Gene Expression Profiling of Prokaryotic Samples using Low Input Quick Amp WT Kit

Application Note

Authors

Nilanjan Guha and Becky Mullinax

Abstract

Agilent's Low Input Quick Amp Labeling WT (LIQA WT) Kit provides a reliable method to amplify and label whole transcripts for the robust generation of cRNA prior to microarray hybridization. In this application note, we describe a method to perform microarray analysis of two key prokaryotic organisms, *Escherichia coli and Mycobacterium tuberculosis*, a GC rich (65.6 % GC content) organism, using the LIQA WT kit. The study demonstrates that the LIQA WT kit can be used successfully to label prokaryotic transcripts from these two organisms, and yield high quality microarray data. The data shows strong correlation ($R^2 > 0.9$) between the replicates of the microarrays for both one- and two-color assays and excellent assay sensitivity as demonstrated by a linear RNA spike-in concentration response. Differential expression analysis of the gene expression in *E. coli* upon treatment with paraquat, a potent inducer of oxidative stress, revealed several key genes as significantly regulated. These genes include superoxide dismutase and fumarase C, both known to be strongly induced by paraquat (Liochev and Fridovich, *Proc. Natl. Acad. Sci.* USA, Vol. 89, pp.5892-5896, July 1992). Overall the data demonstrate that high quality biologically relevant expression data can be generated using the LIQA WT kit in conjunction with Agilent's catalog *E. coli* and custom *M. tuberculosis* arrays.

Agilent Technologies

Methods

Microarrays

Microarray analysis was performed using Agilent's *E. coli* Microarray Kit 8x15K, P/N G4813A, (Agilent Microarray Design ID 020097) and custom-made *M. tuberculosis* 8x15K microarrays (Agilent Microarray Design ID 034585). The *M. tuberculosis* custom microarrays were designed by a collaborator using Agilent eArray, a web-based application that enables users to design microarray probes to any organism and to create custom microarrays. The custom *M. tuberculosis* array contains 5,180 unique probes, out of which 5,120 are *M. tuberculosis* specific with most of these probes are replicated three times. The remaining 60 probes are Agilent control probes, similar to those of the *E. coli* array.

Oxidative stress induction of *E. coli*

10 ml of an overnight saturated *E. coli* culture were added to each of two flasks containing 100 ml of minimal medium. Cultures were incubated at 37 °C with shaking at 200 rpm until an OD_{600} of 1.2 was achieved. Paraquat was added to a final concentration of 300 µM, and the cultures were further incubated for 15 minutes then chilled on ice for 15 minutes. The cells were then pelleted and frozen at –80 °C.

Preparation of test RNA and RNA spike-ins

E. coli RNA was isolated using the Absolutely RNA Miniprep kit (Agilent, Catalog# 400800), and *M. tuberculosis* RNA (provided by a collaborator) was isolated using the Qiagen RNeasy kit. RNA quality was assessed on the Agilent 2100 Bioanalyzer System (P/N G2943C) using the Agilent RNA 6000 Nano kit (P/N 5067-1511) with the Prokaryote Total RNA Assay in the 2100 Expert software. All samples gave RIN (RNA Integrity Number) values of >9. One-color and two-color RNA spike-ins were added from the Agilent One-Color RNA Spike-In Kit (P/N 5188-5282) and Two-Color RNA Spike-In Kit (P/N 5188-5279), respectively. The spike-mixes were diluted in the provided dilution buffer through serial dilutions and added to the total RNA as recommended for 50 ng input RNA: 2 µl of a 1:20000 dilution of the Agilent One-Color RNA Spike-Mix for the one-color assays and 2 µl of a 1:51200 dilution of Spike A mix (for Cy3 labeled samples) or Spike B mix (for Cy5 labeled samples) for the two-color assays.

RNA labeling and hybridization

Agilent's Low Input Quick Amp WT Labeling Kit generates high yields of cyanine labeled cRNAs using a linear amplification protocol. The WT primer mix provided in the kit contains a mixture of oligo dT- and random nucleotide-based T7 promoter primers, thus enabling robust generation of cRNA from templates of both eukaryotic and prokaryotic origin. The Agilent spike-in controls are also simultaneously labeled and amplified with the RNA samples of interest. The method requires low RNA inputs in the range of 25 ng to 100 ng of total RNA per reaction.

The labeling and hybridizations were performed as per the Agilent One-Color Microarray-Based Exon Analysis-LIQA WT Labeling Protocol (publication G4140-90042, Version 1.0, November 2010) and the Agilent Two-Color Microarray-Based Exon Analysis-LIQA WT Labeling Protocol (publication G4140-90052, Version 1.0, November 2010). For all experiments, 50 ng of input total RNA were used. cRNA yields and specific activities were measured using a NanoDrop ND 1000 spectrophotometer (Thermo Scientific).

Microarray washing and scanning

All washes were performed as recommended in the above referenced publications with the exception that the number and length of washing steps were increased for *M. tuberculosis* arrays because the high GC content of the organism required more stringent wash conditions. The *M. tuberculosis* arrays were washed through two wash cycles, with each wash cycle consisting of a five minute wash in Gene Expression Wash Buffer 1 and a one minute wash in Gene Expression Wash Buffer 2. All washes with Wash Buffer 2 were performed at elevated temperature (37 °C) for both *E. coli* and *M. tuberculosis* as per Agilent one-color and two-color microarray protocols. This optimized wash condition ensured minimal non-specific binding and low background levels. The slides were scanned with the Agilent High Resolution C Scanner (G2656CA) using the microarray assay scanning protocol (one-color or two-color) and raw microarray image files were generated.

Agilent Feature Extraction Image Analysis Software (Version 10.7.3) was used to extract raw microarray data image files. Data visualization and analysis was performed using Agilent GeneSpring GX (Version 12.0) Software.

Microarray data analysis

The microarray data were analyzed for gene expression using the Gene Expression workflow in GeneSpring GX using the Advanced Analysis Workflow setting. Default flag settings were used to make the detection calls. Signal intensities for each probe were normalized to the 75th percentile without baseline transformation for one-color assays. The two-color microarray assays, including the self vs. self experiments (for monitoring dye bias), were uploaded as Agilent Expression two-color experiment type in GeneSpring and processed without baseline transformation. Dye swap arrays were identified for the control vs. treated experiments. The software was used to visualize QC metrics, and the signal values were transferred to SpotFire for correlation analysis.

Differential expression analysis was performed using the 'Filter on Volcano Plot' option under the Analysis tab in GeneSpring. Analysis was carried out an entity list consisting of "detected" probes only. The analysis was carried out using this filtered entity list, using a Ttest unpaired statistical method with Benjamini Hochberg FDR method; p-values were computed asymptotically. Output from GeneSpring includes a table consisting of Probe Names, p-values, corrected p-values and Fold change between the selected conditions along with the plot upon completion of the statistical tests.

Results

Assay sensitivity

Agilent one-color and two-color spike-in controls are used to monitor the microarray workflow from sample amplification and labeling to microarray processing. Data in Figure 1 A and B represent the log of the green processed signal for each spike-in transcript against the log of the relative concentration of the Agilent One-Color RNA Spike-In Mix for the *E. coli* and *M. tuberculosis* microarrays, respectively. The concentration response plot demonstrates that the linear range of the assay spans at least 4 orders of magnitude and that the assay has a high sensitivity to low-abundance transcripts. The Agilent Two-Color Spike-In Mix contains two spike-in mixtures (Spike A and Spike B) that each contain 10 spike-in transcripts at concentrations that generate predetermined ratios in the final two-color assay. Spike A mix was labeled with Cyanine 3 and Spike B mix was labeled with Cyanine 5. The graphs in Figure 1 C and D represent the observed vs. expected log ratio for each spike-in transcript for *E. coli* and *M. tuberculosis* two-color assays, respectively. The correlation coefficient of greater than 0.99 shows the high correlation of observed vs. expected ratios for each transcript across the ten-transcript mixture. The slopes of the line fits (both > 0.89) demonstrate very low compression for both assays.

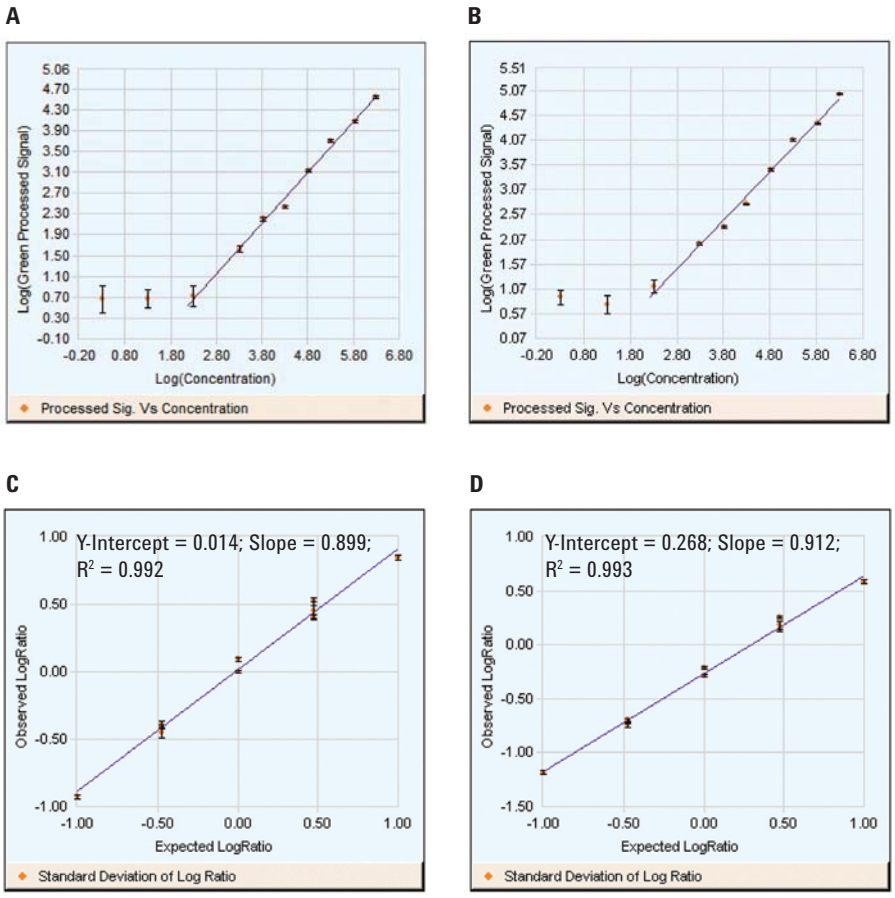

Figure 1. *Processed signal versus relative concentration plots for* E. coli *one-color (A) and* M. tuberculosis *one-color (B) demonstrate the ability of the LIQA WT labeling kit to generate labeled cRNA from very low abundance transcripts to high expressers (4 logs of dynamic range). The expected versus observed log ratios for* E.coli *two-color (C) and* M. tuberculosis *two-color (D) show high correlation of ratios for each transcript across the ten-transcript mixture as demonstrated by correlation coefficient of 0.99, and show the accuracy of the assay as demonstrated by the slope values >0.89.*

Assay reproducibility

We assessed the reproducibility of gene expression predictions for technical replicates of *E. coli* and *M. tuberculosis* arrays. For both one- and two-color microarrays, two different conditions were compared: paraquat-treated vs. control for *E. coli* and two different strains of *M. tuberculosis*. In the two color studies for *E. coli,* the control samples were labeled with Cy3 and the paraquat-treated samples were labeled with Cy5.

All of the microarray assays (one-color and two-color for both *E. coli* and *M. tuberculosis*) were found to be highly reliable, with reproducibility demonstrated by a correlation of R^2 greater than 0.9 for signals across technical replicates. In addition to demonstrating the reproducibility of the assay, the correlation plots also demonstrate the wide dynamic range of detection of the assay, with transcript signals spanning greater than 4 orders of magnitude. The two-color arrays were also evaluated for dye bias by labeling the same RNA with Cy3 and Cy5 and co-hybridizing (self-self) and by performing dye-swap experiments with the control and paraquat-treated *E. coli* RNA. In both cases fewer than 1% of probes exhibited dye bias (data not shown), ensuring high confidence in the differential expression calls.

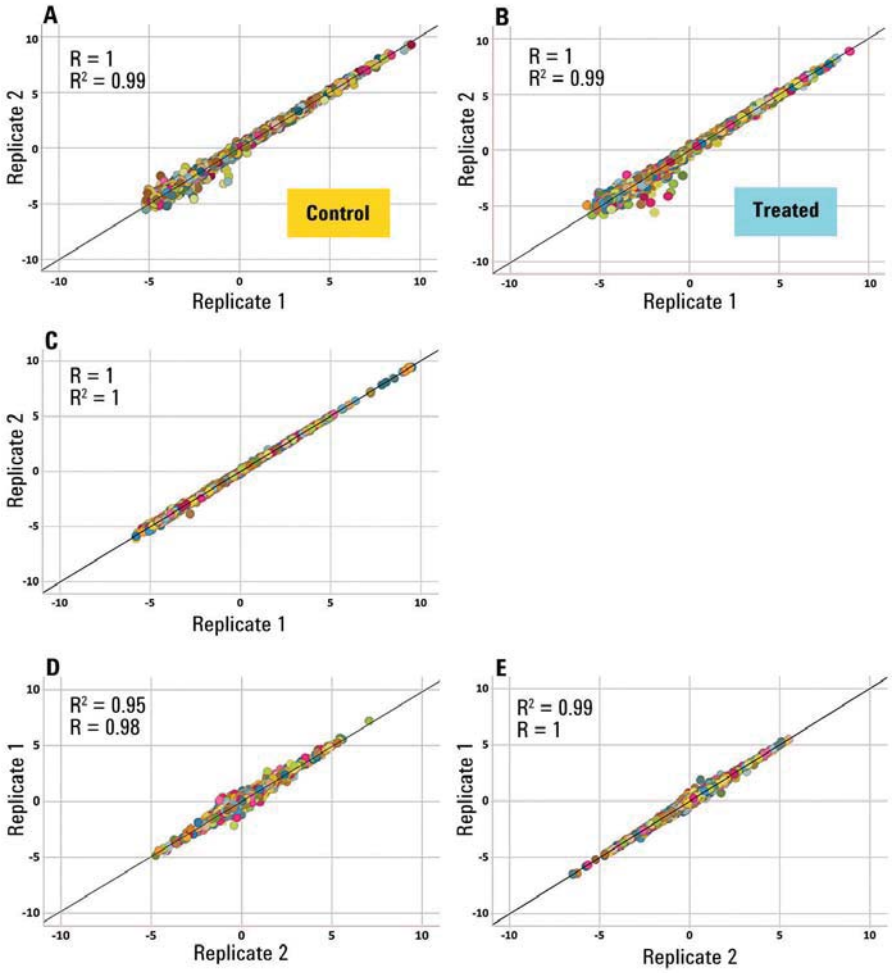

Figure 2. *Correlation of microarray results based on biological probes using the 75th percentile-normalized green processed signals (gProcessedSignals) show the reproducibility of signal intensities between two technical replicates in the one-color assay,* E. coli *Control (A),* E. coli *Treated (B) and* M. tuberculosis *(C); and in the two-color assay,* E. coli *(D) and* M. tuberculosis *(E).*

Differential regulation of *E. coli* genes upon paraquat treatment

Paraquat (1, 1′-dimethyl-4, 4′-bipyridinium) causes oxidative stress in *E. coli* by increasing the intracellular levels of O_2–. The high levels of O_2– in turn lead to enhanced expression and activity of key genes involved in antioxidant activities. Differential gene expression analysis of *E. coli* treated with paraquat using the one-color microarray assay revealed approximately 1,200 of 6,467 entities to be significantly differentially expressed at a twofold cut-off (corrected p-value of < 0.05). GeneSpring allows users to filter genes based on both the significance and magnitude of differential expression. This comparison is visualized using volcano plots, which plot the $\log_2$ fold change values and $-\log_{10}$ p-values for all the measured entities. Figure 3 shows the volcano plot of the gene analyzed in the *E.coli* experiment and highlights those that are significantly affected by paraquat treatment. Similar results were obtained with the two-color assay (data not shown).

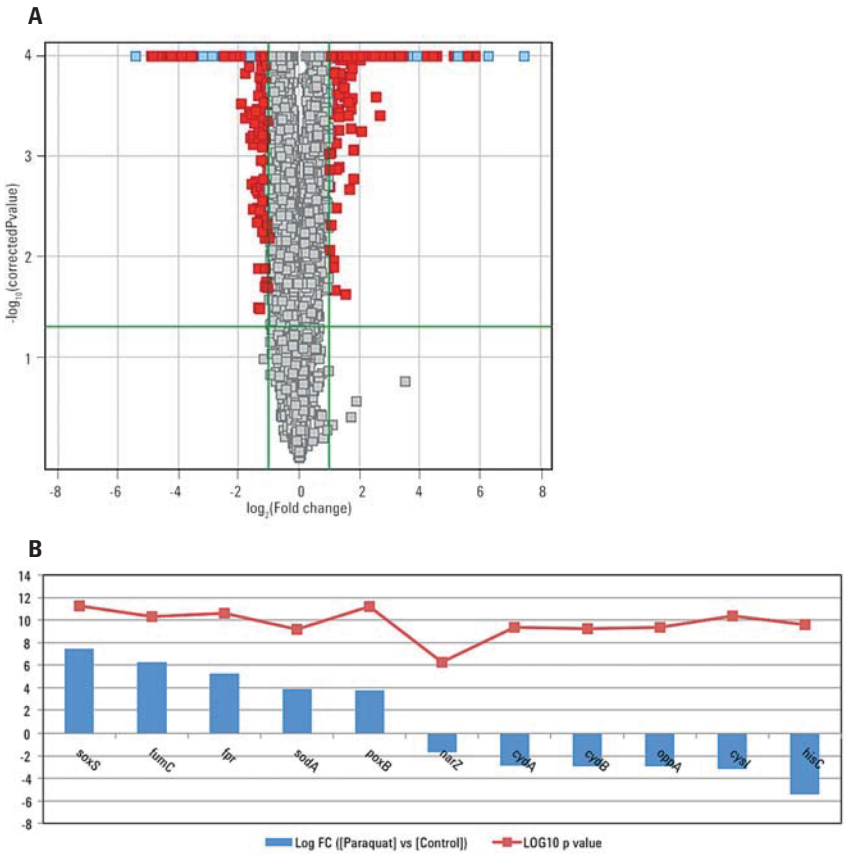

Figure 3. *Volcano plot showing differential gene expression of* E. coli *upon paraquat treatment (A). Entities highlighted in red meet the criteria for differential expression (FC>2; p <0.05). Selected biologically relevant genes are highlighted in blue, and differential expression results are represented on the graph (B).*

Differential gene expression observed in *M. tuberculosis*

Differential expression analysis between the two different strains of *M. tuberculosis* using one-color microarray analysis revealed 1920 of 5120 total biological entities satisfying a corrected p-value cut-off of 0.05 and fold change cut-off of 2.0 (Figure 4). Similar results were obtained with the two-color assay (data not shown). These results demonstrate that Agilent custom arrays can be used to generate statistically relevant data with high confidence in a high GC content organism.

Conclusion

The LIQA WT kit efficiently generates high amounts of labeled cRNA from prokaryotic samples with high specific activities for one-color or two-color gene expression analysis. The kit can also be successfully employed to profile RNA from a GC-rich genome. The labeling method reliably and reproducibly enables detection of biological transcripts with a dynamic range of at least four orders of magnitude. GeneSpring software facilitates data analysis to identify the key genes that are differentially regulated. The Agilent LIQA WT labeling kit and catalog and custom prokaryotic arrays coupled with GeneSpring provide for analysis of prokaryotic gene expression with high confidence.

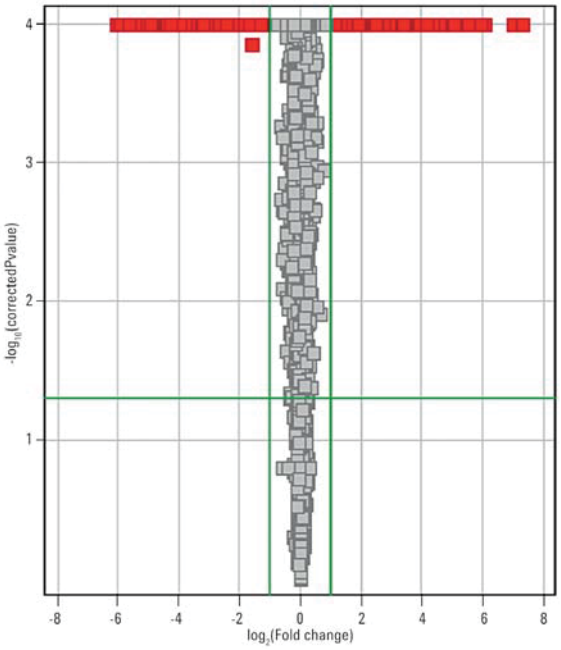

Figure 4. *Volcano plot showing differential gene expression pattern among two different strains of* M. tuberculosis*. Entities highlighted in red meet criteria for differential expression (FC>2; p<0.05).*

Learn more:
www.agilent.com/genomics/
prokaryoteprotocol

For Research Use Only. Not for use in diagnostic procedures.

PR7000-0178

Published in USA, April 12, 2016
Publication Number 5991-0879EN

Agilent Technologies