

Linear mRNA amplification from as little as 5 ng total RNA for global gene expression analysis

Alan Dafforn, Pengchin Chen, Glenn Deng, Michael Herrler, Dawn Iglehart, Sriveda Koritala, Susan Lato, Susheela Pillarisetty, Reshma Purohit, Martin Wang, Shenglong Wang, and Nurith Kurn

Reprinted with permission from *BioTechniques* 37:854-857 (November 2004)

Gene expression analysis has become an invaluable tool for understanding gene function and regulation. However, global expression analysis requires large RNA quantities or RNA preamplification. We describe an isothermal messenger RNA (mRNA) amplification method, Ribo-SPIA™, which generates micrograms of labeled cDNA from 5 ng of total RNA in 1 day for analysis on arrays or by PCR quantification. Highly reproducible GeneChip® array performance ($R^2 > 0.95$) was achieved with independent reactions starting with 5–100 ng Universal Human Reference total RNA. Targets prepared by the Ribo-SPIA procedure (20 ng total RNA input) or the Affymetrix Standard Protocol (10 µg total RNA) perform similarly, as indicated by gene call concordance (86%) and good correlation of differential gene expression determination ($R^2 = 0.82$). Accuracy of transcript representation in cDNA generated by the Ribo-SPIA procedure was also demonstrated by PCR quantification of 33 transcripts, comparing differential expression in amplified and nonamplified cDNA ($R^2 = 0.97$ over a range of nearly 10^6 in fold change). Thus Ribo-SPIA amplification of mRNA is rapid, robust, highly accurate and reproducible, and sensitive enough to allow quantification of very low abundance transcripts.

INTRODUCTION

Gene expression analysis has become a standard tool for understanding gene function in both normal and disease states. Application of array-based methods has, however, been limited by the time-consuming and complex sample preparation required and the large sample size needed for efficient hybridization to the array. Furthermore, many of the samples of most interest, such as those obtained from laser microdissection, cell sorting, or needle biopsies, are very small, yielding far less than

the microgram quantities of RNA required for expression analysis on arrays (1).

Several methods have been developed to enable global gene expression analysis of these small samples. A method based on T7 amplification was first described by Eberwine and colleagues (2) and is now recommended by Affymetrix for sample preparation for GeneChip® arrays (http://www.affymetrix.com/support/technical/technotes/smallv2_technote.pdf). In this method, two or more rounds of T7 RNA polymerase transcription are used to generate enough complementary RNA (cRNA) for mi-

croarray-based analysis. Although capable of reasonable representation of transcripts in the nonamplified material (3,4), this approach is rather lengthy, tedious, and requires highly skilled operators. Furthermore, different numbers of amplification cycles are currently used for preparation of hybridization targets from samples of various sizes, leading to complication in comparison of experiments. **PCR-based methods** for global amplification have also been described (5,6), but are also somewhat cumbersome and subject to **concerns** about fidelity of transcript representation because **of the exponential nature of PCR.** Thus, there is a need for a rapid, simple, general procedure for global gene expression analysis suitable for all total RNA samples, both large and in the low nanogram range.

Expression analysis by quantitative PCR requires less sample preparation. However, multiplexed or miniaturized quantitative PCR analysis can be limited by single-molecule sensitivity limits when the sample is divided into a large number of aliquots. Preamplification of the sample then becomes necessary to generate statistically significant numbers of molecules in each aliquot and thus to enable high-order multiplexing and miniaturization of quantitative PCR.

We describe a novel RNA-based, single primer, isothermal amplification, Ribo-SPIA™, and simple fragmentation and labeling methods designed to circumvent these limitations. The combined procedures generate microgram amounts of amplified cDNA ready for hybridization to GeneChip arrays, or spotted arrays, from 5 ng of starting total RNA in 1 day. Amplified cDNA for quantitative PCR analysis can be generated in under 5 h.

MATERIALS AND METHODS

Materials

Universal Human Reference (UHR) total RNA from Stratagene (La Jolla, CA, USA), human spleen and placenta total RNA from Ambion (Austin, TX, USA), and human skeletal muscle total RNA from BD Biosciences Clontech (Palo Alto, CA, USA) were used. RNA quality was determined using a Bioanalyzer (Agilent, Palo Alto, CA, USA). PCR primers were purchased from Qiagen (Valencia, CA, USA).

Ribo-SPIA

All Ribo-SPIA (NuGEN™ Technologies, San Carlos, CA, USA) amplifications utilized Ovation™ Aminoallyl or Ovation™ Biotin kits (both from NuGEN Technologies), according to manufacturer's instructions (<http://www.nugeninc.com>).

Amplification and Labeling by In Vitro Transcription

For comparison, the same RNA samples were labeled using the Enzo® BioArray™ High Yield™ RNA Transcript Labeling Kit (Affymetrix, Santa Clara, CA, USA). All in vitro transcription (IVT)-based labeling reactions were performed

by Expression Analysis (Durham, NC, USA), an Affymetrix-authorized GeneChip service provider.

Quantitative PCR

Primer pairs for 39 genes representing a wide range of expression in skeletal muscle were designed using Primer Express® software (Applied Biosystems, Foster City, CA, USA). UHR RNA and skeletal muscle total RNA were reverse-transcribed into cDNA and amplified using Ribo-SPIA reagents. Expression levels before and after amplification were quantified by real-time PCR using an MJ Opticon® (MJ Research, Waltham, MA, USA) and the QuantiTect® SYBR® Green PCR Kit (Qiagen). Results for six genes were filtered from the final graph (see Figure 3A). These appeared not to be expressed in UHR control combined cell lines, as indicated by high threshold cycles (Ct) values (>35) for both nonamplified and amplified cDNA. These genes were expressed in muscle and were amplified by Ribo-SPIA.

GeneChip Analysis

Samples labeled using either the Ovation Biotin kit or the Affymetrix standard protocol were hybridized to HG-U133A GeneChip arrays (Affymetrix), stained with streptavidin-phycoerythrin with antibody amplification and scanned following the manufacturer's protocols, except only 2.5 µg per chip of Ribo-SPIA product was hybridized (compared to 10 µg cRNA). Denaturation for 2 min at 99°C before hybridization was followed by hybridization for 20 h. Affymetrix

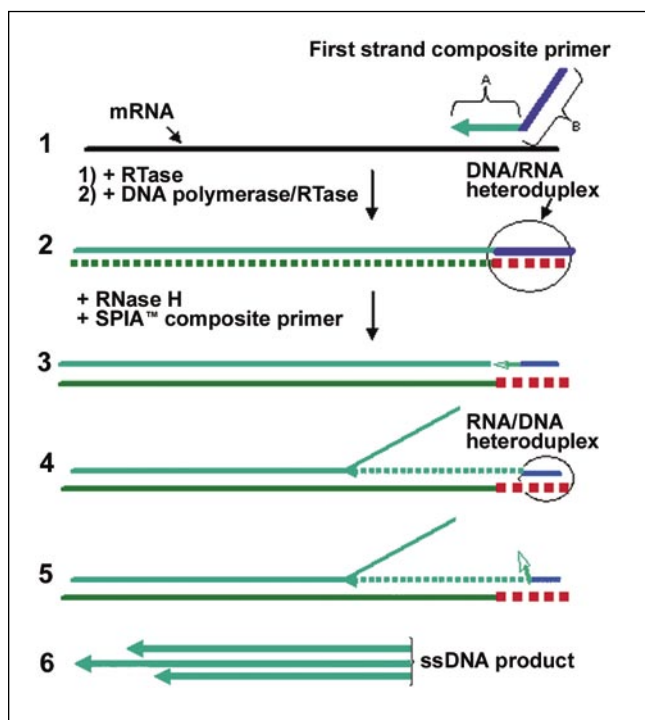


Figure 1. Outline of the Ribo-SPIA amplification process. Individual steps are described in Results and Discussion.

samples were amplified, hybridized, and scanned by Expression Analysis. Array data was analyzed by MAS5 software (Affymetrix).

RESULTS AND DISCUSSION

Ribo-SPIA, a rapid, isothermal RNA amplification process, is outlined in Figure 1. In the first step, a DNA-RNA chimeric primer hybridizes to the 3'-poly(A) tail of messenger RNA (mRNA) to form complex 1 (in Figure 1), where primer segment B is RNA, and is extended by a reverse transcriptase. This strand is copied by a DNA polymerase with reverse transcriptase activity to give double-stranded cDNA with an RNA-DNA heteroduplex at one end. Amplification of this cDNA is accomplished by a single primer, isothermal linear amplification, SPIA. Initially, RNase H unmasks the priming site by digesting RNA in the heteroduplex, revealing a single-stranded DNA sequence that is complementary to the chimeric SPIA primer. The SPIA primer binds to this site (Figure 1, complex 3) and is extended by a strand-displacing DNA polymerase. Once this extension reaction begins, RNase H again digests the RNA introduced by the primer at the 5' end of the new strand (Figure 1, complex 4), thus revealing part of the priming site to which the RNA portion of a new primer molecule can bind. Binding and extension of new molecules of primer (Figure 1, complex 5) leads to a continuous isothermal generation of multiple DNA copies complementary to the original mRNA sequence.

The continuous nature of the reaction makes it possible to rapidly generate microgram amounts of DNA from <5 ng of total RNA input. The reaction gives a product size distribution with most material below 2 kb for optimal performance on arrays designed for 3'-biased IVT methods (Figure 2). The Ovation Aminoallyl kit incorporates aminoallyl-dUTP, allowing subsequent reaction with dyes such as CyTM3/Cy5 for gene expression profiling on spotted arrays. Expression analysis of specific genes may also be carried out directly by various quantitative PCR methods such as TaqMan[®] or SYBR[®] Green detection. Finally, the Ovation Biotin kit

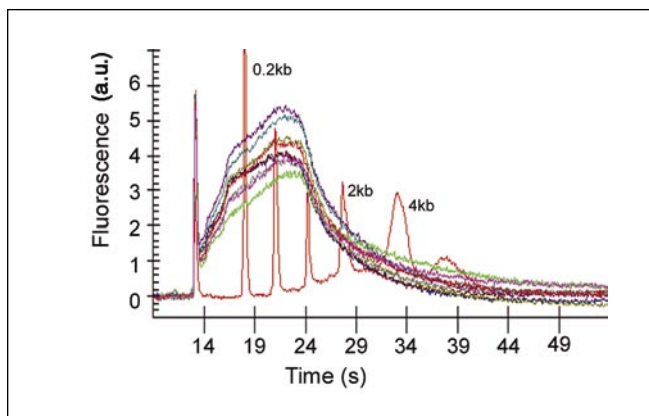


Figure 2. Bioanalyzer traces for nine replicate Ribo-SPIA products. The internal standard peaks are at 0.2, 0.5, 1, 2, 4, and 6 kb (plus a spiked oligonucleotide at 13 s). Peak amplitude varies slightly, depending on the amount of sample loaded. a.u., arbitrary units.

Table 1. Quality Control Metrics for GeneChips

Parameter	Ribo-SPIA Amplified	T7 Labeled
% Present Call	53.6 ± 2.7	51.5 ± 1.7
Scale Factor	1.66 ± 0.56	0.63 ± 0.15
Signal	669 ± 69	336 ± 123
Raw Q	3.53 ± 0.46	3.32 ± 0.51
3'/5' GAPDH	1.92 ± 0.17	0.87 ± 0.07
3'/5' β-actin	20.3 ± 5.0	1.08 ± 0.09

Quality control metrics are compared for three replicate GeneChip arrays with standard T7 samples and 13 replicates with Ribo-SPIA samples.

Table 2. Effect of Total RNA input on Signal Correlation Coefficient R^2 Between Duplicate GeneChip Arrays

	1 ng	5 ng	20 ng	25 ng	100 ng
1 ng	0.97	0.97	0.96	0.94	0.93
5 ng		0.96	0.99	0.97	0.95
20 ng			0.99	0.98	0.96
25 ng				0.98	0.98
100 ng					0.99

Signal correlation between duplicate GeneChip arrays as a function of total RNA input into Ribo-SPIA sample amplification. Correlation coefficients R^2 are given between arrays run with product prepared with the Universal Human Reference (UHR) total RNA inputs indicated.

provides reagents for fragmentation and biotin labeling of a cDNA product not containing aminoallyl groups, giving material suitable for hybridization to GeneChip arrays. The entire process from sample to array can be completed in a single day.

Accuracy and linearity of the amplification was assessed by quantitative PCR. Differential expression of 33 genes present in UHR total RNA and human skeletal muscle RNA was compared in nonamplified or amplified cDNA. Figure 3A shows an excellent linear relationship for \log_2 of relative expression levels before or after amplification (expressed as difference in Ct) over a range of nearly 20 Ct, or 6 orders of magnitude, with $R^2 = 0.97$. Thus Ribo-SPIA technology provides a reliable representation of changes in transcript abundance in nonamplified mRNA.

Preamplification of total mRNA by the Ribo-SPIA procedure before quantitative PCR analysis may prove especially useful in multiplexed microfluidic applications, so as to ensure that all aliquots contain a sufficient number of template molecules for accurate quantification. Since this approach, like IVT, is 3'-biased, careful selection of suitable sequences for amplification is necessary.

Performance was further explored using GeneChip arrays. Results were compared to those obtained from the same RNA samples by a T7-based IVT protocol. Hybridization of 2.5 μ g of labeled and fragmented Ribo-SPIA product gave results roughly equivalent to 10 μ g of cRNA prepared using the Enzo labeling kit. Three independent IVT labelings of UHR total RNA, run by an independent laboratory starting with 10 μ g of total RNA per amplification, gave an average reported Present Call = 51.5% ± 1.7%. Thirteen Ribo-SPIA amplifications with 20 ng total RNA input per

amplification gave Present Call = 53.6% ± 2.7%. Table 1 compares quality control metrics for the two methods. Most parameters are in the acceptable range, with even the GAPDH 3'/5' ratio in the recommended range under 3. High values for β-actin are also seen in material amplified by the Small Sample Affymetrix procedure (see technote smally2 referenced above for results with a different reference RNA) and are attributed to the length of the actin mRNA. Call concordance between the two methods was 86%, calculated by comparing 17,348 genes, which were either present or absent on all T7 arrays, to those on three randomly chosen Ribo-SPIA arrays. Thus the two methods detect comparable numbers of transcripts.

Reproducibility was assessed by duplicate amplifications and hybridizations over a broad range of total RNA

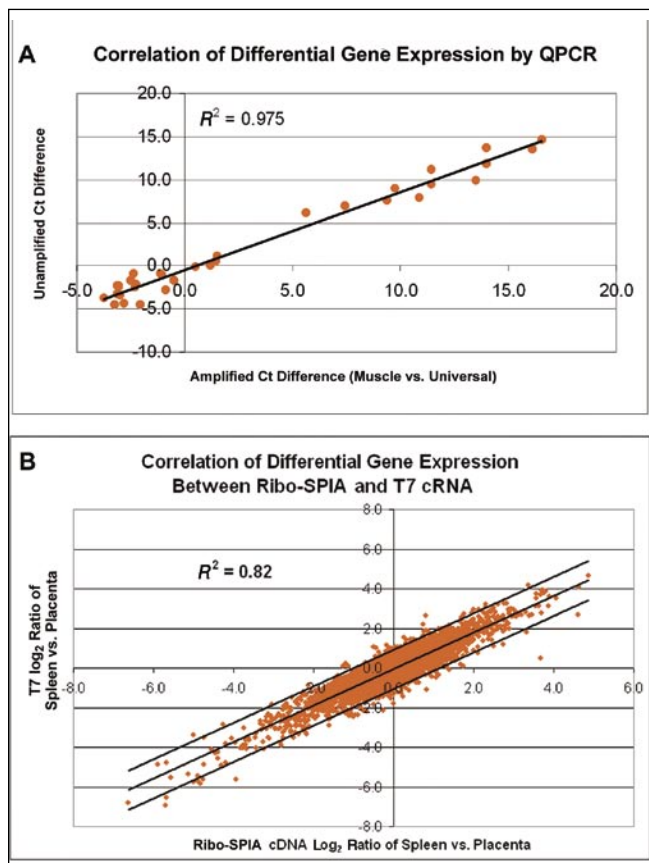


Figure 3. Differential gene expression correlations between methods. (A) Correlation between differential expression in skeletal muscle versus Universal Human Reference (UHR) total RNA measured before or after Ribo-SPIA amplification. Log₂ differential expression is represented as the difference between threshold cycles (Ct) measured by quantitative PCR for each gene. Results after Ribo-SPIA amplification (horizontal axis) are compared to results from starting cDNA before amplification. Each point is an average of results from duplicate Ribo-SPIA reactions. (B) Correlation between differential expression measured on GeneChip arrays for 20 ng samples amplified using Ribo-SPIA or 10 μg samples labeled with the normal Affymetrix protocol. Log₂ differential expression in human placenta versus spleen is plotted for the two methods. The linear least squares fit is shown, along with parallel lines reflecting discrepancies of a factor of 2 between the two methods. Each point is an average of values from triplicate arrays for each method. Genes shown were called present on all six arrays.

input (Table 2). Duplicate amplification reactions and GeneChip array analysis with the same RNA input gave excellent signal correlations, with $R^2 = 0.96-0.99$ over the 1–100 ng input range. Hybridization with targets generated from different total RNA inputs over the 5–100 ng range gave $R^2 = 0.95-0.99$, nearly independent of input. Even correlation between 1 and 100 ng was 0.93. Call concordance between independent amplifications ranged from 91% for target generated from 20 ng total RNA input to 86% for 1 versus 100 ng total RNA input. For comparison, call concordance between duplicate T7 IVT targets prepared from samples containing 10 μg total RNA was 92%. Reliable amplification is clearly possible over a broad range of total RNA input.

Accuracy and linearity of detection of individual transcripts on GeneChip arrays was assessed by comparing differential expression between two methods, since differential expression is the criterion most directly relevant to most applications. Differential expression between placenta and spleen RNA was evaluated using either the Ribo-SPIA method with 20 ng input or the Affymetrix IVT labeling with 10 μg input. Log₂ relative expression for the two tissues was then plotted for the two methods, as shown in Figure 3B. Comparison between methods gives an excellent correlation, with $R^2 = 0.82$ over three orders of magnitude. In most cases, fold change correlates within a factor of two between the two methods, as indicated by guide lines in the figure.

Overall, the Ribo-SPIA isothermal amplification offers results similar to T7-based sample preparation for the GeneChip. Comparable numbers of transcripts are detected, reproducibility within each method is similar, and differential expression is highly correlated between the two methods. This conclusion is reinforced by quantitative PCR analysis showing that differential expression of transcripts is preserved after Ribo-SPIA amplification. Other applications such as sample preparation for spotted arrays and preamplification for quantitative PCR have been demonstrated. Extensions of the technology to even greater sensitivity and application to automated systems are also envisioned.

ACKNOWLEDGMENTS

We are grateful to Joe Don Heath for consultation and to Liz Robertson for assistance in preparation of the manuscript and figures.

COMPETING INTERESTS STATEMENT

The authors are employed by NuGEN Technologies, Inc., the manufacturer of the kit described in this report.

REFERENCES

- Emmert-Buck, M.R., R.F. Bonner, P.D. Smith, R.F. Chuaqui, Z. Zhuang, S.R. Goldstein, R.A. Weiss, and L.A. Liotta. 1996. Laser cap-

-
- ture microdissection. *Science* 274:998-1001.
2. **Van Gelder, R.N., M.E. von Zastrow, A. Yool, W.C. Dement, J.D. Barchas, and J.H. Eberwine.** 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. USA* 87:1663-1667.
 3. **Luzzi, V., M. Mahadevappa, R. Raja, J.A. Warrington, and M.A. Watson.** 2003. Accurate and reproducible gene expression profiles from laser capture microdissection, transcript amplification, and high density oligonucleotide microarray analysis. *J. Mol. Diagn.* 5:9-14.
 4. **Wilson, C.L., S.D. Pepper, Y. Hey, and C.J. Miller.** 2004. Amplification protocols introduce systematic but reproducible errors into gene expression studies. *BioTechniques* 36:498-506.
 5. **Zhumabayeva, B., A. Chenchik, P.D. Siebert, and M. Herrler.** 2004. Disease profiling arrays: reverse format cDNA arrays complementary to microarrays. *Adv. Biochem. Eng. Biotechnol.* 86:191-213.
 6. **Seth, D., M.D. Gorrell, P.H. McGuinness, M.A. Leo, C.S. Lieber, G.W. McCaughan, and P.S. Haber.** 2003. SMART amplification maintains representation of relative gene expression: quantitative validation by real time PCR and application to studies of alcoholic liver disease in primates. *J. Biochem. Biophys. Methods* 55:53-66.

Address correspondence to Alan Dafforn, NuGEN Technologies Inc, 821 Industrial Road, Unit A, San Carlos, CA 94070, USA. e-mail: adafforn@nugeninc.com