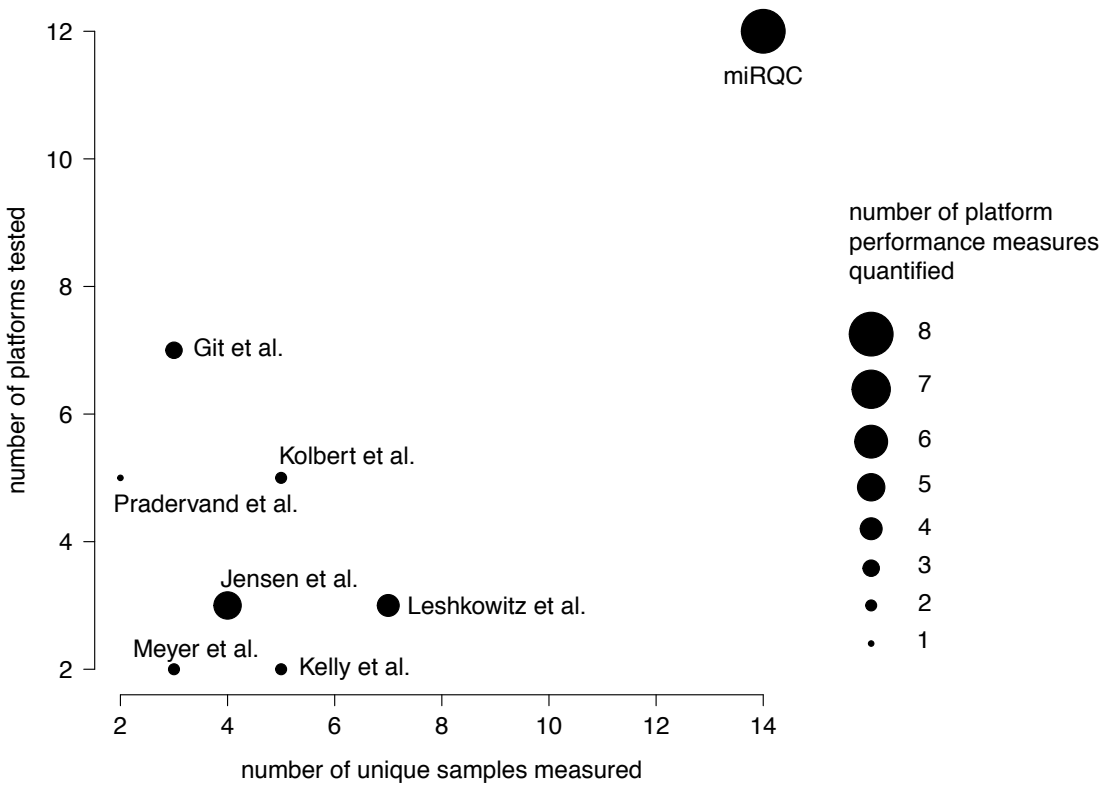
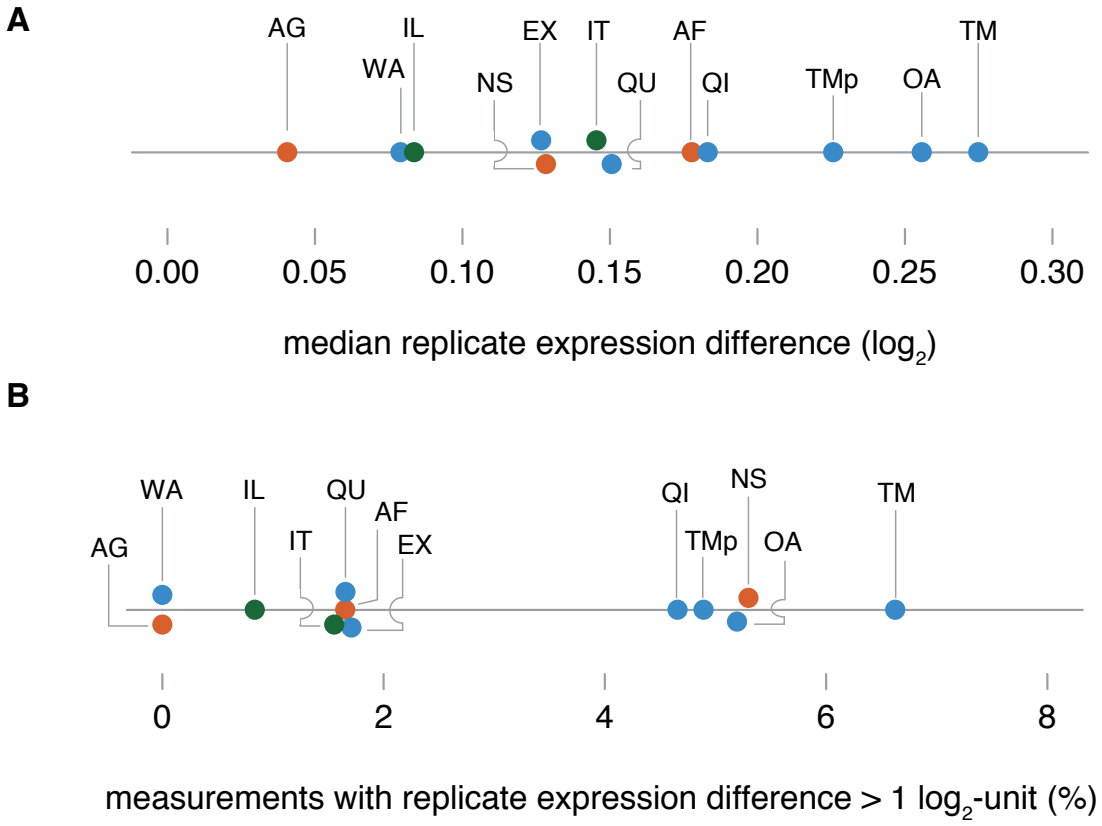


Supplementary Figure 1

Platforms, performance measures and samples measured by other studies.

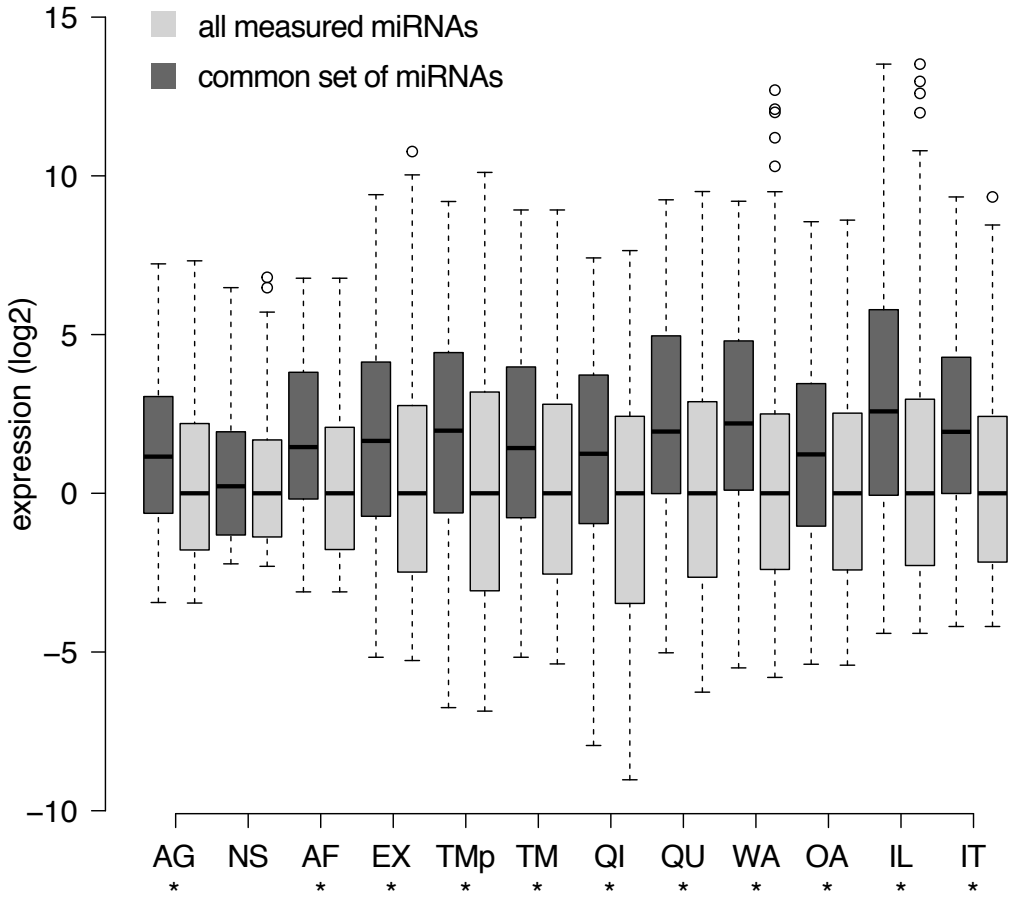
The number of platforms, the number of unique samples and the number of platform performance measures evaluated by other miRNA expression platform performance studies and the miRQC study.



Supplementary Figure 2

Platform reproducibility.

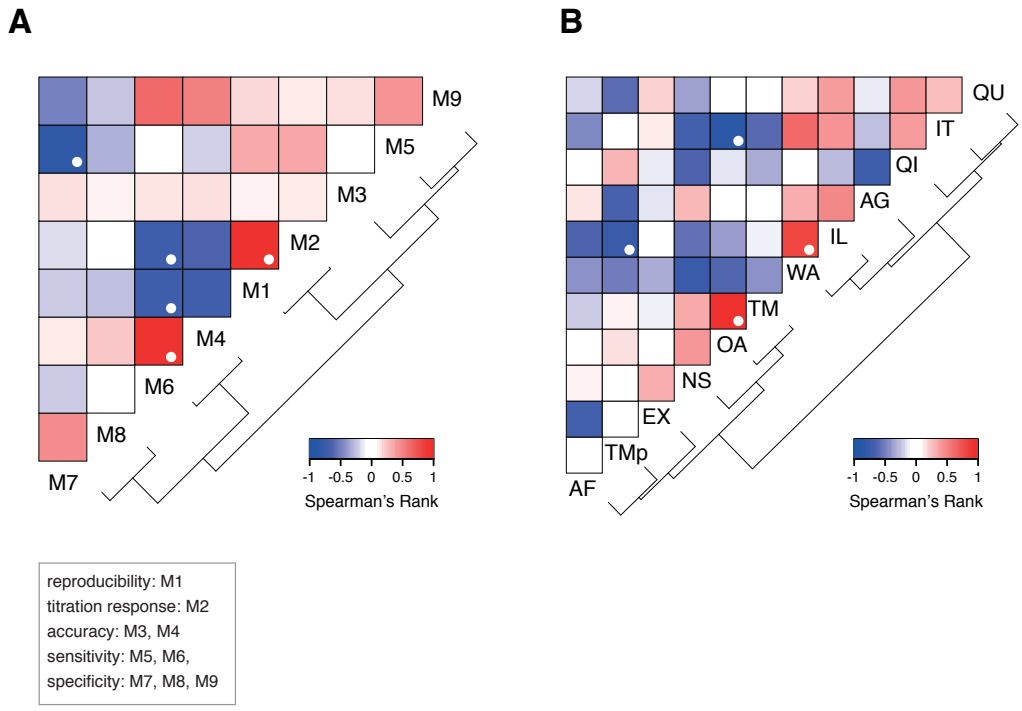
(A) Median replicate expression difference for each platform. (B) Percentage of measurements with a replicate expression difference > 2-fold.



Supplementary Figure 3

Relative miRNA expression levels.

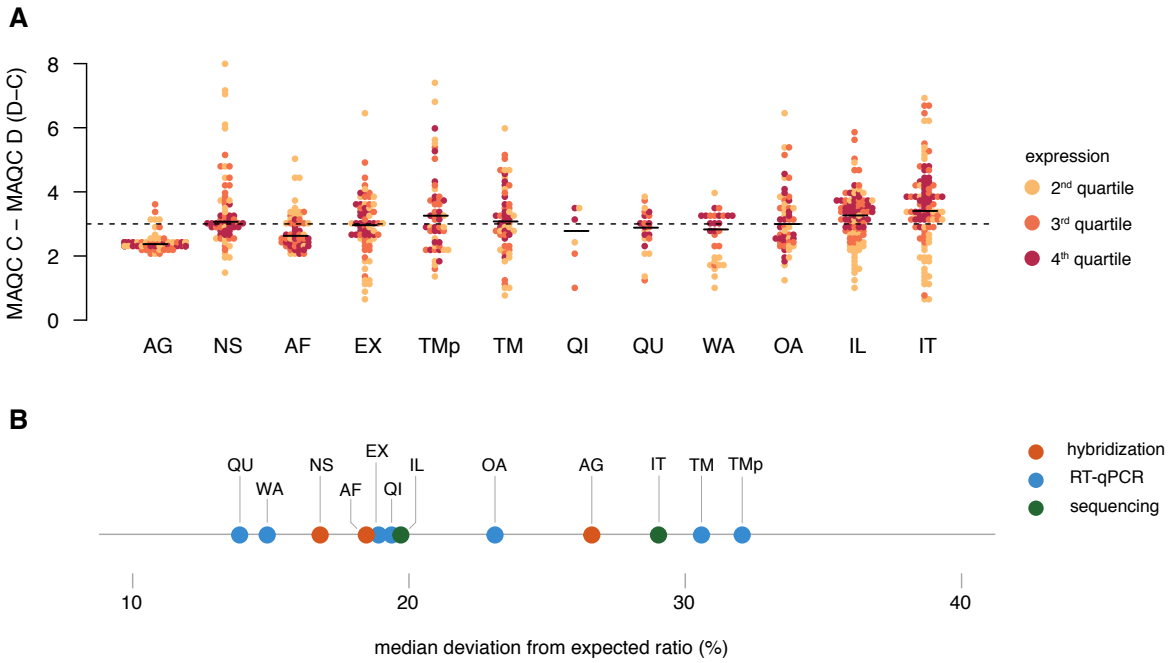
Relative expression levels for miRNAs in the common set and all miRNAs measured for each platform in miRQC A. Platforms for which the difference between both sets is significant are indicated by *.



Supplementary Figure 4

Metric and platform correlation matrix.

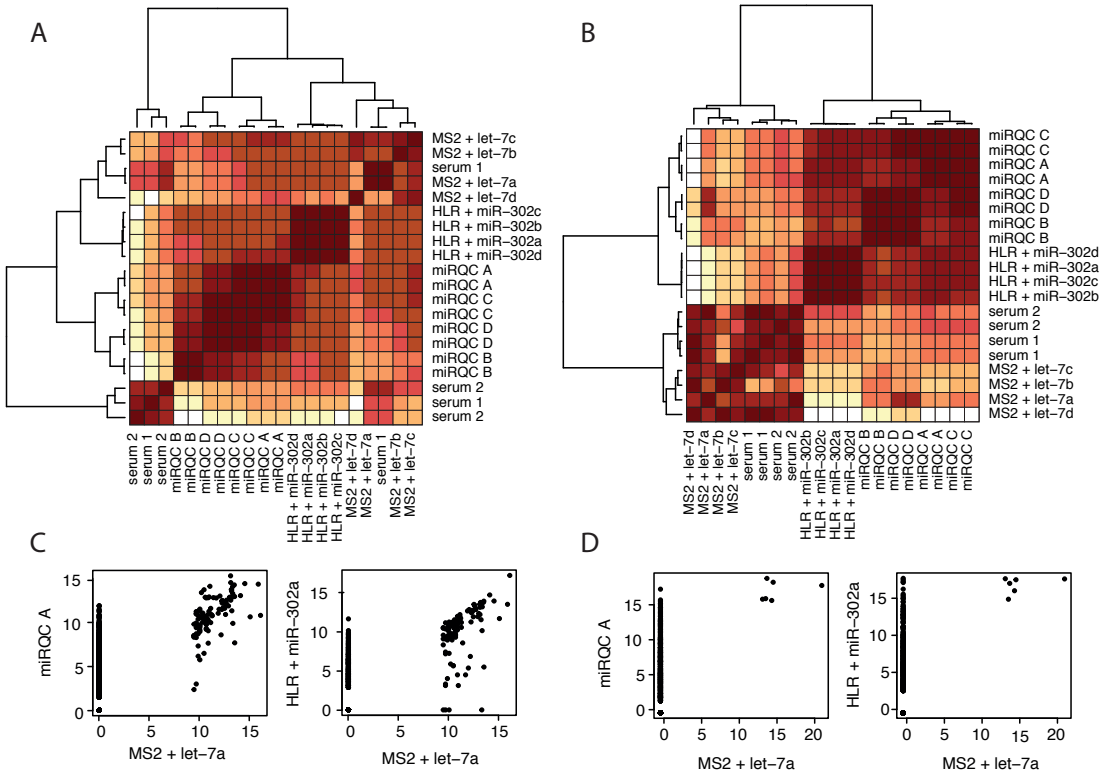
(A) Hierarchically clustered correlation matrix for different performance metrics. Where necessary, metrics were transformed so that a higher metric value corresponds to a better performance (see Online Methods). M1: reproducibility derived from ALC-value, M2: titration response (AUC-value), M3: accuracy derived from miRQC samples, M4: accuracy derived from low-copy templates. M5: detection rate in miRQC samples, M6: detection rate in serum RNA samples, M7: specificity derived from detection rate in MS2 phage RNA, M8: specificity (synthetic miRNAs) derived from number of mismatches with cross-reactivity, M9: specificity (synthetic miRNAs) derived from amount of cross-reactivity. Each square in the heatmap represents a Spearman's Rank correlation value between 2 metrics, calculated across all platforms. Positive correlations are color-coded red, negative correlations are color-coded blue. Correlations were calculated for all possible metric combinations and the resulting correlation matrix was hierarchically clustered. As this matrix is symmetric, only one half of the matrix is shown as a heatmap. This analysis reveals correlated performance metrics e.g. M4-M6: positive correlation; M1-M6: negative correlation. (B) Hierarchically clustered correlation matrix for different performance metrics. Each square in the heatmap represents a Spearman's Rank correlation value between 2 platforms, calculated across all metrics. This analysis reveals which platforms have similar or inverse performance across all measures e.g. OA-TM: similar performance; IL-TMp: inverse performance. Significant (Spearman's Rank p-value < 0.05) correlations in both heatmaps are marked by a white dot.



Supplementary Figure 5

Platform accuracy.

(A) Fold change (MAQC C/D or MAQC D/C) for MAQC A or MAQC B specific miRNAs. The dotted line indicates the expected 3-fold expression change. Individual miRNAs are color-coded according to their expression level. Only miRNAs with expression levels in the 2nd, 3rd or 4th quartile are included. (B) Median deviation from the expected C/D or D/C ratio for each platform taking into account only those miRNAs with expression levels in the 2nd, 3rd or 4th quartile. The median deviation is calculated as $(2^{\text{median}(|\log_2 3 - \log_2 |C/D||)} - 1) \times 100$.

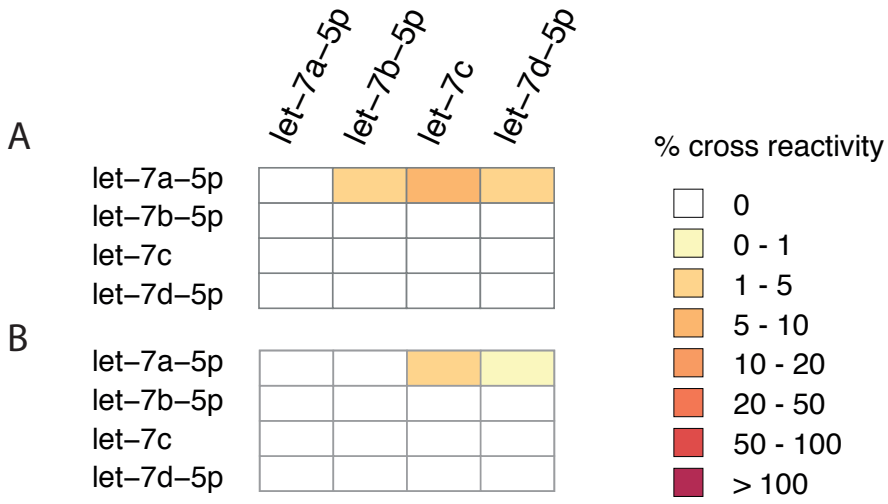


Supplementary Figure 6

Sample correlation clustering.

We evaluated whether contamination of the MS2 libraries might explain the unexpected high miRNA detection rate in MS2 for the IT platform. Sample correlation clustering using IT data reveals that the MS2 samples are clustering together with other high-RNA-content samples. Similar analysis with IL data did not reveal such correlation. In addition, only the abundant miRNAs from the high-RNA-content samples are detected in the IT MS2 samples, suggestive of contamination. Results obtained for the IT platform with the MS2 samples (Figure 4C-E) should therefore be interpreted with caution.

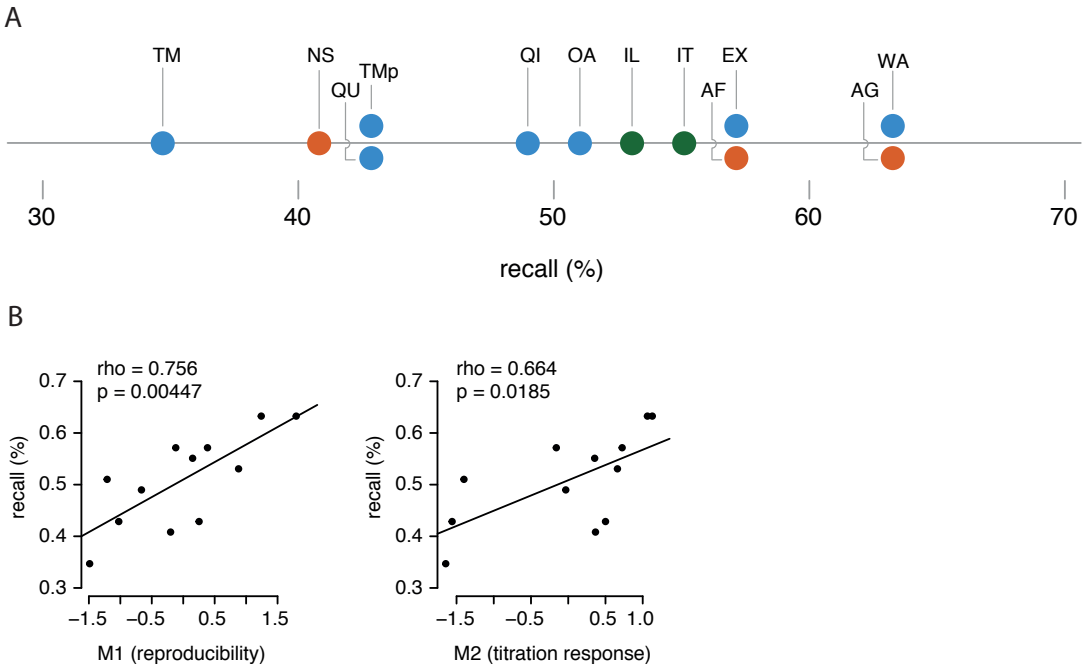
(A) Sample correlation clustering using miRNA expression data generated by the IT platform. (B) Sample correlation clustering using miRNA expression data generated by the IL platform. (C) Expression correlation between MS2 + let-7a and two representative high-RNA-content samples (miRQC A and HLR + miR-302a) for the IT platform. (D) Expression correlation between MS2 + let-7a and two representative high-RNA-content samples (miRQC A and HLR + miR-302a) for the IL platform.



Supplementary Figure 7

Cross reactivity between let-7 family members.

Synthetic miRNAs are indicated in columns, the corresponding measured miRNA signal in rows. Expression values for the perfect match are not indicated in the heatmap (white boxes on diagonal), increasing color intensity represents degree of cross-reactivity. Cross-reactivity was calculated relative to the exact match for each miRNA. (A) Data from IL platform, allowing 1 mismatch during read mapping. (B) Data from IL platform, allowing 0 mismatches during read mapping.



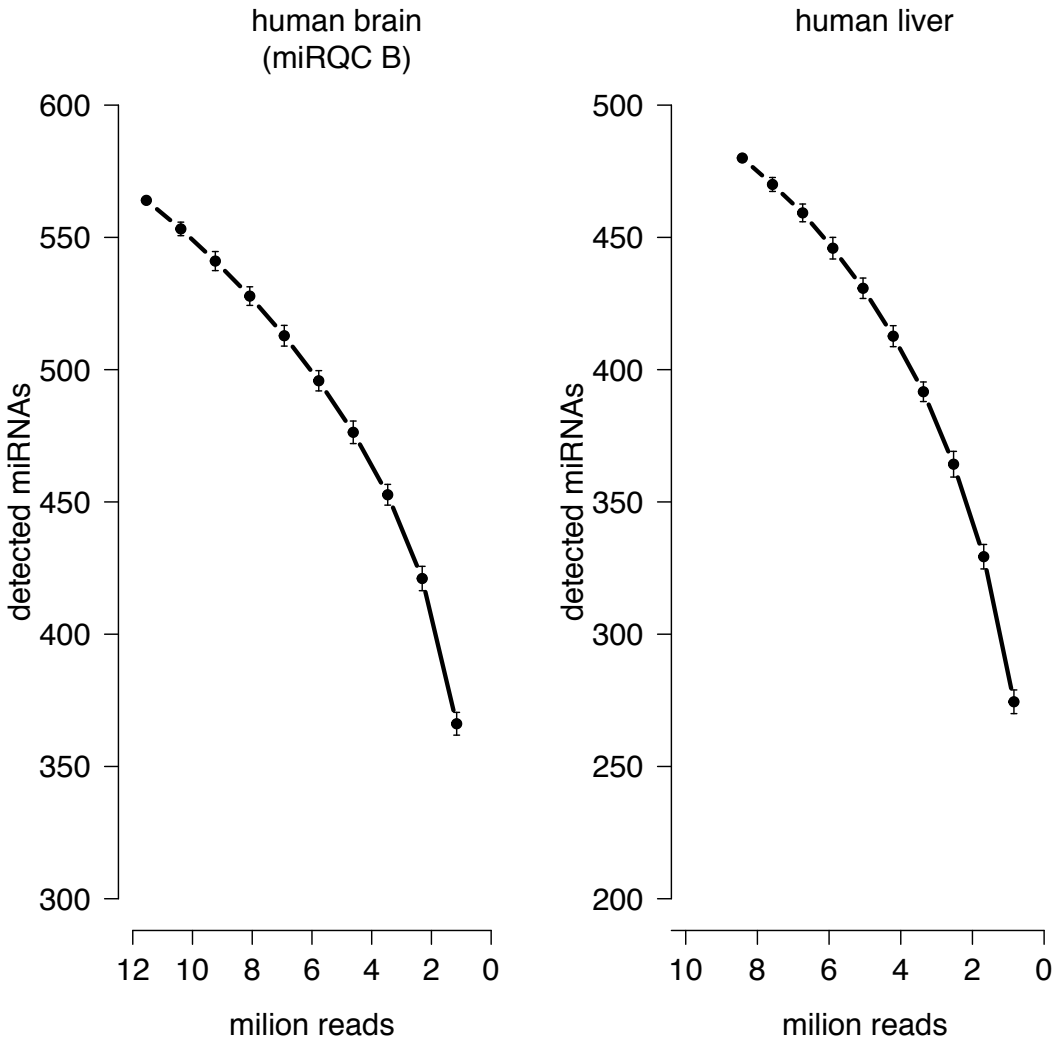
Supplementary Figure 8

Differential miRNA expression.

(A) Platform recall rates, defining the fraction of true differential miRNAs that are retrieved. Recall rate was calculated as:

recall = t_p / t_p + t_n

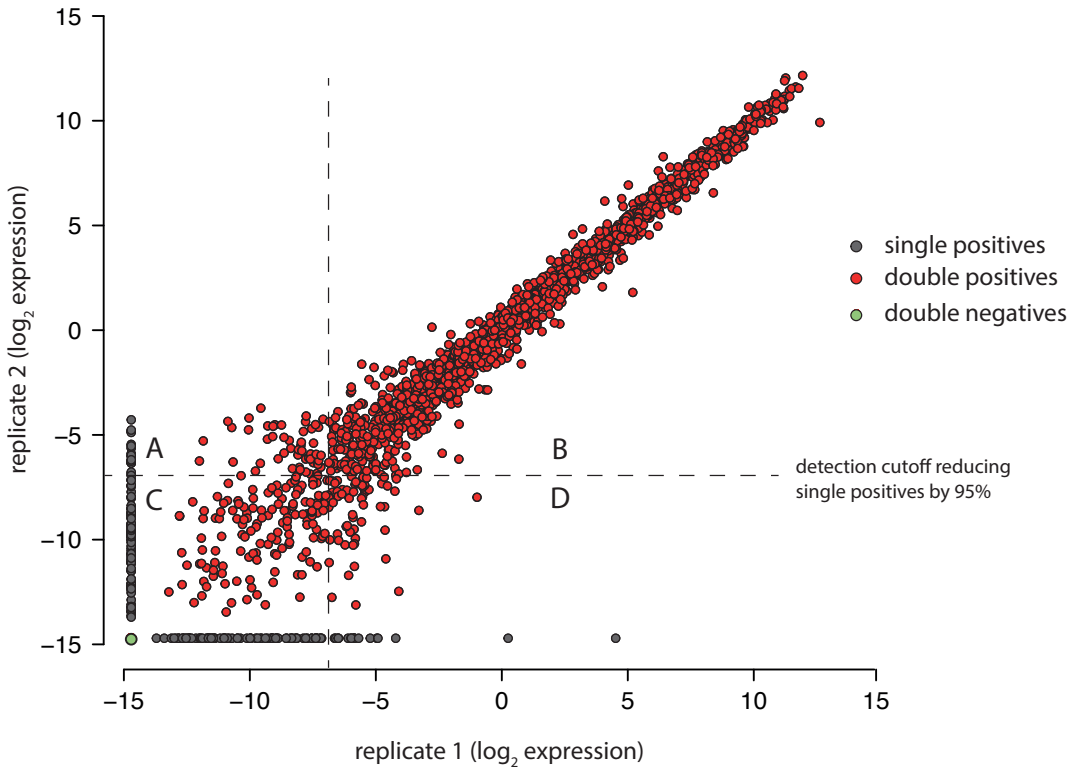
where t_p = true positives and f_n = false negatives. To determine the number of true positives and false negatives, a core set of truly differentially expressed miRNAs was defined as those miRNAs being identified as differentially expressed by at least 2 different technologies (i.e. PCR, hybridization and sequencing). To evaluate which platform performance metrics underlie differences in recall rates, recall rates were correlated to z-score transformed metrics (M1-M9, see online methods). Not unexpectedly, 2 metrics, M1: reproducibility and M2: titration response, showed a significant positive correlation (B). As such, platforms that have high reproducibility and good titration response show higher recall rates.



Supplementary Figure 9

Impact of sequencing depth on the number of detected miRNAs

Impact of sequencing depth on the number of detected miRNAs in 2 meta-samples consisting of replicates from human brain (miRQC B) or human liver. Meta-samples were generated by combining replicates of miRQC B (n = 2, Illumina platforms) and replicates of the liver sample (n = 4, Illumina platform) in order to have a sufficiently high number of reads for subsampling. From these samples, we repeatedly (n = 100) sampled 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% of the reads and calculated the average number of detected miRNAs. MiRNA detection rate was then plotted in function of the number of reads. As expected, detection rate drops when the total number of reads decreases. The fact that there is no plateau phase when considering all reads suggests that detection rate could be further increased by increasing read depth. As these are complex tissues, results for single cell types are likely to be very different as less complex samples would require fewer reads to reach detection saturation.



Supplementary Figure 10

Schematic representation illustrating how platform detection cutoff was established.

When plotting replicate expression values, 3 fractions are defined: single positives (miRNAs detected in only one of the replicates), double positives (miRNAs detected in both replicates) and double negatives (miRNAs detected in none of the replicates). As single positives represent unreliable measurements, detection cutoff should be as such that it maximally reduces this fraction and has a minimal effect on sporadic outliers caused by technical measurement errors. We therefore defined the detection cutoff as the expression level at which the fraction of single positives is reduced by 95% (dotted lines). Applying a detection cutoff divides the 2-dimensional space into four quadrants (A, B, C and D on the plot). Quadrants A and D represents those miRNAs that are expressed above detection cutoff in only ½ replicates, quadrant C represents miRNAs expressed below detection cutoff in both replicates and quadrant B represent miRNAs expression above detection cutoff in both replicates.

	miRBase accession	miRNA		miRBase accession	miRNA		miRBase accession	miRNA		miRBase accession	miRNA		miRBase accession	miRNA
1	MIMAT0000062	hsa-let-7a-5p	50	MIMAT0000259	hsa-miR-182-5p	99	MIMAT0000617	hsa-miR-200c-3p	148	MIMAT0001629	hsa-miR-329			
2	MIMAT0000063	hsa-let-7b-5p	51	MIMAT0000261	hsa-miR-183-5p	100	MIMAT0000646	hsa-miR-155-5p	149	MIMAT0001635	hsa-miR-452-5p			
3	MIMAT0000064	hsa-let-7c	52	MIMAT0000262	hsa-miR-187-3p	101	MIMAT0000680	hsa-miR-106b-5p	150	MIMAT0002170	hsa-miR-412			
4	MIMAT0000065	hsa-let-7d-5p	53	MIMAT0000266	hsa-miR-205-5p	102	MIMAT0000681	hsa-miR-29c-3p	151	MIMAT0002171	hsa-miR-410			
5	MIMAT0000066	hsa-let-7e-5p	54	MIMAT0000267	hsa-miR-210	103	MIMAT0000684	hsa-miR-302a-3p	152	MIMAT0002176	hsa-miR-485-3p			
6	MIMAT0000067	hsa-let-7f-5p	55	MIMAT0000268	hsa-miR-211-5p	104	MIMAT0000686	hsa-miR-34c-5p	153	MIMAT0002178	hsa-miR-487a			
7	MIMAT0000068	hsa-miR-15a-5p	56	MIMAT0000269	hsa-miR-212-3p	105	MIMAT0000688	hsa-miR-301a-3p	154	MIMAT0002805	hsa-miR-489			
8	MIMAT0000069	hsa-miR-16-5p	57	MIMAT0000271	hsa-miR-214-3p	106	MIMAT0000690	hsa-miR-296-5p	155	MIMAT0002809	hsa-miR-146b-5p			
9	MIMAT0000072	hsa-miR-18a-5p	58	MIMAT0000272	hsa-miR-215	107	MIMAT0000691	hsa-miR-130b-3p	156	MIMAT0002811	hsa-miR-202-3p			
10	MIMAT0000073	hsa-miR-19a-3p	59	MIMAT0000275	hsa-miR-218-5p	108	MIMAT0000703	hsa-miR-361-5p	157	MIMAT0002812	hsa-miR-492			
11	MIMAT0000074	hsa-miR-19b-3p	60	MIMAT0000414	hsa-let-7g-5p	109	MIMAT0000705	hsa-miR-362-5p	158	MIMAT0002814	hsa-miR-432-5p			
12	MIMAT0000075	hsa-miR-20a-5p	61	MIMAT0000416	hsa-miR-1	110	MIMAT0000707	hsa-miR-363-3p	159	MIMAT0002818	hsa-miR-496			
13	MIMAT0000076	hsa-miR-21-5p	62	MIMAT0000417	hsa-miR-15b-5p	111	MIMAT0000715	hsa-miR-302b-3p	160	MIMAT0002819	hsa-miR-193b-3p			
14	MIMAT0000077	hsa-miR-22-3p	63	MIMAT0000419	hsa-miR-27b-3p	112	MIMAT0000717	hsa-miR-302c-3p	161	MIMAT0002820	hsa-miR-497-5p			
15	MIMAT0000080	hsa-miR-24-3p	64	MIMAT0000420	hsa-miR-30b-5p	113	MIMAT0000718	hsa-miR-302d-3p	162	MIMAT0002875	hsa-miR-504			
16	MIMAT0000081	hsa-miR-25-3p	65	MIMAT0000421	hsa-miR-122-5p	114	MIMAT0000722	hsa-miR-370	163	MIMAT0002876	hsa-miR-505-3p			
17	MIMAT0000082	hsa-miR-26a-5p	66	MIMAT0000422	hsa-miR-124-3p	115	MIMAT0000723	hsa-miR-371a-3p	164	MIMAT0002888	hsa-miR-532-5p			
18	MIMAT0000083	hsa-miR-26b-5p	67	MIMAT0000425	hsa-miR-130a-3p	116	MIMAT0000724	hsa-miR-372	165	MIMAT0003161	hsa-miR-493-3p			
19	MIMAT0000084	hsa-miR-27a-3p	68	MIMAT0000426	hsa-miR-132-3p	117	MIMAT0000726	hsa-miR-373-3p	166	MIMAT0003164	hsa-miR-544a			
20	MIMAT0000085	hsa-miR-28-5p	69	MIMAT0000427	hsa-miR-133a	118	MIMAT0000727	hsa-miR-374a-5p	167	MIMAT0003180	hsa-miR-487b			
21	MIMAT0000086	hsa-miR-29a-3p	70	MIMAT0000428	hsa-miR-135a-5p	119	MIMAT0000728	hsa-miR-375	168	MIMAT0003233	hsa-miR-551b-3p			
22	MIMAT0000087	hsa-miR-30a-5p	71	MIMAT0000429	hsa-miR-137	120	MIMAT0000729	hsa-miR-376a-3p	169	MIMAT0003247	hsa-miR-582-5p			
23	MIMAT0000090	hsa-miR-32-5p	72	MIMAT0000430	hsa-miR-138-5p	121	MIMAT0000732	hsa-miR-378a-3p	170	MIMAT0003258	hsa-miR-590-5p			
24	MIMAT0000091	hsa-miR-33a-5p	73	MIMAT0000431	hsa-miR-140-5p	122	MIMAT0000733	hsa-miR-379-5p	171	MIMAT0003283	hsa-miR-615-3p			
25	MIMAT0000092	hsa-miR-92a-3p	74	MIMAT0000432	hsa-miR-141-3p	123	MIMAT0000735	hsa-miR-380-3p	172	MIMAT0003329	hsa-miR-411-5p			
26	MIMAT0000093	hsa-miR-93-5p	75	MIMAT0000433	hsa-miR-142-5p	124	MIMAT0000738	hsa-miR-383	173	MIMAT0003340	hsa-miR-542-5p			
27	MIMAT0000094	hsa-miR-95	76	MIMAT0000434	hsa-miR-142-3p	125	MIMAT0000753	hsa-miR-342-3p	174	MIMAT0003393	hsa-miR-425-5p			
28	MIMAT0000095	hsa-miR-96-5p	77	MIMAT0000435	hsa-miR-143-3p	126	MIMAT0000754	hsa-miR-337-3p	175	MIMAT0003885	hsa-miR-454-3p			
29	MIMAT0000097	hsa-miR-99a-5p	78	MIMAT0000436	hsa-miR-144-3p	127	MIMAT0000757	hsa-miR-151a-3p	176	MIMAT0004502	hsa-miR-28-3p			
30	MIMAT0000098	hsa-miR-100-5p	79	MIMAT0000437	hsa-miR-145-5p	128	MIMAT0000758	hsa-miR-135b-5p	177	MIMAT0004597	hsa-miR-140-3p			
31	MIMAT0000099	hsa-miR-101-3p	80	MIMAT0000438	hsa-miR-152	129	MIMAT0000759	hsa-miR-148b-3p	178	MIMAT0004604	hsa-miR-127-5p			
32	MIMAT0000100	hsa-miR-29b-3p	81	MIMAT0000439	hsa-miR-153	130	MIMAT0000761	hsa-miR-324-5p	179	MIMAT0004605	hsa-miR-129-2-3p			
33	MIMAT0000101	hsa-miR-103a-3p	82	MIMAT0000440	hsa-miR-191-5p	131	MIMAT0000763	hsa-miR-338-3p	180	MIMAT0004613	hsa-miR-188-3p			
34	MIMAT0000102	hsa-miR-105-5p	83	MIMAT0000441	hsa-miR-9-5p	132	MIMAT0000764	hsa-miR-339-5p	181	MIMAT0004614	hsa-miR-193a-5p			
35	MIMAT0000222	hsa-miR-192-5p	84	MIMAT0000445	hsa-miR-126-3p	133	MIMAT0000765	hsa-miR-335-5p	182	MIMAT0004675	hsa-miR-219-2-3p			
36	MIMAT0000227	hsa-miR-197-3p	85	MIMAT0000446	hsa-miR-127-3p	134	MIMAT0000770	hsa-miR-133b	183	MIMAT0004682	hsa-miR-361-3p			
37	MIMAT0000231	hsa-miR-199a-5p	86	MIMAT0000447	hsa-miR-134	135	MIMAT0000771	hsa-miR-325	184	MIMAT0004683	hsa-miR-362-5p			
38	MIMAT0000242	hsa-miR-129-5p	87	MIMAT0000449	hsa-miR-146a-5p	136	MIMAT0000772	hsa-miR-345-5p	185	MIMAT0004692	hsa-miR-340-5p			
39	MIMAT0000243	hsa-miR-148a-3p	88	MIMAT0000450	hsa-miR-149-5p	137	MIMAT0000773	hsa-miR-346	186	MIMAT0004697	hsa-miR-151a-5p			
40	MIMAT0000244	hsa-miR-30c-5p	89	MIMAT0000451	hsa-miR-150-5p	138	MIMAT0001075	hsa-miR-384	187	MIMAT0004700	hsa-miR-331-5p			
41	MIMAT0000245	hsa-miR-30d-5p	90	MIMAT0000454	hsa-miR-184	139	MIMAT0001080	hsa-miR-196b-5p	188	MIMAT0004763	hsa-miR-488-3p			
42	MIMAT0000250	hsa-miR-139-5p	91	MIMAT0000455	hsa-miR-185-5p	140	MIMAT0001339	hsa-miR-422a	189	MIMAT0004780	hsa-miR-532-3p			
43	MIMAT0000251	hsa-miR-147a	92	MIMAT0000456	hsa-miR-186-5p	141	MIMAT0001341	hsa-miR-424-5p	190	MIMAT0004784	hsa-miR-455-3p			
44	MIMAT0000252	hsa-miR-7-5p	93	MIMAT0000458	hsa-miR-190a	142	MIMAT0001412	hsa-miR-18b-5p	191	MIMAT0004901	hsa-miR-298			
45	MIMAT0000253	hsa-miR-10a-5p	94	MIMAT0000459	hsa-miR-193a-3p	143	MIMAT0001532	hsa-miR-448	192	MIMAT0004929	hsa-miR-190b			
46	MIMAT0000254	hsa-miR-10b-5p	95	MIMAT0000460	hsa-miR-194-5p	144	MIMAT0001536	hsa-miR-429	193	MIMAT0004945	hsa-miR-744-5p			
47	MIMAT0000255	hsa-miR-34a-5p	96	MIMAT0000461	hsa-miR-195-5p	145	MIMAT0001541	hsa-miR-449a	194	MIMAT0004953	hsa-miR-873-5p			
48	MIMAT0000256	hsa-miR-181a-5p	97	MIMAT0000462	hsa-miR-206	146	MIMAT0001545	hsa-miR-450a-5p	195	MIMAT0004954	hsa-miR-543			
49	MIMAT0000258	hsa-miR-181c-5p	98	MIMAT0000510	hsa-miR-320a	147	MIMAT0001627	hsa-miR-433	196	MIMAT0004958	hsa-miR-301b			

Supplementary Table 1

Set of 196 miRNAs measured by all platforms.

Sample ID	Total read count	Index sequence
1	4363123	ATCACG
2	5946829	CGATGT
3	6635467	TTAGGC
4	4988449	TGACCA
5	6879391	ACAGTG
6	7256512	GCCAAT
7	5633009	CAGATC
8	5849163	ACTTGA
9	1815607	CTTGTA
10	2100107	AGTCAA
11	2067222	AGTTCC
12	2454150	ATGTCA
13	1485	CCGTCC
14	937	GTAGAG
15	1447	GTCCGC
16	389	GTGAAA
17	6176	GTGGCC
18	4828	GTTTCG
19	12438	CGTACG
20	10636	GAGTGG

Supplementary Table 2

Number of reads and index sequence per sample for the Illumina small-RNA sequencing platform.

Sample ID	Total read count
1	1585560
2	1209244
3	2843049
4	1984837
5	645355
6	1630677
7	850776
8	1457397
9	2260966
10	2798307
11	3359332
12	1483355
13	522921
14	2698264
15	4180486
16	1002584
17	522921
18	598730
19	312408
20	545402

Supplementary Table 3

Number of reads per sample for the Ion-Torrent small-RNA sequencing platform.

Supplementary Note 1

Exiqon qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Platform detection cutoff was set at 37 cycles according to the manufacturer's instructions.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 536 unique double positives were detected while the percentage of single positives was 10.58 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 15.3 $\log_2$-units.

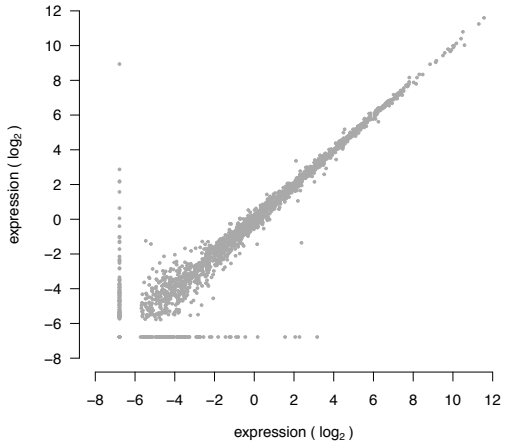


Figure 1: Replicate miRNA expression correlation plot.

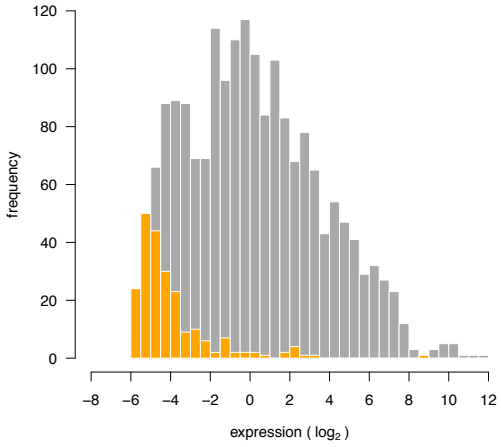
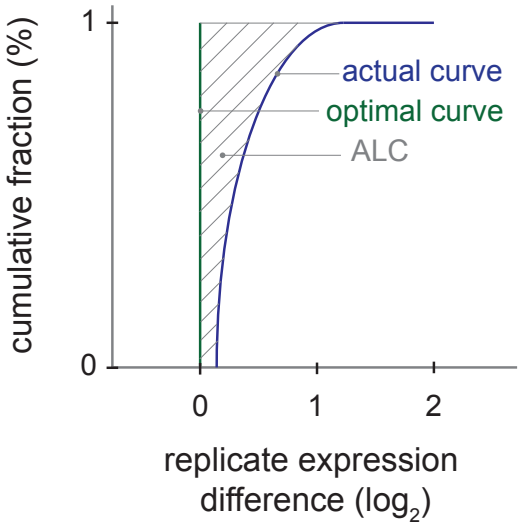


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.289, equivalent to a mean 1.222 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



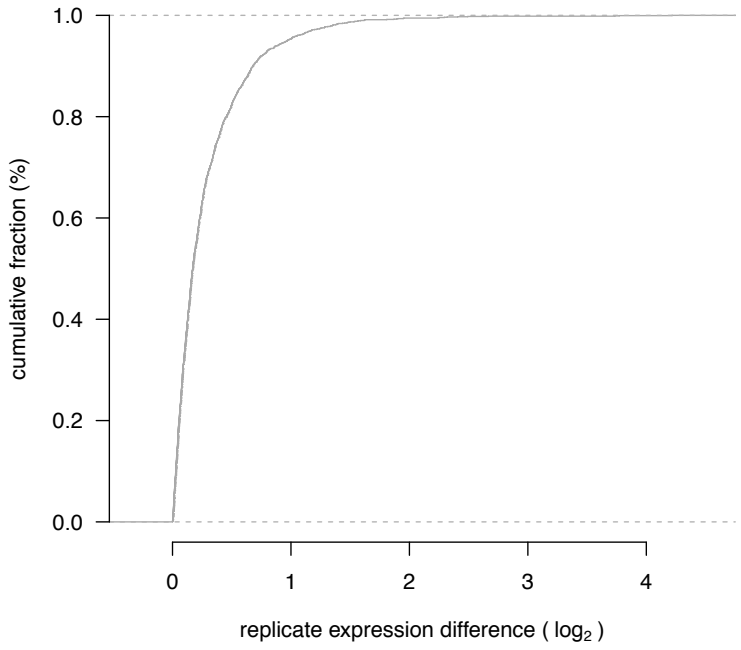
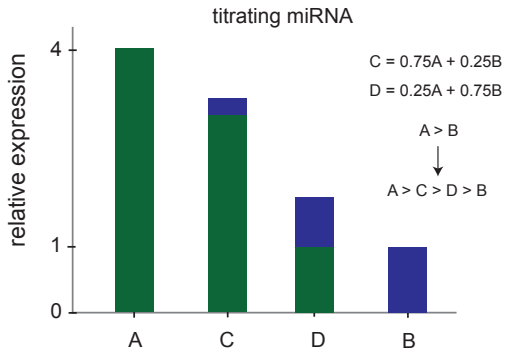


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

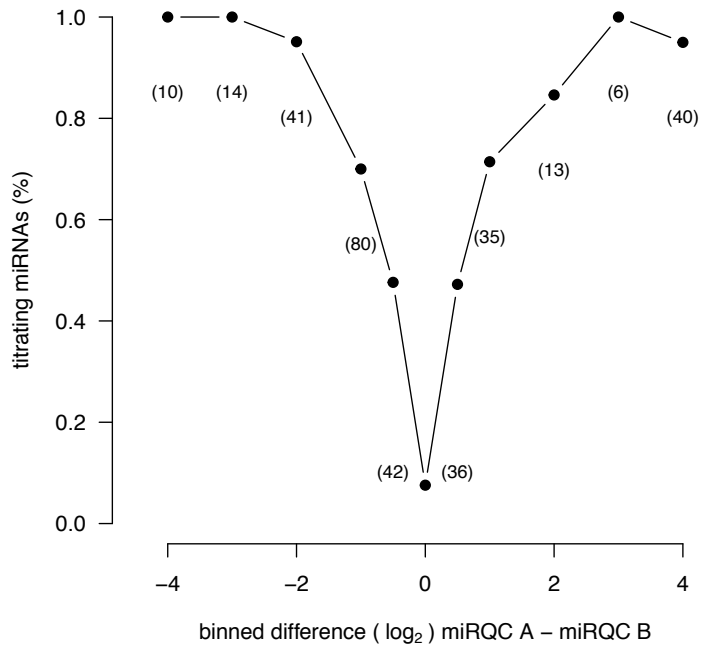
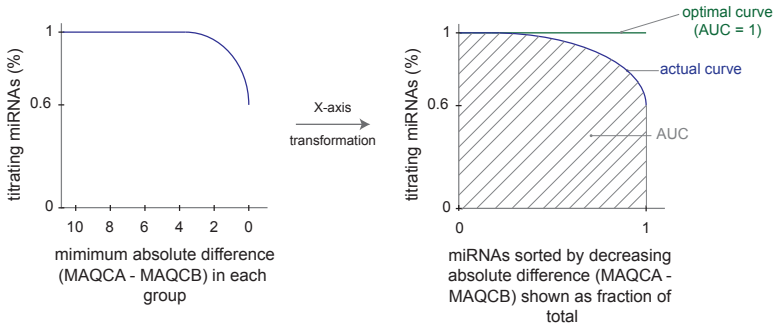


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.858.

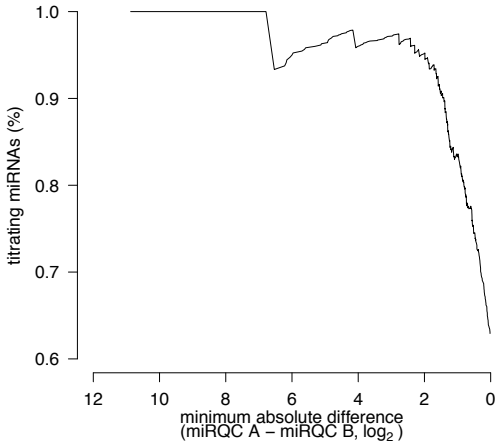


Figure 5: Titration response in function of absolute difference.

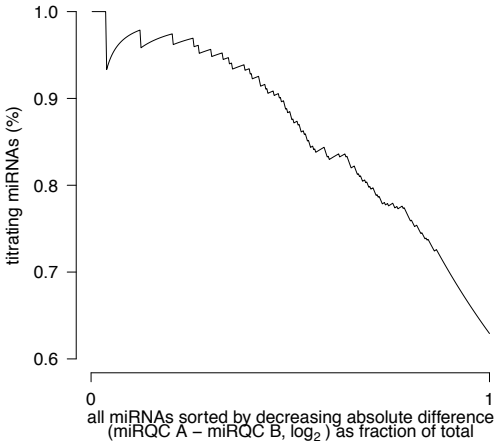


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

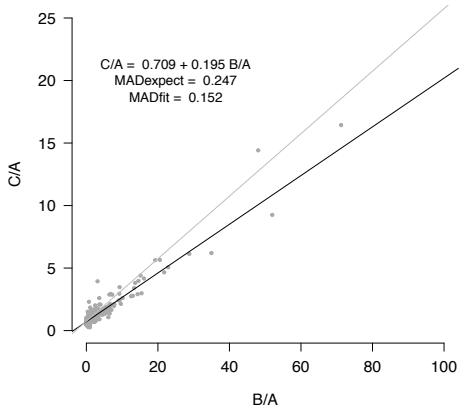


Figure 7: Titration response linearity.

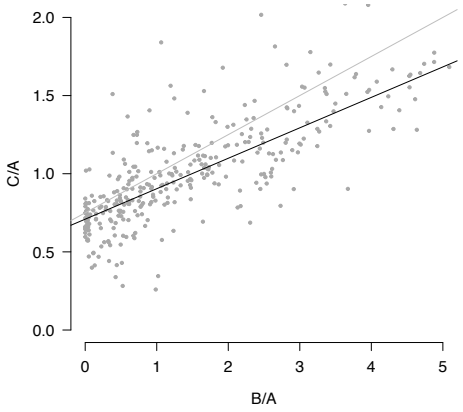


Figure 8: Figure 7 zoomed.

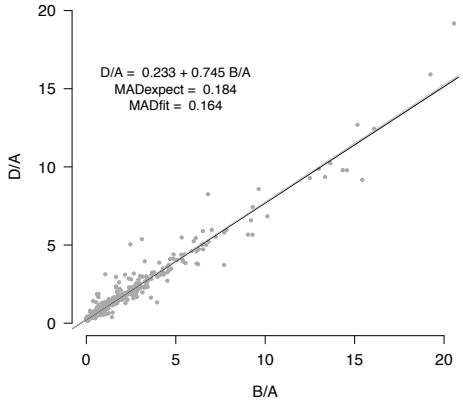


Figure 9: Titration response linearity.

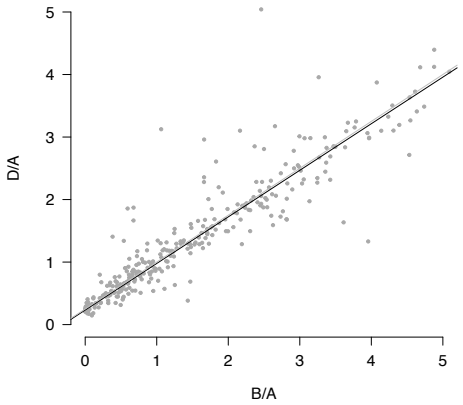


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 0% of all off-target combinations with a median relative cross-reactivity of 0%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	0.0	0.0
let-7b-5p	0.0	100.0	0.0	0.0
let-7c	0.0	0.0	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 8.02%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
5.2	12.1	6.9	7.8

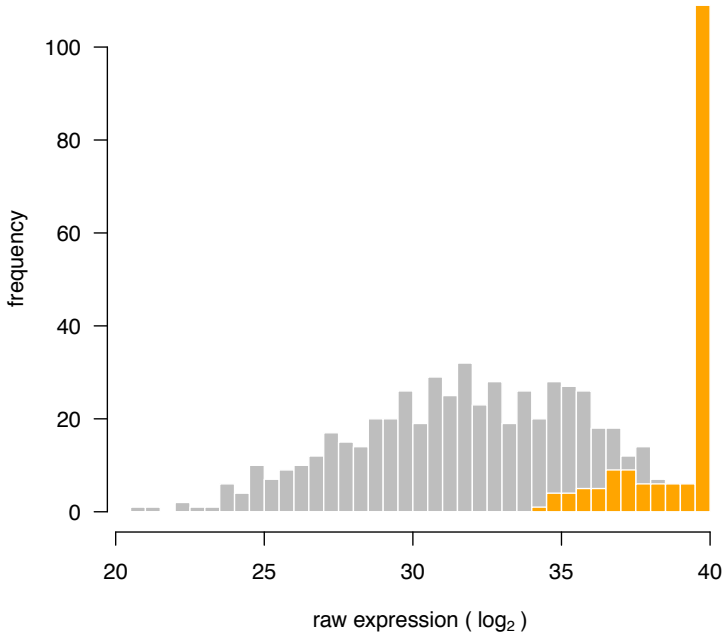


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 75 miRNAs were detected in all 4 samples while 132 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.653, equivalent to a mean 1.572 fold replicate expression difference.

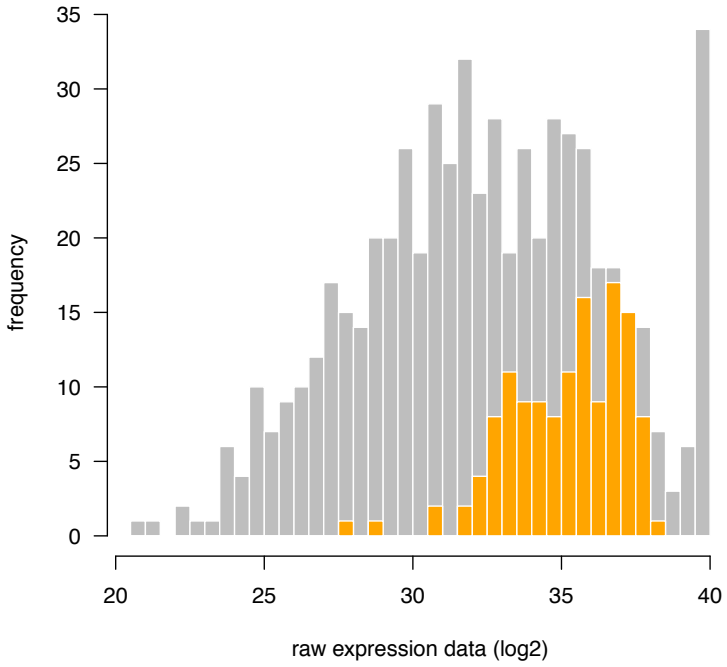


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

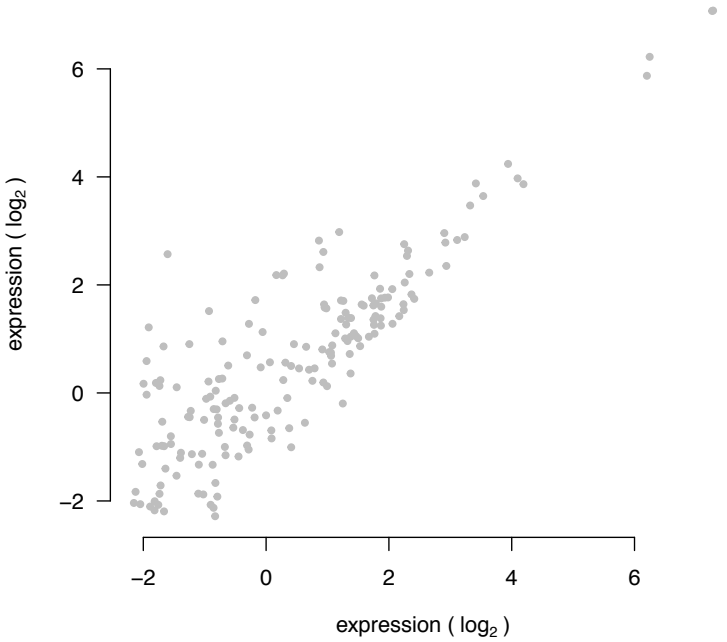


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 25 miRNAs were significantly upregulated and 24 miRNAs were significantly downregulated.

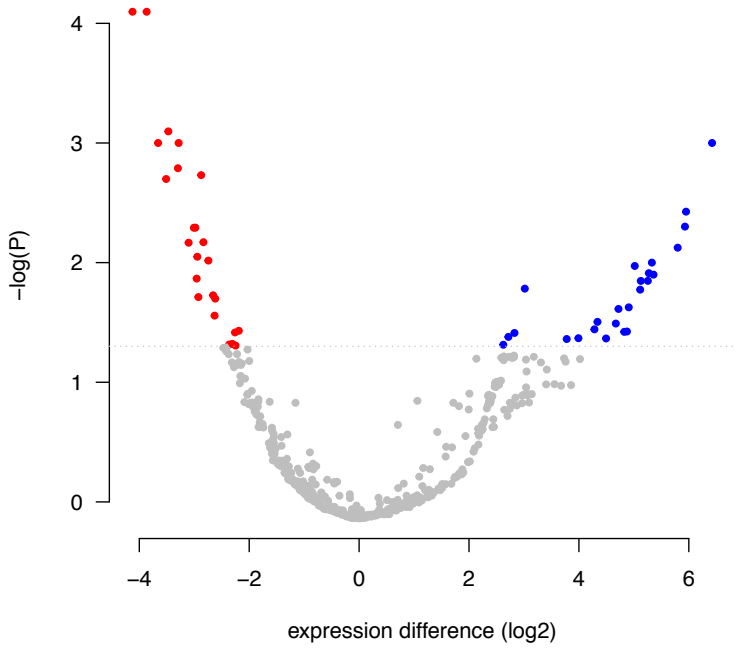


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	536
	fraction single positives (%)	10.58
	expression range (log2-units)	15.3
	ALC	0.289
titration	AUC titration response	0.858
	MADexpect (D/A)	0.247
	MADfit (D/A)	0.152
	MADexpect (C/A)	0.184
	MADfit (C/A)	0.164
specificity	off-target combinations with cross reactivity (%)	0
	median relative cross-reactivity (%)	0
non-template control	positive miRNAs (%)	8.02
serum miRNAs	detected miRNAs	132
differential expression	significant down	24
	significant up	25

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 2

Life Technologies OpenArray qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Platform detection cutoff was set at 25 cycles according to the manufacturer's instructions.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 394 unique double positives were detected while the percentage of single positives was 7.14 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 13.5 $\log_2$-units.

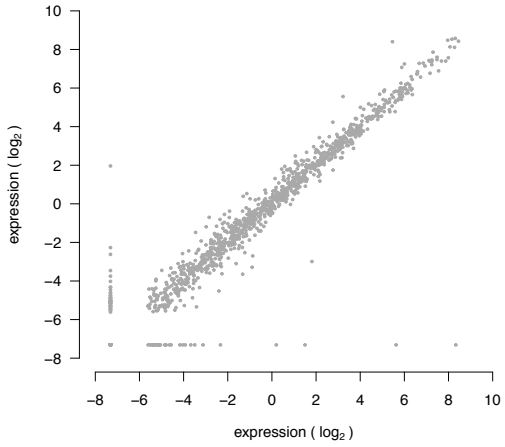


Figure 1: Replicate miRNA expression correlation plot.

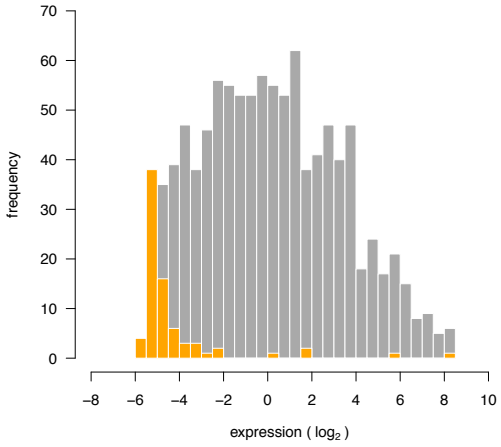
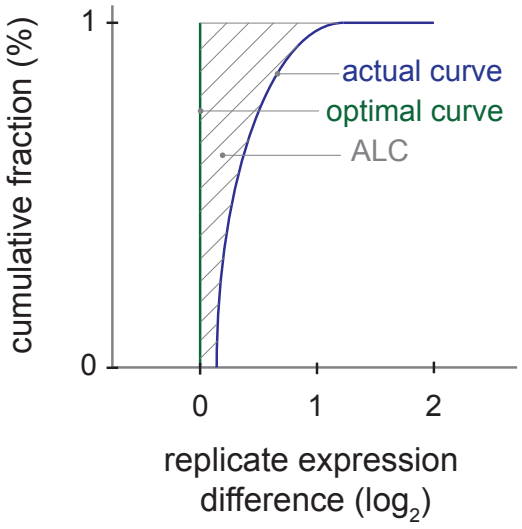


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.361, equivalent to a mean 1.284 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



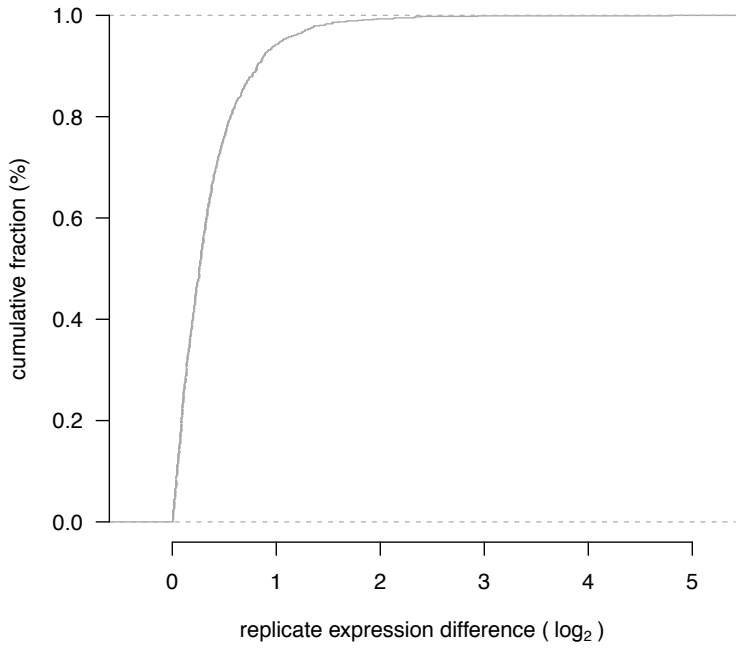
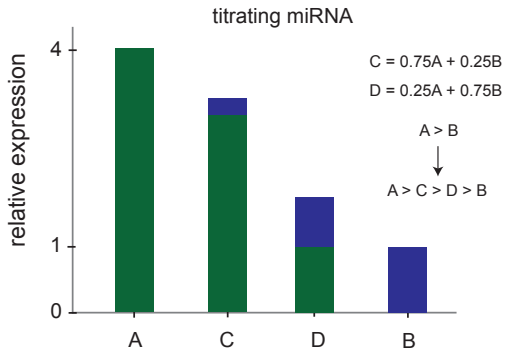


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

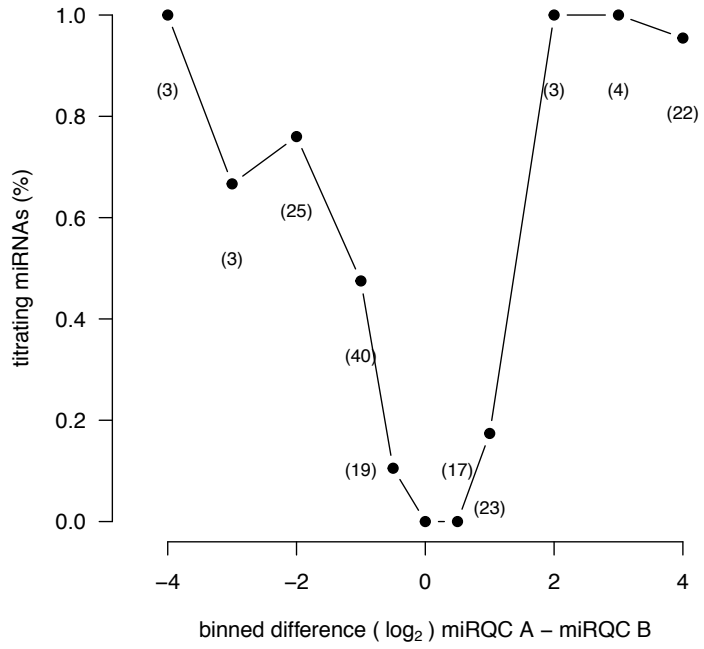
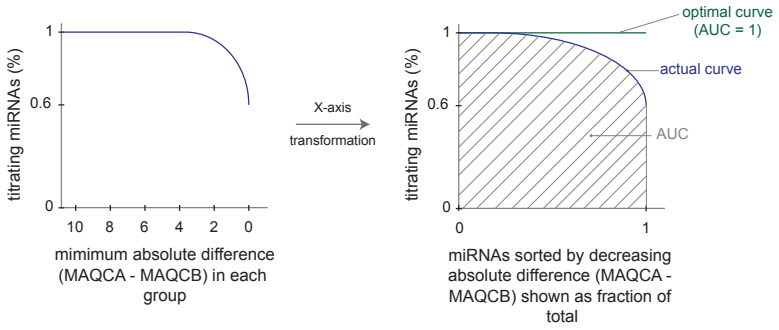


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.684.

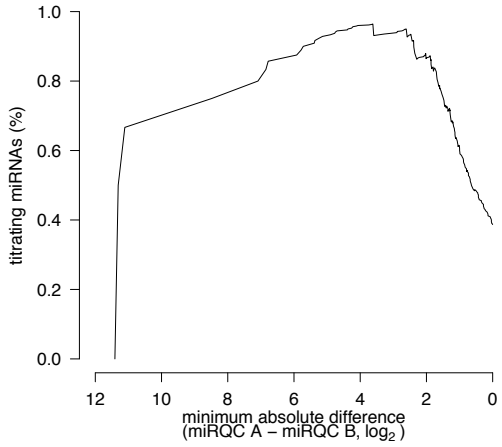


Figure 5: Titration response in function of absolute difference.

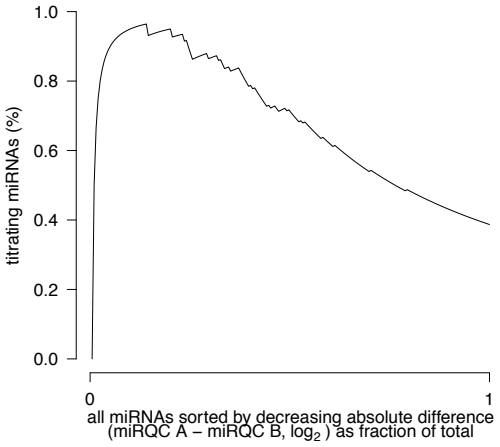


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

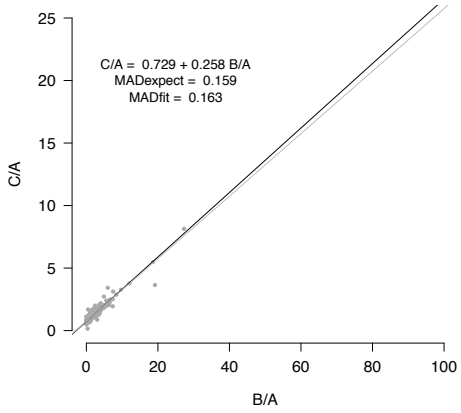


Figure 7: Titration response linearity.

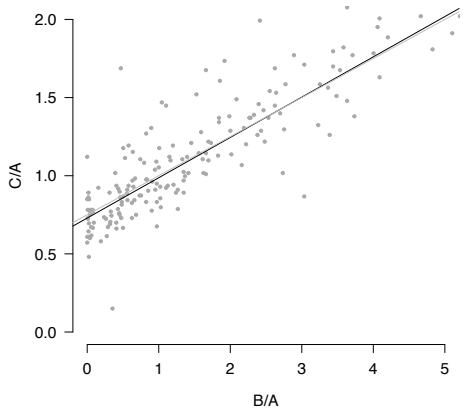


Figure 8: Figure 7 zoomed.

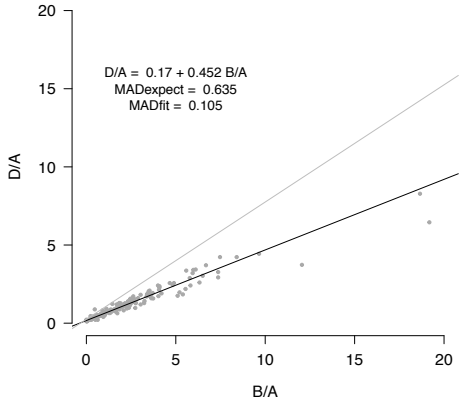


Figure 9: Titration response linearity.

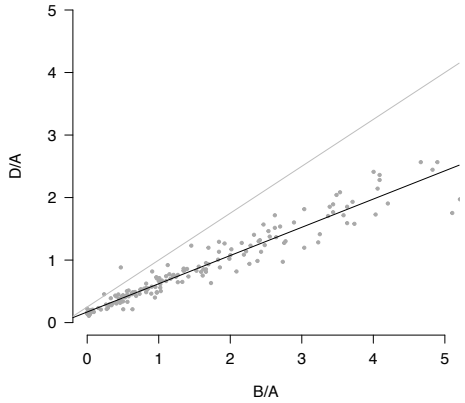


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 4.2% of all off-target combinations with a median relative cross-reactivity of 85.3%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	0.0	0.0
let-7b-5p	0.0	100.0	85.3	0.0
let-7c	0.0	0.0	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 6.79%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 14) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
24.6	1.3	0.8	0.5

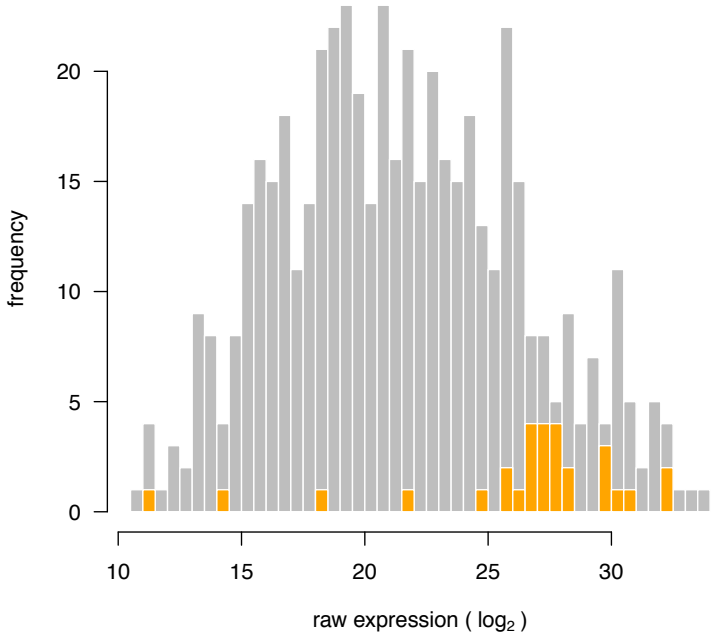


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 44 miRNAs were detected in all 4 samples while 61 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.667, equivalent to a mean 1.588 fold replicate expression difference.

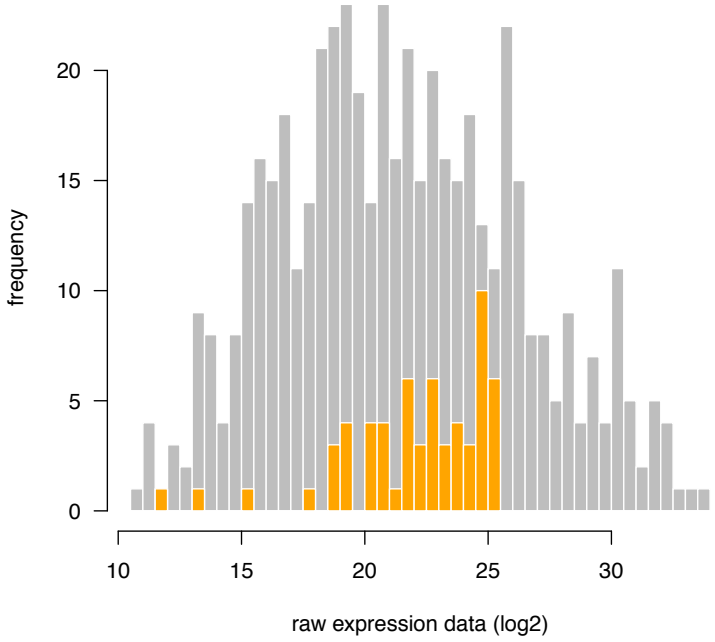


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

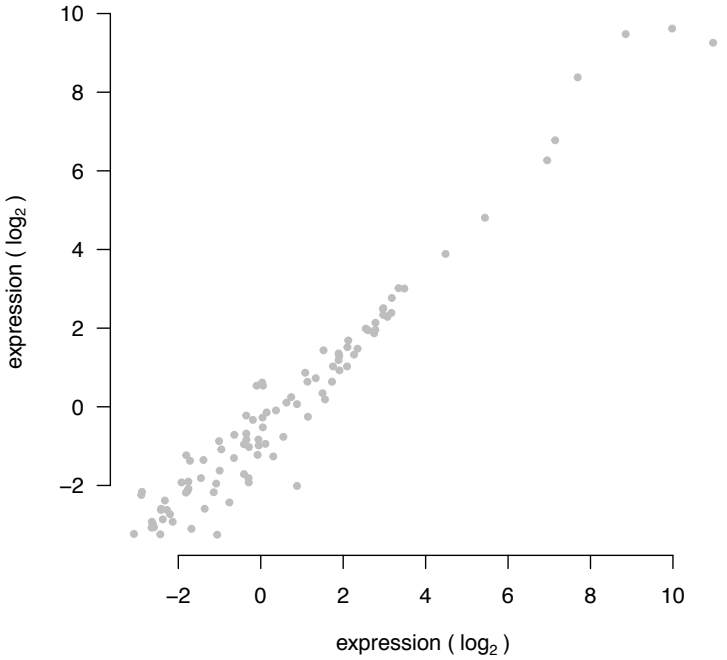


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 18 miRNAs were significantly upregulated and 13 miRNAs were significantly downregulated.

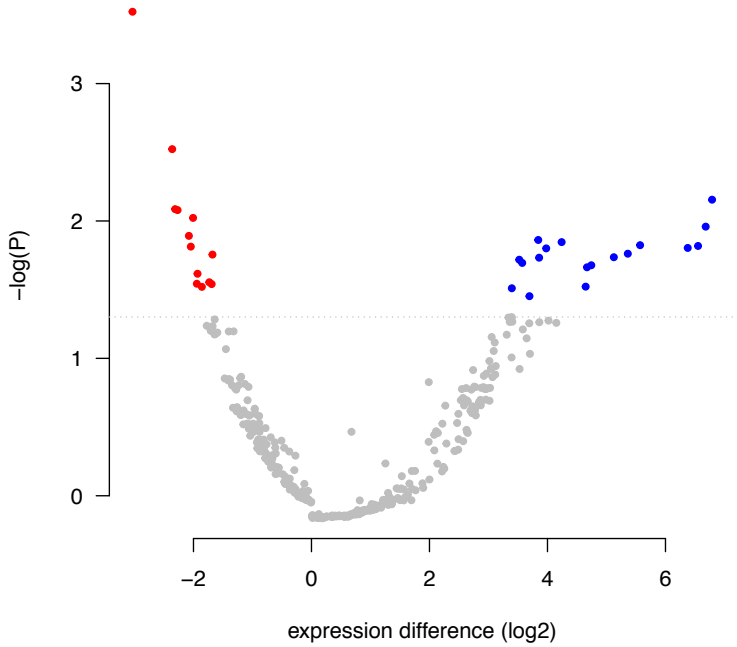


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	394
	fraction single positives (%)	7.14
	expression range (log2-units)	13.5
	ALC	0.361
titration	AUC titration response	0.684
	MADexpect (D/A)	0.159
	MADfit (D/A)	0.163
	MADexpect (C/A)	0.635
	MADfit (C/A)	0.105
specificity	off-target combinations with cross reactivity (%)	4.2
	median relative cross-reactivity (%)	85.3
non-template control	positive miRNAs (%)	6.79
serum miRNAs	detected miRNAs	61
differential expression	significant down	13
	significant up	18

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 3

Life Technologies Taqman Cards qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA.

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Platform detection cutoff was set at 35 cycles according to the manufacturer's instructions.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 436 unique double positives were detected while the percentage of single positives was 13.18 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 13.1 $\log_2$-units.

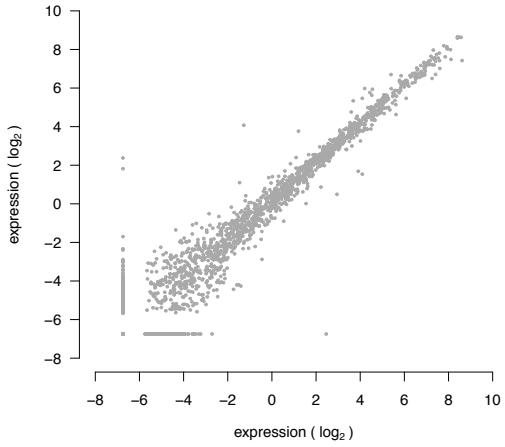


Figure 1: Replicate miRNA expression correlation plot.

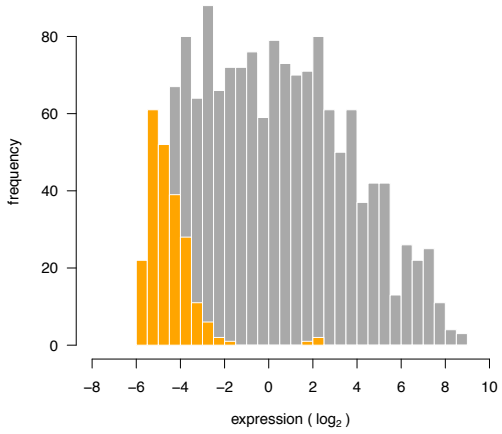


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.487, equivalent to a mean 1.402 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014

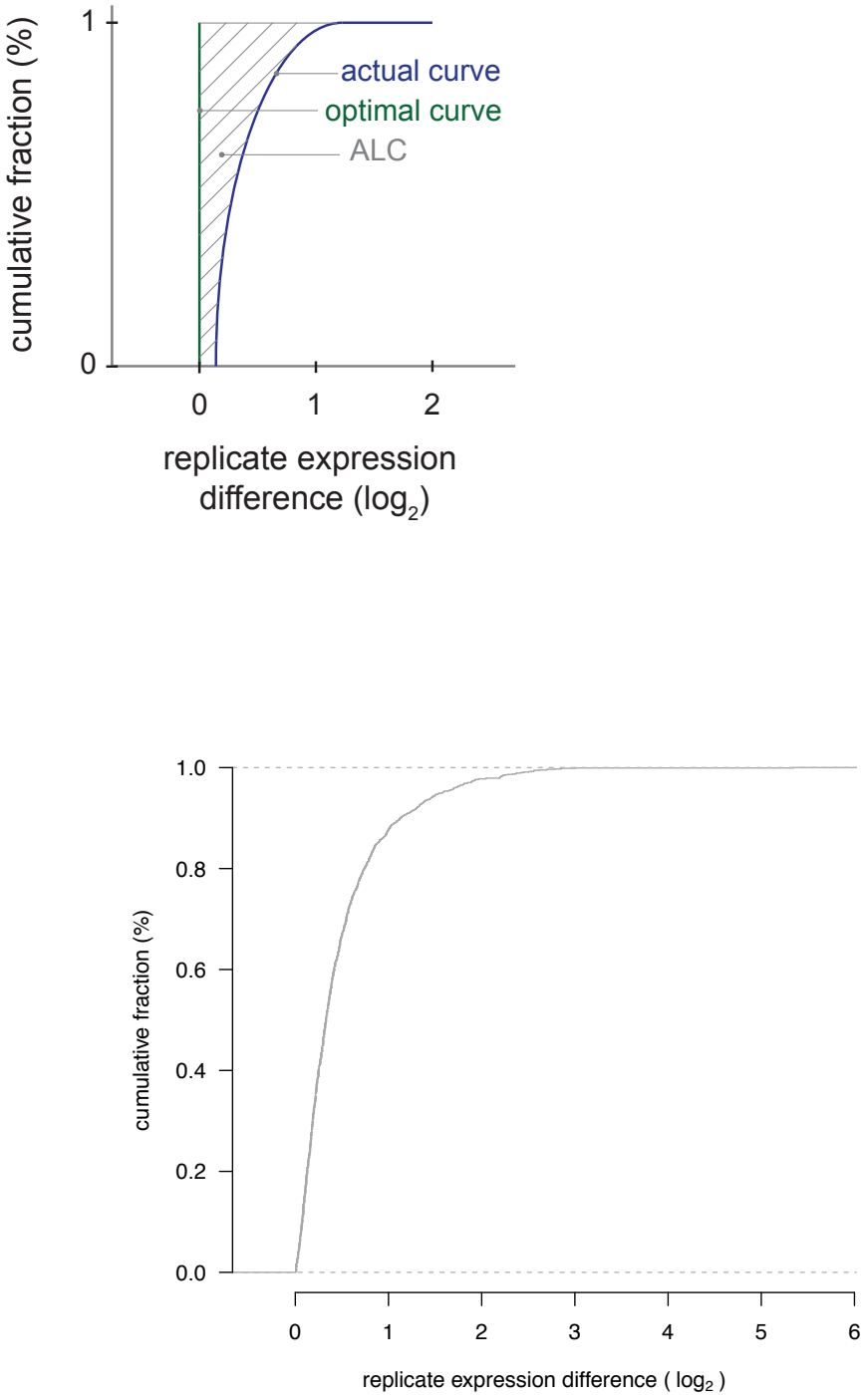
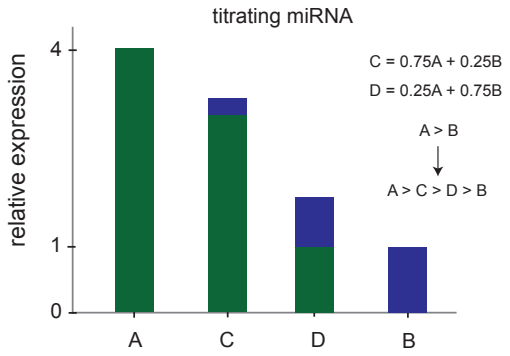


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

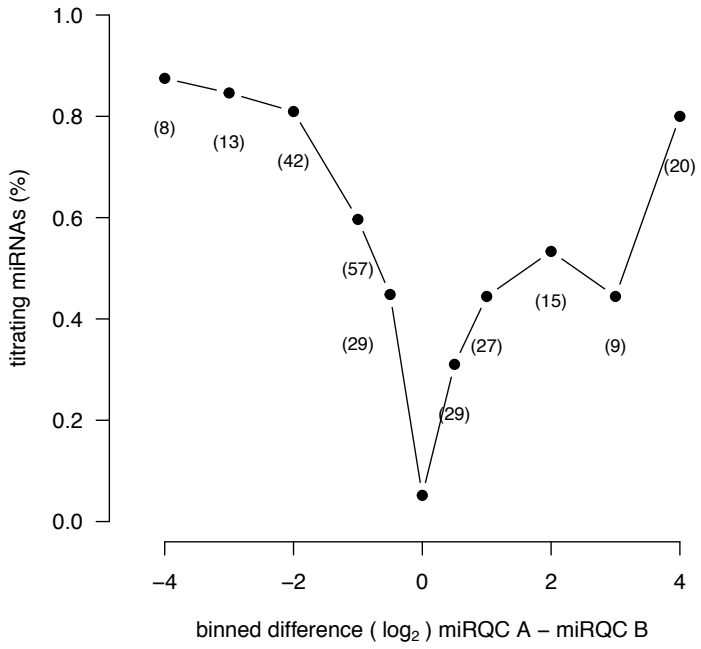
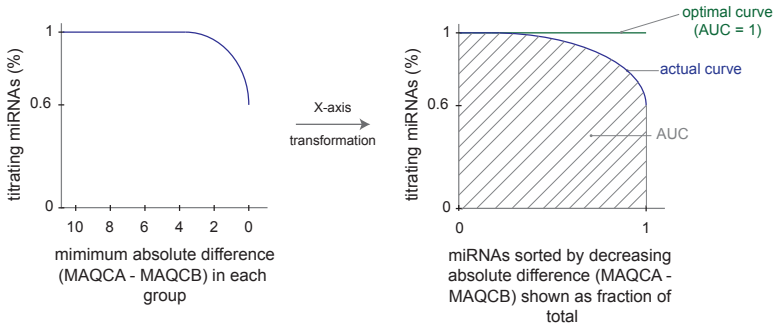


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.695.

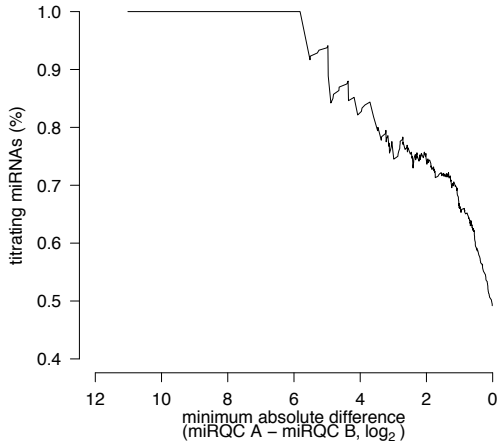


Figure 5: Titration response in function of absolute difference.

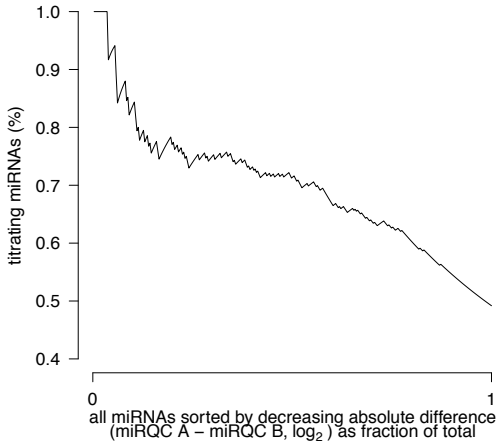


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

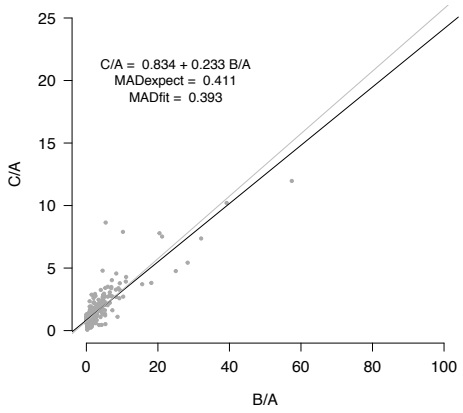


Figure 7: Titration response linearity.

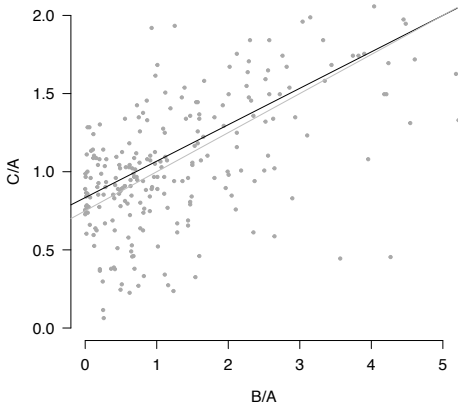


Figure 8: Figure 7 zoomed.

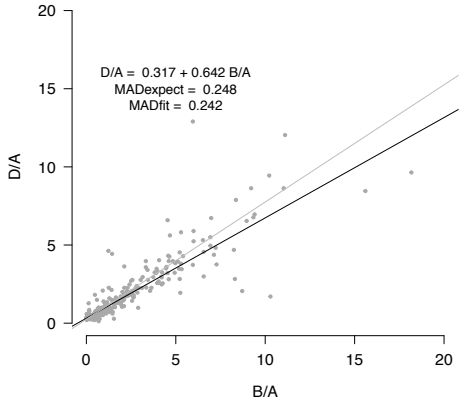


Figure 9: Titration response linearity.

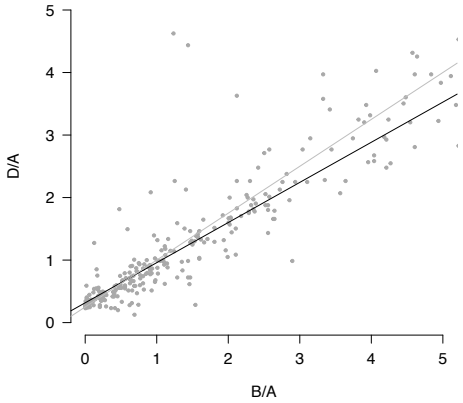


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 4.2% of all off-target combinations with a median relative cross-reactivity of 92%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	0.0	0.0	0.0	0.0
let-7b-5p	0.0	100.0	92.0	0.0
let-7c	0.0	0.0	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 3.21%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 14) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
3.2	4.1	2.3	3.2

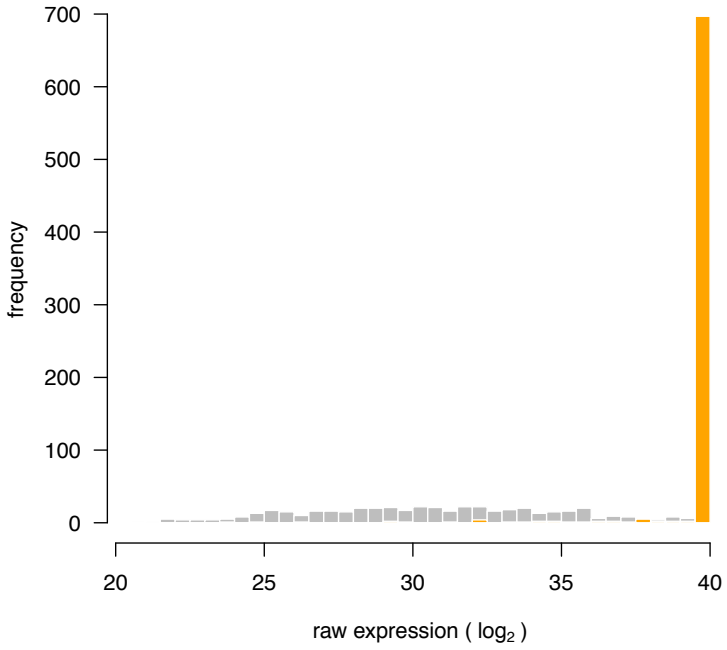


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 25 miRNAs were significantly upregulated and 13 miRNAs were significantly downregulated.

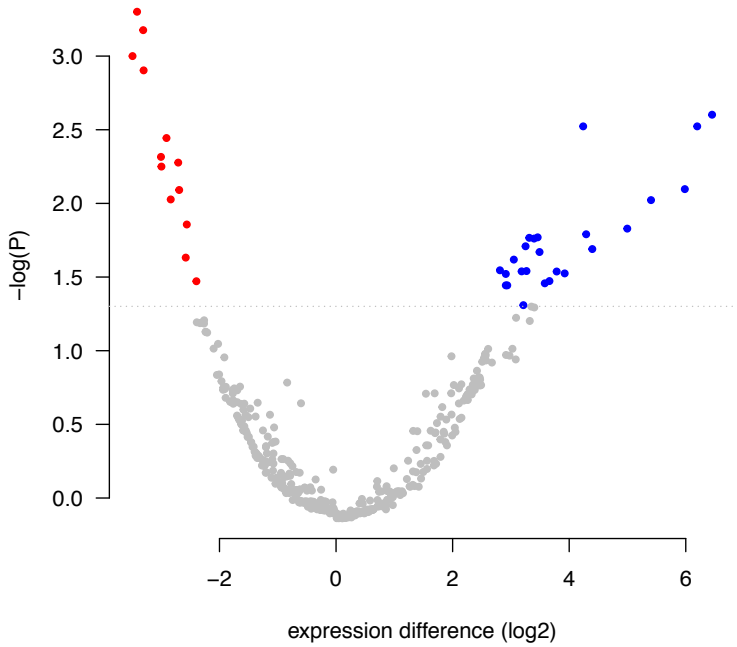


Figure 12: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

8 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	436
	fraction single positives (%)	13.18
	expression range (log2-units)	13.1
	ALC	0.487
titration	AUC titration response	0.695
	MADexpect (D/A)	0.411
	MADfit (D/A)	0.393
	MADexpect (C/A)	0.248
	MADfit (C/A)	0.242
specificity	off-target combinations with cross reactivity (%)	4.2
	median relative cross-reactivity (%)	92
non-template control	positive miRNAs (%)	3.21
serum miRNAs	detected miRNAs	NA
differential expression	significant down	13
	significant up	25

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 4

Life Technologies TaqMan qPCR with pre-amplification

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Platform detection cutoff was set at 32 cycles according to the manufacturer's instructions.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 491 unique double positives were detected while the percentage of single positives was 6.89 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 16 $\log_2$-units.

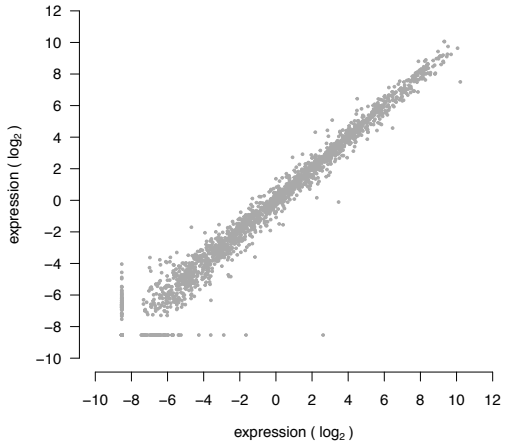


Figure 1: Replicate miRNA expression correlation plot.

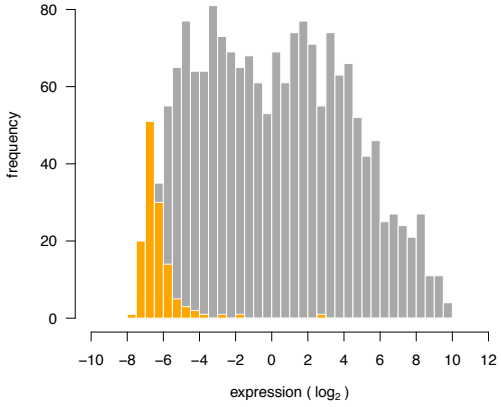
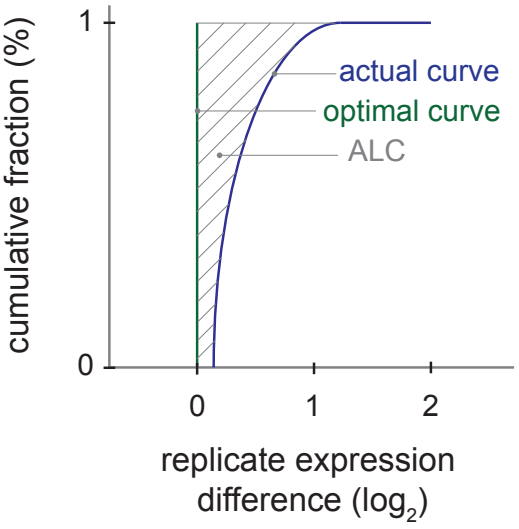


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.41, equivalent to a mean 1.329 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



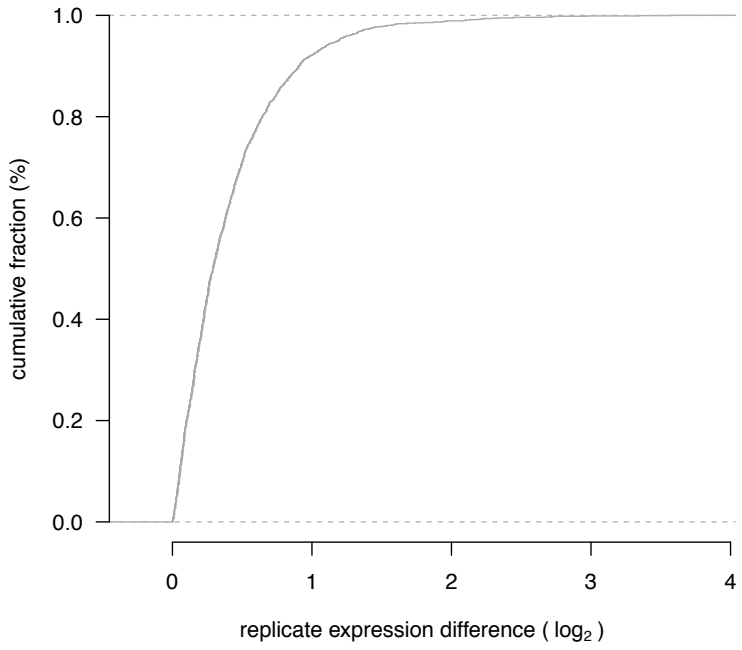
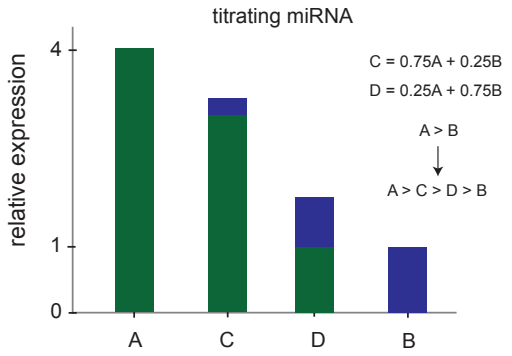


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

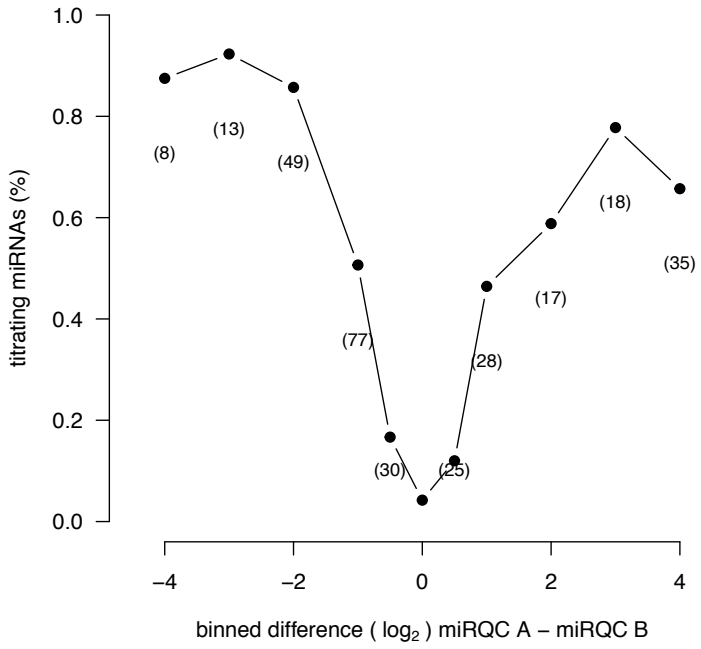
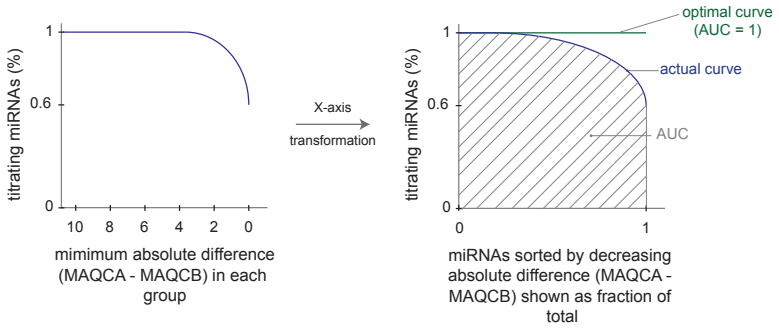


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.665.

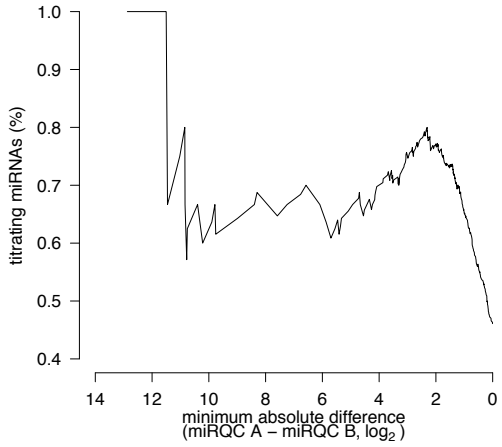


Figure 5: Titration response in function of absolute difference.

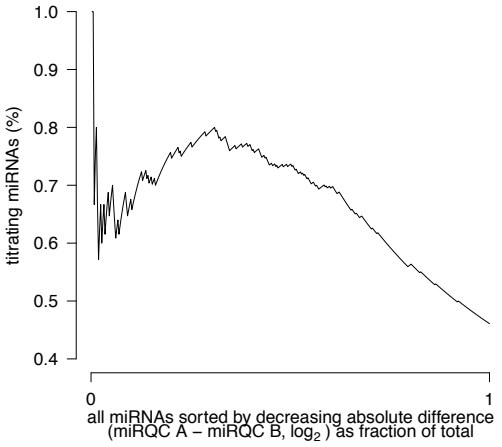


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

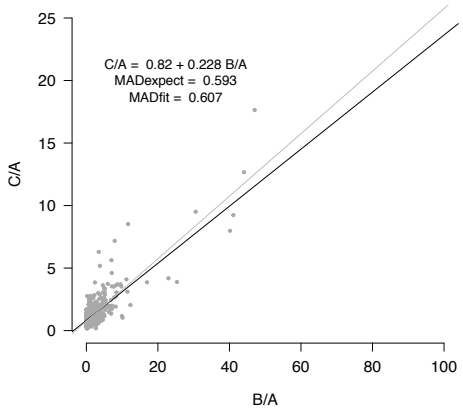


Figure 7: Titration response linearity.

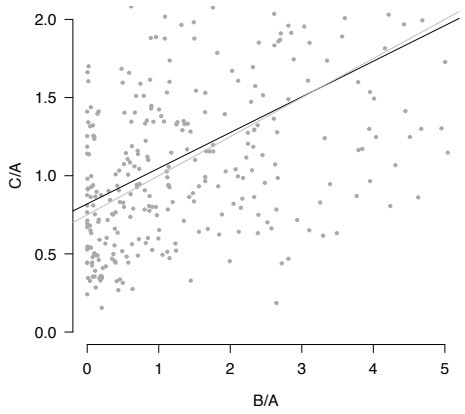


Figure 8: Figure 7 zoomed.

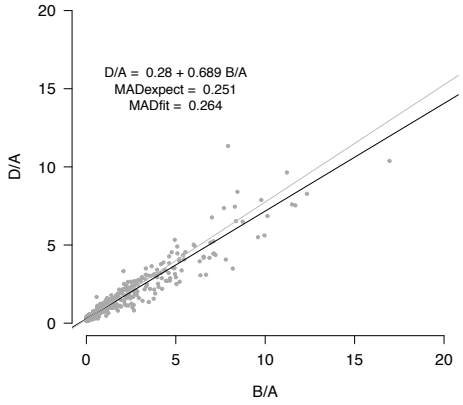


Figure 9: Titration response linearity.

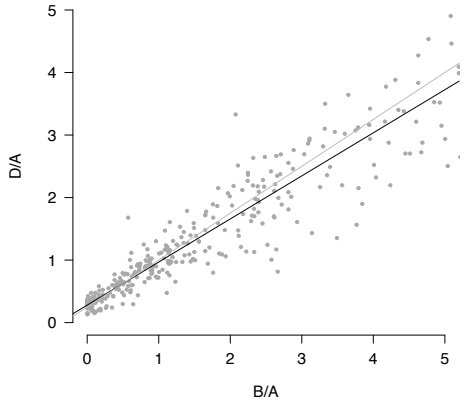


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 41.7% of all off-target combinations with a median relative cross-reactivity of 0.5%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.3
miR-302b-3p	0.0	100.0	0.0	0.1
miR-302c-3p	0.2	0.0	100.0	0.8
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	0.0	0.0
let-7b-5p	0.0	100.0	232.9	1.5
let-7c	0.0	0.4	100.0	0.5
let-7d-5p	9.2	0.0	0.6	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 2.7%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
2.9	2.2	2.9	2.9

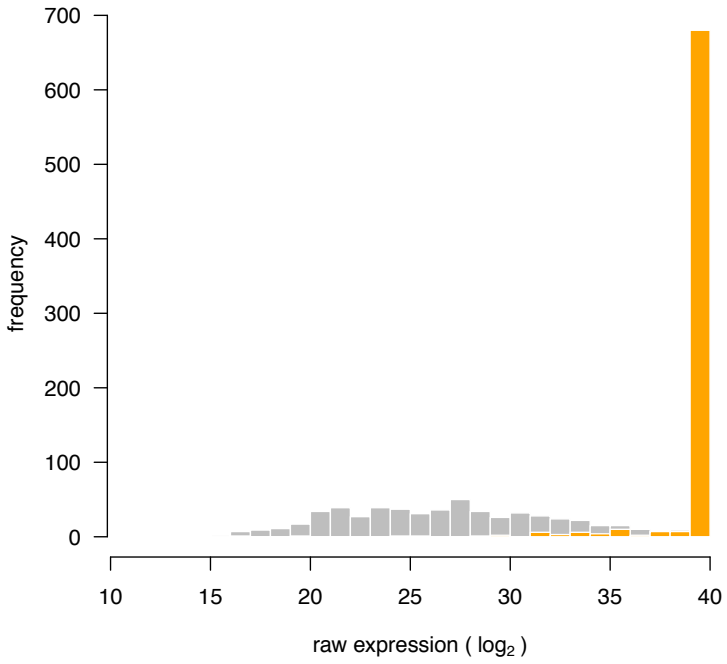


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 105 miRNAs were detected in all 4 samples while 137 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.536, equivalent to a mean 1.45 fold replicate expression difference.

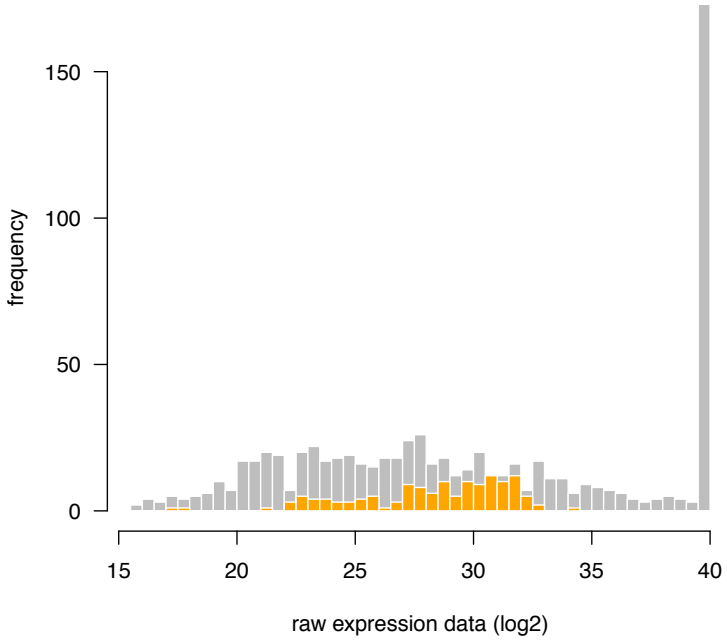


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

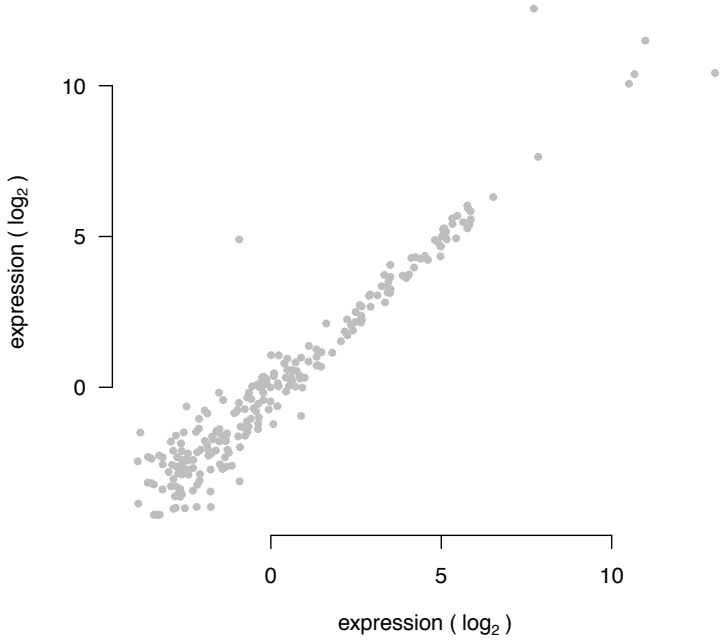


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 28 miRNAs were significantly upregulated and 26 miRNAs were significantly downregulated.

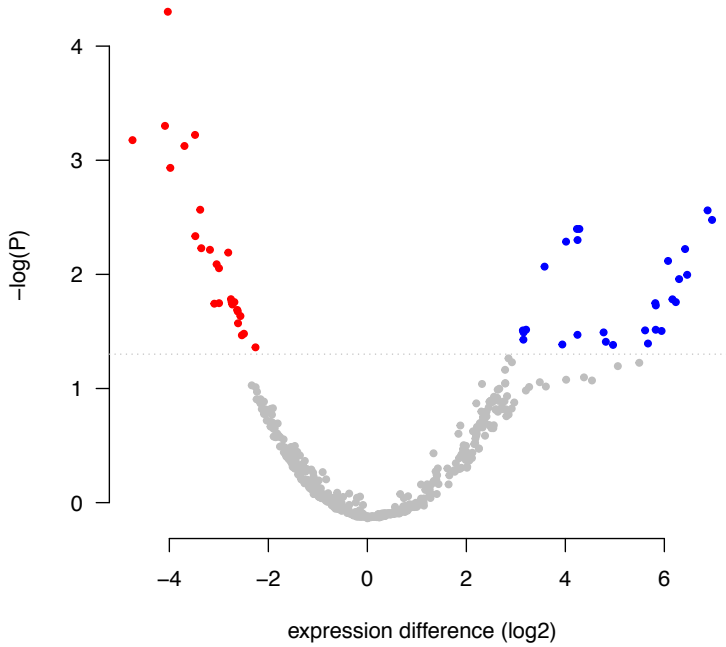


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	491
	fraction single positives (%)	6.89
	expression range (log2-units)	16
	ALC	0.41
titration	AUC titration response	0.665
	MADexpect (D/A)	0.593
	MADfit (D/A)	0.607
	MADexpect (C/A)	0.251
	MADfit (C/A)	0.264
specificity	off-target combinations with cross reactivity (%)	41.7
	median relative cross-reactivity (%)	0.5
non-template control	positive miRNAs (%)	2.7
serum miRNAs	detected miRNAs	137
differential expression	significant down	26
	significant up	28

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 5

Qiagen qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 34.35 cycles.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 363 unique double positives were detected while the percentage of single positives was 2.7 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 17 $\log_2$-units.

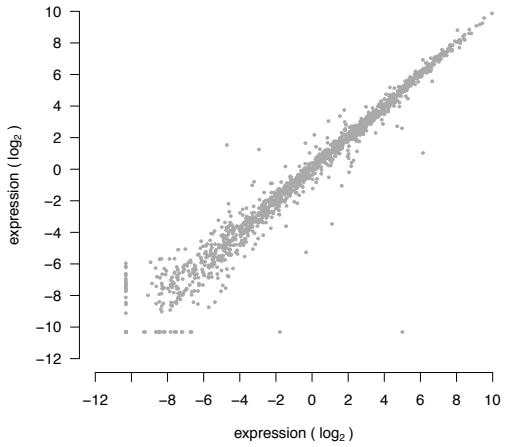


Figure 1: Replicate miRNA expression correlation plot.

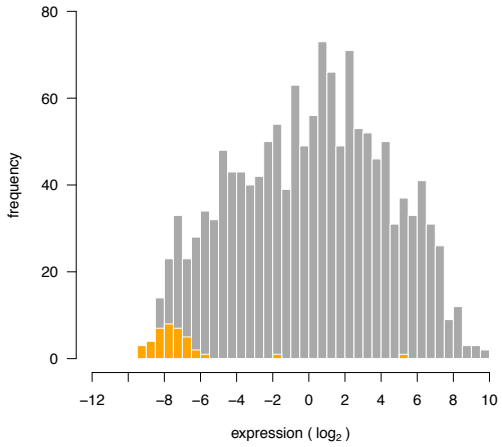
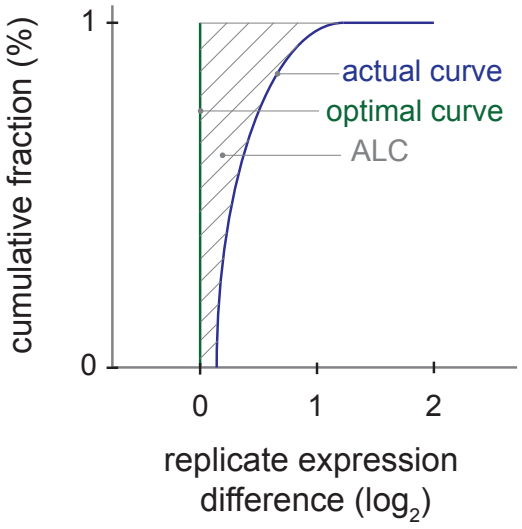


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.385, equivalent to a mean 1.306 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



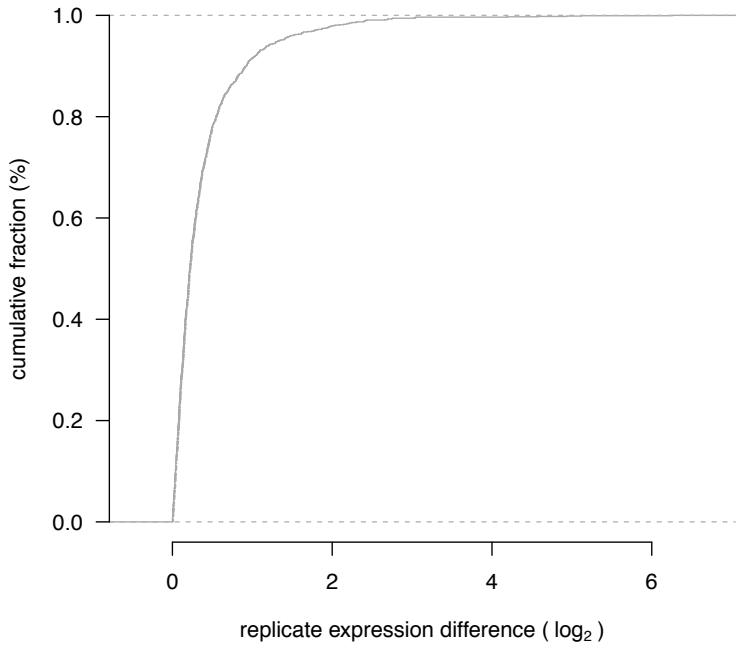
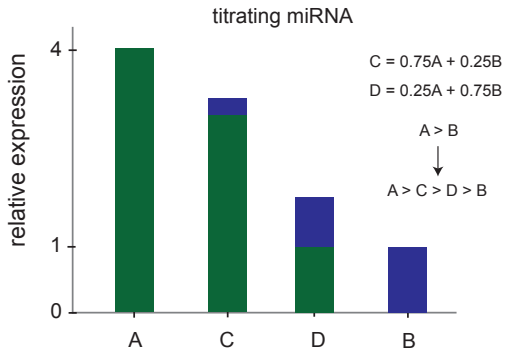


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

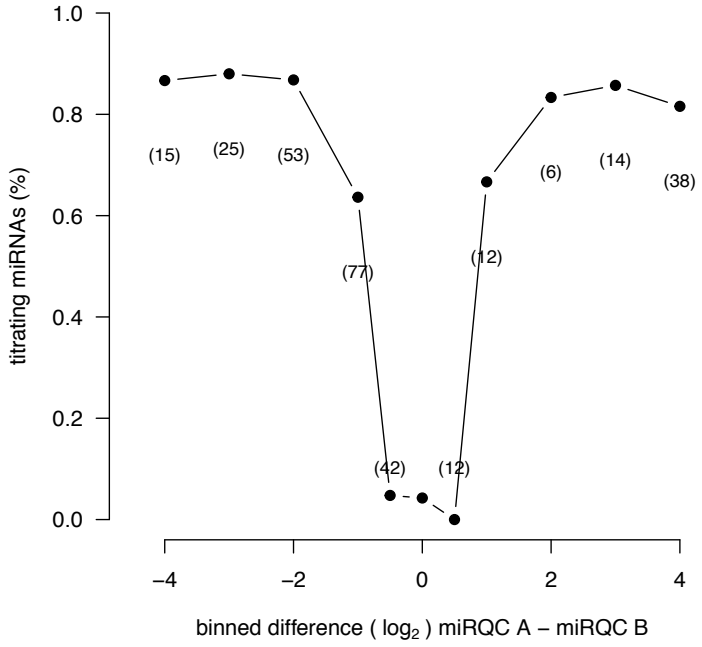
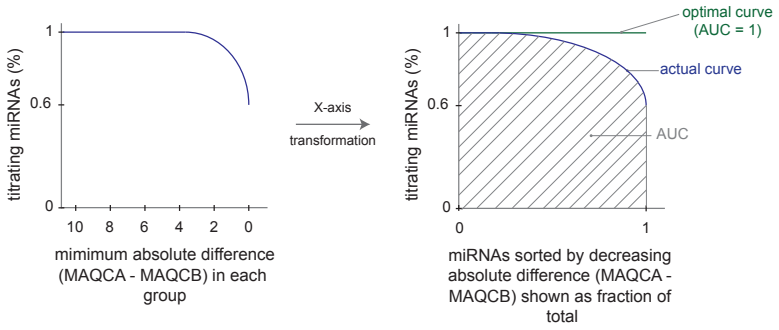


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.784.

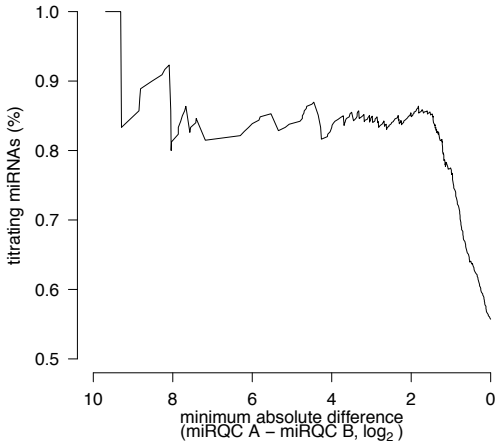


Figure 5: Titration response in function of absolute difference.

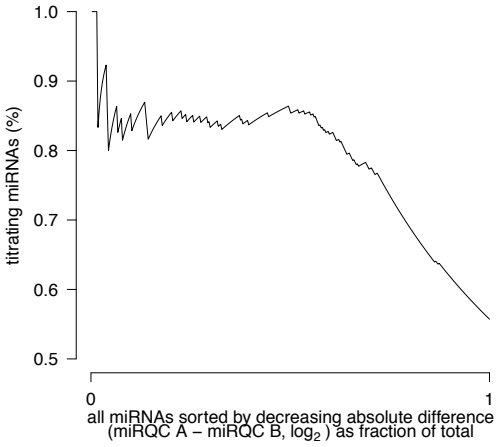


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

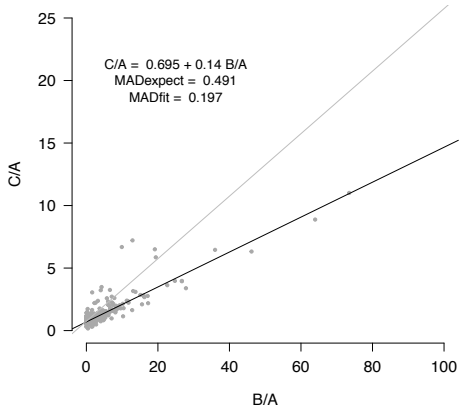


Figure 7: Titration response linearity.

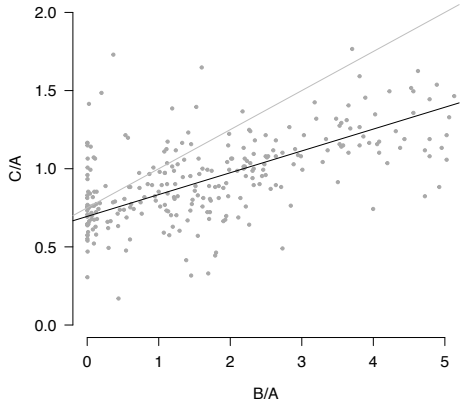


Figure 8: Figure 7 zoomed.

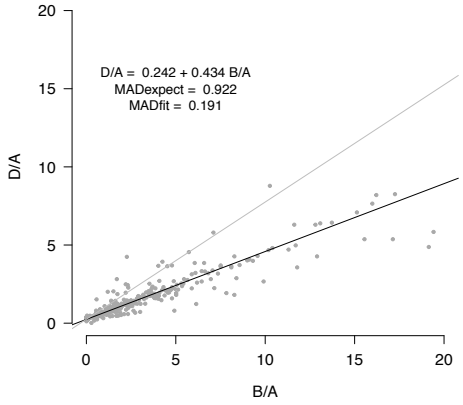


Figure 9: Titration response linearity.

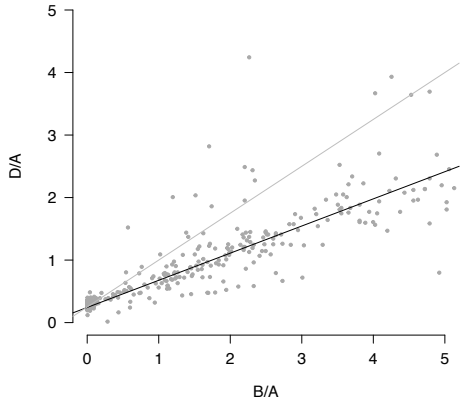


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 50% of all off-target combinations with a median relative cross-reactivity of 0.5%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.1	0.1	0.0
miR-302b-3p	0.0	100.0	0.9	0.0
miR-302c-3p	0.5	0.0	100.0	0.0
miR-302d-3p	0.1	0.1	0.1	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	0.0	3.1
let-7b-5p	0.0	100.0	1.1	0.0
let-7c	1.2	3.0	100.0	0.0
let-7d-5p	0.4	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 8.4%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
7.4	9.1	9.1	8.0

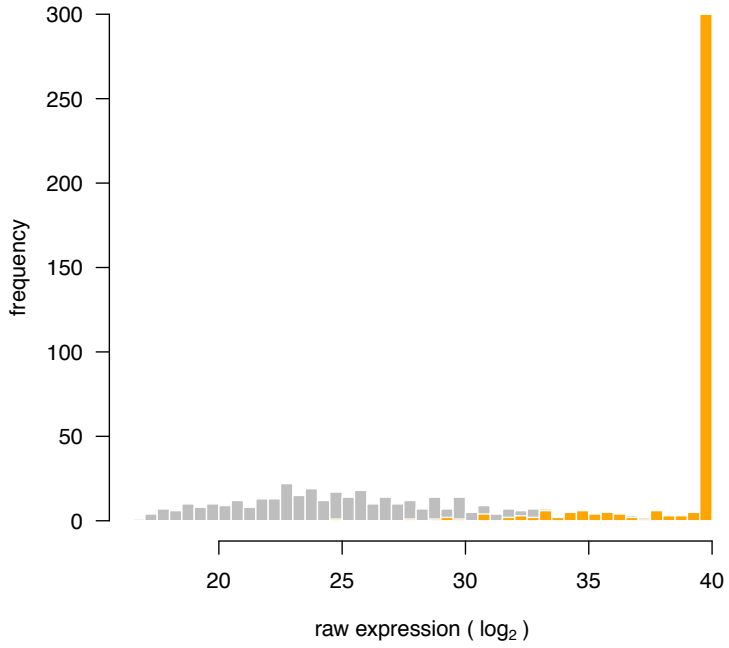


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 55 miRNAs were detected in all 4 samples while 94 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.778, equivalent to a mean 1.715 fold replicate expression difference.

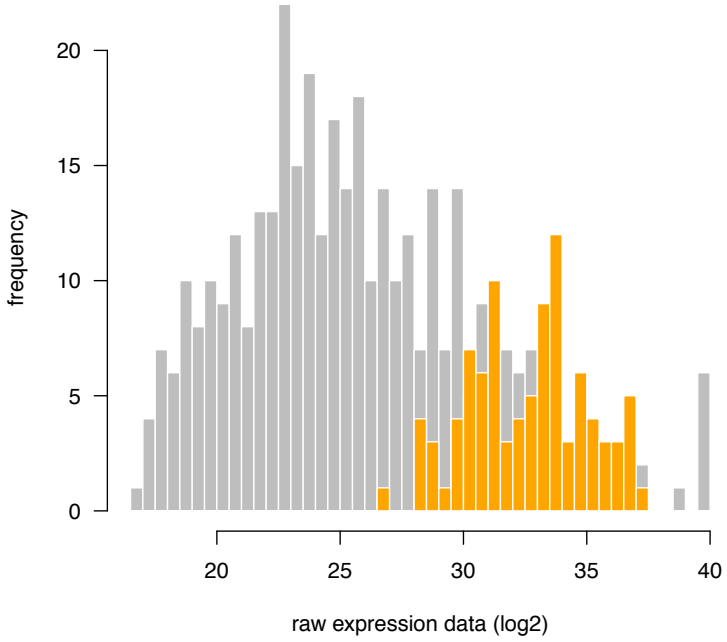


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

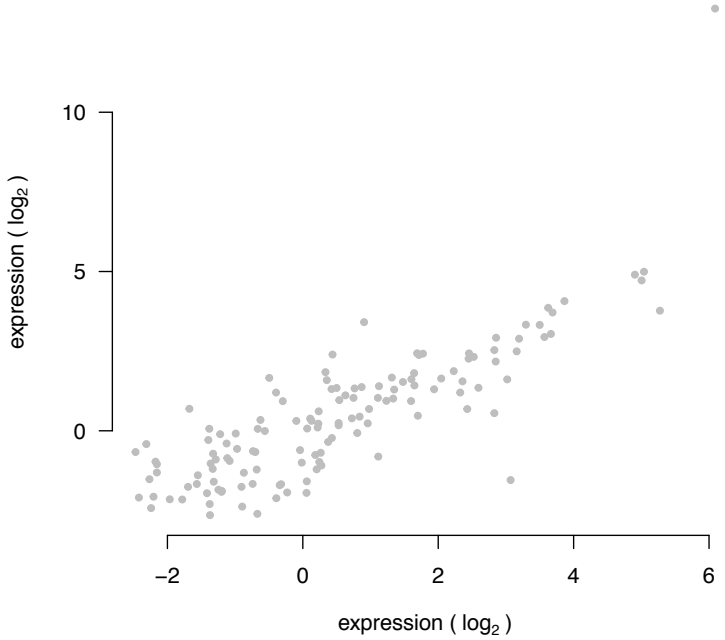


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 29 miRNAs were significantly upregulated and 16 miRNAs were significantly downregulated.

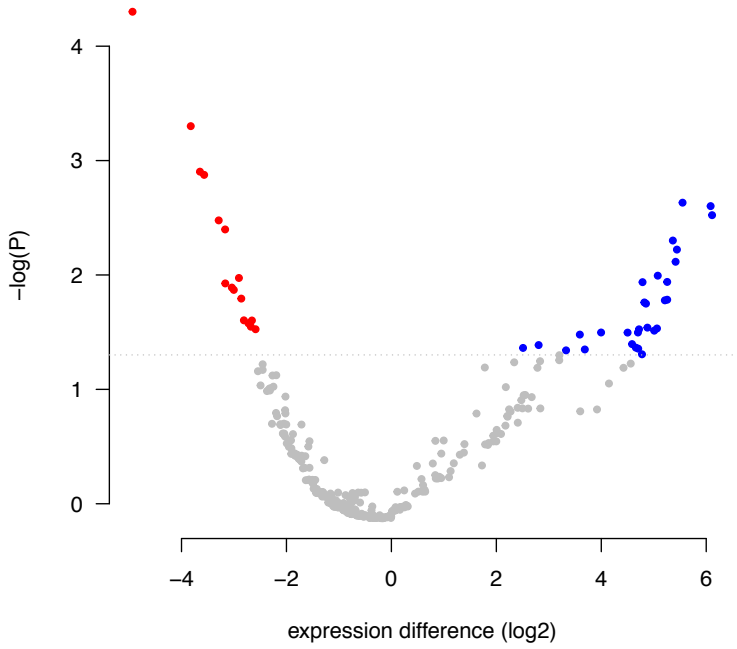


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	363
	fraction single positives (%)	2.7
	expression range (log2-units)	17
	ALC	0.385
titration	AUC titration response	0.784
	MADexpect (D/A)	0.491
	MADfit (D/A)	0.197
	MADexpect (C/A)	0.922
	MADfit (C/A)	0.191
specificity	off-target combinations with cross reactivity (%)	50
	median relative cross-reactivity (%)	0.5
non-template control	positive miRNAs (%)	8.4
serum miRNAs	detected miRNAs	94
differential expression	significant down	16
	significant up	29

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 6

Quanta Biosciences qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 32.61 cycles.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 417 unique double positives were detected while the percentage of single positives was 3.24 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 15.7 $\log_2$-units.

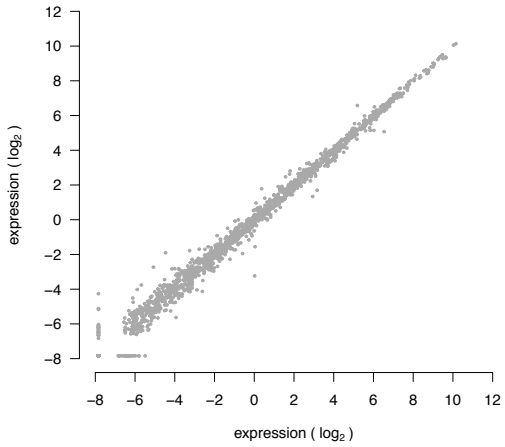


Figure 1: Replicate miRNA expression correlation plot.

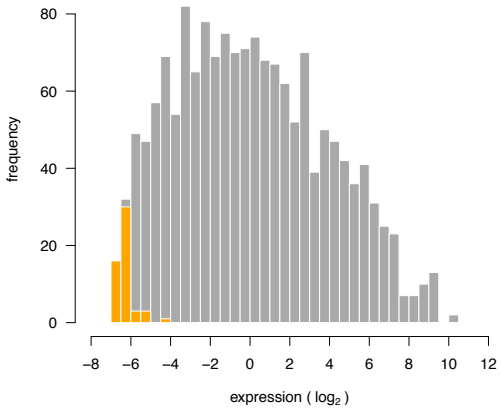
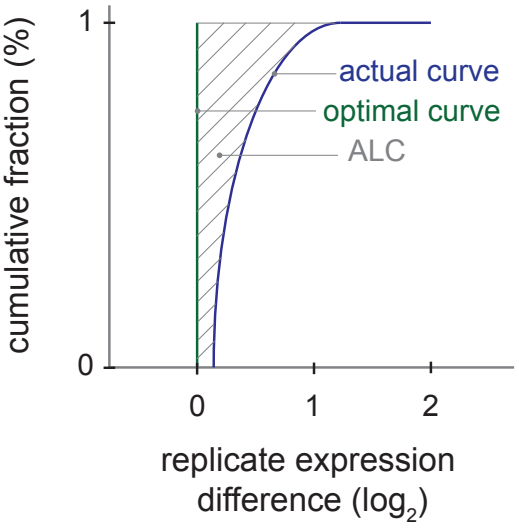


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.249, equivalent to a mean 1.188 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



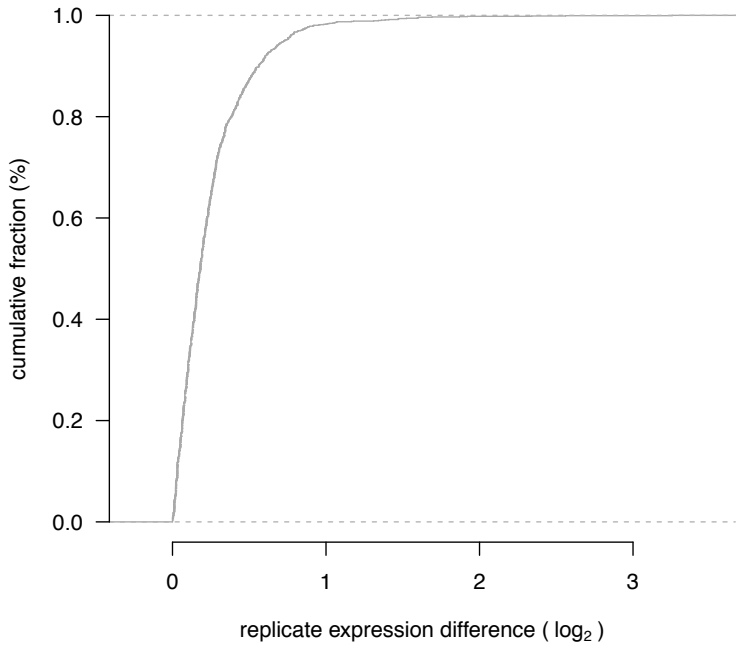
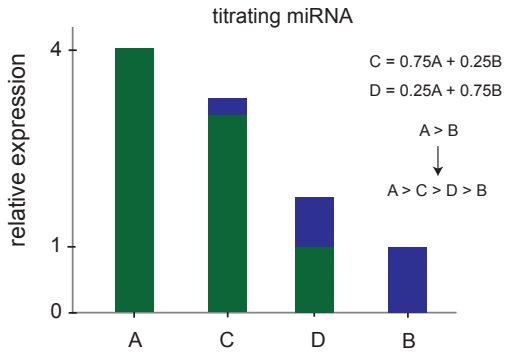


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

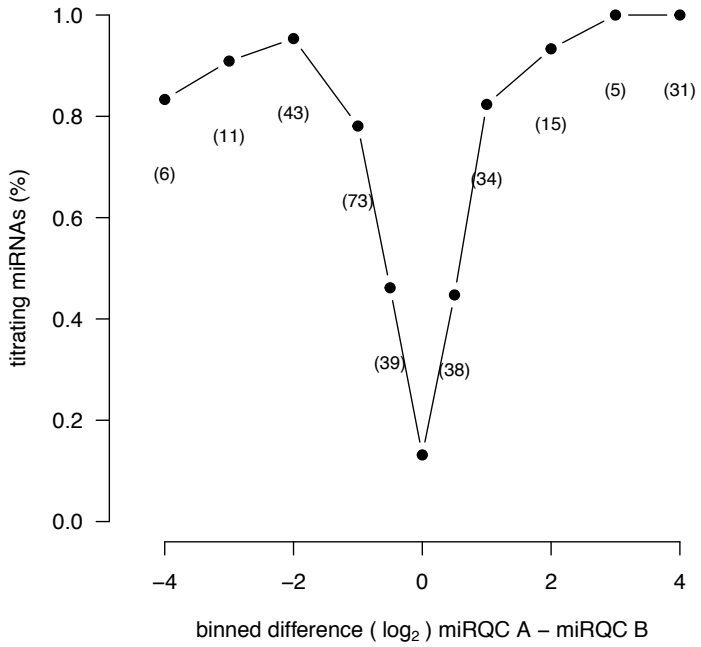
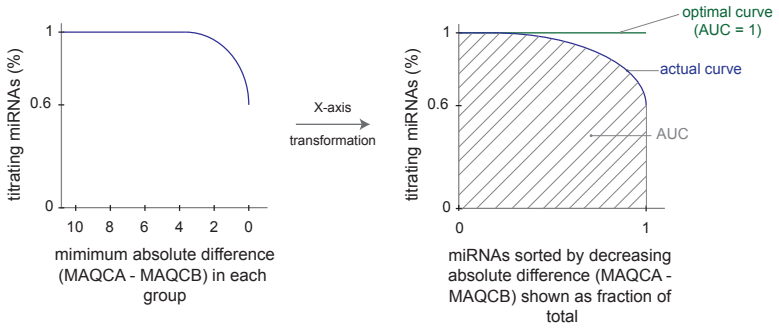


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.866.

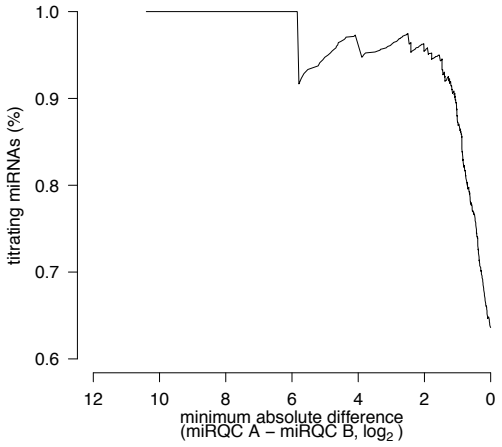


Figure 5: Titration response in function of absolute difference.

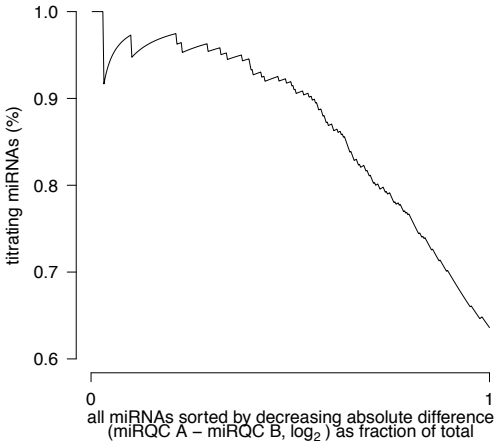


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

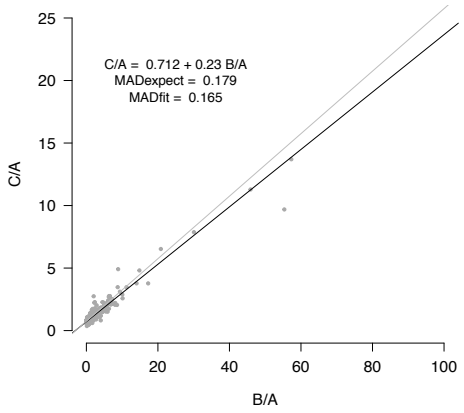


Figure 7: Titration response linearity.

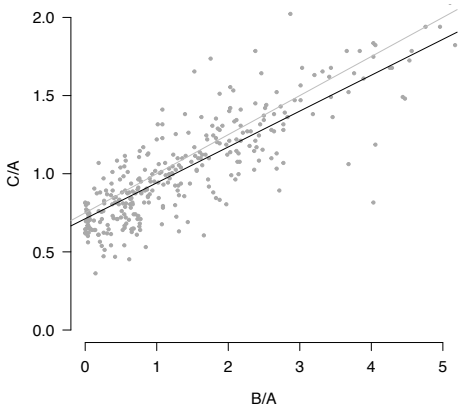


Figure 8: Figure 7 zoomed.

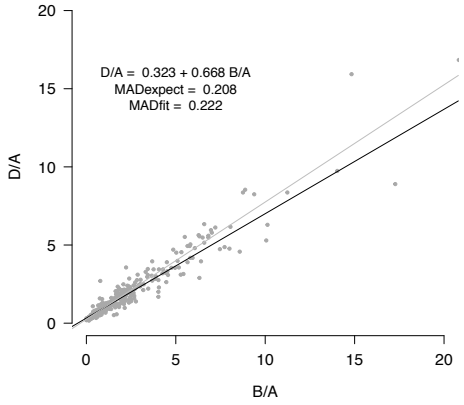


Figure 9: Titration response linearity.

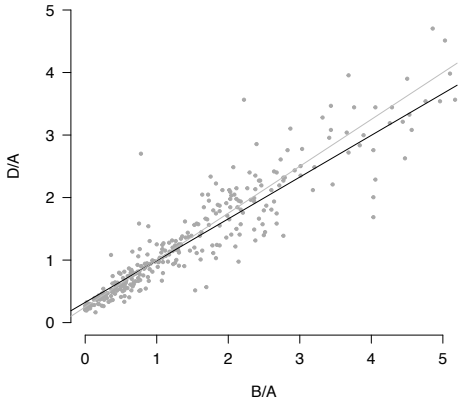


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 16.7% of all off-target combinations with a median relative cross-reactivity of 11.1%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	14.7	0.0
let-7b-5p	0.0	100.0	7.0	0.0
let-7c	23.8	7.6	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 2.88%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
2.9	2.9	2.4	3.4

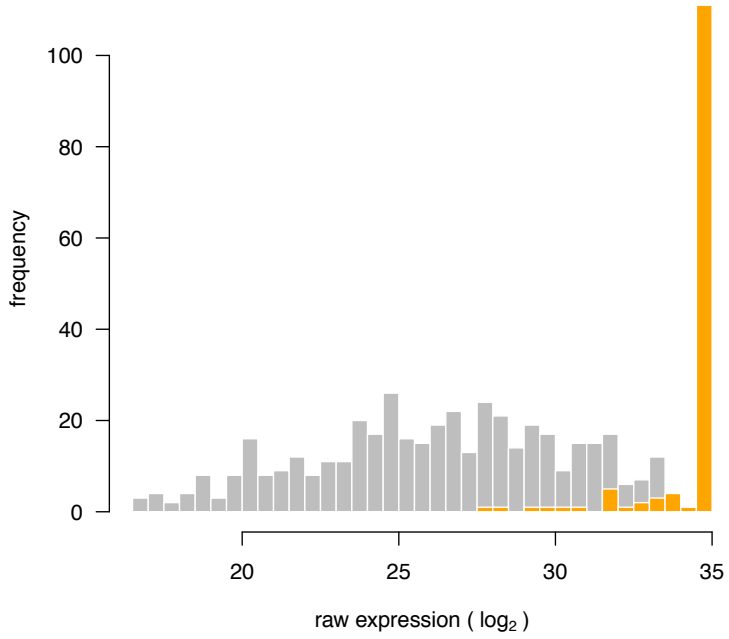


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 58 miRNAs were detected in all 4 samples while 71 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.465, equivalent to a mean 1.38 fold replicate expression difference.

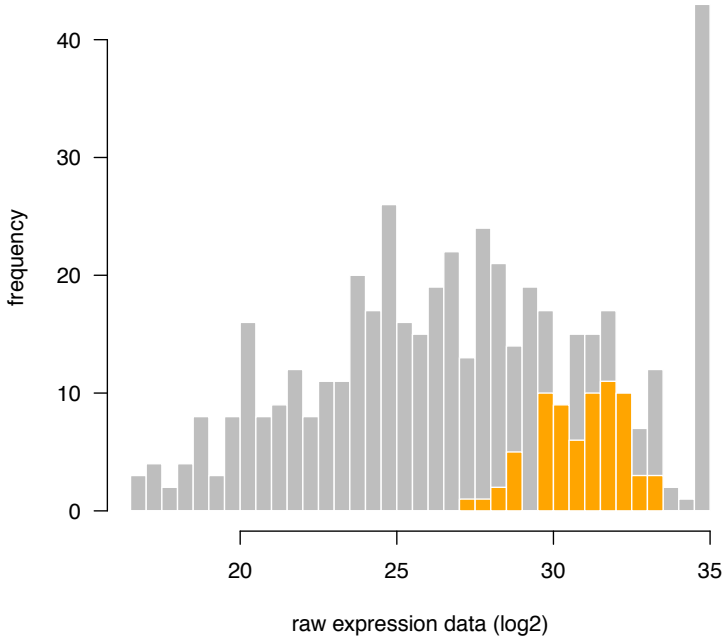


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

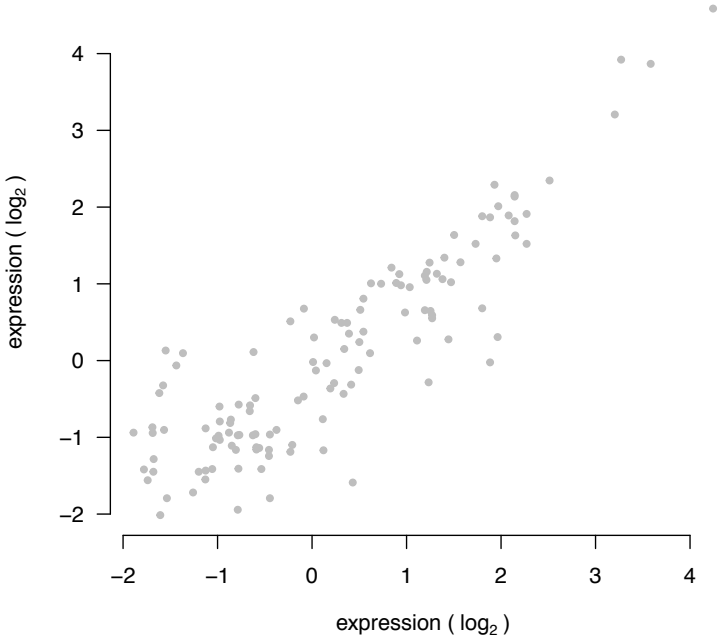


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 44 miRNAs were significantly upregulated and 22 miRNAs were significantly downregulated.

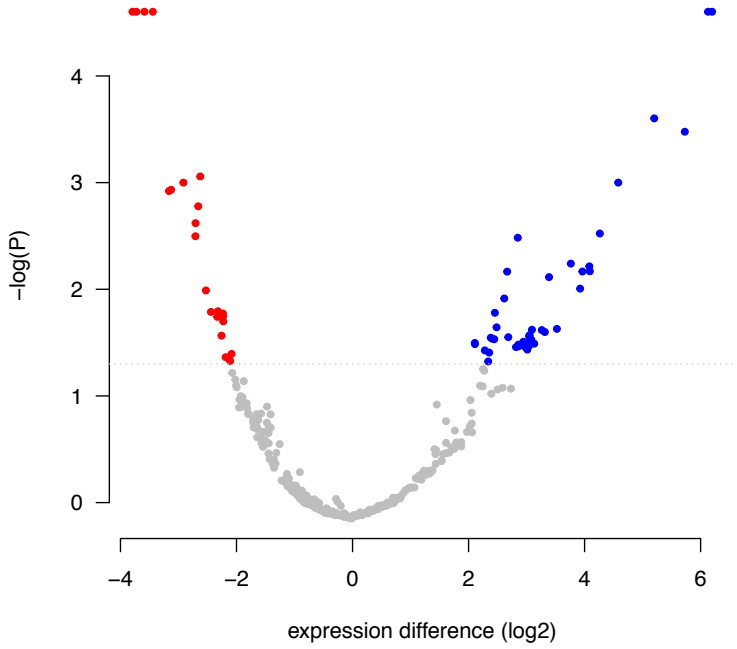


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	417
	fraction single positives (%)	3.24
	expression range (log2-units)	15.7
	ALC	0.249
titration	AUC titration response	0.866
	MADexpect (D/A)	0.179
	MADfit (D/A)	0.165
	MADexpect (C/A)	0.208
	MADfit (C/A)	0.222
specificity	off-target combinations with cross reactivity (%)	16.7
	median relative cross-reactivity (%)	11.1
non-template control	positive miRNAs (%)	2.88
serum miRNAs	detected miRNAs	71
differential expression	significant down	22
	significant up	44

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 7

WaferGen qPCR

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 28.5 cycles.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 920 unique double positives were detected while the percentage of single positives was 2.59 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 16.1 $\log_2$-units.

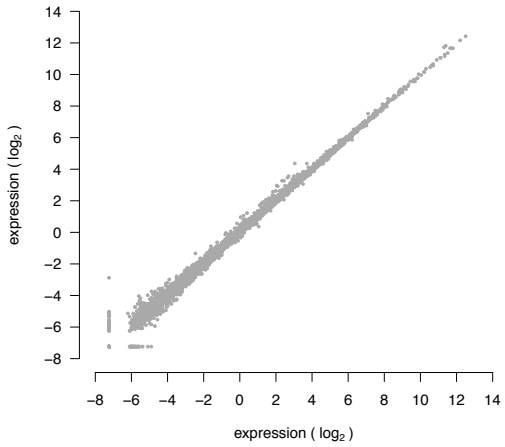


Figure 1: Replicate miRNA expression correlation plot.

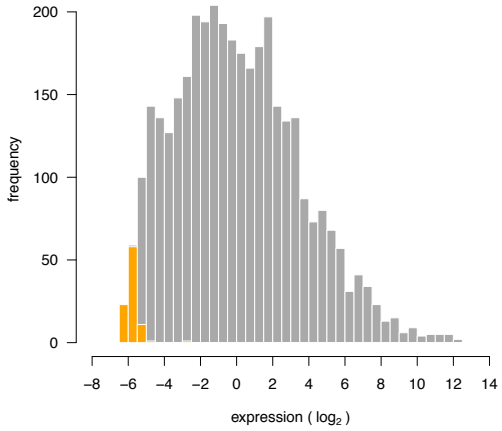
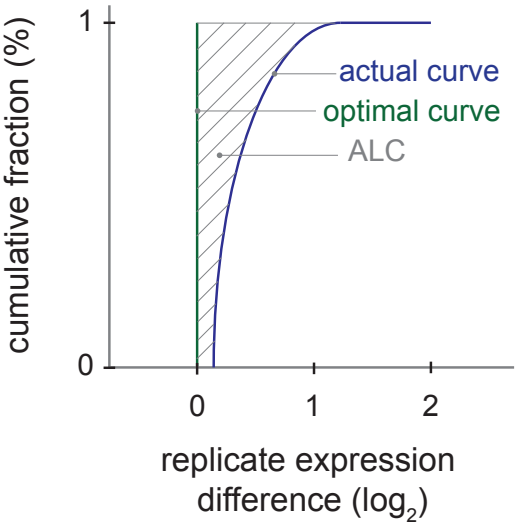


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.159, equivalent to a mean 1.117 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



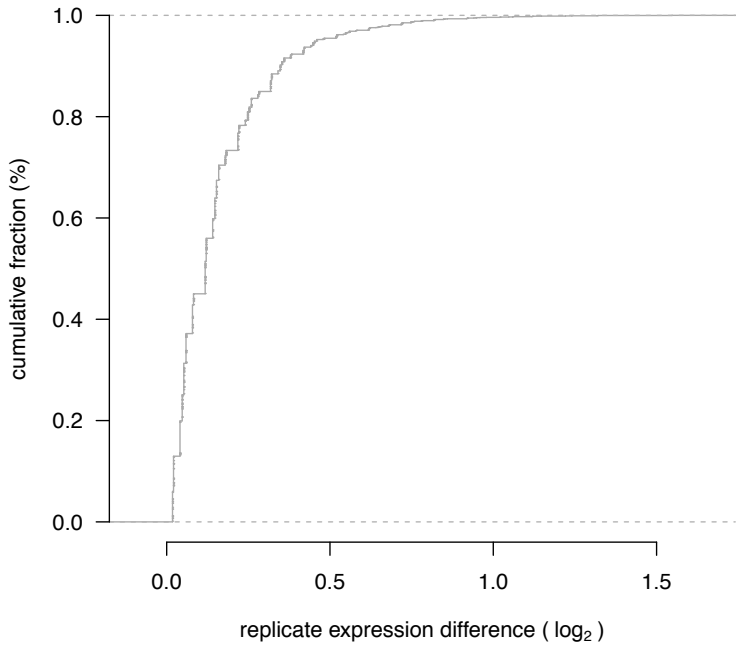
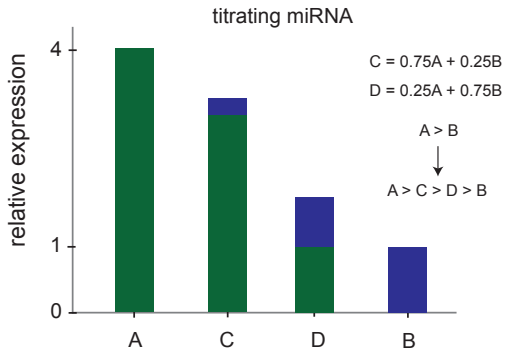


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

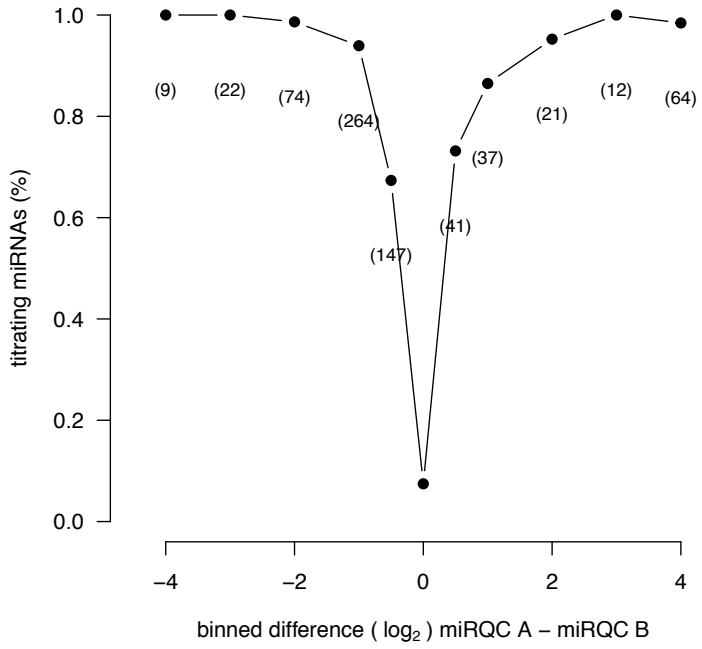
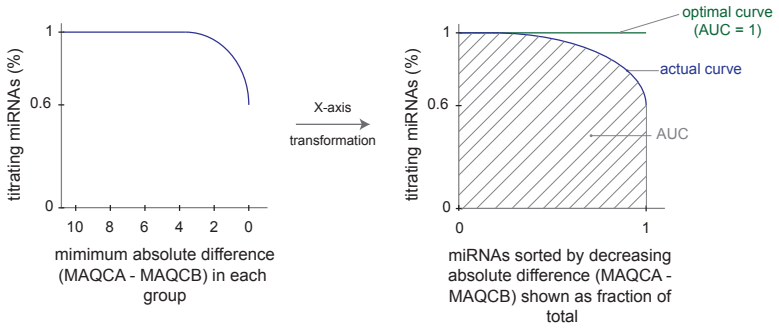


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.933.

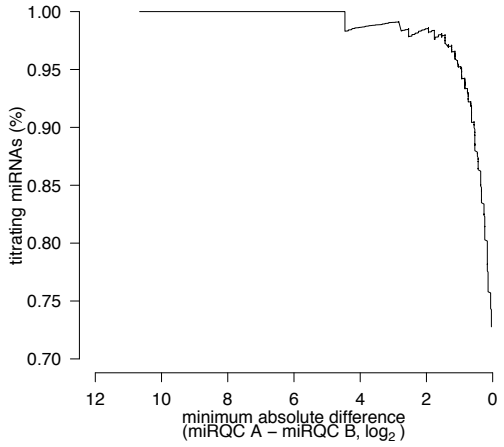


Figure 5: Titration response in function of absolute difference.

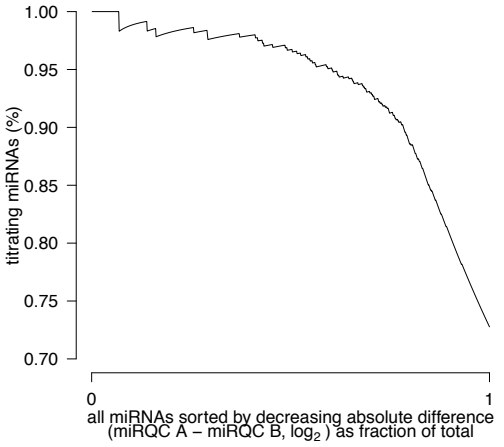


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

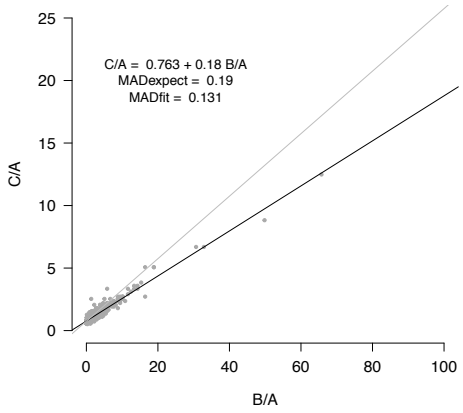


Figure 7: Titration response linearity.

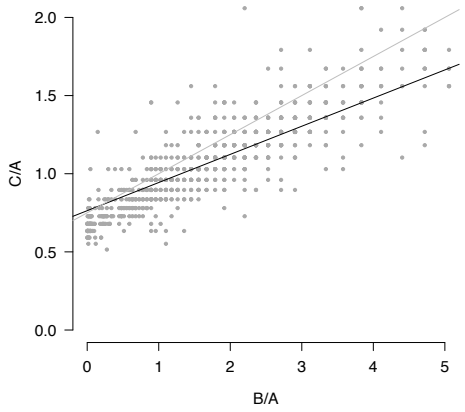


Figure 8: Figure 7 zoomed.

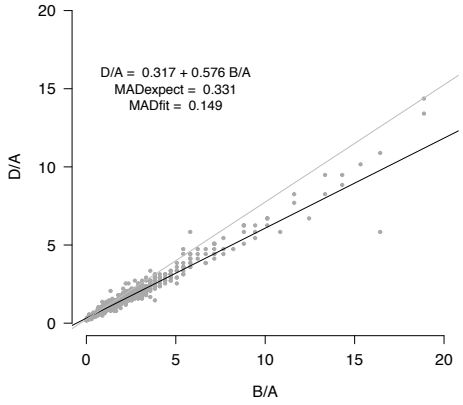


Figure 9: Titration response linearity.

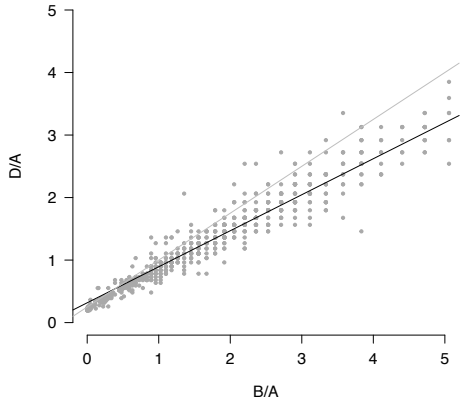


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 54.2% of all off-target combinations with a median relative cross-reactivity of 7.2%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	7.2	5.5	100.0	5.7
miR-302d-3p	2.3	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	35.4	40.6	14.4
let-7b-5p	33.0	100.0	43.5	0.0
let-7c	23.3	4.4	100.0	0.0
let-7d-5p	1.0	1.5	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 10.79%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
9.5	11.1	11.3	11.3

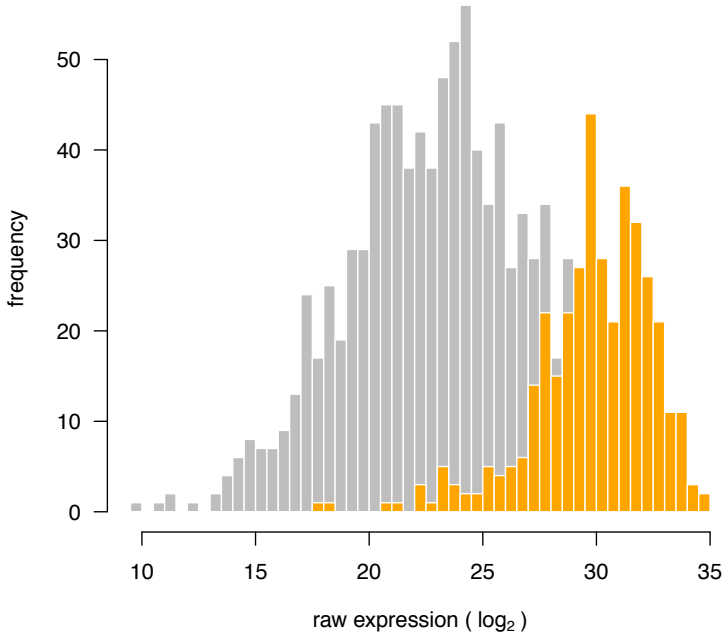


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 22 miRNAs were detected in all 4 samples while 34 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.521, equivalent to a mean 1.435 fold replicate expression difference.

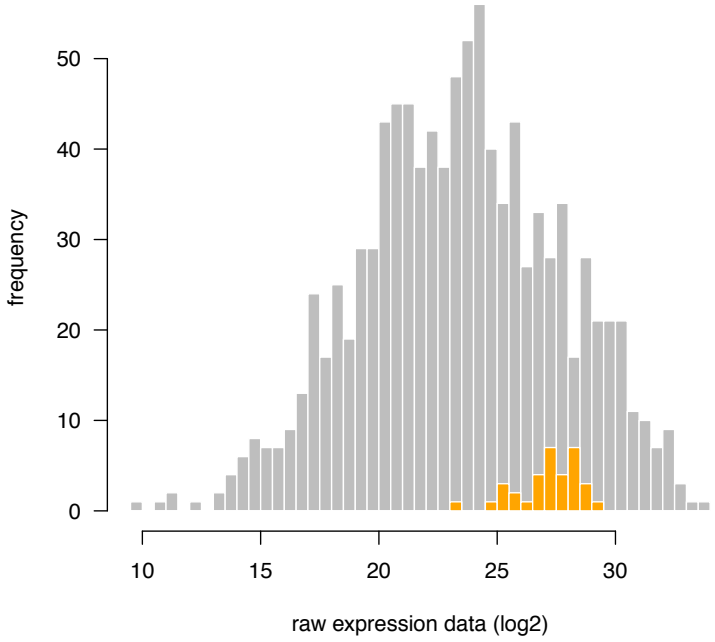


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

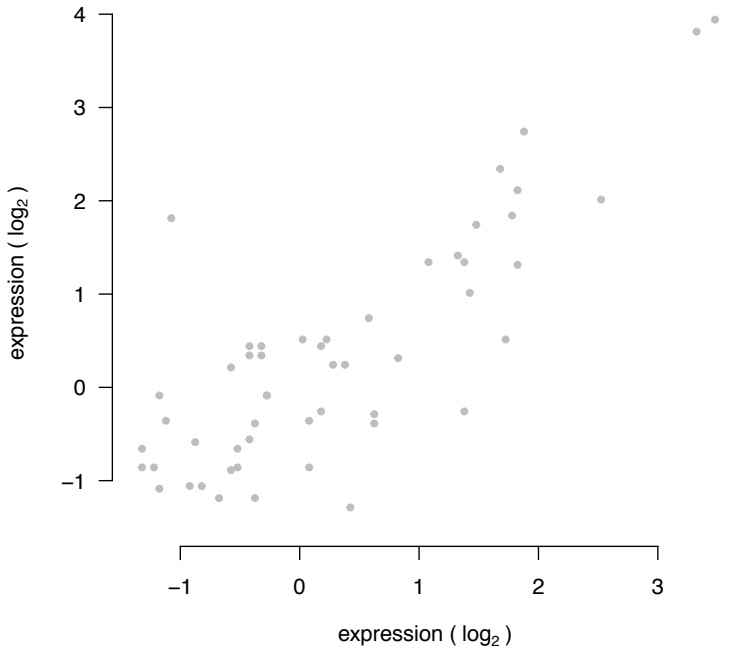


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 73 miRNAs were significantly upregulated and 50 miRNAs were significantly downregulated.

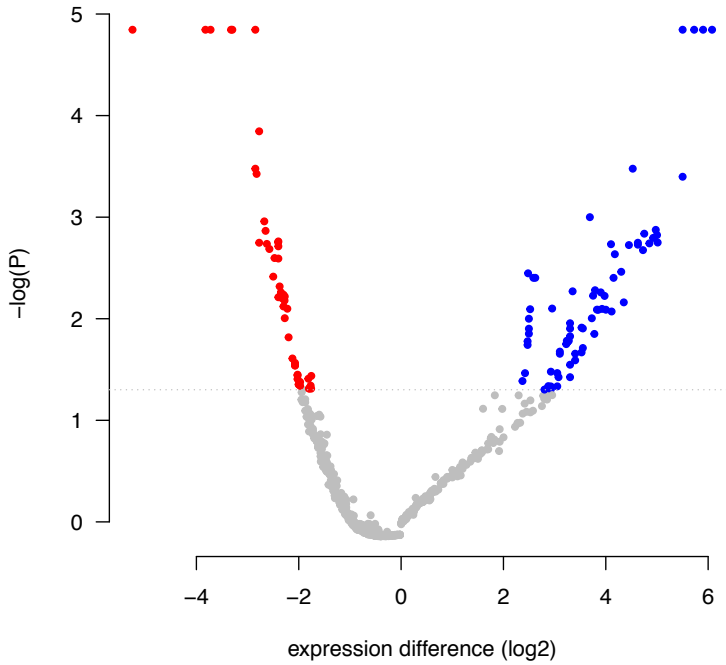


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	920
	fraction single positives (%)	2.59
	expression range (log2-units)	16.1
	ALC	0.159
titration	AUC titration response	0.933
	MADexpect (D/A)	0.19
	MADfit (D/A)	0.131
	MADexpect (C/A)	0.331
	MADfit (C/A)	0.149
specificity	off-target combinations with cross reactivity (%)	54.2
	median relative cross-reactivity (%)	7.2
non-template control	positive miRNAs (%)	10.79
serum miRNAs	detected miRNAs	34
differential expression	significant down	50
	significant up	73

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 8

Affymetrix microarray

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA.

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 2.766.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 617 unique double positives were detected while the percentage of single positives was 14.33 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 9.6 $\log_2$-units.

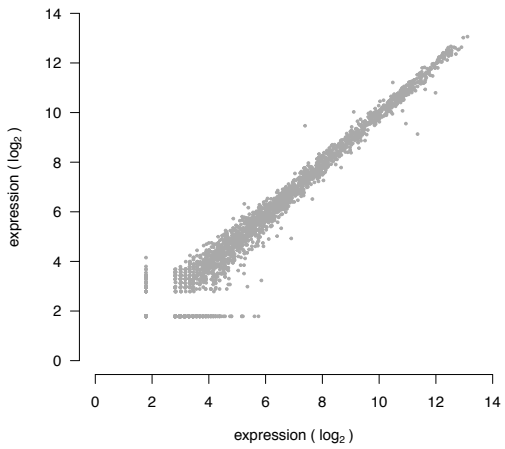


Figure 1: Replicate miRNA expression correlation plot.

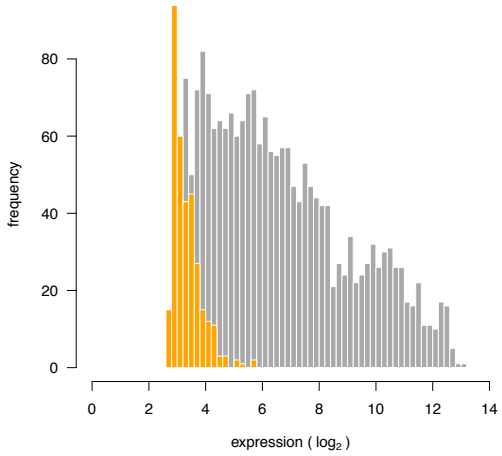
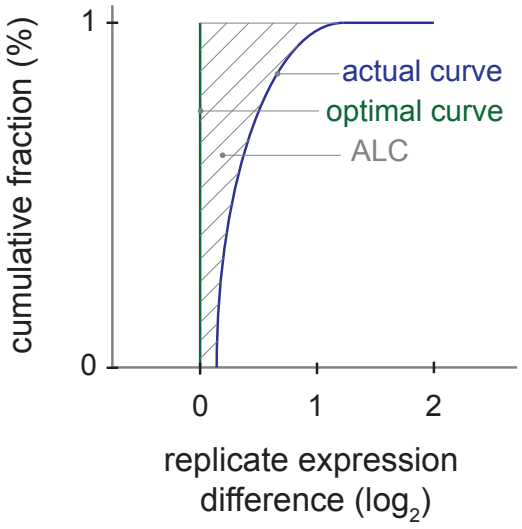


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.274, equivalent to a mean 1.209 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



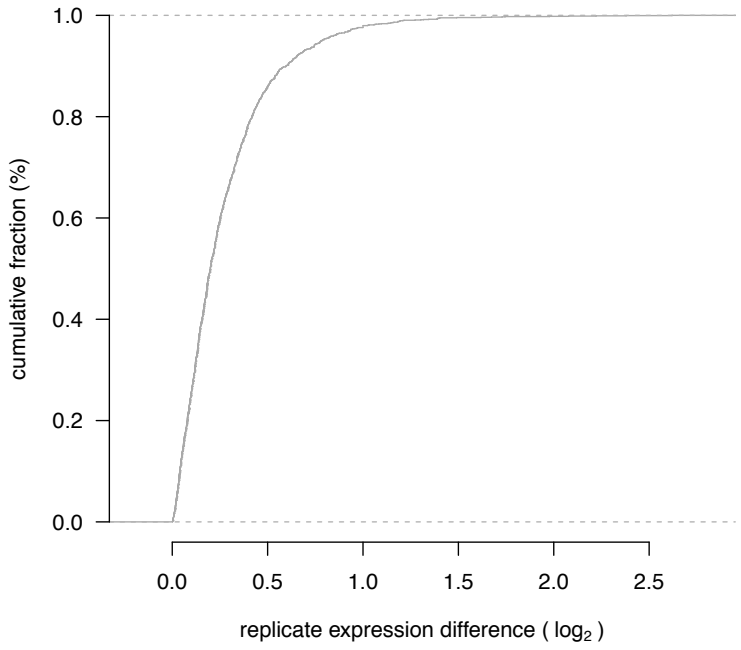
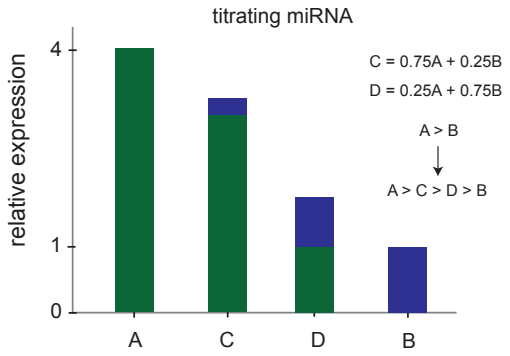


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

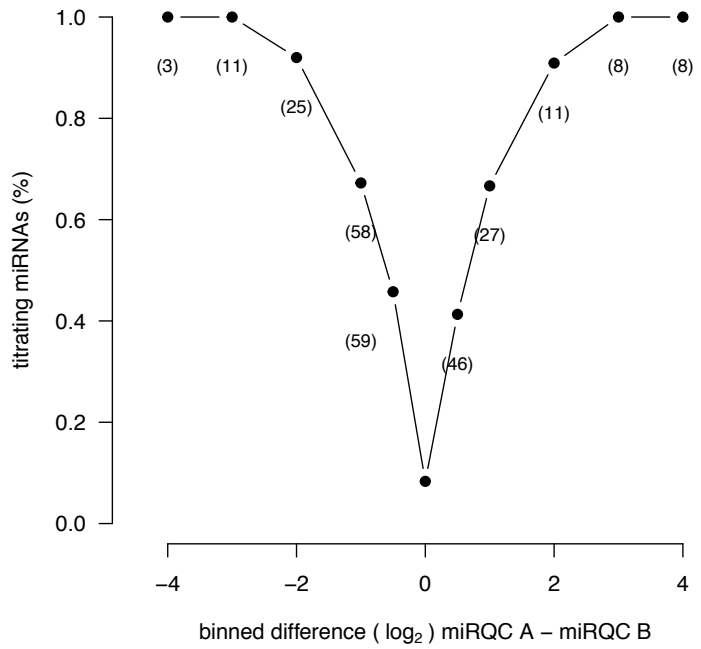
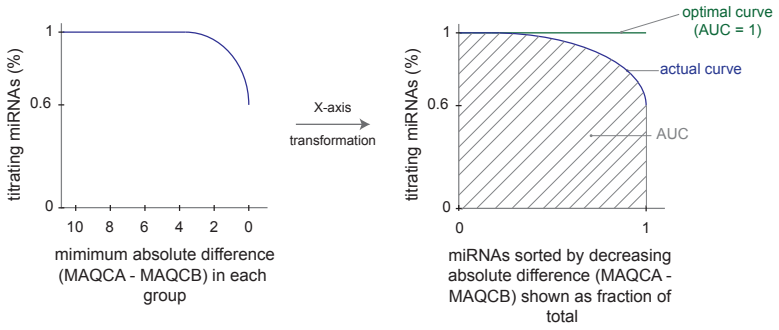


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.729.

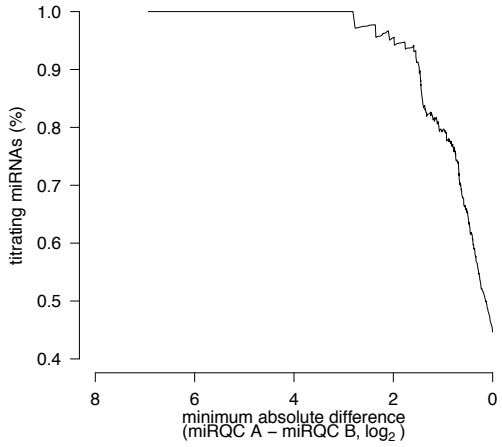


Figure 5: Titration response in function of absolute difference.

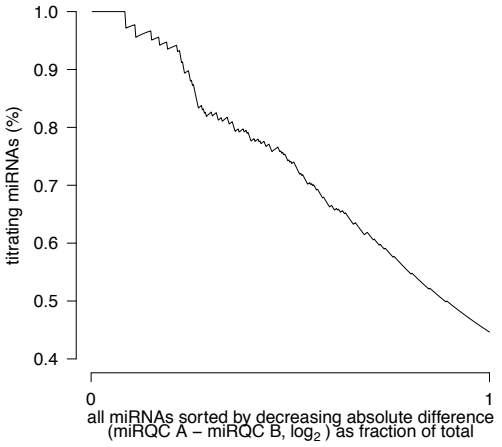


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

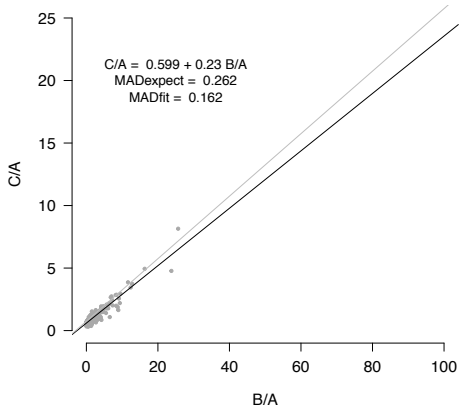


Figure 7: Titration response linearity.

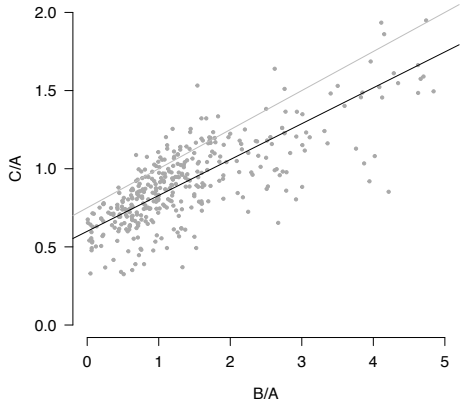


Figure 8: Figure 7 zoomed.

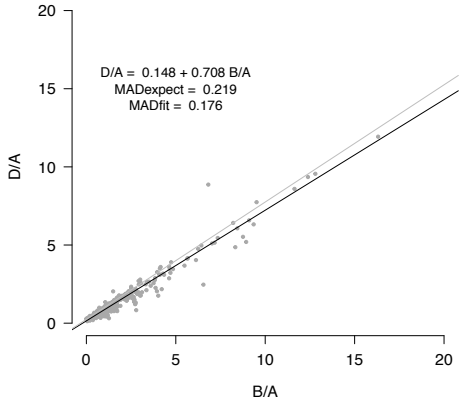


Figure 9: Titration response linearity.

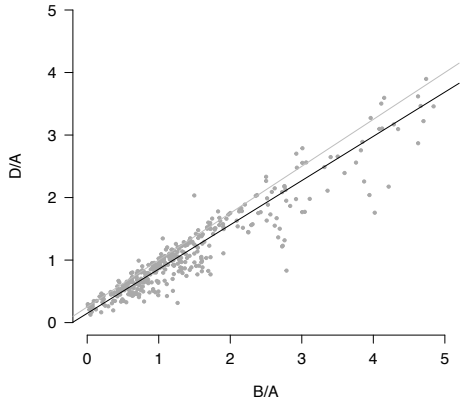


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 50% of all off-target combinations with a median relative cross-reactivity of 1.8%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	0.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	6.2	61.5	6.0
let-7b-5p	0.2	100.0	12.4	0.4
let-7c	0.2	54.4	100.0	0.3
let-7d-5p	1.1	0.3	2.4	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 3.2%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 14) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
1.1	2.1	6.2	3.4

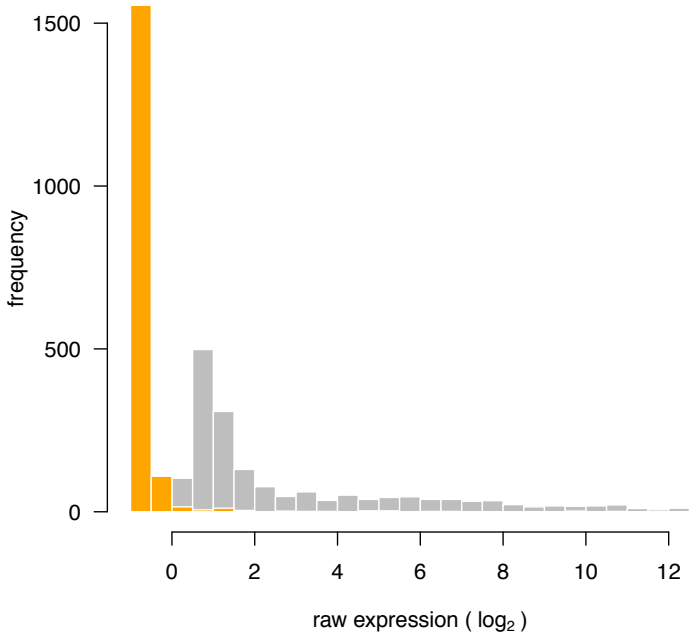


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 59 miRNAs were significantly upregulated and 35 miRNAs were significantly downregulated.

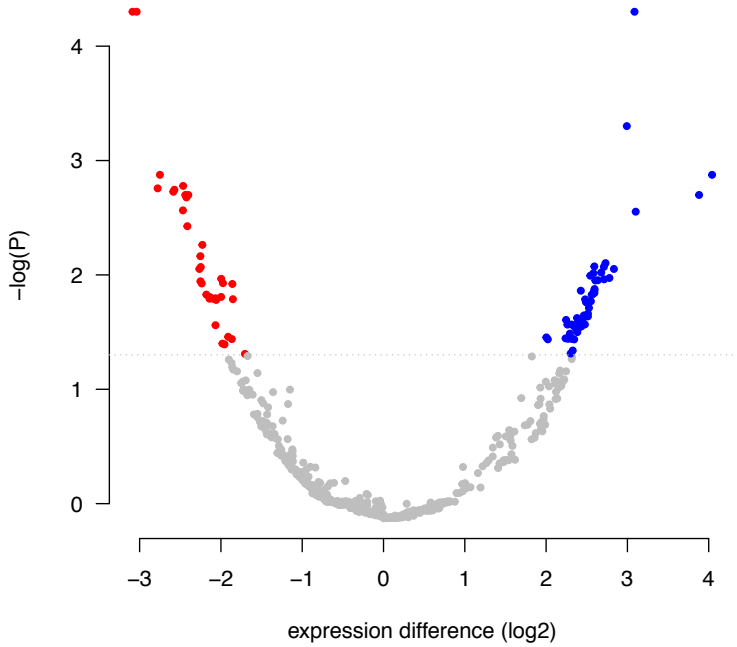


Figure 12: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

8 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	617
	fraction single positives (%)	14.33
	expression range (log2-units)	9.6
	ALC	0.274
titration	AUC titration response	0.729
	MADexpect (D/A)	0.262
	MADfit (D/A)	0.162
	MADexpect (C/A)	0.219
	MADfit (C/A)	0.176
specificity	off-target combinations with cross reactivity (%)	50
	median relative cross-reactivity (%)	1.8
non-template control	positive miRNAs (%)	3.2
serum miRNAs	detected miRNAs	NA
differential expression	significant down	35
	significant up	59

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 9

Agilent microarray

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here -5.389.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 496 unique double positives were detected while the percentage of single positives was 2.02 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 10.9 $\log_2$-units.

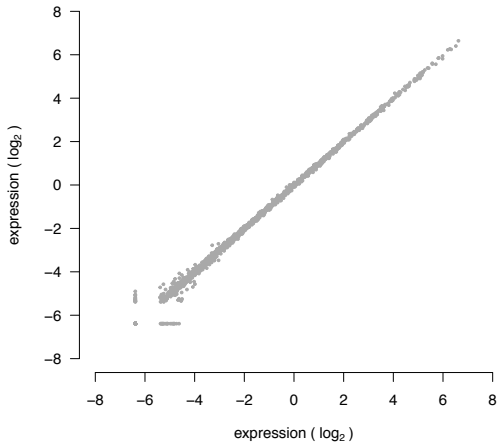


Figure 1: Replicate miRNA expression correlation plot.

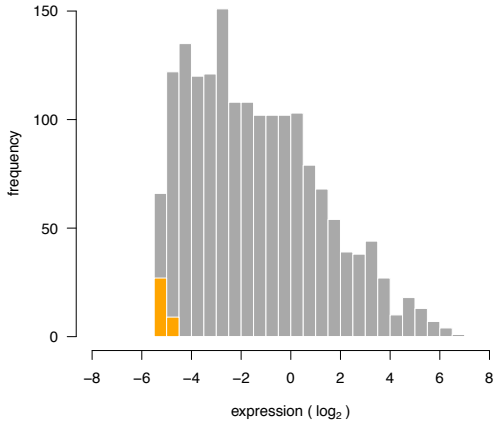
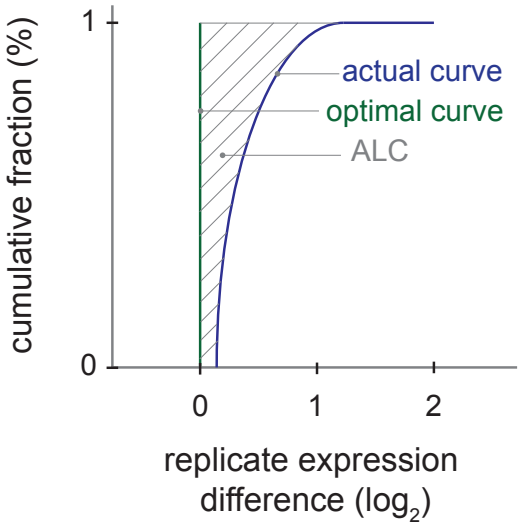


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.064, equivalent to a mean 1.045 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



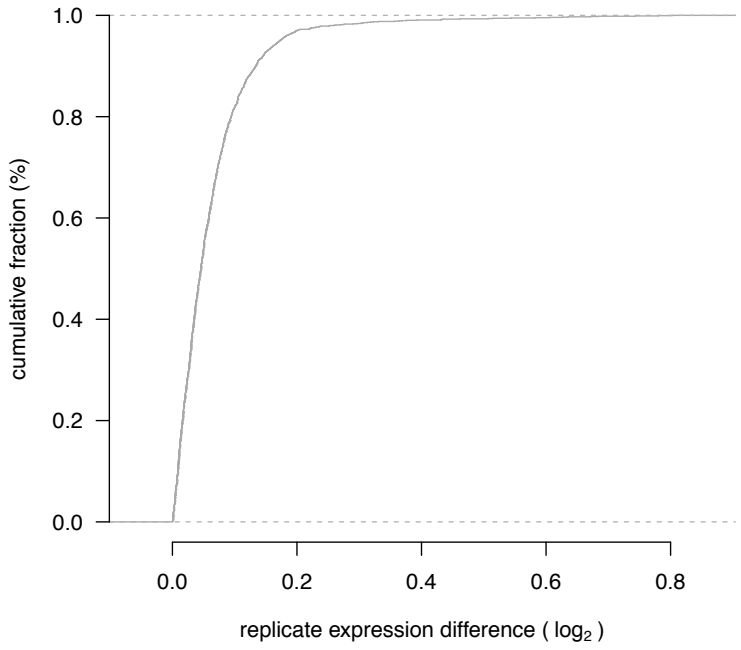
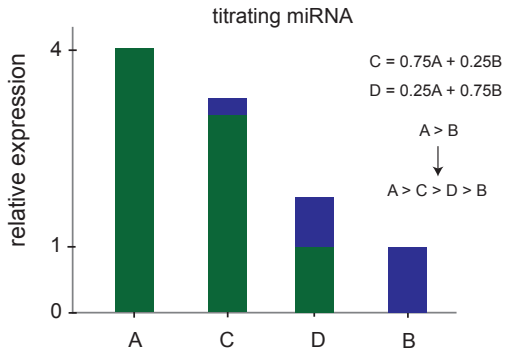


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

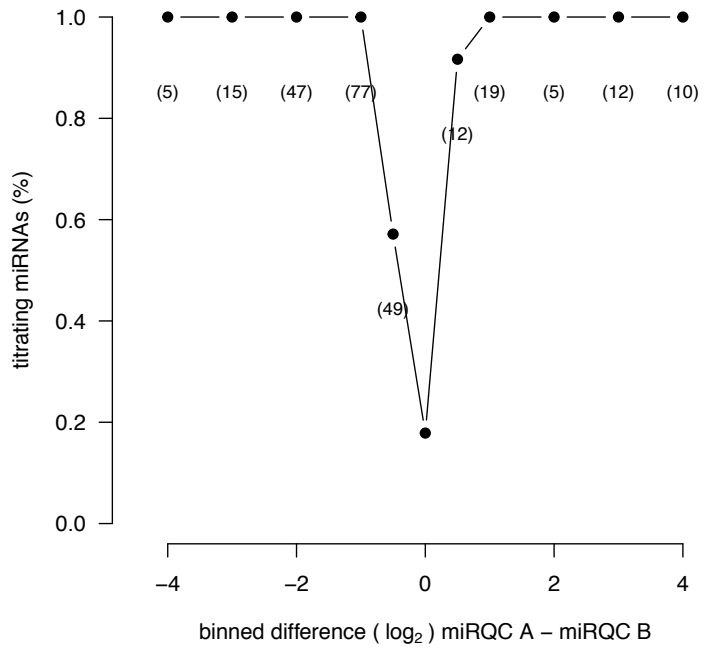
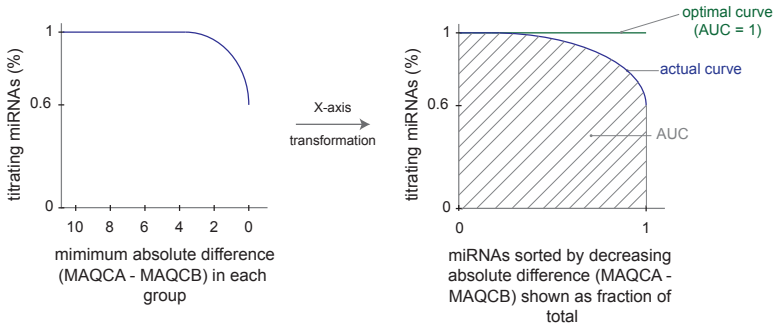


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.947.

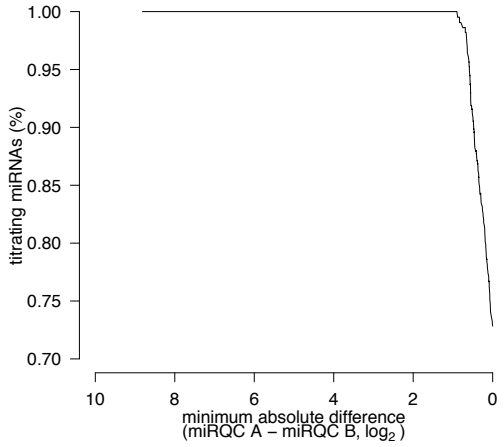


Figure 5: Titration response in function of absolute difference.

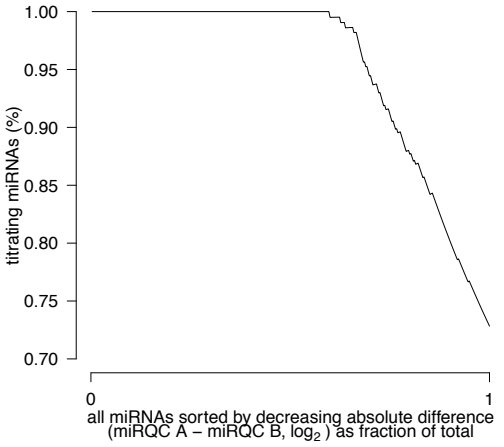


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

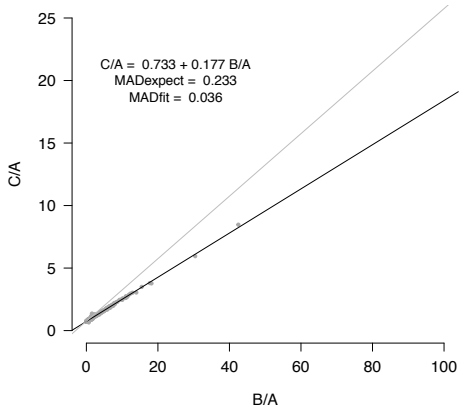


Figure 7: Titration response linearity.

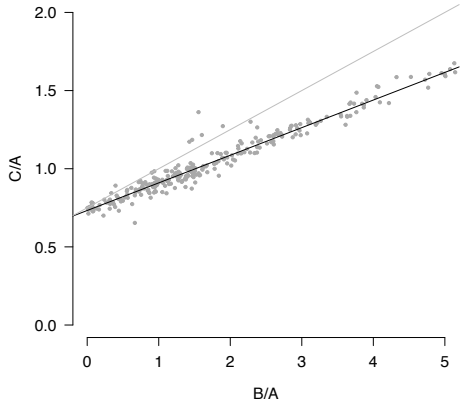


Figure 8: Figure 7 zoomed.

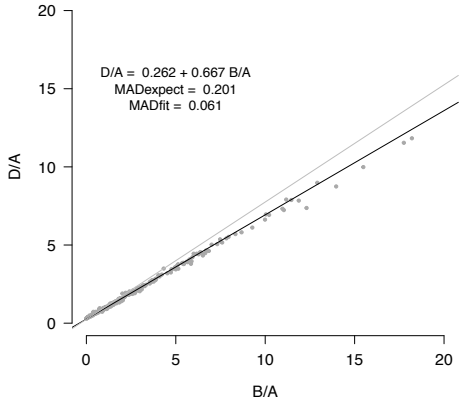


Figure 9: Titration response linearity.

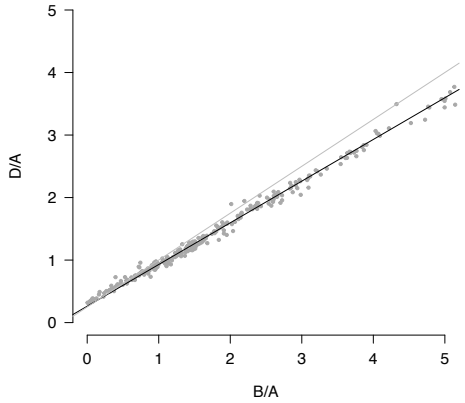


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 25% of all off-target combinations with a median relative cross-reactivity of 20.9%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	19.1	72.8	22.7
let-7b-5p	0.0	100.0	11.6	0.0
let-7c	0.0	52.9	100.0	0.0
let-7d-5p	4.3	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 5.14%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
4.6	5.2	5.4	5.2

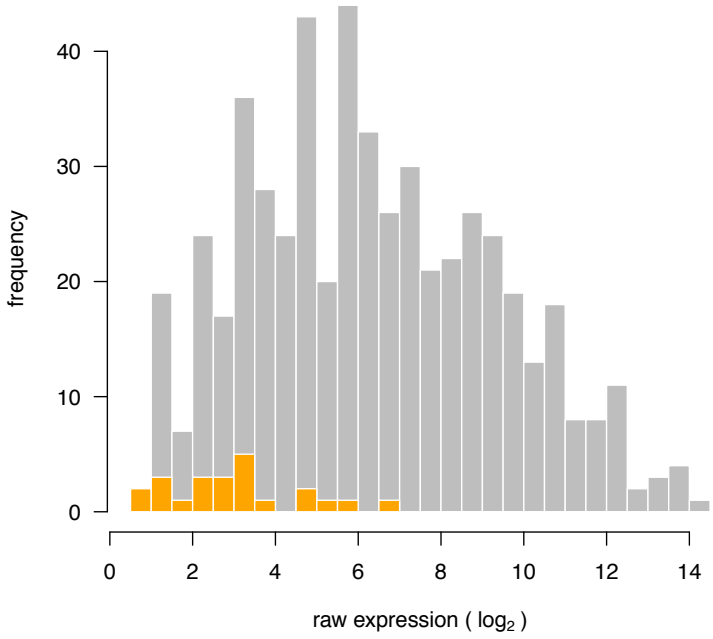


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 50 miRNAs were detected in all 4 samples while 57 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.19, equivalent to a mean 1.141 fold replicate expression difference.

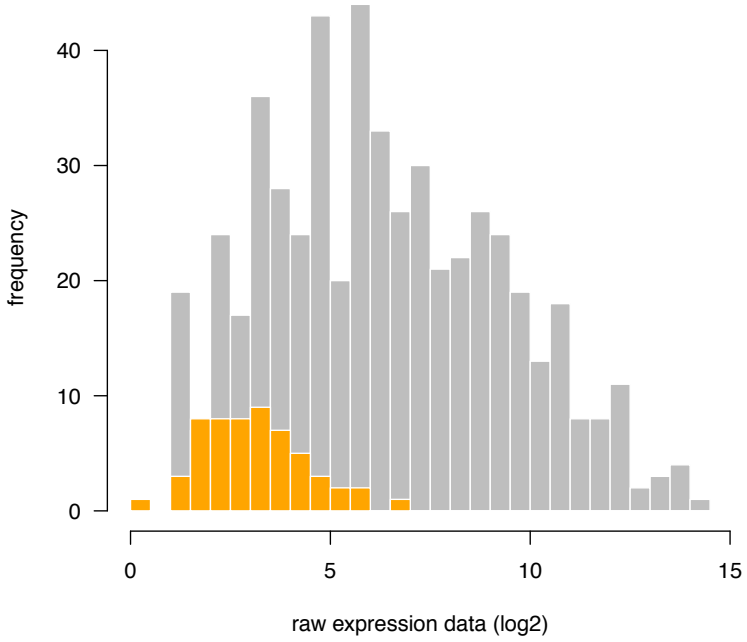


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

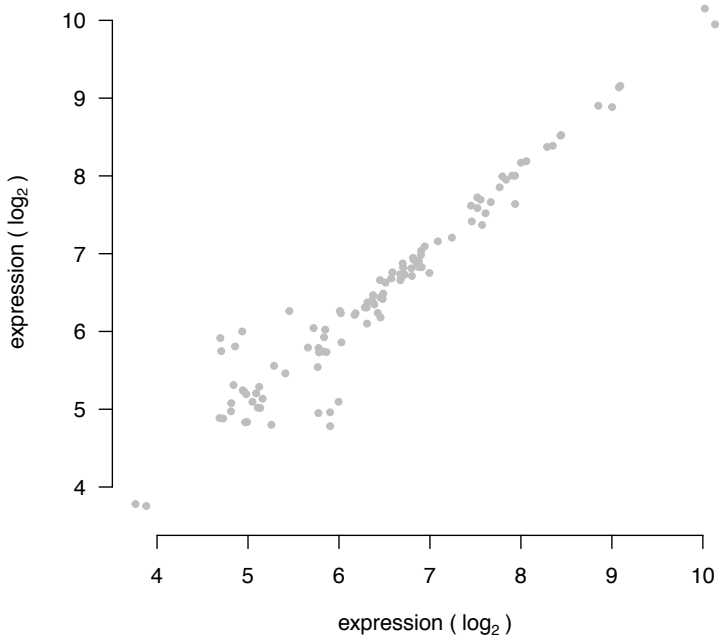


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 38 miRNAs were significantly upregulated and 35 miRNAs were significantly downregulated.

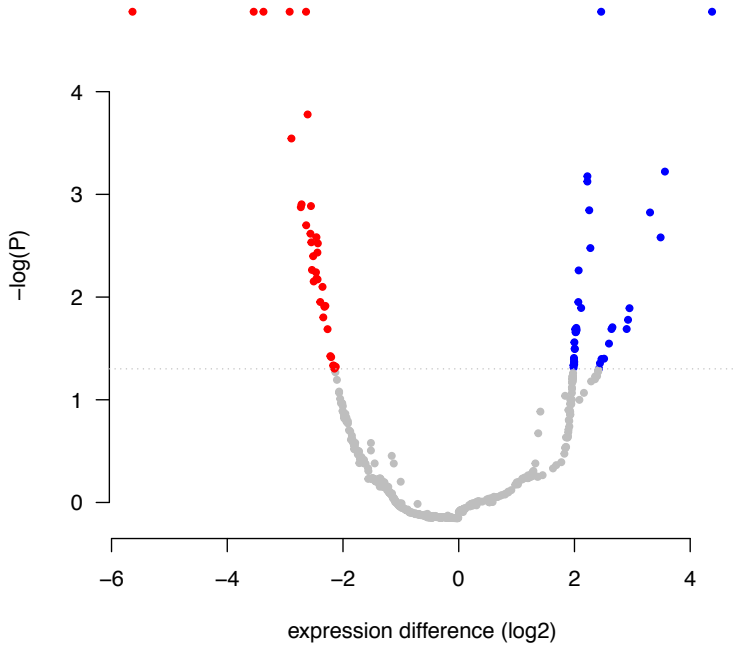


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	496
	fraction single positives (%)	2.02
	expression range (log2-units)	10.9
	ALC	0.064
titration	AUC titration response	0.947
	MADexpect (D/A)	0.233
	MADfit (D/A)	0.036
	MADexpect (C/A)	0.201
	MADfit (C/A)	0.061
specificity	off-target combinations with cross reactivity (%)	25
	median relative cross-reactivity (%)	20.9
non-template control	positive miRNAs (%)	5.14
serum miRNAs	detected miRNAs	57
differential expression	significant down	35
	significant up	38

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 10

Nanostring

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA.

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Platform detection cutoff was determined based on internal negative control probes according to the manufacturer's instructions.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 364 unique double positives were detected while the percentage of single positives was 14.04 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 13.7 $\log_2$-units.

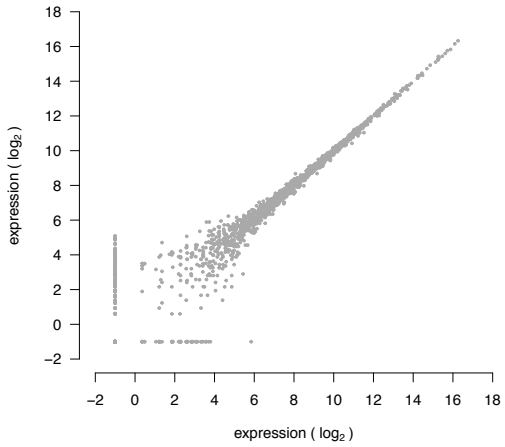


Figure 1: Replicate miRNA expression correlation plot.

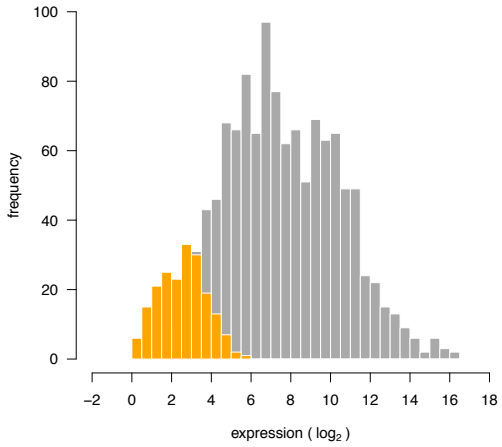
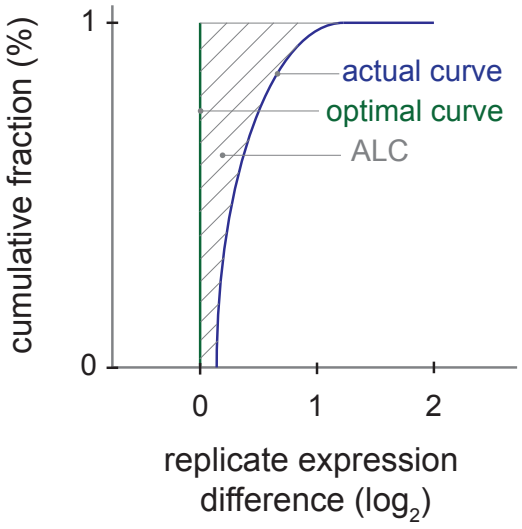


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.319, equivalent to a mean 1.247 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



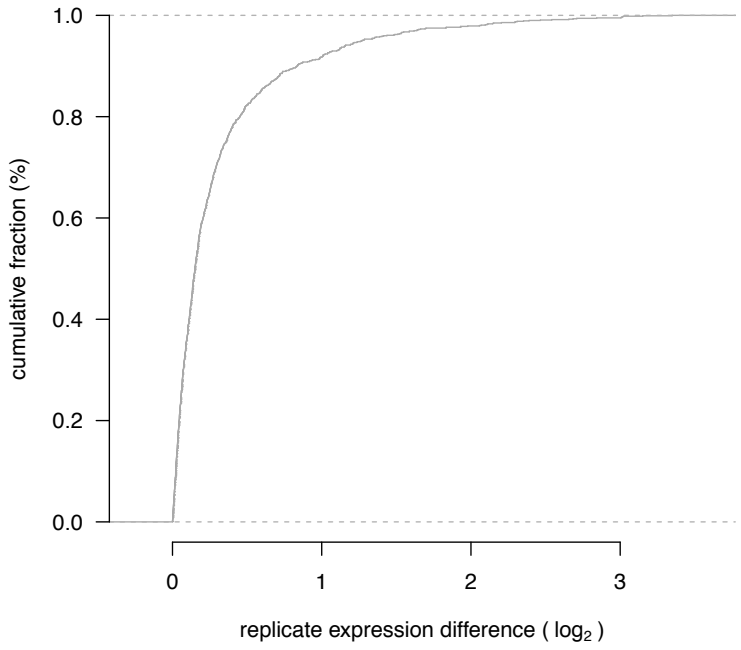
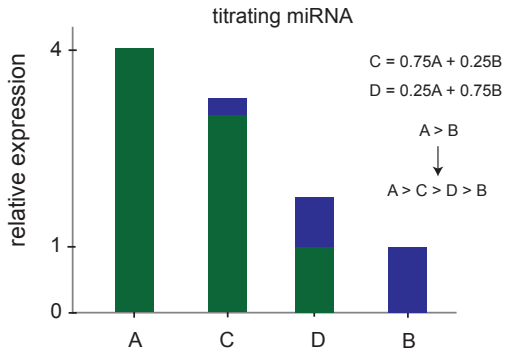


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

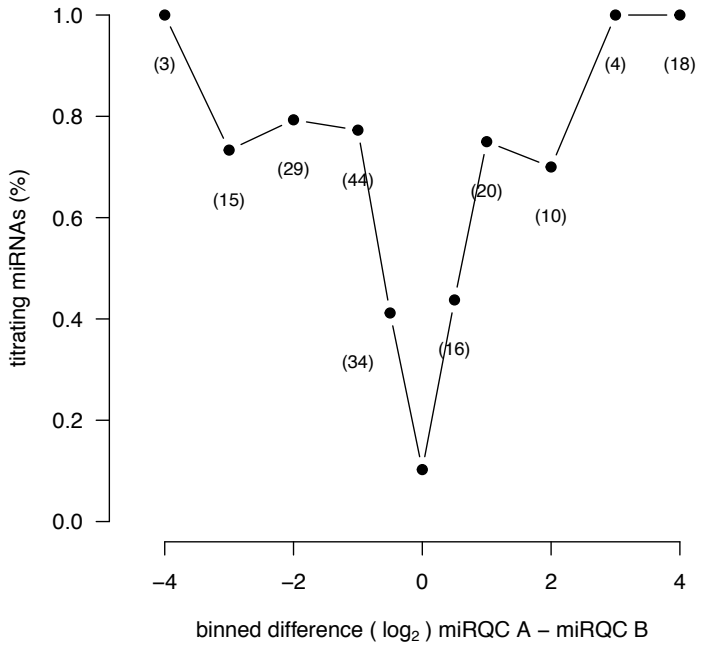
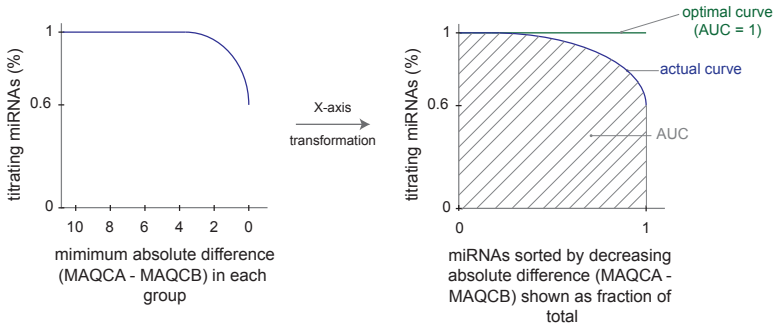


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.806.

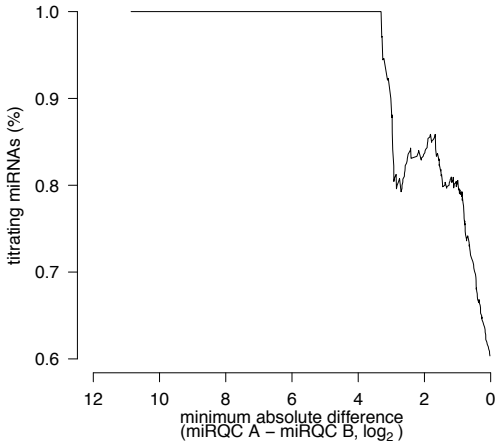


Figure 5: Titration response in function of absolute difference.

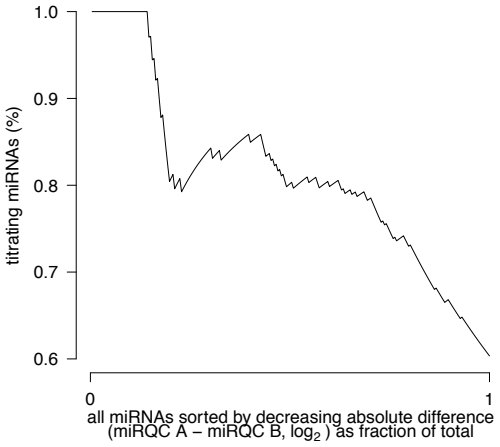


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

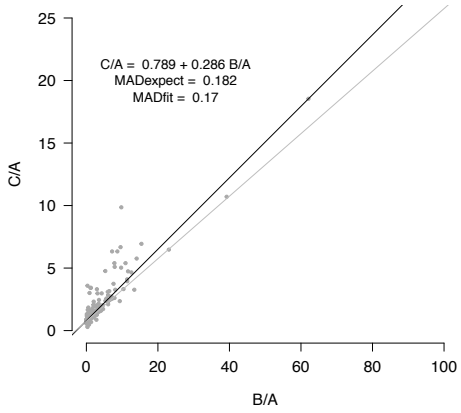


Figure 7: Titration response linearity.

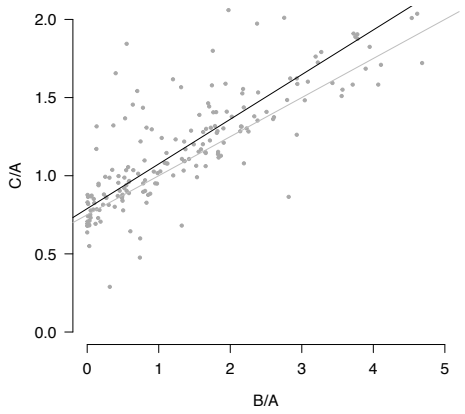


Figure 8: Figure 7 zoomed.

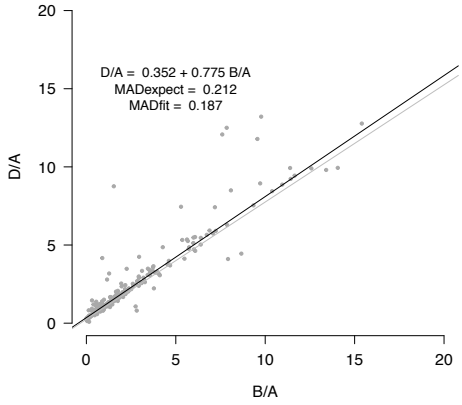


Figure 9: Titration response linearity.

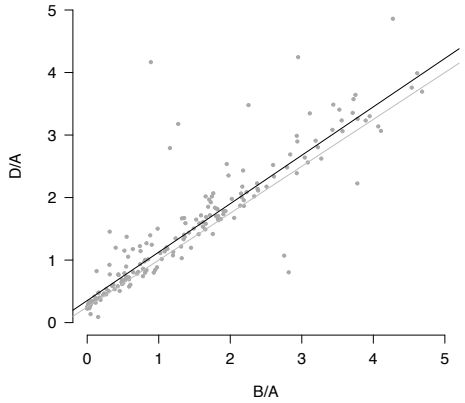


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 8.3% of all off-target combinations with a median relative cross-reactivity of 7.8%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	10.3	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	0.0	0.0	0.0
let-7b-5p	0.0	100.0	0.0	0.0
let-7c	0.0	5.2	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 1.1%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 14) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
1.4	0.3	2.2	0.6

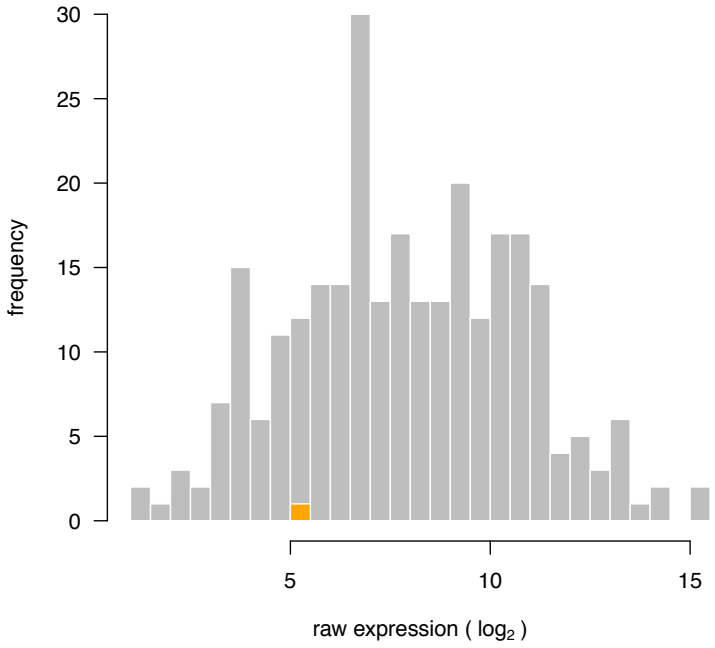


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 18 miRNAs were significantly upregulated and 10 miRNAs were significantly downregulated.

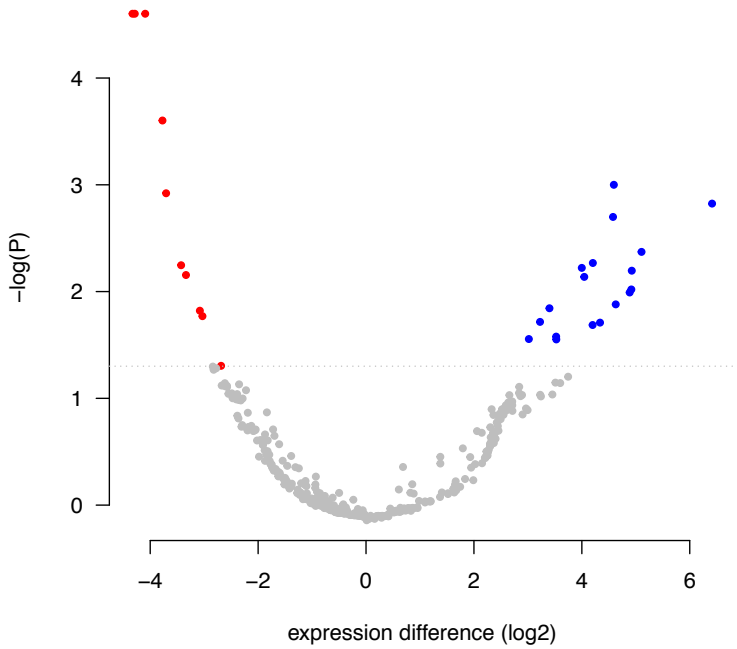


Figure 12: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

8 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	364
	fraction single positives (%)	14.04
	expression range (log2-units)	13.7
	ALC	0.319
titration	AUC titration response	0.806
	MADexpect (D/A)	0.182
	MADfit (D/A)	0.17
	MADexpect (C/A)	0.212
	MADfit (C/A)	0.187
specificity	off-target combinations with cross reactivity (%)	8.3
	median relative cross-reactivity (%)	7.8
non-template control	positive miRNAs (%)	1.1
serum miRNAs	detected miRNAs	NA
differential expression	significant down	10
	significant up	18

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 11

Illumina small RNA sequencing

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 4 reads.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 627 unique double positives were detected while the percentage of single positives was 9.23 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 17 $\log_2$-units.

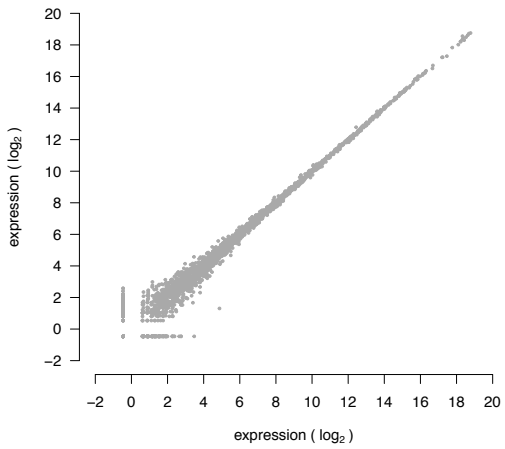


Figure 1: Replicate miRNA expression correlation plot.

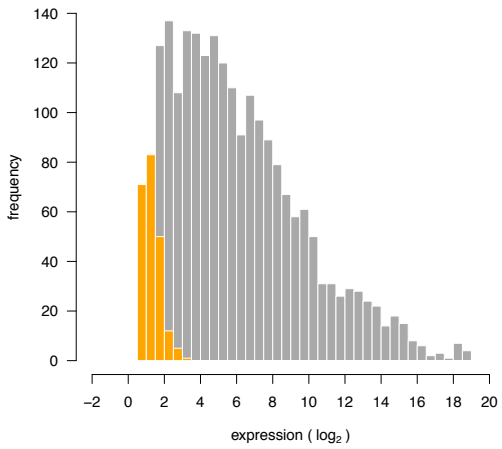
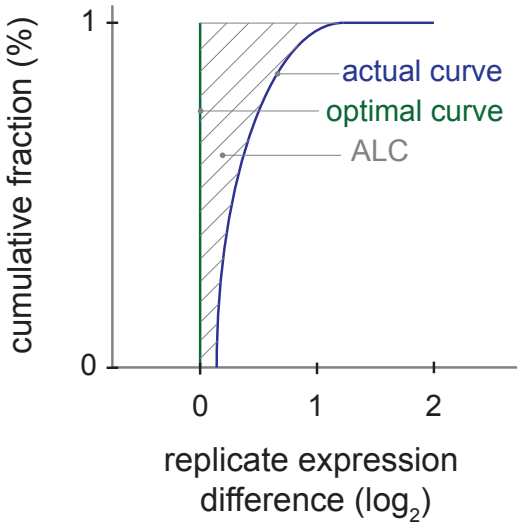


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.225, equivalent to a mean 1.169 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



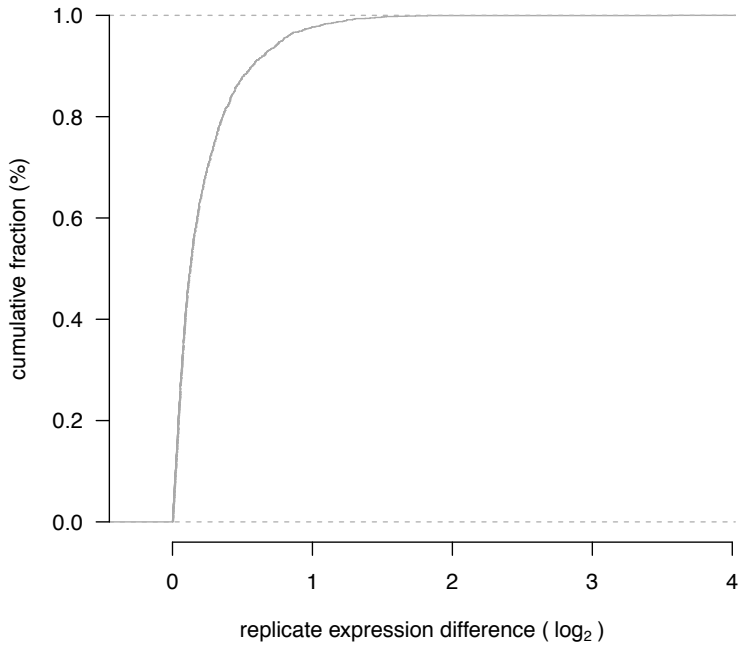
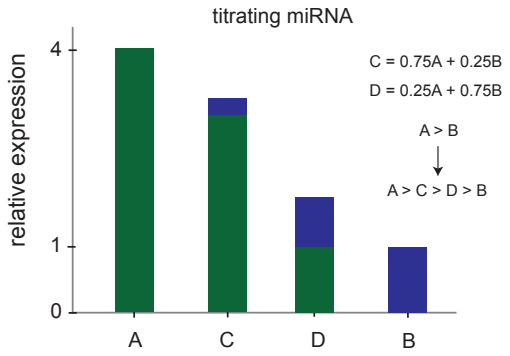


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

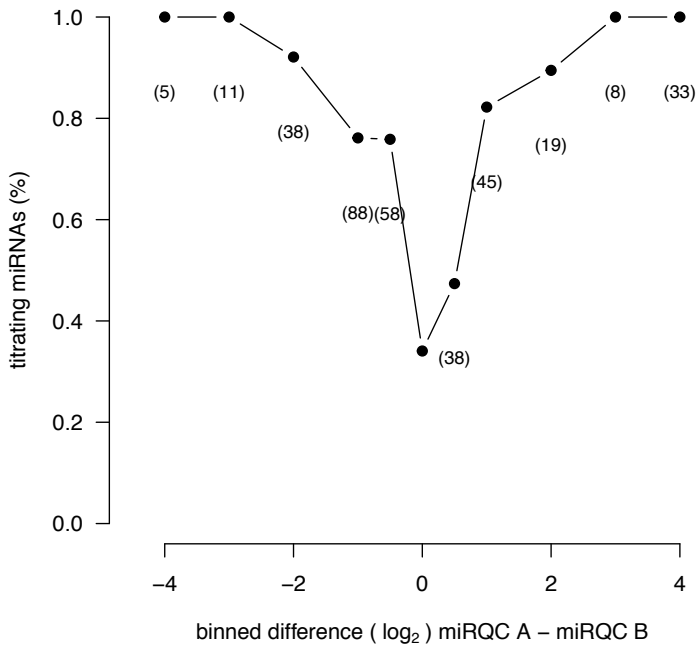
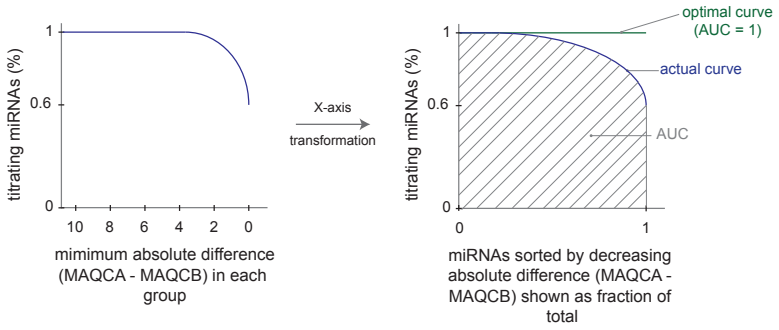


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.878.

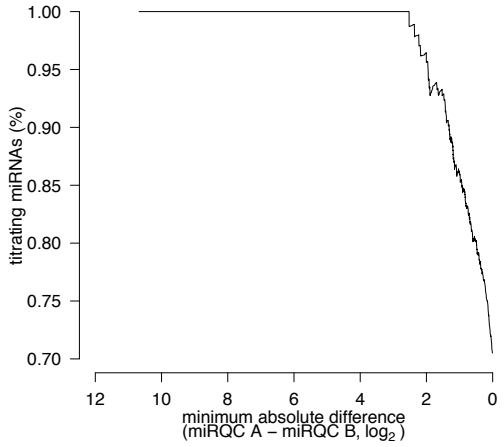


Figure 5: Titration response in function of absolute difference.

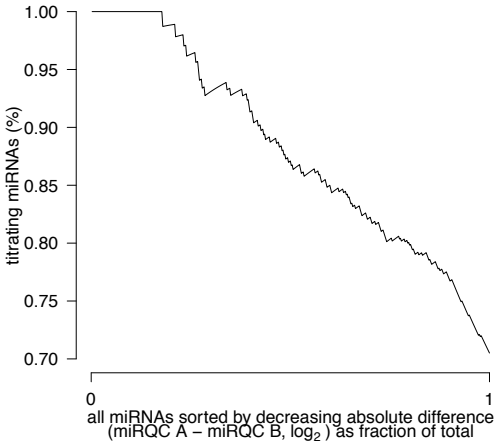


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

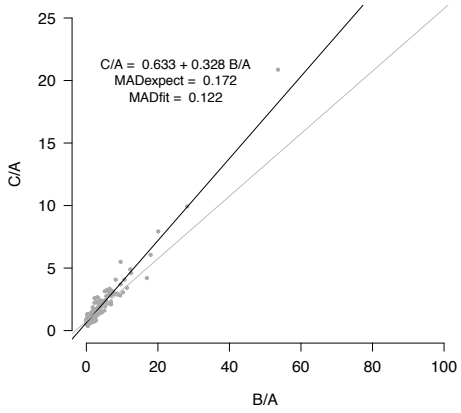


Figure 7: Titration response linearity.

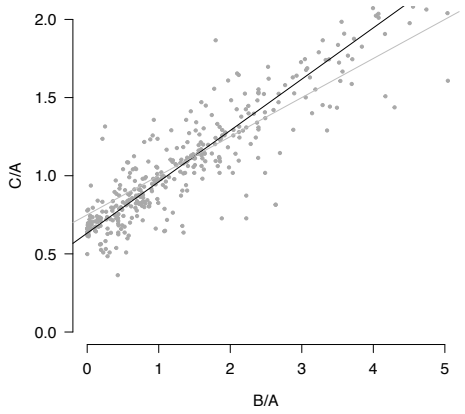


Figure 8: Figure 7 zoomed.

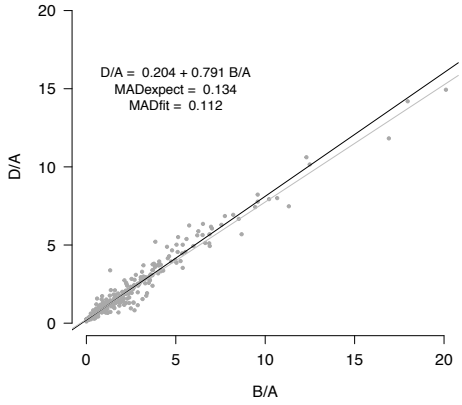


Figure 9: Titration response linearity.

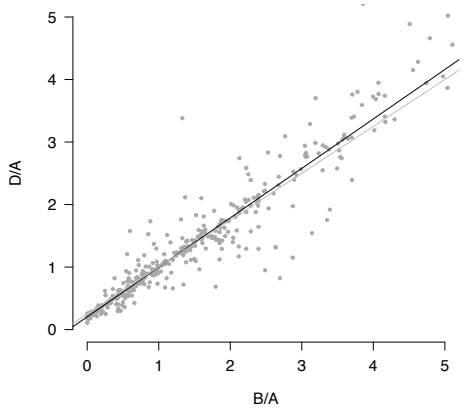


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 12.5% of all off-target combinations with a median relative cross-reactivity of 1.7%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
miR-302a-3p	100.0	0.0	0.0	0.0
miR-302b-3p	0.0	100.0	0.0	0.0
miR-302c-3p	0.0	0.0	100.0	0.0
miR-302d-3p	0.0	0.0	0.0	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
let-7a-5p	100.0	1.7	7.9	1.4
let-7b-5p	0.0	100.0	0.0	0.0
let-7c	0.0	0.0	100.0	0.0
let-7d-5p	0.0	0.0	0.0	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 0.92%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
0.8	1.1	1.1	0.6

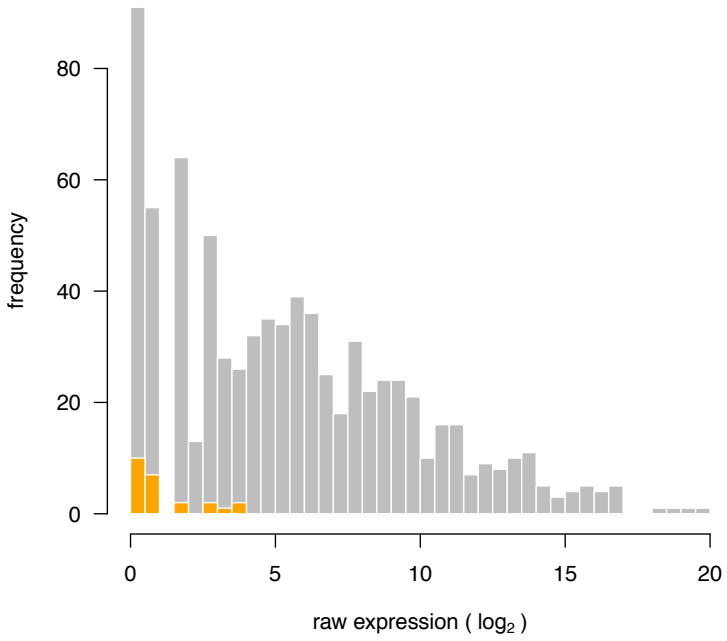


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 28 miRNAs were detected in all 4 samples while 43 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.501, equivalent to a mean 1.415 fold replicate expression difference.

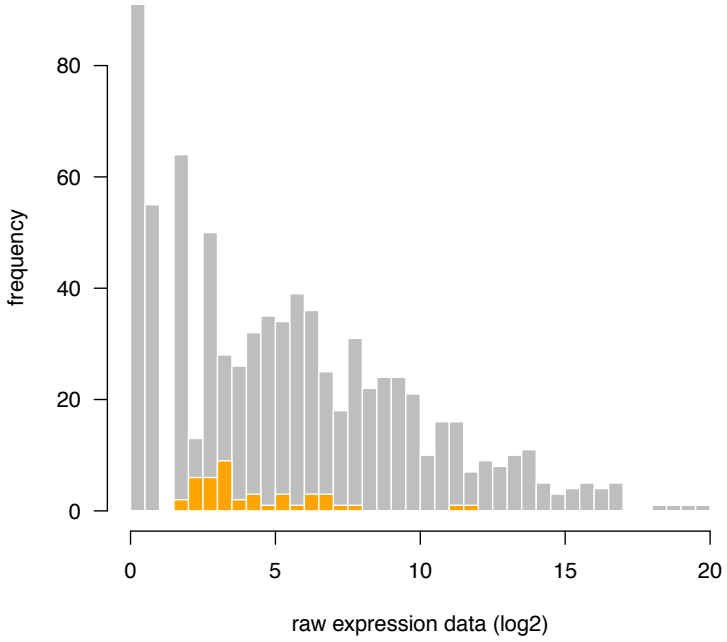


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

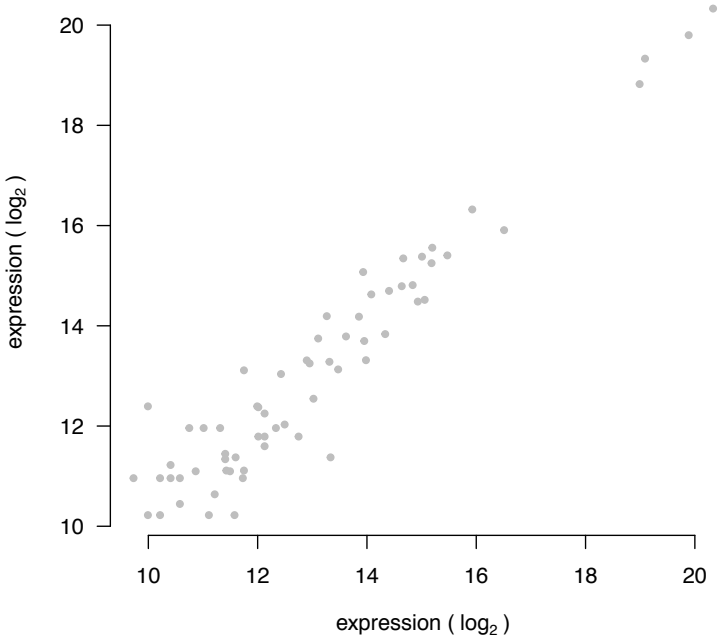


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 41 miRNAs were significantly upregulated and 29 miRNAs were significantly downregulated.

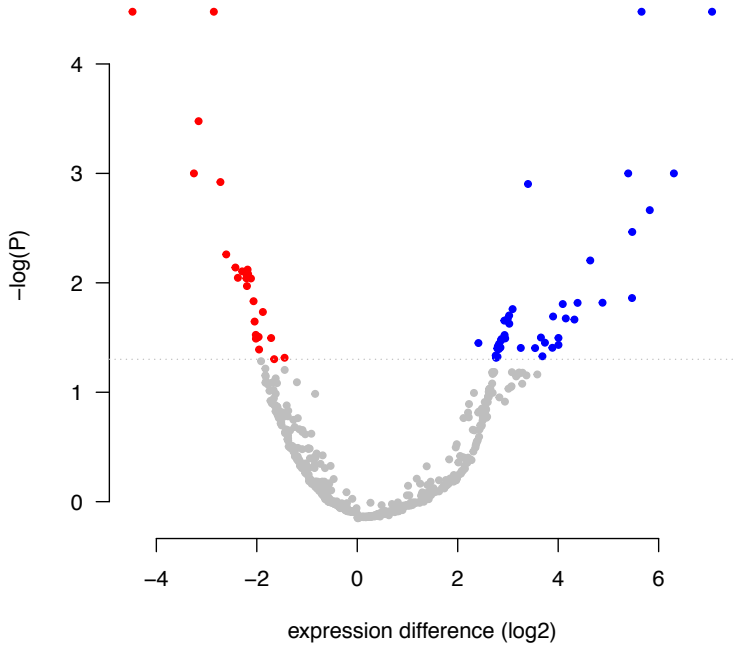


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	627
	fraction single positives (%)	9.23
	expression range (log2-units)	17
	ALC	0.225
titration	AUC titration response	0.878
	MADexpect (D/A)	0.172
	MADfit (D/A)	0.122
	MADexpect (C/A)	0.134
	MADfit (C/A)	0.112
specificity	off-target combinations with cross reactivity (%)	12.5
	median relative cross-reactivity (%)	1.7
non-template control	positive miRNAs (%)	0.92
serum miRNAs	detected miRNAs	43
differential expression	significant down	29
	significant up	41

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 12

Ion Torrent small RNA sequencing

Nature Methods: doi:10.1038/nmeth.3014

1 Sample set

The miRNA Quality Control (miRQC) study was performed using a set of 16 (mandatory) and 4 (optional) standardized positive and negative control samples to evaluate different aspects of platform performance. An overview of all 20 samples is provided in Table 1. Sample numbers will be used throughout the report.

sample number	sample name	spike	spike concentration
1	miRQC A	-	-
2	miRQC A	-	-
3	miRQC B	-	-
4	miRQC B	-	-
5	miRQC C	-	-
6	miRQC C	-	-
7	miRQC D	-	-
8	miRQC D	-	-
9	liver	miR-302a-3p	5e6
10	liver	miR-302b-3p	5e6
11	liver	miR-302c-3p	5e6
12	liver	miR-302d-3p	5e6
13	MS2 phage	let-7a-5p	5e6
14	MS2 phage	let-7b-5p	5e6
15	MS2 phage	let-7c	5e6
16	MS2 phage	let-7d-5p	5e6
17	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
18	serum	miR-10a-5p	6e4
		let-7a-5p	6e4
		miR-302a-3p	6e4
		miR-133a	6e4
19	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4
20	serum	miR-10a-5p	30e4
		let-7a-5p	12e4
		miR-302a-3p	3e4
		miR-133a	1.2e4

Table 1: Sample overview. The concentration of synthetic miRNAs in the liver and MS2 phage RNA is given as number of molecules per ug RNA, the concentration in serum RNA is given as number of molecules per 10 ul serum RNA

Nature Methods: doi:10.1038/nmeth.3014

2 Platform detection cutoff

Expression values from the four miRQC samples were pooled (samples 1, 3, 5 and 7) and compared to their respective replicates (samples 2, 4, 6 and 8). Missing expression values were imputed by replacing them with the lowest expression value of the respective miRNA minus 1 $\log_2$-unit. From this data we defined the single positive fraction (miRNAs detected in only one of the replicates) and the double positive fraction (miRNAs detected in both replicates). The detection cutoff was defined as such that it reduces the single positive fraction by 95%, here 5 reads.

3 Platform reproducibility

After applying the detection cutoff, platform reproducibility was visualized by means of a correlation plot (Figure 1) including both the double positives and single positives. The expression distribution of both the double and single positives is shown in Figure 2. A total of 651 unique double positives were detected while the percentage of single positives was 12.53 %. The expression values of the double positives were subsequently used to calculate the expression range. In order not to have outliers overestimate this range, 0.5% of the highest and lowest expressed miRNAs were removed, retaining 99% of the double positives. This results in an expression range of 12.6 $\log_2$-units.

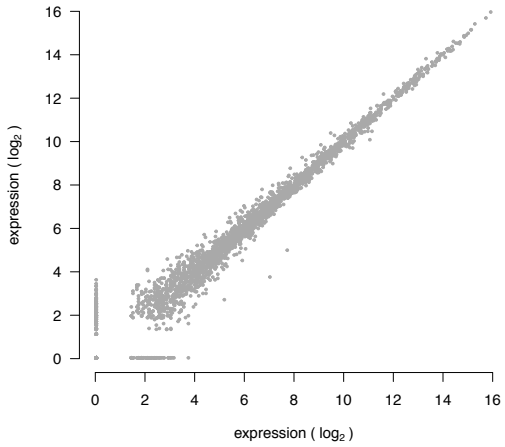


Figure 1: Replicate miRNA expression correlation plot.

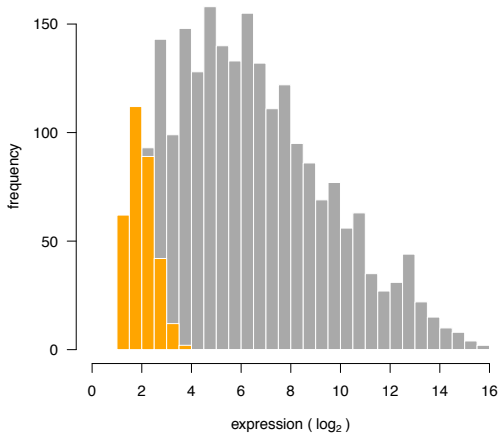
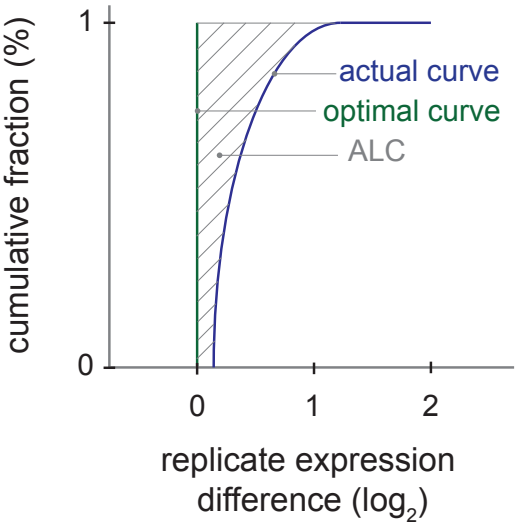


Figure 2: Distribution of single (orange) and double positive (grey) replicates.

Platform reproducibility was calculated based in the ALC-value as shown in the schematic below. We first calculated the absolute value of the expression difference between each replicate (taking into account only the double positives) and plotted the cumulative distribution of this difference (Figure 3). Reproducibility was then quantified as the area left of the cumulative distribution curve (ALC, as shown in the schematic). This area is equivalent to the mean replicate expression difference. The lower the ALC-value, the closer the actual curve resembles the optimal curve. Here, this area is 0.291, equivalent to a mean 1.223 replicate expression fold-change.

Nature Methods: doi:10.1038/nmeth.3014



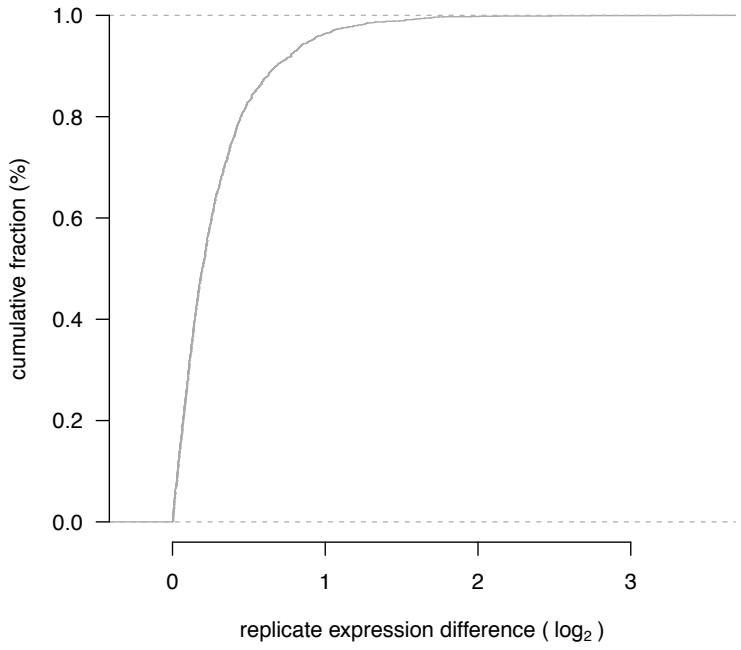
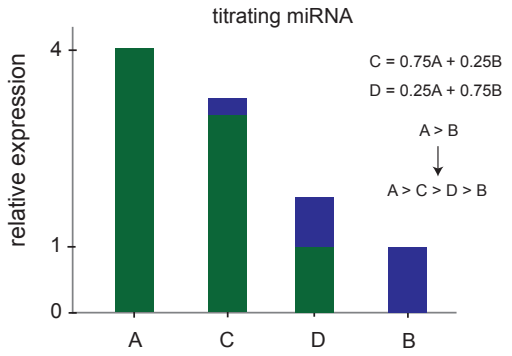


Figure 3: Cumulative distribution of replicate expression difference.

Nature Methods: doi:10.1038/nmeth.3014

4 Titration response

Using miRQC samples A, B, C and D (samples 2,4,6,8), miRNA titration response was calculated and evaluated in function of miRNA fold change. The scematic below illustrates the expression profile of a titrating miRNA with a given expression difference between miRQC A and B.



The titration response of all miRNAs is shown in Figure 4. miRNA fold changes are binned and the percentage of titrating miRNAs within each bin is plotted. The number of miRNAs per bin is listed in the plot.

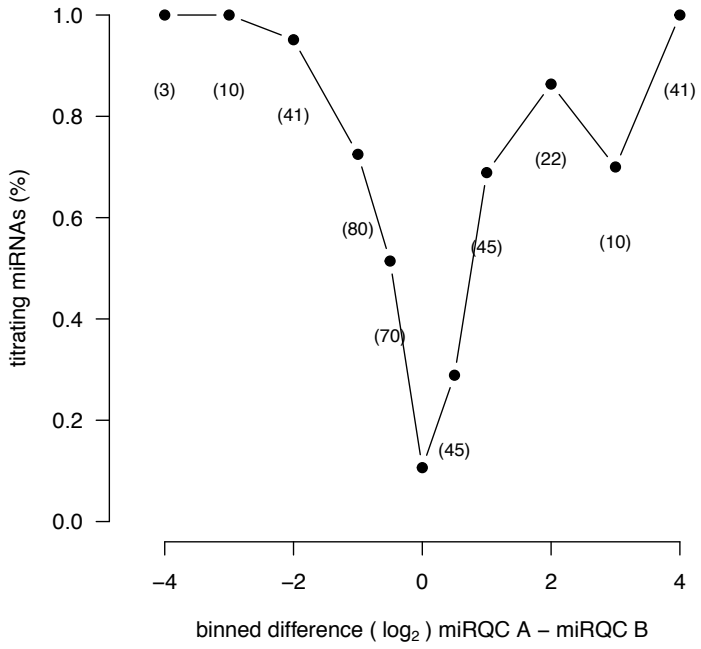
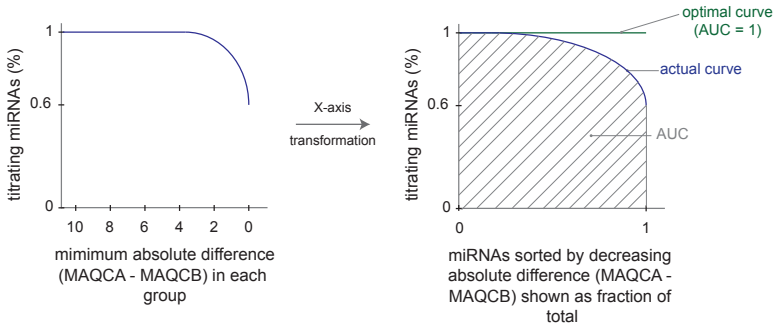


Figure 4: Titration response. Titrating assays in function of binned fold change.

Nature Methods: doi:10.1038/nmeth.3014

In order to better quantify the titration response, data was transformed as indicated in the schematic below.



First, miRNAs were sorted according to decreasing absolute difference between miRQC A and miRQC B whereby the percentage of titrating miRNAs was calculated for an increasing miRNA group size (with group size ranging from 1 to the total number of miRNAs). This percentage was plotted against the minimal expression difference (miRQC A - miRQC B) in each group (Figure 5). This representation allows to evaluate the global titration response in the data set. To express the titration response in a single measure, the percentage of titrating miRNAs was plotted in function of the groups size where group size is shown as a fraction of the total size (e.g. the largest group containing all miRNAs) (Figure 6). The area under this curve (AUC) was used as a measure of overall titration response and should be compared to the AUC for a perfect titration response curve which is equal to 1. Here, the AUC is 0.824.

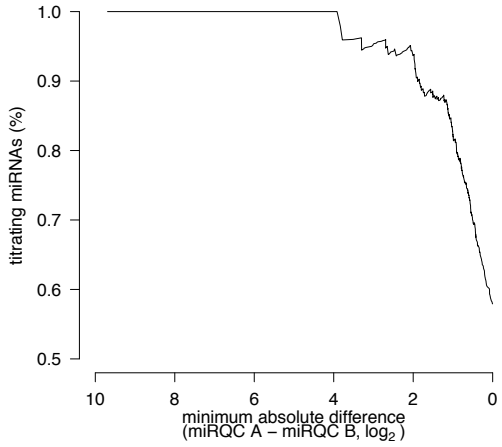


Figure 5: Titration response in function of absolute difference.

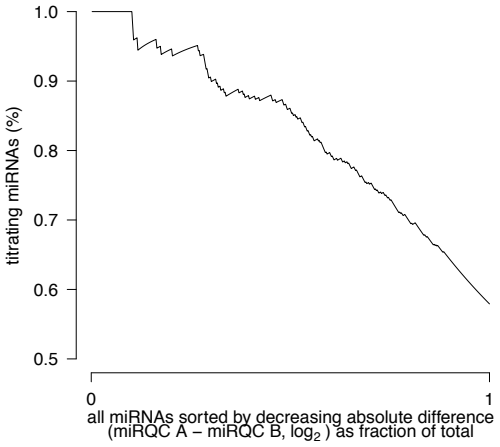


Figure 6: Titration response in function of fraction of total.

To assess titration response linearity, D/A and C/A expression ratios were plotted in function of B/A expression ratios (Figure 7-10) where the expected relation between D/A and B/A is defined by the following function

$$D/A = 0.25 + 0.75\ B/A$$

Nature Methods: doi:10.1038/nmeth.3014

$C/A = 0.75 + 0.25\ B/A$

Robust regression was applied to derive the intercept and slope of the linear regression function. The expected function is plotted in grey, the fitted regression line in black. The deviation between the theoretical function and the actual datapoints or between the fitted function and the actual datapoints is scored based on the median absolute deviation (MAD)

$MAD_{expect} = \text{median}(|y_{measured} - y_{expect}|)/0.675$

$MAD_{fit} = \text{median}(|y_{measured} - y_{fit}|)/0.675$

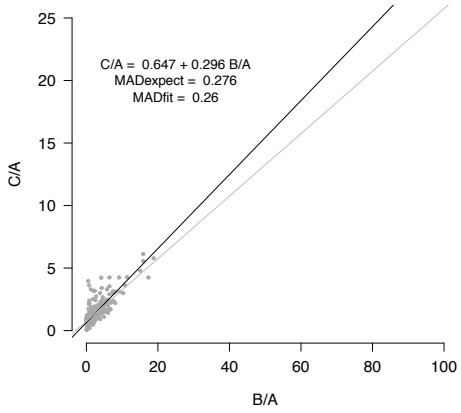


Figure 7: Titration response linearity.

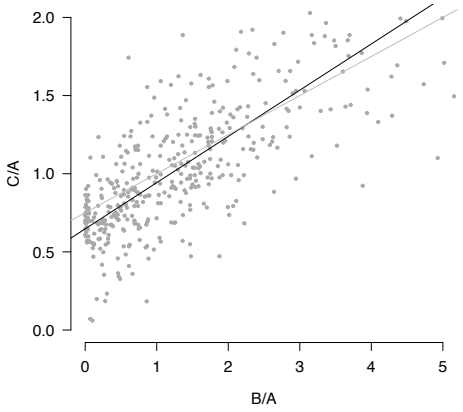


Figure 8: Figure 7 zoomed.

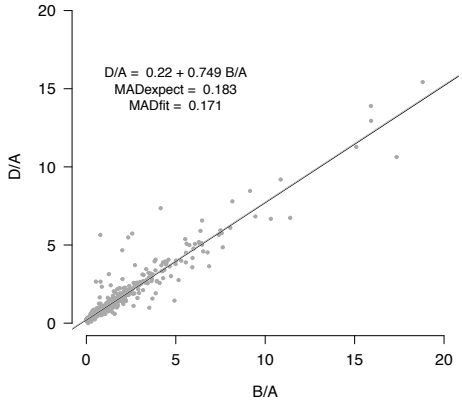


Figure 9: Titration response linearity.

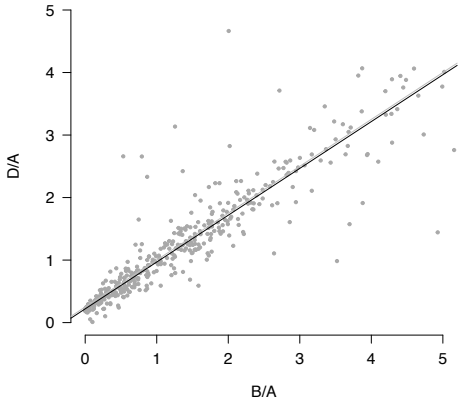


Figure 10: Figure 9 zoomed.

Nature Methods: doi:10.1038/nmeth.3014

5 Specificity

In order to quantify platform specificity, synthetic miRNAs from the let-7 and miR-302 family were spiked in MS2 phage RNA and total human liver RNA, respectively. Cross-reactivity among let-7 and miR-302 family members was evaluated for the respective spike-in experiments. For each sample, the signal was scaled relative to the perfect match. The results are listed in Table 2 and Table 3. Synthetics spikes are listed in columns, measured relative miRNA levels for all corresponding family members in rows. Cross reactivity was detected in 66.7% of all off-target combinations with a median relative cross-reactivity of 6%.

Table 2: miR-302 spike-in experiment

	miR-302a-3p	miR-302b-3p	miR-302c-3p	miR-302d-3p
1	100.0	5.9	0.0	13.2
2	0.0	100.0	4.1	0.0
3	0.0	0.0	100.0	0.0
4	5.5	6.9	6.1	100.0

Table 3: let-7 spike-in experiment

	let-7a-5p	let-7b-5p	let-7c	let-7d-5p
1	100.0	44.8	109.5	14.9
2	13.7	100.0	7.2	1.6
3	0.9	0.7	100.0	0.0
4	3.3	0.0	1.5	100.0

Nature Methods: doi:10.1038/nmeth.3014

6 Non-template control

To assess aspecific miRNA detection, the number of positive (expression above defined cutoff) miRNAs in the MS2 phage RNA samples (samples 13-16) were evaluated, excluding those miRNAs detecting the synthetic spikes. The percentage of positive miRNAs (relative to the number of unique double positives, section 2) for each of the MS2 phage RNA samples is shown in Table 4. The mean percentage of positive miRNAs is 13.21%.
The distribution of expression values for all positive miRNAs in a representative MS2 phage RNA sample (sample 13) is shown together with the expression values for miRQC A (sample 1) in Figure 11.

Table 4: percentage of positive miRNAs

MS2_1	MS2_2	MS2_3	MS2_4
14.8	12.0	19.4	6.8

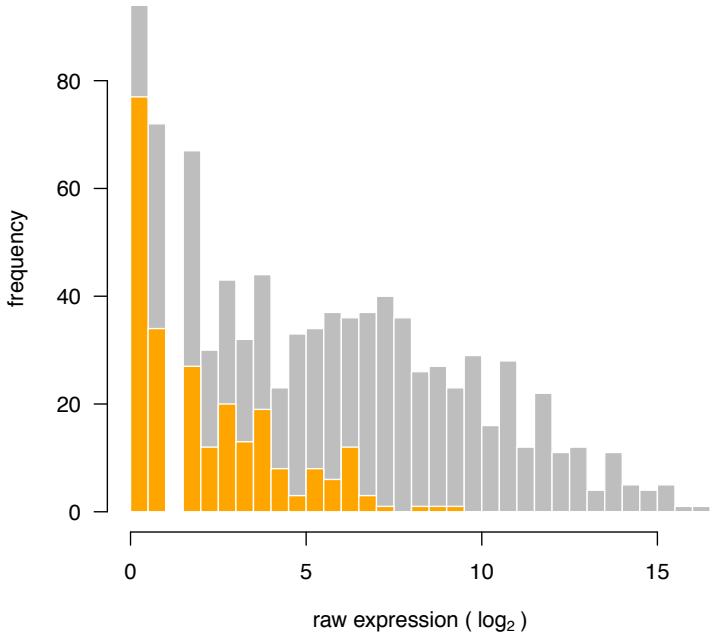


Figure 11: Distribution of expression values for positive miRNAs in the MS2 phage RNA sample (orange bars) and miRQC sample A (grey)

Nature Methods: doi:10.1038/nmeth.3014

7 Expression profiling of serum miRNAs

To evaluate the platform's capacity to detect miRNAs in RNA isolated from human serum, four replicate serum RNA samples (samples 17-20) were quantified. In total, 31 miRNAs were detected in all 4 samples while 69 miRNAs were detected in at least 2 samples. The distribution of the raw expression values for miRNAs detected in all 4 samples (calculated as the mean raw expression in all four serum samples) is plotted together with the distribution of raw miRNA expression values of miRQC A (sample 1) (Figure 12).
To assess reproducibility of low-copy miRNA expression values, normalized miRNA expression values from both serum samples (samples 17 and 19) were pooled and compared to the replicates (samples 18 and 20). Expression correlation for double positives is shown in Figure 13. Based on the double positives, reproducibility was quantified by means of the ALC (see section 2). This value is 0.892, equivalent to a mean 1.856 fold replicate expression difference.

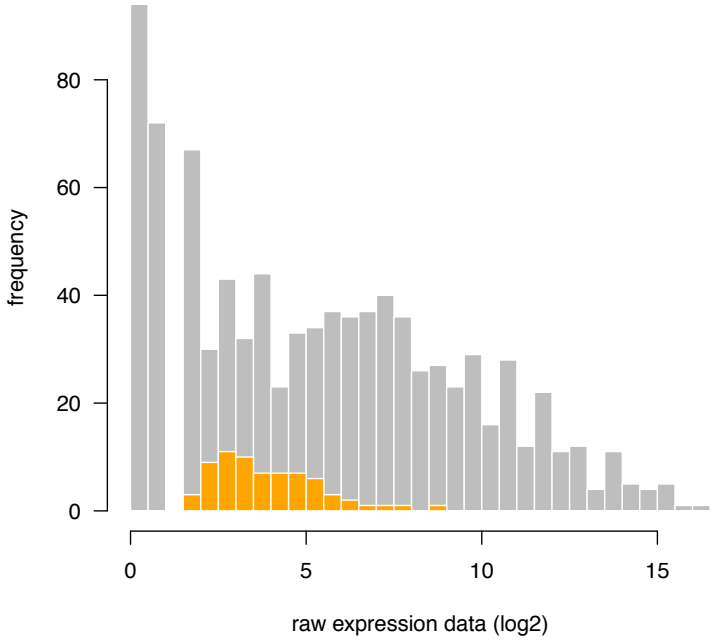


Figure 12: Distribution of miRNA expression values (log2) for miRQC A (grey bars) and serum RNA (orange bars)

Nature Methods: doi:10.1038/nmeth.3014

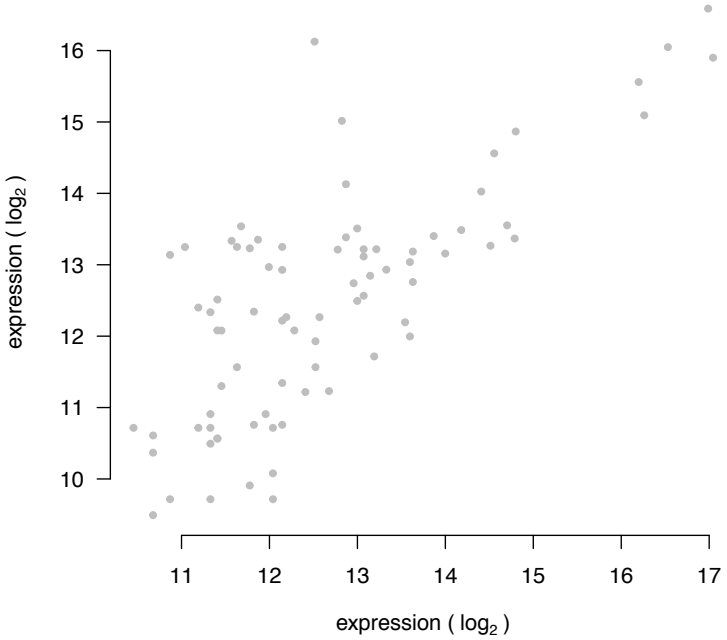


Figure 13: Reproducibility of measured miRNA expression values in replicate serum samples

Nature Methods: doi:10.1038/nmeth.3014

8 Differential miRNA expression

The capacity to detect differential miRNA expression was assessed by comparing miRQC A + miRQC C (group 1) to miRQC B + miRQC D (group 2). Missing miRNA expression values were imputed based on the lowest expression value of the respective miRNA minus one $\log_2$-unit. P-values were calculated using Rank Products with 1000 permutations. Results are visualized in a volcano plot (Figure 14). Significant miRNAs were defined as having a pfp-value (percentage-false-positives) < 0.05. In total, 37 miRNAs were significantly upregulated and 35 miRNAs were significantly downregulated.

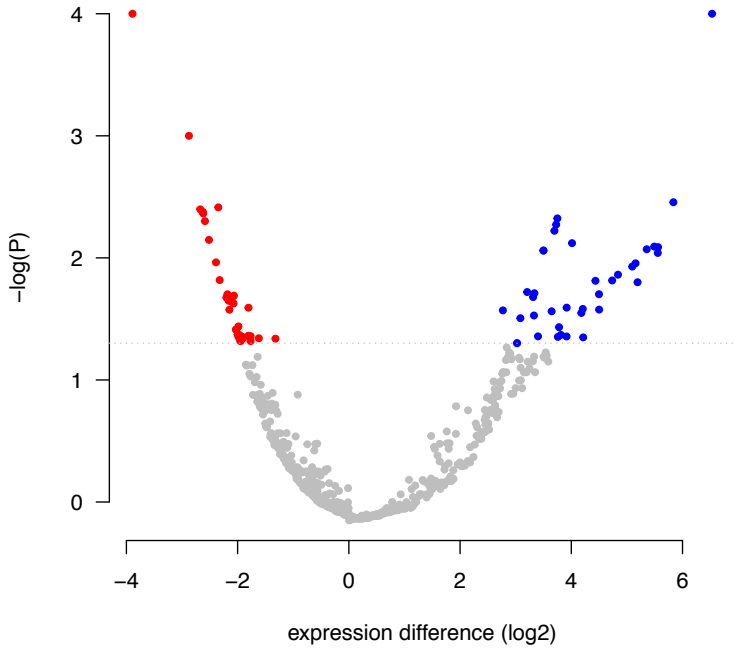


Figure 14: Volcano plot.

Nature Methods: doi:10.1038/nmeth.3014

9 Summary table

Table 5 summarizes all performance parameters for the different experiments. Performance parameter values that could not be calculated (because of missing data) are listed as NA.

Table 5: summary table

experiment	parameter	value
reproducibility	unique double positives	651
	fraction single positives (%)	12.53
	expression range (log2-units)	12.6
	ALC	0.291
titration	AUC titration response	0.824
	MADexpect (D/A)	0.276
	MADfit (D/A)	0.26
	MADexpect (C/A)	0.183
	MADfit (C/A)	0.171
specificity	off-target combinations with cross reactivity (%)	66.7
	median relative cross-reactivity (%)	6
non-template control	positive miRNAs (%)	13.21
serum miRNAs	detected miRNAs	69
differential expression	significant down	35
	significant up	37

Nature Methods: doi:10.1038/nmeth.3014

Supplementary Note 13

This section provides a more elaborate discussion supplied by each vendor and deals with platform performance and miRQC study design.

Exiqon

Serum sample quality

The call-rate seen in the serum-samples in the miRQC study is lower than what we typically see in other serum samples at Exiqon[1]. This difference may come from differences in RNA isolation methods or quality of the serum. Alternative extraction methods may provide higher call-rates from the same samples on all platforms[2,3]. However, this is equal for all platforms – and thus the comparative differences in call-rates should be valid.

Call rate

When comparing serum call rates between different platforms it is important to bear the platform specificity in mind. It is evident from the experiment with samples containing 4 synthetic spike-ins which are members of a closely related microRNA family, that less specific platforms will have signal from more assays per microRNA. Thus, the total call-rate may for less specific platforms be higher than for highly specific platforms because the call rate for the former include false positive signals caused by cross-reactivity between microRNA family members – most notable for the 4 spike-ins added to all serum samples.

The Exiqon platform high specificity is obtained with LNA™ PCR primers to provide true signals and in addition further specificity of signal is obtained through the melting curve analysis based filtering which removes some data points. The consequence of this is lowering the call-rate but increasing the data quality.

Single positives.

The occurrence of single positives is a by-product of the data-filtering applied in pre-analysis data handling at Exiqon. The filtering involves removing signals where the amplicon melting point either is somewhat off relative to the expected value or where the amplicon melting curve appear to have shoulders indicating by-products of the amplification. The result of the filtering is both removal of single positives as well as generation of new single positives which Exiqon considers to be an active exploitation of the potential of SYBR Green for analysis of amplification products and generating trustworthy results.

MS2

Exiqon recommends running a negative control such as MS2 (or better, a mock RNA isolation). This should be used to remove unspecific signal. The background subtraction has not been performed in this paper.

New platform

After the finalization of the experimental part of this paper, Exiqon has launched a new version of the platform. This improved platform uses new mastermix reagents, and some assays have been re-designed for improved function. The result is a platform with less background signal, combined with even better sensitivity and improved mismatch discrimination.

Quanta Biosciences

The qScript™ microRNA Quantification System from Quanta Biosciences can accommodate a wide range of RNA sample types and RNA input amounts in both the cDNA synthesis step and in the qPCR step depending on the specific requirements of different end-user applications. This was nicely demonstrated in the miRQC study where the qScript™ microRNA System was able to successfully profile the entire range of RNA samples provided, including the most challenging serum samples, and performed comparatively well in every area of the study.

To reliably detect and quantify low abundant microRNAs or to verify the absence of specific microRNAs the system must be both highly sensitive and specific, two of the criteria extensively examined in the study. Following an initial profiling experiment, results can be validated by using more or less cDNA template in each qPCR reaction as needed depending on the relative abundance of the microRNAs of interest. In cases where the microRNAs are abundant, the specificity of individual assays can be increased by adding less cDNA template (i.e. 0.1 ng or less) to the qPCR reaction and by increasing the annealing temperature of the PCR cycling conditions. In cases where the microRNAs are absent or their abundance is low, higher amounts of cDNA template (i.e. 10 ng or more) can be added to the qPCR reactions. In addition, we have found that the specificity of some microRNA assays can be increased by raising the temperature of the reverse transcriptase reaction (i.e. from 42 ˚C to 45 ˚C or higher). Adding more cDNA template to the qPCR will typically result in an increase in both the specific and non-specific amplification signals and in these cases it becomes critical to be able to distinguish them. This can be done by preparing samples without the addition of poly(A) polymerase. To verify and validate a specific assay signal there must be a significant difference in the qPCR results between cDNA samples prepared with and without poly(A) polymerase. The qScript microRNA System is unique among the other study participants in providing an assay format that allows direct measurement of assay background and detection of false positive assay signals. In addition, amplicon melt profiles can be used to detect the presence of non-specific amplification products. Ideally this analysis can be performed for each microRNA assay and in each sample.

Quanta Biosciences was pleased to take part in the miRQC study. As a result of this work, new and interesting methods have been developed that allow the analysis and comparison of data from very disparate instrument and reagent systems. This study will thus serve as an important reference to end-users of these technologies and provide a better understanding of the specific strengths and weaknesses of the various microRNA detection and quantification systems. Overall, the qScript™ microRNA Quantification System from Quanta Biosciences performed exceedingly well in every test and offers both high quality and high value compared to the other technology platforms that participated in the study.

Qiagen

For miScript® PCR System testing, plate one of the three-plate Human miRNome miScript miRNA PCR Array (MIHS-3216Z) was used. Plate one represents a panel of the most well characterized miRNAs that are annotated in miRBase. At the time the study commenced, over 1400 additional miScript bench-validated assays covering miRNAs annotated through miRBase V18 were available. Also, disease- and pathway-focused panels of miRNA assays (miScript miRNA PCR Arrays), designed to profile the most relevant and cutting-edge areas of science are available. Furthermore, the miScript PCR System is compatible with a wide range of real-time PCR instruments including microfluidics-based instruments.
The miScript PCR System workflow used for this study omitted an optional preamplification, using the miScript PreAMP PCR Kit and associated miScript PreAMP Primer Mixes. This module can be incorporated into the standard workflow to enable profiling of otherwise unsuitably low RNA amounts or to enhance detection of extremely rare targets. Using a highly multiplex, PCR-based preamplification approach, up to 400 miRNA-specific cDNA targets can be amplified in one reaction, and the total assay coverage matches that of the miScript qPCR assay coverage. Although the miScript performed extremely well on the serum samples in this study, a preamplification step would have added additional extremely low abundance miRNAs to the identifiable targets in these samples.

Wafergen Biosystems

The WaferGen SmartChip system demonstrates excellent qPCR technical performance in a high-density qPCR array. In particular, titration response, reproducibility, and titration accuracy distinguish the WaferGen platform as best in class for technical performance among qPCR platforms. (Figures 2A-I)

The study methodology is designed to find weaknesses in the platform performance, and does highlight some difficult cases for both WaferGen and other vendor platforms. In particular, specificity among closely related species of miRNA is not ideal, and is similar to other platforms. (Figure 4)

With a 100nL nominal well volume, and up to 5184 distinct reaction wells per chip, the SmartChip platform has a demonstrable performance advantage over other qPCR systems in terms of titration response linearity, reproducibility of measurements, and quantitative accuracy for high to medium expression level targets (Supplemental Figures 2A and 2B). While most targets can be reliably measured without the expense and complexity of pre-amplification, as demonstrated, the platform does not limit the use of other reagents, and is compatible with both SyBr and probe-based chemistries. The MyDesign platform, with its open architecture makes use of this flexibility in enabling the use of the preferred chemistry particular to the project at hand.

Measuring up to 1296 miRNAs on a single chip in quadruplicate with excellent reproducibility and robust quantification performance provides the ability to perform discovery experiments in a high throughput manner at low cost per

sample. With the unmatched flexibility of the MultiSample NanoDispenser, WaferGen offers an adaptable and configurable platform for running both preprinted and user-defined content or alternate chemistries with the promise of robust, repeatable results.

Affymetrix

Cross comparison of platforms is an extremely difficult task. The miRNA quality control study was limited in analytical power. Utilization of a single method to analyze different platforms, even within hybridization platforms, is almost impossible and will favor platforms over other platforms. While we applaud the study organizers for their efforts to make cross platform miRNA comparisons, we must make several notes as they pertain to the Affymetrix miRNA solution. Affymetrix latest design, the 3.0 miRNA was not utilized in this study. Further, data normalization techniques that can be performed by our latest software (Expression Console) were not employed which is our typical recommendation. In addition, the study used detection cutoffs which we highly discourage as they artificially reduced the ability to detect low-level miRNAs on our platform. In other studies[4] and in practical use the inter- and intra-reproducibility of the Affymetrix platform has been shown to be >0 .95 correlation, the dynamic range is > 4.8 $\log_{10}$ and the limit of detection is set at 1.0 amol. Lastly, the study does not point out that the Affymetrix platform is currently the only platform capable of detecting pre-miRNA in addition to mature miRNA in over 153 organisms.

Agilent

The Agilent miRNA microarray platform utilizes hybridization as a means to measure the expression profile of mature miRNAs in a highly multiplexed assay. The version of the microarray used in this study measures over 1200 miRNAs from miRBase v16. The labeling strategy utilized directly labels miRNA without any RNA or signal amplification, and labeled miRNAs are then hybridized to the highly specific microarray probes under stringent conditions. The direct labeling process, stringent hybridization conditions and multiple replicate microarray probes per miRNA are key factors in the performance of the Agilent miRNA microarray.

The performance strengths of the Agilent microarray platform as identified in this study relate directly to the labeling and hybridization strategy, as well as to the high quality microarrays manufactured by Agilent. Good system reproducibility, as clearly identified in the study (Figure 2E-I), is critical for obtaining high quality reliable miRNA profiling results. Agilent's reproducibility also, in part, explains the platforms exemplary performance in measuring an accurate titration response (Figure 2A-2D). Titration response is critical as it demonstrates the ability to consistently detect small expression changes across samples.

In those portions of the study where accuracy and sensitivity were measured without the use of serum samples, the Agilent platform performed remarkably well, whether looking at the accuracy of log ratio results across the titration

samples, or the sensitivity based on specific detection of miRNAs in complex samples. Sensitivity and accuracy in serum samples, however, is heavily dependent on the method of RNA extraction. Based on the data obtained from the serum samples and our previous experience with serum samples, we conclude that the serum samples provided by the study coordinator may not have been extracted using a methodology that is optimal for the Agilent platform, as we know that RNA isolation from serum is critical. It also appears that some specific labeling inhibition occurred in these samples, based on the spike-in results.

Overall, we believe that the results of this study highlight some key advantages in the measurement of miRNA expression using the Agilent platform. The excellent consistency in measuring low level differential expression of many miRNAs from a small amount of RNA with a simple and straightforward workflow make the Agilent platform particularly ideal for studies where detection of low level differential expression for multiple miRNAs is required. Agilent's platform is also well suited to almost any study where reliable miRNA profiling is desired across a broad range of samples.

Life Technologies (TM, TMp and OA)

Life Technologies would like to thank the authors for organizing the miRQC study and for their invaluable guidance and discussions on the interpretation of the results.

Life Technologies would also like to acknowledge Genome Explorations (Memphis, TN, USA), who ran the experiments with the TaqMan® MicroRNA Array Cards (TM/TMp), for their kind cooperation.

We are delighted to see that our TaqMan® platforms for profiling were found to be among the best on a broad set of metrics, such as accuracy, quantitation of low abundant miRNAs, sensitivity, detection rate in serum, specificity, and number of false positives.

Nanostring

As a leading provider of highly multiplexed digital genomic analysis products, NanoString Technologies was delighted to participate in the microRNA Quality Control (miRQC) study to demonstrate the performance of the nCounter® system. The NanoString nCounter technology is a highly multiplexed, amplification free, direct digital detection assay that delivers highly reproducible, specific, accurate and sensitive results for detecting RNA, miRNA or DNA targets over a wide dynamic range. Several performance characteristics measured in the miRQC study highlight the strengths of the nCounter platform:

- Accuracy – the nCounter platform provided highly accurate results (Fig. 3A).
- Zero false positive results – the nCounter platform was the only platform in the miRQC study that generated no false positives (Fig 4C).
- Specificity – the nCounter platform was one of the most specific platforms when differentiating between highly homologous miRNA families (Fig. 4E).

NanoString believes that the miRQC study data combined with the detection methodology of the nCounter system should provide users with confidence that their results reflect the true underlying biology of their samples. Ease-of-use and cost per sample, two factors not addressed in the study, are also key decision factors when choosing a platform, and both are key strengths of the nCounter system.

Since the processing of the miRQC samples two significant improvements have been made to the nCounter Analysis System and the nCounter miRNA Expression Assay that we believe would have positively impacted the miRQC data had they been available at the time of sample processing.

1. An update to the nCounter Prep Station was released in April 2013 which provides new purification protocols that offer both improved reproducibility and sensitivity. The new protocol further optimizes the binding of probe-miRNA target complexes to the cartridge surface thereby increasing read counts. As the nCounter system is a digital technology increasing read counts improves both sensitivity and the reproducibility of low expressing targets. Data presented in the nCounter Tech Note entitled "nCounter® System Enhancement Provides Improved Fold-Change Sensitivity and Increases the Number of Detectable Genes", highlights that the number of detectable miRNAs can increase significantly with an accompanying increase in statistically significant fold change sensitivity.
2. The Tech Note released in June 2013 entitled "nCounter miRNA Analysis in Plasma and Serum Samples" discusses the current challenges associated with miRNA studies in blood plasma and serum samples, points out steps in the processing of collected blood which can have an impact on sample quality and elucidates the ways in which variables in sample preparation can be controlled to produce reliable data using nCounter® miRNA Assays.

Both the Tech Notes described above are available for download at nanostring.com

The NanoString nCounter® Human v2 miRNA Expression Assay Kit used in the miRQC study profiles 800 human miRNAs from miRBase v.18. NanoString also offers miRNA kits for Mouse, Rat and Drosophila, plus custom À La Carte miRNA Panels are available for performing larger validation studies on smaller sets of miRNAs. NanoString's miRGE assays allow mRNAs and miRNAs to be analyzed in the same reaction. All nCounter miRNA Assay Kits have been shown to provide highly concordant data between fresh frozen and FFPE samples.

References

1. Blondal, T. *et al.* Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods* **59,** S1–6 (2013).
2. McAlexander, M. A., Phillips, M. J. & Witwer, K. W. Comparison of Methods for miRNA Extraction from Plasma and Quantitative Recovery of RNA from Cerebrospinal Fluid. *Front Genet* **4,** 83 (2013).
3. Andreasen, D. *et al.* Improved microRNA quantification in total RNA from clinical samples. *Methods* **50,** S6–9 (2010).

4. Kolbert, C. P. *et al.* Multi-platform analysis of microRNA expression measurements in RNA from fresh frozen and FFPE tissues. *PLoS ONE* **8,** e52517 (2013).