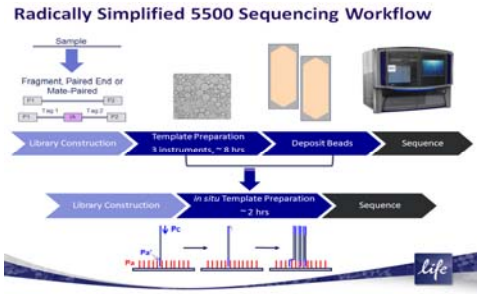
WildFire: A Simple Monoclonal Colony Generation Technology without Emulsion PCR

Zhaochun Ma, Raymond Lee, Jonathan Erikson, Swati Goyal and Kai Lao,
Life Technologies,.850 Lincoln Centre Drive, Foster City, CA 94404, USA

ABSTRACT / INTRODUCTION

We have developed a new massive parallel monoclonal template generation technology, called "WildFire ", where sequencing libraries are in-situ isothermal amplified directly on the surface of a 5500 flowchip. Sequencing libraries are added directly to the 5500-series Genetic Analyzer flowchip, whose surfaces have been coated with a special library-adaptor capture oligonucleotide. A DNA polymerase reaction mix is added, and in a single isothermal step for 30 minutes, single templates are amplified in-situ on the flowchip. The net density of sequencing-colonies created in this manner far exceeds anything currently utilized in next-generation sequencing, reaching 1.1 million colonies per mm^2 per flowchip surface. During in-situ amplification, the capture oligonucleotide is "consumed ", and each individual nucleic-acid fragment "spreads " (like a WildFire) inside the flowchip until reaching an adjacent library fragment(s). When the individually-growing fragments "meet ", the amplification step terminates, because all of the surface-bound primer was consumed. These "self-assembled ", spatially resolved, monoclonal colonies, are then sequenced by SOLiD™ chemistry. The resulting colony-sequencing reads maintain the same high accuracy as our bead-based method. Full genomes (from bacterial to human), exomes (human), and transcriptomes (human) have now been sequenced using WildFire technology. Here we also demonstrate a novel reverse sequencing method with only SOLiD™ forward probes, called WildFire Paired-End sequencing (wPE) using a template switch strategy.



MATERIALS AND METHODS

All targeted capture enrichments (singleplex and multiplex) were done with Ion TargetSeq™ Exome Enrichment Kit probe-sets according to the manufacturer's protocols. DNA was obtained from the Coriell Institute and 5500xl W, 5500xl Genetic Analyzer libraries were constructed using kits from Life Technologies using 5500 SOLiD® Fragment Library Barcode Adaptors 1-16 Kit with and without the 5500xl W conversion module respectively. Two types of library samples including NA12878 and HuRef were used in the preparation of 4 barcoded libraries using an AB Library Builder™ instrument , followed by pippin prep for size selection. This was followed by the Ion TargetSeq™ (v2) exome enrichment. An aliquot of the library pool was used for template preparation on FlowChip using the Wildfire "template walking" protocol Sequencing on 5500xl W Genetic analyzer was done in single-lane runs for 50×50 paired-end tag. Data analysis, including mapping and targeted resequencing analysis and variant calling, was done with LifeScope™ software.

RESULTS

Figure 1. 5500xl W in-situ Template Preparation Is a Fast Isothermal Amplification Generating Ultra High Density Colonies

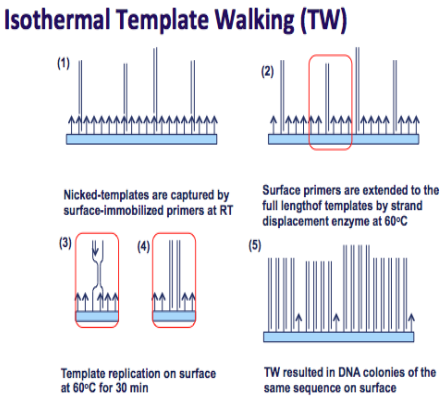
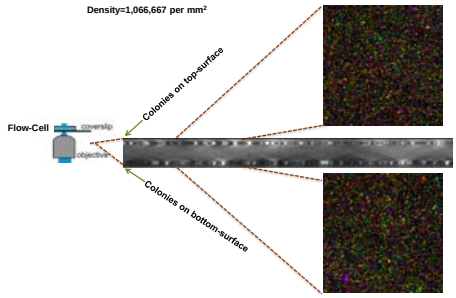
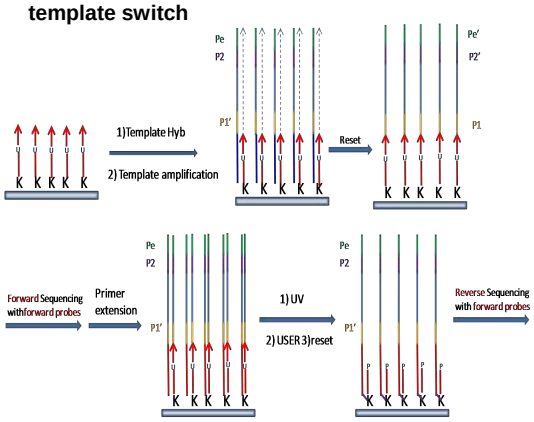


Figure 2. 5500xl W Colonies "Self-Assemble" on Both Sides of Flow-Chip, Further Doubling Output and Reducing the Reagent Cost



Note: dual surface format available outside US only

Figure 3. WildFire Pair-end Reads(wPE) by template switch



K: photo cross-linking reagent, U: deoxyuridine

Table 1. A Wildfire 50×50 Paired End (wPE) Run Generated 21.3 Gb Mapped /Lane (570 top only layer panels)

Filtered Throughput	38 GB
Mapped Throughput	21.34 GB
Total F3 Tag Mapped	276M
Total F5 Tag Mapped	199M
Throughput (Gb/day)	9.5
Both F3 and F5 Tag Mapped %	55%
F3 Tag Mapped only	73%
F5 Tag Mapped only	53%
Panels scanned	570
Reads in Target	249M (64.92%)
Reads off Target	134M (35.07%)
Percent of target bp not covered	4.3%
Average Depth of Coverage within targets	216X
Percent of target bp covered at >=20X	87.81%
dbSNP heterozygous concordance	96.18%
dbSNP homozygous concordance	99.40%

CONCLUSIONS

WildFire technology greatly improves NGS workflow, increases throughput, and significantly decreases net cost-per-genome. A novel pair-end sequencing method was also developed by using a template switch strategy. With top layer only, a single lane of 50X50bp PE run, we generated over 21Gb mapped data with an average depth of 216X, >95% of target bases covered and >87% covered at 20X depth with a db heterozygous SNP concordance of 96%.

.

Acknowledgement

We thank Devin Do and Kamini Varma for providing TargetSeq™ exome library.

For Research Use Only. Not intended for any animal or human therapeutic or diagnostic use.