Parallel Copy Number Variation and Sequence Variation Detection Using Amplicon-based Sequencing on the GS Junior System

Introduction

Large genomic duplications and deletions have been recognized as pathogenic mutations for many years, for example in mental retardation disorders or Duchenne muscular dystrophy (DMD), where single or multiple exons are deleted or duplicated, respectively. Copy number changes are also common in tumors. Detection of these changes shows promise for the diagnosis of a disease and also for therapeutic or prognostic purposes. The determination of gene dosage, in combination with the sequence variation detection enabled using 454 Sequencing Systems could prove highly useful in the future of clinical medicine.

Endocrine diseases, like congenital adrenal hyperplasia, multiple endocrine neoplasia type 1 and SHOX (short stature homeobox-containing gene) deficiency, exhibit partial or complete deletions of the corresponding genes. SHOX deficiency is a frequent cause of short stature. Deletions of the SHOX gene or of the pseudoautosomal region 3’ of SHOX are the most important mutations (80%), whereas missense and nonsense mutations are found in 20% of individuals with SHOX deficiency.

For life science research only.
Not for use in diagnostic procedures.
Methods

There are two methods currently used to detect all of the mutations in these genes: the multiplex ligation-dependent probe amplification (MLPA) for detection of deletions and duplications of single genes or gene parts and Sanger-based sequencing for the detection of point mutations and small deletions or insertions. We propose here a single assay that combines the functions of both of these assays into a single sequencing experiment.

Here we present a two-step multiplex PCR-based protocol to detect point mutations and copy number variations simultaneously using the 454 Life Sciences Roche GS Junior System. A similar approach was published recently using the GS FLX System (1). We collaborated with JSI Medical Systems GmbH (Kippenheim, Germany) to develop a software tool for copy number variation detection.

We analyzed six genes (SHOX, MEN1, RET, CASR, GCK and HNF1A) including 48 point mutations and small deletions and insertions, as well as four large deletions and duplications to validate the multiplex PCR approach. All mutations were confirmed by Sanger-based sequencing and MLPA analysis respectively.

Library Preparation

A universal tailed amplicon sequencing design was used for the generation of the amplicon library (Figure 1). Each multiplex PCR mix includes specific fragments for the analyzed genes, as well as control fragments from different chromosomes as references for the copy number analysis, similar to an MLPA probe mix (Figure 2). PCR fragments were amplified using the FastStart High Fidelity PCR System (Roche), according to the manufacturer’s protocol. PCR reactions were purified using 0.8x AMPure XP (Beckman Coulter, Fullerton, CA) and PureLink columns (Invitrogen, Carlsbad, CA) as a second purification step. PCR products were quantified using the Quant-iT PicoGreen kit (Invitrogen, Carlsbad, CA) and pooled in an equimolar ratio.

Figure 1: Universal tailed amplicon multiplex PCR. 1) PCR target specific primers are tagged at the 5’end with a universal sequence (20bp). This tag is targeted in the second PCR. 2) For the second PCR, the fusion primer is tagged by the 454 Sequencing System’s Primer A or Primer B (adaptor and key) plus MID (Multiplex Identifier) for sample identification. 3) The third PCR using Cy5-labeled primers is for quality control of the PCR products.

Figure 2: Multiplex PCRs of the MEN1 gene. A) The entire coding region of the MEN1 gene was amplified in two multiplex PCR reactions. Overlapping PCR products were generated for large exons. Each multiplex PCR reaction includes control fragments for copy number variation analysis. B) Capillary electrophoresis pattern (Beckman CEQ 8800) from a reference sample. Peaks marked by C represent the control fragments. C) Capillary electrophoresis pattern (Beckman CEQ 8800) from a test sample with deletion of exons 5-10 of the MEN1 gene. Peaks marked by an asterisk represent a reduced peak area and peak height and therefore a reduced copy number.
Methods

emPCR Amplification and Sequencing

emPCR amplification was performed as described in the GS Junior Titanium emPCR Amplification Method Manual – Lib-A. 0.8 copies per bead were used for this library. After emPCR amplification, the enriched library was sequenced according to the Sequencing Method Manual using the GS Junior Titanium Sequencing kit.

Data Analysis

Sequence variation and copy number variation analyses were performed with the SEQUENCE Pilot software (JSI Medical Systems, Kippenheim, Germany). For analyzing copy number variations, the median coverage for all control PCR fragments (forward and reverse sequence) is calculated separately for each multiplex PCR mix. In the next step, the relative coverage for each target PCR product is calculated against the median coverage of the control PCR products. These calculations were used for both the reference sample and every test sample to obtain the relative coverage value for each target PCR product. The third step of the calculation is the comparison of the relative coverage from the test sample to the reference sample (Figure 3).

Results

In this study, 138,098 high-quality reads with an average read length of 370 bp were generated with one GS Junior sequencing run. More than 99% of the target fragments were covered by more than 50 reads, with a 40x coverage being sufficient to detect heterozygotes with high confidence. All point mutations, as well as the deletions and insertions of a few base pairs previously detected by Sanger-based sequencing, could be confirmed with the GS Junior System (Table 1).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of mutations</th>
<th>Detection rate (GS Junior)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RET</td>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>MEN1</td>
<td>9</td>
<td>100%</td>
</tr>
<tr>
<td>GCK</td>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>HNF1A</td>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>CASR</td>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>SHOX</td>
<td>3</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1 Detection rate of the GS Junior System. Mutations analyzed include 7 deletions (1-29 bp), 4 insertions (1-5 bp), 1 indel (2 bp deletion/4 bp insertion), 36 missense/nonsense mutations as well as 3 large deletions and insertions respectively.
Copy number variations of single or multiple exons could also be detected by MLPA, as well as by the GS Junior System (Figures 4 and 5).

**Results**

This study shows the application of the GS Junior System for the simultaneous screening of point mutations and copy number variations of multiplex PCR fragments. One GS Junior sequencing run can replace the combination of Sanger-based sequencing and MLPA analysis in a streamlined and cost-effective workflow.

**Conclusions**

This study shows the application of the GS Junior System for the simultaneous screening of point mutations and copy number variations of multiplex PCR fragments. One GS Junior sequencing run can replace the combination of Sanger-based sequencing and MLPA analysis in a streamlined and cost-effective workflow.

**For life science research only. Not for use in diagnostic procedures.**

RESTRICTION ON USE: Purchaser is only authorized to use 454 Sequencing System Instruments with PicoTiterPlate devices supplied by 454 Life Sciences Corporation and in conformity with the operating procedures contained in the 454 Sequencing System manuals and guides.

License Disclaimer information is subject to change or amendment. For current information on license disclaimers for a particular product, please refer to www.454.com.

454, 454 LIFE SCIENCES, 454 SEQUENCING, GS FLX, GS JUNIOR, EMPCR, PICOTITERPLATE, PTP, and FASTSTART are trademarks of Roche. Other brands or product names are trademarks of their respective holders.

---

**References:**

**Important Note:**
Data included in this article are the sole responsibility of the authors that have published them, and these authors are responsible for following the respective local regulations for assay setup and validation.

Regarding the specific assay(s) described, Roche was neither involved in establishing the experimental conditions nor in defining the criteria for the performance of the specific assays. Roche therefore cannot take any responsibility for performance or interpretation of results obtained for the biological target parameter(s) described by the authors or other users using a similar experimental approach.

Potential users are informed to be aware of and in accordance with local regulations for assay validation and the scope of use for the involved Roche products, and to ensure that their use is valid in the countries where the experiments are performed.

---

**Figure 4** Copy number analysis of the MEN1 deletion (screenshots).
The red dotted line marks the predefined 65% limit for deletions. The copy number variation analysis was done with the Next Generation Sequencing CNV module and the MLPA module of the SEQUENCE Pilot software (JSI Medical Systems GmbH, Kippenheim, Germany). A) The two MEN1 multiplex PCR mixes were quantified separately. B) Combined copy number analysis of the GS Junior System reads. The CNV analysis revealed a deletion of exons 5 to 10. C) MLPA analysis of the same test sample with a deletion of MEN1 exons 5-10. Note that the MLPA probe mix contained no probes for exons 4-6. So the deletion was called only for exons 7-10.

**Figure 5:** Deletion and duplication of the SHOX gene (screenshots).
The SHOX gene, PAR1 region, X-chromosome specific fragments, as well as the control fragments, were amplified using six independent multiplex PCR reactions. The red dotted lines mark the pre-defined 65% limit for deletions and the 130% limit for duplications. A) Combined copy number analysis of the GS Junior System reads for the SHOX gene. B) MLPA analysis for the same test sample as above indicating the duplication of a part of the PAR1 region. C) Combined copy number analysis of the GS Junior System reads for a test sample with deletion of the entire SHOX gene. D) MLPA analysis for the entire SHOX gene deletion.