Journal of Molecular Diagnostics, Vol. 7, No. 1, February 2005

Reproducibility, Fidelity, and Discriminant Validity of mRNA Amplification for Microarray Analysis from Primary Hematopoietic Cells

Liang Li,* Joe Roden,† Bruce E. Shapiro,† Barbara J. Wold,‡ Smita Bhatia,§ Stephen J. Forman,* and Ravi Bhatia*

From the Divisions of Hematology and Bone Marrow Transplantation and Pediatrics,§ City of Hope National Medical Center, Duarte; the Machine Learning Systems Group,† Jet Propulsion Laboratory, Pasadena; and the Division of Biology,‡ California Institute of Technology, Pasadena, California*

Analysis of gene expression in clinical samples poses special challenges, including limited RNA availability and poor RNA quality. Quantitative information regarding reliability of RNA amplification methodologies applied to primary cells and representativeness of resulting gene expression profiles is limited. We evaluated four protocols for RNA amplification from peripheral blood mononuclear cells. Results obtained with 100 ng or 10 ng of RNA amplified using two rounds of cDNA synthesis and *in vitro* transcription were compared with control 2.5-μg RNA samples processed using a single round of *in vitro* transcription. Samples were hybridized to Affymetrix HG-U133A arrays. Considerable differences in results were obtained with different protocols. The optimal protocol resulted in highly reproducible gene expression profiles from amplified samples ($r = 0.98$) and good correlation between amplified and control samples ($r = 0.94$). Using the optimal protocol dissimilarities of gene expression between mononuclear cells from a normal individual and a patient with myelodysplastic syndrome were primarily maintained after amplification compared with controls. We conclude that small variations in methodology introduce considerable distortion of gene expression profiles obtained after RNA amplification from clinical samples and too strong a focus on a very small number of genes picked from an array analysis could be unduly influenced by seemingly acceptable methodologies. However, it is possible to obtain reproducible and representative results using optimized protocols. *(J Mol Diagn 2005, 7:48–56)*

Gene expression analysis using microarray technology is being increasingly used to investigate mechanisms of normal and malignant hematopoiesis, and for the pathophysiological, diagnostic, and prognostic classification of hematological and other malignancies.[1–5] However there are several problems associated with effective implementation of this technology including heterogeneity of cellular composition, often requiring isolation of homogenous subpopulations of cells, and limitations in the size of clinical samples that reduces the quantity of RNA available for analysis. Therefore, mRNA amplification is generally required for analysis of clinical samples.

A technique capable of amplifying small amounts of mRNA without significantly distorting relative mRNA levels is needed to meet the challenge presented by small clinical samples. Linear amplification based on cDNA synthesis and T7 RNA polymerase based *in vitro* transcription (IVT) has been reported to maintain representation of mRNA levels.[6,7] Thus, mRNA amplification using a protocol described by Van Gelder and colleagues[6] was shown to faithfully maintain relative mRNA levels with 1 μg of poly(A)+ or 10 μg of total RNA as starting material,[7,8] and has become the basis for the standard labeling protocol for samples hybridized on Affymetrix GeneChip microarrays. It is also possible to amplify even smaller amounts of RNA by using additional cycles of reverse transcriptase (RT) and IVT.[9,10] Baugh and colleagues[11] reported that the above protocol, in addition to amplifying target mRNA, generated template-independent products that could compromise the specific activity of amplified products. They described a modified amplification protocol that minimized generation of template-independent products by reducing the amount of primer and using small cDNA synthesis volumes. In addition, several kits for RNA amplification based on the above principles are commercially available. These different methodologies have been widely used for RNA amplification and gene expression profile analysis.[12–14] Although they allow analysis of gene expression profiles from tiny samples,

Supported by grants from the National Aeronautics and Space Administration/National Cancer Institute Biomolecular Systems Research Program, the Leukemia and Lymphoma Society of America Translational Research Program, and the General Clinical Research Center (grant 5M01 RR00043).

Accepted for publication September 2, 2004.

A part of the research described in this study was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Address reprint requests to Ravi Bhatia M.D., Division of Hematology and Bone Marrow Transplantation, City of Hope National Medical Center, 1500 East Duarte Rd., Duarte, CA 91010. E-mail: rbhatia@coh.org.

Table 1. Comparison of Different Protocols for RNA Amplification

	Items	AB	AP	HP	AR
1st round					
FSS	Primer	Oligo dT-T7 specified by Affymetrix*		Oligo dT-T7 specified by Baugh et al†	Undisclosed
	Amount of primer	5 pmol/μl		0.5 pmol/μl	Undisclosed
	Volume	20 μl		<10 μl	20 μl
SSS	Volume	150 μl	150 μl	75 μl	50 μl
	T4 DNA polymerase	5 minutes	5 minutes	15 minutes	Undisclosed
	Clean-up	PCI extraction	PCI extraction	PCI extraction and chromatography	Column purification
IVT	Incubation	37°C for 4 hours		42°C for 9 hours	37°C for 5 hours
	Reagents	From Ambion kit		Various sources	Provided by kit
2nd round					
FSS	Primer		Random hexamer		Undisclosed
	Amount of primer	50 ng/μl	50 ng/μl	50 ng/μl	Undisclosed
	Volume	20 μl	20 μl	10 μl	20 μl
SSS	Primer		Same as primers used in FSS of 1st round		
	Amount of primer	100 pmol	100 pmol	5 pmol (100 ng)	Undisclosed
	Volume	150 μl	150 μl	75 μl	50 μl
	T4 DNA polymerase	5 minutes	5 minutes	15 minutes	Undisclosed
	Clean-up		Same as 1st round		
IVT	Incubation		37°C for 5 hours		
	Reagents	BioArray high-yield RNA transcript labeling kit (Enzo Life Science)			

*5′-GCATTAGCGGCCGCGAAATTAATACGACTCACTATAGGGAGA(T)21V-3′
†5′-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG(T)24-3′
FSS, first strand synthesis; SSS, second strand synthesis; IVT, *in vitro* transcription.

the concern is that bias introduced during the amplification process will result in significant changes in the gene expression profile. Moreover, it has been reported that fidelity is more at risk when smaller amounts of starting RNA are used, as might be expected.[11] This problem may become even more prominent when RNA amplification is applied to clinical samples because of poorer RNA quality resulting from reduced mRNA content and losses associated with sample handling and processing. Quantitative information regarding the reliability of RNA amplification from small numbers of primary cells obtained from clinical samples is not available. Moreover, in the absence of direct comparisons the relative efficacy of the different available protocols is not clear.

An optimal mRNA amplification method should provide reproducible results, maintain fidelity of gene expression profile compared with nonamplified controls, and retain capacity to discriminate differences in gene expression profiles between two different samples. Here we compared the effectiveness, reproducibility, and fidelity of several currently available methods of RNA amplification when applied to small amounts of RNA derived from peripheral blood mononuclear cells. The protocols tested include: 1) the Affymetrix recommended protocol (AB); 2) a modified protocol (HP) reported by Baugh and colleagues[11] based on perceived limitations of the AB protocol that uses smaller reaction volumes and lower primer concentrations, and uses reagents from different sources for the first IVT (*http://mcb.harvard.edu/hunter/Protocols/aRNAprotocol.pdf*); 3) a protocol similar to the AB protocol using reagents for first round IVT based on Baugh and colleagues[11] instead of the Ambion Megascript kit (Ambion, Austin, TX) used in the AB protocol (AP); and 4) the Arcturus RiboAmp RNA amplification kit (AR) (Arcturus, Carlsbad, CA). Samples were hybridized to Affymetrix GeneChip HG-U133A microarrays.

Materials and Methods

Sample Collection and RNA Extraction

Peripheral blood samples were collected from two normal donors and one t-MDS patient using protocols approved by the Human Subjects Committee of City of Hope. Mononuclear cells were separated by Ficoll Hypaque density gradient separation. Total RNA was isolated using the RNeasy mini kit (Qiagen, Hilden, Germany). The integrity of the total RNA was checked by formaldehyde-denaturing agarose gel electrophoresis and labeling with SYBR Gold (Molecular Probes, Eugene, OR). All RNA samples showed sharp ribosomal bands and a ratio of intensity of 28S and 18S ribosomal RNA bands of ~2:1. Samples were processed within 2 hours of being drawn.

RNA Amplification

Four different RNA amplification protocols were used and are summarized in Table 1. Results obtained with 2.5 μg of RNA (controls) were compared with results obtained after amplification of 100 ng of RNA from the same samples. The 2.5-μg RNA samples were processed using one round of cDNA synthesis using RT and IVT. The 100-ng RNA samples were amplified using two rounds of RT and IVT, respectively. All reagents were from Invitrogen (Carlsbad, CA) except where separately noted. The final IVT reactions for the synthesis of biotin-labeled cRNA were performed for all protocols using a BioArray

High-Yield RNA transcript labeling kit (Enzo Life Science, Farmingdale, NY). Biotin-labeled cRNA was quantified by absorbance at 260 nm and analyzed by formaldehyde-denaturing agarose gel electrophoresis. Control and amplification reactions with different protocols were repeated at least three times.

Baugh's Modified Protocol (HP)

For control 2.5-μg samples, RT was performed in first-strand buffer with 0.1 μg of T7-$(dT)_{21}$ primer [5′-GCATTAGCGGCCGCGAAATTAATACGACTCACTATAGGGAGA-$(T)_{21}$V, V = A, C, and G] for 1 hour at 42°C. Second-strand synthesis (SSS) was performed in 75-μl total volume with 40 U of DNA polymerase I, 2 U of RNase H, and 10 U of *Escherichia coli* DNA ligase in second-strand buffer by adding 65 μl of an ice-cold SSS premix to the RT reaction and incubating at 16°C for 2 hours. Double-stranded (ds) cDNA was polished with T4 DNA polymerase for 15 minutes and purified by phenol-chloroform extraction followed by chromatography on a BioGel P-6 column (Bio-Rad, Hercules, CA). Biotin-labeled cRNA was obtained through IVT using a BioArray High-Yield RNA transcript labeling kit. For 100-ng samples, the first round of RT and IVT were as for 2.5-μg samples, except that IVT reactions were performed in 40-μl reaction volumes using 160 U of T7 RNA polymerase (Promega, Madison, WI), 7.5 mmol/L each of the GTP, ATP, CTP, and UTP, 60 U of RNase inhibitor in Ampliscribe buffer (Epicenter, Madison, WI) at 42°C for 9 hours. For the second round of amplification, RT was in 10-μl reaction volume with 0.5 μg of random hexamers for 20 minutes at 37°C, 20 minutes at 42°C, 10 minutes at 50°C, and 10 minutes at 55°C. For the second round of SSS, 1 U of RNase H was added to the heat-inactivated RT reaction at 37°C for 30 minutes, followed by denaturation at 95°C for 2 minutes, and snap-cooling on ice. T7-$(dT)_{21}$ primer (0.1 μg) was then added to the chilled cDNA and the SSS reaction primed by incubation for 10 minutes at 42°C followed by snap-cooling on ice. Ice-cold SSS premix (65 μl) (as in the first round but without ligase) was added and incubation and polishing were performed as in the first round. Purification of ds cDNA and IVT to generate biotin-labeled cRNA was as in the first round.

Modified Affymetrix Protocol (AP)

An Affymetrix recommended protocol was modified based on results reported by Baugh and colleagues.[11] Processing of 2.5-μg samples follows the standard protocol recommended by Affymetrix. This protocol uses a 20-μl RT reaction with 200 U of SuperScript II reverse transcriptase and 100 pmol of T7-$(dT)_{24}$ primer [5′-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-$(dT)_{24}$] in first-strand buffer at 42°C for 1 hour. SSS was performed in a 150-μl reaction with the same condition as HP. ds cDNA was polished with 20 U of T4 DNA polymerase at 16°C for 5 minutes. cDNA was purified by Phase Lock gels (PLG) (Eppendorf, Hamburg, Germany), phenol/chloroform extraction, and ethanol precipitation. Synthesis of biotin-labeled cRNA was performed through IVT. For amplification of 100-ng RNA samples, the first round of RT was as above with the exception that first round IVT was performed as with the HP protocol. For the second round of amplification, RT was in 20 μl with 1 μg of random hexamers (Promega) for 2 hours at 42°C followed by RNase H treatment at 37°C for 20 minutes and deactivation of RNase H by heating to 95°C for 5 minutes. T7-$(dT)_{24}$ primer (100 pmol/L), as used in the first round of amplification, was added to the RT reaction and the SSS was primed by incubating at 70°C for 10 minutes followed by snap-cooling on ice. Ice-cold SSS premix (128 μl) containing 40 U of DNA polymerase I in second-strand buffer was added and incubation and polishing were performed as in the first round. After purification of ds cDNA, biotin-labeled cRNA was obtained through IVT.

Standard Affymetrix Protocol (AB)

Processing of 2.5-μg RNA samples was as described above for the AP protocol. For 100-ng samples requiring two rounds of amplification, the Megascript T7 IVT kit (Ambion, Austin, TX) was used for the first round IVT reaction. Other details of this protocol are as described for the AP protocol.

Arcturus RNA Amplification Kit (AR)

The RiboAmp RNA amplification kit (Arcturus, Carlsbad, CA) was used following the manufacturer's instructions. For 2.5-μg samples, RT was in 20 μl (10 μl of RNA in H_2O, 1 μl of primer A, 7 μl of first-strand master mix, and 2 μl of first-strand enzyme mix) at 42°C for 45 minutes followed by 20 minutes at 37°C with 2 μl of first-strand nuclease mix. SSS was primed by adding 1 μl of primer B and 30 μl of SSS mix (29 μl of second-strand master mix and 1 μl of second-strand enzyme mix) to the heat-inactivated RT reaction for 5 minutes at 25°C, 10 minutes at 37°C, and 5 minutes at 70°C. IVT was performed as for the other protocols. For amplification of 100-ng RNA samples, the first round of RT was as with 2.5-μg samples, and IVT was performed in 40-μl reactions (16 μl of ds cDNA, 8 μl of IVT buffer, 12 μl of IVT master mix, and 4 μl of IVT enzyme mix) at 42°C for 3 hours using reagents from the RiboAmp RNA amplification kit. For the second round amplification, RT was done in the same conditions as for the first round except that primer B was used for first-strand cDNA synthesis, and primer A for SSS. Biotin-labeled RNA was obtained by IVT. All of the purification steps in this protocol are performed using the columns provided with the kit.

Fragmentation, Microarray Hybridization, and Scanning

Fragmentation, hybridization, washing, staining, and scanning were all performed using procedures recommended in the GeneChip Expression Analysis Technical Manual (Affymetrix, Santa Clara, CA) at the DNA Array

Core Facility, University of California, Irvine. Labeled cRNA (10 μg) was used in each hybridization on HG-U133A oligonucleotide microarray (Affymetrix). Affymetrix gene chip images were scanned at 3-μm resolution by using a GeneArray Scanner (Hewlett-Packard, Palo Alto, CA). The expression value of each gene was calculated using Affymetrix Microarray Suite software version 5.0.

Data Analysis

Data from each chip was scaled to a target signal intensity of 500. Signal intensity values derived from MAS 5.0 software were transformed to log2 for further analysis. Noise analysis was performed by comparing the means and standard deviations of log2-transformed signal intensities from three replicates. For reproducibility analysis correlation coefficients of signal intensity values between replicate experiments were calculated. For fidelity analysis correlation coefficients of mean values for control and amplified samples were calculated. To analyze the effect of amplification on differential gene expression between two different samples, a comparison of *t*-scores for individual genes was used [*t*-scores = $(X_{i,normal} - X_{i,t\text{-}MDS})/[(s^2_{normal}/n_{normal}) + (s^2_{t\text{-}MDS}/n_{t\text{-}MDS})]$, where X = mean across replicates of log2-transformed signal intensities for a single gene, i, in either the normal or t-MDS RNA sample;s^2 = variance in replicates of log2-transformed signal intensities for a single gene, averaged across all genes; and n = number of replicates]. Hierarchical clustering analysis was performed using GeneSpring 6.0 software (Silicon Genetics, Redwood City, CA).

Results

We evaluated the efficacy, reproducibility, and fidelity of four methods of RNA amplification applied to small amounts of RNA from primary clinical samples. Ten μg of labeled cRNA is required for hybridization on a HG-U133A microarray. All of the protocols tested here could generate sufficient amounts of cRNA for hybridization from 2.5 μg as well as 100 ng of RNA samples (Figure 1). The highest cRNA yield was obtained with AB and AP protocols, and the lowest yield with the AR protocol with both one and two rounds of IVT. This could be related in part to a greater degree of difficulty associated with handling small amounts of material with the AR protocol.

We compared the distribution of means and standard deviations of signal intensities of all genes on the arrays for 2.5-μg and 100-ng RNA samples processed using the different protocols. The data were plotted as scatterplots of mean signal intensity versus variance as well as histograms of means and of variance. For all four protocols 2.5-μg RNA samples processed with one round of IVT had a more compact distribution of standard deviations than 100-ng RNA samples processed with two rounds of IVT (Figure 2). This suggests that results obtained with amplified samples are slightly less consistent and reproducible than with control samples, as expected, although the differences were not large.

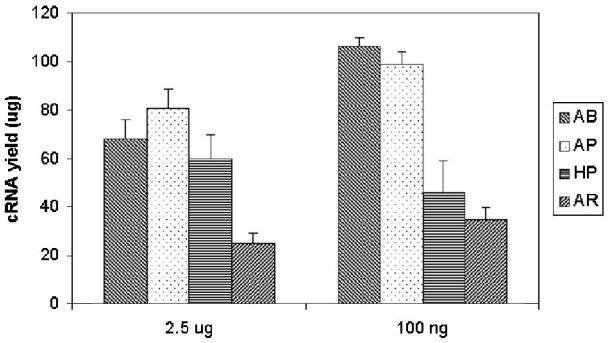


Figure 1. cRNA yields after RNA amplification. cRNA yields after processing of 2.5 μg and 100 ng of total RNA from normal peripheral blood mononuclear cells, with one and two rounds of IVT, respectively, are shown. All protocols generated sufficient amounts of labeled cRNA for hybridization on HG-U133A chips (10 μg) from both 2.5 μg and 100 ng of total RNA. The highest yields were obtained with the AB and AP protocols and the lowest yield with the AR protocol.

To quantify the reproducibility of the amplification reactions, correlation coefficients of signal intensities of individual genes from replicate amplifications and hybridizations were analyzed (Figure 3). One round of amplification did not introduce significant stochastic bias for any of the protocols studied, with the AB and AP protocols providing the best reproducibility. Amplification of 100 ng of RNA using two rounds of IVT resulted in a modest reduction of reproducibility with the AP, HP, and AR protocols. However, the AB protocol yielded the same reproducibility for amplified samples as for controls (r_{mean} = 0.98; n = 3). Evaluation of the concordance of positive and negative calls between replicates yielded similar results with a high degree of concordance in control samples, reduced concordance in amplified samples, and with the best results being obtained with the AB protocol (data not shown).

To measure the effect of amplification on fidelity, correlation coefficients of gene expression profiles obtained from 2.5-μg and 100-ng RNA samples were calculated (Figure 4). With all four protocols correlation between 2.5-μg and 100-ng RNA samples was lower than that between replicate 2.5-μg or 100-ng RNA samples. The reduced correlation may be related to a decrease in the number of present calls detected after RNA amplification (Figure 5). Of the four protocols studied, the AB and AR protocols provided the best correlation (r = 0.94) between control and amplified samples. We conclude that optimized RNA amplification techniques that use two rounds of IVT can produce results similar to those obtained with controls, albeit with detectable reductions in fidelity.

The AB protocol therefore demonstrated overall superiority when cRNA yield, reproducibility, and fidelity are considered together. This protocol was then used to amplify an even smaller 10-ng RNA sample. The protocol was modified from that used for 100-ng RNA samples in that the reaction volume of RT was reduced fourfold. Amplification of 10-ng RNA samples resulted in a further, yet still modest, reduction in reproducibility (r = 0.95, n = 3) and fidelity (r = 0.92 compared with 2.5 μg and r = 0.96 compared with 100-ng RNA samples).

We next evaluated whether differences in gene expression profiles of two different RNA samples, one derived from

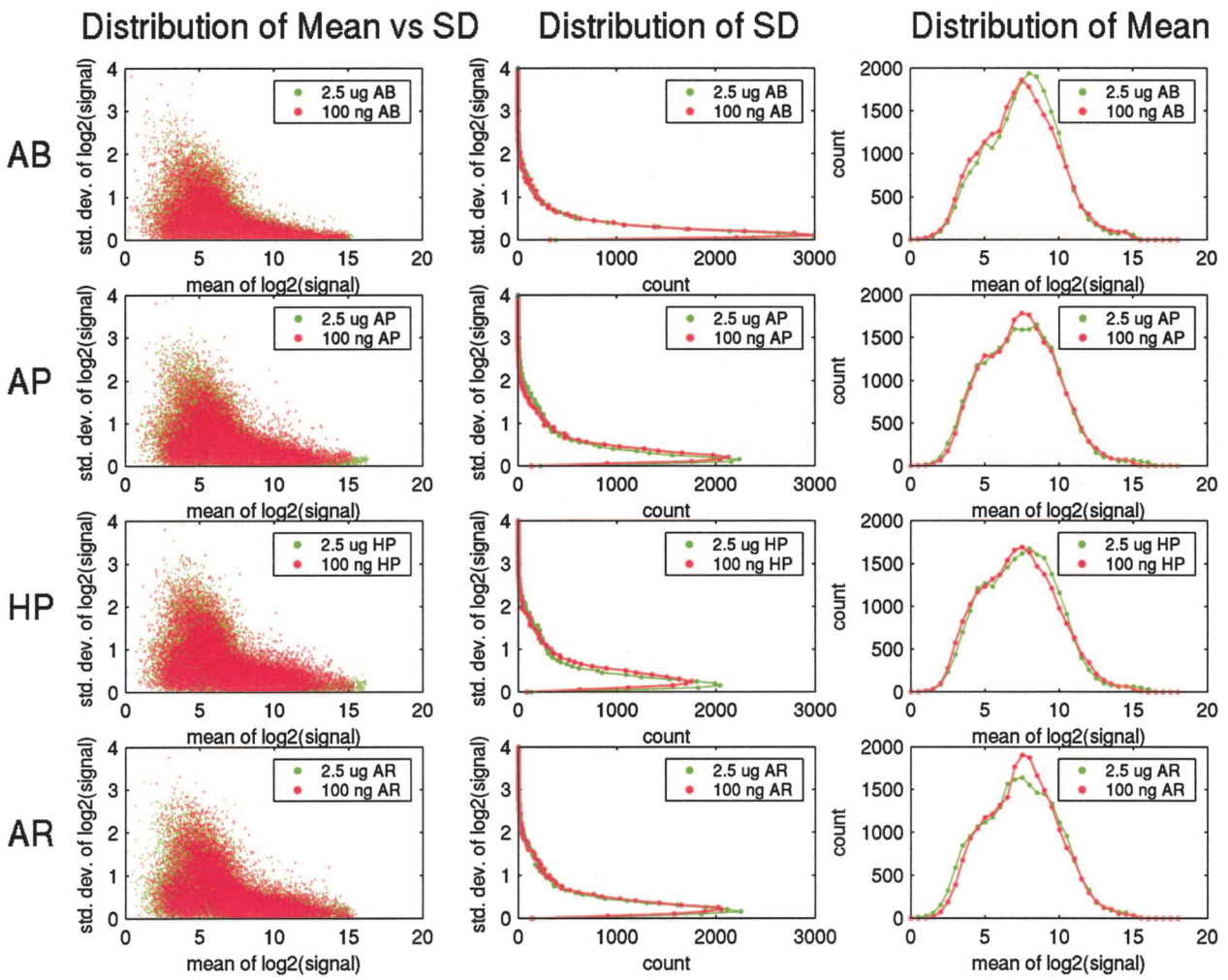
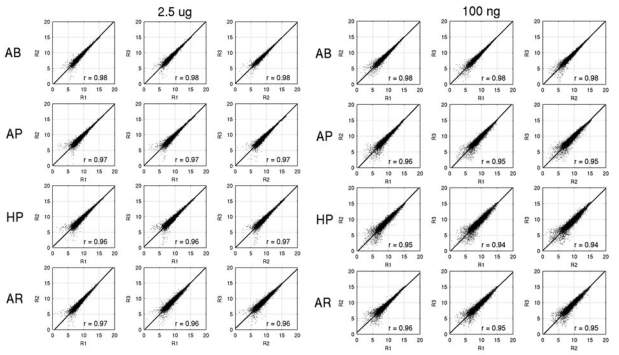


Figure 2. Noise analysis of gene expression profiles after RNA amplification. The distribution of the means and SD of signal intensities of all genes after processing of 2.5 μg of RNA with one round of amplification (green) and 100 ng of RNA with two rounds of amplification (red) is shown. Results represent three replicate amplifications and hybridizations. A slightly more compact distribution of standard deviations was observed after one round of amplification compared with two rounds of amplification for all four protocols.

normal mononuclear cells and the other mononuclear cells from a patient with therapy-related myelodysplasia, were maintained after amplification with the AB protocol. The correlation coefficients of mean signal intensities of gene expression of the two samples showed similar dissimilarity with 100-ng samples (r = 0.86) as with 2.5-μg RNA samples (r = 0.87) (Figure 6A). We also calculated gene-specific t-scores for difference in signal intensities between the two different samples for both 100-ng RNA and 2.5-μg RNA samples. These t-scores are linearly related to the log of ratios of mean signal intensities for the two samples, but have the added advantage that replicates are used and statistical significance is considered. t-Scores obtained from amplified and control samples showed good correlation (Figure 6B). Lists of genes with the highest difference in t-scores obtained with 2.5-μg and 100-ng RNA samples were compared (Table 2). Seven of the top ten genes overexpressed in t-MDS samples were common to 2.5-μg and 100-ng samples, and all of the top 10 genes from 2.5-μg samples were present within the top 15 genes for 100-ng samples. Four of the top ten genes with reduced expression in t-MDS samples were common to 2.5-μg samples and 100-ng samples, and 7 of the top 10 genes from 2.5-μg samples were present within the list of top 20 genes for 100-ng samples. We also performed clustering analysis to assess the effect of amplification on our ability to detect genes differentially expressed between the two samples. A hierarchical clustering algorithm was used. We found that genes expressed differentially between the normal and t-MDS samples clustered well in amplified samples and controls (Figure 7). These results indicate that although some distortion of data does occur after amplification, it is possible to obtain reliable and representative data regarding differences in gene expression between two samples after RNA amplification using the AB protocol.

Discussion

RNA amplification is increasingly used for analysis of gene expression in clinical specimens, but quantitative information regarding the reliability of resultant gene ex-



Amount	Protocol	1 vs 2	1 vs 3	2 vs 3	r_{mean}
2.5 µg	AB	0.98	0.98	0.98	0.98
	AP	0.97	0.97	0.97	0.97
	HP	0.96	0.96	0.97	0.96
	AR	0.97	0.96	0.96	0.96
100 ng	AB	0.98	0.98	0.98	0.98
	AP	0.96	0.95	0.95	0.95
	HP	0.95	0.94	0.94	0.94
	AR	0.96	0.95	0.95	0.95

Figure 3. Comparison of gene expression profiles from replicate amplification reactions. Scatterplots comparing signal intensities of present calls from three replicate amplifications and correlation coefficients of signal intensity values are shown. Present calls were defined as being present in at least one replicate. For 2.5-μg total RNA samples processed using one round of IVT the AB protocol provided the best reproducibility. Amplification of 100 ng of total RNA using two rounds of IVT was associated with reduced reproducibility with the AP, HP, and AR protocols. However the AB protocol yielded similar reproducibility as was seen with one round of amplification.

pression profiles is lacking. In this study we investigated the reproducibility and reliability of four different protocols for amplification of small quantities of RNA applied to primary cells from clinical samples. We have also inves-

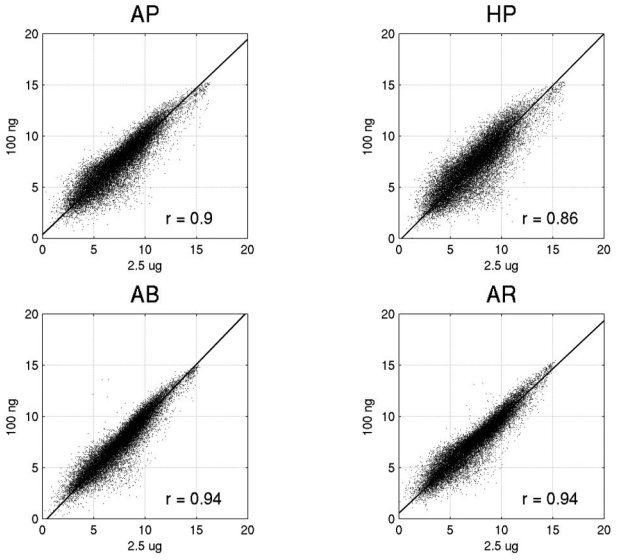


Figure 4. Comparison of gene expression profiles obtained after amplification of 2.5 μg and 100 ng of total RNA. Scatterplots comparing mean signal intensities of individual genes (all calls) obtained after processing of 2.5 μg of total RNA with one round of amplification and 100 ng of RNA with two rounds of amplification are shown. Correlation coefficients of mean signal intensity values between one and two rounds of amplification were best with the AB protocol.

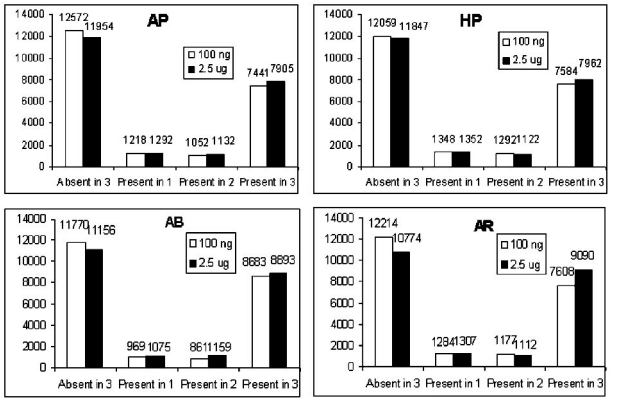


Figure 5. Effect of amplification on the number of present calls detected. The number of present calls detected in three replicate samples after processing of 2.5 μg of total RNA with one round of amplification and 100 ng of RNA with two rounds of amplification are shown. Two rounds of amplification resulted in a small decrease in the number of present calls. The AB protocol demonstrated the least reduction in numbers of present calls after two rounds of amplification.

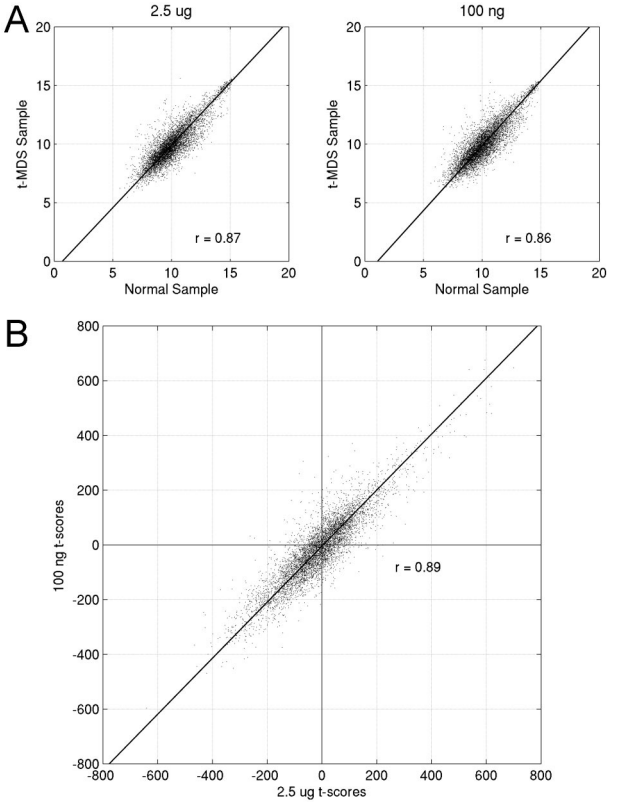


Figure 6. Effect of amplification on discrimination of differences in gene expression profiles of two samples. Gene expression profiles of peripheral blood mononuclear cell RNA samples from a normal individual (sample 1) and a patient with t-MDS (sample 2) obtained with 2.5 μg and 100 ng of total RNA using the AB protocol were compared. Genes reported present in all replicates were included. **A:** Scatterplots comparing mean signal intensities of individual genes obtained with 2.5 μg and 100 ng of total RNA from sample 1 and sample 2 are shown. Correlation coefficients of mean signal intensities of the two samples were maintained after amplification from 100 ng of total RNA ($r = 0.86$) compared with 2.5 μg of total RNA ($r = 0.87$). **B:** Scatterplots of gene-specific t scores comparing mean signal intensities of individual genes for the two RNA samples obtained with 2.5 μg and 100 ng of total RNA. The correlation coefficient for gene-specific t scores for 2.5 μg and 100 ng of total RNA samples was 0.89.

Table 2. Lists of Genes with the Highest *t*-Scores Comparing Normal and t-MDS Samples with and without Amplification

	Increased expression in t-MDS sample				Reduced expression in t-MDS sample			
	100 ng		2.5 μg		100 ng		2.5 μg	
	Probe	Gene	Probe	Gene	Probe	Gene	Probe	Gene
1	208353_x_at	ANK1	213515_x_at	HBG2	208450_at	LGALS2	208450_at	LGALS2
2	213515_x_at	HBG2	208691_at	TFRC	221841_s_at	KLF4	209728_at	*HLA-DRB4*
3	216054_x_at	MYL4	209585_s_at	MINPP1	214329_x_at	*TNFSF10*	209670_at	TRA@
4	206937_at	*SPTA1*	218194_at	*DKFZP566 E144*	211339_s_at	*ITK*	203932_at	*HLA-DMB*
5	207332_s_at	TFRC	207332_s_at	TFRC	209670_at	TRA@	213017_at	*ABHD3*
6	209585_s_at	MINPP1	209122_at	*ADFP*	213906_at	*MYBL1*	219528_s_at	*BCL11B*
7	203031_s_at	*UROS*	208353_x_at	ANK1	201506_at	*TGFBI*	221841_s_at	KLF4
8	216379_x_at	CD24	216054_x_at	MYL4	208891_at	*DUSP6*	219947_at	CLECSF6
9	209771_x_at	CD24	209771_x_at	CD24	220784_s_at	*UTS2*	201739_at	*SGK*
10	209894_at	*LEPR*	216379_x_at	CD24	219947_at	CLECSF6	204642_at	*EDG1*

tigated the ability to discriminate between gene expression profiles of two different samples after amplification. To our knowledge, this is the first report of a systematic assessment of reproducibility and fidelity of gene expression profiles after amplification of small RNA samples from clinical samples.

A primary requirement for any amplification protocol is a high degree of reproducibility. Reproducibility was assessed by evaluating the variance between individual mean expression levels, by evaluating the correlation coefficients of mean expression levels, and by evaluating the concordance of present and absent calls between replicate samples processed using the same protocol. All protocols we tested yielded reproducible gene expression profiles when control amounts of RNA and one cycle of IVT were used. However with the AP, HP, and AR protocols a small reduction in reproducibility was found after two rounds of IVT. In contrast the AB (Affymetrix recommended) protocol resulted in the most reproducible gene expression profiles for both control and amplified samples, and did not result in a decrease in reproducibility after amplification. Because the major difference between the AB and AP amplification protocols is the source of reagents used in first IVT, these results indicate that the quality-controlled and optimized Ambion MegaScript kit may have contributed to the improved results. However difficulties in specimen handling may have contributed to increased variability in results with another optimized reagent kit (AR).

An important consideration is the extent to which the amplification process distorts and introduces bias in the gene expression profile. Gene expression profiles obtained after amplification of 100-ng RNA samples with two rounds of IVT maintained similarity to control 2.5-μg RNA samples, but the similarity was less than that seen within replicates. In addition the concordance for present and absent calls between different protocols was lower than that seen with replicates (data not shown). Distortion in gene expression profiles appeared to be related, at least in part, to loss of a small fraction of positive calls, as determined by the MAS5.0 software, after amplification. Changes in signal intensities of certain gene clusters after amplification may relate at least in part the reduced size and 5′ complexity of the amplified products (0.1 ~ 1.5 kb) compared to unamplified control products (0.2 ~ 9 kb). The reduced 5′ complexity of the amplified product could affect hybridization to certain probe pairs and the calculated signal intensities. The highest degree of correlation of mean expression levels of individual genes between controls and amplified samples was seen with the AB and AR protocols indicating that these protocols resulted in the best fidelity in the gene expression profiles after amplification. A further decrease of fidelity was observed when amplifying a smaller amount of RNA (10 ng RNA) using the modified AB protocol (data not shown), although the results still captured similar groups of putatively up-regulated or down-regulated RNAs as were found from 100-ng samples. Therefore, misleading results may be obtained if care is not taken to standardize the amount of RNA used and the amplification protocol used throughout all experimental conditions. We suspect that reliable amplification from different starting amounts of RNA may require a separate protocol optimization.

The AB protocol provided the best results in terms of reliability of yield, reproducibility, and fidelity. Baugh and colleagues[11] reported that IVT reaction catalyzed by T7 RNA polymerase in the presence of oligo-(dT) could generate template-independent high-molecular weight RNA that could compromise the specific activity of amplified products. Modification of the protocol to reduce the amount of primer carried over from RT to IVT increased the number of present calls more than 10-fold. Although we also observed template-independent amplification on agarose gel electrophoresis after one round of amplification with the AP protocols (data not shown), when the RNA amplification products were applied to microarrays the number of present calls did not differ between Affymetrix protocol (AP) and Baugh's modified protocol (HP). Therefore template-independent amplification did not affect sensitivity in our hands. In fact, the correlation coefficient between amplified and control samples was higher with the AP than the HP protocol ($r = 0.90$ versus $r = 0.84$). Baugh and colleagues[11] also noted that the Affymetrix protocol was limited by loss of 5′ complexity related to the random-primed RT reaction used in the second round of amplification. They reported that inclusion of the T4 single-strand nucleic acid-binding protein (T4gp32) in RT reactions helped maintain 5′ complexity in the amplified products. However in our studies the presence of T4gp32 significantly reduced the efficiency

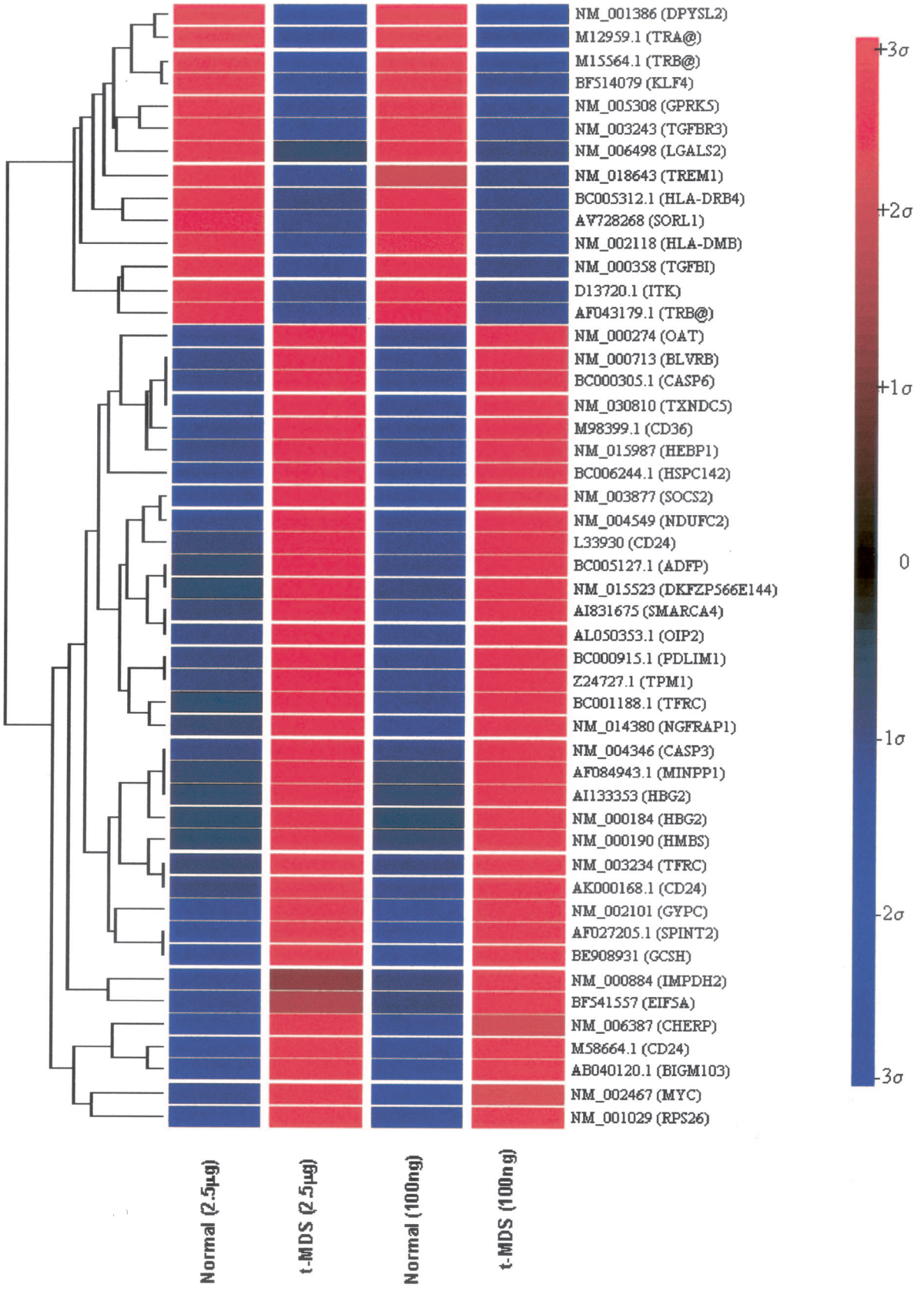


Figure 7. Clustering of genes between control and amplified normal and t-MDS samples. Results of hierarchical clustering of genes differentially expressed between normal and t-MDS samples. Analysis was restricted to genes differentially expressed in the 2.5-μg sample (more than fivefold difference in expression, $P < 0.02$). Each row represents a single gene and the columns represent replicate 2.5 μg and 100 ng of normal and t-MDS RNA samples. Data are colored based on how far the gene is overexpressed (red) or underexpressed (blue) relative to a normalized expression level of 1, in terms of the SE of the measurement.

of RT reaction and resulted in poor yield of amplified RNA, and this step was subsequently omitted. Therefore discrepancies between our results and those of Baugh and colleagues[11] may be related be the use of T4gp32 or could be explained by differences in RNA source and sample handling. We also tested the performance of the Nugen Ovation Biotin RNA amplification kit, which relies on a RNA-based single primer isothermal amplification methodology. However the performance of this methodology in terms of sensitivity, noise, reproducibility, and fidelity was poorer than that of other protocols tested here (results not shown).

Once the AB protocol was determined to be the most suitable one for amplification of total RNA from peripheral blood mononuclear cells, it was important to determine the degree to which biases introduced by RNA amplification would affect our ability to consistently detect differences between two different RNA samples. As long as the biases introduced are reproducible, systematic, and not too severe, amplification should be acceptable for most gene expression surveys. Our comparison of expression profiles of normal and t-MDS cells with and without amplification with the AB protocol showed that differences between the two samples detected with larger amounts of starting RNA and a single round of IVT were primarily maintained after amplification. Although correlation between amplified and nonamplified samples was imperfect, the ability to detect genes differentially expressed between the two samples as determined by independent methods of analysis was well maintained after amplification.

In summary, we have shown that the four protocols evaluated here performed differently when applied to the amplification of small RNA samples from clinical specimens. Rather minor differences in methodology and materials introduced considerable variability in gene expression profiling results. In addition, too strong a focus on a very small number of genes picked from an array analysis could be unduly influenced by seemingly acceptable methodologies. It also seems likely that differences in protocols become more critical for clinical samples where the amount and quality of starting material is especially challenging.

Acknowledgments

We are grateful to Dr. J. Denis Heck and Kim Nguyen from the DNA Array Core Facility, University of California, Irvine, for microarray analysis and for helpful discussions.

References

1. Park DJ, Vuong PT, de Vos S, Douer D, Koeffler HP: Comparative analysis of genes regulated by PML/RAR alpha and PLZF/RAR alpha in response to retinoic acid using oligonucleotide arrays. Blood 2003, 102:3727–3736
2. Yagi T, Morimoto A, Eguchi M, Hibi S, Sako M, Ishii E, Mizutani S, Imashuku S, Ohki M, Ichikawa H: Identification of a gene expression signature associated with pediatric AML prognosis. Blood 2003, 102: 1849–1856
3. Rozovskaia T, Ravid-Amir O, Tillib S, Getz G, Feinstein E, Agrawal H, Nagler A, Rappaport EF, Issaeva I, Matsuo Y, Kees UR, Lapidot T, Lo Coco F, Foa R, Mazo A, Nakamura T, Croce CM, Cimino G, Domany E, Canaani E: Expression profiles of acute lymphoblastic and myeloblastic leukemias with ALL-1 rearrangements. Proc Natl Acad Sci USA 2003, 100:7853–7858
4. Terskikh AV, Miyamoto T, Chang C, Diatchenko L, Weissman IL: Gene expression analysis of purified hematopoietic stem cells and committed progenitors. Blood 2003, 102:94–101
5. Qian Z, Fernald AA, Godley LA, Larson RA, Le Beau MM: Expression profiling of CD34+ hematopoietic stem/progenitor cells reveals distinct subtypes of therapy-related acute myeloid leukemia. Proc Natl Acad Sci USA 2002, 99:14925–14930
6. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH: Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci USA 1990, 87:1663–1667
7. Mahadevappa M, Warrington JA: A high-density probe array sample preparation method using 10- to 100-fold fewer cells. Nat Biotechnol 1999, 17:1134–1136
8. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL: Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 1996, 14:1675–1680
9. Luo L, Salunga RC, Guo H, Bittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR, Erlander MG: Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat Med 1999, 5:117–122
10. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM: High-fidelity mRNA amplification for gene profiling. Nat Biotechnol 2000, 18:457–459
11. Baugh LR, Hill AA, Brown EL, Hunter CP: Quantitative analysis of mRNA amplification by in vitro transcription. Nucleic Acids Res 2001, 29:E29
12. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, Payette T, Pistone M, Stecker K, Zhang BM, Zhou YX, Varnholt H, Smith B, Gadd M, Chatfield E, Kessler J, Baer TM, Erlander MG, Sgroi DC: Gene expression profiles of human breast cancer progression. Proc Natl Acad Sci USA 2003, 100:5974–5979
13. Oshima Y, Ueda M, Yamashita Y, Choi YL, Ota J, Ueno S, Ohki R, Koinuma K, Wada T, Ozawa K, Fujimura A, Mano H: DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. Leukemia 2003, 17:1990–1997
14. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, Koeffler HP: Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. Blood 2002, 100: 3553–3560