Evaluating the consistency of differential expression of microRNA detected in human cancers

Xue Gong1#, Ruihong Wu1#, Hongwei Wang1, Xinwu Guo1, Dong Wang1,2, Yunyan Gu1, Yuanyu Zhang1, Wenyuan Zhao1, Lixin Cheng1, Chenguang Wang1§, Zheng Guo1,2§

1 College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150086, China
2 Bioinformatics Centre, School of Life Science, University of Electronic Science and Technology of China, Chengdu, 610054, China
# contributed equally, § Corresponding author

Running title: Consistency of differential expression of miRNA

Keywords: Cancer/Consistency/Disease biomarker/Function/miRNA

Abbreviation list: DE = Differentially Expressed; miRNA = microRNA; Benjamini and Hochberg = BH.

Grant support: This work was supported by National Natural Science Foundation of China (grant No: 30770558, 30970668, 81071646), Excellent Youth Foundation of Heilongjiang Province (grant No. JC200808), Natural Science Foundation of Heilongjiang Province of China (grant: QC2010012) and Scientific Research Fund of Heilongjiang Provincial Education Department (NO: 11541156).

Requests for reprints:
Zheng Guo, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China 150086. Phone: (86 451) 8661-5933; Fax: (86 451) 8666-9617. Email: guoz@ems.hrbmu.edu.cn

Chenguang Wang, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China 150086. Phone: (86 451) 8661-5933; Fax: (86 451) 8666-9617. Email: cgwang1231@hotmail.com
Abstract

Differential expression of microRNA (miRNA) is involved in many human diseases and could potentially be used as a biomarker for disease diagnosis, prognosis and therapy. However, inconsistency has often been found among differentially expressed (DE) miRNAs identified in various studies when using miRNA arrays for a particular disease such as a cancer. Before broadly applying miRNA arrays in a clinical setting it is critical to evaluate inconsistent discoveries in a rational way. Thus, using datasets from two types of cancers, our study demonstrates that the DE miRNAs detected from multiple experiments for each cancer exhibit stable regulation direction. This result also indicates that miRNA arrays could be used to reliably capture the signals of the regulation direction of DE miRNAs in cancer. We then assumed that two DE miRNAs with the same regulation direction in a particular cancer play similar functional roles if they regulate the same set of cancer-associated genes. Based on this hypothesis, we proposed a score to assess the functional consistency between DE miRNAs separately extracted from multiple studies for a particular cancer. We showed although lists of DE miRNAs identified from different studies for each cancer were highly variable, they were rather consistent at the level of function. Thus, the detection of DE miRNAs in various experiments for a certain disease tends to be functionally reproducible and capture functionally related differential expression of miRNAs in the disease.
**Introduction**

MicroRNAs (miRNAs) compose a class of small endogenous non-coding RNA capable of regulating gene expression, either by inhibiting translation or promoting mRNA degradation (1). It has been suggested that differential expression of miRNA is involved in the initiation and progression of human cancer and thus may serve as a disease marker to improve diagnosis, prognosis and therapy for cancer (2-5). Based on miRNA arrays, numerous studies have been performed to identify differentially expressed (DE) miRNAs between cancer and normal samples (3, 6-10). However, for a particular cancer, DE miRNA generated from multiple studies are often highly inconsistent. For example, no DE miRNAs detected from six studies for prostate cancer were shared by all of the studies (11). Because the reproducibility of discovered biomarkers is of fundamental importance for validation (12, 13), this problem must be resolved in a rational way before full use of miRNA arrays in biological researches is possible.

The irreproducibility problem may not reflect lack of reliability of the technology platform used but may reflect a lack of understanding regarding a correct validation process for biomarkers (14). It is becoming clear that, in high-throughout experiments for finding biomarkers for a disease, the irreproducibility problem could be induced by, among other factors, diverse biological factors such as biological variation and heterogeneity of molecular changes in the disease (15-17). On the other hand, it is also becoming clear that the diverse molecular changes (18, 19) including expression changes of miRNAs in cancer tend to be redundant regarding function (20). Specifically, diverse expression changes of miRNAs in cancer may primarily affect
cancer-associated pathways by targeting cancer genes in these pathways (9, 21, 22). If this is true, different DE miRNAs generated from different studies for a cancer may be consistent in function, as seen for biomarker discovery based on the mRNA array platform (16, 23). Thus, to evaluate the reproducibility of DE miRNA discovery in different studies for a particular cancer, one reasonable approach involves taking into account the functional relationship (16, 24), rather than simply counting the overlaps. That is, we can design scores to evaluate the reproducibility of biomarker discovery using a biological assumption on functional relations such as expression correlation (16) or functionally similarity (24) between biomarkers. Notably, the underlying biological assumption is statistically testable: if a score is significantly higher than expected by chance, then it can provide evidence supporting that the assumption can partially explain the apparently inconsistent biomarkers.

Here, based on multiple miRNA datasets for two cancer types, we demonstrated that the regulation direction (up-regulated or down-regulated in cancer samples compared with normal samples) of DE miRNA detected in different datasets for each cancer were highly consistent. This result indicates that the miRNA arrays could be applied to reliably capture signals of the regulation direction of DE miRNA in cancer (25). Importantly, this result also suggests that the DE miRNA in a particular cancer has a consistent up or down regulation pattern. Upon analysis of different samples for each cancer, the miRNA displaying the most pronounced differential expression varied greatly, which could be attributed to biological variation of the molecular change of miRNA in the samples (26). Based on the biological assumption that two DE miRNAs with significantly overlapped targets may have a similar function (27), we proposed a score entitled POF (Percentage of Overlaps of Function-related miRNAs)
in order to evaluate the functional consistency of DE miRNAs for a cancer. For each of the two cancers, respectively, we found that most of the POF scores between top-ranked DE miRNAs detected from different studies were rather high, suggesting that the DE miRNAs detected from different studies may point to functionally important differential expression of miRNA in the disease. Then, using another independent dataset for colon cancer, we validated the model of functional links between DE miRNAs extracted from two other datasets for this cancer.
Materials and Methods

Datasets

We analyzed five large datasets for two cancer types (colon and gastric). For each dataset, the number of samples of each state (cancer or normal) was no less than 20. The detailed information was described in Table 1. We referred to each dataset using the following nomenclature: cancer type followed by the total number of samples. Three datasets were collected for colon cancer: two datasets (Colon99 (6) and Colon168 (7)) were used to analyze the consistency of DE miRNAs and extract the functional link models, and another dataset (Colon108 (8)) was used to validate the functional link models. The Colon108 dataset included 38 technical replicates with highly reproducible signals (8). For each sample with replicates in this dataset, the expression value of each miRNA was represented by the average of quantile normalized values for this miRNA across the replicates (8).

miRNA targets and cancer-associated genes

Currently, Targetscan (28), miRanda (29) and PicTar (30) are among the most widely used miRNA-target interaction prediction algorithms. Thus, we used them for analyses separately. Considering that miRNA targets predicted by multiple algorithms might be more reliable (31), we also used miRNA-targets interactions appearing in at least two of the eight data sources (see Supplementary Data).

A total of 2104 cancer-associated genes were collected from F-Census which compiles cancer genes from eight data sources (24).
Data preprocessing and DE miRNA selection

For the datasets of Colon99, Colon168, Gastric41 and Gastric353, the raw data was preprocessed as follows: a) the median background signal intensity was subtracted from the signal for each probe; b) signal values less than 1 were replaced with 1; c) log2-transformation was applied; d) probes absent in more than 20% of samples were deleted and other missing values were imputed using the R package kNN imputation function (32); e) signal values of replicate probes of a miRNA were averaged; f) Quantile normalization was performed (33). For Colon108, the quantile normalized data of the original study were used. For all platforms, we annotated probes to mature miRNA using miRbase version 16 (34).

Two-sample $t$-test was performed to select the DE miRNAs. The $P$ values were adjusted by the Benjamini and Hochberg (BH) correction procedure to account for multiple tests with FDR < 5% (35). For comparing the top $n$ DE miRNAs from different datasets, we ranked miRNAs according to the $t$-test $P$ values to select the most significant $n$ ones. The DE miRNAs selected for each dataset were listed in Supplemental Table S1-S5.

Consistency of regulation direction between two lists of DE miRNA

For each cancer, when comparing results found from multiple datasets generated on different platforms, we only analyzed the miRNAs presented in all of the datasets. Using the binomial distribution model, we calculated the probability of observing from $N$ randomly selected miRNAs at least $m$ non-DE miRNAs with consistent regulation directions across two datasets as follows:
\[ P = \sum_{i=m}^{N} C_i^N (P_e)^i (1 - P_e)^{N-i} \]

in which \( P_e \) is the random probability that one non-DE miRNA with the same regulation direction exists across two datasets. For each cancer, we roughly defined non-DE miRNA as miRNA that was not detected as DE miRNA at the FDR level of 10% in either of the datasets for this cancer. The \( P_e \) for each cancer was close to 0.5.

**The PO score**

The PO (Percentage of Overlaps) metric is frequently used to evaluate the consistency of different lists (26, 36, 37). Clearly, in order to be utilized as a disease marker for a disease, a DE miRNA should have a steady regulation pattern across different datasets. In this study we strictly defined DE miRNA, shared by two lists of DE miRNA only to be thus named if it also exhibited the same direction change across the corresponding two datasets (16). Suppose \( O \) miRNAs are shared by list 1 with length \( l_1 \) and list 2 with length \( l_2 \), then the PO score from list 1 to list 2 is \( \text{PO}_{12} = O / l_1 \), and the score from list 2 to list 1 is \( \text{PO}_{21} = O / l_2 \).

To assess the significance of a PO score for two DE miRNA lists with lengths \( l_1 \) and \( l_2 \), respectively, we did a random experiment to test the null hypothesis that the observed PO score would be expected by chance for two random miRNA lists (with lengths of \( l_1 \) and \( l_2 \)) extracted from miRNAs presented in both datasets. This process was repeated 10,000 times. The significance level of an observed PO score was defined as the percentage of the 10,000 random scores no less than the observed score.

**The POF score**
We considered two miRNAs to be functionally similar if they (1) had the same regulation direction, (2) shared significantly more target genes than expected by chance, and (3) shared at least one cancer-associated target gene. The significance of the number of targets shared by two miRNAs was calculated according to the cumulative hypergeometric distribution model:

\[
P = 1 - \sum_{i=0}^{k-1} \frac{C_n^i C_{M-i}^{N-n}}{C_M^N}
\]

in which \(N\) is the number of all targets of all miRNAs, \(M\) and \(n\) are the numbers of targets for the two miRNAs, respectively, and \(k\) is the number of targets shared by these two miRNAs. The \(P\) values were corrected by the Benjamini and Hochberg method for multiple testing (35).

Next, we proposed the POF (Percentage of Overlaps of Function-related miRNAs) score to evaluate the functional consistency between two miRNA lists. Suppose \(O_{f12}\) (or \(O_{f21}\)) denotes the number of miRNAs in list 1 (or list 2) not shared but functionally similar to at least one miRNA in list 2 (or list 1); the POF score from list 1 to list 2 (or from list 2 to list 1) is then \(\text{POF}_{12}=(O+O_{f12})/l_1\) (or \(\text{POF}_{21}=(O+O_{f21})/l_2\)). POF is then determined as \(\text{POF}=(\text{POF}_{12}+\text{POF}_{21})/2\).

We tested two null hypotheses for an observed POF score by random experiments. The first random experiment, referred to as ‘targets randomization’, was used to test the null hypothesis that the score can be expected by chance when no prior biological knowledge of miRNA-target interaction is used. Here, we assigned genes randomly derived from the human genome to miRNAs while keeping the number of targets for each miRNA unchanged. The second random experiment, referred to as ‘miRNAs
randomization’, was applied to test the null hypothesis that the observed POF score can be expected for non-DE miRNAs lists with the same lengths as the DE miRNAs lists. Here, approximately, we defined non-DE miRNAs for a cancer as the remaining miRNAs after excluding the miRNAs detected to be potentially significant with a rather loose FDR cutoff of 10% in either dataset for the cancer. Then, we randomly selected two non-DE miRNA lists with the same lengths as the DE miRNAs lists and calculated their POF score. This process was repeated 10,000 times. The significance level for an observed POF score was defined as the percentage of the randomization scores no less than the observed value.

**Functional link model of DE miRNAs**

For two lists of DE miRNAs separately selected from two datasets for a cancer, we constructed a network model by linking every two miRNAs between the lists if they are functionally similar according the criteria described above.

For a new list of DE miRNAs detected from a new dataset for a cancer, we can test whether it is associated with a previously extracted model for the same disease. We can do this validation based on the same assumption for defining functionally similar miRNAs. That is, for each miRNA in the new list, if it shows the same regulation direction across all datasets and shares significantly more targets with at least one set of common targets of two linked miRNAs in the model, including at least one cancer-associated gene, then it is defined as a functionally similar miRNA with the model.

**Pathway enrichment analysis**
The hypergeometric distribution model was used to find the KEGG (38) pathways enriched with common targets of two DE miRNAs. If we suppose that $M$ is the total number of the common targets of two miRNA, $N$ is the total number of genes in human genome, $n$ is the number of genes annotated in a pathway, and $k$ is the number of the common targets annotated in this pathway, we then calculate the probability of observing by chance at least $k$ targets annotated in this KEGG pathway as follows:

$$ P = 1 - \sum_{i=0}^{k-1} \frac{\binom{i}{n} \binom{M-i}{N-n}}{\binom{M}{N}} $$

The raw $P$ values were adjusted by the BH correction procedure (FDR<5%) (35).
Results

Consistent regulation direction of DE miRNA in a particular cancer

In order to investigate the consistency of DE miRNAs across different datasets, we analyzed two datasets for each of the two solid cancers described in Table 1. In each dataset, we selected DE miRNAs by $t$-test at the FDR level of 5%. Considering that different platforms have different degrees of coverage of all human miRNAs, we focused on analyzing the miRNAs presented in both datasets for a cancer.

In the two datasets for colon cancer, 22 and 49 DE miRNAs were selected from the 110 miRNAs presented in both datasets, respectively, and they shared 14 DE miRNAs(Fig. 1A), in which 13 have the same regulation direction (up-regulated or down-regulated in cancer samples compared with normal samples) across these two datasets. Thus, the PO score (see Materials and Methods) was 0.59 from the short list to the long list and was 0.27 from the opposite direction. These two scores were not high, but were significantly larger than the expected scores of 0.32 and 0.15 in the two directions ($P=0.0010$). In the two datasets for gastric cancer, 78 and 18 DE miRNAs were identified from the 155 miRNAs presented in both datasets and they shared 15 DE miRNAs (Fig. 1B). The PO score was 0.83 from the short list to the long list while it was only 0.19 from the opposite direction; both values were significantly larger than expected by chance ($P<0.0001$). Notably, the longer list of DE miRNAs included approximately half of all miRNAs detected in these two cancer types, indicating that miRNAs are widely altered in cancer.
We then investigated the consistency of the regulation direction of the DE miRNAs identified from various studies for colon and gastric cancers, respectively. For the two colon cancer datasets, 13 (92.9%) of the 14 DE miRNAs detected in both datasets exhibited the same up or down regulation across the datasets \((P=9.16\times10^{-4}, \text{binominal distribution test})\). For the 49 DE miRNAs found in the dataset Colon99, after excluding the 14 overlapping DE miRNAs, 13 exhibited tentative evidence of differential expression with \(P\) values less than 0.1 in the dataset Colon168, and all of them exhibited the same direction change across the two datasets \((P=1.22\times10^{-4}, \text{binominal distribution test})\). This result indicates that many miRNAs with tentative evidence of differential expression in the dataset Colon168 may actually be differentially expressed in colon cancer. Similarly, for gastric cancer, all of the 15 overlapping DE miRNAs showed the same regulation directions across the two datasets \((P=3.05\times10^{-5}, \text{binominal distribution test})\). For the 78 DE miRNAs found in the dataset Gastric353, after excluding the 15 overlapping DE miRNAs, 28 exhibited \(P\) values less than 0.1 in the dataset Gastric41 and 27(96.4%) of them showed the same regulation directions across the two datasets \((P=1.08\times10^{-7}, \text{binominal distribution test})\). The results showed that the regulation direction of the DE miRNAs detected in both datasets for each of the two cancers were highly consistent, indicating that miRNA arrays could reliably capture the signals of the regulation direction of differentially expressed miRNAs in a particular cancer.

**Functional consistency of top-ranked DE miRNA for a particular cancer**

Researchers are often interested in the most significant DE miRNAs. However, for a particular cancer, the top ranked \(n\) most significant DE miRNAs separately detected from various studies are usually highly inconsistent. For example, when \(n\) was equal
to 10, only two (hsa-miR-106b and -10b) were shared by two lists for colon cancer and three (hsa-miR-181b, -21 and -218) were shared for gastric cancer. The corresponding PO scores were only 0.2 and 0.3. When $n$ was equal to 20, the PO scores were 0.2 and 0.45 for colon and gastric cancer, respectively. Then, we proposed the POF score (see Materials and Methods) to evaluate the functional consistency of two lists of the top-ranked $n$ DE miRNAs for each cancer based on the assumption that different DE miRNAs may disturb the same cancer-associated pathways through their common targets. We calculated the POF scores using the miRNA targets predicted by TargetScan (28), miRanda(29) and PicTar (30), respectively. Also, we calculated the scores using miRNA targets documented in at least two of the eight data sources (see Materials and Methods). We found that all the results were quite similar (Table 2), indicating that functional analysis between miRNAs based on their non-random target overlap is rather robust against a certain level of false targets predicted for miRNAs. Hence, we only described the results based on the targets predicted by TargetScan in the following text.

Between the lists of the top 10 (or 20) DE miRNAs detected separately from the two datasets for colon cancer, the POF score was 0.90 (or 0.83). We tested two null hypotheses for these observed high POF scores by random experiments (see Materials and Methods). For the ‘targets randomization’ experiment, the average of 10,000 random POF scores for either the top 10 or 20 miRNAs was 0.20. In 10,000 of the random POF scores, no one exceeded the observed scores ($P$ values<10E-4, Table 2), we thus rejected the null hypothesis that the observed POF score can be expected by chance when no prior biological knowledge of miRNA-target interaction is used. For the ‘miRNAs randomization’ experiment, we approximately defined non-DE
miRNAs as those miRNAs not detected as DE miRNA using a loose FDR control of 10% in either dataset. By randomly extracting two lists of 10 or 20 non-DE miRNAs 10,000 times, the average of random POF scores was 0.44 or 0.47, significantly smaller than the corresponding observed scores. Thus, we rejected the null hypothesis that the observed POF score can be expected for non-DE miRNAs lists with the same lengths as the DE miRNAs lists (Table 2). Similarly, the POF scores for the lists of the top 10 or 20 miRNAs separately detected from the two datasets for gastric cancer was 0.95 or 0.98, both of which were significantly higher than those obtained using the above two kinds of random experiments (Table 2).

**Functional link model of DE miRNAs for a particular cancer**

We constructed the model of functional relationships between the lists of the top 10 DE miRNAs (Fig. 2A) separately selected from the two datasets for colon cancer (see Materials and Methods). Based on this model, we explained some functional links. For example, hsa-miR-106a and -20a were detected as up-regulated DE miRNAs in Colon168 but not in Colon99. However, they may play similar function with hsa-miR-106b which was detected as an up-regulated DE miRNA in both datasets. All these three miRNAs belong to the miR-17 family and regulate the same set of targets including 15 known tumor suppressor genes such as *CDKN1A, PIK3R1, RB1, STK11* and *TGFBR2*. These common targets were significantly enriched in some cancer-associated pathways such as “TGF-beta signalling pathway” ($P=2.47E-5$), “mTOR signalling pathway” ($P=3.16E-5$) and “Cell cycle” ($P=4.91E-3$). For another example, hsa-miR-199b-5p was detected as an up-regulated DE miRNA only in Colon99, but it may play the same role with hsa-miR-135b up-regulated in Colon168 as they co-targeted significantly more targets including one tumor suppressor gene.
(MAP2K4). These common targets were significantly enriched in some cancer-associated pathways such as “ErbB signalling pathway” \((P=2.90E-4)\). Similarly, hsa-miR-30a was detected as a down-regulated miRNA only in Colon99, but it may play the same role with hsa-miR-30c detected as a down-regulated DE miRNA in Colon168. Both of them were from the miR-30 family and shared the same targets including 36 known oncogenes such as \(ABL1\), \(BCL2\), \(KRAS\) and \(PDGFRB\). These common targets were significantly enriched in some cancer-associate pathways such as “Axon guidance” \((P=2.48E-6)\), “Focal adhesion” \((P=1.19E-3)\) and “MAPK signalling pathway” \((P=3.28E-3)\).

Similarly, we constructed the model of functional relationships between the two lists of the top 10 DE miRNAs (Fig. 2B) separately selected from the two datasets for gastric cancer. In this model, hsa-miR-93 was detected as an up-regulated DE miRNA exclusively in dataset Gastric353. However, it might play a similar regulatory role with hsa-miR-25 which is in the same cluster with hsa-miR-93 and was detected as an up-regulated DE miRNA in dataset Gastric41 as they shared 127 targets \((P=8.72E-12)\). Their common targets were significantly enriched in the cancer-associated “TGF-beta signalling pathway” \((P=2.28E-8)\). For another example, hsa-miR-181c was up-regulated in Gastric353 but they may play the same function with hsa-miR-181b detected to be up-regulated in Gastric41. These two miRNAs belong to the miR-181 family and they regulate the same set of targets including 14 known tumor suppressor genes such as \(ATM\), \(CACNA2D2\) and \(CYLD\). The common targets were significantly enriched in some cancer related pathways such as “T cell receptor signalling pathway” \((P=1.26E-3)\) and “Apoptosis” \((P=4.89E-2)\). Similarly, hsa-miR-148b and hsa-miR-212, identified as down-regulated DE miRNAs separately
in dataset Gastric353 and Gastric41, might also play a similar regulatory role in
cancer as they shared significantly more common targets\( (P=2.40\times10^{-10}) \) including 10
cancer associated genes such as \textit{DDX6}, \textit{SMAD2}, and \textit{NFAT5}. These common targets
were significantly enriched in some cancer related pathways such as the “Wnt
signalling pathway” \( (P=2.51\times10^{-5}) \) and the “TGF-beta signalling pathway”
\( (P=4.80\times10^{-6}) \).

We were able to collect another independent dataset (Colon108) for colon cancer.
Here, we used this dataset to validate the functional link model constructed for the
two lists of the top 10 DE miRNAs separately detected from the previous two datasets
(Fig. 2A). We only analyzed the 101 miRNAs presented in all of the three datasets,
among which 62 miRNAs were differentially expressed in Colon108. For the 16
miRNAs in the model, 11 were differentially expressed in Colon108. Among these 11
DE miRNAs, 10 (hsa-miR-1, -106a, -106b, -10b, -135b, -20a, -21, -30a, -30c and -92a)
exhibited the same regulation directions across the three datasets, and this was highly
unlikely to happen by chance \( (P=5.86\times10^{-3}, \text{ binominal distribution test}) \). In the opposite
direction, 8 (80\%) of the top most significant 10 miRNAs detected from Colon108
were included in or functionally linked with the model, which was significantly more
than expected by chance according to both the ‘targets randomization’ and the
‘miRNAs randomization’ experiments (both \( P \) values<0.0001). Among these 8 DE
miRNAs, three are the members of this model, including one up-regulated miRNA
(hsa-miR-135b) and two down-regulated miRNAs (hsa-miR-10b and -30a). The other
five miRNAs (hsa-miR-7, -147, -195, -9 and -95) were functionally associated with
the model (see Materials and Methods). Parallely, we constructed the functional link
model for the two lists of top 20 miRNAs separately from the previous two datasets.
(see Supplementary Fig. S1). Similar validation results were observed for this model. Twenty of the 29 miRNAs in the model were differentially expressed in Colon108. Among these 20 DE miRNAs, 18 exhibited the same regulation directions across the three datasets ($P=2.01E-4$, binominal distribution test). For the top most significant 20 DE miRNAs detected from Colon108, 15 (75%) were included in or functionally linked with the model, which was also significant according to both two kinds of experiments (both $P$ values<0.0001). Therefore, for this cancer type, the DE miRNAs from the three datasets were rather reproducible. Note that for gastric cancer, we were unable to find another large independent datasets to do the same analysis.
Discussion

Our results demonstrated that miRNA arrays can reliably detect the signals of the regulation direction of DE miRNA in cancer, in a manner similar to the study of differential gene expression in diseases based on microarrays (25). Our results also suggest that miRNAs are widely altered in a particular cancer and show a consistent up or down regulation pattern, as part of the primary pathogenesis or subsequent response in cancer progression. However, for a particular cancer, the top-ranked DE miRNAs detected from different datasets varied greatly. It has been suggested that, due to the limited statistical powers in the presence of large variations of gene expression in a complex disease, lists of differentially expressed genes detected from various microarray experiments may be highly inconsistent but may generally comprise true discoveries (16, 39). Also, based on miRNA arrays, the apparently irreproducible DE miRNAs detected from different experiments may all partially capture some important differential expression of miRNAs in the disease. Nevertheless, based on the POF score, we have obtained evidence that top-ranked DE miRNAs separately detected from different studies for a cancer tend to regulate genes in the same cancer-associated pathways. This result is consistent with the observation that most human miRNA tends to play a principal role in important functions in a redundant manner (20). Notably, when comparing data produced from different platforms with different degrees of coverage of all human miRNAs, we only analyzed the miRNAs commonly detected by the platforms. We could expect that results produced by the same platform could reach higher reproducibility than results produced by different platforms which may have extra between-platform variation.
Notably, we have also investigated the six datasets for prostate cancer reviewed by Gandellini et al. (11). Among the five publicly available datasets, three datasets included only 8, 4, and 7 normal samples, respectively, which might be too small to get reliable results. Thus, we only analyzed the dataset (denoted as Prostate76) from (40) and the dataset (denoted as Prostate80) from (41). From the 83 miRNAs presented in both datasets, 22 and 8 DE miRNAs were detected for the two datasets, respectively and they shared only one miRNA (hsa-miR-221) (see Supplementary Results). We could not observe significant consistency between these two DE miRNA lists (see Supplementary Results). As suggested by Gandellini et al. (11), discrepancy between the two datasets could be explained by the difference in their approach of dissection, preservation of specimens. In addition, the phenotypic heterogeneity within each dataset might contribute to the inconsistence. Among the 40 stage T2a/b prostate cancer samples in Prosate80, 20 showed chemical relapse within two years and 20 did not show chemical relapse within ten years and 16 miRNAs were differentially expressed between these two groups (41). Among the 60 tumors in Prostate76, 35 were organ-confined and 17 showed extraprostatic extension and changed miRNA abundance was also observed across these two groups of patients (40). Thus, when we find inconsistent results from different studies for a disease, we need to carefully consider possible influences of diverse factors such as difference in experimental design, biological variation and heterogeneity of the disease.

One limitation of our analysis is that we studied the function of miRNAs through their targets predicted by tools that face the problem of a high false positive rate (42). However, we found our results were rather robust against a certain level of false targets predicted for miRNAs. Nevertheless, it is generally acknowledged that all
targets of miRNA comprise a large battery for all biological conditions (43). Thus, further study is needed to achieve a more comprehensive understanding of targets active in carcinogenesis utilizing large-scale matched miRNA and mRNA arrays (40, 44-46). One issue frequently of concern to researchers is the desire to find DE miRNAs shared by multiple cancer types as well as cancer type-specific DE miRNAs. For example, Volinia et al. compared six cancer types and picked 21 up-regulated miRNAs associated with more than three cancer types, suggesting the complexity and potential for the understanding of common mechanisms apparent in different cancers (9). Various groups have attempted to identify DE miRNAs unique to one or several cancer types. For example, Bandyopadhyay et al. defined cancer type-specific miRNA to be DE miRNA found in fewer than 2 cancer types (47). Johnson et al. concluded that let-7 was lung cancer specific as it was found to be differentially expressed in lung cancer but not in colon and breast cancer (48). However, as demonstrated in this study, for a particular cancer various studies also tend to generate different yet reliable DE miRNAs. Thus, it is possibly misleading to define cancer type-specific DE miRNAs in such a simple way, as the samples used in current miRNA array experiments would be insufficient to generate a full list of DE miRNAs for a particular cancer. In our future studies we plan to analyse arrays for additional cancer types, and moreover attempt to identify the cancer type-specific DE miRNAs by observing DE miRNAs detected for a variety of cancer types while taking into consideration the opposite regulation direction for different cancer types. For example, in this study, we found that hsa-miR-221 consistently up-regulated in both of the two gastric cancer datasets while it consistently down-regulated in both of the two prostate cancer datasets.
In the past few years, miRNA-expression profiling of human tumors has identified many miRNA biomarkers associated with diagnosis, progression, prognosis and treatment for cancers (3). Evaluating the reproducibility of the DE miRNAs found from different studies for a cancer should be an indispensable step towards the pre-clinical research. For example, because a specific miRNA can targets multiple genes involved in multiple cancer-associated pathways, it has been proposed that miRNAs could be drug targets in cancer therapy (5). Certainly, the best drug target is the kind of miRNAs consistently disturbed in different cohorts of patients. However, DE miRNAs identified from different cohorts of patients for a particular cancer are often highly inconsistent and thus hold back the utility of miRNAs biomarkers. Our study suggested that miRNAs playing roles in one cohort of patients may have functionally similar miRNAs playing the same roles in other cohorts of patients. In other words, each of the functional similar miRNAs may potentially play an important role in tumorigenesis for a part of patients. In this situation, “multiplex RNA inhibition targeting” strategy is expected for efficient drug design (49). Thus, the reproducibility evaluation of the DE miRNAs in different cohort of patients could provide a valuable guide for the drug target design. Future work is warranted to suggest combinations of DE miRNAs for efficient drug design by considering their overall sample coverage for a cancer.

ACKNOWLEDGEMENTS

We thank the two anonymous referees for their constructive advices and comments to improve this work.
References


# Tables

## Table 1 - Summary of the five datasets analyzed in this study

<table>
<thead>
<tr>
<th>Dataset</th>
<th>(T:N) §</th>
<th>Platform</th>
<th>Source of Normal sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon99 (6)</td>
<td>78:21</td>
<td>OSU_CCC 11k v2</td>
<td>normal colonic mucosa **</td>
</tr>
<tr>
<td>Colon168 (7)</td>
<td>84:84</td>
<td>OSU_CCC v2.0</td>
<td>adjacent non-tumor tissue</td>
</tr>
<tr>
<td>Colon108 (8)</td>
<td>80:28</td>
<td>Illumina v1</td>
<td>adjacent non-tumor tissue</td>
</tr>
<tr>
<td>Gastric41 (9)</td>
<td>20:21</td>
<td>OSU_CCC 11k v2</td>
<td>adjacent non-tumor tissue</td>
</tr>
<tr>
<td>Gastric353 (10)</td>
<td>184:169</td>
<td>OSU_CCC v3.0</td>
<td>adjacent non-tumor tissue</td>
</tr>
</tbody>
</table>

* Each dataset was denoted by the following nomenclature: cancer type followed by the total number of samples.

§ T, the number of tumor samples; N, the number of normal samples.

** normal samples were described as being from the normal colonic mucosa, thus it is unclear whether they were from the patients.

## Table 2 - Scores for the top n (10 or 20) most significant DE miRNAs

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Method</th>
<th>PO F</th>
<th>e-POF1*</th>
<th>e-POF2*</th>
<th>PO F</th>
<th>e-POF1</th>
<th>e-POF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon168 vs. Colon99</td>
<td>TargetScan</td>
<td>0.90</td>
<td>0.20(&lt;0.0001)</td>
<td>0.43(0.0004)</td>
<td>0.83</td>
<td>0.20(&lt;0.0001)</td>
<td>0.46(0.0004)</td>
</tr>
<tr>
<td>miRanda</td>
<td>0.90</td>
<td>0.20(&lt;0.0001)</td>
<td>0.50(0.0006)</td>
<td>0.85</td>
<td>0.20(&lt;0.0001)</td>
<td>0.50(&lt;0.0001)</td>
<td></td>
</tr>
<tr>
<td>PicTar</td>
<td>0.80</td>
<td>0.20(&lt;0.0001)</td>
<td>0.48(0.0039)</td>
<td>0.78</td>
<td>0.20(&lt;0.0001)</td>
<td>0.47(0.0003)</td>
<td></td>
</tr>
<tr>
<td>Integrated</td>
<td>0.90</td>
<td>0.20(&lt;0.0001)</td>
<td>0.47(0.0004)</td>
<td>0.85</td>
<td>0.20(&lt;0.0001)</td>
<td>0.49(&lt;0.0001)</td>
<td></td>
</tr>
</tbody>
</table>

| Gastric353 vs. Gastric41 | TargetScan | 0.95  | 0.30(<0.0001) | 0.44(<0.0001) | 0.98  | 0.45(<0.0001) | 0.47(<0.0001) |
| miRanda        | 0.95  | 0.30(<0.0001) | 0.50(0.0004) | 0.98  | 0.45(<0.0001) | 0.50(<0.0001) |
| PicTar         | 0.95  | 0.30(<0.0001) | 0.44(0.0004) | 0.95  | 0.45(<0.0001) | 0.48(<0.0001) |
| Integrated     | 0.95  | 0.30(<0.0001) | 0.47(0.0004) | 0.98  | 0.45(<0.0001) | 0.49(<0.0001) |

* Integrated, miRNA-target interactions were in at least two of eight miRNA target sources.

* expected POF score (p-value) estimated by doing ‘targets randomization’ 10,000 times.

§ expected POF score (p-value) estimated by doing ‘miRNAs randomization’ 10,000 times.
Figure legends

Figure 1 - Consistence analysis between different DE miRNA lists for each cancer. A, Overlap between the two lists of DE miRNAs extracted from Colon99 and Colon168. In the overlap, one miRNA (miR-192, italic) showed inconsistent deregulation directions across these two datasets. B, Overlap between the two lists of DE miRNAs extracted from Gastric41 and Gastric353. The DE miRNAs were selected by $t$-test followed by BH correction (FDR < 5%).

Figure 2 - The function link model between the two lists of the top ranked 10 DE miRNAs for each cancer. For two lists of DE miRNAs for a cancer, we linked every two functionally similar miRNAs between the lists to construct the model (see Materials and Methods). A, The model between the top ranked 10 DE miRNAs from Colon99 (cycle and diamond) and the top ranked 10 DE miRNAs from Colon168 (rectangle and diamond); B, The model between the top ranked 10 DE miRNAs from Gastric353 (cycle and diamond) and the top ranked 10 DE miRNAs from Gastric41 (rectangle and diamond). For a cancer, the up-regulated module consists of up-regulated miRNAs (left panel) and the down-regulated module consists of down-regulated miRNAs (right panel).
Figure 2

A  Up-regulated module in Colon cancer

B  Up-regulated module in Gastric cancer
Molecular Cancer Therapeutics

Evaluating the consistency of differential expression of microRNA detected in human cancers

Xue Gong, Ruihong Wu, Hongwei Wang, et al.

Mol Cancer Ther Published OnlineFirst March 11, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/1535-7163.MCT-10-0837

Supplementary Material
Access the most recent supplemental material at:
http://mct.aacrjournals.org/content/suppl/2011/03/13/1535-7163.MCT-10-0837.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.