

PRIMARY RESEARCH

Open Access

# A miRNA expression signature that separates between normal and malignant prostate tissues

Jessica Carlsson<sup>1,2,5</sup>, Sabina Davidsson<sup>4,5</sup>, Gisela Helenius<sup>3</sup>, Mats Karlsson<sup>3</sup>, Zelmina Lubovac<sup>2</sup>, Ove André<sup>4</sup>, Björn Olsson<sup>2</sup> and Karin Klinga-Levan<sup>1\*</sup>

## Abstract

**Background:** MicroRNAs (miRNAs) constitute a class of small non-coding RNAs that post-transcriptionally regulate genes involved in several key biological processes and thus are involved in various diseases, including cancer. In this study we aimed to identify a miRNA expression signature that could be used to separate between normal and malignant prostate tissues.

**Results:** Nine miRNAs were found to be differentially expressed ( $p < 0.00001$ ). With the exception of two samples, this expression signature could be used to separate between the normal and malignant tissues. A cross-validation procedure confirmed the generality of this expression signature. We also identified 16 miRNAs that possibly could be used as a complement to current methods for grading of prostate tumor tissues.

**Conclusions:** We found an expression signature based on nine differentially expressed miRNAs that with high accuracy (85%) could classify the normal and malignant prostate tissues in patients from the Swedish Watchful Waiting cohort. The results show that there are significant differences in miRNA expression between normal and malignant prostate tissue, indicating that these small RNA molecules might be important in the biogenesis of prostate cancer and potentially useful for clinical diagnosis of the disease.

## Background

Prostate cancer is the most common type of cancer in men and accounted for 36% of all male cancer cases in Sweden during 2009. This type of cancer is a heterogeneous disease where some men have an aggressive lapse, while others have a slower development [1]. During the last years, prostate specific antigen (PSA) has been used as a biological marker for this disease. However, since the false positive rate for PSA values is very high, there is an urgent need for new and improved markers [2].

The first microRNA (miRNA) was discovered in 1993 by Ambros and colleagues while they were performing a genetic screen in *Caenorhabditis elegans*. They identified a gene, later named *lin-4*, which does not code for a protein but rather for a 22 nucleotide long RNA molecule. It was shown that the function of this small RNA is to repress the expression of the mRNA *lin-14*, by binding to the 3'UTR of the gene [3]. Later it was

discovered that miRNAs is a class of small RNAs (18-24 nt), that regulate gene expression post-transcriptionally and they have been found in plants, animals and DNA viruses [4-9]. MicroRNAs play a key role in the regulation of genes involved a diverse range of biological processes including development, cell proliferation, differentiation and apoptosis [10,11]. Approximately 1048 human miRNAs have been identified to date (miR-Base release 16) [12] and it is believed that miRNAs regulate about 30% of all protein coding human genes [13,14].

Since many miRNAs are differentially expressed between normal and malignant tissues, as shown in e.g. breast and pancreatic cancer, miRNA expression profiles have potential as tools for diagnosis and prognosis of cancer [15-20]. It has been shown that expression profiles of miRNAs could be used to classify and correctly diagnose even poorly differentiated tumor samples with higher accuracy than mRNAs. Lu *et al.*, investigated tumors with histologically uncertain cellular origin for which a clinical diagnosis was established by anatomical context (colon, ovary, lung, breast and diffuse large B

\* Correspondence: karin.klinga.levan@his.se

<sup>1</sup>Systems Biology Research Centre - Tumor Biology, School of Life Sciences, University of Skövde, Skövde, Sweden

Full list of author information is available at the end of the article

cell lymphoma) and showed that miRNA expression profiles could classify 12 out of 17 samples correctly while, when using mRNA expression, only one out of 17 samples was correctly classified. In addition, biomarker sets consisting of just a few miRNAs were informative enough to differentiate between tissue types [16].

Several attempts to find a miRNA expression profile for prostate cancer has been made during the last years but the results have been inconclusive. At present there are many conflicting results in the literature where results often segregate between different data sets, which can be due to study design, sample collection methods or the sensitivity and specificity of the different platforms used. Even though the results are conflicting, several studies indicate that it is possible to find a miRNA expression signature that can separate between normal and malignant prostate tissues [19,21-26].

In this study we aimed to identify a diagnostic miRNA expression signature, i.e. a set of miRNAs with expression profiles that consistently differ between normal and malignant prostate tissues. If such an expression signature can be identified and shown to have high classification accuracy, then it can potentially serve as the basis for a future diagnostic tool for prostate cancer.

## Results

In this study we included malignant prostate tissue and adjacent normal prostate tissue from twenty patients of the Swedish Watchful Waiting cohort, which consists of men with localized prostate cancer diagnosed by transurethral resection of the prostate (Table 1). The expression of 667 unique miRNAs was analyzed using the TaqMan<sup>®</sup> MicroRNA Array Set v2.0 from Applied Biosystems and miRNAs that were differentially expressed between the malignant and adjacent normal prostate tissues were identified by a paired Student's *t*-test. In total, 30 miRNAs were found to be differentially expressed at the 0.0001 significance level. When a more stringent *p*-value of 0.00001 was applied, a subset of nine differentially expressed miRNAs was identified. When the Benjamini-Hochberg correction was performed on the *p*-values, the differential expression of all the nine miRNAs was still significant at the *p* < 0.001 level (Table 2).

Hierarchical clustering of the two sets of differentially expressed miRNAs was performed, showing that both these miRNA expression signatures could be used to separate between the normal and malignant prostate tissues, with the exception of three and two misplaced samples, respectively (Figure 1, Figure 2). The PCA analysis performed on the smaller expression signature, including nine miRNAs, confirmed the results from the hierarchical clustering (Figure 3).

A PCA analysis of all 667 unique miRNAs was performed to find subgroups among the 19 malignant

**Table 1 Patient material**

Sample	Age	Gleason score	%	WHO
1	80	7	20	2
2	84	6	15	2
3	79	6	60	2
4	83	7	60	2
5	83	6	15	1
6	74	6	5	2
7	86	6	2	3
8	69	7	60	2
9	85	6	10	2
10	78	7	20	2
11	76	8	70	3
12	73	7	80	2
13	75	10	40	3
14	71	9	20	2
15	74	7	15	3
16	79	10	40	3
17	71	9	60	3
18	78	9	60	3
19	85	7	20	3
20	91	9	80	3

Sample stands for the code of the patient (ranging from 1 to 20) and the Gleason score is the total Gleason score from two different tumor areas in the same patient. WHO shows the tumor grading according to the WHO classification scale and % the percentage of tumor tissue.

samples included in the analyses. By gradually decreasing the *p*-value until groups emerged, we found a threshold value (*p* < 0.017) that could be used to identify a set of 16 miRNAs, which arranged the samples into four groups. We then evaluated these groups for correspondence with Gleason scores and found that, with the exception of three samples, the groups represented the four Gleason scores included in the study (Figure 4).

To analyze whether the clear separation between the normal and malignant tissues was due to the fact that the same sample set was used for deriving the expression signature and for testing its performance, we performed a generalization test using cross-validation. We randomly chose 28 samples (14 malignant and the 14 corresponding normal samples) and used them for identification of differentially expressed miRNAs (*p* < 0.0001). The expression signature of these miRNAs was then used to cluster the remaining 10 samples. This selection and clustering procedure was repeated 15 times (15-fold cross-validation). In each repetition we evaluated the hierarchical clustering by setting a threshold resulting in three clusters, since we had one outlier in many of the clusterings. A perfect separation between the normal and malignant tissues was found in four of the repetitions. In four cases, one sample was misplaced (error rate 10%), which means that one normal sample was placed in a cluster where the majority of samples were

**Table 2 Differentially expressed miRNAs**

miRNA	+/-	p-value	Reported no. of times	Adjusted p-value
<b>MIR126*</b>	-	<b>7.68e-07</b>	<b>11</b>	<b>4.50e-04</b>
<b>MIR34A*</b>	-	<b>1.19e-06</b>	<b>13</b>	<b>6.09e-04</b>
<b>MIR622</b>	+	<b>1.52e-06</b>	<b>9</b>	<b>3.88e-04</b>
<b>MIR195</b>	-	<b>3.06e-06</b>	<b>7</b>	<b>8.55e-04</b>
<b>MIR26A</b>	-	<b>4.05e-06</b>	<b>6</b>	<b>4.50e-04</b>
<b>MIR30D</b>	+	<b>4.35e-06</b>	<b>9</b>	<b>4.50e-04</b>
<b>MIR29A*</b>	-	<b>4.48e-06</b>	<b>4</b>	<b>6.09e-04</b>
<b>MIR425*</b>	+	<b>5.65e-06</b>	<b>6</b>	<b>4.50e-04</b>
<b>MIR342-3P</b>	+	<b>6.11e-06</b>	<b>4</b>	<b>6.09e-04</b>
MIR345	-	1.07e-05	2	1.77e-03
MIR24	-	1.08e-05	3	6.09e-04
MIR484	-	1.12e-05	2	1.16e-03
MIR497	+	1.13e-05	4	6.09e-04
MIR30E	+	1.22e-05	3	6.09e-04
MIR93	+	1.29e-05	3	1.64e-03
MIR221*	-	1.37e-05	1	1.64e-03
MIR774	+	1.44e-05	2	1.64e-04
MIR152	-	2.34e-05	0	1.37e-03
MIR27B	-	2.47e-05	4	8.62e-04
MIR101*	+	3.34e-05	2	4.69e-04
MIR145*	-	3.51e-05	1	4.69e-04
MIR143	-	4.65e-05	1	1.40e-03
MIR100	-	5.12e-05	1	6.09e-04
MIR519B-3P	+	5.12e-05	2	1.40e-03
MIR200C*	+	6.66e-05	1	3.88e-04
MIR501-5P	+	6.80e-05	0	1.12e-03
MIR200C	+	6.84e-05	2	3.88e-04
MIR30A	+	6.84e-05	0	1.64e-03
MIR191	+	7.85e-05	1	6.09e-04
MIR30B	-	9.43e-05	1	2.06e-03

MicroRNAs differentially expressed between normal and malignant tissue samples of the prostate with  $p < 0.0001$ . The nine miRNAs in bold text were also significant at the 0.00001 level. The signs +/- indicate if the miRNAs were up- or down-regulated in the malignant tissue compared to the adjacent normal prostate tissue. The reported number of times stands for the number of repetitions in which the miRNA was reported as differentially expressed in the 15-fold cross-validation test.

malignant, or vice versa. In three cases, two samples were misplaced, in three cases, three samples were misplaced, and in one case, four samples were misplaced (Figure 5). Thus, the average error rate was 15%. We also analyzed how many times the nine most significant miRNAs were identified as differentially expressed in the 15 repetitions of the generalization test, and found

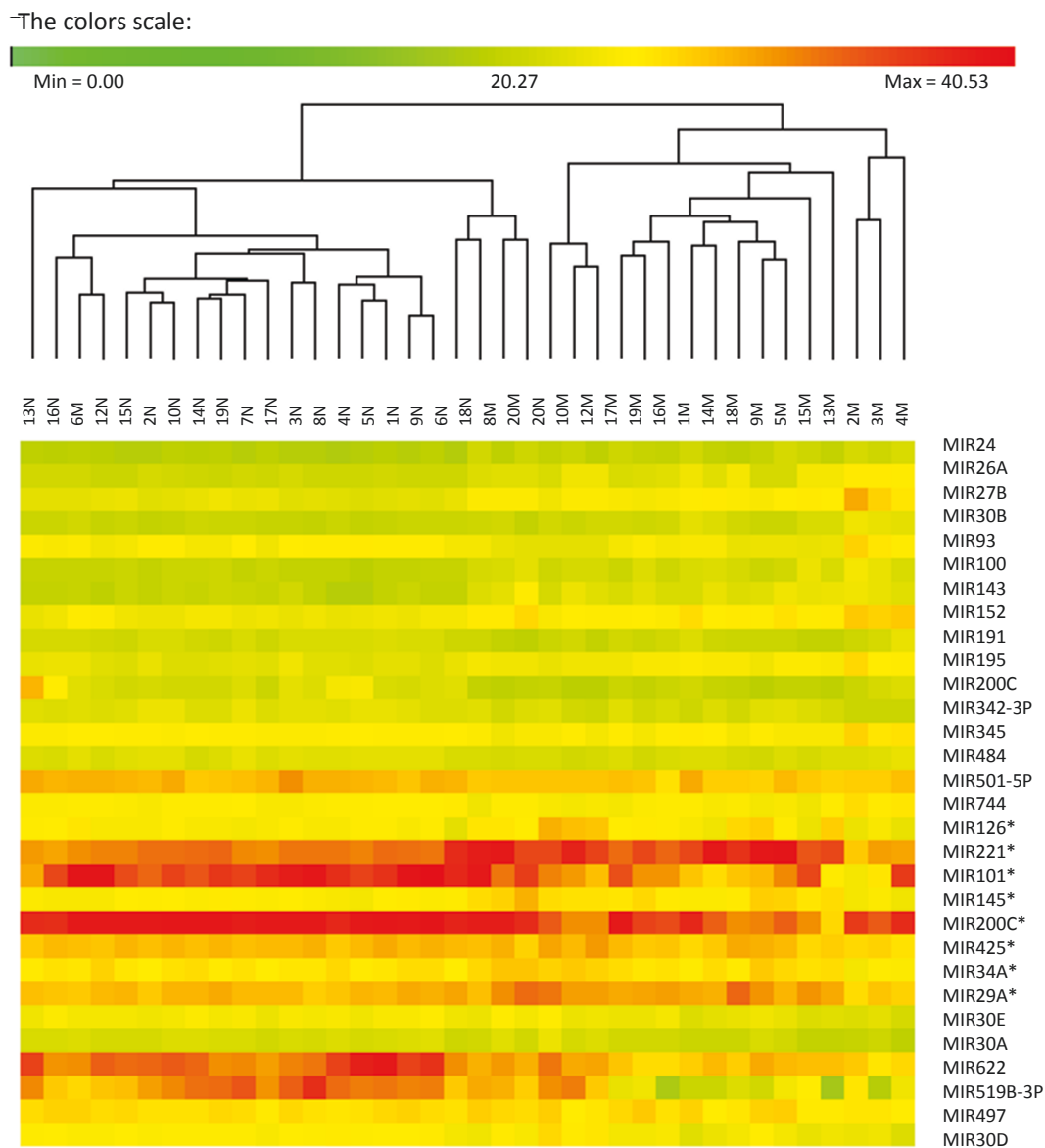
that this number ranged from four to 13 (Table 2). A permutation test was also performed, where the cross-validation procedure was repeated with randomly permuted class labels, but this did not produce any meaningful clusterings that could be interpreted in terms of classification accuracy.

## Discussion

In the present study, we aimed to find miRNAs with expression profiles that consistently differ between normal and malignant prostate tissues. We randomly selected 20 cases from the well-defined Swedish Watchful Waiting cohort (Table 1) and investigated the miRNA expression in malignant and the adjacent normal tissue in each individual in order to get a matched control for each case. 19 samples were used in the subsequent analyses since one sample had to be excluded due to a technical error in handling the qPCR data.

Differentially expressed miRNAs were detected by applying a paired Student's t-test. The test revealed that 30 miRNAs were differentially expressed at a  $p < 0.0001$  significance level and nine miRNAs at a more stringent level ( $p < 0.00001$ ). When a paired Wilcoxon test was applied to the data, 18 miRNAs were identified as differentially expressed ( $p < 0.0001$ ) between the normal and malignant samples. All except five of these miRNAs were also detected as differentially expressed by the paired Student's t-test and seven of the most differentially expressed from the t-test were also detected as differentially expressed by the Wilcoxon test (See Additional file 1). Due to the largely overlapping results, we decided to proceed with the results from the t-test under the assumption that the data used in the study is approximately normally distributed.

The two sets consisting of nine and 30 differentially expressed miRNAs were further analyzed in order to find out if they could be used as expression signatures to correctly separate between the normal and malignant tissues. Hierarchical clustering was performed on the larger expression signature, including 30 miRNAs, and the analysis revealed that 16 out of 19 malignant samples were correctly classified, as well as all of the normal samples (Figure 1). The hierarchical clustering analysis was also applied for the smaller expression signature consisting of nine miRNAs. Using this expression signature, 18 of the 19 malignant samples and 18 of the 19 normal samples were correctly classified. The reason for one of the malignant samples being misclassified was probably that it belongs to a low grade tumor (Gleason score of 6). Its miRNA expression pattern may therefore be more similar to normal tissues than to the fully transformed malignant tumors with a higher Gleason scores. However, all other GS 6 malignant samples were correctly classified. The malignant sample that was

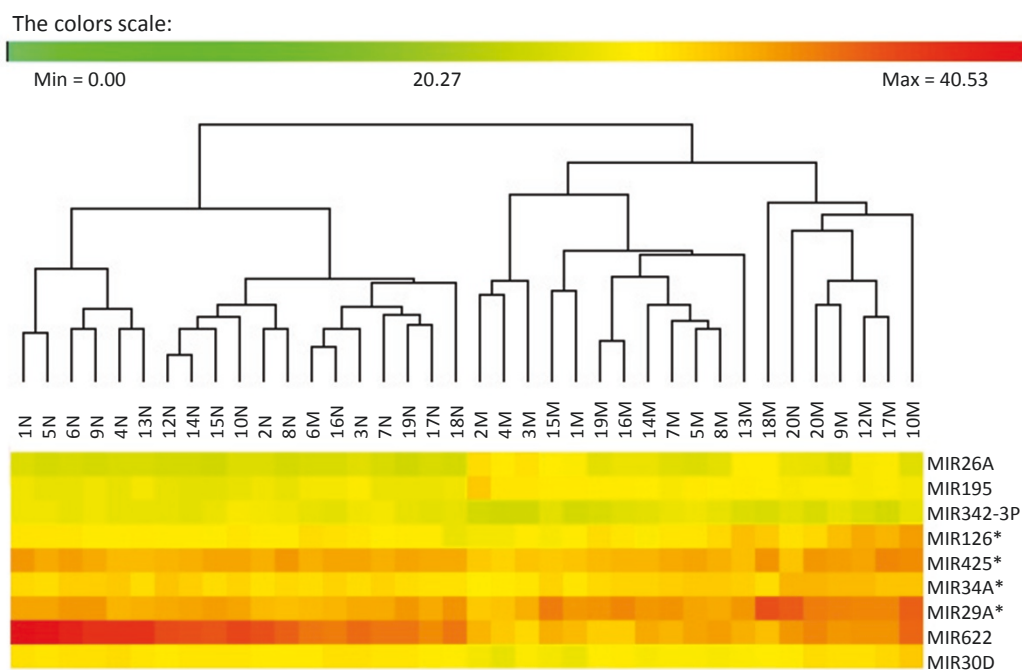


**Figure 1 miRNA expression signature consisting of 30 differentially expressed miRNAs.** Differentially expressed miRNAs ( $p < 0.0001$ ) were clustered and the results show that the expression profiles of these 30 miRNAs could be used to separate the normal (N) from the malignant (M) tissue samples with the exception for three malignant samples (6M, 8M and 20M).

misclassified had a low percentage of tumor cells (5%), which also might explain why it was placed in the normal cluster. On the other hand, the sample containing only 2% of tumor cells (7M) was correctly classified within the malignant cluster, which may indicate in some cases, the expression signature also works for tissues with a low percentage of tumor cells. A normal sample was also misclassified (20N), and we hypothesize that this might be due to that the normal tissue surrounding the tumor area might have been affected by the tumor, a phenomena called TINT (tumor indicating normal tissue) that has begun to be discussed within the

prostate cancer area recently [27]. The PCA analysis of the nine differentially expressed miRNAs confirmed the results from the clustering analysis, as the same samples (6M and 20N) were misclassified using this test (Figure 2, Figure 3).

A miRNA expression signature that will be used for clinical purposes should include a limited number of miRNAs due to practical and economic reasons. Thus, in addition to our initial set of 30 differentially expressed miRNAs we also analyzed a reduced expression signature including the nine most differentially expressed miRNA genes. We found that the smaller



**Figure 2 miRNA expression signature consisting of nine differentially expressed miRNAs.** Differentially expressed miRNAs ( $p < 0.00001$ ) were clustered and the results show that the expression profiles of these nine miRNAs could be used to separate between the normal (N) and malignant (M) tissue samples with the exception of one normal sample (20N) and one malignant sample (6M).

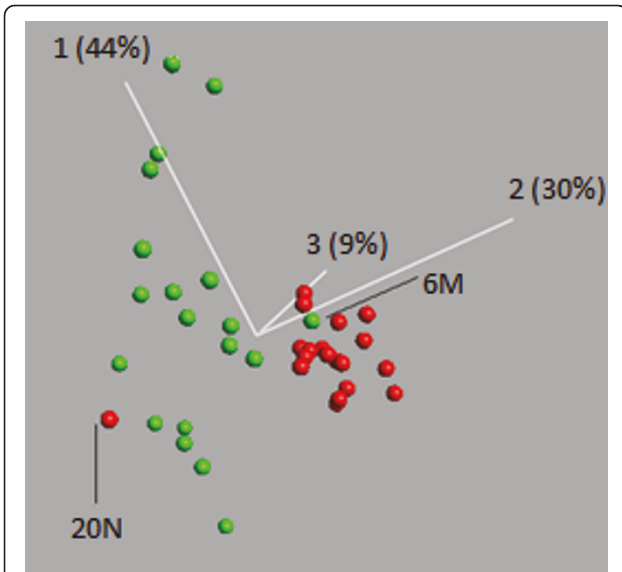
expression signature could classify normal and malignant samples more correctly than the larger signature, which is probably due to the more stringent approach when detecting differentially expressed miRNAs (a lower  $p$ -value).

Six of the miRNAs (MIR26A, MIR126\*, MIR195, MIR30D, MIR29A\* and MIR342-3P) included in the smaller expression signature, have previously been described to be involved in the development of prostate cancer [19,22-25,28,29]. The expression of MIR126\* has been investigated in prostate cancer in two previous studies and the results from these studies correspond well with our results as the miRNA was downregulated in both studies [23,29]. Porkka *et al.*, investigated the expression of three of the miRNAs included in our expression profile, MIR195, MIR26A and MIR29A\* [22]. Their results correspond well with the results from our study since we found that all these three miRNAs were downregulated in malignant prostate tissues. MIR26A and MIR30D have also been shown to be downregulated in malignant prostate tissue in another study [24] while in two other studies MIR26A together with MIR195 were found to be upregulated in malignant prostate tissues [19,25]. The results from the previous studies validate our results that MIR126\*, MIR195, MIR26A, MIR29A\* and MIR30D are differentially expressed in prostate cancer.

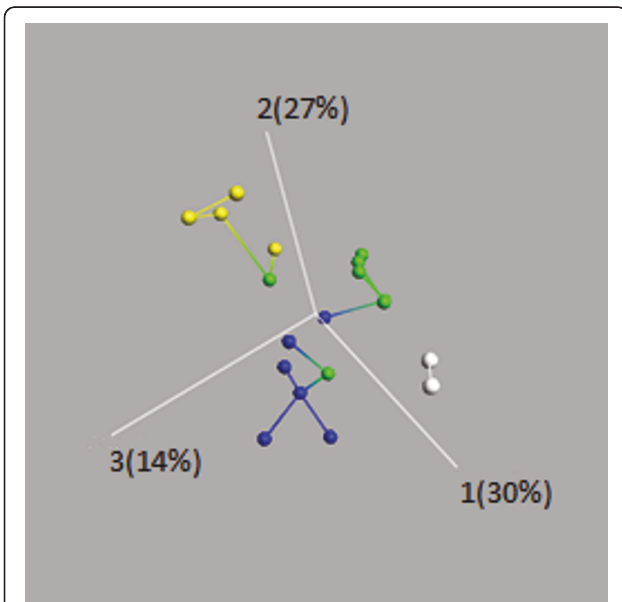
Three of the miRNAs in the signature (MIR26A, MIR126\* and MIR34A\*) have experimentally validated target genes (Table 3). *SLC45A3*, a target gene of MIR126\* encodes a prostate specific antigen called prostein [29]. There are five validated target genes of MIR26A, *SMAD1*, *PLAG1*, *TGFBR2*, *SERBP1* and *EZH2*, and one validated target gene, *NOTCH1*, of MIR34A\* [19,30-35]. These target genes are involved in pathways related to e.g. cell growth and proliferation. None of the miRNAs in the small expression signature seems to be prostate specific and many of them are differentially expressed in several other diseases, such as lung cancer and leukemias [36-41], which indicates that these miRNAs might be important in general cancer development. This means that the expression profile from a single miRNA within this nine miRNA expression signature may not be reliable for diagnosis of prostate cancer specifically. However, the combination of the expression profiles of all nine miRNAs could potentially be prostate specific and thus be used for diagnostic purposes, even in cases where prostate samples are replaced by other cell types, for example circulating tumor cells [42,43].

In order to test the generality of the expression signature, we randomly chose 28 samples (14 malignant and the 14 corresponding normal tissues) to find a new set of differentially expressed miRNAs, which was then used to cluster the remaining 10 samples. When

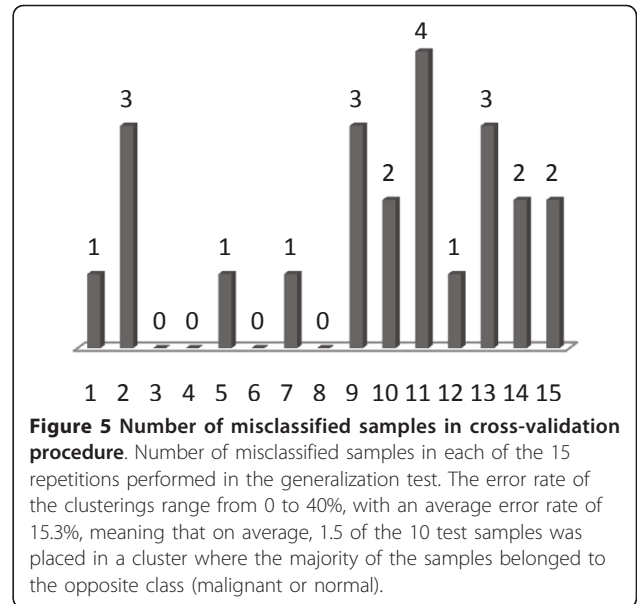




**Figure 3 Principal component analysis of the nine miRNA expression signature.** Principal component analysis of the nine miRNAs differentially expressed at the 0.00001 significance level. Only two samples (20N and 6M) were misplaced. The percentages on the axes describe the amount of variance that is picked up by the principal components in that direction. Green: malignant tissue sample, red: normal tissue sample.



**Figure 4 Principal component analysis to find subgroups within the malignant samples.** A set of 16 miRNAs could be used to separate the malignant samples according to Gleason scores (GS) with the exception of three samples: one GS 6 sample placed in the GS 7 group, one GS 7 sample placed in the GS 6 group, and one GS 7 sample placed in the GS 9 group. In the figure, one GS 7 sample is hidden behind another GS 7 sample. The strings between samples correspond to their nearest neighbors. Blue: GS 6, green: GS 7, yellow: GS 9, white: GS 10.



**Figure 5 Number of misclassified samples in cross-validation procedure.** Number of misclassified samples in each of the 15 repetitions performed in the generalization test. The error rate of the clusterings range from 0 to 40%, with an average error rate of 15.3%, meaning that on average, 1.5 of the 10 test samples was placed in a cluster where the majority of the samples belonged to the opposite class (malignant or normal).

repeated 15 times, this analysis indicated that regardless of the selection of 28 randomly chosen samples, at least one of the nine most differentially expressed miRNAs from our first analysis was identified as differentially expressed (Table 2). We also analyzed how many times each of the nine differentially expressed miRNAs in our signature were chosen as differentially expressed within this analysis (Table 2). We saw that MIR34A\* and MIR126\* followed by MIR622 and MIR30D are the miRNAs that are differentially expressed in most of the repetitions performed (87%, 73%, 60% and 60% of the repetitions, respectively).

No clear tendencies for how different Gleason scores cluster together were found in the hierarchical clustering analysis when using the set of nine miRNAs. We therefore performed a PCA analysis of all 667 miRNAs to find subgroups within the malignant samples. The result from this analysis was a set of 16 miRNAs that could be used to

**Table 3 Validated target genes**

miRNA	Validated target genes
MIR126*	SLC45A3 [29]
MIR34A*	NOTCH1 [35]
MIR622	-
MIR195	-
MIR26A	SERBP1, SMAD1, PLAG1, TGFBR2, EZH2 [19,30-34]
MIR30D	-
MIR29A*	-
MIR425*	-
MIR342-3P	-

Validated target genes for the nine most significantly differentially expressed miRNAs. Only three of the miRNAs currently have validated target genes.  
 - No validated target genes

classify the samples into four subgroups, which largely corresponded with Gleason scores, since only three of the 19 samples were misplaced. These results indicate that it may be possible to find a miRNA expression signature that can be used to aid tumor classification according to Gleason scores, which could be a useful complement to the manual classification performed by pathologists today. To obtain a more certain result regarding the correspondence between miRNA expression signatures and Gleason score, a more thorough study needs to be performed, focusing on this relationship.

## Conclusions

To conclude, we have shown that a miRNA expression signature consisting of nine miRNAs could separate between the normal and malignant prostate tissues with high accuracy. This separation seems to be achievable also on unseen samples, since a cross-validation test was performed and yielded similar results (85% of samples correctly classified). We have also showed that subgroups in the malignant data, revealed by miRNA expression profiles, show high concordance with Gleason scores. The miRNA signature proposed in this study needs to be evaluated in a larger patient material and preferably with another method, such as *in situ* hybridization. The results show that there are significant differences in miRNA expression between normal and malignant prostate, indicating that these small RNA molecules might be important in the biogenesis of prostate cancer and potentially also useful for clinical diagnosis of the disease.

## Methods

### Patient material

Patients were recruited from the population-based Swedish Watchful Waiting cohort [44], consisting of 1256 men with localized prostate cancer. These men had symptoms of benign prostatic hyperplasia (lower urinary tract symptoms) and were subsequently diagnosed with prostate cancer through transurethral resection (TUR-P). All men in this study were determined at the time of diagnosis to have clinical stage T1a or T1b, Mx, and Nx (small tumor, no metastases and no lymph node involvement), according to the staging system of the 2002 American Joint Committee on Cancer, called Classification of Malignant Tumors (TNM) [45]. The prospective follow-up time of this cohort is now up to 30 years. This study includes samples from men who were diagnosed at the University Hospital in Örebro (1977-1991) and at four centers in the southeast region of Sweden: Kalmar, Norrköping, Linköping, and Jönköping (1987-1999). The study was approved by the ethical committee in the Uppsala-Örebro region (Application

number M58-05). The material consisted of formalin fixed paraffin embedded (FFPE) malignant prostate tissues from 20 cases and the adjacent normal tissue from each case, i.e. in total 40 paired samples. We collected cases randomly within each category of Gleason score (6-10) to get an equal distribution of histological differentiation between low grade (6-7) and high grade (8-10) Gleason scores. In addition, the tumor material consisted of different percentages of tumor cells in order to reflect the clinical reality (Table 1).

### MicroRNA qPCR arrays

The TaqMan<sup>®</sup> MicroRNA Array Set v2.0 from Applied Biosystems was used in this study (Applied Biosystems, Foster City, CA, USA). It consists of two cards (Card A and Card B) containing 364 TaqMan<sup>®</sup> MicroRNA assays plus 20 control assays per card, which enables quantification of 667 unique human miRNAs in total. Card A contains miRNAs that tend to be functionally defined, and are broadly and/or highly expressed. The miRNAs in card B are narrowly expressed and/or expressed at low levels and are usually not functionally defined.

### RNA extraction and cDNA preparation

Malignant and adjacent normal tissue areas on the paraffin blocks were marked by a pathologist prior to punching out 3-4 cores (ø 0.6 mm) using the Tissue Micro Array equipment (Pathology devices, Westminster, USA). The Recover All Total Nucleic Acid Isolation Kit optimized for FFPE samples (Ambion, Foster City, CA, USA) was used to extract total RNA. A reverse transcription reaction of 4-10 ng of total RNA was performed using the TaqMan<sup>®</sup> MicroRNA reverse transcription kit and Megaplex<sup>™</sup> RT primers, human pool v2.0 (Applied Biosystems). Subsequently, the cDNA samples were pre-amplified using Megaplex<sup>™</sup> PreAmp primers and TaqMan<sup>®</sup> Preamp master mix (Applied Biosystems).

### Quantitative PCR

The pre-amplified cDNA samples were diluted in a 0.1 X TE Buffer (pH 8.0) before use in the qPCR reaction. The diluted pre-amplified cDNA was mixed with TaqMan<sup>®</sup> PCR master mix II (No AmpErase UNG, Applied Biosystems) and run in a 40 cycle qPCR reaction on the TaqMan<sup>®</sup> MicroRNA A and B Cards. All reactions were performed on the Applied Biosystems 7900 HT system.

### Data analysis

Raw Ct-values (Cycle threshold, i.e. the number of cycles where the amount of amplified cDNA crosses a defined threshold) were calculated using the SDS

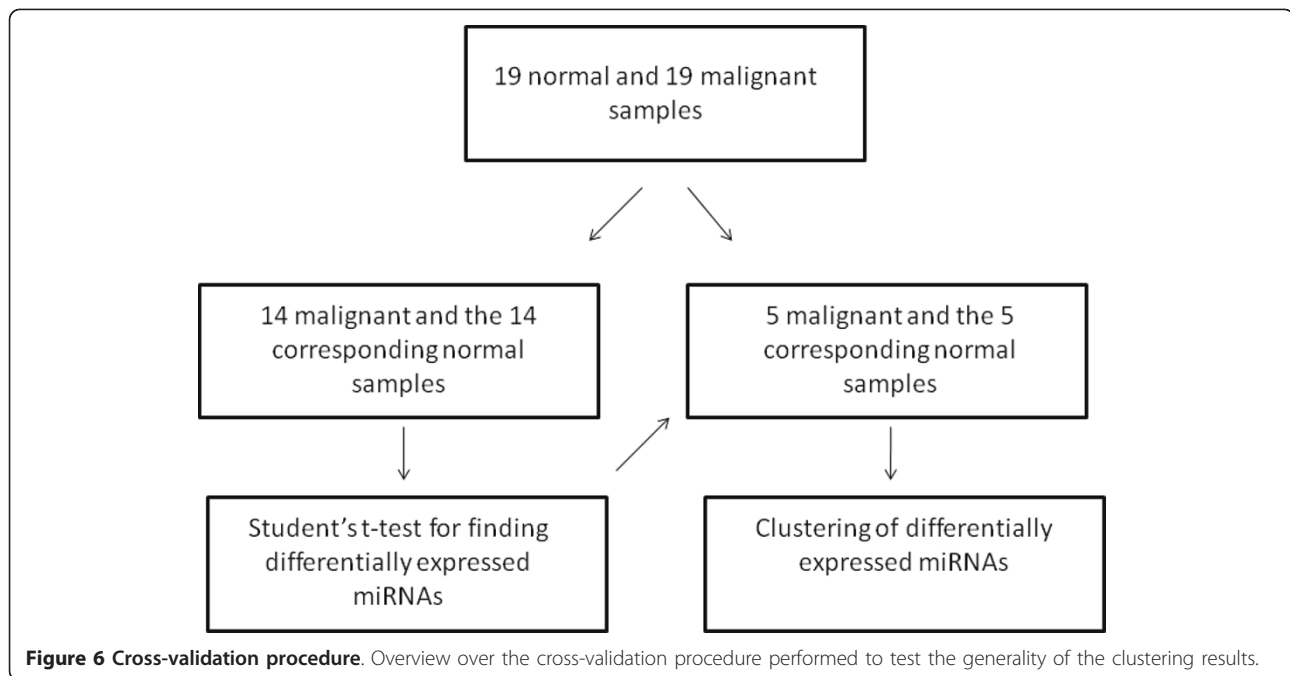
software (Applied Biosystems), applying manually selected thresholds for each miRNA (see Additional file 2). Due to a technical error in the handling of qPCR data, one sample (sample 11) had to be excluded from further analyses.

All statistical analyses were performed in the programming software R [46]. The raw Ct values were normalized using qPCRNorm quantile normalization, which is a data-driven normalization strategy for high-throughput qPCR data [47]. To select miRNAs to be included in the expression signature, differentially expressed miRNAs were detected by applying a paired Student's t-test on the normalized data ( $p < 0.0001$  and  $p < 0.00001$ ) and a Benjamini Hochberg multiple testing correction (included in the multtest R package) was applied to the  $p$ -values. A paired Wilcoxon test was also applied for comparison (see Additional file 1).

To test the accuracy of the miRNA expression signatures, hierarchical clustering analysis was performed on the differentially expressed miRNAs using the Permut-Matrix clustering tool [48]. For measurement of similarity between expression profiles, Euclidean distance was applied. Clustering was done using the average linkage rule, which means that the distance between two clusters is represented by the average of all pairwise distances between the objects contained in the two clusters.

To analyze the generality of the miRNA expression signature, the following cross-validation test of generalization was applied. A set of 14 paired samples (malignant and normal from the same case) was randomly selected and a new classification signature generated by identifying the differentially expressed miRNAs on these 28 samples using a paired Student's t-test ( $p < 0.0001$ ). The remaining 10 samples (five normal and five malignant) were then clustered based on the expression values of the selected miRNAs, and the separation between normal and malignant samples was recorded. This procedure was repeated 15 times and the average classification accuracy of the 15 clusterings was calculated (Figure 6).

A principal component analysis (PCA) was performed on the miRNA expression signatures, using Omics Explorer, Version 2.0 Beta (Qlucore AB, Lund, Sweden), to confirm the results from the hierarchical clusterings. An unsupervised PCA analysis was also performed to find subgroups within the malignant tissues. In this analysis, all miRNAs were initially used and the  $p$ -value was gradually decreased until subgroups within the data were revealed. These groups were then evaluated for correspondence with the *a priori* known groups, i.e. normal versus malignant and the different Gleason scores.



**Figure 6 Cross-validation procedure.** Overview over the cross-validation procedure performed to test the generality of the clustering results.



## Additional material

**Additional file 2: microRNA qPCR array raw data.** Raw data from microRNA qPCR card A and B (XLS-format).

### List of abbreviations used

cDNA: complementary DNA; DNA: deoxyribonucleic acid; miRNA: microRNA; mRNA: messenger RNA; nt: nucleotide; PCA: principal component analysis; qPCR: quantitative polymerase chain reaction; RNA: ribonucleic acid; TE: Tris-EDTA; UTR: untranslated region

### Acknowledgements and Funding

This work has been supported by the Swedish Knowledge Foundation through the Industrial PhD programme in Medical Bioinformatics at Corporate Alliances, Karolinska Institute, Lions cancer research foundation, Nyckelfonden, and Örebro county council research committee.

### Author details

<sup>1</sup>Systems Biology Research Centre - Tumor Biology, School of Life Sciences, University of Skövde, Skövde, Sweden. <sup>2</sup>Systems Biology Research Centre - Bioinformatics, School of Life Sciences, University of Skövde, Skövde, Sweden. <sup>3</sup>Department of Laboratory Medicine, Örebro University Hospital, Örebro, Sweden. <sup>4</sup>Department of Urology, Örebro University Hospital, Örebro, Sweden. <sup>5</sup>School of Health and Medical sciences, Örebro University, Örebro, Sweden.

### Authors' contributions

JC carried out the laboratory work, performed all analyses and drafted the manuscript. SD participated in the laboratory work. GH participated in the design of the study and the laboratory work and supervised the project. MK carried out the pathological marking of the tissues and supervised the project. ZL supervised the project and helped drafted the manuscript. OA, BO and KKL participated in the design of the study, supervised the project and helped drafted the manuscript. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 11 January 2011 Accepted: 27 May 2011

Published: 27 May 2011

### References

1. Socialstyrelsen: **Cancer Incidence in Sweden 2009**. Stockholm: Official statistics of Sweden; 2010.
2. De Angelis G, Rittenhouse HG, Mikolajczyk SD, Blair Shamel L, Semjonow A: **Twenty Years of PSA: From Prostate Antigen to Tumor Marker**. *Rev Urol* 2007, **9**(3):113-123.
3. Lee RC, Feinbaum RL, Ambros V: **The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14***. *Cell* 1993, **75**(5):843-854.
4. Stahlhut Espinosa CE, Slack FJ: **The role of microRNAs in cancer**. *Yale J Biol Med* 2006, **79**(3-4):131-140.
5. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T: **Identification of novel genes coding for small expressed RNAs**. *Science* 2001, **294**(5543):853-858.
6. Lau NC, Lim LP, Weinstein EG, Bartel DP: **An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans***. *Science* 2001, **294**(5543):858-862.
7. Lee RC, Ambros V: **An extensive class of small RNAs in *Caenorhabditis elegans***. *Science* 2001, **294**(5543):862-864.
8. Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, et al: **Identification of virus-encoded microRNAs**. *Science* 2004, **304**(5671):734-736.
9. Sullivan CS, Ganem D: **MicroRNAs and viral infection**. *Mol Cell* 2005, **20**(1):3-7.
10. Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM: **bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila***. *Cell* 2003, **113**(1):25-36.
11. Chen CZ, Li L, Lodish HF, Bartel DP: **MicroRNAs modulate hematopoietic lineage differentiation**. *Science* 2004, **303**(5654):83-86.
12. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ: **miRBase: microRNA sequences, targets and gene nomenclature**. *Nucleic Acids Res* 2006, **34**(Database issue):D140-144.
13. 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**(1):15-20.
14. Rajewsky N: **microRNA target predictions in animals**. *Nat Genet* 2006, **38** Suppl:S8-13.
15. Bloomston M, Frankel WL, Petrocca F, Volinia S, Alder H, Hagan JP, Liu CG, Bhatt D, Taccioli C, Croce CM: **MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis**. *JAMA* 2007, **297**(17):1901-1908.
16. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, et al: **MicroRNA expression profiles classify human cancers**. *Nature* 2005, **435**(7043):834-838.
17. Gaur A, Jewell DA, Liang Y, Ridzon D, Moore JH, Chen C, Ambros VR, Israel MA: **Characterization of microRNA expression levels and their biological correlates in human cancer cell lines**. *Cancer Res* 2007, **67**(6):2456-2468.
18. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, et al: **MicroRNA gene expression deregulation in human breast cancer**. *Cancer Res* 2005, **65**(16):7065-7070.
19. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, et al: **A microRNA expression signature of human solid tumors defines cancer gene targets**. *Proc Natl Acad Sci USA* 2006, **103**(7):2257-2261.
20. Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, et al: **MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype**. *Genome Biol* 2007, **8**(10):R214.
21. Tong AW, Fulgham P, Jay C, Chen P, Khalil I, Liu S, Senzer N, Eklund AC, Han J, Nemunaitis J: **MicroRNA profile analysis of human prostate cancers**. *Cancer Gene Ther* 2009, **16**(3):206-216.
22. Porkka KP, Pfeiffer MJ, Waltering KK, Vessella RL, Tammela TL, Visakorpi T: **MicroRNA expression profiling in prostate cancer**. *Cancer Res* 2007, **67**(13):6130-6135.
23. Mattie MD, Benz CC, Bowers J, Sensinger K, Wong L, Scott GK, Fedele V, Ginzinger D, Getts R, Haqq C: **Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies**. *Mol Cancer* 2006, **5**:24.
24. Ozen M, Creighton CJ, Ozdemir M, Iltmann M: **Widespread deregulation of microRNA expression in human prostate cancer**. *Oncogene* 2008, **27**(12):1788-1793.
25. Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, et al: **Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer**. *Cancer Res* 2008, **68**(15):6162-6170.
26. Schaefer A, Jung M, Mollenkopf HJ, Wagner I, Stephan C, Jentzmik F, Miller K, Lein M, Kristiansen G, Jung K: **Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma**. *Int J Cancer* 2010, **126**(5):1166-1176.
27. Halin SHP, Adamo H, Wikström P, Bergh A: **Tumor indicating normal tissue could be a new source of diagnostic and prognostic markers for prostate cancer**. *Expert opinion* 2010.
28. Prueitt RL, Yi M, Hudson RS, Wallace TA, Howe TM, Yfantis HG, Lee DH, Stephens RM, Liu CG, Calin GA, et al: **Expression of microRNAs and protein-coding genes associated with perineural invasion in prostate cancer**. *Prostate* 2008, **68**(11):1152-1164.
29. Musiyenko A, Bitko V, Barik S: **Ectopic expression of miR-126\*, an intronic product of the vascular endothelial EGF-like 7 gene, regulates protein translation and invasiveness of prostate cancer LNCaP cells**. *J Mol Med* 2008, **86**(3):313-322.
30. Luzi E, Marini F, Sala SC, Tognarini I, Galli G, Brandi ML: **Osteogenic differentiation of human adipose tissue-derived stem cells is modulated by the miR-26a targeting of the SMAD1 transcription factor**. *J Bone Miner Res* 2008, **23**(2):287-295.

31. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: **Prediction of mammalian microRNA targets.** *Cell* 2003, **115**(7):787-798.
32. Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G: **Identification of human microRNA targets from isolated argonaute protein complexes.** *RNA Biol* 2007, **4**(2):76-84.
33. Wong CF, Tellam RL: **MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis.** *J Biol Chem* 2008, **283**(15):9836-9843.
34. Sander S, Bullinger L, Klapproth K, Fiedler K, Kestler HA, Barth TF, Moller P, Stilgenbauer S, Pollack JR, Wirth T: **MYC stimulates EZH2 expression by repression of its negative regulator miR-26a.** *Blood* 2008, **112**(10):4202-4212.
35. Fukuda Y, Kawasaki H, Taira K: **Exploration of human miRNA target genes in neuronal differentiation.** *Nucleic Acids Symp Ser (Oxf)* 2005, **49**: 341-342.
36. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T, *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.** *Cancer Cell* 2006, **9**(3):189-198.
37. Saumet A, Vetter G, Bouttier M, Portales-Casamar E, Wasserman WW, Maurin T, Mari B, Barbry P, Vallar L, Friederich E, *et al*: **Transcriptional repression of microRNA genes by PML-RARA increases expression of key cancer proteins in acute promyelocytic leukemia.** *Blood* 2009, **113**(2):412-421.
38. Soon PS, Tacon LJ, Gill AJ, Bambach CP, Sywak MS, Campbell PR, Yeh MW, Wong SG, Clifton-Bligh RJ, Robinson BG, *et al*: **miR-195 and miR-483-5p Identified as Predictors of Poor Prognosis in Adrenocortical Cancer.** *Clin Cancer Res* 2009, **15**(24):7684-7692.
39. Huse JT, Brennan C, Hambardzumyan D, Wee B, Pena J, Rouhanifard SH, Sohn-Lee C, le Sage C, Agami R, Tuschl T, *et al*: **The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis in vivo.** *Genes Dev* 2009, **23**(11):1327-1337.
40. Visone R, Pallante P, Vecchione A, Cirombella R, Ferracin M, Ferraro A, Volinia S, Coluzzi S, Leone V, Borbone E, *et al*: **Specific microRNAs are downregulated in human thyroid anaplastic carcinomas.** *Oncogene* 2007, **26**(54):7590-7595.
41. Montag J, Hitt R, Opitz L, Schulz-Schaeffer WJ, Hunsmann G, Motzkus D: **Upregulation of miRNA hsa-miR-342-3p in experimental and idiopathic prion disease.** *Mol Neurodegener* 2009, **4**:36.
42. Panteleakou Z, Lembessis P, Sourla A, Pissimissis N, Polyzos A, Deliveliotis C, Koutsilieris M: **Detection of circulating tumor cells in prostate cancer patients: methodological pitfalls and clinical relevance.** *Mol Med* 2009, **15**(3-4):101-114.
43. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Brian KC, Allen A, *et al*: **Circulating microRNAs as stable blood-based markers for cancer detection.** *Proc Natl Acad Sci USA* 2008, **105**(30):10513-10518.
44. Johansson JE, Andren O, Andersson SO, Dickman PW, Holmberg L, Magnuson A, Adami HO: **Natural history of early, localized prostate cancer.** *JAMA* 2004, **291**(22):2713-2719.
45. Sobin L, Gospodarowicz M, Wittekind C: **TNM Classifications of Malignant Tumours.** Blackwell Publishing Ltd; Seventh 2010.
46. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**(10): R80.
47. Mar JC, Kimura Y, Schroder K, Irvine KM, Hayashizaki Y, Suzuki H, Hume D, Quackenbush J: **Data-driven normalization strategies for high-throughput quantitative RT-PCR.** *BMC Bioinformatics* 2009, **10**:110.
48. Caraux G, Pinloche S: **PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order.** *Bioinformatics* 2005, **21**(7):1280-1281.

doi:10.1186/1475-2867-11-14

**Cite this article as:** Carlsson *et al*: A miRNA expression signature that separates between normal and malignant prostate tissues. *Cancer Cell International* 2011 **11**:14.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

