

## RESEARCH ARTICLE

# MicroRNA Expression Profile Analysis Reveals Diagnostic Biomarker for Human Prostate Cancer

Dong-Fu Liu<sup>&</sup>, Ji-Tao Wu<sup>&</sup>, Jian-Ming Wang, Qing-Zuo Liu, Zhen-Li Gao, Yun-Xiang Liu<sup>\*</sup>

### Abstract

Prostate cancer is a highly prevalent disease in older men of the western world. MicroRNAs (miRNAs) are small RNA molecules that regulate gene expression via posttranscriptional inhibition of protein synthesis. To identify the diagnostic potential of miRNAs in prostate cancer, we downloaded the miRNA expression profile of prostate cancer from the GEO database and analysed the differentially expressed miRNAs (DE-miRNAs) in prostate cancerous tissue compared to non-cancerous tissue. Then, the targets of these DE-miRNAs were extracted from the database and mapped to the STRING and KEGG databases for network construction and pathway enrichment analysis. We identified a total of 16 miRNAs that showed a significant differential expression in cancer samples. A total of 9 target genes corresponding to 3 DE-miRNAs were obtained. After network and pathway enrichment analysis, we finally demonstrated that miR-20 appears to play an important role in the regulation of prostate cancer onset. MiR-20 as single biomarker or in combination could be useful in the diagnosis of prostate cancer. We anticipate our study could provide the groundwork for further experiments.

**Keywords:** Prostate cancer - microRNA - network analysis - pathway enrichment

*Asian Pacific J Cancer Prev*, 13, 3313-3317

### Introduction

Prostate cancer is a highly prevalent disease in older men of the western world (Chan et al., 2004; Linton and Hamdy, 2004). It is estimated that 241,740 men will be diagnosed with and 28,170 men will die of prostate cancer in 2012 in the United States (Siegel et al., 2012). Although the age-adjusted rate of cancer deaths has decreased steadily in the past 10 years, prostate cancer remains the second leading cause of cancer deaths in men after lung cancer (Shen and Abate-Shen, 2010). The morbidity and mortality of prostate cancer is principal caused of its propensity to metastasize to other tissue, such as lung, liver and bone (Bubendorf et al., 2000; Logothetis and Lin, 2005).

MicroRNAs (miRNAs) were discovered in 1993 by Victor Ambros et al. during a study of the gene *lin-14* in *C. elegans* development (Lee et al., 1993). They are short, non-coding RNAs with an average of 22 nucleotides in length that usually bind to partially complementary sites in the 3'-untranslated region (UTR) of their mRNA targets (Lee et al., 1993; Wightman et al., 1993). They regulate gene expression by mRNA cleavage and at posttranslational level by translational suppression and play important roles in various biological and metabolic processes (Bartel, 2004; Min and Yoon, 2010). Associations of miRNA expression and tumorigenesis have been observed in a variety of human malignancies

including colon (Cummins et al., 2006), pancreas (Lee et al., 2007), liver (Murakami et al., 2006) and breast (Iorio et al., 2005; Sempere et al., 2007). It is estimated that about 1000 miRNAs exist in vertebrate genome (Berezikov et al., 2005). A total of 940 mature miRNAs have been identified in human genome according to the miRBase (Kozomara and Griffiths-Jones, 2011).

Several miRNAs expression profiles have been reported regarding Prostate cancer (Lu et al., 2005; Mattie et al., 2006; Porkka et al., 2007; Ozen et al., 2008). However, there is still an urgent need to identify new miRNAs as diagnostic biomarkers for prostate cancer. In this study, we collected miRNA expression microarray dataset of prostate cancer from GEO database and identified differentially expressed miRNAs in cancerous tissue compared with normal tissue. After network analysis and pathway enrichment analysis, we finally demonstrated that miR-20 may play an important role in the regulation of prostate cancer onset.

### Materials and Methods

#### *Affymetrix microarray data*

We extracted the miRNA expression profile from the study of Wach et al. (2012), which were deposited in GEO (Gene Expression Omnibus) database (ID: GSE23022). The study was carried out to identify and characterize the diagnostic potential of miRNAs in prostate cancer. A

Department of Urology, Yantai Yuhuangding Hospital, Yantai, China <sup>&</sup>Equal contributors <sup>\*</sup>For correspondence: [yunxiangliu0718@hotmail.com](mailto:yunxiangliu0718@hotmail.com)

total of 40 chips were available, including 20 chips from moderately differentiated prostate cancer and 20 chips from adjacent noncancerous tissue. These tissues were prepared from prostatectomy specimens from men with so far untreated prostate cancer between 1994 and 1999. None of patients had detectable distant metastases at the time of surgery.

#### Pathway data

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals (Kanehisa, 2002). The PATHWAY database records networks of molecular interactions in the cells, and variants of them specific to particular organisms (<http://www.genome.jp/kegg/>). Total 130 pathways, involving 2287 genes, were collected from KEGG.

#### Data preprocessing

The probe-level data in CEL files were converted into expression measures and performed background correction and quartile data normalization by the robust multiarray average (RMA) (Irizarry et al., 2003) algorithm with defaulted parameters in R affy package (Gautier et al., 2004; Team, 2011).

#### Differentially expressed miRNA (DE-miRNA) analysis

The t-test and wilcox test were used to identify miRNAs that were significantly differentially expressed between prostate cancer tissue and noncancerous tissue, respectively. Then, we selected the overlapping DE-miRNAs of these two methods as the final result. The p-value adjusted by the Benjamin and Hochberg (BH) (Benjamini, 1995) method based on the multtest package (van der Laan et al., 2004) of 0.05 was used as the cut-off criterion.

#### Obtaining target genes for DE-miRNAs

TargetSan (Garcia et al., 2011), miRanda (John et al., 2004) and PicTar (Krek et al., 2005) databases were used to retrieve the target genes of DE-miRNAs. They use both miRNA sequences and 3'UTR of protein-coding mRNA sequences as input files generally in fasta format and determine their binding ability by calculating the minimum free energy for hybridization. After the prediction of putative miRNA target genes by each tool with the default parameters, we extracted those target genes shared by at least 2 of these 3 tools to obtain a more solid result.

#### Network analysis and pathway enrichment analysis

The STRING (Search Tool for the Retrieval of Interacting Genes) (Szklarczyk et al., 2011) database provides both experimental and predicted interaction information. Version 9.0 of STRING covers more than 1100 completely sequenced organisms. All associations are provided with a probabilistic confidence score, which is derived by separately benchmarking groups of associations against the manually curated functional classification scheme of the KEGG database. Each score represents a rough estimate of how likely a given

association describes a functional linkage between two proteins that is at least as specific as that between an average pair of proteins annotated on the same 'map' or 'pathway' in KEGG. We used the STRING database to annotate functional interactions between DE-miRNA target genes and other genes by calculated their confidence score.

To functionally classify these genes in the interaction network, we performed pathway enrichment analysis by mapping these genes to KEGG database (Kanehisa, 2002). The count number larger than 2 and FDR less than 0.01 were chosen as cut-off criterion.

## Results

### Differentially expressed miRNA analysis between prostate cancer and healthy control

We obtained publicly available microarray dataset GSE23022 from GEO database. The t-test and wilcox test were used to identify the miRNAs specifically differentially expressed between prostate cancer tissue and noncancerous controls, with multiple testing correction. At an adjusted p-value of 0.05, 16 miRNAs showed a significant differential expression (Table 1).

### Target genes of DE-miRNAs Obtainment

Since miRNAs play important roles of posttranscriptional regression by targeting mRNAs, we studied the function of the DE-miRNAs by identifying

**Table 1. MiRNAs Differentially Expressed in Cancer Cells as Compared to Normal Tissue**

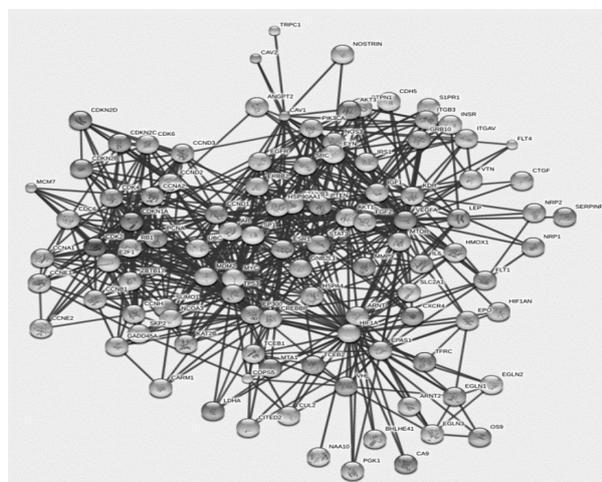
miRNA	adjusted p-value (t-test, tumor vs. normal)	adjusted p-value (wilcox, tumor vs. normal)
hsa-miR-101	0.022071647	0.007359301
hsa-miR-106a	0.022071647	0.008866526
hsa-miR-141	0.044533961	0.038343985
hsa-miR-145	0.022071647	0.007359301
hsa-miR-17	0.014657577	0.007359301
hsa-miR-182	0.022071647	0.00614075
hsa-miR-20a	0.0099099	0.009384458
hsa-miR-20b	0.0099099	0.007359301
hsa-miR-214	0.017938825	0.007359301
hsa-miR-221	0.0095711	0.002469005
hsa-miR-222	0.00078771	0.000783475
hsa-miR-320a	0.047370804	0.028271251
hsa-miR-375	0.0097405	0.007359301
hsa-miR-720	0.022071647	0.019736888
hsa-miR-768-3p	0.022993146	0.019736888
hsa-miR-93	0.0097405	0.007359301

**Table 2. The Predicted Target Genes of DE-miRNAs**

miRNA	Target Gene
hsa-miR-17	NCOA3
hsa-miR-17	VEGFA
hsa-miR-222	CDKN1B
hsa-miR-222	ESR1
hsa-miR-20	VEGFA
hsa-miR-20	CDKN1A
hsa-miR-20	NCOA3
hsa-miR-20	HIF1A
hsa-miR-20	CAV1

**Table 3. The Enriched KEGG Pathways**

Term	Description	Count	FDR
hsa05200	Pathways in cancer	46	3.21E-28
hsa04110	Cell cycle	29	8.78E-22
hsa05215	Prostate cancer	22	4.47E-16
hsa05211	Renal cell carcinoma	18	8.55E-13
hsa05222	Small cell lung cancer	17	4.33E-10
hsa05214	Glioma	15	1.54E-09
hsa04115	p53 signaling pathway	15	4.74E-09
hsa05218	Melanoma	15	8.88E-09
hsa04510	Focal adhesion	22	1.29E-08
hsa05219	Bladder cancer	12	6.67E-08
hsa05212	Pancreatic cancer	13	2.43E-06
hsa05223	Non-small cell lung cancer	11	2.02E-05
hsa05220	Chronic myeloid leukemia	12	4.80E-05
hsa05213	Endometrial cancer	10	2.02E-04
hsa04630	Jak-STAT signaling pathway	13	0.01276

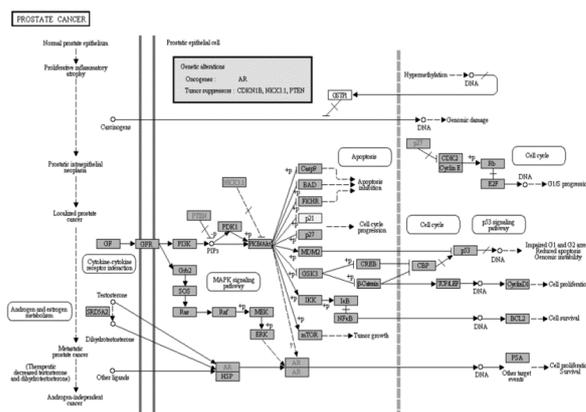


**Figure 1. Interaction Network Constructed Between DE-miRNA and Their Interactive Genes**

putative target genes. Target genes were retrieved from TargetScan, miRanda and PicTar databases. Finally, we obtained 9 target genes corresponding to 3 miRNAs (Table 2). Among the 9 target genes, 4 (bolded ones) were validated previously by experiments and the rest could be putative ones.

*Interaction network construction of DE-miRNA target genes*

We mapped the DE-miRNA target genes to STRING database and screened significant interactions with score larger than 0.9. By integrating these relationships above, we constructed an interaction network between DE-miRNAs and their interactive genes (Figure 1). Only 3 of the above 9 target genes existed in this network: HIF1A (hypoxia inducible factor 1 A), VEGFA (vascular endothelial growth factor A) and CDKN1A (cyclin dependent kinase inhibitor 1A). The average degrees of these three genes were 44, 43 and 33. The average degree is the average number of edges connecting all the nodes in the network. Higher values for average degree indicates a better connected network and is likely more robust. This result suggests that these three genes were hub nodes in the network and play critical roles in the prostate cancer. Pathway enrichment analysis of genes in the interaction network



**Figure 2. Pathway Map of Prostate Cancer.** The gray node p21 is DE-miRNA target gene

We performed pathway enrichment analysis by mapping genes in the interaction network to KEGG database. A total of 16 pathways were enriched with the strict cutoff criterion, including pathways in cancer, cell cycle and prostate cancer (Table 3). Figure 2 shows the KEGG map of prostate cancer. We could find that the DE-miRNA target gene of CDKN1A (cyclin-dependent kinase inhibitor 1A, p21) is an oncogene of prostate cancer. As CDKN1A is the target gene of hsa-miR-20, we concluded that miRNA 20 might play an important role in the regulation of prostate cancer onset.

**Discussion**

In this study, we analysed the expression profile of 678 human miRNAs in 20 matched tissue samples of histologically confirmed prostate cancer tissue and adjacent nonmalignant tissue downloaded from GEO database. A total of 16 miRNAs displayed a significant differential expression in cancerous tissue compared to noncancerous tissue. Of these DE-miRNAs, we identified that miRNA-20 may play an important role in the regulation of prostate cancer onset.

To date, differential expressions of miRNAs in prostate cancer have been published by several studies, however, the results of these studies are not the same. A total 16 differentially expressed miRNAs were identified in our study and 15 of them (miR-101, miR-106a, miR-141, hsa-miR-145, miR-17, miR-182, miR-20a, miR-20b, miR-214, miR-221, miR-222, miR-375, miR-720, miR-768-3p and miR-93) are in concordance with at least one other publication (Porkka et al., 2007; Ambs et al., 2008; Ozen et al., 2008; Tong et al., 2009; Schaefer et al., 2010; Wach et al., 2012).

Two previous studies using these same expression data have been published. The first was the initial description of the population and the array data (Wach et al., 2012), in which the authors identified a total of 25 miRNAs whose expression differed between the groups and indicated that MiRNAs as single biomarkers or in combination could be useful in the diagnosis of prostate cancer. The second study reanalyzed the original data (Schaefer et al., 2010), and identified a miRNA for normalization in miRNA expression studies of prostate cancer. Our current study reanalyzes the original data and adds interaction network

analysis which provides information on the relationships between DE-miRNA target genes and other genes. With the strict cut-off score of 0.9, 3 target genes (HIF1A, VEGFA and CDKN1A) were identified in the network and were shown hub nodes of the network. The importance of a gene is often dependent on how well it associates with other genes in a network. Studies suggest that more centralized genes in the network are more likely to be key drivers to proper cellular function than peripheral genes (nodes) (Horvath et al., 2006).

HIF1A is a key transcription factor that has been implicated in promoting tumor cell survival, proliferation and invasion following the onset of tumor hypoxia (Semenza, 2003). Increased expression of HIF1A in Prostate cancer cells has been correlated with faster tumor growth and higher metastatic potential (Hao et al., 2004; Kimbro and Simons, 2006). HIF1A expression has also been observed to increase as prostate tumors progressed from androgen-dependent to androgen-independent states (Zhong et al., 1998). VEGFA is a member of the VEGF growth factor family which promotes endothelial cell proliferation, survival and migration via binding to 2 specific tyrosine kinase receptors (Ferrara et al., 2003). Up-regulation of VEGF has been associated with significantly increased risk of prostate cancer in two small case control studies (McCarron et al., 2002; Sfar et al., 2006). CDKN1A functions as a regulator of cell cycle G1 phase arrest in response to a variety of stress stimuli. CDKN1A has a great impact on the cell cycle of prostate cancer cells and may play a role in the cancer cells in a p53-independent pathway (Wang et al., 2005).

By mapping the genes in the interaction network to KEGG pathways, we concluded that miRNA-20 whose target gene is CDKN1A might play an important role in prostate cancer onset. The human miR-20 located on chromosome 13q31, undergoes loss of heterozygosity in several different cancers, including prostate cancer. It increased apoptosis in A549 lung cancer cells and promoted osteogenic differentiation of human mesenchymal stem cells by co-regulating BMP signaling (Zhang et al., 2011). MiR-20 regulates cell growth via suppression of E2F1 expression (O'Donnell et al., 2005) and appears to be over-expressed in colon, pancreas and prostate tumors while being down-regulated in breast cancer tumors (Volinia et al., 2006).

In conclusion, we have used the miRNA expression profile downloaded from GEO database to identify miRNAs that differentially expressed in prostate cancerous tissue compared to noncancerous tissue. Our analysis identified several DE-miRNA target genes might play crucial roles in prostate cancer onset, including HIF1A, VEGFA and CDKN1A. Further, we demonstrated that miR-20 might play an important role in the regulation of prostate cancer onset. MiR-20 as single biomarker or in combination could be useful in the diagnosis of prostate cancer. We anticipate our study could provide groundwork for further experiments.

## Acknowledgements

This research was supported by grants from Department

of Science and Technology of Shandong Province (Grant NO. BS2011YY063), and Department of Health of Shandong Province (Grant NO. 2011QZ030).

## References

- Ambs S, Prueitt RL, Yi M, et al (2008). Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Res*, **68**, 6162-70.
- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-97.
- Benjamini YH, Y (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the royal statistical society Series B (Methodological)*, **57**, 289-300.
- Berezikov E, Guryev V, van de Belt J, et al (2005). Phylogenetic shadowing and computational identification of human microRNA genes. *Cell*, **120**, 21-4.
- Bubendorf L, Schopfer A, Wagner U, et al (2000). Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Hum Pathol*, **31**, 578-83.
- Chan JM, Jou RM, Carroll PR (2004). The relative impact and future burden of prostate cancer in the United States. *J Urol*, **172**, S13-6; discussion S7.
- Cummins JM, He Y, Leary RJ, et al (2006). The colorectal microRNAome. *Proc Natl Acad Sci U S A*, **103**, 3687-92.
- Ferrara N, Gerber HP, LeCouter J (2003). The biology of VEGF and its receptors. *Nat Med*, **9**, 669-76.
- Garcia DM, Baek D, Shin C, et al (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsi-6 and other microRNAs. *Nat Struct Mol Biol*, **18**, 1139-46.
- Gautier L, Cope L, Bolstad BM, Irizarry RA (2004). affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, **20**, 307-15.
- Hao P, Chen X, Geng H, et al (2004). Expression and implication of hypoxia inducible factor-1alpha in prostate neoplasm. *J Huazhong Univ Sci Technolog Med Sci*, **24**, 593-5.
- Horvath S, Zhang B, Carlson M, et al (2006). Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci USA*, **103**, 17402-7.
- Iorio MV, Ferracin M, Liu CG, et al (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Res*, **65**, 7065-70.
- Irizarry RA, Hobbs B, Collin F, et al (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249-64.
- John B, Enright AJ, Aravin A, et al (2004). Human MicroRNA targets. *PLoS Biol*, **2**, e363.
- Kanehisa M (2002). The KEGG database. *Novartis Found Symp*, **247**, 91-101; discussion -3, 19-28, 244-52.
- Kimbro KS, Simons JW (2006). Hypoxia-inducible factor-1 in human breast and prostate cancer. *Endocr Relat Cancer*, **13**, 739-49.
- Kozomara A, Griffiths-Jones S (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res*, **39**, D152-7.
- Krek A, Grun D, Poy MN, et al (2005). Combinatorial microRNA target predictions. *Nat Genet*, **37**, 495-500.
- Lee EJ, Gusev Y, Jiang J, et al (2007). Expression profiling identifies microRNA signature in pancreatic cancer. *Int J Cancer*, **120**, 1046-54.
- Lee RC, Feinbaum RL, Ambros V (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75**, 843-54.

- Linton DK, Hamdy FC (2004). Early diagnosis and surgical management of prostate cancer. *Ann Urol (Paris)*, **38**, 137-47.
- Logothetis CJ, Lin SH (2005). Osteoblasts in prostate cancer metastasis to bone. *Nat Rev Cancer*, **5**, 21-8.
- Lu J, Getz G, Miska EA, et al (2005). MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834-8.
- Mattie MD, Benz CC, Bowers J, et al (2006). Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol Cancer*, **5**, 24.
- McCarron SL, Edwards S, Evans PR, et al (2002). Influence of cytokine gene polymorphisms on the development of prostate cancer. *Cancer Res*, **62**, 3369-72.
- Min H, Yoon S (2010). Got target? Computational methods for microRNA target prediction and their extension. *Exp Mol Med*, **42**, 233-44.
- Murakami Y, Yasuda T, Saigo K, et al (2006). Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene*, **25**, 2537-45.
- O'Donnell KA, Wentzel EA, Zeller KI, et al (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, **435**, 839-43.
- Ozen M, Creighton CJ, Ozdemir M, Ittmann M (2008). Widespread deregulation of microRNA expression in human prostate cancer. *Oncogene*, **27**, 1788-93.
- Porkka KP, Pfeiffer MJ, Waltering KK, et al (2007). MicroRNA expression profiling in prostate cancer. *Cancer Res*, **67**, 6130-5.
- Schaefer A, Jung M, Miller K, et al (2010). Suitable reference genes for relative quantification of miRNA expression in prostate cancer. *Exp Mol Med*, **42**, 749-58.
- Schaefer A, Jung M, Mollenkopf HJ, et al (2010). Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma. *Int J Cancer*, **126**, 1166-76.
- Semenza GL (2003). Targeting HIF-1 for cancer therapy. *Nat Rev Cancer*, **3**, 721-32.
- Sempere LF, Christensen M, Silahtaroglu A, et al (2007). Altered MicroRNA expression confined to specific epithelial cell subpopulations in breast cancer. *Cancer Res*, **67**, 11612-20.
- Sfar S, Hassen E, Saad H, et al (2006). Association of VEGF genetic polymorphisms with prostate carcinoma risk and clinical outcome. *Cytokine*, **35**, 21-8.
- Shen MM, Abate-Shen C (2010). Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev*, **24**, 1967-2000.
- Siegel R, Naishadham D, Jemal A (2012). Cancer statistics, 2012. *CA Cancer J Clin*, **62**, 10-29.
- Szklarczyk D, Franceschini A, Kuhn M, et al (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*, **39**, D561-8.
- Team RDC (2011). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.
- Tong AW, Fulgham P, Jay C, et al (2009). MicroRNA profile analysis of human prostate cancers. *Cancer Gene Ther*, **16**, 206-16.
- van der Laan MJ, Dudoit S, Pollard KS (2004). Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Stat Appl Genet Mol Biol*, **3**, Article15.
- Volinia S, Calin GA, Liu CG, et al (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA*, **103**, 2257-61.
- Wach S, Nolte E, Szczyrba J, et al (2012). MicroRNA profiles of prostate carcinoma detected by multiplatform microRNA screening. *Int J Cancer*, **130**, 611-21.
- Wang Y, Shao C, Shi CH, et al (2005). Change of the cell cycle after flutamide treatment in prostate cancer cells and its molecular mechanism. *Asian J Androl*, **7**, 375-80.
- Wightman B, Ha I, Ruvkun G (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell*, **75**, 855-62.
- Zhang JF, Fu WM, He ML, et al (2011). MiRNA-20a promotes osteogenic differentiation of human mesenchymal stem cells by co-regulating BMP signaling. *RNA Biol*, **8**.
- Zhong H, Agani F, Baccala AA, et al (1998). Increased expression of hypoxia inducible factor-1alpha in rat and human prostate cancer. *Cancer Res*, **58**, 5280-4.