EPICENTRE® Biotechnologies

Application of Nextera™ technology to RNA-seq library preparation

Nextera™ technology is fast becoming the method of choice for the preparation of second-generation DNA-sequencing libraries. Other library-preparation workflows are limited by multistep protocols, sample loss, lack of automation capability, and labor costs. Nextera technology counters many of these bottlenecks and offers a simplified procedure applicable to multiple second-generation sequencing platforms. Here we describe the adaptation of the Nextera library-preparation procedure to RNA-seq library preparation.

Most second-generation sequencers adhere to a common library-preparation procedure, with minor variations. This procedure includes: (i) DNA fragmentation (sonication, nebulization or shearing); (ii) DNA end-polishing or A-tailing; and (iii) platform-specific adaptor ligation. Typically, this results in significant sample loss and limited throughput. To streamline the library-preparation procedure, Epicentre has developed Nextera technology, which is capable of simultaneously fragmenting and tagging DNA in a single-tube reaction in less than 2h—using only 50 ng of DNA—with the ability to prepare sequencer-ready, platform-specific libraries. Nextera libraries can also incorporate barcodes, enabling multiplexed sequencing on a single instrument run, resulting in significant cost savings.

Method overview

Nextera technology simultaneously fragments and tags target DNA in a single-tube reaction via a 5-min, *in vitro* transposition reaction. Platform-specific adaptors and optional barcoded adaptors are added, and the library is enriched, by limited-cycle PCR. The molecular-weight distribution of libraries prepared using Nextera technology can be controlled for the read-length requirements of different sequencing platforms by selecting the appropriate reaction buffer (see http://www.epibio.com/nextera).

Nextera kits were designed to prepare libraries from genomic DNA; however, with minor modifications, the technology can be adapted for RNA-seq (cDNA) library preparation. A random-primed method is used to synthesize first-strand cDNA from polyadenylated (poly(A)$^+$) mRNA. Following second-strand synthesis, double-strand (ds) cDNA is used as input for the Nextera "tagmentation" (fragmentation and tagging) reaction. Illumina bPCR-compatible sequences are added to the tagmented DNA, and di-tagged fragments are enriched by limited-cycle PCR. This method generates nondirectional libraries (**Fig. 1**).

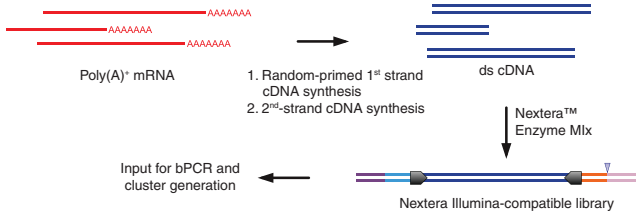

Figure 1 | Schematic overview of the RNA-seq library-preparation method.

Deep sequencing of Illumina-compatible cDNA libraries

We compared RNA-seq libraries constructed using Nextera technology (as shown in **Fig. 1**) to those prepared by a standard procedure. Briefly, for the standard procedure, mRNA was fragmented following poly(A) selection and used as a template for first-strand cDNA synthesis by random-hexamer priming. Next, second-strand synthesis was performed using a commercially available system. After reaction cleanup, Illumina sequencing adaptors were ligated onto the ds cDNA, and fragments of the appropriate size were extracted.

RNA from the same source (human cell line ECC-1) was used to construct all libraries, and each library was sequenced on a different lane of the same Illumina flow cell. Custom Nextera sequencing primers and standard Illumina sequencing primers were used together. Approximately 25 million aligned 36-bp reads were obtained from each lane. **Figure 2a** shows that there is good correlation ($R = 0.878$) between Nextera and standard RNA-seq libraries. The libraries are also of sufficient complexity, as evidenced by the unique reads (over 87%) and the high number of genes represented by >10 reads.

Fraz Syed

EPICENTRE Biotechnologies, Madison, Wisconsin, USA.
Correspondence should be addressed to F. S. (fraz.syed@epibio.com).

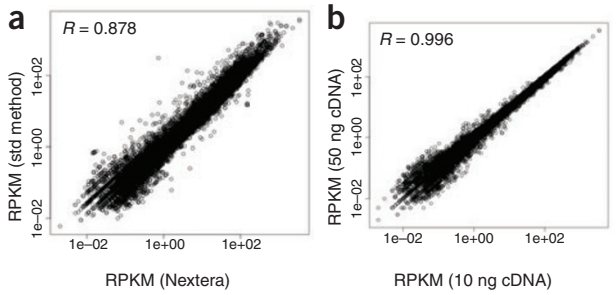

Figure 2 | Correlation of relative gene expression. (**a**) Libraries prepared using Nextera technology versus a standard method. (**b**) Nextera libraries prepared from 50 ng and 10 ng of cDNA.

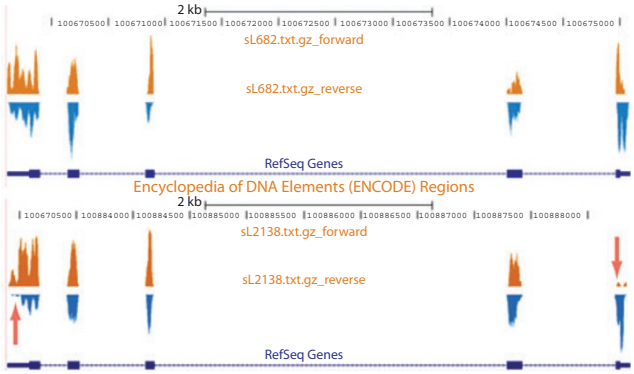

Figure 3 | Library coverage and complexity for standard and Nextera libraries. Expected drops in coverage at the most distal ends of the cDNA are shown with red arrows.

To determine the impact of sample-input amount on library complexity, two Nextera libraries constructed using 50 and 10 ng cDNA, respectively, were sequenced. Approximately 25 million aligned 36-bp reads were obtained per lane of the same Illumina flow cell. **Figure 2b** shows excellent correlation of gene-expression levels between the two Nextera libraries; high library complexity (over 87% unique reads) was also obtained, even with 10 ng input cDNA.

As a control, sequence data obtained from the Nextera library method and the standard method were aligned to a reference genome (**Fig. 3**). Both nondirectional libraries displayed similar sense and antisense transcript coverage. However, since two transposome insertions are required to produce a functional sequencing template, the Nextera library exhibited a slight decrease in coverage of the extreme ends of the cDNA. Complete coverage of linear ds cDNA can be obtained by including one of the adaptor sequences in the PCR or cDNA synthesis primers.

Conclusions

The Nextera library-preparation procedure offers many advantages over current methods, including significant time and cost savings. The Nextera procedure, a single-tube reaction, is scalable and requires less starting DNA than other procedures. As shown here, the Nextera procedure can be readily adapted for RNA-seq library preparation. Deep sequencing of the cDNA libraries shows that Nextera-generated libraries produce accuracy, coverage and bias comparable to those of libraries produced by standard methods. As little as 10 ng of cDNA can be used as input for Nextera library preparation without sacrificing library complexity and coverage. Recently, a similar approach used Nextera technology for deep sequencing of libraries covering the complete retroviral genomes of human and simian immunodeficiency viruses[1].

ACKNOWLEDGMENTS

Sequencing and data analysis were performed in collaboration with K.T. Varley, J. Gertz and R. Myers, Hudson Alpha Institute for Biotechnology, Huntsville, Alabama, USA.

1. Bimber, B.N. *et al.* Whole-genome characterization of human and simian immunodeficiency virus intra host diversity by ultradeep pyrosequencing. *J. Virol.* **84**, 12087–12092 (2010).