Suitability of Stratagene reference RNA for analysis of lymphoid tissues

Karen Dybkær¹, Guimei Zhou¹, Javeed Iqbal¹, David Kelly¹, Li Xiao¹, Simon Sherman¹, Francesco d'Amore², and Wing C. Chan¹

¹University of Nebraska Medical Center, Omaha, NE, USA and ²University Hospital of Aarhus, Aarhus, Denmark

BioTechniques 37:470-474 (September 2004)

We evaluated a lymphoid RNA standard prepared in our laboratory for spotted microarrays against the Universal Human Reference standard from Stratagene. Our goal was to determine if the Stratagene standard, which contains only two lymphoid cell lines out of a pool of 10 human cancer cell lines, had acceptable gene coverage to serve as a comprehensive standard for gene expression profiling of lymphoid tissues. Our lymphoid standard was prepared from thymus, spleen, tonsil, and cell lines representing immature B cells, plasma cells, and natural killer (NK) cells, thus covering the entire spectrum of lymphoid cells and most stromal elements present in specialized lymphoid tissues. The two standards were co-hybridized on oligonucleotide microarrays containing 17,260 genes, and both had fluorescence intensities above background for approximately 85% of the genes. Despite the limited representation of lymphoid cells in the Stratagene standard, only 4.2% genes exhibited expression differences greater than 2-fold including only 0.35% with differences greater than 4-fold. Although the lymphoid standard reflected a more comprehensive representation of immune system-associated genes, the Stratagene standard has the advantage of being commercially available, enabling easier comparison across laboratories and allowing comparative studies across a long period of time.

INTRODUCTION

Spotted DNA microarrays printed with cDNA or oligonucleotides typically require co-hybridization of a sample and a reference standard, so relative expression level for each individual gene can be readily normalized and related across multiple experiments (1–3).

The ideal reference standard should be abundant and easy to obtain, so that extracted RNA could be aliquoted, stored, and used through an extended period of time. It should have expression of all genes with a relative abundance similar to those expected in the sample of interest. This is obtainable when comparing gene expression levels between two conditions, where only a single parameter separates the otherwise identical samples making the untreated state reference (4). When investigating different aspects of a certain system including distinct cell types and tissues, it is more complicated, and a mixture of RNA from different cell lines is frequently used (5–7). Since we wanted to perform gene ex-

pression profiling on both benign and malignant conditions of the lymphoid system, it was important to utilize a reference RNA standard providing a comprehensive spectrum of transcripts at concentration levels suitable for evaluating gene expression changes in various lymphoid tissues and cells. The commercially available Universal Human Reference RNA from Stratagene (La Jolla, CA, USA) consists of equal quantities of total RNA from 10 different cancer cell lines representing brain, breast, skin, testis, cervix, liver, lipocytes, macrophage, but only one lymphoblastic leukemia cell line representing T cells and one plasmacytoma cell line representing B cells. Thus, we were concerned that this standard would not provide sufficient coverage for our microarray analysis of the lymphoid system. Therefore, we empirically evaluated the coverage and expression levels of lymphoid system expressed genes of the Stratagene RNA standard and a RNA standard prepared in our laboratory consisting of a mixture of RNA isolated from thymus, sec-

ondary lymphoid tissues, and a number of lymphoid cell lines.

MATERIALS AND METHODS

Fabrication of Human Oligonucleotide Microarrays

Oligonucleotide probes (60-mers) were designed for each target gene (Compugen, Rockville, MD, USA) and manufactured by Sigma-Genosys (The Woodlands, TX, USA). Oligonucleotides were resuspended (30 μ M) in 3 $\times$ standard saline citrate (SSC) and spotted by the DNA Microarray Facility at the Eppley Cancer Center, University of Nebraska Medical Center (UNMC) onto poly-L-lysine-coated slides using a MagnaSpotter robot (BioAutomation, Dallas, TX, USA) with a 12-pin print head (Telechem, Sunnyvale, CA, USA). After spotting, DNA was cross-linked to the slides by UV irradiation (350 mJ/cm²), blocked by succinic anhydride treatment, and rinsed in ethanol. A total of 18,705 oligonucleotides were spotted, representing 17,260 different genes.

Lymphoid RNA Standard

A lymphoid reference RNA was prepared from human cell lines and lymphoid tissues. The cell lines U-266-Myeloma (representing plasma cells), Nalm-6 (representing pre-B cells), and NK-92 and SNK-6 [both representing natural killer (NK) cells] were harvested, and total RNA was extracted with TRIzol[®] (Invitrogen, Carlsbad, CA, USA) and purified with RNeasy[®] Mini Columns (Qiagen, Valencia, CA, USA). Normal thymus specimens, ruptured spleens, and reactive tonsils were obtained from the UNMC and Children's Hospital, with Institutional Review Board (IRB) approval. Tissue was quickly transported on ice, cut into small pieces in a Petri dish in cold RPMI-1640, and the cells were liberated by a glass tissue homogenizer (Bellco, Vineland, NJ, USA) prior to RNA extraction. Thymus contained immature T cells, tonsil and spleen contained mainly mature B and T cells, and all tissues contained stromal elements. None of the individual RNA fractions

exceeded 30% of the total RNA pool. The fractions were pooled, aliquoted, and stored as ethanol precipitates at -80°C .

Probe Labeling

Indirect labeling using aminoallyl dye coupling was performed as described by DeRisi with slight modifications (<http://derisilab.ucsf.edu/microarray/protocols.html>). Briefly, first strand cDNA was synthesized from 30 μg total RNA and labeled with aminoallyl-dUTP, and the cDNA products were purified on Microcon[®]30 columns and coupled with N-hydroxysuccinimidyl ester CyTM5 or Cy3 dyes (Amersham Biosciences, Piscataway, NJ, USA). In initial assays, labeling efficiency and total amount of labeled cDNA was determined, and reproducibility was ensured by measuring the absorption at 532 nm (Cy3), 635 nm (Cy5), and 260 nm (cDNA), respectively. Typically 1.6–2.0 μg cDNA was generated from 30 μg total RNA after labeling and purification and was used in its entirety for hybridization.

Array Hybridization and Scanning

Samples were mixed in 6 μL deionized water and 10 μg each of poly(dA), human cot-1 DNA, and yeast transfer RNA (tRNA) was added to a final volume of 45 μL containing 4.4 $\times$ SSC, 4.1 $\times$ Denharts, and 50% formamide. Hybridization was performed in the dark for 15 h at 42°C in a Corning hybridization chamber (Corning, Acton, MA, USA). Arrays were washed once in 2 $\times$ SSC, 0.1% sodium dodecyl sulfate (SDS), 3 min in 2 $\times$ SSC, 0.1% SDS, 3 min in 1 $\times$ SSC, 0.01% SDS, and 2 min in 0.2 $\times$ SSC, all at room temperature with stirring in foil-covered wash chambers (TeleChem, Sunnyvale, CA, USA). Arrays were dried by spinning and read with a ScanArrayTM 4000 scanner (PerkinElmer Life and Analytical Sciences, Boston, MA, USA) and images analyzed using Genepix 4.0 software (Axon Instruments, Union City, CA, USA).

Analysis

Triplicate co-hybridizations were performed with lymphoid (Cy5) and Stratagene (Cy3) standards. Features

Figure 1. Self-self hybridizations of Stratagene and lymphoid RNA standards. (A) Self-self hybridization of the Stratagene RNA standard with a Pearson correlation coefficient of 0.98. (B) Self-self hybridization of the lymphoid RNA standard with a Pearson correlation coefficient of 0.96. (C) Correlation plot of two of the three repeated hybridizations with a Pearson correlation coefficient of 0.92.

obscured by blemishes or dust-specks were flagged, and features with signal/background ratios below 1 were removed prior to analysis. Local background was subtracted from individual spots, and intensities for each channel

were global median normalized. The C5/Cy3 ratio was $\log_2$ transformed, and the weighted average was calculated for each gene from the triplicate hybridizations according to the formula

$$\mu = \Sigma(x_i / \delta_i^2) / \Sigma(1 / \delta_i^2)$$

where each data point x_i is weighted inversely by its own variance δ_i^2 . The standard deviation of the weighted average was expressed as $\delta_\mu = \sqrt{1 / \Sigma(1 / \delta_i^2)}$ (8). The weighted average and standard deviation were only calculated for genes where ratios between lymphoid and Stratagene RNA standard were unidirectional for at least two out of the three hybridizations.

RESULTS AND DISCUSSION

Genomic DNA was suggested (1) and experimentally evaluated (2,3) as a universal reference in microarray studies. Pooled RNA reference was superior to a genomic DNA reference when evaluating transcriptional response to oxidative stress. Differences in hybridization kinetics between double-stranded genomic DNA and a RNA-based reference was suggested to mediate the increased variability observed for the genomic DNA reference when hybridized to cDNA microarrays (3). In contrast, others find no statistical difference between the two standards when directly compared by co-hybridization onto spotted cDNA microarrays (2). Since the use of genomic DNA as reference is still controversial, we chose to evaluate only RNA references consisting of pools of RNA.

Evaluation of Oligonucleotide Microarray

To ensure no bias was introduced during incorporation of either fluorophores used for labeling, self-self hybridizations of Stratagene and lymphoid RNA standards (Figure 1, A and B) were performed. Fluorescence intensities for both channels were very similar with Pearson correlation coefficients (r) above 0.95, illustrating an unbiased incorporation of Cy3 and Cy5. Reproducibility was evaluated by plotting Cy5/Cy3 ratios of individual genes from two out of the three independent hybridizations of Stratagene versus lymphoid standards against each other (Figure 1C). Repeats showed Pearson correlation coefficients around 0.90. The ratio of Cy5/Cy3 for the repetitive spots of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was statis-

Figure 2. Differentially expressed genes with 2- to 28-fold differences in expression level between the lymphoid and Stratagene standard. The genes are organized in functional groups, and their expression levels are illustrated by color. Red color is used for genes with the higher expression levels in the lymphoid standard than the Stratagene standard. Green color is assigned to genes with the higher expressing values in the Stratagene standard versus the lymphoid standard.

tically comparable by Student's *t*-test statistics for unpaired data between repeated microarrays. Thus, unbiased labeling and a high degree of reproducibility of hybridization were demonstrated.

Differentially Expressed Genes

With 85% of the printed spots yielding fluorescence intensities 2-fold above the average empty-spot intensities for both standards, more than 15,000 different genes/expressed sequence tags (ESTs) were available for analysis. Using a 2-fold difference as determinant for differentially expressed genes, we obtained 796 genes or 4.2%. Of these, 92% were between 2- to 4-fold differentially expressed, 6% were between 4- and 8-fold, and 2% were differentially expressed by more than 8-fold. The highest observed difference was 28-fold higher expression in the lymphoid standard for a cDNA clone (gi:10437228 entrez). Seventy-five percent of the differentially expressed genes had higher expression levels in the lymphoid than in the Stratagene standard.

Differentially expressed genes could be subdivided into categories that correlated well with the cellular composition of the two RNA standards (Figure 2). Transcripts encoding leukocyte-specific surface antigens and DNA repair genes involved in T cell receptor and immunoglobulin gene rearrangements were highly present in the lymphoid standard. Stromal components of tonsil and spleen and several genes involved in regulating apoptosis had enhanced expression in the lymphoid standard.

For the Stratagene standard, many genes involved in cell cycle and cell proliferation were present at higher transcript levels than in the lymphoid standard, consistent with higher proliferative rates of cell lines than normal cells and tissues. A large group of genes with obvious embryonic-restricted expression patterns were overexpressed in the Stratagene standard, including the γ globins, G and A (9). These transcripts probably originated from the embryonic carcinoma cell line. Testis and brain specific transcripts were also present at higher levels in the Stratagene standard reflecting the origin of the cell lines of

the standard and hence the differences in RNA concentrations.

In our study, most of the genes were expressed at modest levels in both standards, making them fairly comparable. With 597 transcripts being more highly expressed in the lymphoid versus the Stratagene standard, the former has an advantage when working within the lymphoid system. However, only 47 of the transcripts are differentially expressed by more than 4-fold in the lymphoid standard and only 19 of these have a known biological function. In conclusion, the Stratagene standard performed very well in our experiments, and it is commercially available so that experiments by different laboratories using this standard are more readily compared. Thus, the advantages of the Stratagene standard may outweigh the minor deficiencies, and it will provide adequate coverage for the vast majority of transcripts.

ACKNOWLEDGMENTS

This work was supported in part by National Cancer Institute (NCI) grant no. CA84967, Department of Health and Human Services. K.D. is a fellow of the Danish Cancer Society.

COMPETING INTERESTS STATEMENT

The authors declare no conflicts of interest.

REFERENCES

1. Eisen, M.B. and P.O. Brown. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* 303:179-205.
2. Weil, M.R., T. Macatee, and H.R. Garner. 2002. Toward a universal standard: comparing two methods for standardizing spotted microarray data. *BioTechniques* 32:1310-1314.
3. Kim, H., B. Zhao, E.C. Snedrud, B.J. Haas, C.D. Town, and J. Quackenbush. 2002. Use of RNA and genomic DNA references for inferred comparisons in DNA microarray analyses. *BioTechniques* 33:924-930.
4. Cook, K.L. and G.S. Saylor. 2003. Environmental application of array technology: promise, problems and practicalities. *Curr. Opin. Biotechnol.* 14:311-318.
5. Alizadeh, A.A., M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, et al. 2000. Distinct types of diffuse large B-cell lympho-

ma identified by gene expression profiling. *Nature* 403:503-511.

6. Rosenwald, A., G. Wright, A. Wiestner, W.C. Chan, J.M. Connors, E. Campo, et al. 2003. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell* 3:185-197.
7. Ross, D.T., U. Scherf, M.B. Eisen, C.M. Perou, C. Rees, P. Spellman, et al. 2000. Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* 24:227-235.
8. Loos, A., C. Glanemann, L.B. Willis, X.M. O'Brien, P.A. Lessard, R. Gerstmeier, et al. 2001. Development and validation of corynebacterium DNA microarrays. *Appl. Environ. Microbiol.* 67:2310-2318.
9. Papayannopoulou, T., S. Kurachi, B. Nakamoto, E.D. Zanjani, and G. Stamatoyannopoulos. 1982. Hemoglobin switching in culture: evidence for a humoral factor that induces switching in adult and neonatal but not fetal erythroid cells. *Proc. Natl. Acad. Sci. USA* 79:6579-6583.

Received 6 February 2004; accepted 2 June 2004.

Address correspondence to Wing C. Chan, Department of Pathology and Microbiology, 983135 Nebraska Medical Center, Omaha, NE 68198-3135, USA. e-mail: jchan@unmc.edu