



## Next-generation sequencing library preparation: simultaneous fragmentation and tagging using *in vitro* transposition

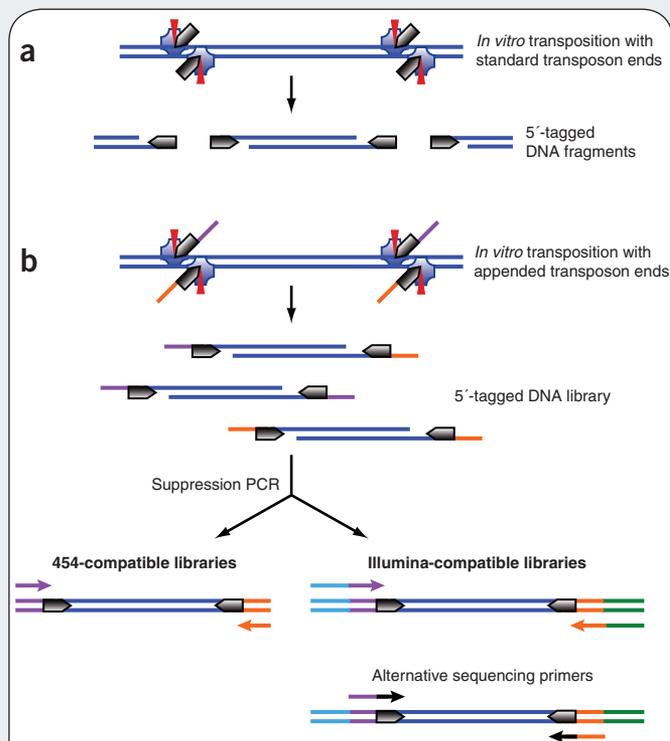
The advent of next-generation sequencing has made possible genome analysis at previously unattainable depth. Roche, Illumina and Life Technologies, among others, have developed well-established platforms for deep sequencing. Regardless of the instrument, one of the bottlenecks for next-generation sequencing is the amount of time and resources required for template and library preparation. Here we describe Epicentre's Nextera™ technology (covered by issued and/or pending patents), which counters this bottleneck and simplifies the sample preparation procedure.

At present, next-generation sequencing platforms use slightly different technologies for sequencing, such as pyrosequencing, sequencing by synthesis or sequencing by ligation. However, most platforms adhere to a common library preparation procedure, with minor modifications, before a 'run' on the instrument. This procedure includes fragmenting the DNA (sonication, nebulization or shearing), followed by DNA repair and end polishing (blunt end or A overhang) and, finally, platform-specific adaptor ligation. This process typically results in considerable sample loss with limited throughput. To streamline the workflow, increase throughput and reduce sample loss, Epicentre has developed Nextera™ technology, a transposon-based method for preparing fragmented and tagged DNA libraries in as little as 4 hours. This flexible, scalable and efficient technique can be used to generate libraries for multiple sequencing platforms.

### Method overview

Nextera technology uses *in vitro* transposition to prepare sequencer-ready libraries (Fig. 1). In a classic transposition reaction, transposases catalyze the random insertion of excised transposons into DNA targets with high efficiency. During cut-and-paste transposition, a transposase makes random, staggered double-stranded breaks in the target DNA and covalently attaches the 3' end of the transferred transposon strand to the 5' end of the target DNA. The transposase and transposon complex, also referred to as a Transposome™ complex, inserts an arbitrary DNA sequence at the point of insertion. We have discovered that the entire complex is not necessary for insertion, and free transposon ends are sufficient for integration.

When free transposon ends are used in the reaction, the target DNA is fragmented and the transferred strand of the transposon end oligonucleotide is covalently attached to the 5' end of the target fragment (Fig. 1a). The size distribution of the fragments can be controlled

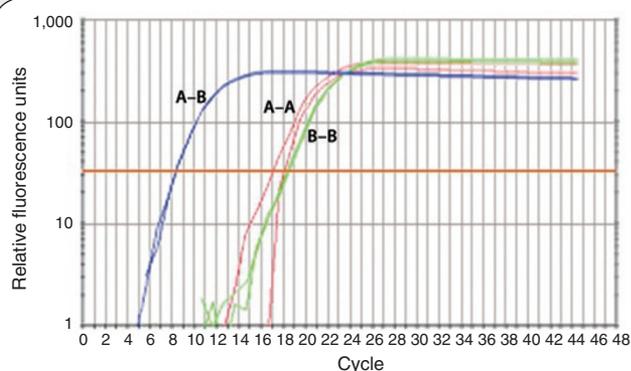


**Figure 1** | Overview of Nextera fragmentation and tagging technology. (a,b) Using *in vitro* transposition, the starting DNA template is randomly fragmented and tagged at sites indicated by red arrows, using standard transposon ends (a) or transposon ends appended to unique adaptor sequences (colored bars) (b). After suppression PCR, the library can be amplified and sequenced using the appropriate platform-specific primers.

Fraz Syed, Haiying Grunenwald & Nicholas Caruccio

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

## APPLICATION NOTES



**Figure 2** | Enrichment of A-B-tagged DNA fragments. Transposon ends were modified to contain two unique Roche/454-compatible tags (A and B), and a library was prepared following the scheme in **Figure 1b**. qPCR was performed using A-A, B-B and A-B primer pairs for 45 cycles. The horizontal orange line is the fluorescence threshold.

by changing the amounts of transposase and transposon ends (data not shown). Exploiting transposon ends with appended sequences results in DNA libraries that can be used in high-throughput sequencing (**Fig. 1b**).

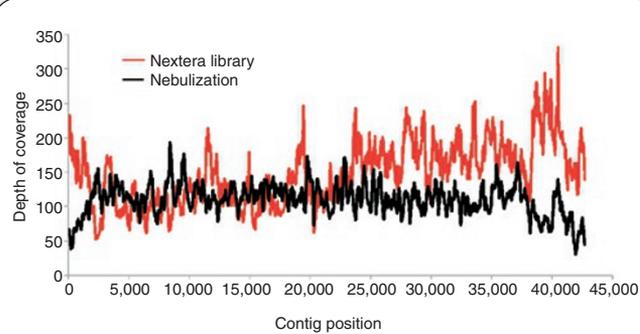
### Creating tagged libraries

Nextera technology can be used to create single- and dual-tagged libraries. To generate platform-specific libraries, complementary tags can be added to the 5' and 3' ends of the fragmented DNA. Once the DNA is labeled at the 5' end, a complementary tag can be added to the 3' end by extension with a strand-displacing polymerase. PCR amplification with a single primer confirms tagging at both ends of the DNA fragments and faithful reproduction of input DNA sizes (data not shown). Thus, single-primer PCR amplifies the genomic library and produces double-stranded DNA fragments with complementary adaptors.

In addition to the complementary tags, two independent tags can be added to the fragmented DNA by appending to the transposon end sequence an engineered adaptor sequence. After extension, the sequencing adaptors enable amplification by emulsion PCR, bridge PCR and other methods. The amplified library can be subsequently sequenced.

### Enrichment of dual-tagged fragment libraries

We used T7 bacteriophage genomic DNA to demonstrate dual tagging and enrichment of fragments containing both tags (**Fig. 2**). We modified transposon ends to contain Roche/454-compatible tags (A and B). After the tagging and fragmenting reaction, we heat-inactivated the transposase and performed limited-cycle PCR (suppression PCR). We analyzed a 1:100 diluted sample by quantitative PCR (qPCR). A mixed population of tagged DNA fragments was obtained, containing the desired A-B-tagged DNA, as well as DNA fragments with either A or B tags at both ends. However, qPCR showed enrichment of the A-B-tagged fragments in the amplified sample.



	Nebulization	Nextera library
Assembled contigs:	1 (42,679 bp)	1 (42,680 bp)
Percentage error:	0.48%	0.41%
Proportion Q40-plus bases:	0.999	0.998
Mapped reads:	99.48%	96.19%
Average coverage:	110	125

**Figure 3** | Comparison of nebulization and Nextera-generated libraries. The Nextera library, enriched in DNA fragments tagged with Roche/454 A and B adaptors, was sequenced on a GS FLX instrument (Roche). A control library prepared from the same starting DNA, using nebulization, was also sequenced. Depth of coverage across the contig (top) and a summary of data (bottom) are shown for each library. Q40 is a quality score denoting the probability of a wrong base call at 1 in 10,000.

### Nebulization versus *in vitro* transposition

We performed deep sequencing of the Nextera-enriched dual-tagged library and of a control library prepared by nebulization and the manufacturer's recommended protocol. Contig assembly, coverage and accuracy of the Nextera library data were comparable to those for the control library produced using nebulization (**Fig. 3**).

### Conclusions

The current library preparation methods for next-generation sequencing are time-consuming and prone to considerable sample loss. Even before library preparation, the recovered DNA must be purified and end-polished. Epicentre's Nextera technology offers many advantages over current library preparation methods, such as a streamlined workflow that can result in substantial savings in time and cost. The method is scalable and requires less starting DNA than current procedures. As described here, Nextera technology adapts *in vitro* transposition, a powerful technique that can simultaneously fragment and tag genomic DNA, by using optimized transposases and incorporating engineered free transposon ends. By additional manipulation, libraries containing complementary or independent adaptor sequences can also be efficiently constructed and amplified before sequencing on most next-generation sequencing platforms.

### ACKNOWLEDGMENTS

Sequencing data were generated in collaboration with Hilary Morrison, Jessica Welch, Ekaterina Andreischeva and William Reznikoff, Marine Biological Laboratory, Woods Hole, Massachusetts, USA.

This article was submitted to *Nature Methods* by a commercial organization and has not been peer reviewed. *Nature Methods* takes no responsibility for the accuracy or otherwise of the information provided.