Microarray-Based Detection of miRNA: Applying Novel Statistical Methods for Evaluation of Platform Performance and Assessment of Data Quality

Xiaojun Di and Zinaida A. Sergueeva
Covance Genomic Laboratory, Seattle, Washington

Presented at the
CHI Biomarker World Congress
Philadelphia, Pennsylvania
4–6 May 2010
Microarray-Based Detection of miRNA: Applying Novel Statistical Methods for Evaluation of Platform Performance and Assessment of Data Quality

Xiaojun Di and Zinaida A. Sergeeva, Covance Genomic Laboratory, Seattle, Washington

Introduction

Microarray-based detection of short 22–25 nucleotides in length RNA polycules that function to modulate the activity of specific mRNA targets and play important roles in a wide range of biological processes (1, 2). MicroRFAs can be an ideal choice for biomarker discovery in cancer tissues because: (i) miRNA expression is frequently dysregulated in cancer (3); (ii) expression patterns of miRNAs in human cancer appear to be tissue-specific; (iii) and (iv) microRNAs are more stable than mRNAs. In general, miRNA expression patterns are known to effectively characterize a tumor’s origin of development and represent potential biomarkers for both the diagnosis and prognosis of human cancer.

The most efficient and cost-effective way to simultaneously interrogate the expression of multiple miRNAs is via microarray technology. Several commercial platforms are now available for research use. After an initial study of microarray-based discovery of microRNAs, Affymetrix and Agilent and an array-like qPCR assay from ABI (data not shown) we have selected the single-color miRNA platform developed by Agilent Technologies (Santa Clara, CA, USA). This technology provides both sequence and size discrimination for mature miRNAs and generates results that are highly correlated with real-time quantitative PCR results (4). Recently published data on the concordance among digital gene expression, microarrays and qPCR for microRNA profiling reported similar findings (5).

Materials and Methods

RNA Samples: Human brain and kidney total RNA (Ambion, Inc.) sample mixtures were made at the following mass ratios: 100% brain, 75% brain / 25% kidney, 50% brain / 50% kidney and 25% brain / 75% kidney according to vendor’s protocol. Data were extracted and summarized using Agilent Feature Extraction Software. Data was background subtracted, normalized by the average 75th percentile of sample-wise signals and then log-transformed (a global constant was added to the log-expression to accommodate for a small number of negative signals).

Experimental Design and Analysis

Any assay with potential to detect single or multiple biomarkers for clinical or diagnostic purposes should be validated. The assay’s day-to-day, operator-to-operator reproducibility, accuracy, precision and functional sensitivity need to be measured and criteria of acceptance to be well defined. We have employed these triplex-replicates of each brain kidney sample mixtures that were run in two different days.

For assay reproducibility we calculated the cumulative percentage of miRNAs with CV under a given threshold (10%) was used to measure precision.

Accuracy assessment was done indirectly due to lack of standard reference. We measured accuracy by comparing observed brain to kidney signal ratios to the expected nominal ratio for the miRNAs with significant differential expression. For miRNAs with brain-up-regulated expression, the 30:70 brain to kidney ratio (brain-upregulated and kidney-upregulated) was used. First, for any RNA mixture sample other than 100% brain, we calculated signal ratios of brain-upregulated miRNAs. Then these ratios were regressed against the ratios between 100% kidney and 100% brain using a linear model, and estimated slope values were used to calculate the slope CV.

For any RNA mixture sample other than 100% brain, we calculated signal ratios of brain-upregulated and kidney-upregulated. First, for any RNA mixture sample other than 100% brain, we calculated signal ratios of brain-upregulated miRNAs. Then these ratios were regressed against the ratios between 100% kidney and 100% brain using a linear model, and estimated slope values were used to calculate the slope CV.

From the literature on tissue-specific miRNAs (6) we have selected seven miRNAs that show high expression level in brain: hsa-miR-124, hsa-miR-129*, hsa-miR-129-3p, hsa-miR-129-5p, hsa-miR-21-3p, hsa-miR-21-5p and hsa-miR-223-3p. These miRNAs have homoeogenous sequences but no detectable level of expression in brain.

As for miRNAs with no detected expression in brain tissue (brain-non-specific) we have chosen the following ones: hsa-miR-199a-3p, hsa-miR-199a-5p, hsa-miR-199b-5p, hsa-miR-499-5p, hsa-miR-214, hsa-miR-114*, hsa-miR-115a, hsa-miR-116a, hsa-miR-118a and kidney- miR-K12-10b (viral miRNA).

Summary

Overall the Agilent miRNA assay provides good reproducibility, precision, accuracy and functional sensitivity to be used for miRNA-based biomarkers screening.

Results and Discussion

Quantitative reproducibility: Pearson correlation coefficients (R) were calculated for sample replicates within/between days and within/between operators. We observed R 0.960 across the board which indicates high level of quantitatively reproducibility (see Figure 1 as an example).

Qualitative reproducibility: We observed that greater than 90% of the detected miRNAs are concords among replicates regardless of the sample type which is a strong evidence of data reproducibility (see Figure 2 as an example).

Precision: The cumulative percentage of miRNAs with less than 10% CV in signal measurement was observed to be greater than 80% for all RNA mixtures except 100% kidney (see Figure 3). This test is indicative of high relative sensitivity of the Agilent miRNA assay performance.

Accuracy: While the brain-up-regulated miRNAs the estimated slopes were very close to expected nominal values with minimum compression and high accuracy (Figure 4b), the slopes for brain-down-regulated miRNAs were notably off and compressed (Figure 4a). This difference in behavior is in miRNA expression profiling we can not explain.

Functional sensitivity: The analysis was done on the selected miRNAs with known tissue characteristics using similar approach to accuracy estimation. As shown in Figure 5 and 6, brain-specific miRNAs generally showed a good correlation between the expected and observed signals in ratio mixtures with respectable relative accuracy. The three miRNAs we have used for brain-specific discrimination showed very little if any correlation. Most of the non-brain-specific miRNA also demonstrated a linear response (Figure 6) with the exception of the non-brain-specific hsa-miR-129 (Figure 5d) that does not have any tissue-specific expression in non-cancerous tissues (7).

References