Microfluidic Amplification as a Tool for Massive Parallel Sequencing of the Familial Hypercholesterolemia Genes

Silke Hollants,1 Egbert J.W. Redeker,2 and Gert Matthijs1*

BACKGROUND: Familial hypercholesterolemia (FH) is an autosomal dominant disorder that affects cholesterol metabolism and is an important risk factor for heart disease. Three different genes were causally linked to this disorder: LDLR (low density lipoprotein receptor), APOB (apolipoprotein B including Ag(x) antigen), and PCSK9 (proprotein convertase subtilisin/kexin type 9). We evaluated a new amplicon preparation tool for resequencing these genes on next generation sequencing (NGS) platforms.

METHODS: For the 3 genes, 38 primer pairs were designed and loaded on the Fluidigm Access Array, a microfluidic array in which a PCR was performed. We amplified 144 DNA samples (73 positive controls and 71 patient samples) and performed 3 sequencing runs on a GS FLX Titanium system from Roche 454, using pyrosequencing. Data were analyzed with the SeqNext module of the Sequence Pilot software.

RESULT: From the 38 amplicons, 37 were amplified successfully, without any further optimization. Sequencing resulted in a mean coverage of the individual amplicons of 71-fold, 74-fold, and 117-fold for the 3 runs, respectively. In the positive controls, all known mutations were identified. In 29% of the patient samples, a pathogenic point mutation or small deletion/insertion was found. Large rearrangements were not detectable with NGS, but were picked up by multiplex ligation-dependent probe amplification.

CONCLUSIONS: Combining a microfluidic amplification system with massive parallel sequencing is an effective method for mutation scanning in FH patients, which can be implemented in diagnostics. For data analysis, we propose a minimum variant frequency threshold of 20% and a minimum coverage of 25-fold.

Familial hypercholesterolemia (FH) is an autosomal dominant disorder that affects cholesterol metabolism. With a prevalence of 1 in 500 for heterozygotes, FH is among the most common inherited disorders in humans. FH is clinically characterized by increased levels of total cholesterol and LDL cholesterol, leading to skin and tendon xanthomas, arcus cornea, and premature coronary artery disease. Homozygotes, with a prevalence of 1 in a million, have a more severe clinical presentation (1).

Three different genes have been causally linked to autosomal dominant hypercholesterolemia (MIM #143890): the low density lipoprotein receptor gene (LDLR) (MIM #606945), which encodes a cell surface receptor that removes LDL from plasma (1); the apolipoprotein B (including Ag(x) antigen) gene (APOB) (MIM #107730), which encodes the main ligand for the LDL receptor (2); and the proprotein convertase subtilisin/kexin type 9 gene (PCSK9) (MIM #607786), which encodes a serum protease of which overexpression reduces LDLR protein concentrations by a mechanism that is still unclear (3, 4). Most cases of autosomal dominant hypercholesterolemia are caused by a mutation in the LDLR gene, for which more than 1100 variants have been reported. These mutations are listed in the LDLR database of University College London (http://www.ucl.ac.uk/ldlr/LOVDv.1.1.0/) (5). The number of mutations detected in the APOB and PCSK9 genes is substantially lower. Genetic variants in other genes also have been described, but their pathogenic nature is less clear (6, 7).

Identification of the mutation causing the disorder is important for treatment of patients and their
relative patients and mutation carriers receive cholesterol-lowering treatments, mainly to reduce the risk of cardiovascular disease. Mutation analysis is often limited to the LDLR gene for which all 18 exons and a part of the promoter region of the LDLR gene are analyzed for (point) mutations and deletions or duplications and is mostly complemented with an analysis of the hot spots of exons 26 and 29 of the APOB gene, which contain recurrent mutations. Only a few laboratories also screen for mutations in (the 12 exons of) PCSK9. Most laboratories have used mutation detection techniques like denaturing gradient gel electrophoresis, denaturing HPLC, or direct Sanger sequencing for the detection of substitutions, small insertions, and deletions. To detect larger rearrangements, including exonic deletions, a quantitative method like multiplex ligation-dependent probe amplification (MLPA) is essential. Recently, so-called “massive parallel sequencing” platforms have been introduced for the high-throughput sequencing of DNA, which provides dramatically increased speed and sequence capacity and reduces per nucleotide sequencing costs (8–10).

These next generation sequencing (NGS) methods have rapidly gained importance, especially in gene identification in a research context. We sought to develop methods for the high-throughput and parallel sequencing of patient samples in a clinical setting.

In this study, we applied an amplicon-sequencing technology for targeted resequencing of the 3 genes (LDLR, APOB, and PCSK9) involved in FH. We used the GS FLX Titanium system from Roche 454, which uses pyrosequencing to produce sequence reads with a mean length of 400 bp. To prepare the amplicon library, we tested and evaluated the Access Array System from Fluidigm, a new PCR-based enrichment technology. Here we describe the approach and evaluation of the method and provide primer sets and conditions for a swift introduction of this novel mutation-scanning method.

Materials and Methods

PRIMER DESIGN

Twenty pairs of primers were designed to cover a part of the promoter region and all 18 exons of the LDLR gene (RefSeq NM_000527.3), including exon–intron boundaries. To sequence the ligand-binding domain in exon 26 and 29 of the APOB gene (RefSeq NM_000384.2), 3 primer pairs were designed covering codons 3450–3618 from exon 26 and codons 4336–4437 from exon 29. A part of the promoter and all 12 exons of PCSK9 (RefSeq NM_174936.2) including their exon–intron boundaries were covered by 15 primer pairs. In total we generated 38 fragments that varied in size from 321 bp to 355 bp (See Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol58/issue4). We aimed at developing amplicons of approximately the same length to get an optimal sequencing result on the GS FLX.

All primers were designed as described in the Access Array System user guide from Fluidigm. In brief, use of Primer3 was advised (http://frodo.wi.mit.edu/primer3/) (11) with changes to some parameters to get more stringent primer conditions: the “Product Size Ranges” parameter was set on 320–350; “Primer Melting Temperatures” were put on 59.0 °C as a minimum, 60.0 °C as the optimum, and 61.0 °C as a maximum; and “Max Poly-X” was changed to 3. To predict the PCR products we used the UCSC (University of California Santa Cruz) In-Silico PCR tool (http://genome.ucsc.edu/) (12) and to additionally verify them we used the SNPCheck v2.0 program (https://ngrp.manchester.ac.uk/SNPCheckV2/snpcheck.htm).

Besides the target-specific primers, sample-specific primers were designed containing a GS FLX adaptor sequence and a bar-coding sequence or multiplex identifier (MID). These sample-specific primers bind to the target-specific primers through an additional and complementary M13 universal sequence that was added to the 5′ end of the target-specific primers and to the 3′ end of the sample-specific primers.

DNA SAMPLES

The DNA samples used in this study were divided into 3 groups. The first group consisted of 48 patients with previously characterized mutations in the LDLR or APOB gene. These samples were previously tested at the Academic Medical Center in Amsterdam. The second group comprised 25 positive control samples for LDLR or APOB, available at the Center for Human Genetics in Leuven, and the third group comprised 71 new patient samples, which were supplied for molecular diagnostic testing but had not been previously analyzed. The DNA was extracted from peripheral blood leukocytes by use of standard techniques.

FLUIDIGM ACCESS ARRAY

The Fluidigm Access Array is a microfluidic chip on which one can load 48 patient samples and 48 primer pairs. The outcome is a pool of 48 fragments per patient sample. With the incorporation of a unique identifier or barcode for each sample and the necessary sequencing adaptors, it is possible to process all samples simultaneously on the sequencing platform.

Three access arrays were used to amplify the DNA samples. On the first array the 48 positive control samples from Amsterdam were loaded, on the second array the 25 positive control samples of the second group together with 23 new samples were injected, and on the third array the 48 remaining
new patient samples from the third group were loaded. For each experiment, 100 ng DNA per sample was used as input for the system.

The experiments were performed according to the 4-Primer Amplicon Tagging Protocol from the manufacturer. Briefly, the target-specific primers were injected in the primer inlets and the sample-specific primers with their unique MID were loaded together with the DNA samples and the PCR reagents in the sample inlets. The primers and the DNA mixture were then combined in the reaction chambers in the chip. After PCR, 10 μL of the samples were collected from their original wells, now containing a pool of 38 amplicons.

VERIFICATION OF THE PCR PRODUCTS
Before running the samples on the GS FLX machine, we verified the amplification of the fragments by an additional labeling PCR that allowed visualization of the PCR products on a capillary DNA sequencer. The amplification mixture of 25 μL included 1× PCR buffer (platinum Taq, Invitrogen) 200 μmol/L dNTPs, 1.5 mmol/L MgCl₂, 0.5 U platinum Taq DNA polymerase (Invitrogen), 0.2 μmol/L of each FAM-labeled A-adaptor primer and reverse M13 primer, and 0.5 μL of the Access Array PCR product. Thermal cycling consisted of 10 cycles of the PCR program recommended by the manufacturer. From this PCR product, 3 μL was combined with a mixture of 14.625 μL Hi-Di formamide and 0.375 μL GeneScan 500 ROX Size Standard (Applied Biosystems). After a denaturation step at 95 °C for 3 min, we loaded the samples on an ABI 3130xl Genetic Analyzer (Applied Biosystems) to perform fragment analysis and analyzed the data with the GeneMapper v4.0 software (Applied Biosystems).

POOLING AND SEQUENCING OF THE SAMPLES
Amplified patient samples derived from the Fluidigm Access Arrays were measured with the Quant-IT Picogreen dsDNA assay kit (Invitrogen) to compose equimolar pools, each containing a mixture of 8 samples. For each Access Array, 6 pools were obtained, and these were sequenced on one quarter of a GS FLX Titanium PicoTiterPlate device.

MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION
All samples, except the 48 new patient samples that were amplified on the third array, had previously been screened for large rearrangements in the LDLR gene. Therefore, to complete the analysis in all patients, an MLPA was performed on these patients with the Salsa MLPA P062-C1 LDLR kit from MRC-Holland.

DATA ANALYSIS
Sequence reads were mapped against Ensembl reference sequences ENST00000252444 (LDLR), ENST00000233242 (APOB), and ENST00000302118 (PCSK9) with the SeqNext module of the Sequence Pilot software (JSI Medical Systems) to detect variants in the patient samples and in the positive control samples. SeqNext sorted the reads for each sample according to its MID, and variants were called only if the variant frequency was more than 10%. If the variant frequency was more than 20%, the variants were listed in the “distinct table,” otherwise they appeared in the “other table” provided by the SeqNext software. When a pathogenic mutation was detected in a new patient sample, this variant was confirmed with Sanger sequencing.

The GS Amplicon Variant Analyzer software from Roche was used to calculate the coverage—the number of times a base was sequenced—for each sample and each fragment. If a fragment was covered <25 times (a predetermined value that should be used for every amplicon-resequencing experiment), the sequence analysis was repeated with conventional Sanger sequencing to achieve a reliable sequence for all fragments.

RESULTS
SEQUENCING OUTPUT AND COVERAGE
For each Access Array, amplicon sequencing of the 48 samples was performed on one quarter of a GS FLX Titanium run. This resulted in 126 247, 145 207, and 241 268 high-quality sequencing reads for the 3 arrays, respectively. The difference in number of reads between the 3 runs was caused by different percentages of enriched beads obtained after the emulsion PCR amplification. Roche recommends a percentage between 5% and 20%, with the best sequencing results with 8% enriched beads. For our third experiment, with the highest amount of reads, we obtained 8.03% enriched beads. The percentages of enriched beads for the 2 other experiments lay closer to 20%.

From the 144 DNA samples in total, only 3 samples from the first array did not give any sequencing read. These 3 samples had failed to amplify, and on the basis of PCR analysis of 8 randomly chosen FH fragments we determined that the samples were degraded (see online Supplemental Fig. 1). These samples were therefore excluded from further calculations. All other samples and amplicons were amplified successfully, except amplicon 6B from the PCSK9 gene, which corresponds to part of exon 6. The primer pair designed for this fragment did not result in an amplification product in any of the 3 experiments, and the set was therefore excluded from all further calculations. In addition, we repeatedly observed very few sequencing reads in
the reverse direction for exon 9 of PCSK9. The sequencing efficiencies for the other primer pairs did not significantly differ between the forward and reverse directions.

A mean coverage of 71-, 74-, and 117-fold was obtained from each run, respectively. With a preset value of 25-fold minimum coverage required for each amplicon, 4.20%, 3.10%, and only 0.28% (5 of 1776 fragments) of the fragments of the respective arrays were not sufficiently covered and had to be resequenced by Sanger sequencing. The distribution of the coverage from all 3 experiments is illustrated in Fig. 1. Notably, more than 99% of the fragments of the third experiment had a coverage of 40-fold or more.

**MUTATION DETECTION**

The SeqNext module of the Sequence Pilot software was used to identify the variants in the DNA samples. First the positive control samples with known mutations were analyzed in a blinded way. In these samples all types of mutations were included, such as single base pair substitutions and small deletions and/or insertions from 1 up to 23 bp and 2 exonic deletions (see online Supplemental Table 2). Of the 70 control samples, 64 samples had a point mutation or small insertion or deletion in the LDLR gene and 4 had 1 of the 2 common mutations in the APOB gene. No samples were available with a mutation in the PCSK9 gene. All mutations were correctly retrieved, apart from the 2 exonic deletions (deletion of exons 7–8 in 1 patient and exons 7–10 in the other sample). Hence the sensitivity of the assay, as calculated from this series, is \( \geq 96\% \) (at 95% confidence) for point mutations and small insertions and deletions (13).

Subsequently, 71 new patient samples were analyzed for pathogenic mutations. Besides the identification of multiple known polymorphisms in each sample, pathogenic point mutations or small deletions/insertions were found in 22 samples (Table 1). Twenty mutations were listed as pathogenic in the LDLR database or were well-known previously reported APOB variants. Furthermore, a novel frame-shift mutation was identified that results in a premature stop codon. One patient was heterozygous for 2 probable pathogenic mutations: 1 in the LDLR gene that was also reported in the LDLR database and 1 in the APOB gene. This novel APOB mutation was considered pathogenic on the basis of the phylogenetic comparison and the Grantham score; moreover, another mutation in the same codon of the gene was previously described in the literature (14). In the group of 71 new patients, 2 had a familial link to 2 other patients in this cohort. Both individuals (patient 37 and 44) effectively carried the familial mutation. Hence, we found 1 or more mutations in 20 of 69 independent samples (29% of the cases). In total, 17 mutations were found in LDLR (81%) and 4 in APOB (19%). No mutations were detected in PCSK9. In addition, large rearrangements in the LDLR gene were also identified in this cohort by use of MLPA: in 5 samples (7%) an exonic deletion was discovered. In total, 25 independent patients (36%) turned out to be heterozygous for a mutation in 1 of the FH-associated genes.

**Discussion**

The development of massive parallel-sequencing techniques has increased the sequence capacity enormously for both research and diagnostics. In a diagnostic setting, the result is that many patient samples can now be combined in a single sequencing run. Because of that, a lot of PCR reactions have to be performed in parallel and amplicon preparation has become the new bottleneck. In addition, with the rise of NGS, the need for high-throughput amplification methods has grown immensely.

In this study we evaluated a novel parallel amplification technique, the Access Array System from Fluidigm, which can combine 48 patient samples with 48 primer pairs to produce 2304 PCR products in 1 experiment. Familial hypercholesterolemia was chosen as a model, and primer pairs were designed for the 3 genes linked to this disease. In practice, the amplicon mixture of 1 array was loaded on one quarter of a GS FLX Titanium run. In total 70 positive control samples and 71 patient samples were sequenced and further analyzed to detect pathogenic mutations.

Our results showed that the Access Array is an appropriate system for producing an amplicon library for sequencing on the GS FLX platform. Only 3 of 144 DNA samples did not amplify, owing to the bad quality of the DNA. The procedure for designing primers turned out to be efficient. The 38 primer pairs that were developed were directly used in the system. One amplicon (amplicon 6B from PCSK9) has to be redesigned and a second amplicon (amplicon 9 from PCSK9) has to be optimized to reduce the difference in sequencing efficiency between the forward and reverse strand. An additional advantage of this technique is the direct incorporation of the necessary sequencing adaptors and the patient-specific MIDs, which makes a second PCR to attach these sequences redundant and therefore reduces the workload. It is estimated that the whole process of amplification and sequencing takes one-third of the time, and, in our hands, the consumable costs per sample were reduced to about one-sixth the cost compared with performing individual PCRs followed by Sanger sequencing. A rough estimate of the entire costs, including manpower, showed that (only) 9 patient samples per run are sufficient to economically use an Access Array and apply NGS instead of (manual) Sanger sequencing (the depreciation costs of the in-
Fig. 1. Coverage distribution plots of FH patient samples amplified with the Fluidigm Access Array and sequenced on the GS FLX Titanium system.
Instruments were not included in this calculation. Such a calculation will of course vary according to the precise laboratory setting, but given that a full run contains 48 samples, the system almost certainly reduces the diagnostic costs.

We wanted to develop a comprehensive assay for FH and wanted to amplify all exons for the screening. In practice, this meant that if a fragment was not sufficiently covered by use of the NGS platform, it was reanalyzed by Sanger sequencing before the result was reported to the referring physician. Hence, the coverage in NGS analysis is important to avoid additional Sanger analysis. The coverage is also related to the sensitivity: we have to be sure to pick up every mutation in a sample and therefore the minimum coverage needed for each fragment has to be well considered. The lower the coverage, the greater the chance one or both alleles will fail to be sequenced and a mutation thereby missed. However, from a cost-effectiveness point of view, the coverage cannot be too high because increasing the coverage uses more capacity of the GS FLX run, thereby increasing the cost per sample.

We decided to set the necessary minimum coverage at 25-fold. It has been previously calculated that a 38-fold coverage is required to detect heterozygous variants with a probability of 99.9% when the mutation is present in at least 25% of the reads [see De Leeneer et al. (15)]. From our experience we learned that a heterozygous variant is not always present in 25% of the reads, and therefore in our opinion one should re-

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gene</th>
<th>Location</th>
<th>cDNA variant</th>
<th>Protein variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>LDLR</td>
<td>Exon 12</td>
<td>c.1775G&gt;A</td>
<td>p.Gly592Glu</td>
</tr>
<tr>
<td>4</td>
<td>APOB</td>
<td>Exon 29</td>
<td>c.13153C&gt;T</td>
<td>p.Arg4385Cys</td>
</tr>
<tr>
<td>10</td>
<td>LDLR</td>
<td>intron 8</td>
<td>c.1187→10G&gt;A</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.429C&gt;T</td>
<td>p.Cys143X</td>
</tr>
<tr>
<td>13</td>
<td>APOB</td>
<td>Exon 26</td>
<td>c.10579C&gt;T</td>
<td>p.Arg3527Trp</td>
</tr>
<tr>
<td>19</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.662_681dup</td>
<td>p.Glu228ThrfsX44</td>
</tr>
<tr>
<td>20</td>
<td>LDLR</td>
<td>Exon 10</td>
<td>c.1567G&gt;A</td>
<td>p.Val523Met</td>
</tr>
<tr>
<td>21</td>
<td>APOB</td>
<td>Exon 26</td>
<td>c.10580G&gt;A</td>
<td>p.Arg3527Gln</td>
</tr>
<tr>
<td>22</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.530C&gt;T</td>
<td>p.Ser177Leu</td>
</tr>
<tr>
<td>28</td>
<td>LDLR</td>
<td>intron 16</td>
<td>c.2390→2A&gt;G</td>
<td>—</td>
</tr>
<tr>
<td>29</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.662A&gt;G</td>
<td>p.Asp221Gly</td>
</tr>
<tr>
<td>30</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.429C&gt;A</td>
<td>p.Cys143X</td>
</tr>
<tr>
<td>31</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.429C&gt;A</td>
<td>p.Cys143X</td>
</tr>
<tr>
<td>32</td>
<td>LDLR</td>
<td>Exon 9</td>
<td>c.1257C&gt;A</td>
<td>p.Tyr419X</td>
</tr>
<tr>
<td>33</td>
<td>LDLR</td>
<td>Exon 7–8</td>
<td>c.941→1186+?del</td>
<td>—</td>
</tr>
<tr>
<td>34</td>
<td>LDLR</td>
<td>Intron 3</td>
<td>c.314→1G&gt;A</td>
<td>—</td>
</tr>
<tr>
<td>35</td>
<td>LDLR</td>
<td>Prom-exon1</td>
<td>c.1→7_67+?del</td>
<td>—</td>
</tr>
<tr>
<td>36</td>
<td>APOB</td>
<td>Exon 26</td>
<td>c.10580G&gt;A</td>
<td>p.Arg3527Gln</td>
</tr>
<tr>
<td>37b</td>
<td>APOB</td>
<td>Exon 26</td>
<td>c.10580G&gt;A</td>
<td>p.Arg3527Gln</td>
</tr>
<tr>
<td>44b</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.429C&gt;A</td>
<td>p.Cys143X</td>
</tr>
<tr>
<td>47</td>
<td>LDLR</td>
<td>Exon 9</td>
<td>c.1247G&gt;A</td>
<td>p.Arg416Gln</td>
</tr>
<tr>
<td>51</td>
<td>LDLR</td>
<td>Exon 9</td>
<td>c.1257C&gt;A</td>
<td>p.Tyr419X</td>
</tr>
<tr>
<td>54</td>
<td>LDLR</td>
<td>Exon 7–8</td>
<td>c.941→1186+?del</td>
<td>—</td>
</tr>
<tr>
<td>55</td>
<td>LDLR</td>
<td>Exon 7–10</td>
<td>c.941→1586+?del</td>
<td>—</td>
</tr>
<tr>
<td>56</td>
<td>LDLR</td>
<td>Exon 7–10</td>
<td>c.941→1586+?del</td>
<td>—</td>
</tr>
<tr>
<td>62</td>
<td>LDLR</td>
<td>Exon 9</td>
<td>c.1257C&gt;A</td>
<td>p.Tyr419X</td>
</tr>
<tr>
<td>70</td>
<td>LDLR</td>
<td>Exon 4</td>
<td>c.429C&gt;A</td>
<td>p.Cys143X</td>
</tr>
</tbody>
</table>

* These patients are family members of other patients in this list.

Table 1. Mutations in the 71 FH patient samples.

722 Clinical Chemistry 58:4 (2012)
lax this criterion to 20%. Thus, we calculated, using a cumulative binomial distribution, that even a minimum coverage of 21-fold is sufficient to detect a heterozygous variant with a probability of 99.9% that is present in at least 20% of the reads. So based on these theoretical considerations and on the experimental data from our study demonstrating that all mutations were found, our recommendation for a minimum coverage threshold would be 25-fold. In this way we build in a slightly higher level of certainty without inflating the costs.

The threshold for the minimum variant frequency of 20% is an arbitrary number. The SeqNext software offers the possibility to work with different cutoffs: variants with a frequency above 10% are listed in the “other table.” A careful analysis of this table indicated that it is necessary to put the minimum variant frequency below the 20% that was originally put forward in the analysis. For example, the 20-bp duplication in the LDLR gene (c.662_681dup) was found in 2 samples (sample 20 of the positive controls and patient sample 19); the minimum variant frequency was 20%/12% (forward/reverse) in the positive control sample and 14%/21% (forward/reverse) in the patient sample, indicating that the original criteria for detection were met in only one direction. The SeqNext software, however, demands a variant frequency of 20% or higher in both directions (only for deletions and duplications), no matter what proportion, to put the variant in the “distinct table.” So these 2 mutations would have been missed had there not been the possibility to set a second cutoff of 10% through which these variants were listed in the “other table.”

We found a mutation detection rate of 36% in 69 FH families, similar to the detection rate of 36.5% that was described in a study from the UK (16). In our cohort of samples the APOB mutation c.10580G>A was one of the most common, accounting for 8% of the detected mutations, and comparable to the 12% reported in the UK study. We found more mutations in exon 4 of LDLR (27%) than the UK study (18%) but add the caution that their cohort was larger than ours; they tested 232 samples. In our collection of samples all kinds of mutations were present: 77% substitutions, 4% small rearrangements (<100 bp), and 19% large rearrangements (>100 bp); this distribution is slightly different from the one reported in the article about the LDLR database (5), for which the numbers were 65%, 24%, and 11% respectively. We found mutations only in the LDLR and APOB genes, and none in the PCSK9 gene. This result might suggest that mutations in PCSK9 are not an important cause of FH in our population. Similar results have been reported in other populations (17, 18). In contrast, in a French series of patients, 12.5% PCSK9 gene mutations were found in samples without an LDLR or APOB mutation (3).

In conclusion, the proposed workflow of combining a microfluidic amplification system like the Fluidigm Access Array with massive parallel sequencing on the Roche GS FLX is an effective method for mutation scanning in FH patients and could be implemented in a diagnostic setting. Our study demonstrated that all point mutations and small deletions and insertions can be detected if a well-considered minimal coverage threshold is set. For large rearrangements, the use of an additional detection method like MLPA is still obligatory. In theory it should be possible to identify large rearrangements with NGS, but we think that this may not be possible with the Fluidigm workflow. The number of PCR cycles in the protocol is too high to warrant a linear PCR amplification of the fragments. On the other hand, the capacity of the Access Array makes it possible to easily include additional genes in the diagnostic test. We did this with PCSK9. However, for diagnostic purposes it would be useful if the arrays were also available in other formats, such as combinations of fewer patient samples with more amplicons. Such format flexibility would be advantageous to implement the technique for multigenic disorders or to perform sequencing runs on other short read chemistries. Indeed, Fluidigm also offers a protocol for amplicon tagging for the Illumina next generation sequencers for which the proposed workflow is very similar to the one for the GS FLX.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
 Honoraria: None declared.
Research Funding: G. Matthijs (as principal investigator), EU-funded TECHGENE project (2009-223143) and EuroGentest project (FP6-512148).
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We are indebted to Liesbet Vliegen, Vanessa Blys, Harry Cuppens, and Ioris Vermeech for the Genomics Core sequencing services. We thank Geneviève Michils, Marijke Spaepen, and Anniek Corveleyn for critically reading the manuscript and Chris Mattocks for advice on the statistics.
References


