

Use of the EZ::TN™ Transposon Insertion System for High Throughput Full-Length cDNA Sequencing at the NIH Intramural Sequencing Center (NISC)

Yuriy Shevchenko, Robert Blakesley, Gerry Bouffard, Keith Wetherby, Eric Green and Jeffrey Touchman
NIH Intramural Sequencing Center, National Institutes of Health, Gaithersburg, MD

Introduction

In addition to producing over two million genomic sequence reads per year, the NIH Intramural Sequencing Center (NISC) also participates in the sequencing program of the Mammalian Gene Collection (MGC) project. Briefly, the goal of this project is to generate collections of human and other mammalian full-length cDNAs that are sequenced to high accuracy, then to make the clones freely available to the scientific community.¹ This article provides a brief description of the high throughput full-length cDNA sequencing activities at NISC. Results of these efforts will be published elsewhere. Progress reports are also available through the MGC Web site (<http://www.ncbi.nlm.nih.gov/MGC>).

Strategy and Methods

Sequencing of MGC cDNA clones, whose inserts range in size from 300 bp to over 6 kb and are relatively high in GC content, represents a challenge. After evaluating several available methods, including primer walking,² concatenated cDNA sequencing³ and multiple transposon-based systems, we decided to use the Tn5-based EZ::TN™ Transposon Insertion System because of its high performance in our setting. This system is robust, gives a reasonably uniform distribution of random transposon insertions, is easily adaptable to 96-well format, and is compatible with antibiotic selection in the MGC cDNA vector.

Generation *In Vitro* of cDNA Insertion Subclones

Transposon insertion reactions using the EZ::TN System are simultaneously set up for all cDNA clones in a 96-well reaction tray. First, a transposon "brew" containing EZ::TN Reaction Buffer, EZ::TN <KAN-2> Transposon DNA, and EZ::TN Transposase is dispensed into each well of the tray. Next, purified cDNA clones (target DNAs) are transferred into separate wells of the tray using a Microlab 4200 pipetting robot (Hamilton). Transposon insertion reactions are carried out by incubating the tray in a thermal cycler (MJ Research) for 2 hours at 37°C. Reactions are terminated by the addition of EZ::TN Stop Solution, followed by incubation at 70°C for 10 minutes.



The EZ::TN™ Transposon Insertion System is an effective component of a high throughput sequencing pipeline at the NIH Intramural Sequencing Center.

Aliquots of the completed insertion reactions are then used for transformation of chemically competent *E. coli*. First, insertion reactions are diluted by a constant factor to eventually achieve about 100 colonies per plate. Next, transformation reactions are performed in a 96-well format with an aliquot of the diluted DNAs. Last, the entire volume of each transformation is plated on a its own 90-mm Petri dish containing appropriate antibiotics. Despite the variability in DNA concentration and insert size between samples, the EZ::TN System produces fairly uniform numbers of colonies for each cDNA clone using our optimized conditions.

Sequencing of cDNA Insertion Subclones

EZ::TN Transposon-containing insertion subclones are entered into our standard sequencing pipeline, thereby benefiting from considerable economies of scale. Colonies are picked with a QPix automated picker (Genetix) and transferred to medium in 96-well growth blocks. After overnight growth, DNAs are isolated in 96-well format using the Concert96 Plasmid Purification System (Life Technologies) and a second Microlab 4200 pipetting robot. For each subclone DNA, paired sequencing reactions containing BigDye Terminator chemistry (Applied Biosystems) are set up with the first pipetting robot. Each reaction of the pair utilizes a primer complementary to one or the other

transposon end, facing into the inserted DNA. The cycling parameters are: 96°C for 1 min, followed by 35 cycles of 96°C for 10 sec; 55°C for 10 sec; and 60°C for 4 min. Completed reaction products are purified, and then analyzed on AB3700 capillary electrophoresis instruments (Applied Biosystems). Success rates for sequencing reactions with transposon-specific insertions are similar to those obtained with other samples in our production pipeline.

All the sequencing data for each cDNA clone are assembled with the use of the conventional program, phrap. We usually see a uniform coverage of the entire cDNA insert with individual reads from the ends of transposon inserts.

Project Status

The ability of the EZ::TN System to be used for processing multiple DNA samples in a batch mode makes it an effective component of a high throughput sequencing pipeline. This transposon-based sequencing system permits parallel independent processing of a large number of clones and is easily scalable to meet the growing demands of the project.

At present, we have successfully integrated transposon-based full-length cDNA sequencing into the main NISC sequencing pipeline and our MGC effort is on track to generate 4000-5000 high-quality cDNA sequences in the coming year.

References

1. Strausberg *et al.* (1999) *Science* **286**, 455.
2. Yu *et al.* (1997) *Genome Res* **7**(4), 353.
3. Andersson *et al.* (1997) *DNA Seq* **7**(2), 63.

EZ::TN™ <KAN-2> Insertion Kit

EZI982K 10 Reactions

EZ::TN™ <TET-1> Insertion Kit

EZI921T 10 Reactions

EZ::TN™ <DHFR-1> Insertion Kit

EZI912D 10 Reactions

Each kit contains the specific EZ::TN™ Transposon, EZ::TN™ Transposase, Buffers and two unlabeled sequencing primers.

Contact EPICENTRE or your local distributor to discuss discounts for bulk quantities.