



Identification of key miRNAs in prostate cancer progression based on miRNA-mRNA network construction



Giulia Dal Santo^{a,b}, Marco Frasca^b, Gloria Bertoli^a, Isabella Castiglioni^c, Claudia Cava^{a,*}

^a Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), Via F. Cervi 93, Segrate-Milan, 20090 Milan, Italy

^b Department of Computer Science, Università degli Studi di Milano, Via Celoria 18, 20133 Milano, Italy

^c Department of Physics "Giuseppe Occhialini", University of Milan-Bicocca Piazza dell'Ateneo Nuovo, 20126 Milan, Italy

ARTICLE INFO

Article history:

Received 10 November 2021

Received in revised form 3 February 2022

Accepted 3 February 2022

Available online 7 February 2022

Keywords:

miRNA
Prostate cancer
Network
Pathway
Gleason score

ABSTRACT

Prostate cancer (PC) is one of the major male cancers. Differential diagnosis of PC is indispensable for the individual therapy, i.e., Gleason score (GS) that describes the grade of cancer can be used to choose the appropriate therapy. However, the current techniques for PC diagnosis and prognosis are not always effective.

To identify potential markers that could be used for differential diagnosis of PC, we analyzed miRNA-mRNA interactions and we build specific networks for PC onset and progression. Key differentially expressed miRNAs for each GS were selected by calculating three parameters of network topology measures: the number of their single regulated mRNAs (NSR), the number of target genes (NTG) and NSR/NTG. miRNAs that obtained a high statistically significant value of these three parameters were chosen as potential biomarkers for computational validation and pathway analysis.

20 miRNAs were identified as key candidates for PC. 8 out of 20 miRNAs (*miR-25-3p*, *miR-93-3p*, *miR-122-5p*, *miR-183-5p*, *miR-615-3p*, *miR-7-5p*, *miR-375*, and *miR-92a-3p*) were differentially expressed in all GS and proposed as biomarkers for PC onset. In addition, "Extracellular-receptor interaction", "Focal adhesion", and "microRNAs in cancer" were significantly enriched by the differentially expressed target genes of the identified miRNAs. *miR-10a-5p* was found to be differentially expressed in GS 6, 7, and 8 in PC samples.

3 miRNAs were identified as PC GS-specific differentially expressed miRNAs: *miR-155-5p* was identified in PC samples with GS 6, and *miR-142-3p* and *miR-296-3p* in PC samples with GS 9.

The efficacy of 20 miRNAs as potential biomarkers was revealed with a Random Forest classification using an independent dataset. The results demonstrated our 20 miRNAs achieved a better performance (AUC: 0.73) than miRNAs selected with Boruta algorithm (AUC: 0.55), a method for the automated feature extraction.

Studying miRNA-mRNA associations, key miRNAs were identified with a computational approach for PC onset and progression. Further experimental validations are needed for future translational development.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Prostate cancer (PC) is the first leading cause of cancer in males as reported in Cancer Statistics 2020 [1]. As the development and progression of PC is highly heterogeneous further investigations are needed for precision medicine and personalized treatments in PC samples [2].

Gleason score plays a key role in predicting patient outcomes as it correlates with malignancy and aggressiveness of PC. Indeed, it is used for patient risk stratification in clinical practice [3]. Gleason score ≤ 6 represents the group with the best overall prognosis (grade group 1). Gleason score 7 is divided into Gleason 3 + 4 (grade group 2) and Gleason 4 + 3 (grade group 3), as evidence showed poorer prognosis in Gleason 4 + 3 tumors compared to Gleason 3 + 4. Gleason score 8 tumors represent grade group 4, and Gleason score 9 or 10 comprise the worst prognostic group (grade group 5) [3].

* Corresponding author.

E-mail address: claudia.cava@ibfm.cnr.it (C. Cava).

Currently, many studies identified single molecules as biomarkers for diagnosis and prognosis of cancer including PC [4–6]. The molecules are detected from high-throughput experimental data (e.g. microarray, and next-generation sequencing) using statistical or advanced computational approaches, for example, screening differentially expressed molecules between disease and control groups [4,5].

Prostate-specific antigen (PSA) is an example of a molecule translated from basic research to clinical practice as it is currently quantified for PC disease prediction. However, in clinical practice, although the level of PSA in serum and Magnetic Resonance Imaging (MRI) techniques are commonly used for PC screening, their sensitivity and specificity are not optimal to avoid unnecessary biopsy [7,8].

The occurrence of a disease is often the effect of multi-level systems rather than the alteration of a single molecule. Therefore, the detection of a single molecular biomarker from groups of patients with the same disease is often not reliable in a set of heterogeneous patients because of the complexity of disease [9]. In line with this scenario, a solution to resolve this issue, is the identification of network biomarkers, which include the altered molecules (nodes) as well as their relations (edges) [10].

microRNAs (miRNAs) are a class of post-transcriptional regulators deregulated in many cancers including PC [11]. Previous studies showed that miRNAs could regulate mRNAs through complementary sequences in the 3' untranslated region (3' UTR) of mRNA target. Based on the intensity of homology to the 3'UTR complementary sequence, miRNAs can cause the inhibition or degradation of mRNA target. In addition, this interaction could eventually modify the cross-talk among pathways and the biological processes of cellular activities [12,13].

Since miRNAs can be both oncogenes and tumor suppressor genes, the identification of complex genetic interactions including mRNA-miRNAs as biomarkers for early differential diagnosis, prognosis and theranostics in PC is of clinical interest [11].

Zhang et al. investigated the distribution of mRNA-miRNA interactions and they found that their distribution follows a power law, which suggested that miRNAs with more targets are few [14]. In addition, they identified a set of genes uniquely regulated by a specific miRNA. A statistical analysis on the topological characteristics of miRNA biomarkers in the network showed that miRNA biomarkers were able to regulate genes independently. In line with this observation, Yan et al. [15] demonstrated that specific miRNAs able to regulate a high percentage of transcription factors were more likely biomarkers.

These studies [14,15] proposed two topological features to characterize miRNA biomarkers based on miRNA-mRNA network: the number of single-line regulation (NSR) and transcription factor gene percentage (TFP). NSR can quantify the tendency of a miRNA to be a candidate biomarker for a specific disease, as it indicates the number of mRNA independently regulated by a specific miRNA, and a recent study demonstrated that miRNAs with higher NSR were biomarkers for cancer management [16]. Another measure to characterize miRNA biomarkers is TFP, defined as the percentage of transcription factor genes targeted by a unique miRNA [17].

In a previous study, a network-based approach called POMA (Pipeline of Outlier MicroRNA Analysis) was applied to explore the topological feature of miRNA-mRNA network to predict miRNA biomarkers for diagnosis and prognosis [18]. The approach is based on the integration of structural interaction of miRNA-mRNA network and miRNA/mRNA expression data, using NSR and TFP [18]. They identified differentially expressed miRNAs and mRNAs and they selected miRNAs with high NSR and TFP as PC biomarkers [18].

Based on these previous studies [16–18], in a recent study the number of targeted genes (NTG) for each miRNA, was also pro-

posed to measure the regulatory interactions of miRNAs in the network [19]. They identified key miRNAs considering miRNA-mRNA networks for predicting the development and progression of PC [19].

We extended our study and proposed a bioinformatic approach to identify common and specific miRNAs for the different PC Gleason scores. We identified PC Gleason score-specific miRNAs characterized by network topological measures and we proposed as PC biomarkers miRNAs with high NSR, NTG and NSR/NTG. The aim was to develop a method able to efficiently combine miRNAs and mRNAs in order to identify a reduced number of miRNAs as diagnostic biomarkers for the Gleason score.

2. Materials and methods

2.1. Data collection and pre-processing

Gene and microRNA expression levels of PC samples were collected from the Cancer Genome Atlas (TCGA) dataset. TCGAbiolinks package was used to download and pre-process the data (Access date: 14 October 2021) [20]. We considered for this study miRNA and gene expression profiles derived from the same patients. We grouped the samples based on the Gleason score which allows patients with PC to be divided into 5 different groups considering the aggressiveness of the tumor.

Experimentally validated miRNA-mRNA interactions were established using miRTarBase platform, a database of microRNA-target interactions validated *in vitro* by reporter assay, western blot, microarrays and next generation sequencing expresses (NGS) (Access date: 1 July 2021) [21]. miRBaseConverter was used for converting the name of miRNAs in different miRBase versions (Access date: 1 July 2021) [21,22].

2.2. Differential expression analysis

Differential expression analysis was performed among different PC sample groups, i.e., Gleason score ≤ 6 vs normal samples. Differentially expressed miRNAs and genes were identified using the TCGAanalyze_DEA function of the TCGAbiolinks package [20]. The criterion for the identification of differentially expressed miRNAs and genes was $|\text{LogFC}| \geq 1$ and adjusted p-value < 0.01 .

2.3. Network construction

We developed a miRNA-mRNA network consisting of differentially expressed miRNAs for each Gleason score. The human miRNA-mRNA network was built integrating validated miRNA-mRNA interactions from miRTarBase. Overall, a total of 5 networks based on Gleason score of PC samples were identified. To measure the regulatory interactions, we calculated the three parameters NSR, NTG, and NSR/NTG for each differentially expressed miRNA in the different PC Gleason scores. Fig. 1 shows an example of NSR and NTG for miRNAs.

Differentially expressed miRNAs specific for each Gleason score with significantly high NSR, NTG, and NSR/NTG values (p-value < 0.05 , Wilcoxon signed-rank test), as well as those miRNAs shared by more Gleason score were selected as key biomarkers for the development and progression of PC. Wilcoxon signed-rank test was used to select miRNAs with a significantly high NSR, NTG, and NSR/NTG value.

2.4. Validation dataset

In order to verify if differentially expressed genes obtained with TCGA dataset are consistent with other independent datasets, we

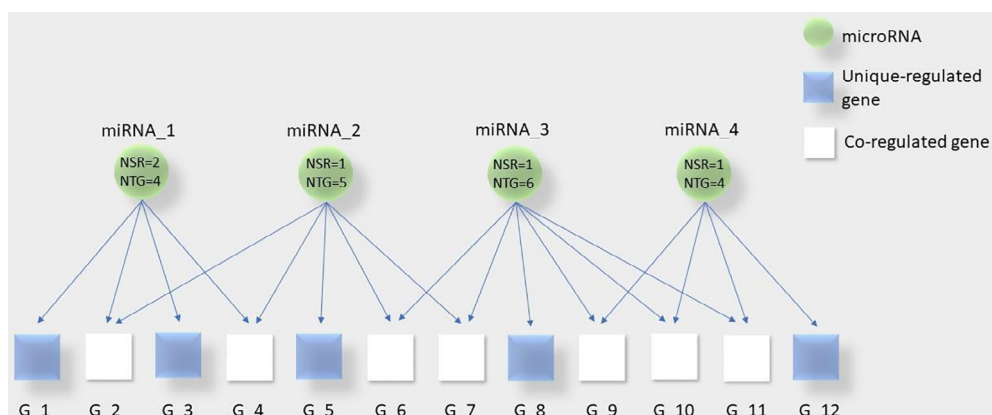


Fig. 1. Description of number of single-line regulation (NSR) and number of targeted genes (NTG) for each differentially expressed miRNA. We reported in blue squares the genes regulated by a unique miRNA and in a white square the genes regulated by more miRNAs. miRNAs are represented with a green circle. As example, the gene 1 (G_1) is regulated by a single miRNA (miRNA_1), while the gene 2 (G_2) is regulated by two miRNAs (miRNA_1 and miRNA_2). NSR of miRNA_1 is 2 as genes G_1 and G_3 are specifically regulated by that miRNA alone. NTG of miRNA_1 is 4 as it regulates 4 genes (G_1, G_2, G_3 and G_4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

considered validation datasets from 5 Gene Expression Omnibus (GEO): GSE118038, GSE21036, GSE45604, GSE46738, and GSE26367. We performed a differential expression analysis between normal samples and PC samples with different Gleason scores ($FDR < 0.05$ and $|\log FC| \geq 1$), using GEO2R.

2.5. Pathway analysis

We performed a pathway analysis using the R-package, clusterProfiler [23]. Hypergeometric distribution for KEGG pathways and p -value < 0.01 were considered for the statistical test. We studied the biological processes of differentially expressed miRNAs analyzing the pathways that are over-represented by differentially expressed target genes.

2.6. Assessing diagnostic relevance of miRNA

To investigate the efficacy of selected miRNAs as potential diagnostic markers for prostate cancer, we set up a classification problem to predict a given GS score (independent variable) in a one-vs-all setting, using the differentially expressed miRNAs for that GS score as input features (dependent variables). We used the dataset GSE118038 (Access date: 1 July 2021). A Random Forest classifier [24] with 500 trees was used, typically among the best performing features-based classifiers. As baseline, we also selected features (miRNAs) using the Boruta algorithm, a state-of-the-art method for the automated feature extraction [25]. This algorithm was allowed to automatically select for each GS score the set of most relevant miRNAs among the entire set of miRNAs in the dataset. The classifier was validated in a 10-fold cross validation setting, reporting the performance averaged across folds.

2.7. Assessing prognostic relevance of miRNA

We performed a survival analysis using the software UALCAN [26] using miRNA expression profiles from TCGA. Statistical significance of survival difference between groups was calculated with the p -value identified from log-rank test [26]. The patients were divided into groups with high and low expression of the miRNA. Specifically, high expression patients exhibit expression value $>$ 3rd quartile of the expression distribution.

3. Results

3.1. Differential analysis in PC progression

We selected 493 PC samples from TCGA that contain both gene and miRNA expression profiles and grouped the PC samples based on Gleason score in 5 grade groups: PC patients with GS 6, GS 7 (3 + 4), GS 7 (4 + 3), GS 8, and GS ≥ 9 . We also identified 52 normal samples that contain both gene and miRNA expression profiles from normal prostate tissues. Table 1 shows PC samples from TCGA divided by Gleason score.

The results obtained by differential expression analysis are presented in Table 2.

3.2. PC Gleason score-specific miRNA network

We generated five PC Gleason score-specific networks based on differentially expressed miRNAs and mRNAs.

NSR, NTG and NSR/NTG parameters for each differentially expressed miRNA were calculated to characterize the regulatory pattern of the miRNAs in the network and to identify miRNAs that could have a crucial role in the progression of PC

Table 3 reports the number of miRNAs with a significant high value of NSR, NTG, and NSR/NTG, and miRNAs that obtained a significant value for all three parameters. Supplementary file 1 shows the miRNAs for each NSR, NTG and NSR/NTG.

Overall, we obtained 20 miRNAs that showed a differential expression in PC samples with significantly high NSR, NTG and NSR/NTG values.

We identified 8 differentially expressed miRNAs (*miR-25-3p*, *miR-93-3p*, *miR-122-5p*, *miR-183-5p*, *miR-615-3p*, *miR-7-5p*, *miR-375*, and *miR-92a-3p*) in all Gleason score samples with significantly high NSR, NTG and NSR/NTG values.

Table 1
Number of samples for each Gleason score.

Gleason Score	N° of samples
6 (3 + 3, 2 + 4)	45
7 (3 + 4)	145
7 (4 + 3)	100
8 (4 + 4, 3 + 5, 5 + 3)	64
9, 10 (4 + 5, 5 + 4, 5 + 5)	139
	Tot.493

Table 2
Differential expressed genes between prostate cancer and normal samples.

Comparison	Up-regulated genes	Down-regulated genes	Tot
GS 6 vs normal samples	676	944	1620
GS 7 (3 + 4) vs normal samples	713	1017	1730
GS 7 (4 + 3) vs normal samples	952	1266	2218
GS 8 vs normal samples	1095	1288	2383
GS >= 9 vs normal samples	1123	1206	2329

miR-10a-5p was differentially expressed in PC samples with GS 6, GS 7 (3 + 4), GS 7 (4 + 3) and GS 8 with high statistically significant values of NSR, NTG and NSR/NTG values. In addition, its high expression in PC was associated with a poor prognosis through a survival analysis (Fig. 2).

Differentially expressed *miR-18a-3p* and *miR-18a-5p* in GS 6, GS 7 (4 + 3), GS 8 and GS 9 PC samples have a significant high value of NSR, NTG and NSR/NTG. *miR-423-3p* was differentially expressed in GS 7 (3 + 4), GS 7 (4 + 3), GS 8 and GS 9 PC samples with significantly high NSR, NTG and NSR/NTG values.

miR-197-3p was differentially expressed in GS 6, GS 8 and GS 9 PC samples with significant high NSR, NTG and NSR/NTG values. *miR-221-3p* was differentially expressed in GS 7 (4 + 3), GS 8, and GS 9 with significantly high NSR, NTG and NSR/NTG values.

miR-21-5p was differentially expressed in GS 8, and GS 9 with significantly high NSR, NTG and NSR/NTG values. *miR-320a* was differentially expressed in GS 6, and GS 7 (3 + 4) with significantly high NSR, NTG and NSR/NTG values. *miR-9-5p* was differentially expressed in GS 6, and GS 9 with significantly high NSR, NTG and NSR/NTG values.

PC grade group-specific differentially expressed miRNAs were also identified: *miR-155-5p* was identified in PC samples with GS 6. *miR-142-3p* and *miR-296-3p* in PC samples with GS 9.

We considered validation datasets from 5 GEO datasets: GSE118038, GSE21036, GSE45604, GSE46738, and GSE26367. We applied a differential expression analysis on GEO datasets considering normal samples and PC samples with different Gleason

scores. Table 4 shows the number of normal and PC primary tumor for each GEO dataset.

9 out of 20 miRNAs (*miR-10a-5p*, *miR-183-5p*, *miR-18a-5p*, *miR-21-5p*, *miR-221-3p*, *miR-25-3p*, *miR-375*, *miR-7-5p*, and *miR-93-3p*) that obtained a statistically significant value of NSR, NTG and NSR/NTG were differentially expressed in at least one of 5 GEO datasets.

For Gleason score 6, 2 miRNAs (*miR-183-5p*, and *miR-375*) were differentially expressed in GSE21036 and GSE26367.

As the GEO datasets do not contain the information of Gleason score 3 + 4 and 4 + 3 for PC samples with Gleason score 7, we could not distinguish the two groups. For Gleason score 7: 3 miRNAs (*miR-183-5p*, *miR-375*, and *miR-221-3p* in GSE21036; *miR-375*, *miR-183-5p*, *miR-221-3p* in GSE45604; *miR-183-5p*, and *miR-375* in GSE26367) out of 14 (*miR-423-3p*, *miR-10a-5p*, *miR-25-3p*, *miR-93-3p*, *miR-122-5p*, *miR-183-5p*, *miR-320a*, *miR-615-3p*, *miR-7-5p*, *miR-375*, *miR-92a-3p*, *miR-18a-5p*, *miR-221-3p*, and *miR-18a-3p*) were differentially expressed in PC patients with Gleason score 7 in the GEO datasets.

For Gleason score 8 8 miRNAs (*miR-10a-5p*, *miR-183-5p*, *miR-18a-5p*, *miR-21-5p*, *miR-221-3p*, *miR-25-3p*, *miR-375*, and *miR-7-5p*), out of 15 miRNAs (*miR-423-3p*, *miR-10a-5p*, *miR-18a-5p*, *miR-25-3p*, *miR-93-3p*, *miR-221-3p*, *miR-122-5p*, *miR-183-5p*, *miR-197-3p*, *miR-18a-3p*, *miR-21-5p*, *miR-615-3p*, *miR-7-5p*, *miR-375*, and *miR-92a-3p*) were differentially expressed in PC samples with Gleason score 8 in GEO datasets.

Specifically, in GSE118038 *miR-10a-5p*, *miR-18a-5p*, *miR-221-3p*, *miR-183-5p*, and *miR-21-5p*; in GSE21036 *miR-18a-5p*, *miR-221-3p*, *miR-183-5p*, *miR-7-5p*, and *miR-375*; in GSE46738 *miR-25-3p*, *miR-183-5p*, *miR-21-5p*, and *miR-375*; in GSE26367 *miR-25-3p*, *miR-183-5p*, and *miR-375*.

5 out of 17 differentially expressed miRNAs in Gleason score >= 9 of TCGA data were differentially expressed in Gleason score >= 9 of GEO datasets (*miR-25-3p*, *miR-221-3p*, *miR-183-5p*, and *miR-375* in GSE21036; *miR-93-3p*, and *miR-221-3p* in GSE45604). The reduced number of differentially expressed miRNAs can be due to the low number of samples with Gleason score 9 in GEO dataset.

Fig. 3 summarizes the results obtained in TCGA data.

Table 3

Differentially expressed miRNAs for each prostate cancer grade group with a significant high value of number of single-line regulation (NSR), number of targeted genes (NTG) and NSR/NTG. miRNAs that have all 3 significant parameters were presented in the last column.

Gleason Score (GS)	# miRNAs (NSR)	# miRNAs NTG	# miRNAs NSR/NTG	# miRNAs NSR ∩ NTG ∩ NSR/NTG
GS 6	27	27	26	<i>miR-10a-5p</i> , <i>miR-18a-5p</i> , <i>miR-25-3p</i> , <i>miR-93-3p</i> , <i>miR-122-5p</i> , <i>miR-183-5p</i> , <i>miR-197-3p</i> , <i>miR-18a-3p</i> , <i>miR-9-5p</i> , <i>miR-320a</i> , <i>miR-155-5p</i> , <i>miR-615-3p</i> , <i>miR-7-5p</i> , <i>miR-375</i> , <i>miR-92a-3p</i>
GS 7 (3 + 4)	22	22	23	<i>miR-423-3p</i> , <i>miR-10a-5p</i> , <i>miR-25-3p</i> , <i>miR-93-3p</i> , <i>miR-122-5p</i> , <i>miR-183-5p</i> , <i>miR-320a</i> , <i>miR-615-3p</i> , <i>miR-7-5p</i> , <i>miR-375</i> , <i>miR-92a-3p</i>
GS 7 (4 + 3)	24	26	27	<i>miR-423-3p</i> , <i>miR-10a-5p</i> , <i>miR-18a-5p</i> , <i>miR-25-3p</i> , <i>miR-93-3p</i> , <i>miR-221-3p</i> , <i>miR-122-5p</i> , <i>miR-183-5p</i> , <i>miR-18a-3p</i> , <i>miR-21-5p</i> , <i>miR-615-3p</i> , <i>miR-7-5p</i> , <i>miR-375</i> , <i>miR-92a-3p</i>
GS 8	35	31	32	<i>miR-423-3p</i> , <i>miR-10a-5p</i> , <i>miR-18a-5p</i> , <i>miR-25-3p</i> , <i>miR-93-3p</i> , <i>miR-221-3p</i> , <i>miR-122-5p</i> , <i>miR-183-5p</i> , <i>miR-197-3p</i> , <i>miR-18a-3p</i> , <i>miR-21-5p</i> , <i>miR-615-3p</i> , <i>miR-7-5p</i> , <i>miR-375</i> , <i>miR-92a-3p</i>
GS 9	33	36	36	<i>miR-423-3p</i> , <i>miR-18a-5p</i> , <i>miR-25-3p</i> , <i>miR-93-3p</i> , <i>miR-221-3p</i> , <i>miR-296-3p</i> , <i>miR-122-5p</i> , <i>miR-183-5p</i> , <i>miR-197-3p</i> , <i>miR-18a-3p</i> , <i>miR-9-5p</i> , <i>miR-21-5p</i> , <i>miR-142-3p</i> , <i>miR-615-3p</i> , <i>miR-7-5p</i> , <i>miR-375</i> , <i>miR-92a-3p</i>

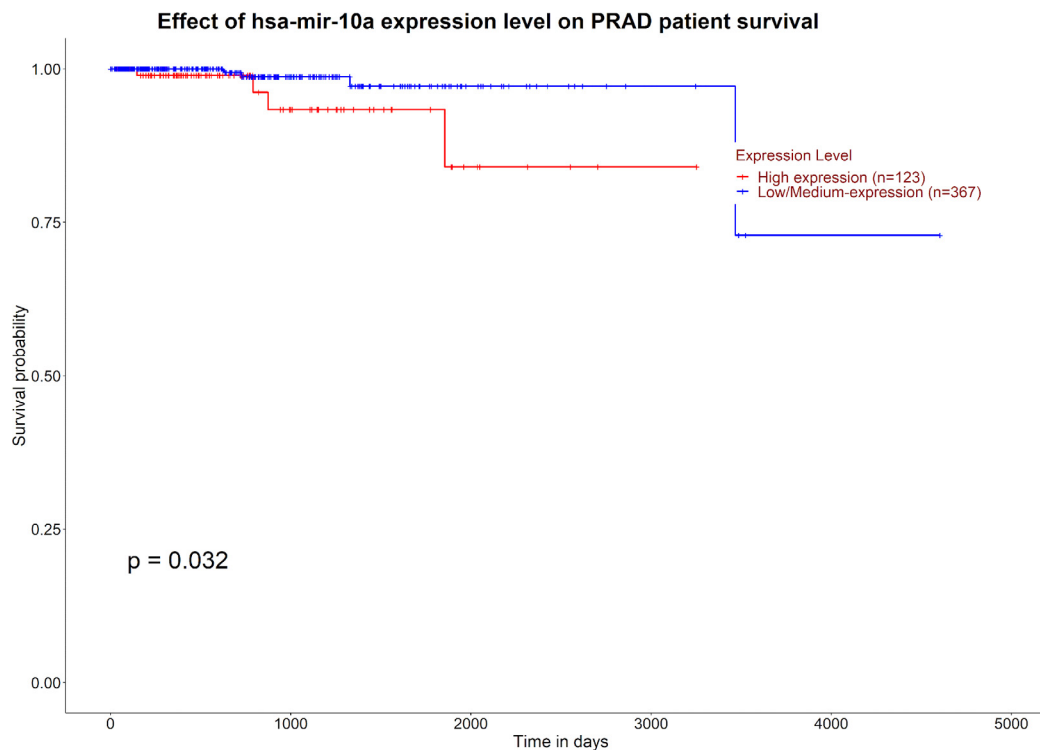


Fig. 2. Survival analysis applied to *miR-10a-5p*. The high expression of miRNA shows a poor overall survival in patients with prostate cancer.

Table 4

Number of normal and primary tumor tissues for GSE118038, GSE21036, GSE45604, GSE46738, and GSE26367 considered for the analysis of differentially expressed miRNAs. The primary tumors are divided by Gleason score (GS).

DATASET	NORMAL	PRIMARY TUMOR	GS ≤ 6	GS 7	GS 8	GS 9o 10	GS NA
GSE118038	37	33	0	11	12	9	1
GSE21036	28	99	57	33	6	3	0
GSE45604	10	50	15	25	6	4	0
GSE46738	4	53	15	13	21	4	0
GSE26367	11	173	46	86	29	3	9

3.3. Pathway analysis

In order to investigate the biological role of 20 miRNAs, we performed a pathway analysis. The analysis revealed for each miRNA the pathways enriched with differentially expressed genes for each Gleason score that are targets of that miRNA. Table 5 reported the pathways over-represented by differentially expressed genes that are targets of the 20 miRNAs.

We explored the pathways enriched with the targets of the 8 differentially expressed miRNAs (*miR-25-3p*, *miR-93-3p*, *miR-122-5p*, *miR-183-5p*, *miR-615-3p*, *miR-7-5p*, *miR-375*, and *miR-92a-3p*) presented all Gleason score.

Differentially expressed genes that are targets of *miR-25-3p* are enriched in the pathway “Mucin type O-glycan biosynthesis” in Gleason score 7 (3 + 4), “Oocyte meiosis”, “Mucin type O-glycan biosynthesis”, and

“Fanconi anemia” in Gleason score 7 (4 + 3) and GS 8, “Oocyte meiosis” in GS > 9.

Differentially expressed targets of *miR-93-3p* are enriched in the pathway “ECM-receptor interaction” in GS 7 (3 + 4) and (4 + 3).

miR-122-5p is enriched in the pathways “Mucin type O-glycan biosynthesis”, and “Glycosphingolipid biosynthesis - ganglio series”.

The pathways “Arrhythmogenic right ventricular cardiomyopathy”, “ECM-receptor interaction”, “Hypertrophic cardiomyopathy”,

and “Dilated cardiomyopathy” are enriched with differentially expressed targets of *miR-183-5p* in GS 6; “Progesterone-mediated oocyte maturation” and “Oocyte meiosis” are enriched with differentially expressed targets of *miR-183-5p* in GS 7 (4 + 3), GS 8 and GS < 9.

miR-615-3p significantly regulates the pathways “ECM-receptor interaction”, “Focal adhesion” and

“Viral carcinogenesis” in GS 6; “Viral carcinogenesis” in GS 7 (3 + 4); “Focal adhesion”, “Viral carcinogenesis”,

“ECM-receptor interaction”, “Cell cycle”, “Systemic lupus erythematosus”, and “PI3K-Akt signaling” in GS 7 (4 + 3); “Viral carcinogenesis” and “Cell cycle” in GS 8; “Viral carcinogenesis” in GS > 9.

Differential expressed genes in Gleason score 6, 7 (3 + 4), 7 (4 + 3), 8, and > 9, targets of *miR-375* are enriched in the pathway: “MicroRNAs in cancer”.

PC grade group-specific differentially expressed miRNAs were also identified: *miR-155-5p* was identified in PC samples with GS 6. *miR-142-3p* and *miR-296-3p* in PC samples with GS 9.

In addition, we analyzed the pathways regulated by miRNAs specific for Gleason score. For *miR-155-5p*, specific for Gleason score 6, its differentially expressed genes are enriched in the pathway “Neutrophil extracellular trap formation”. For *miR-196a-5p*, specific for Gleason score 7 (4 + 3) its differentially expressed genes are enriched in the pathway “Cell cycle”. 3 pathways “Human T-

	6	34	43	8	910
miR-25-3p					
miR-93-3p					
miR-122-5p					
miR-183-5p					
miR-615-3p					
miR-7-5p					
miR-375					
miR-92a-3p					
miR-10a-5p					
miR-18a-5p					
miR-423-3p					
miR-18a-3p					
miR-221-3p					
miR-197-3p					
miR-320a					
miR-9-5p					
miR-21-5p					
miR-155-5p					
miR-296-3p					
miR-142-3p					

Fig. 3. 20 differentially expressed miRNAs in TCGA data grouped by Gleason score with a significant number of single-line regulation (NSR), number of targeted genes (NTG) and NSR/NTG. The cells colored in blue indicate that the miRNA is differentially expressed in that Gleason score. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cell leukemia virus 1 infection”, “Cushing syndrome”, and “Endocrine resistance” are enriched with differential expressed genes in Gleason score 8 targets of *miR-331-3p* specific for Gleason score 8. Differentially expressed genes in Gleason score 9 and 10, targets of *miR-335-5p*, are enriched in 2 pathways “Hypertrophic cardiomyopathy”, and “Wnt signaling pathway”.

3.4. miRNAs as potential diagnostic markers

To evaluate the diagnostic ability of the 20 miRNAs identified we classified the samples of the PC dataset GSE118038. The dataset GSE118038 was used, since both normal and primary tumor classes are almost equally represented.

We achieved a good performance: 0.83 as AUC for the classification GS 7 vs others, 0.68 for GS 8 vs other and 0.69 for G9 vs others using the 20 miRNAs. In addition, our miRNA signature achieved a better performance than Boruta algorithm. In Table 6 the accuracy and area under receiver operating characteristic (ROC) curve were reported for each GS score separately.

4. Discussion

In clinical practice, there is an increasing of screening and monitoring strategies for PC [3]. However, they lack accuracy to show the dynamical changes during PC progression [3]. miRNAs are important regulators of expression in several biological processes including pathways associated to PC, which makes them as appealing biomarkers for PC personalized medicine [11]. It is widely known that altered miRNAs are associated with different diseases, including PC [11].

In this study, we explored key miRNAs as candidate biomarkers for PC development and progression. We performed a computational approach based on system biology to integrate translational informatics with medicine. We investigated miRNA and gene expression levels from TCGA and we divided PC samples based on Gleason score. Then a differential expression analysis from the comparison between PC samples with a specific Gleason score and normal samples is applied, and we identified differentially expressed miRNAs and genes. For each differentially expressed

miRNA the mRNA targets have been studied, and we built a miRNA-mRNA network specific for each Gleason score. Network topological measures (NSR, NTG, and NSR/NTG) were calculated for each miRNA in the networks, yielding a list of interesting miRNAs with a potential role in PC. Compared with a previous study [19] that considered solely localized PC and metastatic PC samples, our study analyzed miRNAs associated with the development and progression processes of PC based on different Gleason score of cancer. Specifically, NSR, NTG, and NSR/NTG for each differentially expressed miRNA in a specific Gleason score were calculated for the first time to measure the regulatory power of miRNAs. Overall, we found 20 differentially expressed miRNAs in TCGA data with a role in PC and with a statistically significant value of NSR, NTG and NSR/NTG. We used GEO independent datasets to validate the differential expression of these miRNAs and we found that 9 out of 20 miRNAs were differentially expressed also in at least one of 5 GEO datasets considered. To investigate the efficacy of 20 miRNAs as potential tool for differential diagnosis of PC, we performed a Random Forest classification using an independent dataset. We compared our 20 miRNAs with miRNAs selected by the Boruta algorithm, a well-known state-of-the-art method for the automated feature extraction. The results demonstrated that our 20 miRNAs are at least as informative as the boruta features for the prediction of the Gleason score.

We found 8 differentially expressed miRNAs (*miR-615-3p*, *miR-7-5p*, *miR-375*, *miR-92a-3p*, *miR-25-3p*, *miR-93-3p*, *miR-122-5p*, and *miR-183-5p*) in all Gleason score samples with statistically significant high value of NSR, NTG and NSR/NTG that could be responsible of PC onset.

The high expression of *miR-615-3p* was already associated with poor outcome in PC patients indicating its oncogenic role [27]. In addition, functional studies demonstrated that *miR-615-3p* increases cell viability, apoptosis, and proliferation [28]. Pathway analysis in our study, reported its association with “Extracellular-receptor interaction” and “Focal adhesion”.

miR-7-5p has been proposed as a tumor suppressor miRNA. In particular, in those PC samples with a lymphatic dissemination phenotype, this miRNA is upregulated [29]. Indeed, in PC, as well as in breast cancer, hepatocarcinoma and glioblastoma multiforme,

Table 5
Number of pathways enriched with differentially expressed genes (DEGs) that are targets of the 20 miRNAs involved in prostate cancer.

miRNA	GS 6	GS 7 (3 + 4)	GS 7 (4 + 3)	GS 8	GS 9/10
<i>miR-10a-5p</i>	1) Bile secretion (DEGs: 24)	(DEGs: 27)	(DEGs: 39)	1) Cell cycle (DEGs: 38)	1) Cell cycle (DEGs: 38)
<i>miR-122-5p</i>	1) Mucin type O-glycan biosynthesis 2) Glycosphingolipid biosynth - ganglio series (DEGs: 46)	1) Mucin type O-glycan biosynthesis 2) Glycosphingolipid biosynth - ganglio series (DEGs: 48)	(DEGs: 64)	(DEGs: 63)	(DEGs: 61)
<i>miR-142-3p</i>	(DEGs: 23)	(DEGs: 24)	(DEGs: 32)	(DEGs: 36)	(DEGs: 27)
<i>miR-155-5p</i>	(DEGs: 47)	(DEGs: 53)	1) Melanoma 2) Regulation of actin cytoskeleton 3) Cellular senescence 4) AGE-RAGE signaling pathway in diabetic complications 5) Insulin resistance 6) PI3K-Akt signaling pathway 7) Pancreatic cancer 8) Focal adhesion (DEGs: 68)	(DEGs: 67)	(DEGs: 65)
<i>miR-183-5p</i>	1) Arrhythmogenic right ventricular cardiomyopathy 2) ECM-receptor interaction 3) Hypertrophic cardiomyopathy 4) Dilated cardiomyopathy (DEGs: 17)		1) Progesterone-mediated oocyte maturation 2) Oocyte meiosis (DEGs: 26)	1) Progesterone-mediated oocyte maturation 2) Oocyte meiosis (DEGs: 23)	1) Progesterone-mediated oocyte maturation 2) Oocyte meiosis (DEGs: 27)
<i>miR-18a-3p</i>	(DEGs: 25)	(DEGs: 19)	(DEGs: 32)	(DEGs: 29)	(DEGs:33)
<i>miR-18a-5p</i>	1) Focal adhesion (DEGs: 13)	1) Focal adhesion (DEGs: 13)	1) Nitrogen metabolism (DEGs: 21)	1) Nitrogen metabolism (DEGs: 19)	1) Nitrogen metabolism (DEGs: 23)
<i>miR-197-3p</i>	(DEGs: 25)	(DEGs: 30)	(DEGs: 35)	(DEGs: 39)	(DEGs: 38)
<i>miR-21-5p</i>	1) Prostate cancer 2) MicroRNAs in cancer (DEGs): 36	1) MicroRNAs in cancer (DEGs: 37)	1) MicroRNAs in cancer 2) Prostate cancer 3) Hypertrophic cardiomyopathy 4) Dilated cardiomyopathy 5) Endocrine resistance (DEGs: 51)	1) MicroRNAs in cancer 2) Hypertrophic cardiomyopathy 3) Dilated cardiomyopathy 4) Prostate cancer 5) Bladder cancer 6) Melanoma 7) Hepatitis C 8) Hepatitis B 9) Arrhythmogenic right ventricular cardiomyopathy 10) Endocrine resistance (DEGs: 52)	1) MicroRNAs in cancer 2) Prostate cancer 3) Hepatitis B 4) Endocrine resistance 5) Melanoma 6) Hepatitis C 7) Small cell lung cancer 8) Bladder cancer 9) Epstein-Barr virus infection 10) Chemical carcinogenesis - receptor activation (DEGs: 57)
<i>miR-221-3p</i>	1) MAPK signaling 2) MicroRNAs in cancer (DEGs:24)	1) MAPK signaling (DEGs:27)			
<i>miR-25-3p</i>	(DEGs: 32)	1) Mucin type O-glycan biosynth. (DEGs:25)	(DEGs: 40) 1) Oocyte meiosis 2) Mucin type O-glycan biosynth. 3) Fanconi anemia (DEGs: 44)	(DEGs:38) 1) Oocyte meiosis 2) Mucin type O-glycan biosynth. 3) Fanconi anemia (DEGs:40)	(DEGs: 40) 1) Oocyte meiosis (DEGs: 49) (DEGs: 19)
<i>miR-296-3p</i>	(DEGs: 14)	(DEGs:15)	1) Viral carcinogenesis 2) Systemic lupus erythematosus 3) Alcoholism 4) Neutrophil extracellular trap formation 5) Central carbon metabolism in	(DEGs:20)	

Table 5 (continued)

miRNA	GS 6	GS 7 (3 + 4)	GS 7 (4 + 3)	GS 8	GS 9/10
<i>miR-320a</i>	1) MicroRNAs in cancer 2) Adherens junction (DEGs: 29)	1) Neutrophil extracellular trap formation (DEGs: 36)	cancer 6) Melanoma (DEGs: 21) 1) MicroRNAs in cancer 2) Neutrophil extracellular trap formation 3) Viral carcinogenesis 4) Systemic lupus erythematosus 5) Adherens junct. (DEGs:42)	(DEGs:45)	(DEGs:39)
<i>miR-375</i>	1) MicroRNAs in cancer (DEGs: 31)	1) MicroRNAs in cancer (DEGs:34)	1) MicroRNAs in cancer (DEGs: 39)	1) MicroRNAs in cancer (DEGs:40)	1) MicroRNAs in cancer (DEGs:37)
<i>miR-423-3p</i>	(DEGs: 10)	(DEGs:13)	1) Fatty acid metabolism (DEGs:18)	1) Fatty acid metabolism (DEGs: 16)	1) Fatty acid metabolism (DEGs: 16)
<i>miR-615-3p</i>	1) ECM-receptor interaction 2) Focal adhesion 3) Viral carcinogenesis (DEGs:38)	1) Viral carcinogenesis (DEGs:35)	1) Focal adhesion 2) Viral carcinogenesis 3) ECM-receptor interaction 4) Cell cycle 5) Systemic lupus erythematosus 6) PI3K-Akt signaling (DEGs:64)	1) Viral carcinogenesis 2) Cell cycle (DEGs:53)	1) Viral carcinogenesis (DEGs: 59)
<i>miR-7-5p</i>	(DEGs: 43)	(DEGs: 35)	(DEGs:51)	(DEGs: 52)	(DEGs: 55)
<i>miR-9-5p</i>	1) Endocrine resistance (DEGs: 28)	(DEGs: 27)	1) Endocrine resistance (DEGs: 38)	(DEGs: 37)	1) Endocrine resistance (DEGs: 39)
<i>miR-92a-3p</i>	(DEGs:64)	(DEGs:60)	(DEGs:99)	(DEGs:105)	(DEGs: 104)
<i>miR-93-3p</i>	(DEGs: 9)	1) ECM-receptor inter. (DEGs: 10)	1) ECM-receptor inter. (DEGs: 17)	(DEGs:14)	(DEGs:15)

Table 6

Area Under Curve (AUC) and accuracy were reported as obtained from the classification of prostate cancer samples from GSE118038 using 20 miRNAs and Boruta algorithm.

GS score	selected miRNAs		Boruta miRNAs	
	AUC	Accuracy	AUC	Accuracy
7	0.829	0.814	0.655	0.771
8	0.677	0.828	0.417	0.743
9	0.692	0.871	0.579	0.814
Average	0.732 ± 0.083	0.838 ± 0.030	0.55 ± 0.121	0.776 ± 0.036

this miRNA has a role in metastasis, inhibiting proliferation, invasion, and migration of the cell by targeting PI3K/Akt, FAK and KLF4 expression [29].

miR-375 is a well-known PC-associated miRNA, as it has been proposed as a biomarker for PC diagnosis and prognosis [30,31]. Its overexpression is associated to poor overall survival [32] and to chemo-resistance [33]. Pathway analysis found its association with “microRNAs in cancer”.

Urinary exosomal *miR-92a-3p* has been found in PC subjects compared to control subjects [34]. This miRNA is able to regulate SOX4 expression and seems to be regulated by its target [35]. In PC, *miR-92a* is downregulated and its loss increases cell viability, migration, and invasion [35].

miR-25-3p part of the *miR-106b-25* cluster, was previously observed to be upregulated in prostate primary tumor and distant metastases [36]. A previous study of correlation between *miR-25-3p* and Gleason score showed a positive association [28]. The pathway analysis reported that pathway “Mucin type O-glycan biosynthesis” is regulated by *miR-25-3p*.

Elevated levels of *miR-93-3p* were reported to be predictive biomarkers for triple negative breast cancer (TNBC) suggesting it as non-invasive test. In addition, it seems to play a role in the chemoresistance in TNBC regulating Wnt/ β -catenin signaling by reducing SFRP1 [37]. Our analysis found that the pathway “Extracellular-receptor interaction” is regulated by *miR-93-3p* in G7.

miR-122-5p was reported to be a tumor suppressive gene in different cancers. However, the role in PC has not been elucidated [38,39]. Its potential mechanism of action could be the regulation of its target DUSP4. *miR-122-5p* could inhibit the migration and metastasis by downregulating DUSP4 [40]. The pathway analysis revealed that “Mucin type O-glycan biosynthesis” and “Glycosphingolipid biosynthesis - ganglio series” are regulated by *miR-122-5p* in G6 and GS7 (3 + 4).

In breast cancer *miR-183-5p* regulates a known tumor suppressor gene PTEN, promoting cell proliferation in a regulatory network that involves FOXO3a and cyclin-dependent kinase inhibitors p27 and p21 [41,42].

miR-10a-5p was found to be differentially expressed in GS 6, 7, and 8. Moreover, in our study its high expression was correlated with a poor prognosis through a survival analysis in PC samples. Previous studies showed an association between PC and *miR-10a-5p* [43,44].

In addition, in our study we identified PC Gleason score-specific differential expressed miRNAs: *miR-155-5p* was identified only in PC samples with GS 6, *miR-142-3p* and *miR-296-3p* only in PC samples with GS 9.

Although there are studies that associated *miR-155-5p* with PC, there are not in the literature studies that showed an association of this miRNA with Gleason score 6 [45–47]. *miR-142-3p* has been reported upregulated in urine samples of metastatic PC patients, but no association with GS 9 has been previously demonstrated [48]. *miR-296-3p* plays a role in tumour cell resistance to natural killer cell downregulating ICAM-1. Although no association has

been previously reported with GS 9 *miR-296-3p* promotes PC metastasis correlating it with a more aggressive phenotype [49].

In the era of high throughput data, translational informatics and in-silico studies offer great opportunities to extract knowledge from multi-omics data. In contrast to wet-lab experiments, big data such as transcriptome and epigenomic are essential for the feature selection and the accurate discovery of disease signatures. Previous studies analyzed miRNA-mRNA interactions [10,11]. However, few studies have focused on miRNAs with stronger regulatory power and miRNAs that independently regulate a gene. The strength of our study is the feature selection based on miRNA expression levels, miRNA-mRNA interactions, and network topological measures. The study of miRNA expression levels allows us to select miRNAs that are altered by the disease, and miRNA-mRNA interactions consider their functional synergism. Network topological measures (NSR and NTG) quantify the regulatory role of single miRNAs in the network.

There are some limitations in our study that should be considered in subsequent works. Firstly, genes in the miRNA-mRNA network were considered as having equal importance. However, a gene could have a more prominent role in the disease etiology, and a correlation analysis could be used to test the accuracy of the interactions [50]. Secondly, to better characterize the heterogeneity of PC, other clinical conditions could be included such as invasion and metastasis. Thirdly, a wet lab validation for future translation application should be carried out.

In conclusion, our study suggested 20 miRNAs as potential biomarkers involved in the regulation of target genes crucial for the onset and progression of PC. 8 out of 20 miRNAs (*miR-615-3p*, *miR-7-5p*, *miR-375*, *miR-92a-3p*, *miR-25-3p*, *miR-93-3p*, *miR-122-5p*, and *miR-183-5p*) seem to be common biomarkers in all Gleason score and proposed as biomarkers for PC onset. The pathways “Extracellular-receptor interaction”, “Focal adhesion”, and “microRNAs in cancer” were significantly enriched by the differentially expressed target genes of several of identified miRNAs.

miRNAs were identified as PC Gleason score-specific miRNAs: *miR-155-5p* was identified in PC samples with Gleason score 6, and *miR-142-3p* and *miR-296-3p* in PC samples with Gleason score \geq 9.

CRedit authorship contribution statement

Giulia Dal Santo: Methodology, Data curation, Writing – review & editing. **Marco Frasca:** Methodology, Data curation, Writing – review & editing. **Gloria Bertoli:** Writing – review & editing. **Isabella Castiglioni:** Writing – review & editing. **Claudia Cava:** Conceptualization, Methodology, Data curation, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7–30. <https://doi.org/10.3322/caac.21590>.
- [2] Haffner MC, Zwart W, Roudier MP, True LD, Nelson WG, Epstein JI, et al. Genomic and phenotypic heterogeneity in prostate cancer. *Nat Rev Urol.* 2021;18(2):79–92. <https://doi.org/10.1038/s41585-020-00400-w>.
- [3] Sehn JK. Prostate Cancer Pathology: Recent Updates and Controversies. *Mo Med.* 2018 Mar–Apr;115(2):151–155.
- [4] Iwata T, Sedukhina AS, Kubota M, Oonuma S, Maeda I, Yoshiike M, et al. A new bioinformatics approach identifies overexpression of GRB2 as a poor prognostic biomarker for prostate cancer. *Sci Rep.* 2021;11(1). <https://doi.org/10.1038/s41598-021-85086-9>.
- [5] Marzec J, Ross-Adams H, Pirrò S, Wang J, Zhu Y, Mao X, et al. The transcriptomic landscape of prostate cancer development and progression: an integrative analysis. *Cancers (Basel).* 2021 Jan 19;13(2):345. <https://doi.org/10.3390/cancers13020345>.
- [6] Cava C, Zoppis I, Mauri G, Ripamonti M, Gallivanone F, Salvatore C, et al. Combination of gene expression and genome copy number alteration has a prognostic value for breast cancer. *Annu Int Conf IEEE Eng Med Biol Soc.* 2013;2013:608–11. <https://doi.org/10.1109/EMBC.2013.6609573>.
- [7] Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol.* 2004;17(3):292–306. <https://doi.org/10.1038/modpathol.3800054>.
- [8] Damascelli A, Gallivanone F, Cristel G, Cava C, Interlenghi M, Esposito A, et al. Advanced imaging analysis in prostate MRI: building a radiomic signature to predict tumor aggressiveness. *Diagnostics (Basel).* 2021;11(4):594. <https://doi.org/10.3390/diagnostics11040594>.
- [9] Cava C, Bertoli G, Castiglioni I. In silico identification of drug target pathways in breast cancer subtypes using pathway cross-talk inhibition. *J Transl Med.* 2018;16(1):154. <https://doi.org/10.1186/s12967-018-1535-2>.
- [10] Cava C, Bertoli G, Castiglioni I. Portrait of Tissue-Specific Coexpression Networks of Noncoding RNAs (miRNA and lncRNA) and mRNAs in Normal Tissues. *Comput Math Methods Med.* 2019;3(2019):9029351. <https://doi.org/10.1155/2019/9029351>.
- [11] Cava C, Bertoli G, Colaprico A, Bontempi G, Mauri G, Castiglioni I. In-Silico Integration Approach to Identify a Key miRNA Regulating a Gene Network in Aggressive Prostate Cancer. *Int J Mol Sci.* 2018;19(3):910. <https://doi.org/10.3390/ijms19030910>.
- [12] Cava C, Colaprico A, Bertoli G, Bontempi G, Mauri G, Castiglioni I. How interacting pathways are regulated by miRNAs in breast cancer subtypes. *BMC Bioinform.* 2016;17(Suppl 12):348. <https://doi.org/10.1186/s12859-016-1196-1>.
- [13] Cava C, Manna I, Gambardella A, Bertoli G, Castiglioni I. Potential Role of miRNAs as Theranostic Biomarkers of Epilepsy. *Mol Ther Nucleic Acids.* 2018;7(13):275–90. <https://doi.org/10.1016/j.mtn.2018.09.008>.
- [14] Zhang W, Zang J, Jing X, Sun Z, Yan W, Yang D, et al. Identification of candidate miRNA biomarkers from miRNA regulatory network with application to prostate cancer. *J Transl Med.* 2014;12(1). <https://doi.org/10.1186/1479-5876-12-66>.
- [15] Yan W, Xu L, Sun Z, Lin Y, Zhang W, Chen J, et al. MicroRNA biomarker identification for pediatric acute myeloid leukemia based on a novel bioinformatics model. *Oncotarget.* 2015;6(28):26424–36. <https://doi.org/10.18632/oncotarget.4459>.
- [16] Lin Y, Qian F, Shen L, Chen F, Chen J, Shen B. Computer-aided biomarker discovery for precision medicine: data resources, models and applications. *Brief Bioinform.* 2019;20(3):952–75. <https://doi.org/10.1093/bib/bbx158>.
- [17] Lin Y, Wu W, Sun Z, Shen L, Shen B. MiRNA-BD: an evidence-based bioinformatics model and software tool for microRNA biomarker discovery. *RNA Biol.* 2018;15(8):1093–105. <https://doi.org/10.1080/15476286.2018.1502590>.
- [18] Lin Y, Chen F, Shen Li, Tang X, Du C, Sun Z, et al. Biomarker microRNAs for prostate cancer metastasis: screened with a network vulnerability analysis model. *J Transl Med.* 2018;16(1). <https://doi.org/10.1186/s12967-018-1506-7>.
- [19] Lin Y, Miao Z, Zhang X, Wei X, Hou J, Huang Y, et al. Identification of Key MicroRNAs and Mechanisms in Prostate Cancer Evolution Based on Biomarker Prioritization Model and Carcinogenic Survey. *Front Genet.* 2021;11. <https://doi.org/10.3389/fgene.2020.596826>.
- [20] Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data e71. *Nucleic Acids Res.* 2016;44(8):e71. <https://doi.org/10.1093/nar/gkv1507>.
- [21] Huang H-Y, Lin Y-C-D, Li J, Huang K-Y, Shrestha S, Hong H-C, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 2020. <https://doi.org/10.1093/nar/gkz896>.
- [22] Xu T, Su N, Liu L, Zhang J, Wang H, Zhang W, et al. miRBaseConverter: an R/Bioconductor package for converting and retrieving miRNA name, accession, sequence and family information in different versions of miRBase. *BMC Bioinform.* 2018;19(S19). <https://doi.org/10.1186/s12859-018-2531-5>.
- [23] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–7. <https://doi.org/10.1089/omi.2011.0118>.
- [24] Breiman L. Random Forests. *Machine Learning* 2001;45(1):5–32. <https://doi.org/10.1023/A:1010933404324>.
- [25] Kursa MB, Jankowski A, Rudnicki WR. Boruta - a system for feature selection. *Fundam Inf* 2010;101(4):271–85.
- [26] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BSVK, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia.* 2017;19(8):649–58. <https://doi.org/10.1016/j.neo.2017.05.002>.
- [27] Laursen EB, Fredsøe J, Schmidt L, Strand SH, Kristensen H, Rasmussen AKI, et al. Elevated miR-615-3p expression predicts adverse clinical outcome and promotes proliferation and migration of prostate cancer cells. *Am J Pathol.* 2019;189(12):2377–88. <https://doi.org/10.1016/j.ajpath.2019.08.007>.
- [28] Pudova EA, Krasnov GS, Nyushko KM, Kobelyatskaya AA, Savvateeva MV, Poloznikov AA, et al. miRNAs expression signature potentially associated with lymphatic dissemination in locally advanced prostate cancer. *BMC Med Genomics.* 2020;13(S8). <https://doi.org/10.1186/s12920-020-00788-9>.
- [29] Kalinowski FC, Brown RAM, Ganda C, Giles KM, Epis MR, Horsham J, et al. microRNA-7: a tumor suppressor miRNA with therapeutic potential. *Int J Biochem Cell Biol.* 2014;54:312–7. <https://doi.org/10.1016/j.biocel.2014.05.040>.
- [30] Kachakova D, Mitkova A, Popov E, Popov I, Vlahova A, Dikov T, et al. Combinations of serum prostate-specific antigen and plasma expression levels of let-7c, miR-30c, miR-141, and miR-375 as potential better diagnostic biomarkers for prostate cancer. *DNA Cell Biol.* 2015;34(3):189–200. <https://doi.org/10.1089/dna.2014.2663>.
- [31] Szczyrba J, Nolte E, Wach S, Kremmer E, Stöhr R, Hartmann A, et al. Downregulation of Sec23A protein by miRNA-375 in prostate carcinoma. *Mol Cancer Res.* 2011;9(6):791–800. <https://doi.org/10.1158/1541-7786.MCR-10-0573>.
- [32] Huang X, Yuan T, Liang M, Du M, Xia S, Dittmar R, et al. Exosomal miR-1290 and miR-375 as prognostic markers in castration-resistant prostate cancer. *Eur Urol.* 2015;67(1):33–41. <https://doi.org/10.1016/j.eururo.2014.07.035>.
- [33] Wang Y, Lieberman R, Pan J, Zhang Qi, Du M, Zhang P, et al. miR-375 induces docetaxel resistance in prostate cancer by targeting SEC23A and YAP1. *Mol Cancer.* 2016;15(1). <https://doi.org/10.1186/s12943-016-0556-9>.
- [34] Rodríguez M, Bajo-Santos C, Hessvik NP, Lorenz S, Fromm B, Berge V, et al. Identification of non-invasive miRNAs biomarkers for prostate cancer by deep sequencing analysis of urinary exosomes. *Mol Cancer.* 2017;16(1). <https://doi.org/10.1186/s12943-017-0726-4>.
- [35] Liao G, Xiong H, Tang J, Li Y, Liu Y. MicroRNA-92a Inhibits the Cell Viability and Metastasis of Prostate Cancer by Targeting SOX4 153303382095935. *Technol Cancer Res Treat.* 2020;19. <https://doi.org/10.1177/1533033820959354>.
- [36] Martens-Uzunova ES, Jalava SE, Dits NF, van Leenders GJLH, Møller S, Trapman J, et al. Diagnostic and prognostic signatures from the small non-coding RNA transcriptome in prostate cancer. *Oncogene* 2012;31(8):978–91. <https://doi.org/10.1038/onc.2011.304>.
- [37] Li H-Y, Liang J-L, Kuo Y-L, Lee H-H, Calkins MJ, Chang H-T, et al. miR-105/93-3p promotes chemoresistance and circulating miR-105/93-3p acts as a diagnostic biomarker for triple negative breast cancer. *Breast Cancer Res.* 2017;19(1). <https://doi.org/10.1186/s13058-017-0918-2>.
- [38] Xu Z, Liu G, Zhang M, Zhang Z, Jia Y, Peng Li, et al. miR-122-5p Inhibits the Proliferation, Invasion and Growth of Bile Duct Carcinoma Cells by Targeting ALDOA. *Cell Physiol Biochem.* 2018;48(6):2596–606. <https://doi.org/10.1159/000492702>.
- [39] Xu X, Gao F, Wang J, Tao L, Ye J, Ding Li, et al. MiR-122-5p inhibits cell migration and invasion in gastric cancer by down-regulating DUSP4. *Cancer Biol Ther.* 2018;19(5):427–35. <https://doi.org/10.1080/15384047.2018.1423925>.
- [40] Hu N, Tian Y, Song Y, Zang L. miR-122-5p suppresses the oncogenesis of PTC by inhibiting DUSP4 expression. *Mol Med Rep.* 2021;23(5):368. <https://doi.org/10.3892/mmr.2021.12007>.
- [41] Jin L, Luo Y, Zhao YC, Tao H. MiR-183-5p Promotes Tumor Progression of Osteosarcoma and Predicts Poor Prognosis in Patients. *Cancer Manag Res.* 2021;27(13):805–14. <https://doi.org/10.2147/CMAR.S285909>.
- [42] Sarver AL, Li L, Subramanian S. MicroRNA miR-183 functions as an oncogene by targeting the transcription factor EGR1 and promoting tumor cell migration. *Cancer Res.* 2010;70(23):9570–80. <https://doi.org/10.1158/0008-5472.CAN-10-2074>.
- [43] Worst TS, Previti C, Nitschke K, Diessl N, Gross JC, Hoffmann L, et al. miR-10a-5p and miR-29b-3p as Extracellular Vesicle-Associated Prostate Cancer Detection Markers. *Cancers (Basel).* 2019;12(1):43. <https://doi.org/10.3390/cancers12010043>.
- [44] Ku A, Fredsøe J, Sørensen KD, Borre M, Evander M, Laurell T, et al. High-throughput and automated acoustic trapping of extracellular vesicles to identify microRNAs with diagnostic potential for prostate cancer. *Front Oncol.* 2021;11. <https://doi.org/10.3389/fonc.2021.631021>.
- [45] Liao G, Ma H, Li Y, Sheng Y, Chen C. Selenium nanoparticles inhibit tumor metastasis in prostate cancer through upregulated miR-155-5p-related pathway. *Biosci Biotechnol Biochem.* 2021;85(2):287–96. <https://doi.org/10.1093/bbb/zbab089>.
- [46] Yao L-Y, Ma J, Zeng X-M, Ou yang J. MicroRNA-155-5p inhibits the invasion and migration of prostate cancer cells by targeting SPOCK1 1. *Oncol Lett.* 2020;20(6):1. <https://doi.org/10.3892/ol.2020.12215>.
- [47] Daniunaite K, Dubikaityte M, Gibas P, Bakavicius A, Rimantas Lazutka J, Ulys A, et al. Clinical significance of miRNA host gene promoter methylation in prostate cancer. *Hum Mol Genet.* 2017;26(13):2451–61. <https://doi.org/10.1093/hmg/ddx138>.
- [48] Shin S, Park YH, Jung SH, Jang SH, Kim MY, Lee JY, et al. Urinary exosome microRNA signatures as a noninvasive prognostic biomarker for prostate

- cancer. NPJ Genom Med. 2021;6(1):45. <https://doi.org/10.1038/s41525-021-00212-w>.
- [49] Liu X, Chen Q, Yan J, Wang Y, Zhu C, Chen C, et al. MiRNA-296-3p-ICAM-1 axis promotes metastasis of prostate cancer by possible enhancing survival of natural killer cell-resistant circulating tumour cells e928. Cell Death Dis. 2013;4(11):e928. <https://doi.org/10.1038/cddis.2013.458>.
- [50] Nersisyan S, Galatenko A, Galatenko V, Shkurnikov M, Tonevitsky A, Andrés-León E. miRGTf-net: Integrative miRNA-gene-TF network analysis reveals key drivers of breast cancer recurrence. PLoS ONE 2021;16(4):e0249424. <https://doi.org/10.1371/journal.pone.0249424>.