

RESEARCH ARTICLE

Open Access

Gene expression profiling of human whole blood samples with the Illumina WG-DASL assay

Mary E Winn^{1,2}, Marian Shaw^{2,3}, Craig April⁵, Brandy Klotzle⁵, Jian-Bing Fan⁵, Sarah S Murray^{2,3,4} and Nicholas J Schork^{2,3,4*}

Abstract

Background: Microarray-based gene expression analysis of peripheral whole blood is a common strategy in the development of clinically relevant biomarker panels for a variety of human diseases. However, the results of such an analysis are often plagued by decreased sensitivity and reliability due to the effects of relatively high levels of globin mRNA in whole blood. Globin reduction assays have been shown to overcome such effects, but they require large amounts of total RNA and may induce distinct gene expression profiles. The Illumina whole genome DASL assay can detect gene expression levels using partially degraded RNA samples and has the potential to detect rare transcripts present in highly heterogeneous whole blood samples without the need for globin reduction. We assessed the utility of the whole genome DASL assay in an analysis of peripheral whole blood gene expression profiles.

Results: We find that gene expression detection is significantly increased with the use of whole genome DASL compared to the standard IVT-based direct hybridization. Additionally, globin-probe negative whole genome DASL did not exhibit significant improvements over globin-probe positive whole genome DASL. Globin reduction further increases the detection sensitivity and reliability of both whole genome DASL and IVT-based direct hybridization with little effect on raw intensity correlations. Raw intensity correlations between total RNA and globin reduced RNA were 0.955 for IVT-based direct hybridization and 0.979 for whole genome DASL.

Conclusions: Overall, the detection sensitivity of the whole genome DASL assay is higher than the IVT-based direct hybridization assay, with or without globin reduction, and should be considered in conjunction with globin reduction methods for future blood-based gene expression studies.

Background

Peripheral whole blood is an attractive source of mRNA for the identification, examination, and development of disease biomarkers via microarray-based gene expression [1]. In fact, many studies have explored the utility of gene expression patterns in whole blood for the purposes of classifying or predicting clinical conditions [2-4]. However, the sensitivity and specificity of microarray assays using peripheral whole blood are reduced due to the relatively high proportion of globin mRNA present in total RNA, which obscures the detection of transcripts expressed at low levels in whole blood [5,6]. While globin reduction assays have been shown to overcome these

effects when used in conjunction with Affymetrix microarrays [7] and the standard Illumina direct hybridization assay [8,9], globin reduction assays require large amounts of total RNA [7], fail to completely eliminate globin transcripts [7], and may induce distinct gene expression profiles [10]. Consequently, methods of developing blood-based gene expression biomarker panels that do not involve globin reduction are needed. Developing a microarray-based gene expression assay that does not rely on globin reduction or other methods of sample fractionation, such as the isolation of PBMCs or other cell types from the blood, should reduce sample variability introduced by sample handling and preparation. This will result in a more accurate reflection of the transcriptome at the time of blood draw, and will reduce time and cost.

There are ways to eliminate the need for globin reduction including 1.) the removal of globin probes from the

* Correspondence: nschork@scripps.edu

²Scripps Genomic Medicine and Scripps Translational Science Institute, The Scripps Research Institute, La Jolla, CA 92037, USA

Full list of author information is available at the end of the article

microarray; and 2.) the elimination of globin transcript amplification. Originally developed for the profiling of partially degraded and fixed RNA samples, the highly sensitive and reproducible Illumina cDNA-mediated annealing, selection, extension and ligation (DASL) assay [5,11] uses random priming and a modifiable oligo pool for cDNA synthesis. Random priming in conjunction with PCR amplification may allow for the increased detection of low abundance transcripts. In addition, removing globin-specific oligos from the DASL Assay Oligo Pool (DAP) should decrease noise associated with the high abundance of globin mRNA transcripts and potentially eliminate the necessity of globin reduction. Currently, the DAP is available with and without globin-specific oligos. In order to assess the need for globin reduction with the Illumina DASL assay, we compared microarray gene expression profiles of peripheral blood total RNA and globin-reduced RNA amplified via *in vitro* transcription (IVT)-based direct hybridization, DASL with globin-specific oligos, and DASL without globin-specific oligos.

Methods Summary

Peripheral whole blood samples were collected from eight human donors in PAXGene blood RNA tubes. RNA was isolated after freezing and storage and then prepared for gene expression analysis using the Illumina Human-Ref8 v3.0 Beadchip. Alpha and beta globin were reduced from a portion of the total RNA using the GLOBINclear assay (Ambion, Austin, TX, USA). Two methods of microarray target preparation were examined: Illumina IVT-based direct hybridization (IVT) and Illumina Whole-Genome DASL (WG-DASL) (Figure 1). The differences between IVT and WG-DASL are outlined in Table 1. Two DASL Assay Oligo pools (DAP) were utilized for DASL target preparation: the DASL Assay Oligo Pool with globin probes (DAP +) and the DASL Assay Oligo Pool without globin probes (DAP-). Comparisons involving the number of genes whose expression levels were detected and the actual levels of expression of the genes were made across the different platforms. A more complete description of the methods is provided in the Methods section.

Results

Comparison between IVT and WG-DASL with and without globin reduction

Following target amplification as outlined in Figure 1, samples were hybridized with the Illumina Human-Ref8 v 3.0 following the manufacturer's instructions. Each target preparation method was assessed for performance by the number of probes detected as present (Detection p -value < 0.05) (Figure 2). Probes are generally detected as present if the probe intensity is significantly increased in comparison to the array background intensity. As noted, high

levels of background due to the presence of globin transcripts in whole blood are known to decrease the number of significantly detected probes. The WG-DASL target preparation method significantly improved detection sensitivity compared to IVT (p -value = 2.13×10^{-9} from an analysis of variance (ANOVA)). Globin reduction decreased probe detection variability with both IVT and WG-DASL target preparation methods. The removal of globin probes from the DASL assay oligo pool (DAP-) resulted in a moderate increase in the number of probes detected but had no significant effect on detection variability (p -value = 0.680, ANOVA) as compared to the DAP+ target preparation method. Overall, 8677 probes were detected across all samples by the five target preparation methods (Figure 3), but only 867 probes were detected by IVT alone. 2604 probes were detected by WG-DASL alone.

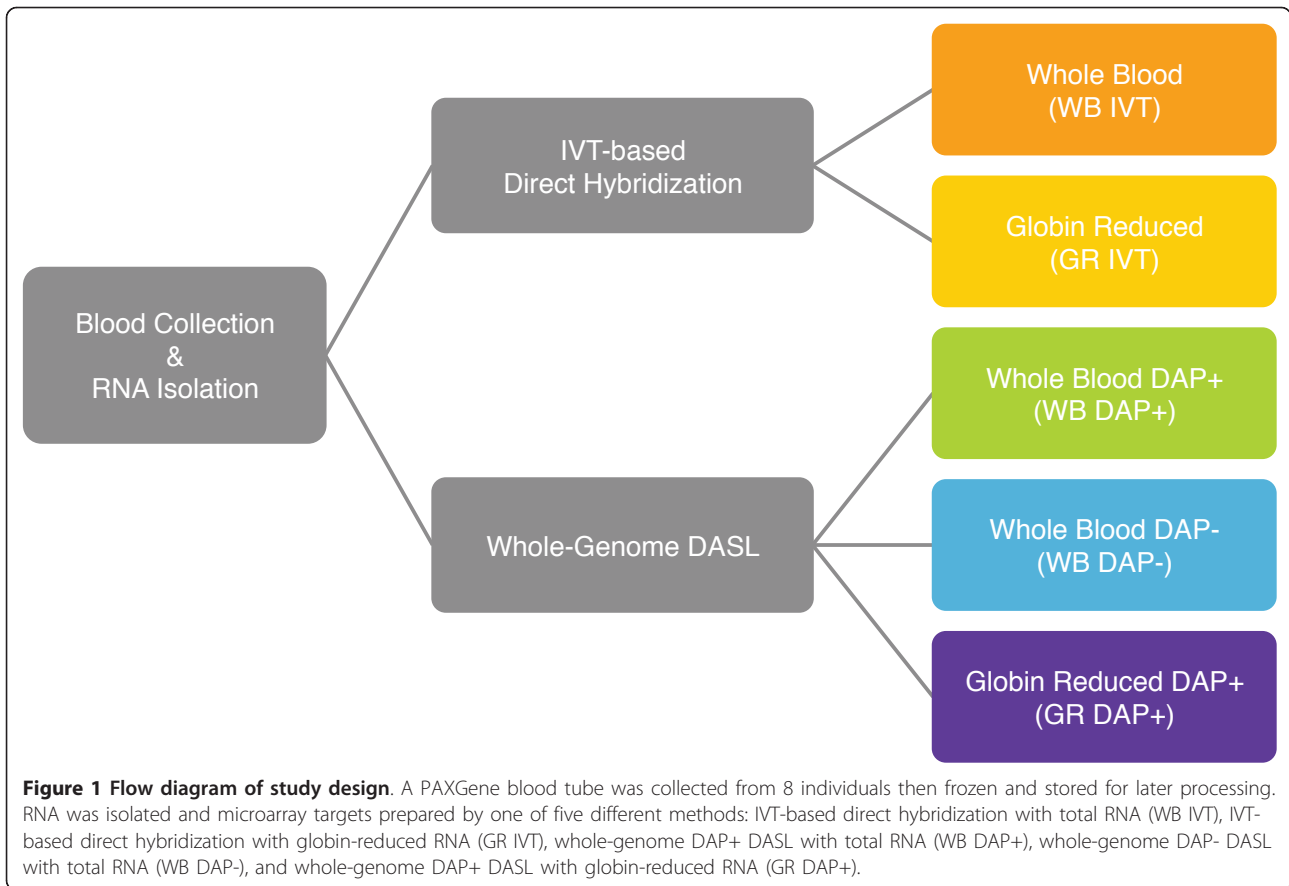
NanoDrop Spectrophotometer 260/280 ratios were moderately decreased following globin reduction with an average ratio equal to 2.06 prior and 1.97 post globin reduction (Table 2). However, raw intensity correlations indicate that whole and globin-reduced blood yield similar expression profiles with both IVT and DASL DAP+ assays. Overall raw intensity values increased in globin reduced samples (Figure 4A, C-D) despite the failure of GLOBINclear to completely eliminate the two most abundant globin transcripts, hemoglobin alpha (HBA2) and hemoglobin beta (HBB). The removal of globin probes from the DASL Assay Oligo Pool (DAP-) (Figure 4B) had little effect on gene expression profiles compared to DAP+ ($R^2 = 0.993$) despite the near complete elimination of HBA2 and HBB.

Expression patterns maintained across target preparation methods

IVT target amplification is approximately linear while WG-DASL is approximately exponential, making it difficult to compare expression intensities directly. Thus, it was important in our analyses that the sample-to-sample relations are maintained among each target preparation method. Despite the differences in target amplification, sample relations were preserved across the five target preparation methods as shown by unsupervised hierarchical clustering (Figure 5). For example, with both IVT and WG-DASL, expression profiles for Sample 3 and Sample 7 exhibited the greatest differences from the other six samples, while for the IVT or WG-DASL whole blood RNA clustered separately from globin reduced RNA.

Discussion

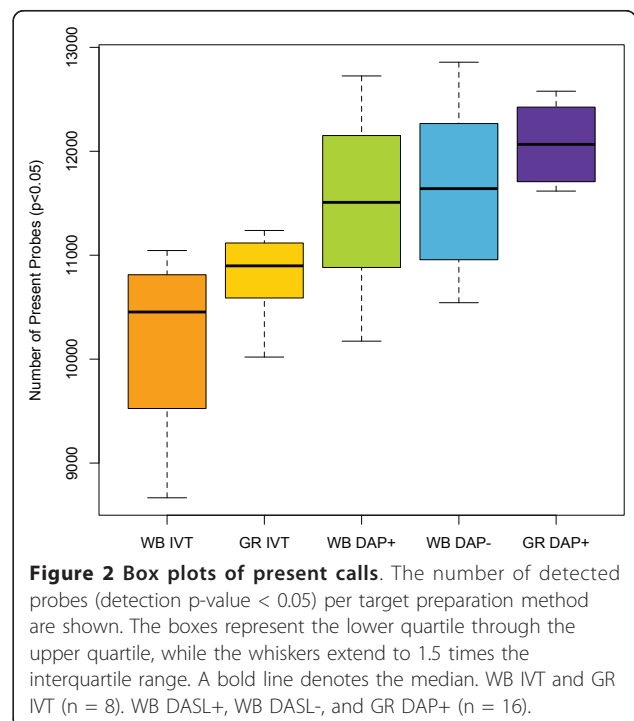
The analysis of whole blood, microarray-based gene expression profiles is often hindered by low sensitivity and high variability due to high levels of globin mRNA transcripts. These issues have been addressed by the

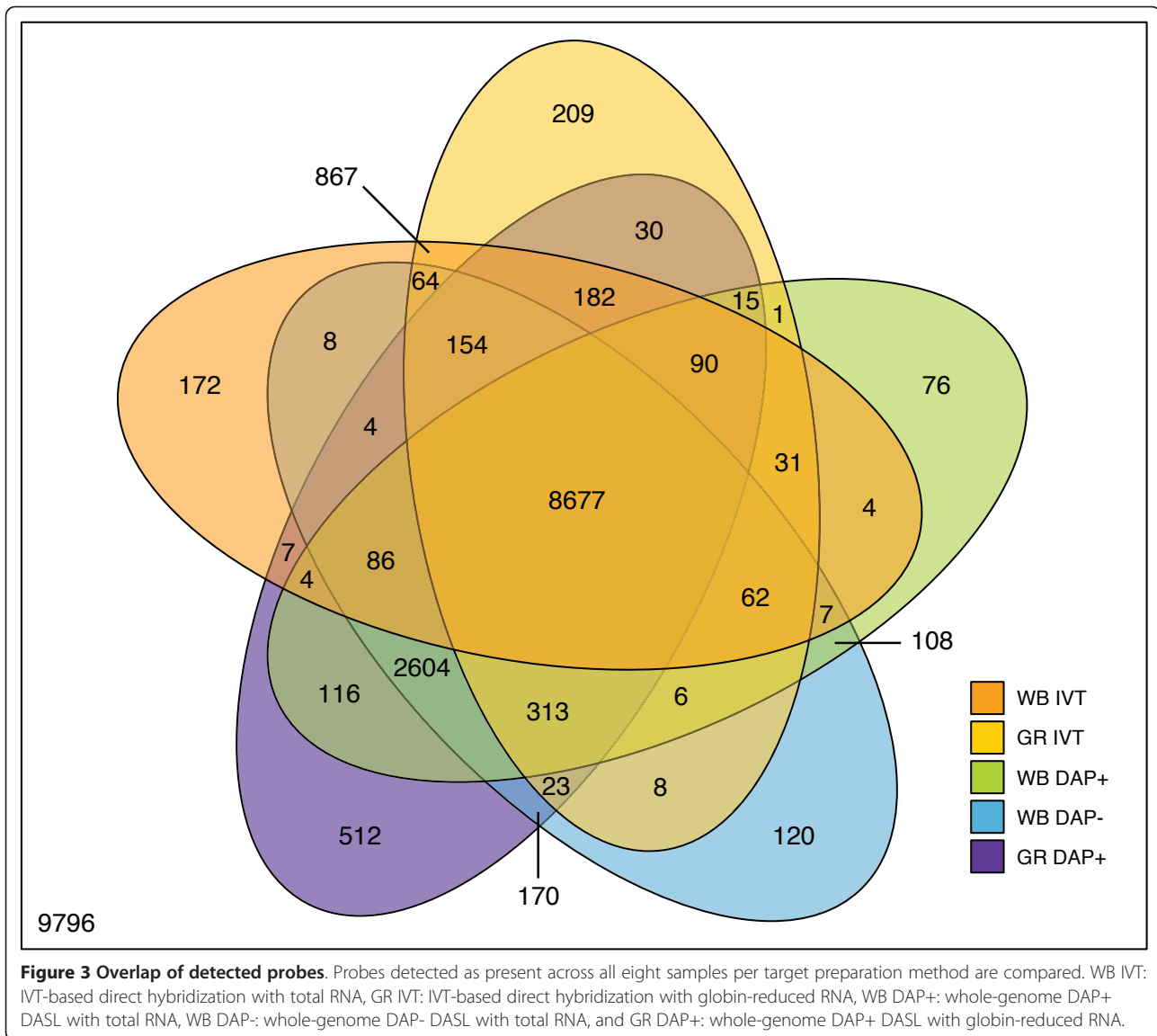


development of globin-reduction methods, which specifically target and remove globin transcripts prior to array hybridization. However, studies have shown that globin reduction, like other methods of sample fractionation, may alter expression profiles [10], require large amounts of sample input, increase sample variability [7], and lead to increased costs. Thus, the ability to assay whole blood without sample fractionation or globin reduction may result in improved gene expression profile quality and decrease cost.

Table 1 Summary of IVT and WG-DASL Methods

	IVT	WG-DASL
Target Preparation Protocol Name	<i>In Vitro</i> Transcription	cDNA-mediated annealing, selection, extension and ligation
Total RNA Input Amount	50-100 ng	10-200 ng
Priming Method	Reverse Transcription off polyA tail	Poly(T) and random priming with biotinylated nonamers
Amplification	<i>In Vitro</i> Transcription (Linear)	PCR (Exponential)
Hybridization	Illumina BeadChip	





Here we describe the utility of a highly sensitive, whole-genome assay in the assessment of whole blood gene expression. Our results suggest that gene expression detection sensitivity is significantly increased with the whole-genome cDNA-mediated annealing, selection, extension and ligation (WG-DASL) assay as compared to IVT-based direct hybridization (IVT). The increased detection sensitivity of WG-DASL may be due to, 1.) random priming allowing for cDNA synthesis along the length of mRNA transcripts, or 2.) the ability to produce larger amounts of cDNA with PCR amplification. Regardless, attempts to further improve detection sensitivity and decrease expression variability through the selective removal of globin probes from the DASL assay oligo pool (DAP-) did not exhibit any large improvements over globin-probe positive DASL (DAP+). Our study also confirms the positive effect

of globin reduction on microarray quality when used in conjunction with the Illumina BeadChip and standard IVT-based hybridization [9], while showing that the positive effect of globin reduction extends to WG-DASL as well. However, as shown by unsupervised hierarchical clustering analysis, globin reduction appears to mildly influence gene expression profiles produced by both IVT and WG-DASL assays. Whether this is due to the induction of a globin reduction-specific profile [10], reduced RNA quality due to globin reduction, or the result of decreased noise is unknown, and should be taken into consideration while planning blood-based gene expression experiments.

Conclusions

Overall, our results suggest that the detection sensitivity of the WG-DASL assay is higher than the IVT-based

Table 2 RNA quality as assessed by 260/280 ratio

Sample ID	Before Amplification	Amplified	1st Globin Reduction	2nd Globin Reduction	Pooled Globin Reduction	Diluted to 20 ng/ul	Average
C00023 (tRNA)	2.03	2.01				2.07	2.04
C00023 (GC RNA)	1.98	1.99	2.02	1.89	2.00	1.94	1.97
C00027 (tRNA)	2.06	2.01				2.06	2.04
C00027 (GC RNA)	1.98	2.01	2.02	1.95	1.94	1.89	1.97
C00169 (tRNA)	2.04	2.12				1.89	2.02
C00169 (GC RNA)	1.96	1.95	1.88	1.90	2.02	1.84	1.93
C00179 (tRNA)	2.04	1.99				2.12	2.05
C00179 (GC RNA)	1.95	1.98	1.90	2.01	2.01	2.02	1.98
C00275 (tRNA)	2.03	1.95				2.14	2.04
C00275 (GC RNA)	1.99	1.99	1.87	1.91	1.93	2.01	1.95
C00304 (tRNA)	2.05	1.99				2.25	2.10
C00304 (GC RNA)	1.98	2.00	1.89	2.11	1.91	1.96	1.98
C00311 (tRNA)	2.04	2.10				2.12	2.09
C00311 (GC RNA)	2.02	2.00	2.04	1.93	1.97	2.01	2.00
C00342 (tRNA)	2.01	2.07				2.21	2.10
C00342 (GC RNA)	2.03	2.03	1.97	1.99	1.93	1.97	1.99

RNA quality was assessed before and after globin reduction as well as before and after amplification. tRNA: total RNA; GC RNA: GLOBINclear treated RNA or globin reduced RNA.

direct hybridization assay, with or without globin reduction, and should be considered in conjunction with globin reduction methods for future blood-based gene expression studies. However, further investigation into the ability of the WG-DASL assay to distinguish between disease populations using whole blood is needed, as our study was not designed to address such issues.

Methods

Blood collection and RNA isolation

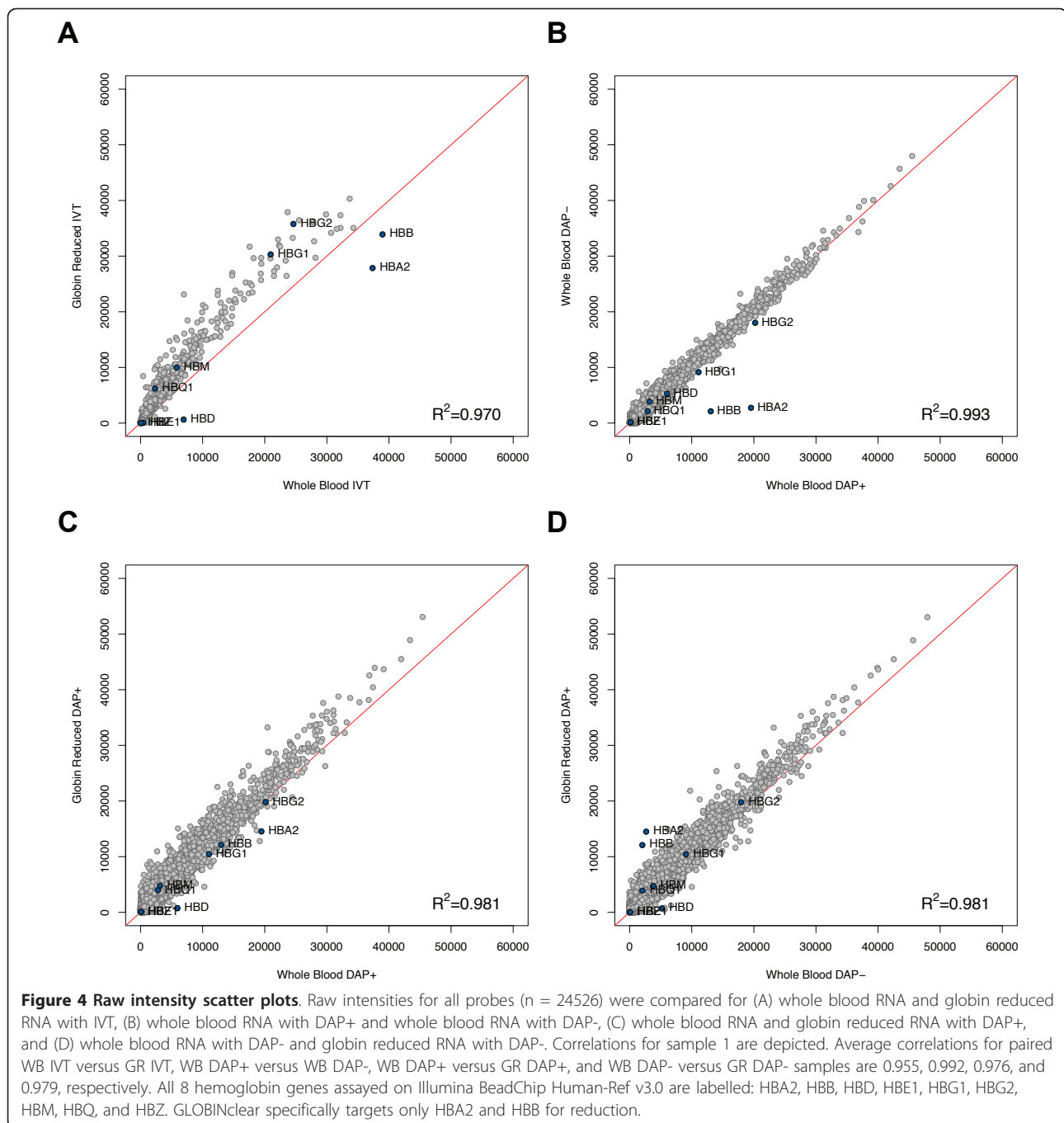
For each sample, 2.5 ml whole blood was collected in a PAXgene Blood RNA collection tube (Qiagen, Valencia, CA, USA) and stored frozen at -80°C prior to RNA isolation. RNA isolation was performed using the PAX-Gene Blood RNA Isolation System (Qiagen, Valencia, CA, USA). RNA quantity and quality were assessed by NanoDrop[®] Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) before and after globin reduction as well as before and after RNA amplification. For the 8 samples isolated, the total RNA yield ranged from 5.8 - 13.8 ug (average 7.9 ug +/- 1.0 ug), while A260/A280 ratios revealed all samples appeared to be of sufficient quality for microarray analysis (1.93 - 2.10) (Table 2), despite a moderate decrease in quality following globin reduction.

Globin Reduction

Alpha and beta globin mRNA were reduced from a portion of the total RNA samples using the GLOBINclear[™] Human kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions with the recommended start quantity of 2 µg of total RNA. Each sample was processed twice then globin-reduced RNA pooled prior to RNA amplification and hybridization.

RNA amplification and hybridization

Whole blood total RNA and globin-reduced samples were assayed at both Scripps Genomic Medicine (La Jolla, CA, USA) and Illumina (San Diego, CA, USA) for IVT and DASL-based labelling, hybridization, and scanning, respectively (Table 1). Briefly, the WG-DASL method utilizes biotinylated random nonamer and oligo (dT) primers to convert 10-200 ng input RNA to cDNA. The biotinylated cDNA is then immobilized to a streptavidin-coated solid support and annealed to a pool of gene-specific oligonucleotides (DAP) for extension and ligation followed by PCR amplification with a biotinylated and a fluorophore-labeled universal primer. Finally, the single-stranded PCR products are eluted and hybridized to an Illumina BeadChip. For this study, 250 ng and 100 ng input RNA were utilized for IVT and DASL, respectively.



Gene expression analysis was performed on all whole blood RNA and globin-reduced samples using Human-Ref8 v3.0 Beadchips (Illumina, San Diego, CA, USA) containing 24,526 probes. All arrays were scanned with the Illumina BeadArray Reader and read into Illumina GenomeStudio[®] software (version 1.1.1). Individual samples were assayed once for all IVT analyses and twice for all DASL analyses. Given the limited amount of mRNA, replicates were only performed for the DASL assay due to its relative novelty as compared to the IVT

assay. All replicates were highly correlated (average $R^2 = .9925$). All raw data is available on the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>, [GSE 28064]).

Microarray data analysis

Raw intensities values were exported from GenomeStudio[®] software (version 1.1.1) for data processing and analysis in R (<http://www.R-project.org>) and Bioconductor (<http://www.bioconductor.org>) [12]. Data quality and

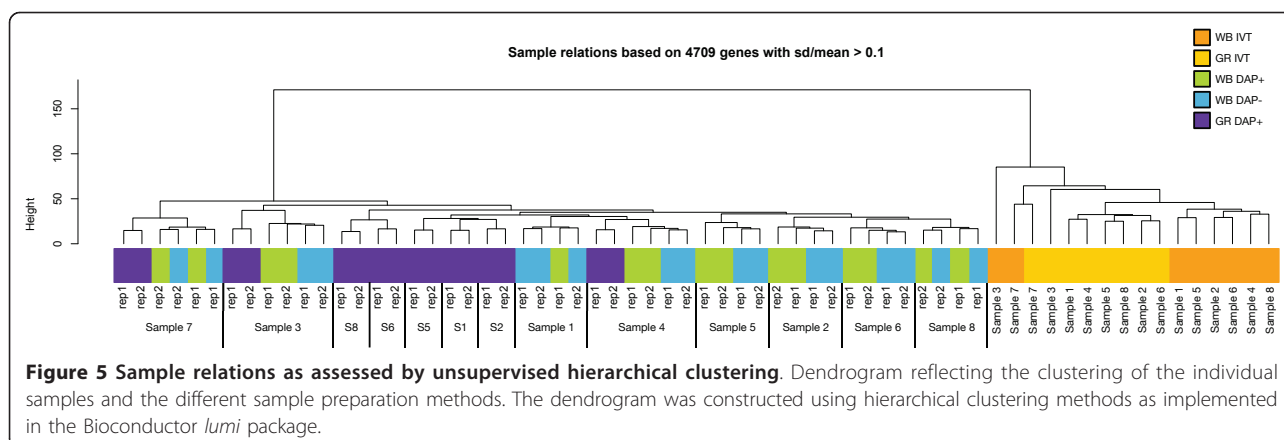


Figure 5 Sample relations as assessed by unsupervised hierarchical clustering. Dendrogram reflecting the clustering of the individual samples and the different sample preparation methods. The dendrogram was constructed using hierarchical clustering methods as implemented in the Bioconductor *lumi* package.

sample relations were assessed using the Bioconductor *lumi* package [13]. Probes with a Detection *p*-value less than 0.05 were considered present. Analysis of Variance (ANOVA) was used to assess the consistency of present/absent calls across the different sample preparation methods. Correlation coefficients were calculated from the raw intensity levels to assess the similarity of expression profiles.

Abbreviations

cDNA: complementary deoxyribonucleic acid; DAP+: DASL Assay Oligo Pool with globin probes; DAP-: DASL Assay Oligo Pool without globin probes; DASL: cDNA-mediated annealing, selection, extension and ligation; GR: RNA following globin reduction by GLOBINclear; HBA2: hemoglobin, alpha 2; HBB: hemoglobin, beta; HBD: hemoglobin, delta; HBE1: hemoglobin, epsilon; HBG1: hemoglobin, gamma A; HBG2: hemoglobin, gamma G; HBM: hemoglobin, mu; HBQ: hemoglobin, theta 1; HBZ: hemoglobin, zeta; IVT: in vitro-transcription; mRNA: messenger RNA; WB: total RNA from peripheral whole blood; WG: whole genome.

Acknowledgements

This work was supported in part by the following research grants: U19 AG023122-05; R01 MH078151-03; N01 MH22005; U01 DA024417-01; P50 MH081755-01; R01 AG030474-02; N01 MH022005; R01 HL089655-02; R01 MH080134-03; U54 CA143906-01; UL1 RR025774-03 as well as the Price Foundation and Scripps Genomic Medicine.

Author details

¹Graduate Program in Biomedical Sciences, Department of Medicine, University of California at San Diego, La Jolla, CA 92093, USA. ²Scripps Genomic Medicine and Scripps Translational Science Institute, The Scripps Research Institute, La Jolla, CA 92037, USA. ³Scripps Health, La Jolla, CA 92037, USA. ⁴Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA. ⁵Illumina Inc., San Diego, CA 92121, USA.

Authors' contributions

MEW participated in the design of the study, performed all data analysis, and drafted the manuscript. MS carried out the IVT-based microarray assays. CA and BK carried out the DASL-based microarray assays. JF and SSM participated in the design of the study. NJS conceived of and participated in the design and coordination of the study and helped draft the manuscript. All authors read and approved the final manuscript.

Received: 19 April 2011 Accepted: 15 August 2011
 Published: 15 August 2011

References

- Rockett JC, Burczynski ME, Fornace AJ, Herrmann PC, Krawetz SA, Dix DJ: **Surrogate tissue analysis: monitoring toxicant exposure and health status of inaccessible tissues through the analysis of accessible tissues and cells.** *Toxicol Appl Pharmacol* 2004, **194**:189-199.
- Hoang LT, Lynn DJ, Henn M, Birren BW, Lennon NJ, Le PT, Duong KTH, Nguyen TTH, Mai LN, Farrar JJ, Hibberd ML, Simmons CP: **The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults.** *J Virol* 2010, **84**:12982-12994.
- Lin D, Hollander Z, Ng RT, Imai C, Ignaszewski A, Balshaw R, Freue GC, Wilson-McManus JE, Qasimi P, Meredith A, Mui A, Triche T, McMaster R, Keown PA, McManus BM: **Whole blood genomic biomarkers of acute cardiac allograft rejection.** *J Heart Lung Transplant* 2009, **28**:927-935.
- Takahashi M, Hayashi H, Watanabe Y, Sawamura K, Fukui N, Watanabe J, Kitajima T, Yamanouchi Y, Iwata N, Mizukami K, Hori T, Shimoda K, Ujike H, Ozaki N, Iijima K, Takemura K, Aoshima H, Someya T: **Diagnostic classification of schizophrenia by neural network analysis of blood-based gene expression signatures.** *Schizophr Res* 2010, **119**:210-218.
- Fan J, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, McBride C, Li H, Fu X, Oliphant A, Barker DL, Chee MS: **A versatile assay for high-throughput gene expression profiling on universal array matrices.** *Genome Res* 2004, **14**:878-885.
- Wright C, Bergstrom D, Dai H, Marton M, Morris M, Tokiwa G, Wang Y, Fare T: **Characterization of globin RNA interference in gene expression profiling of whole-blood samples.** *Clin Chem* 2008, **54**:396-405.
- Vartanian K, Slotke R, Johnstone T, Casale A, Planck SR, Choi D, Smith JR, Rosenbaum JT, Harrington CA: **Gene expression profiling of whole blood: comparison of target preparation methods for accurate and reproducible microarray analysis.** *BMC Genomics* 2009, **10**: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.
- Tian Z, Palmer N, Schmid P, Yao H, Galdzicki M, Berger B, Wu E, Kohane IS: **A practical platform for blood biomarker study by using global gene expression profiling of peripheral whole blood.** *PLoS ONE* 2009, **4**:e5157.
- Liu J, Walter E, Stenger D, Thach D: **Effects of globin mRNA reduction methods on gene expression profiles from whole blood.** *J Mol Diagn* 2006, **8**:551-558.
- April C, Klotzle B, Royce T, Wickham-Garcia E, Boyaniwsky T, Izzo J, Cox D, Jones W, Rubio R, Holton K, Matulonis U, Quackenbush J, Fan J: **Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples.** *PLoS ONE* 2009, **4**:e8162.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**:R80.

13. Du P, Kibbe WA, Lin SM: lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008, **24**:1547-1548.

doi:10.1186/1471-2164-12-412

Cite this article as: Winn *et al.*: Gene expression profiling of human whole blood samples with the Illumina WG-DASL assay. *BMC Genomics* 2011 **12**:412.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

