Reproducibility of Alternative Probe Synthesis Approaches for Gene Expression Profiling with Arrays

Suzanne D. Vernon, Elizabeth R. Unger,
Mangalathu Rajeevan, Irina M. Dimulescu,
Rosane Nisenbaum, and Catherine E. Campbell

From the Centers for Disease Control and Prevention, United States Public Health Service, Department of Health and Human Services, Atlanta, Georgia

Before gene expression profiling with microarray technology can be transferred to the diagnostic setting, we must have alternative approaches for synthesizing probe from limited RNA samples, and we must understand the limits of reproducibility in interpreting gene expression results. The current gold standard of probes for use with both microarrays and high-density filter arrays are synthesized from 1 μ g of purified poly(A)+ RNA. We evaluated two approaches for synthesizing cDNA probes from total RNA with subsequent hybridization to high-density filter arrays: 1) reverse transcription (RT) of 5 μ g total RNA and 2) RT-polymerase chain reaction (RT-PCR) of 1 μ g total RNA, using the SMART system. The reproducibility of these two approaches was compared to the current gold standard. All three methods were highly reproducible. Triplicate experiments resulted in the following concordance correlation coefficients to evaluate reproducibility: 0.88 for the gold standard, 0.86 for cDNA probe synthesized by RT from total RNA, and 0.96 for the SMART cDNA probe synthesized from total RNA. We also compared the expression profile of 588 genes for the total RNA methods to that obtained with the gold standard. Of 150 positive genes detected by the gold standard, 97 (65%) were detected by cDNA probe synthesized by RT of total RNA, and 122 (81%) were detected by the SMART cDNA probe. We conclude that SMART cDNA probe produces highly reproducible results and yields gene expression profiles that represent the majority of transcripts detected with the gold standard. (*J Mol Diag* 2000, 2:124–127)

It is often not possible to obtain clinical samples that will yield sufficient quantities of purified mRNA for array-based applications. Thus alternative methods for making labeled cDNA probes are required to transfer the power of arrays to clinical and epidemiological research settings.

Synthesis of labeled cDNA probes from preparations of total RNA is the most promising alternative to purified mRNA for array applications. Indeed, CLONTECH Laboratories (Palo Alto, CA) and Research Genetics (Huntsville, AL), both manufacturers of filter arrays, include protocols in the user manuals for the preparation of cDNA probes from 0.5–10 μ g of total RNA. Polymerase chain reaction (PCR)-based cDNA methods for amplification from limited amounts of RNA are also being used for differential gene expression profiling. The SMART PCR cDNA synthesis method (CLONTECH Laboratories) was used in gene expression profiling experiments to produce cDNA libraries from total RNA that were representative of the mRNA.¹ *In vitro* transcription of heterogeneous cDNA synthesized from total RNA² has been modified to incorporate biotin into the probe and adapted to high-density arrays.³ Key to these alternative approaches is a reproducible and representative synthesis of complex cDNA probes derived from the mRNA population present in the original sample.

Determining the reproducibility of various cDNA labeling methods with subsequent hybridization to an array is important if we ultimately want to use the arrays to associate specific gene transcripts with a disease process or use arrays as population-based disease screening tools. We evaluated two approaches for synthesizing cDNA probes from total RNA with subsequent hybridization to high-density filter arrays: 1) reverse transcription (RT) of 5 μ g total RNA and 2) RT-PCR of 1 μ g total RNA with the SMART system. The reproducibility of these two approaches was compared to the current gold standard, cDNA probe synthesized from 1 μ g of purified poly(A)+ RNA.

The most effective way to represent a gene expression profile is to synthesize a labeled cDNA from a purified mRNA template. Most published procedures create gene expression profiles by synthesizing cDNA probes with 1 μ g of purified mRNA from at least 100 μ g of total RNA.

Accepted for publication June 1, 2000.

Address reprint requests to Dr. Elizabeth R. Unger, Centers for Disease Control and Prevention, 1600 Clifton Road, MSG18, Atlanta, GA 30333. E-mail: eunger@cdc.gov.

Materials and Methods

RNA Preparations

Total RNA was extracted from monolayers of a cervical cancer epithelial cell line (CaSki; ATCC) by the modified guanidinium thiocyanate method.⁴ After DNase treatment, 800 μg total RNA was used to purify 5 μg poly(A)+ RNA with an Oligotex mRNA kit (Qiagen, Santa Clarita, CA). For DNase treatment of smaller total RNA aliquots, we treated 10 μg of total RNA with 1 U DNase I (GenHunter Corporation, Nashville, TN) for 15 minutes at room temperature and usually recovered 80% of the total RNA. All RNA samples were examined for the absence of DNA and RNA degradation by denaturing agarose gel electrophoresis and were quantified by UV spectrophotometry.

Digoxigenin-Labeled cDNA Probes

Labeling methods were based on some of the optimized conditions for chemiluminescent detection of digoxigenin-labeled cDNA probes,⁵ with the following modifications. For the poly(A)+ cDNA probe, 1 μg of poly(A)+ RNA was reverse transcribed for 55 minutes at 42°C, using the reverse transcriptase, buffer, and oligo d(T) primers from the Superscript II preamplification kit (Life Technologies, Gaithersburg, MD) and a dig-11-dUTP nucleotide mix (Roche Molecular Biochemicals, Indianapolis, IN). The reaction was stopped with heat at 70°C for 15 minutes, followed by 2 U RNase H (Roche Molecular Biochemicals) treatment. One microliter of the 20- μl reverse transcription (RT) reaction was evaluated for dig-11-dUTP incorporation by denaturing acrylamide electrophoresis. The remaining 19 μl of the preparation was used for hybridization. This digoxigenin-labeled probe will be referred to as the poly(A) probe in subsequent sections.

Digoxigenin-labeled cDNA probe from total RNA was synthesized as described above, with the following modifications. Five micrograms of total RNA was reverse transcribed with oligo d(T) for 2 hours at 42°C, with an additional 200 U Superscript II added after the first hour. The labeled cDNA was treated with 2 U each RNase H and RNase A (Roche Molecular Biochemicals) for 20 minutes at 37°C. One microliter of the 20- μl RT reaction was evaluated for dig-11-dUTP incorporation by denaturing acrylamide electrophoresis. The remaining 19 μl of the preparation was used for hybridization. This type of digoxigenin-labeled probe will be referred to as the total RNA probe in subsequent sections.

Digoxigenin-labeled double-stranded cDNA probes were synthesized using the SMART PCR cDNA synthesis kit and the Advantage cDNA PCR kit (CLONTECH Laboratories). The first-strand cDNA synthesis was done as specified in the manufacturers' user manual and included 1 μg total RNA, the CDS synthesis primer, the SMART II oligonucleotide, and 200 U Superscript II. The double-stranded cDNA was amplified with Advantage cDNA polymerase mix, PCR primer (from the SMART-PCR cDNA synthesis kit), and dig-11-dUTP nucleotide mix. A 1- μg total RNA sample without RT was included to

monitor DNA contamination. The PCR conditions were 95°C for 1 minute, followed by 15 cycles of 95° for 5 seconds, 65°C for 5 seconds, and 68°C for 6 minutes. One microliter of the 100- μl PCR reaction was evaluated for dig-11-dUTP incorporation by agarose gel electrophoresis. The no RT control sample was also evaluated for dig-11-dUTP incorporation in this manner. Fifty microliters of the labeled PCR product was used in the hybridization. This type of digoxigenin-labeled probe will be referred to as the SMART probe in subsequent sections.

Hybridization and Chemiluminescent Detection

All digoxigenin-labeled probes were hybridized to the Atlas Human cDNA expression array (CLONTECH). Atlas cDNA expression arrays contain 588 carefully selected cDNA fragments ranging in size from 200 to 600 bp arrayed in duplicate on nylon membranes. Hybridization and chemiluminescent detection of the arrays were done as previously described,⁵ except for the following modifications for each of the various probes. Hybridization was extended to 48 hours at 42°C for the total RNA probes. Hybridization of the SMART probes was overnight at 44°C. Poly(A) and total RNA probes were denatured at 95°C for 3 minutes. The SMART probes were denatured by boiling for 10 minutes. All filters were prehybridized with 100 $\mu\text{l}/\text{cm}^2$ Dig Easy Hyb (Roche Molecular) containing 0.5 $\mu\text{g}/\text{ml}$ cot DNA and poly dA (amounting to 10 ml of prehybridization solution) and hybridized with 50 $\mu\text{l}/\text{cm}^2$ Dig Easy Hyb containing 0.5 $\mu\text{g}/\text{ml}$ cot DNA and the specified volume of probe. For Atlas membranes, this corresponds to 5 ml of hybridization solution per membrane.

Image Acquisition

The chemiluminescent signal generated by CDP-Star was detected by a 1-hour exposure to Lumi Film (Roche Molecular Biochemicals). The film was scanned on a flatbed scanner, and tagged image file formats were generated. These array image files were subsequently loaded into BioNumerics (Applied Maths, Kortrijk, Belgium) for gene intensity quantification. Gene hybridization intensities were determined after background subtraction. Because of the variability in background level between the array images, each array image was normalized to itself. This was done in BioNumerics by defining the lowest intensity blank control on the filter to zero, and the highest intensity positive control on the filter to 100%. The program quantified all gene intensities according to the set calibration. To compensate for experimental variability, we calculated a cutoff value (COV) to determine positive and negative values. The COV for each image was determined using the average of the three negative controls on the filter image with the lowest intensity plus one SD.

Data Analysis

We assessed the reproducibility of the gene hybridization intensities for triplicate readings of each cDNA labeling

method, using two approaches: the Pearson correlation and the concordance correlation.⁶ The Pearson correlation measures a linear relationship between a pair of readings, and the concordance correlation measures the level of agreement between two or more readings in relation to the 45° identity line. Because the Pearson correlation is limited to a pairwise calculation, experiment 1 was compared to experiment 2, then experiment 1 to experiment 3. The lowest of these two Pearson coefficients is shown. All 588 normalized numerical gene values for each image were used in the calculations. Statistical analysis used BioNumerics, Microsoft Excel, and SAS (SAS Institute, Cary, NC).

Results

To determine the reproducibility of the various cDNA probe synthesis procedures, we performed three separate labeling reactions for each method on the same RNA preparation. Figure 1 graphically depicts the log base 2 of the hybridization intensities for the triplicate experiments of each labeling method. For each graph, the results of experiment 1 are plotted on the x axis, and the results of experiment 2 and 3 are plotted on the y axis. If all results were in perfect agreement, the points would fall on the 45° identity line. The two parallel lines indicate the position of twofold differences in intensity.

The three poly(A) probes gave concordant results for 84% of the 588 genes (150 positive genes and 342 negative genes). There were 96 discrepant genes (positive in only one or two experiments). The reproducibility of this method was high (Pearson correlation coefficient 0.94, concordance correlation coefficient 0.88). The three total RNA probes were concordant for 73% of the 588 genes (108 positive genes and 322 negative genes). There were 158 discrepant genes by this method. The reproducibility of total RNA probes was also good (Pearson correlation coefficient 0.90, concordance correlation coefficient 0.86). The three SMART probes were concordant for 79% of the 588 genes (212 positive genes and 250 negative genes). There were 126 discrepant genes detected. The reproducibility of the SMART probe method was very good, as reflected by nearly perfect coefficients (Pearson correlation coefficient 0.96, concordance correlation coefficient 0.96).

To evaluate the representation of the two probes synthesized from total RNA, we compared the gene expression profile of the total RNA and SMART probes to that generated by the poly(A) probes. Of the 150 positive genes detected with the three poly(A) probes (the gold standard), 97 (65%) were detected with all total RNA probes and 122 (81%) with all SMART probes. Of the 342 genes not detected with the poly(A) probes, 293 (86%) were also negative with the total RNA probes and 225 (66%) were negative with the SMART probes.

Discussion

Microarray technology has matured to the point where understanding the limits of reproducibility is important in

Figure 1. The log base 2 values of the hybridization intensities for the triplicate experiments are plotted on a graph for each labeling method. For each scatter plot, the results of experiment 1 compared to experiment 2 are plotted in blue, and the results of experiment 1 compared to experiment 3 are plotted in red. If all results were in perfect agreement, the points would fall on the 45° identity line. The two solid parallel lines indicate the position of twofold differences in intensity. The two dashed perpendicular lines indicate the COVs. Scatter plots are shown for the three poly(A) probe experiments (A), the three total RNA probe experiments (B), and the three SMART probe experiments (C).

interpreting results. Repeat analysis of gene expression in the same sample with the same technique should ideally yield identical profiles. Evaluating how well replicate assays achieve this ideal allows empirical determination of the criteria for differential gene expression. For example, if replicate assays reliably reproduce gene expression levels within a twofold difference, then a greater than twofold difference in gene expression between samples could be interpreted as differential expression.

Reproducibility is affected by all steps of the assay. We evaluated the impact of two different labeling methods for probes from total or poly(A)+ RNA isolated from the same source by performing each labeling, hybridization and image analysis in triplicate. Reproducibility for each method was evaluated graphically (Figure 1). For the 588-gene array, triplicate assays for the poly(A) probe gave concordant results for 83.6% of the genes. The total RNA probes gave the lowest number of concordant results (73.1%). In each case, a few genes fell outside of

the twofold cutoff lines, but a greater number were detected in only one or two experiments. Concordance values for these alternative probe synthesis approaches with high-density filter arrays have not previously been reported. Mahadevappa and Warrington³ report 83% concordance for 1779 genes assayed with a fluorescent microchip array.

Concordance was measured using two different correlation coefficients. The reproducibility index⁶ is designed to "evaluate the agreement between two or more readings from the same sample by measuring the variation from the 45° line through the origin (the concordance line)." Other coefficients, such as the Pearson correlation coefficient, measure a linear relationship but are limited to a pairwise comparison and do not detect scatter from the concordance line. In these comparisons, the Pearson and reproducibility index followed a similar trend. The range of the reproducibility index was slightly greater, suggesting that it may be a more refined measure of concordance.

Neither coefficient reflects the true significance of the discrepant values. From a technical point of view, there are a number of steps in the procedure that can be the source of these discrepancies. From a biological perspective, the complexity of the mRNA template and the efficiency of the enzymatic reactions may also be the explanation for these discrepant values. There have been attempts to preserve mRNA complexity by increasing the efficiency of the RT reaction.⁷ Amplification can also be a source of variability in the experimental procedure.⁸ All three probe labeling approaches we used here involve an RT step, and all three probes had similar levels of discrepant gene detection. This argues that the discrepant values are due to sample complexity and RT, rather than amplification. The Smart PCR cDNA method was originally developed to produce a high-quality, full-length cDNA for library construction. Recently, the Smart PCR cDNA synthesis was used to confirm differentially expressed genes identified on microarrays.¹

Having established the reproducibility of each method, we compared the gene expression profile for each of the total RNA labeling methods to that obtained for the poly(A) probe. Purified mRNA from a total RNA sample has served as the gold standard for gene expression analysis in most array applications. The SMART probe detected 81% of the genes detected with the poly(A) probe, compared to 65% for the total RNA probe. In fact, the SMART probes detected more genes than did the

poly(A) probes (212 *versus* 150). An apparent increase in representation for a total RNA sample relative to poly(A)+ RNA was described for an *in vitro* transcription labeling system.³ Because poly(A)+ RNA requires extensive purification, loss may occur during all of the isolation steps. The 1 µg poly(A)+ RNA used in the hybridization is derived from approximately 100 µg of total RNA. In contrast, 1 µg total RNA template for the SMART probe represents 0.01 µg poly(A)+ RNA. The PCR method clearly boosts the sensitivity of detection. The expression of those genes detected only by the SMART probes will have to be validated.

We conclude that gene expression analysis using high-density filter arrays with probes generated from either poly(A)+ RNA or total RNA shows a reproducibility similar to that reported for fluorescent microarrays. The SMART probe synthesis method is extremely reproducible and efficient. For situations with limited RNA, the SMART probe synthesis from only 1 µg total RNA will give reproducible and representative gene expression profiles. Because only a fraction of the initial cDNA reaction is subsequently labeled, this approach generates additional material that can be used for further testing and validation.

References

1. Endege WO, Steinmann KE, Boardman LA, Thibodeau SN, Schlegel R: Representative cDNA libraries and their utility in gene expression profiling. *Biotechniques* 1999, 26:542-550
2. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH: Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA* 1990, 87:1663-1667
3. Mahadevappa M, Warrington JA: A high-density probe array sample preparation method using 10- to 100-fold fewer cells. *Nature Biotech* 1999, 17:1134-1136
4. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform. *Anal Biochem* 1987, 162:156-159
5. Rajeevan MS, Dimulescu IM, Unger ER, Vernon SD: Chemiluminescent analysis of gene expression on high-density filter arrays. *J Histochem Cytochem* 1999, 47:337-342
6. Lin LI-K: A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989, 45:255-268
7. Zhang J, Byrne CD: Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. *Biochem J* 1999, 15:337:231-241.
8. Brail LH, Jang A, Billia F, Iscove NN, Klamut HJ, Hill RP: Gene expression in individual cells; analysis using global single cell reverse transcription polymerase chain reaction (GSC RT-PCR). *Mutat Res* 1999, 406:45-54