Identification of a three miRNA signature as a novel potential prognostic biomarker in patients with bladder cancer

Fei Peng1,*, Hui Li2,*, Hailang Xiao3, Ling Li3, Yan Li4 and Yi Wu1

1Department of Laboratory, The First Affiliated Hospital of Hunan Normal University, Hunan Provincial People’s Hospital, Changsha, China
2Reproductive Department, Xiangya Hospital, Central South University, Changsha, China
3Nephrology Department, The First Affiliated Hospital of Hunan Normal University, Hunan Provincial People’s Hospital, Changsha, China
4Department of Paediatrics, The First Affiliated Hospital of Hunan Normal University, Hunan Provincial People’s Hospital, Changsha, China

*These authors have contributed equally to this work

Correspondence to: Yi Wu, email: wuyi19701210@outlook.com

Keywords: BLCA; miRNA; prognosis; biomarker; TCGA

Received: November 23, 2016 Accepted: June 20, 2017 Published: November 06, 2017

Copyright: Peng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

There is not a good biomaker that is closely related to survival time for bladder cancer (BLCA). The aim of the study is to identify a miRNA signature that could predict prognosis in BLCA patients according to the data from The Cancer Genome Atlas (TCGA). A total of 377 BLCA patients were finally enrolled in the study. The three miRNA signature was identified by Multivariate Cox proportional hazards analyses that common clinical variables were controlled. The three microRNA signature showed greater predicting prognosis capacity for predicting 5-year survival in BLCA with an AUC of 0.664, 0.681 and 0.668 in Train set, Test set and Total set respectively. Furthermore, there was a significant difference between high score and low score in Total set (P=3e-05), Test set (P=0.00435) and Train set (P=0.00143), respectively. Therefore, these results provided a new prospect for prognostic biomarker of BLCA.

INTRODUCTION

Bladder cancer is the ninth most common malignancies worldwide [1]. 76,960 new cases (4.6% of all new cancer cases) and 16,390 deaths of bladder cancer were estimated in the United States during 2016 [2]. It brings a heavy economic burden to our society. In the USA, the annual national cost of bladder cancer care has outpaced inflation in recent years and is expected to reach $US5.25 billion ($US4.91 billion, 2010 values) in 2020 [3]. Bladder cancer has different biological characteristics. Superficial tumors are presented in 70% newly diagnosed patients which are usually not life threatening but tend to recur. Other 30% patients present with muscle-invasive status have a high risk of death from distant metastases. The invasive cystoscopy is the most common methods to detect suspicious bladder cancer because of the low sensitivity of other diagnostic tests like urine cytology [4]. Specific and sensitive biomarkers that can provide reliable information for tumor diagnosis or the prognosis may allow a more precise assessment and achieve a better-targeted effective therapy for bladder cancer. It is also of great importance to help illuminate the mechanism of bladder cancer development and progression.

MicroRNAs (miRNAs) are a family of short (typically 18–25 nucleotides), single-stranded and highly conserved non-coding RNAs that play an important role in the regulation of expression and function of eukaryotic genomes. They can bind to the 3'-untranslated region of target mRNAs through base pairing, resulting in mRNA degradation or translational repression [5]. Studies showed that miRNAs play a critical role in a variety of biological processes, such as cellular proliferation, apoptosis, migration, and angiogenesis [6–9]. All these
processes are strongly related to tumor formation and development, which indicate a significant role of miRNAs in the pathogenesis of cancers. miR-17-92 is found overexpressed in human lung cancers and enhanced cell proliferation [10]. Five miRNAs (miR-185, miR-20a, miR-210, miR-25 and miR-92b) were upregulated in gastric cancer and might serve as a diagnostic marker [11]. Other studies also revealed the relationship between miRNAs and other kinds of cancers, such as breast cancer [12], colorectal cancer [13] and pancreatic cancer [14]. Several studies have demonstrated the value of miRNAs as a prognostic marker in bladder cancer [15, 16]. However, most of the studies were carried out based on small number of patients or focused on just a specific miRNA. Here, we aimed to conduct the research using the dataset retrieved from The Cancer Genome Atlas (TCGA, http://cancergenome.nih.gov/) to identify a panel of miRNA signature which could predict prognosis in bladder cancer.

RESULTS

Characters of the datasets

A total of 437 miRNA expression profiles (level 3 data) were obtained from TCGA. In miRNA profiles, there are 1881 human miRNAs in BLCA samples (n = 418) and normal tissues (n = 19). The clinical data for those patients obtained from TCGA were available. According to included criteria, a total of 377 BLCA patients were finally enrolled in the study. For subsequent analysis, we randomly divided the total patients into the training set (n = 189) and testing set (n = 188) respectively. There was no significant difference on the clinical covariates between the two sets (P > 0.05). (Table 1).

Identification of differentially expressed miRNAs in BLCA patients

Analysis of miRNA expression profiles in BLCA patient tissues (n = 418) compared with normal tissues (n = 19) identified a total of 392 differentially expressed miRNAs (logFC > 1 or logFC < -1, P < 0.05 after FDR adjustment). Of these, 295 miRNAs were overexpressed in BLCA patients (Supplementary Table 1), and 97 miRNAs were downexpressed in BLCA patients. (Supplementary Table 2).

The miRNA signature risk score as an independent indicator for BLCA prognosis

The risk score for each patient was calculated based on the 3 miRNAs. we developed a three microRNA signature. The miRNAs expression level was as the log2 reads per million of total aligned miRNA reads. The prognostic score was calculated as follows: Prognostic-

The univariate Cox regression analyses showed that Pathologic stage (P = 7.137E-07), Pathology T stage (P = 0.0002), Pathology N stage (P = 2.697E-07) and the three microRNA signature (P = 9.815E-07) were significantly related with survival time of BLCA patients, and the multivariate Cox regression analyses revealed that miRNA model scores was an independent prognostic factor (Table 2). In addition, the cut-off value was 14.05851 according to Train set and there was a significant difference between high score and low score in Total set (P = 3e-05), Test set (P = 0.00435) and Train set (P = 0.00143), respectively. (Figure 2). Morever, we found high score in the three microRNA signature had a shorten survival time in all groups (Supplementary Figure 3-11) except Female group (P = 0.194) (Supplementary Figure 1) and Radiation therapy group and (P = 0.317) (Supplementary Figure 2). However, there was a significant difference in the remaining groups (Supplementary Figure 3-11)

Furthermore, The three microRNA signature showed greater predicting prognosis capacity for predicting 5-year survival in BLCA with an AUC of 0.664, 0.681 and 0.668 in Train set, Test set and Total set respectively (Figure 3). We also evaluated the three microRNA signature on the each subgroup of clinical characteristics and found it was significantly predictive for predicting 5-year survival in all groups (Supplementary Figure 12-22).

Target prediction and functional enrichment of the three microRNA signature in BLCA

The numbers of the target genes of the three miRNAs were 1034 and 1892, which were predicted by two data sets using miRanda and Targetscan. A total of 853 target genes were included in the two data sets (Supplementary Tables 3). We performed enrichment analyses to elucidate the biological function of target genes of the three microRNA signature. Finally, Gene ontology (GO) analysis revealed that there were 35 of the proteins were associated with biological process (BP) (Supplementary Tables 4, 46 of the proteins with cellular component (CC) (Supplementary Tables 4), and 56 of the proteins with molecular function (MF) (Supplementary Tables 4), respectively. The top ten enriched functional analysis was shown in Figure 4. The top of enriched
biological process, cellular component and molecular function were protein phosphorylation, cytoplasm and protein binding respectively. Therefore, a total of 48 KEGG pathways were enriched by the three microRNA signature (Supplementary Tables 5). The top enriched KEGG pathway was the MAPK signaling pathway (Figure 4).

**DISCUSSION**

MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression by binding to the 3' untranslated region of their target mRNAs [17]. Moreover, miRNAs have been recognized as important intervention targets and predictive tools for various
human cancers because of the stability and convenience of miRNA detection [18, 19]. Xu et al [20] found miR-33b-3p can improve the efficacy of chemotherapies for treating lung cancers. Cecene et al [21] indicated that miR-195 can be a clinically useful biomarker in Turkish breast cancer patients. However, there is not a excellent indicator to predict the survival time for BLCA patients. Hence, we made a three miRNA signature as to evaluate the prognosis of BLCA patients.

In this study, we applied Multivariate Cox proportional hazards analyses that after controlling for common clinical variables to obtain the miRNA signature which can predict the prognosis of BLCA patients. After screening the best miRNA signature by the method of AIC, a three miRNA signature consisting of hsa-mir-337, hsa-mir-3913-2 and hsa-mir-497 was identified as an independent predictor for prognosis of BLCA. There was a significant difference between high score and low score in all groups except the groups of radiation therapy, non-white and female. Conclusion should be cautious to confirm in the radiation therapy group due to only 19 patients are obtained in this group. Due to different clinical treatments could affect the relapse-free survival (RFS) and other survival endpoints of the patients, the conclusion that treated with Radiation therapy doesn’t make much sense. In addition, the score of the three miRNA signature was not related to the survival time of female group and non-white group in BLCA patients according to our analysis. As for 5-year survival rate, the three miRNA signature have a good performance in all groups. The AUC of ROC curve for the three miRNA signature predicting 5-year survival rate were 0.664, 0.681 and 0.668 in the Training set, Testing set and Total set, respectively. However, the AUC value of ROC curve was not large that indicated the three miRNA signature for predicting 5-year survival needs to improve the accuracy and the specificity. Furthermore, we found only the score of three miRNA signature had a close association with the survival time of BLCA patients through the multivariate Cox regression analysis. For other clinical characteristics, no significant difference was found in the multivariate Cox regression analysis.

Table 2: The association of clinical factors and the miRNA signature score with survival time in BLCA patients

| variables | Univariate analysis | Multivariate analysis |
|-----------|---------------------|----------------------|
|           | HR(95%CI)           | P value              | HR(95%CI) | P value |
| Age(>65 vs ≤65) | 2.270 | 0.0116 | 1.197 | 0.4263 |
| Pathologic stage | 1.821 | 7.137E-07 | 1.25 | 0.1738 |
| Pathology T stage | 1.621 | 0.0002 | 1.285 | 0.1070 |
| Pathology N stage | 1.585 | 2.6971E-07 | 1.285 | 0.1070 |
| Gender(Male vs Female) | 0.819 | 0.2986 | | |
| Radiation therapy(Yes vs No) | 0.951 | 0.9123 | | |
| race(white vs Non-white) | 1.221 | 0.6547 | | |
| miRNA model scores(High score vs Low score) | 7.101 | 9.815E-07 | 5.972 | 1.50E-05 |

Figure 1: Three miRNA expressed in the score of BLCA patients.
Figure 2: Kaplan–Meier curve for three miRNA signature and the survival time of BLCA patients. (A) There is a significant difference between high score and low score in the survival time of BLCA patients in the Training set ($P=0.00143$) (B) There is a significant difference between high score and low score in the survival time of BLCA patients in the Testing set ($P=0.00435$) (C) There is a significant difference between high score and low score in the survival time of BLCA patients in the Total set ($P=3\times10^{-5}$).

Figure 3: ROC curves for the three miRNA signature in predicting 5-year survival rate in BLCA patients. (A) The ROC curve for predicting 5-year survival had an AUC of 0.664 in the Training set. (B) The ROC curve for predicting 5-year survival had an AUC of 0.681 in the Testing set. (C) The ROC curve for predicting 5-year survival had an AUC of 0.668 in the Total set.
analysis. That is to say, the three miRNA signature can be regarded as the only independent predictor for the patients of BLCA which means our miRNA signature is closely related to the survival time of BLCA. Wang et al. [22] indicated that hsa-miR-337-3p expression did not dramatically affect gastric cancer cell proliferation, but transfection of the hsa-miR-337-3p mimic did reduce gastric cancer cell invasion capacity. Jiang et al. [23] found that MiR-497 directly targeted 3′-UTR of Nrdp1 mRNA to inhibit its translation.

A total of 853 target genes were included in the two database of miRanda and Targetscan. Enrichment and function analysis was made, because we want to figure out the progression how a normal cell developed into a cancer cell. In enrichment and function analysis of DAVID database, our results shows that the top of enriched biological process, cellular component and molecular function were protein phosphorylation, cytoplasm and protein binding, respectively. As for protein phosphorylation, many studies have indicated it was closely related to cancer progress [24–25]. For KEGG pathway analysis, the most important signaling pathways was MAPK. It participate in the process of cancer development, such as endometrial cancer [26], lung cancer [27]. In addition, Kakkar et al. [28] showed that MAPK pathways play a important role for altered signalling in DNT pathogenesis and targeted therapies. Liu et al. [29] found that B7-H1 regulates p38 MAPK activation via association with DNA-PKcs.

Although the three miRNA signature is a excellent predictor for survival time of BLCA patients, some limitation is also existed. Firstly, there are only 377 patients of BLCA included in our study from TCGA that needs a validation or more cases to confirm. Secondly, the AUC of ROC curve in all groups is between 0.5 and 0.7 that means the three miRNA signature didn’t perform very well in predicting 5-year survival rate in BLCA patients. In addition, the three miRNA signature was suitable for most subgroups of BLCA patients. However, there was no association between the three miRNA signature and two subgroups of female group and non-white group.

**Figure 4:** The top ten of GO term and pathway by target genes of three miRNA signature in BLCA patients. (A) GO BP categories of three miRNA signature. (B) GO CC categories of three miRNA signature. (C) GO MF categories of three miRNA signature. (D) KEGG pathway of three miRNA signature.
MATERIALS AND METHODS

Data sources and screening

The miRNA expression profiles (level 3 data) and corresponding clinical data for BLCA patients were obtained from TCGA data portal. Both the miRNA profiles data and clinical data of BLCA are publically available and open-access. These patients’ extended demographics were characterized by the TCGA. The patients were included in the study to meet the following criteria: (1) patients with fully characterized (clinical data and miRNA profiles) tumors; (2) patients with at least 1 month of survival time.

Identification of differentially expressed miRNAs between BLCA and normal tissue

To identify miRNAs differentially expressed between BLCA and normal tissues, the raw counts of miRNA expression obtained from the TCGA dataset that consist of 418 BLCA samples and 19 normal tissue. One miRNA expression filter was miRNAs expressed in at least two normal or tumor samples. The expression differences were characterized by logFC(|log 2 fold change|>1) and associated P-values(P<0.05). The analysis was performed using the R package of edgeR [30]. Furthermore, the R package of pheatmap was used to draw the Heatmap.

Survival analysis and ROC curve of prognostic model

First of all, the expression level of each miRNA was log2 transformed for further analysis. 377 BLCA patients were random divided into Training set and Testing set. Univariate Cox proportional hazards regression with significance level set as 0.001 was performed to find out the miRNAs significantly associated with survival time in the Training set. Multivariate Cox proportional hazards analyses were also conducted to evaluate the independent prognostic value of the miRNA signature after controlling for common clinical variables in the Training set. Akaike information criterion [31][AIC] was used to choose the best miRNA signature which was not only associated with survival time in Training set but also in Testing set. we used the median value of miRNA signature in Training set as cut off value and divided Training set, Testing set and Total set into high score group and low score group, respectively. Kaplan-Meier curves were used to estimate the survival for BLCA patients with high score or low score and ROC curve was adopt to predict the 5-year survival rate for BLCA patients with high score or low score. Chi-square test was used to evaluate the difference between Training set and Testing set and P value was set as 0.05. All analysis were performed using R (Packages: survival and survival ROC).

Target prediction and enrichment analysis

The target genes of miRNAs were predicted by two programs including miRanda and Targetscan. The target genes were selected by miRanda and Targetscan that the final target genes were selected were included in all the two data sets). The enrichment analysis of these target genes was analyzed using DAVID online analysis [32] (https://david.ncifcrf.gov/). The gene sets containing less than 5 genes overlapping were removed from the DAVID analysis, and analysis for significance was determined when P values(P<0.05).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This work was not supported by any institution.

REFERENCES

1. May M. Statistics: attacking an epidemic. Nature. 2014; 509:S50–S51.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7–30.
3. Santos F, Dragomir A, Zakaria AS, Kassouf W, Aprikian A. Predictors of costs associated with radical cystectomy for bladder cancer: a population-based retrospective cohort study in the province of Quebec, Canada. J Surg Oncol. 2016; 113:223-228.
4. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. Lancet. 2009; 374:239–249.
5. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005; 120:15–20.
6. Suárez Y, Sessa WC. MicroRNAs as novel regulators of angiogenesis. Circ Res. 2009; 104:442–454.
7. Xu P, Guo M, Hay BA. MicroRNAs and the regulation of cell death. Trends Genet. 2004; 20:617–624.
8. Carleton M, Cleary MA, Linsley PS. MicroRNAs and cell cycle regulation. Cell Cycle. 2007; 6:2127–2132.
9. Zhang B, Pan X, Cobb GP, Anderson TA. microRNAs as oncogenes and tumor suppressors. Dev Biol. 2007; 302:1–12.
10. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, Yatabe Y, Kawahara K, Sekido Y, Takahashi T. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res. 2005; 65:9628–9632.
11. Zhou X, Zhu W, Li H, Wen W, Cheng W, Wang F, Wu Y, Qi L, Fan Y, Chen Y, Ding Y, Xu J, Qian J, et al. Diagnostic
value of a plasma microRNA signature in gastric cancer: a microRNA expression analysis. Sci Rep. 2015; 5:11251.

12. Wu X, Somlo G, Yu Y, Palomares MR, Li AX, Zhou W, Chow A, Yen Y, Rossi JJ, Gao H, Wang J, Yuan YC, Frankel P, et al. De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. J Transl Med. 2012; 10:42.

13. Suto T, Yokobori T, Yajima R, Morita H, Fujii T, Yamaguchi S, Altan B, Tsutsumi S, Asao T, Kuwano H. MicroRNA-7 expression in colorectal cancer is associated with poor prognosis and regulates cetuximab sensitivity via EGFR regulation. Carcinogenesis. 2015; 36:338–345.

14. Chitkara D, Mittal A, Mahato RI. miRNAs in pancreatic cancer: therapeutic potential, delivery challenges and strategies. Adv Drug Deliv Rev. 2015; 81:34–52.

15. Feng Y, Liu J, Kang Y, He Y, Liang Y, Yang P, Yu Z. miR-19a acts as an oncoembryonic microRNA and is up-regulated in bladder cancer. J Exp Clin Cancer Res. 2014; 33:67.

16. Zhang J, Wang S, Han F, Li J, Yu L, Zhou P, Chen Z, Xue S, Dai C, Li Q. MicroRNA-542-3p suppresses cellular proliferation of bladder cancer cells through post-transcriptionally regulating survivin. Gene. 2016; 579:146–152.

17. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell. 2009; 139:642-55.

18. Kong YW, Ferland-McCollough D, Jackson TJ, Bushell M. MicroRNAs in cancer management. Lancet Oncol, 2012; 13: e249–58.

19. Corsini LR, Bronte G, Terrasi M, Amodeo V, Fanale D, Fiorentino E, Cicero G, Bazan V, Russo A. The role of microRNAs in cancer: diagnostic and prognostic biomarkers and targets of therapies. Expert Opin Ther Tar, 2012; 16: S103–9.

20. Xu S, Huang H, Chen YN, Deng YT, Zhang B, Xiong XD, Yuan Y, Zhu Y, Huang H, Xie L, Liu X. DNA damage responsive miR-33b-3p promoted lung cancer cells survival and cisplatin resistance by targeting p21WAF1/CIP1. Cell Cycle. 2016; 15:2920-2930.

21. Cecene G, Ak S, Eskiler GG, Demirdogen E, Erturk E, Gokgoz S, Polatkan V, Egeli U, Tunca B, Tezcan G, Topal U, Tolunay S, Tasdelen I. Circulating miR-195 as a Therapeutic Biomarker in Turkish Breast Cancer Patients. Asian Pacific Journal of Cancer Prevention, 2016; 17:4241-4246.

22. Wang Z, Wang J, Yang Y, Hao B, Wang R, Li Y, Wu Q. Loss of has-miR-337-3p expression is associated with lymph node metastasis of human gastric cancer. Journal of Experimental & Clinical Cancer Research 2013, 32:76.

23. Liu X, Wu X, Cao S, Harrington SM, Yin P, Mansfield AS, Dong H. Mir-497 promotes metastasis of colorectal cancer cells through Nrdp1 inhibition. Tumor Biol. (2015) 36:7641–7647.

24. Steen HC, Kotredes KP, Nogusa S, Harris MY, Balachandran S, Gamero AM. Phosphorylation of STAT2 on Serine-734 negatively regulates the IFNα-induced antiviral response. 2016. pii: jcs. 185421.

25. Abhishek K, Sardar AH, Das S, Kumar A, Ghosh AK, Singh R, Saini S, Mandal A, Verma S, Kumar A, Purkait B, Dikhit MR, Das P. Phosphorylation of translation initiation factor 2 alpha in Leishmania donovani under stress is necessary for parasite survival. Mol Cell Biol. 2016. pii: MCB. 00344-16.

26. Wang J, Sun X, Zhang H, Wang Y, Li Y. MPA influences tumor cell proliferation, migration, and invasion induced by RANKL through PRB involving the MAPK pathway in endometrial cancer. Oncology reports 33.2 (2015): 799-809.

27. Yao J, Ma C, Gao W, Liang J, Liu C, Yang H, Yan Q, Wen Q. Fentanyl induces autophagy via activation of the ROS/ MAPK pathway and reduces the sensitivity of cisplatin in lung cancer cells. Autophagy. 2016; 19:20.

28. Kakkar A, Majumdar A, Kumara A, Tripathib M, Pathaka P, Sharma MC, Suria V, Chandrac SP, Sarkara C. Alterations in BRAF gene, and enhanced mTOR and MAPK signaling in dysembryoplastic neuroepithelial tumors (DNTs). Epilepsy Research. 2016; 141-151.

29. Liu X, Wu X, Cao S, Harrington SM, Yin P, Mansfield AS, Donga H. B7-H1 antibodies lose antitumor activity due to activation of p38 MAPK that leads to apoptosis of tumor-reactive CD8+ T cells. Sci Rep. 2016; 6: 36722.

30. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England). 2010; 26:139-140.

31. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic biology. 2004; 53:793-808.

32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009; 4:44-57.