Journal of Cellular Biochemistry 103:556–563 (2008)

Comparison of RNA Amplification Methods and Chip Platforms for Microarray Analysis of Samples Processed by Laser Capture Microdissection

Elena Caretti,[1] Karthik Devarajan,[2] Renata Coudry,[3] Eric Ross,[2] Margie L. Clapper,[3] Harry S. Cooper,[4] and Alfonso Bellacosa[1,5]*

[1]Human Genetics Program, Fox Chase Cancer Center, Philadelphia, Pennsylvania
[2]Department of Biostatistics, Fox Chase Cancer Center, Philadelphia, Pennsylvania
[3]Cancer Prevention and Control Program, Fox Chase Cancer Center, Philadelphia, Pennsylvania
[4]Department of Pathology, Fox Chase Cancer Center, Philadelphia, Pennsylvania
[5]Laboratory of Developmental Therapeutics, Institute Regina Elena, Rome, Italy

Abstract Laser capture microdissection (LCM) permits isolation of pure cell populations from which RNA can be extracted, amplified, and subjected to microarray analysis, allowing information to be obtained on the gene expression profile of defined cell types. To avoid amplification artifacts and detect genes expressed at different levels, it is important to optimize the choice of both RNA amplification step and microarray platform. We captured by LCM the same colon cancer biopsy and conducted a cross comparison of distinct RNA amplification methods and different chip platforms. We tested two RNA amplification methods with different chemistry: the one-cycle Ovation™ system (NuGEN) and the two-cycle Ribo OA™ method (Arcturus). We also compared two different whole genome platforms, based on Affymetrix technology: the U133 plus 2.0 and the X3P array, with probe sets closer to the 3′ end of transcripts. After RNA amplification, microarray analysis, and data normalization, we investigated reproducibility and correlation of different methods and arrays. Our results indicate that the Arcturus Ribo OA method is superior for both array choices, especially in combination with X3P arrays, showing the lowest variance and Spearman correlation of 0.986. The quicker NuGEN procedure, when coupled with X3P arrays, also yielded excellent results (correlation of 0.951). These observations will be useful for planning large-scale analyses of LCM-dissected clinical samples. J. Cell. Biochem. 103: 556–563, 2008.

Key words: laser capture microdissection; RNA amplification; microarray analysis

Microarray technology can provide large amounts of information about cellular gene expression for both normal and diseased states. However, cellular heterogeneity of the tissues very often presents a challenge to gene expression profiling of specific cell types. For example, neoplastic tissue samples are usually contaminated with surrounding stroma cell types or infiltrating lymphocytes (as a result of the immune response in vivo). The heterogeneity in tissues, which physiologically contain different proportions of many cell types, may contaminate the gene expression measurements, and in turn significantly confound statistical analysis. Thus, sample heterogeneity can result in the identification of differentially expressed genes that may be unrelated to the cell type being studied or in the identification of irrelevant genetic changes [Lahdesmaki et al., 2005].

Laser capture microdissection (LCM) is an emergent technique that can be used to obtain purified cell populations from the most heterogeneous tissue, and therefore derive precise information on the gene expression profile of defined cell types. By using LCM it is possible to identify normal or pathologic cells of interest, which are subsequently captured using a laser-based technology. This approach has been very

Grant sponsor: National Cancer Institute; Contract Grant numbers: N01 CN-95037, CA-06927.

*Correspondence to: Dr. Alfonso Bellacosa, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111.
E-mail: alfonso.bellacosa@fccc.edu

Received 28 March 2007; Accepted 24 April 2007
DOI 10.1002/jcb.21426

successful in DNA based studies, because of the high stability of DNA. However, microarray studies require RNA, which is less stable than DNA and it is much more challenging to avoid degradation during the dissecting process [Wang et al., 2006].

Intrinsic characteristics of LCM allow isolation of only small amounts of total RNA (generally a few nanograms for each sample), thus rendering necessary an RNA amplification step prior to microarray analysis [Upson et al., 2004; Luzzi et al., 2005; McClain et al., 2005; Schindler et al., 2005]. In fact, higher quantities are generally required to perform hybridization on arrays, starting from a few micrograms to many micrograms, depending on protocols and arrays.

It is therefore important to optimize the choice of both the RNA amplification step and microarray platform, in order to detect low-expression genes and avoid artifacts. To date, several groups have conducted detailed comparisons of different RNA amplification methods [Goff et al., 2004; Upson et al., 2004]. However, there are no data presently in the literature on the cross comparison between different methods for sample amplification combined with different microarray platforms, leaving the choice of the optimal combination of amplification protocol and array type an open problem.

In this article, we processed replicate samples laser-captured from the same colon adenocarcinoma biopsy with two RNA amplification methods characterized by completely different biochemistry, the one-cycle Ovation™ biotin system (NuGEN), and the two-cycle Ribo OA™ amplification kit (Arcturus). RNA amplified with the NuGEN and Arcturus amplification systems was hybridized to two different array platforms covering the whole genome, both based on Affymetrix technology: the U133 plus 2.0 and the new generation array X3P, which has been designed with probe sets closer to the 3′ end (within 300 bases) of the transcripts (the purpose of the new array is to improve the gene expression profiling of partially degraded RNAs, which are very often obtained after LCM, for example, from paraffin-embedded samples).

For the replicates of different combinations of amplification protocol and array type, we obtained measurements of reproducibility and correlation. This led us to formulate a recommendation on the optimal combination of amplification method and microarray platform for analysis of LCM-derived RNAs. These observations will be useful for planning future studies of LCM-dissected clinical samples.

METHODS

LCM, RNA Extraction, and Amplification

A colon cancer biopsy collected during surgery conducted at Fox Chase Cancer Center was snap-frozen and subsequently processed for cryo-sections. Sections were placed on slides, fixed, and stained with hematoxilin and eosin according to standard protocols. A total of 120 crypts, corresponding approximately to 1,200 cells [Upson et al., 2004], were collected for each section using LCM (Fig. 1). Cells were incubated 1 h at 42°C in 50 μl of extraction buffer, then used for RNA extraction according to PicoPure (Arcturus) RNA isolation kit procedures. From each dissection of 120 crypts, approximately 1 ng of total RNA was obtained, in agreement with previous studies [Upson et al., 2004] and immediately treated with RNAse-in (Arcturus).

RNA samples were prepared from several dissections and checked for quality on the Agilent 2100 Bioanalyzer. Eight samples that showed presence of distinct peaks corresponding to intact 28S and 18S ribosomal RNAs were included in the analysis, while the others were discarded (Fig. 2). In order to ensure exactly the same starting conditions prior to processing for amplification and hybridization onto arrays, the eight RNA samples were pooled together and then aliquoted. To render our analysis statistically significant, we performed four replicates with NuGEN Ovation™ biotin system and four with Arcturus Ribo OA™.

With the NuGEN protocol, RNAs were reverse transcribed to cDNAs, then they were amplified during the so-called SPIA amplification, a linear isothermal DNA amplification process [Dafforn et al., 2004], and finally amplified products, consisting of single-strand DNA (ssDNA), were biotin labeled and fragmented according to the manufacturer's guidelines. Yields of amplified products for the four samples processed with the NuGEN protocol were: 4.7, 4.5, 4.3, and 5.3 μg.

Four samples were processed with the Arcturus Ribo OA™, a T7 RNA polymerase based technology, following the manufacturer's recommendations. Briefly, after two rounds of

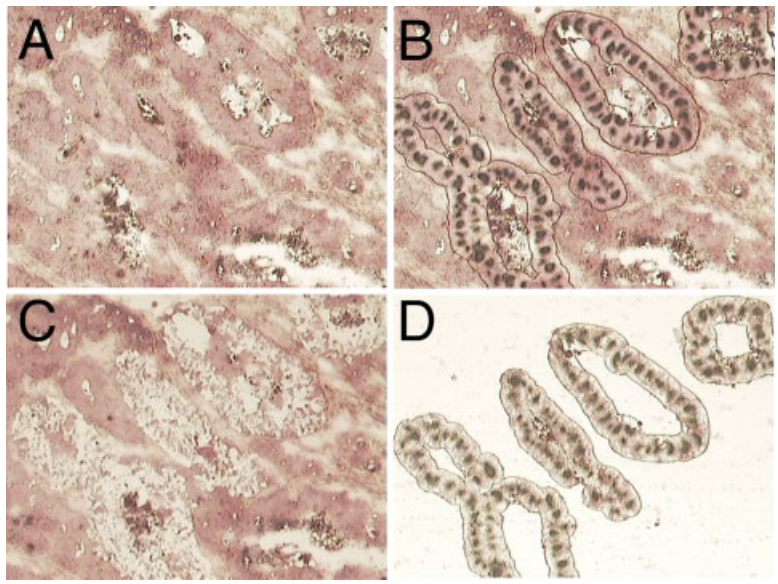

Fig. 1. A colon adenocarcinoma cryo-section before (**A**) and after (**B**–**D**) laser capture. Captured cancer cells are in (D).

reverse transcription separated by one cycle of RNA amplification by T7 RNA polymerase-based in vitro transcription [Phillips and Eberwine, 1996; Wang et al., 2000; Feldman et al., 2002], a second in vitro transcription, that generates labeled amplified RNA (aRNA), was conducted with Affymetrix reagents from the one-cycle kit. Yields of amplified products for the four samples processed with the Arcturus Ribo OA^{TM} protocol were: 98, 90, 89.2, and 104.9 μg.

Multiple intermediate steps, from the starting total RNA to the final hybridization product, were checked for quality on the Agilent 2100 Bioanalyzer, including ssDNA after cleanup and after fragmenting (for NuGEN processing), as well as aRNA from the second round of amplification, before and after cleanup, and fragmented-labeled final aRNA (for Arcturus processing).

Hybridization and Microarray Analysis

A total of 2.2 μg of ssDNA labeled and fragmented with the NuGEN kit was used for each sample in the hybridization reaction. The hybridization cocktails were prepared according to the manufacturer instructions, using 2.2 μl of acetylated BSA 50 mg/ml, 2.2 μl herring sperm DNA 10 mg/ml, 22 μl 100%

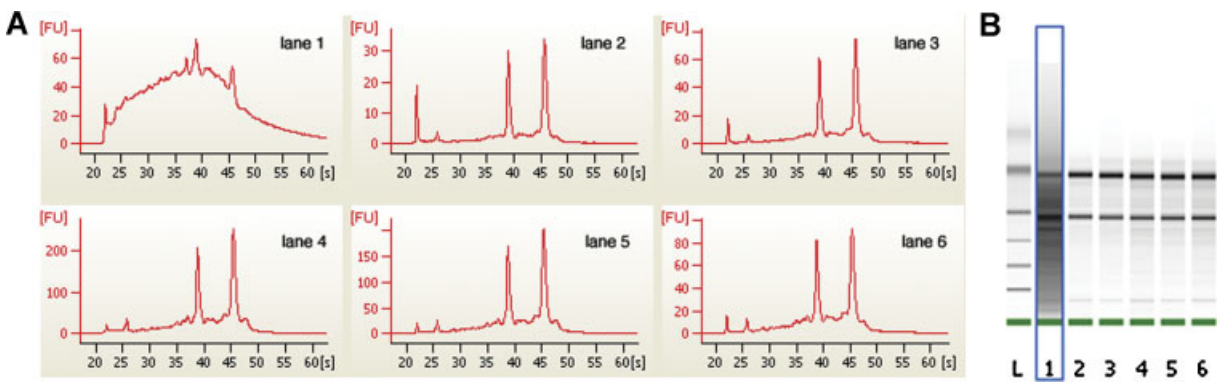

Fig. 2. Quality control of the integrity of LCM-derived RNAs prior to amplification. **A**: Plots of Fluorescence Units versus time (in seconds), and **B**: gel-like image, depicting the electrophoretic runs on the Agilent 2100 Bioanalyzer. Sample in **lane 1** shows evidence of degradation whereas samples in **lanes 2**–**6** show distinct peaks corresponding to intact 28S and 18S ribosomal RNAs.

DMSO, 11 μl 20× eukaryotic hybridization controls (bioB, bioC, bioD, cre) and 3.7 μl control oligonucleotide B2 3 nM, 110 μl 2× hybridization buffer and water to a final volume of 220 μl. Hybridization cocktails were denaturated at 99°C for 2 min, then placed at 45°C for 5 min and finally spun at 14,000 rpm for 5 min before loading them on arrays. A total of 200 μl of the hybridization cocktails was loaded on arrays after a short GeneChip pre-hybridization in 1× hybridization buffer for 10 min in a rotating oven at 45°C and 60 rpm. Hybridization lasted 18 h in a rotating oven at 45°C and 60 rpm.

aRNAs obtained with Arcturus two-round amplification and labeled with Affymetrix one-cycle in vitro transcription were fragmented at 95°C for 30 min in a fragmentation buffer (Affymetrix module for sample cleanup). Fifteen micrograms of each sample were used for hybridization cocktail preparation, which includes the same components as the cocktail prepared for NuGEN samples, with the only differences being in the amount of the sample (15 μg instead of 2.2 μg) and the suggested total volume (300 μl instead of 220 μl). Hybridization cocktails were denatured at 99°C for 5 min (instead of 2 min for NuGEN), then placed at 45°C for 5 min and finally spun at 14,000 rpm for 5 min before loading them on arrays. The pre-hybridization step was conducted in the same way as described above and then the hybridization cocktail (200 μl) was loaded onto the Affymetrix arrays. Hybridization lasted 16 h in a rotating oven at 45°C and 60 rpm.

After hybridization, cocktails were removed from arrays and the latter were completely filled with Wash A solution (Affymetrix). They were subsequently washed inside the Affymetrix Fluidic Station FS-450 with non-stringent Wash A and stringent Wash B solutions. They were then stained with antibody solution (2 mg/ml BSA, 0.1 mg/ml goat IgG stock, 3 μg/ml biotinylated antibody, 300 μl 2× stain buffer and 266.4 μl water) and SAPE solution (2 mg/ml BSA, 10 μg/ml streptavidin phycoerythrin, 600 μl 2× stain buffer and 540 μl water). The fluidic scripts for 11 μm feature size array used were EukGE-WS2-v4 for all arrays loaded with ssDNA obtained with NuGEN, while EukGE-WS2-v5 for arrays loaded with aRNAs generated with Arcturus RiboOA and labeled with Affymetrix IVT. Finally, arrays were placed inside the Affymetrix GeneChip Scanner 3000 for data acquisition. To rule out the possibility of problems during hybridization, we evaluated internal controls (signals for β-actin and GADPH), mean signal intensities and background values.

RESULTS

We studied the quality and reproducibility of signal intensities from Affymetrix U133 plus 2.0 and X3P arrays, each using two different RNA amplification kits, NuGEN and Arcturus. Gene expression measurements were obtained from a colon sample for each array-kit combination. There were eight technical replicates in all, two for each combination. The U133 plus 2.0 array contains 54,675 probe sets while the X3P array contains 61,359 probe sets. The raw CEL files for replicates across the two kits from each array type were pre-processed using RMA [Irizarry et al., 2003] as well as Affymetrix's MAS5 algorithm. We obtained M versus A plots for the two replicates within each array-kit combination to assess data quality. Here, M is the difference in $\log_2$ expression values and A is the average of $\log_2$ expression values [Dudoit et al., 2002]. In essence, we plot the variability in replicate observations versus mean intensity on the $\log_2$ scale. An M versus A plot for normalized data should show a point cloud about the $M = 0$ axis. These are shown in Figure 3(A–D) for RMA-processed data and Figure 4(A–D) for MAS5-processed data, along with the corresponding normalization curve fitted using LOWESS [Cleveland and Devlin, 1988]. In each case, we see a point cloud roughly symmetric about the $M = 0$ axis suggesting good data quality. We also observe that the variability (range of the M scale) is consistently higher for MAS5-processed data relative to RMA.

For replicates in each array-kit combination, we assessed the reproducibility of the data using Spearman's rank correlation and "tightness" of data, measured as the proportion of probe sets that are within 30% of each other in magnitude. For both RMA- and MAS5-processed data, the X3P-Arcturus combination resulted in the highest correlation indicating high reproducibility. Tables I and II present these measures, respectively, for RMA processed data and for MAS5-processed data for each combination.

Additionally, we utilized the call information obtained from Affymetrix's MAS5 software for

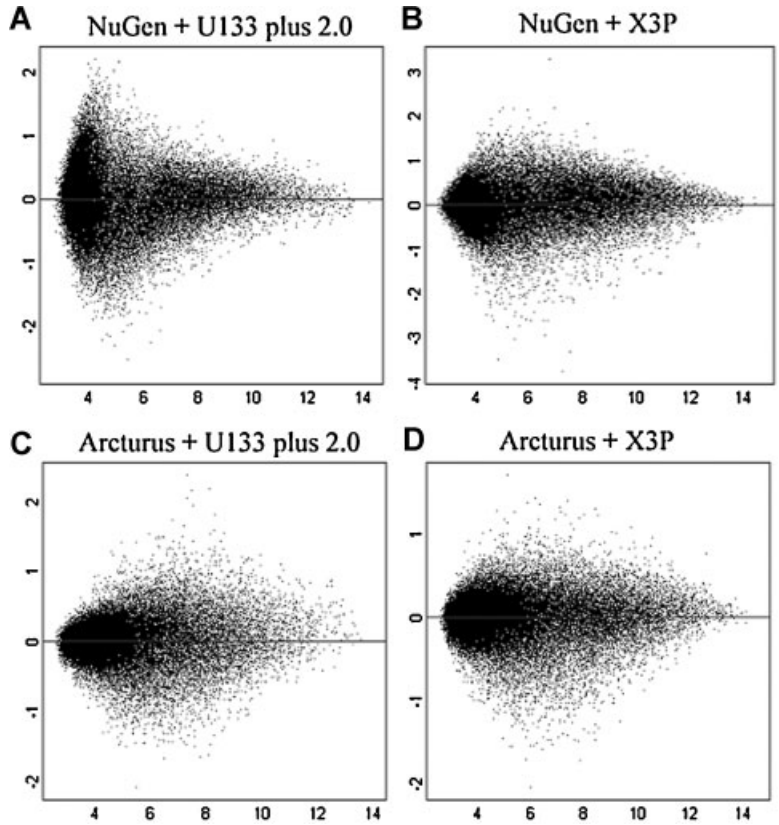

Fig. 3. MA plots obtained with RMA-processed data. **A**: NuGEN-U133 plus 2.0, (**B**) NuGEN-X3P, (**C**) Arcturus-U133 plus 2.0, and (**D**) Arcturus-X3P.

the two replicates in each array-kit combination. The X3P-NuGen combination gave the highest mean percentage of "Present" calls (i.e., expressed genes) (48.01%) followed by X3P-Arcturus (44.54%). We measured the concordance in calls between replicates using Kendall's tau measure. A value close to unity in magnitude indicates high concordance in the pairs being compared. The X3P-Arcturus combination resulted in the highest concordance in calls. Tables III and IV present the percentage of "Present" calls and Kendall's tau, respectively, for each array-kit combination.

We also compared the expression profiles between NuGen and Arcturus in terms of Spearman's rank correlation. We computed these correlations using both RMA as well as MAS5 pre-processed data for all probe sets and each array type. In all cases, the correlation between the mean intensities was around 0.8. The range of correlations between the replicates as well as the correlation of mean intensities between replicates for each case are displayed in Table V.

Based on these analyses, our results have been unidirectional, indicating best data quality as well as the highest reproducibility using the X3P array and the Arcturus kit. This is corroborated by the consistency of our results across the two pre-processing methods applied.

DISCUSSION

Laser capture microdissection is an innovative technology developed a decade ago. It has been described as a rapid one-step procurement under direct microscopic visualization of selected human cell populations from a section of complex, heterogeneous tissue [Emmert-Buck et al., 1996].

Its application in microarray studies is even more recent and during the last couple of years it has greatly increased, generating a remarkable body of literature very quickly.

LCM and microarrays together have demonstrated to be a very useful tool to study gene expression in many different cancer types, such as breast cancer [Cowherd et al., 2004] or gastric

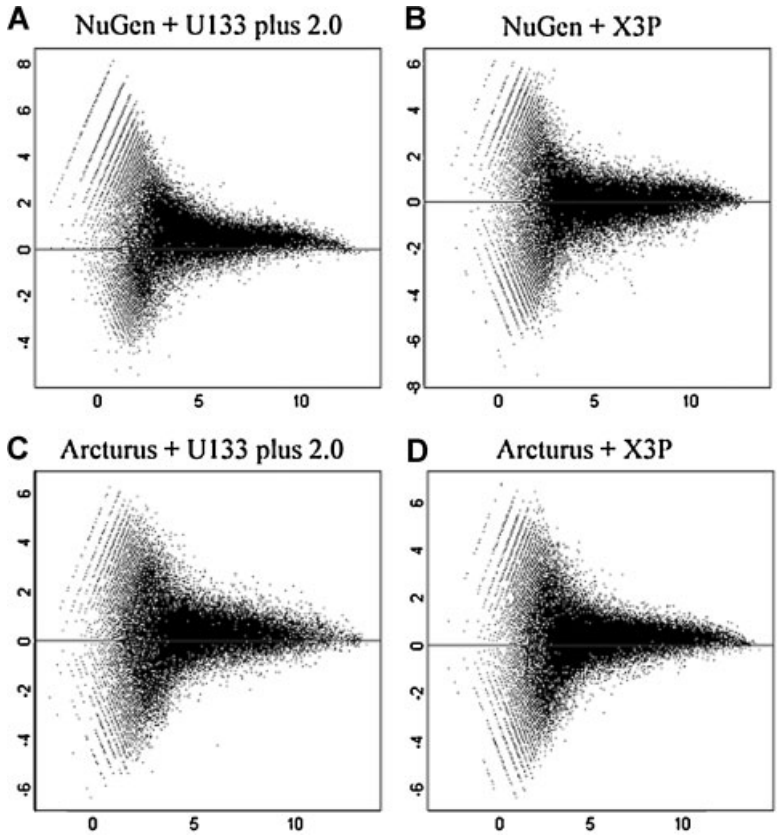

Fig. 4. MA plots obtained with MAS5-processed data. **A**: NuGEN-U133 plus 2.0, (**B**) NuGEN-X3P, (**C**) Arcturus-U133 plus 2.0, and (**D**) Arcturus-X3P.

cancer [Espina et al., 2004]. A premium feature of all these studies is the ability to perform microarrays on pure cell populations increasing accuracy and data quality.

Recently, LCM has been associated with other techniques such as proteomics [Cowherd et al., 2004; Espina et al., 2004; Wu et al., 2005] or immunohistochemistry (immuno-LCM) [Buckanovich et al., 2006]. In this case, immuno-LCM has been used to purify specific cell populations from the tumor microenvironment, investigating molecular events within

TABLE I. Spearman's Rank Correlation Between Replicate Samples From RMA- and MAS5-Processed Data

	RMA		MAS5	
Correlation (Spearman)	U133 plus 2.0	X3P	U133 plus 2.0	X3P
NuGEN	0.882	0.951	0.852	0.880
Arcturus	0.980	0.986	0.878	0.900

TABLE II. "Tightness" of Data Between Replicate Samples From RMA- and MAS5-Processed Data

	RMA		MAS5	
Tightness of replicates (30%)	U133 plus 2.0	X3P	U133 plus 2.0	X3P
NuGEN	67.4	73.7	26.1	38.0
Arcturus	86.9	90.8	36.1	35.0

TABLE III. Percentage of "Present" Calls for Each Array-Kit Combination From MAS5-Processed Data

P (%) Calls—average (range)	U133 plus 2.0	X3P
NuGEN	44.43 (40.51,48.34)	48.01 (47.35,48.67)
Arcturus	41.30 (40.58, 42.02)	44.54 (44.03, 45.04)

TABLE V. Spearman's Rank Correlation Between NuGen and Arcturus Samples From RMA- and MAS5-Processed Data for U133 plus 2.0 and X3P Arrays

Correlation	U133 plus 2.0	X3P
RMA	0.73–0.98 (0.80)	0.78–0.99 (0.81)
MAS5	0.77–0.88 (0.80)	0.76–0.90 (0.80)

specific cellular compartments in the tumor microenvironment.

Due to the increasing number of applications of LCM in many fields, it is important to evaluate different methods for sample processing and to compare different platforms, in order to establish the more reproducible and less variable protocol for microarray analysis of LCM material.

CONCLUSION

In our test conducted on replicate LCM-derived samples, Arcturus Ribo OA™ amplification protocol performed better than NuGEN Ovation™ biotin system giving a correlation of 0.986 in combination with X3P arrays and of 0.980 in combination with Human U133 plus 2.0 arrays, while NuGEN achieved values of 0.951 and 0.882, respectively. We also demonstrated that the X3P arrays performed better in every combination analyzed and we can argue that this higher performance might be even more remarkable for LCM-derived poor quality RNAs. While the Arcturus Ribo OA method appears to be superior in both array choices, the quicker and, at the moment, more expensive NuGEN procedure when coupled with X3P arrays yielded excellent results. This should be taken into account when time considerations are deemed of higher priority than overall quality and cost. Future comparisons with additional platforms, such as the oligonucleotide-based arrays from Agilent and other companies, will allow a determination of the best system for microarray analysis of LCM material.

Finally, RMA is a pre-processing method that includes background correction, normalization and summarization. The normalization part of the procedure takes into account the information from all arrays so that inter-array comparisons can be made. The scaling approach used in the MAS5 algorithm does not account for inter-array variability.

In summary, we believe that the present article provides a useful benchmark for optimization of RNA amplification, array platform and statistical analysis for LCM-based microarray studies.

TABLE IV. Concordance in Calls Between Replicates Using Kendall's Tau Measure From MAS5-Processed Data

Correlation (Spearman/Kendall, P calls)	U133 plus 2.0	X3P
NuGEN	0.753	0.793
Arcturus	0.785	0.807

ACKNOWLEDGMENTS

We thank Andres Klein-Szanto for critical reading of the manuscript, R. Sonlin for expert secretarial assistance, and the Population Studies, Cytogenetics, LCM, and Biostatistics Facilities at Fox Chase Cancer Center. This work was supported by grants N01 CN-95037 and core grant CA-06927 from the National Cancer Institute and by an appropriation from the Commonwealth of Pennsylvania.

REFERENCES

Buckanovich RJ, Sasaroli D, O'Brien-Jenkins A, Botbyl J, Conejo-Garcia JR, Benencia F, Liotta LA, Gimotty PA, Coukos G. 2006. Use of immuno-LCM to identify the in situ expression profile of cellular constituents of the tumor microenvironment. Cancer Biol Ther 5:635–642.

Cleveland WS, Devlin SJ. 1988. Locally-weighted regression. An approach to regression analysis by local fitting. J Am Stat Assoc 83:596–610.

Cowherd SM, Espina VA, Petricoin EF III, Liotta LA. 2004. Proteomic analysis of human breast cancer tissue with laser-capture microdissection and reverse-phase protein microarrays. Clin Breast Cancer 5:385–392.

Dafforn A, Chen P, Deng G, Herrler M, Iglehart D, Koritala S, Lato S, Pillarisetty S, Purohit R, Wang M, Wang S, Kurn N. 2004. Linear mRNA amplification from as little as 5 ng total RNA for global gene expression analysis. Biotechniques 37:854–857.

Dudoit S, Yang YH, Callow MJ, Speed TP. 2002. Statistical methods for identifying genes with differential expression in replicated cDNA microarray experiments. Stat Sin 12:111–139.

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA. 1996. Laser capture microdissection. Science 274:998–1001.

Espina V, Dettloff KA, Cowherd S, Petricoin EF III, Liotta LA. 2004. Use of proteomic analysis to monitor responses to biological therapies. Expert Opin Biol Ther 4:83–93.

Feldman AL, Costouros NG, Wang E, Qian M, Marincola FM, Alexander HR, Libutti SK. 2002. Advantages of mRNA amplification for microarray analysis. Biotechniques 33:906–912. 914.

Goff LA, Bowers J, Schwalm J, Howerton K, Getts RC, Hart RP. 2004. Evaluation of sense-strand mRNA amplification by comparative quantitative PCR. BMC Genomics 5:76–84.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249–264.

Lahdesmaki H, Shmulevich L, Dunmire V, Yli-Harja O, Zhang W. 2005. In silico microdissection of microarray data from heterogeneous cell populations. BMC Bioinformatics 6:54–68.

Luzzi VI, Holtschlag V, Watson MA. 2005. Gene expression profiling of primary tumor cell populations using laser capture microdissection, RNA transcript amplification, and GeneChip microarrays. Methods Mol Biol 293:187–207.

McClain KL, Cai YH, Hicks J, Peterson LE, Yan XT, Che S, Ginsberg SD. 2005. Expression profiling using human tissues in combination with RNA amplification and microarray analysis: Assessment of Langerhans cell histiocytosis. Amino Acids 28:279–290.

Phillips J, Eberwine JH. 1996. Antisense RNA amplification: A linear amplification method for analyzing the mRNA population from single living cells. Methods 10: 283–288.

Schindler H, Wiese A, Auer J, Burtscher H. 2005. cRNA target preparation for microarrays: Comparison of gene expression profiles generated with different amplification procedures. Anal Biochem 344:92–101.

Upson JJ, Stoyanova R, Cooper HS, Patriotis C, Ross EA, Boman B, Clapper ML, Knudson AG, Bellacosa A. 2004. Optimized procedures for microarray analysis of histological specimens processed by laser capture microdissection. J Cell Physiol 201:366–373.

Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. 2000. High-fidelity mRNA amplification for gene profiling. Nat Biotechnol 18:457–459.

Wang H, Owens JD, Shih JH, Li MC, Bonner RF, Mushinski JF. 2006. Histological staining methods preparatory to laser capture microdissection significantly affect the integrity of the cellular RNA. BMC Genomics 7:97–104.

Wu MS, Lin YS, Chang YT, Shun CT, Lin MT, Lin JT. 2005. Gene expression profiling of gastric cancer by microarray combined with laser capture microdissection. World J Gastroenterol 11:7405–7412.