JMD CME Program

*Journal of Molecular Diagnostics, Vol. 8, No. 5, November 2006*
*Copyright © American Society for Investigative Pathology*
*and the Association for Molecular Pathology*
*DOI: 10.2353/jmoldx.2006.060021*

# Technical Advance

# Effects of Globin mRNA Reduction Methods on Gene Expression Profiles from Whole Blood

Jinny Liu,* Elizabeth Walter,†‡ David Stenger,* and Dzung Thach* on behalf of Epidemic Outbreak Surveillance

*From the Center for Biomolecular Science and Engineering,* Naval Research Laboratory, Washington, DC; Texas A&M University Systems,† College Station, Texas; and the United States Air Force Surgeon General Headquarters,‡ Falls Church, Virginia*

**Excessive globin mRNA in whole blood RNA decreases transcript detection sensitivity and increases signal variation on microarrays. Hence, methods based on peptide nucleic acid inhibitory oligos and biotinylated DNA capture oligos have been developed to reduce globin mRNA. However, there is limited information about the effects of these two methods on gene expression profiles. Thus, we systematically compared the facility and effects of the two globin reduction methods on profile measurements from Jurkat cell line RNA with or without spiked globin mRNA and human blood RNA isolated using PAXgene collection tubes. We showed that the methods were efficient at increasing the sensitivity of transcript detection without loss of specificity, but neither method could recover a profile equivalent to that of an identical RNA sample without globin mRNA excesses. The capture oligo method had slightly better transcript detection sensitivity for cell line RNA, lowered signal variation for PAXgene RNA, and more similar profiles to controls than the inhibitory method. However, the capture method required larger amounts of initial high-quality RNA to yield sufficient cRNA amounts, and its procedures were more complex and time consuming than the inhibitory method. These results inform the selection of methods suitable for multicenter surveillance of gene expression profiles.** ***(J Mol Diagn 2006, 8:551–558; DOI: 10.2353/jmoldx.2006.060021)***

Gene expression profile studies of human blood samples in the context of epidemiological surveillance of immunological signatures are confronted by numerous challenges. Previously, we found that increases in hemoglobin from whole blood PAXgene samples of healthy and ill patients contributes to decreased sensitivity of detection of transcripts on Affymetrix microarrays,[1] and others have shown that spiking increasing amounts of globin mRNA into total RNA from cell lines results in decreasing detection of transcripts (Affymetrix technical note: Globin Reduction Protocol: A Method for Processing Whole Blood RNA Samples for Improved Array Results. Santa Clara, CA). To curtail this problem, two methods for globin mRNA reduction have been commercially developed. The method from Affymetrix uses the specific binding of peptide nucleic acid (PNA) oligos to the 3′ end of globin mRNA to inhibit reverse transcription during cDNA synthesis.[2] The method from Ambion, removes globin mRNA from total RNA via hybridization with biotinylated DNA oligos that specifically capture globin mRNA followed by binding with streptavidin magnetic beads.

Both methods seem to offer particular pros and cons depending on the intended application. The established target preparation process normally involves RNA isolation, concentration, cDNA synthesis, and *in vitro* transcription to generate amplified cRNA.[1,3] Theoretically, the PNA-based process seems to fit better into the established target preparation process because it is performed in the same tube as the cDNA synthesis reaction,

Supported in part by the Defense Threat Reduction Agency, Headquarters United States Air Force Surgeon General's Office, Office of Naval Research, and the Naval Research Laboratory.

The opinions and assertions contained herein are the private ones of the authors and are not to be construed as official or reflecting the views of the Department of Defense.

The Epidemic Outbreak Surveillance Consortium is an Air Force Medical Service initiative comprised of: Sponsorship: P. Demitry; Executive Board: R. Holliday; Operational Board: D. Stenger and E. Walter; Technical Advisors and Collaborators: B. Agan, L. Daum, D. Metzgar, D. Niemeyer, and K. Russell; Research and Clinical Staff: M. Archer, R. Bravo, N. Freed, J. Fuller, J. Gomez, K. Gratwick, M. Jenkins, M. Jesse, B. Johnson, E. Lawrence, B. Lin, C. Meador, H. Melgarejo, K. Mueller, C. Olsen, D. Pearson, F. Stotler, D. Thach, Z. Wang, D. Watson, S. Worthy, G. Vora, and C. Myers; Operations Support Staff: K. Grant and C. James.

Accepted for publication July 12, 2006.

Address reprint requests to Dzung Thach, Center for Biomolecular Science and Engineering, Code 6920, Naval Research Laboratory, 4555 Overlook Ave., SW, Bldg. 30, Washington, DC 20375. E-mail: dzung.thach@nrl.navy.mil.

whereas the DNA-based method is a subprotocol that is inserted after total RNA concentration and before cDNA synthesis, therefore taking more time. However, the stability of PNA in the long term is unknown and the method requires taking measures to prevent PNA aggregation and precipitation. Accordingly, the manufacturer recommends that a control RNA sample with spiked-in globin mRNA be run with each batch to ensure that the globin reduction process occurred. In contrast, DNA oligomeric chemistry is more established. Both methods have been shown by their manufacturers to increase detection of transcripts and reduce biological variation between samples. However, it is not known how these methods would transform a gene expression profile, how they compare in practicality and outcomes, or to what extent either method can generate a profile similar to that of samples without high levels of globin mRNA.

Therefore, we systematically compared the facility and effects of the two globin reduction methods on gene expression profile measurements from Jurkat cell line RNA with or without spiked globin mRNA and from whole blood RNA isolated from PAXgene tubes. Our results will enable decisions on which method is suitable for incorporation into large-scale multicenter surveillance of gene expression profiles.

## Materials and Methods

### Sample Collection

With approval of the institutional review board at Lackland Air Force Base (San Antonio, TX) and after informed consent, ~25 ml of blood, filling 10 PAXgene tubes (Preanalytix, Hombrechtikon, Switzerland), were drawn from one healthy volunteer. Blood was drawn into tubes by standard protocol (Preanalytix product circular: PAXgene Blood RNA Tube. Hombrechtikon, Switzerland). All tubes were maintained at room temperature for 2 hours, frozen at −20°C, and shipped on dry ice to the Naval Research Laboratory in Washington, DC, for processing. Tubes were stored at −80°C before RNA isolation.

### Sample Processing

Blood collection and RNA isolation was performed using the PAXgene blood RNA system (Qiagen, Valencia, CA), which consists of an evacuated tube for blood collection and a kit for isolation of total RNA from whole blood.[4] All of the Jurkat cell RNA with or without spiked-in globin mRNA were purchased from Affymetrix (Santa Clara, CA). The isolated RNA underwent further cleanup and concentration followed by globin reduction or no reduction control procedures and was amplified, labeled, and interrogated on the HG-U133 plus 2.0 GeneChip microarrays (Affymetrix).

### Total RNA Isolation from Blood

Frozen PAXgene tubes were thawed at room temperature for 2 hours followed by total RNA isolation as described in the PAXgene blood kit handbook (Preanalytix product insert: PAXgene Blood RNA Kit Handbook. Hombrechtikon, Switzerland) but modified by increasing proteinase K from 40 to 80 μl (>600 mAU/ml) per sample, extending the 55°C incubation time from 10 to 30 minutes, and passing through a QIAshredder spin column (Qiagen). The QIAshredder column shears genomic DNA, thus further facilitating tight pellet formation. The optional on-column DNase digestion was not performed. Purified total RNA was stored at −80°C.[1,3]

### Total RNA Cleanup and Concentration

For more complete removal of DNA from purified PAXgene RNA, replicate RNA samples were pooled followed by in-solution DNase treatment using the DNA-free kit (Ambion, Austin, TX) but without the addition of DNase inactivation reagent because DNase and other contaminants will be removed in the subsequent step. After DNase treatment, RNA was further purified and concentrated using the RNAeasy MinElute Cleanup kit (Qiagen). Subsequently, 1 μl from each sample was run on the Bioanalyzer 2100 (Agilent, Palo Alto, CA) for assessment of RNA quality, whereas the NanoDrop spectrophotometer (ND-1000) was used for quantitation. Usage of the bioanalyzer was analogous to capillary gel electrophoresis. This resulted in electropherograms displaying florescent intensity versus time, which correlates with the amount of RNA versus the size of RNA, respectively.

### Globin Reduction and Target Preparation

The effects of globin mRNA were alleviated by using the Globinclear (Ambion) or the GeneChip Globin Reduction (Affymetrix) kit, but with modifications. For the Globinclear procedure, biotinylated globin-capture DNA oligos were added to 5 μg of total RNA and globin mRNA were removed by streptavidin magnetic beads. Then the remaining globin-reduced total RNA was purified using magnetic beads and eluted in 30 μl of water. One μl of RNA was used for bioanalyzer measurement and the remaining RNA was concentrated to 8 μl using a Speed Vac (Thermo Electron Corporation, Waltham, MA) at room temperature. For the GeneChip globin reduction procedure, 5 μg of total RNA in 9 μl of BR5 from the RNAeasy MinElute Cleanup eluant was used for cDNA synthesis, whereas the column that came with the globin reduction kit was not used. PNA oligos for inhibition of reverse transcription of globin mRNA were purchased from Applied Biosystems, Foster City, CA. Control mRNA (Affymetrix) at known concentrations was spiked into the samples. These controls were from the *lys*, *phe*, *thr*, and *dap* genes of the prokaryote *Bacillus subtilis* and were used because such targets would be absent in eukaryotic samples. All subsequent steps were as described in the GeneChip Expression Analysis Technical Manual, version 701021 Rev. 3.

### Database Integration

Laboratory data contained information about the processing of samples from blood in PAXgene tubes to cRNA target preparation, as well as bioanalyzer and NanoDrop measurements. Electropherograms were analyzed by the Biosizing software expert software B.02.01 (Agilent) to output 28S/18S intensity ratios and RNA integrity number (RIN) quality control metrics, whereas the NanoDrop outputs RNA quantity and 260/280 ratios. Report files summarizing the quality of target detection for an array were generated by GeneChip operating software version 1.2 (Affymetrix). JMP (SAS, Cary, NC) was used to join these various data tables together into a metadata table. For gene expression data, signal values were calculated using the Microarray Suite 5.0 algorithm with and without scaling to test it effects on various downstream analytical methods.

### Statistical Analysis

Statistical quality control and relations among metadata variables and microarray results were analyzed in JMP. Analysis of variance and multidimensional scaling of gene expression data were performed in Arraytools 3.3.0 beta 3a developed by Richard Simon and Amy Lam (*http://linus.nci.nih.gov/BRB-ArrayTools.html*). Heat maps and dendrograms were graphed using dChip 1.3.[5,6]

## Results

### Quality of RNA, Globin Reduction, and Target Preparation

Three types of RNA samples were used to study the effects of globin reduction methods on gene expression profiles: Jurkat RNA isolated from Jurkat cell line (J), Jurkat RNA with globin mRNA spiked-in (JG), or PAXgene RNA from whole blood (B). Three techniques were tested: using biotinylated globin capture DNA oligos (D), using globin reverse transcription inhibitory PNA oligos (P), or no globin reduction treatment as technical control (C) (Figure 1).

The same lot of Jurkat and Jurkat-globin RNA was used throughout. After globin reduction with capture oligos, 77 to 95% of Jurkat and 80 to 91% of Jurkat-globin RNA was recovered (Table 1, second versus first row); the lowered recovery trend for the latter likely reflected globin mRNA removal. This lowered amount of RNA for cDNA synthesis probably accounted for the subsequently lowered cRNA yields for the capture oligo technique (Table 1, third row). Electropherograms of cRNA from Jurkat-globin RNA undergoing globin reduction via capture or inhibitory oligo technique resulted in smaller globin peaks compared with no reduction control (Figure 2A). They also became similar to traces of cRNA from Jurkat RNA without any spiked-in globin mRNA, which were similar regardless of technique (Figure 2A, only JC is shown). These results suggested that for high-quality RNA from a cell line the globin reduction techniques were

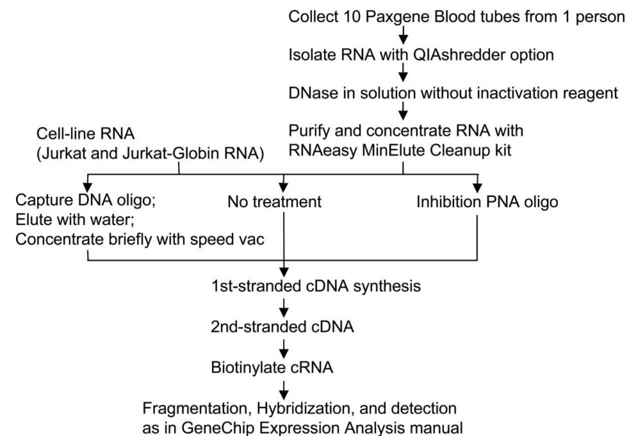


**Figure 1.** Diagram summarizing the experimental design, sample treatments, and important modifications. See the Materials and Methods section for corresponding details on sample collection and processing; total RNA isolation, cleanup, and concentration; and globin reduction and target preparation.

mostly efficient and did not change the cRNA population size distribution.

PAXgene RNA was pooled from tubes collected from one volunteer. The RIN value of RNA in the eluant of the MinElute column was 8.2 to 9.3. After globin reduction with capture oligos, 75 to 80% of PAXgene RNA was recovered (Table 1, second versus first row), but it now has RIN values of 7.2 to 7.6, suggesting slight RNA degradation. Electropherograms of cRNA from PAXgene RNA undergoing globin reduction via capture or inhibitory oligo technique resulted in smaller globin peaks compared with no reduction control (Figure 2B). However, the cRNA from the inhibitory oligo method showed smaller size distribution than the other methods. These results indicated that, for high-quality RNA, the globin reduction techniques were mostly efficient but could induce particular changes in RNA quality parameters such as RNA recovery amount, cRNA yield, RIN, and cRNA size distribution.

### Quality of Microarray Measurements

There was no difference in scale factors (SFs) of gene expression signal distributions among the cell line groups, suggesting that the majority of signals were unchanged among groups. Thus, scaled signal values were used throughout unless specified. Graphs of signals of the four spiked-in mRNA controls versus known concentrations showed the expected linear relationship. Background was highest in the Jurkat-globin RNA control group (Table 1). The ratios of signals from the 3′ to the 5′ region of *GAPDH* and *Actin* transcripts were all below 3, indicating good RNA quality and efficient *in vitro* transcription.

Similar results were observed for PAXgene RNA samples. Background was lower in the PAXgene RNA control group than in the Jurkat-globin control group, but the 3′/5′ *GAPDH* and *Actin* ratios showed a higher trend than the cell line RNA groups (Table 1), probably because of differences in sample type, collection, and handling. No

**Table 1.** Quality Metrics of RNA, Globin Reduction, Target Preparation, and Microarray Results among RNA Type and Technical Treatment

| | Jurkat RNA | | |
|---|---|---|---|
| Treatment | Capture DNA oligo | Inhibition PNA oligo | Control |
| Starting material (μg) | 4 | 4 | 4 |
| Yields after treatment | 3.56 ± 0.41* | 4 | 4 |
| Adjusted cRNA yield | 71.13 ± 5.41 | 96.4 ± 30.66 | 113.47 ± 40.77 |
| Scale factors | 4.50 ± 1.38 | 3.98 ± 0.62 | 4.42 ± 0.52 |
| Background | 64.21 ± 12.46 | 68.47 ± 11.30 | 60.91 ± 3.71 |
| 3′/5′ GAPDH | 1.06 ± 0.04 | 1.05 ± 0.03 | 1.09 ± 0.07 |
| 3′/5′ Actin | 1.33 ± 0.15 | 1.23 ± 0.06 | 1.31 ± 0.03 |
| Present calls (%) | 46.8 ± 1.18 | 45.5 ± 0.62 | 44.8 ± 1.65 |

Data in each column are from triplicates.
*STDEV.

(*table continues*)

other differences were detected. These results suggest that the microarray measurements were of sufficient quality for inclusion in further statistical analysis.

## *Effects of Globin Reduction on Microarray Detection of Transcripts*

Jurkat RNA showed the highest percentage of transcript detected, with no differences among globin reduction methods or controls [Table 1, present calls (%)]. For Jurkat-globin RNA, removal of globin mRNA using either the capture or the inhibitory method increased the percentage of transcripts detected by 9 or 5%, respectively, approaching those of Jurkat RNA alone. The capture oligo method resulted in more transcripts detected than the inhibitory oligo method (*t*-test, $P < 0.05$).

Table 2 shows the cross-tabulation of transcript detection calls between Jurkat-globin RNA with globin removal treatments versus Jurkat RNA alone as the true detection calls. Treatment of Jurkat-globin RNA with the capture or inhibitory oligo method increased concordant present calls by 8 and 5% and decreased false negative calls by 7 and 4% but increased false positive calls only by 1 and 0.22%, respectively, relative to no treatment control. The sensitivities for detection of transcripts from Jurkat-globin RNA after capture, inhibitory, or no treatment were 85.9% [=38.64/(38.64 + 6.33)], 79.3%, and 68.8%, whereas the specificities were 95.3% [=50.0/(2.49 + 50.0)], 96.7%, and 97.1%, respectively. These results suggested that the capture oligo method enabled the highest sensitivity of detection of transcript (*t*-test, $P < 0.01$) without significant loss of specificity.

For PAXgene RNA, removal of globin mRNA increased the percentage of transcripts detected by 6.5% for the capture oligo method and 5.8% for the inhibitory oligo method compared with no treatment control, but with no difference in transcript detection between the two globin reduction methods [Table 1, present calls (%)]. Thus, by comparing the percentages of transcripts detected among various groups and the agreement between transcripts detected from cell line RNA with and without globin, we determined that the globin reduction methods specifically increased transcript detection.

## *Effects of Globin Reduction Techniques on Transcript Variation*

Jurkat RNA with no treatment control showed similar coefficient of variation (CV) as those for Jurkat-globin RNA with no treatment (Figure 3A, JC versus JGC), suggesting

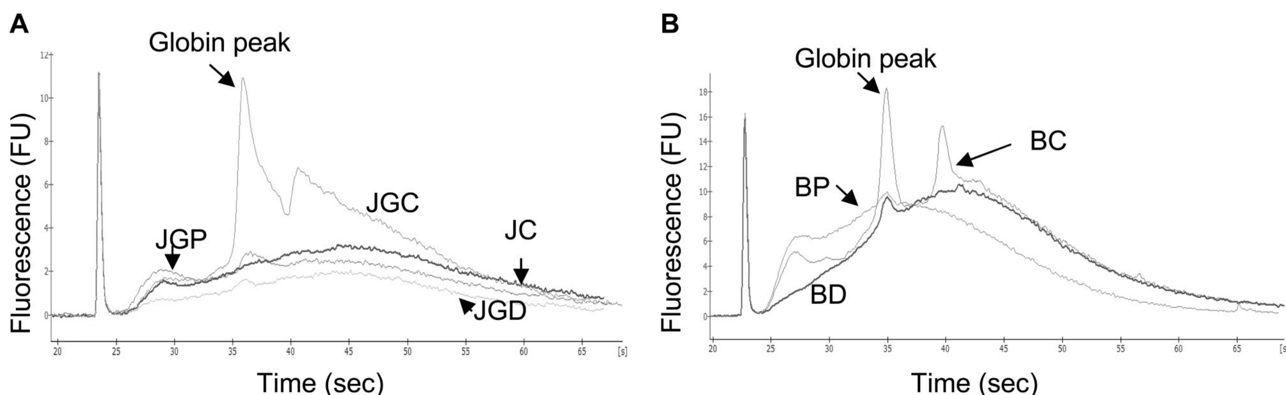


**Figure 2.** The globin reduction techniques efficiently reduce globin overabundance from cRNA derived from cell line or PAXgene RNA. Electropherograms of cRNA derived from Jurkat, Jurkat-globin, or PAXgene blood RNA that underwent various technical treatments. **A:** Traces of cRNA derived from Jurkat-globin RNA treated with DNA capture oligo (JGD), with PNA inhibitory oligo (JGP), or no treatment control (JGC), and from Jurkat RNA with no treatment control (JC). **B:** Traces of cRNA derived from PAXgene RNA treated with DNA capture oligo (BD), PNA inhibitory oligo (BP), or no treatment control (BC).

**Table 1.** *Continued*

| Jurkat-globin RNA | | | PAXgene blood RNA | | |
|---|---|---|---|---|---|
| Capture DNA oligo | Inhibition PNA oligo | Control | Capture DNA oligo | Inhibition PNA oligo | Control |
| 4 | 4 | 4 | 5 | 5 | 5 |
| 3.43 ± 0.24 | 4 | 4 | 3.71 ± 0.32 | 5 | 5 |
| 58.33 ± 2.91 | 107.93 ± 29.99 | 124.27 ± 30.96 | 25.87 ± 3.91 | 30.61 ± 17.05 | 41.18 ± 7.76 |
| 5.13 ± 1.06 | 5.10 ± 0.50 | 5.41 ± 0.89 | 7.78 ± 1.82 | 7.40 ± 1.71 | 10.68 ± 0.71 |
| 56.06 ± 3.18 | 70.90 ± 5.86 | 86.6 ± 4.22 | 57.59 ± 3.19 | 61.27 ± 5.58 | 54.27 ± 5.17 |
| 1.06 ± 0.07 | 1.09 ± 0.10 | 1.14 ± 0.02 | 1.70 ± 0.11 | 3.59 ± 1.86 | 2.25 ± 0.11 |
| 1.25 ± 0.01 | 1.17 ± 0.05 | 1.05 ± 0.03 | 2.55 ± 0.30 | 5.94 ± 3.74 | 3.16 ± 0.26 |
| 41.53 ± 0.83 | 37.4 ± 0.7 | 32.37 ± 1.56 | 39.33 ± 1.38 | 38.53 ± 2.39 | 32.77 ± 1.39 |

that the high amount of spiked globin mRNA did not account for increases in CVs. However, if one unnecessarily performed globin reduction methods on Jurkat RNA alone, then the CV increased (Figure 3A; JC, JP, JD), likely because of the additional steps. The capture oligo method increased the CV more than the inhibitory method probably because of the greater number of steps needed in the former method. On the other hand, if one appropriately performed globin reduction steps with Jurkat-globin RNA, then the CV was lowered compared with those of Jurkat RNA alone undergoing globin reduction (Figure 3A; JGD versus JD, JGP versus JP). These results indicated that the presence of globin mRNA mitigates the increase in CV due to additional processing steps of globin reduction.

Different from the cell line RNA above, PAXgene RNA that underwent the inhibitory method showed slightly higher CV than the other treatments (Figure 3B, BP versus BD and BC). Pearson correlation coefficients among triplicates for each of the nine groups ranged from 0.985 to 0.997 (data not shown) and showed the same trends as the CV graphs in Figure 3. Thus, comparisons of measures of variation suggest that the globin reduction techniques affect transcript signal variation in a manner that is dependent on the type of RNA.

## *Effects of Globin Reduction on Gene Expression Profiles*

Multidimensional scaling indicated that triplicates were tightly grouped and Jurkat RNA profiles with different technical conditions clustered close to one another, whereas clusters for Jurkat-globin RNA with different technical conditions were more distanced from each other than those for Jurkat RNA (Figure 4A, J versus JG). In addition, the capture oligo-treated Jurkat-globin RNA cluster located closest to the Jurkat RNA clusters (Figure 4A; JGD and JD, JC, CP). Similar results were observed from hierarchical clustering analysis (Figure 4C). These results suggested that globin removal from Jurkat-globin RNA using the capture oligo method resulted in gene expression profiles most similar to that of Jurkat RNA. PAXgene RNA also formed three clusters of triplicate corresponding to each technical condition, with the triplicates undergoing either globin reduction methods being closer to each other (Figure 4, B and D). These results suggested that the different methods induced distinct gene expression profiles, but globin reduction brought the profiles closer to the original profile than without treatment, particularly with the capture oligo method.

## *Effects of Globin Reduction on Specific Transcript Signals*

The HG-U133 plus 2.0 microarray has ~54,120 probe sets to detect more than 47,000 human transcripts. Class comparison analysis resulted in 8614 probe sets that showed a difference in expression level in at least one of the cell line RNA groups. The signals from these probe sets are shown in the sample clustered heat-map in Figure 5A. Group I 490 probe sets showed differences between Jurkat-globin RNA versus Jurkat RNA alone, regardless of method. Group II represented 3833 probe sets that were similarly at low signal in Jurkat-globin RNA

**Table 2.** Cross-Tabulation of Detection Calls between Jurkat-Globin RNA with Various Treatments versus Jurkat RNA

| | | Jurkat-globin RNA | | | | | |
|---|---|---|---|---|---|---|---|
| | | Capture DNA oligo | | Inhibition PNA oligo | | Control | |
| | Detection calls | P* (%) | A† (%) | P (%) | A (%) | P (%) | A (%) |
| Jurkat RNA | P (%) | 38.64 ± 0.67‡ | 6.33 ± 1.09 | 35.43 ± 0.62 | 9.26 ± 1.39 | 30.64 ± 1.24 | 13.88 ± 1.84 |
| | A (%) | 2.49 ± 0.48 | 50.00 ± 1.05 | 1.72 ± 0.30 | 50.90 ± 1.31 | 1.50 ± 0.23 | 51.13 ± 1.36 |

The overall call concordance excluding margin calls between Jurkat and JG RNA was tabulated.
*Present calls.
†Absent calls.
‡STDEV.

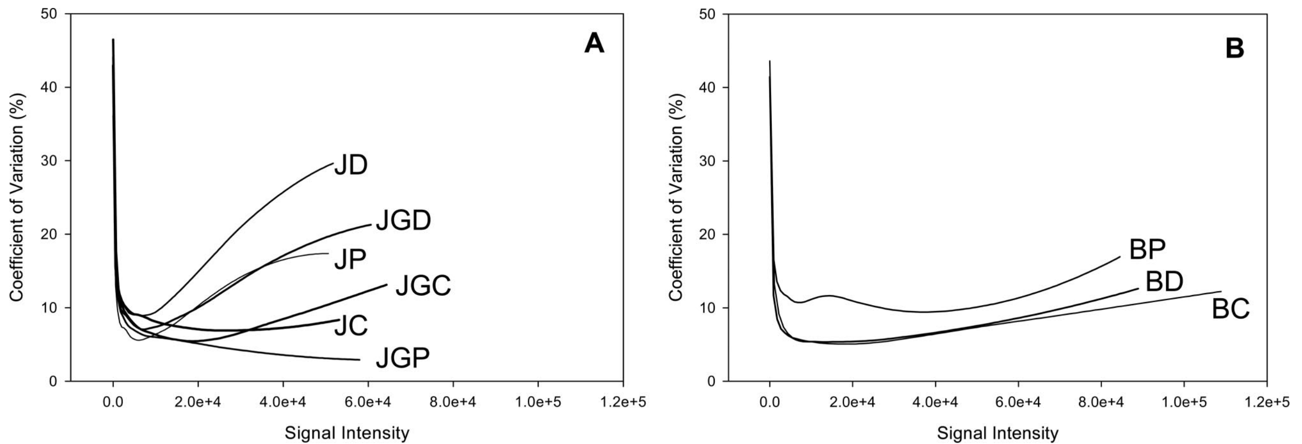


**Figure 3.** The globin reduction techniques affect on transcript signal variation is dependent on RNA type. The overabundance of globin mRNA mitigates the increase in transcript signal variation because of the additional steps required for globin reduction. Coefficient of variation versus signals from all probe sets for various RNA types with different technical treatments. Curves shown are from Loess 2-degree freedom fitting of the signal data points. **A:** Cell line RNA. Jurkat (J) or Jurkat-globin (JG) RNA treated with capture DNA oligo (JD, JGD), inhibitory PNA oligo (JP, JGP), or no globin reduction control (JC, JGC). **B:** PAXgene blood RNA treated with capture DNA oligo (BD), inhibitory PNA oligo (BP), or no treatment control (BC).

with capture oligo method and all Jurkat RNA but were at higher signals in Jurkat-globin RNA with inhibitory method and at highest signals in Jurkat-globin RNA control. These probe sets probably represent false positives for the inhibitory and control treatments and might be attributable to unspecific binding of globin mRNA to these probes. Group III represented 1476 high signaled probe sets in Jurkat RNA samples but low signals in all Jurkat-globin RNA samples. These false negative probe sets could result from increased background from globin mRNA overabundance. Group IV represented 1380 probe sets that could be revealed after globin RNA reduction with the capture oligo protocol but not the inhibitory protocol or after no treatment. Group V represented

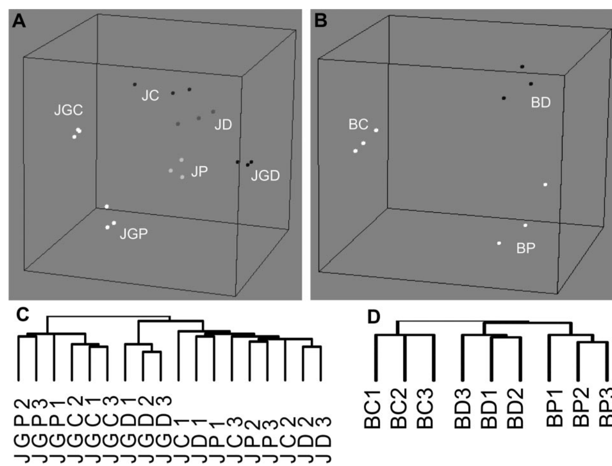


**Figure 4.** Different treatments induced distinct gene expression profiles, but globin reduction brought the profiles closer to the original profile than without treatment, particularly with the capture oligo method. Unsupervised cluster analysis of gene expression profiles from various RNA types with different technical treatments. Logged signals from all probe sets were used. **A** and **B:** Multidimensional scaling of unscaled signals. **C** and **D:** Hierarchical clustering using center correlation and average linkage. In **A** and **C** cell line RNA, Jurkat (J) or Jurkat-globin (JG) RNA were treated with capture DNA oligo (JD, JGD), inhibitory PNA oligo (JP, JGP), or no globin reduction control (JC, JGC). In **B** and **D**, PAXgene blood RNA was treated with capture DNA oligo (BD), inhibitory PNA oligo (BP), or no treatment control (BC).

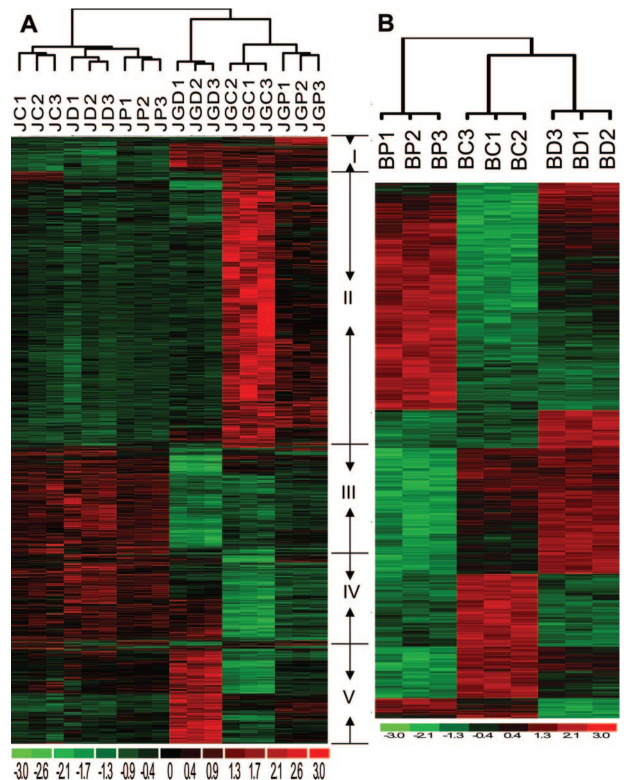


**Figure 5.** Patterns of changes in differential transcript signals induced by RNA types and treatment methods. Heat-maps of probe sets with differing signals among groups. Probe sets and samples were clustered using center correlation and average linkage. Only dendrograms from clustering of samples are shown on top of each panel. To select from 54,120 probe sets for those with differing signals, analysis of variance with random variance model was performed using *P* values $\leq 0.001$ as significant. **A:** Heat map and sample clustering dendrogram of 8614 probe sets with significant differences in signals among the cell line RNA groups. Jurkat (J) or Jurkat-globin (JG) RNA treated with capture DNA oligo (JD, JGD), inhibitory PNA oligo (JP, JGP), or no globin reduction control (JC, JGC). I to V based on JGD expression pattern. **B:** Heat map and sample clustering dendrogram of 1988 probe sets with differing in signals among the PAXgene blood RNA groups treated with capture DNA oligo (BD), inhibitory PNA oligo (BP), or no treatment control (BC).

1435 unique higher signal probe sets resulting from the capture oligo protocol and probably represent false positive for this method. Class comparison analysis resulted in 1988 probe sets that showed differences in expression levels among PAXgene RNA treatment groups (Figure 5B). These results more specifically characterized the number of probe sets, the magnitude and direction, and patterns of changes in gene expression profiles induced by RNA types and treatment methods.

## Discussion

We have systematically evaluated and compared the effects of two globin reduction methods on gene expression profile measurement outcomes from cell line and PAXgene whole blood RNA. We showed that the globin reduction techniques were mostly efficient at reducing globin mRNA from cRNA but can induce RNA type-dependent changes in quality of RNA during sample preparation and quality of microarray results. Importantly, we showed that the methods were efficient at increasing the sensitivity of transcript detection without loss of specificity. However, the globin reduction steps introduce particular transformations to the original gene expression profiles, and even after globin removal, one could not recover a profile equivalent to that of an identical RNA sample without globin mRNA excesses. Comparing the methods, we found that the capture oligo method generally gave better gene expression profile outcomes, but it had certain practical difficulties and complexities compared with the inhibitory oligo method.

Highly pure RNA is required for successful implementation of the capture oligo method probably because of the harsh condition for hybridization of the oligo to globin mRNA. Of particular interest, we found that RNA isolated using the PAXgene blood kit after in-solution DNase treatment was not of sufficient quality to be compatible with the capture oligo method. Therefore, additional clean-up and concentration subprocedures of ethanol precipitation, or alternatively RNeasy Minelute column, were tested. The latter resulted in sufficiently pure PAXgene RNA and, with at least 5 μg of pure PAXgene RNA, was able to yield enough cRNA for hybridization onto microarray. However, the cRNA yields from the capture oligo method were still lower than the other methods regardless of RNA type. The lowered cRNA yield could be attributable to the capture oligo method inevitably degrading RNA because the RIN value generally decreased by 1 after this method, or attributable to the RNA going into cDNA synthesis in water instead of BR5. On the other hand, our preliminary results indicated that as low as 1 μg of DNase-digested PAXgene RNA treated with inhibitory oligo was able to amplify enough cRNA (~20 μg) for microarray hybridization. Thus, both globin reduction methods required additional starting RNA from PAXgene blood kit because of loss attributable to the extra clean-up and concentration steps, and the capture oligo method required even more because of loss from its numerous steps and low cRNA yield. Some of this additional starting RNA requirement is partially offset by recent target preparation developments from companies such as Affymetrix and Ambion that can amplify the transcriptome more efficiently by optimizing labeling molecule and enzyme conditions.

Previous studies noted that globin reduction reduced interindividual variations of expression profiles from whole blood samples (Affymetrix technical note: Globin Reduction Protocol: A Method for Processing Whole Blood RNA Samples for Improved Array Results. Santa Clara, CA), whereas herein, we considered only contribution of techniques and RNA type to variation of expression profiles. For cell line RNA, we found that the usage of globin reduction with sample RNA that needs globin removal buffered the increases in variation due to additional steps of globin reduction. Unnecessary usage of globin reduction with RNA that does not need globin removal resulted in the most varying expression signals (Figure 3A), probably because the oligos nonspecifically bind to nonglobin mRNA. The effects of globin reduction methods on profile variation were different between cell line RNA versus PAXgene RNA. The inhibitory method with Jurkat-globin RNA gave the lowest CV, whereas it gave the highest CV with PAXgene RNA compared with control (Figure 3B). This could be attributable to the reduction of cRNA size from inhibitory treatment (Figure 2B) and may also account for the high *GAPDH* and *Actin* 3′/5′ ratios (Table 1) for this group.

Importantly, we showed that the globin reduction methods significantly improved the sensitivity of transcript detection while maintaining relatively stable specificity. Using results from cell line RNA, we quantitatively determined the sensitivity of detection of Jurkat-globin RNA for the capture oligo, inhibitory oligo, and control methods to be 85.9, 79.3, and 68.8%, and the specificities were 95.3, 96.7, and 97.1%, respectively. The capture oligo method improved the percent present calls by 9 and 6.5%, whereas the inhibitory oligo method improved by 5 and 5.8% for Jurkat-globin and PAXgene RNA, respectively. For Jurkat-globin RNA, improvements in sensitivity seemed to be related to decreasing in background effects (Table 1) but were not the case for PAXgene RNA. For PAXgene RNA control (Table 1), the scale factor showed the highest magnitude trend, whereas background was lowest, suggesting that improvements in sensitivity for PAXgene RNA is via other mechanisms that increased signals. We showed that the globin reduction steps introduce particular transformations to the original gene expression profiles, and after globin removal, one could not recover a profile similar to RNA without globin mRNA overabundance controls. Quantitative clustering analysis showed that the capture oligo method resulted in profiles closest to controls (Figure 4, A and C).

In summary, both globin methods improved sensitivity of transcript detection compared with no treatment controls. The capture oligo method had slightly better sensitivity of transcript detection for cell line RNA, lowered signal variation for PAXgene RNA, and gene expression profiles closer to controls than the inhibitory oligo method. However, capture oligo required larger amounts of initial RNA that was of high-quality RNA to yield sufficient cRNA amounts, and its procedures are more com-

plex and time consuming than the inhibitory method. In addition, the capture oligo method separates globin mRNA from the sample, thus allowing for usage with non-3′ bias techniques downstream, such as direct chemical labeling of RNA or interrogation on exon microarrays, whereas the inhibitory method is only compatible with subsequent 3′ bias protocols because it only blocks globin mRNA reverse transcription.

These results have significant impact on population level protocols. For instance, in populations with a diversity of diseases and phenotypes, one may not obtain uniformly high-quality RNA even using the PAXgene tubes; thus, the capture oligo method would not be compatible. Alternatively, if the population sample has wide ranges of globin mRNA amounts, the inhibitory method may not be effective for samples with very high percentage of globin mRNA. Also, even though one gains in sensitivity with the globin reduction methods, this comes at a cost of requiring higher amounts of initial total RNA, which may not be available in certain circumstances. Thus, we have determined, evaluated, and compared advantages and disadvantages of using or not using certain globin reduction methods and suggested some recommendations. Our findings should help others in multicenter study design and progress us closer toward gene expression profile surveillance.

## Acknowledgments

We thank the sample donors.

## References

1. Thach DC, Agan BK, Olsen C, Diao J, Lin B, Gomez J, Jesse M, Jenkins M, Rowley R, Hanson E, Tibbetts C, Stenger DA, Walter E: Surveillance of transcriptomes in basic military trainees with normal, febrile respiratory illness, and convalescent phenotypes. Genes Immun 2005, 6:588–595
2. Debey S, Zander T, Brors B, Popov A, Eils R, Schultze JL: A highly standardized, robust, and cost-effective method for genome-wide transcriptome analysis of peripheral blood applicable to large-scale clinical trials. Genomics 2006, 87:653–664
3. Thach DC, Lin B, Walter E, Kruzelock R, Rowley RK, Tibbetts C, Stenger DA: Assessment of two methods for handling blood in collection tubes with RNA stabilizing agent for surveillance of gene expression profiles with high density microarrays. J Immunol Methods 2003, 283:269–279
4. Rainen L, Oelmueller U, Jurgensen S, Wyrich R, Ballas C, Schram J, Herdman C, Bankaitis-Davis D, Nicholls N, Trollinger D, Tryon V: Stabilization of mRNA expression in whole blood samples. Clin Chem 2002, 48:1883–1890
5. Li C, Hung Wong W: Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2001, 2:RESEARCH0032.1–0032.11
6. Li C, Wong WH: Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci USA 2001, 98:31–36