Short Technical Report
Probe Generation Directly from Small Numbers of Cells for DNA Microarray Studies

C.C. Xiang, M. Chen, O.A. Kozhich[1], Q.N. Phan, J.M. Inman[1], Y. Chen[2], and M.J. Brownstein

NIMH/NIH, Bethesda, [1]National Cancer Institute at Frederick, Frederick, and [2]National Human Genome Research Institute, NIH, Bethesda, MD, USA

ABSTRACT

Recently, we described a technique that allows us to prepare probes for expression profiling from 0.5–1 μg RNA without template or signal amplification. However, we were unable to use this method to study cells harvested by needle biopsy, cell sorting, or laser capture microdissection. Here we give a new protocol for amplifying RNA with multiple reaction cycles and preparing fluorescent probes from approximately 10 cells. We use random 9-mers with a T3 RNA polymerase recognition sequence on the 5′ end for every round of cDNA synthesis except the first. The latter is primed with oligo(dT) with a T7 RNA polymerase recognition sequence on the 5′ end. Results were highly reproducible and reliable, and the products generated using our method seemed comparable to those produced using the RiboAmp RNA kit when both were used to do two cycles of amplification. To test our method's utility, we lysed cells directly into reverse transcription buffer containing RNase inhibitor and performed three rounds of RNA amplification. The expression profiles of mouse C2 and NIH 3T3 cells obtained with 11 232-element arrays using amplified RNAs were similar to those seen when probes were prepared from unamplified templates.

INTRODUCTION

DNA microarrays permit one to look at the expression of thousands of genes in parallel, and they have been used for numerous basic and clinical studies in the last few years (4,8,9,13). To take full advantage of such arrays, one would like to be able to profile small numbers of cells or even single cells. Labeling probes by direct labeling of total RNA requires a minimum of about 20 μg template, and analysis of small tissue samples is impossible using this method (6,7). We have recently described a technique that allows us to prepare probes from 0.5–1 μg RNA without signal or template amplification (15). For many experiments, this is adequate, but to study small samples, it is not.

In the past several years, techniques have been developed that are much less RNA intensive than ours (1,2,10,11). These methods depend on one or two rounds of template amplification, and they still require a good deal of starting material. For example, Wang et al. (14) described a procedure that allowed them to label 10 ng total RNA (made from a concentrated stock solution) with two rounds of RNA amplification mediated by T7 RNA polymerase (T7). Since a typical mammalian cell has only 5–10 pg RNA, this corresponds to the amount of RNA in 1000–2000 cells. Recently, Puskas et al. (12) compared two RNA amplification methods—T7-based in vitro transcription (one round amplification starting with 3 μg total RNA) and PCR-based amplification (25 cycles starting with 50 ng total RNA)—to a standard direct labeling protocol. They found that both amplification techniques give reproducible microarray data, but

they were not without problems. The false-positive and false-negative rates were higher than those seen with the conventional labeling protocol.

To the best of our knowledge, no quantitative array studies have been undertaken with fewer than 1000 cultured or laser-captured cells. Typically, one round of T7 amplification gives a 100- to 1000-fold increase in RNA, and two rounds give a 10 000-fold increase (5,14). Thus, as suggested earlier, two rounds should not generate enough template to label probe using the technique that we described earlier this year (15). Instead, three or more rounds of amplification would be needed. Here we show that several consecutive rounds of amplification can be performed using the T3N9 primer described above and that DNA probes prepared from approximately 10 cells give reliable expression profiles.

MATERIALS AND METHODS

Microarray Fabrication

cDNA microarrays with a total of 11 232 elements were printed on poly-L-lysine coated slides. The cDNAs used were provided by Bento Soares (University of Iowa, IA, USA). Plasmids were extracted from the bacteria using QIAprep® Turbo kits and a BioRobot™ 8000 (both from Qiagen, Valencia, CA, USA). The cDNA inserts were amplified with modified M13 primers (M13F, 5′-GTTGTAAA-ACGACGGCCAGTG-3′, and M13R, 5′-CACACAGGAAACAGCTATG-3′) and purified with MultiScreen® PCR plates (Millipore, Bedford, MA, USA).

The PCR products were diluted in 50% DMSO to an average concentration of 200 ng/µL. These products (5 µL each) were transferred to 384-well plates (Genetix, St. James, NY, USA) and then printed using an OmniGrid® arrayer (GeneMachines, San Carlos, CA, USA). The printed slides were aged for a week and then post-processed before hybridization. For detailed descriptions of coating glass with poly-L-lysine and post-processing the printed slides, please visit http://cmgm.stanford.edu/pbrown/mguide/index.html.

RNA Sample Preparation

Total RNA was extracted from mouse C2 and NIH 3T3 cells using TRIZOL® reagent (Invitrogen, Carlsbad, CA, USA). Amplified RNAs were prepared from total RNA or from lysed cells. For total RNA, we used 1 µg (in 2 µL) total RNA in the first round of amplification and 1 µg amplified RNA for the second and subsequent rounds. The RNA template was added to a 0.2-mL PCR tube containing 6 µL 5× first-strand buffer (Invitrogen), 31 µL RNase-free water (Invitrogen), and 1 µL RNase inhibitor (Promega, Madison, WI, USA). The volume was decreased to 23 µL in a vacuum centrifuge (SpeedVac®; Thermo Savant, Holbrook, NY, USA) before initiating the first-strand cDNA synthesis.

Cells were first collected and washed in cold 1× PBS. Then, they were counted and diluted to final concentrations of 5000 or 50 000 cells/mL. Two microliters of cell suspension (approximately 10 or 100 cells) were added to a 0.5-mL tube containing 6 µL 5× first-strand buffer (Invitrogen), 31 µL RNase-free water, and 1 µL RNase inhibitor. The cells were lysed in a cup sonicator (Misonix, Farmingdale, NY, USA), and the particulate matter was pelleted in a microcentrifuge at 16 000× *g* for 15 min at 4°C. The supernatant was transferred to a 0.2-mL PCR tube, and the volume was reduced to 23 µL as described above. DNase I (0.5 µL; Ambion, Austin, TX, USA) was added, and the genomic DNA was digested at 37°C for 30 min, following which the DNase was inactivated at 75°C for 5 min.

For first-strand synthesis, 1 µL of a 100 pmol/µL solution of T7dT primer (5′-GGCCAGTGAATTGTAATACG-ACTCACTATAGGGAGGCGGTTTT-TTTTTTTTTTTTTTTTT-3′; Operon, Alameda, CA, USA) was added to 23 µL RNA- or DNase-treated cell lysate. The RNA was denatured at 70°C for 10 min and chilled on ice for 10 min. dNTPs (1 µL 10 mM; Amersham Biosciences, Piscataway, NJ, USA), 3 µL 0.1 mM DTT (Invitrogen), and 2 µL SuperScript II™ reverse transcriptase (Invitrogen) were added to the tubes and incubated at 42°C for 2 h. For the second-strand cDNA synthesis, 81 µL RNase-free water, 30 µL 5× second-strand buffer (Invitrogen), 3 µL 10 mM dNTPs, 1 µL *E. coli* DNA ligase (Invitrogen), 4 µL *E. coli* DNA polymerase I (Invitrogen), and 1 µL RNase H (Invitrogen) were added to bring the total volume to 150 µL, and the tubes were incubated at 16°C for 2 h. At the end of this time, 2 µL T4 DNA polymerase (Invitrogen) were added, and the samples were incubated at 16°C for 5 min. Phase Lock Gel™ (Eppendorf, Westbury, NY, USA) and phenol-chloroform-IAA (Invitrogen) were used to extract the cDNA products, and Microcon-30® columns (Millipore) were then employed to concentrate and further purify them. The cDNAs were dried with a SpeedVac and resuspended in 4.5 µL RNase-free water. RNA was transcribed from the DNA template with a MEGA Script™ T7 reagents (Ambion), according the manufacturer's instructions, and purified with an RNeasy® Mini kit (Qiagen).

For second and subsequent rounds of amplification, we used a T3N9 primer (5′-GCGCGAAATTAACCCTCACTA-AAGGGAGAGGGGNNNNNNNNN-3′; Invitrogen) to drive first-strand cDNA synthesis. Second-strand cDNA synthesis and in vitro transcription were performed as described above.

When the RiboAmp™ kit (Arcturus, Mountain View, CA, USA) was employed to amplify purified samples of total RNA, the reagents were used according to the manufacturer's directions.

Probe Labeling with Amine-Modified Random Primers

Probes were synthesized from 5 µg total RNA, 2 µg amplified material produced from purifed RNA, or the entire product from lysed cells. The labeling method was described in detail earlier (15). Briefly, the RNA (15.5 µL) was combined with amine-modified random primer (AC6T hexamer; Sigma-Genosys, The Woodlands, TX, USA; 2 µg/µL, 2 µL) and RNase inhibitor (5 U/µL, 1 µL), incubated at 70°C for 10 min, and chilled on ice for 10 min. The primer/RNA solution was then added to the RT mixture [6 µL 5× first-strand buffer; 0.6 µL 50× aminoallyl dUTP/dNTPs (25 mM dATP, dGTP, and dCTP, 15 mM dTTP, and 10 mM aminoallyl dUTP); 3 µL 0.1 M DTT; 2 µL SuperScript II reverse transcriptase] and incubated at 42°C for 2 h. The reaction was terminated with 10 µL 0.5 M EDTA, and the RNA was hydrolyzed with 10 µL 1 M NaOH at 65°C for 30 min. The solution was neutralized with 10 µL 1 M HCl, and MinElute™ PCR purification kits (Qiagen) were used to purify the products. These were concentrated to 9 µL in a SpeedVac, and then 1 µL 1 M sodium bicarbonate, pH 9.3, was added to the cDNA solution, followed by 14.5 µL dye solution [NHS-ester Cy3 or Cy5 (Amersham Biosciences), 62.5 µg/µL in DMSO]. The resulting solution was mixed by pipetting it up and down several times; the tubes were wrapped in aluminum foil and incubated at room temperature for 1 h on an orbital shaker (USA Scientific, Ocala, FL, USA). The labeling reaction was stopped with 4.5 µL 4 M hydroxylamine hydrochloride (Sigma, St. Louis, MO, USA). The contents of the tubes were mixed, briefly centrifuged, and incubated for 30 min at room temperature in the dark. The probes were purified using QI-Aquick™ PCR purification kits (Qiagen). The products were partially dried in a vacuum centrifuge, and the volumes were adjusted to 23 µL with water. Then, 4.5 µL 20× SSC, 2 µL poly(A) (10 mg/mL), and 0.6 µL 10% SDS were added, and the probes were denatured at 100°C for 2 min. The solutions were pipetted onto arrays, coverslips were applied, and the slides were placed in a hybridization chamber (Corning, Corning, NY, USA). They were incubated in a 65°C water bath for 16–24 h, washed with 0.5× SSC, 0.01% SDS, followed by 0.06× SSC at room temperature (10 min each). The slides

were then placed in 50-mL Falcon® tubes and centrifuged for 5 min at 200× *g* at room temperature.

Array Scanning and Data Analysis

The arrays were scanned with a GenePix® 4000A scanner (Axon, Foster City, CA, USA) at 10 µm resolution. The PMT voltage settings were varied to obtain the maximum signal intensities with less than 1% probe saturation. The resulting images were analyzed using IPLab (Fairfax, VA, USA) and ArraySuite (NHGRI, Bethesda, MD, USA) software. The log$_2$ mean intensity was calculated to generate scatter plots for the self/self comparisons of probes prepared from amplified C2 RNA. To compare data from different rounds of amplification with T3N9 and to compare our method to the Arcturus RiboAmp method, C2/3T3 ratios were determined, all of the ratios were assigned quality scores, and elements with quality lower than 1 were removed from the dataset (3,15). The remaining ratios were then ranked from lowest to highest. The top 10% and bottom 10% of ranked genes were selected, and only those that were concordant in duplicate experiments with unamplified total RNA were considered "true" outliers. Then, the top and bottom 10% of ranked genes identified in each amplification experiment were compared to the list of genes obtained with unamplified probes.

RESULTS AND DISCUSSION

The amplification method we used is diagrammed in Figure 1. Complementary DNA synthesis from total RNA is primed by oligo-dT, on the 5′ end of which is a T7 RNA polymerase promoter sequence (T7dT). The RNA strands are replaced with DNA by RNase H, DNA polymerase I, and *E. coli* DNA ligase. Antisense RNA is then transcribed from the dsDNA template by T7 RNA polymerase. Up to this point, our method is rather similar to others that have been described in the literature. Now it diverges from them. DNA synthesis from transcribed RNA is primed by random 9-mers with T3 RNA polymerase promoter sequences

on their 5′ ends (T3N9). dsDNA is prepared as described above, and RNA is generated by T3 RNA polymerase. Steps 4 and 5 can be repeated as many times as one wants. The resulting products are not 3′-end biased.

First, we compared our amplification method to the one marketed by Arcturus, the RiboAmp RNA kit, which is recommended by the manufacturer for amplifying RNA from laser-captured cells. Since the instructions provided indicate that the kit can only be used for two rounds of amplification, we used it this way and employed our method to do two rounds (one T7dT primed and one T3N9 primed) as well. We amplified both C2 and 3T3 RNAs and labeled 2 µg amplified product as well as 5 µg unamplified C2 and 3T3 RNA. C2 probes from the Arcturus amplification, our amplification, and the unamplified samples were combined with the corresponding 3T3 probes, and
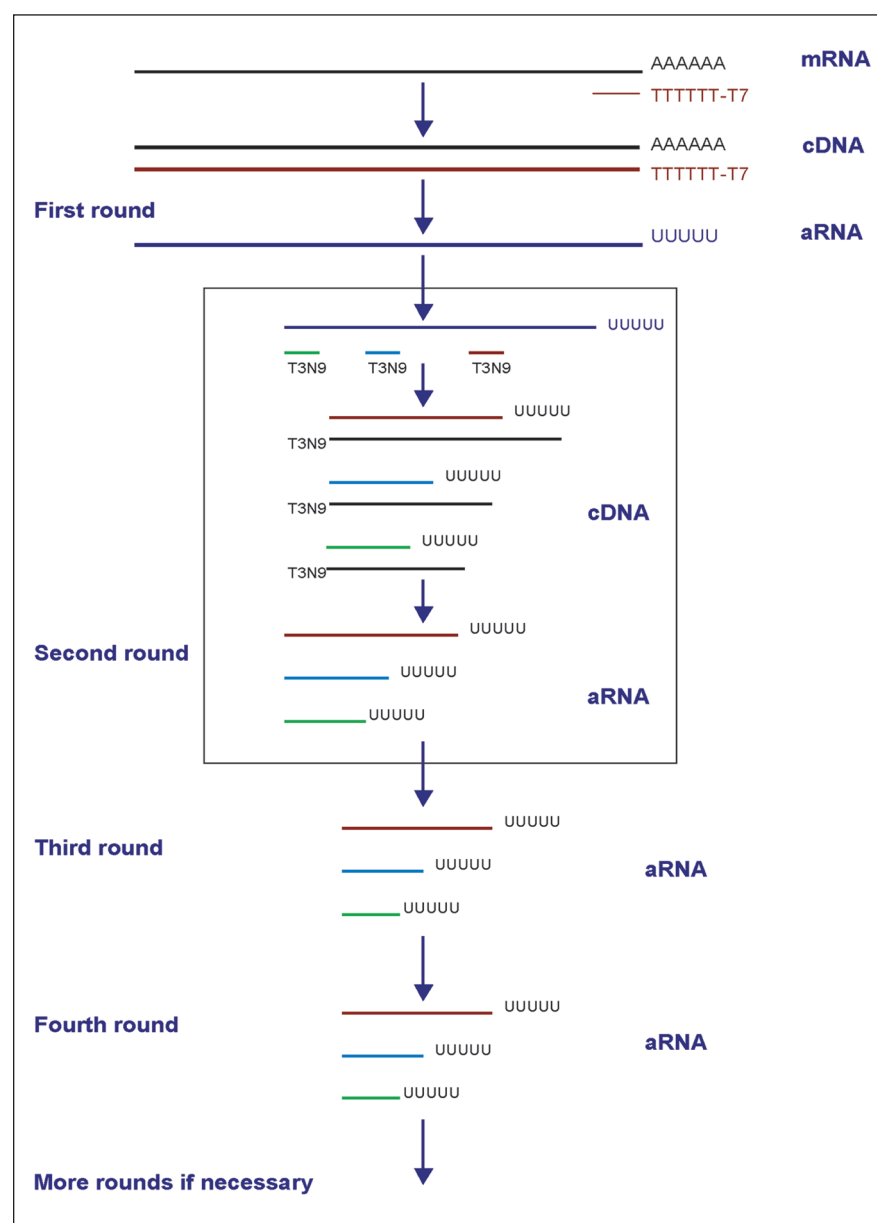

Figure 1. RNA amplification strategy. First-strand synthesis of the cDNA used for the first-round amplification is primed with T7dT. This is followed by second-strand replacement reactions (see Materials and Methods). For the second and subsequent rounds of amplification, reverse transcription of the first strand is primed with T3N9. This step can be repeated several times.

the probe pairs were hybridized to 11 232-element cDNA arrays. After the data were filtered and normalized, 544 genes were identified as "true outliers" based on three experiments with unamplified total RNA (see the "Array Scanning and Data Analysis" section for details). Using probes prepared from RNAs that had undergone two rounds of amplification, we found 394 consistent outliers using our method and 397 with the Arcturus RiboAmp method. In the first round of amplification, our method and the Arcturus method gave similar results—445 and 446 consistent

outliers, respectively. Of these, about 18% were not on the list of "true outliers" discovered when unamplified RNA was used. This result was encouraging, but as mentioned earlier, two rounds of synthesis would surely not permit us to profile small numbers of cells, and we wondered whether we could safely do additional cycles.

To answer this question, we amplified two samples of C2 RNA in parallel with one another a total of six times. At the end of each synthetic cycle, we made Cy5- and Cy3-labeled probes from the products, combined these, and

used them to develop 11 232-element arrays. As shown in Figure 2, analysis of the signals observed revealed excellent correlation coefficients at each consecutive step.

This was not a terribly stringent analysis, however, and we designed a more rigorous one. We amplified both C2 and 3T3 RNAs up to six times using T3N9 to drive first-strand synthesis in each round except the first, when T7dT was employed instead. Probes were made from amplified C2 (Cy3-labeled) and 3T3 (Cy5-labeled) RNAs following each round of amplification. These
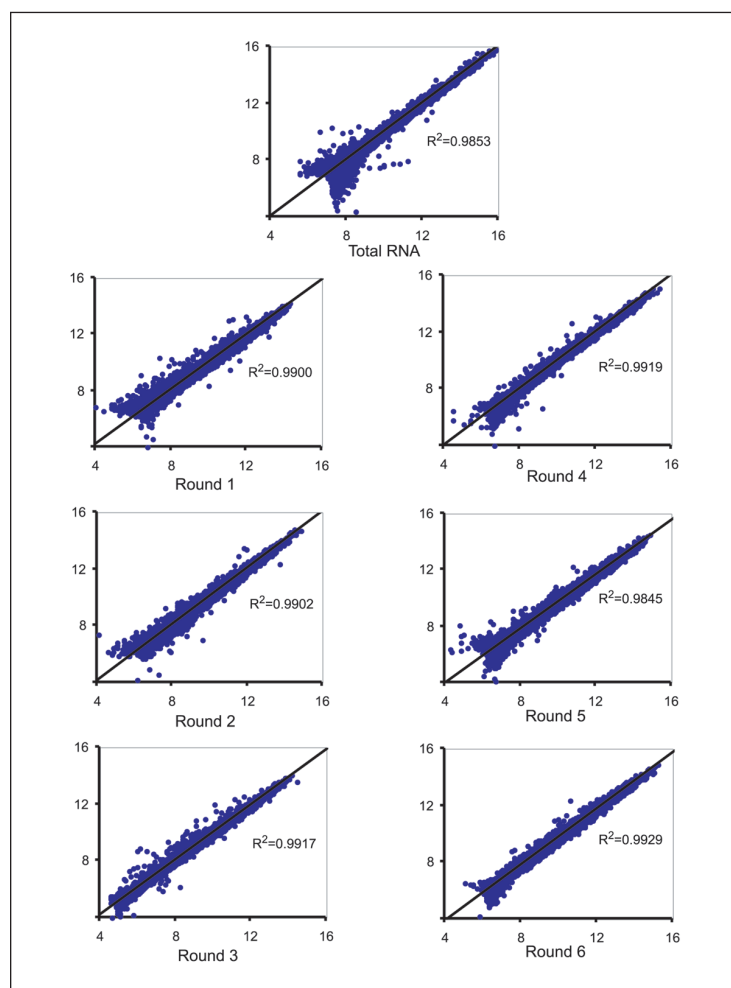

Figure 2. Reproducibility of RNA amplification with T3N9: self versus self experiments. A single RNA sample was divided in half and amplified up to six rounds. Following each round, the samples were used to prepare Cy3- or Cy5-labeled probes, which were combined and applied to an 11 323-element array. DNA used as a template for the first round of amplification was synthesized from 1 µg mouse C2 total RNA. The reverse transcription was primed by T7dT. Subsequent rounds of amplification were primed by T3N9 (see Materials and Methods and Figure 1). Probes made from unamplified total RNA ($R^2 = 0.9853$), and RNA that had been amplified 1, 2, 3, 4, 5, and 6 times gave excellent correlation coefficients—$R^2 = 0.9900$, 0.9902, 0.9917, 0.9919, 0.9845, and 0.9929, respectively.
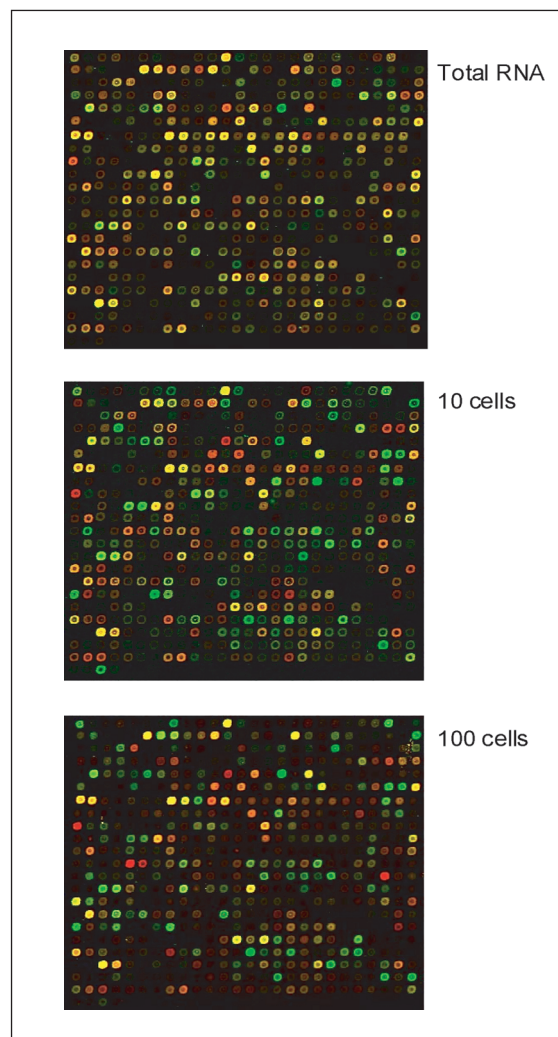

Figure 3. Probes from a small number of cells. Partial images of 11 232-element arrays. Cy5-labeled probe was made from 5 µg mouse 3T3 total RNA used as reference. Amplified and unamplified C2 RNAs were labeled with Cy3. Top, Probes made from 5 µg each mouse C2 and 3T3 total RNA; middle, Cy3 probe made from 10 mouse C2 cells after three rounds of RNA amplification; bottom, Cy3 probe made from 100 mouse C2 cells after three rounds of RNA amplification.

were combined, and 11 232-element arrays were developed with probes from unamplified total RNAs and amplified RNAs. The filtered ratios were ranked as described above. In duplicate experiments with unamplified RNA, 602 "true outliers" were found. When arrays were developed using probes made from RNAs that had been amplified 1 to 6 times, respectively, 513, 512, 498 469, 442, and 404 concordant outliers were found. Thus, after four rounds, we were still able to detect 75% of "true outliers" we had seen with unamplified probes. As many other researchers have observed, a single round of amplification creates some bias—that is, the profiles obtained are quite similar when amplified and unamplified probes are compared, but not identical. In three subsequent amplification steps with T3N9, however, little additional distortion is introduced. More than 91% of the differentially expressed genes found in rounds one and four were concordant. This is not to say that products from rounds 1 and 4 are identical, even though consecutive rounds of amplification give very similar products. To compare two samples, it is best to amplify both the same number of times.

To profile 10 and 100 cells, we prepared DNase-treated lysates from mouse C2 cells, as described in the Materials and Methods section, and performed three rounds of amplification on the RNA in these extracts. We then made probes from the products, developed arrays with them, and compared the results to those obtained with probes made from purified C2 cell total RNA. As shown in Figure 3, the resulting probes gave remarkably similar results.

We conclude that the amplification method we have developed, used in tandem with the labeling technique we described earlier, can be used to make probes from a small number of cells for expression profiling. Refining the technique should allow us to study single cells.

ACKNOWLEDGMENTS

REFERENCES

1. **Baugh, L.R., A.A. Hill, E.L. Brown, and C.P. Hunter.** 2001. Quantitative analysis of mRNA amplification by *in vitro* transcription. Nucleic Acids Res. *29*:e29.
2. **Buchman, G.W., D.M. Schuster, and A. Rashtchian.** 1993. Selective RNA amplification—a novel method using dump-containing primers and uracil DNA glycosylase. PCR Methods Appl. *3*:28-31.
3. **Chen, Y., V. Kamat, E.R. Dougherty, M.L. Bittner, P.S. Meltzer, and J.M. Trent.** 2002. Ratio statistics of gene expression levels and applications to microarray data analysis. Bioinformatics *18*:1207-1215.
4. **DeRisi, J., L. Penland, P.O. Brown, M.L. Bittner, P.S. Meltzer, M. Ray, Y., Chen, Y.A. Su, and J.M. Trent.** 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet. *14*:457-460.
5. **Dixon, A.K., P.J. Richardson, R.D. Pinnock, and K. Lee.** 2000. Gene-expression analysis at the single-cell level. Trends Pharmacol. Sci. *21*:65-70.
6. **Duggan, D., M. Bitter, Y. Chen, P. Meltzer, and J. Trent.** 1999. Expression profiling using cDNA microarrays. Nat. Genet. Suppl. *21*:10-14.
7. **Eisen, M.B. and P.O. Brown.** 1999. DNA arrays for analysis of gene expression. Methods Enzymol. *303*:179-205.
8. **Heller, R.A., M. Schena, A. Chai, D. Shalon, T. Bedilion, J. Gilmore, D.E. Woolley, and R.W. Davis.** 1997. Discovery and analysis of inflammatory disease-related genes using cDNA microarray. Proc. Natl. Acad. Sci. USA *94*:2150-2155.
9. **Khan, J., R. Simon, M. Bittner, Y. Chen, S.B. Leighton, T. Pohida, P.D. Smith, Y. Jiang, et al.** 1998. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res. *58*:5009-5013.
10. **Luo, L., R.C. Salunga, H. Guo, A. Bittner, K.C. Joy, J.E. Galindo, H. Xiao, K.E. Rogers, et al.** 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat. Med. *5*:117-122.
11. **Phillips, J. and J.H. Eberwine.** 1996. Antisense RNA amplification: a linear amplification method for analyzing the mRNA population from single living cells. Methods *10*:283-288.
12. **Puskas, L.G., A. Zvara, L. Hackler, Jr., and P. Van Hummelen.** 2002. RNA amplification results in reproducible microarray data with slight ratio bias. BioTechniques *32*:1330-1340.
13. **Schena, M., D. Shalon, R.W. Davis, and P.O. Brown.** 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*:484-487.
14. **Wang, E., L.D. Miller, G.A. Ohnmacht, E.T. Liu, and F.M. Marincola.** 2000. High-fidelity mRNA amplification for gene profiling. Nat. Biotechnol. *18*:457-459.
15. **Xiang, C.C., O.A. Kozhich, M. Chen, J.M. Inman, Q.N. Phan, and M.J. Brownstein.** 2002. An improved method to label probes for DNA microarray work: amine-modified random primers. Nat. Biotechnol. *20*:738-742.

Address correspondence to:

Dr. Michael J. Brownstein
Laboratory of Genetics
NIMH/NHGRI
National Institutes of Health
Building 36, Room 3D06
Bethesda, MD 20892, USA
e-mail: mike@codon.nih.gov