Development of a new reference standard for microarray experiments

Francesco Gorreta[1], Dagania Barzaghi[1], Amy J. VanMeter[1], Vikas Chandhoke[1], and Luca Del Giacco[1,2]

BioTechniques 36:1002-1009 (June 2004)

Often microarray studies require a reference to indirectly compare the samples under observation. References based on pooled RNA from different cell lines have already been described (here referred to as RNA-R), but they usually do not exhaustively represent the set of genes printed on a chip, thus requiring many adjustments during the analyses. A reference could also be generated in vitro transcribing the collection of cDNA clones printed on the microarray in use (here referred to as T3-R). Here we describe an alternative and simpler PCR-based methodology to construct a similar reference (Chip-R), and we extensively test and compare it to both RNA-R and T3-R. The use of both Chip-R and T3-R dramatically increases the number of signals on the slides and gives more reproducible results than RNA-R. Each reference preparation is also evaluated in a simple microarray experiment comparing two different RNA populations. Our results show that the introduction of a reference always interferes with the analysis. Indeed, the direct comparison is able to identify more up- or down-regulated genes than any reference-mediated analysis. However, if a reference has to be used, Chip-R and T3-R are able to guarantee more reliable results than RNA-R.

INTRODUCTION

The introduction of microarray technology has allowed investigators to analyze and compare the expression of thousands of genes simultaneously. A simple experiment design requires only two samples: a query RNA and its control. Spotted arrays allow the use of two different fluorochromes at the same time, so that both the sample and control can be hybridized and compared on the same slide (1). The simultaneous use of two dyes allows reduction of artifacts due to uneven hybridization, local background differences, and slide-to-slide variations (2–4). On the other hand, not all gene expression studies allow a direct comparison among samples (e.g., lack of control RNA, large number of samples), and the introduction of a common reference is required (5,6). The introduction of a reference in the experiment design allows investigators to compare a large number of samples. Furthermore, it can be used as an internal standard for quality control purposes.

A reference can be created by pooling RNA from the different samples in analysis, or it can be unrelated to them (e.g., commercial RNA from a combination of different cell lines) (5,7–11). Whichever choice is made, a reference should represent as many genes printed on the slide as possible. This is crucial because signals not significantly different from the background, even if only in one of the two channels, are commonly filtered out from the data set. Low or no signal spots, even if quantified, are associated with a high grade of variability (12), which negatively affects the quality of the normalization process and data analysis (7). Some analysis protocols compute adjustments in the determination of the dyes ratio if a gene is not detected in the reference but detected in the sample of interest by substituting the reference intensity with a function of the threshold level on which the filtering step is based (13). A reference able to represent the majority of genes printed on the slide would reduce the number of adjustments, along with the associated error.

Since RNA-based references depend on gene expression, they can hardly be representative of all genes printed on a microarray. To overcome this problem, Kim et al. (7) tried to generate a genomic DNA-based reference in order to represent the entire genome. However, the same authors demonstrated that a reference obtained by pooling different RNA preparations has still more advantages than the genomic one. A third approach is to generate a reference based on the collection of clones printed on the chip. The only protocol in literature to prepare such a reference (referred to here as T3-R) has been described by Sterrenburg et al. (14).

Here we provide a faster, easier, and cheaper PCR-based method for the preparation of a reference based on the cDNA clones printed on the chip (Chip-R). We also extensively tested and compared Chip-R to a commercially available RNA reference (RNA-R) and to T3-R by analyzing: (*i*) the number of spots detected by self versus self-hybridizations and (*ii*) the reproducibility of the hybridizations across replicates. Finally, in order to evaluate the impact of the introduction of a reference on gene expression analyses, we compared two RNA samples directly on and through different references.

[1]George Mason University, Manassas, VA, USA and [2]Dipartimento di Biologia, Università degli Studi di Milano, Milan, Italy

MATERIALS AND METHODS

Array Fabrication

A microarray was constructed from 5376 human cDNA clones (Research Genetics, Huntsville, AL, USA). The complete list of genes with accession numbers is published at http://www.gmu.edu/centers/genomics/research/keys.html. cDNA inserts were amplified directly from clones in culture using GF200F (5′-CTGCAAGGCGATTAAGTTGGGTAAC-3′) and GF200R (5′-GTGAGCGGATAACAATTTCACACAGGAAACAGC-3′) plasmid universal primers. Amplification products were purified using Multiscreen® PCR plates (Millipore, Billerica, MA, USA), dried, and resuspended in 30 µL of 3× standard saline citrate (SSC). The collection of amplified cDNAs was printed on poly-L-lysine-coated slides in a single replicate using GeneMachines™ Model OGR-03 OmniGrid Microarrayer (Genomic Solutions, Ann Arbor, MI, USA) with SMP3 pins (TeleChem International, Sunnyvale, CA, USA). Negative controls consisting of no-template PCR amplifications ($n = 67$) were also printed on the microarray.

Agarose Gel Analysis and POPO-3 Staining

Aliquots of all PCR amplifications were analyzed by agarose gel electrophoresis in order to monitor the yield and specificity of the reactions. POPO™-3 (Molecular Probes, Eugene, OR, USA) DNA staining was also performed in triplicate to monitor the number of spots actually containing PCR products. Briefly, each slide was incubated at room temperature for 4 min in a 10,000-fold diluted stain solution. The microarray was then washed twice with 1× Tris-EDTA (TE) (for 1 and 3 min, respectively). The slides were centrifuged dry at 85× *g* for 3 min at room temperature and scanned in the Cy™3 channel using ScanArray® Express HT (PerkinElmer Life and Analytical Sciences, Boston, MA, USA). The intensity of a spot was considered significant when higher than the median local background plus two standard deviations calculated in each subarray. A total of 162 clones (3% of the total) failed either the amplification or printing procedure and were subtracted from the analysis.

Chip-R and T3-R Preparation

All clones selected for the custom microarray were pooled and replicated in a single 96-well plate containing 200 µL/well of LB broth in the presence of a selective antibiotic. After overnight growth, 50 µL from each well were harvested and pooled. Plasmids were extracted from the pool using the QIAprep® Spin Miniprep kit (Qiagen, Valencia, CA, USA), and inserts were amplified with GF200F and GF200R universal primers. To generate Chip-R, PCR was performed in the presence of a mixture of dNTPs containing 25 mM dATP, dCTP, and dGTP, 15 mM dTTP (Fisher Scientific International, Hampton, NH, USA), and 10 mM aminoallyl-dUTP (aa-dUTP; Sigma, St. Louis, MO, USA) for successive labeling. An aliquot of the reaction was monitored on agarose gel. The reaction mixture was cleaned using QIAquick® PCR Purification kit (Qiagen) before the final coupling reaction with Cy dyes (Amersham Biosciences, Piscataway, NJ, USA).

To generate a single specific probe, PCR was performed on one clone using the same conditions used for Chip-R. Half of the reaction mixture was digested at both ends with *Eco*RI and *Not*I. Before the labeling reaction, the probe was purified from agarose gel.

To generate T3-R, the pool of PCR products was transcribed in vitro using the MAXIscript® T3 Kit (Ambion, Austin, TX, USA) according to the manufacturer's recommendations. Because our library is composed of different cloning vectors, we took into account that 1.5% of the clones did not contain the T3 promoter sequence, and we excluded those clones from the analysis. After DNase treatment, the RNA was cleaned using Microcon® YM-30 Centrifugal Filter Devices (Millipore) and then reverse-transcribed in the presence of aa-dUTP.

RNA-R Preparation

We used the commercially available pool of total RNA from 10 different human cell lines (Universal Human Reference RNA; Stratagene, La Jolla, CA, USA) for the RNA-based reference. RNA was then amplified using the MessageAmp™ aRNA Kit (Ambion) according to the manufacturer's recommendations. The aRNA quality was also monitored by electrophoresis on agarose gels, and the average size was evaluated with the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA).

Sample Preparation

An aliquot of 1.5 µg of total RNA extracted from the T24 cell line was first DNase treated (DNA-*free*™; Ambion), then amplified using the MessageAmp aRNA Kit. The amplified RNA (aRNA) was quantified, and its quality was monitored by agarose gel analysis and the Agilent 2100 Bioanalyzer.

Sample Labeling

An aliquot of 4 µg aRNA from both the reference RNA-R and T24 cell line was reverse-transcribed and labeled according to The Institute for Genomic Research (TIGR) protocol (http://www.tigr.org/tdb/microarray/protocolsTIGR.shtml). Briefly, aRNA was heated at 70°C for 10 min in a total volume of 18.5 µL in the presence of 6 µg of random hexamer primers (Invitrogen, Carlsbad, CA, USA). The reaction was cooled on ice for 1 min, and then 6 µL 5× first strand buffer (Invitrogen), 3 µL 0.1 M dithiothreitol (DTT), 0.6 µL 50× dNTPs mixture containing aa-dUTP, and 400 U SuperScript™ II reverse transcriptase (Invitrogen) were added to the reaction. The mixture was incubated at 42°C overnight. After RNA hydrolysis, the cDNA was purified using Microcon YM-30 Centrifugal Filter Devices according to the manufacturer's instructions. Aminoallyl-labeled samples from cell lines, RNA-R, Chip-R, and T3-R were coupled with the Cy dye esters according to TIGR's protocol.

Prehybridization

Spotted cDNA microarrays were first rehydrated in a humidity chamber, denatured at 95°C for 4 s, and then cDNA was cross-linked by UV to the slide surface.

Slides were incubated at 45°C for 45 min in prehybridization buffer containing 5× SSC, 0.1% sodium dodecyl sulfate (SDS), and 1% bovine serum albumin (BSA; Sigma). After a wash in Milli-Q® water (Millipore), the slides were dipped in isopropanol and air-dried.

Hybridization

The labeled probes were vacuum-dried and resuspended in the hybridization buffer (25% formamide, 5× SSC, 0.1% SDS). Samples were denatured at 95°C for 3 min and applied to a prehybridized microarray slide. The microarray slide was incubated at 45°C overnight in a sealed hybridization chamber (CMT™ hybridization chamber; Corning, Corning, NY, USA). Slides were washed twice in 1× SSC, 0.2% SDS (10 min at 45°C), twice in 0.1× SSC, 0.1% SDS (10 min at 45°C), twice in 0.1× SSC (10 min at 45°C), rinsed in Milli-Q water, and dried by brief centrifugation.

Image Acquisition and Image Processing

All hybridized slides were scanned by confocal laser scanner ScanArray Express HT at 75% of photomultiplier tube (PMT), 75% laser power (LP), and 10 µm of pixel resolution. Images were acquired by ScanArray Express 2.0 software (GSI Lumonics, Billerica, MA, USA) and processed with QuantArray® 3.0 software (PerkinElmer Life and Analytical Sciences) in order to measure the intensity and the local background for both Cy3 and Cy5 channels of each spot.

Data Analysis

Local background value, as provided by QuantArray software, was subtracted from the intensity level of each gene in each channel. A filtering step was performed in each channel separately, using thresholds equal to $\mu_{NC} + 1\sigma_{NC}$ (with μ_{NC} and σ_{NC} denoting the mean and standard deviation of the background-corrected intensity levels of the negative controls printed on the slide). Genes that passed the filtering step in both channels were selected for further analyses and normalization. When genes passed the filtering step in only one channel, we substituted the intensity of the filtered channel with the value of μ_{NC} for that channel as described in Epstein et al. (13). Genes that did not pass the threshold in both channels (undetected genes) were excluded from further analyses. The Cy5/Cy3 ratio was then normalized by its median.

For indirect comparison of the T24 control (C) and T24 treated (T) samples through the use of references, we calculated the ratio for each gene as described in Kim et al. (7). Briefly, the ratio for each gene was calculated after per-chip normalization as follows:

$$T/C = (T/R^1)_{slide1} \times (R^2/C)_{slide2}$$

where R^1 and R^2 are the intensity values for the reference in the two slides.

Statistical Analyses

Statistical tests and graphs were generated using Prism version 3.03 software (GraphPad Sofware, San Diego, CA, USA).

Real-Time Reverse Transcription PCR

Reverse transcription reactions were performed using 2 µg DNase-treated total RNA (DNA-*free*) from both T24 (C) and T24 (T). Reactions were heated at 70°C for 5 min in a total volume of 12.5 µL in the presence of 100 ng of random hexamers and then cooled at room temperature for 1 min. After a 30-s centrifugation at 10,000× *g*, 4 µL 5× first strand buffer, 2 µL 0.1 M DTT, 0.5 µL 25 mM dNTPs mixture, and 200 U of SuperScript II reverse transcriptase were added to the reaction. The mixture was then incubated at 42°C for 2 h. The resulting cDNA was stored frozen (-80°C) until assayed by real-time PCR. Real-time PCR mixtures contained 1 µL of the reverse transcribed sample, 200 nM of each forward and reverse primers, and 1× iQ™ SYBR® Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA). Reactions were carried out in a total volume of 15 µL and were performed in a 96-well format in the iCycler™ iQ Real-Time Detection System (Bio-Rad Laboratories). Amplification of 18S rRNA was performed in parallel with the genes of interest as an internal standard (15). For both the gene of interest and 18S, two independent PCR were performed from the same reverse transcribed sample.

RESULTS

Chip-R and T3-R Preparation

All 5376 human cDNA clones printed on the chip were inoculated and pooled in a single 96-well plate as shown in Figure 1. After overnight growth, all 96 wells were pooled, and a single plasmid DNA extraction was performed. For the production of Chip-R, the pool of plasmids was used as template for PCR using the GF200F and GF200R library universal primers in presence of aa-dNTPs. The resulting products were labeled with Cy dyes. For the production of T3-R, as described in Sterrenburg et al. (14), the pool of PCR product was transcribed in vitro into RNA using T3 RNA polymerase. After DNase treatment, the sample was cleaned up and reverse-transcribed in the presence of aa-dUTP for the labeling (Figure 1). Part of the original cDNA clones (1.5%) could not be transcribed in vitro because the vector lacked the T3 promoter sequence. These clones have been excluded from further analyses.

Chip-R, T3-R, and RNA-R Spot Detection and Self versus Self Test

To evaluate the ability of each reference to represent the set of printed genes and to investigate the reproducibility of the hybridizations, we performed three independent labeling reactions for each dye, followed by a self versus self test. Figure 2 shows the number of spots detected in each channel using 4 µg of each reference preparation. No statistical differences were exhibited between Chip-R and T3-R. In fact, out of 5214 cDNA spots analyzed, Chip-R and T3-R were able to hybridize an average of 97% of the total number of spots, while RNA-R was only able to hybridize 52%. In order to evaluate the level of sensitivity, different starting amounts of Chip-R have been used, and comparable results have been obtained using 4, 2, 1, or 0.5 µg of

Table 1. Self versus Self Hybridizations of Chip-R, T3-R, and RNA-R References

	Distribution's Standard Deviation		
	Chip-R	**T3-R**	**RNA-R**
Slide 1	0.24	0.20	0.27
Slide 2	0.20	0.18	0.28
Slide 3	0.20	0.20	0.29
Triplicates	**0.19**	**0.17**	**0.25**

Standard deviations of normalized ratios distributions. Results from a single slide experiment and from the triplicates.

PCR product (data not shown).

In the self versus self test, each gene is expected to have a normalized Cy5/Cy3 ratio equal or close to one. Table 1 summarizes the standard deviation of the different distributions. Again, we obtained comparable results between Chip-R and T3-R, which exhibited a narrower distribution in respect to RNA-R with a lower number of genes in the tails of the distribution (0.02, 0.06, and 0.4% of the detected genes had a ratio greater than 2 or lower than 0.5 for Chip-R, T3-R, and RNA-R, respectively). Moreover, the percentage relative error of the normalized intensity of each detected spot across the triplicates was measured for each reference in each channel separately. T3-R and Chip-R exhibited similar distributions, while RNA-R showed higher variability (Figure 3).

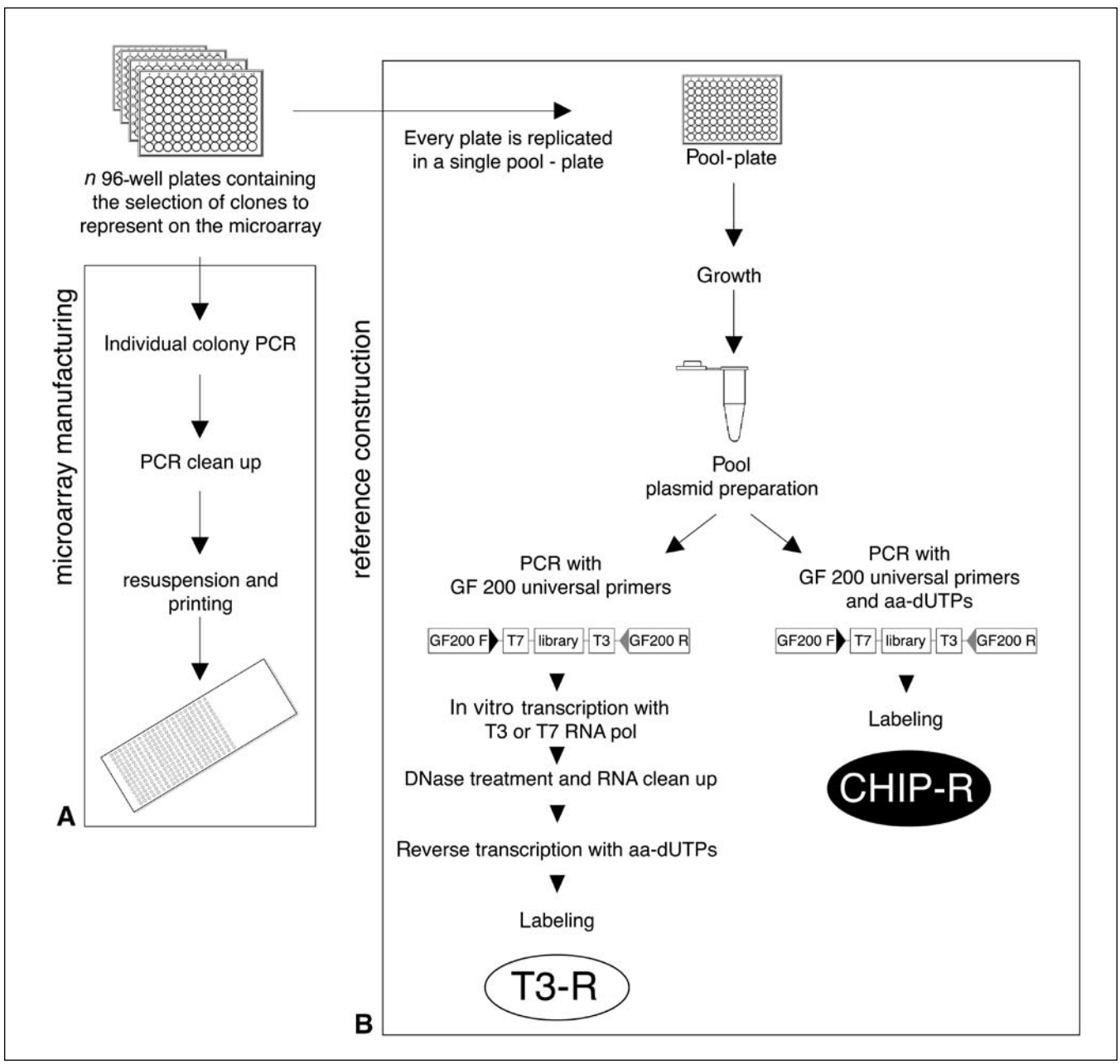

Figure 1. Production scheme of clone-based references. (A) Basic scheme of cDNA microarray manufacturing: cDNA inserts are amplified by PCR, cleaned up, and printed on slides. (B) Parallel reference production: T3-R and Chip-R. aa-dUTP, aminoallyl-dUTP.

Single-Probe Preparation and Cross-Hybridization

One potential disadvantage of using Chip-R and T3-R is that both the reference and the cDNA spotted contain part of the vector's polylinker, and this might lead to cross-hybridization. In order to determine the impact of this phenomenon, we performed three hybridization experiments using a single gene probe amplified and labeled using the same protocol as described for Chip-R. This specific probe contained 270 bp belonging to the vector. As a result, nonspecific spots were detected on each slide (1 ± 1% of the total), but only the specific clone was simultaneously identified in all three replicates. To better determine the influence of the polylinker nucleotides, the specific probe was digested at both ends in order to remove the plasmid sequence before hybridization. Similar results were obtained. Only the correct spot was detected simultaneously in three replicates, and Student's *t*-test analysis showed no significant difference in the number of spots detected on each slide (0.9 ± 0.3% for the fragment only) or in the average intensity of those spots ($P < 0.05$).

References Evaluation and Real-Time Reverse Transcription PCR Validation

The objective of our study was to evaluate how the introduction of a reference in a microarray gene expression analysis may affect the final outcome and determine which reference guarantees results closer to a direct comparison. T24 cells were treated with an anti-tumorigenic compound for 24 h (T) and compared to untreated cells (C). Total RNA was extracted from T and C samples, labeled, and directly compared by co-hybridization in triplicate. A list of genes up- and down-regulated by the treatment was generated (more than 2.5-fold, $P < 0.05$). In parallel, both T and C samples were co-hybridized with Chip-R, T3-R, and RNA-R. Each experiment was performed three independent times in order to reduce the variations due to experimental artifacts. Indirect comparison of T and C through the use of each reference was performed as described by Kim et al. (7), and lists of up- and down-regulated genes for each reference were generated. The resulting lists were then compared (Figure 4).

Table 2. Real-Time PCR Confirmations for Three Genes Present in More than One Microarray Generated List

	Microarray Analysis				
	Direct Comparison	Using Chip-R	Using T3-R	Using RNA-R	Real-Time PCR
Gene 1	0.17	0.27	0.33	0.39	0.09
Gene 2	0.27	0.40	0.37	0.31	0.41
Gene 3	4.1	N.S.	N.S.	4.6	12.5

T24 cells (T) to untreated cells (C) ratio calculated using microarray or real-time data.
N.S., not significant.

Real-time PCR validation has been performed on three randomly selected genes present in more than one list. Table 2 reports the T/C ratio resulting from the different approaches. In order to identify potential false positives, we focused on genes present only in one of the lists. In the direct comparison, only 3 out of 16 genes were not in common with any other list, and only one was identified by real-time analysis as being a false positive. Five genes were tested from each list obtained using Chip-R, T3-R, and RNA-R. All were pinpointed as false positives in both Chip-R and RNA-R, while only two were false positives in the T3-R experiment (Figure 4).

Double Strand (Chip-R) and Single Strand (T3-R) References

We further investigated if the intro-

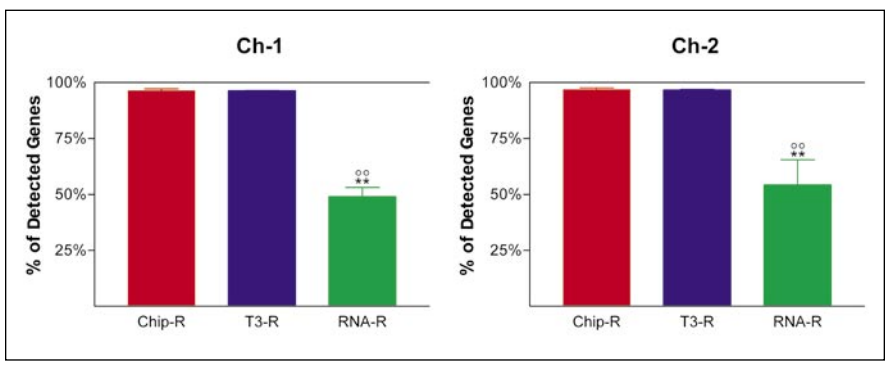

Figure 2. Percentage of detectable genes for each reference. Average of genes detected after three self versus self tests for Chip-R, T3-R, and RNA-R. One-way analysis of variance (ANOVA) with Bonferroni correction was performed to compare the three groups. Error bars represent standard deviations. A total of 5214 genes were analyzed on each microarray. **, $P < 0.01$ versus Chip-R; °°, $P < 0.01$ versus T3-R.

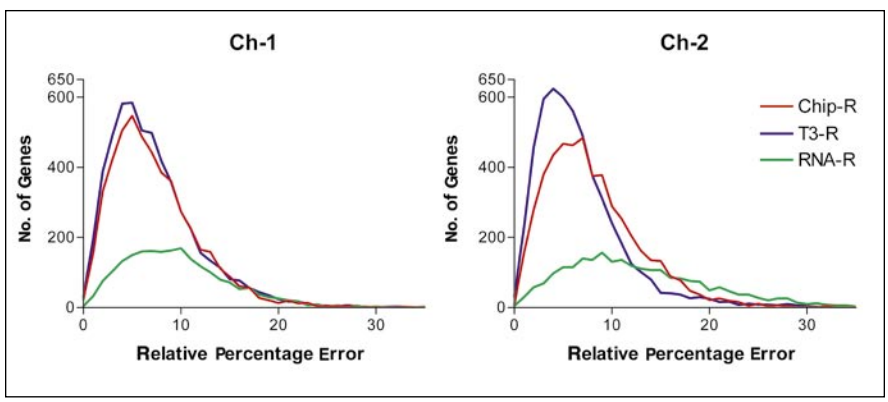

Figure 3. Evaluation of reference variability. Frequency distributions of the percentage relative error across triplicates for the different references in each channel.

duction of a double strand DNA reference, such as Chip-R, might affect gene expression analysis. In this study, the T24 C sample was hybridized on the cDNA chip with either the T24 T sample, RNA-R, single-stranded T3-R, or with the double-stranded Chip-R. We investigated if T24 control-normalized intensities were affected by the different nature of the references used. After performing per channel median intensity normalization, we calculated the average intensity and standard deviation of the T24 C sample in each gene across replicates for each set of hybridizations. A Student's *t*-test was performed to evaluate if the T24 control sample values obtained after hybridization with Chip-R differed from the values obtained after hybridizing with T24 treated samples. Only 0.6% of the genes were significantly different ($P < 0.01$). The same results were obtained after hybridization with T3-R (0.6% $P < 0.01$), while a slight increase of variation was shown after hybridization with RNA-R (1.2% $P < 0.01$). These observations suggest that the use of a double strand reference does not introduce additional variation to the analysis.

DISCUSSION

Two-color microarray technology allows for direct comparison of a sample to its control on the same slide. However, some experimental designs lack sufficient amounts of control RNA to hybridize all samples. In these circumstances, it is possible to compare all samples through a common reference.

Several approaches to generate a reference have been described. Efforts have been made to synthesize a universal genomic reference with the aim to represent all genes printed on the chip independently by their expression level (7). Other approaches are RNA-based, and the reference can be either a pool of RNAs from different samples or a commercial RNA preparation (e.g., the RNA-R in this study) (5,7–11). An alternative strategy is to generate a reference based on the clones spotted on the array (14).

Here we developed a new method to generate a reference (Chip-R) specifically designed on our custom cDNA microarray. Even if library-based, we believe that both T3-R and Chip-R can be potentially utilized for oligonucleotide arrays. The preparation of Chip-R can be performed during microarray manufacturing, and because this procedure is PCR-based, it allows the production of large amounts of reference with little expense. Once the clones have been pooled, it is possible to scale-up the production of Chip-R by performing PCR in 96-well plates and using purification columns for large volumes. With this strategy, it is possible to obtain enough material for hundreds of hybridizations. Compared to the T3-R method, Chip-R does not require additional primer design, in vitro RNA transcription, DNase treatment, and reverse transcription steps before the labeling reaction (14).

The main difference between the two preparations is that T3-R results in a single strand reference, while Chip-R is a double strand. The major concern in designing Chip-R was that the two strands of DNA from the reference could potentially hybridize not only to the spot, but also reanneal to themselves or hybridize to the query, thereby sequestering the sample. Our data however, showed no differences in the final

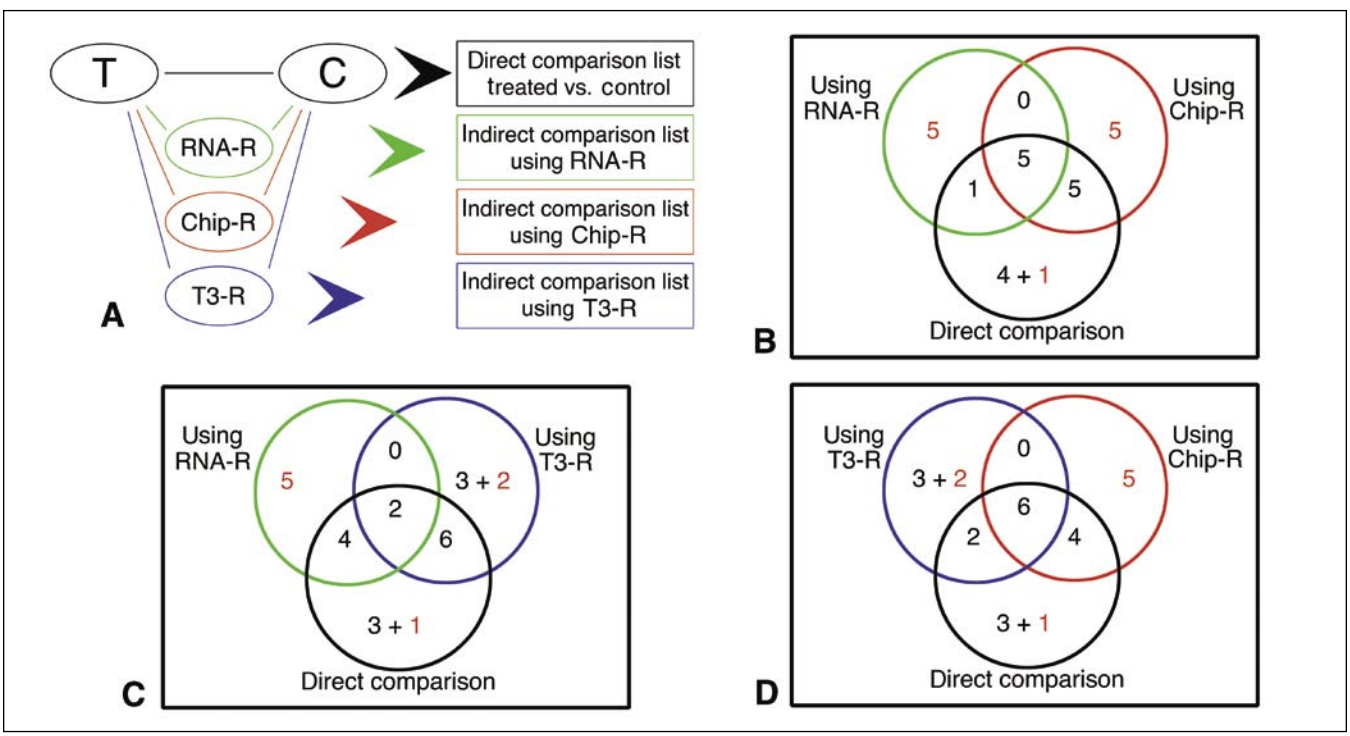

Figure 4. Hybridizations scheme and comparison among lists. (A) Representation lists generation for the direct and indirect comparison of two biological samples. (B–D) Comparison among lists of up- or down-regulated genes obtained by the direct comparison of two RNA populations and by their indirect comparison through different references. Red, genes identified as false positives by real-time PCR.

results obtained using either T3-R or Chip-R. A possible explanation is that the amount of double-stranded DNA spotted on the slide is extremely in excess, and the kinetics of the reaction favors the binding of the Chip-R to the spot rather than to the sample probe.

The use of either Chip-R or T3-R significantly increased the number of genes detected on the slides when compared to RNA-R. This is important because signals not significantly different from the background are commonly filtered out from the data set or are associated with high variability (12). An increase in the number of detected genes in the reference permits a better quantification because less computational adjustments are needed in the analysis protocol to determine the sample to reference ratio (13).

The self versus self test showed that Chip-R and T3-R had a lower percentage of genes in the tail of the distribution (ratio > 2 or < 0.5) than RNA-R. Chip-R and T3-R also exhibited less variability across replicates than RNA-R. This is an essential characteristic for a reference, because the indirect comparison of samples through it assumes that its values are constant across slides. The advantage of Chip-R and T3-R references is to provide only specific probes corresponding to the sequences printed on the chip. Furthermore, the advantage of Chip-R is to provide probes with the same exact length of the DNA spotted on the chip, potentially increasing the specificity of the annealing and reducing cross-hybridization signals. In fact, the previously reported T3-R reference preparation was performed reverse transcribing the RNA using random primers, with consequent reduction of the average size of the molecules (14). We also show that the presence of plasmid sequences in the reference does not lead to significant cross-hybridization.

Finally, we evaluated how the introduction of a reference affects gene expression analysis between different RNA populations and which reference showed results closer to the direct comparison. Our results showed that, in respect to the direct comparison, the introduction of a reference always decreases the number of identified up- and down-regulated genes. If possible, a direct comparison is always the best experiment design. In fact, it is able not only to identify more misregulated genes but also generates a lower number of false positives. When a reference was introduced, T3-R and Chip-R were able to identify more genes than RNA-R, with a lower percentage of false positives. In respect to T3-R, Chip-R was able to identify a comparable number of genes, but presented a higher percentage of false positives (33% for Chip-R and 21% for T3-R).

All of our data suggest that a reference specifically designed on the cDNA microarray in use is the better choice when the experiment design does not allow a direct comparison between all samples. Moreover, Chip-R is quick, inexpensive to prepare, and guarantees results comparable to the only other clone-based reference available.

ACKNOWLEDGMENTS

This work was funded under Commonwealth Technology Research Fund grant no. SE2002-02. The authors are grateful to Dr. Debra L. Bemis (Columbia University, NY) for providing RNA and for help in editing the manuscript.

REFERENCES

1. **Weil, M.R., T. Macatee, and H.R. Garner.** 2002. Toward a universal standard: comparing two methods for standardizing spotted microarray data. BioTechniques *32*:1310-1314.
2. **Dudley, A.M., J. Aach, M.A. Steffen, and G.M. Church.** 2002. Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. Proc. Natl. Acad. Sci. USA *99*:7554-7559.
3. **Colantuoni, C., G. Henry, S. Zeger, and J. Pevsner.** *2002.* Local mean normalization of microarray element signal intensities across an array surface: quality control and correction of spatially systematic artifacts. BioTechniques *32*:1316-1320.
4. **Martinez, M.J., A.D. Aragon, A.L. Rodriguez, J.M. Weber, J.A. Timlin, M.B. Sinclair, D.M. Haaland, and M. Werner-Washburne.** 2003. Identification and removal of contaminating fluorescence from commercial and in-house printed DNA microarrays. Nucleic Acids Res. *31*:e18.
5. **Alizadeh, A., M. Eisen, R.E. Davis, C. Ma, H. Sabet, T. Tran, J.I. Powell, L. Yang, et al.** 1999. The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. Cold Spring Harb. Symp. Quant. Biol. *64*:71-78.
6. **Eisen, M.B. and P.O. Brown.** 1999. DNA arrays for analysis of gene expression. Methods Enzymol. *303*:179-205.
7. **Kim, H., B. Zhao, E.C. Snesrud, B.J. Haas, C.D. Town, and J. Quackenbush.** 2002. Use of RNA and genomic DNA references for inferred comparisons in DNA microarray analyses. BioTechniques *33*:924-930.
8. **Alizadeh, A.A., M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, et al.** 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*:503-511.
9. **Scherf, U., D.T. Ross, M. Waltham, L.H. Smith, J.K. Lee, L. Tanabe, K.W. Kohn, W.C. Reinhold, et al.** 2000. A gene expression database for the molecular pharmacology of cancer. Nat. Genet. *24*:236-244.
10. **Ross, D.T., U. Scherf, M.B. Eisen, C.M. Perou, C. Rees, P. Spellman, V. Iyer, S.S. Jeffrey, et al.** 2000. Systematic variation in gene expression patterns in human cancer cell lines. Nat. Genet. *24*:227-235.
11. **Yang, I.V., E. Chen, J.P. Hasseman, W. Liang, B.C. Frank, S. Wang, V. Sharov, A.I. Saeed, et al.** 2002. Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biol. *3*:62.1-62.13.
12. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat. Genet. *32(Suppl)*:496-501.
13. **Epstein, C.B., W. Hale IV, and R.A. Butow.** 2001. Numerical methods for handling uncertainty in microarray data: an example analyzing perturbed mitochondrial function in yeast. Methods Cell Biol. *65*:439-452.
14. **Sterrenburg, E., R. Turk, J.M. Boer, G.B. van Ommen, and J.T. den Dunnen.** 2002. A common reference for cDNA microarray hybridizations. Nucleic Acids Res. *30*:e116.
15. **Grace, M.B., C.B. McLeland, S.J. Gagliardi, J.M. Smith, W.E Jackson III, and W.F. Blakely.** 2003. Development and assessment of a quantitative reverse transcription-PCR assay for simultaneous measurement of four amplicons. Clin. Chem. *49*:1467-1475.

Received 16 December 2003; accepted 12 March 2004.

Address correspondence to:

Francesco Gorreta
George Mason University
10900 University Blvd
MS4D7 RM 303W
Manassas, VA 20110-2201, USA
e-mail: fgorret1@gmu.edu