



## OPEN Role of circulating MicroRNAs in prostate cancer diagnosis and risk stratification in the MCC Spain study

Inés Gómez-Acebo<sup>1,2,3,15</sup>✉, Sara Valero-Dominguez<sup>1,2</sup>, Javier Llorca<sup>1</sup>, Jessica Alonso-Molero<sup>1,2,3</sup>, Thalía Belmonte<sup>4</sup>, Gemma Castaño-Vinyals<sup>3,5,6,7</sup>, Ana Molina-Barceló<sup>8</sup>, Rafael Marcos-Gragera<sup>3,9,10,11</sup>, Manolis Kogevas<sup>3,5,6,7</sup>, Paz Rodríguez-Cundín<sup>12</sup>, Juan Alguacil<sup>3,13</sup>, Beatriz Pérez-Gómez<sup>3,14</sup>, Marina Pollán<sup>3,14</sup> & Trinidad Dierssen-Sotos<sup>1,2,3</sup>

To identify circulating microRNAs (miRNAs) associated with prostate cancer and to develop predictive models capable of distinguishing cases from controls and stratifying patients by Gleason risk categories (low, intermediate, and high risk). This case-control study included 203 prostate cancer cases and 54 population-based controls. Serum samples were analyzed by RT-qPCR (performed at QIAGEN Genomic Services). Total RNA was extracted from 200  $\mu$ l of serum using the miRNeasy Serum/Plasma Advanced Kit and reverse-transcribed with the miRCURY LNA RT Kit. A panel of 46 candidate miRNAs was profiled, and feature selection was performed using LASSO penalization. Logistic regression models were used to estimate age- and covariate-adjusted odds ratios (ORs) with 95% confidence intervals (CIs) for the association between miRNA expression and prostate cancer risk. Predictive performance was assessed using repeated 5-fold cross-validation with bootstrap resampling (10 repetitions; 1000 resamples), and summarized using the area under the receiver operating characteristic curve (AUC) with bias-corrected 95% CIs. Fourteen miRNAs were significantly associated with prostate cancer. Notably, miR-199a-5p (OR = 1.89, 95% CI: 1.13–3.15;  $p = 0.015$ ) and miR-150-5p (OR = 0.20, 95% CI: 0.06–0.63;  $p = 0.006$ ) showed consistent differential expression across all Gleason risk categories, with miR-199a-5p overexpressed and miR-150-5p underexpressed, suggesting a potential role in disease progression. miR-145-5p, miR-182-5p, and miR-93-5p were significantly associated with prostate cancer in both the overall model and in low- and intermediate-risk strata, highlighting their potential relevance in early-stage disease. In contrast, miR-24-3p was exclusively overexpressed in high-risk prostate cancer (OR = 2.93, 95% CI: 1.43–5.98;  $p = 0.003$ ), indicating a possible link with aggressive tumor phenotypes. Predictive models demonstrated strong discriminatory performance, particularly for the low-risk group (AUC = 0.930, 95% CI: 0.882–0.979), followed by the intermediate-risk (AUC = 0.806, 95% CI: 0.728–0.883) and high-risk groups (AUC = 0.752, 95% CI: 0.658–0.848). The overall model achieved an AUC of 0.824 (95% CI: 0.756–0.892), reflecting robust performance in distinguishing cases from controls. This study identifies key circulating miRNAs associated with prostate cancer and demonstrates their potential in predictive models for risk stratification. The strongest discriminatory performance was observed in low-risk tumors (AUC = 0.930), followed by intermediate- (AUC = 0.806) and high-risk (AUC = 0.752) categories. These findings support the use of miRNAs as non-invasive biomarkers for diagnosis and personalized management of prostate cancer.

**Keywords** Prostate cancer, Gleason score, MiRNA, Diagnosis, Risk stratification

<sup>1</sup>University of Cantabria, Santander, Spain. <sup>2</sup>IDIVAL, Santander, Spain. <sup>3</sup>Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Institute of Health Carlos III, Madrid, Spain. <sup>4</sup>IUOPA, University of Oviedo and ISPA (Health Research Institute of Asturias), Oviedo, Spain. <sup>5</sup>Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain. <sup>6</sup>Department of Medicine and Life Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain. <sup>7</sup>Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain. <sup>8</sup>Cancer and Public Health Unit, Foundation for the Promotion of Health and Biomedical Research (FISABIO-Salud Pública) in the

Valencia Region, Valencia, Spain. <sup>9</sup>Epidemiology Unit and Girona Cancer Registry, Catalan Institute of Oncology, Directorate Plan of Oncology, Girona Biomedical Research Institute Dr. Josep Trueta (IDIBGI-CERCA), Girona, Spain. <sup>10</sup>Josep Carreras Leukaemia Research Institute, Girona, Spain. <sup>11</sup>Department of Medical Sciences, Medical School, University of Girona, Girona, Spain. <sup>12</sup>Preventive Medicine, Hospital Universitario Marqués de Valdecilla, Santander 39011, Spain. <sup>13</sup>Natural Resources, Health and Environment Research Centre (RENSMA), Universidad de Huelva-Campus El Carmen, Huelva, Spain. <sup>14</sup>National Centre for Epidemiology, Carlos III Institute of Health, Madrid, Spain. <sup>15</sup>Facultad de Medicina, Universidad de Cantabria, Avda. Herrera Oria s/n, Santander 39011, Spain. ✉email: ines.gomez@unican.es

## Background

In 2022, prostate cancer was the second most frequently diagnosed cancer among men worldwide, following lung cancer, and ranked as the fifth leading cause of cancer-related mortality<sup>1,2</sup>. In Spain, data from 2024 indicate that prostate cancer is also the most common cancer in men and the third leading cause of cancer-related deaths<sup>3</sup>. These statistics highlight the critical importance of early detection and diagnosis as key public health priorities<sup>4</sup>.

While the risk factors for prostate cancer are not fully understood, associations have been identified with non-modifiable factors such as age, race, family history, and specific genetic mutations<sup>5,6</sup>. Inherited genetic predisposition has gained relevance, with mutations in genes such as *BRCA1*, *BRCA2*, *HPC1*, *HPCX* and *HOXB13* being recognized as significant contributors to aggressive prostate cancer risk<sup>7–9</sup>. Additionally, certain lifestyle habits, including excessive alcohol consumption, tobacco use, sleep disturbances, and unhealthy dietary patterns (e.g., vitamin D deficiency, high fat and red meat intake, or low consumption of vegetables), have been implicated as potential risk factors<sup>1,5,6,10</sup>.

One of the primary challenges in reducing prostate cancer mortality is late-stage diagnosis and the lack of effective responses to conventional therapies<sup>11</sup>. Advanced-stage tumors often develop resistance to hormonal therapies, limiting treatment options and worsening prognosis<sup>12</sup>. Consequently, early diagnosis not only improves disease prognosis but also enables less invasive treatments, positively impacting patients' quality of life<sup>13,14</sup>.

Currently, the main screening tools for prostate cancer are the digital rectal examination (DRE) and the measurement of prostate-specific antigen (PSA) in blood<sup>13</sup>. However, PSA testing has limitations due to its low specificity, as elevated levels can result from benign conditions such as prostatic hyperplasia or prostatitis, leading to false positives<sup>14,15</sup>. Moreover, these methods often lead to overdiagnosis and overtreatment of indolent cases, causing unnecessary complications for patients<sup>16,17</sup>. As a result, less invasive alternatives such as magnetic resonance imaging and the use of biomarkers in serum and urine have been explored to enhance diagnostic accuracy<sup>15,18</sup>.

MicroRNAs (miRNAs) are small, non-coding RNA molecules, ranging from 17 to 27 nucleotides, that regulate gene expression by binding to untranslated regions of messenger RNA (3' UTR or, less frequently, 5' UTR), modulating post-transcriptional processes<sup>19,20</sup>. These molecules are involved in key biological mechanisms such as cellular proliferation, differentiation, and apoptosis and are estimated to regulate between 30% and 60% of human genes<sup>21,22</sup>. Depending on the context, miRNAs can act as tumor suppressors or oncogenes<sup>20,21</sup>.

In prostate cancer, miRNAs play a central role by regulating genes involved in tumor progression, androgen receptor signaling, and apoptotic pathways<sup>23–25</sup>. Dysregulation of miRNA expression has been associated with prostate cancer development and progression<sup>19</sup>. Moreover, these molecules can be detected in biological fluids such as plasma, urine, or serum, making them promising non-invasive biomarkers for disease diagnosis, prognosis, and treatment<sup>4,26</sup>.

Recent studies have demonstrated the utility of miRNAs in risk stratification based on Gleason categories, highlighting their ability to differentiate low-, intermediate-, and high-risk tumors, thereby supporting personalized therapeutic strategies<sup>27–29</sup>. Nonetheless, significant gaps persist in the current literature. For example, Yuan et al. emphasized the need for more detailed analyses of miRNA expression profiles across distinct prostate cancer (PCa) subtypes, using control populations matched for age, sex, and demographic background, along with the identification of differentially expressed miRNAs associated with specific clinical phenotypes<sup>29</sup>. Similarly, Martínez-González et al. underscored the importance of validating circulating miRNA signatures in larger, more diverse cohorts, using standardized methodologies and rigorous statistical frameworks<sup>28</sup>. Together, these findings highlight the need for further studies that robustly assess the clinical relevance of candidate miRNAs across well-defined risk categories.

In this context, the primary objective of the present study is to evaluate the association of 46 circulating miRNAs with prostate cancer status and their distribution across different Gleason risk groups, using plasma samples collected within a multicenter case-control study. Supplementary Table 1 provides a detailed overview

Group of analysis	Number of patients (Percentage)
Control Group	54 (21.01%)
Low Risk Group (Gleason score less than 7)	42 (16.34%)
Intermediate Risk Group (Gleason score equal to 7)	110 (42.80%)
High Risk Group (Gleason score greater than 7)	51 (19.84%)

**Table 1.** Distribution of patients by risk group based on Gleason score.

of the selected miRNAs, including their putative biological functions and previously reported associations with prostate cancer.

## Methods

### Study population

The MCC-Spain study is a multicenter case-control study that recruited 1115 incident prostate cancer cases and 1562 controls from six Spanish provinces between 2008 and 2013. Cases were men aged 27 to 85 years with histologically confirmed prostate cancer diagnosed between September 2008 and December 2012. Recruitment took place in 10 hospitals across six regions of Spain (Asturias, Barcelona, Cantabria, Huelva, Madrid and Valencia). Cases resided in the hospital catchment area for at least 6 months prior to diagnosis and were followed until 2018 to assess their vital and disease-free status. Recruitment and follow-up protocols have been described in detail elsewhere<sup>30,31</sup>.

Control subjects were men without a history of prostate cancer residing in the same hospital catchment areas. Controls were frequency-matched to cases by age (in 5-year groups) and geographic area and were randomly selected from general practitioner lists at participating primary healthcare centers. Physicians contacted controls to invite them to participate. Exclusion criteria included severe physical impairments or communication difficulties that prevented participation in the interviews.

For the present analysis, 203 prostate cancer cases and 54 population-based controls with available miRNA data were included. Written informed consent was obtained from all participants. The study protocol was approved by the ethics committees of all participating institutions, and the database was registered with the Spanish Data Protection Agency (Registration No. 2102672171). Additional details regarding ethical approval and data availability are available at <http://www.mccspain.org>.

All methods were carried out in accordance with relevant guidelines and regulations.

### Data collection

Face-to-face interviews conducted by trained personnel gathered data on potential risk factors, including age, educational level, family socioeconomic status, race, body mass index (BMI), dietary intakes (assessed using a validated food frequency questionnaire and expressed in g/day and total energy intake in kcal/day), family history of prostate cancer, smoking habits, and leisure-time physical activity.

Clinical information, including clinical-pathological stage, PSA levels, and Gleason score, was obtained from medical records. Patients were stratified into three risk groups based on Gleason score: low-risk (Gleason score < 7), intermediate-risk (Gleason score = 7), and high-risk (Gleason score > 7).

### Biological samples

Blood samples were collected from both cases and controls in the early morning, under fasting conditions. The samples were first centrifuged at 3,000×g for 20 min at 10 °C, followed by a second centrifugation at 15,000×g for 10 min at the same temperature to remove cellular debris. The resulting serum samples were stored at –80 °C until further use.

### Validation of MiRNAs based on previous studies

A total of 257 participants were randomly selected for this phase based on two criteria: blood sample collection occurred before treatment initiation (including any surgical intervention), and participants were from provinces included in the MCC-Spain study. This phase included 54 population-based controls and prostate cancer cases stratified by Gleason score. A total of 46 miRNAs were analyzed, selected from previously published signatures between 2020 and 2024<sup>19,28,32–59</sup> (see Supplementary Table 2). Differential expression analysis was conducted for all 46 miRNAs, with further selection of 14 miRNAs based on LASSO regression models.

The miRNA analysis was conducted through quantitative real-time PCR (RT-qPCR) at QIAGEN Genomic Services. Serum samples were thawed on ice and centrifuged at 3000×g for 5 min at 4 °C. From each sample, 200 µl of serum were transferred to a FluidX tube, and 60 µl of Buffer RPL (containing 1 µg of carrier RNA and an RNA spike-in template mix) were added. The mixture was vortexed for 1 min, incubated for 7 min at room temperature, and combined with 20 µl of Buffer RPP. Total RNA was extracted using the miRNeasy Serum/Plasma Advanced Kit and eluted into a final volume of 50 µl.

Next, 2 µl of total RNA were reverse-transcribed into complementary DNA (cDNA) in 10 µl reactions using the miRCURY LNA RT Kit (QIAGEN). The cDNA was diluted 50× and assayed via qPCR using the custom miRCURY LNA PCR panel and the SYBR Green miRCURY LNA master mix (QIAGEN). Negative controls without RNA templates from the RT step were included. Amplification was performed on a LightCycler<sup>®</sup> 480 Real-Time PCR System (Roche) using 384-well plates. Amplification curves were analyzed with the Roche LC software using the second derivative method to determine Cq values and analyze melting curves.

The expression levels of let-7 d-5p and let-7i-5p miRNAs, detected in all samples, were used as normalization references. The TMM normalization method<sup>60</sup> was applied, and the resulting values were log<sub>2</sub>-transformed. Additionally, a custom normalization approach was implemented using hsa-let-7 d-5p and hsa-let-7i-5p, based on the average expression levels across all detected assays ( $n = 348$  samples).

Rigorous quality control procedures were applied. Individual reactions were excluded based on the following criteria: (i) the presence of more than one melting temperature for the amplified product; (ii) a melting temperature deviating from database values; (iii) low amplification efficiency. In addition, RT-qPCR analyses were performed in duplicate (and triplicate when feasible) to ensure reliability and minimize bias.

miRNA	TMM cases*	TMM controls*	log <sub>2</sub> (FC)	Fold change*	Odds ratio (95% CI)**	P**	Odds ratio (95% CI)***	P***
miR-101-3p	1.261	1.501	-0.240	0.847	0.41(0.19–0.89)	0.024	0.43 (0.19–0.96)	0.039
miR-139-5p	-3.970	-4.154	0.184	1.136	0.71(0.48–1.05)	0.086	0.70 (0.47–1.04)	0.074
miR-141-3p	-5.716	-5.408	-0.308	0.808	0.78(0.56–1.11)	0.166	0.80 (0.56–1.14)	0.211
miR-145-5p	-1.700	-2.306	0.606	1.522	1.49(0.80–2.77)	0.208	1.53 (0.79–2.95)	0.209
<b>miR-146a-5p</b>	<b>-0.164</b>	<b>-0.087</b>	<b>-0.077</b>	<b>0.948</b>	<b>0.44(0.21–0.94)</b>	<b>0.033</b>	<b>0.47 (0.21–1.02)</b>	<b>0.057</b>
miR-150-5p	-0.258	0.316	-0.574	0.672	0.71(0.43–1.18)	0.186	0.70 (0.41–1.19)	0.186
<b>miR-182-5p</b>	<b>-6.085</b>	<b>-5.728</b>	<b>-0.357</b>	<b>0.781</b>	<b>0.66(0.44–0.97)</b>	<b>0.034</b>	<b>0.69 (0.46–1.02)</b>	<b>0.066</b>
<b>miR-199a-5p</b>	<b>-3.465</b>	<b>-4.267</b>	<b>0.802</b>	<b>1.744</b>	<b>1.89(1.13–3.15)</b>	<b>0.015</b>	<b>1.77 (1.02–3.08)</b>	<b>0.043</b>
miR-24-3p	2.253	1.805	0.448	1.364	2.10(0.72–6.12)	0.173	1.92 (0.62–5.96)	0.260
<b>miR-26b-5p</b>	<b>0.038</b>	<b>-0.128</b>	<b>0.166</b>	<b>1.122</b>	<b>2.98(1.15–7.70)</b>	<b>0.024</b>	<b>3.30 (1.23–8.90)</b>	<b>0.018</b>
<b>miR-93-5p</b>	<b>1.524</b>	<b>1.421</b>	<b>0.103</b>	<b>1.074</b>	<b>4.37(1.25–15.31)</b>	<b>0.021</b>	<b>4.28 (1.12–16.41)</b>	<b>0.034</b>

**Table 2.** Multivariate analysis: MiRNA signature in cases vs. controls. Notes: • TMM (Trimmed Mean of M-values): Method used for miRNA expression normalization. • Fold change (FC): Ratio of expression between cases and controls (FC > 1 indicates overexpression; FC < 1 indicates underexpression). log<sub>2</sub>(FC) is the base-2 logarithmic transformation of FC. • Crude (unadjusted) vs. Adjusted (\*\* and \*\*\*): The unadjusted model does not control for any confounders; the second model is adjusted for all miRNAs included in the final set; the third model also adjusts for age, BMI, educational level, socioeconomic status, and smoking.

## Statistical analysis

The following analysis was carried out for the overall model (all cases vs. controls) and for submodels stratified by Gleason classification (low, intermediate and high).

dCq values from controls were compared with each case group using Student's t-test without further adjustments. Results are reported as log Fold Change (logFC), p-values, and the False Discovery Rate (FDR), calculated using the Benjamini-Hochberg method<sup>61</sup>. A positive logFC indicates that the miRNA is overexpressed in cases, while a negative logFC reflects underexpression. The greater the absolute logFC value, the larger the difference between cases and controls.

To assess the predictive potential of the miRNAs, LASSO penalized logistic regression models were constructed<sup>62</sup>. This technique allows for the inclusion of a large number of predictors while generating parsimonious models by shrinking the coefficients of less relevant variables to zero. The analysis started with an initial model that included all 46 miRNAs analyzed during the validation phase. The regularization parameter ( $\lambda$ ) was selected via 10-fold cross-validation. Subsequently, the LASSO-selected miRNAs were included in a logistic regression model. This approach was adopted to capture potential synergistic relationships among miRNAs, which might not individually reach significance but could collectively contribute to the model's discriminatory capacity. The full list of miRNAs included in each model is detailed in Supplementary Table 1.

Results from logistic regression models are presented as odds ratios (OR) with 95% confidence intervals (CI). The area under the ROC curve (AUROC) and its bias-corrected 95% CI were obtained via bootstrap. Each model was then re-estimated, adjusting for age, BMI, educational level, socioeconomic status, and smoking status. Race was not included, as nearly all participants were Caucasian. Family history was excluded due to collinearity.

To assess model performance, we implemented a repeated cross-validation procedure consisting of 10 independent repetitions of 5-fold cross-validation. In each repetition, an AUROC was obtained and its bias-corrected 95% CI was derived from internal bootstrap estimation. As a result, 10 AUROC values were generated per model. Cross-validation results are presented as the mean of these 10 AUROCs along with their 95% CI.

All statistical analyses were performed using Stata 18/SE (StataCorp, College Station, TX, USA), while all figures were generated using R version 4.4.2 with the ggplot2 and cowplot libraries.

## Results

### Characteristics of the study population

Supplementary Table 3 summarizes the characteristics of the 203 prostate cancer cases and 54 controls included in the study. Significant differences were observed in key variables. Controls were overrepresented in Cantabria and Asturias, while most cases were concentrated in Barcelona ( $p < 0.001$ ). The mean age of controls was slightly higher than that of cases ( $67.39 \pm 9.23$  vs.  $65.18 \pm 6.65$  years,  $p = 0.048$ ). Additionally, a higher percentage of controls belonged to higher socioeconomic levels (53.70% vs. 30.54%,  $p = 0.006$ ). Family history of prostate cancer was more frequent among cases (23.47% vs. 11.54%,  $p = 0.060$ ). No significant differences were observed in BMI, red meat consumption, physical activity, or smoking habits between the groups.

Table 1 presents the classification of prostate cancer cases by Gleason risk categories. Most cases were classified as intermediate-risk (42.80%) or high-risk (19.84%) based on the Gleason score. The remaining 16.34% of cases fell into the low-risk category (Gleason < 7).

### MiRNA signatures in prostate cancer

Table 2 presents the multivariate analysis comparing miRNA expression between prostate cancer cases and controls. Among the miRNAs associated with an increased risk of prostate cancer, miR-93-5p exhibited the

miRNA	TMM cases*	TMM controls*	log <sub>2</sub> (FC)	Fold change*	Odds ratio (95% CI)**	P**	Odds ratio (95% CI)***	P***
miR-145-5p	-1.557	-2.306	0.749	1.681	3.50(1.30–9.43)	0.013	8.40 (2.03–34.78)	0.003
miR-146a-5p	-0.243	-0.087	-0.156	0.898	0.04(0.01–0.28)	0.001	0.01 (0.00–0.19)	0.002
miR-150-5p	-0.457	0.316	-0.773	0.585	0.20(0.06–0.63)	0.006	0.12 (0.03–0.58)	0.008
<b>miR-182-5p</b>	<b>-6.359</b>	<b>-5.728</b>	<b>-0.631</b>	<b>0.646</b>	<b>0.26(0.11–0.62)</b>	<b>0.002</b>	<b>0.15 (0.04–0.58)</b>	<b>0.006</b>
miR-199a-5p	-3.419	-4.267	0.848	1.800	2.07(0.80–5.37)	0.135	1.96 (0.54–7.12)	0.305
miR-27a-3p	1.895	1.485	0.410	1.329	2.42(0.48–12.33)	0.287	2.79 (0.36–21.38)	0.324
miR-331-3p	-4.057	-4.362	0.305	1.235	2.69(0.68–10.58)	0.158	7.46 (1.08–51.72)	0.042
miR-423-3p	-1.090	-1.517	0.427	1.344	2.03(0.36–11.44)	0.423	2.06 (0.20–21.70)	0.546
<b>miR-93-5p</b>	<b>1.656</b>	<b>1.421</b>	<b>0.235</b>	<b>1.177</b>	<b>11.00(1.84–65.90)</b>	<b>0.009</b>	<b>16.21 (1.79–146.44)</b>	<b>0.013</b>

**Table 3.** MiRNA signature: controls vs. Low-Risk group (Gleason < 7). Notes: • TMM (Trimmed Mean of M-values): Method used for miRNA expression normalization. • Fold change (FC): Ratio of expression between cases and controls (FC > 1 indicates overexpression; FC < 1 indicates underexpression). log<sub>2</sub>(FC) is the base-2 logarithmic transformation of FC. • Crude (unadjusted) vs. Adjusted (\*\* and \*\*\*): The unadjusted model does not control for any confounders; the second model is adjusted for all miRNAs included in the final set; the third model also adjusts for age, BMI, educational level, socioeconomic status, and smoking.

strongest association in the overall model (OR = 4.37, 95% CI: 1.25–15.31;  $p = 0.021$ ), and even stronger in the low-risk group (OR = 11.00, 95% CI: 1.84–65.90;  $p = 0.009$ ), followed by miR-26b-5p (OR = 2.98, 95% CI: 1.15–7.70;  $p = 0.024$ ) and miR-199a-5p (OR = 1.89, 95% CI: 1.13–3.15;  $p = 0.015$ ). These miRNAs demonstrated significant overexpression in cases compared to controls. Conversely, miR-101-3p (OR = 0.41, 95% CI: 0.19–0.89;  $p = 0.024$ ), miR-146a-5p (OR = 0.44, 95% CI: 0.21–0.94;  $p = 0.033$ ), and miR-182-5p (OR = 0.66, 95% CI: 0.44–0.97;  $p = 0.034$ ) were significantly associated with a lower risk, suggesting their underexpression in prostate cancer cases compared to controls.

In addition to the model adjusted for all miRNAs included in the final panel, a more comprehensive model was developed that further adjusted for age, BMI, educational level, socioeconomic status, and smoking status (see the last two columns in Table 2). Overall, these additional adjustments did not substantially alter the direction or significance of the associations, indicating that the identified miRNA signatures were robust even after accounting for key demographic and lifestyle factors. For instance, miR-93-5p remained the strongest risk-associated miRNA, whereas miR-101-3p, miR-146a-5p, and miR-182-5p continued to exhibit protective effects. These findings suggest that the observed miRNA–disease relationships are not merely confounded by age, BMI, educational level, socioeconomic status, or smoking behavior, thereby reinforcing the potential utility of these miRNAs as independent biomarkers for prostate cancer risk.

In the low-risk group (Gleason < 7, Table 3), the fully adjusted logistic regression model—incorporating all miRNAs in the final panel as well as age, BMI, educational level, socioeconomic status, and smoking status—identified miR-93-5p as having the strongest association with increased risk (OR = 11.00, 95% CI: 1.84–65.90;  $p = 0.009$ ), followed by miR-145-5p (OR = 3.50, 95% CI: 1.30–9.43;  $p = 0.013$ ). Conversely, miR-146a-5p (OR = 0.04, 95% CI: 0.01–0.28;  $p = 0.001$ ), miR-150-5p (OR = 0.20, 95% CI: 0.06–0.63;  $p = 0.006$ ), and miR-182-5p (OR = 0.26, 95% CI: 0.11–0.62;  $p = 0.002$ ) demonstrated significant protective associations, suggesting underexpression in cases compared to controls. miR-199a-5p showed a non-significant association with increased risk (OR = 2.07, 95% CI: 0.80–5.37;  $p = 0.135$ ). These findings underscore the robustness of the miRNA panel in distinguishing low-risk prostate cancer, even after adjusting for key demographic and lifestyle factors.

In the intermediate-risk group (Gleason = 7, Table 4), the logistic regression model—incorporating all miRNAs in the final panel as well as age, BMI, educational level, socioeconomic status, and smoking status—revealed that miR-199a-5p remained significantly associated with an increased risk of prostate cancer (OR = 2.12, 95% CI: 1.19–3.77;  $p = 0.011$ ), alongside miR-26b-5p (OR = 3.31, 95% CI: 1.09–10.10;  $p = 0.035$ ). Conversely, miR-139-5p exhibited a significant protective association (OR = 0.57, 95% CI: 0.36–0.90;  $p = 0.017$ ), despite a modest fold change (1.043), suggesting that small differences in expression may still be biologically relevant. These findings underscore the continued importance of these miRNAs in distinguishing intermediate-risk disease, even after controlling for key demographic and lifestyle factors.

In the high-risk group (Gleason > 7, Table 5), the logistic regression model incorporating all miRNAs in the final panel as well as age, BMI, educational level, socioeconomic status, and smoking status identified miR-24-3p as the only miRNA significantly associated with an increased risk of prostate cancer (OR = 2.93, 95% CI: 1.43–5.98;  $p = 0.003$ ). This marker displayed notably higher expression in cases compared to controls. In contrast, miR-199a-5p and miR-150-5p showed non-significant trends toward increased and decreased risk, respectively. These results suggest that miR-24-3p may play a particularly important role in the pathogenesis of more aggressive prostate cancer.

Figure 1 presents a Venn diagram illustrating the distribution of miRNAs across Gleason risk categories (Low, Intermediate, and High Risk). In the Low-Risk group (Gleason < 7), miR-146a-5p, miR-27a-3p, miR-331-3p, and miR-423-3p are uniquely identified, suggesting their potential specificity for early-stage prostate cancer. In contrast, miR-139-5p, miR-141-3p, miR-26b-5p, and miR-96-5p are exclusive to the Intermediate-Risk group (Gleason = 7), indicating their involvement in this stage of the disease. Several miRNAs are shared between risk categories, highlighting their broader role in prostate cancer progression. Notably, miR-145-5p, miR-182-5p,

miRNA	TMM cases*	TMM controls*	log <sub>2</sub> (FC)	Fold change*	Odds ratio (95% CI)**	P**	Odds ratio (95% CI)***	P***
miR-139-5p	-4.093	-4.154	0.061	1.043	0.57(0.36–0.90)	0.017	0.53 (0.33–0.85)	0.009
miR-141-3p	-5.777	-5.408	-0.369	0.774	0.69(0.47–1.02)	0.060	0.67 (0.45–1.00)	0.052
miR-145-5p	-1.758	-2.306	0.548	1.462	1.71(0.92–3.17)	0.089	<b>1.96 (1.02–3.78)</b>	<b>0.043</b>
miR-150-5p	-0.233	0.316	-0.549	0.683	0.66(0.40–1.11)	0.115	0.60 (0.35–1.05)	0.073
miR-182-5p	-6.107	-5.728	-0.379	0.769	0.72(0.48–1.08)	0.114	0.73 (0.48–1.12)	0.148
<b>miR-199a-5p</b>	<b>-3.431</b>	<b>-4.267</b>	<b>0.836</b>	<b>1.785</b>	<b>2.12(1.19–3.77)</b>	<b>0.011</b>	<b>1.93 (1.04–3.59)</b>	<b>0.038</b>
<b>miR-26b-5p</b>	<b>0.064</b>	<b>-0.128</b>	<b>0.192</b>	<b>1.142</b>	<b>3.31(1.09–10.10)</b>	<b>0.035</b>	<b>3.84 (1.14–12.90)</b>	<b>0.030</b>
miR-93-5p	1.471	1.421	0.050	1.035	2.37(0.93–6.05)	0.072	2.27 (0.82–6.26)	0.114
miR-96-5p	-6.400	-5.761	-0.639	0.642	0.71(0.49–1.02)	0.065	0.73 (0.49–1.08)	0.112

**Table 4.** MiRNA signature: controls vs. Intermediate-Risk group (Gleason = 7). Notes: • TMM (Trimmed Mean of M-values): Method used for miRNA expression normalization. • Fold change (FC): Ratio of expression between cases and controls (FC > 1 indicates overexpression; FC < 1 indicates underexpression). log<sub>2</sub>(FC) is the base-2 logarithmic transformation of FC. • Crude (unadjusted) vs. Adjusted (\*\* and \*\*\*): The unadjusted model does not control for any confounders; the second model is adjusted for all miRNAs included in the final set; the third model also adjusts for age, BMI, educational level, socioeconomic status, and smoking.

miRNA	TMM cases*	TMM controls*	log <sub>2</sub> (FC)	Fold change*	Odds ratio (95% CI)**	P**	Odds ratio (95% CI)***	P***
miR-150-5p	-0.148	0.316	-0.464	0.725	0.64 (0.39–1.04)	0.074	0.68 (0.40–1.17)	0.164
miR-199a-5p	-3.574	-4.267	0.693	1.617	1.31 (0.82–2.11)	0.255	1.29 (0.78–2.11)	0.321
<b>miR-24-3p</b>	<b>2.317</b>	<b>1.805</b>	<b>0.512</b>	<b>1.426</b>	<b>2.93 (1.43–5.98)</b>	<b>0.003</b>	<b>2.73 (1.29–5.77)</b>	<b>0.009</b>

**Table 5.** MiRNA signature: controls vs. High-Risk group (Gleason > 7). Notes: • TMM (Trimmed Mean of M-values): Method used for miRNA expression normalization. • Fold change (FC): Ratio of expression between cases and controls (FC > 1 indicates overexpression; FC < 1 indicates underexpression). log<sub>2</sub>(FC) is the base-2 logarithmic transformation of FC. • Crude (unadjusted) vs. Adjusted (\*\* and \*\*\*): The unadjusted model does not control for any confounders; the second model is adjusted for all miRNAs included in the final set; the third model also adjusts for age, BMI, educational level, socioeconomic status, and smoking.

and miR-93-5p are common to both the Low- and Intermediate-Risk groups, suggesting a function in early to intermediate disease transition. Meanwhile, miR-150-5p and miR-199a-5p are the only miRNAs present in all three Gleason categories, reinforcing their relevance as potential biomarkers across different stages of disease progression. In the High-Risk group (Gleason > 7), miR-24-3p is uniquely found, suggesting a potential role in advanced or aggressive prostate cancer.

### Predictive performance of miRNA-based models

Using LASSO regression, a set of 15 miRNAs was identified as key contributors across four predictive models, highlighting their potential as non-invasive biomarkers for prostate cancer stratification based on Gleason risk categories (Fig. 2). Each model demonstrated discriminative power through the area under the curve (AUC), varying according to risk classification:

#### General model (AUC = 0.824, 95% CI: 0.756–0.892)

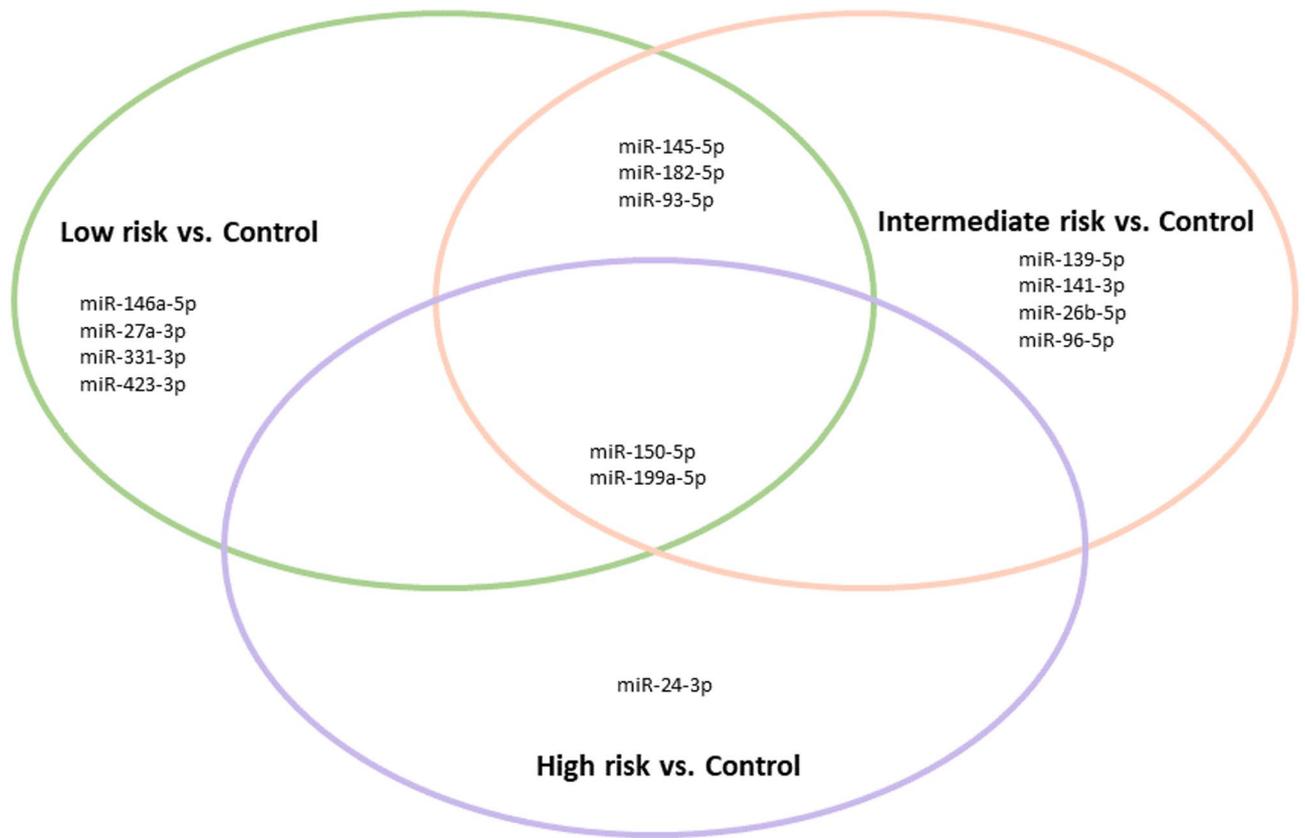
This model provided an overall risk assessment by incorporating 10 miRNAs, including miR-101-3p, miR-139-5p, miR-141-3p, miR-145-5p, miR-146a-5p, miR-150-5p, miR-182-5p, miR-199a-5p, miR-24-3p, and miR-26b-5p. These findings support the potential of miRNA-based models in distinguishing prostate cancer across different Gleason risk groups.

#### Low-Risk model (AUC = 0.930, 95% CI: 0.882–0.979)

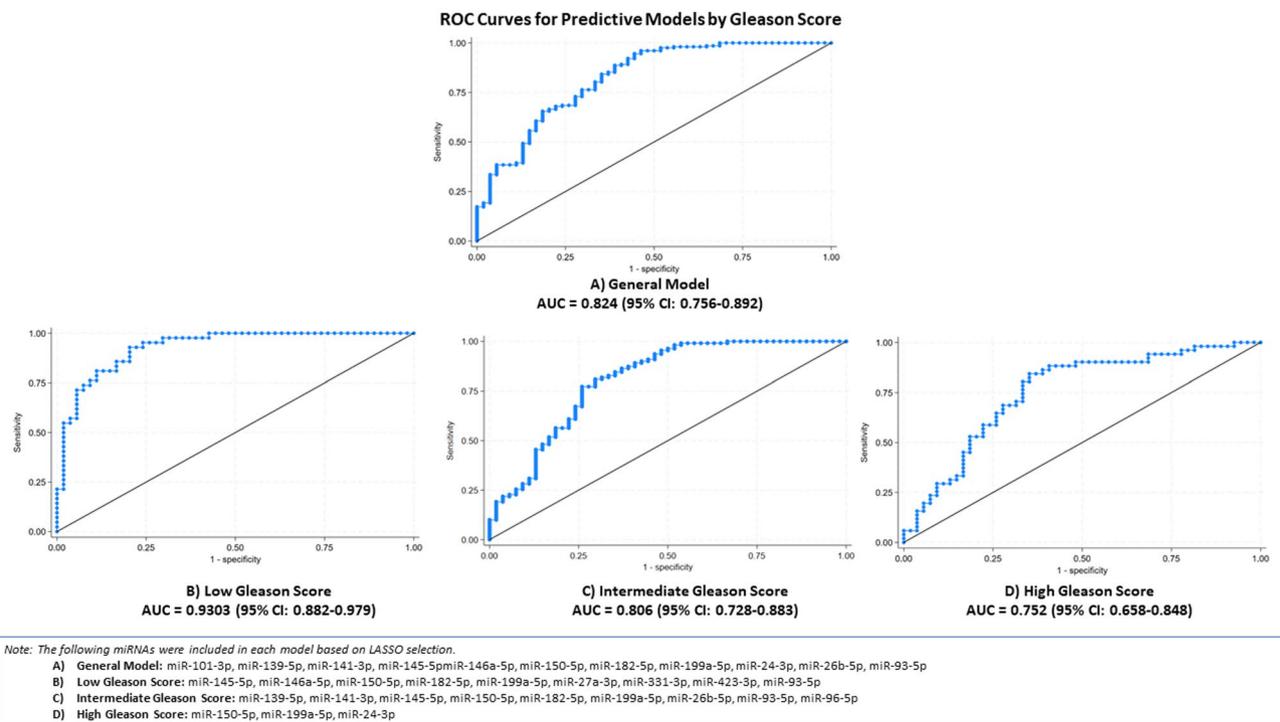
Achieving the highest predictive accuracy, this model included 10 miRNAs: miR-145-5p, miR-146a-5p, miR-150-5p, miR-182-5p, miR-199a-5p, miR-27a-3p, miR-331-3p, miR-423-3p, and miR-93-5p. The results suggest that low-risk tumors exhibit less biological heterogeneity, making them more predictable based on miRNA signatures.

#### Intermediate-Risk model (AUC = 0.806, 95% CI: 0.728–0.883)

This model included 7 miRNAs—miR-139-5p, miR-141-3p, miR-145-5p, miR-150-5p, miR-182-5p, miR-199a-5p, and miR-26b-5p—highlighting their relevance in Gleason 7 tumors, where disease progression is more variable.



**Fig. 1.** Venn Diagram Depicting miRNA Distribution in Gleason Risk Categories.



**Fig. 2.** Predictive Performance of miRNA-Based Models Across Gleason Risk Groups. (A) General model comparing controls vs. all cases. (B) Comparing controls vs. low-risk cases (Gleason <7). (C) Comparing controls vs. intermediate-risk cases (Gleason =7). (D) Comparing controls vs. high-risk cases (Gleason >7).

### High-Risk model (AUC = 0.752, 95% CI: 0.658–0.848)

With a more moderate AUC, this model comprised three miRNAs: miR-150-5p, miR-199a-5p, and miR-24-3p. These markers may aid in distinguishing aggressive Gleason >7 tumors, where heterogeneity and molecular complexity increase.

Complementary to these findings, a stratified 10-fold cross-validation logistic regression analysis was performed to directly compare controls with each risk group. The model discriminating between controls and low-risk prostate cancer (Gleason <7) incorporated nine miRNAs (hsamir1455p, hsamir146a5p, hsamir1505p, hsamir1825p, hsamir199a5p, hsamir27a3p, hsamir3313p, hsamir4233p, and hsamir935p) and achieved a mean AUC of 0.8445 (95% CI, 0.75–0.91). For the comparison between controls and intermediate-risk patients (Gleason =7), a different set of nine miRNAs (hsamir1395p, hsamir1413p, hsamir1455p, hsamir1505p, hsamir1825p, hsamir199a5p, hsamir26b5p, hsamir935p, and hsamir965p) yielded a mean AUC of 0.773 (95% CI, 0.6633–0.8557). In the model distinguishing controls from high-risk patients (Gleason >7), three miRNAs (hsamir1505p, hsamir199a5p, and hsamir243p) were used, resulting in a mean AUC of 0.7846 (95% CI, 0.6351–0.8396). Collectively, these results indicate that specific miRNA signatures confer moderate to excellent discriminatory power across varying Gleason risk categories, with particularly robust performance in identifying low-risk prostate cancer. The convergence of findings from both LASSO-based models and cross-validated logistic regression analyses supports the potential utility of these miRNA panels as non-invasive biomarkers for prostate cancer risk stratification, warranting further validation in independent cohorts.

## Discussion

This study highlights the potential of circulating miRNAs as non-invasive biomarkers for the diagnosis and risk stratification of prostate cancer.

### Predictive models

Four predictive models developed in this study identified 15 circulating miRNAs associated with prostate cancer, each reflecting distinct molecular signatures across Gleason risk categories. The low-risk model demonstrated the highest performance (AUC =0.930), followed by the intermediate-risk (AUC =0.806) and the high-risk disease (AUC =0.752), underscoring the models' ability to differentiate tumors by aggressiveness. Notably, the cross-validated AUCs were only slightly lower than those from the initial LASSO models, supporting the robustness and internal validity of the predictive the models (see Supplementary Table 4).

miR-150-5p and miR-199a-5p emerged as robust biomarkers, consistently present in all models, while miR-93-5p was associated with low- and intermediate-risk categories, while miR-24-3p was uniquely identified in the high-riskgroup. Protective miRNAs, such as miR-146a-5p and miR-150-5p, were especially prominent in general and the low risk models, suggesting a relevant role in less aggressive disease stages.

Although early miRNA research tended to focus on individual biomarkers, recent studies support the utility of multi-miRNA signatures. For instance, Gandellini et al. (2021) identified a three-miRNA panel (miR-511-5p, miR-598-3p, and miR-199a-5p) that enhanced risk prediction in prostate cancer patients under active surveillance<sup>39</sup>. Bergez-Hernández et al. (2022) developed predictive models combining multiple miRNAs (miR-145-5p and miR-146a-5p) with clinical variables, achieving high diagnostic accuracy for distinguishing prostate cancer from benign prostatic disease<sup>40</sup>. These findings align with our results, particularly for miR-199a-5p, miR-146a-5p, and miR-145-5p, which were also included in our predictive models. Such overlaps underscore the relevance of these miRNAs as robust biomarkers and highlight the potential of multi-miRNA signatures to refine risk stratification, improve diagnostic accuracy, and guide clinical decision-making in prostate cancer management.

### MiRNAs overexpressed in prostate cancer patients

Through both global and Gleason-stratified analyses, we identified miRNAs associated with increased risk.

Among cancer-associated miRNAs, miR-199a-5p emerged as a robust marker, showing significantly higher expression in cases across the overall, low-risk, and intermediate-risk groups. In this study, miR-199a-5p was significantly overexpressed in cases, with levels 73% higher in cases compared to controls, establishing it as a robust marker for prostate cancer risk. These findings align with reports suggesting a pro-oncogenic role for miR-199a-5p, particularly in advanced disease. For instance, Alwhaibi et al. (2022) demonstrated that miR-199a-5p expression increases with Akt1 suppression, promoting epithelial-to-mesenchymal transition through the TGFβ pathway. Similarly, high miR-199a-5p levels have been associated with higher Gleason scores, progression to metastasis, and aggressive prostate cancer phenotypes<sup>63</sup>. However, other studies suggest an opposing role. Hu et al. (2023) identified low miR-199a-5p levels in prostate cancer patients, linking its downregulation to failure in inhibiting HIF-1α, a key regulator of angiogenesis and tumor survival<sup>64,65</sup>. This discrepancy may stem from differences in patient populations, disease stages, or tissue-specific expression patterns. Biologically, miR-199a-5p regulates key pathways such as PI3 K/AKT and TGFβ, which are critical for cell proliferation, angiogenesis, and tumor progression<sup>63,66</sup>. The dual nature of miR-199a-5p—acting as an oncogene in some contexts and a tumor suppressor in others—underscores the complexity of its role in prostate cancer and highlights the necessity for further research to clarify its functional mechanisms and potential as a therapeutic target.

Numerous studies have identified miR-145-5p as a tumor suppressor in prostate cancer. Its overexpression has been linked to the regulation of critical pathways such as WIP1, SOX2, and PI3 K/AKT, which are involved in cell cycle control and differentiation, limiting tumor proliferation and invasion<sup>67</sup>. Additionally, miR-145-5p overexpression induces apoptosis and reduces the migratory capacity of tumor cells through the modulation of genes like *TNFSF10* and the inhibition of epithelial-to-mesenchymal transition<sup>68</sup>. Conversely, low levels of miR-145-5p in advanced prostate cancer tissues (Gleason ≥8) have been associated with negative regulation of the p53 pathway, contributing to tumor progression and uncontrolled cell proliferation<sup>69,70</sup>. Its downregulation

has also been shown to increase *MYO6* expression, promoting epithelial-to-mesenchymal transition and further exacerbating tumor aggressiveness<sup>71</sup>. Together, these studies position miR-145-5p as a key tumor suppressor in advanced stages of prostate cancer. In contrast, we observed a significant overexpression of miR-145-5p in low-risk prostate cancer (Gleason <7) compared to controls. This result, which diverges from the existing literature, raises the possibility of a context-dependent dual function of miR-145-5p. In early tumor stages, miR-145-5p overexpression may promote initial tumor cell proliferation through mechanisms not yet fully understood<sup>72</sup>. Epigenetic alterations, such as localized hypomethylation of the miR-145 promoter, or interactions with key pathways like PI3 K/AKT or WIP1, might contribute to this effect. Importantly, while other studies have highlighted miR-145-5p's role in suppressing prostate cancer cell motility, they have not directly addressed its influence on early tumor development or risk stratification. These discrepancies underscore the complexity of miR-145-5p's function and highlight the need for further studies to explore its role in early-stage prostate cancer and its potential as a biomarker.

In our study, miR-93-5p was overexpressed in prostate cancer cases compared to controls, in both low- and intermediate-risk groups. These findings underscore the clinical significance of miR-93-5p as a potential biomarker for prostate cancer, particularly in stratifying patients at low and intermediate risk. Its overexpression may reflect its involvement in early tumorigenic processes and its potential to promote disease progression, even in patients with seemingly lower-grade disease. Consistent with our findings, previous studies have identified miR-93-5p as being overexpressed in prostate cancer. Martínez-González et al. (2021) observed elevated levels of this miRNA in both tissue and plasma samples, correlating its overexpression with increased disease aggressiveness, particularly in tumors with Gleason scores >7<sup>28</sup>. Other studies have associated miR-93-5p with reduced overall survival and higher recurrence rates within five years post-treatment, emphasizing its role in poor prognostic outcomes<sup>42,73,74</sup>. Biologically, miR-93-5p exerts its effects by regulating key genes in the TGF- $\beta$  signaling pathway, which promotes cell migration and inhibits apoptosis<sup>74,75</sup>. It also targets tumor suppressor genes such as *FRMD6*, as well as genes involved in cell proliferation and DNA replication, including *TCF7*, *E2 F2*, *RRM2*, and *PKMYT1*, thereby driving cellular proliferation and metastasis<sup>75,76</sup>. This dual role in promoting tumor growth and resistance to apoptosis highlights miR-93-5p's critical contribution to prostate cancer progression and its potential as a therapeutic target.

Finally, in our study, miR-24-3p was significantly overexpressed in the high-risk group (Gleason >7) compared to controls. This finding suggests an important role for miR-24-3p in aggressive prostate cancer phenotypes, reinforcing its potential as a biomarker for high-risk disease. Supporting our results, miR-24-3p has been reported to act as an oncogene in prostate cancer by promoting cell proliferation, migration, and invasion. High levels of this miRNA downregulate the tumor suppressor gene *SOCS6*, driving tumor progression<sup>77</sup>. Additionally, Zhang et al. (2021) observed miR-24-3p overexpression in serum samples from prostate cancer patients, suggesting its potential as a non-invasive biomarker for aggressive disease<sup>42</sup>. Conversely, some studies have reported miR-24-3p underexpression in tissue samples from prostate cancer patients, associating its loss with disruptions in cell growth, invasion, and apoptosis through the regulation of targets such as *HOXA11-AS*<sup>78</sup>. Similar findings were noted by Lin et al. (2020) and Zhang et al. (2021), emphasizing its tumor-suppressive role in certain tissue environments<sup>42,71</sup>. These conflicting observations between tissue and serum highlight the complex and context-dependent behavior of miR-24-3p in prostate cancer. Biologically, miR-24-3p's dual role may depend on the tumor microenvironment and the biological matrix in which it is analyzed. Its overexpression in serum and aggressive tumors supports its oncogenic potential, whereas its underexpression in tissue may reflect compensatory mechanisms or tumor-specific heterogeneity. These findings underscore the need for further research to clarify its role across different biological contexts and explore its utility as a diagnostic and prognostic biomarker in high-risk prostate cancer.

In our study, miR-26b-5p was found to be overexpressed in prostate cancer patients, correlating with an intermediate-risk classification (Gleason score 7) compared to controls. These findings contrast with most previous studies, which have reported lower miR-26b-5p expression in prostate cancer and in patients with higher disease risk<sup>43,79</sup>. miR-26b-5p has been described as a negative regulator of prostate cell malignancy<sup>37</sup>, and its downregulation has been associated with increased tumor progression. To date, we have not identified studies in the literature reporting miR-26b-5p overexpression in prostate cancer similar to our findings. This discrepancy may be attributed to differences in tumor heterogeneity, sample types, or methodological approaches for miRNA quantification. Additionally, the expression of miR-26b-5p may be influenced by specific clinicopathological factors, including tumor microenvironment dynamics and inflammatory responses. Notably, miR-26b-5p has been observed to interact with multiple signaling pathways and genes related to prostate cancer, including p53, ERK/MAPK, TGF- $\beta$ , IGF1R, PTEN, and RB1<sup>79,80</sup>. Nevertheless, our results highlight the need for further studies to investigate the role of miR-26b-5p in specific patient subgroups and its potential function in the biology of intermediate-risk tumors.

### MiRNAs underexpressed in prostate cancer patients

Our results indicate that higher miR-101-3p expression is associated with a lower risk of prostate cancer, suggesting its potential role as a biomarker for disease detection. Consistently, a study by Duca RB et al. (2021) reported miR-101-3p overexpression in blood samples from prostate cancer patients, with even higher levels observed in those with metastatic disease<sup>81</sup>. However, conflicting evidence exists, as other studies have reported miR-101-3p downregulation in prostate cancer tissue and tumor cells<sup>82,83</sup>, with its reduced expression being associated with poorer prognosis<sup>82</sup>. This discrepancy may stem from differences in sample origin (blood vs. tissue), tumor heterogeneity, or methodological variations in miRNA quantification. Notably, miR-101-3p has been implicated in inflammatory pathways by promoting the upregulation of IL-6, IL-8, TNF- $\alpha$ , CD80, and CD86<sup>84</sup>. Additionally, it regulates key proteins such as TRB1 and Ezh2, as well as genes like SUB1, which are involved in the MAPK signaling pathway<sup>83–85</sup>. These findings highlight the complex regulatory role of miR-101-

3p in prostate cancer and underscore the need for further studies to elucidate its function in different biological contexts and disease stages.

The present study provides further evidence supporting the tumor suppressor role of miR-146a-5p in prostate cancer, particularly in early stages of the disease. Our results demonstrated protective associations in both the global analysis and the low-risk group (Gleason < 7). Specifically, lower expression levels of miR-146a-5p were observed in cases compared to controls, emphasizing its potential as a biomarker for early detection and risk stratification in prostate cancer. Our findings are consistent with those of Fredsøe et al. (2020), who also reported low plasma levels of miR-146a-5p in prostate cancer patients<sup>54</sup>. This alignment strengthens the hypothesis of miR-146a-5p's involvement in prostate tumor suppression. However, discrepancies across studies—such as Puhka et al. (2022), who observed overexpression of miR-146a-5p in urinary extracellular vesicles from patients with more aggressive tumors (Gleason  $\geq$  8)<sup>86</sup> suggest a context-dependent role for this miRNA. These differences underline the importance of considering the biological matrix and clinical context when interpreting miRNA levels. At the molecular level, miR-146a-5p regulates apoptosis, inflammation, and oxidative stress through genes like *EGFR*, *TRAF6*, and *IRAK1*, which are implicated in tumor progression<sup>58,59,87,88</sup>.

Similarly, miR-182-5p was underexpressed in cases—particularly in the low-risk group—compared to controls, consistent with reports linking it to less aggressive tumor phenotypes. For example, Baumann et al. (2019) observed that although miR-182 was elevated in prostate cancer epithelium compared to benign tissue, higher levels were inversely associated with recurrence, supporting its potential protective role in low-risk prostate cancer<sup>89</sup>. However, other studies report conflicting findings. Souza et al. (2022) and Bidarra et al. (2019) identified elevated miR-182-5p expression in both tissue and plasma samples of prostate cancer patients, with higher levels associated with advanced pathological stages and metastasis<sup>46,90</sup>. These results suggest a pro-oncogenic role in more aggressive phenotypes. Kurul et al. (2019) also highlighted the dual role of miR-182-5p, noting its potential as a diagnostic biomarker but emphasizing its variable function depending on disease stage<sup>91</sup>. Such findings underline the complex and potentially bifunctional nature of miR-182-5p, which may act as a tumor suppressor in early stages while promoting progression in advanced disease.

miR-150-5p showed underexpression in low-risk cases. This finding aligns with previous research by Yu et al. (2018) and Păunescu et al. (2019), which demonstrated that miR-150-5p suppresses prostate cancer cell proliferation and invasion by targeting *MAP3 K12*<sup>92</sup> and is downregulated in both plasma and tissue samples of prostate cancer patients<sup>55</sup>. However, other studies report conflicting results. Liu et al. (2015) associated high miR-150 expression with poorer overall and disease-free survival, while Zhao et al. (2016) demonstrated that miR-150 promotes prostate cancer cell proliferation and invasion by acting on p53<sup>93,94</sup>. These contradictions highlight the complex role of miR-150-5p, suggesting that its function may vary depending on the disease stage or cellular context. Biologically, miR-150-5p is involved in regulating key pathways like apoptosis and cellular signaling, potentially explaining its dual functionality in prostate cancer progression.

miR-139-5p was underexpressed in the intermediate-risk group, consistent with its established role as a tumor suppressor. Previous research demonstrated that miR-139-5p inhibits cell proliferation, migration, and epithelial-mesenchymal transition by targeting *SOX5*<sup>95</sup>. Furthermore, Nam et al. (2020) found that miR-139-5p induces autophagy by regulating mTOR and Beclin-1, leading to apoptosis in prostate cancer cells<sup>96</sup>. These mechanisms emphasize the potential therapeutic and prognostic value of miR-139-5p in managing prostate cancer. Biologically, its ability to modulate critical pathways associated with tumor progression reinforces its significance as a promising biomarker for disease management.

### Strengths and limitations

This study has several strengths. Firstly, the inclusion of prostate cancer cases from 10 hospitals spanning six regions of Spain provides substantial clinical and biological variability, thereby enhancing the generalizability of the findings. This geographic and institutional diversity supports the robustness of the results and aligns with the need for representative cohorts in biomarker research. Second, the use of penalized regression (LASSO) for model development is a key methodological strength. LASSO allows for the selection of parsimonious models by minimizing overfitting and alpha error while avoiding researcher subjectivity in variable selection. This method has been widely validated as more reliable than traditional approaches like stepwise regression, particularly in high-dimensional datasets such as miRNA profiles. Third, the study's design, including Gleason score stratification and the inclusion of cases across a range of risk categories (low, intermediate, and high), ensures comprehensive analyses. The stratified approach allows for the identification of miRNAs associated with distinct disease stages, a critical element in understanding the biological complexity of prostate cancer. Finally, the discriminatory power of the models, as demonstrated by the area under the ROC curve (AUC), reflects a strong predictive capacity, particularly for low- and intermediate-risk cases. The AUC values are highest for low-risk cases (0.930), followed by intermediate-risk cases (0.806) and high-risk cases (0.752). This trend underscores the models' ability to effectively differentiate early or less aggressive stages of prostate cancer, which is crucial for timely intervention and improved patient outcomes.

Despite the strengths of this study, there are some limitations. First, the selection of miRNAs was exclusively based on previously published studies, which ensures the inclusion of well-established biomarkers but may omit emerging candidates from exploratory analyses or high-throughput screenings; however, this approach enhances reproducibility, a critical element in biomarker research. Second, although the sample size in certain subgroups—particularly those at high risk—is small, and only 54 healthy controls were included, this scenario is common in miRNA studies due to the high costs and financial constraints involved, as noted by Zhang et al. (2022)<sup>33</sup>. While a larger cohort would increase statistical power, the chosen sample size aligns with typical practice in this field. Third, our models were developed and validated using the same dataset (with internal stratified cross-validation), which introduces the possibility of overfitting. Although  $\lambda$  was selected globally and not within a nested CV framework, we attempted to mitigate overfitting through rigorous validation.

However, external validation in independent cohorts is needed to confirm the models' performance. Fourth, even though our models demonstrate strong performance, their real-world effectiveness in clinical practice (for instance, improving sensitivity and specificity in screening programs) requires further investigation. Finally, conducting the study in a single country limits generalizability to other populations with different ethnic and cultural backgrounds, highlighting the need to validate these findings in diverse contexts. These limitations do not diminish the robustness of our results; rather, they indicate clear opportunities for future research to further strengthen the evidence base.

## Conclusion

This study highlights the potential of circulating miRNAs as non-invasive biomarkers for prostate cancer diagnosis and risk stratification. Among the 15 identified miRNAs, miR-199a-5p and miR-150-5p stood out as consistently overexpressed differentially expressed across all Gleason risk categories, reinforcing their potential role in prostate cancer progression. Additionally, miR-145-5p, miR-182-5p, and miR-93-5p were present in both the general model (cases vs. controls) and in low- and intermediate-risk categories, suggesting their relevance in early-stage disease. In contrast, miR-24-3p was uniquely associated with high-risk prostate cancer and the general model, indicating its potential involvement in aggressive tumor phenotypes.

The predictive models demonstrated strong discriminatory power, particularly for low- and intermediate-risk cases, underscoring their ability to differentiate early-stage tumors, which is crucial for timely intervention.

## Data availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. For data requests, please contact Inés Gómez-Acebo at ines.gomez@unican.es.

Received: 27 February 2025; Accepted: 6 May 2025

Published online: 20 May 2025

## References

- Bergengren, O. et al. 2022 Update on prostate Cancer epidemiology and risk Factors-A systematic review. *Eur. Urol.* **84**, 191–206 (2023).
- WHO. G. Cancer Today. (2022). <https://gco.iarc.who.int/today/>
- AECC, A. española contra el cáncer. Dimensiones del cáncer | AECC Observatorio. (2025). <https://observatorio.contraelcancer.es/explora/dimensiones-del-cancer>
- Mugoni, V., Ciani, Y., Nardella, C. & Demichelis, F. Circulating RNAs in prostate cancer patients. *Cancer Lett.* **524**, 57–69 (2022).
- Gómez-Acebo, I. et al. Risk model for prostate Cancer using environmental and genetic factors in the Spanish Multi-Case-Control (MCC) study. *Sci. Rep.* **7**, 8994 (2017).
- Papantoniou, K. et al. Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. *Int. J. Cancer.* **137**, 1147–1157 (2015).
- Almeeri, M. N. E., Awies, M. & Constantinou, C. Prostate cancer, pathophysiology and recent developments in management: A narrative review. *Curr. Oncol. Rep.* **26**, 1511–1519 (2024).
- Pritchard, C. C. et al. Inherited DNA-Repair gene mutations in men with metastatic prostate Cancer. *N Engl. J. Med.* **375**, 443–453 (2016).
- Urabe, F. et al. Prostate cancer and liquid biopsies: clinical applications and challenges. *Int. J. Urol. Off J. Jpn Urol. Assoc.* **31**, 617–626 (2024).
- Romaguera, D. et al. Consumption of ultra-processed foods and drinks and colorectal, breast, and prostate cancer. *Clin. Nutr.* **40**, 1537–1545 (2021).
- Sekhoacha, M. et al. Prostate Cancer review: genetics, diagnosis, treatment options, and alternative approaches. *Mol. Basel Switz.* **27**, 5730 (2022).
- Wade, C. A. & Kyprianou, N. Profiling prostate Cancer therapeutic resistance. *Int. J. Mol. Sci.* **19**, 904 (2018).
- Berenguer, C. V., Pereira, F., Câmara, J. S. & Pereira, J. A. M. Underlying features of prostate Cancer-Statistics, risk factors, and emerging methods for its diagnosis. *Curr. Oncol. Tor. Ont.* **30**, 2300–2321 (2023).
- Eickelschulte, S. et al. Biomarkers for the detection and risk stratification of aggressive prostate Cancer. *Cancers* **14**, 6094 (2022).
- Boehm, B. E., York, M. E., Petrovics, G., Kohaar, I. & Chesnut, G. T. Biomarkers of aggressive prostate Cancer at diagnosis. *Int. J. Mol. Sci.* **24**, 2185 (2023).
- Farha, M. W. & Salami, S. S. Biomarkers for prostate cancer detection and risk stratification. *Ther. Adv. Urol.* **14**, 17562872221103988 (2022).
- Light, A., Elhage, O., Marconi, L. & Dasgupta, P. Prostate cancer screening: where are we now? *BJU Int.* **123**, 916–917 (2019).
- Cereser, L., Evangelista, L., Giannarini, G. & Girometti, R. Prostate MRI and PSMA-PET in the primary diagnosis of prostate Cancer. *Diagn. Basel Switz.* **13**, 2697 (2023).
- Azani, A. et al. MicroRNAs as biomarkers for early diagnosis, targeting and prognosis of prostate cancer. *Pathol. - Res. Pract.* **248**, 154618 (2023).
- Gupta, J. et al. Angiogenesis and prostate cancer: MicroRNAs comes into view. *Pathol. Res. Pract.* **248**, 154591 (2023).
- Khan, M. M., Sharma, V. & Serajuddin, M. Emerging role of MiRNA in prostate cancer: A future era of diagnostic and therapeutics. *Gene* **888**, 147761 (2023).
- Menon, A., Abd-Aziz, N., Khalid, K., Poh, C. L. & Naidu, R. MiRNA: A promising therapeutic target in Cancer. *Int. J. Mol. Sci.* **23**, 11502 (2022).
- Bilal, M., Javaid, A., Amjad, F., Youssif, T. A. & Afzal, S. An overview of prostate cancer (PCa) diagnosis: potential role of MiRNAs. *Transl Oncol.* **26**, 101542 (2022).
- Fabris, L. et al. The potential of MicroRNAs as prostate Cancer biomarkers. *Eur. Urol.* **70**, 312–322 (2016).
- Sreekumar, A. & Saini, S. Role of MicroRNAs in neuroendocrine prostate Cancer. *Non-Coding RNA.* **8**, 25 (2022).
- Rafikova, G., Gilyazova, I., Enikeeva, K., Pavlov, V. & Kzhyshkowska, J. Prostate cancer: genetics, epigenetics and the need for immunological biomarkers. *Int. J. Mol. Sci.* **24**, 12797 (2023).
- Hoey, C. et al. Circulating MiRNAs as non-invasive biomarkers to predict aggressive prostate cancer after radical prostatectomy. *J. Transl Med.* **17**, 173 (2019).
- Martínez-González, L. J. et al. Identification of MicroRNAs as viable aggressiveness biomarkers for prostate Cancer. *Biomedicine* **9**, 646 (2021).
- Yuan, F., Hu, Y., Lei, Y. & Jin, L. Recent progress in MicroRNA research for prostate cancer. *Discov Oncol.* **15**, 480 (2024).

30. Castaño-Vinyals, G. et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *Gac Sanit.* **29**, 308–315 (2015).
31. Alonso-Molero, J. et al. Cohort profile: the MCC-Spain follow-up on colorectal, breast and prostate cancers: study design and initial results. *BMJ Open.* **9**, e031904 (2019).
32. Seputra, K. P., Purnomo, B. B., Susianti, H., Kalim, H. & Purnomo, A. F. miRNA-21 as reliable serum diagnostic biomarker candidate for metastatic progressive prostate cancer: Meta-analysis approach. *Med. Arch. Sarajevo Bosnia Herzeg.* **75**, 347–350 (2021).
33. Zhang, W. T., Zhang, G. X., Zhao, R. Z. & Gao, S. S. The potential diagnostic accuracy of Circulating MicroRNAs for prostate cancer: A meta-analysis. *Actas Urol. Esp.* **46**, 138–149 (2022).
34. Stafford, M. Y. C., Willoughby, C. E., Walsh, C. P. & McKenna, D. J. Prognostic value of miR-21 for prostate cancer: a systematic review and meta-analysis. *Biosci. Rep.* **42**, BSR20211972 (2022).
35. Nitusca, D. et al. Diagnostic value of microRNA-375 as future biomarker for prostate Cancer detection: A Meta-Analysis. *Med. Kaunas Lith.* **58**, 529 (2022).
36. Greco, F. et al. The potential role of MicroRNAs as biomarkers in benign prostatic hyperplasia: A systematic review and Meta-analysis. *Eur. Urol. Focus.* **5**, 497–507 (2019).
37. Rana, S., Valbuena, G. N., Curry, E., Bevan, C. L. & Keun, H. C. MicroRNAs as biomarkers for prostate cancer prognosis: a systematic review and a systematic reanalysis of public data. *Br. J. Cancer.* **126**, 502–513 (2022).
38. Karadag, A. et al. Identification of MiRNA signatures and their therapeutic potentials in prostate cancer. *Mol. Biol. Rep.* **48**, 5531–5539 (2021).
39. Gandellini, P. et al. Prediction of grade reclassification of prostate Cancer patients on active surveillance through the combination of a Three-miRNA signature and selected clinical variables. *Cancers* **13**, 2433 (2021).
40. Bergez-Hernández, F. et al. Expression analysis of MiRNAs and their potential role as biomarkers for prostate Cancer detection. *Am. J. Mens Health.* **16**, 15579883221120989 (2022).
41. Schitcu, V. H. et al. MicroRNA dysregulation in prostate Cancer. *Pharmacogenomics Pers. Med.* **15**, 177–193 (2022).
42. Zhang, S. et al. MicroRNA panel in serum reveals novel diagnostic biomarkers for prostate cancer. *PeerJ* **9**, e11441 (2021).
43. Giglio, S. et al. A preliminary study of micro-RNAs as minimally invasive biomarkers for the diagnosis of prostate cancer patients. *J. Exp. Clin. Cancer Res. CR.* **40**, 79 (2021).
44. Koh, Y. et al. Urine Cell-Free MicroRNAs in localized prostate Cancer patients. *Cancers* **14**, 2388 (2022).
45. Hasanoğlu, S. et al. Investigating differential MiRNA expression profiling using serum and urine specimens for detecting potential biomarkers for early prostate cancer diagnosis. *Turk. J. Med. Sci.* **51**, 1764–1774 (2021).
46. Bidarra, D. et al. Circulating MicroRNAs as biomarkers for prostate Cancer detection and metastasis development prediction. *Front. Oncol.* **9**, 900 (2019).
47. Lyu, J. et al. Discovery and validation of serum MicroRNAs as early diagnostic biomarkers for prostate Cancer in Chinese population. *BioMed. Res. Int.* **9306803** 2019 (2019).
48. Ibrahim, N. H., Abdellateif, M. S., Kassem, S. H. A., Abd El Salam, M. A. & El Gammal, M. M. Diagnostic significance of miR-21, miR-141, miR-18a and miR-221 as novel biomarkers in prostate cancer among Egyptian patients. *Andrologia* **51**, e13384 (2019).
49. Damodaran, M. et al. Differentially expressed miR-20, miR-21, miR-100, miR-125a and miR-146a as a potential biomarker for prostate cancer. *Mol. Biol. Rep.* **48**, 3349–3356 (2021).
50. Abramovic, I. et al. MiR-182-5p and miR-375-3p have higher performance than PSA in discriminating prostate Cancer from benign prostate hyperplasia. *Cancers* **13**, 2068 (2021).
51. Canatan, D. et al. Use of MicroRNAs as biomarkers in the early diagnosis of prostate cancer. *Acta Bio-Medica Atenei Parm.* **93**, e2022089 (2022).
52. Rajendiran, S. et al. MicroRNA-940 as a potential serum biomarker for prostate Cancer. *Front. Oncol.* **11**, 628094 (2021).
53. Cai, S. et al. Single-molecule amplification-free multiplexed detection of Circulating MicroRNA cancer biomarkers from serum. *Nat. Commun.* **12**, 3515 (2021).
54. Fredsøe, J. et al. Profiling of Circulating MicroRNAs in prostate Cancer reveals diagnostic biomarker potential. *Diagn. Basel Switz.* **10**, 188 (2020).
55. Paunescu, I. A. et al. Biomarker potential of plasma MicroRNA-150-5p in prostate Cancer. *Med. Kaunas Lith.* **55**, 564 (2019).
56. Urabe, F. et al. Large-scale Circulating MicroRNA profiling for the liquid biopsy of prostate Cancer. *Clin. Cancer Res. Off J. Am. Assoc. Cancer Res.* **25**, 3016–3025 (2019).
57. Cruz-Burgos, M. et al. Unraveling the role of EV-Derived miR-150-5p in prostate Cancer metastasis and its association with High-Grade Gleason scores: implications for diagnosis. *Cancers* **15**, 4148 (2023).
58. Kookli, K. et al. Role of microRNA-146a in cancer development by regulating apoptosis. *Pathol. - Res. Pract.* **254**, 155050 (2024).
59. Dobrijević, Z. et al. Diagnostic properties of miR-146a-5p from liquid biopsies in prostate cancer: A meta-analysis. *Pathol. Res. Pract.* **262**, 155522 (2024).
60. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
61. Benjamini, Y. & Hochberg, Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* **25**, 60–83 (2000).
62. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).
63. Alwhaibi, A. et al. Regulation of Let-7a-5p and miR-199a-5p expression by Akt1 modulates prostate Cancer Epithelial-to-Mesenchymal transition via the transforming growth Factor-β pathway. *Cancers* **14**, 1625 (2022).
64. Hu, S. et al. miR-199a/214 cluster enhances prostate cancer sensitiveness to nimotuzumab via targeting TBL1XR1. *Kaohsiung J. Med. Sci.* **39**, 1178–1189 (2023).
65. Zhong, J. et al. Downregulation of miR-199a-5p promotes prostate adeno-carcinoma progression through loss of its Inhibition of HIF-1α. *Oncotarget* **8**, 83523–83538 (2017).
66. Tseng, J. C. et al. ROR2 suppresses metastasis of prostate cancer via regulation of miR-199a-5p-PIAS3-AKT2 signaling axis. *Cell. Death Dis.* **11**, 376 (2020).
67. Luo, B. et al. microRNA-145-5p inhibits prostate cancer bone metastatic by modulating the epithelial-mesenchymal transition. *Front. Oncol.* **12**, 988794 (2022).
68. Sun, J., Deng, L. & Gong, Y. MiR-145-5p Inhibits the Invasion of Prostate Cancer and Induces Apoptosis by Inhibiting WIP1. *J. Oncol.* **4412705** (2021). (2021).
69. Huang, Z. G. et al. MiRNA-145-5p expression and prospective molecular mechanisms in the metastasis of prostate cancer. *IET Syst. Biol.* **15**, 1–13 (2021).
70. Ji, S. et al. miR-145-5p inhibits neuroendocrine differentiation and tumor growth by regulating the SOX11/MYCN Axis in prostate cancer. *Front. Genet.* **13**, 790621 (2022).
71. Lin, Y. et al. Identification of key MicroRNAs and mechanisms in prostate Cancer evolution based on biomarker prioritization model and carcinogenic survey. *Front. Genet.* **11**, 596826 (2020).
72. Manvati, S. et al. miR-145 supports cancer cell survival and shows association with DDR genes, methylation pattern, and epithelial to mesenchymal transition. *Cancer Cell. Int.* **19**, 230 (2019).
73. Wang, C. et al. Increased expression of microRNA-93 correlates with progression and prognosis of prostate cancer. *Med. (Baltim).* **99**, e18432 (2020).

74. Yang, Y., Jia, B., Zhao, X., Wang, Y. & Ye, W. miR-93-5p May be an important oncogene in prostate cancer by bioinformatics analysis. *J. Cell. Biochem.* **120**, 10463–10483 (2019).
75. Jafari, N. et al. Novel plasma exosome biomarkers for prostate cancer progression in co-morbid metabolic disease. *Adv. Cancer Biol. - Metastasis.* **6**, 100073 (2022).
76. Sun, M. et al. Activation of the HNRNPA2B1/miR-93-5p/FRMD6 axis facilitates prostate cancer progression in an m6A-dependent manner. *J. Cancer.* **14**, 1242–1256 (2023).
77. Lin, Y. et al. miR-24-3p stimulates migration, invasion and proliferation of prostate cancer cells by targeting suppressor of cytokine signaling 6. *Int. J. Clin. Exp. Pathol.* **11**, 1803–1810 (2018).
78. Cheng, Y. et al. LncRNA HOXA11-AS promotes cell growth by sponging miR-24-3p to regulate JPT1 in prostate cancer. *Eur. Rev. Med. Pharmacol. Sci.* **25**, 4668–4677 (2021).
79. Cochetti, G. et al. Different levels of serum MicroRNAs in prostate cancer and benign prostatic hyperplasia: evaluation of potential diagnostic and prognostic role. *OncoTargets Ther.* **9**, 7545–7553 (2016).
80. Lin, Y., Qi, X., Chen, J. & Shen, B. Multivariate competing endogenous RNA network characterization for cancer MicroRNA biomarker discovery: a novel bioinformatics model with application to prostate cancer metastasis. *Precis Clin. Med.* **5**, pbac001 (2022).
81. Duca, R. B. et al. MiR-19b-3p and miR-101-3p as potential biomarkers for prostate cancer diagnosis and prognosis. *Am. J. Cancer Res.* **11**, 2802–2820 (2021).
82. Gu, Z. et al. Inhibition of MicroRNA miR-101-3p on prostate cancer progression by regulating Cullin 4B (CUL4B) and PI3K/AKT/mTOR signaling pathways. *Bioengineered* **12**, 4719–4735 (2021).
83. Lin, Y. et al. Biomarker MicroRNAs for prostate cancer metastasis: screened with a network vulnerability analysis model. *J. Transl Med.* **16**, 134 (2018).
84. Niespolo, C. et al. Tribbles-1 expression and its function to control inflammatory cytokines, including Interleukin-8 levels are regulated by MiRNAs in macrophages and prostate Cancer cells. *Front. Immunol.* **11**, 574046 (2020).
85. Qiu, X. et al. Targeting Ezh2 could overcome docetaxel resistance in prostate cancer cells. *BMC Cancer.* **19**, 27 (2019).
86. Puhka, M. et al. Exploration of extracellular vesicle MiRNAs, targeted mRNAs and pathways in prostate cancer: relation to disease status and progression. *Cancers* **14**, 532 (2022).
87. Huang, Y. et al. Upregulation of miR-146a by YY1 depletion correlates with delayed progression of prostate cancer. *Int. J. Oncol.* **50**, 421–431 (2017).
88. Xu, B. et al. Hsa-miR-146a-5p modulates androgen-independent prostate cancer cells apoptosis by targeting ROCK1. *Prostate* **75**, 1896–1903 (2015).
89. Baumann, B. et al. Association of high miR-182 levels with Low-Risk prostate Cancer. *Am. J. Pathol.* **189**, 911–923 (2019).
90. Souza, M. F. et al. MiR-182-5p modulates prostate Cancer aggressive phenotypes by targeting EMT associated pathways. *Biomolecules* **12**, 187 (2022).
91. Kurul, N. O. et al. The association of let-7c, miR-21, miR-145, miR-182, and miR-221 with clinicopathologic parameters of prostate cancer in patients diagnosed with low-risk disease. *Prostate* **79**, 1125–1132 (2019).
92. Yu, J., Feng, Y., Wang, Y. & An, R. Aryl hydrocarbon receptor enhances the expression of miR-150-5p to suppress in prostate cancer progression by regulating MAP3K12. *Arch. Biochem. Biophys.* **654**, 47–54 (2018).
93. Dezhong, L. et al. miR-150 is a factor of survival in prostate cancer patients. *J. BUON Off J. Balk. Union Oncol.* **20**, 173–179 (2015).
94. Zhao, Y., Zhu, Y. & Song, J. MiR-150 promotes the cell invasion of prostate cancer cells by directly regulating the expression of p53. <https://www.ingentaconnect.com/contentone/govi/pharmaz/2016/00000071/00000009/art00008> (2016). <https://doi.org/10.1691/ph.2016.6645>
95. Yang, B. et al. Downregulation of miR-139-5p promotes prostate cancer progression through regulation of SOX5. *Biomed. Pharmacother Biomedicine Pharmacother.* **109**, 2128–2135 (2019).
96. Nam, R. K. et al. MicroRNA-139 is a predictor of prostate cancer recurrence and inhibits growth and migration of prostate cancer cells through cell cycle arrest and targeting IGF1R and AXL. *Prostate* **79**, 1422–1438 (2019).

## Acknowledgements

Some samples included in this study were provided by the Basque Biobank [www.biobancovasco.org](http://www.biobancovasco.org) and the Biobanco La Fe (B.0000723) and they were processed following standard operation procedures with the appropriate approval of the Ethical and Scientific Committees.

## Author contributions

IGA and JL were responsible for data acquisition and curation. IGA and JL secured funding for the study. IGA and JL supervised the research. IGA and SVD drafted the first version of the manuscript. All authors contributed to the study design and conception, reviewed previous versions of the manuscript, and approved the final version.

## Funding

This study was partially funded by the “Accion Transversal del Cancer,” approved on the Spanish Ministry Council on the 11 October 2007, Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PS09/00773, PS09/01286, PS09/01903, PS09/02078, PS09/01662, PI11/01889, PI11/02213, PI12/00488, PI12/01270, PI12/00715, PI14/01219, PI14/0613, PI17/01388 and PI18/00171), Fundación Marqués de Valdecilla (API 10/09), the ICGC International Cancer Genome Consortium CLL [The ICGC CLL-Genome Project was funded by Spanish Ministerio de Economía y Competitividad (MINECO) through the Instituto de Salud Carlos III (IS-CIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)], the Junta de Castilla y León (LE22 A10-2), the Consejería de Salud of the Junta de Andalucía (PI-0571-2009, PI-0306-2011, and salud201200057018 tra), the Conselleria de Sanitat of the Generalitat Valenciana (AP\_061/10), the Recercaixa (2010 ACUP 00310), the Regional Government of the Basque Country, the Consejería de Sanidad de la Región de Murcia, by the European Commission grants FOOD-CT-2006-036224-HIWATE, the Spanish Association Against Cancer (AECC) Scientific Foundation, by the Catalan Government—Agency for Management of University and Research Grants (AGAUR) grants 2017SGR723 and 2014SGR850, the Fundación Caja de Ahorros de Asturias, and the University of Oviedo. ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the “Centro de Excelencia Severo Ochoa 2019–2023” Program (CEX2018-000806-S) and support from the Generalitat de Catalunya through the CERCA Program. AP-C was supported by the MINECO (Ministry of Economy in Spain) Grant no. PRE2019-089038, fellowship.

## Declarations

### Competing interests

The authors declare no competing interests.

### Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by CEIC 2018.280 and 2008/3123/I. The patients/participants provided their written informed consent to participate in this study.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-01373-9>.

**Correspondence** and requests for materials should be addressed to I.G.-A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025