GENOMICS INFORMATICS PROTEOMICS METABOLOMICS

ATCTGATCCTTCTGAACGGAACTAATTTCAA
GAATCTGATCCTTGAACTACCTTCCAAGGTG

Analysis of Gene Expression Using Agilent Microarrays and RNA Isolated from Human Whole Blood

Authors
Christopher Rizzo
Agilent Technologies
Little Falls, DE USA

Athene Chan, Diane Ilsley-Tyree
Agilent Technologies
Santa Clara, CA USA

Abstract

Agilent's dual-mode Gene Expression Microarray platform is a comprehensive and robust system that couples the ease and simplicity of one-color experimental design with the capability of two-color sensitivity for detection of small gene expression changes—all in a single platform. The Agilent platform now includes a method for the analysis of differential gene expression from human whole blood, which is compatible with the PAXgene™ Blood RNA System (PreAnalytiX GmbH, Hombrechtikon, CH). The PAXgene Blood RNA System consists of a blood collection tube and nucleic acid purification kit and is intended for the collection, storage, and transport of blood and stabilization of intracellular RNA for subsequent isolation and purification of intracellular RNA. Combined with the PAXgene System, Agilent's amplification and labeling technology, in conjunction with the high sensitivity and specificity of 60-mer oligonucleotide microarrays, offer a total solution for gene expression in whole blood without the need for additional steps or kits to remove globin transcripts. Titration experiments using Agilent RNA spike-in control targets demonstrate that the sensitivity of detection for low abundance or poorly expressed transcripts with RNA isolated from whole blood compares favorably with RNA isolated from other cell and tissue sources.

Introduction

Blood is an attractive tissue for clinical and biological research studies because samples can be easily collected and multiple measurements can be made from the same patient. Furthermore, a wide range of analyses can be made by gene expression profiling of peripheral blood samples. These include: prognosis, diagnosis, and management of diseases; monitoring of pathological and biochemical responses to toxicodynamic and pharmacodynamic factors; and determination of genetic factors potentially affecting all of the above. As one example, a recent presentation at the Annual Meeting for the Society of Toxicology demonstrated that sub-toxic and non-toxic doses of hepatotoxicants can be monitored using gene expression signatures in blood as surrogate markers prior to the onset of detectable liver damage[1]. This application note describes the use of the Agilent Technologies

portfolio of products with the PAXgene Blood RNA System to facilitate the analysis of gene expression profiles from whole blood.

There are two major challenges that can hinder RNA profiling from blood samples. The first of these is the relative instability of the *in vitro* cellular RNA profile after blood collection, whether due to RNA degradation or gene induction. The second is the presence of interfering components such as globin mRNA, genomic DNA, and other small molecules that can inhibit downstream enzymatic reactions. The high abundance of globin transcript in whole blood is particularly troublesome as it can inhibit the synthesis of other labeled transcripts, thereby creating bias in the downstream microarray results.

Agilent's solution for whole blood gene expression analysis is designed to be compatible with the PreAnalytiX PAXgene Blood RNA System (PreAnalytiX, GmbH, catalog number 762134) which stabilizes the gene transcription profile at the point of sample collection and provides highly purified cellular RNA[3]. The Blood RNA Isolation kit contains an optimized Resuspension Buffer and Proteinase K to facilitate the extraction of RNA from whole blood. An on-column DNase treatment is performed to reduce genomic DNA for especially sensitive applications.

The high-purity RNA isolated with the PAXgene Blood RNA Kit can then be used directly in amplification and labeling reactions using Agilent's Low RNA Input Linear Amplification Kit. The kit uses a patented, high-yield, single tube procedure to generate amplified fluorescently-labeled cRNA that is suitable for both one-color and two-color gene expression applications. Only a single round of amplification is required and no additional kits or procedures are needed to remove globin message or protein (Figure 1). Agilent's microarray platform for gene expression includes empirically selected 60-mer oligonucleotide probes for maximum specificity, optimized high-stringency protocols, and RNA spike-in controls. Agilent's Feature Extraction software electronically generates a QC report to validate the results from every experiment, and Feature Extraction output files can be exported to GeneSpring or Rosetta Resolver data analysis products.

Materials and Methods

Isolation of Total RNA from Whole Blood

HeLa cells were obtained from American Type Culture Collection (ATCC, Manassas, VA, catalog number CCL-2) and total RNA was purified using Agilent's Total RNA Isolation Mini Kit (5185-6000). Human spleen RNA was purchased from Ambion, Inc. (Austin, TX, catalog number 7852).

All blood was collected from healthy volunteer donors. 2.5 mls of whole blood was collected into each PAXgene Blood RNA Tube using standard methods. Samples were handled as recommended by the manufacturer prior to processing. RNA was isolated from blood samples using the PAXgene Blood RNA Kit (PreAnalytix GmbH, Switzerland, product number 762134). RNA quality and integrity were determined using the Eukaryote Total RNA Nano 6000 assay (Agilent RNA 6000 Nano LabChip Kit, part number 5065-4476) on the Agilent Technologies 2100 Bioanalyzer (part number G2940CA) and quantified by measuring A_{260nm} on a UV/Vis spectrophotometer (ND-1000, NanoDrop Technologies, Rockland, DE).

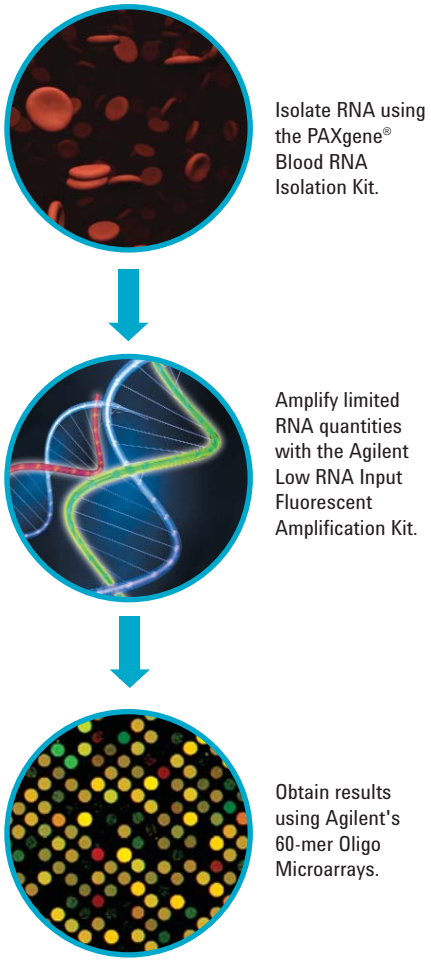

Figure 1. Steps involved in the whole blood gene expression analysis workflow.

Generation of Fluorescently Labeled cRNA Targets

Total RNA isolated from whole blood was amplified and fluorescently labeled using the Agilent Technologies Low Input Linear Amplification Kit (part number 5184-3523) following the detailed protocol described in the kit manual (publication number 5185-5818). Each reaction contained 200 or 500 ng total RNA from whole blood and 2 µl (34 pg) of RNA spike-in control (diluted as described in the product insert, 5188-5279). Cyanine 3- and cyanine 5-labeled CTPs were obtained from PerkinElmer/NEN Life Sciences (Boston, MA, catalog numbers

NEL580 & NEL581 respectively). Spike-in mix A was included in the cyanine 3 reactions and spike-in mix B in the cyanine 5 reactions. The labeled cRNA was purified using Qiagen's RNeasy Kit (part number 74104). Mass yields and specific activities of the labeled cRNA targets were determined by measuring the absorbance spectra on a UV/Vis spectrophotometer. Quality of the targets was further assessed using the mRNA Nano 6000 assay on the Agilent 2100 Bioanalyzer. Alternatively, the fluorescently labeled cRNA transcripts were generated with a comparable kit from a competitor, following the manufacturer's recommended protocol.

Titration of Spike-In Targets to Determine Sensitivity

Spike-in titration experiments were performed using serial 2-fold dilutions of the Spike A and Spike B mixes. Each spike-in mix contains 10 different transcripts spanning a 200-fold mass dynamic range. A range of 2 to 137 pg of the transcript mix was mixed with 500 ng total RNA isolated from HeLa cells, spleen, or donor whole blood, amplified and fluorescently labeled, and hybridized to Agilent microarrays.

Hybridization and Scanning of Microarrays

For each hybridization, 750 ng of cyanine 3- and 750 ng of cyanine 5-labeled cRNA were fragmented and hybridized to an Agilent Technologies Human 1A (V2) Gene Expression Microarray (part number G4110B), using the Agilent Gene Expression hybridization kit (part number 5188-5242) as described in the Two-Color Microarray-Based Gene Expression Analysis v4.0 manual (publication number G4140-90050). Following hybridization, all microarrays were washed as described in the manual and scanned using Agilent's dual laser DNA microarray scanner (part number G2565A). The scans were converted to data files with Agilent's Feature Extraction software (Version 8.5). Data were analyzed using Microsoft Access and Spotfire.

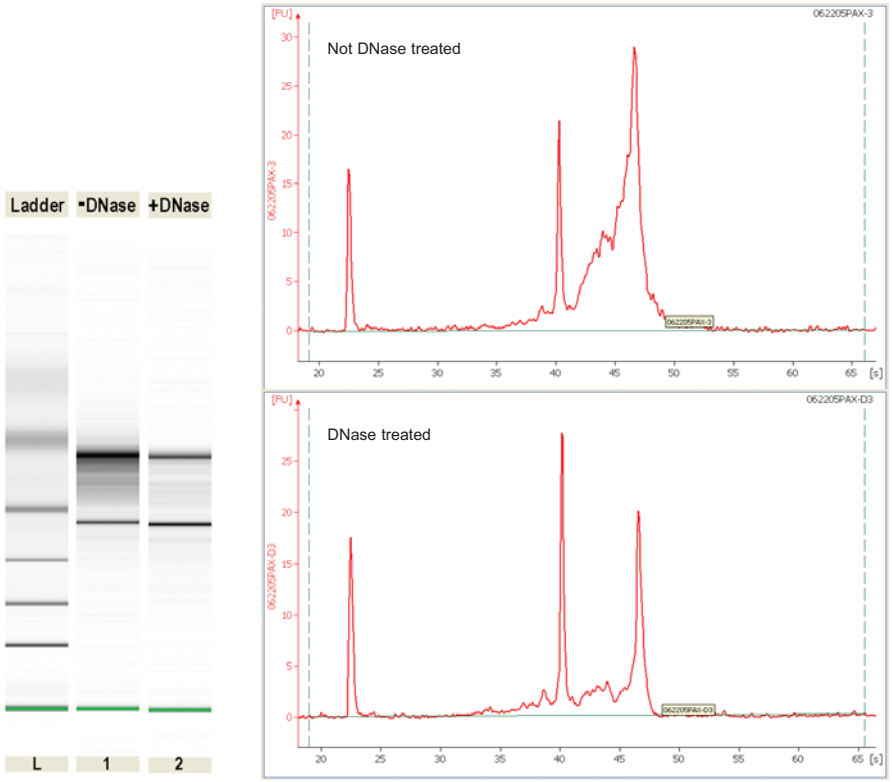

Figure 2. Agilent 2100 Bioanalyzer gel image and electropherograms of total RNA isolated from whole blood. The gel image shows the ladder in lane L, RNA not treated with DNase in Lane 1, and DNase treated RNA in Lane 2. The RNA Integrity Number (RIN) for the DNase treated sample was 9.0.

Results

Isolation and Purification of RNA from Whole Blood

High quality starting material is essential for successful microarray experiments. RNA preparations are normally analyzed spectrophotometrically to measure nucleic acid concentration and assess purity. Typical RNA yields from 2.5 ml of whole blood range from 4 to 15 µg and can vary widely among donors. The ratio of absorbance at 260 nm and 280 nm (A260/280) for RNA isolated from blood using the PAXgene kit was consistently 1.9 - 2.0, indicative of RNA that is essentially free of contaminating protein. The RNA was further assessed for quality and integrity using Agilent's 2100 Bioanalyzer[2]. The top panel of Figure 2 shows a broad 28S peak, with elevated baseline between the 18S and 28S peaks that is indicative of genomic DNA contamination. The RNA was then treated with DNase, resulting in the removal of the contaminating genomic DNA (bottom panel). The gel images and traces shown here illustrate that higher purity RNA product can be obtained with the PreAnalytiX kit when using the optional on-column DNase treatment. Moreover, removal of genomic DNA from total RNA samples prior to their preparation as cRNA targets has

previously been shown to have a positive impact on microarray performance.

High Yield and Specific Activity of Labeled cRNA

Fluorescently labeled cRNAs were analyzed spectrophotometrically on a ND-1000 Spectrometer (NanoDrop Technologies, Rockland, DE). Yield and specific activity of cyanine 3- and cyanine 5-labeled cRNA targets were comparable to those obtained for other tissues and cell lines at this input level. Therefore, the use of whole blood as a starting material for generating labeled cRNA shows no deleterious effects when using the Low Input Linear Amplification kit reagents and protocols.

Sensitivity and Dynamic Range

Other microarray platforms require the removal of globin mRNA prior to microarray hybridization to decrease noise and increase sensitivity in whole blood samples. To determine if there was a change in the ability to detect low abundance or poorly expressed transcripts in RNA isolated from whole blood relative to RNA isolated from cell lines or tissues, synthetic RNA spike-in titration experiments were performed. The Agilent Spike-In kit contains a mixture of 10 different synthetic transcripts that span a 200-fold mass dynamic range. Increasing amounts of the spike-in mixture (2-137 pg, total mass of all 10 transcripts) were mixed with 500 ng total RNA isolated from healthy blood donors, HeLa cells, or spleen. The same total RNA was used in each reaction, and equal amounts of spike-in mix A and spike-in mix B were added to the cyanine 3 and cyanine 5 amplification reactions, respectively. The data from the titration experiment are shown in Figures 3 and 4. Figure 3 shows the

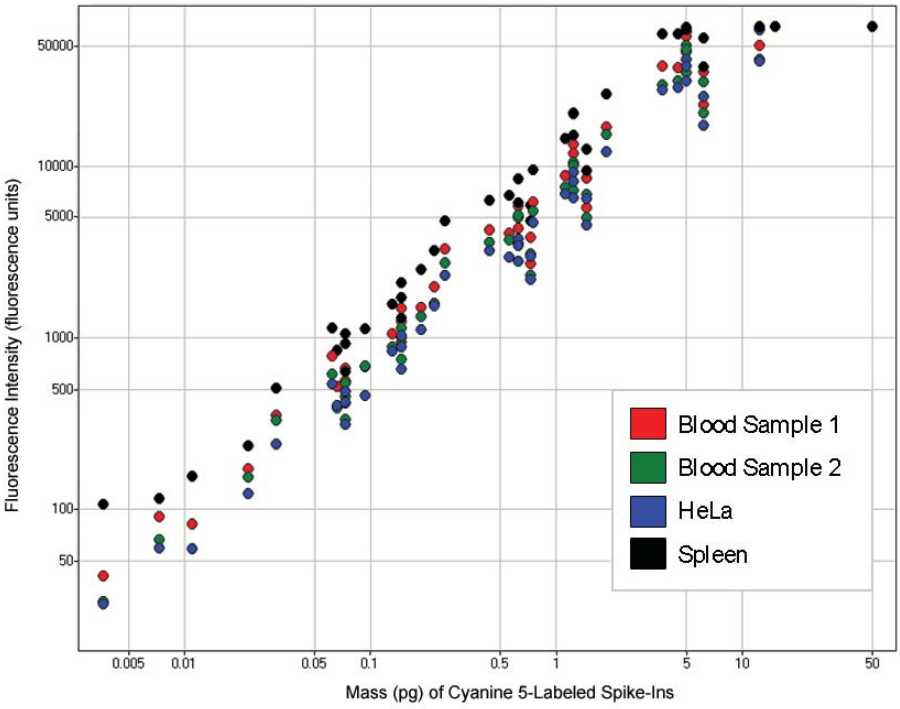

Figure 3. Dose response curve of red (cyanine 5) signal intensity versus mass of spike-in transcripts.

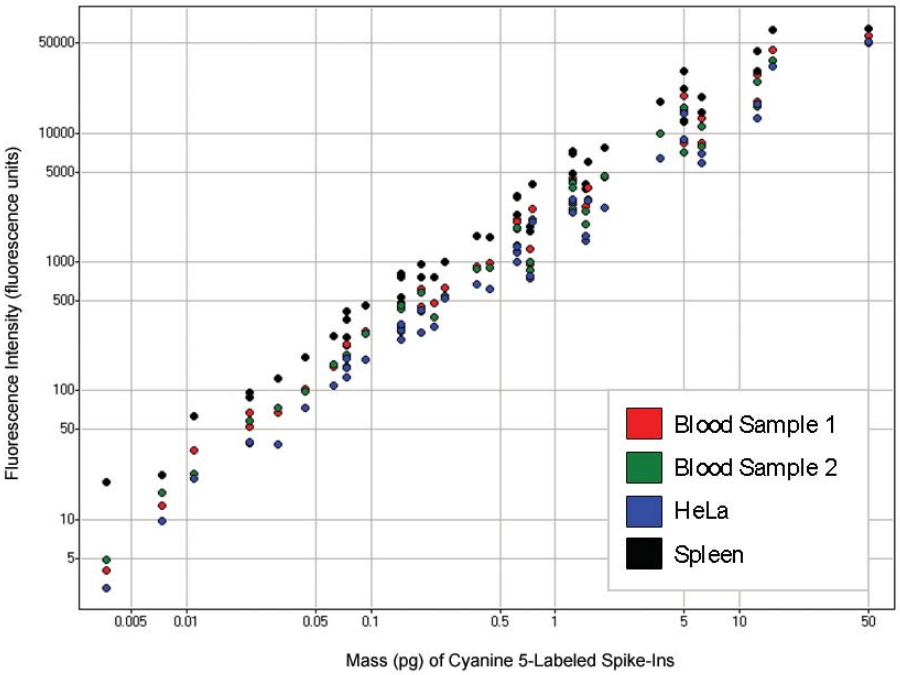

Figure 4. Dose response curve of green (cyanine 3) signal intensity versus mass of spike-in transcripts.

Table 1. E1a slopes: Expected vs. Observed

Sample/Spike-in	2pg	4pg	17pg	34pg	137pg
Blood Sample 1	0.9901	0.9570	0.9413	0.9456	0.9411
Blood Sample 2	0.9446	0.9460	0.9425	0.9584	0.9483
HeLa	0.9367	0.9217	0.9229	0.9201	0.9581
Spleen	0.8751	0.9383	0.9369	0.9419	0.9403

Table 1. Accuracy of log ratios. Expected versus observed log ratios for the 4 samples were plotted as a function of spike-in mass. Slope values were taken from Plot X from the eQC report generated by Agilent's feature extraction software.

cyanine 5 fluorescent signal intensity as a function of increasing mass of each of the 10 synthetic transcript for each of the 4 sample types. Each point represents the average of 30 replicate probes across the microarray. For all 4 RNA samples, the linear signal intensity range is quite similar. This dose response curve demonstrates that the detection of transcripts at low concentrations is comparable between different sample types with no reduction in sensitivity when using RNA from whole blood. Similar results are shown for cyanine 3 labeled samples (Figure 4). The Agilent two-color RNA spike-in controls are designed to measure the accuracy of fold change detection of the platform. Analysis of expected versus observed log ratios of the spike-in controls (Table 1) when included with each of the 4 sample types over a range of spike-in masses indicates accurate detection of even low copy number targets regardless of sample type. The slope of the expected vs observed log ratios for even the lowest spike-in mass input is greater than 0.94 for the whole blood samples. These results suggest that sensitivity on the Agilent platform does not change as a function of sample type, even for whole blood samples.

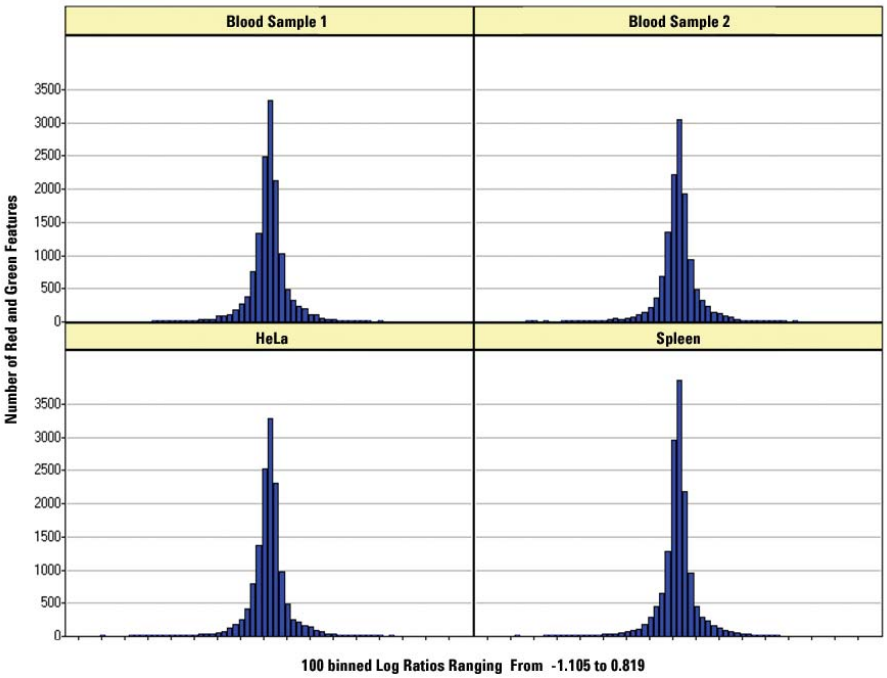

Figure 5. This figure shows the number of features, for each of the four sample types, distributed within binned log ratios ranging between -0.0107 to 0.0270. For each sample, the mass of spike-in transcripts was 34 pg. The x-axis represents the Red vs Green Log Ratio for each experiment, separated into 100 bins ranging from -1.105 to 0.819. The y-axis is the number of features within each binned log ratio. Mean and standard deviation of log ratios for the four sample types are shown in the Table 2.

Sample	Mean of Log Ratio	Standard Deviation of Log Ratio
Blood Sample 1	-0.0024	0.1228
Blood Sample 2	0.0011	0.1221
HeLa	-0.0003	0.1224
Spleen	-0.0024	0.1260

Table 2. Mean and standard deviation of log ratios for each sample type.

System Noise

To measure system noise, self versus self experiments were analyzed. In a self-comparison experiment the expected log ratio for each feature is zero. Figure 5 shows the distribution of log ratios from selected microarrays (34 pg of the spike-in mixes added). For each of the sample types, the ratios were tightly clustered around the expected value of zero (Table 2). Figure 6 shows the fluorescence intensity plots for the same set of microarrays. The red and green intensity values lie within a very narrow range along the diagonal, with increased variability only at the very low signal intensity region, reflecting very little noise in the system performance. These results indicate that regardless of the type of sample source, including whole blood, overall system noise of the Agilent platform is very low. Agilent's robust and sensitive gene expression solution can be used with any sample type, without the need for special treatments, such as globin reduction, to reduce the complexity of the sample.

To determine if the Agilent Low Input Linear Amplification kit contributes to the system performance without the need for globin reduction, a side by side comparison was performed between Agilent's kit and an amplification and labeling kit supplied by another vendor. The distribution of the cRNA products generated by each kit were analyzed using the Bioanalyzer and RNA 6000 Nano chips. In these experiments 200 ng of total RNA isolated from whole blood using the PAXgene kit was used as input for amplification and labeling reactions using cyanine 5-CTP and either Agilent's kit or the Competitor's kit. Representative electropherograms of the cyanine 5-labeled cRNAs are shown

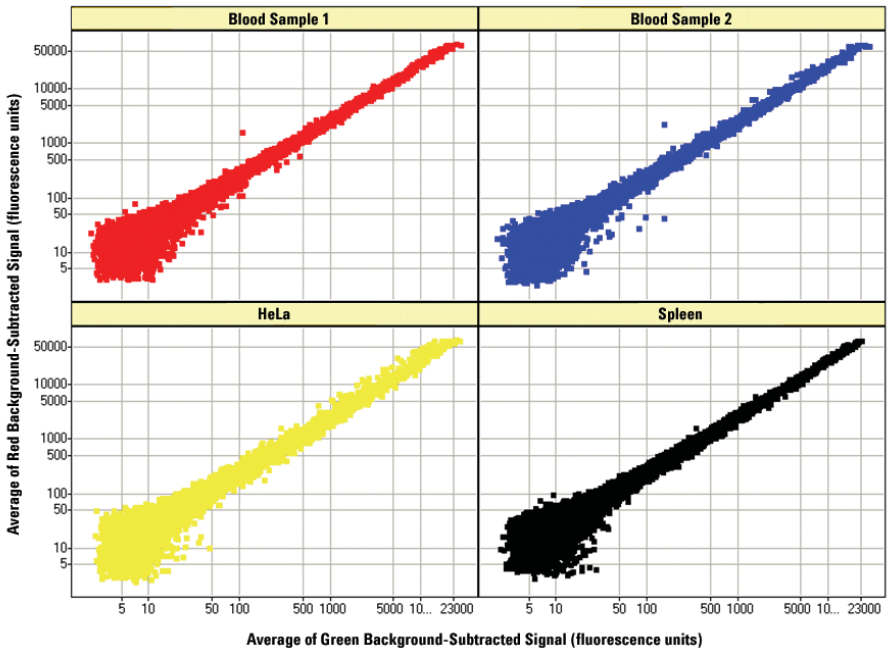

Figure 6. Intensity plots from self vs self hybridizations using fluorescently labeled cRNA derived from human whole blood, human spleen, and HeLa cell total RNA.

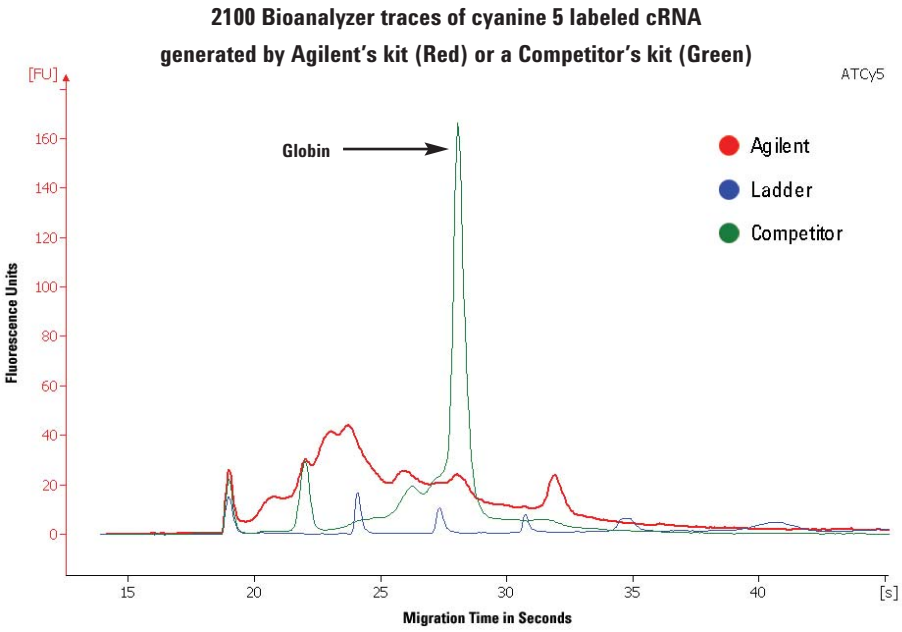

Figure 7. Comparison of cRNA profiles from Agilent's Low Input Linear Amplification kit and a competitor's cRNA labeling kit.

in Figure 7. The Agilent kit generated fluorescently labeled cRNA with a broad distribution of cRNA and no distinct peaks. This cRNA distribution is similar to those observed with other cell line and tissue samples (data not shown). In contrast, cRNA produced with the Competitor's kit contained a major peak

at approximately 650 nucleotides most likely corresponding to globin mRNA. This peak represents the majority of the cRNA products generated. cRNA generated from other non-blood samples showed a broad distribution of cRNA, with no distinct peaks (data not shown).

Conclusions

The Agilent solution for gene expression profiling of whole blood samples provides researchers with a convenient system to obtain accurate, reliable, and consistent results. Key benefits of the system:

- Efficient RNA purification and low input linear amplification stabilizes the gene expression profile and eliminates the need for erythrocyte lysis and globin message reduction while delivering highly purified RNA with little genomic DNA contamination.
- High yield and specific activity cRNA targets with fewer steps and no globin reduction.
- The same broad dynamic range of transcript detection for microarrays hybridized with whole blood samples as achieved with other RNA sources.
- A complete workflow with downstream oligonucleotide microarrays, scanners, and analysis tools optimized for integrated performance.

The ability to work with challenging samples such as whole blood is an important aspect of Agilent's new dual-mode Gene Expression Platform which provides for both one- and two-color detection of changes in gene expression. Experiments are currently in progress to demonstrate the utility of the One-Color Platform for applications using whole blood, thereby providing the flexibility to select either mode of detection using Agilent microarrays. This flexibility continues with Agilent's multiple options for microarray design, feature density, and the customization provided by our eArray web-based application. The two detection modes share a very similar workflow, including the labeling and hybridization protocols, scanner hardware, and data analysis software (Feature Extraction and GeneSpring GX). In conclusion, Agilent's dual-mode platform provides the performance and flexibility suited to address any experimental challenge in microarray-based analysis of gene expression, and when combined with the PAXgene Blood RNA System, the platform delivers a reliable, integrated approach for gene expression analysis from whole blood.

References

1.) Agilent Technologies Informational Workshop, Annual Meeting for the Society of Toxicology, San Diego, CA March 8th, 2006

2.) O. Mueller, S. Lightfoot, and A. Schroeder. "RNA Integrity Number (RIN) - Standardization of RNA Quality Control", Agilent Technologies, Inc. publication 5989-1165EN, www.agilent.com/chem

3.) L. Rainen, U. Oelmueller, S. Jurgensen, R. Wyrich, C. Ballas, J. Schram, C. Herdman, D. Bankaitis-Davis, N. Nicholls, D. T. Rollinger, and V. Tryon. "Stabilization of mRNA Expression in Whole Blood Samples" J. Clin Chem. 48:11, pp. 1883-1890 (2002)."

For more information

Learn more:
www.agilent.com/chem/dualmode

Buy online:
www.agilent.com/chem/store

Find an Agilent customer center in your country:
www.agilent.com/chem/contactus

U.S. and Canada
1-800-227-9770
agilent_inquiries@agilent.com

Europe
info_agilent@agilent.com

Asia Pacific
adinquiry_aplsca@agilent.com

About Agilent's Integrated Biology Solutions

Agilent Technologies is a leading supplier of life science research systems that enable scientists to understand complex biological processes, determine disease mechanisms, and speed drug discovery. Engineered for sensitivity, reproducibility, and workflow productivity, Agilent's integrated biology solutions include instrumentation, microfluidics, software, microarrays, consumables, and services for genomics, proteomics, and metabolomics applications.

Printed in the USA April 27, 2006
5989-4899EN

Agilent Technologies