Send Orders for Reprints to reprints@benthamscience.net

438 *Current Pharmaceutical Biotechnology,* **2014**, *15*, **438-444**

Predicting Cancer-Related MiRNAs Using Expression Profiles in Tumor Tissue

Hsiuying Wang*

Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

Abstract: MicroRNAs (miRNAs) are small, noncoding RNAs with important functions in development, cell differentiation, and regulation of cell cycle and apoptosis. Many studies have now shown that miRNAs are involved in the initiation and progression of cancers. In this study, procedures based on the relative R-squared method (RRSM) are proposed to investigate miRNA-mRNA regulatory relationships between 114 miRNAs and 16063 mRNAs for different organic tissues. These procedures are based on comparing the expression profiles in tumor tissue and those in normal tissues, or based on the expression profiles in tumor tissue only. The analyzed results are used to predict high-confident miRNAs for tumor development and their targets. This study predicts many high-confident miRNAs which are associated with colon cancer, prostate cancer, pancreatic cancer, lung cancer, breast cancer, bladder cancer and kidney cancer, respectively.

Keywords: Cancer, microarray expression profile, microRNA, normal tissue, the relative R-squared method, tumor tissue.

INTRODUCTION

MicroRNA (miRNA) is a short non-coding RNA around 22 nt, which suppresses gene expressions via translational suppression or is a factor involved in mRNA degradation by binding to 3'-untranslated regions (3'UTR). It has been estimated that miRNAs regulate ~30% of human genes [1]. The first miRNA was discovered in 1993 by Victor Ambros, Rosalind Lee and Rhonda Feinbaum during a study into development in the nematode Caenorhabditis elegans (C. elegans) regarding the gene lin-14 [2]. Previous studies indicated that "MiRNAs play a key role in diverse biological processes, including development, cell proliferation, differentiation and apoptosis. Accordingly, altered miRNA expression is likely to contribute to human disease, including cancer" [3]. As a result, miRNAs may be useful tools for characterizing specific cancers and for determining patient prognoses [4].

Calin *et al*. [5] published the first study to link miR-15 and miR-16 to cancer in 2002. Subsequent reports have shown that miRNAs are altered in many cancers, and they can initiate carcinogenesis or drive progression [6].

Accurate miRNA target prediction can help understand miRNA regulatory mechanisms. Several target prediction computational algorithms for motifs complementary predictions have been developed, for example, miRanda [7] and TargetScan [1, 8], but they show poor overlap between their predicted results [9]. In addition to sequence motifs complementary predictions, since miRNA expression profile can help identify human solid tumors, it is used for discovering miRNA targets [10]. However, investigating miRNA expression profiles can become computationally complicated when multiple miRNAs and their effects across multiple tissues are to be considered [11].

To overcome this difficulty, statistical methods have been proposed to build up a network of associations between the miRNAs and their target mRNAs [11-13]. Huang *et al*. [12] established a method, GenMiR++, using Bayesian variation analysis to explore miRNA targets. However, it is complicated and requires extensive calculations. In order to provide a more effective approach, Wang and Li [11] proposed the relative R-squared method to select high-confidence targets of miRNAs, which is easy to interpret and less computationally expensive.

Hsieh and Wang [13] called the relative R-squared method as RRSM. The analyses in these previous studies were performed to predict miRNA targets across different tissues, such as colon tissues, lung tissues and kidney tissues simultaneously. In this study, I intend to explore miRNA targets focusing on a particular organic tissue to discover the miRNAs which are related to the development of a particular cancer. Although the RRSM has been established for miRNA target prediction, it has not been used to predict high-confident miRNAs associated with a specific cancer. Thus, in this study, I propose procedures based on the RRSM to find the miRNAs which are associated with cancers, and also investigate the performance of these procedures in predicting cancer-related miRNAs.

The other statistical approaches to predict high-confident miRNA targets include correlation analysis, GenMiR++ etc. Since the comparisons of these methods with the RRSM have been provided in literature, in this paper, we mainly focus on using the RRSM to predict high-confident miRNAs associated with cancer and do not focus on comparing these methods. In addition to the above mentioned methods, a more flexible approach is to replace the linear regression model used in the RRSM by other models, such as non-

*Address correspondence to this author at the Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan; Tel: 886-3-5712121 ext. 56813; Fax: 886-3-5728745; E-mail: wang@stat.nctu.edu.tw

1389-2010/14 $58.00+.00

linear regression models. Since there are many non-linear regression models, it needs to develop a methodology to select suitable non-linear regression models in a future study.

METHOD

Using Expression Profiles in Tumor Tissues and in Normal Tissues

The RRSM is established based on a relative instead of an absolute statistical point of view and it provides an efficient approach for miRNA target prediction [13]. To select high-confident miRNAs which are associated with a cancer development, first I apply the RRSM to find miRNA targets using expression profiles in tumor tissues. This is the one of the proposed methods to select high-confident targets. Many studies have been developed for investigating miRNA expression profiles to predict cancer-related miRNAs by comparing expression profiles in tumor tissue with expression profiles in normal tissue [14-17]. Therefore, to compare the expression profiles in tumor tissue with those in normal tissues, I also apply the RRSM to select miRNA targets in normal tissue. After obtaining miRNA targets in tumor tissue and miRNA targets in normal tissue respectively, the targets, which are selected using the tumor tissue, but are not selected using the normal tissue, are regarded as high-confident targets. This is also the one of the proposed methods to select high-confident targets.

To perform the RRSM, two thresholds p_0 and s for the p-value and the relative r-squared value need to be set in the RRSM. The details of applying the RRSM refer to the previous studies [11, 13]. In this study, the regression model used in the RRSM is the linear regression model. The programs for the RRSM are available on http://www.stat.nctu.edu.tw/hwang/website_wang%20new.htm [13].

To adopt the procedures, the steps of performing the first proposed procedure by comparing the tumor tissues with the normal tissues are given as follows.

Procedure 1

Predict miRNA targets with the RRSM by comparing the expression profiles in a tumor tissue and those in a normal tissue.

Step 1. Set thresholds p_0 and s.

Step 2. Perform the RRSM using the expression profiles in the tumor tissue to select the miRNA targets.

Step 3. Perform the RRSM using the expression profiles in the normal tissue to select the miRNAs targets.

Step 4. The miRNA targets selected in **Step 2** excluding the miRNA targets selected in **Step 3** are the high-confident miRNA targets obtained by comparing this tumor tissue with this normal tissue, and the corresponding miRNAs are the high-confident miRNAs that are associated with this tumor development.

The flowchart of Procedure 1 is briefly described in (Fig. **1**).

Analysis of Expression Profiling in Tumor Tissues

In addition to the above proposed method by comparing the expression profiles in the tumor tissue and those in the normal tissue, on the contrary, I may explore high-confident miRNAs mainly using expression profiles in the tumor tissue instead of comparing the result with the results using normal tissue. The procedure of performing the method using the tumor tissue is given in Procedure 2.

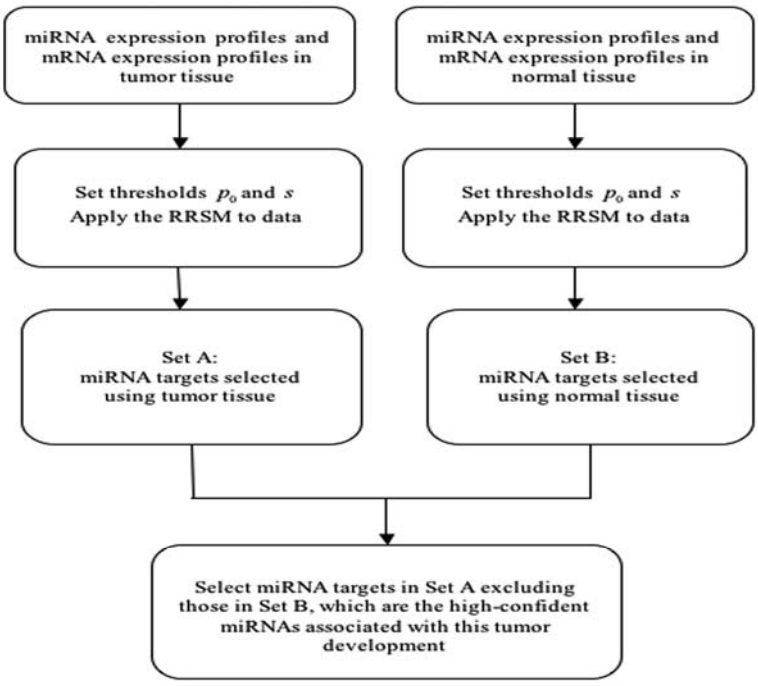

Fig. (1). Flowchart of Procedure 1.

Procedure 2

Predict miRNAs targets with the RRSM using the expression profiles in a tumor tissue.

Step 1. Set thresholds p_0 and s.

Step 2. Perform the RRSM using the expression profiles in the tumor tissue to select the miRNA targets.

Step 3. The miRNA targets selected in **Step 2** are the high-confident targets and the corresponding miRNAs are the high-confident miRNAs that are associated with this tumor development.

The flowchart of Procedure 2 is given in Fig. (**2**).

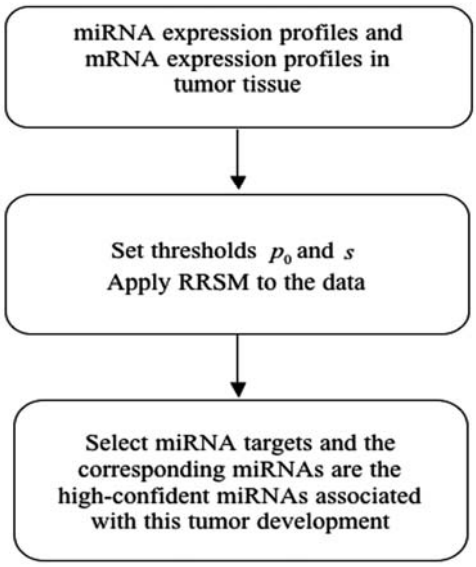

Fig. (2). Flowchart of Procedure 2.

RESULT

I adopt the data sets used in Huang *et al.* (2007) [18] and Hsieh and Wang (2011) [13] to predict high-confident miRNAs for tumor development. This data set includes the miRNA and mRNA expression data for 114 human miRNAs and 16063 mRNAs across a mixture of 88 normal and cancerous tissue samples of 11 organs [13]. The mRNA expression profiles that were published in 2005 consist of two microarray platforms, GPL80 and GPL98 [19]. A data set was filtered from the data to include 6387 potential target pairs, which covers 890 unique mRNAs, because some miRNAs have the same mRNAs as their potential targets [18]. The samples of 11 organs include 67 samples from tumor tissues and 21 samples from normal tissues. Since the first method is to discover the high-confident miRNAs comparing the expression profiles in tumor tissue with the expression profiles in normal tissue, I mainly consider the tissues in with tumor samples and normal samples in this data set. As a result, I present the predicted high-confident miRNAs which are associated with colon tumor, prostate tumor, pancreatic tumor, lung tumor, breast tumor, bladder tumor and kidney tumor, respectively.

Tables **1** and **2** list the predicted high-confident miRNAs. The predicted miRNA targets are given in the supplementary material.

Colon Cancer

In 2003, Michael *et al*. [20] published the first study of miRNAs in colon cancer, identifying miR-143 and miR-145 as novel dysregulated miRNAs in colon cancer. In addition, MiR-155 and miR-21 expression levels are significantly correlated with colorectal cancer [18]. MiR-1 can have a tumor suppressor function in colorectal cancer by directly downregulating MET oncogene [21].

In this study, to find high-confident miRNAs for colon tumor development, I first normalize the data and apply Procedure 2 using the thresholds $p_0 = 0.3$ and $s = 0.999$ and to find 52 high-confident miRNAs using expression profiles in tumor tissue. MiR-143, miR-145, miR-155, miR-21 and miR-1 are included in these 52 miRNAs (miRNAs in Tables 1 and 2 for colon tissue). After comparing the results using expression profiles in the tumor tissue and those using expression profiles in the normal tissue by Procedure 1, 42 high-confident miRNAs are selected in these 52 miRNAs (Table **1**).

Prostate Cancer

Prostate cancer is the most prevalent strain of cancer in men. MicroRNA-21 is shown to promote apoptosis resistance and invasion in prostate cancer cells [22]. MiR-205 transcription is shown to be commonly repressed in prostate cancer [23].

In this study, I first use the raw data and apply Procedure 2 using the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 35 high-confident miRNAs using the expression profiles in tumor tissue. MiR-21 and miR-205 are included in these 35 miRNAs (miRNAs in Tables **1** and **2** for prostate tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 24 high-confident miRNAs are selected in these 35 miRNAs (Table **1**).

Pancreatic Cancer

Pancreatic cancer is the fourth leading cause of cancer-related death in the United States and has the poorest overall survival rate among all human cancers because of late diagnosis and absence of screening tools [24]. The expression patterns of miR-21 and miR-200 are shown to be related to pancreatic cancer [25, 26].

In this study, to find high-confident miRNAs for pancreatic tumor development, I first normalize the data and apply Procedure 2 with the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 57 high-confident miRNAs using the expression profiles in tumor tissue. MiR-200a, miR-200b, miR-200c, and miR-21 are included in these 57 miRNAs (miRNAs in Tables **1** and **2** for pancreatic tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 54 high-confident miRNAs are selected in these 57 miRNAs (Table **1**).

Table 1. High-confident miRNAs selected by Procedure 1.

Tissue	High-confident miRNAs
Colon	miR-124a, miR-125b, miR-146, miR-1, miR-30b, miR-137, miR-19a, miR-203, miR-33, miR-219, miR-223, miR-193, miR-200a, miR-206, miR-216, miR-34a_(sub_1),miR-199a*, miR-30a, miR-96_(sub_1), miR-9, miR-30d, miR-30c, miR-24, miR-23b, miR-23a, miR-21, miR-19b, miR-183, miR-182, miR-145, miR-140, miR-139, miR-135, miR-133a, miR-125a, miR-101, miR-155, miR-200b, miR-200c, miR-30e, miR-34b, miR-34c_(sub_1)
Prostate	miR-124a, miR-125b, miR-29b_(sub_2), miR-218, miR-199a*, miR-194, miR-29c, miR-96_(sub_1), miR-9, miR-29a, miR-27b, miR-23b, miR-23a,miR-21, miR-183, miR-182, miR-138, miR-135, miR-128b, miR-128a, miR-125a, miR-200b, miR-200c, miR-27a
Pancreatic	miR-124a, miR-125b, miR-99a, miR-1, miR-30b, miR-137, miR-100, miR-181a,miR-181b, miR-203, miR-204_(sub_1), miR-205, miR-200a, miR-206, miR-218,miR-211, miR-215, miR-194, miR-190, miR-30a, miR-96_(sub_1), miR-9, miR-30d, miR-30c, miR-27b, miR-26a, miR-24, miR-23b, miR-23a, miR-22,miR-21, miR-192, miR-183, miR-182, miR-181c, miR-145, miR-144, miR-143,miR-138, miR-135, miR-133a, miR-128b, miR-128a, miR-125a, miR-107, miR-103, miR-26b_(sub_1), miR-101, miR-155, miR-200b, miR-200c, miR-27a, miR-30e, miR-99b
Lung	miR-124a, miR-125b, miR-99a, miR-146, miR-30b, miR-100, miR-18, miR-19a,miR-203, miR-33, miR-200a, miR-34a_(sub_1), miR-199a*, miR-194, miR-30a, miR-9, miR-30d, miR-30c, miR-27b, miR-24, miR-23b, miR-23a, miR-22,miR-19b, miR-182, miR-152, miR-148, miR-141, miR-138, miR-135,miR-125a, miR-107, miR-103, miR-155, miR-200b, miR-200c, miR-27a,miR-30e, miR-34b, miR-34c_(sub_1), miR-99b
Breast	miR-124a, miR-125b, miR-29b_(sub_2), miR-181a, miR-181b, miR-19a, miR-203, miR-204_(sub_1), miR-223, miR-33, miR-193, miR-218, miR-34a_(sub_1), miR-211, miR-190, miR-29c, miR-96_(sub_1), miR-9,miR-29a, miR-27b, miR-26a, miR-24, miR-23b, miR-23a, miR-22, miR-19b,miR-183, miR-182, miR-181c, miR-152, miR-148, miR-145, miR-144, miR-141,miR-139, miR-135, miR-128b, miR-128a, miR-125a, miR-26b_(sub_1), miR-101, miR-155, miR-200b, miR-200c, miR-27a, miR-302, miR-34b, miR-34c_(sub_1), miR-93_(sub_1)
Bladder	miR-124a, miR-125b, miR-16, miR-29b_(sub_2), miR-92, miR-146, miR-30b,miR-142-3p, miR-137, miR-106a, miR-17-5p, miR-181a, miR-181b, miR-19a,miR-203, miR-204_(sub_1), miR-205, miR-223, miR-33, miR-200a, miR-218,miR-211, miR-215, miR-199a*, miR-195, miR-194, miR-190, miR-30a, miR-29c, miR-9, miR-30d, miR-30c, miR-29a, miR-27b, miR-26a, miR-24,miR-23b, miR-23a, miR-22, miR-21, miR-20_(sub_1), miR-19b, miR-192, miR-183, miR-182, miR-181c, miR-15a, miR-153, miR-152, miR-148,miR-145, miR-144, miR-141, miR-140, miR-139, miR-135, miR-133a,miR-128b, miR-128a, miR-125a, miR-26b_(sub_1), miR-106b, miR-155,miR-17-3p, miR-200b, miR-200c, miR-25, miR-27a, miR-30e, miR-32
Kidney	miR-125b, miR-29b_(sub_2), miR-146, miR-137, miR-181a, miR-181b,miR-199a_(sub_1), miR-199b,miR-204_(sub_1), miR-223, miR-33, miR-206,miR-34a_(sub_1), miR-211, miR-215, miR-199a*, miR-29c, miR-96_(sub_1), miR-29a, miR-27b, miR-24, miR-23b, miR-23a, miR-22, miR-21, miR-192,miR-181c, miR-145, miR-125a, miR-101, miR-200b, miR-200c, miR-27a,miR-302, miR-34b, miR-34c_(sub_1), miR-93_(sub_1)

Table 2. High-confident miRNAs selected by Procedure 2 excluding miRNAs selected by Procedure 1.

Tissue	High-confident miRNAs
Colon	miR-204_(sub_1), miR-205, miR-218, miR-211, miR-26a, miR-22, miR-144,miR-143, miR-26b_(sub_1), miR-93_(sub_1)
Prostate	miR-146, miR-203, miR-205, miR-33, miR-24, miR-22, miR-144, miR-143, miR-133a, miR-101, miR-301
Pancreatic	miR-146, miR-219, miR-193
Lung	miR-223, miR-218, miR-128b, miR-128a
Breast	miR-146, miR-205, miR-194, miR-133a, miR-99b
Bladder	miR-34b
Kidney	miR-194, miR-153, miR-140, miR-139, miR-135, miR-99b

Lung Cancer

Lung cancer is the leading cause of cancer mortality worldwide, and 80% of lung cancers are non-small cell lung cancers [27]. MiRNAs function as tumor suppressors or oncogenes in lung cancer. Epidermal growth factor receptor (EGFR) signaling and *EGFR* mutations have been a major

focus of lung cancer studies conducted during the past 5 years. Several recent studies have uncovered a relationship between the EGFR signaling pathway and miRNAs [28]. Weiss *et al.* [29] showed that miR-128b is a direct regulator of *EGFR*. Zheng *et al*. [30] determined the levels of miRNAs by real-time RT-PCR in 74 lung cancer patients and 68 age-matched cancer-free controls, and found that the levels of miR-155, miR-197, and miR-182 in the plasma of lung cancer including stage I patients were significantly elevated compared with controls

In this study, I first use the raw data and apply Procedure 2 with the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 45 high-confident miRNAs using the expression profiles in the tumor tissue. MiR-128b, miR-155 and miR-182 are included in these 45 miRNAs (miRNAs in Tables 1 and 2 for lung tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 41 high-confident miRNAs are selected in these 45 miRNAs (Table **1**).

Breast Cancer

Previous studies have shown that a number of miRNAs are deregulated in human breast cancer [31, 32]. It has been shown that human miR-9 expression levels are reduced in many breast cancer samples due to hypermethylation an epigenetic modification [33, 34]. Also, miR-155 was found to be over expressed in breast cancer [35].

In this study, I first use the raw data and apply Procedure 2 with the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 54 high-confident miRNAs in the tumor tissue. MiR-9 and miR-155 are included in these 54 miRNAs (miRNAs in Tables **1** and **2** for breast tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 49 high-confident miRNAs are selected in these 54 miRNAs (Table **1**).

Bladder Cancer

Bladder cancer is a common urologic cancer that may have the highest recurrence rate of any malignancy [36]. Ichim *et al*. [37] identified 7 miRNAs (miR-145, miR-30a-3p, miR-133a, miR-133b, miR-195, miR-125b and miR-199a*) that were significantly downregulated in bladder cancer.

In this study, I first use the raw data and apply Procedure 2 with the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 71 high-confident miRNAs using the expression profiles in the tumor tissue. MiR-145, miR-133a, miR-195, miR-125b and miR-199a* are included in these 71 miRNAs (miRNAs in Tables 1 and 2 for bladder tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 70 high-confident miRNAs in these 71 miRNAs are selected (Table **1**).

Kidney Cancer

The two most common types of kidney cancer are renal cell carcinoma and urothelial cell carcinoma of the renal pelvis. Dutta *et al.* found miR-34a overexpression in renal cell carcinoma [38]. Higher expression of miR-21 is associated with an increase in the stage of renal cancer [39].

In this study, I first use the raw data and apply Procedure 2 with the thresholds $p_0 = 0.3$ and $s = 0.999$ to find 43 high-confident miRNAs using the expression profiles in the tumor tissue. MiR-21 and miR-34a are included in these 43 miRNAs (miRNAs in Tables 1 and 2 for kidney tissue). After comparing the results using expression profiles in the tumor tissue and the results using expression profiles in the normal tissue by Procedure 1, 37 high-confident miRNAs are selected in these 43 miRNAs (Table **1**).

In this study, p_0 and s are selected to be 0.3 and 0.999 because from the previous studies that the S is suggested to be selected to be near to 1 and p_0 is suggested to be selected such that the number of selected miRNAs is close to a specified proportion of the total number of miRNAs. A more descent method for threshold selection can refer to Hsieh and Wang [40]. It is our future work to investigate this method in discovering miRNAs associated with cancers.

COMPARISON

In addition to confirming some selected miRNAs from the literature, I also adopt the Human MicroRNA Disease Database (HMDD) to investigate the performance of the RRSM in selecting the high-confident miRNAs associated with cancers. HMDD provides human microRNA-disease association data, which is manually collected from publications [41]. First, I find the miRNAs from HMDD which are shown to be associated with cancers, and then use these data to calculate the sensitivities and specificities of the RRSM in order to verify the adequacy of RRSM. The sensitivity measures the proportion of actual positives which are correctly identified as such, and the specificity measures the proportion of negatives which are correctly identified as such [42]. To calculate the sensitivity of the RRSM for selecting miRNAs associated with a cancer, I need first to find the set of miRNAs for this cancer, denoted as Set A, which are identified as the true positives in the 114 miRNAs from HMDD, and then find the set of miRNAs, denoted as Set B, which are selected by the RRSM among these true positives. Then the sensitivity of the RRSM for selecting miRNAs associated with this cancer is defined as the number of elements in Set B divided by the number of elements in Set A. The specificity of the RRSM for selecting miRNAs associated with this cancer is defined as

$$\text{specificity} = \frac{\text{the number selected by the RRSM} - \text{the number of elements in Set B}}{114 - \text{the number of elements in Set A}}$$

Since HMDD does not include the prostate cancer data, I present the results for the first two cancers listed in Table **1** which are shown in HMDD. The sensitivity and the specificity of Procedure 1 for the colon cancer are 0.405 and 0.347, respectively; the sensitivity and the specificity of Procedure 2 for the colon cancer are 0.476 and 0.444, respectively. For pancreatic cancer, the sensitivity and the specificity of Procedure 1 are 0.519 and 0.435, respectively; the sensitivity and the specificity of Procedure 2 are 0.519 and 0.612, re-

spectively. In these two cases, Procedure 2 has higher sensitivity and specificity than Procedure 1.

To evaluate the performance of the RRSM on selecting miRNAs associated with a cancer, I also adopt the RRSM to other randomized data sets to select miRNAs associated with this cancer. This randomized data set is not the true expression profile in this tumor tissue. If the RRSM leads to higher sensitivity and specificity values using the true expression profile in this tumor tissue than using this randomized data set, then it is concluded that the RRSM is efficient in selecting miRNAs associated with a cancer. Thus, I calculate the sensitivity and the specificity of the RRSM applying on other randomized data sets in selecting miRNAs associated with colon cancer and pancreatic cancer, respectively. The sensitivity and the specificity of Procedure 1 applying to other data set for the colon cancer are 0.29 and 0.23, respectively; the sensitivity and the specificity of Procedure 2 applying to other data set for the colon cancer are 0.34 and 0.31, respectively. And the sensitivity and the specificity of Procedure 1 applying to other data set for the pancreatic cancer are 0.417 and 0.223, respectively; the sensitivity and the specificity of Procedure 2 applying on other data set for the pancreatic cancer are 0.4 and 0.339, respectively. Since the sensitivity and the specificity of the RRSM applying to the valid data set are significantly higher than those applying to the other randomized data sets, it reveals that the RRSM is an efficient method.

DISCUSSION

In this study, I adopt the RRSM to find high-confident miRNAs associated with colon, prostate, pancreatic, lung, breast and kidney tumor development, respectively. The same threshold is used for RRSM in the 7 organic tissues. A more feasible approach is to adopt different thresholds which depend on the characteristics of an organic tissue to reduce the false positive or false negative rates. It is our future research direction.

In addition, in dealing with the expression profiles in the 7 organic tissues, for the colon tissue and pancreatic tissue, the normalized data are used; for the other organic tissues, the raw data without normalization are used. Hsieh and Wang [13] used the raw data to select miRNA targets across the 88 tissues because the false positive rate of using the raw data is less than that of using the normalized data. Therefore, the raw data are used in this study for most tissues. The reason of using the normalized data for the colon tissue and the pancreatic tissue is that too many or much fewer miRNAs are selected by RRSM under the thresholds $p_0 = 0.3$ and $s = 0.999$ with the raw data. Therefore, for these two cases, the normalized data are used.

Moreover, it is plausible to explore a more suitable regression model as a model in the RRSM instead of a linear regression model to find high-confident miRNAs in a future study. Other models, such as non-linear regression models, can be considered as alternative models to be used in the high-confident miRNAs exploration.

CONFLICT OF INTEREST

The author confirms that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work has been supported by "Aiming for the Top University Program" of the National Chiao Tung University and Ministry of Education, National Science Council, National Center for Theoretical Sciences, Taiwan.

REFERENCES

[1] Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **2005**, *120*(1), 15-20.

[2] Lee, R.C.; Feinbaum, R.L.; Ambros, V. The C-Elegans Heterochronic Gene Lin-4 Encodes Small Rnas with Antisense Complementarity to Lin-14. *Cell*, **1993**, *75*(5), 843-854.

[3] Sassen, S.; Miska, E.A.; Caldas, C. MicroRNA: Implications for cancer. *Virchows Archiv.*, **2008**, *452*(1),1-10.

[4] Foldes-Papp, Z.; Konig, K.; Studier, H.; Buckle, R.; Breunig, H.G.; Uchugonova, A.; Kostner, G.M. Trafficking of Mature miRNA-122 into the Nucleus of Live Liver Cells. *Curr. Pharm. Biotechnol.*, **2009**, *10*(6), 569-578.

[5] Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; Rassenti, L.; Kipps, T.; Negrini, M.; Bullrich, F.; Croce, C.M. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*(24),15524-15529.

[6] Croce, C.M. Causes and consequences of microRNA dysregulation in cancer. *Nat. Rev. Genet.*, **2009**, *10*(10),704-714.

[7] John, B.; Enright, A.J.; Aravin, A.; Tuschl, T.; Sander, C.; Marks, D.S. Human microRNA targets. *PLoS Biol.*, **2004**, *2*(11),1862-1879.

[8] Lewis, B.P.; Shih, I.H.; Jones-Rhoades, M.W.; Bartel, D.P.; Burge, C.B. Prediction of mammalian microRNA targets. *Cell*, **2003**, *115*(7), 787-798.

[9] Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell*, **2009**, *136*(2), 215-233.

[10] Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **2005**, *433*(7027),769-773.

[11] Wang, H.; Li, W.H. Increasing MicroRNA target prediction confidence by the relative R-2 method. *J. Theor. Biol.*, **2009**, *259*(4),793-798.

[12] Huang, J.C.; Morris, Q.D.; Frey, B.J. Bayesian inference of microRNA targets from sequence and expression data. *J. Comput. Biol.*, **2007**, *14*(5), 550-563.

[13] Hsieh, W.J.; Wang, H. Human microRNA target identification by RRSM. *J. Theor. Biol.*, **2011**, *286*, 79-84.

[14] Guan, P.; Yin, Z.H.; Li, X.L.; Wu, W.; Zhou, B.S. Meta-analysis of human lung cancer microRNA expression profiling studies comparing cancer tissues with normal tissues. *J. Exp. Clin. Canc. Res.*, **2012**, *31*.

[15] van Schooneveld, E.; Wouters, M.C.A.; Van der Auwera, I.; Peeters, D.J.; Wildiers, H.; Van Dam, P.A.; Vergote, I.; Vermeulen, P.B.; Dirix, L.Y.; Van Laere, S.J. Expression profiling of cancerous and normal breast tissues identifies microRNAs that are differentially expressed in serum from patients with (metastatic) breast cancer and healthy volunteers. *Breast Cancer Res.*, **2012**, *14*(1).

[16] Li, B.R.; Chang, J.T.; Chu, Y.L.;Kang, H.F.; Yang, J.; Jiang, J.T.; Ma, H.B. Membrane proteomic analysis comparing squamous cell lung cancer tissue and tumour-adjacent normal tissue. *Cancer Lett.*, **2012**, *319*(1), 118-124.

[17] Hamfjord, J.; Stangeland, A.M.; Hughes, T.; Skrede, M.L.; Tveit, K.M.; Ikdahl, T.; Kure, E.H. Differential expression of miRNAs in colorectal cancer: Comparison of paired tumor tissue and adjacent normal mucosa using high-throughput sequencing. *Plos One*, **2012**, *7*(4).

[18] Huang, J.C.; Babak, T.; Corson, T.W.; Chua, G.; Khan, S.; Gallie, B.L.; Hughes, T.R.; Blencowe, B.J.; Frey, B.J.; Morris, Q.D. Using

expression profiling data to identify human microRNA targets. *Nat. Methods*, **2007**, *4*(12), 1045-1049.

[19] Hsieh, W.J.; Lin, F.M.; Huang, H.D.; Wang, H. Investigating microRNA-Target interaction-supported tissues in human cancer tissues based on miRNA and target gene expression profiling. *Plos One*, **2014**, *9*(4), e95697.

[20] Michael, M.Z.; O'Connor, S.M.; Pellekaan, N.G.V.; Young, G.P.; James, R.J. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol. Cancer Res.*, **2003**, *1*(12), 882-891.

[21] Reid, J.F.; Sokolova, V.; Zoni, E.; Lampis, A.; Pizzamiglio, S.; Bertan, C.; Zanutto, S.; Perrone, F.; Camerini, T.; Gallino, G.; Verderio, P.; Leo, E.; Pilotti, S.; Gariboldi, M.; Pierotti, M.A. miRNA Profiling in colorectal cancer highlights miR-1 involvement in MET-dependent proliferation. *Mol. Cancer Res.*, **2012**, *10*(4), 504-515.

[22] Wang, L.; Tang, H.; Thayanithy, V.; Subramanian, S.; Oberg, A.L.; Cunningham, J.M.; Cerhan, J.R.; Steer, C.J.; Thibodeau, S.N. Gene Networks and microRNAs implicated in aggressive prostate cancer. *Cancer Res.*, **2009**, *69*(24), 9490-9497.

[23] Hulf, T.; Sibbritt, T.; Wiklund, E.; Patterson, K.; Song, J.; Stirzaker, C.; Qu, W.; Nair, S.; Horvath, L.; Armstrong, N.; Kench, J.; Sutherland, R.; Clark, S. Epigenetic-induced repression of microRNA-205 is associated with MED1 activation and a poorer prognosis in localized prostate cancer. *Oncogene*, **2013**, *32*(23), 2891-2899.

[24] Ali, S.; Almhanna, K.; Chen, W.; Philip, P.A.; Sarkar, F.H. Differentially expressed miRNAs in the plasma may provide a molecular signature for aggressive pancreatic cancer. *Am. J. Transl. Res.*, **2011**, *3*(1), 28-47.

[25] Ali, S.; Ahmad, A.; Banerjee, S.; Padhye, S.; Dominiak, K.; Schaffert, J.M.; Wang, Z.W.; Philip, P.A.; Sarkar, F.H. Gemcitabine sensitivity can be induced in pancreatic cancer cells through modulation of miR-200 and miR-21 expression by curcumin or its analogue CDF. *Cancer Res.*, **2010**, *70*(9), 3606-3617.

[26] Bao, B.; Ali, S.; Kong, D.J.; Sarkar, S.H.; Wang, Z.W.; Banerjee, S.; Aboukameel, A.; Padhye, S.; Philip, P.A.; Sarkar, F.H. Anti-tumor activity of a novel compound-CDF is mediated by regulating mir-21, mir-200, and pten in pancreatic cancer. *Plos One*, **2011**, *6*(3).

[27] Vansteenkiste, J.; De Ruysscher, D.; Eberhardt, W.E.; Lim, E.; Senan, S.; Felip, E.; Peters, S.; Group, E.G.W. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.*, **2013**, *24*(Suppl 6), vi89-98.

[28] Lin, P.Y.; Yu, S.L.; Yang, P.C. MicroRNA in lung cancer. *Br. J. Cancer*, **2010**, *103*(8), 1144-1148.

[29] Weiss, G.J.; Bemis, L.T.; Nakajima, E.; Sugita, M.; Birks, D.K.; Robinson, W.A.; Varella-Garcia, M.; Bunn, P.A.; Haney, J.; Helfrich, B.A.; Kato, H.; Hirsch, F.R.; Franklin, W.A. EGFR regulation by microRNA in lung cancer: correlation with clinical response and survival to gefitinib and EGFR expression in cell lines. *Ann. Oncol.*, **2008**, *19*(6), 1053-1059.

[30] Zheng, D.L.; Haddadin, S.; Wang, Y.; Gu, L.Q.; Perry, M.C.; Freter, C.E.; Wang, M.X. Plasma microRNAs as novel biomarkers for early detection of lung cancer. *Int. J. Clin. Exp. Pathol.*, **2011**, *4*(6), 575-586.

[31] Iorio, M.V.; Ferracin, M.; Liu, C.G.;Veronese, A.; Spizzo, R.; Sabbioni, S.; Magri, E.; Pedriali, M.; Fabbri, M.; Campiglio, M.; Menard, S.; Palazzo, J.P.; Rosenberg, A.; Musiani, P.; Volinia, S.; Nenci, I.; Calin, G.A.; Querzoli, P.; Negrini, M.; Croce, C.M. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* **2005**, *65*(16),7065-7070.

[32] Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; Prueitt, R.L.; Yanaihara, N.; Lanza, G.; Scarpa, A.; Vecchione, A.; Negrini, M.; Harris, C.C.; Croce, C.M. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*(7),2257-2261.

[33] Lehmann, U.; Hasemeier, B.; Christgen, M.; Muller, M.; Romermann, D.; Langer, F.; Kreipe, H. Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J. Pathol.*, **2008**, *214*(1),17-24.

[34] Hsu, P.Y.; Deatherage, D.E.; Rodriguez, B.A.T.; Liyanarachchi, S.; Weng, Y.I.; Zuo, T.; Liu, J.; Cheng, A.S.L.; Huang, T.H.M. Xenoestrogen-Induced Epigenetic Repression of microRNA-9-3 in Breast Epithelial Cells. *Cancer Res.*, **2009**, *69*(14),5936-5945.

[35] Mattiske, S.; Suetani, R.J.; Neilsen, P.M.; Callen, D.F. The oncogenic role of miR-155 in breast cancer. *Cancer Epidem. Biomar.*, **2012**, *21*(8),1236-1243.

[36] Anastasiadis, A.; de Reijke, T.M. Best practice in the treatment of nonmuscle invasive bladder cancer. *Therapeut. Adv. Urol.*, **2012**, *4*(1),13-32.

[37] Ichimi, T.; Enokida, H.; Okuno, Y.; Kunimoto, R.; Chiyomaru, T.; Kawamoto, K.; Kawahara, K.; Toki, K.; Kawakami, K.; Nishiyama, K.; Tsujimoto, G.; Nakagawa, M.; Seki, N. Identification of novel microRNA targets based on microRNA signatures in bladder cancer. *Int. J. Cancer*, **2009**, *125*(2),345-352.

[38] Dutta, K.K.; Zhong, Y.; Liu, Y.T.; Yamada, T.; Akatsuka, S.; Hu, Q.; Yoshihara, M.; Ohara, H.; Takehashi, M.; Shinohara, T.; Masutani, H.; Onuki, J.; Toyokuni, S. Association of microRNA-34a overexpression with proliferation is cell type-dependent. *Cancer Sci.*, **2007**, *98*(12),1845-1852.

[39] Zaman, M.S.; Shahryari, V.; Deng, G.R.; Thamminana, S.; Saini, S.; Majid, S.; Chang, I.; Hirata, H.; Ueno, K.; Yamamura, S.; Singh, K.; Tanaka, Y.; Tabatabai, Z.L.; Dahiya, R. Up-regulation of microRNA-21 correlates with lower kidney cancer survival. *Plos One*, **2012**, *7*(2).

[40] Hsieh, W.J.; Wang, H. RRSM with a data-dependent threshold for miRNA target prediction. *J. Theor. Biol.*, **2013**, *337C*, 54-60.

[41] Lu, M.; Zhang, Q.; Deng, M.; Miao, J.; Guo, Y.; Gao, W.; Cui, Q. An analysis of human microRNA and disease associations. *Plos One*, **2008**, *3*(10), e3420.

[42] Wang, B.; Wang, X.F.; Xi, Y. Normalizing bead-based microRNA expression data: A measurement error model-based approach. *Bioinformatics*, **2011**, *27*(11), 1506-1512.

Received: June 14, 2013 **Revised: December 10, 2013** **Accepted: May 13, 2014**