

Research article

Open Access

## Measuring microRNAs: Comparisons of microarray and quantitative PCR measurements, and of different total RNA prep methods

Robert A Ach\*, Hui Wang and Bo Curry

Address: Agilent Laboratories, Agilent Technologies, 5301 Stevens Creek Blvd., Santa Clara, CA 95051, USA

Email: Robert A Ach\* - robert\_ach@agilent.com; Hui Wang - hui\_wang@agilent.com; Bo Curry - bo\_curry@agilent.com

\* Corresponding author

Published: 11 September 2008

Received: 28 March 2008

BMC Biotechnology 2008, 8:69 doi:10.1186/1472-6750-8-69

Accepted: 11 September 2008

This article is available from: <http://www.biomedcentral.com/1472-6750/8/69>

© 2008 Ach et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Determining the expression levels of microRNAs (miRNAs) is of great interest to researchers in many areas of biology, given the significant roles these molecules play in cellular regulation. Two common methods for measuring miRNAs in a total RNA sample are microarrays and quantitative RT-PCR (qPCR). To understand the results of studies that use these two different techniques to measure miRNAs, it is important to understand how well the results of these two analysis methods correlate. Since both methods use total RNA as a starting material, it is also critical to understand how measurement of miRNAs might be affected by the particular method of total RNA preparation used.

**Results:** We measured the expression of 470 human miRNAs in nine human tissues using Agilent microarrays, and compared these results to qPCR profiles of 61 miRNAs in the same tissues. Most expressed miRNAs (53/60) correlated well ( $R > 0.9$ ) between the two methods. Using spiked-in synthetic miRNAs, we further examined the two miRNAs with the lowest correlations, and found the differences cannot be attributed to differential sensitivity of the two methods. We also tested three widely-used total RNA sample prep methods using miRNA microarrays. We found that while almost all miRNA levels correspond between the three methods, there were a few miRNAs whose levels consistently differed between the different prep techniques when measured by microarray analysis. These differences were corroborated by qPCR measurements.

**Conclusion:** The correlations between Agilent miRNA microarray results and qPCR results are generally excellent, as are the correlations between different total RNA prep methods. However, there are a few miRNAs whose levels do not correlate between the microarray and qPCR measurements, or between different sample prep methods. Researchers should therefore take care when comparing results obtained using different analysis or sample preparation methods.

### Background

MicroRNAs (miRNAs) are small (~18–24 nucleotides) non-coding RNAs which bind to mRNAs to regulate protein expression, either by blocking translation and/or by

promoting degradation of the mRNA target (reviewed in [1-3]), or alternatively by increasing translation [4,5]. They have been found to be involved in numerous functions such as cell fate determination, cell proliferation,

cell differentiation, and cell death (reviewed in [6,7]). Profiles of miRNAs in various types of tumors have been shown to contain potential diagnostic and prognostic information (reviewed in [8,9]). The number of known miRNAs has rapidly increased in recent years, and currently there are 722 human miRNA sequences reported in the Sanger Institute's miRNA database release 10.0 (miR-Base) [10-12], with potentially many more yet to be reported [13,14].

Several methods for global miRNA profiling are currently in common use. These include quantitative RT-PCR (qPCR) involving stem-loop RT primers combined with TaqMan PCR (Applied Biosystems) analysis [15,16], qPCR with locked nucleic acid primers (Exiqon) [17], qPCR using poly(A) tailing (QIAGEN, Stratagene) [18,19], high-throughput sequencing of small RNA libraries [20], and microarray analysis (for examples, see [21-26]). Typical experimental workflows often involve using different methods of measuring miRNAs at different research stages. For this reason, it is important to know how well the different measurements agree with each other. Several groups have compared microarray profiling results with those obtained by quantitative PCR for either a small number of genes or a small set of samples [21,22,26-30]; however there has been no systematic comparison of larger numbers of miRNAs across a widely diverse range of human tissues using the two methods.

In this study, we compared the relative expression of 61 different miRNAs across nine different human tissues, measured using both Agilent miRNA microarrays and TaqMan qPCR. The Agilent microarray platform features the direct end-labeling and profiling of mature miRNAs from total RNA without any size fractionation or amplification to minimize experimental loss, bias, or variations [31,32]. The labeling reaction is performed under denaturing conditions to provide high labeling yield, minimal sequence bias [26], and consistently reproducible efficiency for every miRNA sequence [31,32]. By incorporating hairpin structures in the microarray probe, base-pairing with the additional nucleotide incorporated during labeling, and empirical melting point-determination, the platform is capable of single-nucleotide discrimination in the miRNA sequences while specifically distinguishing the mature miRNAs from longer RNAs in the total RNA sample [26,31,32]. We chose to compare this microarray system against the Taqman qPCR system in particular, since at the time this work was performed this was the most commonly utilized miRNA qPCR system. We found excellent correlation between the microarray and PCR results for most of the miRNAs. We further examined two of the miRNAs showing low correlations by using spiked-in synthetic RNAs, and found that differen-

tial sensitivity between the two techniques is not the cause of the discrepancy.

Another factor which could potentially affect the results of an miRNA profiling study is the method used to isolate RNA from the biological sample. Both the Agilent microarray system and the TaqMan qPCR systems use total RNA as the starting material; however, it is unclear whether different total RNA preparation methods will yield systematically different miRNA profiling results. In this report, we compared the results of miRNA microarray profiling obtained with three different commonly used total RNA prep methods. We found that the results for most miRNAs were equivalent among the different sample preparation methods, but that measured levels of a small number of miRNAs differed systematically.

## Results and discussion

### **Quantitative RT-PCR and Agilent microarray miRNA profiles correlate strongly**

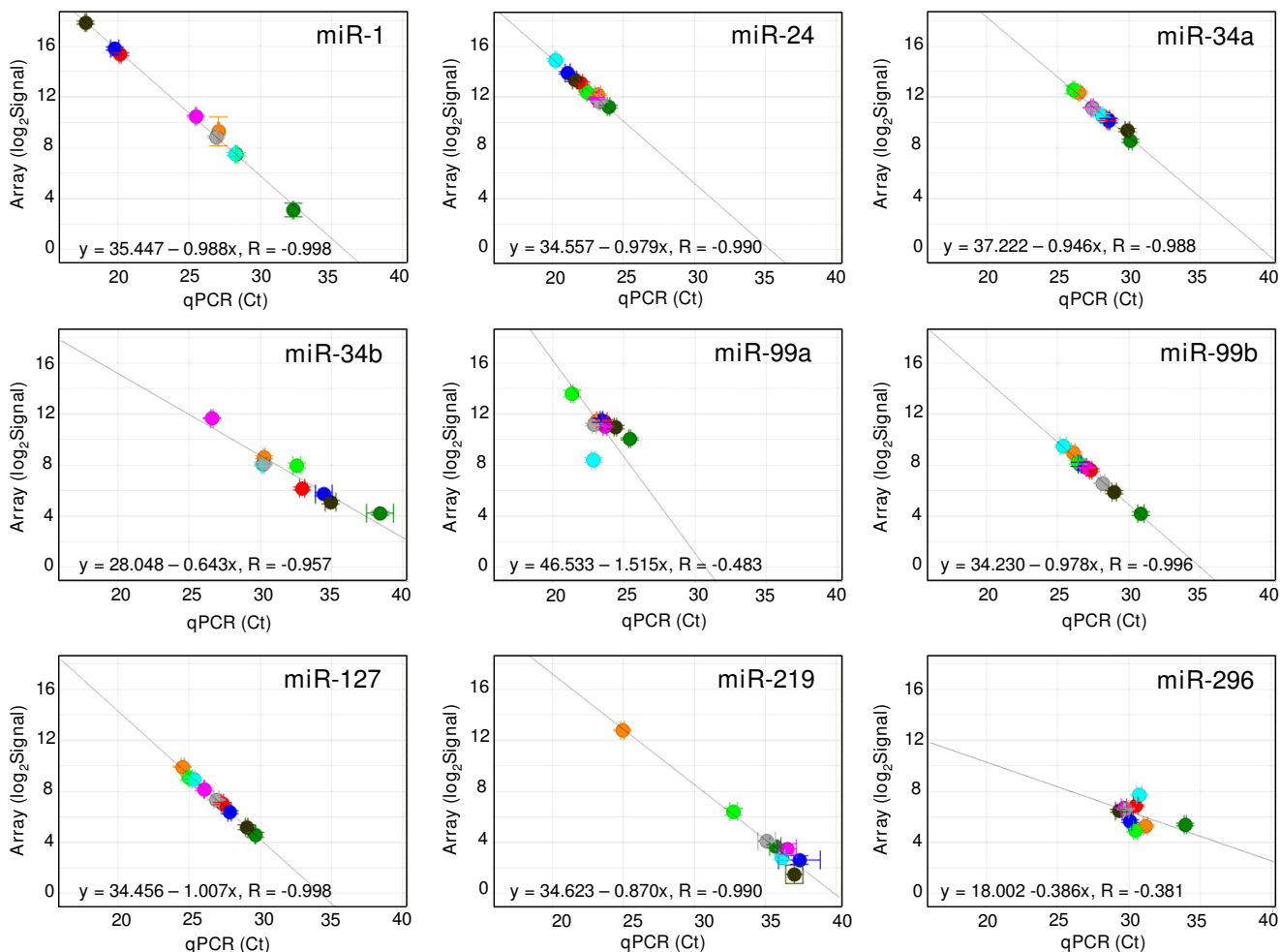
We previously reported that a comparison of Agilent microarray profiling and SYBR green-based quantitative RT-PCR (qPCR) of ten miRNAs in seven different human tissues found the two measurements correlated quite well [26]. To perform a more extensive comparison, we analyzed the expression of 61 human miRNAs in nine different tissues (brain, breast, heart, liver, placenta, testes, ovary, skeletal muscle, thymus), using both Agilent miRNA microarrays and TaqMan stem-loop qRT-PCR [15]. Aliquots of the same RNA samples were used for both the microarray and qPCR measurements. We chose these particular 61 miRNAs for several reasons. First, they represent a wide range of expression levels, as determined in an initial array analysis of some of the tissues. Second, they have wide differences in GC content, ranging from 23% (miR-190) to 68% (miR-328). Third, we chose several miRNAs which had potentially problematic sequences or exhibited atypical behavior during the development of the Agilent microarray platform: two of these did not show as good a linear titration curve as other miRNAs tested in a previous study (miR-126\*, miR-296) [26], and two other miRNAs were previously reported not to be labeled by enzymatic methods similar (but not identical) to that used with the Agilent microarray assay (miR-208, miR-219) [33].

Of the 61 miRNAs examined, only miR-637 was not detected by either method in any of the tissues. The rest of the miRNAs assayed were detected in most or all of the tissues by both methods, with two exceptions: miR-208, expressed only in the heart [34] and at very low levels in skeletal muscle, and miR-138, expressed in the brain, and at lower levels in placenta and thymus (all data is shown in Additional File 1).

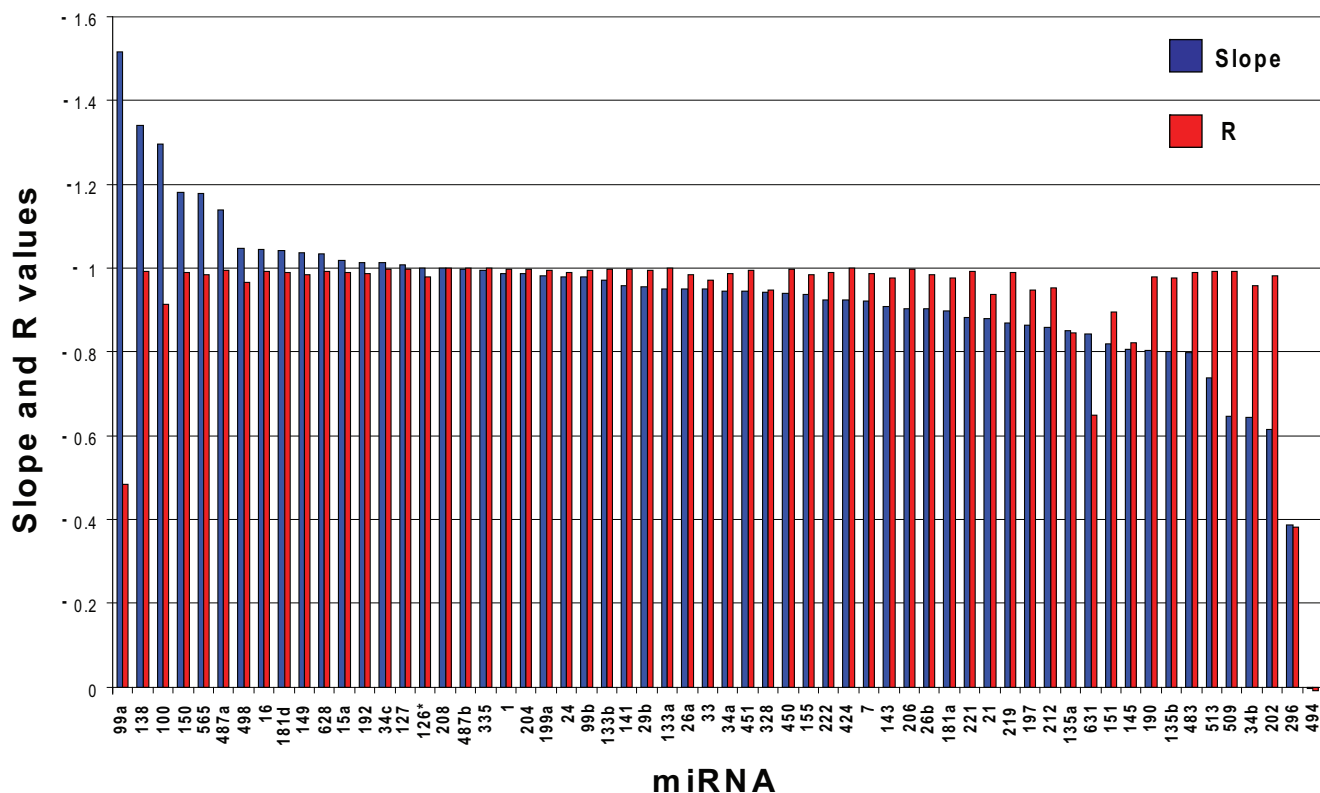
The qPCR and microarray results were compared by plotting the qPCR cycle threshold (Ct) value versus the log<sub>2</sub> of the array signal for each miRNA in all nine tissues (representative plots are shown in Figure 1, with the remaining plots shown in Additional Files 2, 3, 4, 5, 6, 7). These two values should be directly comparable, since both the qPCR Ct value and the log<sub>2</sub> of the microarray signal change by a value of 1 for every 2-fold change in miRNA concentration. If the qPCR and microarray measurements are equivalent, the plots will show a linear correlation (R = -1) with a slope of -1. Figure 2 shows the slopes and the

correlation values for each of the 60 miRNAs. 56 of the 60 miRNAs show correlation values (R) between -0.8 and -1.0, and 50/60 plots have slopes between -1.2 and -0.8. Of the four miRNAs which were selected as potentially problematic in the microarray measurements, only miR-296 did not correlate between the microarray and qPCR assays; miR-208, miR-219, and miR-126\* all gave excellent correlations.

To examine the results for all 60 miRNAs on one plot, we cannot simply plot qPCR Ct values versus microarray sig-



**Figure 1**  
**Comparison of qPCR and microarray miRNA profiling for individual miRNAs in nine human tissues.** Scatter plots are shown for 9 of the 61 miRNAs assayed, with qPCR results (cycle threshold (Ct) values) on the x axes and microarray results (log<sub>2</sub> of the total gene signal) on the y axes. Each data point represents one tissue. All plots are drawn to the same scale. The equations and R values on each plot are for the orthogonally-fitted line. Spot colors indicate the tissue: red = breast, pink = testes, dark blue = heart, light blue = placenta, dark green = liver, light green = ovary, orange = brain, brown = skeletal muscle, and grey = thymus. Tissues where qPCR results were flagged as "undetermined" by ABI software, or where log<sub>2</sub> of the total gene signal on arrays was < 1, were not plotted. Error bars indicate standard deviation (SD) of Ct values for qPCR results and (SD/Mean)\*log<sub>2</sub>e of the signals for the array results. Scatter plots for the remaining 51 miRNAs (one miRNA gave no signals with either qPCR or arrays) are in Additional Files 2, 3, 4, 5, 6, 7.



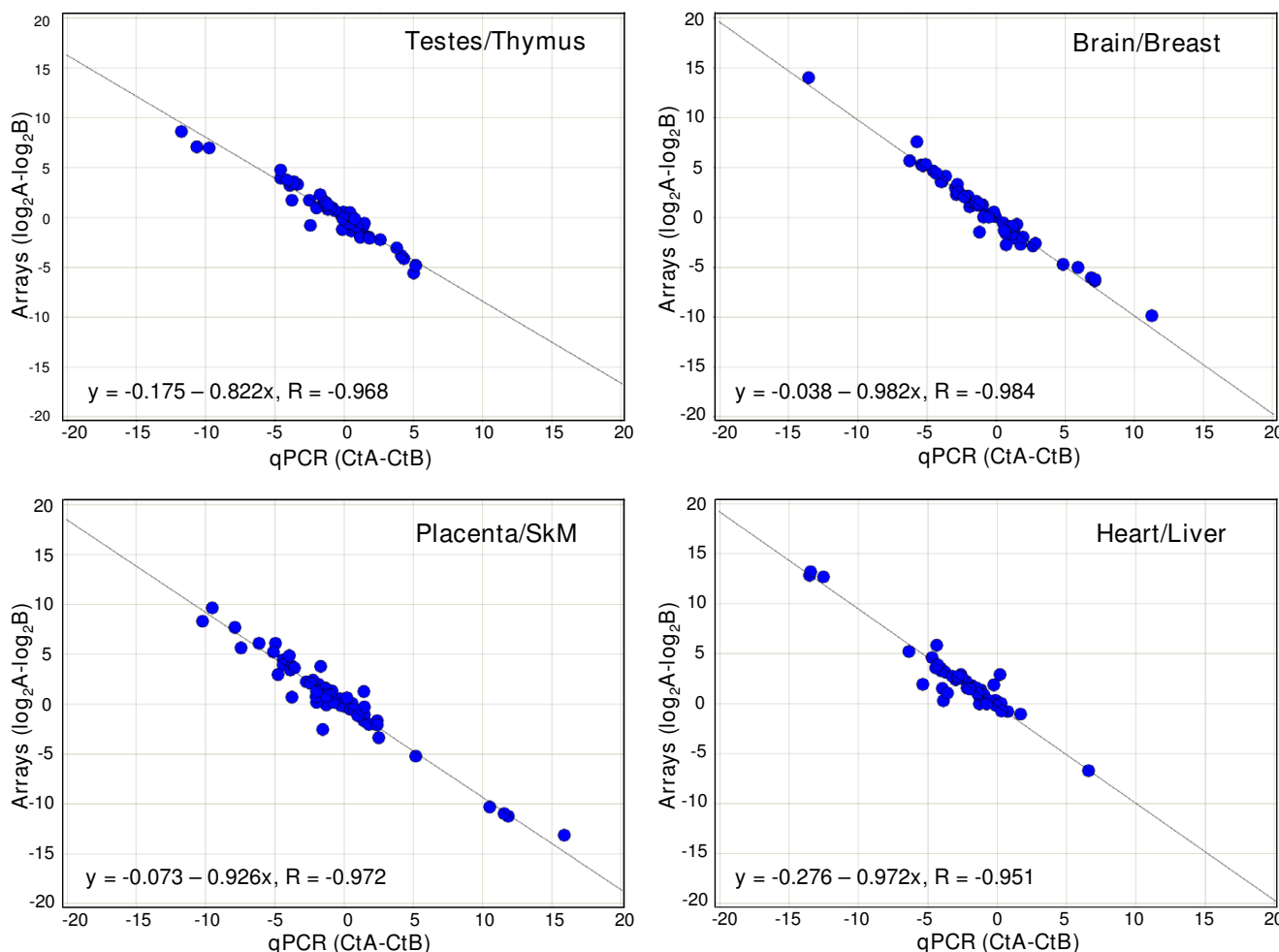
**Figure 2**  
**Slope and R values for all 60 miRNA scatter plots.** The slopes and R values for the orthogonal fit lines for all 60 of the miRNAs plotted in Figure 1 and Additional Files 2, 3, 4, 5, 6, 7 are ordered by the slope values. Slopes are shown in blue and R values are in red.

nals for all miRNAs in all tissues, because both the qPCR and microarray assays have differential sensitivities to different miRNAs. Thus, instead of looking at absolute expression levels, we must look at relative ratios of miRNA expression between two different tissues. To judge the consistency of fold-changes measured by microarray and qPCR platforms, we plotted the ratios of miRNA expression between all 36 possible pairs of tissues as measured by qPCR (Ct(tissue1)-Ct(tissue2)) and by microarrays (log2(signal in tissue1)-log2(signal in tissue2)). Four such plots are shown in Figure 3 (the other 32 plots are shown in Additional Files 8, 9, 10, 11, 12, 13, 14, 15), while Figure 4 shows the slopes and R values for each of the plots for the 36 tissue pairs. The plots all show very good correlation between the qPCR and array ratios, with R values between -0.984 and -0.821. The slopes of the 36 plots vary between -1.05 and -0.793. The intercepts of these fold-change plots (shown in Figures 3 and Additional Files 8, 9, 10, 11, 12, 13, 14, 15) indicate the consistency between fold-changes measured by the two methods. The mean of the intercepts of the line fits for the 36 tissue pairs was 0.00 +/- 0.23 (1 SD) (data not shown).

This level of variability (17%) is comparable to that seen among independent measurements using the same technique.

**Measurements of spike-ins of miRNAs which systematically differ between platforms show linear sensitivity**

miR-494, miR-296, and miR-99a are the three miRNAs that exhibit the most discrepant correlation values and slopes between the qPCR and microarray assays (Fig. 2); however, if the measurement of miR-99a in placenta is omitted, the slope for this miRNA becomes -0.844 with R = -0.929 (see plot in Fig. 1). For miR-494 and miR-296, if one platform were measuring levels of these miRNAs accurately, while the other platform were not, then we might expect a significant divergence from linearity to be observed between the two measurements when adding increasing amounts of synthetic miR-494 or miR-296 RNA into a total RNA sample. To test this, we added 1 zmol to 10 fmols of synthetic miR-296 and miR-494 RNAs to 100 ng of total RNA from liver or placenta, and measured the qPCR and microarray responses (Figure 5).



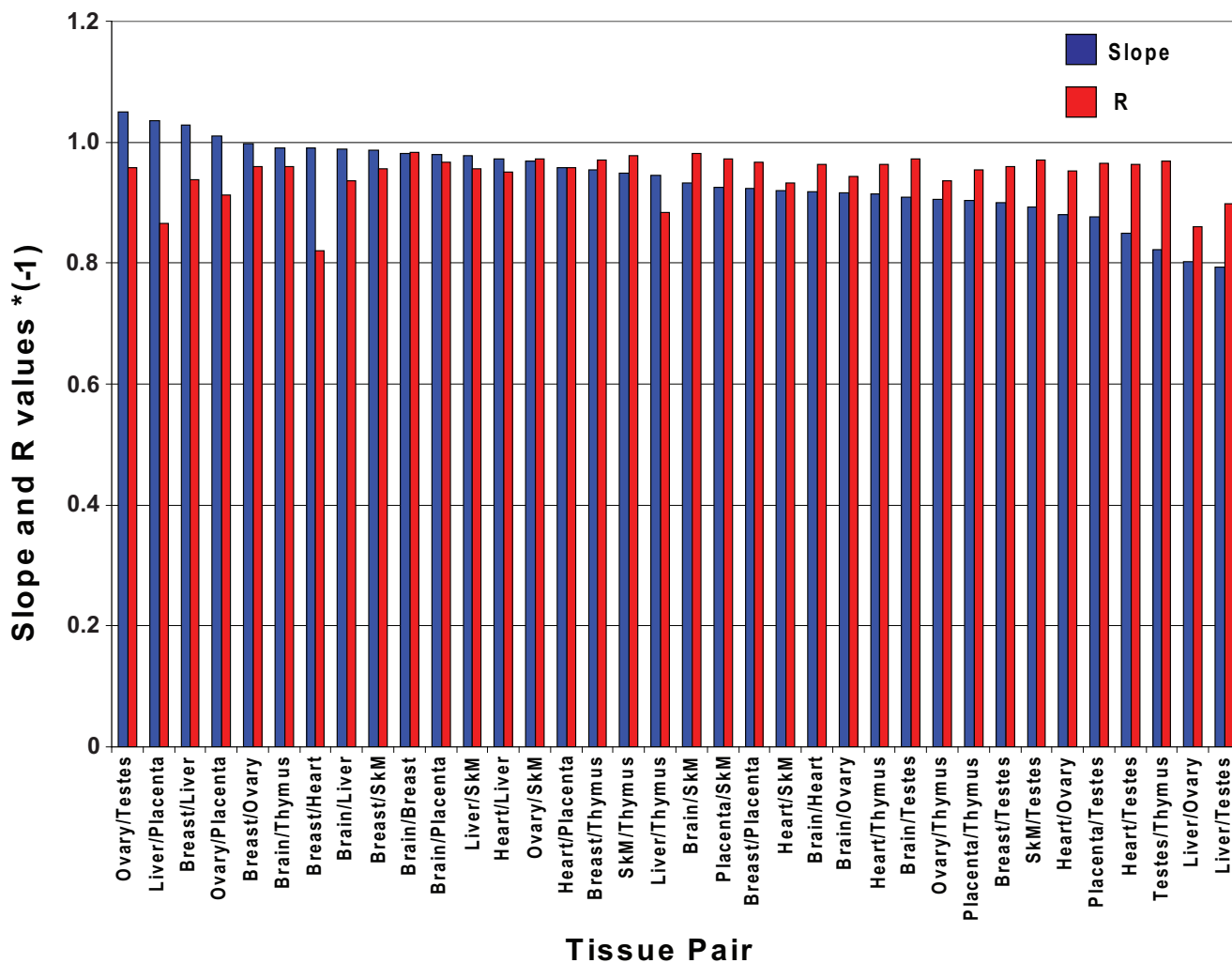
**Figure 3**  
**Comparison of qPCR and microarray miRNA measurements for 60 miRNAs in four tissue pairs.** Scatter plots are shown for miRNA expression ratios in four different tissue pairs, as determined by qPCR (x axis) and microarrays (y axis), where each data point represents one miRNA. The qPCR values are the difference between the Ct values from the two tissues, and the microarray values are the difference between the  $\log_2(\text{total gene signals})$  from the two tissues. The equations and R values on each plot are for the orthogonally-fitted line. Scatter plots for the remaining 32 tissue pairs are shown in Additional Files 8, 9, 10, 11, 12, 13, 14, 15. SkM = Skeletal Muscle.

For both miRNAs, in both tissues, the relation between qPCR measurement and array measurement is linear above a threshold spike-in concentration. The R values of the linear regions are very close to -1, with slopes between -0.842 and -0.935, indicating that both the qPCR and the microarrays are producing sample-responsive and internally consistent measurements of miR-296 and miR-494 at these concentration levels. Below the threshold spike-in levels, the qPCR Ct values and microarray signals are unchanged for miR-494, while for miR-296 the Ct values increase slightly, but the array measurements are unchanged. We conclude that the difference between the two platforms is not due to different sensitivity, since both the microarray and qPCR measurements are capable of

measuring miR-296 and miR-494 accurately above a spike-in concentration threshold. Presumably some type of interference confounds the measurement of endogenous expression levels in the complex sample, on one or both of the platforms.

**Three different total RNA preparation methods show similar yields and quality**

A question which arises when comparing the miRNA profiling results reported in different studies is whether the methods used to isolate RNA from tissue or cell line samples systematically affect the miRNA profiles. To examine whether the miRNA profile of a sample is affected by the type of total RNA prep method used, we prepared one



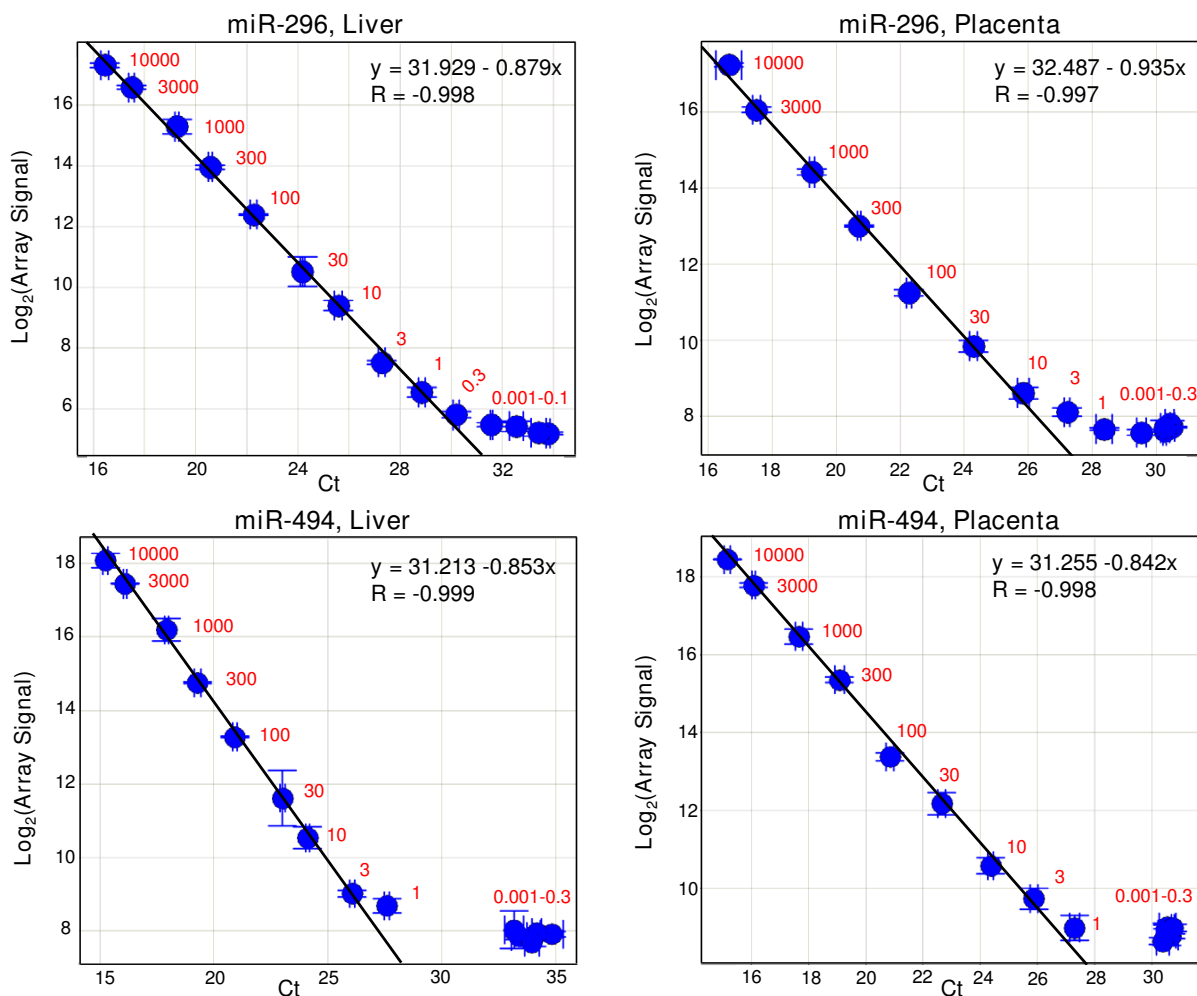
**Figure 4**  
**Slope and R values for all 36 tissue pair plots.** The slopes and R values for the orthogonal fit lines for all 36 possible tissue pairs plotted in Figure 3 and Additional Files 8, 9, 10, 11, 12, 13, 14, 15 are ordered by the slope values. Slopes are shown in blue and R values are in red.

large frozen cell pellet from each of two different human cell lines, HeLa (a cervical carcinoma line) and ZR-75-1 (a breast carcinoma line). We then subdivided these pellets into equal aliquots, and performed total RNA isolation on the aliquots using three different techniques: phenol/guanidinium (TRIzol, Invitrogen) followed by isopropanol precipitation, and two column-based techniques, miRNeasy (QIAGEN) and *mirVana* (ABI). Four to eleven replicate preps were performed on each cell type with each method.

Mean RNA yields, as measured by absorbance at 260 nm, and quality metrics for each prep type are shown in Table 1. The RNA integrities of the preps were analyzed on the Agilent 2100 Bioanalyzer, and all the preps had high quality RNA according to the RIN number [35,36]. This indi-

cates that most of the RNAs in the various preps were intact, with minimal breakdown. However, RIN values do not provide information about non-RNA contaminants, such as organic reagents and DNA. The TRIzol preps showed the lowest mean 260/230 ratios, possibly indicating the presence of some remaining TRIzol reagent in the final product.

Since the absorbance at 260 nm is used to quantitate the amount of RNA for use in the measurement assays, and since the 260:230 ratios can only serve as a crude guideline to possible contaminants, it is important to examine the absorption spectra in more detail (Figure 6). Some of the spectra clearly show the presence of additional peaks between 220–230 nm, indicating the presence of contam-



**Figure 5**  
**Comparison of qPCR and microarray measurements for miR-296 and miR-494 titrations.** Scatter plots are shown for titration of synthetic miR-296 into liver (top left panel) and placenta (top right panel) total RNAs, and miR-494 into liver (lower left) and placenta (lower right) total RNAs. Ct values from qPCR are plotted on the x-axis, while log<sub>2</sub> of the total gene signal from microarray measurements are plotted on the y-axis. Numbers in red show the number of attomoles of spike-in miRNA per 100 ng total RNA. Equations and R values are for the orthogonal line fit of the linear regions of each titration. Error bars indicate standard deviation (SD) of Ct values for qPCR results and (SD/Mean)\*log<sub>2</sub>e for the array results.

**Table 1: List and characteristics of sample preps from HeLa and ZR-75-1 cell pellets.**

Cell Line	Prep Type	No. of Preps	RNA Yield Mean (SD)	260:280 Mean (SD)	260:230 Mean (SD)	RIN Mean (SD)
HeLa	TRIzol	10	45.59 (7.17)	1.89 (0.0406)	1.28 (0.208)	9.9 (0.14)
HeLa	mirVana	8	43.16 (12.01)	1.96 (0.0364)	1.36 (0.323)	9.9 (0.11)
HeLa	miRNeasy	11	34.07 (5.40)	2.08 (0.00874)	1.97 (0.389)	9.9 (0.11)
ZR-75-1	TRIzol	4	23.52 (1.39)	1.84 (0.0432)	0.815 (0.0759)	9.8 (0.075)
ZR-75-1	mirVana	4	16.92 (5.84)	2.01 (0.0356)	1.28 (0.361)	9.1 (0.17)
ZR-75-1	miRNeasy	4	18.37 (1.93)	2.05 (0.0311)	1.87 (0.337)	9.8 (0.050)

The number of individual preps performed using the indicated total RNA prep method and cell lines are shown, as are the mean and standard deviations (in parentheses) of the total RNA yields (in micrograms), 260:280 ratios, 260:230 ratios, and RIN numbers for all the preps done with the same method in each cell type. Preps started with either  $5 \times 10^6$  cells (HeLa) or  $1 \times 10^7$  cells (ZR-75-1).

















