



Published in final edited form as:

Urol Oncol. 2013 November ; 31(8): 1701–1708. doi:10.1016/j.urolonc.2012.06.010.

Plasma microRNA profiles for bladder cancer detection

Liana Adam, M.D., Ph.D.^a, Matthew F. Wszolek, M.D.^a, Chang-Gong Liu, Ph.D.^b, Wang Jing, Ph.D.^c, Lixia Diao, Ph.D.^c, Alexander Zien, Ph.D.^d, Jitao D. Zhang^d, David Jackson, Ph.D.^d, and Colin P.N. Dinney, M.D.^{a,*}

^aDepartment of Urology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

^bDepartment of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

^cDepartment of Bioinformatics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

^dMolecular Health, GmbH, Heidelberg, Germany

Abstract

Background—Bladder cancer (BC) is a burdensome disease with significant morbidity, mortality, and cost. The development of novel plasma-based biomarkers for BC diagnosis and surveillance could significantly improve clinical outcomes and decrease health expenditures. Plasma miRNAs are promising biomarkers that have yet to be rigorously investigated in BC.

Objective—To determine the feasibility and efficacy of detecting BC with plasma miRNA signatures.

Materials and methods—Plasma miRNA was isolated from 20 patients with bladder cancer and 18 noncancerous controls. Samples were analyzed with a miRNA array containing duplicate probes for each miRNA in the Sanger database. Logistic regression modeling was used to optimize diagnostic miRNA signatures to distinguish between muscle invasive BC (MIBC), non-muscle-invasive BC (NMIBC) and noncancerous controls.

Results—Seventy-nine differentially expressed plasma miRNAs (local false discovery rate [FDR] <0.5) in patients with or without BC were identified. Some diagnostically relevant miRNAs, such as miR-200b, were up-regulated in MIBC patients, whereas others, such as miR-92 and miR-33, were inversely correlated with advanced clinical stage, supporting the notion that miRNAs released in the circulation have a variety of cellular origins. Logistic regression modeling was able to predict diagnosis with 89% accuracy for detecting the presence or absence of BC, 92% accuracy for distinguishing invasive BC from other cases, 100% accuracy for distinguishing MIBC from controls, and 79% accuracy for three-way classification between MIBC, NMIBC, and controls.

*Corresponding author. Tel.: +1-713-792-3250; fax: +1-713-794-4824. cdinney@mdanderson.org (Colin P.N. Dinney).

The authors declare no conflict of interest.

Conclusions—This study provides preliminary data supporting the use of plasma miRNAs as a noninvasive means of BC detection. Future studies will be required to further specify the optimal plasma miRNA signature, and to apply these signatures to clinical scenarios, such as initial BC detection and BC surveillance.

Keywords

Bladder cancer; MicroRNAs; Plasma; Diagnostic

1. Introduction

Bladder cancer (BC) follows a dual-track carcinogenesis concept [1,2]. Most BCs are non-muscle-invasive (NMIBC) lesions that frequently recur, but rarely progress to muscle-invasive disease. Despite the tendency for a relatively indolent course, NMIBC has a substantial impact on health care due to the costs incurred from frequent cystoscopic surveillance [3,4]. In contrast, about 20% of BCs are muscle-invasive at diagnosis (MIBC). These tumors, unlike their NMIBC counterparts, arise from severe dysplasia or carcinoma in situ (CIS). MIBC tumors account for the bulk of BC-related mortality, which amounts to approximately 15,000 deaths per year in the United States, or about 20% of the annual BC incidence [5].

Developing novel noninvasive diagnostics for BC detection and surveillance is imperative to both decrease cost and improve patient outcomes via earlier detection. All the currently approved US Food and Drug Administration markers have limitations. More specifically, inadequate sensitivity and specificity compared with cystoscopy leads to false negative and false positive results, respectively [6–8]. MiRNAs have been suggested as promising alternative biomarkers for detecting cancer, informing prognosis, and monitoring treatment response [9,10]. From a biological standpoint, miRNAs may be superior predictive markers compared with protein, DNA, or mRNA. A single miRNA may regulate hundreds of target mRNAs that are frequently grouped in a specific biological pathway. Consequently, a focused miRNA signature may provide comparable prognostic information several orders of magnitude greater than mRNAs [11,12]. From a practical viewpoint, miRNAs are more stable than mRNAs or proteins and are less subject to degradation during sample processing. Thus, miRNAs are more suitable for analysis in formalin-fixed paraffin-embedded tissues, urine, serum, or plasma [12,13].

As a proof of concept study, this study investigates the potential of plasma miRNAs in detecting BC. To this end, a comparison of all known and predicted noncoding RNA species in the plasma of patients with and without BC were analyzed with the goal of identifying a miRNA fingerprint associated with different disease states. If plasma miRNA signatures accurately represent BC status, additional studies specifically examining potential clinical applications, such as initial hematuria workup or surveillance of known BC patients will be conducted.

2. Materials and methods

2.1. Patient selection and sample processing

The study design was approved by the University of Texas MD Anderson Cancer Center Institutional Review Board (protocol: LAB09-0149). Informed consent was obtained per the participating institution's protocol. Whole blood samples were prospectively collected from BC patients ($n = 20$) before surgery, and from control patients without urologic malignancies ($n = 18$). The blood samples from the BC and control patients were collected at MD Anderson Cancer Center, Houston, TX, and Municipal Hospital, Tisisoara, Romania, respectively. Non-muscle-invasive bladder cancer (NIMBC) was defined as Ta-T1, and MIBC was defined as T2-T4 per 2002 AJCC/UICC TNM system. Grade was designated per 1973 WHO Guidelines.

Whole blood (5–8 ml) was collected in an ethylene diamine tetracetic acid (EDTA) tube. The sample was centrifuged twice at 4°. Plasma (supernatant after second centrifugation) was then stored at -80°C . Samples were frozen within 4 hours of collection. Total RNA was isolated per a commercially available kit adhering to the manufacturer's protocol (mirVana; miRNA Isolation Kit, Applied Biosystems, Carlsbad, CA). RNA hybridization was performed with a custom-made noncoding RNA array (MD Anderson Cancer Center, Houston, TX) containing 9,600 miRNAs in duplicate, as previously described [14].

2.2. Reverse transcriptase polymerase chain reaction analysis for quality control assessment

Reverse transcriptase polymerase chain reaction was performed with the TaqMan miRNA Reverse Transcription Kit (Applied Biosystems, Carlsbad, CA) per manufacturer's instructions. Total RNA (0.2 μg) was used for cDNA amplification by using an arbitrarily primed multicolor detection system (Applied Biosystems). miR-150 and miR-192 expression levels were assessed to ensure that the isolated RNA material from cell-free plasma contained amplifiable miRNA (data not shown) on all BC and control samples [14]. Assays were performed in triplicate, and miRNA expression levels were calculated by using the comparative cycle threshold (Ct) method. The fold change was calculated by using the 2^{-Ct} method. U6 snRNA was used as the endogenous control.

2.3. Statistical analysis

All statistical analyses were performed in the statistics system R by using Bioconductor, additional public packages, and custom programming [15].

Since quantile normalization, a common approach for preprocessing microarray data, failed to achieve approximate constant variance over the range of expression levels, variance-stabilizing transformation was used instead [16]. To assess signal strength and reliability, the Pearson correlation coefficient (r) of the duplicate measurements was computed for each miRNA.

Clustering and principal-component analysis were used for exploratory (i.e., “unsupervised”) analysis [17,18]. T -tests and extensions thereof, specifically shrinkage t -

tests, followed by false discovery rate (FDR) control were used to identify differentially expressed miRNAs that correlated with the disease state (NMIBC or MIBC) [19,20]. To develop systems for predicting diagnosis, several machine learning methods were applied, including random forests of classification trees, nearest shrunken centroids, and regularized logistic regression (LR) [18,21]. For each of these methods, an appropriate technique was used to estimate the generalization performance of the obtained classifiers, namely bootstrapping and leave-one-out cross-validation (LOO-CV). In a post-processing step, the importance of each miRNA was extracted to the resultant classifiers [22].

To address the inherent difficulties in identifying differentially expressed miRNAs from thousands of candidates in relatively few patient samples, local FDRs, regularization (more specifically, penalization or shrinkage), and cross-validation (CV) were used to ensure meaningful results and minimize the potential for overfitting.

2.4. Pathway enrichment analysis

A Kyoto Encyclopedia of Genes and Genomes (KEGG)-based pathway enrichment analysis with use of Diana-miRPath software for the gene targets predicted by Diana microT Pic-Tar, and TargetScan was performed to elucidate potential pathways involved in differentially expressed miRNAs (<http://diana.cslab.ece.ntua.gr/>).

3. Results

3.1. Clinical and demographic information

Table 1 summarizes the clinical and demographic characteristics for the control patients. One control patient was treated for a basal cell carcinoma of the skin, 1 had a benign rectal polyp, and 1 had a uterine fibroid. None were diagnosed with malignant carcinomas. Table 2 summarizes the clinical and demographic characteristics for the NMIBC and MIBC groups.

3.2. Unsupervised analysis

Initial analysis focused on whether plasma miRNA expression levels can discriminate BC from controls. Principal component analysis (Fig. 1A) and clustering (Fig. 1B) were used for exploratory analysis. The resultant groupings of samples were subsequently reproduced and verified. The results of both approaches suggest that factors other than the presence or absence of BC influence plasma miRNA expression profiles. This was not unexpected as many coexisting conditions, such as metabolic, inflammatory or infectious processes could affect plasma miRNA expression. Further for BC patients, NMIBC and MIBC did not cluster together. To this end, supervised analysis was required.

3.3. Supervised analysis: Clustering of differentially expressed microRNAs

Ten differentially expressed miRNAs between controls and BC patients were identified with high confidence (moderated *t*-test, tail-based FDR <10%). Several of the identified miRNAs, such as miR-1290 (Fig. 2A) and miR-92b (Fig. 2B), had expression patterns that correlated with BC disease state. Patient samples were subsequently 'reclustered' based on these 10 selected miRNAs (Fig. 2C).

MIBC samples showed sufficient separation from noncancerous controls. NMIBC samples had a broader distribution, which overlapped with MIBC and/or controls. Given the known heterogeneity in the clinical course of NMIBC, the NMIBC group was dichotomized by pathologic grade into low-grade NMIBC and high-grade NMIBC. Most NMIBC with miRNA expression patterns initially coded as ‘noncancerous’ was subsequently identified as low-grade NMIBC, while almost all the NMIBC previously coded as ‘MIBC’ was subsequently identified as high-grade NMIBC (Fig. 2C).

3.4. Machine learning classification

Seventy-nine potentially discriminatory miRNAs between BC and controls were identified (local FDR <0.5). However, no individual miRNA sufficiently distinguished BC from control. Therefore, machine learning methods were used to identify classifiers consisting of multiple miRNAs (Table 3). To facilitate the process of training classifiers in a dataset consisting of many variables (9,600 candidate miRNAs) and relatively few observations (38 samples), several steps were taken, including using the correlation coefficient r between duplicate measurements (Fig. 3A). Further, miRNAs with r values <0.4 were excluded, and the expression data of the remaining miRNAs was weighted by multiplying them by r . This excluded ‘silent’ miRNAs since measurements predominately consisting of noise are expected to yield low correlation coefficients. In addition to the remaining non-‘silent’ miRNAs, a binary indicator variable encoding patient sex was also used as a feature, as the prevalence of BC is higher in men compared with women.

Bootstrapping and cross validation were used to obtain realistic estimates of predictive outcomes. For instance, regularized LR was trained on all but 1 patient, and a prediction was made for the left-out patient (LOO-CV). This method was cycled and repeated for all patients. The strength of the regularization was determined by maximizing the LOO negative log likelihood. An LR prediction was the estimated probability of the patient being in one class (e.g., BC), given the miRNA measurements. A hard prediction was naturally derived by applying a threshold of 0.5. For BC (both MIBC and NMIBC) vs. controls, this hard prediction yielded 90% sensitivity and 89% specificity (Fig. 3B, red circle). Changing the threshold can trade decreased sensitivity for increased specificity, or vice versa. For instance, applying a 0.8 threshold yielded 75% sensitivity at 100% specificity, hence preventing any false alarm in the LOO-CV (Fig. 3B, orange circle). For MIBC vs. others (NMIBC and controls), 80% sensitivity and 96% specificity were achieved (Fig. 3C). One hundred percent accuracy was achieved for identifying MIBC vs. controls (Fig. 3D).

3.5. Diagnostically useful miRNAs

Each miRNA’s relative contribution to the LR classifier was computed. The 40 most diagnostically useful miRNAs were determined. First, for distinguishing BC from control samples (Fig. 4A); second, for MIBC vs. other (NMIBC and controls) samples (Fig. 4B). Several miRNAs, such as miR-541, miR-200b, miR-566, miR-487, and miR-148b, were up-regulated in the plasma of BC patients, whereas the expression of other miRNAs, such as miR-25, miR-92a, -92 b, miR-302, and miR-33b, were significantly higher in control patients. These results suggest that a miRNA “fingerprint” may be associated with the presence of BC.

4. Discussion

This study suggests a diagnostic application for plasma miRNAs in BC. LR was the most accurate statistical method for predicting diagnosis with 89% accuracy for detecting the presence or absence of BC, 92% accuracy for distinguishing invasive BC from other cases, 79% accuracy for 3-way classification, and 100% accuracy for distinguishing MIBC from controls. The preliminary results of plasma miRNA favor comparably to routinely used urinary biomarkers, such as urine cytology, BTA, NMP-22, ImmunoCyt, and UroVysion [6–8]. In future studies, plasma miRNA will need to be evaluated in specific clinical scenarios, such as initial evaluation for hematuria, or for surveillance in known BC patients, and will also need to be directly compared with other noninvasive markers like urine cytology.

Although the results are promising, this study has several limitations. First, the sample sizes were small. The lack of larger sample sizes limited the statistical power to detect differences between study cohorts, and potentially resulted in an inability of plasma miRNA profiles to distinguish NMIBC from normal samples or MIBC. For example, FGFR3-related miR-99/100, previously shown to be downregulated in NMIBC tumors, showed higher expression in MIBC patients than in NMIBC patients, but these differences were not statistically significant [23]. Apart from inherent differential miRNA expression between NMIBC and MIBC, an alternative explanation could be that changes in tumor-derived plasma miRNAs may be proportional to tumor volume. The small number of patients also influenced our statistical analysis. An unsupervised approach would have been ideal to minimize overfitting and overly optimistic results; however, alternative statistical methods (local FDRs, regularization, cross-validation) were used in a supervised analysis to ensure reliable and accurate results.

The control group was not ideally matched for age (younger than BC group) or gender (more female than BC group). To address these confounding factors, miR-1290 and miR-92b were separately assessed based on gender (data not shown). Both miRNAs showed a consistent differential expression between control and BC patients regardless of gender (low miR-1290 and miR-92b in both male and female BC patients). Further, gender was included as a covariate so its potential effects on miRNA expression are already taken into account in the logistic regression modeling. Regarding the influence of age on miRNA expression, both miR-1290 and miR-92b trended toward higher expression levels with older age (data not shown). Since the BC group was older and had lower miRNA miR-1290 and miR-92b, it is unlikely that age alone could explain the differential miRNA expression patterns between BC patients and controls. A normal ‘healthy’ control population could also have been used in this study. Given the wide array of medical conditions in the control group, however, it is less likely that a systematic bias influencing miRNA expression exists. Geographic distribution between BC patients and controls (all controls were from Romania and all BC were from the United States) may also be a source of bias. In addition to including larger sample sizes, future studies will need stringent inclusion criteria to optimally match BC patients and controls to minimize the aforementioned potential sources of bias and ensure that differences in miRNA profiles are predominantly a reflection of BC disease state.

The most diagnostically useful miRNAs had expression levels generally lower in BC patients compared with non-cancerous controls. This finding adds additional complexity to the analysis as it confirms that other phenomena contribute to the circulating miRNA milieu. The expression of these particular miRNAs was inversely correlated with advanced stages of BC, which supports their utility as a marker of BC disease state.

It was quite intriguing that miR-33b and miR-92b were down-regulated in the plasma of patients with BC. miRNAs typically bind to their mRNA targets at the 3' untranslated regions (UTRs), and subsequently trigger mRNA degradation or inhibition of protein translation [14]. Pathway enrichment analysis (data not shown) revealed that many of the predicted miRNA targets were involved in critical pathways known to affect BC progression, including the tumor growth factor-beta signaling pathway [24]. Furthermore, analysis of the potential binding targets for miR-92 and miR-33 predicted 3 potential binding sites for miR-92b in the CD69 3'UTR, and 1 unique site for miR-33b in both the CD96 and CTLA-4 3'UTRs. Importantly, CD69 protein is expressed by activated T cells, including natural killer (NK) cells; CD96 is expressed by NK cells and is important for cell adhesion of NK cells to their target cells; and CTLA-4 is expressed primarily by activated T cells and dendritic cells [25–27]. Furthermore, miR-33 has recently been associated with macrophages, which are key constituents of tumor stroma known to both promote tumor progression and destruction depending on the their phenotype [28–30].

It is possible that certain adaptive or innate immunologic responses may be the plasma “sources” of some miRNAs, such as miR-92b and miR-33b, and release them into the systemic circulation as part of a homeostatic mechanism. In this scenario, inciting events, such as the onset of malignancy, may activate immune responses, including subsets of T cells associated with the down-regulation of miR-92b and miR-33b. This “immunosurveillance theory,” first proposed by Paul Ehrlich in the early 1900s and subsequently further developed in the 1970s, currently includes the concept of tumor immuno-editing, which is believed to continue during tumor development [31]. Both innate and adaptive immunity are believed to be involved in tumor biology, and they both can promote tumor progression as well as mediate tumor destruction [29,30].

5. Conclusions

To our knowledge, this study represents the first attempt at utilizing plasma miRNA for BC detection. Looking beyond the limitations of the current study design, plasma miRNA profiles seem to perform favorably compared with routinely used urine assays. This study provides preliminary data to support the feasibility and efficacy of this approach, and will serve as a foundation for further studies both to specify the optimal plasma miRNA signature and to apply these signatures to clinical scenarios, such as initial BC detection and BC surveillance.

Acknowledgments

The authors thank T. Nicola for kindly providing plasma samples from noncancerous patients, S. Ajibode for technical assistance, and C. Calin for helpful discussions on the manuscript. This work was supported by Molecular Health, GmbH, through a sponsored research agreement with the UT MD Anderson Cancer Center and by the

National Institutes of Health through MD Anderson's Cancer Center Support grant CA016672, NIH/NCI GU SPORE (P50 CA91846-04) grant.

References

1. Dinney CP, McConkey DJ, Millikan RE, et al. Focus on bladder cancer. *Cancer Cell*. 2004; 6:111–116. [PubMed: 15324694]
2. Wu XR. Urothelial tumorigenesis: A tale of divergent pathways. *Nat Rev Cancer*. 2005; 5:713–725. [PubMed: 16110317]
3. Botteman MF, Pashos CL, Redaelli A, et al. The health economics of bladder cancer: A comprehensive review of the published literature. *Pharmacoeconomics*. 2003; 21:1315–1330. [PubMed: 14750899]
4. Noyes K, Singer EA, Messing EM. Healthcare economics of bladder cancer: Cost-enhancing and cost-reducing factors. *Curr Opin Urol*. 2008; 18:533–539. [PubMed: 18670280]
5. Siegel R, Naishadham D, Jemal A. Cancer statistics. *CA Cancer J Clin*. 2012; 2012:10–29.
6. Konety BR. Molecular markers in bladder cancer: A critical appraisal. *Urol Oncol Semin Orig Investig*. 2006; 24:326–337.
7. van Rhijn BWG, van der Poel HG, van der Kwast TH. Cytology and urinary markers for the diagnosis of bladder cancer. *Eur Urol Supplements*. 2009; 8:536–541.
8. Tilki D, Burger M, Dalbagni G, et al. Urine markers for detection and surveillance of non-muscle-invasive bladder cancer. *Eur Urol*. 2011; 60:484–492. [PubMed: 21684071]
9. Cortez MA, Bueso-Ramos C, Ferdin J, et al. MicroRNAs in body fluids—the mix of hormones and biomarkers. *Nat Rev Clin Oncol*. 2011; 8:467–477. [PubMed: 21647195]
10. Hui A, How C, Ito E, et al. Micro-RNAs as diagnostic or prognostic markers in human epithelial malignancies. *BMC Cancer*. 2011; 2011:11. 500.
11. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer*. 2006; 6:857–866. [PubMed: 17060945]
12. Bartel DP. MicroRNAs: Target recognition and regulatory functions. *Cell*. 2009; 136:215–233. [PubMed: 19167326]
13. Cortez MA, Calin GA. MicroRNA identification in plasma and serum: A new tool to diagnose and monitor diseases. *Expert Opin Biol Ther*. 2009; 9:703–711. [PubMed: 19426115]
14. Liu CG, Calin GA, Volinia S, et al. MicroRNA expression profiling using microarrays. *Nat Protoc*. 2008; 3:563–578. [PubMed: 18388938]
15. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol*. 2004; 5:R80. [PubMed: 15461798]
16. Huber W, Von Heydebreck A, Sültmann H, et al. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*. 2002; 8:S96–S104.
17. Gehlenborg N, O'Donoghue SI, Baliga NS, et al. Visualization of omics data for systems biology. *Nat Methods*. 2010; 7:S56–S68. [PubMed: 20195258]
18. Duda, RO.; Hart, PE.; Stork, DG. Pattern classification. California: John Wiley and Sons; 2001.
19. Opgen R, Strimmer K. Accurate ranking of differentially expressed genes by a distribution-free shrinkage. *Stat Appl Genet Mol Biol*. 2007; 6 (Article 9).
20. Strimmer K. A unified approach to false discovery rate estimation. *BMC Bioinformatics*. 2008; 9:303. [PubMed: 18613966]
21. Zhu J, Hastie TK. Logistic regression and the import Vector machine. *J Comput Graph Statist*. 2005; 1:185–120.
22. Zien, A.; Kramer, N.; Ratsch, G., et al. The feature importance ranking measure, European Conference on Machine Learning and Principle and Practice of Knowledge Discovery in Databases; September 7–11, 2009; Bled, Slovenia. p. 694-709.
23. Catto JW, Miah S, Owen HC, et al. Distinct microRNA alterations characterize high- and low-grade bladder cancer. *Cancer Res*. 2009; 69:8472–8481. [PubMed: 19843843]

24. Castillejo A, Rothman N, Murta-Nascimento C, et al. TGFB1 and TGFBR1 polymorphic variants in relationship to bladder cancer risk and prognosis. *Int J Cancer*. 2009; 124:608–613. [PubMed: 19004027]
25. Sancho D, Gómez M, Sánchez-Madrid F. CD69 is an immunoregulatory molecule induced following activation. *Trends Immunol*. 2005; 26:136–140. [PubMed: 15745855]
26. Fuchs A, Colonna M. The role of NK cell recognition of nectin and nectin-like proteins in tumor immunosurveillance. *Semin Cancer Biol*. 2006; 16:359–366. [PubMed: 16904340]
27. Laurent S, Carrega P, Saverino D, et al. CTLA-4 is expressed by human monocyte-derived dendritic cells and regulates their functions. *Hum Immunol*. 2010; 71:934–941. [PubMed: 20650297]
28. Horie T, Ono K, Horiguchi M, et al. MicroRNA-33 encoded by an intron of sterol regulatory element-binding protein 2 (Srebp2) regulates HDL in vivo. *Proc Natl Acad Sci USA*. 2010; 107:17321–17326. [PubMed: 20855588]
29. Ostrand-Rosenberg S. Immune surveillance: A balance between pro-tumor and anti-tumor immunity. *Curr Opin Genet Dev*. 2008; 18:11–18. [PubMed: 18308558]
30. Bunt SK, Yang L, Sinha P, et al. Reduced inflammation in the tumor microenvironment delays the accumulation of myeloid-derived suppressor cells and limits tumor progression. *Cancer Res*. 2007; 67:10019–10026. [PubMed: 17942936]
31. Burnet FM. The concept of immunological surveillance. *Prog Exp Tumor Res*. 1970; 13:1–27. [PubMed: 4921480]

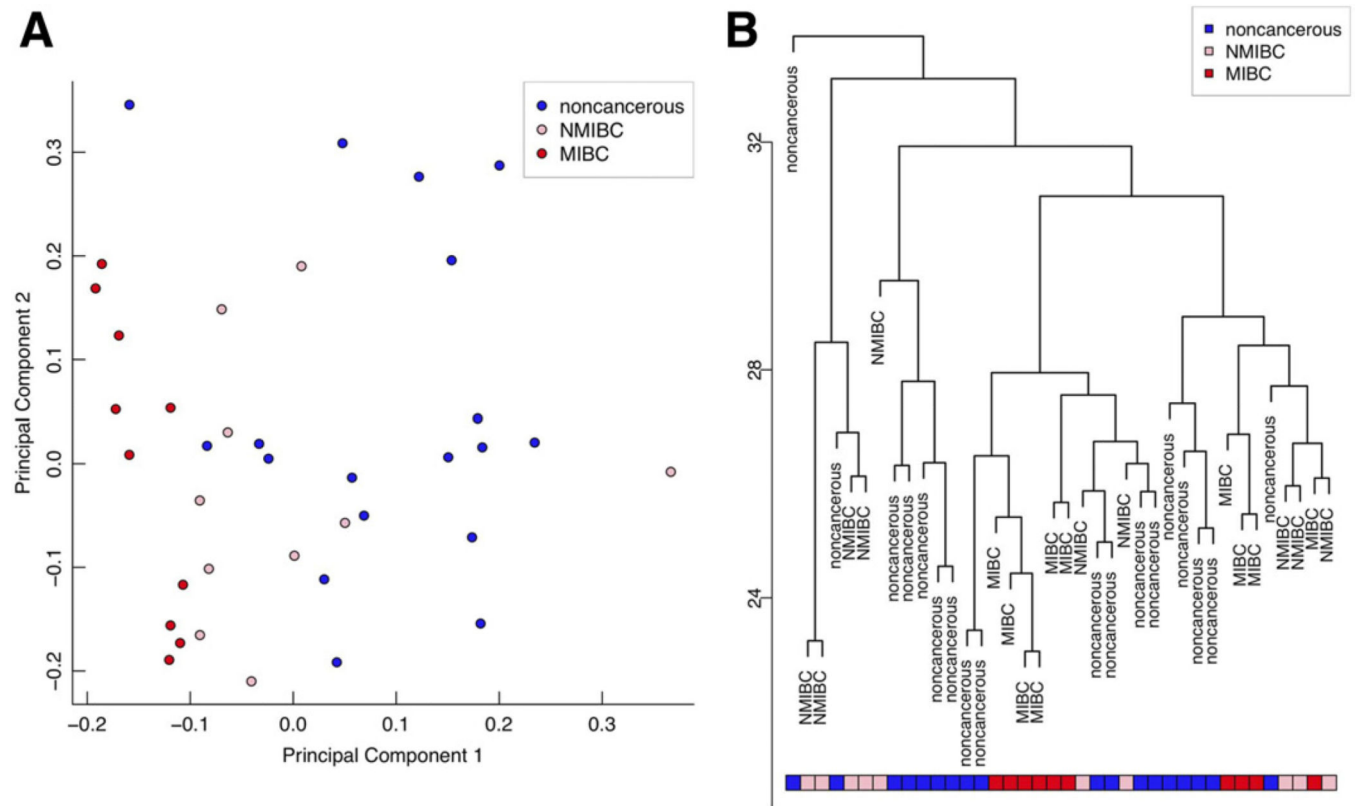


Fig. 1.

Principal component analysis and hierarchical clustering of samples based on expression levels of all 9,600 assayed sequences. The expression profiles of human miRNAs represented by (A) the first two principal components or by (B) unsupervised clustering cannot clearly distinguish BC from control samples.

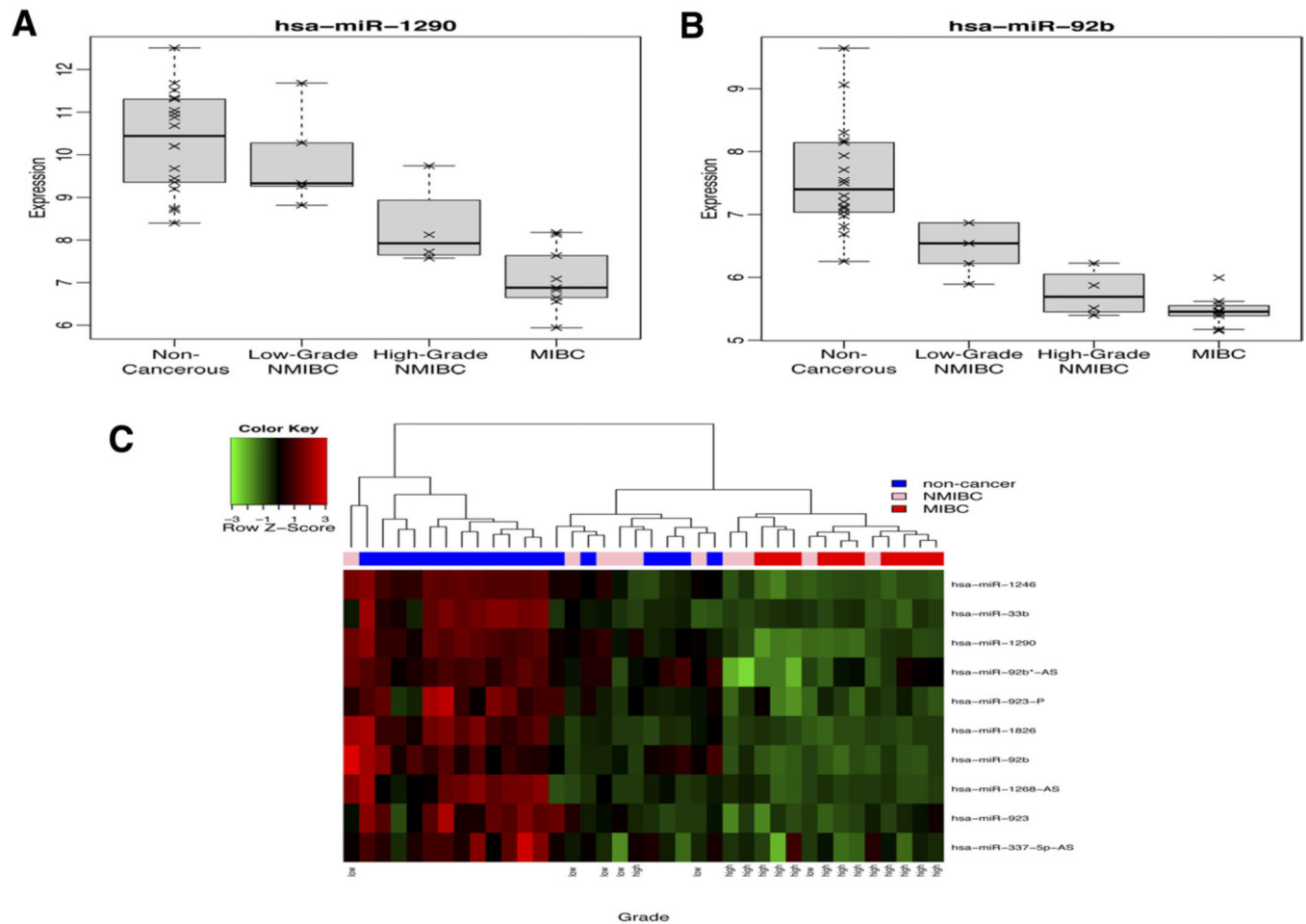
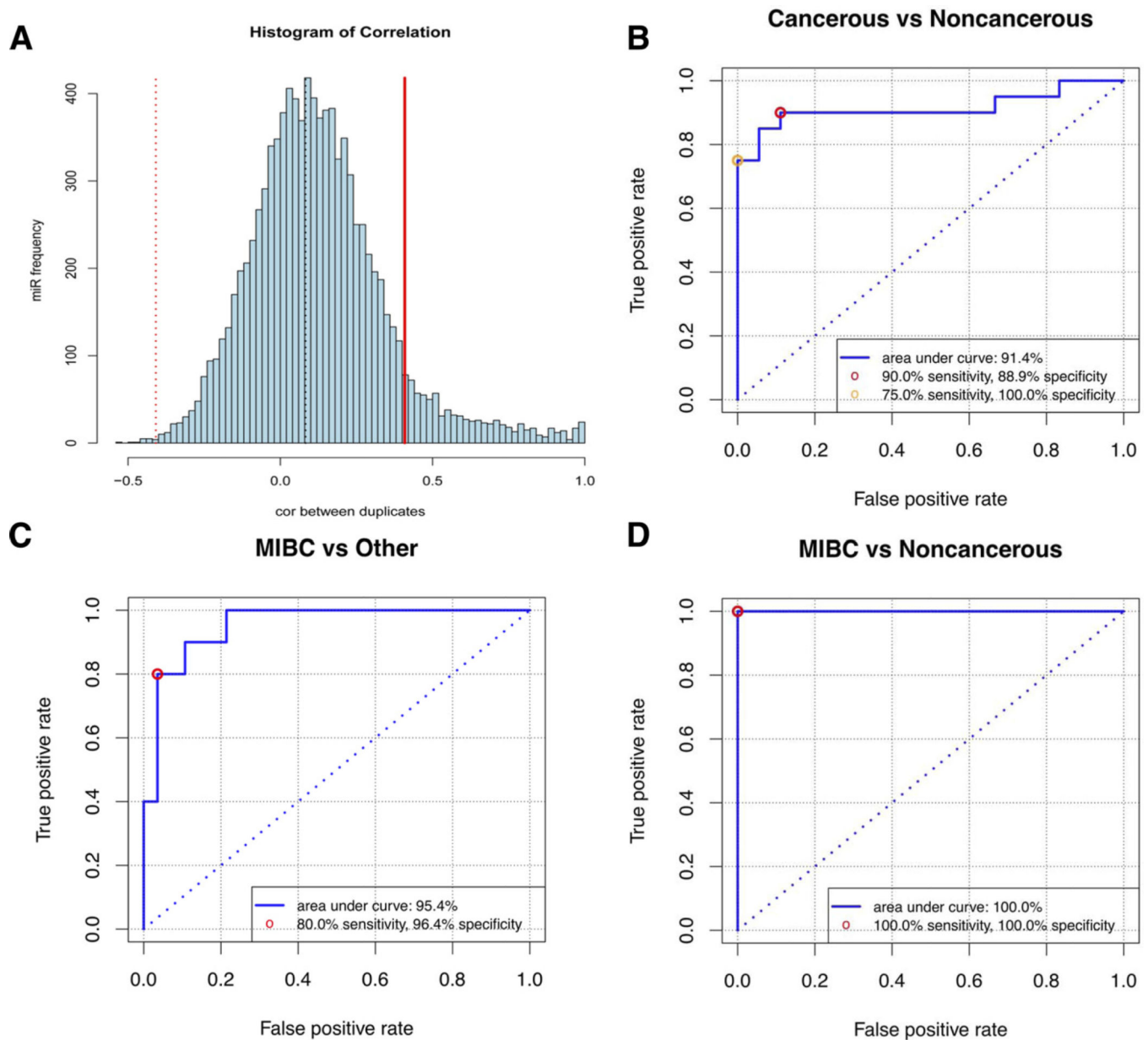


Fig. 2. Differentially expressed miRNAs correlate with various disease states. (A) miR-1290 and (B) miR-92b expression correlate with pathologic grade and stage; (C) clustering of samples using discriminative miRNAs reflects histologic grade and disease state.

**Fig. 3.**

LR analysis results. (A) Correlation plot between each duplicate miRNA value. (B) LOO-ROC curve for LR classifier for cancerous (MIBC and NMIBC) vs. controls or (C) MIBC vs. other (NMIBC and controls). The red circle corresponds to the natural probability threshold of 0.5. (D) LOO-ROC curve for LR classifier for MIBC vs. controls. Dotted line represents random prediction.

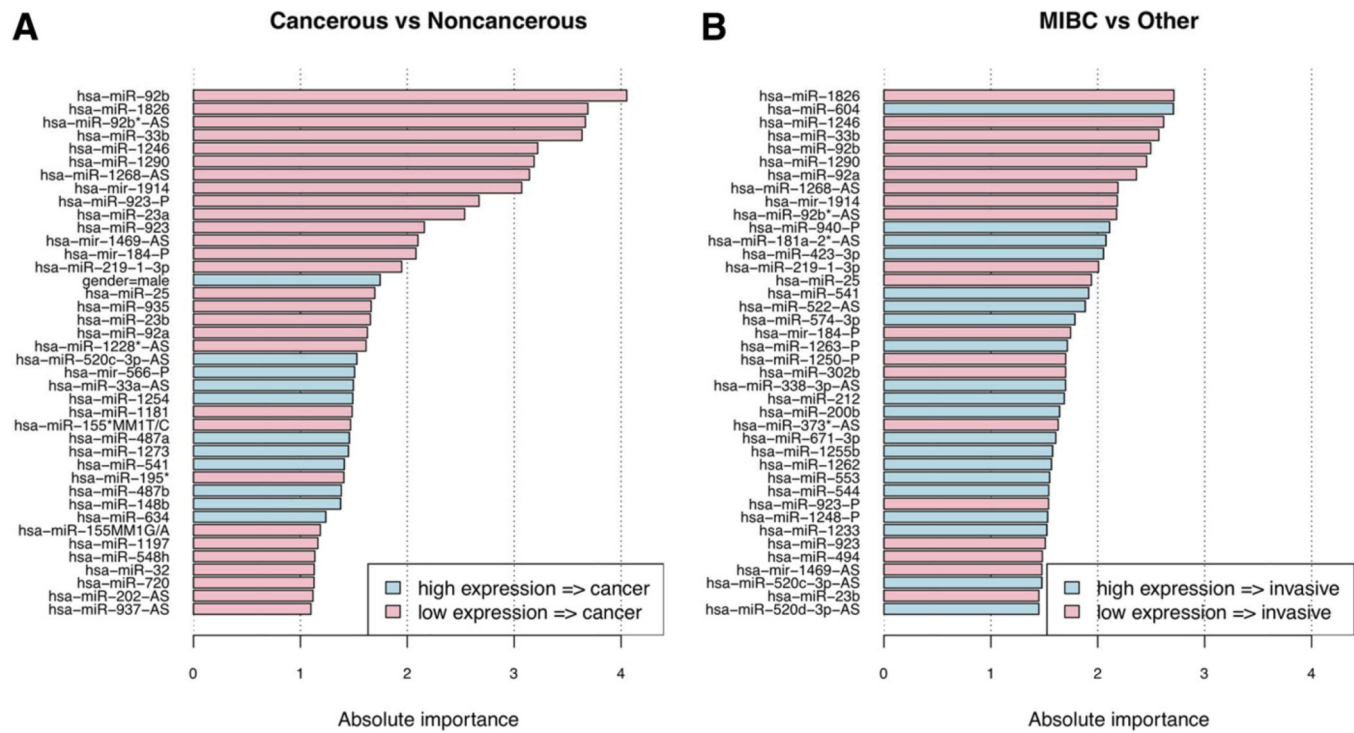


Fig. 4. The 40 most important features as determined from trained classifiers. (A) Cancerous (MIBC and NMIBC) vs. controls; (B) MIBC vs. other (NMIBC and controls).

Table 1

Clinical and demographic characteristics for controls (n = 18)

Median age (range) (years)	42 (18–84)
Sex, no. (%)	
Male	7 (39)
Female	11 (61)
Diagnosis, no. (%)	
Gastrointestinal	9 (50)
Genitourinary	4 (22)
Endocrine	2 (11)
Other	3 (17)

Table 2

Clinical and demographic characteristics of bladder cancer patients

	Non-Muscle invasive (n = 10)	Muscle invasive (n = 10)
Median age (range) (years)	68 (57–81)	62 (52–73)
Sex, no. (%)		
Male	8 (80)	7 (70)
Female	2 (20)	3 (30)
Stage, no. (%)		
Ta	7 (70)	
T1	3 (30)	
T2		7 (70)
T3		1 (10)
T4		2 (20)
Grade, no. (%)		
1	1 (10)	0 (0)
2	5 (50)	0 (0)
3	4 (40)	10 (100)
Concomitant CIS		
Yes	0 (0)	6 (60)
No	10 (100)	4 (40)
Histology		
TCC	10 (100)	6 (60)
TCC with aberrant differentiation	0 (0)	3 (30)
Small cell	0 (0)	1 (10)
Intravesicle chemotherapy		
Yes	7 (70)	4 (40)
No	3 (30)	6 (60)
Systemic chemotherapy		
Yes	0 (0)	5 (50)
No	10 (100)	5 (50)

Table 3

Cross-validation accuracy for logistic regression

	No. of cases	No. of CV errors *	Accuracy *	Sensitivity *	Specificity *	auROC
Cancerous vs other	20 vs 18	4	89%	90%	89%	91%
Invasive vs other	10 vs 28	3	92%	80%	96%	95%
Invasive vs non-cancerous	10 vs 18	0	100%	100%	100%	100%

auROC = area under the receiver operating characteristics curve.

* Binary predictions obtained by a 0.5 significance threshold.