Published in final edited form as:

Cell Rep. 2019 December 17; 29(12): 4212–4222.e5. doi:10.1016/j.celrep.2019.11.078.

Comparison of Reproducibility, Accuracy, Sensitivity, and Specificity of miRNA Quantification Platforms

Paula M. Godoy[1,2], **Andrea J. Barczak**[2], **Peter DeHoff**[3,4], **Srimeenakshi Srinivasan**[3,4], **Alton Etheridge**[5], **David Galas**[5], **Saumya Das**[6,8], **David J. Erle**[2,7,8], **Louise C. Laurent**[3,4,8,9,*]

[1]Division of Medical Oncology, Department of Medicine and Department of Developmental Biology, Washington University in Saint Louis, 4518 McKinley Ave., CB 8069, St. Louis, MO 63110, USA

[2]Lung Biology Center, University of California, San Francisco, UCSF Box 2922, San Francisco, CA 94143, USA

[3]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Diego, 9500 Gilman Drive, Mail Code 0695, La Jolla, CA 92093-0695, USA

[4]Sanford Consortium for Regenerative Medicine, 2880 Torrey Pines Scenic Drive, La Jolla, CA 92037, USA

[5]Pacific Northwest Research Institute, 720 Broadway, Seattle, WA 98122, USA

[6]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA 02114, USA

[7]Cardiovascular Research Institute, University of California, San Francisco, UCSF Box 2922, San Francisco, CA 94143, USA

[8]These authors contributed equally

[9]Lead Contact

SUMMARY

Given the increasing interest in their use as disease biomarkers, the establishment of reproducible, accurate, sensitive, and specific platforms for microRNA (miRNA) quantification in biofluids is of high priority. We compare four platforms for these characteristics: small RNA sequencing (RNA-seq), FirePlex, EdgeSeq, and nCounter. For a pool of synthetic miRNAs, coefficients of variation for technical replicates are lower for EdgeSeq (6.9%) and RNA-seq (8.2%) than for FirePlex

*Correspondence: llaurent@ucsd.edu.

AUTHOR CONTRIBUTIONS

P.M.G., A.J.B., S.D., D.J.E., and L.C.L. designed the study. P.M.G. isolated RNA from all plasma samples, generated cDNA libraries, and prepared samples for sequencing. D.J.E., A.J.B., and P.M.G. handled all correspondence with HTG Molecular. S.D. was the main correspondent with Abcam. L.C.L., P.D., and S.S. worked with NanoString to process samples on nCounter. P.M.G. analyzed the data. A.E. and D.G. generated spike-in data, and P.D. processed the data. P.M.G., D.J.E., S.D., and L.C.L. wrote the paper. All authors contributed to manuscript editing.

DECLARATION OF INTERESTS

S.D. is a founding member of Dyrnamix, which did not play any role in the funding, design, or interpretation of this study.

SUPPLEMENTAL INFORMATION

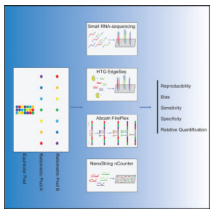
Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.11.078.

(22.4%); nCounter replicates are not performed. Receiver operating characteristic analysis for distinguishing present versus absent miRNAs shows small RNA-seq (area under curve 0.99) is superior to EdgeSeq (0.97), nCounter (0.94), and FirePlex (0.81). Expected differences in expression of placenta-associated miRNAs in plasma from pregnant and non-pregnant women are observed with RNA-seq and EdgeSeq, but not FirePlex or nCounter. These results indicate that differences in performance among miRNA profiling platforms impact ability to detect biological differences among samples and thus their relative utility for research and clinical use.

In Brief

Using pools of synthetic RNA oligonucleotides and standardized extracellular RNA samples, Godoy et al. compare small RNA sequencing to three targeted miRNA quantification platforms to evaluate reproducibility, bias, specificity and sensitivity, and accuracy. Each platform has strengths and limitations important to consider for biomarker discovery, clinical validation, and broad clinical use.

Graphical Abstract



INTRODUCTION

MicroRNAs (miRNAs) are a class of small (18- to 22-nt) non-coding RNAs with known roles in gene regulation (Bartel, 2004). miRNAs can be released from cells into the extracellular space and have been detected in all tested biological fluids (Godoy et al., 2018; Sohel, 2016). As potential indicators of tissue function, extracellular miRNAs have been proposed as possible prognostic and diagnostic biomarkers for a variety of diseases and for monitoring response to therapy (Das et al., 2019). qPCR, microarrays, and small RNA sequencing (RNA-seq) are measurement methods that are commonly used to study miRNA expression in tissues. Systematic comparisons between these methods have demonstrated their utility for research studies (Mestdagh et al., 2014; Giraldez et al., 2018; Yeri et al., 2018), but each of them has limitations that may impact their usefulness for quantification of extracellular miRNAs for clinical use. In particular, small RNA-seq is excellent for discovery studies but is less useful for high-throughput or rapid turnaround applications, while low sensitivity and long turnaround time are the major limitations for microarrays, and qPCR is not easily scalable to large numbers of miRNAs. Recently, several platforms have been developed specifically to address these gaps. Understanding differences in reproducibility, bias, and ability to detect biological differences across these platforms is important for selection of methods for translation of initial discovery-based exRNA (extracellular RNA) studies to large-scale clinical validation and actual clinical use. Here, we use synthetic miRNA pools and exRNA from plasma (Table S1) to build upon previous

studies by comparing a previously assessed small RNA-sequencing protocol (Giraldez et al., 2018) to three relatively novel platforms: HTG Molecular's EdgeSeq miRNA Whole Transcriptome Assay (EdgeSeq), Abcam's FirePlex (FirePlex), and NanoString's nCounter (nCounter).

For all four miRNA quantification platforms, the numerical readout for each interrogated miRNA correlates with its abundance in the tested sample. However, how these platforms measure miRNA varies widely (Table S2). EdgeSeq, FirePlex, and nCounter are targeted platforms, which detect only those miRNAs for which target-specific probes are included in the assay. Small RNA-seq, on the other hand, is a discovery platform, which captures small RNA sequences with a 5' phosphate and 3' hydroxyl group by adding common 5' and 3' adapter sequences in a non-sequence-specific manner. The small RNA-seq method we use here was optimized for low-input samples and was shown to be less biased than other widely used commercial small RNA-sequencing methods (Giraldez et al., 2018). This was primarily achieved by modifying both the 5' and 3° adapters to include four degenerate nucleotides on the ends that are ligated to the RNA molecule (Jayaprakash et al., 2011).

EdgeSeq is a multiplexed nuclease protection assay with next-generation sequencing readout (Girard et al., 2016). First, probes containing sequences complementary to 2,083 specific miRNAs and flanking sequences for downstream amplification are incubated with the miRNA-containing sample. Probes that successfully hybridize to their cognate miRNA in the sample are protected from nuclease digestion, amplified with the addition of barcodes, and then sequenced. Therefore, the output for EdgeSeq is read count, as in small RNA-seq, but unlike small RNA-seq, the number of reads reflects the quantity of probes that were bound by miRNAs and protected from digestion.

The remaining two methods use probes and fluorescent reporters. The Multiplex Circulating FirePlex miRNA Assay (Ab-cam) is based on gel microparticle technology (Chapin et al., 2011). The FirePlex hydrogel particles contain a central region that binds specific miRNAs based on complementarity, as well as two separate end regions with differing fluorescent intensities that serve as a barcode for the central analyte region. Bound miRNAs are ligated to universal adapters and then eluted from the hydrogels for amplification by PCR using biotinylated primers specific for the universal adapters. Following amplification, the now-biotinylated miRNA targets are rehybridized to the hydrogel particles, and a fluorescent reporter specific for biotin is used for quantitative detection of fluorescence on a flow cytometer. Analysis of the fluorescence attributed to the biotin-specific reporter (representative of relative target miRNA abundance) combined with the unique fluorescent barcodes on each end of the hydrogel (specific signature for the miRNA target) allows for multiplexed detection of up to 68 individual miRNAs per assay.

The nCounter platform relies on hybridization of miRNAs to probes conjugated to unique fluorescent barcodes and can potentially assay up to 800 different targets at once (Geiss et al., 2008; Denaro et al., 2017). Unlike the other platforms we tested, nCounter does not require an amplification step and counts the total number of fluorescent barcodes to determine the quantity of miRNA molecules in the sample. EdgeSeq and FirePlex can use

isolated RNA or crude biofluid as input, while small RNA-seq and nCounter require isolated RNA.

miRNA quantification using each of the tested platforms involves several steps, each of which can display preferences for certain RNA molecules, resulting in differences in the efficiencies with which miRNAs are detected. For example, during small RNA-seq library preparation, adapters ligate more efficiently to some miRNAs than others, resulting in bias (Jayaprakash et al., 2011; Hafner et al., 2011), whereas in hybridization-based assays, efficiency of probe binding varies in a sequence-specific manner, leading to cross-hybridization (Wu et al., 2005). Additionally, incorporation of incorrect nucleotides can occur during amplification or sequencing, leading to alignment errors or cross-hybridization. These target-specific biases preclude using signal strength as a direct measure of the abundance of a particular miRNA and can lead to differences in the ability to detect and reproducibly quantify specific miRNAs.

The NIH-supported Extracellular RNA Communication Consortium (ERCC) was launched to establish foundational knowledge and technologies for extracellular RNA research (Das et al., 2019). Here, we report the results of an ERCC-supported miRNA analysis platform comparison that examined reproducibility, bias, specificity, and relative quantification using both defined pools of synthetic miRNAs and exRNA from pooled human plasma samples. The use of synthetic miRNA pools allowed us to assess performance using complex mixtures of miRNAs at known concentrations. The use of plasma exRNA samples allowed us to compare performance using a clinically relevant sample type and to assess the ability of two platforms to assay miRNAs directly, without RNA isolation.

RESULTS

Reproducibility Across Technical Replicates for Synthetic miRNA Pools

Three pools of synthetic miRNAs (see STAR Methods) were analyzed with each of the four platforms. The first pool, referred to as the equimolar pool, contained 759 synthetic human miRNAs and 393 synthetic non-human RNA oligonucleotides at the same molar concentration. The other two pools, referred to as ratiometric pools A and B, each contained 286 human miRNAs and 48 non-human miRNAs at different concentrations, with the absolute concentrations of individual miRNAs varying over a 10-fold range within each pool. The relative concentrations of a given miRNA between pool A and pool B varied from 1:10 to 10:1. To assess the reproducibility of each platform, we examined the coefficient of variation (CV) of each miRNA's signal intensity across technical replicates for RNA-seq, EdgeSeq, and FirePlex (Tables S3 and S4). Only miRNAs considered to be detectable were included in the analysis (see STAR Methods). Technical replicates were not performed by NanoString, and therefore, reproducibility could not be assessed for the nCounter assay. For the equimolar pool, the median CV was higher for FirePlex (22.4%) than for small RNA-seq (8.2%) and EdgeSeq (6.9%). CV decreased as signal increased for RNA-seq and EdgeSeq, but not for FirePlex (Figure 1A). Ratiometric pools A and B showed similar CVs as the equimolar pool, although CVs decreased as signal intensities increased for all platforms, including FirePlex (Figure S1; Tables S4 and S5). Overall, we concluded that technical reproducibility was higher for small RNA-seq and EdgeSeq than for FirePlex.

Assessing the Bias Associated with Each Platform

Equal quantities of two different miRNAs can result in different signal intensities due to detection bias. Determining detection bias for a set of miRNAs requires a comparison between the amounts of these miRNAs in a sample and the signal intensities associated with each miRNA. In biological samples, the miRNA concentrations are usually not known. To accurately assess the bias for each platform, we used the equimolar and ratiometric synthetic pools. For each pool, we calculated the expected signal intensity based on the known concentration of each component miRNA that is considered detectable and quantified the detection bias as the ratio of observed to expected counts (see STAR Methods).

Small RNA-seq exhibited the most bias, with many target RNA sequences displaying a substantially lower than expected signal ($\log_2$ detection bias <0) and a relatively small number of target RNA sequences showing a markedly higher than expected signal ($\log_2$ detection bias >0; Figure 1B). With small RNA-seq, only 31% of the miRNAs in the equimolar pool had signals that were within 2-fold of the median signal. EdgeSeq had the least bias (76% within 2-fold of the median signal), while nCounter (47%) and FirePlex (57%) were intermediate. Results obtained with the ratiometric pools were similar to results obtained with the equimolar pool (Figure 1B). Therefore, although the version of small RNA-seq that we used has lower bias than some other widely used methods, it nonetheless exhibited substantially more bias than the other three platforms.

Relationships between detection bias and miRNA sequences differed between platforms. For small RNA-seq, EdgeSeq, and to a lesser extent FirePlex, the GC content of an RNA sequence correlated with the detection bias (Figure 2). For EdgeSeq, miRNAs that were least efficiently detected all had low GC content (<35%). There was minimal if any evidence of an association between bias and GC content for nCounter (Figure 2). We also explored associations between the bias and the identity of the 5′ or 3′ nucleotide of the target miRNA for each of the platforms (Figure S2). In EdgeSeq, signal intensities of miRNAs with a 5′ cytosine (n = 64) were higher than those with a 5′ uracil (n = 200) (Bonferroni-adjusted $p = 1.5 \times 10^{-3}$). Both small RNA-seq and nCounter had higher signals for miRNAs with 3′ guanines (n = 218 and n = 85, respectively) compared to 3′ uracils (n = 365, n = 175, respectively; $p = 4.2 \times 10^{-3}$ and $p = 1.6 \times 10^{-2}$, respectively). Overall, GC content and the identities of the 5′ and 3′ nucleotides had effects on signal intensity that differed between platforms, but these factors are not sufficient to accurately predict or adjust for bias.

We next examined whether detection biases were consistent within platforms and compared biases across platforms. Biases were generally consistent within platforms, as demonstrated by correlating detection bias determined with one pool to detection bias determined with another pool (Figure S3). As expected, biases were less well correlated between platforms (Figure 3). The highest correlations were between EdgeSeq and small RNA-seq (R = 0.38) and EdgeSeq and FirePlex (R = 0.29). We conclude that detection biases for specific miRNAs differ substantially between platforms.

Specificity and Sensitivity Analysis

We used data from the equimolar pool to determine the ability of each platform to distinguish between synthetic miRNAs that were present from those that were absent. Depending upon the platform, false-positive signals for miRNAs not present in samples might be caused by a variety of phenomena, including incorrect nucleotide incorporation during reverse transcription or DNA amplification, sequencing errors, cross-hybridization to non-cognate miRNAs or other probes, or auto-fluorescence. All platforms showed some overlap between the distribution of signals for miRNAs that were present in the pool compared with those that were absent (Figure 4A). As one means to assess false positives, we calculated the proportion of miRNAs that were absent from the pool but had signals higher than the fifth percentile of miRNAs that were present in the pool. This proportion was lowest for small RNA-seq (31/2,081 miRNAs, 1.5%), intermediate for FirePlex (1/21, 4.8%), and nCounter (22/376, 5.9%), and highest for EdgeSeq (146/1,632, 8.9%). By other metrics, separation between present and absent miRNAs was also best for small RNA-seq (ratio of median signal for present to median signal for absent = 1,750; area under the receiver operating characteristic curve [AUC] = 0.99), intermediate for EdgeSeq (ratio = 728, AUC = 0.97) and nCounter (ratio = 1078, AUC = 0.94), and least for FirePlex (ratio = 125, AUC = 0.81) (Figures 4A and 4B). Overall, small RNA-seq was superior to the other platforms by each of these measures of sensitivity and specificity.

We investigated whether false-positive signals could be related to cross-detection of miRNAs with similar sequences within the synthetic pool (Figure S4). For EdgeSeq, the relatively small set of absent miRNAs with sequence similarity to present miRNAs did tend to have higher signals than other absent miRNAs. This was also observed with small RNA-seq but was not evident for nCounter. We were unable to assess FirePlex's ability to distinguish closely related sequences, since the smaller set of probes did not include any designed to recognize absent miRNAs that were similar in sequence to those present in the pool. Of the other three platforms, nCounter displayed the least evidence for cross-detection, whereas EdgeSeq and to a lesser extent RNA-seq showed some evidence of cross-detection of closely related sequences.

Relative Quantification of miRNAs

These four platforms are typically used for relative quantification (i.e., comparing the level of any particular miRNA between samples, such as case versus control). We used the ratiometric pools to assess each platform's accuracy for relative quantification (Figure 5). Each of the four platforms provided reasonably good estimates of ratios for most miRNAs, although all platforms were inaccurate for certain miRNAs. Root-mean-square error (RMSE) for miRNA log ratios for RNA-seq (0.45), EdgeSeq (0.47), FirePlex (0.58), and nCounter (0.46) were quite similar. This evaluation with synthetic miRNA pools indicates that the four platforms behave similarly well for relative quantification.

Analysis of Reproducibility and Complexity in Plasma Samples

To evaluate the reproducibility of each platform with biologically relevant samples, we analyzed exRNA isolated from human male plasma samples in quadruplicate using small RNA-seq and in triplicate using EdgeSeq. We could not evaluate reproducibility of purified

exRNA on FirePlex, because RNA from these samples was not included in the FirePlex assay, nor could we evaluate the reproducibility of purified exRNA on nCounter, because technical replicates were not performed by NanoString. Because EdgeSeq and FirePlex can also take as input a small volume of crude biofluid, we also analyzed exRNA directly from plasma in duplicate for EdgeSeq and triplicate for FirePlex. Reproducibility between technical replicates for RNA isolated from plasma was worse than that observed with the synthetic equimolar pool for small RNA-seq and EdgeSeq (Figure S5; Tables S5 and S6), likely due to the lower average concentration of each miRNA in the plasma RNA samples compared to the synthetic pools. For isolated RNA, the median CV was higher with small RNA-seq (33.4%) than with EdgeSeq (14.4%). As previously reported for RNA-seq (Srinivasan et al., 2019), CVs decreased as signal increased, but overall CVs remained higher for the plasma RNA samples than for the synthetic pools.

For crude plasma, the median CV was lower with EdgeSeq (17.8%) than with FirePlex (43.2%) (Figure S5). Signal intensities between data generated from isolated RNA versus crude plasma were moderately well correlated for EdgeSeq (R = 0.62; Figure S5). To assess whether the use of crude biofluid biased against miRNAs carried in certain subcompartments in plasma, we inspected signals for miRNAs previously found to be differentially associated with five different subcompartments in human serum: $CD63^+$ extracellular vesicles, $CD81/CD9^+$ EVs, $AGO2^+$ ribonucleoproteins (RNPs), high-density lipoprotein (HDL), and the lipoprotein-free fraction (LFF) (Srinivasan et al., 2019). We found no obvious systematic differences in signal according to the assigned subcompartment (Figure S5), suggesting that EdgeSeq is capable of detecting miRNAs that are preferentially associated with each subcompartment within crude plasma samples.

To assess whether the complexity of the RNA sample affects the measurement of miRNAs, we compared the relative signal intensities of spike-in synthetic RNA sequences where possible. We could not perform this analysis for two of the platforms; EdgeSeq has only one positive internal control, and the FirePlex panel we used did not contain any positive internal controls. For nCounter, there were nine internal positive controls, and for these, we compared relative signal intensities between datasets generated from the synthetic equimolar pool and RNA isolated from the male plasma pool. For small RNA-seq, 58 synthetic exogenous small RNA sequences were spiked in to a water- only sample and RNA isolated from the male plasma pool, and relative signal intensities were compared between the resulting datasets. The signal intensities correlated extremely well for both nCounter (R = 0.99) and small RNA-seq (R = 0.98, Pearson correlation) (Figure S6).

Analysis of Extracellular miRNA in Pregnant Female Plasma Samples

We next compared miRNA signals across platforms using the mean normalized signal intensities of pregnant female plasma samples from two donors and observed moderate correlation between small RNA-seq and EdgeSeq (R = 0.68), small RNA-seq and FirePlex (0.78), and EdgeSeq and FirePlex (0.74) but weaker correlation with nCounter (small RNA-seq R = 0.43, EdgeSeq R = 0.53, FirePlex = 0.63) (Figure S7). In the plasma samples, the weaker correlation between nCounter and the other three platforms appeared to be due to reduced sensitivity, as the nCounter measurements for the large majority of the miRNAs

were near the lower limit of detection (median for negative control probes = 16.5, median for all human miRNA probes = 18). In addition, the differences in the patterns of bias between nCounter and the other methods (as seen with the synthetic pools in Figure 3) may also contribute to the weaker correlation.

To assess whether there was agreement in results among the different platforms when comparing biologically distinct samples, we examined the expression of a cluster of 50 placenta-specific miRNAs in chromosome 19 known to be specifically expressed during pregnancy (Ouyang et al., 2014). For all platforms, we compared signals for these miRNAs in pregnant female plasma samples versus a pool of non-pregnant female plasma (see STAR Methods). For all platforms, these miRNAs generally yielded signals that were near the lower end of the range for all miRNAs, suggesting that the placenta-associated miRNAs present in pregnant female plasma are generally less abundant than many other miRNAs found in both pregnant female and non-pregnant female plasma. A statistically significant increase in the expression of these miRNAs in the plasma of pregnant women compared to non-pregnant women was observed for 13/13 miRNAs detected by small RNA-seq ($p = 1.2 \times 10^{-4}$, Mann-Whitney test) (Figure 6; Tables S5 and S6). EdgeSeq was also able to detect differences (10/11 miRNAs increased, $p = 9.8 \times 10^{-4}$), although fold differences tended to be smaller than with RNA-seq. FirePlex (2/5 miRNAs increased, $p = 0.41$) and nCounter (4/12 miRNAs increased, $p = 0.69$) did not show a significant difference in placenta-associated miRNAs. We conclude that RNA-seq and, to a lesser extent, EdgeSeq were able to detect differences in relatively low abundance miRNAs of placental origin, whereas FirePlex and nCounter were not.

DISCUSSION

Two major objectives of Phase 1 of the ERCC were to identify methods for robust and reproducible quantification of extracellular miRNAs in biological fluids and to establish their utility as biomarkers. Here, we explored four miRNA measurement platforms, three of which could serve as potential alternatives to small RNA-seq for relative quantification of extracellular miRNA. The platforms, which utilize diverse technologies, were selected based on rapid turnaround time and ease of use, properties that are attractive for biomarker assays. The four platforms evaluated in this study included a small RNA-sequencing protocol optimized for low-input samples that had previously been compared to the more widely known commercially available small RNA-seq library preparation kits (Giraldez et al., 2018), EdgeSeq, FirePlex, and nCounter. A recent report compared the performance of small RNA-seq, FirePlex, EdgeSeq, and QIAGEN miRNome on standardized samples of brain, liver, and placenta (Yeri et al., 2018). To our knowledge, a systematic study using extracellular RNA, crude biofluid, and synthetic RNA mixes that allow for accurate evaluation of assay performance has not been previously performed. Overall, our findings (summarized in Table 1) demonstrated that each platform had specific advantages and drawbacks that need to be taken into consideration when selecting a technology for exRNA studies, particularly those aimed at development and clinical application of extracellular miRNA biomarkers.

Our study evaluated the performance of each platform with respect to reproducibility, bias, specificity, and relative quantification using standardized samples. Our samples consisted of three pools of synthetic miRNAs, one of which contained the same concentration of each component miRNA (the equimolar pool) and two that contained different concentrations of the component miRNAs, designed such that the relative concentration of individual miRNAs between these two ratiometric pools ranged from 1:10 to 10:1. While these pools were critical for determining the bias, accuracy, and specificity of each platform, we also analyzed exRNA from pooled healthy male plasma, pooled healthy non-pregnant female plasma, and pregnant female plasma samples from two individual donors. These plasma exRNA samples allowed us to compare how the platforms performed on biological samples.

Small RNA-seq is the only platform we studied that does not require hybridization, making it the only truly non-targeted platform. Given that small RNA-seq will capture any small RNA of acceptable (>17 nt and less than ~30 nt used here) length with a 5′ phosphate and 3′ hydroxyl group, it enables measurement of all RNA biotypes, which may be advantageous when studying clinically useful biofluids in which miRNAs comprise a small fraction of the exRNA, such as urine (Godoy et al., 2018). However, the turnaround time for small RNA-seq is slow, requiring at least a week for RNA isolation, cDNA library generation, sequencing, and analysis. Moreover, the reproducibility of small RNA-seq is strongly and negatively affected when different RNA isolation and library preparation protocols are used (Giraldez et al., 2018; Srinivasan et al., 2019). Even when the same protocols are used, reproducibility of measurements of plasma miRNAs with signal intensities less than 32 normalized counts was poor, potentially making it difficult to reliably quantify low-abundance miRNAs across samples. Small RNA-seq is also impeded by bias introduced during adapter-ligation due to sequence-dependent T4 RNA ligase bias (Jayaprakash et al., 2011; Hafner et al., 2011). Although less biased than other commercially available small RNA-sequencing kits (Giraldez et al., 2018), our small RNA-seq protocol exhibited higher bias compared to the three other tested platforms. This bias was partially dependent on GC content. Small RNA-seq detected a small number of reads that mapped to miRNAs not present in the synthetic equimolar pool, which may have been caused by errors introduced during PCR or sequencing or caused by contamination. Although most of these detected “not present” miRNAs had very low signal intensities, a few had greater than 500 normalized counts. Although all platforms could detect known differences in the synthetic ratiometric pools, small RNA-seq showed better detection of expected differences between biological samples than the other platforms, with the largest number of placenta-specific miRNAs having significantly higher levels of expression in pregnant female plasma compared to non-pregnant female plasma.

EdgeSeq is a high-throughput semi-automated assay that can process up to 96 samples in 24 hours, excluding the time for sequencing, with the multiplexing limit determined by the number of sequencing barcodes available. Additionally, its ability to use as little as 25 mL of crude plasma as input makes EdgeSeq an attractive platform for clinical use. Software for the alignment of fastq files and generation of miRNA counts is built into the same instrument that processes the samples, excluding the need for extra computational resources. For both crude plasma and isolated RNA technical replicates, EdgeSeq was more reproducible than small RNA-seq. One important advantage of EdgeSeq is its ability to

directly assay miRNAs in small volumes of biofluid. Assays using crude plasma (25 μL) were only slightly less reproducible than those using isolated RNA (extracted from 200 mL plasma). Comparisons of results from crude plasma and isolated RNA suggest that there are no major differences in the ability of EdgeSeq to detect miRNAs from these two starting materials. On the other hand, the correlations between results from crude plasma and isolated RNA were not high enough to allow us to recommend direct comparisons between data derived from these two input types. EdgeSeq did not suffer as strongly from detection bias as small RNA-seq; for the synthetic oligonucleotide pools, most miRNAs were detected at signal intensities near the expected signal. Although sources of bias in EdgeSeq are unknown, very low GC content was strongly associated with low signal intensity. EdgeSeq profiling of the equimolar pool showed high signal for several miRNAs not present in the equimolar pool, in extreme cases reaching signal intensities greater than the maximum signal intensity achieved by miRNAs in the pool. Although there was evidence of some signal from cross-hybridization, sequence similarity did not seem to explain most of the false-positive signals. These signals may be attributable to other factors like incomplete digestion of unbound probes or cross-hybridization of probes in the assay. Placenta-specific miRNAs were detected at increased levels in pregnant female plasma compared to non-pregnant female plasma by EdgeSeq, but fold differences were modest compared to small RNA-seq, likely due to EdgeSeq's lower sensitivity for quantification of low-abundance miRNAs. This is consistent with small RNA-seq showing slightly better performance in relative quantification for the ratiometric pools (Figure 5).

Like EdgeSeq, FirePlex can use as input a small volume (20 mL) of crude plasma. The number of miRNAs that can be analyzed in a single FirePlex assay is smaller than for the other methods, but FirePlex miRNA panels are easily customized, which offers the advantage of only focusing analysis on miRNAs of interest. Processing of samples takes ~1.5 days and requires the use of a flow cytometer for signal detection. Of all tested platforms, FirePlex was the least reproducible for the synthetic samples but had similar reproducibility as small RNA-seq for the plasma samples. FirePlex had low bias, with median signal intensities close to the expected signal. The source of detection bias in FirePlex is unknown and was only slightly affected by GC content. A small number of miRNAs not present in the synthetic equimolar pool were detected at levels comparable to those present in the pool, but because there were no FirePlex probes for sequences within an edit distance of 4 from these false positives, we could not determine whether sequence similarity contributed to false positives. Although we evaluated reproducibility of technical replicates of crude plasma, we did not perform the FirePlex assay using RNA isolated from the same crude plasma sample, so we could not compare the reproducibility of crude plasma versus purified exRNA for this platform. For FirePlex, all of the placenta-specific miRNAs were detected with low signal, and thus, the failure to detect differences may be due to poor reproducibility of measurements near the threshold of detection. At the time of publication, a newer version of FirePlex is now marketed that has increased the dynamic range of the assay, but we did not test the performance of this newer version.

While nCounter does not accept crude plasma as input, it can assay up to 800 miRNAs, making it a higher-throughput assay than FirePlex. The turnaround time for the nCounter assay is ~2 days. Like EdgeSeq, software for fluorescence detection, demultiplexing,

background subtraction, and reporting of signal intensities, is incorporated into the processor. One limitation to our study is that technical replicates were not performed by Nano-String, so reproducibility could not be assessed for nCounter. nCounter had less bias than small RNA-seq but more bias than the other two platforms. GC content did not seem to have any effect on the detection bias, and the detection bias for nCounter was not correlated with any other platform, suggesting unique biases. nCounter involves direct detection and eliminates potential bias associated with amplification. However, our analysis indicates that the other aspects of the assay are subject to bias and that results from nCounter and the other three platforms tested cannot be used to directly infer absolute miRNA quantities. We observed that for all of the plasma exRNA samples, nCounter yielded a much smaller fraction of miRNAs detected above background compared to the other three platforms, and thus, the failure to detect differences between pregnant and non-pregnant serum could be due to lower sensitivity of this platform.

Our study reveals differences in performance that are relevant when selecting a platform for a specific application. Small RNA-seq is well suited for discovery applications, had the best ability to distinguish between miRNAs that were present in or absent from the synthetic pool, and was best able to meet the biological challenge of detecting placenta-associated miRNAs in pregnant female plasma. However, due to throughput and turnaround time considerations, small RNA-seq may not be practical for clinical use. EdgeSeq offers the ability to detect a large number of miRNAs in either isolated RNA samples or crude plasma, had relatively low detection bias, and had modest success in detecting placenta-associated miRNAs in pregnant female plasma. FirePlex, with a smaller number of probes per assay, is best suited for more targeted analyses and like EdgeSeq can be used with crude plasma as well as isolated RNA, but this platform did not offer clear advantages in terms of reproducibility, bias, sensitivity, specificity, or relative quantification compared with the other methods and could not detect placenta-associated miRNAs in pregnant female plasma. For nCounter, we were unable to assess reproducibility, but we found that the absence of an amplification step did not reduce bias below that seen with the other two probe-based platforms, and we were unable to detect placenta-associated miRNAs in pregnant female plasma.

The results from our analysis of placenta-associated miRNAs in pregnant and non-pregnant female plasma are largely consistent with those found in a previous study comparing the number of differentially expressed miRNAs among brain, placenta, and liver using data generated from small RNA-seq, EdgeSeq, FirePlex, and a qPCR assay from QIAGEN (miRnome) (Yeri et al., 2018). In that study, of all miRNAs detected as differentially expressed across tissue types using small RNA-seq, ~80% were also differentially expressed using EdgeSeq and ~70% using FirePlex. In our study, small RNA-seq and EdgeSeq were also the most concordant in identifying biological differences between sets of samples. Specifically, EdgeSeq and small RNAseq were able to identify differential expression of placenta-associated miRNAs in pregnant and non-pregnant female plasma samples, while FirePlex was not. A difference between the two studies is that FirePlex was able to distinguish among the three tissue types in the Yeri et al. study, but was not able to distinguish between pregnant and non-pregnant female plasma in our study. The likely reason for this is that the miRNA profiles differ more widely among the tissue types (which

are composed only of RNA from the cognate tissue) than between pregnant and non-pregnant female plasma (noting that the estimated fraction of exRNA originating from the placenta in pregnant female plasma is only ~3.4%–15.4% in the 2nd–3rd trimesters) (Koh et al., 2014).

With the exception of small RNA-seq, samples were analyzed by the manufacturers of the platforms, and the resulting data were sent to us for analysis. It is important to recognize that different versions of these platforms may have different performance characteristics. For example, we observed that the version of the FirePlex assay used here was superior for relative quantification but was less reproducible than an earlier version tested with some of the same samples (data not shown). The analyses that we report here provide a useful comparison of available platforms for miRNA quantification and highlight limitations that should be considered when developing future technologies. Our results also highlight the fact that platform-specific issues related to reproducibility, bias, sensitivity, and specificity must be taken into account when interpreting the results of extracellular miRNA analyses.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Louise Laurent (llaurent@ucsd.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Biofluid samples were labeled with study identifiers; no personally identifiable information was shared among participating laboratories.

Human plasma from 10 healthy male and 10 healthy female donors 21–45 years of age were collected, processed, and combined to create a male pool and a non-pregnant female pool (Human Non-Pregnant Female Plasma Pool #2) by the laboratory of Dr. Ionita Ghiran at Beth Israel Deaconess Medical Center (BIDMC). The BIDMC IRB approved the protocol to consent participants and collect samples (Giraldez et al., 2018). Blood was collected from a peripheral vein using a 19 g butterfly needle with K_2EDTA as the anticoagulant at room temperature and centrifuged at 500 *xg* for 10 min (Giraldez et al., 2018). The supernatant was removed and re-centrifuged at 2,500 xg for 10 minutes. The plasma was divided into 1 mL aliquots and stored at −80°C until exRNA isolation was performed. The pooled male plasma from BIDMC was used to assess reproducibility for small RNA-seq, EdgeSeq, and FirePlex (Table S1).

Pregnant female plasma from two healthy pregnant donors (36–37 years of age) during the 2nd and 3rd trimester and non-pregnant female plasma from 10 healthy non-pregnant female donors (22–25 years of age) were collected under an IRB protocol approved by the Human Research Protections Programs at the University of California San Diego (UCSD). Blood was collected from a peripheral vein using a 19 g butterfly needle with K_2EDTA as the anticoagulant at room temperature and centrifuged at 2,000 *xg* for 20 min to remove cells and cell debris. Pregnant female plasma was divided into 1 mL aliquots. For the non-

pregnant female plasma pool (Human Non-Pregnant Female Plasma Pool #1), plasma samples were pooled using equal volumes from each donor and divided into 1.5 mL aliquots. All aliquots were stored at −80°C until exRNA isolation was performed.

For the pregnant versus non-pregnant female plasma comparisons, pregnant female plasma samples from UCSD were used to quantify expression of placenta-associated miRNAs on all four platforms, pooled non-pregnant female plasma from UCSD (Human Non-Pregnant Female Plasma Pool #1) was analyzed using small RNA-seq, EdgeSeq, and nCounter, and pooled non-pregnant female plasma from BIDMC (Human Non-Pregnant Female Plasma Pool #2) was analyzed using FirePlex (Table S1).

METHOD DETAILS

Synthetic RNA Pools—The three pools of synthetic miRNAs used in this study were distributed to members of the NIH Extracellular RNA Communication Consortium (ERCC) and have been described previously (Giraldez et al., 2018; Table S7). The equimolar pool was produced by combining 962 human and non-human synthetic miRNAs from the miRXplore Universal Reference (Miltenyi Biotech) with a custom set of 190 other RNA oligonucleotides at equal molar concentrations (total of 1,152 RNA oligonucleotides, including 759 that correspond to known human miRNA sequences). The ratiometric pools contain 334 synthetic miRNAs, 286 of which correspond to human miRNA sequences, and were generated by IDT.

RNA Isolation—Total RNA from the pool of healthy human male plasma was isolated as follows (Giraldez et al., 2018). 6 mL of QIAzol Lysis Reagent was added to 1.2 mL of plasma. After vortexing and incubating for 5 minutes at room temperature, 200 μL of chloroform was added, followed by vigorous shaking for 15 s. Samples were incubated for 3 minutes at room temperature and centrifuged for 15 minutes at 12,000 x g at 4°C. The upper aqueous phase was transferred to a new tube where 1.5 volumes of 100% ethanol was added. 700 μL of the mixture was added to an assembled RNeasy MinElute spin column and centrifuged for 15 s at 1,000xg at room temperature. This step was repeated until the rest of the sample had been loaded. The spin column was washed and centrifuged three times: the first wash was with 700 mL Buffer RWT and centrifuged for 15 s at 8,000 x g at room temperature, second with 500 μL Buffer RPE and centrifuged for 15 s at 8,000 x g at room temperature, and third wash was with 500 μL of fresh 80% ethanol and centrifuged for 2 minutes at ≥ 8,000 x g at room temperature. The lid of the spin column was opened and spun at full speed spin for 5 minutes at room temperature to remove residual ethanol. RNA was extracted from the column by applying 30 μL of RNase-free water directly to the column and centrifuging for 1 minute at 100 x g and for another minute at full speed. 5 μL was aliquoted into 1.5 μL and frozen at −80°C.

Total RNA from the pools of non-pregnant female plasma (Human Non-Pregnant Female Plasma Pool #1) and the two pregnant female plasma samples (for placenta-associated miRNA expression analysis) was isolated as follows: Frozen plasma samples were thawed on ice, centrifuged at 2,000 x *g* for 5 minutes at 4°C, and the QIAGEN miRNEasy Micro Kit was used to isolate RNA from 200 mL of plasma according to the manufacturer's

protocol except that 1 μL of Qiazol and 180 μL of chloroform were used (Godoy et al., 2018).

Small RNA-Sequencing—For small RNA-sequencing of the synthetic equimolar and ratiometric pools, a total molar concentration of 10 femtomoles were used as the starting input. For small RNA-sequencing of the healthy human male plasma, 2.1 μL of isolated RNA was used as input, corresponding to ~84 μL of plasma. For small RNA-sequencing of the pool of non-pregnant female plasma (Human Non-Pregnant Female Plasma Pool #1) and the two pregnant female plasma samples (for placenta-associated miRNA expression analysis), 5 mL of isolated RNA was used as input, corresponding to 200 μL of plasma. RNA from the pool of healthy human male plasma was run in quadruplicate; RNA from all female plasma samples was run in duplicate.

Libraries were sequenced using a modified version of the TruSeq Small RNA Library Prep (previously described as 4N protocol D in Giraldez et al., 2018). To ligate the 3’−4N adaptor, 1 mL of4 mM 3’−4N adaptor and 5 μL of RNA were added to each tube containing dried 15% PEG (prepared beforehand) and incubated at 70°C for 2 minutes. Samples were immediately placed on ice. To each sample, 2 μL of Ligation Buffer (HML), 1 μL of RNase Inhibitor, and 1 mL of T4 RNA Ligase 2, Deletion Mutant were added and incubated at 28°C for 1 hour. To remove residual 3’−4N adaptor, 1 μL of1 μg/μL E. Coli SSBP was added and incubated on ice for 10 minutes. Then, 1 mL of 5’ deadenylase was added and incubated at 30°C for 45 minutes. Lastly, 1 μL of RecJF was added and incubated at 37°C for 45 minutes. To ligate the 5’−4N adapters, 1 μL of 10 μM 5’−4N adaptor was added to a separate tube and incubated at 70°C for 2 minutes. To this tube, 1 μL of 10 μM ATP and 1 mL of T4 RNA Ligase was added. All 3 μL were transferred to the 3’−4N adaptor-ligated RNA from the previous step. Mixture was incubated at 28°C for 1 hour and placed on ice for 2 minutes. Samples were concentrated down from 14 μL to6 μL in preparation for reverse transcription and amplification. 1 μL of RNA RT primer was added to each tube and incubated at 70°C for 2 minutes and on ice for 2 minutes. 2 μL of 5X First Strand Buffer, 0.5 μL of 12.5 μM dNTP Mix, 1 mL of 100μM DTT, 1 mL of RNase Inhibitor, and 1 μL of SuperScript II were added to each sample and incubated at 50°C for 1 hour. For PCR amplification of the library, 8.5 μL of ultra pure water, 25 μL of PCR Mix, and 2 μL of RNA PCR primer were added to each tube. 2 μL of a unique PCR Primer Index is added to each tube for barcoding. PCR thermocycler program was as follows: 98°C for 30 s, 20 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 15 s, followed by 72°C for 10 minutes, and an infinite hold at 4°C. Libraries were run on the Bioanalyzer High Sensitivity DNA Chip for quality control (expected peak at 152 bp) and size-selected using the Pippin Prep per manufacturer’s recommendation.

Synthetic equimolar pools were sequenced to a median read depth of 12,569,524 for total reads and 11,230,216 for miRNA reads, 11,130,839 and 5,987,818 for synthetic ratiometric pools, 23,262,217 and 2,367,895 for human male plasma, 17,823,762 and 2,072,402 for human female plasma, and 11,901,781 and 3,314,000 for pregnant female plasma samples. The libraries corresponding to the synthetic pools and the pool of healthy human male plasma were sent to the laboratory of Muneesh Tewari at the University of Michigan, Ann Arbor for sequencing on an Illumina HiSeq 4000. For all other plasma samples, sequencing

was done at the Center for Advanced Technology at the University of California, San Francisco on a HiSeq 2500. Fastq files were shared and aligned via the extracellular RNA processing toolkit (exceRpt) pipeline (Rozowsky et al., 2019). Sequences smaller than 17 nt were removed prior to alignment. No mismatches were allowed when aligning to the synthetic pools; 1 mismatch was allowed for the human plasma libraries. Small RNA-seq results for the synthetic pools and pools of healthy human male plasma have been previously reported (Giraldez et al., 2018).

HTG EdgeSeq Assay—We shipped 40 nM stocks of the synthetic equimolar and synthetic ratiometric pool to HTG Molecular that were then diluted to 1 pM in HTG Lysis Buffer just prior to processing. 25 μL of the dilution was added to each well for a total molar concentration of 0.025 femtomoles. Crude plasma from the pool of healthy human male plasma, human non-pregnant female plasma pool #1, and two pregnant female plasma samples was prepared by adding 40 μL of sample to 40 mL of HTG Plasma Lysis Buffer followed by addition of 8 μL of Proteinase K, and incubated for 180 minutes. 25 μL were used as the final volume for each replicate, corresponding to 13 μL of plasma. For RNA from the pool of healthy human male plasma, 5 mL of RNA (corresponding to ~200 μL of plasma) isolated as described in Godoy et al. (2018) were added to 20 μL of HTG Lysis Buffer per reaction. For all reactions, the final volume was 25 μL. All samples were run on the HTG EdgeSeq Processor using the HTG EdgeSeq miRNA WT assay, which included 2,102 total probes, 2,083 of which were designed to recognize human miRNAs. Three technical replicates were run for each of the three synthetic samples, the pool of human male plasma, and RNA isolated from the human male plasma. Nine technical replicates of pregnant female plasma and non-pregnant female plasma were run. Barcodes and adapters were added to processed samples using 16 cycles of PCR. Libraries were sequenced on an Illumina MiSeq.

Abcam FirePlex Assay—We shipped RNA isolated from the human male plasma and crude pregnant female plasma from 2 donors, along with the synthetic pools and pools of human male and female plasma (Human Non-Pregnant Female Plasma Pool #2) to Abcam for processing with the FirePlex assay. This assay included 131 total probes, all of which were designed to recognize human miRNAs. miRNAs were included in the panel based on their presence in the synthetic pools and plasma small RNA-sequencing libraries and covered a broad range of expression. 5 of the placenta-specific miRNAs were also included in the panel. Two separate panels were run in order to accommodate all 131 total probes, as FirePlex is limited to 68 fluorescent barcodes per panel. The first panel contained 66 miRNA probes, the second contained 65. 6 miRNAs were present in both panels; for these miRNAs only the signal intensity for the miRNA in the first panel was kept. For each panel, 0.57 femtomoles of the synthetic equimolar pool was added to each well. For the ratiometric pool, 0.33 femtomoles were added. 20 μL of crude plasma was used for each replicate. All samples were run in triplicate.

NanoString nCounter Assay—The nCounter assay included 828 total probes, 798 of which were designed to recognize human miRNAs. The target concentration for each miRNA in the synthetic pools was 10 attamoles. Therefore, for the equimolar synthetic pool,

which contained ~1200 RNA sequences, we sent the manufacturer 3 picomoles total at a concentration of 12 femtomoles/mL. For each ratiometric pool, which contained 334 RNA sequences, 1.5 picomoles total at a concentration of 3.4 femtomoles/mL were sent. For the exRNA samples isolated from pregnant female and non-pregnant female plasma using the RNA isolation method described in Godoy et al. (2018), we sent the equivalent of the amount contained in 20 μL of plasma, as recommended by the manufacturer. The manufacturer (Nano-string) analyzed the samples on the nCounter Human v3 miRNA panel.

Detectable miRNA Criteria—For small RNA-sequencing, we mapped reads to the complete set of synthetic RNA oligonucleotides in each pool. The equimolar pool contained 164 synthetic RNAs that were outside of the miRNA size range (< 18 nt or > 30 nt) or were not 5’phosphorylated and these were excluded from the analysis. Only 759 of the 988 detectable miRNAs were human miRNAs; however, non-human miRNAs are still considered detectable for small RNA-seq as they are the appropriate length and contain the appropriate end modifications. Non-human miRNAs consisted of rat, mouse, and viral miRNAs. For the other platforms, detectable miRNAs were defined as those which were targeted by a probe in the assay.

QUANTIFICATION AND STATISTICAL ANALYSIS

Signal Intensity—For small RNA-seq and EdgeSeq, signal intensities were determined from the numbers of mapped reads. For FirePlex and nCounter, the manufacturer adjusted the signal intensities by subtracting background fluorescence from raw signal intensities. For all analyses, signal intensities were quantile normalized using the normalizeQuantile function in the R library limma (Ritchie et al., 2015).

Analysis of Reproducibility—To assess reproducibility of the synthetic miRNA pools and the human male plasma samples, coefficient of variation (CV) was calculated as a percentage using the cv function in the R library raster (Hijmans, 2019). For analysis of the relationship between CV and signal intensity, miRNAs were then divided into 5 equal-sized groups for the synthetic pools and 4 equal-sized groups for the plasma samples according to median signal intensity across replicates. All plots were produced using ggplot2 (Wickham 2009).

Analysis of Bias—To analyze bias, we calculated bias as the ratio of observed signal to expected signal. For the equimolar pool, the expected signal was defined as the total signal intensity (sum of averaged normalized signal intensities across technical replicates for all detectable miRNAs) divided by the number of detectable miRNAs. For the ratiometric pools, the expected value for any given miRNA was made proportional to the relative concentration of that miRNA (e.g., a miRNA at 10X had a signal intensity 10 times greater than a miRNA at 1X). We used the Spearman method (cor function in R library stats) to correlate GC content and detection bias. To evaluate whether there was a relationship between the 5’ or 3’ nucleotide and normalized signal in the synthetic equimolar pool, we performed Mann-Whitney tests between each combination of nucleotides using the wilcox.test function in the R library stats. P values were corrected using the Bonferroni method using the p.adjust function in the R library stats (R Core Team 2018). Correlation of

bias in the synthetic equimolar pools (observed/expected) across platforms was calculated using the Pearson method.

Analysis of Sensitivity and Specificity—For small RNA-seq, “absent miRNAs” were defined as all miRNAs in miRbase (version 21) that were not present in the synthetic pool. We only considered miRNAs in the synthetic pools that were human. For EdgeSeq, FirePlex, and nCounter, “absent miRNAs” were defined as all miRNAs that had probes but were not present in the synthetic pool.

To generate receiver operator characteristic curves, the true positive rate and false positive rate were calculated using only the signal intensities from the synthetic equimolar pool for each platform. We set each threshold as the expected value for each platform multiplied by all numbers from 0 to 200 with an increment of 0.05. The number of true positives was calculated as the number of miRNAs in the synthetic equimolar pool with a signal intensity greater than or equal to the threshold. Similarly, the number of false negatives was equal to the number of miRNAs with a signal intensity less than the threshold. In the same manner, the number of false positives and true negatives were calculated for miRNAs not in the synthetic equimolar pool. The true positive rate is calculated as the number of true positives divided by the sum of the number of true positives and false negatives. The false positive rate is calculated as the number of false positives divided by the sum of the number of false positives and true negatives. The area under the curve was calculated using the AUC function in the R library DescTools (Signorell et al., 2019).

We calculated the minimum Levenshtein edit distance between each absent miRNA and any miRNA present in the pool using the Levenshtein python package.

Analysis of Relative Quantification—The expected count for each miRNA was calculated by dividing the sum of all miRNA signal intensities by the relative number of detectable miRNAs. For example, in EdgeSeq, there were 242 overlapping miRNAs between the ratiometric pool and the EdgeSeq probe set. However, because miRNAs were present at varying concentrations (e.g., 139 at 1X, 14 at 1.5X, 14 at 2X), the sum of signal intensities was not divided by 242 but by the sum of the number of miRNAs multiplied by the relative concentration (e.g., 139*1 + 14 * 1.5 + 14*2 +...). Once the expected count was calculated, it was adjusted for each miRNA based on its relative concentration.

The expected ratio for each miRNA was calculated as the concentration of a miRNA in pool A divided by the concentration in pool B, unless the concentration in B was higher than in A, in which case the expected ratio was the inverse. The observed ratio for each miRNA was calculated from quantile normalized intensities.

The best fit line was calculated using the lm function in the R library stats, and the correlation coefficient and p value were calculated using the Pearson method using the cor function in the R library stats.

Analysis of Spike-In Signal Intensities—2.5×10^{-18} moles of 58 RNA oligonucleotides ranging from 16 – 70 nucleotides with sequences that do not align to

human were spiked-in to RNA isolated from human plasma and water prior to library generation. Libraries were generated using the 4N protocol B as previously described (Giraldez et al., 2018) with the following exceptions: the 5° adaptor was substituted with the adaptor from the 4N C protocol, and PippinHT gels were run using marker 30G and the range of 134–162bp was collected. Sequencing was done on an Illumina NextSeq 500. Fastq files were processed as described above and were additionally aligned to the sequences corresponding to the spike-ins. For nCounter, 9 sequences are included in the assay as internal positive controls. These were considered as spike-ins. Raw signal intensities from spike-ins were quantile normalized, averaged, and $\log_2$-transformed. For small RNA-sequencing, signal intensities from the water only sample were compared to RNA isolated from human male plasma. For nCounter, signal intensities from the synthetic equimolar sample were compared to RNA isolated from human male plasma. No water only control was included in the nCounter assay due to limited space. Correlation coefficients were calculated using the Pearson method.

Placenta-Associated miRNA Expression in Plasma—The 50 miRNAs arising from chr19:53666679–53788773 (hg19) were considered to be placenta-associated miRNAs (Ouyang et al., 2014). Technical and biological replicates were quantile normalized and $\log_2$-transformed signal intensities from pregnant female and non-pregnant female samples were compared for all of the 50 miRNAs that were detected in each platform. Since 9 technical replicates were run for the non-pregnant female plasma and pregnant female plasma, we randomly chose 3 replicates for this analysis. For small RNA-seq, EdgeSeq, and nCounter, the non-pregnant female comparator was the pool of healthy non-pregnant female plasma collected at UCSD. For Fireplex, the comparator was the pool of healthy non-pregnant female plasma collected at BIDMC. In a previous run of FirePlex, both pools were run simultaneously and had high correlation coefficients ($R > 0.92$), suggesting similar miRNA expression levels.

DATA AND CODE AVAILABILITY

All normalized signal intensities for all samples across all assays are provided in the supplement. The accession number for data generated from small RNA-sequencing of the maternal and non-maternal plasma samples reported in this paper is dbGaP: phs001892.v1.p1 (https://www.ncbi.nlm.nih.gov/gap/). Data generated from small RNA-sequencing of the synthetic RNA pools and the pool of healthy human male plasma were deposited by Giraldez et al. (2018) and are available on the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) under accession numbers GEO: GSE94584 (equimolar), GSE94585 (ratiometric pool A and pool B), and GSE94582 (human male plasma). Samples corresponding to this study have the following prefix: 4N_D.Lab1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This publication is part of the NIH Extracellular RNA Communication Consortium paper package and was supported by the NIH Common Fund's exRNA Communication Program. We would like to thank Ionita Ghiran for the pools of healthy human male and female plasma and Aileen Fernando, Fabian Flores, and the UCSD Clinical and Translational Research Institute for assisting with collection of the pregnant female and non-pregnant female samples used in this study. We would like to thank Prescott Woodruff, who is a co-principal investigator on the U01HL126493 grant supporting this study, and William Thistlethwaite for help with data deposition. This work was supported by the NIH Extracellular RNA Communication Program (grant U01HL126493 to P.G.W. and D.J.E., grants UH3TR000906 and U01HL126494 to L.C.L., grant UH3TR000901 to S.D., and grant U01HL126496 to D.G.).

REFERENCES

Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297. [PubMed: 14744438]

Chapin SC, Appleyard DC, Pregibon DC, and Doyle PS (2011). Rapid microRNA profiling on encoded gel microparticles. Angew. Chem. Int. Ed. Engl. 50, 2289–2293. [PubMed: 21351338]

Das S, Ansel KM, Bitzer M, Breakefield XO, Charest A, Galas DJ, Gerstein MB, Gupta M, Milosavljevic A, McManus MT, et al.; Extracellular RNA Communication Consortium (2019). The Extracellular RNA Communication Consortium: establishing foundational knowledge and technologies for extracellular RNA research. Cell 177, 231–242. [PubMed: 30951667]

Denaro M, Ugolini C, Poma AM, Borrelli N, Materazzi G, Piaggi P, Chiarugi M, Miccoli P, Vitti P, and Basolo F (2017). Differences in miRNA expression profiles between wild-type and mutated NIFTPs. Endocr. Relat. Cancer 24, 543–553. [PubMed: 28830935]

Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T, et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat. Biotechnol. 26, 317–325. [PubMed: 18278033]

Giraldez MD, Spengler RM, Etheridge A, Godoy PM, Barczak AJ, Srinivasan S, De Hoff PL, Tanriverdi K, Courtright A, Lu S, et al. (2018). Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. Nat. Biotechnol. 36, 746–757. [PubMed: 30010675]

Girard L, Rodriguez-Canales J, Behrens C, Thompson DM, Botros IW, Tang H, Xie Y, Rekhtman N, Travis WD, Wistuba II, et al. (2016). An expression signature as an aid to the histologic classification of non-small cell lung cancer. Clin. Cancer Res. 22, 4880–4889. [PubMed: 27354471]

Godoy PM, Bhakta NR, Barczak AJ, Cakmak H, Fisher S, MacKenzie TC, Patel T, Price RW, Smith JF, Woodruff PG, and Erle DJ (2018). Large differences in small RNA composition between human biofluids. Cell Rep. 25, 1346–1358. [PubMed: 30380423]

Hafner M, Renwick N, Brown M, Mihailovic A, Holoch D, Lin C, Pena JT, Nusbaum JD, Morozov P, Ludwig J, et al. (2011). RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. RNA 17, 1697–1712. [PubMed: 21775473]

Hijmans RJ (2019). raster: geographic data analysis and modeling. R package version 28–19. https://cran.r-project.org/web/packages/raster/index.html.

Jayaprakash AD, Jabado O, Brown BD, and Sachidanandam R (2011). Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res. 39, e141. [PubMed: 21890899]

Koh W, Pan W, Gawad C, Fan HC, Kerchner GA, Wyss-Coray T, Blumenfeld YJ, El-Sayed YY, and Quake SR (2014). Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. Proc. Natl. Acad. Sci. USA 111, 7361–7366. [PubMed: 24799715]

Mestdagh P, Hartmann N, Baeriswyl L, Andreasen D, Bernard N, Chen C, Cheo D, D'Andrade P, DeMayo M, Dennis L, et al. (2014). Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. Nat. Methods 11, 809–815. [PubMed: 24973947]

Ouyang Y, Mouillet JF, Coyne CB, and Sadovsky Y (2014). Review: placenta-specific microRNAs in exosomes - good things come in nano-packages. Placenta 35 (Suppl), S69–S73. [PubMed: 24280233]

R Core Team (2018). R: A language and environment for statistical computing (R Foundation for Statistical Computing). https://www.R-project.org/.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47. [PubMed: 25605792]

Rozowsky J, Kitchen RR, Park JJ, Galeev TR, Diao J, Warrell J, Thistlethwaite W, Subramanian SL, Milosavljevic A, and Gerstein M (2019). exceRpt: a comprehensive analytic platform for extracellular RNA profiling. Cell Syst. 8, 352–357.e3. [PubMed: 30956140]

Signorell A, Aho K, Alfons A, Anderegg N, Aragon T, Arppe A, Baddeley A, Barton K, Bolker B, Borchers HW, et al. (2019). DescTools: Tools for descriptive statistics. R package version 0.99.28.

Sohel MH (2016). Extracellular/circulating microRNAs: release mechanisms, functions and challenges. Achiev. Life Sci. 10, 175–186.

Srinivasan S, Yeri A, Cheah PS, Chung A, Danielson K, De Hoff P, Filant J, Laurent CD, Laurent LD, Magee R, et al. (2019). Small RNA Sequencing across diverse biofluids identifies optimal methods for exRNA isolation. Cell 177, 446–462.e16. [PubMed: 30951671]

Wickham H (2009). ggplot2: Elegant Graphics for Data Analysis (Springer Science & Business Media).

Wu C, Carta R, and Zhang L (2005). Sequence dependence of cross-hybridization on short oligo microarrays. Nucleic Acids Res. 33, e84. [PubMed: 15914663]

Yeri A, Courtright A, Danielson K, Hutchins E, Alsop E, Carlson E, Hsieh M, Ziegler O, Das A, Shah RV, et al. (2018). Evaluation of commercially available small RNASeq library preparation kits using low input RNA. BMC Genomics 19, 331. [PubMed: 29728066]

Highlights

- Four miRNA profiling methods are compared on synthetic and biological samples
- Small RNA-seq is the most accurate, sensitive, and specific
- EdgeSeq is the most reproducible and has the least detection bias
- nCounter is less sensitive than small RNA-seq, EdgeSeq, and FirePlex

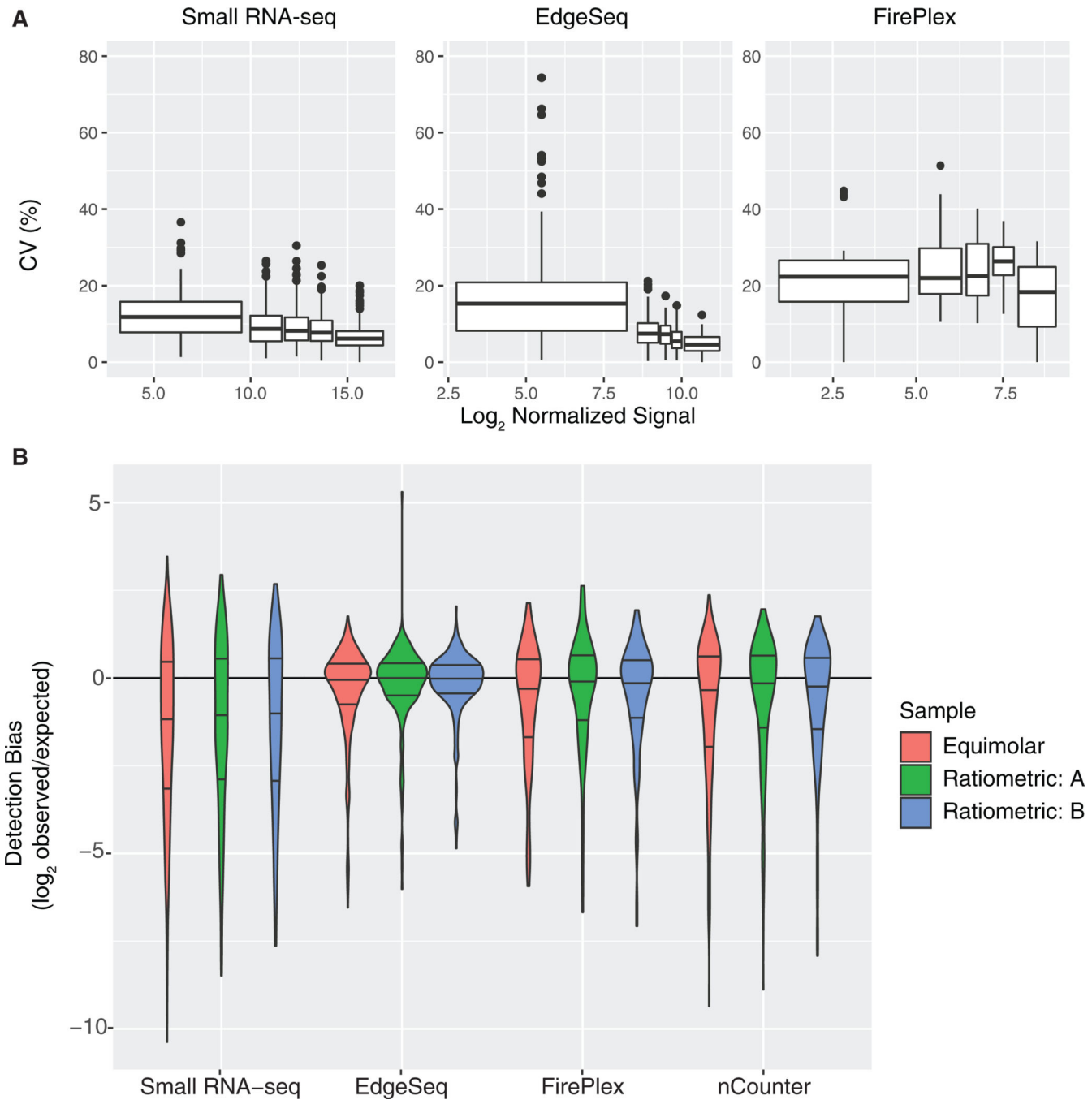


Figure 1. Reproducibility and Detection Bias Determined Using a Synthetic Equimolar Pool

(A) Coefficient of variation for technical replicates (small RNA-seq N = 4, EdgeSeq N = 3, and FirePlex N = 3) expressed as a percentage as a function of median signal. Each boxplot represents ~20% of the total number of detectable miRNAs, grouped by ascending expression (for total number of detectable miRNAs: small RNA-seq N = 988, EdgeSeq N = 451, and FirePlex N = 103). Boxes represent median and interquartile ranges, whiskers represent 1.5 times the interquartile range, and dots represent outliers.

(B) The observed signal to expected signal (detection bias) is plotted for each platform for the equimolar sample (red), the ratiometric A pool (green), and the ratiometric B pool (blue). The width of the violin plot represents the density of miRNAs. The lines represent the median (middle) and interquartile ranges (top and bottom).
See also Figure S1.

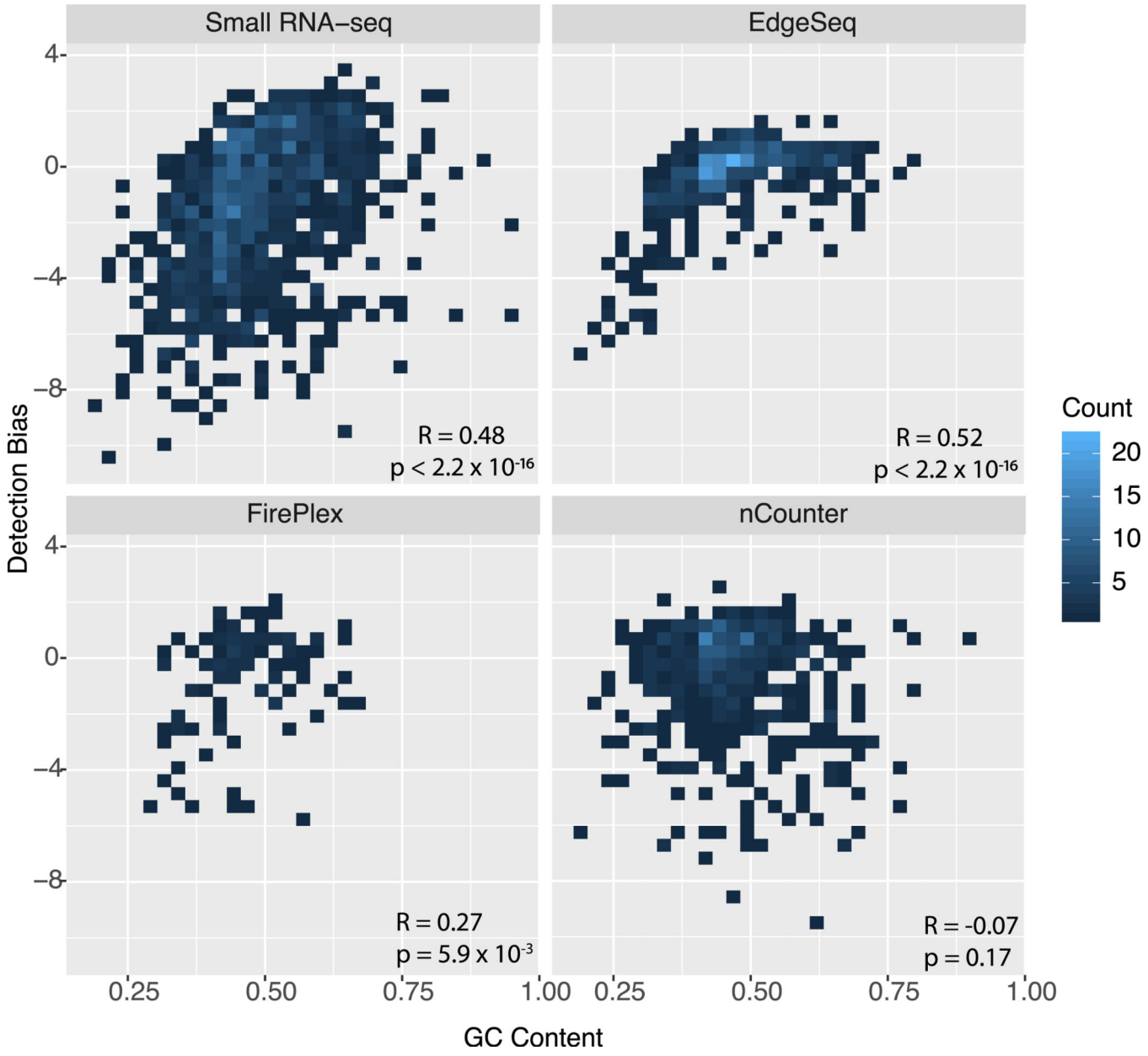


Figure 2. Relationships between Bias and GC Content

Detection bias is plotted as a function of GC content. Lighter blue colors represent a higher density of miRNAs. Correlation coefficients and p values were calculated using the Pearson method.

See also Figure S2.

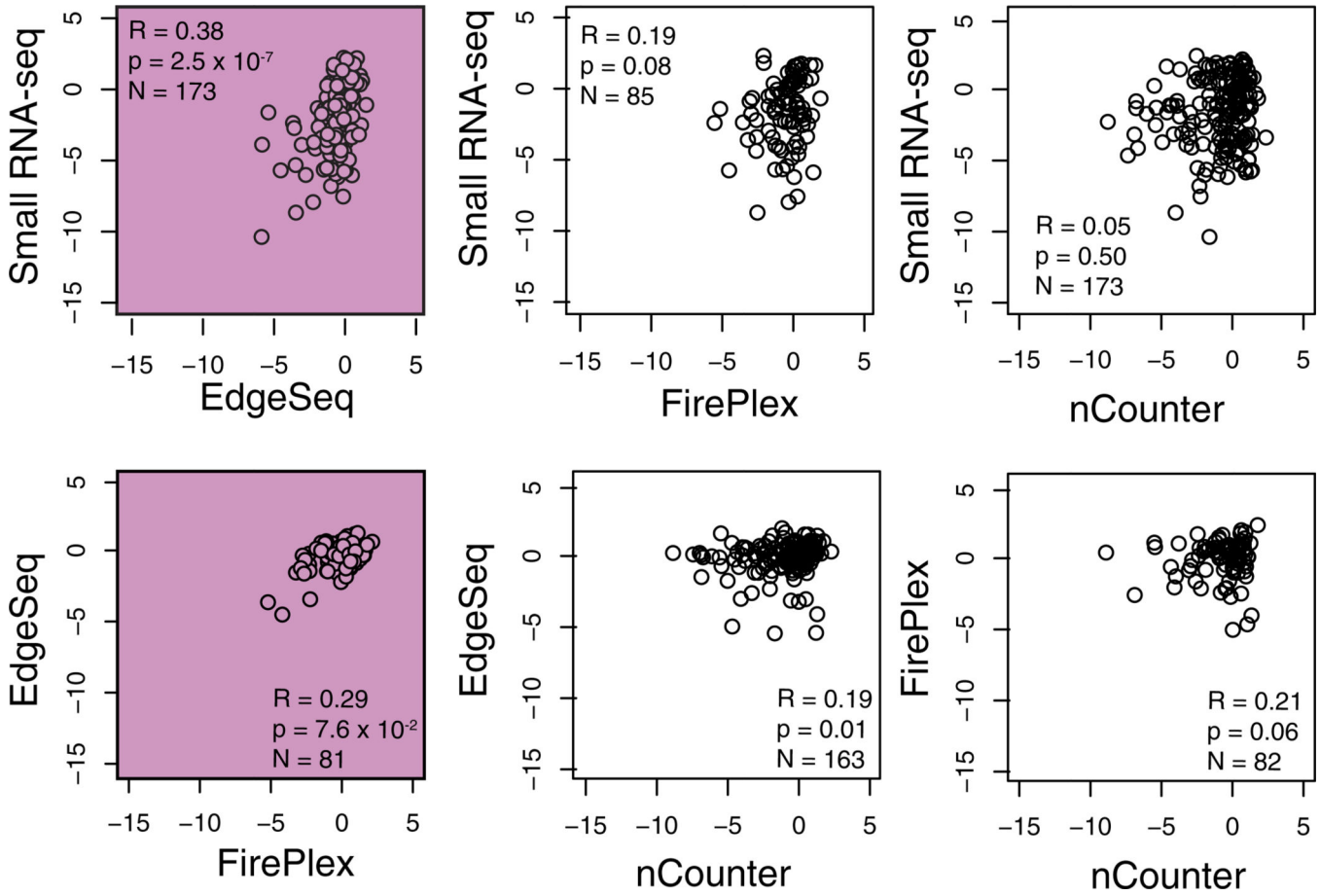


Figure 3. Detection Bias Comparison across Platforms

Each point represents a pairwise comparison of the detection bias for a miRNA between the equimolar pool for the different platforms. Correlation coefficients and p values are calculated using the Pearson method.

See also Figure S3.

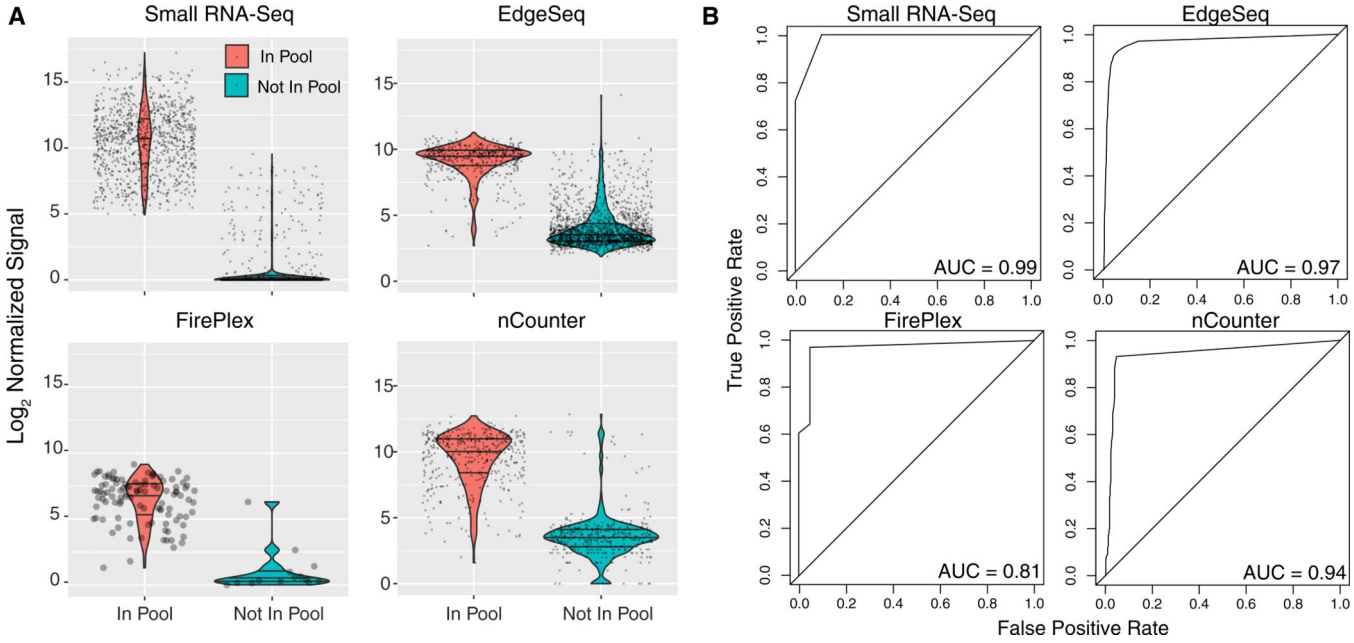


Figure 4. Sensitivity and Specificity of Each Platform as Assessed Using the Synthetic Equimolar Pool

(A) Normalized and $\log_2$-transformed signal is plotted for detectable miRNAs in the synthetic equimolar pool (small RNA-seq n = 758, EdgeSeq n = 451, FirePlex n = 50, and nCounter n = 422) and those not in the synthetic equimolar pool (small RNA-seq n = 2,081, EdgeSeq n = 1,624, FirePlex n = 15, and nCounter n = 376).

(B) Receiver operating characteristic curves.

See also Figure S4.

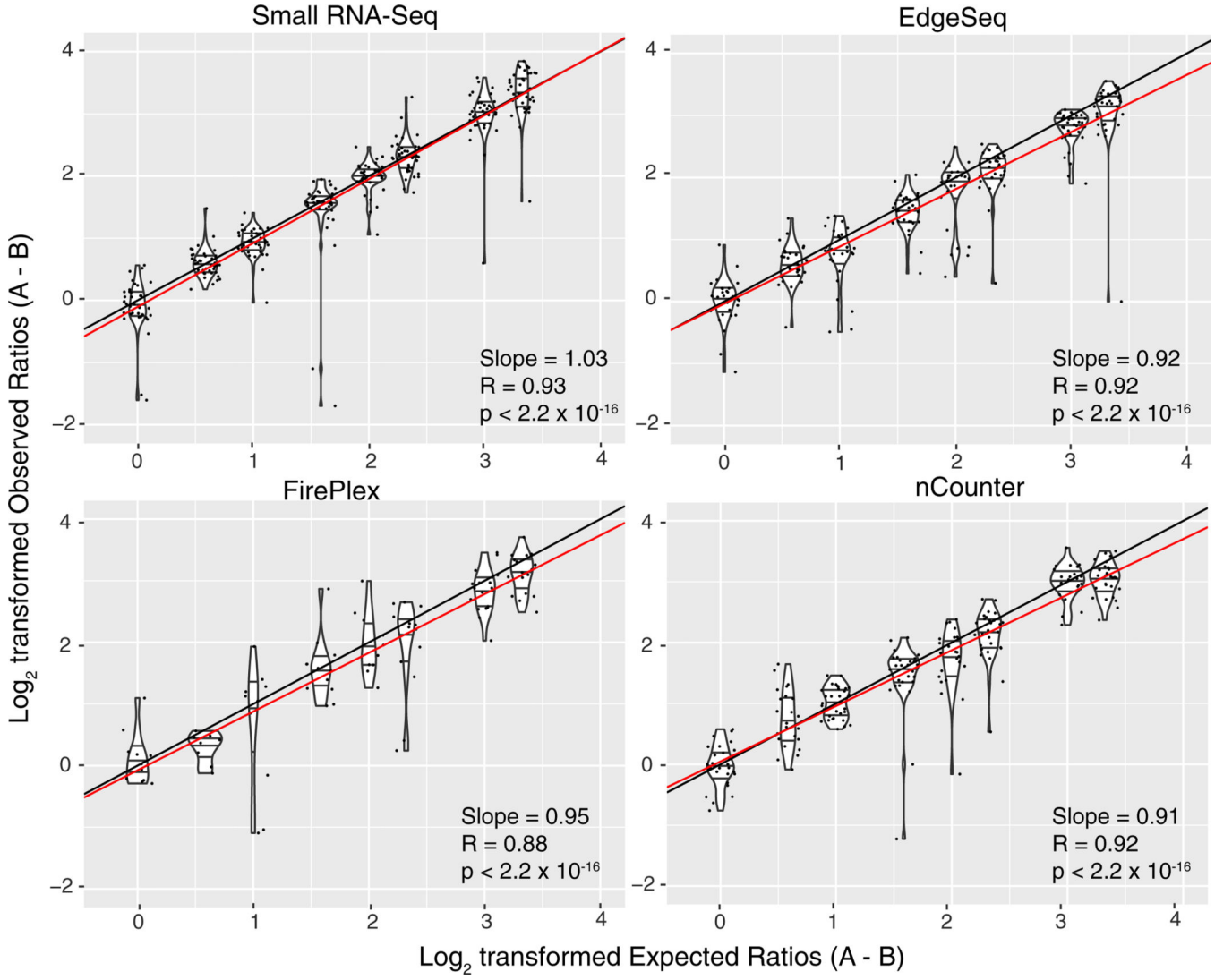


Figure 5. Accuracy of Relative Quantification Determined by Pools of Synthetic miRNAs at Ratiometric Concentrations

The expected fold change of a miRNA in pool A and pool B is plotted against the observed fold change, represented here as a $\log_2$-transformed ratio. Each violin plot represents the distribution of observed ratios for a particular expected ratio. The lines represent the median (middle) and the interquartile range (top and bottom). The dots represent the actual ratios. The black line is the line of identity, and the red line is the best-fit line as calculated using linear regression.

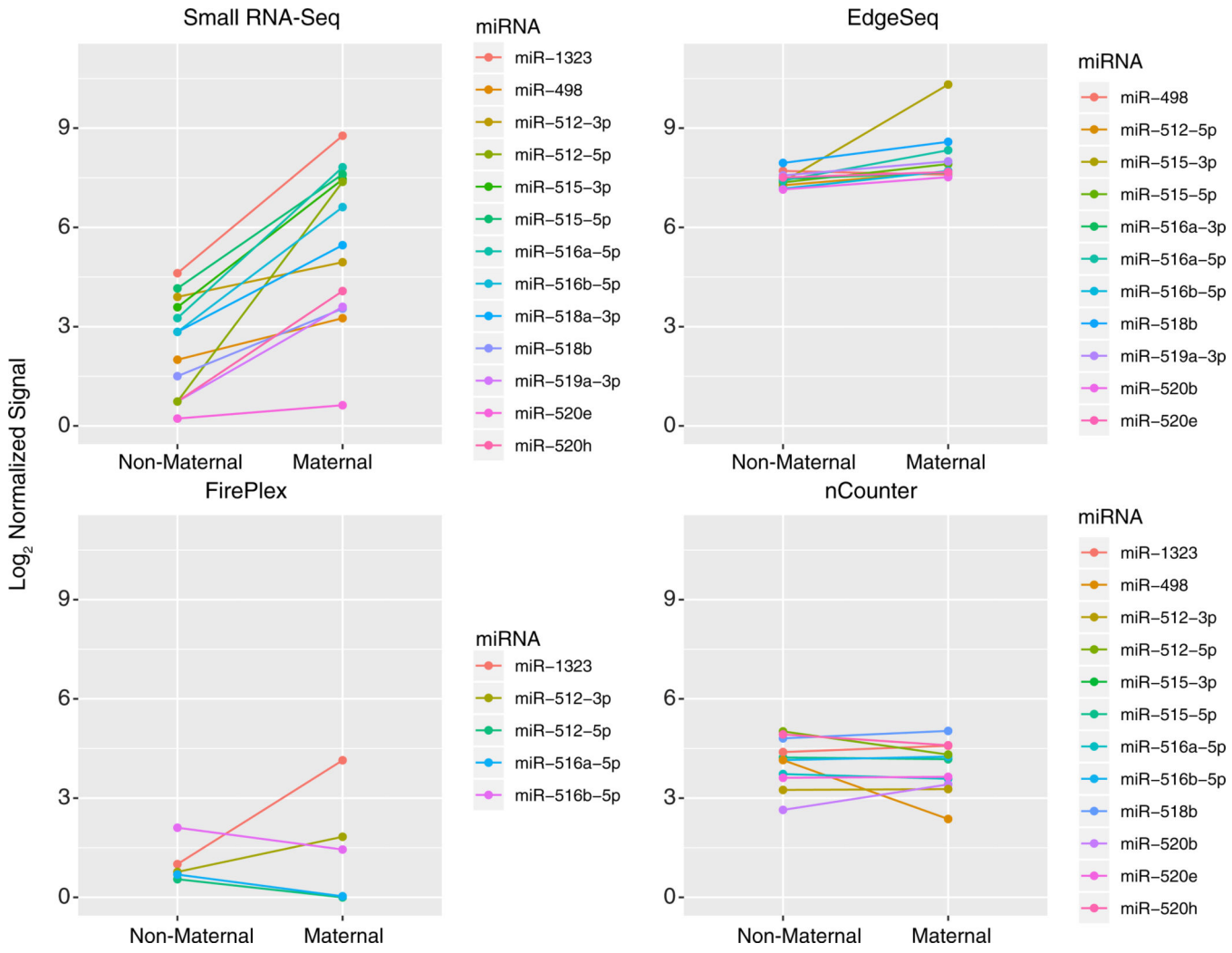


Figure 6. Expression of Placenta-Associated miRNAs in Non-pregnant Female Plasma and Pregnant Female Plasma

Each pair of connected dots represents miRNA signals for non-pregnant female plasma (pooled samples) and pregnant female plasma (mean of two pregnant female plasma samples assayed separately). See also Figures S5 and S7.

Table 1.

Summary of Performance Characteristics

	Small RNA-Seq	EdgeSeq	FirePlex	nCounter
Reproducibility of synthetic equimolar samples (% CV)	8.2	6.9	22.4	NA
Reproducibility of RNA isolated from plasma (% CV)	33.4	14.4	NA	NA
Reproducibility of crude plasma (% CV)	NA	17.8	43.2	NA
Bias (% miRNAs within 2-fold of the median)	31	76	57	47
Specificity and sensitivity (AUC)	0.99	0.97	0.94	0.81
Relative quantification of ratiometric equimolar samples (RMSE)	0.45	0.47	0.58	0.46
Relative quantification of maternal miRNAs (number of differentially expressed placenta- associated miRNAs)	13/13	10/11	2/5	4/12

NA, not applicable.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
HTG EdgeSeq miRNA WTA (1x96)	HTG Molecular Diagnostics, Inc.	NA
FirePlex Custom miRNA Panel	Abcam	NA
nCounter	NanoString	NA
miRNeasy Micro Kit	QIAGEN	Cat#217084
RNeasy Mini Kit	QIAGEN	Cat#74104
TruSeq Small RNA Kit	Illumina	Cat#RS-200-0012
T4 RNA Ligase 2, Deletion Mutant	NEB	Cat#M0242L
E. Coli Single Stranded Binding Protein	Promega	Cat#M3011
5' Deadenylase	NEB	Cat#M0331
RecJF	NEB	Cat#M0264
SuperScript II	Invitrogen	Cat#18064-014
Deposited Data		
Small RNA-sequencing of maternal and non-maternal plasma	dbGaP	phs001892.v1.p1
Small RNA-sequencing of the equimolar pool of synthetic RNA oligonucleotides	GEO	GSE94584
Small RNA-sequencing of the ratiometric pools (A and B) of synthetic RNA oligonucleotides	GEO	GSE94585
Small RNA-sequencing of the pool of healthy human male plasma	GEO	GSE94582
Oligonucleotides		
Equimolar pool of synthetic RNA oligonucleotides	Table S7	Giraldez et al., 2018
Ratiometric pools A and B of synthetic RNA oligonucleotides	Table S7	Giraldez et al., 2018
3'-4N Adaptor: 5'-/5rApp/(N1:25252525)(N1)(N1)(N1)TGGAATTCTCGGGTGCC AAGG/3ddC/-3oo'	IDT	N/A
5'-4N Adaptor: 5'rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrCr (N:2525252525)r(N)r(N)r(N)r(N)-3'	IDT	N/A
Software and Algorithms		
limma	Ritchie et al., 2015	https://bioconductor.org/packages/release/bioc/html/limma.html
raster	Hijmans, 2019	https://cran.r-project.org/web/packages/raster/index.html
ggplot2	Wickham 2009	https://cran.r-project.org/web/packages/ggplot2/index.html
DescTools	Signorell et al., 2019	https://cran.r-project.org/web/packages/DescTools/index.html
extracellular RNA processing toolkit (exceRpt) pipeline	Rozowsky et al., 2019	genboree.org