Automation of a protocol for rapidly analyzing gene expression on a large scale will yield faster results at less cost

June 22 2012

Figure 1: The automated CAGE cDNA preparation system for rapid genome-wide gene expression analysis. © 2012 Masayoshi Itoh, RIKEN Omics Science Center

To generate an overall view of cell function, molecular biologists build simultaneous expression, or activity, profiles of thousands of genes. Gene expression begins with a process called transcription, during which the DNA sequence encoding a gene is copied into RNA. The information contained in the transcript is then translated into a chain of amino acids that folds up to form a functional protein molecule.

Building on a technique called cap analysis of gene expression (CAGE) to identify and analyze transcription start sites and their expression levels , a research team in Japan, led by Masayoshi Itoh of the RIKEN Omics Science Center in Yokohama, has developed a large-scale technique for analyzing gene expression that increases productivity eight-fold compared to previous methods.

The original CAGE protocol involves 17 steps including a process that 'traps' the ends of the original transcripts, numerous extractions using organic solvents, and centrifugation to create complementary DNA (cDNA) libraries. Itoh and his colleagues simplified and eliminated the PCR amplification step to reflect the original expression profile. The prepared libraries consisted of millions of DNA molecules, each corresponding to the transcription start site. The researchers then determined the sequence of these cDNA molecules, and the amount of each one, using a single molecule sequencer called the HeliScope Genetic Analysis System.

Since each step in the original protocol was performed manually, the technique was labor intensive and it took approximately six weeks to prepare 96 cDNA libraries. Itoh and colleagues currently have two of the new systems in place that allow them to generate 192 cDNA libraries every eight days. The automation cuts sequencing costs because it involves less manual work, and the system can be used with the three most common DNA sequencing platforms.

Itoh and his colleagues achieved this improvement by adapting CAGE so that all the steps are completed by an automated system consisting of a robotic manipulator arm equipped with an 8-channel liquid-handling device (Fig. 1). They simplified the protocol further by using magnetic beads to purify the cDNA. The automated method is significantly faster than the original protocol and reliably produces highly accurate sequences. Since constant supervision is not required, two or more

systems can be easily operated simultaneously.

"We are currently developing the automation system for the next generation Illumina HiSeq2000 sequencing platform," says Itoh. "This involves improving the steps that produce and trap the transcription start site cDNAs, and we aim to establish the new library preparation system in the first half of 2012."

More information: Itoh, M., Kojima, M., Nagao-Sato, S., Saijo, E., Lassmann, T., Kanamori-Katayama, M., Kaiho, A., Lizio, M., Kawaji, H., Carninci, P., et al. Automated workflow for preparation of cDNA for cap analysis of gene expression on a single molecule sequencer. PLoS ONE 7, e30809 (2012). www.plosone.org/article/info %3Adoi%2F10.1371%2Fjournal.pone.0030809

Shiraki T., Kondo S., Katayama S., Waki K., Kasukawa T., Kawaji H., Kodzius R., Watahiki A., Nakamura M., Arakawa T., Fukuda, S., Sasaki, D., Podhajska, A., Harbers, M., Kawai, J., Carninci, P., Hayashizaki, Y. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proceedings of the National Academy of Sciences 100, 15776–15781 (2003). www.pnas.org/content/100/26/15776.full

Kanamori-Katayama M., Itoh M., Kawaji H., Lassmann T., Katayama S., Kojima, M., Bertin, N., Kaiho, A., Ninomiya, N., Daub, C. O., Carninci, P., Forrest, A. R. R., Hayashizaki, Y. Unamplified cap analysis of gene expression on a single molecule sequencer. Genome Research 21, 1150–1159 (2011). genome.cshlp.org/content/21/7/1150.short

Provided by RIKEN