

RESEARCH

Open Access



# Impact of piR\_004530 reactivation in lung cancer: implications for recurrence and survival of lung squamous cell carcinoma patients

Yangyi He<sup>1,2</sup>, Melissa Acosta-Plasencia<sup>1</sup>, David Sánchez-Lorente<sup>3,4,5</sup>, Nuria Viñolas<sup>4,5,6</sup>, Daniel Martínez<sup>4,5,7</sup>, Tania Díaz<sup>1</sup>, Antonio Altuna-Coy<sup>1</sup>, Risha Na<sup>1</sup>, Yi Liu<sup>1</sup>, Marc Boada<sup>3,4,5</sup>, Angela Guirao<sup>3,4,5</sup>, Laureano Molins<sup>3,4,5</sup>, Ramón M. Marrades<sup>4,5,8,9</sup> and Alfons Navarro<sup>1,4,5\*</sup>

## Abstract

**Background** PIWI-interacting RNAs (piRNAs) are germline-characteristic small noncoding RNAs whose reactivation has been recently observed during carcinogenesis because of the germline reactivation program, which can be considered a hallmark of cancer. We evaluated the prognostic impact of hsa\_piR\_004530 reactivation in non-small cell lung cancer (NSCLC) and studied its functional role in cell lines and organoids.

**Methods** A total of 243 NSCLC resected patients were analyzed. Hsa\_piR\_004530 expression was quantified in tumor ( $n = 243$ ) and normal tissue ( $n = 31$ ) using qRT-PCR. Tumor recurrence, disease-free survival (DFS), and overall survival (OS) were used as clinical endpoints in survival analysis. Cox regression models were generated, and decision curve analysis for assessment of clinical benefit on disease prognosis was used. The effect of hsa\_piR\_004530 overexpression was assessed in NSCLC cell lines by cell migration, invasion, proliferation, apoptosis, and colony formation analysis. Patient-derived organoids from SCC patients were generated, and cell viability was evaluated after piRNA overexpression.

**Results** Hsa\_piR\_004530 became reactivated on the tumor compared to normal tissue and was overexpressed in SCC patients. The prognosis analysis in the whole cohort showed that patients with high levels of hsa\_piR\_004530 had shorter DFS and OS. However, the subanalysis by histology revealed that the true prognostic impact of hsa\_piR\_004530 was specifically on SCC patients. These results became validated in an additional cohort. SCC patients with high hsa\_piR\_004530 had a higher relapse rate and shorter DFS and OS. Hsa\_piR\_004530 emerged as an independent prognostic factor in the multivariate analysis. Since the SCC patients who received adjuvant treatment had the best postsurgical prognosis, we focused on the role of hsa\_piR\_004530 on non-treated patients, which allowed us to identify a high-risk group where the piRNA ameliorated patients' risk stratification, highlighting superior clinical benefit in postsurgical relapse prediction. Hsa\_piR\_004530 overexpression was associated with increased migration, invasion, and colony formation. Moreover, the overexpression induced stem cell gene activation and correlated with higher size spheroid formation. In patient-derived organoids, the overexpression boosted organoid cell viability.

**Conclusions** Hsa\_piR\_004530 became reactivated in NSCLC, where it played a role as an oncogene, and its higher levels correlated with disease recurrence and shorter survival in SCC patients.

\*Correspondence:  
Alfons Navarro  
anavarroponz@ub.edu  
Full list of author information is available at the end of the article



© The Author(s) 2025. **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/>.

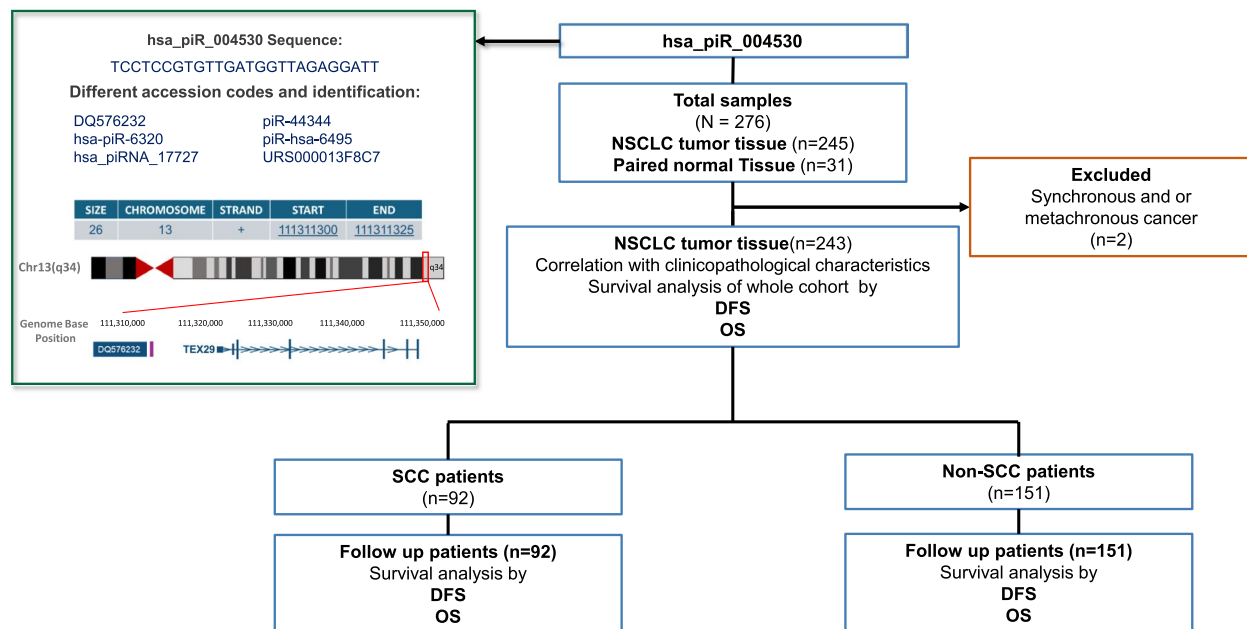
**Keywords** NSCLC, Lung squamous cell carcinoma, PIWI-interacting RNAs, piRNAs, Recurrence, Hsa\_piR\_004530, TEX29

## Background

Despite advancements in modern surgical techniques, targeted therapies, and immunotherapy, lung cancer remains one of the malignancies with the poorest prognosis, making it the leading cause of cancer-related deaths [1, 2]. In 2022, 2.5 million lung cancer cases were diagnosed, and 1.8 million lung cancer-related deaths were reported worldwide [3]. Moreover, only around 16% of patients diagnosed with lung cancer are still alive after 5 years [4, 5]. Most patients (85%) are diagnosed with non-small cell lung cancer (NSCLC). When NSCLC patients are clinically diagnosed with nonmetastatic disease, complete surgical resection remains the primary curative approach, especially for early stages (I–II) [6]. However, recurrence after surgery can develop in one-third of resected patients [7] despite the administration of adjuvant or neoadjuvant treatments [8]. At the histological level, the two most frequent histological subtypes of NSCLC are adenocarcinoma (ADK, 45%) and squamous cell carcinoma (SCC, 20%) [9, 10]. The behavior and treatment of both subtypes are quite different. Historically, NSCLC has been treated according to a “uniform” strategy without taking into account the histological subtype, but recent advances in sequencing technology and an improved biological understanding have enhanced a new era of targeted therapy for ADK, initially by the implementation of therapeutic strategies for patients harboring mutations in molecular targets such as epidermal growth factor receptor (*EGFR*) or anaplastic lymphoma kinase (*ALK*) [11–13] and more recently for other target genes such as hepatocyte growth factor receptor (*MET*) or Kirsten rat sarcoma viral oncogene homolog (*KRAS*) between others [14, 15]. Unfortunately, SCC continues to have a poor overall prognosis due to the lack of targeted therapies and the scarcity of reliable prognostic markers [16]. SCC is a histological subtype mostly associated with male smoking patients [9] that often arises in the central part of the lungs, especially near the bronchial tubes [17]. At the molecular level, SCC is markedly different from ADK, lacking mutations in *EGFR* or *ALK* but exhibiting enrichment in *TP53* mutations in most patients. The identification and targeting of driver mutations in patients with lung SCC have exhibited a delay [18, 19], and the effectiveness of immunotherapy can be variable [20]. Accordingly, treatment of SCC remains previously stagnant, often involving surgical resection and platinum-based adjuvant therapy in early stages and requiring further work to identify reliable prognostic markers.

Small noncoding RNAs have been recognized as reliable prognostic markers since their expression is usually deregulated in cancer and correlates with patient outcomes [21, 22]. Most of the past studies have focused on microRNAs, leaving other groups such as PIWI-interacting RNAs (piRNAs) understudied and poorly analyzed as prognostic biomarkers [23]. The main reason that initially discouraged studies on piRNAs was the fact that they were described first as germline-specific RNAs [24–27]. Moreover, supporting this, the biogenesis of piRNAs was mediated by PIWI proteins, an Argonaute subfamily of proteins whose expression was mostly limited to the germline cells [28–31]. However, it has been recently reported that both PIWI proteins and piRNAs can be detected in tumor cells, including NSCLC [32, 33]. Their deregulation is more marked in dedifferentiated cells, which aims to postulate that they are involved in the regulation of cancer stem cells [34, 35]. These piRNAs, with a length of 26–32 nt in humans, were associated mainly with the regulation of genome stability and protection against transposons [24], but also other functions have been described such as the regulation of gene expression acting in a similar way to miRNAs [36]. Their deregulation in cancer has been poorly explored [37, 38], and few of them have been identified as prognostic biomarkers [32, 39–41].

In the present study, we focused on the analysis of hsa\_piR\_004530 (piR\_004530), a piRNA identified in one of the first studies of piRNA profiles [26], which is closely located in the genome to testis-expressed protein 29 (*TEX29*, also known as *C13orf16*) [42] (Fig. 1), a germline gene whose hypomethylation was related to worst overall survival (OS) in lung cancer [43]. Moreover, we confirmed a positive correlation between *TEX29* and piR\_004530 in tumor tissue (Additional file 1: Fig. S1A). Additionally, the reactivation of both genes was observed following treatment with 5-aza-2-deoxycytidine (5-Aza-dC), which inhibits DNA methyltransferases (Additional file 1: Fig. S1B), but the specific overexpression of piR\_004530 did not induce significant changes in the expression of *TEX29* (Additional file 1: Fig. S1C). Therefore, since the reactivation of the “germline program” has been proposed as a mechanism for promoting carcinogenesis [44], we hypothesized that piR\_004530 is a germline piRNA that becomes reactivated during the tumorigenesis process and could act as a potential biomarker for NSCLC prognosis. In the present work, we studied the



**Fig. 1** Information about piR\_004530 sequence, synonyms and genomic location, and REMARK diagram of the study. DFS, disease-free survival; OS, overall survival

reactivation of piR\_004530 in NSCLC and evaluated its role as a relapse biomarker.

## Methods

### Patients samples

Tumor and paired normal tissue from 245 adult patients diagnosed with NSCLC who underwent complete surgical resection at the Hospital Clinic of Barcelona from May 2007 to October 2019 were prospectively collected. After obtention, the samples were immediately immersed in RNAlater® (Thermo Fisher Scientific, Waltham, MA, USA) and stored at  $-80^{\circ}\text{C}$  until processing.

### Cell lines

Cryopreserved samples of lung cancer cell lines HCC-15, SK-MES-1, HCC44, A549 (DSMZ, Braunschweig, Germany), and H23 (CRL-5800, ATTC, Manassas, VA, USA) and the normal immortalized lung cell line BEAS2B (ECACC, Sigma-Aldrich, St. Louis, MO, USA) were received in our laboratory and passaged for less than 6 months. HCC-15, HCC44, and H23 were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium (Thermo Fisher Scientific), while SK-MES-1 and A549 were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific). All of them were supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific). BEAS2B was cultured using

BEGMTM-2 BulletKit (Lonza, Basel, Switzerland). All cell lines were grown at  $37^{\circ}\text{C}$  in 5%  $\text{CO}_2$  and 95% relative humidity.

### 5-Aza-dC treatment

For methylation analysis, the SK-MES-1 cell line ( $4 \times 10^5$  cells) was plated 24 h prior to treatment. The next day, cells were treated with  $100\text{-}\mu\text{M}$  5-aza-dC or dimethyl sulfoxide (DMSO) (vehicle control) daily for 5 days and were harvested, and gene expression was assessed.

### RNA extraction, mRNA and piRNA quantification

Total RNA was isolated from tissue and cell lines using TRIzol® reagent (Thermo Fisher Scientific) following the manufacturer's instructions. RNA was quantified using NanoDrop ND-1000 device (Thermo Fisher Scientific). PiR\_004530 (5'-TCCTCCGTGTTGATGGTTAGAGGATT-3') was quantified using Custom TaqMan Small RNA Assay (assay ID: CTCE4DG; Thermo Fisher Scientific) as per the manufacturer's protocol. The reverse transcription reaction was carried out using 10 ng of RNA and the high-capacity cDNA Reverse Transcription Kit® (Thermo Fisher Scientific). Real-time PCR was performed using a 7500 real-time PCR system (Applied Biosystems, Thermo Fisher Scientific). And hsa-miR-191 (assay ID: 002299, TaqMan microRNA Assay, Thermo Fisher Scientific) was used as an endogenous control. We performed mRNA quantification of TEX29 (Hs00381012\_m1),

SOX2 (Hs01053049\_s1), OCT4 (Hs04260367\_gH), NANOG (Hs02387400\_g1), CD133 (Hs01009259\_m1), CD44 (Hs01075864\_m1), c-MYC (Hs00153408\_m1), CDKN1B (Hs00153277\_m1), and GAPDH (Hs99999905\_m1) as previously described using TaqMan expression assays [45]. Relative expression was calculated using the  $2^{-\Delta\Delta C_t}$  method.

#### Protein purification and Western blot

Protein purification was performed as previously described [45]. Thirty micrograms (30  $\mu$ g) of total protein from SK-MES-1 and HCC-15 cells—either transfected with the piR\_004530 mimic or control mimic—were separated by electrophoresis on 4–12% Bolt™ Bis–Tris Plus Mini Protein Gels (Invitrogen) and transferred onto nitrocellulose membranes. Membranes were blocked and incubated with a mouse monoclonal anti-Sox2 antibody (66,411–1-Ig, Proteintech) at a 1:1000 dilution and a mouse monoclonal anti- $\beta$ -actin antibody (clone A5441, Sigma-Aldrich, Germany) at the same dilution. A purified goat anti-mouse IgG (H + L) HRP-conjugated secondary antibody (clone 31,431, Invitrogen) was used at a 1:2000 dilution. Chemiluminescent signal for Sox2 was detected using the SuperSignal™ West Femto Substrate (Pierce Biotechnology, Boston, MA, USA), while detection of  $\beta$ -actin was performed using the SuperSignal™ West Pico Substrate (Pierce Biotechnology). Band intensities were visualized with the ChemiDoc™ Imaging System and quantified using Quantity One software (Bio-Rad, Barcelona, Spain) according to the manufacturer's instructions. All results were normalized to  $\beta$ -actin levels.

#### Cell transfection

On day 0,  $1.15 \times 10^5$  SK-MES-1 cells or  $1.25 \times 10^5$  HCC-15 or H23 cells were seeded in 24-well plates. On day 1, the cells were transfected with 100 nM of 3' terminal 2'-O-methyl piR\_004530 (5'-UCC UCC GUG UUG AUG GUU AGA GGA UU-3') or with 3' terminal 2'-O-methyl RNA control (5'-UCA CAA CCU CCU AGA AAG AGU AGA-3') using Lipofectamine 3000 (Thermo Fisher Scientific) following the manufacturer's instructions. For piR\_004530 inhibition experiments, the same methodology was used, but the cells were transfected with an ASO anti-piR\_004530 (5'-AATCCTCTA ACCATCAACACGGAGGA-3') or ASO control (5'-GATACCAGGGACATACGCTTGATCCTAGC-3'). To verify the overexpression or silencing of the piRNA, piR\_004530 expression was evaluated by real-time PCR 48 h after transfection.

#### Cell migration assay

Cell migration was measured by wound-healing assay by plating  $1.15 \times 10^5$  SK-MES-1 cells or  $1.25 \times 10^5$  HCC-15

or H23 cells in a 24-well plate. Twenty-four hours after plating, cells were transfected with piR\_004530 or control RNA as previously described. Eighteen hours after transfection, the cell monolayer was scraped in a straight line to create a "scratch" with a p10 pipet tip. Differences between cells growing in HCC-15, SK-MES-1, and H23 were assessed with images taken at 0, 6, 24, 30, 48, 72, and 96 h in IX53 inverted microscope using CellSense Entry 1.7 software (Olympus). The migration area ( $\mu\text{m}^2$ ) was assessed at different time points until the scratch was closed (HCC-15 and H23 cells 96 h, SK-MES-1 cells 48 h), with the 0-h time being set at 0%. Areas were calculated using ImageJ V1.54f.

#### Cell invasion assay

Invasion capacity was measured using QCM ECMatrix Cell Invasion Assay (ECM550, Sigma-Aldrich, Merck Life Science) following the manufacturer's instructions. After 18 h of transfection,  $2 \times 10^5$  cells were placed in serum-free media and seeded on the top of the invasion chamber. The lower chamber was filled with 10% fetal bovine serum media. At 24 h, membranes were stained, and images were obtained to visualize the proportion of invading cells. For quantification, membranes were immersed in 10% acetic acid to dissolve the cell stain solution. The absorbance of the solution was measured at 560 nm using the VersaMax™ Microplate Reader (Molecular Devices, San Jose, CA, USA).

#### MTS assay

Cell proliferation was analyzed using CellTiter 96® Aqueous One Solution Cell Proliferation Assay (MTS assay) (Promega, Madison, WI, USA). Eighteen hours after transfection,  $1 \times 10^3$  cells were seeded in 96-well plates, and proliferation was analyzed at 0, 24, 48, 72, and 96 h. For the analysis, cells were cultured at 37 °C and 5%CO<sub>2</sub> for 30 min with 20  $\mu$ l of MTS, and then absorbance at 490 nm was recorded using a VersaMax™ Microplate Reader.

#### Caspase 3/7 assay

Apoptosis was measured using Caspase-Glo® 3/7 Assay System (Promega) as per manufacturer's protocol. Eighteen hours after transfection,  $1 \times 10^4$  cells were seeded on white-coated 96-well plates. An equal volume of Caspase-Glo 3/7 reagent was added to each well and incubated for 1 h at room temperature (RT). Relative light intensity was measured in each well using a Veritas Microplate Luminometer (Turner BioSystems, Sunnyvale, CA, USA).

#### Clonogenic assays (2D)

A colony formation assay was performed to measure clonogenic capacity for 15 days. Obtained colonies were

fixed with 4% paraformaldehyde for 30 min at RT and stained with 0.5% crystal violet (Sigma-Aldrich) for 30 min. Then, the plate was air-dried for 24 h, and images were obtained. For quantification, methanol was added to each well to solubilize the crystal violet, and absorbance was measured at 595 nm using the VersaMax™ Microplate Reader.

### Spheroid formation assay

Cultured 2000 cells were seeded into a 96-well plate coated with agarose (1.5%) and cultured with stem cell media composed of DMEMF12, 1×B27 supplement (without vit. A), EGF (20 ng/ml), bFGF (20 ng/ml), and 1% P/S. Cells were maintained in suspension for 6 days until self-aggregated into spheroids. Cell size and morphology were determined using ImageJ V1.54f.

### Patient-derived lung cancer organoid

Patient-derived lung cancer organoids were generated from three different SCC patients. The first was a 61-year-old white male smoker with TNM stage IIB (T1N1M0), who harbored TP53 (c.713GT) and PTEN (c.724GT) mutations; his tumor tissue was PD-L1 positive (20%) as assessed by immunohistochemistry with the mouse anti-PD-L1 clone 22C3 antibody. The second patient was a 63-year-old white male, former smoker, diagnosed with SCC at TNM stage IA3 (T1N0M0); no clinically relevant mutations were detected, although a PIK3CA gain was identified. The third patient was a 77-year-old white male smoker with SCC at TNM stage IA1 (T1aN0M0), and no clinically relevant mutations were detected. Resected tumor tissue from all patients was collected within 6 h of surgery and processed as previously described [46]. Briefly, each tissue was mechanically dissociated with a sterile scalpel into smaller pieces followed by incubation with collagenase type II and DNase I (Merck) in DMEM/F12 (Thermo Fisher Scientific) with intermittent agitation on an orbital shaker at 37 °C for 2 h. After enzymatic digestion, the suspensions were passed through 100-μm cell strainer and centrifuged. Red blood cell lysis buffer (Sigma Aldrich) was added to lyse the erythrocytes for 5 min at RT. The resulting cells were resuspended in cold Matrigel (Corning, NY, USA), and 30-μl drops were seeded into pre-warmed 48-well culture plates at 37 °C for 10–20 min. Upon complete solidification, 400 μl of organoid culture media from [47] was added to each well, and plates were placed into incubators at 37 °C with a humidity level and 5% CO<sub>2</sub>. Organoid culture medium contained advanced DMEM/F12 supplemented with 100 μg/ml penicillin/streptomycin, 2-mM GlutaMAX 100×, 10-mM HEPES, 1:50 B-27 (Gibco, Thermo Fisher Scientific), 5-mM nicotinamide, 1.25-mM N-acetylcysteine, 500-nM A83-01, and 500-nM SB202190 (Sigma Aldrich,

Merck); 500 ng/ml R-spondin 1 and 100 ng/ml noggin (PeproTech, London, UK); and 100 ng/ml FGF-7, 25 ng/ml FGF-10, and 5-μM Y-27632 (StemCell, Vancouver, Canada).

### Organoid nucleofection

Organoids were mechanically dissociated using glass Pasteur pipettes. After cell counting,  $2 \times 10^5$  cells were resuspended with Lonza P3-Primary cell nucleofection Solution, and 1 μl of piR\_004530 or control was added. Then, the solution was transferred to 100-μl nucleofection cuvette (V4XP-3012, Lonza, Basel, Switzerland). Nucleofection was performed using Lonza 4D Nucleofector with X unit using the program EA125. After nucleofection, organoid culture medium supplemented with 10-μM Y-27632 was added to the nucleofection cuvette. The cell mixture was then recovered and cultured in Matrigel drops as previously indicated.

### Determination of organoid viability

Around  $1 \times 10^3$  nucleofected cells were seeded in a 10-μl drop of Matrigel in a white opaque 96-well plate. At day 3, after 30 min of plate equilibration to RT, CellTiter-Glo 3D (Promega) reagent was added to each well in a volume of 1:1. The plate was shaken for 5 min and incubated for 25 min at RT. The luminescence was measured in an Orion II Microplate Luminometer.

### Expression arrays analysis

Global gene expression in 10 tumor tissues (3 cases not expressing piR\_004530 vs 7 cases expressing it) was analyzed by Affymetrix GeneChip Human Gene 2.1 ST array in the Genomics Core Facility from IDIBAPS. Raw.cel files were provided which were used for further analysis. Bioinformatic analysis was performed as previously described [48] using R version 4.1.1 and Bioconductor packages.

### Statistical analyses

To verify that the data followed a normal distribution, Shapiro–Wilk and Kolmogorov–Smirnov tests were used. Since data was not normally distributed ( $p < 0.001$  in both tests), we used the Mann–Whitney *U*-test to compare differences between two groups in continuous variables and Kruskal–Wallis for more than two groups in the correlation with clinicopathological characteristics. The chi-square test was used to compare differences in proportions between two groups.

Median follow-up of the cohort was estimated using the reverse Kaplan–Meier method using *prodlm* package of R. The primary endpoints analyzed were disease-free survival (DFS) and overall survival (OS). As a secondary endpoint, time to relapse (TTR) was also analyzed.

DFS was calculated from the time of surgery to the date of relapse or death or last follow-up, OS was calculated from the time of surgery to the date of death by any cause or last follow-up, and TTR was calculated from the time of surgery to the date of relapse or last follow-up. Kaplan–Meier curves for DFS, OS, and TTR were drawn and compared by means of a log-rank test. Univariate and multivariate Cox regression analysis was conducted. All factors with a  $p$ -value  $< 0.1$  in the univariate analysis were included in the Cox multivariate regression analyses. Decision curve analysis (DCA) [49] was used to evaluate the clinical benefit of piR\_004530 in TTR and patient's risk stratification, using 48 months (4-year survival) as the survival time endpoint. The analysis was performed using *dcurves* package of R.

The optimal cutoff of piR\_004530 expression for DFS and OS was identified using X-Tile v.3.6.1 (Yale University 2003–2005) [50]. Patients that  $\log_{10}(2^{-[(\text{piR}_{004530} \text{ Ct} - \text{miR}_{-191} \text{ Ct}) - 12.5]}) > 1.19$  were classified as high.

For in vitro studies, at least three technical and three biological replicates have been performed in each in vitro experiment. A one-sample  $t$ -test is used to compare with the control condition.

All statistical analyses were performed using IBM SPSS Statistics v29, GraphPad Prism 10, and R v4.1.1. The level of significance was set to  $p < 0.05$ .

# Results

## Patient characteristics and prognostic predictors

The final analysis included 243 patients with a mean age of 67 years (range 33–85) and predominantly consisted of males (73.3%). Patients tended to be stage I (62.1%) and ADK (56%). A total of 72 (29.6%) patients received adjuvant therapy. None of the patients received neoadjuvant treatment. The median follow-up time was 84 months (interquartile range [IQR]: 65–99). A total of 96 (39.5%) patients experienced recurrence after surgery, and 56 of them had distant metastasis. The main clinicopathological characteristics and their impact on DFS and OS of the whole cohort are summarized in Table 1. Survival analysis was also conducted separately on SCC ( $n = 92$ ) and non-SCC ( $n = 151$ ) patients since the impact on prognosis of the piRNA analyzed differed between both histological subtypes. The REMARK diagram of the study is shown in Fig. 1.

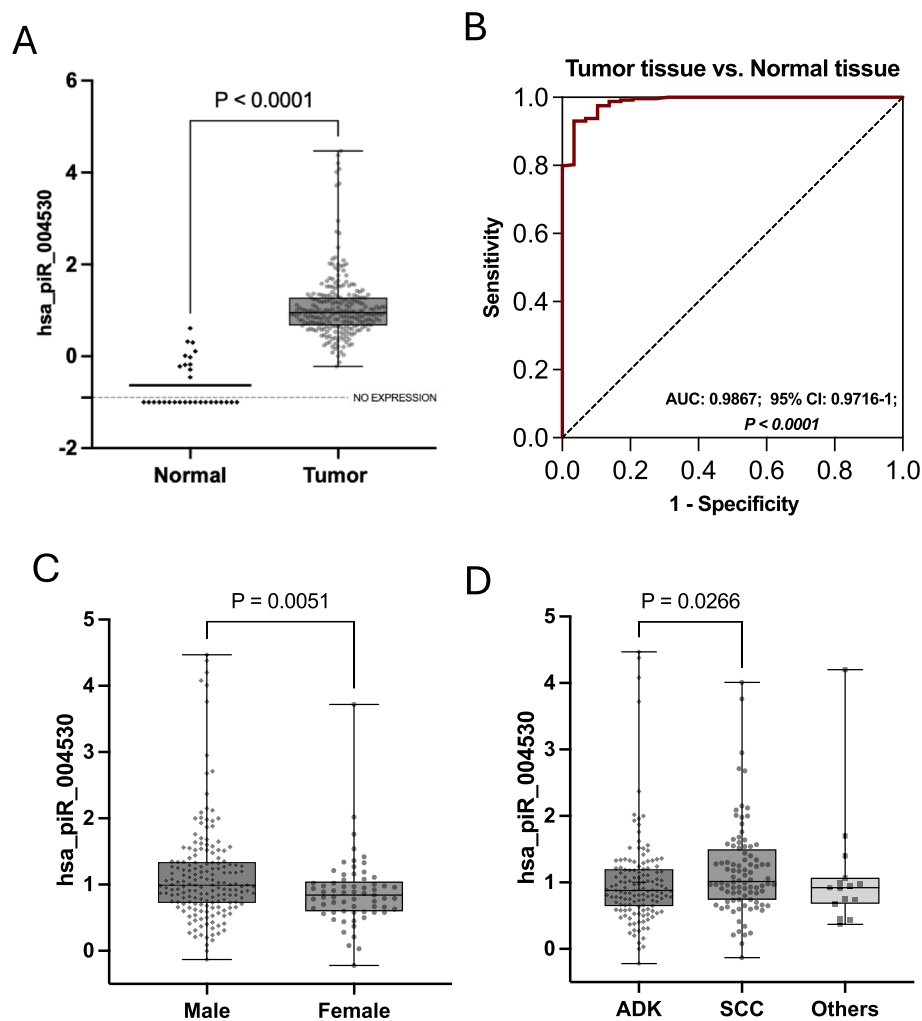
Most SCC patients were male (90.2%), and all of them were smokers (41.3% current and 58.7% former). The median follow-up time was 82 (IQR: 64–102) months. Patients tended to be stage I (58.7%). A total of 30 (32.6%) patients experienced disease recurrence after surgery. A total of 29 (31.5%) patients received adjuvant therapy, and this was the unique factor impacting DFS and OS in SCC, where patients

**Table 1** Patient's main clinicopathological characteristics

| Characteristics                           | N=243       | DFS   | OS    |
|-------------------------------------------|-------------|-------|-------|
| Sex                                       |             |       |       |
| Male                                      | 178 (73.3%) | 0.063 | 0.037 |
| Female                                    | 65 (26.7%)  |       |       |
| Median age, years (25th, 75th percentile) | 67 (61,74)  |       |       |
| < =65                                     | 103 (42.2%) | 0.479 | 0.571 |
| > 65                                      | 140 (57.6%) |       |       |
| ECOG PS                                   |             |       |       |
| 0                                         | 132 (54.3%) | 0.117 | 0.279 |
| 1                                         | 66 (27.2%)  |       |       |
| 2                                         | 45 (18.5%)  |       |       |
| Pathological stage                        |             |       |       |
| I                                         | 150 (61.7%) | 0.049 | 0.035 |
| II                                        | 63 (26%)    |       |       |
| III                                       | 30 (12.3%)  |       |       |
| Tumor size                                |             |       |       |
| < =30 mm                                  | 117 (48.2%) | 0.544 | 0.796 |
| 31–50 mm                                  | 73 (30%)    |       |       |
| 51–70 mm                                  | 36 (14.8%)  |       |       |
| > 70 mm                                   | 17 (7%)     |       |       |
| Lymph node involvement                    |             |       |       |
| N0                                        | 178 (73.3%) | 0.056 | 0.204 |
| N+                                        | 65 (26.7%)  |       |       |
| Histology                                 |             |       |       |
| Adenocarcinoma                            | 136 (56%)   | 0.638 | 0.589 |
| Squamous cell carcinoma                   | 92 (37.9%)  |       |       |
| Others                                    | 15 (6.1%)   |       |       |
| Smoking history                           |             |       |       |
| Current smoker                            | 93 (38.3%)  | 0.636 | 0.222 |
| Former smoker                             | 123 (50.6%) |       |       |
| Never smoker                              | 27 (11.1%)  |       |       |
| Type of surgery                           |             |       |       |
| Lobectomy                                 | 191 (78.6%) | 0.791 | 0.943 |
| Pneumonectomy                             | 15 (6.2%)   |       |       |
| Bilobectomy                               | 7 (2.9%)    |       |       |
| Segmentectomy                             | 14 (5.8%)   |       |       |
| Atypical resection                        | 16 (6.6%)   |       |       |
| Adjuvant treatment                        |             |       |       |
| Yes                                       | 72 (29.6%)  | 0.498 | 0.367 |
| Relapse                                   |             |       |       |
| Yes                                       | 96 (39.5%)  | -     | -     |

receiving adjuvant treatment had a better DFS (median survival time: 51.6 months vs not reached;  $p = 0.004$ ) and OS (66.4 months vs not reached;  $p = 0.008$ ) than other patients (Additional file 1: Fig. S2A).

Non-SCC patients were 62.3% male and 37.1% female and mostly ADK (90.1%). The median follow-up time was 86 (IQR: 65–99) months. Patients tended



**Fig. 2** piR\_004530 expression analysis in tumor and paired normal tissue and correlation with main clinicopathological characteristics **A** Box plot analysis of piR\_004530 in tumor and paired normal tissue. **B** ROC curve analysis showing the capacity of piR\_004530 to discriminate between tumor and normal tissues. **C** Differential expression of piR\_004530 between male and female patients. **D** Differential expression of piR\_004530 according to main histological subtypes. ADK, adenocarcinoma; SCC, squamous cell carcinoma

to be stage I (64.2%). A total of 66 (43.7%) patients experienced postsurgical recurrence. A total of 43 (28.5%) patients received adjuvant therapy. Large tumor size, positive lymph node involvement, and higher stage were related to shorter DFS ( $p=0.004$ ,  $p<0.001$ , and  $p=0.007$ , respectively) and shorter OS ( $p=0.08$ ,  $p=0.003$ , and  $p=0.014$ , respectively) (Additional file 1: Fig. S2B).

#### piR\_004530 expression in patient samples

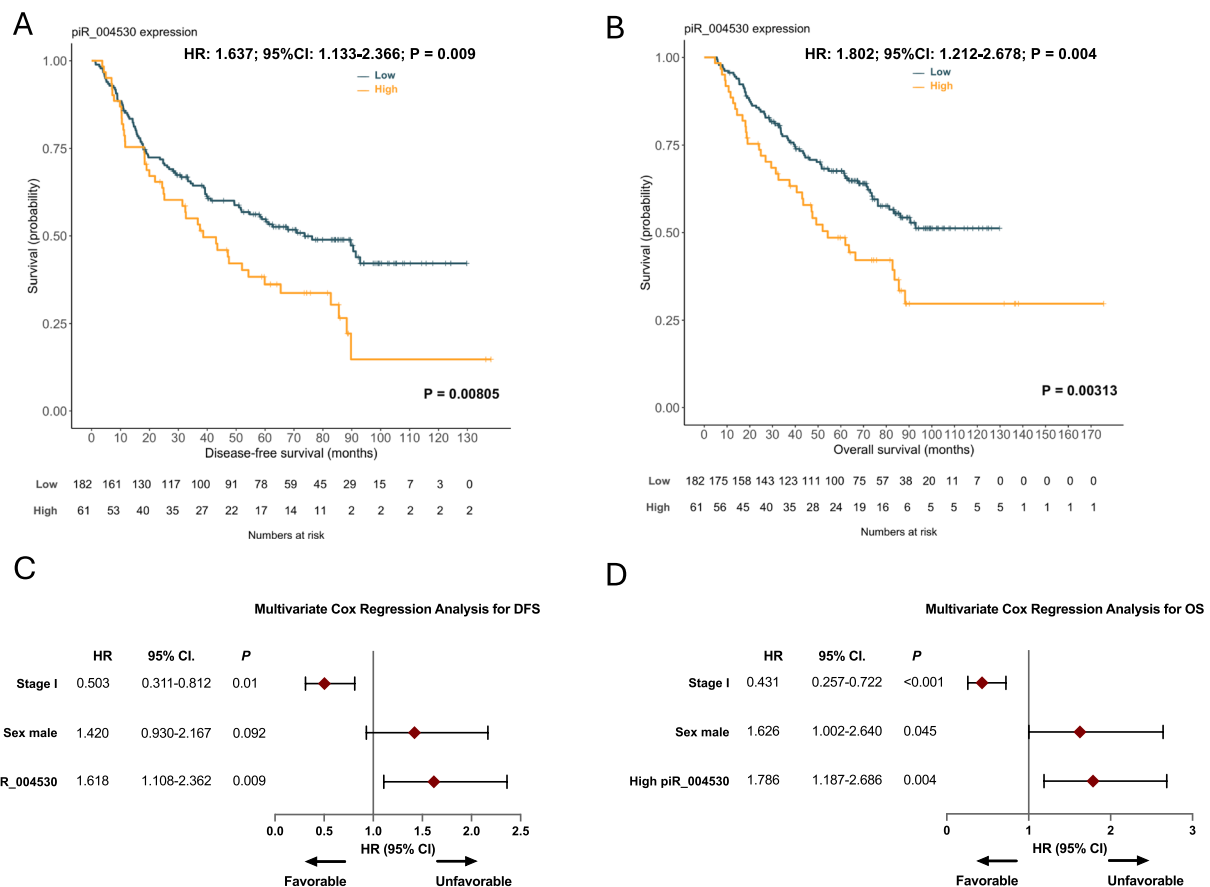
The analysis of the expression levels of piR\_004530 in tumor and paired normal tissue showed a significant upregulation in tumor tissue ( $p<0.0001$ , Fig. 2A). Remarkably, in 63% of the normal tissue samples, the piR\_004530 was not detectable. ROC curve analysis

highlighted the ability of piR\_004530 to discriminate tumor from normal tissue (area under the ROC curve [AUC]: 0.9867, 95% confidence interval [CI]: 0.9716–1,  $p<0.001$ ; Fig. 2B).

The correlation of piR\_004530 expression with the main clinicopathological characteristics showed that it is significantly downregulated in females ( $p=0.0051$ , Fig. 2C) and upregulated in SCC patients in comparison to ADK ( $p=0.0266$ , Fig. 2D).

#### Patients with higher levels of piR\_004530 had the worst prognosis after surgery

To explore the potential prognostic impact of piR\_004530 expression, we determined the optimal cut-off for piR\_004530 using X-tile, which identified the



**Fig. 3** Impact of piR\_004530 levels on patient outcome in the whole cohort. **A** Kaplan–Meier plot of the DFS of NSCLC cases stratified by the cutoff value of piR\_004530. *P*-value (bottom right) was calculated using the log-rank test. The univariate Cox analysis is included at the top of the graph. **B** Kaplan–Meier plot of the OS of NSCLC cases stratified by the cutoff value of piR\_004530. *P*-value (bottom right) was calculated using the log-rank test. The univariate Cox analysis is included at the top of the graph. **C** Forest plot of multivariate Cox regression analysis for DFS of patients with NSCLC. **D** Forest plot of multivariate Cox regression analysis for OS of patients with NSCLC. DFS, disease-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval

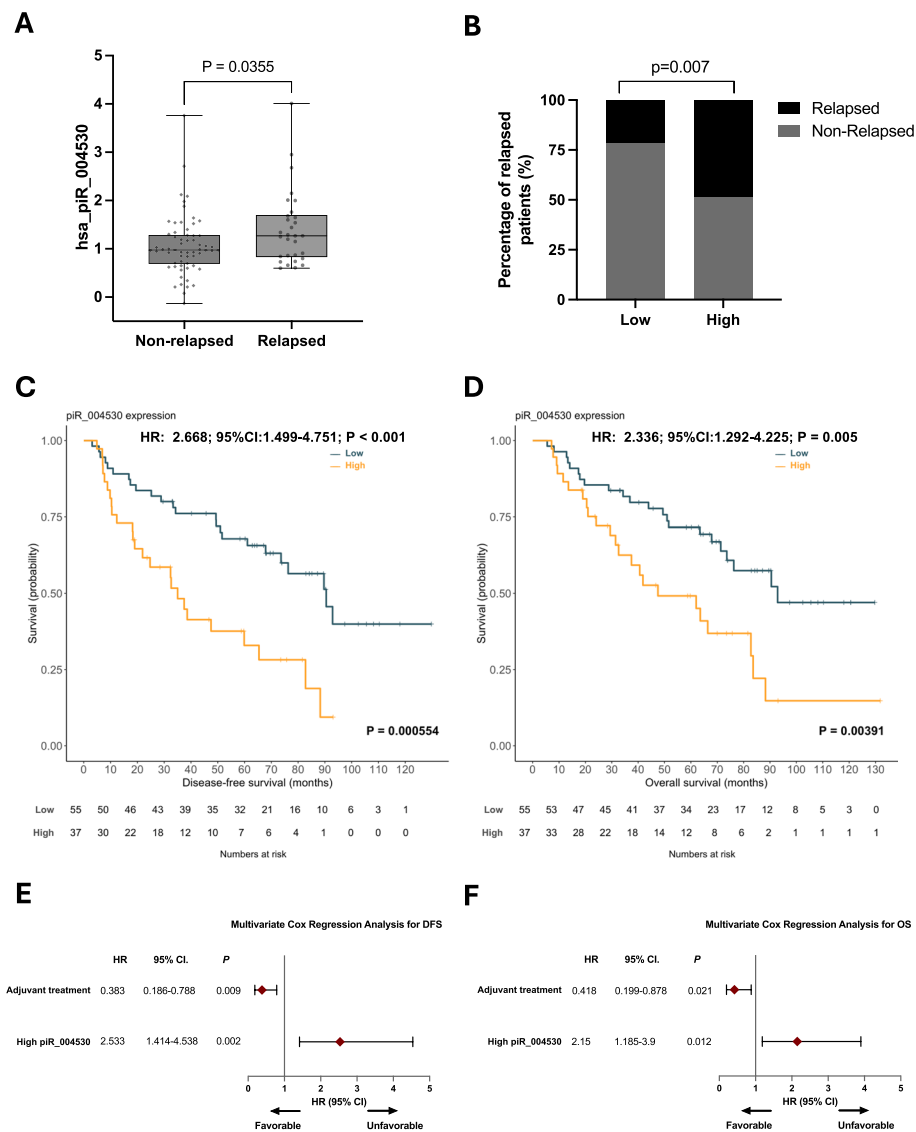
75th percentile as the best discriminant point for survival analysis. Patients were classified into low ( $n=182$ , 75%) or high ( $n=61$ , 25%) expression. Patients with high expression had shorter DFS (38.6 vs. 73.6 months,  $p=0.008$ ; Fig. 3A) and OS (54.3 months vs. not reached,  $p=0.003$ ; Fig. 3B) than patients with low piR\_004530 expression.

Univariate Cox proportional regression analysis (Fig. 3A, B) corroborated the shorter DFS (hazard ratio [HR]: 1.637; 95% CI: 1.133–2.366;  $p=0.009$ ) and shorter OS (HR: 1.802; 95% CI: 1.212–2.678;  $p=0.004$ ) of patients with high piR\_004530 expression levels. Moreover, in the multivariate Cox regression analysis, piR\_004530 expression levels emerged as an independent prognostic factor for DFS (HR: 1.618; 95% CI: 1.108–2.362;  $p=0.013$ ) and OS (HR: 1.786; 95% CI: 1.187–2.686;  $p=0.005$ ) adjusted for patient's sex and stage (Fig. 3C, D).

### SCC patients with higher levels of piR\_004530 displayed significantly worse outcomes after surgery

Since we observed significant differences in the expression levels between SCC ( $n=92$ ) and non-SCC patients ( $n=151$ ) (Fig. 2D), we decided to evaluate the prognostic role in each histological group independently (shown in REMARK diagram in Fig. 1).

In SCC patients, interestingly, we observed significant differences in piR\_004530 expression levels according to relapse (Fig. 4A), where patients who relapsed after surgery had higher piRNA levels ( $p=0.0355$ ). When patients were classified in low ( $n=55$ , 60%) or high ( $n=37$ , 40%) expression, patients with high expression had a higher relapse rate ( $p=0.007$ , Fig. 4B) and had shorter DFS (35.033 vs. 90.5 months,  $p<0.001$ , Fig. 4C) and shorter OS (13.526 months vs. not reached,  $p=0.004$ , Fig. 4D) than patients with low piR\_004530 expression.



**Fig. 4** piR\_004530 expression impacts relapse and survival of squamous cell carcinoma patients (SCC). **A** Box plot analysis of piR\_004530 levels between relapsed and non-relapsed SCC patients

Supportively, univariate Cox proportional regression analysis (Fig. 4C, D) confirmed that high piR\_004530 levels were related to shorter DFS ( $HR: 2.668$ , 95%  $CI: 1.499-4.751$ ,  $p < 0.001$ ) and shorter OS ( $HR: 2.336$ , 95%  $CI: 1.292-4.225$ ,  $p = 0.005$ ).

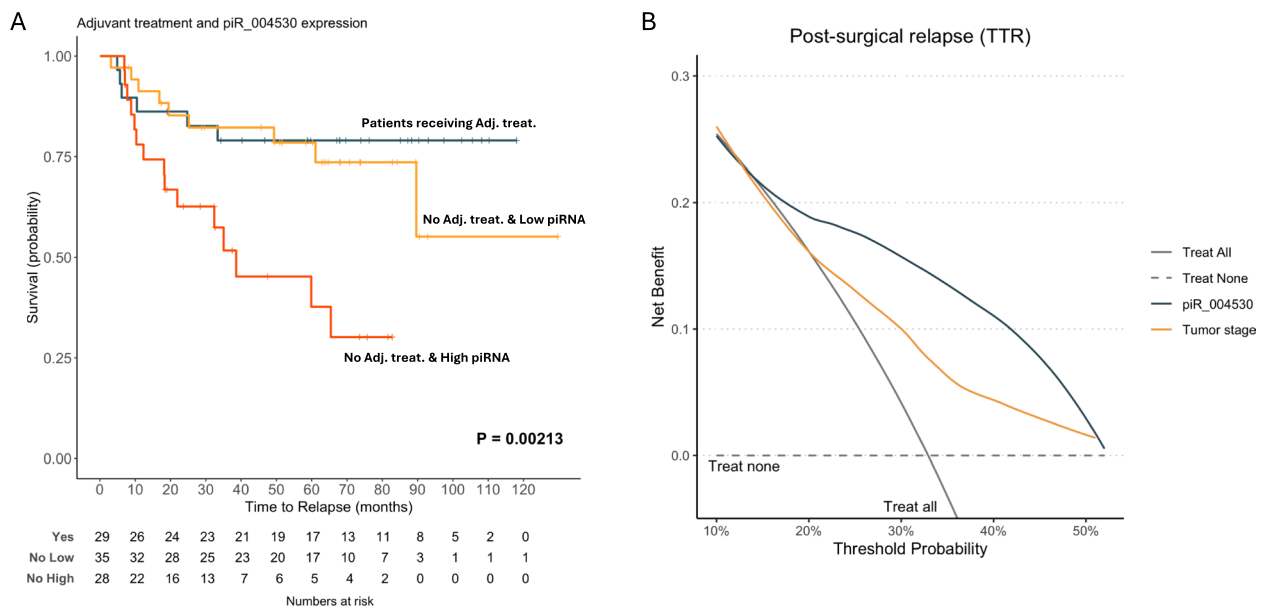
Moreover, in the multivariate Cox regression analysis, piR\_004530 expression levels emerged as an independent prognostic factor for DFS ( $HR: 2.533$ ; 95%  $CI: 1.414-4.538$ ;  $p = 0.002$ ) and OS ( $HR: 2.15$ ; 95%  $CI: 1.185-3.9$ ;  $p = 0.012$ ) adjusted for adjuvant treatment (Fig. 4E, F).

In non-SCC patients (ADK,  $n = 136$ ; other histology,  $n = 15$ ), no significant differences were observed either for DFS or for OS according to piR\_004530 levels

(Additional file 1: Fig. S3A). Neither the analysis of ADK nor that of other histology independently showed any marked prognostic value (Additional file 1: Fig. S3B, C) highlighting their clinical significance in the outcome of SCC patients.

#### Consistency of piR\_004530 prognostic trends in a validation patient cohort

To support the findings obtained in our initial cohort, we analyzed an independent set of 74 NSCLC patients who underwent surgical resection at our institution between November 2019 and October 2023 (Additional file 1: Fig. S4). Due to the impact of the COVID-19 pandemic



**Fig. 5** piR\_004530 identifies patients at high risk of relapse between patients not receiving adjuvant treatment. **A** Kaplan–Meier plot of TTR of SCC cases stratified into three groups: 1) Patients receiving adjuvant treatment, 2) patients not receiving adjuvant treatment and with low levels of piR\_004530, and 3) patients not receiving adjuvant treatment and with high levels of piR\_004530. *P*-values were calculated using the log-rank test. **B** Decision curve analysis of the control (tumor stage) and the piR\_004530-fitted models for TTR of SCC patients not receiving adjuvant treatment

during this period, the number of available samples was limited, which reduced the size and completeness of this validation series. Nevertheless, we observed a similar trend: patients with high piR\_004530 expression showed shorter DFS (24.767 months vs. not reached,  $p=0.00887$ ), particularly among SCC cases (24.1 months vs. not reached,  $p=0.00356$ ) (Additional file 1: Fig. S5). While in OS the results did not reach significance in the whole cohort, the patients with high piR\_004530 had shorter OS in both the whole cohort and in the SCC group (SCC: 48.5 months vs. not reached,  $p=0.0528$ ).

Although these results are based on a smaller and single-center cohort, they are consistent with the prognostic value observed in the main cohort and support the relevance of piR\_004530 expression as a potential biomarker in NSCLC.

#### piR\_004530 entails superior clinical benefit in the prediction of postsurgical relapse for non-treated patients

Since in SCC patients the piRNA expression and the administration of adjuvant treatment emerged as independent prognostic factors in the multivariate analysis, we decided to integrate both factors to evaluate whether the combination resulted in a superior risk stratification marker for postsurgical recurrence which

could be of utility to improve post-surgery treatment decisions. For this analysis, we focused on the secondary endpoint TTR. Kaplan–Meier analysis for TTR unveiled that patients receiving adjuvant treatment after surgery (a clinical decision based on clinico-pathological characteristics, including disease stage) had longer TTR (Fig. 5A) than patients not receiving treatment. In the group of patients not receiving adjuvant treatment, we observed that while patients with low piR\_004530 did not show significant differences in relation to TTR with the group of treated patients, the group of non-treated patients with high piR\_004530 had significantly shorter TTR ( $p=0.00213$ , Fig. 5A). Moreover, this observation was also confirmed for DFS ( $p=0.000271$ ) and OS ( $p=0.00154$ ) (Additional file 1: Fig. S6). These results suggest that piR\_004530 could be useful in identifying a high-risk group among patients, who, based on clinical characteristics, were not given adjuvant treatment after surgery. This group may benefit from adjuvant treatment despite the initial decision. To corroborate this hypothesis, a DCA analysis was applied to assess the clinical net benefit of piR\_004530 on the prediction of postsurgical recurrence. Strikingly, multivariate models fitted for piR\_004530 resulted in superior prediction of SCC patients' TTR over tumor stage (control) (Fig. 5B).

### piR\_004530 overexpression in lung cancer cell lines promotes migration and invasion

We decided to conduct in vitro analysis to confirm the role of the piR\_004530 as a biomarker for relapse. The expression analysis of piR\_004530 in several lung cancer cell lines showed that the piRNA was underexpressed in two ADK cell lines and upregulated in one SCC cell line in comparison to the normal immortalized lung cell line BEAS2B (Additional file 1: Fig. S7A). We overexpressed piR\_004530 in two SCC cell lines, HCC-15 and SK-MES-1 (Additional file 1: Fig. S7B), and in one ADK cell line H23 and evaluated whether the cell migration and invasion capacities of these cell lines were compromised. Overexpression of piR\_004530 was associated with increased migration capacity in all cell lines, as observed in the wound-healing assay (Fig. 6A, B, C). Significant differences were observed at 24 h and 30 h for SK-MES-1 ( $p < 0.001$  and  $p < 0.001$ , respectively) and for H23 ( $p = 0.0167$  and  $p = 0.0138$ , respectively) and at 48, 72, and 96 h for HCC-15 ( $p = 0.013$ ,  $p < 0.001$ , and  $p < 0.001$ , respectively). We also confirmed that inhibition of piR\_004530 using an ASO (Additional file 1: Fig. S7C) produced the opposite effect on migration, with reduced piR\_004530 levels associated with decreased cell migration (Additional file 1: Fig. S8).

Additionally, the overexpression of piR\_004530 also impacted cell invasion in the three cell lines analyzed. The cells overexpressing piR\_004530 had a major invasion capacity through the extracellular matrix compared to control cells (Fig. 6D, E, F). In SK-MES-1, the cell staining quantification showed an increase of 65% ( $p = 0.0175$ ); HCC-15 showed a rise of 15% ( $p = 0.0240$ ), and H23 showed a rise of 38% ( $p = 0.036$ ) compared to control cells.

### Evaluation of the effect of piR\_004530 overexpression on cell viability and apoptosis of lung cancer cells

The analysis of cell viability by measuring metabolic activity (MTS assay) showed minor or no differences between control and overexpressing piR\_004530 cells (Additional file 1: Fig. S9A). Only a modest significant increase in proliferation can be observed in SK-MES-1 in the group overexpressing piR\_004530, where a trend to significance was observed at 96 h ( $p = 0.0504$ ) and a significant difference was detected at 120 h ( $p = 0.0395$ ).

The apoptosis analysis showed that in two of the three cell lines analyzed, the apoptosis became reduced after piwiRNA overexpression. In SK-MES-1 and H23, a mean reduction of 25.1% ( $p = 0.0151$ ) and 14.1% ( $p = 0.0145$ ) in caspase 3/7 activity could be appreciated at 24 h, respectively, while no significant differences in apoptosis were observed in HCC-15 cell line (Fig. 6G, H, I).

### piR\_004530 overexpression enhanced colony formation

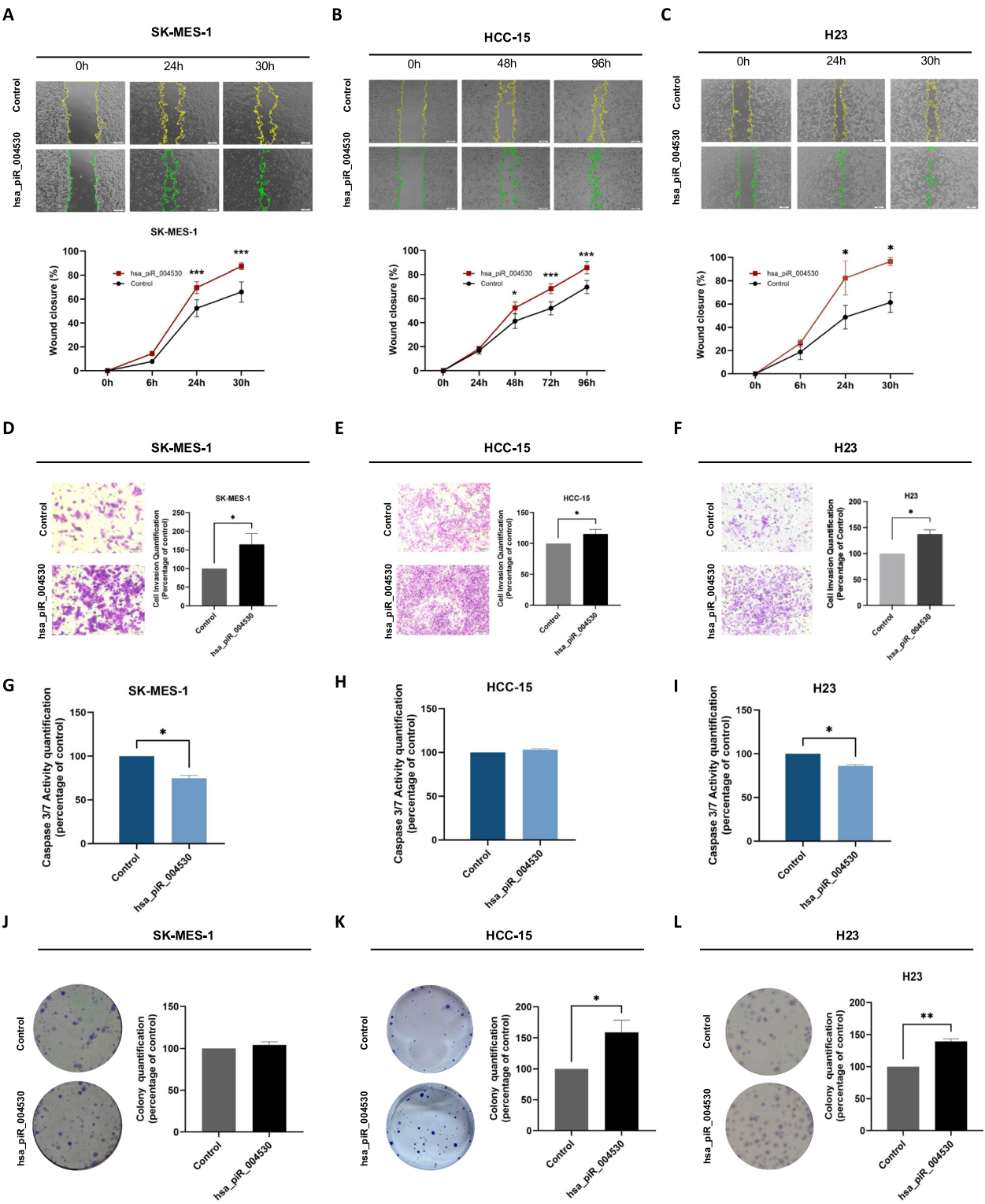
To assess whether piR\_004530 impacted the ability of single cells to survive and grow into colonies, we performed a colony formation assay. The overexpression of piR\_004530 enhanced colony formation in two of the three cell lines analyzed, HCC-15, H23 cells, but not in SK-MES-1 (Fig. 6J, K, L). After 15 days, piR\_004530-overexpressing HCC-15 and H23 cells formed 59.17% ( $p = 0.0339$ ) and 39.58% ( $p = 0.0032$ ) more colonies, respectively, than controls.

### piR\_004530 overexpression induced stemness

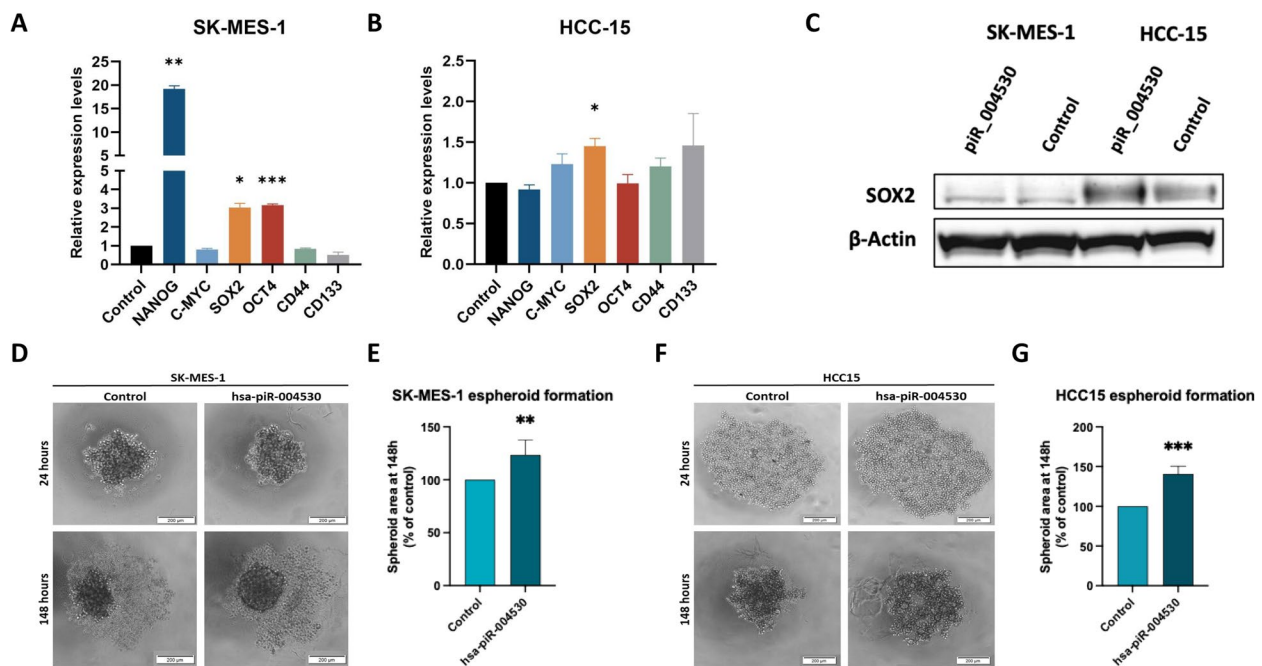
Since it has been previously shown that the piRNA-PIWI axis is more active in undifferentiated cells and stem cells [35], we decided to analyze whether the overexpression of our piRNA induced the expression of different stem cell markers, including *SOX2*, *OCT4*, *NANOG*, *CD133*, *CD44*, and *C-MYC*. The overexpression of piR\_004530 correlated with significant upregulation of *SOX2* mRNA levels in both cell lines ( $p = 0.0115$  for SK-MES-1 and

(See figure on next page.)

**Fig. 6** In vitro analysis in lung cancer cell lines to determine the functional role of piR\_004530. **A** Representative image of wound-healing assay in SK-MES-1 cell line transfected with piR\_004530 or control RNA. Photographs were taken at 0, 24, and 30 h. Calculated wound healing areas from three independent replicates are shown in the graph. **B** Representative image of wound-healing assay in HCC-15 cell line transfected with piR\_004530 or control RNA. Photographs were taken at 0, 48, and 96 h. Calculated wound healing areas from three independent replicates are shown in the graph. **C** Representative image of wound-healing assay in H23 cell line transfected with piR\_004530 or control RNA. Photographs were taken at 0, 24, and 30 h. Calculated wound healing areas from three independent replicates are shown in the graph. **D** SK-MES-1, **E** HCC-15, and **F** H23 cell lines. Representative images of the well of the control RNA or piR\_004530-transfected cells are shown on the left. Quantitative analysis of invasion by dissolving crystal violet-stained cells in 10% acetic acid and colorimetric reading of OD at 560 nm are shown at the right. Data is presented as a ratio of control. Apoptosis analysis by measuring caspase 3/7 activity in **G** SK-MES-1, **H** HCC-15, and **I** H23 control and piR\_004530 transfected cells. Data is shown as a percentage of control. Colony formation (2D) assay in **J** SK-MES-1, **K** HCC-15, and **L** H23 cell line for 15 days. A representative image is shown on the left, and quantification by solubilization of crystal violet through absorbance measurement of three independent replicates is shown on the right. All data represented in bar and line plots are the average of at least three independent replicates, and error bars represent SEM. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$



**Fig. 6** (See legend on previous page.)



**Fig. 7** hsa\_piR\_004530 promotes the stemness in lung SCC cell lines. Expression levels of NANOG, C-MYC, SOX2, OCT4, CD44, and CD133 in **A** SK-MES-1 or **B** HCC-15 cells transfected with hsa\_piR\_004530 in comparison with control RNA. \*\* $p < 0.01$  and \*\*\* $p < 0.001$ . **C** Western blot analysis of SOX2 in transfected SK-MES-1 and HCC-15 cells with hsa\_piR\_004530 mimic or control. **D** Representative bright-field images at 24 and 148 h of spheroid formation assay using SK-MES-1 cells transfected with hsa\_piR\_004530 mimic or control. **E** Bar graph showing spheroid area at 148 h in cells transfected with hsa\_piR\_004530 mimic versus control in SK-MES-1 cells. **F** Representative bright-field images at 24 and 148 h of spheroid formation assay using HCC-15 cells transfected with hsa\_piR\_004530 mimic or control. **G** Bar graph showing spheroid area at 148 h in cells transfected with hsa\_piR\_004530 mimic versus control in HCC-15 cells. All data represented are the average of at least three independent replicates, and error bars represent SEM. \* $p < 0.05$ ,  $p < 0.01$  (\*\*),  $p < 0.001$  (\*\*\*)

$p = 0.0423$  for HCC-15; Fig. 7A, B). Moreover, in the SK-MES-1, significant upregulation of *NANOG* ( $p = 0.0013$ ) and *OCT4* ( $p = 0.0009$ ) was observed (Fig. 7A). Moreover, we performed protein analysis of SOX2 levels since this stem cell marker was overexpressed in both cell lines. Protein levels of SOX2 increased after piR\_004530 overexpression, especially in the HCC-15 cell line (Fig. 7C).

Finally, we performed spheroid formation assays as a functional approach to evaluate stem cell properties after piR\_004530 overexpression. These assays assess the self-renewal capacity and stemness potential of the cells, providing biologically relevant functional validation. The overexpression of piR\_004530 was associated with the formation of larger spheroids in both cell lines. In SK-MES-1 cells, the spheroid area at 148 h was 127.79% that of control cells ( $p = 0.0095$ ), and in HCC-15 cells, it was 153.55% ( $p = 0.0001$ ) (Fig. 7D, E, F, G).

#### piR\_004530 overexpression boosted cell viability of patient-derived SCC organoids

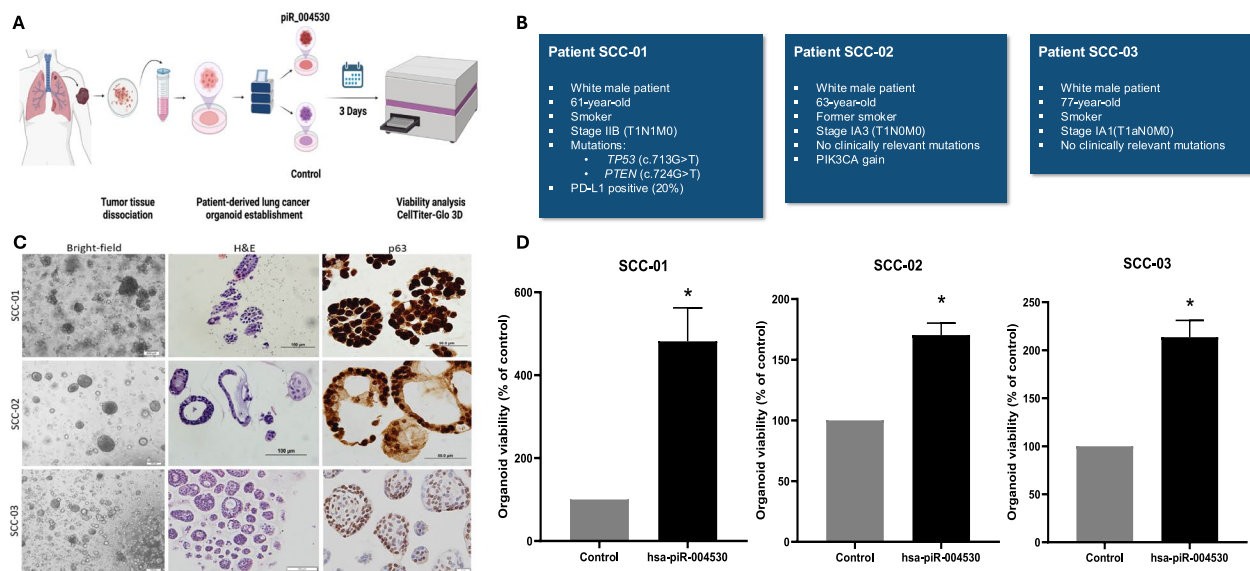
To validate the impact of piR\_004530 in 3D models, we grew patient-derived lung cancer organoids from tumor tissue of three resected SCC patients, and using

nucleofection, we overexpressed the piRNA and evaluated cell viability at 3 days (Fig. 8A). The main patient characteristics of the patient of origin of the organoid cell lines are summarized in Fig. 8B. Figure 8C shows a representative image of each cultured organoid, a hematoxylin staining of embedded organoids, and the p63 staining (SCC marker) for each case.

The nucleofection experiment showed that in all cases, the overexpression of the piR\_004530 increased the viability of cultured organoids (Fig. 8D,  $p < 0.05$ ), demonstrating that the piR\_004530 overexpression impacted 3D growth of primary patient samples. We also confirmed that inhibition of piR\_004530 using an ASO produced the opposite effect, with reduced piR\_004530 levels associated with decreased viability of cultured organoids (Additional file 1: Fig. S9B).

#### Identification of differential expressed genes between expressing and not expressing piR\_004530 patients

To preliminarily address how piR\_004530 exerts its oncogenic effect, we performed an analysis of expression arrays using previously generated data comparing



**Fig. 8** hsa\_piR\_004530 overexpression boosted cell viability of patient-derived lung SCC organoids. **A** Study workflow for obtention of patient-derived lung cancer organoids from NSCLC resected patients and nucleofection procedure used to analyze the role of hsa\_piR\_004530 overexpression and its impact on organoid viability (created with BioRender.com). **B** Main clinicopathological characteristics of the patients of origin of the used organoids. **C** Representative bright-field images of organoid growth in Matrigel, formalin-fixed paraffin-embedded (FFPE) sections of organoids stained with hematoxylin–eosin, and immunohistochemistry of p63, an SCC marker of FFPE sections of organoids used in the study. **D** Comparison of cell viability at 3 days between organoids nucleofected with hsa\_piR\_004530 or control RNA. All data represented are the average of at least three independent replicates, and error bars represent SEM

patients with high piR\_004530 expression ( $n=7$ ) to those without piRNA expression ( $n=3$ ). From this, we identified a set of 115 differentially expressed genes (Additional file 1: Fig. S10A and Additional file 2: Table S1). With the differentially expressed genes (DEG) identified, we performed a gene enrichment analysis for Gene Ontology terms to identify the biological processes affected by the re-expression of our piRNA. A high number of the DEG identified were related to metabolic processes (GO:0031323, GO:0060255, GO:0043170, GO:0019222; GO:0080090, GO:0009893). There was also an enrichment in genes associated with cell communication (GO:0007154) and cell surface receptor signaling pathways (0007166). Moreover, an enrichment in genes related to response to different stimuli and to stress (GO:0050896, GO:0051716, GO:0042221, GO:0033554, GO:0006950) was also identified (Additional file 1: Fig. S10B).

With the idea that the piRNA could act as a micro-RNA binding to mRNA targets and inhibiting them, we performed a negative correlation analysis to identify genes whose expression decreased with the reactivation of the piRNA (Additional file 1: Fig. S10C and Additional file 3: Table S2). These genes may represent potential direct targets of this piRNA, and between them, the most significantly correlated gene was SERPINB6 ( $r = -0.938$ ,  $p = 0.0000612$ ). We have

also analyzed the prognostic role of SERPINB6 using Kaplan–Meier plotter data, and the lower expression correlated with shorter OS of lung cancer patients ( $p = 0.0038$ ; Additional file 1: Fig. S10D). Additionally, other highly negatively correlated genes and lncRNAs have been identified in the top 5, including LMNTD2-AS, GRID2IP, LOC202181, or GAK. Low GAK levels were also observed to correlate with shorter OS in lung cancer using Kaplan–Meier plotter database ( $p < 0.001$ ). No data on Kaplan–Meier plotter database was available neither for GRID2IP nor for the lncRNAs.

Moreover, since our hypothesis on the selection of this piRNA was related to reactivation of germline programs, we performed an additional analysis to identify other reactivated germline genes in cases re-expressing the piRNA. First, we identified genes positively correlated with the expression of our piRNA in the array data (Additional file 4: Table S3). Secondly, we cross-referenced the list obtained with the Cancer-Testis Antigens Database (CTDatabase) [51]. Using this database, we identified several genes that correlated positively with piR\_004530 expression that has been consistently described as germline-reactivated genes in tumors, including MAGEB1, ARX, ZNF165, TEX101, GAGE13, SPAG4, ODF1, SYCE1, AKAP3, or SPANXC. Moreover, we have also identified in the list other genes that, despite not being included in the

CTDatabase, are with high probability germline genes such as CATSPER4, other members of the TEX family (TEX2, TEX22, TEX45, TEX261), other members of the SPAG family (SPAG7, SPATA3), PRSS21, ADAM1A, DEFB108B, and TSKS.

## Discussion

In resected NSCLC patients, the recurrence rate remains high (39.5% in our cohort) despite being treated with potentially curative surgery [52]. Adjuvant treatment has been proven to improve the prognosis of resected patients with N1/N2 NSCLC [53]; however, its efficacy in stage I patients remains controversial [54, 55]. Some small tumors are highly aggressive and, despite being detected at a clinically favorable stage, can progress to widespread, fatal disease. Moreover, the treatment does not differ across histological types, but the biological behavior of tumors varies significantly [56]. With the discussion reaching the cellular level, we must consider distinguishing NSCLC based on cellular characteristics, and this entails beginning our research by initially separating the most frequent histological subtypes.

In the present study, we decided to investigate the potential role of piR\_004530 expression as a biomarker for NSCLC outcome in resected patients (clinical stages I–II) based on the idea that this germinal noncoding RNA, which is closely located in the genome to *TEX29* [42] (six copies of this piRNA are located upstream of *TEX29*), could become reactivated during the carcinogenesis process and could be used as a prognostic biomarker [44]. First, we corroborated that *TEX29* expression positively correlated with piR\_004530 (Additional file 1: Fig. S1A), and, secondly, we verified that both genes become reactivated after DNA methyltransferase inhibition by 5-Aza-dC treatment (Additional file 1: Fig. S1B). Moreover, *TEX29* was previously identified in the top 10 upregulated genes in CD133+ human prostate cells [57], which allowed us to hypothesize that its expression was related to stemness regulation, as is likely the case for our piRNA. All this makes us think that piR\_004530 becomes reactivated together with *TEX29* and could be part of a germline reactivation program related to the induction of tumorigenesis. Moreover, in lung ADK patients, it was observed that *TEX29* hypomethylation was associated with shorter OS [43]. Additionally, using Kaplan–Meier plotter [58] (*TEX29* probe 236853\_at), we observed that NSCLC patients expressing high levels of *TEX29* had shorter OS than patients with low levels (median survival 63.4 vs. 85 months,  $p=0.002$ ; Additional file 1: Fig. S11). However, a limitation of our study is that we did not directly assess the methylation status of *TEX29* or its correlation with piR\_004530 expression in our patient cohort. Although our in vitro

results support the hypothesis of co-regulation by DNA methylation, this mechanism should be confirmed in future studies including methylation profiling of patient samples.

After studying piR\_004530 expression in 276 samples (88% tumor tissues), we observed that this piRNA seems to be reactivated in tumor tissue, since in most paired normal tissue samples we were not able to detect its expression. In this line, we have previously shown that the PIWIL1 protein, involved in the biogenesis of piRNAs, becomes reactivated in the tumor tissue of some NSCLC patients, and this reactivation is associated with worse patient outcomes [32]. Both results corroborate that the PIWI/piRNA pathway reactivation, in this case exemplified by piR\_004530, is an oncogenic event. Moreover, the higher levels of piR\_004530 in tumor tissue were associated with shorter DFS and OS, indicating that we have identified a promising relapse biomarker. Although the analysis in the whole cohort showed in a significant way that higher levels of the piRNA were associated with worse prognosis, the subanalysis by histological groups showed that the effect of the piRNA was more relevant in the SCC subgroup. In this group, the piRNA expression was significantly higher than that in other histologies and was also significantly higher in relapsed patients. Moreover, while in SCC patients the higher levels were associated in a stronger way with shorter DFS and OS, the differences did not remain significant in the non-SCC subgroup. As a further step to support our findings, we performed a preliminary validation using an independent cohort of patients resected at our center between 2019 and 2023. Although the number of available samples was limited due to restricted access to clinical material during the COVID-19 pandemic, the observed results were consistent with those from our initial cohort, particularly in SCC patients, reinforcing the potential prognostic value of piR\_004530 in NSCLC. We acknowledge, however, that this second cohort was derived from the same institution, and that external and prospective multicenter validation will be necessary to confirm the generalizability of our findings.

Given the well-established association between smoking and lung squamous cell carcinoma, we evaluated whether smoking history could confound the prognostic significance of piR\_004530. We first assessed piR\_004530 expression across different smoking status groups (never, former, and current smokers) and found no significant differences. Additionally, in a multivariate Cox regression model that included smoking status, age, sex, and tumor stage, piR\_004530 remained an independent prognostic factor. These results suggest that the prognostic value of piR\_004530 is not confounded by smoking history and

may serve as a robust biomarker regardless of tobacco exposure.

Additionally, since we observed a significant difference in piRNA expression between male and female patients—with lower levels in the latter group—we decided to evaluate whether the piRNA exhibits sex-specific effects on the survival of lung cancer patients. However, in both sexes, higher levels of the piRNA were associated with worse survival (Additional file 1: Fig. S12). Therefore, no sex-dependent effect on survival was observed.

Although the TNM classification system is the most relevant clinical factor used to decide postsurgical strategies in lung cancer patients [59], this does not account for the significant differences in survival rates between different histopathological cell types. Fukui T. et al. analyzed the prognosis after surgery of 1587 patients with ADK (77%) and 472 with SCC (23%) and reported significant survival differences between ADK and SCC patients. Specifically, they observed that stage I SCC patients exhibited markedly worse prognoses compared to ADK patients [60]. Back to our cohort, we discovered that among patients in stage IA, those with SCC exhibited significantly poorer OS compared to those with ADK. Moreover, when we performed the univariate and multivariate analysis, the unique clinical factor that impacted the survival of SCC patients was the administration of adjuvant treatment after surgery, showing that the treated patients had longer OS than non-treated patients, which included most stage I patients. With this context, we decided to evaluate whether our piRNA could be of utility to discriminate patients at high risk of relapse who are not initially considered for adjuvant treatment only taking into account clinical characteristics. Interestingly, we observed that patients not receiving adjuvant treatment with higher piRNA levels had more recurrence and shorter survival, which suggests that the piRNA could be used as a complementary tool to choose patients tributaries of receiving adjuvant treatment after surgery in this group. Supporting this, the DCA analysis showed that the piRNA had a superior net benefit compared to the tumor stage in predicting postsurgical relapse. In this line, Hsiao S. H. et al. have shown that platinum-based adjuvant chemotherapy provides a survival benefit for patients with surgically resected early-stage SCC, but they did not observe significant differences for ADK [61]. Under the current TNM classification system that does not separate by histology, certain SCC patients with poor prognoses are grouped together with ADK patients with good prognoses, and these groups do not receive subsequent treatments. In other tumors, such as esophageal cancer, the last classification system analyzes early-stage patients with different histology separately, which improves patient outcomes [62]. While this approach

would undoubtedly increase complexity, it could potentially lead to more tailored and effective treatment strategies. Therefore, it would be crucial to treat all SCC patients at high risk of relapse independently of the stage, and piR\_004530 can be of utility in this labor.

To confirm that the role as a biomarker of piR\_004530 was related to its oncogenic potential, we performed several in vitro analyses in lung cancer cell lines and in patient-derived lung cancer organoids. The overexpression of the piRNA in the lung cancer cell lines promoted migration and invasion. Moreover, in two of the cell lines, SK-MES-1 and H23, reduced apoptosis, while not significant differences were observed for HCC-15 cell line. Additionally, colony formation was enhanced in HCC15 and H23 cell lines, showing that the overexpression of the piwiRNA impacts clonogenic capacity and long-term cell growth. The potential explanation of the differences observed between cell lines can be associated with intrinsic differences in the molecular characteristics of the different cell lines. However, to our knowledge, this is the first time that the oncogenic role of piR\_004530 has been reported. However, in the literature, other piRNAs have been previously identified with a pathogenic role in NSCLC [63], including some of the most studied ones such as piR-651. PiR-651 has been described also as upregulated in NSCLC and playing a function promoting proliferation, inhibiting apoptosis, and even regulating cell migration [64, 65] in a similar way to our piRNA.

Since it has been previously shown that the PIWI/piRNA pathway is involved in the control of stemness [66], we decided to evaluate whether the overexpression of piR\_004530 had some impact on stem cell markers and in 3D spheroid formation. Cancer stem cells tend to express stem cell markers such as *SOX2*, *NANOG*, *CD133*, and *CD44*, among others [63]. Interestingly, we observed that in both cell lines, an upregulation of the *SOX2* gene was observed (we also validated this at protein level), and an upregulation of *NANOG* and *OCT4* was also observed in the SK-MES-1 cell line. *SOX2* has been described to be related to SCC histology [67], and its downregulation inhibits proliferation and induces apoptosis in NSCLC cell lines [68, 69], which is in line with the effect observed with our piRNA. Moreover, in relation to *OCT4* and *NANOG* upregulation, it has been reported that in both cases, higher levels of *OCT4* [70] or *NANOG* [71] were associated with worse OS in NSCLC patients. Additionally, its coexpression has been related to the induction of cancer stem cell properties and malignancy [72]. The reactivation of these stem cell genes after induction of piR\_004530 is in line with the promotion of a more aggressive phenotype. Moreover, we observed that piR\_004530 overexpression was associated with the formation of larger

spheroids, which confirms the functional relevance of this piRNA in regulating stemness. Recently, it has been argued that piRNAs may play a potential role in cancer stem cells and contribute to the tumorigenesis process [35]. Although a limited number of studies have evaluated the role of piRNAs in cancer stem cell regulation, some piRNAs have been identified as upregulated, such as piR\_823 [73] and piR-932 [74] in breast cancer stem cells, or MW557525, which is necessary to reprogram fibroblasts to cancer stem cells [75].

Finally, we validated the oncogenic role of the piRNA in patient-derived samples using organoids, which are stem cell-based models [46]. We generated patient-derived lung cancer organoids from three SCC patients and evaluated cell viability after the overexpression of the piRNA. In all cases, the piRNA overexpression boosted cell viability, demonstrating that the piRNA is involved in the regulation of the cancer stem cells and the regulation of the organoid growth in a 3D context.

To gain preliminary insight into the potential mechanism through which piR\_004530 exerts its oncogenic effects, we performed an exploratory analysis using previously generated expression array data. By comparing patients with high piR\_004530 expression to those lacking its expression, we identified a set of DEGs, many of which were enriched in biological processes related to metabolism, cell communication, and stress response. Additionally, a correlation analysis revealed several negatively associated genes, including SERPINB6 and GAK, both of which showed significant prognostic value in lung cancer with lower levels associated with shorter OS. SERPINB6 is an intracellular serine protease inhibitor that regulates proteolytic activity, protecting cells from damage due to uncontrolled protease action. By inhibiting specific serine proteases, it contributes to maintaining cellular homeostasis and preventing excessive stress or apoptosis. In various cancers, SERPINB6 is often downregulated [76], potentially leading to enhanced protease activity and promoting tumor cell invasion, migration, and progression. Its loss may also facilitate extracellular matrix degradation and impair the cellular stress response, contributing to a more aggressive phenotype [77]. In lung cancer specifically, lower SERPINB6 expression has been associated with increased tumor aggressiveness and a poorer prognosis [78]. These findings suggest that piR\_004530 may exert part of its function through the repression of specific mRNA targets, such as SERPINB6 and GAK. While these results offer valuable clues, we acknowledge that elucidating the precise molecular mechanism of piR\_004530's action requires further investigation and functional validation, which we propose as a focus for future studies.

Finally, we also identified positively correlated genes with our piRNA, and some of them have been previously validated as cancer-testis tumor-reactivated genes. In this sense, our data support the hypothesis that piR\_004530 is part of a germline-associated transcriptional program that becomes reactivated in some patients during the carcinogenesis process.

## Conclusions

In summary, we have shown that piR\_004530 became reactivated in NSCLC and played a role as an oncogene, enhancing cell viability and promoting migration and invasion in both cell lines and patient-derived organoids. Moreover, this piRNA is a promising biomarker for SCC patients, since its higher levels are associated with a higher probability of disease recurrence after surgery and shorter OS. Although further investigation in additional cohorts of patients is warranted to validate these findings, we believe that we have identified a promising biomarker that can serve as a useful tool to guide treatment decisions regarding adjuvant therapy in SCC patients. Since no other studies detail in-depth research on this piRNA, we will continue to explore the relevant mechanisms involved in the oncogenic effect of this piRNA in NSCLC.

## Abbreviations

|            |                                            |
|------------|--------------------------------------------|
| NSCLC      | Non-small cell lung cancer                 |
| ADK        | Adenocarcinoma                             |
| SCC        | Squamous cell carcinoma                    |
| EGFR       | Epidermal growth factor receptor           |
| ALK        | Anaplastic lymphoma kinase                 |
| MET        | Hepatocyte growth factor receptor          |
| KRAS       | Kirsten rat sarcoma viral oncogene homolog |
| piRNAs     | PIWI-interacting RNAs                      |
| piR_004530 | Hsa_piR_004530                             |
| TEX29      | Testis-expressed protein 29                |
| OS         | Overall survival                           |
| RPMI       | Roswell Park Memorial Institute            |
| DMEM       | Dulbecco's Modified Eagle Medium           |
| FBS        | Fetal bovine serum                         |
| RT         | Room temperature                           |
| TTR        | Time to relapse                            |
| DFS        | Disease-free survival                      |
| IQR        | Interquartile range                        |
| AUC        | Area under the ROC curve                   |
| CI         | Confidence interval                        |
| HR         | Hazard ratio                               |

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-025-04342-1>.

Additional file 1: Figures S1-S12. Fig. S1. Correlation analysis between TEX29 and piR\_004530 expression in tumor tissue from 75 NSCLC patients. Expression analysis of piR\_004530 and TEX29 after 5 days of 100uM 5-AZA-dC treatment, a DNA methyl transferase inhibitor. Expression analysis of TEX29 levels in cell lines and organoids transfected with piR\_004530 mimic or control. Fig. S2. Clinicopathological factors impacting DFS and OS in the SCC and non-SCC patients. Fig. S3. Analysis of the impact of piR\_004530 expression levels on DFS and OS in Non-SCC patients, ADK

patients, and other patients. Fig. S4. Main clinical and pathological characteristics of the independent NSCLC validation cohort ( $n = 74$ , 2019–2023). The median follow-up time was 46 months. *This table supports the survival analyses shown in Fig. S5.* Fig. S5. Kaplan–Meier survival analysis in an independent cohort of 74 NSCLC patients. Patients were stratified by piR\_004530 expression levels. Disease-free survival and overall survival in the whole validation cohort. DFS and OS in the SCC subgroup of the validation cohort. High piR\_004530 expression was associated with poorer outcomes, particularly in SCC patients. Log-rank  $p$ -values are indicated in each panel. Due to the limited sample size of the validation cohort, these results should be interpreted with caution. Fig. S6. Kaplan–Meier plot of a DFS or b OS of SCC cases stratified in 3 groups: 1) patients receiving adjuvant treatment; 2) patients not receiving adjuvant treatment and with low levels of piR\_004530; and 3) patients not receiving adjuvant treatment and with high levels of piR\_004530.  $P$ -values were calculated using the Log-rank test. Fig. S7. A piR\_004530 expression in several lung cancer cell lines in comparison to the normal immortalized cell line BEAS2B. HCC44, H23, and A549 were derived from ADK patients, while HCC-15 and SK-MES-1 were derived from SCC patients. B Expression levels of piR\_004530 in control and piR\_004530 overexpressed group of SK-MES-1 and HCC-15. C Expression levels of piR\_004530 in control and piR\_004530 silenced group of SK-MES-1 and HCC-15.  $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.0001$ . Fig. S8. Wound healing assay using ASO to inhibit piR\_004530 expression. A Representative image of wound-healing assay in SK-MES-1 cell line transfected with anti-piR\_004530 or control ASO. Photographs were taken at 0, 6, 24, 30 and 48 h. Calculated wound healing areas from 3 independent replicates are shown in the graph. B Representative image of wound-healing assay in HCC-15 cell line transfected with anti-piR\_004530 or control ASO. Photographs were taken at 0, 6, 24, 30, 48 and 72 h. Calculated wound healing areas from 3 independent replicates are shown in the graph. C Comparison of cell viability at 3 days between organoids nucleofected with ASO-piR\_004530 or ASO Control. All data represented are the average of at least 3 independent replicates and error bars represent SEM.  $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.001$ . Fig. S9. MTS analysis in SK-MES-1, HCC15 and H23 control and piR\_004530 cells transfected cells. Cell viability was measured using MTS at 0, 24, 48, 72, 96 and 120 h. Fig. S10. A Hierarchical cluster analysis of the top 100 differential expressed genes identified between patients expressing and not expressing piR\_004530. B Enrichment analysis for Gene Ontology terms using the DEG identified. C Hierarchical clustering based on the correlation value of the top 100 genes negatively correlated with piR\_004530 expression. D Expression correlation graphs of the top 5 negatively correlated genes and the Kaplan–Meier Plot of OS according to SERPINB6 and GAK expression. Fig. S11. Kaplan–Meier Plot of OS according to *TEX29* expression. Data was obtained from the Kaplan–Meier Plotter - lung cancer cohort. For the analysis, the *TEX29* Affymetrix probe 236853\_at was used. The auto-select best cutoff option was used, and 1411 patients were included in the analysis. Fig. S12. Kaplan–Meier plot of DFS and OS according piR\_004530 expression stratified by sex. DFS according piR\_004530 expression in Female patients; OS according piR\_004530 expression in Female patients; DFS according piR\_004530 expression in Male patients; and OS according piR\_004530 expression in Male patients.  $P$ -values were calculated using the Log-rank test.

Additional file 2: Table S1. List of differentially expressed genes obtained by array analysis between the group of patients with high piR\_004530 expression to those without piRNA expression.

Additional file 3: Table S2. List of genes negatively correlated with the expression of our piRNA in the array data.

Additional file 4: Table S3. List of genes positively correlated with the expression of our piRNA in the array data.

## Acknowledgements

We thank all the patients who enrolled in this study. Fig. 8A was created with BioRender.com.

## Authors' contributions

YH: performed research, data analysis, and prepared the manuscript draft. MAP performed research on organoid and spheroid models. TD, AAC, RN, YL, performed research. DSL, NV, DM, MB, AG, LM, and RMM: contributed to the

acquisition of clinical data and performed critical evaluation of the manuscript. AN designed and directed the study, performed the statistical analysis, and coordinated throughout the manuscript writing. All authors read and approved the final manuscript.

## Funding

The Ministry of Economy and Competition (MINECO) co-financed with the European Union FEDER funds (SAF2017-88606-P, 2017), SEPAR-AstraZeneca Ayudas Investigación PII Oncología 2021, Becas SEPAR 2022 (no. de Proyecto 1326–2022), and Becas SEPAR 2023 (no. de Proyecto 1482–2023). Becas SEPAR 2023, 1482–2023, SEPAR-AstraZeneca Ayudas Investigación PII Oncología 2021, Becas SEPAR 2022, 1326–2022, Ministry of Economy and Competition (MINECO) co-financed with the European Union FEDER funds, SAF2017-88606-P.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Declarations

### Ethics approval and consent to participate

The protocol for patient sample collection was supervised and approved by the Clinical Research Ethics Committee of the Hospital Clinic of Barcelona (project approval number HCB/2013/8230) and the Bioethics Commission of the University of Barcelona (project approval number IRB00003099) (Barcelona, Spain). Written informed consent was obtained from each participant in accordance with the Declaration of Helsinki. Clinical trial number is not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Human Anatomy Unit, Molecular Oncology and Embryology Laboratory, Department of Surgery and Surgical Specializations, Faculty of Medicine and Health Sciences, Universitat de Barcelona (UB), Casanova 143, 08036 Barcelona, Spain. <sup>2</sup>School of Basic Medical Sciences, Chengdu University, Chengdu 610106, China. <sup>3</sup>Department of Thoracic Surgery, Hospital Clínic de Barcelona, University of Barcelona, 08036 Barcelona, Spain. <sup>4</sup>Thoracic Oncology Unit, Hospital Clinic, Barcelona, Spain. <sup>5</sup>Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), C. Villarreal, 170, 08036 Barcelona, Spain. <sup>6</sup>Department of Medical Oncology, Institut Clínic de Malalties Hemato-Oncològiques (ICMHO), Hospital Clínic de Barcelona, 08036 Barcelona, Spain. <sup>7</sup>Department of Pathology, Hospital Clínic de Barcelona, University of Barcelona, 08036 Barcelona, Spain. <sup>8</sup>Department of Pneumology, Institut Clínic Respiratori (ICR), Hospital Clínic de Barcelona, University of Barcelona, 08036 Barcelona, Spain. <sup>9</sup>Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, 28029 Madrid, Spain.

Received: 20 January 2025 Accepted: 13 August 2025

Published online: 07 October 2025

## References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin.* 2021;71(1):7–33.
2. Kadasah SF. Prognostic significance of glycolysis-related genes in lung squamous cell carcinoma. *Int J Mol Sci.* 2024;25(2):1143.
3. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74(3):229–63.
4. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one

- of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*. 2018;391(10125):1023–75.
5. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49.
6. Remon J, Soria J-C, Peters S. Early and locally advanced non-small-cell lung cancer: an update of the ESMO Clinical Practice Guidelines focusing on diagnosis, staging, systemic and local therapy. *Ann Oncol*. 2021;32(12):1637–42.
7. Schuchert MJ, Normolle DP, Awais O, Pennathur A, Wilson DO, Luketich JD, et al. Factors influencing recurrence following anatomic lung resection for clinical stage I non-small cell lung cancer. *Lung Cancer*. 2019;128:145–51.
8. Fink-Neuboeck N, Lindenmann J, Porubsky C, Fediuk M, Anegg U, Maier A, et al. Hazards of recurrence, second primary, or other tumor at ten years after surgery for non-small-cell lung cancer. *Clin Lung Cancer*. 2020;21(4):333–40.
9. Rami-Porta R, Bolejack V, Giroux DJ, Chansky K, Crowley J, Asamura H, et al. The IASLC Lung Cancer Staging Project: the new database to inform the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2014;9(11):1618–24.
10. Schabath MB, Cote ML. Cancer progress and priorities: lung cancer. *Cancer Epidemiol Biomarkers Prev*. 2019;28(10):1563–79.
11. Hendriks L, Kerr K, Menis J, Mok T, Nestle U, Passaro A, et al. Oncogene-addicted metastatic non-small-cell lung cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2023;34(4):339–57.
12. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci*. 2004;101(36):13306–11.
13. Cardarella S, Johnson BE. The impact of genomic changes on treatment of lung cancer. *Am J Respir Crit Care Med*. 2013;188(7):770–5.
14. De Luca A, Maiello MR, D’Alessio A, Pergameno M, Normanno N. The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opin Ther Targets*. 2012;16(sup2):S17–27.
15. Derman BA, Mileham KF, Bonomi PD, Batus M, Fidler MJ. Treatment of advanced squamous cell carcinoma of the lung: a review. *Translational lung cancer research*. 2015;4(5):524.
16. Chen JW, Dhahbi J. Lung adenocarcinoma and lung squamous cell carcinoma cancer classification, biomarker identification, and gene expression analysis using overlapping feature selection methods. *Sci Rep*. 2021;11(1):13323.
17. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489(7417):519.
18. Santos ES, Hart L. Advanced squamous cell carcinoma of the lung: current treatment approaches and the role of afatinib. *Onco Targets Ther*. 2020;9305–21.
19. Friedlaender A, Banna G, Malapelle U, Pisapia P, Addeo A. Next generation sequencing and genetic alterations in squamous cell lung carcinoma: where are we today? *Front Oncol*. 2019;9:166.
20. Rodak O, Peris-Díaz MD, Olbromski M, Podhorska-Okolów M, Dziegiel P. Current landscape of non-small cell lung cancer: epidemiology, histological classification, targeted therapies, and immunotherapy. *Cancers (Basel)*. 2021;13(18):4705.
21. Zhang C, Sun C, Zhao Y, Wang Q, Guo J, Ye B, et al. Overview of MicroRNAs as diagnostic and prognostic biomarkers for high-incidence cancers in 2021. *Int J Mol Sci*. 2022;23(19):11389.
22. Zhu X, Kudo M, Huang X, Sui H, Tian H, Croce CM, et al. Frontiers of MicroRNA signature in non-small cell lung cancer. *Frontiers in Cell and Developmental Biology*. 2021;9: 643942.
23. Taverna S, Masucci A, Cammarata G. PIWI-RNAs small noncoding RNAs with smart functions: potential theranostic applications in cancer. *Cancers (Basel)*. 2023;15(15):3912.
24. Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, et al. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev*. 2006;20(13):1732–43.
25. Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev*. 2006;20(13):1709–14.
26. Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*. 2006;442(7099):199–202.
27. Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*. 2006;442(7099):203–7.
28. Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*. 2007;316(5825):744–7.
29. Berrier AL, Yamada KM. Cell–matrix adhesion. *J Cell Physiol*. 2007;213(3):565–73.
30. Gan B, Chen S, Liu H, Min J, Liu K. Structure and function of eTudor domain containing TDRD proteins. *Crit Rev Biochem Mol Biol*. 2019;54(2):119–32.
31. Weick E-M, Miska EA. piRNAs: from biogenesis to function. *Development*. 2014;141(18):3458–71.
32. Navarro A, Tejero R, Viñolas N, Cordeiro A, Marrades RM, Fuster D, et al. The significance of PIWI family expression in human lung embryogenesis and non-small cell lung cancer. *Oncotarget*. 2015;6(31):31544.
33. Jian Z, Han Y, Li H. Potential roles of PIWI-interacting RNAs in lung cancer. *Front Oncol*. 2022;12: 944403.
34. Liang D, Yang Y, Liu Y. The role Hiwi gene in the maintenance of lung cancer stem cell populations. *Neoplasma*. 2013.
35. Garcia-Borja E, Siegl F, Mateu R, Slaby O, Sedo A, Busek P, et al. Critical appraisal of the piRNA-PIWI axis in cancer and cancer stem cells. *Bio-marker Research*. 2024;12(1):15.
36. Ozata DM, Gainetdinov I, Zoch A, O’Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet*. 2019;20(2):89–108.
37. Cheng J, Guo J-M, Xiao B-X, Miao Y, Jiang Z, Zhou H, et al. piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clin Chim Acta*. 2011;412(17–18):1621–5.
38. Law PT-Y, Qin H, Ching AK-K, Lai KP, Co NN, He M, et al. Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. *J Hepatol*. 2013;58(6):1165–73.
39. Zhang W, Zheng Z, Wang K, Mao W, Li X, Wang G, et al. piRNA-1742 promotes renal cell carcinoma malignancy by regulating USP8 stability through binding to hnRNP and thereby inhibiting MUC12 ubiquitination. *Exp Mol Med*. 2023;55(6):1258–71.
40. Cordeiro A, Navarro A, Gaya A, Díaz-Beyá M, Gonzalez-Farré B, Castellano JJ, et al. PiwiRNA-651 as marker of treatment response and survival in classical Hodgkin lymphoma. *Oncotarget*. 2016;7(29):46002–13.
41. He Y, Altuna-Coy A, Acosta-Plasencia M, Molins L, Sánchez-Lorente D, Martínez D, et al. Diagnostic and prognostic value of hsa\_piR\_022710, hsa\_piR\_019822, and hsa\_piR\_020840 in early-stage non-small-cell lung cancer: implications for recurrence and survival in squamous cell carcinoma patients. *Int J Mol Sci*. 2025;26(7):2870.
42. Bellil H, Ghieh F, Hermel E, Mandon-Pepin B, Vialard F. Human testis-expressed (TEX) genes: a review focused on spermatogenesis and male fertility. *Basic and clinical andrology*. 2021;31(1):9.
43. Wang R, Zhu H, Yang M, Zhu C. DNA methylation profiling analysis identifies a DNA methylation signature for predicting prognosis and recurrence of lung adenocarcinoma. *Oncol Lett*. 2019;18(6):5831–42.
44. Bruggeman JW, Koster J, van Pelt AM, Speijer D, Hamer G. How germline genes promote malignancy in cancer cells. *BioEssays*. 2023;45(1):2200112.
45. Acosta-Plasencia M, Castellano JJ, Díaz T, He Y, Marrades RM, Navarro A. Discovering genes and microRNAs involved in human lung development unveils IGFBP3/miR-34a dynamics and their relevance for alveolar differentiation. *Stem Cell Res Ther*. 2024;15(1):263.
46. Acosta-Plasencia M, Castellano JJ, Taylor AM, He Y, Martínez D, Monzo M, et al. Patient-derived organoids as a stem cell model to study lung cancer. 2022.
47. Sachs N, Papaspyropoulos A, Zomer-van Ommen DD, Heo I, Böttinger L, Klay D, et al. Long-term expanding human airway organoids for disease modeling. *EMBO J*. 2019;38(4): e100300.
48. Canals J, Navarro A, Vila C, Canals JM, Díaz T, Acosta-Plasencia M, et al. Human embryonic mesenchymal lung-conditioned medium promotes differentiation to myofibroblast and loss of stemness phenotype in lung adenocarcinoma cell lines. *J Exp Clin Cancer Res*. 2022;41(1):37.

49. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak*. 2008;8:1–17.
50. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10(21):7252–9.
51. Almeida LG, Sakabe NJ, deOliveira AR, Silva MCC, Mundstein AS, Cohen T, et al. CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Res*. 2009;37(suppl\_1):D816–D9.
52. Rajaram R, Huang Q, Li RZ, Chandran U, Zhang Y, Amos TB, et al. Recurrence-free survival (RFS) in surgically-resected non-small cell lung cancer (NSCLC) patients—a systematic literature review and meta-analysis. *Chest*. 2023.
53. Kris MG, Gaspar LE, Chaft JE, Kennedy EB, Azzoli CG, Ellis PM, et al. Adjuvant systemic therapy and adjuvant radiation therapy for stage I to IIIA completely resected non-small-cell lung cancers: American Society of Clinical Oncology/Cancer Care Ontario clinical practice guideline update. *J Clin Oncol*. 2017;35(25):2960–74.
54. Kato H, Ichinose Y, Ohta M, Hata E, Tsubota N, Tada H, et al. A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. *N Engl J Med*. 2004;350(17):1713–21.
55. Pignon J-P, Tribodet H, Scagliotti GV, Douillard J-Y, Shepherd FA, Stephens RJ, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *Database of Abstracts of Reviews of Effects (DARE): Quality-Assessed Reviews [Internet]: Centre for Reviews and Dissemination (UK); 2008*.
56. Wang W, Liu H, Li G. What's the difference between lung adenocarcinoma and lung squamous cell carcinoma? Evidence from a retrospective analysis in a cohort of Chinese patients. *Front Endocrinol (Lausanne)*. 2022;13:947443.
57. Shepherd CJ, Rizzo S, Ledaki I, Davies M, Brewer D, Attard G, de Bono J, Hudson DL. Expression profiling of CD133 + and CD133 — epithelial cells from human prostate Abstract *The Prostate*. 2008;68(9) 1007–1024. <https://doi.org/10.1002/pros.20765>.
58. Györfy B. Transcriptome-level discovery of survival-associated biomarkers and therapy targets in non-small-cell lung cancer. *Br J Pharmacol*. 2024;181(3):362–74.
59. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of Malignant Tumours. *J Thorac Oncol*. 2007;2(8):706–14.
60. Fukui T, Taniguchi T, Kawaguchi K, Fukumoto K, Nakamura S, Sakao Y, et al. Comparisons of the clinicopathological features and survival outcomes between lung cancer patients with adenocarcinoma and squamous cell carcinoma. *Gen Thorac Cardiovasc Surg*. 2015;63:507–13.
61. Hsiao SH, Chen WT, Chung CL, Chou YT, Lin SE, Hong SY, et al. Comparative survival analysis of platinum-based adjuvant chemotherapy for early-stage squamous cell carcinoma and adenocarcinoma of the lung. *Cancer Med*. 2022;11(10):2067–78.
62. Rice TW, Blackstone EH, Rusch VW. of the AJCC Cancer Staging Manual: esophagus and esophagogastric junction. *Ann Surg Oncol*. 2010;17:1721–4.
63. Zheng Y, Wang L, Yin L, Yao Z, Tong R, Xue J, et al. Lung cancer stem cell markers as therapeutic targets: an update on signaling pathways and therapies. *Front Oncol*. 2022;12: 873994.
64. Zhang SJ, Yao J, Shen BZ, Li GB, Kong SS, Bi DD, et al. Role of piwi-interacting RNA-651 in the carcinogenesis of non-small cell lung cancer. *Oncol Lett*. 2018;15(1):940–6.
65. Li D, Luo Y, Gao Y, Yang Y, Wang Y, Xu Y, et al. piR-651 promotes tumor formation in non-small cell lung carcinoma through the upregulation of cyclin D1 and CDK4. *Int J Mol Med*. 2016;38(3):927–36.
66. Rojas-Ríos P, Simonelig M. piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. *Development*. 2018;145(17):dev161786.
67. Yuan P, Kadara H, Behrens C, Tang X, Woods D, Solis LM, et al. Sex determining region Y-box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS ONE*. 2010;5(2): e9112.
68. Xiang R, Liao D, Cheng T, Zhou H, Shi Q, Chuang T, et al. Downregulation of transcription factor SOX2 in cancer stem cells suppresses growth and metastasis of lung cancer. *Br J Cancer*. 2011;104(9):1410–7.
69. Choe C, Kim H, Min S, Park S, Seo J, Roh S. SOX2, a stemness gene, induces progression of NSCLC A549 cells toward anchorage-independent growth and chemoresistance to vinblastine. *Onco Targets Ther*. 2018;6:197–207.
70. Li X, Wang J, Xu Z, Ahmad A, Li E, Wang Y, et al. Expression of Sox2 and Oct4 and their clinical significance in human non-small-cell lung cancer. *Int J Mol Sci*. 2012;13(6):7663–75.
71. Du Y, Ma C, Wang Z, Liu Z, Liu H, Wang T. Nanog, a novel prognostic marker for lung cancer. *Surg Oncol*. 2013;22(4):224–9.
72. Chiou S-H, Wang M-L, Chou Y-T, Chen C-J, Hong C-F, Hsieh W-J, et al. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res*. 2010;70(24):10433–44.
73. Ding X, Li Y, Lü J, Zhao Q, Guo Y, Lu Z, et al. piRNA-823 is involved in cancer stem cell regulation through altering DNA methylation in association with luminal breast cancer. *Frontiers in cell and developmental biology*. 2021;9: 641052.
74. Zhang H, Ren Y, Xu H, Pang D, Duan C, Liu C. The expression of stem cell protein Piwi2 and piR-932 in breast cancer. *Surg Oncol*. 2013;22(4):217–23.
75. Jin L, Zhang Z, Wang Z, Tan X, Wang Z, Shen L, et al. Novel piRNA MW557525 regulates the growth of Piwi2-iCSCs and maintains their stem cell pluripotency. *Mol Biol Rep*. 2022;49(7):6957–69.
76. Shiiba M, Nomura H, Shinozuka K, Saito K, Kouzu Y, Kasamatsu A, et al. Down-regulated expression of SERPIN genes located on chromosome 18q21 in oral squamous cell carcinomas. *Oncol Rep*. 2010;24(1):241–9.
77. Fonseca JAR. Exploring the role of proteolysis in extracellular matrix remodeling: links to chronic obstructive pulmonary disease and lung cancer: Universidade NOVA de Lisboa (Portugal); 2017.
78. Chou R-H, Wen H-C, Liang W-G, Lin S-C, Yuan H-W, Wu C-W, et al. Suppression of the invasion and migration of cancer cells by SERPINB family genes and their derived peptides. *Oncol Rep*. 2012;27(1):238–45.

# Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.