# Quality Assessment and Data Analysis for microRNA Expression Arrays Supplementary Material

D. Sarkar[a]*, R. Parkin[a], S. Wyman[a], A. Bendoraite[a], C. Sather[a], J. Delrow[a], A. Godwin[b], C. Drescher[a], W. Huber[c], R. Gentleman[a], M. Tewari[a]

October 16, 2008

[a]Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA
[b]Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, PA 19111, USA
[c]EBI-EMBL, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK

# 1 Materials and Methods

## 1.1 Experimental Methods

Primary human ovarian surface epithelial (HOSE) cells were obtained from the normal ovaries of women using a modification of the technique described previously (Dyck et al., 1996). Samples were collected from 4 different women (age range 42-80 years; mean age 58 years), of which all were postmenopausal. Patients underwent oophorectomy for several different reasons. Two of the women reported a family history of breast and/or ovarian cancer, three of the women reported postmenopausal bleeding and all were diagnosed with endometrial cancer. In all cases, specimens were taken from normal appearing ovarian surface epithelium, which was confirmed following histopathological review. Two cases were found to contain a benign ovarian cyst upon review. Ovarian tissue specimens were removed either by laparotomy or laparoscopy. The ovarian surface epithelium was scraped from the ovarian surface and the cells were collected by centrifugation at 2000 $g$ for five minutes. The cell pellets were plated into 25 $cm^2$ tissue culture flasks using media described previously (Li et al., 2004). Flasks were incubated at 37$^\circ$C in 5% $CO_2$ without disturbance for two days to allow the cells to colonize. The medium was changed two times per week until cells were confluent, after which cells were trypsinized and transferred to new tissue culture flasks using 0.025% trypsin/0.01% at a ratio of 1:4. RNA was isolated from primary HOSE cells at passage 4 to 7.

Snap frozen ovarian cancer tissue specimens corresponding to serous and endometrioid histologies were obtained from the Pacific Ovarian Cancer Research Consortium Repository. All clinical samples in this study were collected under Institutional Review Board-approved protocols.

*to whom correspondece should be addressed

### 1.1.1 Microarrays

Spotted microRNA arrays were manufactured using a set of 1448 locked nucleic acid (LNA) capture probes purchased from Exiqon, Inc (Woburn, MA) and includes LNA probes for all microRNAs in all organisms as annotated in miRBase Release 8.1 (Griffiths-Jones et al., 2007). The probe set covers over 92% of microRNAs annotated in miRBase 9.0. In addition, 146 proprietary probes (these are referred to as miRPlus) are included that are not part of miRBase. Each probe is spotted in duplicate (spatially separated) on Schott Nexterion's SLIDE E (epoxy) surface. Each slide consists of two independent microRNA arrays, which can be hybridized to different sample pairs. Pre-hybridized arrays were treated following Schott's protocol for the Nexterion's SLIDE surface.

Experimental and reference samples were combined and resuspended in 200 $\mu l$ of 1x miRCURY hybridization buffer (Exiqon, Inc., Woburn, MA), incubated for 5 minutes at 95°C, and subsequently centrifuged for 5 minutes at 14,000xg.

Combined sample pairs were co-hybridized to each array using the 1SureHyb Microarray Hybridization System (Agilent Technologies, Santa Clara, CA). Hybridization chambers were rotated at 4 rpm at 60°C for 16-18 hours using a Robbins Scientific 400 Hybridization Incubator (SciGene Corporation, Sunnyville, CA). Post-hybridized arrays were washed for 2 minutes in 2XSSC + 0.2% SDS at 60°C, followed by 2x washes in 1xSSC at RT for 2 min each, and finishing with a final wash in 0.2XSSC at RT for 2 min. Slides were dried by centrifugation at 200xg for 5 min and scanned at 10 $\mu m$ resolution using a GenePix 4000B scanner (Molecular Devices Corporation, Sunnyvale, CA). Image quantification was performed using GenePix Pro 6.0 Microarray Image Analysis software (Molecular Devices Corporation, Sunnyvale, CA).

### 1.1.2 Spike-ins

We selected forty non-human microRNAs that did not show cross-hybridization with multiple human tissue RNA samples tested on the arrays. Synthetic RNA oligos were synthesized for each non-human microRNA based on their miRBase sequence (Qiagen). The pooled oligos were first tested on the arrays at one femtomole each. Oligos giving signal intensities below 400 were eliminated from the pool and the concentrations of some of the remaining fifteen oligos were adjusted to maximize the range of the signal they provided, mimicking the span of intensities of a biological sample. The final pool of synthetic non-human microRNAs used as spikeins: Cel-miR-55, Osa-miR437, Cel-miR-267, Cel-miR-65, Osa-miR440, Cel-miR-66, Osa-miR442, Cel-miR-37, Cel-miR-39, Cel-miR-238, Cel-miR-248, Osa-miR414, Cel-miR-254, Cel-miR-50, and Cel-miR-54.

The quantity of cel-miR-267 was increased to two femtomole, cel-miR-65 was increased to four femtomole, and osa-miR414 was increased to six femtomole. The other twelve oligos remained at one femtomole in the labeling reaction. The pool of spike-in oligos at their appropriate quantity was added to each labeling reaction in a volume of two microliters. Signal intensity of the spike-in controls was used to assess the labeling efficiency of each sample hybridized on the arrays.

### 1.1.3 Green channel reference

Synthetic human microRNA universal reference pool RNA oligos were synthesized for 454 microRNAs based on sequences in miRBase (Qiagen). All oligos were first pooled and tested on arrays in uniform quantities between 0.01 femtomole and three picomole to determine a general working

| Signal intensity range | Number of miroRNAs | Final input in reaction |
|---|---|---|
| 10,000-60,000 | 61 + 59 | 0.5 femtomole |
| 5000-9,999 | 65 | 1 femtomole |
| 2,000-4,999 | 65 | 2 femtomole |
| 1,000-1,999 | 51 | 2.5 femtomole |
| 0-999 | 97 | 4 femtomole |
| No signal | 21 | NS |

Table S-1: Concentrations used for synthetic oligomers in creating the universal reference. A further 35 microRNAs from the initial pool of 454 were omitted as their signal intensities were weak at all tested concentrations. Some spots on the microarray represent multiple microRNAs (e.g., hsa-miR-518f*-526a, hsa-miR-133a-133b, etc.), giving a total of 448 unique probes for human microRNAs.

concentration. Due to the wide variation of signal intensity in the pool, the oligos were split into eight subpools grouped by signal intensity. Each pool was tested on arrays in several quantities. Thirty-five oligos were not included in the pools due to lack of signal above background at all tested concentrations. An additional subpool of 21 oligos was later removed from the final reference pool because the signal intensity remained weak regardless of increased quantites (Table S-1). This left a final pool of 398 microRNAs.

The universal reference was prepared from the seven subpools so that the appropriate quantity of each oligo was added to the labeling reaction in a volume of seven microliters (the two categories with the highest intensity were combined into a single row in Table S-1). The full list of microRNAs and the concentration at which they exist in the universal reference pool is given in Supplemental Table S1.

In order to assess whether specific properties of the microRNAs were responsible for the differences in intensity we examined two specific quantities associated with each microRNA to look for correlates. In particular we found an association with nucleotide composition, higher intensities (and hence lower concentrations used) were associated with relatively higher G and A content, as well as relatively lower C and U content (Figure S-1).

We also noticed that some of those microRNAs with low intensity had palindromic or near palindromic sequences. For example, hsa-miR-101, with mature sequence `UACAGUACUGUGAUAACUGAAG`, contains this ten nucleotide subsequence `ACAGUACUGU` which is its own reverse complement. We took a cutoff at a 5 nucleotide sequence as our cut-off (similar results are obtained with other choices) and found a consistent pattern of increased percentage of microRNAs with a palindromic sequence as the concentration used increased, and the largest percentage (43%) in those where no signal was obtained.

## 1.2 Applicability to other microarray platforms

The effectiveness of spike-in controls depends on the choice of controls and their concentrations, which need to be tuned to the microarray platform used. This is possible as long as the relevant platform contains a sufficient number of probes that do not cross-hybridize to human microRNAs. Although further experiments are required to confirm suitability, an initial set of candidate microR-

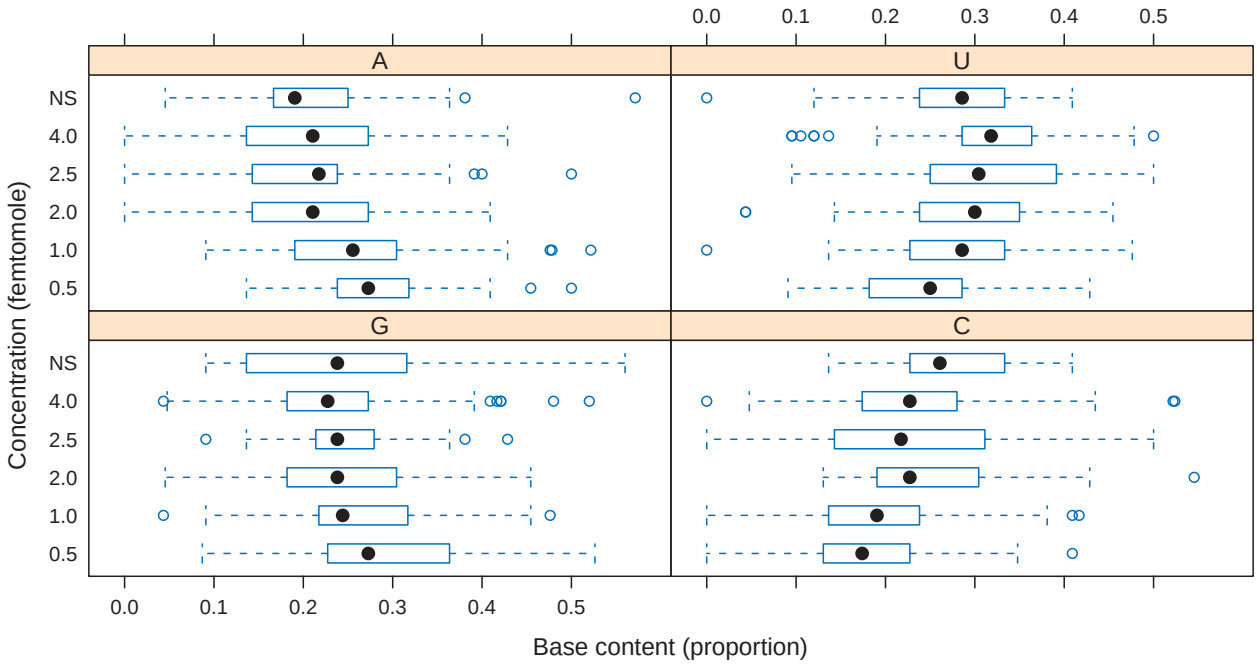


Figure S-1: Association of nucleotide composition with concentration of oligos in the synthetic human microRNA universal reference pool. NS indicates the "no signal" subpool.

NAs can be obtained by computing a similarity measure between each microRNA and the set of known human microRNAs.

We computed such a measure for all mature microRNA sequences available in miRBase 12.0. It is based on the best local alignment of the mature sequence of a microRNA to all human microRNA sequences, using a score function that assigns a score of $+1$ to a match, $-1$ to a mismatch, and $-1$ to a gaps of length one (longer gaps are not allowed). The score of the best alignment can thus be interpreted as the length of the alignment with penalties for mismatches and gaps. MicroRNAs with low similarity to any known human microRNA are natural candidates for spike-ins, provided they are present on the array platform of interest. However, one must also ensure that the spike-ins do not cross-hybridize with each other. Supplemental Table S2 gives a list of potential spike-in candidates for popular commercial arrays, ordering microRNAs available on a platform[1] by increasing similarity with human microRNAs, after removing those that were more similar to a previous candidate on the list than to a human microRNA. Shorter versions of these tables are given in the main text (Tables 2–6). R code to compute the similarity score is provided as a script in the `microRNA` package available from the Bioconductor project (`http://www.bioconductor.org`).

[1]The contents of each platform were inferred from publicly available information posted at vendor websites and GEO. In some cases, the list was not explicit; in other cases, microRNA names were formatted differently and needed to be transformed before comparison. Although we did our best to use all available information, some non-human microRNAs that could be potentially useful as spike-ins may have been omitted in our analysis. Nonetheless, for each of the array platforms described, we identified numerous good candidates for use as spike-in reagents, as listed in Tables 2-6 and Supplemental Table S2.

## 1.3 Quality Assessment

In our experiment, microarray hybridization was performed in two separate batches. Plots designed to highlight the batch effect, such as Figure 2 in the main paper, can serve as diagnostics for assessing different normalization methods, as normalization should remove or at least reduce the evidence of a batch effect. Figure S-2 looks at the pairwise "distance" between all arrays, where the distance between two arrays is defined as the median of the absolute differences in expression for human microRNAs.

It is difficult to account for possible print-tip effects by performing within-array normalization. Such normalization is usually based on the assumption that there is a common distribution of intensities for all print-tip blocks, or that all print-tip blocks are approximately the same with respect to observed intensity values.

In our case, the densities vary more across print-tips than across arrays, presumably due to the small number of expressed spots per print-tip. This makes within-array normalization difficult, but density-based diagnostic plots can make identification of problematic print-tip blocks easier. There do not appear to be any such problematic cases in our data.

We may also look for unusual effects at the level of spots. Figure S-3 shows a scatter plot of the median and MAD (for each spot, across arrays) of log un-normalised expression in the reference channel for spots that correspond to known human microRNAs. We see two spots that have unusually high variability across arrays, and should potentially be discarded from the analysis. In Figure S-4, we look at the expressions for the spots with the highest variability using a false-color image plot, which confirms an odd behaviour for the two spots identified in the location-spread plot.

## 1.4 Normalization

The VSN method (Huber et al., 2002) combines background correction and normalization into one single procedure. For a data matrix $((x_{ki}))$, where $k$ indexes probes and $i$ indexes arrays, it estimates a per-array transformation $h_i$ that is used to transform the $x_{ki}$ values as

$$x_{ki} \mapsto h(x_{ki}) = \mathrm{glog}\left(\frac{x_{ki} - a_i}{b_i}\right) \tag{1}$$

where $a_i$ and $b_i$ are a background offset and a scale parameter for array $i$, and the glog transformation is the so-called generalized logarithm or attenuated logarithm. In our example, VSN normalization is performed with $k$ ranging over probes with known human microRNA targets.

A similar procedure is also used for spike-in VSN normalization. In this case, $k$ ranges over the spike-ins, but the estimated transformations $h_i$ are then applied to all probes. Figure S-5 provides motivation for spike-in normalization.

Figure 4 in the main paper compares normalization methods using the variability across arrays of the spike-ins. For another assessment independent of the spike-ins, Figure S-6 indicates the variability of four U6-snRNA spots (which are believed to have constant expression in all samples) relative to human microRNA spots. In fact, expression of the U6-snRNA spots are often used for normalization; the advantage to using spike-ins is that they provide explicit control on the range of expression values in which we are interested.

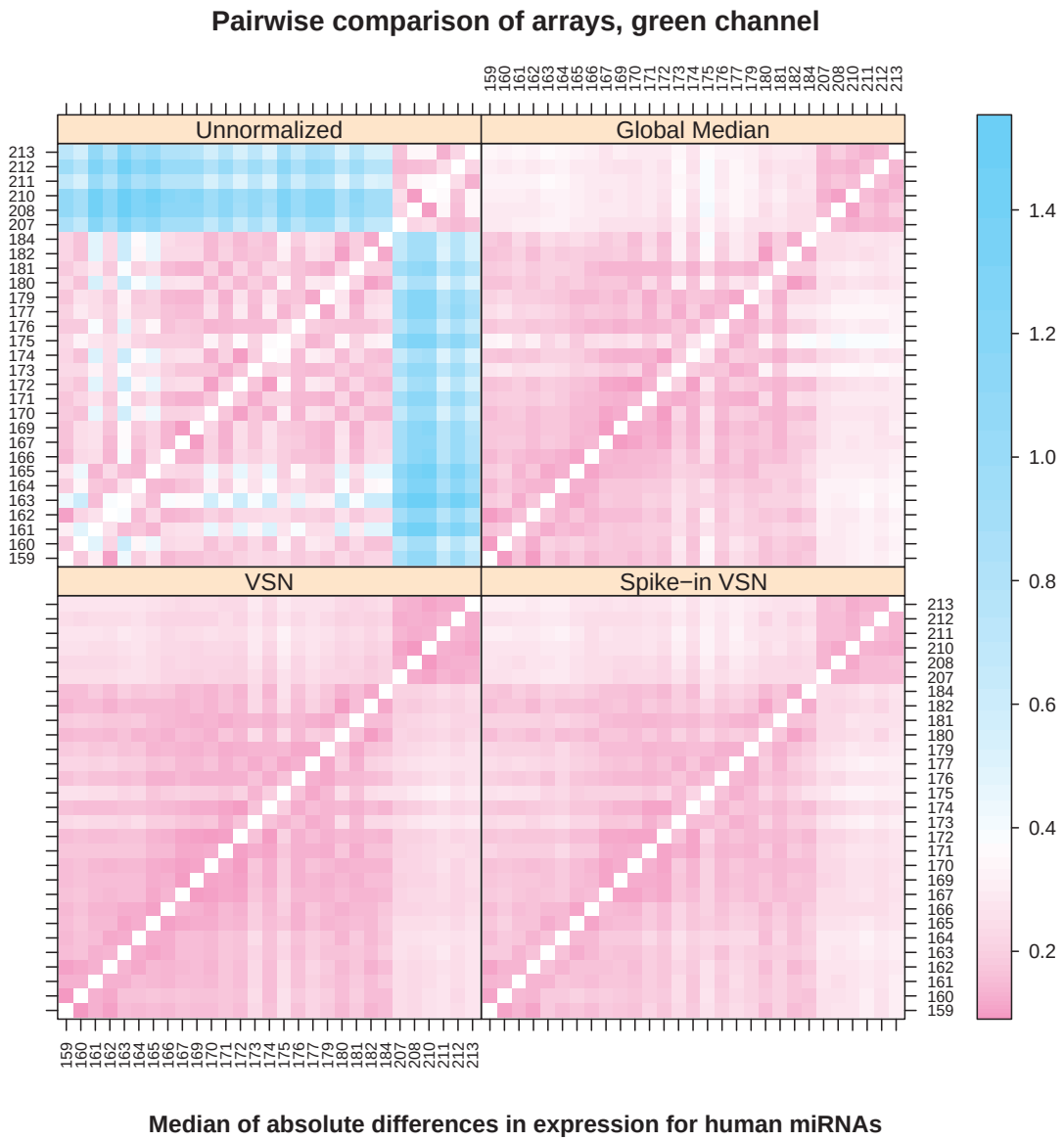


Figure S-2: Pairwise "distance" between arrays, as measured by the median of the absolute differences in expression for human microRNAs in the green channel. By design, these spots should all be expressed, with the same expression in all arrays. In practice, the arrays of the second batch are quite distinct from those in the first; however, all normalization methods decrease this distance considerably, indicating that we have improved going from unnormalized to normalized data. Note that the color breaks are on a log scale, emphasizing variation at the lower end.

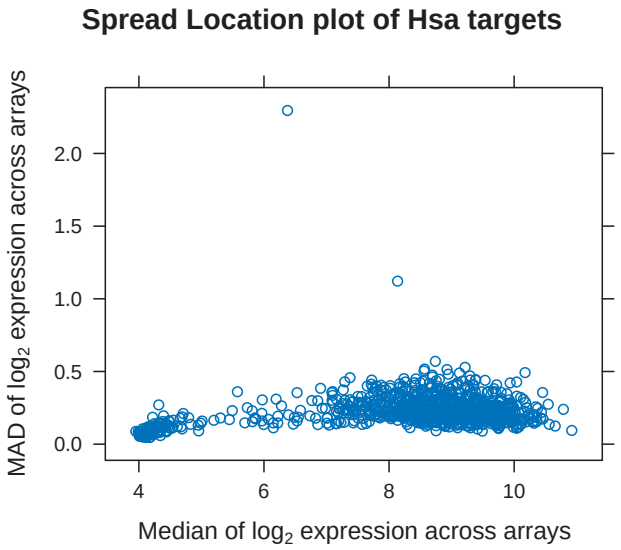


Figure S-3: A scatter plot of spread (MAD) *vs* location (median) across arrays of unnormalized log expression in the reference channel, for human microRNAs. Since the reference channel should be the same on all arrays, we expect their variability to be low. We see a few spots with low median expression (indicating a failure of the reference), but more interestingly, two spots with unusually high MAD.

## 1.5 Comparing normalization methods using the reference channel

In addition to the methods described in the main paper, the reference channel values provides another opportunity to assess the different normalization techniques by comparing the variances of expressions across arrays. Since the reference channel expressions are designed to be constant across arrays, this variability should be small. Figure S-7 compares the variabilities (as measured by the MAD) of each human microRNA across arrays for unnormalized expressions and the three sets of normalized expressions using a scatter-plot matrix.

## 1.6 Comparison with qRT-PCR measurements

We used qRT-PCR to asssess the levels of seventeen microRNAs. These can be used to assess the performance of various normalization methods. We expect the normalized expressions to correlate well with the qRT-PCR measurements, and better normalization schemes to have stronger correlation. The supplementary file `qrtpcr.pdf` plots the qRT-PCR values against expression measures, and Table S-2 summarizes the agreement between the two using observed correlation. The observed correlations are fairly strong for most microRNAs, both for unnormalized and normalized expressions. Analysis of variance (Table S-3) suggests that the amount of correlation generally does not differ significantly between normalization methods.

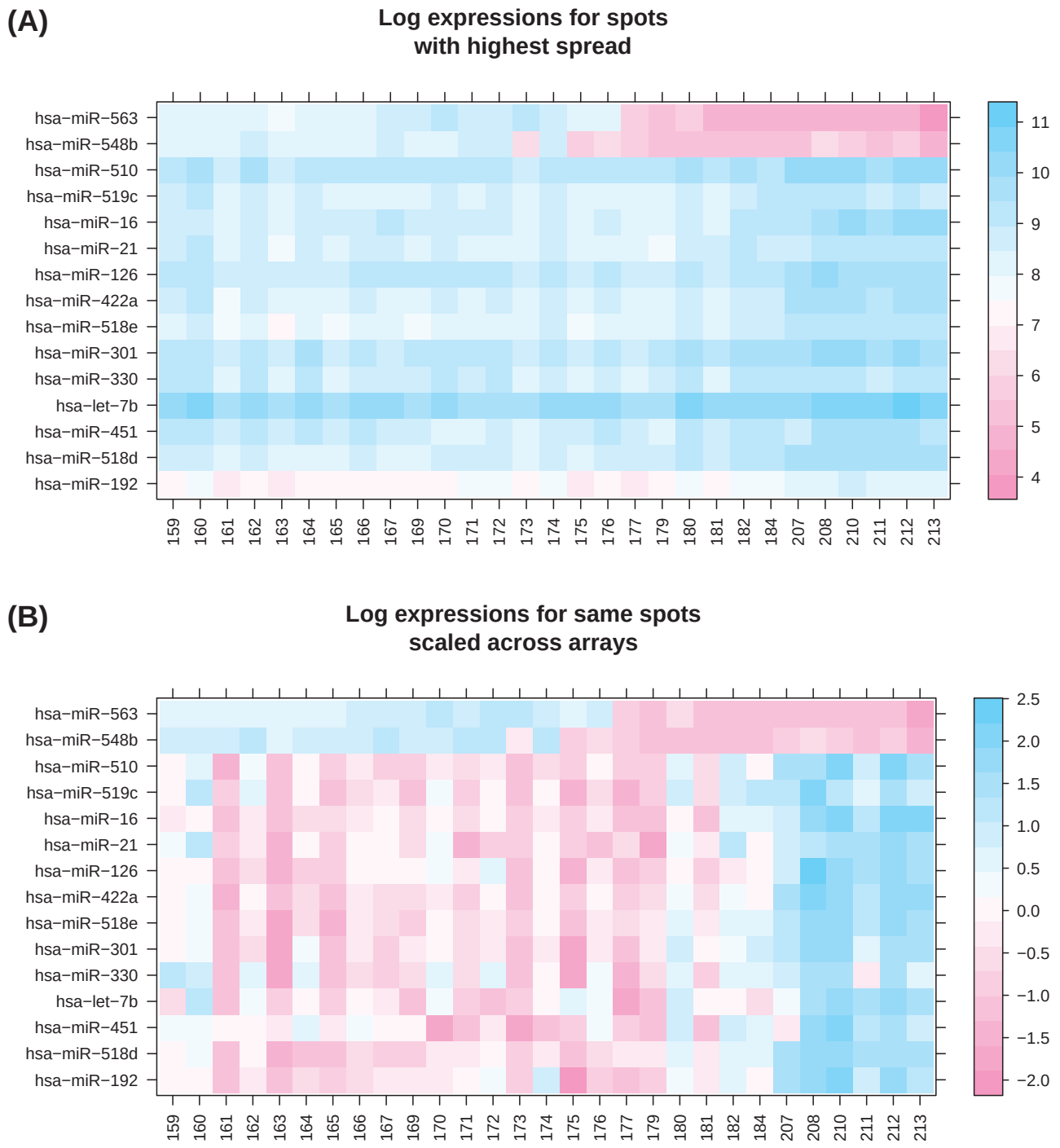


Figure S-4: False color plots showing unnormalized log expression of spots with highest variability. Columns represent arrays and rows represent spots. (A) shows unscaled log expressions, and (B) shows log expressions after centering and scaling across arrays. The rightmost two spots are clearly unusual in a systematic manner (the expression change is related to the order of the arrays). (B) also shows systematic patterns for the other spots; these are consistent with the batch effect seen in Figure 2.

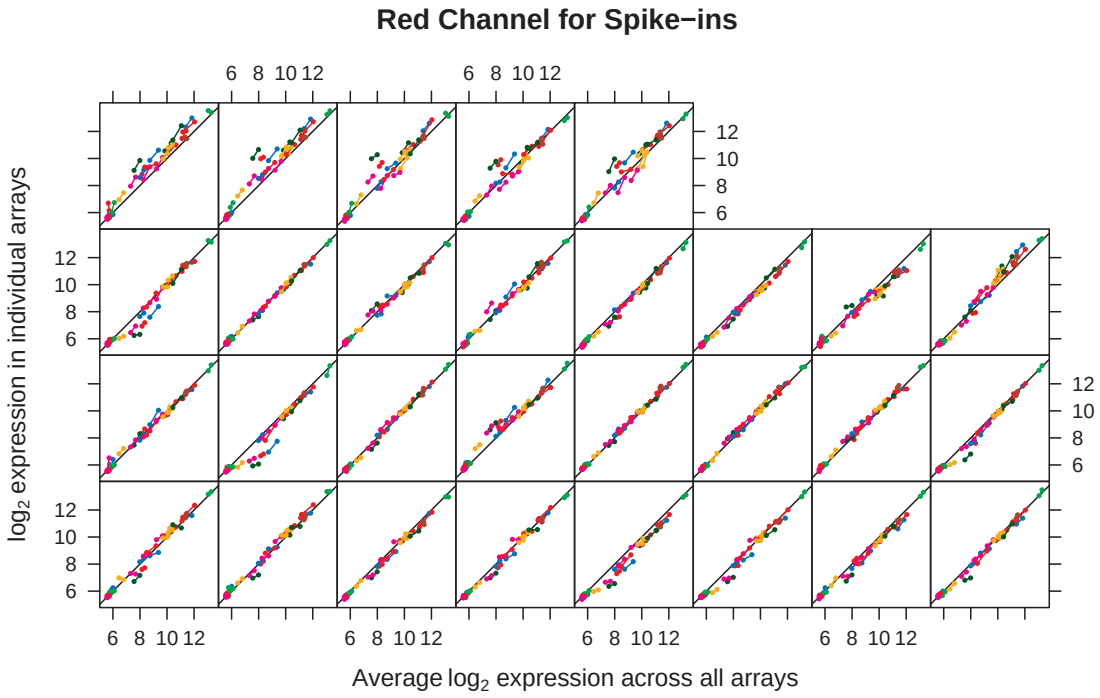


Figure S-5: Log expression values in the red channel for spiked-in microRNAs. Values in each array are plotted against the coresponding average across all arrays. Ideally, all points should fall on the diagonal if no normalization were required. While this is not true, the deviations can be mostly accounted for by a scale and location change. Duplicate spots are shown in the same color and joined by a line. This demonstrates a subtle effect, i.e., the change in expression for two duplicates of a control are small but almost always in the same direction. This suggests a spatial effect, which may also explain the banding pattern in Figure 2. Since the effect is small compared to the range of expressions across spike-in controls, we ignore this spatial aspect in subsequent analysis.

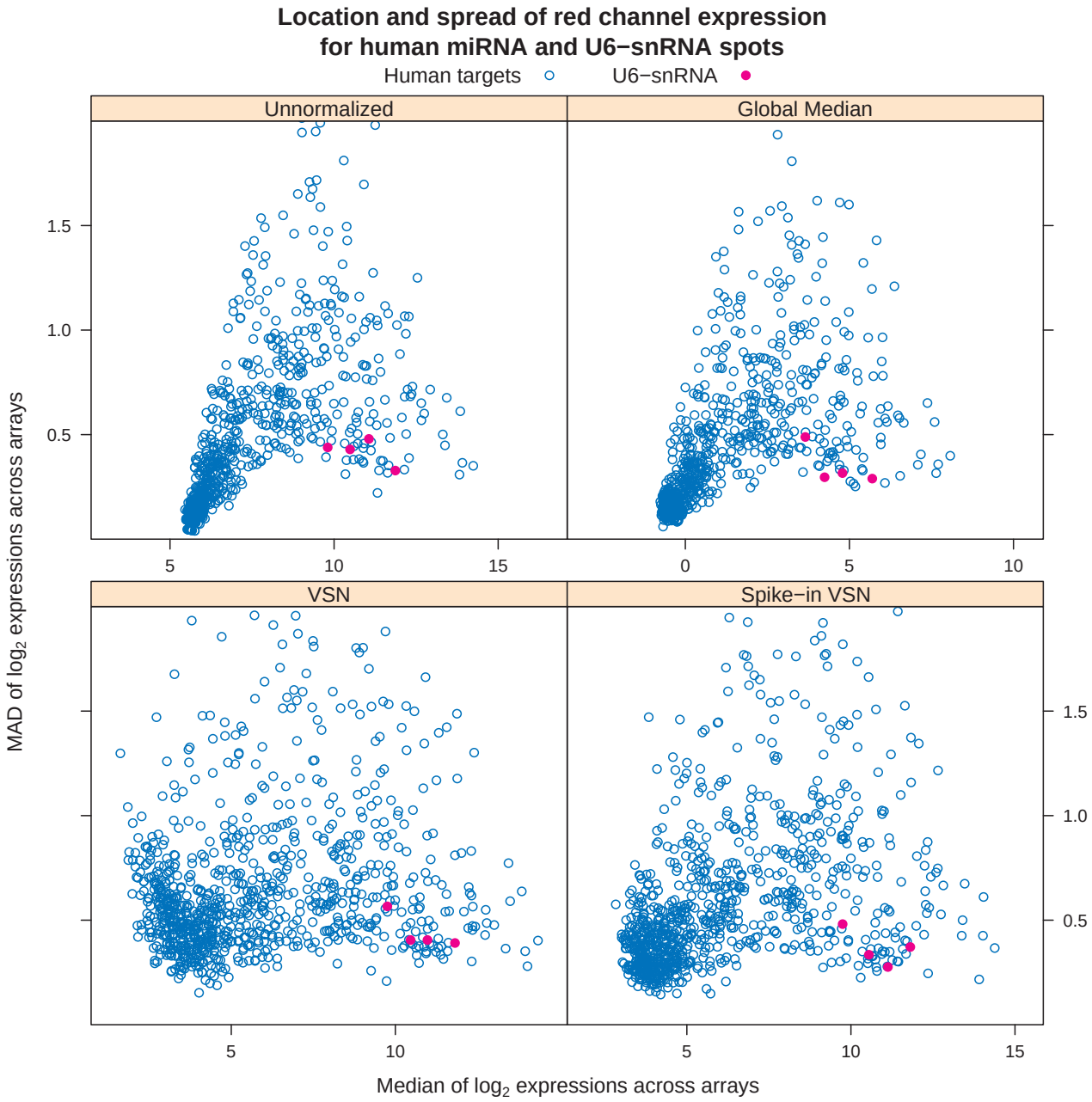


Figure S-6: Spread-location plot for probes on the red channel. This plot is identical to Figure 4 in the main paper, except that instead of the spike-ins, the highlighted points represent U6-snRNA spots, which are also believed to remain consistent across arrays. The relative performance of the methods can be assessed by the criterion of minimizing variability of the U6-snRNA spots.

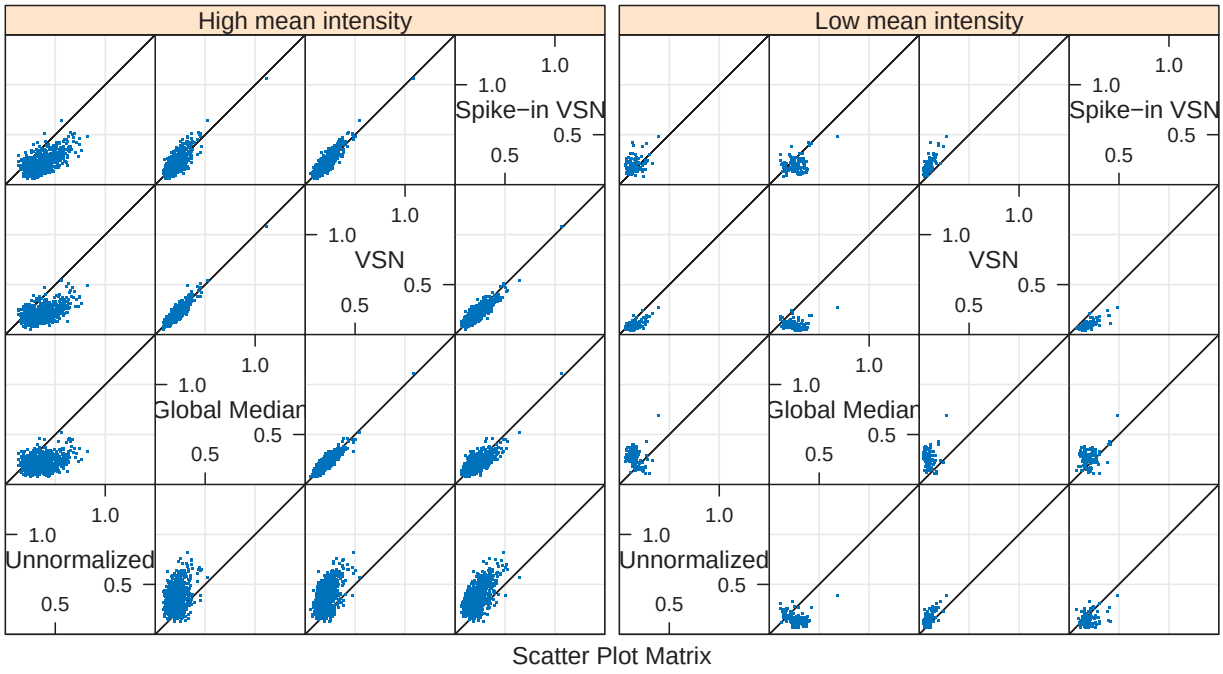


Figure S-7: MAD of each human microRNA across arrays for the four sets of intensities in the reference channel. The right panel plots the microRNAs which have average unnormalized log-expression less than 8, a heuristic cutoff suggested by Figure 3. Global median normalization performs somewhat poorly for low-intensity spots. Note, however, that these are unlikely to be expressed at all, and would be rejected by the non-specific filtering step. This comparison is somewhat unfair to spike-in normalization as it only assumes that the spike-ins should remain constant across arrays, whereas in fact this is true for all probes in the green channel.

| | raw | GM | VSN | SI |
|---|---|---|---|---|
| hsa-miR-146b | −0.826 | −0.823 | −0.792 | −0.783 |
| hsa-miR-205 | −0.897 | −0.872 | −0.932 | −0.927 |
| hsa-miR-200a | −0.900 | −0.895 | −0.899 | −0.899 |
| hsa-miR-182 | −0.917 | −0.911 | −0.929 | −0.925 |
| hsa-miR-96 | −0.828 | −0.802 | −0.832 | −0.839 |
| hsa-miR-100 | −0.844 | −0.854 | −0.846 | −0.804 |
| hsa-miR-155 | −0.591 | −0.600 | −0.566 | −0.513 |
| hsa-miR-422b | −0.833 | −0.820 | −0.859 | −0.855 |
| hsa-miR-200b | −0.888 | −0.881 | −0.889 | −0.889 |
| hsa-miR-370 | −0.615 | −0.604 | −0.650 | −0.647 |
| hsa-miR-519e* | −0.226 | −0.219 | −0.238 | −0.162 |
| hsa-miR-10a | −0.747 | −0.762 | −0.756 | −0.728 |
| hsa-miR-10b | −0.649 | −0.660 | −0.633 | −0.609 |
| hsa-miR-449 | −0.820 | −0.802 | −0.851 | −0.861 |
| hsa-miR-21 | −0.642 | −0.661 | −0.646 | −0.604 |
| hsa-miR-107 | −0.743 | −0.719 | −0.713 | −0.693 |
| hsa-miR-30a-5p | −0.439 | −0.388 | −0.398 | −0.433 |

Table S-2: Pearson correlations between log expression and average $\Delta$Ct in qRT-PCR, using all samples, for 17 microRNAs. Since a higher number of cycles indicates lower initial amount, strong negative correlation indicates agreement.

**Results from Hua et al. (2008)**

Hua et al. (2008) also compare various normalization methods for microRNA microarrays, using correlation with PCR results to quantify performance. We subject their results to a similar analysis of variance, with results given in Table S-4. Although they report print-tip loess normalization as the method that performs best, we find no statistically significant difference between it, median normalization, and VSN.

We have not considered print-tip loess normalization in our analysis, as it does not seem to be applicable to our data; in particular, a substantial proportion of microRNAs appear to be unexpressed in our samples. This seems not to be true for Hua et al. (2008), as indicated by the M vs A plot in their Figure 4.

# 2 Results

## 2.1 Differential Expression

In Tables S-5 through S-8, we present the list of the top 10 differentially expressed microRNAs, as reported by limma (Smyth, 2005), between the serous and endometroid subtypes of ovarian cancer for various normalization schemes. When available, the table is augmented by the result of PCR validation (in the form of a $p$-value).

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Unnormalized (Baseline) | −0.7297 | 0.0469 | −15.5595 |
| Global Median | 0.0078 | 0.0076 | 1.0234 |
| VSN | −0.0013 | 0.0076 | −0.1704 |
| Spike-in VSN | 0.0136 | 0.0076 | 1.7853 |

Table S-3: Summary of a mixed effect model fit with correlations as response, and type of normalization as predictor. More negative correlation indicates better agreement. All seventeen microRNAs are included in a single model, with a random effect for the microRNAs. The `lme4` package (Bates, 2007) was used to fit the model. The first row represents the baseline or intercept, which in this case corresponds to correlation with unnormalized intensities. The remaining rows represent the effect of each normalization method relative to this baseline. VSN normalization shows the strongest correlation on average, but the differences between the four approaches are not statistically significant.

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Printtip loess (Baseline) | 0.4075 | 0.1311 | 3.1091 |
| Median | −0.0838 | 0.1135 | −0.7376 |
| VSN | −0.0475 | 0.1135 | −0.4183 |
| Unnormalized | −0.3575 | 0.1135 | −3.1484 |

Table S-4: Summary of a mixed effect model fit with correlations reported by Hua et al. (2008) as response, and type of normalization as predictor. Higher correlation is better. Correlations for all eight microRNAs listed in Table 1 are included in a single model, with a random effect for the microRNAs. The first row represents the baseline, which here corresponds to correlation with print-tip loess normalized intensities. The remaining rows represent the effect of each normalization method relative to this baseline. The performance of VSN and median normalization are not significantly different.

# References


Douglas Bates. *lme4: Linear mixed-effects models using S4 classes*, 2007. R package version 0.99875-9.

H. G. Dyck, T. C. Hamilton, A. K. Godwin, H. T. Lynch, S. Maines-Bandiera, and N. Auersperg. Autonomy of the epithelial phenotype in human ovarian surface epithelium: changes with neoplastic progression and with a family history of ovarian cancer. *Int J Cancer*, 69(6):429–436, 1996.

S. Griffiths-Jones, H. K. Saini, S. Dongen, and A. J. Enright. miRBase: tools for microRNA genomics. *Nucleic Acids Research*, 2007.

You-Jia Hua, Kang Tu, Zhong-Yi Tang, Yi-Xue Li, and Hua-Sheng Xiao. Comparison of normalization methods with microRNA microarray. *Genomics*, 2008. doi: 10.1016/j.ygeno.2008.04.002.

Wolfgang Huber, Anja von Heydebreck, Holger Sueltmann, Annemarie Poustka, and Martin Vin-

| | logFC.array | t | adj.P.Val | logFC.PCR | pval.PCR |
|---|---|---|---|---|---|
| hsa-miR-422b | 2.15 | 6.430 | 0.000 | 1.878 | 0.000 |
| hsa-miR-146b | 1.53 | 5.712 | 0.000 | 1.842 | 0.005 |
| hsa-miR-625 | 1.24 | 5.289 | 0.001 | Not done | Not done |
| hsa-miR-155 | 1.41 | 4.167 | 0.013 | 3.139 | 0.012 |
| hsa-miR-200a | −1.18 | −3.829 | 0.025 | −0.943 | 0.058 |
| hsa-miR-200b | −1.23 | −3.441 | 0.050 | −0.914 | 0.054 |
| hsa-miR-193a | 0.94 | 3.400 | 0.050 | Not done | Not done |
| hsa-miR-99a | 1.42 | 3.379 | 0.050 | Not done | Not done |
| hsa-miR-148a | −1.01 | −3.216 | 0.056 | Not done | Not done |
| hsa-miR-23b | 1.07 | 3.189 | 0.056 | Not done | Not done |

Table S-5: Table of the top ten hits using Spike-in VSN normalization. This is the same as Table 1 in the main paper.

| | logFC.array | t | adj.P.Val | logFC.PCR | pval.PCR |
|---|---|---|---|---|---|
| hsa-miR-422b | 2.16 | 6.280 | 0.000 | 1.878 | 0.000 |
| hsa-miR-625 | 1.18 | 5.077 | 0.002 | Not done | Not done |
| hsa-miR-146b | 1.47 | 5.005 | 0.002 | 1.842 | 0.005 |
| hsa-miR-155 | 1.37 | 3.920 | 0.022 | 3.139 | 0.012 |
| hsa-miR-34a | −1.07 | −3.797 | 0.022 | Not done | Not done |
| hsa-miR-200a | −1.25 | −3.767 | 0.022 | −0.943 | 0.058 |
| hsa-miR-148a | −1.10 | −3.744 | 0.022 | Not done | Not done |
| hsa-miR-200b | −1.30 | −3.454 | 0.033 | −0.914 | 0.054 |
| hsa-miR-99a | 1.45 | 3.450 | 0.033 | Not done | Not done |
| hsa-miR-652 | −0.53 | −3.415 | 0.033 | Not done | Not done |

Table S-6: Table of top hits using VSN normalization

gron. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18 Suppl. 1:S96–S104, 2002.

N. F. Li, G. Wilbanks, F. Balkwill, I. J. Jacobs, D. Dafou, and S. A. Gayther. A modified medium that significantly improves the growth of human normal ovarian surface epithelial (OSE) cells in vitro. *Laboratory Investigation*, 84(7):923–931, 2004.

Gordon K. Smyth. Limma: linear models for microarray data. In R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, and W. Huber, editors, *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, pages 397–420. Springer, New York, 2005.

| | logFC.array | t | adj.P.Val | logFC.PCR | pval.PCR |
|---|---|---|---|---|---|
| hsa-miR-146b | 1.58 | 6.115 | 0.000 | 1.842 | 0.005 |
| hsa-miR-422b | 1.76 | 6.032 | 0.000 | 1.878 | 0.000 |
| hsa-miR-625 | 1.24 | 5.513 | 0.000 | Not done | Not done |
| hsa-miR-155 | 1.30 | 4.237 | 0.010 | 3.139 | 0.012 |
| hsa-miR-193a | 1.06 | 3.987 | 0.016 | Not done | Not done |
| hsa-miR-222 | 1.09 | 3.867 | 0.019 | Not done | Not done |
| hsa-miR-23b | 1.18 | 3.745 | 0.022 | Not done | Not done |
| hsa-let-7i | 0.76 | 3.451 | 0.041 | Not done | Not done |
| hsa-miR-200a | −1.04 | −3.366 | 0.041 | −0.943 | 0.058 |
| hsa-miR-193b | 0.93 | 3.352 | 0.041 | Not done | Not done |

Table S-7: Table of top hits using global median normalization

| | logFC.array | t | adj.P.Val | logFC.PCR | pval.PCR |
|---|---|---|---|---|---|
| hsa-miR-146b | 1.70 | 6.661 | 0.000 | 1.842 | 0.005 |
| hsa-miR-422b | 1.89 | 6.095 | 0.000 | 1.878 | 0.000 |
| hsa-miR-625 | 1.37 | 5.493 | 0.000 | Not done | Not done |
| hsa-miR-155 | 1.43 | 4.516 | 0.005 | 3.139 | 0.012 |
| hsa-miR-193a | 1.19 | 4.178 | 0.009 | Not done | Not done |
| hsa-miR-222 | 1.22 | 3.857 | 0.019 | Not done | Not done |
| hsa-miR-23b | 1.31 | 3.798 | 0.019 | Not done | Not done |
| hsa-miR-23a | 0.85 | 3.586 | 0.028 | Not done | Not done |
| hsa-miR-193b | 1.05 | 3.480 | 0.032 | Not done | Not done |
| hsa-let-7i | 0.88 | 3.451 | 0.032 | Not done | Not done |

Table S-8: Table of top hits without any normalization