
<tr>
<td>TCFP</td>
<td>17q12</td>
<td>2</td>
<td>962</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>TMPO</td>
<td>12q23.1</td>
<td>9</td>
<td>2701</td>
<td>DCM</td>
</tr>
<tr>
<td>TNCC1</td>
<td>3p21.1</td>
<td>6</td>
<td>695</td>
<td>HCM, DCM</td>
</tr>
<tr>
<td>TNN1</td>
<td>19q13.42</td>
<td>8</td>
<td>842</td>
<td>HCM, recessive DCM, familial restrictive cardiomyopathy</td>
</tr>
<tr>
<td>TNN12</td>
<td>1q32.1</td>
<td>15</td>
<td>1153</td>
<td>HCM, DCM, CSD, AF, SCID</td>
</tr>
<tr>
<td>TPM1</td>
<td>15q22.2</td>
<td>9</td>
<td>1788</td>
<td>HCM, DCM, SCID</td>
</tr>
<tr>
<td>TTN</td>
<td>2q31.2</td>
<td>311</td>
<td>101 515</td>
<td>DCM, HCM, ICD, CSD, arrhythmias</td>
</tr>
<tr>
<td>VCN</td>
<td>10q22.2</td>
<td>21</td>
<td>5282</td>
<td>DCM, HCM</td>
</tr>
</tbody>
</table>

AF indicates atrial fibrillation; ARVC, arrhythmogenic right ventricular cardiomyopathy; CSD, conduction system disease; LVNC, left ventricular noncompaction; SCID, sudden cardiac death; SVT, supraventricular tachycardia; VT, ventricular tachycardia.
Table 2. Patient Characteristics

<table>
<thead>
<tr>
<th>Patient</th>
<th>Sex</th>
<th>Age, y</th>
<th>HTX</th>
<th>NYHA</th>
<th>Disease</th>
<th>Family History</th>
<th>MWT, mm</th>
<th>LVEF, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID3101</td>
<td>Men</td>
<td>46</td>
<td>—</td>
<td>II</td>
<td>DCM</td>
<td>Negative</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>ID3107</td>
<td>Men</td>
<td>21</td>
<td>—</td>
<td>I</td>
<td>HNCM</td>
<td>Positive</td>
<td>20</td>
<td>68</td>
</tr>
<tr>
<td>ID3131</td>
<td>Men</td>
<td>21</td>
<td>—</td>
<td>I</td>
<td>HOCM</td>
<td>Positive</td>
<td>18</td>
<td>71</td>
</tr>
<tr>
<td>ID3180</td>
<td>Women</td>
<td>46</td>
<td>—</td>
<td>I</td>
<td>DCM</td>
<td>Positive</td>
<td>27</td>
<td>65</td>
</tr>
<tr>
<td>ID3214</td>
<td>Women</td>
<td>58</td>
<td>—</td>
<td>II</td>
<td>DCM</td>
<td>Positive</td>
<td>11</td>
<td>45</td>
</tr>
<tr>
<td>ID3236</td>
<td>Men</td>
<td>44</td>
<td>—</td>
<td>II</td>
<td>DCM</td>
<td>Positive</td>
<td>9</td>
<td>38</td>
</tr>
<tr>
<td>ID3283</td>
<td>Women</td>
<td>77</td>
<td>—</td>
<td>II</td>
<td>HNCM</td>
<td>Positive</td>
<td>17</td>
<td>60</td>
</tr>
<tr>
<td>ID3307</td>
<td>Men</td>
<td>63</td>
<td>—</td>
<td>III</td>
<td>DCM</td>
<td>Positive</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>ID3336</td>
<td>Women</td>
<td>36</td>
<td>—</td>
<td>I</td>
<td>DCM</td>
<td>Positive</td>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td>ID3482</td>
<td>Women</td>
<td>11</td>
<td>—</td>
<td>II</td>
<td>HNCM</td>
<td>Positive</td>
<td>16</td>
<td>65</td>
</tr>
</tbody>
</table>

HOCM indicates hypertrophic obstructive cardiomyopathy; HNCM, hypertrophic nonobstructive cardiomyopathy; HTX, history of heart transplantation; NYHA, New York Heart Association functional class; MWT, maximal wall thickness; and LVEF, left ventricular ejection fraction.

NGS Leads to High Sequence Coverage of Cardiomyopathy Genes

Ten consecutive patients with primary cardiomyopathies (5 with HCM and 5 with DCM) were included. In all patients, the genetic pathogenesis was previously unknown. Main patient characteristics are shown in Table 2.

To achieve sufficient sequence coverage of the genes of interest, we performed subgenomic enrichment using oligonucleotide microarrays (for a schematic overview, see Figure 1). Typically, about 0.7 ng of enriched genomic DNA were recovered from the pair of capture arrays from which 0.34 ng were used for NGS using the ABI SOLiD 3G+ sequencing platform. Each sequencing run produced about $3 \times 10^7$ 50-bp-long sequence reads, corresponding roughly to $2 \times 10^9$ bases of total sequence output per patient sample. Next, we analyzed the number of reads that covered the selected coding regions of cardiomyopathy disease genes. Over all patients, 6% to 40% of all filtered reads matched to the target region. The mean number of mappable reads to the region of interest (ROI) was $2.34 \times 10^6$. Mean depth of coverage over all samples was $42 \times$ (range = $34$ to $115$; SEM = $340$) (Figure 2). When taking the mean background depth of coverage into consideration, the ratio of the mean depth of coverage of the ROI to the mean depth of coverage of the background showed that samples were enriched by a factor of 2169 (range = $1345$-fold to $4556$-fold; SEM = $1149$). The mean target coverage...
was 97±1%, demonstrating the high performance of the sequence capture approach. As an example for the efficiency of sequence enrichment, Figure 3 depicts the sequence coverage of part of the desmin gene (DES, MIM 125660) in all 10 patients, whereas Figure 4 shows a randomly selected example showing the sequence coverage of the large titin gene (TTN, MIM 188840; coding length, 100 272 bases). For a comprehensive summary of the obtained sequencing results, see also Table 3.

Sequence Variants and Mutations Within Cardiomyopathy Genes Can Be Reliably Detected by NGS

All successfully mapped sequence reads were analyzed to detect sequence variants, including insertions and deletions. Variants were called automatically, generating a list of sequence variants compared with the reference sequence (hg18; http://genome.ucsc.edu/), which we had annotated with NCBI dbSNP build 128 and an internal database...
Figure 3. Microarray-based target capturing leads to enrichment of selected genomic regions. Shown is an example of the sequence coverage obtained for DES (depicted are exons 2 to 5) in 10 consecutive patients. Array-based enrichment aggregates sequence reads (blue) within the selected genomic regions, resulting in high sequence coverage of selected exons. Intronic regions, which were not enriched, are only poorly covered.
comprising disease mutations derived from HCM and DCM mutation databases (http://angis.org.au/Databases/Heart/ and http://genepath.med.harvard.edu/∼seidman/cg3/index.html). Additionally, all novel nonsynonymous variants were filtered against the March 2010 pilot release of the 1000 Genomes project. All together, 57,548 known or novel variants were detected in the 10 patients. Of these variants, 57,089 (99.2%) are predicted to be noncoding or synonymous, whereas 459 (0.8%) are nonsynonymous, leading to the exchange of 1 or more amino acids (Table 4). We further detected 21,663 small insertions or deletions, of which 809 are located in the ROI. Of all variants, 64.3% were heterozygous, showing 2 alleles at a defined genomic position.

Using our mutation detection workflow, we identified in HCM patient ID3283 a heterozygous microdeletion TT47321263..47321264del within the coding region of myosin-binding protein-3 (MYBPC3) (Figure 5A). The deletion, which is known to cause HCM, leads to a frameshift starting at codon 412 (p.F412fs) and consecutively premature termination of protein translation.25,26 In a second patient with HCM (ID3107),

**Table 3. NGS Run Statistics and Target Coverage**

<table>
<thead>
<tr>
<th>Patient</th>
<th>Total Reads</th>
<th>Filtered Reads</th>
<th>On-Target Reads</th>
<th>Coverage ≥1, % Reads</th>
<th>Coverage ≤10, % Reads</th>
<th>Average Depth of Coverage</th>
<th>Enrichment, Fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID3307</td>
<td>35660379</td>
<td>15158280</td>
<td>2095589</td>
<td>97</td>
<td>91</td>
<td>369.28</td>
<td>1559.14</td>
</tr>
<tr>
<td>ID3180</td>
<td>43623759</td>
<td>13745580</td>
<td>4462848</td>
<td>97</td>
<td>95</td>
<td>786.43</td>
<td>3661.66</td>
</tr>
<tr>
<td>ID3336</td>
<td>39310313</td>
<td>16750585</td>
<td>3088782</td>
<td>97</td>
<td>92</td>
<td>544.30</td>
<td>2079.63</td>
</tr>
<tr>
<td>ID3482</td>
<td>35974866</td>
<td>14593764</td>
<td>836677</td>
<td>97</td>
<td>91</td>
<td>147.44</td>
<td>646.58</td>
</tr>
<tr>
<td>ID3283</td>
<td>29092624</td>
<td>8570555</td>
<td>1630429</td>
<td>97</td>
<td>91</td>
<td>1519.02</td>
<td>2145.47</td>
</tr>
<tr>
<td>ID3101</td>
<td>41229506</td>
<td>16219790</td>
<td>6552519</td>
<td>98</td>
<td>96</td>
<td>1154.67</td>
<td>4556.09</td>
</tr>
<tr>
<td>ID3131</td>
<td>34292542</td>
<td>10930750</td>
<td>2208216</td>
<td>97</td>
<td>92</td>
<td>389.13</td>
<td>2278.35</td>
</tr>
<tr>
<td>ID3107</td>
<td>37564805</td>
<td>9424612</td>
<td>1654445</td>
<td>97</td>
<td>92</td>
<td>291.54</td>
<td>1979.79</td>
</tr>
<tr>
<td>ID3236</td>
<td>17162815</td>
<td>5692160</td>
<td>678629</td>
<td>97</td>
<td>90</td>
<td>119.59</td>
<td>1344.57</td>
</tr>
<tr>
<td>ID3214</td>
<td>6022308</td>
<td>1520326</td>
<td>193349</td>
<td>94</td>
<td>77</td>
<td>34.07</td>
<td>1434.28</td>
</tr>
<tr>
<td>Mean</td>
<td>31993392</td>
<td>11260640</td>
<td>2340148</td>
<td>97</td>
<td>91</td>
<td>361.86</td>
<td>2169</td>
</tr>
</tbody>
</table>
we detected another deletion, CT47313209..47313210del within the MYBPC3 gene, which leads to a frameshift at codon 955 and premature termination of protein translation.25–27 We further identified 2 myosin heavy chain-β (MYH7) missense mutations: G22971762A (exon 5) in HCM patient ID3180 (Figure 6A) and G22968054A in patient ID3482. The first MYH7 mutation is predicted to lead to the exchange of arginine with glutamine at position 143 (p.R143Q), whereas the second leads to the exchange of arginine by a cysteine (p.R453C).28,29 In DCM patient ID3336, we identified a known LMNA nonsense mutation (C154372340T) leading to premature termination of protein translation at amino acid position 321 (p.R321stop) (Figure 6B).23,30 We also detected a variant with uncertain pathogenicity within MYBPC3 (p.R326Q) in patient ID3307.31 In summary,
we detected in 6 patients a known disease mutation for HCM or DCM (Table 5). To test the specificity of variant detection, we sequenced the genomic regions harboring the 6 detected mutations and 44 randomly selected variants (n/H1100550). As shown in Table 6 and Figures 5 through 7, we could successfully verify 6 (100%) of 6 disease mutations and, in total, 43 (86%) of 50 detected variants by capillary sequencing.

To address sensitivity of the presented approach, we screened the coding region with adjacent intronic sequences of MYBPC3 and LMNA by capillary sequencing. From the detected 49 variants, 47 (96%) also were identified by NGS. Only 2 heterozygous variants residing in the intronic region of MYBPC3 were not called by the SNP detection algorithm but were covered with 10 and 13 unique reads, respectively. Further analysis showed that lower-quality indices and not coverage prevented calling of these variants.

Detection of Novel Sequence Variants and Potential Cardiomyopathy Mutations

Besides the detected variants and known cardiomyopathy disease mutations, we identified several novel sequence variants, which are not listed in the NCBI dbSNP or mutation-specific databases. For the 4 patients with HCM in whom we identified a known disease-causing mutation, we detected 75 novel nonsynonymous variants possibly acting as disease modifiers. For 1 patient with HCM (ID3131) in whom we did not find a known mutation, we identified in MYH7 a heterozygous missense mutation T484C leading to the amino acid exchange p.Y162H (Figure 7A). Although this mutation is novel, a mutation at the same amino acid position (p.Y162C) is known to cause HCM.32 For the 3 patients with DCM in whom we could not identify a known disease mutation, 58 novel nonsynonymous variants were detected, possibly causative for DCM. To bioinformatically identify variants that most likely act disease

<table>
<thead>
<tr>
<th>Patient</th>
<th>Mutation</th>
<th>Mutation Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID3180</td>
<td>MYH7 G22971762A</td>
<td>R143Q</td>
</tr>
<tr>
<td>ID3283</td>
<td>MYBPC3 T47321263 ... 47321264del</td>
<td>F412fs</td>
</tr>
<tr>
<td>ID3107</td>
<td>MYBPC3 C747313209 ... 4731210del</td>
<td>P955fs</td>
</tr>
<tr>
<td>ID3482</td>
<td>MYH7 G22968054A</td>
<td>R453C</td>
</tr>
<tr>
<td>ID3336</td>
<td>LMNA C154372340T</td>
<td>R321ter</td>
</tr>
<tr>
<td>ID3307</td>
<td>MYBPC3 G47324447A</td>
<td>R326Q</td>
</tr>
</tbody>
</table>

Table 5. Known Disease Mutations Detected in 10 Patients With Cardiomyopathy

Figure 6. Nonsense mutations within MYH7 and LMNA. A, Sequence alignment of 50-bp SOLiD reads against the reference sequence of MYH7 (top). In HCM patient ID3180, a missense mutation G428A in exon 5 of MYH7 predictably leads to the exchange of arginine with glutamine at position 143 (p.R143Q). A chromatogram confirms the heterozygote missense mutation (bottom). B, Sequence alignment showing a LMNA nonsense mutation in DCM patient ID3336, predictably leading to premature termination of protein translation (top). A chromatogram confirms the heterozygote stop mutation (bottom).
Table 6. Validation of Detected Variants and Mutations

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gene</th>
<th>Reference Position</th>
<th>AA Exchange</th>
<th>Reference Allele</th>
<th>Variant Allele</th>
<th>Confirmed by Sanger Sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID3101</td>
<td>TTN</td>
<td>179352280</td>
<td>S&gt;L</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3101</td>
<td>MYH7</td>
<td>22963165</td>
<td>C905fs</td>
<td>G</td>
<td>—</td>
<td>No</td>
</tr>
<tr>
<td>ID3101</td>
<td>DSG2</td>
<td>27376616</td>
<td>E&gt;K</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3101</td>
<td>PARVB</td>
<td>42726784</td>
<td>W&gt;R</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3101</td>
<td>MYPN</td>
<td>69603927</td>
<td>S&gt;N</td>
<td>G</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>ID3107</td>
<td>LDB3</td>
<td>88483707</td>
<td>—</td>
<td>A</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3107</td>
<td>TEMEM2</td>
<td>73549916</td>
<td>R&gt;H</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3107</td>
<td>TEMEM2</td>
<td>73505380</td>
<td>T&gt;T</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3107</td>
<td>MYBPC3</td>
<td>47313209–47313210</td>
<td>P955fs</td>
<td>CT</td>
<td>—</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3107</td>
<td>TCAP</td>
<td>35075837</td>
<td>—</td>
<td>A</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3131</td>
<td>MYBPC3</td>
<td>47326617</td>
<td>S&gt;G</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3131</td>
<td>MYH7</td>
<td>22971762</td>
<td>R&gt;Q</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3131</td>
<td>TN</td>
<td>179337706</td>
<td>V&gt;M</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3131</td>
<td>MYH6</td>
<td>22931651</td>
<td>V&gt;A</td>
<td>A</td>
<td>G</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3131</td>
<td>TPM1</td>
<td>61138893</td>
<td>—</td>
<td>C</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3214</td>
<td>ILK</td>
<td>6585971</td>
<td>P&gt;L</td>
<td>C</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>ID3214</td>
<td>PARVB</td>
<td>42821201</td>
<td>V&gt;A</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3214</td>
<td>PARVB</td>
<td>42821229</td>
<td>—</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3214</td>
<td>TN</td>
<td>179352280</td>
<td>S&gt;L</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3214</td>
<td>DSP</td>
<td>7525794</td>
<td>R&gt;G</td>
<td>C</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>ID3236</td>
<td>ILK</td>
<td>6585971</td>
<td>P&gt;L</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3236</td>
<td>TEMEM2</td>
<td>73549916</td>
<td>R&gt;H</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3236</td>
<td>TEMEM2</td>
<td>73550029</td>
<td>—</td>
<td>C</td>
<td>G</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3283</td>
<td>DMD</td>
<td>32413115</td>
<td>D&gt;G</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3283</td>
<td>TN</td>
<td>179329196</td>
<td>G&gt;D</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3283</td>
<td>MYBPC3</td>
<td>47321263–47321264</td>
<td>F412fs</td>
<td>TT</td>
<td>—</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3283</td>
<td>DSC2</td>
<td>26903040</td>
<td>I&gt;V</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3283</td>
<td>ACTC1</td>
<td>32868460</td>
<td>—</td>
<td>G</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3307</td>
<td>MYBPC3</td>
<td>47324447</td>
<td>R&gt;Q</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3307</td>
<td>DSG2</td>
<td>27365107</td>
<td>V&gt;I</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3307</td>
<td>PARVB</td>
<td>42726784</td>
<td>W&gt;R</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3307</td>
<td>TN</td>
<td>179163877</td>
<td>P&gt;L</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3307</td>
<td>TRAF6</td>
<td>36473064</td>
<td>—</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3336</td>
<td>LMNA</td>
<td>154372340</td>
<td>R&gt;Stop</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3336</td>
<td>DMD</td>
<td>31406271</td>
<td>R&gt;Q</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3336</td>
<td>HDAC2</td>
<td>114372280</td>
<td>—</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3336</td>
<td>LMNA</td>
<td>154372809</td>
<td>M&gt;T</td>
<td>T</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3336</td>
<td>TNK3</td>
<td>60357396</td>
<td>—</td>
<td>A</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3482</td>
<td>TCAP1</td>
<td>35075837</td>
<td>—</td>
<td>A</td>
<td>C</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3482</td>
<td>MYH7</td>
<td>22968054</td>
<td>R&gt;C</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3482</td>
<td>TNNT2</td>
<td>199599130</td>
<td>E163fs</td>
<td>C</td>
<td>—</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3482</td>
<td>SMYD2</td>
<td>212558105</td>
<td>H&gt;Y</td>
<td>C</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3482</td>
<td>TN</td>
<td>179352280</td>
<td>S&gt;L</td>
<td>G</td>
<td>A</td>
<td>Yes</td>
</tr>
</tbody>
</table>

AA indicates amino acid.
causing, we performed in silico analysis using SIFT. From 188 novel nonsynonymous variants in the 10 patients (58 in patients with DCM without known mutation), 26 were predicted to be damaging (Table 4). For instance, we found a novel variant C209T in 1 patient with DCM (ID3236), leading to the amino acid exchange p.P70L in the integrin-linked kinase gene (ILK) (Figure 7B). Besides the predicted detrimental effect on protein function, this variant is not present in 500 control subjects screened by heteroduplex analysis and cosegregates within the pedigree of the index patient (online-only Data Supplement Figure 1).

**Discussion**

Cardiomyopathies are a genetically heterogeneous group of disorders, hindering comprehensive and affordable genetic diagnostics in clinical practice. We established here high-throughput mutation screening in disease genes for HCM and DCM using microarray-based target enrichment followed by SOLiD NGS, allowing the accurate detection of sequence variants and mutations in multiple disease-relevant genetic loci in parallel.

Comprehensive genetic profiling of patients with cardiomyopathy is largely hindered by the timely and financial efforts needed to genetically screen all known disease genes. Genetic complexity of cardiomyopathies will increase even further because, especially for DCM, the majority of disease genes are not yet defined. Sanger sequencing is currently the gold standard to detect sequence variants within disease genes; however, it faces limitations in throughput and automation. Over the past years, NGS methods have dramatically improved, allowing now the analysis of gigabytes of sequence information in 1 single run. However, in contrast to Sanger sequencing, NGS unselectively analyzes whole genomic DNA, resulting in millions of reads that do not cover the ROIs, in most cases the coding regions of disease genes. To overcome unselective sequencing, we enriched known cardiomyopathy disease regions by microarray-based sequence capturing before NGS. By this approach, we were able to increase the mean depth of coverage of cardiomyopathy genes 2169-fold, allowing detailed and cost-efficient analysis of 1092 disease exons and adjacent intronic regions in 1 NGS run.
Genomic regions with low complexity or GC-rich sequence stretches can be difficult to capture by target enrichment approaches, resulting in low sequence coverage of these ROIs. However, in our study, over all samples, only 3% of ROIs could not be covered, and 91% were covered with at least 10 reads, which allows reliable variant detection on the SOLiD platform. These results will improve even further by optimizing the synthesized capture probes, which usually increase ROI coverage and thereby close the remaining sequencing gaps. Optimization of capture probes and workflow standardization is also likely to reduce interindividual variability of sequence enrichment.

As shown in this article, our target enrichment strategy leads in average to a mean ROI coverage of 412 reads, allowing the reliable detection of sequence variants with high accuracy. Because HCM and DCM show a significant overlap in disease genes, we analyzed genomic DNA from peripheral blood of 5 patients with HCM and 5 with DCM with the same assay, enriched 1092 exons, and identified 1891 sequence variants within these disease regions, of which 349 were nonsynonymous. As shown, we identified mutations in known disease genes in 4 (80%) of the 5 patients with HCM and in 2 (40%) of the 5 patients with DCM. Despite the small sample size, these detection rates are in the expected range for both disorders. To confirm the specificity of our results, we selected 50 variants, including the detected disease mutations, and confirmed their existence by Sanger capillary sequencing. Of all disease mutations, 100% could be confirmed, whereas 86% of all sequence variants, including deletions and insertions, were identified correctly. To address sensitivity, we Sanger sequenced the coding region of MYBPC3 and LMNA in all patients. As shown, our targeted NGS approach also provides good sensitivity (96%), concordant with previous studies using NGS. Only 2 variants were missed by NGS, not because of insufficient coverage, but because of lower-quality indices of sequence reads, a matter likely to be resolved when optimizing variant detection parameters.

The costs of NGS are still substantial in terms of investment in equipment. Nonetheless, the cost per base is substantially lower than for Sanger sequencing, and high parallelism reduces the need for manual handling steps. Hence, the presented NGS-based testing approach seems already competitive in this regard. Although commercially available screening of 4 to 8 HCM/DCM genes usually costs >$4000, the presented approach cost less than $1800 and comprises 47 parameters.

The most pragmatic approach is to compare detected variants against mutation databases, which contain gene mutations already shown to be causative for the respective disease. In our study, we identified within 6 patients with cardiomyopathy established disease-causing mutations. For instance, we identified 2 microdeletions within the MYBPC3 gene, predictably leading to a frameshift and premature termination of protein translation. The detection of insertions and deletions in NGS data has proven to be generally more difficult, particularly because of the relatively short read length typical for most platforms. Accordingly, the false-positive rate of detected insertions and deletions is considerably higher in our study compared to single-nucleotide variations. Here, the combination of paired-end sequencing and advanced assembly and insertion and deletion detection algorithms will further improve performance, although possibly not to levels of sensitivity and specificity that are achievable for single-nucleotide variations.

However, insertions and deletions are still more difficult to detect by other novel genetic testing approaches, such as hybridization-based sequencing chips.

Recent studies identified numerous patients with cardiomyopathy who have more than 1 disease-causing mutation. These mutations can occur in either the same gene (compound heterozygotes) or in different genes (double heterozygotes) in up to 5% of genetic cardiomyopathies. As a consequence of these complex genotypes, the individual risk to develop end-stage disease or sudden cardiac death can strongly increase. As shown, we identified in all patients with HCM and DCM a disease-causing mutation as well as several novel sequence variants that might influence disease onset or progression. However, proving the disease causality of these potential mutations remains challenging. Using bioinformatic tools such as SIFT, one can predict harmful variants under the premise that protein evolution is correlated with protein function. Hence, variants changing conserved residues are more likely to alter or disrupt protein function.

Using SIFT, we could significantly reduce the number of novel nonsynonymous variants that may be disease related. For instance, we found the heterozygous missense mutation Y162H in HCM patient ID3131. Although this mutation is novel, a mutation affecting the same amino acid, Y162C, is known to cause HCM. It was suggested that tyrosine-162 is important for myosin ATPase activity because the tyrosine-162 and glycine-258 are opposite to each other in a domain-domain interface that might be affected when myosin cyclically interacts with actin and ATP. We also detected a novel nonsynonymous variant p.P70L in the ILK gene in a patient with DCM with severely reduced left ventricular ejection fraction. The 21-year-old daughter carrying the same mutation shows a moderately decreased left ventricular ejection fraction. However, as mentioned previously, interpretation of disease contribution remains difficult. From a standpoint of statistical significance, large numbers of cases and controls have to be analyzed to ensure significant associations. Here, deep-resequencing projects like the 1000 Genomes project ambitiously aim to provide more detailed maps of human genetic variation through genotyping at least 1000 human genomes from worldwide populations using NGS technologies. These data will further help to distinguish rare, but benign variants from disease-causing mutations. Additionally, animal models, such as genetically manipulated mice or zebrafish, will contribute to a deeper understanding of the effects of novel sequence variants and compound heterozygosity.
In summary, the presented staged procedure using microarray-based sequence enrichment followed by NGS allowed us to detect cardiomyopathy-causing mutations with high accuracy in a fast and cost-efficient manner that will be suitable for daily clinical practice of genetic testing. Combining our approach with multiplexing NGS will reduce the costs even further for comprehensive diagnosis of multigenic diseases. As pointed out herein, rapid detection of novel gene variants has to be accompanied by additional strategies proving their disease causality.

Acknowledgments

We thank Rouven Nietsch, Sarah Hassel, and Jessica Rudloff for excellent technical assistance.

Sources of Funding

This work was supported by grants from the Bundesministerium für Bildung und Forschung: NGFN-plus (to Drs Meder, Katus, and Rotbauer), NGFN-transfer 01GR0803 (to Drs Meder, Beier, Strom, Pfeuffer, and Korn), 01EX0806, the medical faculty of the University of Heidelberg (to Dr Meder), and the European Union (FP7 INHERITANCE).

Disclosures

Drs Keller, Boisguerin, and Beier, A. Borries, and P. Stähler are employed by febit biomed gmbh.

References

31. Nimmura H, Patton KK, McKenna WJ, Soults J, Maron BJ, Seidman JG, Seidman CE. Sarcomere protein gene mutations in hypertrophic cardio-
Proper genetic diagnosis of patients with primary heart muscle diseases, such as hypertrophic cardiomyopathy (HCM) or dilated cardiomyopathy (DCM), is paramount for individualized patient care, because some genetic forms are associated with disease manifestation at an early age, an overall poor prognosis, or a high incidence of sudden cardiac death. However, comprehensive genetic testing is not offered to all patients because of the genetic heterogeneity of cardiomyopathies as well as limitations in throughput and high costs of current genetic testing. Next-generation sequencing (NGS) technologies can generate gigabases of DNA sequence information in a single sequencing run, now fundamentally changing genomic research. In the present study, we report on the development of a comprehensive genetic testing strategy of patients with cardiomyopathy by an approach using microarray hybridization-based subgenomic enrichment of cardiomyopathy disease genes followed by SOLID NGS in patients with HCM or DCM. With this approach, a large number of disease genes can be assessed by 1 sequencing run within a short time and at acceptable costs. Furthermore, this approach allows identification of potentially disease-causing variants in novel candidate genes of cardiomyopathies. Hence, targeted NGS permits comprehensive genetic screening in patients with inherited cardiomyopathies with implications for clinical practice, potentially enabling physicians to define the individual genetic risk of patients with cardiomyopathy and their relatives.
Targeted Next-Generation Sequencing for the Molecular Genetic Diagnostics of Cardiomyopathies

Benjamin Meder, Jan Haas, Andreas Keller, Christiane Heid, Steffen Just, Anne Borries, Valesca Boisguerin, Maren Scharfenberger-Schmeer, Peer Stähler, Markus Beier, Dieter Weichenhan, Tim M. Strom, Arne Pfeuffer, Bernhard Korn, Hugo A. Katus and Wolfgang Rottbauer

_Circ Cardiovasc Genet._ 2011;4:110-122; originally published online January 20, 2011; doi: 10.1161/CIRCGENETICS.110.958322

_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/4/2/110

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2011/06/21/CIRCGENETICS.110.958322.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Genetics_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation: Cardiovascular Genetics_ is online at:
http://circgenetics.ahajournals.org//subscriptions/
Pedigree of index patient ID3236 carrying the *ILK* p.P70L mutation. The 21-year-old daughter of the index patient also carries the heterozygous *ILK* P70L mutation and has a moderately reduced left ventricular ejection fraction (LVEF). The 23-year-old daughter of the index patient has no signs of DCM and normal LVEF. The 11-year-old son and the 9-year-old daughter report no clinical signs of heart failure. Due to their age they were not accessible (n.a.) for genetic analysis.