

Application Note

Blocking of Globin Reverse Transcription to Enhance Human Whole Blood Gene Expression Profiling

Yasmin Beazer-Barclay, Doug Sinon, Christopher Morehouse, Mark Porter,
and Mike Kuziora Ph.D.
Genomics R&D and Product Management
Gene Logic Inc.

Blocking of Globin Reverse Transcription to Enhance Human Whole Blood Gene Expression Profiling

Gene expression profiling of human blood is becoming increasingly used in the pharmaceutical and diagnostic industries for the discovery and development of clinical biomarkers linked to human disease. The laboratory steps and equipment required to reproducibly isolate white blood cells (WBCs) not only increase the cost but also the elapsed time to RNA isolation. The manipulations required and time to RNA isolation likely result in altered gene expression profiles. To mitigate this issue, methods have been developed to process whole blood prior to RNA isolation and gene expression analysis. The most useful microarray expression profiling data is generated from whole blood samples in which red blood cells (RBCs) are selectively pre-lysed and the RNA is subsequently purified. However, the abundance of globin mRNA transcripts in RBCs and reticulocytes masks the mRNA contributed by WBCs.

Some approaches result in near instantaneous stabilization of RNA from all blood cells, including RBCs and reticulocytes, but do not remove globin transcripts. For example, adding TRIzol® reagent to whole blood allows for a quick and complete lysis of all cells in a guanidinium salt environment thereby inhibiting RNase activity. Similarly, the PAXgene™ Blood RNA Isolation System (PAX) allows for the instant protection of total RNA integrity and stabilization of gene expression patterns without the need for additional equipment, such as a chemical hood. Both methods are easier to implement in a clinical setting than WBC purification. However, for these techniques to be useful for gene expression of WBCs, it is essential to implement a robust globin-blocking strategy to reduce the background of globin message contributed by RBCs.

Gene Logic's globin reduction method utilizes 2'O-methyl modified oligomers as gene-specific "blockers" that greatly reduce the ability of reverse transcriptase to generate globin cDNAs during the sample preparation process (Figure 1). The blocking procedure utilizes a proprietary mixture of oligomers that have been optimized for use with 2-5 µg of total RNA from whole blood.

In this report, we compare microarray data from WBCs (the "gold standard" in terms of low globin mRNA contamination), unblocked whole blood and blood samples in which globin cDNA synthesis is blocked using Gene Logic's globin reduction protocol.

Materials and Methods

Blood from 6 donors was collected into vacutainer tubes containing K3-EDTA (Becton Dickinson) and was then pooled into a single sample. Within 30 minutes of collection, pooled blood was processed as follows:

WBC - Ten ml of pooled blood was utilized in the “RNeasy Midi Protocol for Isolation of Total Cellular RNA from Whole Blood” protocol (Qiagen), termed WBC in following paragraphs.

PAX - The remaining blood was aliquotted into PAXgene™ Blood RNA tubes (PreAnalytiX) and subsequently used for total RNA isolation following the “PAXgene™ Blood RNA Kit” (PreAnalytiX) protocol.

The quality of each RNA preparation was individually confirmed using the Agilent 2100 Bioanalyzer system. For sample amplification and labeling, we used the Affymetrix IVT kit with five 5 µg aliquots of RNA from both the WBC RNA pool and PAXgene™ total RNA pool. An additional six 5 µg aliquots of PAXgene™ total RNA were processed using our globin reduction protocol. At the start of the first strand cDNA synthesis reaction, total RNA is combined with 5 µl of a mixture containing globin mRNA blocking oligonucleotides and the T7-oligo (dT) primer and allowed to anneal at 70°C for 10 minutes. All “blocked” samples were then processed exactly as “unblocked” samples following the standard Affymetrix IVT protocol. All samples were applied to Affymetrix HGU133 2.0 GeneChip® microarrays following the manufacturer’s recommendation for hybridization, washing and scanning.

Results and Discussion

Successful blocking of globin transcription can be easily visualized by gel analysis of the cRNA product following sample amplification and labeling. An agarose gel image of representative samples from WBC, unblocked and globin-blocked PAXgene RNA are shown in Figure 1. cRNA from unblocked blood in lanes 3 and 4 show a distinct band at about 600 bases corresponding to globin cRNA transcripts. The absence of this band in lanes 5 and 6 indicates that addition of the Gene Logic blocking oligonucleotides effectively suppressed

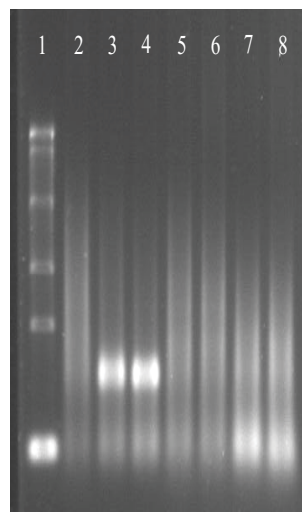


Figure 1. Addition of globin blockers inhibits globin cRNA synthesis. The amplified cRNA from several IVT reactions was assessed via agarose gel electrophoresis (1X MOPS, 1.25% agarose). cRNA from globin templates runs at about 600 bases. Lane1: sizing standard, Lane2: WBC, Lanes 3 & 4: unblocked, Lanes 5&6: blocked.

reverse transcription (RT) of globin cDNAs, greatly reducing the amount of globin template during the IVT amplification step that generates cRNA.

Successful globin blocking is also apparent after examination of MAS5 signal intensities of globin probe sets on Affymetrix HU133 2.0 GeneChip® microarrays (Table 1). The primary effect of the globin blocking protocol is to suppress α -, β -, and γ -globin signals since the blocking oligonucleotides specifically target these globins. The secondary effect on other globins including δ -, ϵ -, θ -, and ζ -globin probe sets are likely to result from cross hybridization of the targeted globins.

Probe Set Name	Type	Mean MAS 5.0 Signal Intensity			Average % Change in Intensity	
		WBC	Unblocked	Blocked	Unblocked Vs. WBC	Blocked vs. WBC
204018_x_at	alpha 2	10903	20070	8603	88 +/- 11 %	(16) +/- 9%
209458_x_at		11497	20204	10058		
211699_x_at		10259	18683	8063		
211745_x_at		12748	23043	10050		
214414_x_at		8432	16761	8522		
217414_x_at		8917	18258	6804		
1562981_at	beta	18	31	41	85 +/- 19%	(31) +/- 106%
209116_x_at		11671	24701	1374		
211696_x_at		10886	19933	2106		
217232_x_at		10200	17452	1661		
206919_at	epsilon 1	12	136	19	1033 +/- 464%	0.13 +/- 82%
217683_at		31	148	13		
204419_x_at	gamma A	395	12954	1523	3008 +/- 243%	264 +/- 31%
204848_x_at		442	12979	1513		
206834_at	delta	275	7127	2268	2193 +/- 9%	1308 +/- 10%
213515_x_at	gamma G	507	10004	1908		
220807_at	theta 1	44	524	898		
206647_at	zeta	13	443	311		

Table 1. MAS 5.0 Signal Reduction in Globin Probe Sets after Blocking.

A Pearson Correlation heat map, generated with log2 transformed MAS 5.0 signal intensities (Figure 2), shows that globin-blocked samples have a much higher correlation to WBC samples (0.899 ± 0.004) than the unblocked samples (0.853 ± 0.005). Additionally, technical replicates among the globin-blocked samples show higher intra-group average correlation (0.920 ± 0.006) than unblocked samples (0.878 ± 0.005) indicating better sample reproducibility. A histogram showing MAS 5.0 signal distribution (Figure 3) also confirms that the array signal characteristics of globin-blocked samples more closely resemble the signal distribution trends observed with WBC than with unblocked samples.

One of the biggest drawbacks to using cRNA derived from unblocked blood on Affymetrix GeneChip® microarrays is the drastic reduction in the number of probe sets called “present” by the MAS 5.0 detection algorithm when compared to WBCs isolated from the same sample. Figure 4A illustrates that the use of globin blockers restores the number of probe sets called present to levels similar to that observed with WBCs. The concordance or agreement in

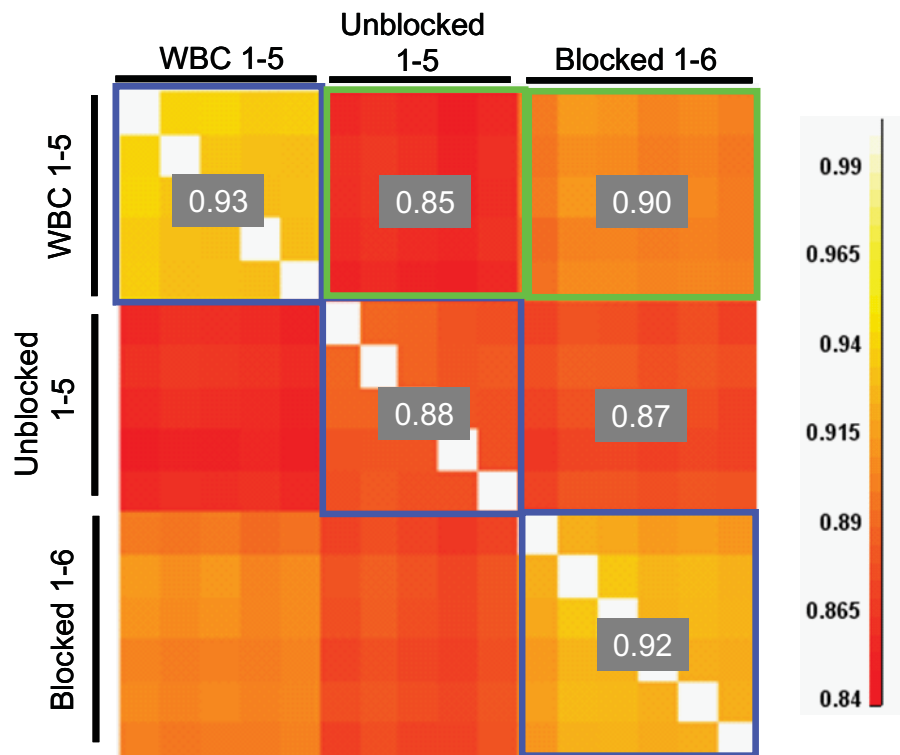


Figure 2. Pearson's Correlation Heatmap. Pearson's Correlation was calculated across log₂ (signal) over all probe sets for each sample. Blocked samples exhibit higher intra-sample correlation and higher correlation to WBC samples than unblocked samples. Mean correlation for each group is indicated numerically within the gray box.

detection call between methods for any given probe set is also significantly improved when comparing WBC to globin-blocked samples as opposed to unblocked (Figure 4B).

Differential expression analysis suggests that the globin-blocking protocol shows fewer significant expression differences to WBC samples as compared to unblocked samples (Figure 5). We observed 9050 probe sets with a p value <0.01 and a fold change of +/- 2 fold when comparing unblocked- to WBC-derived expression data. In contrast, a similar comparison between WBCs and globin-blocked samples revealed 4610 probe sets showing a significant change between methods. The additional 4440 probe sets are generally all down-regulated which supports the observation of reduced % present calls in the unblocked samples. Although there are fewer differences between WBC and globin-blocked samples, these results show many significant changes still exist between WBC and whole blood even with the use of a globin-blocking protocol. The reasons for this may be due to differences in collection methods or differences in cell populations related to the methods for blood processing.

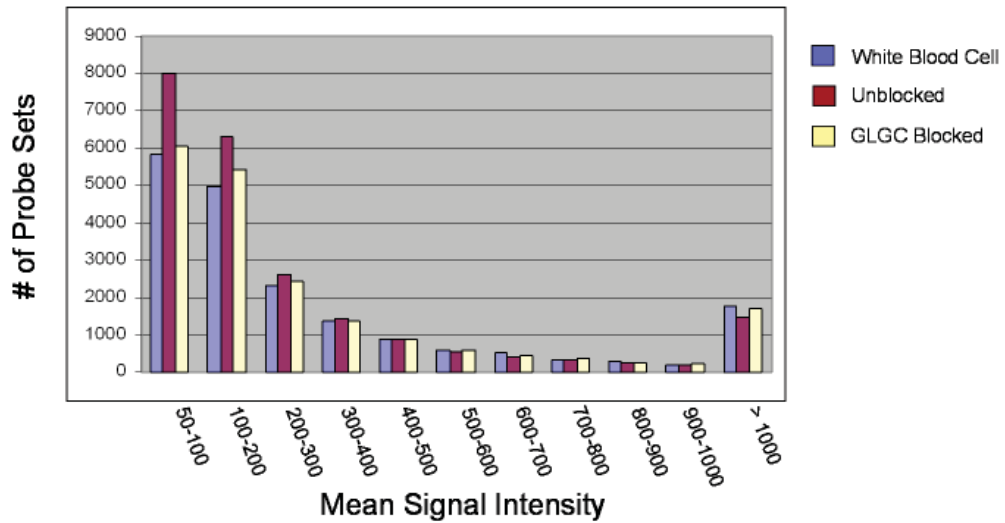


Figure 3. Average MAS 5.0 Signal Distribution. Signal intensities across group replicates were averaged and binned as indicated. The number of probe sets within each bin is plotted. Each protocol resulted in > 30,000 probe sets between 0 and 50 mean signal intensity (data point not shown). Unblocked resulted in ~ 3,000 fewer probe sets in 0 – 50 region than either the WBC or blocking protocols.

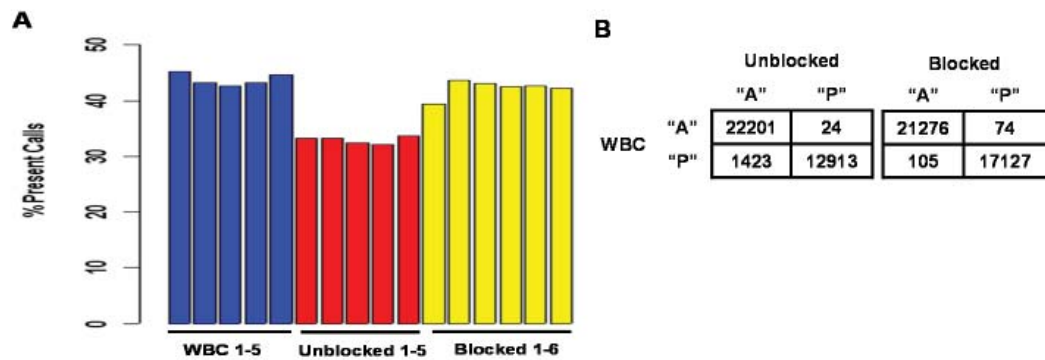


Figure 4. Present Call Comparison. (A) Of the ~55,000 probe sets on the HU133Plus array, the percentage called Present by the MAS5 detection algorithm is indicated for each array in the study. (B) Concordance of detection calls between WBC and Unblocked or Blocked arrays. For the concordance tables, a probe set had to be called either P or A on all replicates within a sample type to be counted.

Conclusions

The high levels of globin mRNA transcripts in whole blood lead to problems with sensitivity and reproducibility of microarray expression data. To overcome these problems, the fractionation of white blood cells prior to sample preparation is often required for the best data quality. The Gene Logic globin-blocking procedure is an easier technical alternative to WBC fractionation yet gives comparable data quality. Our globin-blocking method is also simpler than other commercially available methods since it involves the simple addition of an oligonucleotide mix at the cDNA synthesis step.

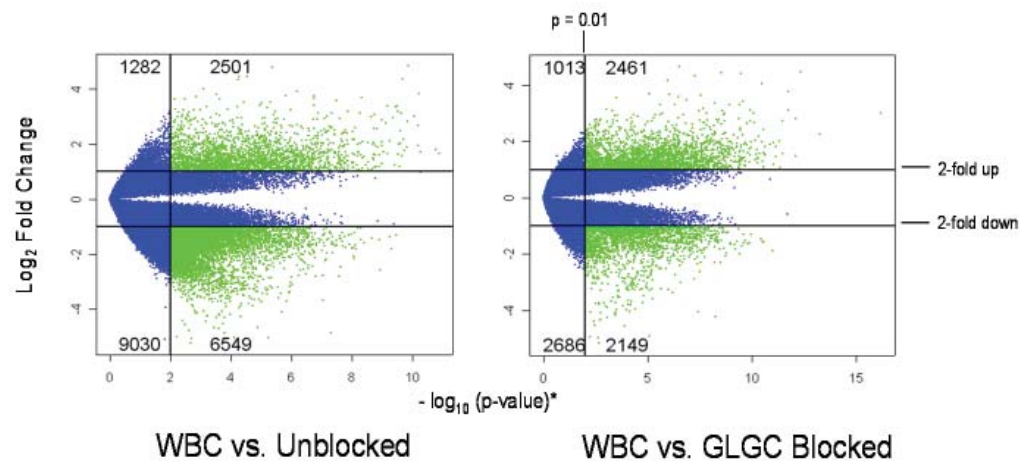


Figure 5. Volcano plots are used to visualize differential expression between WBC and unblocked or globin blocked blood samples. Probe sets indicating a +/- 2 fold change with an associated p-value of ≤ 0.01 between methods are shown in green. The calculated p-values are based on individual Student's two-tailed t-tests of unequal variance of log₂ transformed MAS5 signal values with no corrections for multiple comparisons.

The observed advantages of using the Gene Logic globin-blocking procedure as opposed to whole unblocked blood on array data quality include a greater degree of correlation to white blood cell expression relative to unblocked samples as well as within technical replicates relative to unblocked protocol. A greater concordance to white blood cell samples is also observed based on the number of present calls, agreement in present calls, distribution of signal intensities and fold change/t-test comparisons. Thus, our globin reduction method provides a viable alternative to WBC isolation prior to expression analysis on microarrays.

The use of Gene Logic's blocking method should be applicable to any commercially-available microarray which utilizes labeled cRNA as target. In addition to the human globin-blocking protocol described here, Gene Logic has also developed and validated proprietary globin-blocking oligonucleotide mixtures for use in processing whole blood samples from other species including rat, mouse, canine, and non-human primates.

Products highlighted in this Application Note include the following:

Microarray Data Generation and Analysis™ Services

Gene Logic is a leading producer of commercial Affymetrix microarray data, and has developed proprietary methodologies and software tools for managing the process to the highest standards. Our multidisciplinary staff includes pathologists, biologists, microbiologists, nurses, biostatisticians, and software experts. Our information management systems facilitate proper handling of samples from receipt to delivery. State-of-the-art facilities, automation, and stringent quantitative and qualitative quality control metrics ensure consistency and quality.

Gene Logic offers a wide range of flexible service options:

- Gene expression data generation from tissue, cells, blood, or RNA
- Optimized gene expression profiling of whole blood using a proprietary protocol to reduce globin interference
- RNA isolation from low-yield tissues or samples that are difficult to process
- Custom bioinformatics analysis for biomarker or drug target identification and mechanisms of toxicity
- Custom platform development
- SNP genotyping from DNA, blood, or tissue

Visit Gene Logic's website for more information at www.genelogic.com, or call us at 1.800.GENELOGIC or +1.301.987.1709.



610 Professional Drive • Gaithersburg, Maryland 20879 • Phone: 1.800.GENELOGIC • +1.301.987.1709 • Fax: 301.987.1701
Email: info@genelogic.com • Web: www.genelogic.com

© 2005 Gene Logic Inc. All Right Reserved. Microarray Data Generaation and Analysis is a trademark of Gene Logic Inc. All other trademarks are the property of their respective owner.