Perspective   Open Access

Comparison between microarray hybridization and serial analysis of gene expression

J Kim[*]

Department of Chemistry, Busan National University, Busan, Republic of Korea

[*]Corresponding author. E-mail: kimso123@ac.kr

DECSRIPTION

The main difference between microarray hybridization and Serial Analysis of Gene Expression (SAGE) techniques is that SAGE is a sequencing-based gene expression profiling technique, whereas microarray hybridization requires prior knowledge of the sequences to be analyzed. SAGE can be used to obtain complete transcriptional profiles of expressed genes, albeit unknown genes, in organisms with poorly characterized genomic and expressed sequences. Even after the human genome has been sequenced, predicting all of the encoded genes remains a significant challenge. Long SAGE, a recent SAGE adaptation, allows the derived transcriptase to be used for annotating expressed genes in the genome. SAGE is thus a truly global and unbiased gene expression technique. The decision between SAGE and microarray hybridization is influenced by several factors, including the scope of the genetic screen, the number of samples, the amount of starting material, and the availability of resources such as an automated DNA sequences. A simple two-library comparison requires approximately 106 bases to be sequenced, but with current advances in automated sequencing, this should not be a significant impediment for most laboratories. Using the microarray technique, however, large numbers of samples can be analysed more efficiently. The SAGE technique's Polymerase Chain Reaction (PCR) amplification step reduces the amount of RNA required for sample-limiting experiments when compared to standard microarray techniques.

In studies that compared SAGE and microarray data, there appeared to be a good correlation between the two techniques though one recent study found the SAGE technique to be more quantitatively reproducible. It is also possible to use both techniques simultaneously by screening broadly with the SAGE technique, followed by a more focused custom-made microarray experiment. To date, no software exists for directly comparing SAGE to microarray databases, but a software capable of visualizing metabolic pathways at the genome-wide level using either microarray or SAGE databases has recently been reported. The amount of RNA required ranges between 50 and 500 ng of mRNA or 5 and 50 g of total RNA. SAGE libraries were created using far less material (9 oocytes to 100 000 cells).The construction of a SAGE library from a one-cell mRNA equivalent was recently reported, though the final quality of this library has yet to be demonstrated. SAGE has a distinct advantage due to its sequencing requirement. Its digital database enables direct SAGE library comparisons. The ability to perform virtual Northern blots and query over 100 human SAGE libraries representing three million transcripts *via* the internet represents a powerful resource.

Protocol and Concepts

SAGE employs a number of enzymatic, PCR amplification, purification, and cloning steps, but it is based on two fundamental principles. The first principle is that short oligonucleotide sequences defined by a specific restriction endonuclease (anchoring enzyme, AE) at a fixed distance from the poly(A) tail can uniquely identify mRNA transcripts. A 10-bp sequence tag can theoretically generate different sequence combinations, which is more than enough to differentiate all transcripts derived from the human genome. The second principle is that by concatenating these short oligonucleotides end-to-end, multiple transcripts can be detected per sequencing reaction.