APPLICATION NOTE

Targeted Sequencing of 4000 Exons Using the RainDance Technologies Sequence Enrichment Assay

Introduction

Recent advances in DNA sequencing technologies have improved accuracy and dramatically reduced the cost of DNA sequencing. However, even with the improved efficiency of these second-generation systems, sequencing thousands of whole human genomes across various phenotypes is expensive and time consuming. To exploit the full potential of these new sequencing techniques, a robust method for isolating biologically relevant genomic loci on the megabase scale will be required. This technique has been commonly referred to as targeted sequencing.

The Sequence Enrichment application from RainDance Technologies leverages the sensitivity and specificity of PCR in a novel droplet-based format that avoids the limitations of traditional multi-plex amplification or hybridization methods. The RDT 1000 utilizes RainStorm™ microdroplet-based technology to amplify hundreds to thousands of genomic loci with high specificity and uniformity.

Enrichment is a critical step in an efficient targeted sequencing workflow. This application note will describe the use of the RDT 1000 and the Sequence Enrichment assay for targeted sequencing of 4000 human exons using two of the leading second-generation DNA sequencing platforms.

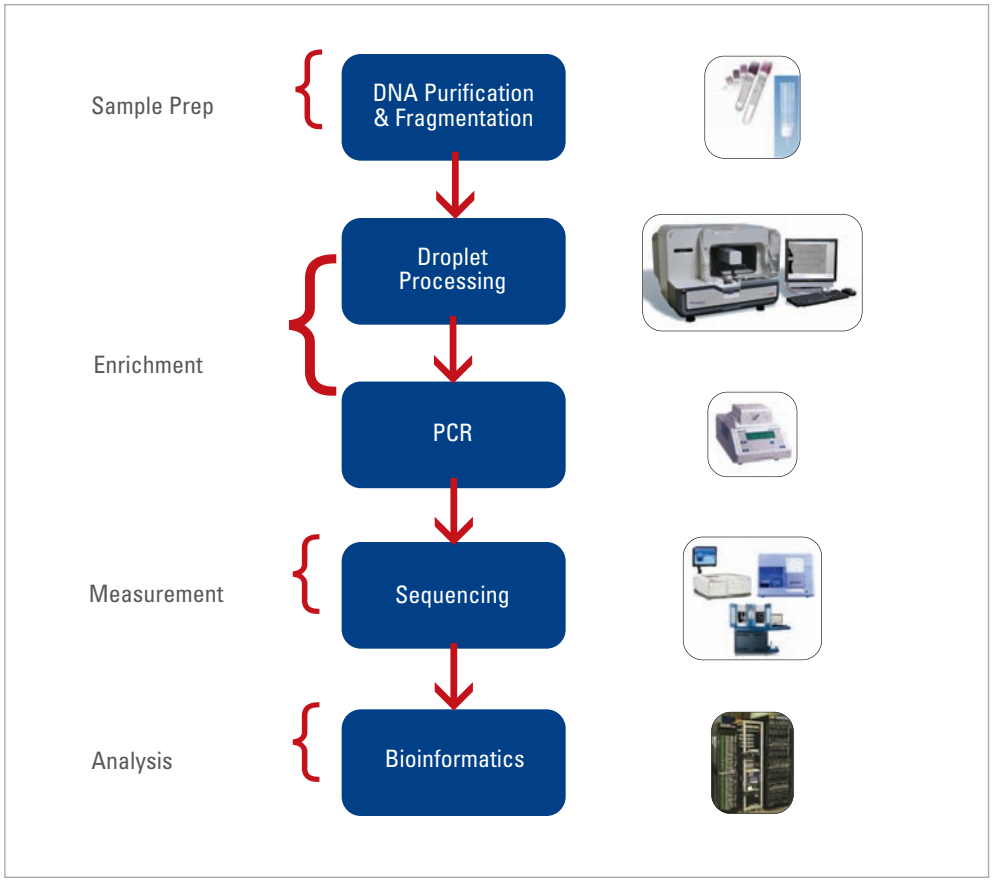


Figure 1. Targeted sequencing workflow. Droplet processing requires less than one hour per sample allowing for the processing of up to eight samples per day including PCR amplification.

RainDance™ Technologies

Methods

Primer Design

A set of 4000 human exons was selected for the purpose of demonstrating feasibility of sequence enrichment using the RDT 1000. The sequences were chosen to represent a wide range of GC content (25-81%), amplicon length (299-659 bp) and primer Tm (55-61° C). The total target sequence represented in the 4000 primer pairs was 1.63 Mb.

Droplet Library Generation

Pooled forward and reverse primers were prepared at equal concentrations for each of the 4000 primer pairs. The primers were reformatted into droplets in a serial process and pooled into a single droplet library emulsion. Aliquots of the droplet library were prepared for use on the RDT 1000 (Figure 2).

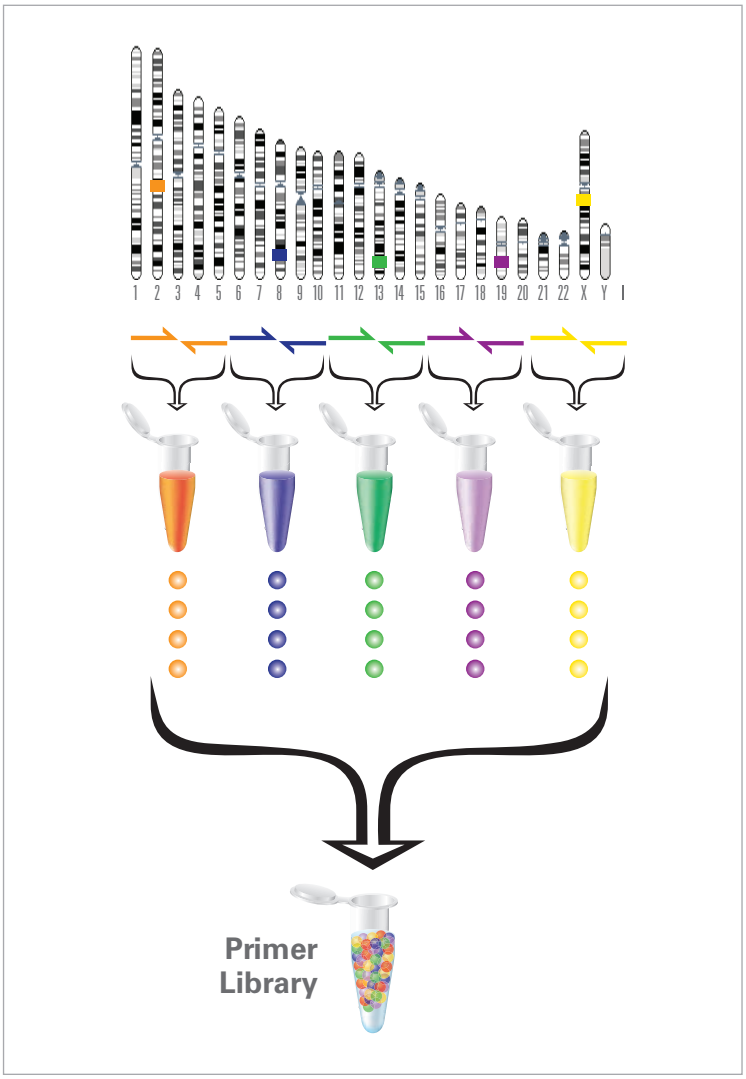
Droplet PCR

2 micrograms of purified human genomic DNA was fragmented to a size of 2-4 Kb. The PCR mix, containing genomic DNA, Platinum® Taq DNA Polymerase High Fidelity (Invitrogen), dNTPs and reaction buffer, was loaded into the RDT 1000 along with the primer library (Figure 3). Droplets were collected into a 0.2 mL PCR tube and amplified using 55 cycles of PCR. Amplification products were recovered by breaking the emulsion and purification using a MiniElute PCR Purification kit (Qiagen).

DNA Sequencing

Purified amplicons were processed for sequencing using both the 454 FLX and Illumina Genome Analyzer platforms using standard sequencing library preparation methods.

Figure 2: Primer library production process.



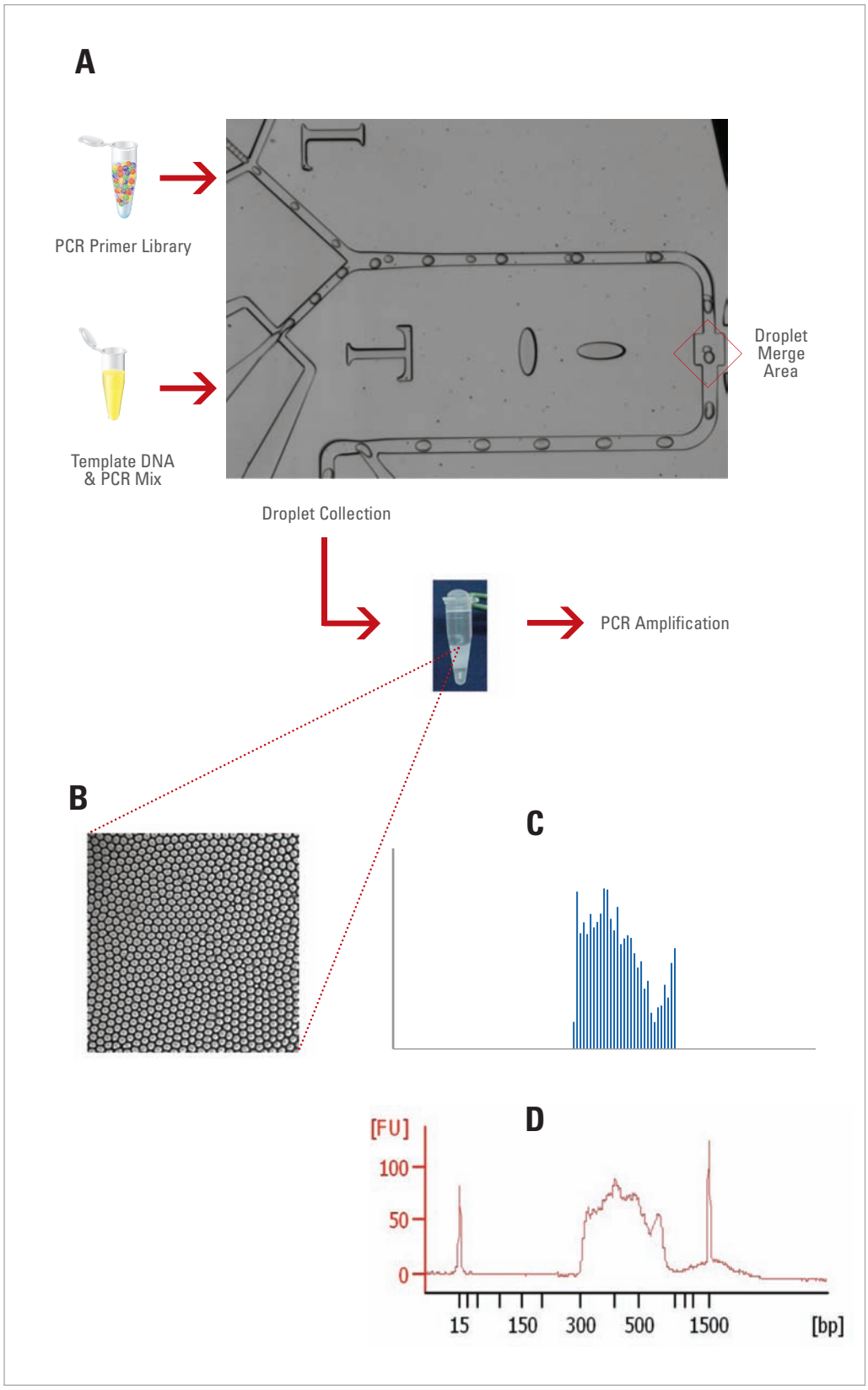


Figure 3. Droplet-based PCR process. PCR reaction mix containing fractionated genomic DNA is loaded into the RDT 1000 sequence enrichment chip along with the PCR primer library. The instrument generates droplets containing a gDNA + PCR mix and pairs them with the primer droplets. The paired droplets flow past an electrode embedded in the chip and are instantly merged together. The resulting PCR droplets are then automatically dispensed into a PCR tube which is transferred by the user to a standard thermal cycler for PCR amplification ***(A)****.*

The uniformity and stability of the droplets are shown in a micrograph of the droplet emulsion ***(B).***

Amplification products are purified by breaking the emulsion and purified using a PCR clean-up kit. An aliquot of the purified products was analyzed by capillary electrophoresis to quantify product yield. The resulting electropherogram is compared to a calculated peak profile based on the predicted size and number of amplicon products in the PCR library ***(C).***

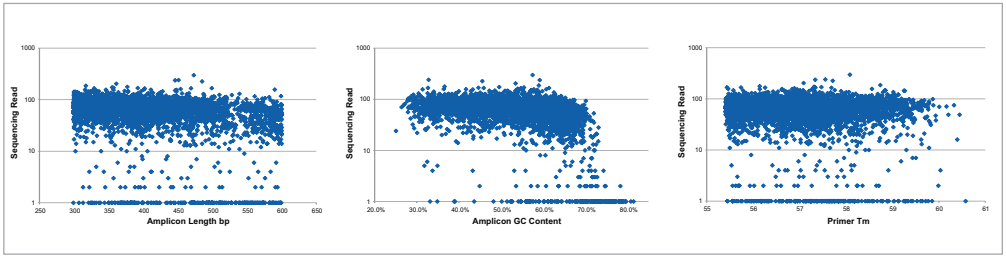
The comparison is used as a QC step prior to DNA sequencing ***(D).***

Results

A summary of the results of sequencing using both 454 FLX (Roche) and Genome Analyzer (Illumina) is shown in Table 1. The percent of target amplicons covered was very high for both sequencing platforms. The relative difference in the representation of target amplicon coverage is shown for both platforms as the percent of amplicons within a 5-fold and 10-fold relative depth of coverage. Distribution of coverage is shown for the 454 FLX data in Figure 5.

Table 1: Targeted Sequencing Results	454 FLX	Illumina GA
Reads Mapping to Target Regions	94%	78%
Percent of Target Amplicons Covered	96.9%	99.7%
Target Representation Bias (<5-fold difference)	89%	81%
Target Representation Bias (<10-fold difference)	93%	93%
SNP Concordance	98.4%	98.5%
Allelic Ratios for SNPs	Mean — 49.5% STD — 9.8%	

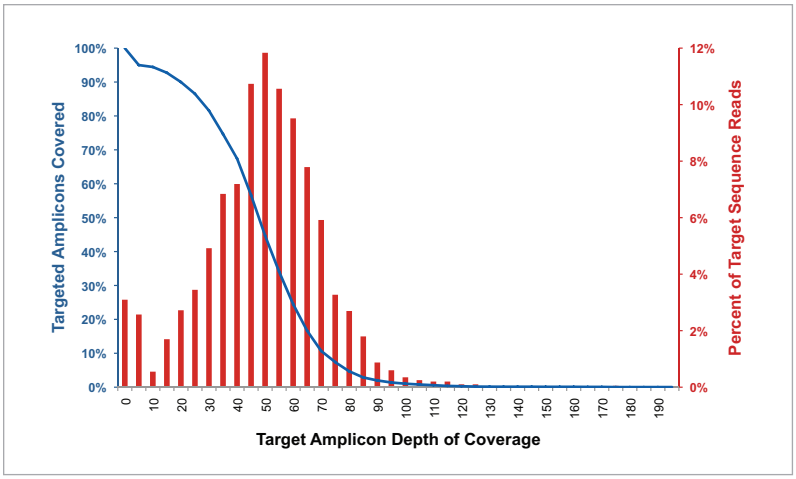
Figure 4. Results from 454 FLX sequencing (coverage as a function of amplicon length, GC content, and primer Tm).



Conclusion

Sequence enrichment using droplet-based PCR provides high specificity and uniform enrichment required to maintain an efficient targeted sequencing workflow and more importantly, provide high-quality sequence data in order to provide high confidence in variant detection required for large-scale targeted resequencing applications. The RDT 1000 and Sequence Enrichment assay function to serve workflows for both short-read and long-read sequencing platforms.

Figure 5. Results of targeted sequencing (454 FLX data shown). The figure shows the distribution of target amplicon read depth (red) and the cumulative amplicon read depth (blue) for all 4000 PCR amplicons. The results demonstrate the uniformity of coverage and high percentage of target amplicons with sufficient depth of coverage for high-confidence SNP analysis.



RainDance Technologies, Inc.
44 Hartwell Avenue
Lexington, MA 02421, U.S.A.
Phone: 781-861-6300
Phone, toll-free: 888-724-6440
Fax: 781-861-1233
www.raindancetech.com
E-mail: info@raindancetech.com

RainDance Technologies