

Research article

Open Access

Microarray analysis in clinical oncology: pre-clinical optimization using needle core biopsies from xenograft tumors

Elizabeth M Goley¹, Soni J Anderson¹, Cynthia Ménard¹, Eric Chuang¹, Xing Lü², Philip J Tofilon² and Kevin Camphausen*¹

Address: ¹Radiation Oncology Branch, National Cancer Institute, National Institutes of Health, DHHS, Bethesda, MD, USA and ²Molecular Radiation Therapeutics Branch, National Cancer Institute, National Institutes of Health, DHHS, Bethesda, MD, USA

Email: Elizabeth M Goley - goleye@mail.nih.gov; Soni J Anderson - andersso@mail.nih.gov; Cynthia Ménard - menardc@mail.nih.gov; Eric Chuang - chuange@mail.nih.gov; Xing Lü - lvxi@mail.nih.gov; Philip J Tofilon - tofilonp@mail.nih.gov; Kevin Camphausen* - camphauk@mail.nih.gov

* Corresponding author

Published: 19 May 2004

Received: 29 January 2004

BMC Cancer 2004, 4:20

Accepted: 19 May 2004

This article is available from: <http://www.biomedcentral.com/1471-2407/4/20>

© 2004 Goley et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: DNA microarray profiling performed on clinical tissue specimens can potentially provide significant information regarding human cancer biology. Biopsy cores, the typical source of human tumor tissue, however, generally provide very small amounts of RNA (0.3–15 µg). RNA amplification is a common method used to increase the amount of material available for hybridization experiments. Using human xenograft tissue, we sought to address the following three questions: 1) is amplified RNA representative of the original RNA profile? 2) what is the minimum amount of total RNA required to perform a representative amplification? 3) are the direct and indirect methods of labeling the hybridization probe equivalent?

Methods: Total RNA was extracted from human xenograft tissue and amplified using a linear amplification process. RNA was labeled and hybridized, and the resulting images yielded data that was extracted into two categories using the mAdb system: "all genes" and "outliers". Scatter plots were generated for each slide and Pearson Coefficients of correlation were obtained.

Results: Results show that the amplification of 5 µg of total RNA yields a Pearson Correlation Coefficient of 0.752 (N = 6,987 genes) between the amplified and total RNA samples. We subsequently determined that amplification of 0.5 µg of total RNA generated a similar Pearson Correlation Coefficient as compared to the corresponding original RNA sample. Similarly, sixty-nine percent of total RNA outliers were detected with 5 µg of amplified starting RNA, and 55% of outliers were detected with 0.5 µg of starting RNA. However, amplification of 0.05 µg of starting RNA resulted in a loss of fidelity (Pearson Coefficient 0.669 between amplified and original samples, 44% outlier concordance). In these studies the direct or indirect methods of probe labeling yielded similar results. Finally, we examined whether RNA obtained from needle core biopsies of human tumor xenografts, amplified and indirectly labeled, would generate representative array profiles compared to larger excisional biopsy material. In this analysis correlation coefficients were obtained ranging from 0.750–0.834 between U251 biopsy cores and excised tumors, and 0.812–0.846 between DU145 biopsy cores and excised tumors.

Conclusion: These data suggest that needle core biopsies can be used as reliable tissue samples for tumor microarray analysis after linear amplification and either indirect or direct labeling of the starting RNA.

Background

Recent studies suggest that DNA microarray profiling performed on clinical specimens may provide information directly applicable to cancer diagnosis and treatment. One application of microarray analysis is aimed at differentiating subgroups of cancers using gene expression profiling, also referred to as class discovery [1-4]. Golub et al. used gene expression profiling of patient leukemia cells to distinguish acute myeloid leukemia (AML) from acute lymphoblastic leukemia (ALL) [1]. Moreover, they showed that subsets of leukemia cells that morphologically appeared to be ALL had a gene expression profile and response to therapy that was more consistent with AML, thus a new class of leukemia was described. Microarray analysis was also used to identify a subset of ALL tumors with a distinct gene expression profile that respond poorly to standard therapy [4]. Subgroup profiles have also been developed for other histologically homogeneous tumors such as diffuse B-cell lymphomas and hereditary breast cancers [2,3]. A diffuse large B-cell lymphoma (DLBCL) tumor cohort was divided into subsets of tumors with distinct gene expression profiles that correlated with overall survival [2]. Hedenfalk et al. compared gene expression profiles of breast tumors from women with and without BRCA1 or BRCA2 mutations. They showed that these two classes of breast tumors displayed different gene expression profiles and created a BRCA1/BRCA2 subclass. Such studies aimed at delineating the gene expression profiles of subtypes of tumors that exist within a purportedly homogeneous tumor population may not only aid in cancer diagnosis, but may also provide novel insight into the genetic mechanisms of oncogenesis.

In addition to cancer diagnosis, gene profiling is being explored as a means of predicting tumor treatment response, a long sought after goal of clinical oncology. Towards this end, a number of studies have related tumor gene expression profiles to treatment outcome and response to a given cytotoxic therapy, a process termed class prediction. Tumor gene expression profiles were generated for a series of patients with esophageal cancer treated with surgery and adjuvant chemotherapy resulting in the identification of an expression profile that correlated with longer survival and possibly tumor chemosensitivity [5]. Similarly, clinical outcomes and gene expression profiles were compared in subsets of patients with longer survival in B-cell lymphomas, AML and breast cancer, which produced gene expression profile that correlated with tumor response in each of these tumor types [6-8].

Thus, initial reports have suggested that incorporating gene expression profiling into clinical trials may provide novel information relevant to both cancer diagnosis and treatment. However, to fully investigate the potential clin-

ical applicability/value of microarray analysis will necessitate the performance of large prospective clinical trials. Such trials will confront a number of confounding variables including the uniform collection and preparation of RNA [9], the stability of the reference and experimental samples over a prolonged period [10] and microarray quality control over time [11]. However, the most significant impediment to performing these studies is likely to be the small size of the tumor biopsy. In a recent study of fifty-five breast biopsies obtained using a 14-gauge needle, the median recovery was only 1.34 μg of RNA (range 100 ng–12.6 μg) [12]. Moreover, Assersohn et al. in their analysis of breast fine needle aspirates reported a mean recovery of 202,500 cells, which would correspond to approximately 100 ng of RNA [13]. Because obtaining multiple biopsies from the same patient will likely be the exception, in most clinical trials the majority of gene expression profiles will have to be generated from single core biopsies that are likely to yield these small amounts of starting RNA material. Thus, in an attempt to optimize and validate procedures for tumor sample sizes relevant to the clinical setting, we have performed a series of microarray based gene expression analyses on core biopsies from human tumor xenograft models. These studies included the evaluation of the representative nature of amplified RNA compared with the original RNA sample, a comparison of direct and in-direct methods of probe labeling and the determination of the minimum starting RNA material needed to perform a valid and representative amplification.

Methods

Tumor xenografts

Three human tumor cell lines were used in this study: U251, a glioblastoma cell line, and two prostate carcinoma cell lines DU145 and LNCAP. Each was obtained from ATCC (Gaithersburg, MD). Xenografts were maintained in SCID mice (Jackson Labs, NH) with U251 and DU145 implanted subcutaneously (sc) and LNCAP consisting of an orthotopic tumor in the prostate gland. Animals were sacrificed and their tumors, average 350 mm³ in size, were immediately biopsied with a 14-gauge (2.1 mm diameter) semi-automatic core biopsy needle (MRI Devices Corp., Germany). Several biopsies were obtained in non-necrotic areas of the tumor. Core biopsies were snap frozen in liquid nitrogen or preserved in RNAlater (Qiagen.)

RNA amplification

Total RNA was isolated using TRIzol[®] reagent (Invitrogen) and purified with RNeasy[®] mini kits (Qiagen) according to the manufacturer's instructions. RNA samples and Universal Human Reference RNA (Stratagene) were amplified one or two rounds using RiboAmp[®] RNA Amplification Kits (Arcturus) per manufacturer's instructions. Starting

RNA quantity varied by experiment, between 0.8 and 10 μg of whole tumor samples. The reference RNA for these experiments was amplified using 10 μg aliquots of Universal Reference. All amplified samples were purified with an RNeasy[®] Kit according to the manufacturer's instructions and samples assessed for purity by agarose gel electrophoresis, and spectrophotometry was used to determine concentration.

Probe preparation

For direct labeling ten micrograms of amplified universal human reference were labeled with cyanine 5-dUTP (Cy5) and 5 μg of amplified U251 RNA samples were labeled with cyanine 3-dUTP (Cy3) using SUPER-SCRIPT II and Oligo (dT) 12–18 (Invitrogen). The method described by Khan et al., was followed [14]. Labeled probes were purified using Micro Bio-Spin 6 Chromatography Columns (Bio-Rad Laboratories) and purified using Microcon-30 spin columns (Millipore, three 400 μl TE washes used). The final elution was taken up to 17 μl with TRIS EDTA.

For indirect labeling of total RNA, 20 μg of both sample and reference were diluted into a total of 12 μl of DEPC water. For amplified RNA, starting material was reduced to 3 μg of both sample and reference, also diluted into a total of 12 μl of DEPC water per sample. For total RNA, 1 μl of 500 ng/ μl oligonucleotide d(T)12–18 (Stratagene Fairplay Kit) was added to each tube for cDNA priming. Amplified samples had 1 μl of 3 $\mu\text{g}/\mu\text{l}$ random primer (Invitrogen) added to each reaction tube. Samples were incubated at 70°C for 10 minutes, and cooled on ice. Each sample then had the following components added: 2 μl of 10 \times StrataScript reaction buffer (Stratagene Fairplay Kit), 1 μl of 20 \times dNTP mix, 1.5 μl of 0.1 M DTT, and 0.5 μl Rnase Block. Subsequent to mixing, 1 μl of 50 U/ μl StrataScript RT was added to each tube and tubes were incubated at 48°C for 25 minutes. Another aliquot of 1 μl StrataScript RT was added to each tube, and tubes were incubated for an additional 35 minutes. The resulting cDNA was purified using a MINElute Kit (Qiagen). Samples were next coupled to either 111 μg of monofunctional dye for total RNA-derived cDNA, or 55 μg of monofunctional dye for amplified RNA-derived cDNA. Reference RNA was always labeled with Cy5; tumor samples were labeled with Cy3.

Microarray hybridization

Microarray Slides were obtained from the Radiation Oncology Sciences Program Microarray Lab at the National Institutes of Health [15]. Slides were 8 k human slides printed on site using a Named Genes clone set from Research Genetics (Huntsville, AL), spotted onto poly-L-lysine coated slides using an OmniGrid arrayer (GeneMachines, San Carlos, CA). Slides were pre-hybridized for at least one hour at 42°C with 40 μl of pre-hybridization

solution consisting of 5 \times SSC, 0.1% SDS, and 1% BSA. Solution was loaded under M Series Lifterslips (Erie Scientific). Pre-hybridization solution was washed off by rapidly plunging the slides in distilled water for 2 minutes, followed by 100% isopropanol for 2 minutes. Slides were allowed to air dry prior to sample hybridization. Cy3 and Cy5 labeled targets were combined together for hybridization after dye-coupled cDNA purification. 1 μl of human COT-1 DNA (Invitrogen) and 1 μl of pd(A)40–60 (Amersham Biosciences) was added to each tube. Targets were denatured at 100°C for 1 minute before snap cooling on ice. 20 μl of pre-warmed (42°C) 2 \times F-Hybridization Buffer (50% formamide, 10 \times SSC, 0.2% SDS) was added to each sample. The combined target/hybridization solution mixture was incubated at 42°C for one minute, mixed and loaded onto microarray slides. Humidity was maintained in each chamber through the addition of 20 μl of DEPC water. Slides were hybridized at 42°C overnight. Post-hybridization washing included: 5 minutes in 2 \times SSC + 0.1% SDS, 5 minutes in 1 \times SSC, 5 minutes in 0.2 \times SSC, and finally 1 minute in 0.05 \times SSC. Slides were dried in a centrifuge set for 25°C at 650 rpm for 3 minutes. Slides were scanned at 10 microns using a Genepix[®] 4000 scanner (Axon Instruments), and images and data were stored in a database (mAdb) maintained by the Center for Information Technology, National Institutes of Health.

Data analysis

Data was extracted into two categories using the mAdb system: "all genes" and "outliers". "All genes" were extracted excluding spots flagged as Bad/Not found, and spots with target diameters less than 50 μm or greater than 300 μm . "Outliers" were defined as spots with a signal to background ratio ≥ 2 , and an overall signal $\geq 1,000$, genes required values in 100% of arrays, and the expression ratio was ≥ 2 or ≤ 0.5 . Spots were included if either channel was $\geq 2,500$ but the other criteria were unmet. Target pixels were 1 SD above the background $\geq 80\%$. After extraction, scatter plots were generated for each slide and Pearson Coefficients of correlation were obtained.

Results

Amplified versus total RNA

Expanding clinical applications of cDNA microarrays necessitate the ability to perform arrays with relatively minute starting amounts of RNA. The most commonly used method to increase the amount of RNA from a clinical specimen to a level suitable for analysis is through RNA amplification [16]. However, whether the gene expression profile generated from amplified RNA is representative of that obtained from total RNA has been the subject of some controversy [17]. To address this issue in our system, two replicates of amplified RNA and unamplified RNA, derived from sc U251 tumors, were

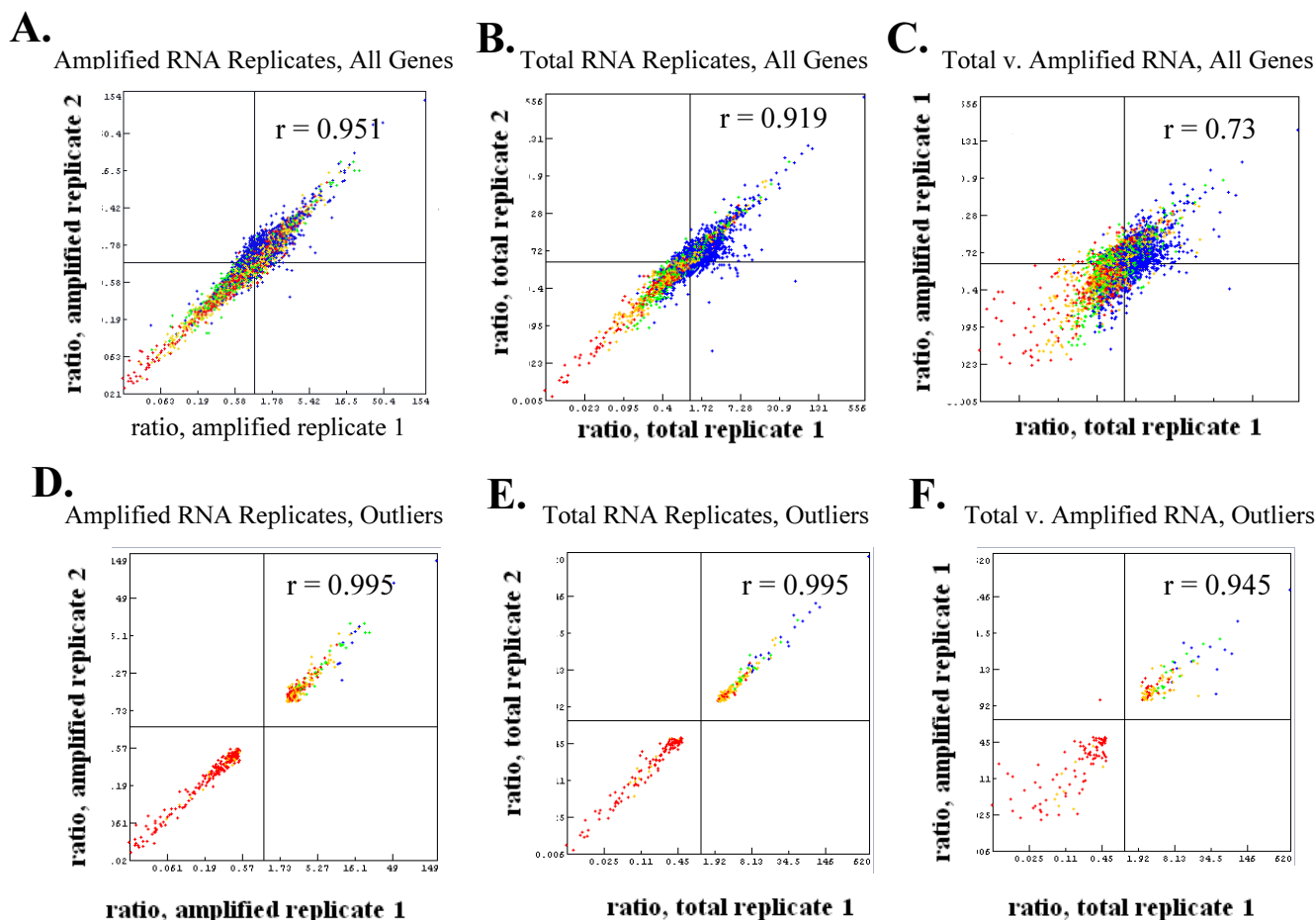


Figure 1
Correlations using amplified and total RNA isolated from U251 tumor. (A) Correlations between gene expression ratios for all genes using amplified RNA replicates 0.951 ($n = 6,841$ genes) and (B) total RNA replicates 0.919 ($n = 6,981$ genes). (D) and (E) represent scatter plots and correlations obtained for amplified 0.995 ($n = 456$ genes) and total replicates 0.995 ($n = 274$ genes), respectively, for outlier gene expression. (C) and (F) Representative correlation between total and amplified RNA from one sample 0.73 for all genes ($n = 6,987$ genes) and 0.945 for outliers ($n = 203$ genes). Extraction methods for all genes and outliers are found in the methods.

hybridized against amplified reference and un-amplified reference RNA, respectively. Scatter plots comparing the total gene expression level of the 2 amplified samples to each other and the 2 original RNA samples to each other (figures 1a and 1b) resulted in high correlation coefficients indicative of the reproducibility of the microarray platform used in these studies. Earlier publications have determined that a threshold of 0.70 for the concordance value is a marker of reproducibility [13,16]. Higher correlations within groups were obtained when just the outliers were compared (Figures 1d and 1e), consistent with prior studies [16]. Comparison of the original RNA samples to the amplified samples from in vivo biopsy cores yielded similar correlation coefficients, and once again, the outlier

genes have a higher coefficient than the all gene group (figure 1c and 1f). The outlier genes obtained for the original RNA and amplified RNA samples were also compared. Replicates for each group were averaged and the number of outliers in each determined (data not shown). In the original RNA sample 263 outliers were detected, 67% of those were also detected within the amplified RNA sample. Therefore, the amplification of the RNA material from a U251 tumor biopsy core maintained the majority of outlier genes found in the un-amplified sample.

RNA detection limits

Although a high concordance existed between our amplified and non-amplified specimens, the amount of starting

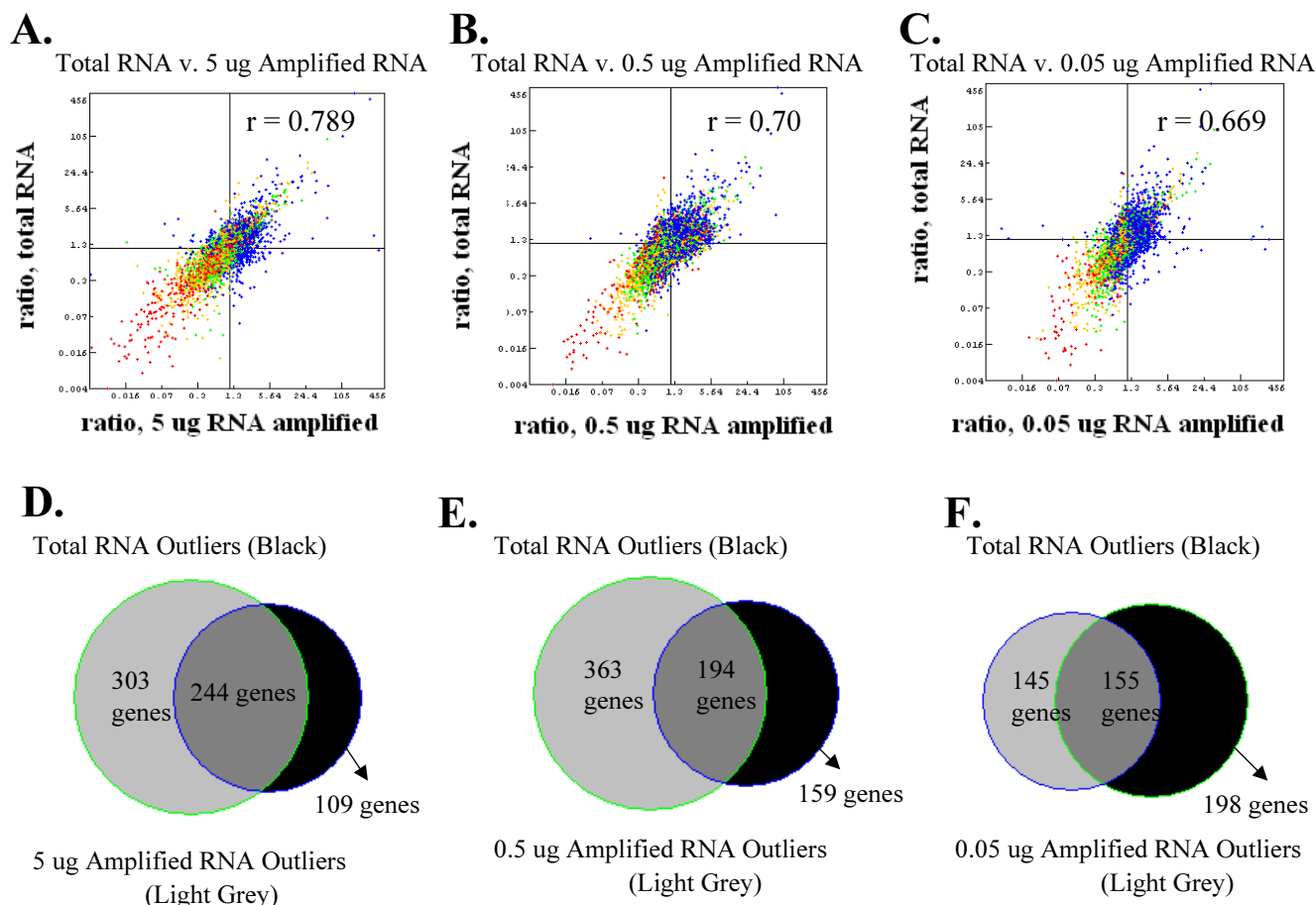


Figure 2

RNA detection limits. (A) Total RNA versus 5 μ g amplified RNA yielded a coefficient of 0.789 after two rounds of amplification ($n = 6,096$ genes) (B) The correlation for all genes between total RNA and a serially diluted 0.5 μ g aliquot that was removed from the 5 μ g sample and amplified 2 rounds yielded a coefficient of 0.7 ($n = 6,102$ genes). (C) The correlation between gene expression ratios for all genes between total RNA and a 0.05 μ g aliquot that has been serially diluted from the 0.5 μ g aliquot yielded a coefficient of 0.669 ($n = 6,114$ genes). (D-F) Venn diagrams depicting outlier concordance between the samples featured in (A-C, respectively). Extraction methods for all genes and outliers are found in the methods.

RNA material for these experiments was greater than what might be expected from a core biopsy from a patient's tumor. We amplified 5 μ g of RNA, whereas typical biopsy core specimens obtained in our clinic using a 14-gauge needle yield 0.3–3 μ g of RNA (unpublished data). To determine the amount of RNA starting material needed to amplify and maintain a high concordance with the original biopsy specimen, serial dilutions (5 μ g, 0.5 μ g, and 0.05 μ g) of LnCaP tumor biopsy material was used for amplification. Orthotopic LnCaP tumors were used in this study to validate that the high correlation seen in the earlier experiments were not specific to sc xenograft tumors. To control for the potential influence of varying the reference to sample RNA ratio, reference RNA was correspond-

ingly serially diluted to 5 μ g, 0.5 μ g and 0.05 μ g of starting material. All samples were then amplified 2 rounds to generate enough material for hybridization. The comparison of 5 μ g of un-amplified RNA versus amplified RNA of this orthotopic tumor model showed similar coefficients to the ones derived from the sc U251 model suggesting that the site of tumor growth does not influence the concordance values (figure 2a). However, the resulting Pearson Coefficients indicate a decreasing similarity to the original specimen as the starting amounts of material are reduced (figure 2a/2b/2c). We also examined outlier concordance by determining the percentage of original RNA outliers detected in the 5 μ g, 0.5 μ g, and 0.05 μ g of amplified samples. Three hundred and fifty three genes were

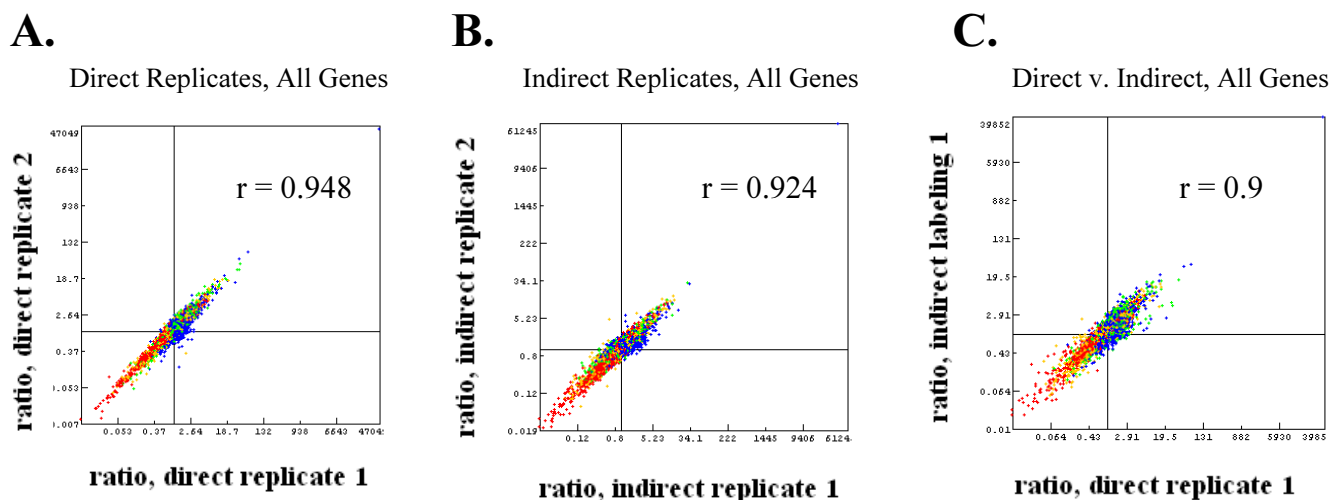


Figure 3
Scatter plot images and corresponding Pearson Coefficients between direct and indirect labeling methods. (A) Scatter plot of direct labeling replicates yielding a Pearson Coefficient of 0.948 for all genes ($n = 7,476$ genes). (B) Scatter plot of indirect labeling replicates yielding a Pearson Coefficient of 0.924 for all genes ($n = 7,474$). (C) A representative image of a comparison of the gene expression ratios between a direct and indirectly labeled slide. The average Pearson Coefficient generated comparing labeling methods was 0.895 for all genes (range 0.883–0.911, $n = 7,474$ genes). Extraction method for all genes is found in the methods.

identified as outliers for 5 μg of un-amplified RNA, while, 547, 557 and 300 genes were outliers in the 5 μg , 0.5 μg , and 0.05 μg amplified specimens, respectively (figure 2d/2e/2f). The overlap in outliers between the 5 μg of original RNA and each amplified specimen was 244 (69%), 194 (55%), 155 (44%) genes. Although there was concordance between the 5 μg amplified and the original specimens for outlier identification, a 44% concordance for the outliers in the amplified 0.05 μg sample compared to the 5 μg original sample is below the 0.70 concordance value cited in the literature as a marker of reproducibility [13].

Direct versus indirect labeling methods

An additional parameter that can determine the amount of starting RNA material needed is the method of fluorophore labeling. The two methods used to fluorescently label cDNA probes are direct labeling and indirect (amino-allyl) labeling. As the indirect method require less starting material and less expensive labeling dye, it was important to determine whether the data generated from these labeling methods was interchangeable. Two replicates of amplified RNA from the U251 sc tumor sample and reference RNA were labeled via each method, microarray analysis performed and scatter plots generated. Pearson Coefficients were generated to determine similarity. For direct labeling, 5 μg of experimental sample and 10 μg of reference was used and for the indirect method 3 μg of both the experimental sample and reference was used. The

reproducibility of the hybridization results for both the direct or indirect method of probe preparation using this microarray platform was indistinguishable for the all gene groups (figure 3a/3b). Similarly, the outlier coefficients from each labeling method were extracted and the directly labeled replicates yielded a Pearson Coefficient of 0.992 for outliers ($n = 647$), and indirectly labeled replicates yielded a Pearson Coefficient of 0.99 for outliers ($n = 588$) (data not shown). When directly comparing the method of probe preparation (figure 3, representative image), a high correlation was observed in the hybridization results. Therefore, while using less amplified RNA starting material, the indirect method of probe preparation yields similar results to directly labeled samples.

Tumor needle core biopsies versus excised tumor

After determining the limits of RNA amplification and the optimal method of probe labeling we simulated a patient biopsy procedure. Three core specimens were taken from a xenograft tumor, representing the patient biopsies, and the remainder of the tumor was harvested to represent the entire gene expression profile. Samples were obtained from two tumor types, DU145 and U251, both grown in the flank of SCID mice. The DU145 samples were flash frozen in liquid nitrogen, whereas the U251 samples were stored in RNAlater. As the use of liquid nitrogen can be cumbersome in the operating room, the reagent RNAlater has been used to preserve RNA specimens at room

Table 1: Pearson Coefficients between cores and corresponding tumor samples

Sample	DUI145 (LN2) Whole Tumor All Genes	DUI145 (LN2) Whole Tumor Outliers	U251 (later) Whole Tumor All Genes	U251 (later) Whole Tumor Outliers
Core 1	0.827	0.979	0.824	0.950
Average	N = 6,785 genes	N = 436 genes	N = 6,880 genes	N = 381 genes
Core 2	0.846	0.979	0.750	0.972
Average	N = 6,883 genes	N = 416 genes	N = 6,786 genes	N = 396 genes
Core 3	0.812	0.970	0.834	0.973
Average	N = 6,772 genes	N = 340 genes	N = 6,840 genes	N = 389 genes

"All Genes" were extracted with the following parameters: Spots flagged as Bad/Not Found were excluded, and spots with target diameters outside of 50 µm and 300 µm were excluded. "Outliers" were extracted with the following parameters: Sgl/Bkg ratio >= 2, signal >= 1,000, override if either channel >= 2,500, target pixels 1 SD about Bkg >= 80%, target diameter btw 50 µm and 300 µm, genes required values in 100% of arrays, and ratio >= 2 in >= 2 arrays (applied symmetrically)

temperature for up to 24 hours. Therefore, we included specimens stored in RNAlater to determine whether the specimens maintained fidelity. All core samples, range of original RNA 0.8 µg to 5 µg, were amplified and 3 µg of amplified RNA was indirectly labeled and hybridized against amplified human reference. Individual cores served as biologic replicates; 2 excisional tumor biopsies were obtained and each was analyzed in duplicate.

Comparison of the DUI145 excisional biopsy cores generated average coefficients of 0.90 for all genes (n = 6,711 genes), and 0.99 for outliers (n = 363 genes)(data not shown). Similarly, U251 excisional biopsies yielded an average coefficient of 0.90 for all genes (n = 7,078 genes) and 0.98 for outliers (n = 494 genes). The similarity between the coefficients obtained for the U251 and DUI145 hybridizations suggests that storage of samples in RNAlater reagent does not influence the results obtained from microarray hybridization. The similarity between the individual cores and their corresponding excisional biopsy samples for both DUI145 and U251 was then assessed according to Pearson Coefficients (Table 1). For each tumor the individual cores had correlations that ranged from 0.75 to 0.85 when compared to the whole tumor specimen for all genes and 0.95–0.97 for the outliers. Therefore, in these *in vivo* model systems the gene expression profile generated from needle core biopsies correlated to the gene expression profile of the whole tumor.

Discussion

As microarray technology enters mainstream usage in clinical oncology for class discovery and class prediction, a major confounder in the generation of gene expression profiles is likely to be the necessity of using small tissue samples (<1 µg). To address this issue, we investigated xenograft tumors to evaluate the similarity between amplified and non-amplified specimens, the minimal

amount of starting material for RNA amplification that retains fidelity, and the method of fluorophore probe labeling. To our knowledge, this is the only study reported to date evaluating all three of these variables in one xenograft experiment. We then simulated a patient needle core biopsy to determine whether the gene expression profile generated from such a biopsy is representative of the tumor.

A number of publications on the fidelity of amplified RNA compared to total RNA generally conclude that the amplified RNA does represent the total RNA sample particularly when comparing the outlier gene pattern [18-25] As noted by Nygaard et al., however, gene expression ratios are not always fully preserved [18]. In these publications, various starting material amounts, methods of analysis, and methods of amplification varied greatly. Feldman et al. recently demonstrated the advantages of mRNA amplification for microarray analysis, however, in their study cultured murine tumor cell lines were used [16]. The data presented here is consistent with those obtained from *in vitro* murine cell lines in that it illustrates that total RNA from human xenografts can be purified, amplified and hybridized using commercially available kits to result in similar gene expression profiles for amplified RNA versus total RNA [16].

After validating that amplified RNA and total RNA preparations yielded similar results for all genes as well as the outliers, we determined the amount of starting material needed to begin the amplification process and retain fidelity. Nygaard et al. recently published that 0.2 µg of total RNA amplified 2 rounds yields an average correlation between amplified and non-amplified arrays ranging from 0.71 to 0.84 [18]. Similarly, Wang started with total RNA in the 0.25–3.0 µg range and demonstrated no affect on the fidelity or reproducibility of amplification compared to the total RNA samples [26]. In each of these

studies the starting material was from cell culture. The results presented here using starting material from human xenograft models agrees with their findings: 0.5 µg of total RNA amplified 2 rounds has a Pearson Coefficient of 0.7 when comparing the amplified RNA and total RNA specimens. Because the amount of material generally obtained from needle core biopsies is typically between 0.3 and 5 µg these results suggest that the amount of amplified RNA generated from patient core biopsies can provide sufficient material for microarray analysis. However, these data also suggest that as the amount of total RNA starting material decreases the coefficient of similarity to the original specimen also decreases, which may contribute to inconsistencies in data interpretation.

The two common labeling methods for microarray analysis are direct and indirect (amino-allyl) labeling [27]. Direct labeling, which directly incorporates labeled nucleotides into cDNA during reverse transcription, is biased towards the incorporation of Cy3 dye. This may be due to either the inability of the reverse transcriptase (RT) enzyme to efficiently incorporate bulky Cy-labeled dUTP or the Cy labeled dye may be insoluble under RT conditions. Indirect labeling overcomes this bias by incorporating the less-bulky amino-allyl modified nucleotides into cDNA followed by the coupling of Cy3 and Cy5 to the amino-allyl groups. Previous comparisons of the direct and indirect labeling methods have shown similar hybridization results, however, they have used total RNA derived from cell culture or large amounts of normal tissue starting material [28-30]. We have expanded upon this data by verifying in a human xenograft model that the results of hybridization comparisons between the direct and indirect labeling methods produced similar results and can therefore be directly compared. As the indirect method uses less amplified starting material to produce a probe, more hybridizations can be performed than with the direct labeling method.

Having established that amplified RNA labeled indirectly from a human xenograft model could be used for microarray analysis, it was then necessary to establish that core biopsies would generate representative expression array profiles compared to excisional biopsy material. Sotiriou et al. using human tumor xenografts obtained a coefficient of 0.87 between two cores from the same tumor using amplified RNA and coefficients of 0.77 and 0.78 between excisional biopsy total RNA and resultant core biopsy amplified RNA [31]. Assersohn et al., also using a human xenograft model, found a correlation of 0.76 between cores and 0.69 between cores and whole tumor extracts [13]. Our findings for the coefficients between cores and the whole tumor were in agreement with these studies (range 0.75–0.85).

As the most common method of biopsy storage is to flash freeze the specimen in liquid nitrogen, the majority of our specimens were prepared in this manner. However, the use of liquid nitrogen can be cumbersome to use in the operating room. Recently, the reagent *RNAlater* has been used to preserve RNA specimens at room temperature for up to 24 hours. To evaluate this storage reagent, U251 tumors were biopsied and stored in both liquid nitrogen and *RNAlater*. Consistent with prior studies, similar coefficients were derived using either method of storage. Thus it appears that snap freezing with liquid nitrogen or the more convenient use of *RNAlater* is equivalent.

Conclusions

The data presented in this paper demonstrate that very small amounts of human xenograft tissue (as low as 0.5 µg) can be amplified generating results that faithfully represent the corresponding total RNA samples. Fifty-five percent of total RNA outliers were detected with this amount of starting material. Because typical core biopsies yield between 0.3–5 µg of starting material, they should provide sufficient product for microarray experiments after amplification has occurred. We have also demonstrated that human xenograft core biopsies ranging from 0.5–5 µg of starting total RNA material yield Pearson Coefficients between 0.750–0.846 when the amplified core material is compared to the whole tumor. Taken together, these experiments demonstrate that core biopsies after linear amplification and either indirect or direct labeling can reliably be used for clinical oncology microarray studies.

Competing interests

None declared.

Authors' contributions

EG aided in study design, obtained and assisted in analyzing the data, and drafted the manuscript. SA aided in generating and analyzing data, and critically reviewed the manuscript. CM provided specimens and expertise in clinical study design, and critically reviewed the manuscript. EC manufactured all slides used in the study, provided expertise in the microarray process, and critically reviewed the manuscript. XL aided in generating and analyzing data. PT aided in data analysis and critically reviewed the manuscript. KC conceived of the study, assisted in data analysis, and aided in drafting and reviewing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the members of the ROSP Microarray Lab for producing the microarray slides used in this study and for providing assistance, and the Advanced Technology Center Microarray Facility for providing training and assistance.

References

- Golub T, Slonim D, Tamayo P, Huard C, Gassenbeek M, Mesirov JP, Coller H, Loh ML, Downing J, Caligiuri M, Bloomfield C, Lander E: **Molecular classification of cancer: class discovery and class prediction by gene expression monitoring.** *Science* 1999, **286**:531-537.
- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Grenier TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Bvd JC, Bostein D, Brown PO, Straudt LM: **Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling.** *Nature* 2000, **403**:503-511.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J: **Gene-expression profiles in hereditary breast cancer.** *New Engl J Med* 2001, **344**:539-548.
- Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR: **Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling.** *Cancer Cell* 2002, **2**:133-143.
- Kihara C, Tsunoda T, Tankaka T, Yamana H, Furukawa Y, Ono K, Kitahara O, Zembutsu H, Yanagawa R, Hirata K, Takagi T, Nakamura Y: **Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles.** *Cancer Res* 2001, **61**:6474-6479.
- Rosenwald A, Wright G, Leroy K, Yu X, Gaulard P, Gascoyne RD, Chan WC, Zhao T, Haioun C, Greiner TC, Weisenburger DD, Lynch JC, Vose J, Armitage JO, Smeland EB, Kvaloy S, Holte H, Delabie J, Campo E, Montserrat E, Lopez-Guillermo A, Ott G, Muller-Hermelink HK, Connors JM, Braziel R, Grogan TM, Fisher RI, Miller TP, LeBlanc M, Chiorazzi M, Zhao H, Yang L, Powell J, Wilson WH, Jaffe ES, Simon R, Klausner RD, Staudt LM: **Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma.** *J Exp Med* 2003, **198**:851-862.
- Okutsu J, Tsunoda T, Kaneta Y, Katagiri T, Kitahara O, Zembutsu H, Yanagawa R, Miyawaki S, Kuriyama K, Kubota N, Kimura Y, Kubo K, Yagasaki F, Higa T, Taguchi H, Tobita T, Akiyama H, Takeshita A, Wang YH, Motoji T, Ohno R, Nakamura Y: **Prediction of chemosensitivity for patients with acute myeloid leukemia, according to expression levels of 28 genes selected by genome-wide complementary DNA microarray analysis.** *Mol Cancer Ther* 2002, **1**:1035-1042.
- Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC, O'Connell P: **Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer.** *Lancet* 2003, **362**:362-369.
- Florell SR, Coffin CM, Holden JA, Zimmermann JW, Gerwels JW, Summers BK, Jones DA, Leachman SA: **Preservation of RNA for Functional Genomic Studies: A Multidisciplinary Tumor Bank Protocol.** *Mod Pathol* 2001, **14**:116-128.
- Bertucci F, Viens P, Tagett R, Nguyen C, Houlgatte R, Birnbaum D: **DNA Arrays in Clinical Oncology: Promises and Challenges.** *Laboratory Invest* 2003, **83**:305-316.
- Nimgaonkar A, Sanoudou D, Butte AJ, Haslett JN, Kunkel LM, Beggs AH, Kohane IS: **Reproducibility of gene expression across generations of Affymetrix microarrays.** *BMC Bioinformatics* 2003, **4**:27.
- Ellis M, Davis N, Coop A, Liu M, Schumaker L, Lee RY, Srikanthana R, Russell CG, Singh B, Miller W, Stearns V, Pannanen M, Tsangaris T, Gallagher A, Liu A, Zwart A, Hayes DF, Lippman ME, Wang Y, Clarke R: **Development and Validation of a Method for Using Breast Core Needle Biopsies for Gene Expression Microarray Analyses.** *Clin Cancer Res* 2002, **8**:1155-1166.
- Assersohn L, Gangi L, Zhao Y, Dowsett M, Simon R, Powles TJ, Liu ET: **The Feasibility of Using Fine Needle Aspiration from Primary Breast Cancers for cDNA Microarray Analyses.** *Clin Cancer Res* 2002, **8**:794-801.
- Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM, Meltzer PS: **Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays.** *Cancer Res* 1998, **58**:5009-5013.
- Chuang YY, Chen Y, Gadiseti C, Chandramouli VR, Cook JA, Coffin D, Tsai MH, DeGraff W, Yan H, Zhao S, Russo A, Liu ET, Mitchell JB: **Gene Expression after Treatment with Hydrogen Peroxidase, Menadione, or t-Butyl Hydroperoxidase in Breast Cancer Cells.** *Cancer Res* 2002, **62**:6246-6254.
- Feldman AL, Costouros NG, Wang E, Qian M, Marincola FM, Alexander HR, Libutti SK: **Advantages of mRNA Amplification for Microarray Analysis.** *BioTechniques* 2002, **33**:906-914.
- Baugh LR, Hill AA, Brown EL, Hunter CP: **Quantitative analysis of mRNA amplification by in vitro transcription.** *Nucleic Acids Res* 2001, **29**:1-9.
- Nygaard V, Loland A, Holden M, Langaas M, Rue H, Liu F, Myklebost O, Fodstad O, Hovig E, Smith-Sorensen B: **Effects of mRNA amplification on gene expression ratios in cDNA experiments by analysis of variance.** *BMC Genomics* 2003, **4**:11.
- Iscove NN, Barbara M, Gu M, Gibson M, Modi C, Winegarden N: **Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA.** *Nat Biotechnol* 2002, **20**.
- Hu L, Wang J, Beggerly K, Wang H, Fuller GN, Hamilton SR, Coombes KR, Zhang W: **Obtaining reliable information from minute amounts of RNA using cDNA microarrays.** *BMC Genomics* 2002, **3**:16.
- Puskas LG, Zvara A, Hackler L, Van Hummelen P: **RNA Amplification Results in Reproducible Microarray Data with Slight Ratio Bias.** *BioTechniques* 2002, **32**:1330-1340.
- 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* 1990, **87**:1663-1667.
- Zhao H, Hastie T, Whitfield ML, Borresen-Dale AL, Jeffrey SS: **Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis.** *BMC Genomics* 2002, **3**:31.
- Bertucci F, Bernard K, Loriod B, Chang YC, Granjeaud S, Birnbaum D, Nguyen C, Peck K, Jordan BR: **Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples.** *Hum Mol Genet* 1999, **8**:1715-1722.
- Polacek DC, Passerini AG, Shi C, Francesco NM, Manduchi E, Grant GR, Powell S, Bischof H, Winkler H, Stoeckert CJ, Davis PF: **Fidelity and enhanced sensitivity of differential transcription profiles following linear amplification of nanogram amounts of endothelial mRNA.** *Physiol Genomics* 2002, **13**:147-156.
- Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM: **High-fidelity mRNA amplification for gene profiling.** *Nat Biotechnol* 2000, **18**:457-9.
- Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J: **A Concise Guide to cDNA Microarray Analysis.** *BioTechniques* 2000, **29**:548-562.
- Manduchi E, Scarce M, Brestelli JE, Grant GR, Kaestner KH, Stoeckert CJ: **Comparison of different labeling methods for two-channel high-density microarray experiments.** *Physiol Genomics* 2002, **10**:169-179.
- Yu J, Othman MI, Farjo R, Zarepari S, MacNee SP, Yoshida S, Swaroop A: **Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays.** *Mol Vis* 2002, **8**:130-137.
- Richter A, Schwager C, Hentze S, Ansoorge W, Hentze MW, Muckenthaler M: **Comparison of Fluorescent Tag DNA Labeling Methods Used for Expression Analysis by DNA Microarrays.** *BioTechniques* 2002, **33**:620-630.
- Sotiriou C, Khanna C, Jazaeri AA, Petersen D, Liu ET: **Core Biopsies Can Be Used to Distinguish Differences in Expression Profiling by cDNA Microarrays.** *J Mol Diagn* 2002, **4**:30-36.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2407/4/20/prepub>