Introduction

The introduction of second generation sequencers has greatly increased the pace at which sequencing data can be obtained. One of the most common applications for high throughput sequencing is sequence variant or mutation detection in genes or genomic regions of interests, and sequence enrichment methods are employed to target the specific regions of the genome. Recent advances in droplet-based microfluidics technologies allow for the efficient and accurate enrichment of DNA regions of interest. This is accomplished by performing millions of single-plex PCR reactions in parallel within picoliter volume droplets inside a single PCR tube. These highly uniform droplets result in the generation of PCR amplicons in which the PCR success rate is high and the yield for all amplicons is very consistent. This results in data in which the majority of sequencing reads map to the regions of interest (specificity), the number of sequencing reads obtained for all regions is uniform (low bias) and the percentage of target regions in which sequencing reads were obtained is very high (completeness). These attributes result in the maximization of the sequencing efficiency of the second generation sequencers allowing researchers to obtain more sequencing data from their samples at a greater depth of coverage.

RainStorm™ Droplet-Based PCR

Primer Library Generation
(a) Identify targeted sequences of interest in the genome.
(b) Design and synthesize forward and reverse primer pairs for each targeted sequence.
(c) Generation of primer pair droplets. A microfluidic chip is used to encapsulate the aqueous PCR primers in inert fluorinated carrier oil with a biopolymer surfactant to generate the equivalent of a picoliter scale test tube compatible with standard molecular biology.
(d) Primer pair droplets are mixed together so that each library element has an equal representation.

Genomic DNA Template Mix Preparation
(e) Genomic DNA is fragmented into 2 to 4 kb fragments and purified.
(f) Purified genomic DNA is mixed together with all of the components of the PCR reaction except the PCR primers.

Primer-Template Merge and PCR
(g) Primer Library droplets (~16pL) are dispensed to the microfluidic chip.
(h) Genomic DNA Template is delivered as an aqueous solution and template droplets (~18pL) are formed within the microfluidic chip. The primer pair droplets and template droplets are then paired together in a 1:1 ratio.
(i) Paired droplets pass through the channel of the microfluidic chip to pass through a merge area where an electric field induces the two discrete droplets to coalesce into a single PCR droplet (~26 pL). Up to 2 million PCR droplets are collected into a single 0.2 ml PCR tube.

The collection of PCR droplets (PCR Library) is processed in a standard thermal cycler for targeted amplification, followed by the emulsion of PCR droplets to release the PCR amplicons into solution for purification and next-generation sequencing.

GWAS Candidate Loci

PCR Primer Libraries were designed to enrich three contiguous regions from RGL2, Sp21, and 19p13. The samples were sequenced using 50 bp single end reads using the Illumina GAIIx system. The coverage (uniquely mapped reads) using the RainDance Sequence Enrichment assay ranged between 97-99%, with previously published results of 37-48% when using hybridization capture.

Gene Networks

Representative sequence coverage (MLH1 gene).

Heterogeneous Samples

Two HapMap DNA samples were mixed together at varying ratios to represent sample pools of increasing complexity as a model of heterogeneous samples such as pooled DNA from multiple individuals or tumor samples with low tumor cellularity. The pooled samples were processed on the RDT 1000 using a PCR Primer Library of 1536 primer pairs representing 700X of target sequence. The resulting PCR products were sequenced to an average depth of 1500X on the Illumina GAIIx. The figure on the right shows good correlation between the expected and observed allele frequency in the samples.

Summary

The combination of single-molecule PCR and large number of replicate independent PCR reactions enabled by the RDT 1000 Sequence Enrichment assay provides superior performance for targeted resequencing-based validation of genomic variation associated with complex diseases. The technique is ideally suited as a method for follow-up to GWAS studies, resequencing of candidate genes and deep resequencing of tumor samples.

References


For Research Use Only. Not for use in diagnostic procedures.