

RESEARCH ARTICLE

Cross Platform Standardisation of an Experimental Pipeline for Use in the Identification of Dysregulated Human Circulating MiRNAs

Helena Kelly¹, Tim Downing², Nina L. Tuite¹, Terry J. Smith^{3,4}, Michael J. Kerin⁵, Róisín M. Dwyer⁵, Eoin Clancy^{3,4}, Thomas Barry^{1*}, Kate Reddington¹

1 Nucleic Acid Diagnostics Research Laboratory (NADRL), Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland, **2** School of Biotechnology, Dublin City University, Dublin, Ireland, **3** Molecular Diagnostics Research Group (MDRG), School of Natural Sciences, National University of Ireland, Galway, Ireland, **4** Biomedical Diagnostics Institute (BDI) Programme, National University of Ireland, Galway, Ireland, **5** Discipline of Surgery, School of Medicine, National University of Ireland Galway, Galway, Ireland

* Thomas.barry@nuigalway.ie



CrossMark
click for updates

 OPEN ACCESS

Citation: Kelly H, Downing T, Tuite NL, Smith TJ, Kerin MJ, Dwyer RM, et al. (2015) Cross Platform Standardisation of an Experimental Pipeline for Use in the Identification of Dysregulated Human Circulating MiRNAs. PLoS ONE 10(9): e0137389. doi:10.1371/journal.pone.0137389

Editor: Partha Mukhopadhyay, National Institutes of Health, UNITED STATES

Received: July 10, 2015

Accepted: August 17, 2015

Published: September 10, 2015

Copyright: © 2015 Kelly et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data discussed in this publication are deposited in NCBI's Gene Expression Omnibus (Edgar et al 2002) and are accessible through GEO Series accession number GSE72080 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72080>) (Line 331–334).

Funding: EC and TS have been supported during the time this work was carried out by Science Foundation Ireland under Grant No. 10/CE/B1821. MJK and RMD are supported by Breast Cancer Research. NLT, TB, and KR are supported under European Union's Seventh Programme for research,

Abstract

Introduction

Micro RNAs (miRNAs) are a class of highly conserved small non-coding RNAs that play an important part in the post-transcriptional regulation of gene expression. A substantial number of miRNAs have been proposed as biomarkers for diseases. While reverse transcriptase Real-time PCR (RT-qPCR) is considered the gold standard for the evaluation and validation of miRNA biomarkers, small RNA sequencing is now routinely being adopted for the identification of dysregulated miRNAs. However, in many cases where putative miRNA biomarkers are identified using small RNA sequencing, they are not substantiated when RT-qPCR is used for validation. To date, there is a lack of consensus regarding optimal methodologies for miRNA detection, quantification and standardisation when different platform technologies are used.

Materials and Methods

In this study we present an experimental pipeline that takes into consideration sample collection, processing, enrichment, and the subsequent comparative analysis of circulating small ribonucleic acids using small RNA sequencing and RT-qPCR.

Results, Discussion, Conclusions

Initially, a panel of miRNAs dysregulated in circulating blood from breast cancer patients compared to healthy women were identified using small RNA sequencing. MiR-320a was identified as the most dysregulated miRNA between the two female cohorts. Total RNA and enriched small RNA populations (<30 bp) isolated from peripheral blood from the same

technological development and demonstration. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

female cohort samples were then tested for using a miR-320a RT-qPCR assay. When total RNA was analysed with this miR-320a RT-qPCR assay, a 2.3-fold decrease in expression levels was observed between blood samples from healthy controls and breast cancer patients. However, upon enrichment for the small RNA population and subsequent analysis of miR-320a using RT-qPCR, its dysregulation in breast cancer patients was more pronounced with an 8.89-fold decrease in miR-320a expression. We propose that the experimental pipeline outlined could serve as a robust approach for the identification and validation of small RNA biomarkers for disease.

Introduction

Micro RNAs (miRNAs) are a class of highly conserved (18–25 nucleotide) small non-coding RNAs that play an important part in the post-transcriptional regulation of gene expression, making them essential to many fundamental pathological and biological processes [1,2] in all forms of life from animals [3], plants [4], even to some viruses [5]. MiRNAs have an important role in the manifestation of a wide range of diseases from autoimmune disorders to cancer [6–9].

The mode of action of miRNA is complex as multiple miRNAs may function together, either by synergy or in competition to target a single mRNA. Conversely, a single miRNA can moderate many targets [10,11]. It is estimated that miRNAs may be responsible for regulating up to 60% of eukaryotic mRNAs [12–14]. The ability to generate miRNA expression profiles is therefore helpful to better understand these processes.

Significant advances have been made in recent years in the field of miRNA analysis and the understanding of their importance to biological processes. These advances have demonstrated the potential utility of some of these miRNAs as biomarkers for disease. Specific miRNA signatures have been identified in tissues [15], circulating in the blood [16] and in cancers such as lung, ovarian and breast [17–19]. While many studies have identified specific miRNAs as candidate markers for disease, there are still a number of challenges to be overcome in order to demonstrate the clinical utility of these targets as true biomarkers of disease. Currently, there are three main methods employed to determine the global expression levels of miRNAs and to predict their potential as biomarkers for disease. These methodologies are: reverse transcription real-time PCR (RT-qPCR) [20,21], microarray [22–24], and small RNA sequencing [25,26].

While RT-qPCR is considered the gold standard for evaluation and validation of miRNA biomarkers [27,28], because it accurately quantifies low-copy miRNAs [29], small RNA sequencing is rapidly becoming routine for the initial identification of potential miRNA disease biomarkers. However, recent studies have demonstrated a lack of standardisation in the repertoire of small RNAs detected by high-throughput sequencing approaches [30,31]. In many cases where putative biomarkers are identified with small RNA sequencing, they are not substantiated when RT-qPCR is used for the validation of these novel miRNA biomarkers for disease [31–33]. Recent publications have highlighted a need for standardisation of upstream variables such as RNA sample preparation and downstream variables such as accurate quantification of enriched small RNAs [34]. Neither spiked-in RNA controls nor adjustment during bioinformatic processing are sufficient to account for technical artefacts and varying experimental protocols [35]. Moreover, measures of relative but not absolute RNA expression are comparable across platforms, but only if precise procedures are used [36]. These experimental design and normalisation strategies need to be addressed in order to overcome the limitations

observed with small RNA sequencing for novel biomarker identification and subsequent validation of those putative biomarkers using RT-qPCR.

Recently, Redshaw et al., compared methods of small RNA isolation and enrichment, and subsequent quantitative analysis of purified miRNAs using two commercially available RT-qPCR kits. An important observation in this study demonstrated that when using RT-qPCR for determining miRNA absolute copy number in a sample, short RNA enrichment protocols (< 200 bases) can reduce the detected copy number of miRNAs by up to 25% when compared to using total RNA [33,34]. This demonstrates that it is not possible to accurately compare absolute miRNA copy number in enriched RNA populations to the total RNA population.

This observation may also have important implications when cross platform analysis between small RNA sequencing and RT-qPCR are performed. For example, prior to the generation of cDNA for deep sequencing, a gel purification step to enrich for small RNA (< 30) bases is typically performed [37]. This pre-enrichment step facilitates the removal of higher molecular weight RNAs including precursor miRNAs. Consequently, a purification step to enrich for small RNA of <30 bases may also be required for RT-qPCR to relate it to high-throughput sequence data. To our knowledge, there are currently no studies described in the literature evaluating this approach.

In this study we present an experimental pipeline for analysis of small RNA populations for both deep sequencing and RT-qPCR that facilitates reliable cross-platform analysis for the identification and validation of biomarkers associated with disease states. As a proof of concept for this approach, we identified a differentially expressed miRNA (miR-320a) in circulating blood among a cohort of healthy females (n = 23) compared to a cohort of female breast cancer patients (n = 14) based on Illumina sequencing data. We have further validated the miR-320a differential expression levels between these cohort groups using RT-qPCR when the experimental pipeline outlined in this study is adhered to.

Materials and Methods

Ethics Statement

Written informed consent was received from all participants in the study. Ethical approval was granted by the Clinical Research Ethics Committee, Galway University Hospital.

Study Population

Peripheral blood samples were obtained from a total of 55 women, with 36 healthy female volunteers and 19 patients with a diagnosis of breast cancer (Table 1). The healthy blood sample donors were interviewed prior to sample collection to ensure there was no current illness or previous history of breast cancer. Small RNA sequencing was carried out on blood samples from 23 healthy controls and 14 breast cancer patients. These samples, together with a further independent cohort of blood samples from 13 healthy controls and 5 breast cancer patients, were used to validate the small RNA sequencing results using RT-qPCR analysis (n = 55 in total).

RNA Isolation and Purification

All whole blood samples for analysis in this study were collected in PAXgene blood tubes (Qiagen, Hilden, Germany). Total RNA, including small RNAs and miRNAs, were isolated and purified from 2.5ml of whole blood using the PAXgene blood miRNA kit (Qiagen, Hilden, Germany) per the manufacturer's instructions [38]. Total purified RNA integrity was analyzed by capillary electrophoresis using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara,

Table 1. Sample cohort set analysed using small RNA sequencing and RT-qPCR.

Sample Type	Age	Type ^a	Subtype ^a	Small RNA Sequencing	RT-qPCR
Cancer	48	Ductal	Luminal A	✓	✓
Cancer	85	Lobular	Luminal A	✓	✓
Cancer	83	Lobular	Luminal A	✓	✓
Cancer	54	Ductal	Luminal A	✓	✓
Cancer	81	Ductal	Luminal A	✓	✓
Cancer	71	Ductal	Luminal A	✓	✓
Cancer	47	Ductal	Luminal A	✓	✓
Cancer	48	Lobular	Luminal A	✓	✓
Cancer	84	Microinvasive	Luminal B	✓	✓
Cancer	65	Inflammatory	Her2	✓	✓
Cancer	52	Ductal	Luminal A	✓	✓
Cancer	60	Ductal	Luminal A	✓	✓
Cancer	57	Ductal	Luminal A	✓	✓
Cancer	68	Ductal	Luminal A	✓	✓
Cancer	81	Ductal	Luminal A	not performed	✓
Cancer	48	Lobular	Luminal A	not performed	✓
Cancer	64	Ductal	Luminal A	not performed	✓
Cancer	70	Ductal	Luminal A	not performed	✓
Cancer	84	Ductal	Luminal A	not performed	✓
Pre-Menopausal Healthy	32	NA	NA	✓	✓
Pre-Menopausal Healthy	26	NA	NA	✓	✓
Pre-Menopausal Healthy	25	NA	NA	✓	✓
Pre-Menopausal Healthy	27	NA	NA	✓	✓
Pre-Menopausal Healthy	24	NA	NA	✓	✓
Pre-Menopausal Healthy	32	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Pre-Menopausal Healthy	26	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Pre-Menopausal Healthy	26	NA	NA	✓	✓
Pre-Menopausal Healthy	28	NA	NA	✓	✓
Pre-Menopausal Healthy	24	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Pre-Menopausal Healthy	30	NA	NA	✓	✓
Post-Menopausal Healthy	52	NA	NA	✓	✓
Post-Menopausal Healthy	57	NA	NA	✓	✓
Post-Menopausal Healthy	70	NA	NA	✓	✓
Post-Menopausal Healthy	62	NA	NA	✓	✓
Post-Menopausal Healthy	51	NA	NA	✓	✓
Post-Menopausal Healthy	56	NA	NA	✓	✓
Post-Menopausal Healthy	72	NA	NA	✓	✓
Pre-Menopausal Healthy	24	NA	NA	not performed	✓
Pre-Menopausal Healthy	30	NA	NA	not performed	✓
Pre-Menopausal Healthy	35	NA	NA	not performed	✓
Pre-Menopausal Healthy	23	NA	NA	not performed	✓
Pre-Menopausal Healthy	22	NA	NA	not performed	✓

(Continued)

Table 1. (Continued)

Sample Type	Age	Type ^a	Subtype ^a	Small RNA Sequencing	RT-qPCR
Pre-Menopausal Healthy	30	NA	NA	not performed	✓
Pre-Menopausal Healthy	26	NA	NA	not performed	✓
Pre-Menopausal Healthy	30	NA	NA	not performed	✓
Post-Menopausal Healthy	53	NA	NA	not performed	✓
Post-Menopausal Healthy	64	NA	NA	not performed	✓
Post-Menopausal Healthy	54	NA	NA	not performed	✓
Post-Menopausal Healthy	59	NA	NA	not performed	✓
Post-Menopausal Healthy	52	NA	NA	not performed	✓

^a NA = Not applicable

doi:10.1371/journal.pone.0137389.t001

California) and concentration and purity were measured by UV absorbance at 260nm using a NanoDrop 2000 (Thermo Scientific, Waltham, Massachusetts, USA). Isolated total RNA was stored at -80°C prior to use.

Small RNA Sequencing

Small RNA library preparation and sequencing was carried out externally by BaseClear (Leiden, Holland) as described [39]. Briefly, small RNAs of lengths below 30 bases were enriched by polyacrylamide gel size selection purification. Illumina HiSeq adaptors were then ligated to the 5' and 3' ends of the small RNAs. These were amplified for 17 PCR cycles using primers for the adaptors. Post-amplification products were run on an agarose gel and the fragments of 90–110 bp, corresponding to the small RNA plus the adaptors, were isolated [40]. The purified retrieved PCR products were assessed for quality and quantity using the Agilent 2100 bioanalyser (Agilent Technologies). After determining quality and sufficient quantity, according to the manufacturer's instructions, the library was sequenced on an Illumina HiSeq 2000 platform.

Read Mapping and Annotation

The single-end 32 base RNA sequence reads generated were processed using the Illumina Casava pipeline tool (v1.8.3). Quality assessment was based on data passing the Illumina Chastity filtering [41]. Reads containing adaptors or PhiX control signal were removed. Quality assessment was based on the remaining reads using FASTQC v0.10.0 [42] to ensure base quality did not decline significantly across the read length. Small RNA 3' adapter (Illumina TruSeq) sequences and low quality bases were removed from the reads using the CLC Genomics Workbench v5.1.5 [43]. Reads smaller than 18 bases and those longer than 25 bases after trimming were excluded. Unique sequence groups were determined for each sample and these RNA sets were annotated as mature RNAs (IsomiRs) allowing for up to a two base variation in length up- or down-stream and up to one mismatch. The expression level of each small RNA was annotated using either miRBase *Homo sapiens*, or miRBase *Mus musculus*, or Ensembl *Homo sapiens* GRCh37 ncRNA database releases 57 and 67. Thus the data produced was analysed separately as a set of annotated miRNAs and a set of unknown small RNAs.

Small RNA Sequencing Expression Normalisation and Testing

Differential RNA expression was evaluated for transcripts with a minimum of one count per million in at least three samples [44] using three normalisation methods: reads per kilobase per million mapped reads (RPKM) [45], trimmed mean of M-values (TMM) [46] and quantile

normalisation [47]. Firstly, the fraction of reads mapping to small RNA or other small RNA transcripts for each sample was normalised within each sample as RPKM. Reads mapping to multiple RNAs were associated with those RNAs. Secondly, count data for each sample was normalised using quantile normalisation, arguably the optimal non-scaling normalisation approach with expression data aggregated in different sequencing runs [48]. This assumed the overall read count distributions were similar but can account for a large volume of zero or low, as well as high read counts [49]. Thirdly, the TMM value between sample pairs as a normalising factor reduced technical variation signals between small RNAs by compensating for differences in read levels between samples [46]. These were computed using R v2.15.0 Bioconductor package EdgeR v3.2.4 [50], which applied a negative binomial distribution model to identify differentially expressed miRNAs. Reads absent in one group but present in the other group required at least 10 reads in total to avoid sequencing artifacts. Multi-dimensional scaling (MDS) of the RPKM, quantile and TMM normalised levels within and between groups did not identify any samples as individual outliers. The miRNA and small RNAs examined for differential expression in this study had to satisfy all three RPKM, TMM and quantile normalisation methods using Welch's t-test with Benjamini-Hochberg corrected two-tailed p-values < 0.05.

Total RNA Standardisation for RT-qPCR

In order to ensure consistency in RNA from sample to sample, each total RNA sample was normalised to a concentration of 100 ng/μl based on UV absorbance at 260 nm (nanodrop).

Small RNA Enrichment for RT-qPCR

Small molecular weight RNA enrichment (approximately 16 to 25 nucleotide bases) was achieved by a process of size selection using polyacrylamide gels. Fifteen μl of total RNA from each sample, at 100 ng/μl, was applied to a 15% polyacrylamide TBE urea gel. This RNA was electrophoresed alongside a size-specific RNA indicator consisting of RNA oligonucleotides (MWG-Biotech AG, Essenberg, Germany) of between 16 and 23 bases in size. These gels were stained using an ethidium bromide bath and imaged by UV transilluminator. The size markers were utilized to indicate the gel region corresponding to the small RNA of interest present in a sample. These indicated regions were then excised and the RNA retrieved, in an elution volume of 15 μl of nuclease free water (Life technologies, Carlsbad, CA, USA), using a small RNA PAGE recovery kit (Zymo Research, Irvine, CA, USA). This process ensured that the volume was identical both pre- and post-enrichment and that the subsequent RT-qPCR reactions were carried out using the same effective volumes of total RNA and enriched small RNA [33].

In order to assess and control for the efficiency of the small RNA enrichment process an external spike-in normalisation control was required. A miRNA of *Arabidopsis* origin (Aly-miR-159) was chosen for this control and was spiked into all total RNA samples prior to size selection at a final concentration of 1.2×10^7 copies per μl as previously described [33]. This miRNA was chosen as an external spike-in control due the lack of sequence homology to human miRNAs published on miRBase [33,51]. This spike in control has the potential to act as an appropriate control for human miRNA studies when added to a sample under investigation prior to purification of nucleic acid.

MiRNA Reverse Transcription

Reverse transcription reactions were performed using the Taqman miRNA Reverse Transcription Kit (Life Technologies) and MMLV reverse transcriptase (Life Technologies) on RNA, both total and enriched, isolated from each individual clinical sample. The reaction mix for the reverse transcription process was comprised of 1 μl of 10x first strand buffer, 1 μl dNTP

nucleotide mix, 2 μ l specific Taqman miRNA reverse transcription assay, 4.5 μ l nuclease free water and 0.5 μ l of MMLV reverse transcriptase per reaction. Template RNA (1 μ l) was then added. The thermocycling protocol used was as follows: 16°C for 30 min, 42°C for 30 min, 85°C for 5 min. Reverse transcription was carried out on the same volumes of enriched small RNA eluant and the corresponding non-enriched samples (total RNA). Reactions were performed on a thermocycler (PTC-200 MJ Research, Minnesota, USA).

All reactions were carried out in triplicate and appropriate controls were incorporated into each run: namely a no reverse transcriptase negative control (RT NEC) and a no template negative control (RT NTC). The produced cDNA was diluted using water (1:50) and stored at -20°C prior to qPCR analysis.

Standard Curves and Quantitative Real Time-PCR

Absolute quantification was achieved for each specific miRNA by the generation of standard curves. Synthetic oligonucleotide representing the specific miRNA under investigation, both human and Arabidopsis, were synthesized by Integrated DNA technologies (Coralville, Iowa, USA). These were diluted to the appropriate concentration ranges and reverse transcribed as outlined above. All cDNA was diluted 1:50 and stored at -20°C prior to qPCR analysis. To control for run to run variation, internal standard curves were included in each qPCR experiment. A line was fit to the data points obtained for each standard dilution using the quantification cycle (Cq) values reported. This standard curve was then used to absolutely quantify the number of copies of the target miRNA present in the sample by comparing the Cq values observed for each sample to the relevant standard curve.

Quantitative PCR reactions were performed in accordance with the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines. All qPCR were performed on a LightCycler480 using the LightCycler480 probes master kit (Roche Diagnostics, Basel, Switzerland) and Taqman miRNA assays (Cat. No. 4427975, Life Technologies). All Taqman assays were available predesigned and purchased from Life Technologies with the exception of the aly-miR-159 assay which was custom designed and also purchased from Life Technologies. A final volume of 20 μ l was used for each reaction mixture. The optimized master mix contained 2 \times LightCycler 480 probes master (6.4 mM MgCl₂), Taqman miRNA assay (1 μ l), template cDNA (1 μ l), spike-in control (1 μ l) and the volume was adjusted with the addition of nuclease-free dH₂O.

The qPCR thermocycling conditions were as follows: 95°C for 10 min, followed by 50 cycles of 95°C for 15 s, 60°C for 1 min, followed by a cooling step at 40°C for 10 s. The temperature ramp rate on the LightCycler 480 instrument was 4.4°C/s while heating and 2.2°C/s while cooling. The LightCycler 480 Software, Version 1.5 (Roche) was used to calculate the quantification cycle (Cq) value. The qPCR runs were controlled using qPCR no template controls (RT NTC) in addition to the reverse transcription controls (RT NEC).

Small RNA Enrichment Process Normalisation

The median quantification cycles observed in the control assay, specifically 27.4 (Total RNA) and 31.78 (enriched RNA) were used to calculate normalisation factors for the total RNA samples and enriched small RNA samples as previously described [52]. This normalisation factor was then used to control for any potential loss of RNA through the processing of the sample. This normalisation factor was applied to the RT-qPCR data obtained for all assays as outlined. The application of the normalisation factor allowed correction for any differences in recovery of RNA after small RNA enrichment, thus removing the potential for inaccurate quantification due to sample processing methods.

Results

Total RNA Isolation and Purification

All isolated total RNA samples were found to have an RNA integrity number (RIN) of 8 or greater. Based on the RIN numbers and quantities of total RNA obtained (Nanodrop), the extracted and purified RNA samples were determined to be of sufficient quality and quantity to enable processing for small RNA sequencing and RT-qPCR.

Global Analysis of MiRNA from Healthy Controls and Breast Cancer Patients

Sequencing by synthesis amplification of the small RNA profile isolated from peripheral blood yielded 175,267 different miRNAs representing 1,211 known annotated miRNAs [53]. A further 138,457 were unannotated small RNAs and 9,227 were identified as putative tRNA fragments. No global difference was observed between the blood samples from healthy control group and breast cancer patients (S1 Fig) and unannotated small RNAs (S2 Fig). Both groups were dominated by miR-451a, which was the most highly expressed miRNA, accounting for 60% (healthy) and 65% (breast cancer) of the total. The next most frequently observed miRNAs were miR-191 (13% of control and 11% of breast cancer) and then miR-484 (3% and 4%).

Differential Expression of Specific MiRNAs from Small RNA Sequencing Data

To test for differential expression of circulating miRNAs between the healthy control group and breast cancer patients, only miRNAs with adequate RPKM-scaled expression levels were examined to remove bias due to differences in sequencing depths. These miRNAs were normalised using both upper quantile and TMM methods: t-tests were performed such that differential expression was considered valid only if it was consistently different using all three (RPKM, quantile and TMM) methods after multiple test correction [50]. Interrogation of the miRNA results using RPKM adjustment alone would have yielded six differentially expressed miRNAs (miR-320a, miR-140, miR-30d, miR-22, miR-191, and miR-150). Of these potentially differentially expressed miRNAs, three (miR-191, miR-150 and miR-30d) were not supported after TMM testing, and only miR-320a and miR-140 were differentially expressed using upper quantile normalisation as well (Table 2).

These two miRNAs showed lower expression in the blood samples from breast cancer patients relative to the healthy control group: miR-320a had the highest fold-change of 2.30, and miR-140 one of 1.72 (Table 3). Overall, the highest 15 miR-320a expression levels were observed in healthy female samples, while the lowest observable expression levels for miR-320a were associated with samples from breast cancer patients (Fig 1). The miR-320a effect size (Cohen's $d = 1.05$) indicated a high probability (89%) of detecting a true differential expression (Table 3). The largest t-test p-value for miR-320a was 0.00187 (RPKM $t = 3.35$), and this was supported by the upper quantile ($p = 0.00002$) and TMM ($p = 0.00032$) normalised values. For these sequencing data, the vast majority of miR-320a isoforms detected were mature transcripts (99.6%) of which 83.1% were a single sequence (AAAAGCTGGGTTGAGAGGGCTAT). 13.8% of miR-320a sequences were identical to this sequence with an additional A nucleotide at the 3' end. The remainder were rare isoforms.

For miR-140, the largest t-test p was <0.006 (TMM t-test = 2.93), and the effect size ($d = 1.19$) suggested high power (95%) to find changes in expression if present (Table 3). RPKM and upper quantile normalised t-test p-values supported a lower miR-140 expression in the blood samples from breast cancer patients relative to healthy control group ($p < 0.001$ and

Table 2. Differential expression patterns separating breast cancer patient from control samples.

miRNA	RPKM values				Upper quantile normalised				TMM normalised			
	t	p-value	Rank	*	t	p-value	Rank	*	t	p-value	Rank	*
miR-320a	3.35	0.00187	5	*	5.1	0.00002	1	*	3.96	0.00032	1	**
miR-140	3.74	0.00068	3	*	3.35	0.00203	2	*	2.93	0.00585	2	**
let-7b	1.35	0.18649	10		3.22	0.003	3	*	2.78	0.00869	3	**
miR-30d	4.96	0.00001	1	*	2.2	0.03727	6		2.35	0.02555	6	
miR-150	2.95	0.00535	6	*	1.97	0.05679	7		0.52	0.60688	16	
miR-22	3.82	0.00048	2	*	1.71	0.09918	8		2.71	0.01072	4	**
miR-425	1.47	0.15133	9		1.05	0.30512	12		1.42	0.16665	12	
miR-191	3.54	0.00113	4	*	0.85	0.40409	14		1.39	0.17443	13	

Differential expression testing for eight miRNAs in breast cancer patients (15) and controls (26) that showed some evidence of differential expression across quantile, RPKM and TMM normalisation methods. Welch t-tests for differential expression using RPKM values, upper quantile normalised rates and TMM normalised levels all supported lower miR-320a and miR-140 expression in breast cancer patients. The two-tailed p-values were Benjamini-Hochberg corrected to adjust for multiple testing and those with $p < 0.05$ are marked *.

doi:10.1371/journal.pone.0137389.t002

$p = 0.002$ respectively, Fig 1). Most miR-140 isoforms detected were mature transcripts (90.7%): 46.0% were a 23-base sequence (TACCACAGGGTAGAACCACGGAC) that differed by an additional 5' T nucleotide and a substitution of C>A at site 14 compared to the second most abundant isoform (40.1%, ACCACAGGGTAGACCCACGGAA). The remainder were rare isoforms.

Let-7b showed slightly lower expression levels in the blood samples from breast cancer patients relative to the healthy control group—however, the smaller effect size (0.43) indicated that the limited sample size here resulted in low power (25%) to detect true expression differences if present. Further work involving sequencing at a higher read depth to mitigate the low let-7b expression levels should evaluate if the upper quantile ($t = 3.22, p = 0.003$) and TMM ($t = 2.78, p < 0.009$) normalised tests are more accurate reflection of expression state differences than the RPKM results ($t = 1.35, p = 0.187$). The vast majority of let-7b isoforms detected were

Table 3. Micro RNA expression counts and variation between breast cancer patient and control samples.

miRNA	Control count		Breast cancer count		Fold change	Difference	Power
	Mean	SD	Mean	SD			
miR-320a	845	528	368	379	2.3	1.05	0.89
miR-140	3,687	1,407	2,141	1,190	1.72	1.19	0.95
let-7b	606	441	424	401	1.43	0.43	0.25
miR-30d	4,137	1,532	2,255	896	1.83	1.55	0.99
miR-150	2,176	1,092	1,388	617	1.57	0.92	0.79
miR-22	3,070	1,301	1,904	647	1.61	1.2	0.95
miR-425	2,485	892	2,087	807	1.19	0.47	0.29
miR-191	30,283	13,812	16,598	10,679	1.82	1.12	0.92

Expression values for eight miRNAs in breast cancer patients (15) and controls (26) that showed some evidence of differential expression. FC stands for the RPKM fold expression change (control/condition). Difference is Cohen's d: the mean change in miRNA expression between groups scaled by the pooled SD (standard deviation), which reflects the effect size. Power indicates the probability of detecting true differential expression for that miRNA given the difference observed, alpha value (0.05) and sample sizes.

doi:10.1371/journal.pone.0137389.t003

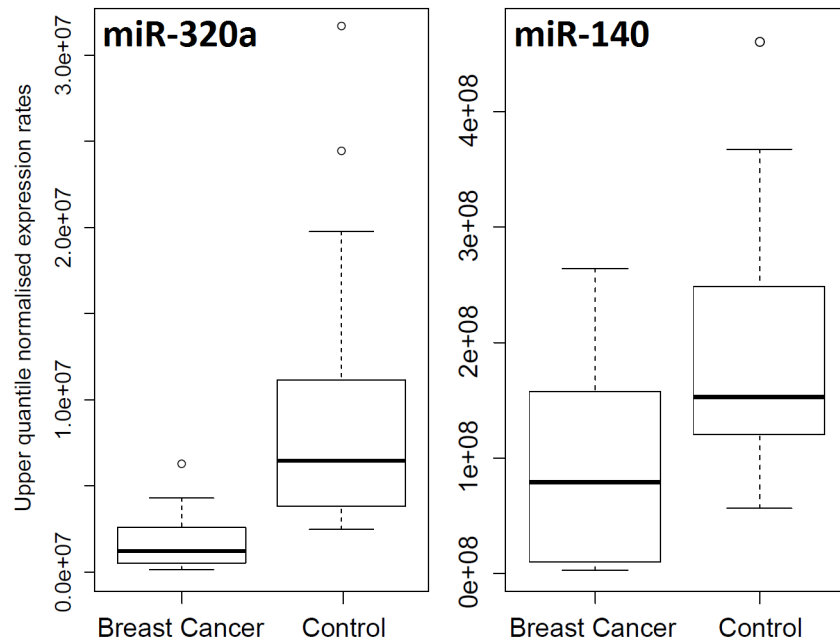


Fig 1. Differential analysis of miRNA between breast cancer and control groups. The distribution of miR-320a and miR-140 upper quantile normalised expression levels between the breast cancer and control groups from Illumina reads. Note that the y-axes are different for each miRNA because of the differing scales.

doi:10.1371/journal.pone.0137389.g001

mature transcripts (95.6%): 73.6% were a single sequence (TGAGGTAGTAGGTTGTGTGAC). All other distinct isoforms were rare (<8%).

Low Expression of Unannotated Small RNAs from Small RNA Sequencing Data

The small RNA profiles consisted of mainly miRNAs—the unannotated small RNAs fraction accounted for just 2.5% (healthy) and 2.6% (breast cancer) of the total. The distribution of expression values was more even than the annotated miRNAome: most highly expressed unannotated small RNA (ENST00000364228) represented 0.66% of the control and 0.93% of the breast cancer unannotated small RNAs. A total of 47 small unannotated RNAs were detected exclusively in the breast cancer samples with a total count level across samples >11. Five had expression levels greater than the mean total counts per gene in the cancer cohort (57.9, [S1 Table](#))—the remaining 42 were rare. Of these five, ENST00000490626 was detected most frequently in the breast cancer cohort (7 out of 14). No small unannotated RNAs were exclusively differentially expressed across the healthy control samples. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [[54](#)] and are accessible through GEO Series accession number GSE72080 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72080>).

MiRNA Target Selection and Small RNA Enrichment

MiR-320a was analysed further using RT-qPCR in an expanded sample set to examine the expression level of miR-320a. Let-7b was also analysed using RT-qPCR because it was previously reported to be differentially expressed between blood of breast cancer patients and healthy controls [[14](#)]. Similarly, miR-195 expression was assessed using RT-qPCR because it was dysregulated in breast cancer patients from the same geographical location as evaluated in

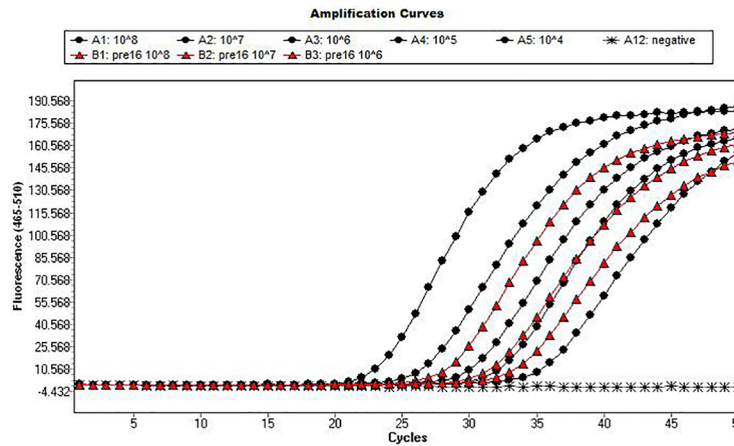


Fig 2. Analysis of precursor and mature miR-16 sequence using real-time PCR. Real-time PCR amplification curves detecting the mature miR16 standard curve 10^8 – 10^4 (circles- decreasing in concentration from left to right) with detection of precursor miR-16 sequence from 10^8 – 10^6 (triangles- decreasing in concentration from left to right) with the no template control, highlighted with diamonds in the FAM channel (465 to 510 nm).

doi:10.1371/journal.pone.0137389.g002

this study [19]. MiR-140 was not evaluated using RT-qPCR because its ranked expression did not partition the groups as clearly as miR-320a. MiR-16 expression levels were analysed with RT-qPCR as a stable endogenous control [18].

Based on the sequence similarity between precursor and mature miRNAs, it was hypothesized that RT-qPCR assays may detect both precursor and mature sequences. The detection of the precursor would directly impact on the quantification of the absolute numbers of mature miRNAs. In order to demonstrate the potential for such an error in quantification through the use of total RNA in the measurement process, we included a synthetic miR-16 precursor sequence as a control template in the miR-16 assay. From this analysis we demonstrated that the miR-16 precursor sequence was also detected by the miRNA assay (Fig 2, S2 Table).

To reduce the potential for non-specific amplification and quantification of precursor sequences, a size selection small RNA enrichment process was carried out on all clinical samples to be analysed using RT-qPCR. This enrichment for RNA of approximately 30 bases or less resulted in a decrease in miRNA copy numbers detected by 4.1 to 11.7-fold ($p < 0.01$, Table 4).

Table 4. Impact of small RNA enrichment on miRNA measurement.

Two-sample T-test	Mean	Stdev	SE Mean	Fold change Total to Small
				Decrease
Total miR-195	6.88E+07	8.34E+07	1.87E+07	
Enriched Small miR-195	5.87E+06	6.00E+06	1.38E+06	11.70
Total miR-16	2.52E+08	2.56E+08	5.72E+07	
Enriched Small miR-16	4.88E+07	4.96E+07	1.14E+07	5.20
Total miR-320a	1.30E+08	1.60E+08	3.58E+07	
Enriched Small miR-320a	3.16E+07	3.80E+07	8.95E+06	4.10
Total Let-7b	1.99E+07	2.66E+07	5.94E+06	
Enriched Small Let-7b	3.64E+06	4.83E+06	1.14E+06	5.50

doi:10.1371/journal.pone.0137389.t004

Standard Curve Generation and RT-qPCR

Absolute quantification was also achieved for each specific miRNA evaluated in this study by the generation of standard curves. Each miRNA target was represented by a standard curve of at least 6 sequential dilutions in order to allow for accurate quantification of the target miRNA in line with current best practices [55]. The standard curve for the miRNA target miR-195 and miR-16 ranged from 10^8 to 10^3 copies. The standard curve for the miRNA target miR-320a and Let-7b ranged from 10^9 to 10^4 copies. The standard curve for the spike-in control miRNA (Aly miR-159) ranged from 10^9 to 10^4 copies (S3 Fig). These standard curves were analyzed in each quantification experiment and demonstrated a high degree of consistency and accuracy with amplification efficiencies of between 1.8 and 2.1. MiR-320a, let-7b, miR-195 and miR-16 were analysed and quantified using RT-qPCR in samples from both the healthy controls and breast cancer patients using both total RNA and enriched small RNA. RT-qPCR quantification was also carried out for the spike-in control miRNA (aly-miR-159) in each sample analysed. The quantification cycle number measurement was consistent across the sample sets with a median of 27.4 and a mean of 28.3 observed for the spike-in control in the total RNA samples. A median of 31.78 and a mean of 31.8 was observed for the spike-in control in the enriched small RNA samples.

An initial validation study was carried out on the RNA isolated from a distinct set of blood samples from a cohort of healthy controls and breast cancer patients. This indicated that the trends observed in the analysis of the small RNA sequencing data were supported in the RT-qPCR analysis of a new sample set. The miR-320a total RNA had a decrease in expression of 2.3-fold ($p = 0.008$) between the breast cancer samples and healthy control group (Table 5, Fig 3A). Let-7b also demonstrated a two-fold decrease in expression levels from breast cancer patients and the healthy control group ($p = 0.03$). MiR-195 and miR-16 showed no notable differences between groups (Table 5, Fig 3A).

The same miRNAs were analysed using enriched small RNA from the same sample set: MiR-320a had 5.45-fold ($p < 0.0001$) and Let-7b 8.5-fold ($p = 0.003$) lower expression levels in the breast cancer blood samples (Table 5, Fig 3B). No differences in expression levels were found for the miR-195 or miR-16 (Table 5, Fig 3B). MiR-320a was the most dysregulated miRNA across all methods applied here: small RNA sequencing, RT-qPCR of total RNA and RT-qPCR of enriched RNA. A further RT-qPCR analysis of miR-320a was carried out on the total and small RNA from the original samples amplified by sequencing. This was combined with the initial data to create an expanded sample set ($n = 55$). The miR-320a expression in total RNA across the expanded sample set showed lower expression in the breast cancer blood samples (2.38-fold, $p = 0.111$) (Table 6, Fig 4A). On enrichment for the small RNA fraction, an 8.89-fold difference ($p = 0.006$) between the healthy control group and breast cancer patients was observed (Table 6, Fig 4B).

Discussion

The quality and quantity of miRNA isolated from peripheral blood samples is dependent on the total RNA extraction and purification methods used. RNA purification methods have a significant impact on experiment reproducibility [56,57] and this can be compounded by differential RNA degradation rates [58], which can subsequently affect small RNA analysis [32,59]. The PAXgene blood miRNA kit was used for total RNA isolation to maintain consistency as well as reduce bias for both small RNA sequencing and RT-qPCR analysis. Expression profiles generated by this collection and purification system are considered accurate [60,61] and it performs well in comparison to other isolation kits and methods [62,63]. This system also allows the samples to be stored in a frozen state for up to two years without a change in the expression

Table 5. Analysis of miRNA dysregulation between healthy and cancer patient samples from the initial sample set.

Total RNA initial sample set					
Two sample T test	Mean	Stdev	SE Mean	P-Value	Fold change Healthy to Cancer
miR-195 Healthy	6.86E+07	8.53E+07	1.38E+07	0.316	1.36
miR-195 Cancer	9.48E+07	8.36E+07	2.16E+07		
miR-16 Healthy	2.51E+08	2.70E+08	4.16E+07	0.933	1.02
miR-16 Cancer	2.56E+08	2.12E+08	5.01E+07		
miR-320a Healthy	1.53E+08	1.83E+08	2.86E+07	0.008	-2.3
miR-320a Cancer	6.63E+07	5.49E+07	1.29E+07		
Let-7b Healthy	2.22E+07	2.93E+07	4.63E+06	0.03	-2.00
Let-7b Cancer	1.11E+07	8.21E+06	1.94E+06		
Enriched small RNA initial sample set					
Two sample T test	Mean	Stdev	SE Mean	P-Value	Fold change Healthy to Cancer
miR-195 Healthy	4.92E+06	5.36E+06	8.69E+05	0.139	1.65
miR-195 Cancer	8.13E+06	7.36E+06	1.90E+06		
miR-16 Healthy	4.76E+07	5.36E+07	8.37E+06	0.729	1.09
miR-16 Cancer	5.18E+07	3.46E+07	8.66E+06		
miR-320a Healthy	4.24E+07	3.84E+07	6.00E+06	0.0001	-5.45
miR-320a Cancer	7.78E+06	1.24E+07	3.10E+06		
Let-7b Healthy	9.15E+06	1.61E+07	2.55E+06	0.003	-8.50
Let-7b Cancer	1.08E+06	7.61E+05	2.03E+05		

doi:10.1371/journal.pone.0137389.t005

profile [61]. In this study, high quality RNA was recovered from all blood samples, showing RIN of 8 and above. A RIN >5 and preferably >8 is considered ideal for small RNA sequencing and RT-qPCR [64].

High quality RNA is essential for small RNA sequencing, and similarly robust methods for processing are also crucial. The small RNA sequencing normalisation methods used here varied in their detection power due to differences in their scaling and ranking algorithms, particularly in relation to assumptions about extreme value and measurement errors [46–48]. This is caused by low miRNA counts and is an ongoing limitation for RNAseq experiment reproducibility [65]. RPKM rescales as a fraction of the total output, but makes no change to the underlying distribution of expression levels. Quantile normalisation assumes errors are global and affects all genes proportionally [66]. TMM reduces the effects of extreme values and normalises genes with comparable expression levels [67]. By comparing RPKM, upper quantile and TMM approaches, we found that these assumptions about the range and distribution of small RNAs expression levels yielded differing numbers of differentially expressed biomarkers. All three methods in our analysis were applied to maximise specificity at a potential loss of sensitivity. Although differential expression testing power and discovery of novel small RNAs would be improved by increasing sequencing depth and provide a larger sample of small RNAs, the alternative of increasing sample size is more pressing because it would have a larger effect on the biological relevance of the observed changes in expression.

After small RNA sequencing normalisation methods were applied, profiling of whole peripheral blood from the healthy control group and breast cancer patients found evidence of lower miR-320a expression in the latter group, consistent with previous observations [68,69]. MiR-320a is a known tumour-suppressor implicated in several cancers including colon. [70–72] and oral squamous cancer [73]. Additionally, it can bind fatty acid synthase and thus limit the spread of some bone cancers [74]. Other recent work suggests that miR-320a expression may exhibit complex effects dependent on the genetic composition of the breast cancer

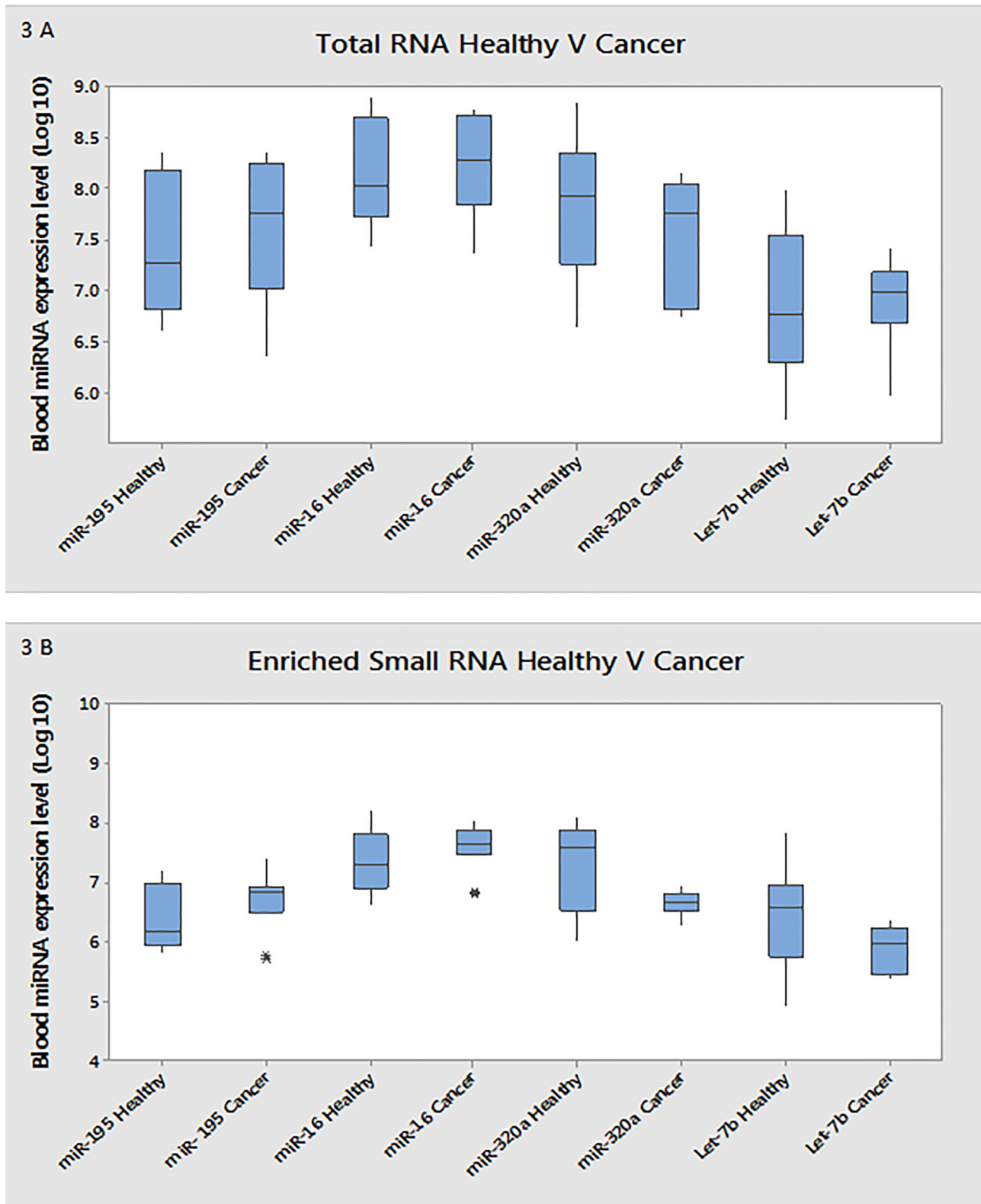


Fig 3. Differential analysis of miRNA isolated from total and enriched RNA from initial sample set. (A) Differential expression of miRNAs between healthy and cancer samples measured in Total RNA from the initial sample set measured using RT-qPCR. A 2.3 fold decrease is observed (with a P value of 0.008), in miR-320a from Healthy ($1.53E +08$) to Cancer ($6.63E +07$) state. Log 10 box-plots indicating changes in expression of miR-195, miR-16 and Let-7b

from healthy to cancer state measured in total RNA using RT-qPCR. (B) Differential expression of miRNAs between healthy and cancer samples measured in enriched small RNA from the initial sample set measured using RT-qPCR. A 5.45 fold decrease is observed (with a P value of 0.0001), in miR-320a from Healthy (4.24E +07) to Cancer (7.78E +06) state. Log 10 box-plots indicating changes in expression of miR-195, miR-16 and Let-7b from healthy to cancer state measured in total RNA using RT-qPCR. *Indicates an outlying value.

doi:10.1371/journal.pone.0137389.g003

subtype, but that its dysregulation is linked to poorer survival outcome [75] as supported consistently across all reported investigations, irrespective of methodology the used.

MiR-140 was also identified as dysregulated in the breast cancer blood samples with a 1.72-fold lower expression based on our sequencing data. However, the signal was inconsistent across samples, indicating that it may have diagnostic value in a subset of breast cancer cases. Previously, miR-140 down-regulation was linked to the promotion of cancer stem cell formation in basal-like early stage breast cancer [76]. MiR-140 has also been found to suppress both the metastasis as well as the growth of tumours in patients with non-small cell lung cancer by targeting the insulin-like growth factor 1 receptor [77]. This supports the proposition that targeting miR-140, its regulators, or mRNA targets could improve survival outcomes by regulating stem cell formation [73].

Prior to small RNA sequencing, all samples were processed and enriched for such that only mature miRNAs sequences were examined. To correlate the results obtained from small RNA sequencing and RT-qPCR, the use of a small RNA enrichment step was investigated in this study. A synthetic oligonucleotide representing the precursor to miR-16 was generated for use as control template in a miR-16 Taqman assay. This synthetic precursor oligonucleotide template was detected by the mature miR-16 RT-qPCR assay, indicating a lack of specificity for the mature miRNA (Fig 2, S2 Table). In order to remove this potential for inaccuracy of quantification of miRNAs using qPCR it may also be necessary to employ a size selection step during RNA purification.

As such, for the purposes of this study miR-320a, let-7b, miR-16 and miR-195 were examined by RT-qPCR using both total RNA and also by developing a pre-RT-qPCR enrichment process for RNA <30 bases. Size selection enrichment for the small RNA population decreased RNA copy number, but may have increased the fraction of mature RNAs (Table 4). The decrease in copy numbers of specific miRNAs observed from total RNA to enriched small RNA samples highlights the potential for error in quantification due to precursor miRNAs in the total RNA. A non-human miRNA (*Arabidopsis* miR-159) was measured before and after small RNA enrichment as a control. This miRNA had a consistent Cq of 27±3 in total RNA and 31±5 in enriched small RNA. The RT-qPCR data was normalised using the median miR-159 expression to reduce potential bias. The data generated for the enriched small RNA consistently showed less variance and were more reliable than those for the total RNA. The steps taken in processing the samples and the data analysis have resulted in an increased accuracy in

Table 6. Analysis of miRNA dysregulation between healthy and cancer patient samples for the complete sample set.

		Total RNA full sample set				
Two-sample T-test	Mean	Stdev	SE Mean	P-Value	Fold change Healthy to Cancer	
Total miR-320a Healthy	1.58E+08	1.83E+08	4.90E+07			
Total miR-320a Cancer	6.63E+07	5.84E+07	2.38E+07	0.111	-2.38	
		Enriched small RNA full sample set				
Two-sample T-test	Mean	Stdev	SE Mean	P-Value	Fold change Healthy to Cancer	
Enriched Small miR-320a Healthy	4.20E+07	4.03E+07	1.12E+07			
Enriched Small miR-320a Cancer	4.72E+06	2.20E+06	9.83E+05	0.006	- 8.89	

doi:10.1371/journal.pone.0137389.t006

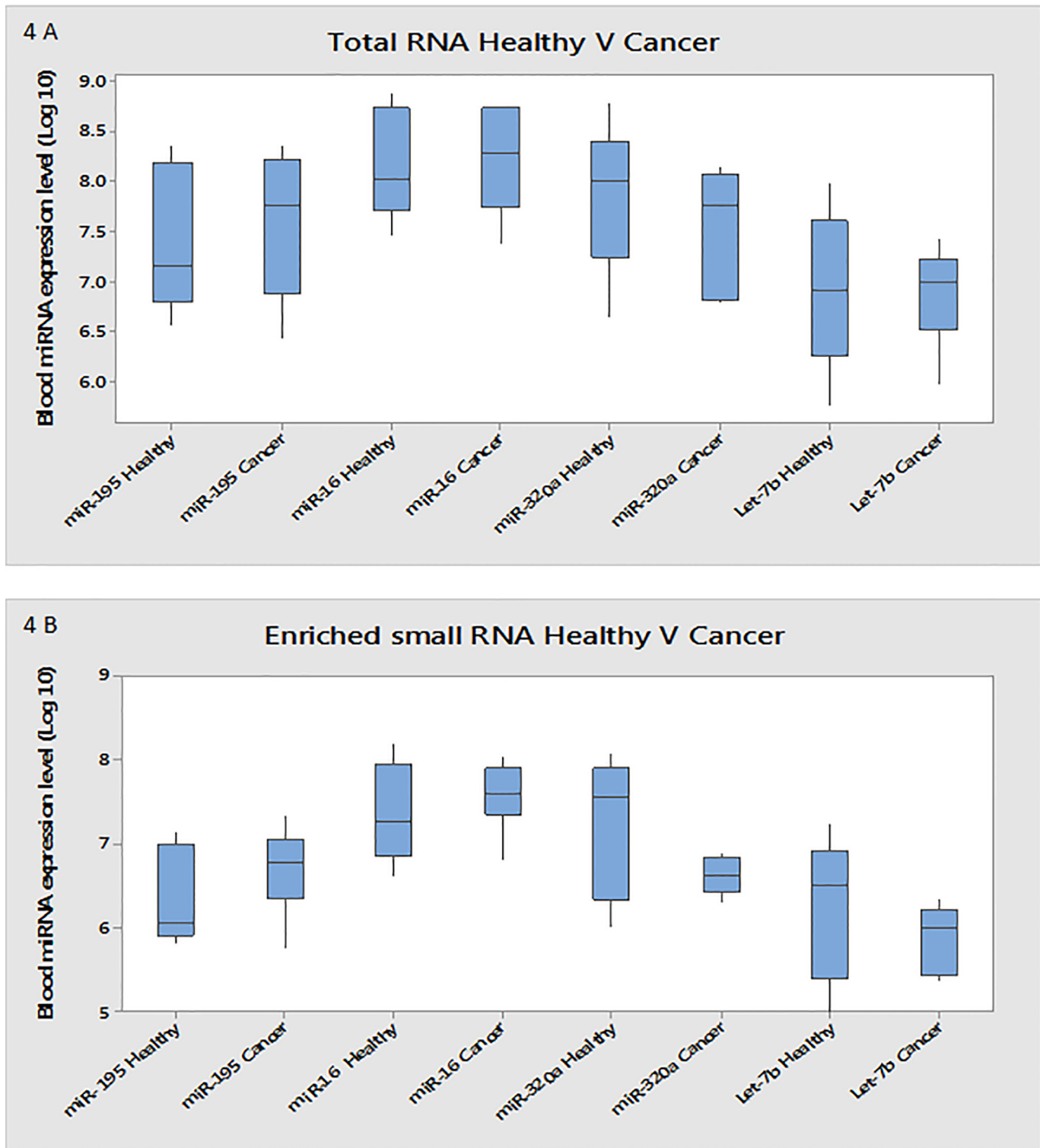


Fig 4. Differential analysis of miRNA isolated from total and enriched RNA from complete sample set. (A) Differential expression of miRNAs between healthy and cancer samples measured in total RNA from the complete sample set measured using RT-qPCR. A 2.38 fold decrease is observed (with a P value of 0.111), in miR-320a from Healthy ($1.58E+08$) to Cancer ($6.63E+07$) state. Log 10 box-plots indicating changes in expression of miR-195, miR-16 and Let-7b from healthy to cancer state measured in total RNA using RT-qPCR. (B) Differential expression of selected miRNAs between healthy and cancer samples measured in enriched small RNA from the complete sample set measured using RT-qPCR. An 8.89 fold decrease is observed (with a P value of 0.006), in miR-320a from Healthy ($4.20E+07$) to Cancer ($4.72E+06$) state. Log 10 box-plots indicating changes in expression of miR-195, miR-16 and Let-7b from healthy to cancer state measured in enriched small RNA using RT-qPCR.

doi:10.1371/journal.pone.0137389.g004

the measurement of mature miRNA targets and thus superior inference power due to lower variation in expression due to experimental procedures.

These standardisation steps facilitated the validation by RT-qPCR relative expression differences first observed in small RNA sequence reads. The size selection step and miR-159 control normalisation indicated an 8.89-fold down-regulation of miR-320a in circulating blood from breast cancer patients relative to the healthy control group. Let-7b also had lower expression in circulating blood from the breast cancer cohort, an observation that was not statistically supported by sequencing results, and no evidence of differential expression was observed using either method for miR-195 or miR-16. “This finding is consistent with the use of miR-16 as a stable endogenous control [78].”

These findings serve to demonstrate the utility of the standardisation and normalisation steps carried out in the discovery and initial validation of potential biomarkers for disease through the comparison of data across platform types. This study developed a suitable RT-qPCR miRNA quantification protocol that produced measurements comparable with small RNA sequencing output. This supports the perspective that relative gene expression levels can be measured in a reproducible manner across platforms [79], so that high-throughput miRNA discovery in pilot cohorts can be validated in larger patient groups using RT-qPCR. Imprecision in RT-qPCR protocols and primer sequence design can yield inaccurate measurements as can GC bias in small RNA sequencing [80] or in TRIzol extraction [81], and due to local sequence artefacts [82]. The reduction of experimental variation in the process of miRNA quantification from sample collection to final interpretation through the use of the steps outlined in this study allow for more robust and comparable findings. By applying these methods to whole peripheral blood RNA from healthy control group and breast cancer patients, miR-320a expression was down regulated using both small RNA sequencing and RT-qPCR. However, the clinical significance of this miRNA as a potential biomarker for breast cancer from circulating blood must now be evaluated further using larger cohort sample sizes.

Conclusion

The advances in the field of circulating miRNAs and their relevance for disease profiling and treatment have dramatically increased interest in their utility as biomarkers. They exhibit many features that make them ideally suited to use as diagnostic and prognostic biomarkers. The lack of consistency in reporting and the difficulty comparing data from one study to another are major stumbling blocks in the advancement of miRNAs as clinically relevant biomarkers [83].

On commencement of this study there was an absence of a standardised protocol for the treatment of RNA samples from isolation from the blood to expression comparison between methods. In order to facilitate cross study comparison of data and to increase the accuracy, robustness and repeatability of findings, the development of such a protocol was important. Overall, our findings emphasise the importance of the use of a standardised protocol in the processing of RNA samples and the development of miRNA profiles.

We have demonstrated an initial proof of concept for a pipeline of controlled steps involving standardisation and normalisation that serves to reduce the potential for bias, inaccuracy and variability in the findings. Through further validation of this pipeline and its application in cross platform studies, it will be possible to accurately compare miRNA profiles both between small RNA sequencing and RT-qPCR. If miRNA quantification and profiling studies used a standardised approach, it may become possible to accurately compare data generated in separate studies. This would decrease the potential for conflicting results and provide a global resource of data that would lead to an enhanced potential for meaningful developments in the field of miRNA profiling.

Supporting Information

S1 Fig. Global miRNA expression profiles of (A) control and (B) breast cancer patient samples. These profiles are comparable with no major differences between patient cohorts. The most highly expressed miRNA (miR-451a) accounted for 60% (control) and 65% (breast cancer) of the total miRNA population.

(PDF)

S2 Fig. Global unannotated small RNA expression profiles of (A) control and (B) breast cancer patient samples. These profiles are comparable with no major differences between cohorts. The percentages correspond to values across all small RNAs only, which accounted for 2.5% (control) and 2.6% (breast cancer) of the total small RNA population. The ten small RNAs shown account for 48.5% (control) and 45.2% (breast cancer) of the total. ENST00000364228 was the most highly expressed in both groups (17.2% of total in controls, 14.0% of total in breast cancer cohort).

(PDF)

S3 Fig. Representative standard curve employed for Aly-miR-159 spike-in control assay.

This spike-in control assay was used for normalisation of RT-qPCR data.

(PDF)

S1 Table. Small unannotated RNAs detected unique to breast cancer patients.

(PDF)

S2 Table. Cq values obtained for mature and precursor miR-16 assay.

(PDF)

Author Contributions

Conceived and designed the experiments: HK TD EC TS NLT TB KR. Performed the experiments: HK TD. Analyzed the data: HK TD RMD EC TB KR. Contributed reagents/materials/analysis tools: TD RMD MJK TB. Wrote the paper: HK TD NLT TS MJK RMD EC TB KR.

References

1. Grosshans H, Filipowicz W (2008) Molecular biology—The expanding world of small RNAs. *Nature* 451: 414–416. doi: [10.1038/451414a](https://doi.org/10.1038/451414a) PMID: [18216846](https://pubmed.ncbi.nlm.nih.gov/18216846/)
2. Silahtaroglu A, Stenvang J (2010) MicroRNAs, epigenetics and disease. *Essays in Biochemistry: Epigenetics, Disease and Behaviour* 48: 165–185.
3. Ambros V (2004) The functions of animal microRNAs. *Nature* 431: 350–355. PMID: [15372042](https://pubmed.ncbi.nlm.nih.gov/15372042/)
4. Zhang BH, Pan XP, Cobb GP, Anderson TA (2006) Plant microRNA: A small regulatory molecule with big impact. *Developmental Biology* 289: 3–16. PMID: [16325172](https://pubmed.ncbi.nlm.nih.gov/16325172/)
5. Cullen BR (2006) Viruses and microRNAs. *Nature Genetics* 38: S25–S30. PMID: [16736021](https://pubmed.ncbi.nlm.nih.gov/16736021/)
6. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, et al. (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Research* 37: D98–D104. doi: [10.1093/nar/gkn714](https://doi.org/10.1093/nar/gkn714) PMID: [18927107](https://pubmed.ncbi.nlm.nih.gov/18927107/)
7. Martins PAD, Leptidis S, Salic K, De Windt LJ (2010) MicroRNA Regulation in Cardiovascular Disease. *Current Drug Targets* 11: 900–906. PMID: [20415652](https://pubmed.ncbi.nlm.nih.gov/20415652/)
8. Farazi TA, Spitzer JI, Morozov P, Tuschl T (2011) miRNAs in human cancer. *Journal of Pathology* 223: 102–115. doi: [10.1002/path.2806](https://doi.org/10.1002/path.2806) PMID: [21125669](https://pubmed.ncbi.nlm.nih.gov/21125669/)
9. Saugstad JA (2010) MicroRNAs as effectors of brain function with roles in ischemia and injury, neuroprotection, and neurodegeneration. *Journal of Cerebral Blood Flow and Metabolism* 30: 1564–1576. doi: [10.1038/jcbfm.2010.101](https://doi.org/10.1038/jcbfm.2010.101) PMID: [20606686](https://pubmed.ncbi.nlm.nih.gov/20606686/)

10. Schickel R, Boyerinas B, Park SM, Peter ME (2008) MicroRNAs: key players in the immune system, differentiation, tumorigenesis and cell death. *Oncogene* 27: 5959–5974. doi: [10.1038/onc.2008.274](https://doi.org/10.1038/onc.2008.274) PMID: [18836476](https://pubmed.ncbi.nlm.nih.gov/18836476/)
11. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* 115: 787–798. PMID: [14697198](https://pubmed.ncbi.nlm.nih.gov/14697198/)
12. Friedman RC, Farh KKH, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* 19: 92–105. doi: [10.1101/gr.082701.108](https://doi.org/10.1101/gr.082701.108) PMID: [18955434](https://pubmed.ncbi.nlm.nih.gov/18955434/)
13. Zhang R, Su B (2009) Small but influential: the role of microRNAs on gene regulatory network and 3' UTR evolution. *Journal of Genetics and Genomics* 36: 1–6. doi: [10.1016/S1673-8527\(09\)60001-1](https://doi.org/10.1016/S1673-8527(09)60001-1) PMID: [19161940](https://pubmed.ncbi.nlm.nih.gov/19161940/)
14. Bartel DP, Chen CZ (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature Reviews Genetics* 5: 396–400. PMID: [15143321](https://pubmed.ncbi.nlm.nih.gov/15143321/)
15. Luo X, Burwinkel B, Tao S, Brenner H (2011) MicroRNA Signatures: Novel Biomarker for Colorectal Cancer? *Cancer Epidemiology Biomarkers & Prevention* 20: 1272–1286. doi: [10.1158/1055-9965.EPI-11-0035](https://doi.org/10.1158/1055-9965.EPI-11-0035) PMID: [21551242](https://pubmed.ncbi.nlm.nih.gov/21551242/)
16. Kosaka N, Iguchi H, Ochiya T (2010) Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. *Cancer Science* 101: 2087–2092. doi: [10.1111/j.1349-7006.2010.01650.x](https://doi.org/10.1111/j.1349-7006.2010.01650.x) PMID: [20624164](https://pubmed.ncbi.nlm.nih.gov/20624164/)
17. Roth C, Kasimir-Bauer S, Pantel K, Schwarzenbach H (2011) Screening for circulating nucleic acids and caspase activity in the peripheral blood as potential diagnostic tools in lung cancer. *Molecular Oncology* 5: 281–291. doi: [10.1016/j.molonc.2011.02.002](https://doi.org/10.1016/j.molonc.2011.02.002) PMID: [21398193](https://pubmed.ncbi.nlm.nih.gov/21398193/)
18. Shapira I, Oswald M, Lovecchio J, Khalili H, Menzin A, et al. (2014) Circulating biomarkers for detection of ovarian cancer and predicting cancer outcomes. *Br J Cancer* 110: 976–983. doi: [10.1038/bjc.2013.795](https://doi.org/10.1038/bjc.2013.795) PMID: [24366298](https://pubmed.ncbi.nlm.nih.gov/24366298/)
19. Mar-Aguilar F, Mendoza-Ramírez JA, Malagón-Santiago I, Espino-Silva PK, Santuario-Facio SK, et al. (2013) Serum Circulating microRNA Profiling for Identification of Potential Breast Cancer Biomarkers. *Disease Markers* 34. doi: [10.3233/DMA-120957](https://doi.org/10.3233/DMA-120957) PMID: [23334650](https://pubmed.ncbi.nlm.nih.gov/23334650/)
20. Chen CF, Ridzon DA, Broomer AJ, Zhou ZH, Lee DH, et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research* 33. PMID: [16314309](https://pubmed.ncbi.nlm.nih.gov/16314309/)
21. Shi R, Chiang VL (2005) Facile means for quantifying microRNA expression by real-time PCR. *Biotechniques* 39: 519–525. PMID: [16235564](https://pubmed.ncbi.nlm.nih.gov/16235564/)
22. Hanna JA, Wimberly H, Kumar S, Slack F, Agarwal S, et al. (2012) Quantitative analysis of microRNAs in tissue microarrays by in situ hybridization. *Biotechniques* 52: 235–+. doi: [10.2144/000113837](https://doi.org/10.2144/000113837) PMID: [22482439](https://pubmed.ncbi.nlm.nih.gov/22482439/)
23. Yin JQ, Zhao RC, Morris KV (2008) Profiling microRNA expression with microarrays. *Trends in Biotechnology* 26: 70–76. doi: [10.1016/j.tibtech.2007.11.007](https://doi.org/10.1016/j.tibtech.2007.11.007) PMID: [18191262](https://pubmed.ncbi.nlm.nih.gov/18191262/)
24. Li W, Ruan KC (2009) MicroRNA detection by microarray. *Analytical and Bioanalytical Chemistry* 394: 1117–1124. doi: [10.1007/s00216-008-2570-2](https://doi.org/10.1007/s00216-008-2570-2) PMID: [19132354](https://pubmed.ncbi.nlm.nih.gov/19132354/)
25. Hafner M, Landgraf P, Ludwig J, Rice A, Ojo T, et al. (2008) Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 44: 3–12. PMID: [18158127](https://pubmed.ncbi.nlm.nih.gov/18158127/)
26. Lopez J, Diallo A, Cruceanu C, Fiori L, Laboissiere S, et al. (2015) Biomarker discovery: quantification of microRNAs and other small non-coding RNAs using next generation sequencing. *BMC Medical Genomics* 8: 35. doi: [10.1186/s12920-015-0109-x](https://doi.org/10.1186/s12920-015-0109-x) PMID: [26130076](https://pubmed.ncbi.nlm.nih.gov/26130076/)
27. Becker C, Hammerle-Fickinger A, Riedmaier I, Pfaffl MW (2010) mRNA and microRNA quality control for RT-qPCR analysis. *Methods* 50: 237–243. doi: [10.1016/j.ymeth.2010.01.010](https://doi.org/10.1016/j.ymeth.2010.01.010) PMID: [20079844](https://pubmed.ncbi.nlm.nih.gov/20079844/)
28. Benes V, Castoldi M (2010) Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods* 50: 244–249. doi: [10.1016/j.ymeth.2010.01.026](https://doi.org/10.1016/j.ymeth.2010.01.026) PMID: [20109550](https://pubmed.ncbi.nlm.nih.gov/20109550/)
29. Pieter M, Nicole H, Lukas B, Ditte A, Nathalie B, et al. (2014) Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nature Methods*.
30. Chevillet JR, Lee I, Briggs HA, He Y, Wang K (2014) Issues and Prospects of microRNA-Based Biomarkers in Blood and Other Body Fluids. *Molecules* 19: 6080–6105. doi: [10.3390/molecules19056080](https://doi.org/10.3390/molecules19056080) PMID: [24830712](https://pubmed.ncbi.nlm.nih.gov/24830712/)
31. Git A, Dvinge H, Salmon-Divon M, Osborne M, Kutter C, et al. (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *Rna-a Publication of the Rna Society* 16: 991–1006.
32. Moldovan L, Batte KE, Trgovcich J, Wisler J, Marsh CB, et al. (2014) Methodological challenges in utilizing miRNAs as circulating biomarkers. *Journal of Cellular and Molecular Medicine* 18: 371–390. doi: [10.1111/jcmm.12236](https://doi.org/10.1111/jcmm.12236) PMID: [24533657](https://pubmed.ncbi.nlm.nih.gov/24533657/)

33. Redshaw N, Wilkes T, Whale A, Cowen S, Huggett J, et al. (2013) A comparison of miRNA isolation and RT-qPCR technologies and their effects on quantification accuracy and repeatability. *Biotechniques* 54: 155–+. doi: [10.2144/000114002](https://doi.org/10.2144/000114002) PMID: [23477383](https://pubmed.ncbi.nlm.nih.gov/23477383/)
34. Huggett J, Dheda K, Bustin S, Zumla A (2005) Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun* 6: 279–284. PMID: [15815687](https://pubmed.ncbi.nlm.nih.gov/15815687/)
35. Risso D, Ngai J, Speed TP, Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology* 32: 896–902. doi: [10.1038/nbt.2931](https://doi.org/10.1038/nbt.2931) PMID: [25150836](https://pubmed.ncbi.nlm.nih.gov/25150836/)
36. Kratz A, Carninci P (2014) The devil in the details of RNA-seq. *Nat Biotech* 32: 882–884.
37. Lu C, Meyers BC, Green PJ (2007) Construction of small RNA cDNA libraries for deep sequencing. *Methods* 43: 110–117. PMID: [17889797](https://pubmed.ncbi.nlm.nih.gov/17889797/)
38. <http://www.prealanalytix.com/product-catalog/blood/ma/products/paxgene-blood-rna-kit/>.
39. Meyer M, Kircher M (2010) Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb Protoc* 2010.
40. Illumina <http://www.illumina.com/applications/sequencing/rna.html>.
41. Bentley D (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456: 53–59. doi: [10.1038/nature07517](https://doi.org/10.1038/nature07517) PMID: [18987734](https://pubmed.ncbi.nlm.nih.gov/18987734/)
42. www.bioinformatics.babraham.ac.uk/projects/fastqc/.
43. <http://www.clcbio.com/>.
44. Lawless N, Foroushani ABK, McCabe MS, O'Farrelly C, Lynn DJ (2013) Next Generation Sequencing Reveals the Expression of a Unique miRNA Profile in Response to a Gram-Positive Bacterial Infection. *Plos One* 8. doi: [10.1371/journal.pone.0057543](https://doi.org/10.1371/journal.pone.0057543) PMID: [23472090](https://pubmed.ncbi.nlm.nih.gov/23472090/)
45. Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628. doi: [10.1038/nmeth.1226](https://doi.org/10.1038/nmeth.1226) PMID: [18516045](https://pubmed.ncbi.nlm.nih.gov/18516045/)
46. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 11. doi: [10.1186/gb-2010-11-3-r25](https://doi.org/10.1186/gb-2010-11-3-r25) PMID: [20196867](https://pubmed.ncbi.nlm.nih.gov/20196867/)
47. Bullard JH, Purdom E, Hansen KD, Dudoit S (2010) Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *Bmc Bioinformatics* 11. doi: [10.1186/1471-2105-11-94](https://doi.org/10.1186/1471-2105-11-94) PMID: [20167110](https://pubmed.ncbi.nlm.nih.gov/20167110/)
48. Garmire LX, Subramaniam S (2012) Evaluation of normalization methods in mammalian microRNA-Seq data. *Rna-a Publication of the Rna Society* 18: 1279–1288.
49. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19: 185–193. PMID: [12538238](https://pubmed.ncbi.nlm.nih.gov/12538238/)
50. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140. doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616) PMID: [19910308](https://pubmed.ncbi.nlm.nih.gov/19910308/)
51. Griffiths-Jones S, Grocock R, van Dongen S, Bateman A, Enright A (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34: D140–144. PMID: [16381832](https://pubmed.ncbi.nlm.nih.gov/16381832/)
52. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences* 105: 10513–10518.
53. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Research* 36: D154–D158. PMID: [17991681](https://pubmed.ncbi.nlm.nih.gov/17991681/)
54. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207–210. PMID: [11752295](https://pubmed.ncbi.nlm.nih.gov/11752295/)
55. Svec D, Tichopad A, Novosadova V, Pfaffl MW, Kubista M (2015) How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomolecular Detection and Quantification*.
56. Kroh E, Parkin R, Mitchell P, Tewari M (2010) Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods* 50: 298–301. doi: [10.1016/j.ymeth.2010.01.032](https://doi.org/10.1016/j.ymeth.2010.01.032) PMID: [20146939](https://pubmed.ncbi.nlm.nih.gov/20146939/)
57. McDonald JS, Milosevic D, Reddi HV, Grebe SK, Algeciras-Schimmich A (2011) Analysis of Circulating MicroRNA: Preanalytical and Analytical Challenges. *Clinical Chemistry* 57: 833–840. doi: [10.1373/clinchem.2010.157198](https://doi.org/10.1373/clinchem.2010.157198) PMID: [21487102](https://pubmed.ncbi.nlm.nih.gov/21487102/)
58. Baker M (2010) MicroRNA profiling: separating signal from noise. *Nature Methods* 7: 687–692. doi: [10.1038/nmeth0910-687](https://doi.org/10.1038/nmeth0910-687) PMID: [20805796](https://pubmed.ncbi.nlm.nih.gov/20805796/)

59. Meyer CA, Liu XS (2014) Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nature Reviews Genetics* 15: 709–721. doi: [10.1038/nrg3788](https://doi.org/10.1038/nrg3788) PMID: [25223782](https://pubmed.ncbi.nlm.nih.gov/25223782/)
60. Patnaik SK, Yendamuri S, Kannisto E, Kucharczuk JC, Singhal S, et al. (2012) MicroRNA Expression Profiles of Whole Blood in Lung Adenocarcinoma. *Plos One* 7. doi: [10.1371/journal.pone.0046045](https://doi.org/10.1371/journal.pone.0046045) PMID: [23029380](https://pubmed.ncbi.nlm.nih.gov/23029380/)
61. Chal V, Vassilakos A, Lee Y, Wright JA, Young AH (2005) Optimization of the PAXgene (TM) blood RNA extraction system for gene expression analysis of clinical samples. *Journal of Clinical Laboratory Analysis* 19: 182–188. PMID: [16170815](https://pubmed.ncbi.nlm.nih.gov/16170815/)
62. Kim J-H, Jin H-O, Park J-A, Chang YH, Hong YJ, et al. (2014) Comparison of three different kits for extraction of high-quality RNA from frozen blood. *SpringerPlus* 3: 76. doi: [10.1186/2193-1801-3-76](https://doi.org/10.1186/2193-1801-3-76) PMID: [24567882](https://pubmed.ncbi.nlm.nih.gov/24567882/)
63. Keller A, Leidinger P, Steinmeyer F, Stähler C, Franke A, et al. (2014) Comprehensive analysis of microRNA profiles in multiple sclerosis including next-generation sequencing. *Multiple Sclerosis Journal* 20: 295–303. doi: [10.1177/1352458513496343](https://doi.org/10.1177/1352458513496343) PMID: [23836875](https://pubmed.ncbi.nlm.nih.gov/23836875/)
64. Fleige S, Pfaffl M (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* 27: 126–139. PMID: [16469371](https://pubmed.ncbi.nlm.nih.gov/16469371/)
65. Luo HE, Li JT, Chia BKH, Robson P, Nagarajan N (2014) The importance of study design for detecting differentially abundant features in high-throughput experiments. *Genome Biology* 15. doi: [10.1186/s13059-014-0527-7](https://doi.org/10.1186/s13059-014-0527-7) PMID: [25517037](https://pubmed.ncbi.nlm.nih.gov/25517037/)
66. Hicks SC, Irizarry RA (2014) When to use Quantile Normalization?
67. Zhou X, Oshlack A, Robinson MD (2013) miRNA-Seq normalization comparisons need improvement. *Rna-a Publication of the Rna Society* 19: 733–734.
68. Cheng ZA, Qiu SB, Jiang L, Zhang AL, Bao WJ, et al. (2013) MiR-320a is Downregulated in Patients with Myasthenia Gravis and Modulates Inflammatory Cytokines Production by Targeting Mitogen-activated Protein Kinase 1. *Journal of Clinical Immunology* 33: 567–576. doi: [10.1007/s10875-012-9834-5](https://doi.org/10.1007/s10875-012-9834-5) PMID: [23196978](https://pubmed.ncbi.nlm.nih.gov/23196978/)
69. Shang C, Zhang H, Guo Y, Hong Y, Liu YH, et al. (2014) MiR-320a down-regulation mediates bladder carcinoma invasion by targeting ITGB3. *Molecular Biology Reports* 41: 2521–2527. doi: [10.1007/s11033-014-3110-0](https://doi.org/10.1007/s11033-014-3110-0) PMID: [24443232](https://pubmed.ncbi.nlm.nih.gov/24443232/)
70. Zhao H, Dong T, Zhou H, Wang L, Huang A, et al. (2013) miR-320a suppresses colorectal cancer progression by targeting Rac1. *Carcinogenesis*.
71. Sun J-Y, Huang Y, Li J-P, Zhang X, Wang L, et al. (2012) MicroRNA-320a suppresses human colon cancer cell proliferation by directly targeting β -catenin. *Biochemical and Biophysical Research Communications* 420: 787–792. doi: [10.1016/j.bbrc.2012.03.075](https://doi.org/10.1016/j.bbrc.2012.03.075) PMID: [22459450](https://pubmed.ncbi.nlm.nih.gov/22459450/)
72. Huang A, Zhao H, Quan Y, Jin R, Feng B, et al. (2014) E2A Predicts Prognosis of Colorectal Cancer Patients and Regulates Cancer Cell Growth by Targeting miR-320a. *PLoS ONE* 9: e85201. doi: [10.1371/journal.pone.0085201](https://doi.org/10.1371/journal.pone.0085201) PMID: [24454819](https://pubmed.ncbi.nlm.nih.gov/24454819/)
73. Wu YY, Chen YL, Jao YC, Hsieh IS, Chang KC, et al. (2014) miR-320 regulates tumor angiogenesis driven by vascular endothelial cells in oral cancer by silencing neuropilin 1. *Angiogenesis* 17: 247–260. doi: [10.1007/s10456-013-9394-1](https://doi.org/10.1007/s10456-013-9394-1) PMID: [24114198](https://pubmed.ncbi.nlm.nih.gov/24114198/)
74. Cheng C, Chen ZQ, Shi XT (2014) MicroRNA-320 inhibits osteosarcoma cells proliferation by directly targeting fatty acid synthase. *Tumor Biology* 35: 4177–4183. doi: [10.1007/s13277-013-1546-9](https://doi.org/10.1007/s13277-013-1546-9) PMID: [24390663](https://pubmed.ncbi.nlm.nih.gov/24390663/)
75. Yang HP, Yu J, Wang L, Ding D, Zhang L, et al. (2014) miR-320a is an independent prognostic biomarker for invasive breast cancer. *Oncology Letters* 8: 1043–1050. PMID: [25120655](https://pubmed.ncbi.nlm.nih.gov/25120655/)
76. Li Q, Yao Y, Eades G, Liu Z, Zhang Y, et al. (2014) Downregulation of miR-140 promotes cancer stem cell formation in basal-like early stage breast cancer. *Oncogene* 33: 2589–2600. doi: [10.1038/onc.2013.226](https://doi.org/10.1038/onc.2013.226) PMID: [23752191](https://pubmed.ncbi.nlm.nih.gov/23752191/)
77. Yuan YF, Shen YX, Xue L, Fan H (2013) miR-140 Suppresses Tumor Growth and Metastasis of Non-Small Cell Lung Cancer by Targeting Insulin-Like Growth Factor 1 Receptor. *Plos One* 8. doi: [10.1371/journal.pone.0073604](https://doi.org/10.1371/journal.pone.0073604) PMID: [24039995](https://pubmed.ncbi.nlm.nih.gov/24039995/)
78. Heneghan HM, Miller N, Lowery AJ, Sweeney KJ, Newell J, et al. (2010) Circulating microRNAs as Novel Minimally Invasive Biomarkers for Breast Cancer. *Annals of Surgery* 251: 499–505. doi: [10.1097/SLA.1090b1013e3181cc1939f](https://doi.org/10.1097/SLA.1090b1013e3181cc1939f) PMID: [20134314](https://pubmed.ncbi.nlm.nih.gov/20134314/)
79. Su ZQ, Labaj PP, Li S, Thierry-Mieg J, Thierry-Mieg D, et al. (2014) A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nature Biotechnology* 32: 903–914. doi: [10.1038/nbt.2957](https://doi.org/10.1038/nbt.2957) PMID: [25150838](https://pubmed.ncbi.nlm.nih.gov/25150838/)

80. Aird D, Ross MG, Chen WS, Danielsson M, Fennell T, et al. (2011) Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology* 12. doi: [10.1186/gb-2011-12-2-r18](https://doi.org/10.1186/gb-2011-12-2-r18) PMID: [21338519](https://pubmed.ncbi.nlm.nih.gov/21338519/)
81. Kim Y, Yeo J, Kim B, Ha M, Kim V (2012) Short Structured RNAs with Low GC Content Are Selectively Lost during Extraction from a Small Number of Cells. *Mol Cell* 46: 893–895. doi: [10.1016/j.molcel.2012.05.036](https://doi.org/10.1016/j.molcel.2012.05.036) PMID: [22749402](https://pubmed.ncbi.nlm.nih.gov/22749402/)
82. VanGuilder HD, Vrana KE, Freeman WM (2008) Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques* 44: 619–626. doi: [10.2144/000112776](https://doi.org/10.2144/000112776) PMID: [18474036](https://pubmed.ncbi.nlm.nih.gov/18474036/)
83. Leidner RS, Li L, Thompson CL (2013) Dampening Enthusiasm for Circulating MicroRNA in Breast Cancer. *PLoS ONE* 8. doi: [10.1371/journal.pone.0057841](https://doi.org/10.1371/journal.pone.0057841) PMID: [23472110](https://pubmed.ncbi.nlm.nih.gov/23472110/)