

## Transposon-Based Strategies for Efficient DNA Sequencing and Functional Genomics

Many of the products highlighted in this issue of the *Forum* offer innovative strategies for cloning a fragment of DNA or making a library of genomic clones. The EZ::TN™ and HyperMu™ Insertion Kits described here facilitate the characterization of cloned DNA by enabling researchers to completely sequence cDNA or genomic clones in plasmid, cosmid, fosmid or BAC vectors without subcloning or primer walking. Each kit exploits the ability of DNA sequences called transposons to “hop” or transpose into another DNA molecule.

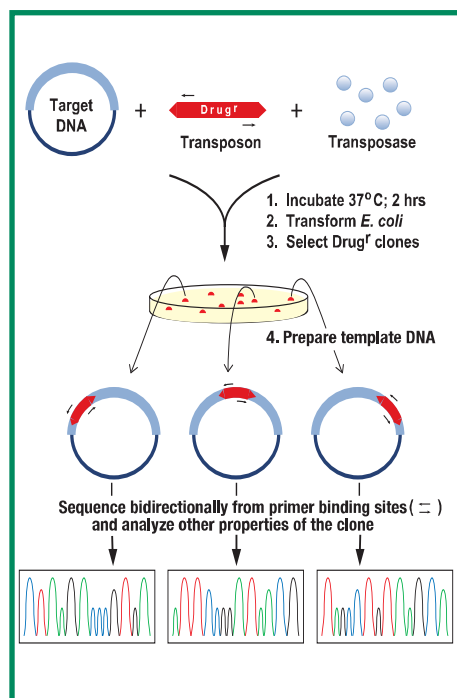
### EZ::TN Insertion Kits

EZ::TN Insertion Kits are based on the hyperactive Tn5 *in vitro* transposition system described by Goryshin and Reznikoff.<sup>1</sup> This system retains the highly random insertion characteristics of Tn5 but has a transposition frequency 1000-fold higher than wild-type Tn5. As shown in Figure 1, all EZ::TN Insertion Kits can be used to randomly insert an EZ::TN™ Transposon, containing a selectable marker and primer binding sites for sequencing, into any DNA molecule *in vitro*. Then, transform *E. coli* with an aliquot of the reaction and select for the marker encoded by the EZ::TN Transposon. Up to millions of independent insertion clones are obtained, each with a single EZ::TN Transposon at a different site. Prepare template DNA from randomly chosen insertion clones and sequence bidirectionally from the ends of the EZ::TN Transposon.

Many researchers, like David Boyd et al., realize that a key benefit to transposon-based sequencing is “to allow for rapid sequencing of the insert”.<sup>2</sup> Transposon-based sequencing is rapid because the method avoids primer walking and subcloning. Even the largest BAC clone can be sequenced without having to make a shotgun library or synthesize new primers. Sequence reads can be generated simultaneously from multiple transposon insertion clones rather than from sequential “walks” of a template.

### HyperMu Insertion Kits

HyperMu Insertion Kits are based on the well-characterized temperate bacteriophage Mu. Like the EZ::TN Insertion Kits,



**Figure 1. The process for generating insertion clones for sequencing, and a myriad of other applications, using an EZ::TN™ or HyperMu™ Insertion Kit.**

these kits simplify and speed up complete sequencing of any cloned DNA that is too large to sequence with a single set of sequencing reactions. The Mu-based transposition system uses a hyperactive enzyme that retains the highly random insertion characteristics of MuA transposase<sup>3</sup> but is at least 50-times more active *in vitro* than the enzyme from other suppliers. Consistently high transposition efficiencies are required to completely sequence a clone, especially those with large inserts.

### High-Throughput cDNA Sequencing

Perhaps the power of using EZ::TN and HyperMu Insertion Kits for sequencing is best exemplified by their use in high-throughput settings. The NIH Intramural Sequencing Center (NISC), for example, needed a method for full-insert sequencing of cDNA clones. Given an average insert size of 2 kb, making a shotgun library of each cDNA clone was impractical. Primer walking also has limitations in this high-throughput environment since the extensive use of oligonu-

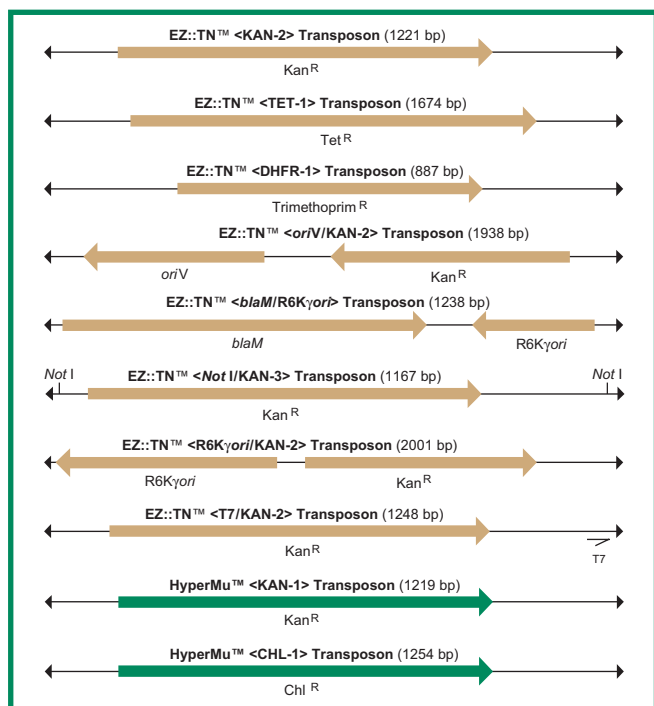
cleotides is not cost effective. Moreover, the serial nature of the process and the difficulties encountered to ensure that a primer and template anneal correctly, make scale-up difficult.<sup>4</sup>

Having eliminated both subcloning and primer walking, the NISC turned to a transposon-based sequencing strategy. Using the EZ::TN Insertion System, the NISC sequenced more than 4200 cDNA clones to produce greater than 8 Mb of high-accuracy sequence (less than one error in 1 million bp). The per-sequence-read cost was very comparable to those associated with shotgun sequencing of large-insert genomic clones. The NISC also noted that the EZ::TN Transposon insertions were highly random, lacking insertional “hot spots” or extensive gaps between insertions.<sup>4</sup>

### Gap Finishing with Transposons

Most large-scale genome sequencing projects have a “finishing group” that is responsible for closing gaps that occur after the initial rounds of sequencing. Gaps are caused by a variety of factors including highly repetitive DNA, AT- and GC-rich regions, or poor quality sequencing traces. “Transposon bombing” as it is referred to by the finishing group at Washington University Genome Sequencing Center is often used to close these gaps and offers a number of advantages over other strategies.<sup>5</sup> Primer walking, for example, generates unusable sequence in regions containing repetitive DNA because the primer anneals to multiple sites within the template. The unique primer binding sites found at the ends of EZ::TN and HyperMu Transposons eliminate this problem. In addition, the transposon order can be determined by physically mapping the transposon insertion site. The resulting data can then be used to confirm the correct order of the assembled sequence.

EPICENTRE offers a variety of EZ::TN Insertion Kits for characterizing cloned DNA. In addition to primer binding sites and a selectable marker each of the EZ::TN Transposons included in these kits (Figure 2) contains features that can be used in gene analysis, proteomics, or RNA research.



**Figure 2. EZ::TN™ and HyperMu™ Transposons each contain a selectable marker, primer binding sites for bidirectional sequencing, and a variety of other useful DNA sequences.**

**Synthesize RNA from Any Region of Cloned DNA**

The EZ::TN™ <T7/KAN-2> Promoter Insertion Kit provides an easy and reliable method to randomly insert a phage T7 RNA polymerase promoter into any target DNA. The transposon does not have a transcription termination sequence so RNA can be produced from chosen insertion clones by *in vitro* transcription using an AmpliScribe™ T7-Flash™ Transcription Kit, or *in vivo* after transformation of *E. coli* that have an inducible T7 RNA polymerase gene.

**References**

- Goryshin, I.Y. and Reznikoff, W.S. (1998) *J. Biol. Chem.* **273**, 7367.
- Boyd, D. et al. (2001) *J. Bacteriol.* **183**, 5725.
- Butterfield, Y.S.N. et al. (2002) *Nucl. Acids Res.* **30**, 2460.
- Shevchenko, Y. et al. (2002) *Nucl. Acids Res.* **30**, 2469.
- Devine, S.E. et al. (1997) *Genome Res.* **7**, 551.
- Hoffman, L.M. and Loomis, K.B. (2000) *EPI-CENTRE Forum* **7**(3), 4.
- Mankouri, H.W. et al. (2002) *Nucl. Acids Res.* **30**, 1103.
- Jendrisak, J. et al. (2002) *EPI-CENTRE Forum* **9**(1), 14.
- Yoon, Y.G. and Koob, M.D. (2003) *Nucl. Acids Res.* **31**, 1407.

[www.epicentre.com/transposomics.asp](http://www.epicentre.com/transposomics.asp)

<b>EZ::TN™ &lt;KAN-2&gt; Insertion Kit</b>	
EZI982K	10 Reactions
<b>EZ::TN™ &lt;TET-1&gt; Insertion Kit</b>	
EZI921T	10 Reactions
<b>EZ::TN™ &lt;DHFR-1&gt; Insertion Kit</b>	
EZI912D	10 Reactions
<b>EZ::TN™ &lt;oriV/KAN-2&gt; Insertion Kit</b>	
EZI02VK	10 Reactions
<b>EZ::TN™ β-Lactamase Fusion Kit</b>	
EZI31BL	10 Reactions
<b>EZ::TN™ In-Frame Linker Insertion Kit</b>	
EZI04KN	10 Reactions
<b>EZ::TN™ &lt;R6K&lt;sub&gt;ori&lt;/sub&gt;/KAN-2&gt; Insertion Kit</b>	
EZI011RK	10 Reactions
<b>EZ::TN™ &lt;T7/KAN-2&gt; Promoter Insertion Kit</b>	
EZI03T7	10 Reactions
<b>HyperMu™ &lt;KAN-1&gt; Insertion Kit</b>	
HMI032K	10 Reactions
<b>HyperMu™ &lt;CHL-1&gt; Insertion Kit</b>	
HMI039C	10 Reactions

**Find Functional Domains or Epitopes of Proteins**

The EZ::TN™ In-Frame Linker Insertion Kit was designed to rapidly and easily produce random 19-amino acid (19 codons; 57 nucleotides) in-frame insertions into genes of expressed proteins for protein engineering, functional analysis, and domain or epitope mapping.<sup>6, 7</sup> The kit features the EZ::TN™ <Not I/KAN-3> Transposon, which contains a kanamycin-resistance marker flanked by *Not I* restriction sites. Kanamycin-resistant insertion clones are digested with *Not I*, ligated, and re-transformed into *E. coli*. Since each resulting clone contains a random 19-codon insertion that can be read in all three reading frames the protein retains its original amino acid sequence on both sides of the insertion site.

**Direct Selection of Genes Encoding Non-Cytoplasmic Proteins**

The EZ::TN™ β-Lactamase Fusion Kit was developed for the direct selection of genes encoding membrane and secreted proteins. The kit features the EZ::TN™ <blaM/R6K<sub>ori</sub>> Transposon, which contains a β-lactamase gene (*blaM*) that lacks both promoter and secretory signal sequences. The transposon is randomly inserted into target DNA, which is then transformed into *E. coli* and selected on ampicillin. Only insertion clones with transcriptional fusions to genes encoding extracytoplasmic proteins will grow. These fusions generate hybrid proteins that

can transport the *blaM* moiety through the inner membrane and confer resistance to ampicillin.

**“Rescue” Plasmid and Mitochondrial DNA**

The EZ::TN™ <R6K<sub>ori</sub>/KAN-2> Insertion Kit facilitates propagation and sequencing of circular DNA molecules, such as plasmids<sup>8</sup> and mitochondrial DNAs,<sup>9</sup> that would not otherwise replicate in *E. coli*. The EZ::TN Transposon containing an *E. coli* R6K<sub>ori</sub> origin of replication and a kanamycin-resistance marker is randomly inserted into target DNA. Insertion clones can then be propagated in TransforMax™ *pir*<sup>+</sup> or *pir*-116 Electrocompetent *E. coli* which express the *pir* gene product, required for replication from the R6K<sub>ori</sub>, and selected on kanamycin plates.

**Integrate CopyControl Capability into BAC and Fosmid Clones**

The EZ::TN™ <oriV/KAN-2> Insertion Kit enables researchers to integrate CopyControl capability into existing single-copy BAC and fosmid clones. Like the CopyControl pCC1 BAC™ and pCC1FOS™ Vectors, BAC and fosmid clones containing the EZ::TN™ <oriV/KAN-2> Transposon can be maintained in TransforMax™ EPI300™ Electrocompetent *E. coli* at single copy to ensure insert stability. When needed, clones can be induced to 10 to 50 copies per cell to maximize the yield and purity of DNA for sequencing, fingerprinting and other applications.