Reliable Expression Profiling Using Low RNA Inputs

Authors

Anne Bergstrom Lucas,
Gary Lin,
Sharoni Jacobs, and
Stephanie Fulmer-Smentek
Agilent Technologies, Inc.,
Santa Clara, CA 95051

Introduction

Gene expression profiling by microarray analysis provides an important avenue for understanding biological mechanisms, classifying tissue and tumor types, and identifying signs for diagnosis and prognosis. Advances in isolating small populations of cells using methods such as laser capture microdissection have resulted in the need to conduct gene expression profiling experiments from increasingly smaller amounts of total RNA. To address this need, we have developed a modified linear amplification procedure that generates sufficient quantities of Cy-labeled cRNA suitable for oligonucleotide microarray experiments. The assay requires only 10 nanograms total RNA input and takes less than 6.5 hours.

Agilent's new Low Input Quick Amp (LIQA) Labeling Kit (5190-2305, 5190-2306) is based on a single-tube procedure employing the AffinityScript Reverse Transcriptase, a mutant MMLV-RT that binds primer-template complexes with 10-fold higher efficiency as compared to wild type MMLV-RT (Arezi and Hogrefe, 2008). The use of the AffinityScript Reverse Transcriptase results in increased cDNA yields and improved sensitivity of the assay. The protocol uses a single round of *in vitro* transcription amplification without purification of the cDNA product to generate labeled cRNA in less than one day and enables gene expression profile comparisons in less than two days.

With the introduction of LIQA, a range of improvements were made to both lower the input requirements and to enhance usability over earlier labeling kits while maintaining the high performance of Agilent gene expression assays. The new LIQA Kits include blends of individual enzymes allowing for more accurate pipetting of low volume reagents. Ease of use was improved by reformulation of the transcription reaction buffer to include a PEG component and nuclease-free water. A comparison of the LIQA Kit to the legacy Quick Amp Labeling Kit, as well as to other gene expression measurement technologies is shown in **Table 1**.

Agilent Technologies

	Agilent LIQA	Agilent Quick Amp	Competitor Microarray Assays	RNA-Seq
Time to Data	1.5 days	1.5 days	1.5-2.5 days	4-5 days
Hands-on Time	~ 6 hours	~ 6 hours	~ 6-6.5 hours	~ 4 hours
Price per Sample	$	$$	$-$$$	$$$-$$$$
Lower RNA Input Requirement	10 ng	200 ng (1-color) 50 ng (2-color)	50 ng	1 µg
Downstream Analysis	Simple	Simple	Simple	Challenging
Dynamic Range	~ 5 logs	~ 5 logs	~ 2-3.5 logs	~ 5-6 logs

Table 1: Comparison of features amongst approaches to RNA preparation and gene expression analysis, including the new Agilent Low Input Quick Amp Labeling Kit.

Materials and Methods

RNA preparation and addition of spike-in controls

The RNA samples used in the Microarray Quality Control study (MAQC Consortium, 2006) were used in these experiments; the Stratagene Universal Human Reference RNA (SUHRR or MAQC_A, 740000) and the Ambion Brain Reference RNA (ABRR or MAQC_B, AM6050). Total RNAs were quantified using a NanoDrop ND1000 spectrophotometer (Nanodrop Technologies, Thermo Scientific). The quality of the RNA was assessed with the Agilent 2100 Bioanalyzer (G2943CA) using the RNA 6000 Nano Kit (5067-1511).

RNA labeling

All sample preparations and hybridizations were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis, Low Input Quick Amp Labeling Protocol User Manual (G4140-90040, version 6.0, December 2009) and the Agilent Two-Color Microarray-Based Gene Expression Analysis, Low Input Quick Amp Labeling Protocol User Manual (G4140 90050, version 6.0, December 2009). Master mixes sufficient for four reactions (5x with overage) or for eight reactions (10x with overage) were prepared.

Microarray hybridization, scanning, and data analysis

Labeled RNA was hybridized to SurePrint G3 Custom GE 8x60K microarrays (G4102A) and to the Whole Human Genome Microarray, 4x44K (G4112F). The same biological content was present on both array formats. Fragmentation mixes were prepared using 600 ng for the 8-pack format and 1.65 µg cRNA for the 4-pack format. Hybridizations were carried out as described in the user manuals and the washed arrays were scanned using the Agilent High Resolution C Scanner (G2565CA) using scanning protocols specific for gene expression microarrays and the format (either 8x60K or 4x44K).

Agilent Feature Extraction Image Analysis Software (Version 10.5.1) was used for data extraction from raw microarray image files. Data visualization and analysis was performed with GeneSpring GX (Version 10.0) software and detected calls were made by employing the default flag settings for Agilent microarrays. Additional data analysis was performed in MS Access and MS Excel and visualized with Spotfire software.

Results/Discussion

Yields using the Low Input Quick Amp (LIQA) Kit

We performed titrations of the two MAQC samples to assess a range of total RNA input levels for the new LIQA Kit. Total RNA input amounts ranging from 1 ng to 500 ng were compared and cyanine-labeled cRNA yields and specific activities were determined by NanoDrop spectrophotometry. The reactions were performed in triplicate, resulting in the average cRNA yields shown in **Figure 1**.

High correlation and dynamic range

Replicate correlations at different total RNA input levels show high correlation and ~ 5 logs dynamic range. Scatter plot signal correlations for technical replicate assays with either 10 ng, 25 ng, or 200 ng of the same MAQC_A (SUHRR) total RNA sample were generated in GeneSpring GX using biological probes that were detected on all three microarrays in the one-color titration experiment described above. Orthogonal fit correlations were greater than 0.99 for the comparisons of the 75th percentile-normalized $\log_2$ signals for the technical replicate arrays. A similar two-color microarray experiment was performed with the MAQC_A sample labeled with Cy3 and the MAQC_B sample labeled with Cy5. Scatter plot $\log_2$ B/A ratio correlations for technical replicate assays with either 10 ng, 25 ng, or 200 ng of the MAQC total RNA samples were generated in GeneSpring GX program using biological probes that were detected on at least two of the three microarrays in the two-color titration experiment. Orthogonal fit correlations were greater than 0.97 for the $\log_2$ B/A ratios. These high correlations spanned a wide dynamic range up to five logs of magnitude, as shown in **Figure 2**. Such reproducible and sensitive data is important for the ability to study low-expressing genes.

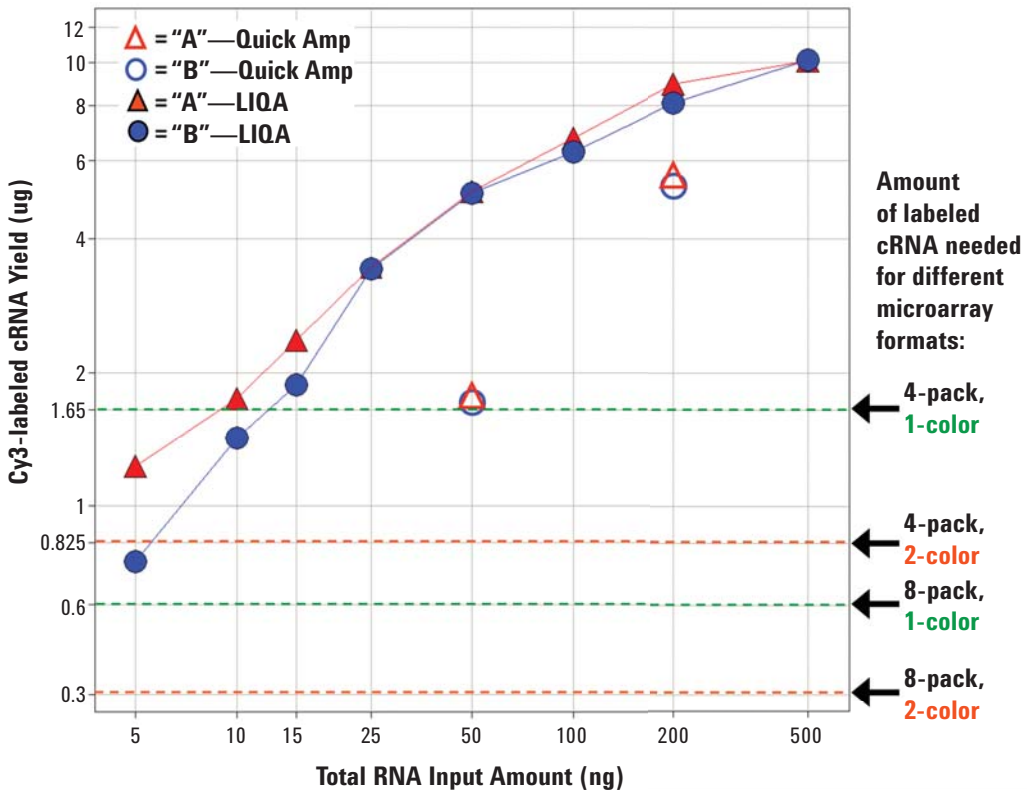

Figure 1: cRNA yields increase as amount of total RNA increases in the Low Input Quick Amp assay as denoted by the filled symbols. The red filled triangles are the MAQC_A sample and the blue filled circles are the MAQC_B sample labeled with LIQA. The corresponding open symbols indicate cRNA yields for 50 ng or 200 ng RNA input into the legacy Quick Amp Kit. Black arrows indicate amount of cyanine-labeled cRNA needed for an 8-pack or a 4-pack one-color or two-color microarray experiment. The cRNA yields were similar for two-color LIQA input titrations experiments using Cy5-CTP (data not shown).

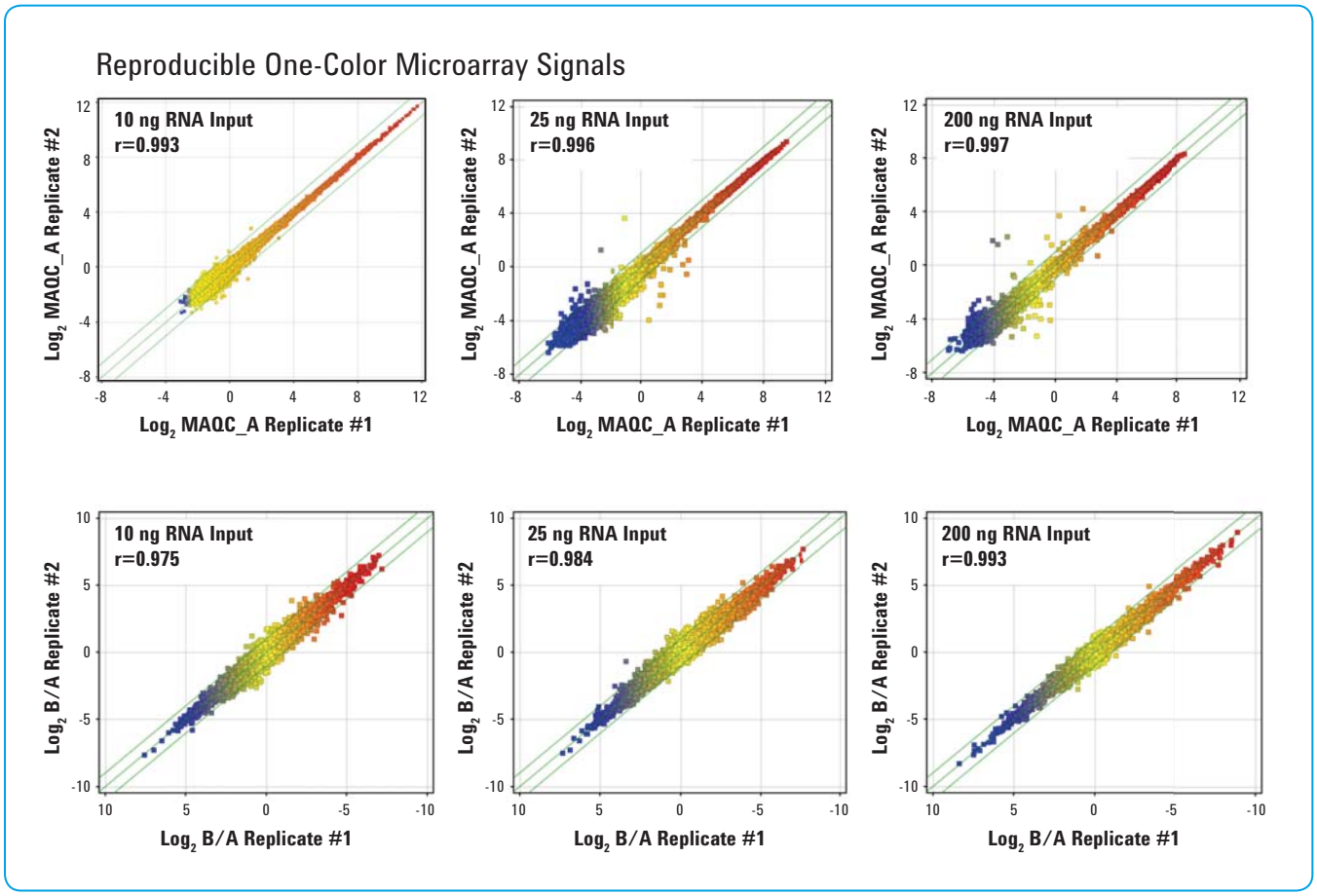

Figure 2: Technical replicate samples show high correlations in both one-color and two-color microarray assays. The one-color 75th percentile-normalized signal correlations are high (>0.99) across the almost five orders of magnitude of microarray signal for all total RNA input ranges. The two-color $\log_2$ ratios of B/A are also highly correlative (>0.97) across the total RNA input ranges tested (10 ng, 25 ng, and 200 ng shown here).

Reproducible and repeatable data

A comparison of the repeatability and the reproducibility can be measured by the variation of the microarray signals across technical replicate experiments. For this analysis we compared the 75th percentile-normalized signals using genes that were detected (well above background and uniform signal) on all three technical replicate SurePrint G3 microarrays for coefficient of variation (CV) calculations. The results are presented in a box-whisker plot in **Figure 3**. As expected, the CVs were lower when signals of the genes detected were higher, resulting in lower inter-quartile boxes and shorter whiskers for the higher total RNA input levels. Nevertheless, reproducible and repeatable data could be obtained with as low as 10 ng of input RNA. Overall the median %CVs for all of the samples shown above are <10% CV, demonstrating that the LIQA assay labels RNA with highly reliable and reproducible efficiency.

LIQA and Quick Amp Kit comparison

We compared results using the LIQA assay or the Quick Amp assay using representative "low" and "medium" RNA input. At 50 ng RNA input, LIQA produces at least three-fold more labeled cRNA as compared to Quick Amp. As shown earlier in **Figure 1**, the yield from 50 ng input into Quick Amp was equivalent to that from the 10 ng input into LIQA and the yields from the 200 ng input into Quick Amp were equivalent to 50 ng input into LIQA.

To compare the repeatability and the reproducibility of Quick Amp and LIQA, we performed CV calculations for those genes detected on all three replicate arrays in this experiment. The results were plotted in a box-whisker plot as shown in **Figure 4.** The performance of the LIQA Kit at the "low" 10 ng input level is comparable to the "low" 50 ng input level in the Quick Amp Kit, and performance at the "medium" 50 ng input level for the LIQA Kit is comparable to the "medium" 200 ng input level into the Quick Amp Kit.

Equivalent detection limits

Both the Quick Amp and LIQA Kits have similar limits of detection. When the Agilent One-Color Spike-In controls are diluted as recommended, the second-most dilute spike-in is present at a concentration of approximately 1 in 6.8 million copies per cell. The Agilent spike-in concentration-response plots from the Agilent QC reports were compared for representative G3 arrays from the different input levels. In **Figure 5**, we show that both kits have the ability to detect transcripts in concentrations as low as 1:6.8 million copies per cell.

Comparison of Quick Amp and LIQA data to orthogonal TaqMan data

One of the outcomes of the MAQC study was TaqMan gene expression measurements for a subset of ~1000 genes. There was sufficient assay signal with both the MAQC_A sample and the MAQC_B sample on the TaqMan platform to make ratios of MAQC_B sample to MAQC_A sample (B/A) for 774 of the 1000 genes assayed. With continued improvements to the sensitivity of the Agilent platform the number of genes detected by microarray has steadily risen over the years to 700 genes, as demonstrated in the experiment comparing the Quick Amp and LIQA Kits. The scatter plots comparing the $\log_2$ ratios of the B/A samples from the TaqMan assay to either the

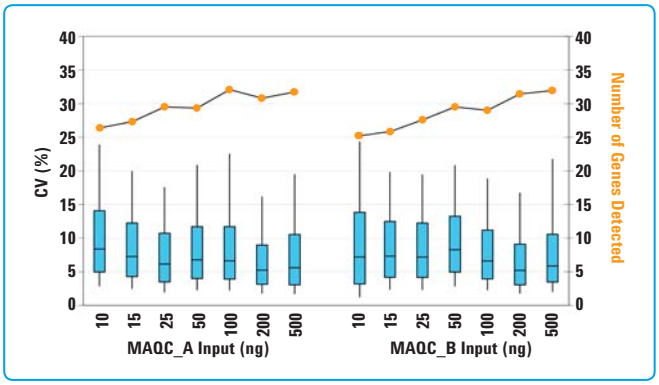

Figure 3: Median inter-array %CVs are well below 10% CV for total RNA input levels ranging from 10 ng to 500 ng. The median %CV is represented by the line in the middle of each box. Edges of the box are the inter-quartile range (25th percentile to 75th percentile) of %CVs. The ends of the whiskers on the boxes denote the 90th percentile and the 10th percentile of %CV. The %CV calculation was based on the number of genes detected across all three replicate G3 arrays and this value is represented on the secondary y-axis as the number of genes detected.

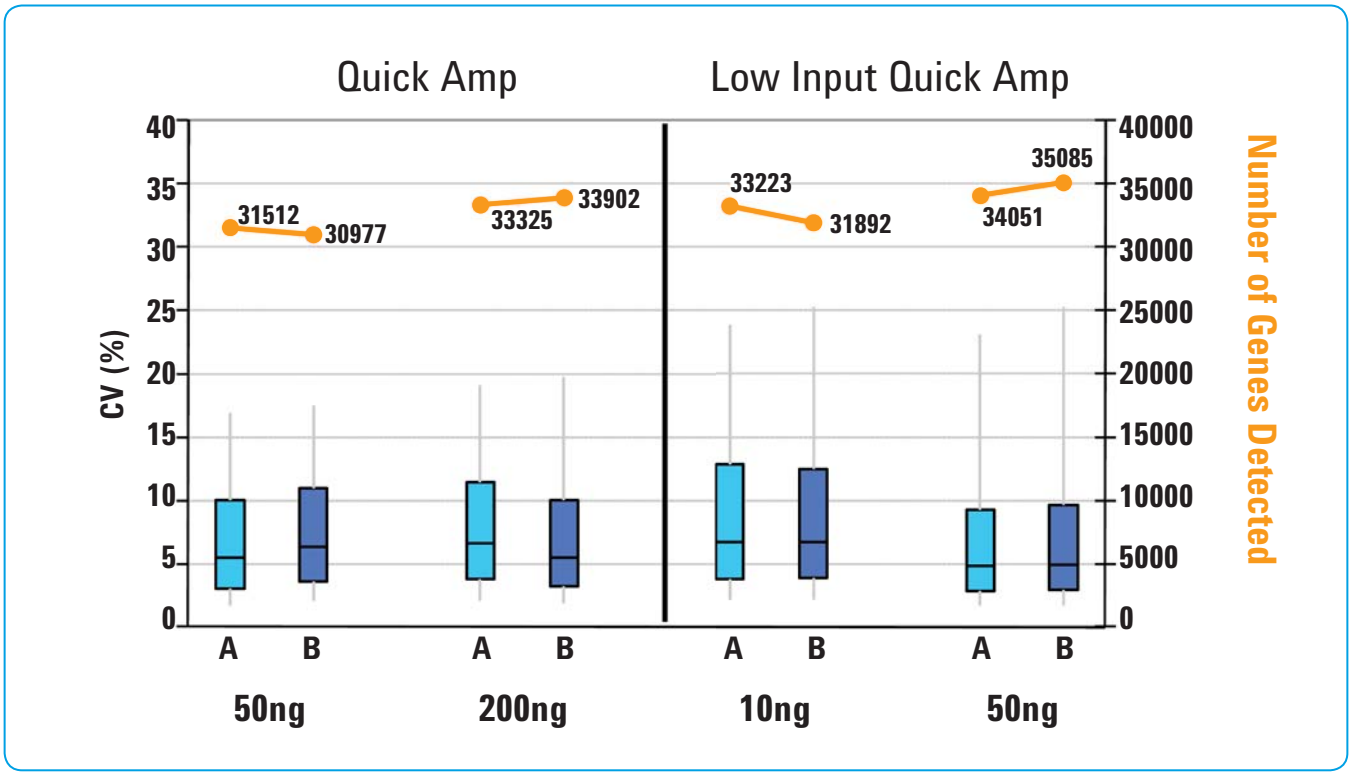

Figure 4: The LIQA Kit detects a higher number of genes with reduced median %CVs, indicating that the new LIQA assay performs equivalently (or better) as compared to the Quick Amp Kit. The median %CV is represented by the line in the middle of each box. Edges of the box are the inter-quartile range (25th percentile to 75th percentile) of %CVs. The ends of the whiskers on the boxes denote the 90th percentile and the 10th percentile of %CV. The %CV calculation was based on the number of genes detected across all three replicate arrays and this value is represented on the secondary y-axis as the number of genes detected.

Quick Amp assay or the LIQA assays demonstrate high correlations (>0.92) between Agilent microarray measurements and the MAQC TaqMan data as shown in **Figure 6**.

The plots in **Figure 6** indicate that the highest $\log_2$ B/A ratios that were generated by the TaqMan platform in the MAQC study were approximately 8.0 as compared to $\log_2$ ratios greater than 12.0 for the Agilent microarray platform. This difference in sensitivity results in the unusual flattening of the shape of the scatter plot comparison of log ratios between the two gene expression measurement platforms and slightly skews the slope values. Even with this skew, the slopes for the microarray measurements as compared to the qPCR measurements are greater than 0.86, indicating that the Agilent microarrays generate gene expression measurements with very little signal compression. This means accurate representation of gene expression levels, even for low- and high-abundance genes.

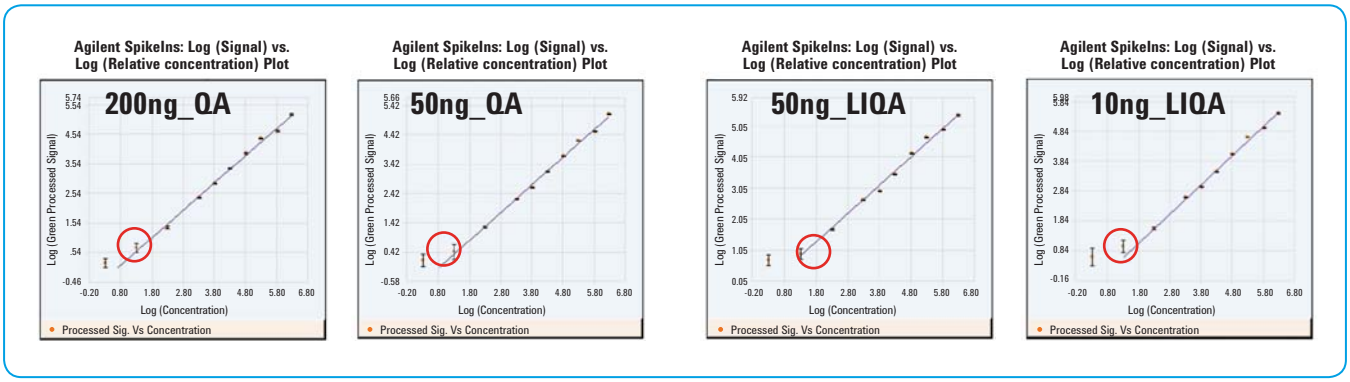

Figure 5: Concentration-response plots demonstrate that the LIQA Kit generates labeled cRNA for a microarray assay that has the same ability to detect rare transcripts as compared to the Quick Amp Kit. In all four of these examples, the signals from the Agilent spike-in controls indicate that the microarray assays span five orders of magnitude dynamic range.

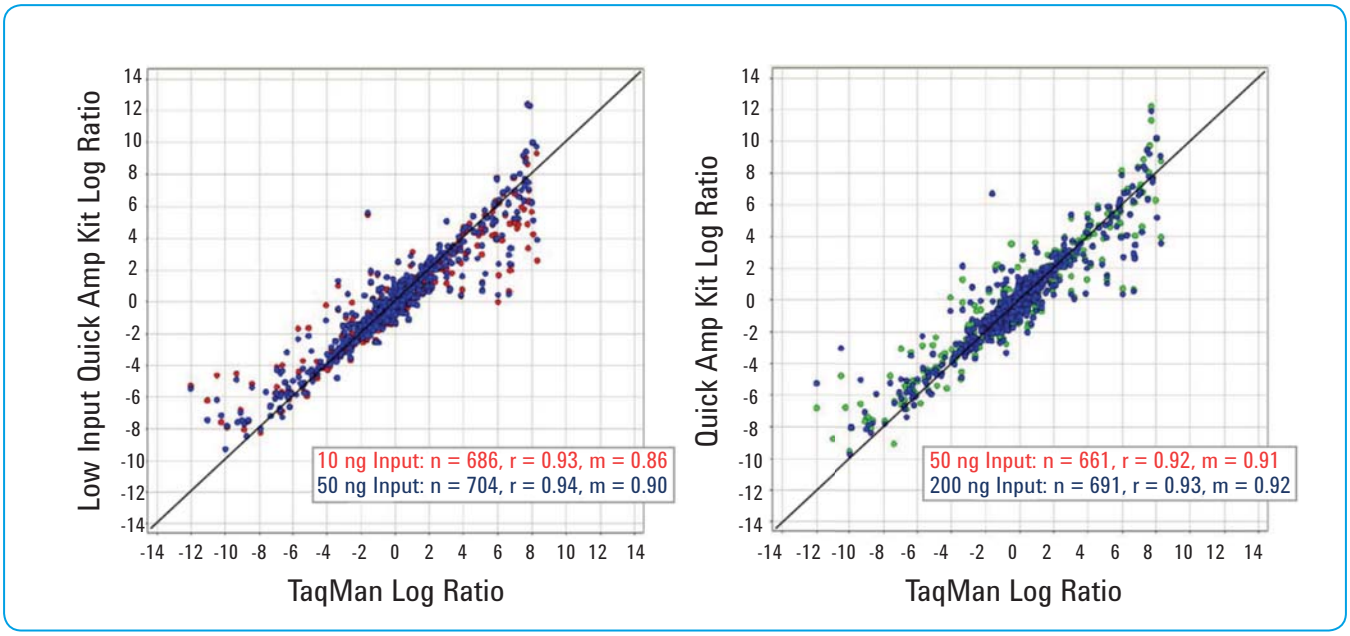

Figure 6: Both the Low Input Quick Amp Kit and the Quick Amp Kit demonstrate very high concordance to the MAQC TaqMan data. The common input level of 50 ng used in both kits is colored blue in the graphs above for easy comparison between the two graphs. In this comparison the LIQA assay was able to generate 704 measurements as compared to 661 measurements for the Quick Amp Kit with the same 50 ng total RNA input amount.

Conclusion

The Low Input Quick Amp (LIQA) Labeling Kit enables gene expression experiments using only a few nanograms of total RNA. Efficient labeling for high quality results take less than one day. Using LIQA and Agilent gene expression microarrays, genes present in amounts equivalent to 1:6.8 million transcripts can be detected reliably, providing quality data for low abundant genes. 25 ng total RNA input generated sufficient cRNA for a 4-pack array and 10 ng total RNA input generated sufficient cRNA for an 8-pack. This kit is highly suitable for gene expression analysis of precious samples where RNA amounts are scarce.

When compared to the legacy Quick Amp Labeling Kit, the LIQA Labeling Kit generates data with similar or improved data quality, but with lower RNA input amounts. The microarray studies using LIQA were able to detect more genes with lower amounts of starting total RNA as compared to those using Quick Amp.

With the introduction of LIQA for RNA labeling, researchers have a tool to gain trustworthy gene expression data quickly, using only small amounts of RNA samples.

References

Arezi, B. and Hogrefe, H. (2008) Novel mutations in Moloney Murine Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Research*, **37**, 473–481.

MAQC Consortium (2006) The MicroArray Quality Control (MAQC) project shows inter- and intra-platform reproducibility of gene expression measurements. *Nature Biotechnology*, **24**, 1151–1161.

For More Information

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

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 (Option 2, Option 3)
sig_sales@agilent.com

Asia Pacific
inquiry_lsca@agilent.com

Europe
info_agilent@agilent.com

Published in USA, April 20, 2010
5990-5550EN

Agilent Technologies