Reproducible and inexpensive probe preparation for oligonucleotide arrays

Yuzhi Zhang, Brendan D. Price, Sotirios Tetrads1, Subrata Chakrabarti, Gautam Maulik and G. Mike Makrigiorgos*

Department of Radiation Oncology, Dana Farber Cancer Institute, Harvard Medical School, Boston, MA, USA and 1University of California School of Dentistry, 10883 Le Conte Avenue, Los Angeles, CA 90095-1668, USA

Received March 1, 2001; Revised May 1, 2001; Accepted May 10, 2001

ABSTRACT

We present a new protocol for the preparation of nucleic acids for microarray hybridization. DNA is fragmented quantitatively and reproducibly by using a hydroxyl radical-based reaction, which is initiated by hydrogen peroxide, iron(II)-EDTA and ascorbic acid. Following fragmentation, the nucleic acid fragments are densely biotinylated using a biotinylated psoralen analog plus UVA light and hybridized on microarrays. This non-enzymatic protocol circumvents several practical difficulties associated with DNA preparation for microarrays: the lack of reproducible fragmentation patterns associated with enzymatic methods; the large amount of labeled nucleic acids required by some array designs, which is often combined with a limited amount of starting material; and the high cost associated with currently used biotinylation methods. The method is applicable to any form of nucleic acid, but is particularly useful when applying double-stranded DNA on oligonucleotide arrays. Validation of this protocol is demonstrated by hybridizing PCR products with oligonucleotide-coated microspheres and PCR amplified cDNA with Affymetrix Cancer GeneChip microarrays.

INTRODUCTION

RNA and DNA probes used for hybridization on oligonucleotide microarrays require fragmentation and labeling (e.g. biotinylation) (1–4). The reproducibility of this process is critical since it affects the overall reliability of microarray testing. In addition, currently recommended protocols for RNA/DNA labeling invariably require expensive reagents that can constitute a significant fraction of the microarray cost (10–30% depending on the amount of material required for hybridization; for example, to obtain the 15–20 µg of biotinylated fragmented target DNA required per hybridization, the labeling kit expense is ~32% of the discounted array price of $250.00). As the expense of purchasing and using microarrays decreases (4), improving the reliability and lowering the cost of the preparation of DNA for hybridization becomes even more important.

Most current methods for nucleic acid labeling utilize labeled deoxynucleotide triphosphates incorporated into cRNA or cDNA during PCR or in vitro synthesis reactions, prior to microarray hybridization (4–7). Often, due to limitations in the amount of starting material, cDNA has to be PCR amplified before use in microarray hybridization (6,8). For some widely used array designs (e.g. the Affymetrix Gene Expression GeneChips) fragmenting and labeling cDNA presents a challenge since the large amounts (15–20 µg) of biotinylated-fragmented target DNA required per hybridization raise significant practical difficulties (vide infra). On the other hand, unless the double-stranded target is efficiently fragmented and labeled, formation of undesirable secondary structures and rehybridization between complementary target sequences hinders sufficient signal generation (1–3). Application of double-stranded cDNA to Affymetrix Gene Expression GeneChips has not been previously reported.

Here we present a simple, highly reproducible and very cost effective protocol for fragmenting and biotinylating complex nucleic acid targets. The method: (i) utilizes a hydroxyl radical-based chemical reaction to reproducibly fragment the target nucleic acid; (ii) employs a biotinylated psoralen analog to generate densely labeled target fragments via photoactivation; (iii) can be used on any nucleic acid (single-stranded or double-stranded DNA or RNA); (iv) is easily scalable to large amounts of nucleic acid. The method is especially useful when double-stranded DNA is hybridized on oligonucleotide microarrays. We demonstrate application of this approach by screening double-stranded PCR products on oligonucleotide-coated microspheres and PCR amplified cDNA on Affymetrix Gene Expression G110 Cancer Arrays.

MATERIALS AND METHODS

Preparation of cDNA and PCR products

To prepare double-stranded cDNA, mRNA from cultured SAOS-2 cells, a p53- cell line, was extracted using oligo(dT)-coated magnetic beads (Dynabeads mRNA Direct kit; Dynal, Lake Success, NY) and was used to synthesize cDNA (Universal Riboclone Synthesis System; Promega, Madison, WI). Double-stranded cDNA (200 ng) was then digested with...
1 U of the 4 bp cutter enzyme Sau3AI (recognition site 5’-GATC-3’; New England Biolabs, Beverly, MA) for 2 h at 37°C and purified with a QIAquick kit (Qiagen, Valencia, CA). Samples were resuspended in 60 µl of ligation buffer for ligation of asymmetric linkers corresponding to Sau3AI restriction sites. The linkers used were a 24mer (5’-AGCCTTCCAGCCTTCCACGCA-3’) and a 12mer (5’-AGTCGCGTGTA-3’). The DNA sample was mixed with the linkers, annealed at 50°C and then cooled to 10°C. An aliquot of 3 µl of T4 DNA ligase (400 U/µl; New England Biolabs) was added and the mixture incubated overnight at 15°C. Amplification was carried out in a Perkin Elmer GeneAmp PCR 9600 system (PE Biosystems, Foster City, CA) as follows: addition of Advantage HF-2 polymerase (Clontech, Palo Alto, CA) and incubation for 5 min at 68°C, then 20 cycles of 1 min at 95°C and 3 min at 68°C, followed by a final extension for 10 min at 68°C. PCR products were purified first on QIAquick columns and then by microbiospin-6 column filtration (Bio-Rad, Hercules, CA) to remove the remaining traces of ethanol eluting from the QIAquick columns. This procedure, used in representational difference analysis (9–11) and in comparative genomic hybridization (12), yields amplified double-stranded cDNA fairly representative of the original material.

To obtain evidence of whether the current procedure results in microarray data adequately representative of the gene expression levels in cdna, the p53- cdNA was ‘spiked’ with graded amounts of a p53-containing plasmid. The 7.1 kb plasmid included the full-length p53 sequence (~1.8 kb) and was prepared as described (13,14).

A 236 bp PCR product was also amplified from the same plasmid and used as a test system with the present protocol. For this, 5’-ACT CAA GGA TGC CCA GGC TG-3’ forward and 5’-CCT ATT GCA AGC AAG GGT TC-3’ reverse primers were used; PCR included 95°C for 1 min, then 25 cycle PCR of 95°C for 30 s and 68°C for 1 min using Advantage-HF2 polymerase (Clontech), followed by 5 min final extention. In some experiments the 236 bp PCR product was end-biotinylated using a biotinylated primer in the PCR reaction. Ultrapure calf thymus genomic DNA was obtained from Sigma (St Louis, MO). The lyophilized DNA was suspended in phosphate buffer overnight and then extensively dialyzed against 0.1 M phosphate buffer, pH 7.0, to remove traces of Tris EDTA that are present in the commercial product (Tris is an inhibitor of the chemical fragmentation as it scavenges hydroxyl radicals).

DNA fragmentation and psoralen-biotinylation

The PCR amplified cDNA, with or without spiked p53, was fragmented using a hydroxyl radical producing reaction, which utilizes hydrogen peroxide, ascorbic acid and iron-EDTA complex (15–18). The protocol described in the hydroxyl radical footprinting method (16,17) was followed, with minor modifications, with chemical reagents (all from Sigma) freshly made up before each experiment. Briefly, 0.4 mM Fe(II) was prepared by dissolving ferrous ammonium sulfate [(NH₄)₂Fe(SO₄)₂·6H₂O] in triply-distilled water. Immediately afterwards, equal volumes of 0.4 mM Fe(II) and 0.8 mM EDTA were mixed to prepare the complex of Fe(II) with EDTA. The reaction with DNA was carried out in a 10-fold dilution of phosphate-buffered saline, pH 7.0 (0.1× PBS).

Depending on the experiment, the final reaction mixture consisted of a total volume of 40 or 80 µl, including 10 µM Fe(II)-EDTA complex, 0.03–0.06% (8.8–17.6 mM) hydrogen peroxide, 1–2 mM sodium ascorbate, and 1 or 10 µg DNA for fragmentation in 0.1× PBS. The cleavage reaction was initiated by rapid mixing and vortexing of all three reagents with DNA in a 1.5 ml Eppendorf tube and continued for 10–30 min at room temperature. The reaction was then stopped by adding thiourea to 10 mM and EDTA to 2 mM. Finally, the fragmented DNA was purified by microbiospin-6 column filtration (Bio-Rad). The same protocol was used to fragment calf thymus genomic DNA as well as the 236 bp PCR products.

To biotinylate the fragmented DNA, a biotinylated psoralen analog (EZ-Link psoralen-PEO-biotin, catalog no. 29986VC; Pierce Chemical Co., Rockford, IL) was used. DNA was first denatured at 95°C for 3 min, immediately placed on ice and psoralen-PEO-biotin was added to 200 µM final concentration. The mixture was irradiated on ice, using long wavelength UVA light (XL-1500 UV Crosslinker; Fischer Scientific) for 30 min. Following crosslinking, the DNA was ethanol precipitated, resuspended in Tris-EDTA and quantitated using spectrophotometry. The fragmentation products were separated by 1.5% agarose gel electrophoresis and stained with ethidium bromide. Experiments were repeated three to five times. This protocol has been shown to lead to a dense biotinylation pattern (1 biotin per 20 bases) and can be used with both DNA and RNA (19,20).

As an alternative, enzymatic digestion was applied to fragment DNA. Aliquots of 10⁻²–10⁻⁴ µ of microcococal nuclease (Sigma) were used to digest cDNA (15 min at 37°C in 5 mM Tris–HCl, 10 mM NaCl, 3 mM CaCl₂, pH 8.2). The reaction was stopped by addition of 5 mM EDTA. This DNA was then biotinylated using psoralen-PEO-biotin, as described above.

Hybridization to oligonucleotide-coated microspheres

To examine whether DNA fragmentation and biotinylation using the present protocol results in DNA that can still be successfully hybridized and detected on oligonucleotide microarrays, hybridization of the 236mer PCR product to oligonucleotide-coated microspheres was first tested as a rudimentary ‘single array element’. Conjugation of a 25mer oligonucleotide, complementary to a central region of the 236mer, to carboxylated microspheres was done by a minor modification of the Luminex protocol (21,22). Briefly, 5 × 10⁶ unlabeled, carboxylated microspheres, 5.5 µm diameter (Molecular Probes, Eugene, OR), were vortexed into a 1.5 ml microcentrifuge tube, dispersed by sonication for 30 s, centrifuged at 8000 g for 1 min and, after supernatant removal, adjusted to 0.1 M 2-(N-morpholino)ethanesulfonic acid, pH 4.5. An aliquot of 1 nmol amino-modified oligonucleotide (5’-amino-ggt-CAG-TCT-ACC-TCC-Cgt-CAT-AAA-A) was added, followed by 2.5 µl of fresh 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide-HCl (Pierce Chemical) from a 10 mg/ml stock. After 1 h at room temperature the microspheres were repeatedly washed by centrifugation and stored at 4°C in the dark.

To hybridize the biotinylated 236mer PCR fragment to microspheres, 0–25 ng DNA was suspended in Tris-EDTA (TE, pH 7.0), heat denatured at 96°C, added to 10 000 microspheres and adjusted to tetramethylammonium chloride hybridization buffer, 50 µl total volume, for 15–60 min at
50°C. Following washing by centrifugation, 12 μl of 10 μg/ml fluorescent streptavidin-Alexa-488 (Molecular Probes) was added to the microspheres for 15 min at room temperature. Samples were then analyzed with a 488 nm flow cytometer at the Dana Farber Flow Cytometry Core Facility. Alternatively, instead of biotinylating the fragmented 236mer via psoralen, the end-biotinylated 236mer was used unfragmented for microsphere hybridization.

**Hybridization of fragmented-biotinylated cDNA to Affymetrix Cancer Arrays**

To test the reliability and reproducibility of the current fragmentation-biotinylation protocol for screening cDNA on oligonucleotide microarrays, the Affymetrix recommended procedure for application of fragmented-biotinylated cRNA to Cancer Arrays was adopted. The cDNA was applied to microarrays (18–20 μg DNA per microarray) and screened at Research Genetics (Huntsville, AL).

**RESULTS AND DISCUSSION**

**Fragmentation using the hydroxyl radical reaction**

Hydroxyl radical produces random, non-specific fragmentation of DNA, which allows the generation of DNA fragments of the size appropriate for hybridization to microarrays (23). Fragmentation of genomic calf thymus DNA by the hydroxyl radical reaction is demonstrated in Figure 1A, for two time points. It is shown that 10 min fragmentation converted 10 μg high molecular weight DNA to fragments mainly ranging from 50 to 200 bp long. The reproducibility of the DNA fragmentation pattern using the present approach was tested by repeating the procedure in more than 10 independent experiments. Fragmentation of successive PCR amplified cDNA samples as well as the 236 bp PCR product (Fig. 1B and C) using this hydroxyl radical-based method was reproducible, giving the same range (50–200 bp) of fragments in independent experiments. In contrast, micrococcal nuclease digestion gave inconsistent results (data not shown). Similar problems have been reported for DNase I (24).

**Flow cytometry following hybridization to microspheres**

Because hydroxyl radical produces DNA base modifications as well as strand breaks (25), the ability of the fragmented DNA to hybridize with complementary sequences was first tested in a simpler system, prior to using the microarrays. Hydroxyl radical-induced fragmentation of DNA prior to psoralen-biotinylation allowed specific hybridization of DNA to oligonucleotide-coated microspheres (Fig. 2, curves 1 and 3). The non-fragmented, end-labeled 236mer, on the other hand, resulted in roughly similar signals (Fig. 2, curve 2). All data were normalized to the value obtained in curve 1, for 50 ng DNA, which was taken to be equal to 100.

Sample fragmentation is a general requirement when larger DNA (e.g. cDNA) is hybridized on oligonucleotide microarrays (2). The data in Figure 2 indicate that the hydroxyl radical reaction adequately serves the purpose. Hybridization of DNA to oligonucleotide-coated, optically encoded microspheres has been proposed as a high throughput method to detect DNA sequences and mutations (21,22,26,27). Therefore, the current protocol may also potentially be used with oligonucleotide-coated arrays of optically encoded microspheres.

**Hybridization to Affymetrix Cancer Arrays**

Figure 3A depicts the spectrum of fluorescent signal intensities obtained when two identically treated, but independent, cDNA samples were fragmented and biotinylated with the current protocol prior to application to Affymetrix Cancer microarrays. Using the company supplied software, 98.4% of the 2044 genes screened were diagnosed as presenting ‘no change’ among the two identical samples, while 1.6% (33 genes) presented significant or marginal decreases/increases in expression levels. The consistency of the signals was similar to that reported when using cRNA on GeneChip microarrays (1–3). Figure 3B depicts the relative signal intensities obtained when the p53–cDNA population was spiked with p53. The signals obtained were proportional to the spiked p53 cDNA (0.2, 0.02,
GeneChip Arrays. (A) For example Pierce EZ-Link psoralen-PEO-biotin, 5 mg for $150.00, 1 d. (B) For example the Affymetrix labeling kit used for fragmentation and labeling of cDNA, 6 PCR reactions for $120.00, b. (C) For example NEN Life Sciences biotin-dCTP, 6 PCR reactions for $120.00, c. (D) For example Gibco Life Technologies BioPrime Kit, 30 reactions for $225.00, with 0.5 µg DNACost of fragmentation and labeling per GeneChip. Table 1. Method Cost of fragmentation and labeling per µg DNA Cost of fragmentation and labeling per GeneChip* Random primer + biotin-dNTP $15.00* $600.00 Biotin-dCTP during PCR, then nuclease $50.00* $2000.00 TdT reaction with fluorescent dideoxynucleotide $25.00* $1000.00 Hydroxyl radical + biotinylated psoralen $0.50* $20.00 (present protocol)

*It was assumed that a total of 40 µg double-stranded cDNA needs to be fragmented and biotinylated followed by purification in order to obtain the final 15–20 µg sample required per array.

1For example Gibco Life Technologies BioPrime Kit, 30 reactions for $225.00, with 0.5 µg labeled product per reaction.

2For example NEN Life Sciences biotin-dCTP, 6 PCR reactions for $120.00, –0.4 µg DNA product following amplification of 1 ng DNA template for 25 cycles.

3For example the Affymetrix labeling kit used for fragmentation and labeling of 1–3 µg PCR product for the p53 genotyping arrays; $620.00 for 25 labeling reactions.

4For example Pierce EZ-Link psoralen-PEO-biotin, 5 mg for $150.00, 1 µg labeled DNA per reaction, 280 reactions total.

The data indicate that by following the present protocol, specific hybridization occurs in a reasonably quantitative manner. To test the signal of a typical housekeeping gene (human integrin) with over-digestion. The decrease in signal was also associated with an ~30% increase in background (not shown). The signal background is a result of autofluoresence of the array surface and non-specific binding of target or stain molecules (Affymetrix GeneChip Expression Analysis Manual, Affymetrix, Santa Clara, CA). As fragmentation increases, the number of nucleic acid molecules able to hybridize specifically decreases and the proportion of very small (<10 bp) nucleic acid molecules which undergo non-specific hybridization increases.

Figure 3D demonstrates microarray signals from two identical cDNA samples fragmented enzymatically, at different times, using micrococcal nuclease instead of the hydroxyl radical reaction. Following fragmentation, these samples were labeled using biotinylated psoralen and screened on two arrays. In contrast to the hydroxyl radical fragmentation data obtained in Figure 3A, analysis of the data in Figure 3D demonstrated that only 78% of the 2044 genes screened were diagnosed as presenting 'no change' among the two identical samples when enzymatic fragmentation was adopted. Taken together, the data in Figure 3C and D indicate that changes in DNA fragmentation or poor reproducibility can affect the ability to detect changes in expression levels. Unlike RNA, which can be efficiently fragmented by boiling (1–3), DNA is relatively resistant to heat-induced strand breaks and has traditionally been fragmented enzymatically (28) or by ultrasonication (24). Enzymatic methods are poorly reproducible, while ultrasonication of DNA requires a powerful and expensive instrument with an ultrasound sterile tip able to be immersed in a 10–50 µl sample. The current protocol provides an improved and practical alternative.

Cost advantage of the current protocol

An advantage of the current DNA fragmentation-biotinylation protocol relative to other methods is cost. Oligonucleotide array hybridization requires application of several micrograms of fragmented-labeled CDNA or RNA. Due to the unavoidable material losses that occur during DNA/RNA labeling and purification procedures, typically double the initial amount of nucleic acid needs to be fragmented and labeled. To apply the recommended 15–20 µg DNA to the Affymetrix Gene Expression arrays, an initial amount of 40 µg DNA was fragmented and labeled here. Table 1 compares the cost per microgram of DNA and per array for preparation of DNA for hybridization.
with the current approach versus the random primer method (6,8,29), incorporation of biotinylated deoxynucleotides during synthesis (1–3) and the terminal deoxynucleotide labeling method (28) used for the Affymetrix genotyping arrays. For the popular Affymetrix Gene Expression arrays in particular, the cost of fragmentation and biotinylation with any of these methods would be exceedingly high. In contrast, a 1–2 order of magnitude cost reduction is achievable using the current protocol.

CONCLUSION

A reproducible, simple and low cost protocol has been developed to prepare DNA for hybridization to oligonucleotide microarrays. The protocol is especially useful when double-stranded DNA is hybridized on microarrays, but can also be applied to single-stranded DNA or RNA. This approach also circumvents problems associated with limited starting material and/or the inconvenience of handling RNA.

ACKNOWLEDGEMENTS

We are most grateful to Rhonda Snyder, Laura Carr and Donna Brown of the Genome Services Department, Research Genetics, for their helpful suggestions and technical assistance regarding the Affymetrix microarray screening. This research was supported in part by NIH grants CA 83234 and CA/HG 90422 to G.M.M.

REFERENCES