A Simple Multiplex PCR Approach for Target Enrichment in Next-Gen Sequencing

Chris Hebel, Xiaochuan Zhou
2575 W. Belfort, Suite 270
Houston, Texas 77054
Tel. 713 664-7087
Fax 713 664-8181

Multiplexing PCR is a simple way to extract genomic regions of interest from various genetic tests such as variant analysis or genotyping. Somatic mutations such as SNPs are unlikely to be best detected using regular whole genome sequencing and genotyping by sequencing (GBS) in any large genome species requires reduction of genome complexity. Therefore, many current genetic test workflows start with multiplexing PCR to extract genetic marker carrying regions from whole genomes before running hybridization, sequencing, or electrophoresis tests to identify the markers.

We have developed a new multiplexing PCR approach with a significantly simplified workflow and significantly improved robustness. When applied to sequencing target enrichment application, the workflow for producing amplified targets involves only one hands-on step and one PCR run. Sequencing library adapters, sample barcodes and molecular tags are all incorporated during this single PCR run. The approach is designed to require low sample input and to produce superior amplicon uniformity and sequence specificity. The approach involves a novel primer design and a proprietary reaction composition. A PCR run consists of two functionally separated reaction phases, namely target capture and library amplification, without any hands-on step in between.

**A Combination of Multiple Innovations**

Omega Primers™ - A novel primer design that fundamentally improves primer specificity and yet is capable of sustaining certain local template variations such as SNPs.

Relay-PCR™ - An elegant PCR method that combines multiple target capture and monoplex PCR amplification into a single reaction.

Molecular Tags - Used to identify individual PCR template molecules, PCR and sequencing induced biases and errors are detected and then removed.

**New Design Principles** - We have developed new approaches to perform quantitative prediction of PCR primer performance and incorporated into design software.

**New Primers Design**

An Omega Primer™ consists of three functional sections:
1. 5p arm functions as an anchor attaching to a DNA template.
2. 3p arms function as an initiator in a polymerase extension reaction.
3. Loop separates the two arms. In a Relay-PCR the loop is used as a priming section for monoplex PCR.

**One Step Workflow**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Sample preparation</td>
</tr>
<tr>
<td>2.</td>
<td>PCR Reaction</td>
</tr>
<tr>
<td>3.</td>
<td>Next-Gen Sequencing</td>
</tr>
</tbody>
</table>

**Sequencing Ready Library**

<table>
<thead>
<tr>
<th>Library</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>Library primer 1</td>
</tr>
<tr>
<td>Sample 2</td>
<td>Library primer 2</td>
</tr>
<tr>
<td>Target</td>
<td>Library primer 3</td>
</tr>
</tbody>
</table>

**Molecular Tagging**

In NGS applications the use of molecular tags (MTag) to label individual genomic DNA templates prior to PCR amplification provides valuable benefits.

1. Original DNA molecules are uniquely identified, counted, and/or analyzed.
2. PCR and sequencing induced biases and errors are detected and then removed, analyzing less positive results.
3. Variant calling accuracy is increased resulting in improved low allele frequency detection limits.

**Results**

Experimental validation of multiplex Omega Primer Relay-PCR

Relay-PCR using Omega Primers was validated on human genomic DNA. PCR product hand in gel image (A) demonstrates the successful amplification of correct products by capturing six individual targets at six different genome locations. PCR product band in gel image (B) demonstrates the successful amplification of all six targets by mixing six pairs of Omega primers plus one pair of common primers in a single tube. Sequences of the PCR products were validated using a custom microarray (data not shown here).

**Expected PCR product length**

<table>
<thead>
<tr>
<th>Primer Type</th>
<th>Expected Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1:</td>
<td>300 bp</td>
</tr>
<tr>
<td>Type 2:</td>
<td>300 bp</td>
</tr>
<tr>
<td>Type 3:</td>
<td>300 bp</td>
</tr>
</tbody>
</table>

**Reference**


**New Design Principles**

- **Specificity Enhancement**
  - In a typical design, each primer is 16 nt long ensuring that a minimum of 12 out of 16 template nucleotides must match in order to achieve a significant binding, the probability of significant binding in a completely random sequence is 
  \[ P = \frac{n}{4^m} \]
  where \( n \) is the number of matches and \( m \) is the length of the primer.

- **Reduced Primer-Primer Dimer Formation**
  - Type 1: Extendable cross hybridization in 3p arm sections. This will produce an amplifiable product and is the most undesirable primer-primer interaction.
  - Type 2: Non-extendable cross hybridization to any section of the primer. This will produce any extension product and will sequence any great product unless tight primer-primer binding is formed which complies with our target primer-template bindings.
  - Type 3: Extendable cross hybridization in 3p arm or loop sections. This will produce any extension product and will sequence any great product unless tight primer-primer binding is formed which complies with our target primer-template bindings.

**Minimized Allele Drop Out in Variant Regions**

There are 99 high allele frequencies (＞99%) SNPs in a human genome of 30 bp long, an average of 1 SNP in every 30 bp. Omega Primers are designed to work well even when they have to be placed in SNP containing regions. For human CNA sequences, Omega Primers 3p arms are designed to tolerate SNPs above a predetermined allele frequency of 2000 genome database as well as user provided pathogenic variants. We can use this approach to design assays of predictable population coverage.

**Diagram**

- Diagram A: A Simple Multiplex PCR Approach for Target Enrichment in Next-Gen Sequencing
- Diagram B: Expected PCR product length