



OPEN ACCESS

EDITED BY

Elisa Frullanti,
University of Siena, Italy

REVIEWED BY

Xinpei Deng,
Sun Yat-sen University Cancer Center
(SYSUCC), China
Zainab Siddiqui,
Era University, India

*CORRESPONDENCE

Michelle E. Pewarchuk,
✉ mpewarchuk@bccrc.ca

RECEIVED 14 April 2025

ACCEPTED 01 August 2025

PUBLISHED 29 August 2025

CITATION

Pewarchuk ME, Souza VGP, Cohn DE, Telkar N,
Stewart GL, Bénard KH, Reis PP, Martinez VD,
Robinson WP and Lam WL (2025) Identification
of oncofetal PIWI-interacting RNAs as potential
prognostic biomarkers in non-small cell
lung cancer.

Front. Genet. 16:1611805.

doi: 10.3389/fgene.2025.1611805

COPYRIGHT

© 2025 Pewarchuk, Souza, Cohn, Telkar,
Stewart, Bénard, Reis, Martinez, Robinson and
Lam. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Identification of oncofetal PIWI-interacting RNAs as potential prognostic biomarkers in non-small cell lung cancer

Michelle E. Pewarchuk^{1,2*}, Vanessa G. P. Souza¹, David E. Cohn^{1,2},
Nikita Telkar^{1,3,4}, Greg L. Stewart^{1,2}, Katya H. Bénard^{1,2},
Patricia P. Reis⁵, Victor D. Martinez^{6,7}, Wendy P. Robinson^{3,4} and
Wan L. Lam^{1,2}

¹British Columbia Cancer Research Institute, Vancouver, BC, Canada, ²Interdisciplinary Oncology Program, University of British Columbia, Vancouver, BC, Canada, ³Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada, ⁴British Columbia Children's Hospital Research Institute, Vancouver, BC, Canada, ⁵Department of Surgery and Orthopedics, Faculty of Medicine, São Paulo State University (UNESP), Botucatu, São Paulo, Brazil, ⁶Department of Pathology and Laboratory Medicine, IWK Health Centre, Halifax, NS, Canada, ⁷Department of Pathology, Dalhousie University, Halifax, NS, Canada

Lung cancer is the leading cause of cancer-related deaths worldwide, with non-small cell lung cancer (NSCLC) accounting for the majority of these cases. Despite advancements in targeted therapies, early detection remains a significant challenge, highlighting the need for novel biomarkers. This study investigates the role of PIWI-interacting RNAs (piRNAs) in lung cancer, specifically focusing on their potential as oncofetal biomarkers in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), the two most common histological subtypes of NSCLC. We hypothesize that piRNAs exhibit oncofetal expression patterns and may contribute to lung cancer development. Through bioinformatics analysis, we identified distinct piRNA profiles in non-neoplastic, malignant, and fetal lung tissues. Among these, 37 piRNAs in LUAD and 46 piRNAs in LUSC displayed oncofetal expression, meaning they were present in tumor tissues but absent in adjacent normal lung tissue. These oncofetal piRNAs showed significant prognostic value in both LUAD and LUSC cohorts, with a specific signature of eight oncofetal piRNAs predicting high-risk patients in LUAD. We validated the robustness of this signature in a separate in-house cohort, which underscores its potential as a prognostic biomarker. Our findings suggest that oncofetal piRNAs could offer new diagnostic and therapeutic opportunities, particularly for early detection.

KEYWORDS

lung cancer, PIWI-interacting RNAs, oncofetal biomarkers, lung adenocarcinoma, lung squamous cell carcinoma, non-coding RNAs

1 Introduction

Lung cancer remains one of the most commonly diagnosed cancers and the leading cause of cancer-related mortality worldwide, making it a priority for research aimed at identifying novel therapeutic targets and biomarkers (Siegel et al., 2024). Lung cancer is broadly classified into two main histological types: small cell lung cancer (SCLC) and non-

small cell lung cancer (NSCLC), with NSCLC accounting for approximately 85% of cases. NSCLC is further divided into adenocarcinoma (LUAD), squamous cell carcinoma (LUSC), and large cell carcinoma (LCC), with LUAD and LUSC being the predominant subtypes (Herbst et al., 2018). These subtypes differ in their cell of origin: LUAD typically arises from distal glandular cells in the alveoli, while LUSC originates from the proximal airway epithelium (Cheung and Nguyen, 2015). Molecular advances have identified key, targetable genomic alterations in NSCLC (particularly LUAD), including *EGFR*, *BRAF*, and *KRAS* mutations, and *ALK* translocations, leading to the development of targeted therapies as well as diagnostic markers (Cancer Genome Atlas Research Network, 2014; Chen et al., 2014; Du et al., 2018; Herbst et al., 2018; Tan and Tan, 2022). However, lung cancer is diagnosed predominantly at advanced stages and these treatments are prone to resistance, resulting in poor survival outcomes. Therefore, strengthening our understanding of the molecular pathogenesis of lung cancer and identifying novel molecules involved in tumorigenesis can drive the development of effective treatments and biomarkers, ultimately improving patient prognosis (Herbst et al., 2018; Cassim et al., 2019).

Increasing evidence supports that oncofetal reprogramming, where tumor cells co-opt signaling pathways typically active only during embryonic development, has significant implications for cancer biology and potential therapeutic strategies (Sharma et al., 2022; Cao et al., 2023). Oncofetal reprogramming has been observed in various cancers where gene expression patterns resembling early developmental stages of the corresponding organ have been identified in tumor profiles (Coulouarn et al., 2005; Huang et al., 2019; Hassan et al., 2009; Feng et al., 2014; Hu and Shivdasani, 2005). Some oncofetal molecules have entered clinical application as early diagnostic markers, therapeutic targets, or prognostic indicators for various tumors. For example, alpha-fetoprotein (AFP) and carcinoembryonic antigen (CEA) serve as diagnostic biomarkers for hepatocellular carcinoma (Mizejewski, 2002; Chen et al., 2020), and colorectal cancer (Das et al., 2017; Feng et al., 2018), respectively.

Small non-coding RNAs (sncRNAs) represent a diverse class of RNA molecules that, despite lacking protein-coding potential, play crucial roles in gene regulation (Fu, 2014). They include microRNAs (miRNAs), transfer RNA-derived fragments (tRFs), small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs) (Fu, 2014). Our group previously demonstrated that miRNAs exhibit oncofetal expression patterns in lung cancer (Cohn et al., 2021). Furthermore, we showed that oncofetal miRNAs regulate NFIB, a transcription factor essential for lung differentiation and maturation during development (Becker-Santos et al., 2016).

While miRNAs have been extensively studied, piRNAs remain largely unexplored in lung cancer, despite their abundance, evolutionary conservation, and established roles in gene regulation and development (Mei et al., 2013). Initially characterized for their function in transposon regulation and germline development (Halic and Moazed, 2009; Iwasaki et al., 2015; Teixeira et al., 2017), piRNAs have shown expression in multiple somatic tissues and cancers, where they may contribute to tumorigenesis. A recent pan-cancer analysis revealed widespread piRNA expression across normal and malignant tissues from

11 organ types, with distinct tissue-specific profiles, particularly in the thyroid and prostate (Martinez et al., 2015). Certain piRNAs have been associated with patient survival, suggesting their potential as cancer biomarkers (Martinez et al., 2015). However, their functional significance in lung cancer remains largely unknown, and few piRNAs have been characterized to date, highlighting the need for further investigation.

We sought to explore whether other ncRNA classes, such as piRNAs, might also exhibit oncofetal characteristics and contribute to lung cancer progression. In this study, we characterized the piRNA landscape in LUAD and LUSC, focusing on their potential role as oncofetal markers. Like miRNAs, we hypothesize that specific piRNAs may display oncofetal expression patterns, potentially reflecting their involvement in cancer development. We utilized a bioinformatics framework previously applied to miRNA discovery to identify and classify piRNAs based on their expression patterns in non-neoplastic, malignant, and fetal lung tissues. Furthermore, we evaluated the association of oncofetal piRNAs with patient survival and established an oncofetal piRNA prognostic signature for LUAD.

2 Methods

2.1 Data collection

This study utilized four cohorts: The Cancer Genome Atlas (TCGA) cohort, which included paired tumor and adjacent adult non-neoplastic lung tissue (ANL) samples for LUAD and LUSC; a fetal lung (FL) cohort comprising human fetal lung samples; and a validation cohort from the British Columbia Cancer Agency (BCCA) containing LUAD and paired ANL tissue samples. Small RNA sequencing (small RNA-seq) data for paired samples (tumor and ANL) in the TCGA LUAD (n = 46 pairs) and TCGA LUSC (n = 45 pairs) cohorts were downloaded as binary alignment map (BAM) files from the TCGA data repository (The Cancer Genome Atlas Research Network et al., 2013; National Cancer Institute, 2019), and processed to extract small RNA expression profiles. The FL cohort consisted of small RNA-seq data for 25 human fetal lung samples, obtained as described in (Cohn et al., 2021) and available in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) repository under accession number GSE175462 (Edgar et al., 2002). The BCCA validation cohort comprised 58 paired LUAD and ANL samples obtained with informed written consent from patients at Vancouver General Hospital, with approval from the University of British Columbia/BC Cancer Agency Research Ethics Board (H15-03060), as previously described (Cohn et al., 2021). Small RNA-seq data for BCCA cohort samples were downloaded from the GEO repository under accession number GSE175462 (Edgar et al., 2002). All cohorts excluded samples with fewer than five million reads from the analysis to ensure data quality and reliability. The clinical characteristics of the patients included in each cohort are summarized in Table 1.

2.2 Evaluation of piRNA expression

FASTQ files were obtained by converting BAM files using SAMtools (v1.17) (Danecek et al., 2021) for all cohorts. Reads

TABLE 1 Demographic and clinical characteristics of patients in TCGA LUSC, TCGA LUAD, and BCCA LUAD cohorts.

Characteristics	TCGA-LUSC (n = 45 pairs)	TCGA-LUAD (n = 46 pairs)	BCCA-LUAD (n = 58)
Median Age (range) years	68 (45–85)	67 (47–85)	70 (45–86)
Sex			
Male	32 (71.11%)	20 (43.48%)	12 (20.69%)
Female	13 (28.89%)	26 (56.52%)	36 (62.07%)
Stage			
IA	5 (11.11%)	10 (21.74%)	18 (31.03%)
IB	16 (35.56%)	9 (19.57%)	18 (31.03%)
II	1 (2.22%)	0	0
IIA	7 (15.56%)	8 (17.39%)	2 (3.45%)
IIB	5 (11.11%)	3 (6.52%)	11 (18.97%)
IIIA	3 (6.67%)	3 (6.52%)	4 (6.90%)
IIIB	0	1 (2.17%)	1 (1.72%)
IV	1 (2.22%)	0	1 (1.72%)
N/A	7 (15.56%)	12 (26.09%)	3 (5.2%)

N/A, not available.

with a mean Phred score below 20 were discarded to ensure high-quality data using Trimmomatic (v. 0.39) (Bolger et al., 2014). The quality-filtered FASTQ files were processed using the online platform miRMaster 2.0 (Fehlmann et al., 2017). Default miRMaster 2.0 settings were applied, with the following modifications: a custom protocol incorporating a 3' adapter sequence of ATCTCGTATGCCGTCTTCTGCTTGT, a minimum read length of 15, a sliding window requiring a quality threshold of 1, the STAR alignment algorithm, and a minimum read stack height of 5. Normalization was performed using reads per million (RPM). For patients with duplicate or triplicate tumor samples, expression data were averaged. Following miRMaster 2.0 processing, additional manual expression filtering was performed. For each sample group, piRNAs were considered 'expressed' if detected at ≥ 1 RPM in at least 10% of samples within each tissue type. piRNAs that did not meet this criterion were classified as 'undetected'. Figures representing piRNA expression were generated using R packages: ggplot2 (v. 3.5.1) for data visualization (Wickham, 2016), UpSetR (v. 1.4.0) for intersection analysis (Conway et al., 2017), and pheatmap (v. 1.0.12) (Kolde, 2019). Hierarchical clustering was performed using Euclidean distance (clustering_distance_cols = "euclidean") to measure dissimilarities between samples and Ward's method (clustering_method = "ward.D") for linkage. The resulting dendrogram was used for heatmap representation.

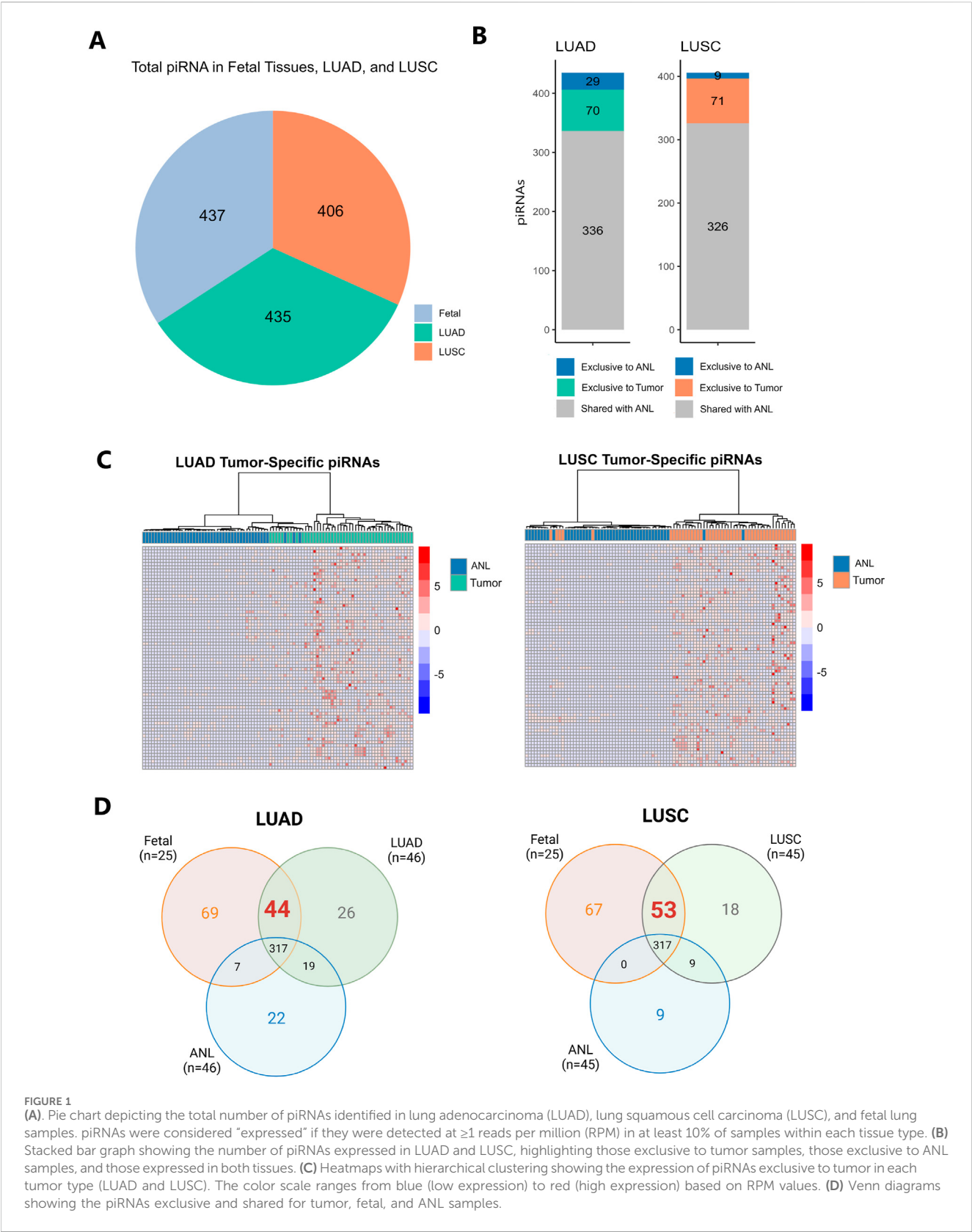
2.3 Differential expression and classification of oncofetal piRNAs

Differential expression was assessed using the Wilcoxon test for each piRNA between two sample groups: tumor vs. ANL and FL vs. ANL. We performed a Wilcoxon test for paired samples for tumor vs. ANL comparisons, whereas, for FL vs. ANL, we used the test for unpaired samples. All statistical analyses were conducted in R Studio (v. 4.3.1) (RStudio Team, 2020) using the wilcox.test function from the stats package (v. 4.4.1) part of R (R Core Team, 2013). The fold change (FC) was calculated by dividing the mean expression of the

first group by that of the second group. The p-values were adjusted using the Benjamini-Hochberg (BH) correction to control the false discovery rate (Benjamini and Hochberg, 1995). piRNAs were considered differentially expressed if they met the following criteria: an adjusted p-value < 0.05 and $FC > 2$. A piRNA was classified as oncofetal if: (a) it was undetected in ANL samples, meaning it did not meet our predefined criteria in ANL (as detailed in section 2.2), and (b) it was overexpressed in both tumor and fetal lung samples compared to ANL (BH $p < 0.05$ and $FC > 2$).

2.4 Construction and validation of oncofetal piRNA prognostic model

Survival analyses were conducted using R Studio (v. 4.3.1) (RStudio Team, 2020) to investigate the association between oncofetal piRNA expression and patient outcomes. All tumor samples from TCGA-LUAD (n = 378) and TCGA-LUSC (n = 470) with available survival information were included in this analysis. Small RNA-seq data were downloaded and processed as described in Sections 2.1 and 2.2. Clinical and survival data for these cohorts were retrieved from the University of California Santa Cruz (UCSC) Xena Browser (Goldman et al., 2020). Survival analysis was also performed on the BCCA-LUAD cohort (n = 58). To assess the prognostic relevance of individual piRNAs, we conducted a univariate Cox proportional hazards regression analysis using the coxph () function from the survival package in R (v. 3.6-4) (Therneau and Grambsch, 2000; Therneau, 2024). The dependent variable was overall survival (OS), defined by survival status and time, while the independent variable was the expression level of each piRNA. The hazard ratio (HR) and p-values were extracted from the Cox model summary, and piRNAs with $p < 0.05$ were considered significantly associated with OS. Following the univariate analysis, a multivariate Cox proportional hazards regression model was constructed to evaluate the combined prognostic value of significant piRNAs. The coxph () function was used, incorporating significant piRNAs as covariates. This model estimated the combined effect of all selected piRNAs on survival.



A risk score was calculated for each patient based on the regression coefficients of the multivariate Cox model and the expression levels of the piRNAs. The formula for the risk score is the sum of the product of each piRNA’s expression level and its corresponding regression coefficient. These risk scores were used to stratify patients into two groups (high and low risk) based on the median risk score.

The survival probability of the two groups (high vs. low risk) was compared using Kaplan-Meier survival curves. The `survfit()` function from the survival package was used to fit the survival curves, and the log-rank test was performed to assess the statistical significance between groups.

2.5 Evaluation of prognostic performance

To evaluate the prognostic accuracy of the risk score, a receiver operating characteristic (ROC) curve was generated using the `survivalROC` package (v. 1.0.3.1) (Heagerty and Saha-Chaudhuri, 2022). The time-dependent ROC analysis was performed at the median OS time. The area under the curve (AUC) was calculated to assess the model's discriminative ability, where AUC = 0.5 indicates no predictive power and AUC = 1.0 indicates perfect discrimination. The concordance index (C-index) was calculated using the `concordance.index()` function from the `survcomp` package (v. 1.54.0) (Haibe-Kains et al., 2008; Schröder et al., 2011) to assess the model's ability to rank patients based on risk scores correctly. The C-index ranges from 0.5 (random prediction) to 1.0 (perfect prediction). To ensure the robustness and stability of the C-index, bootstrap resampling (1,000 iterations) was performed (Efron, 1982). In each iteration, a new dataset was generated by randomly sampling (with replacement) from the original cohort, and the C-index was recalculated. The distribution of bootstrap-derived C-index values was visualized using a density plot, providing insight into variability and confidence in the model's performance.

3 Results

3.1 Comprehensive piRNA expression profiling of human fetal lung tissue, LUSC, and LUAD

To investigate piRNAs potentially involved in both lung development and tumorigenesis, we characterized piRNA expression profiles across human fetal lung tissue, LUAD, LUSC, and matched ANL tissues. piRNAs were considered expressed if they had ≥ 1 RPM in at least 10% of samples within each tissue type (fetal, ANL, and tumor). A total of 437 piRNAs were expressed in fetal lung tissue, 435 in LUAD, and 406 in LUSC (Figure 1A; Supplementary Tables S1–S3). Among the piRNAs identified as expressed, 70 (16.1%) in LUAD and 71 (17.5%) in LUSC were exclusively expressed in tumor samples. Conversely, 29 piRNAs (6.67%) in LUAD and 9 piRNAs (2.22%) in LUSC were uniquely expressed in ANL tissues (Figure 1B). Hierarchical clustering analysis revealed distinct piRNA expression patterns that stratified samples into two clusters corresponding to tumor and ANL types, as demonstrated in the LUAD and LUSC heatmaps (Figure 1C). Further analysis of the tumor-exclusive piRNAs revealed that 31 piRNAs were shared between LUAD and LUSC, while 39 piRNAs were specific to LUAD and 40 were unique to LUSC (Supplementary Table S4). This suggests that some piRNAs may be involved in subtype-specific aspects of cancer biology. In contrast, shared piRNAs between LUAD and LUSC may play critical roles in maintaining essential characteristics of lung cancer. For each subtype of lung cancer, we

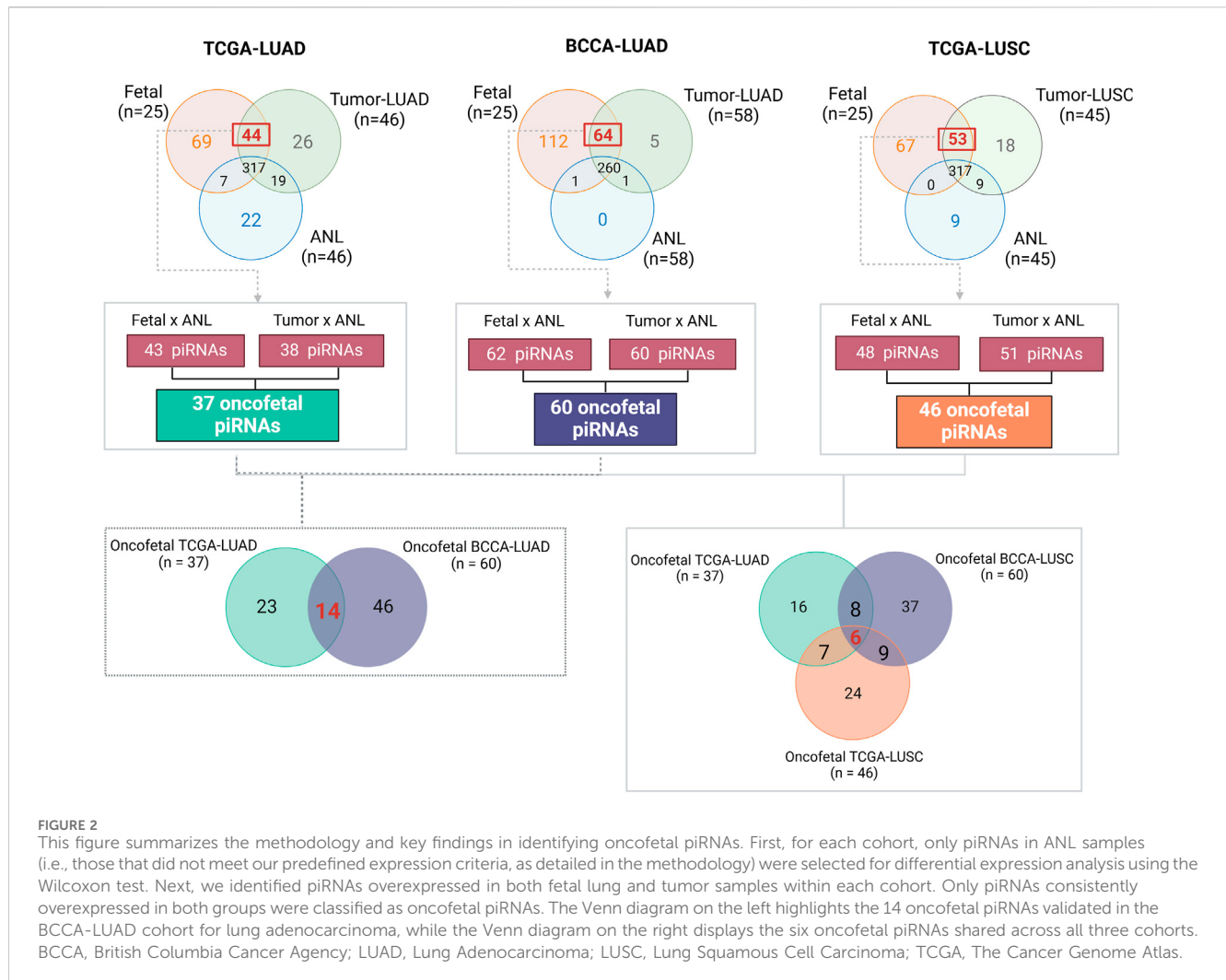
identified specific piRNAs that were expressed in fetal lung tissue but undetected in ANL tissue (indicating they did not meet our expression criteria as defined in the methodology). We found 44 piRNAs in LUAD and 53 in LUSC expressed in both tumor and fetal samples, but undetected in ANL. These piRNAs, exhibiting expression patterns reminiscent of fetal lung, could serve as potential candidates for oncofetal reactivation (Figure 1D). For more details, consult Supplementary Tables S5, 6.

3.2 Oncofetal piRNAs are reactivated in cancer

To identify oncofetal piRNAs, we compared the piRNA expression profiles of human fetal lung samples with those of LUAD, LUSC, and patient-matched non-malignant lung tissue (ANL) (Supplementary Figure S1). We identified 43 piRNAs overexpressed in fetal lung compared to ANL and 38 in tumor compared to ANL (Supplementary Table S7). Of these, 37 were overexpressed in both fetal and tumor samples and were classified as oncofetal piRNAs (Supplementary Table S8). Validation in the BCCA-LUAD cohort confirmed 14 of these 37 oncofetal piRNAs (Supplementary Table S9). In the TCGA-LUSC cohort, 48 piRNAs were overexpressed in fetal lung compared to ANL and 51 in tumor compared to ANL (Supplementary Table S10), leading to the identification of 46 oncofetal piRNAs (Supplementary Table S11). Six oncofetal piRNAs were consistently identified across all three datasets (LUAD, LUSC, and fetal lung), including piR-58186, piR-50725, piR-50446, piR-35175, piR-35174, and piR-30926. Figure 2 provides a detailed methodological overview and summarizes our findings across all datasets. In Supplementary Figure S2, we present heatmaps to illustrate the distinct expression patterns of oncofetal piRNAs across each cohort.

3.3 Oncofetal piRNA prognostic signature predicts high-risk patients in lung adenocarcinoma

Oncofetal piRNAs identified in TCGA-LUAD and TCGA-LUSC were used to construct a prognostic model for each histological subtype. For TCGA-LUAD, univariate Cox regression analysis was performed using 37 oncofetal piRNAs, of which eight showed prognostic significance ($p < 0.05$): piR-41794, piR-44716, piR-44715, piR-33687, piR-34804, piR-33686, piR-33519, and piR-61135 (Supplementary Table S12). Kaplan-Meier (KM) survival analysis stratified patients into high- and low-expression groups based on the median expression level of each oncofetal piRNA to assess survival differences. Among the eight prognostic oncofetal piRNAs, three (piR-33687, piR-33686, and piR-61135) were significantly associated with OS when patients were categorized by high and low expression (Supplementary Figure S3). This discrepancy may arise because Cox regression considers continuous expression values and accounts for time-to-event data, increasing sensitivity to prognostic associations. In contrast, KM analysis relies on dichotomized expression groups, which can reduce statistical power. To evaluate the combined prognostic value of these eight oncofetal piRNAs identified in the univariate analysis,



a multivariate Cox regression analysis was conducted, and a risk score was established as described in the methodology (Supplementary Table S13). Patients were divided into high- and low-risk groups based on median risk score. KM analysis of the oncofetal-piRNA prognostic signature showed a significant difference in OS between high- and low-risk groups ($p < 0.0001$) (Figure 3A). The AUC of the ROC curve was 0.65, demonstrating moderate predictive performance (Figure 3A). The prognostic model was further evaluated using the concordance index (C-index), which measures the agreement between actual OS and model predictions (where a C-index of 0.5 indicates no prognostic value and values closer to 1 indicate better predictive accuracy). The C-index for TCGA-LUAD was 0.63, suggesting moderate predictive accuracy (Supplementary Figure S4A). To validate the model, we applied the oncofetal piRNA prognostic signature to the BCCA cohort. The signature successfully stratified patients into high- and low-risk groups ($p < 0.05$) (Figure 3B). In BCCA-LUAD, the AUC for survival prediction was 0.72, indicating better predictive performance compared to TCGA-LUAD (Figure 3B). The C-index in the validation cohort was 0.65 (Supplementary Figure S4B). When stratifying by tumour stage, the prognostic signature performed significantly in both early and advanced stage tumors with marked separation in early-stage cancers (Supplementary

Figure S5). Furthermore, the piRNA-based model performed comparably to the Tumour, Node, Metastasis (TNM) system, which is a gold standard clinical prognostic marker for NSCLC (Supplementary Figure S6) (Zens et al., 2021).

For TCGA-LUSC, univariate Cox regression analysis identified five prognostic oncofetal piRNAs ($p < 0.05$): piR-34789, piR-55152, piR-55151, piR-34342, and piR-33404 (Supplementary Table S12). KM survival analysis stratified patients into high- and low-expression groups based on the median expression level of each oncofetal piRNA, and found two of these piRNAs (piR-55151 and piR-55152) were significantly associated with OS ($p < 0.05$) (Supplementary Figure S7). However, when evaluating the combined prognostic signature in LUSC, the model could not stratify patients into high- and low-risk groups ($p > 0.05$), indicating that the signature did not have prognostic value in this cohort.

4 Discussion

Among sncRNAs, piRNAs have garnered interest due to their abundance, evolutionary conservation, and known roles in gene regulation and development (Mei et al., 2013). However, their

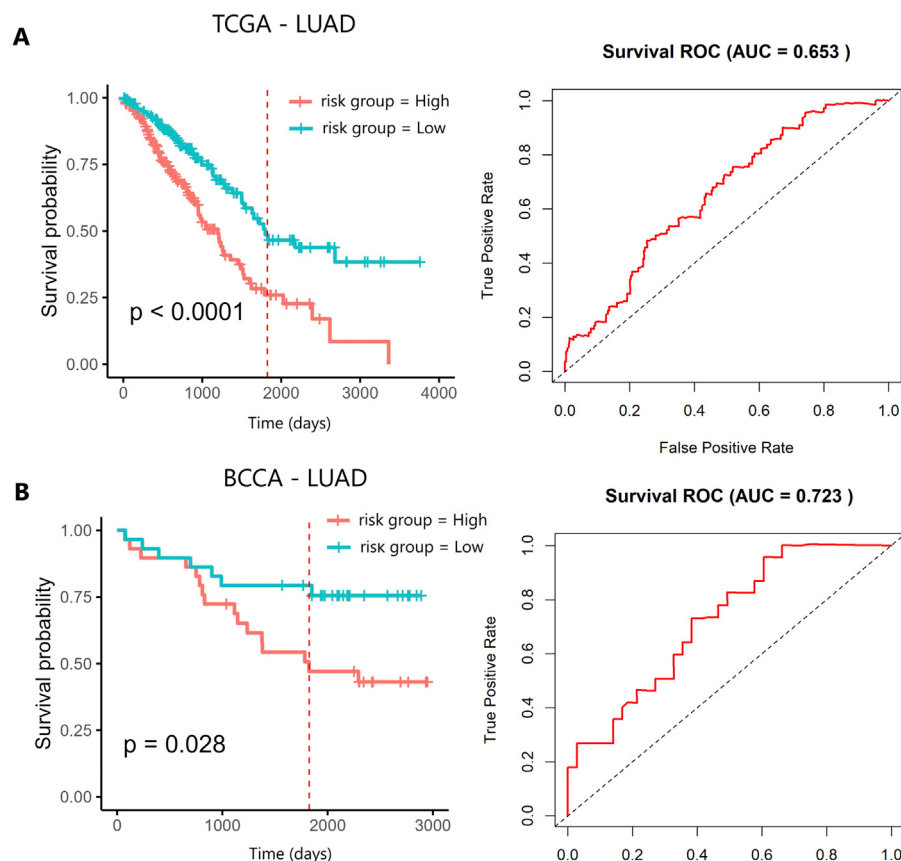


FIGURE 3

Kaplan–Meier survival analysis of the 8-oncofetal piRNA prognostic signature in (A) TCGA-LUAD and (B) BCCA-LUAD. The red vertical line marks the five-year survival threshold. AUC, Area Under the Curve; BCCA, British Columbia Cancer Agency; LUAD, Lung Adenocarcinoma; TCGA, The Cancer Genome Atlas.

functional significance in cancer biology, particularly in lung cancer, remains limited. There have been a handful of piRNAs studied for their association with lung cancers. For example, piR-651, piR-52200, and piR-34971 have been shown to be upregulated in NSCLC, with piR-651 linked to cancer progression (Li et al., 2016), potentially through regulation of cell cycle control proteins such as cyclin D1 and CDK4 (Li et al., 2016). Other piRNAs, including piR-55490, piR-46545, and piR-35127, have been downregulated in lung cancer and demonstrated tumor growth suppression by modulating the AKT/mTOR pathway (Peng et al., 2016; Reeves et al., 2017). Alterations in PIWI proteins, which are central to piRNA biogenesis through slicer activity, may contribute to the deregulated expression of piRNAs in cancer. Changes in expression, such as overexpression (e.g., PIWIL1) or downregulation (e.g., PIWIL2), has been identified in several cancers and are associated with poor prognosis (Navarro et al., 2015; Zhang et al., 2023). Possible mechanisms include disruption of piRNA amplification, aberrant transcriptional silencing of tumor suppressor genes (Zhang et al., 2023), as well as piRNA-independent oncogenic function (Shi et al., 2020; Zhang et al., 2023).

Although piRNAs are implicated in germline tissue maintenance and somatic tissues within various cancers, our knowledge on their presence in fetal tissues is limited and the overlap between these two domains has not been previously

explored until this study. piRNAs are expressed in embryos and fetal germline tissues (Roovers et al., 2015) and the placenta (Martinez et al., 2021a; 2021b), yet their role in somatic fetal tissues, particularly lung tissue, has been understudied. The expression of PIWI pathway proteins, including VASA and PIWIL2, is not prominent in first-trimester embryos but becomes evident in second-trimester fetal ovaries (Roovers et al., 2015). This observation aligns with our findings of a strong piRNA population in second-trimester fetal lung tissue, supporting the relevance of piRNAs in both developmental and cancer contexts.

In adult lung tissues, we identified 435 piRNAs in LUAD and 406 piRNAs in LUSC, with the piRNA populations in these tissues showing distinct expression patterns. Previous studies have reported piRNA populations in various cancers, including liver, colorectal, glioblastoma, and gastric cancers (Bartos et al., 2021; Riquelme et al., 2021). While several piRNAs were unique to each lung cancer subtype, 31 piRNAs were shared between them. These shared piRNAs may be involved in maintaining general characteristics in lung cancer such as genomic instability, epigenetic alterations, and immune evasion (Herbst et al., 2018; Hendriks et al., 2024).

A key finding of this study is the identification of 37 piRNAs in LUAD and 46 piRNAs in LUSC that exhibit oncofetal expression patterns. Oncofetal reprogramming can contribute to characteristics shared between tumorigenesis and fetal development including

rapid cell division, cellular invasion, migration, and angiogenesis (Ma et al., 2010). Pathways necessary in development such as Wnt, Notch, Hedgehog or TGF- β signaling, are commonly affected (Cao et al., 2023). The targets of the oncofetal miRNAs we previously discovered in the LUAD cohort were enriched for Wnt and Hedgehog pathways (Cohn et al., 2021). Similarities in their function to miRNAs and previous identification in regulation of pathways such as Wnt/ β -catenin may indicate the possibility that piRNAs are linked to development-associated pathways (Deng et al., 2024). Although we did not investigate target prediction due to lack of software availability, we found several oncofetal piRNAs correlated with OS, with a signature of eight piRNAs in the TCGA-LUAD cohort showing prognostic value. These piRNAs were used to generate a risk score that could stratify patients into high- and low-risk groups, with worse OS observed in the high-risk group. This signature was validated in the BCCA-LUAD cohort, showing significant stratification of patients into high- and low-risk groups. The validation results were statistically significant, despite the smaller sample size in the BCCA cohort ($n = 58$ vs. $n = 378$ in TCGA). The relative scarcity of deaths in the BCCA cohort, with most samples coming from early-stage lung cancer, did not prevent the signature from demonstrating prognostic potential. These findings suggest that the piRNA signature is a promising predictor of patient outcomes, and further validation in larger cohorts is warranted. In contrast, the LUSC prognostic model failed validation likely due to the cancer's high molecular and immune heterogeneity, distinct cell-of-origin, fewer well-defined driver mutations than LUAD, and limited availability of robust biomarkers. Together, these factors reduce model generalizability and reproducibility (Lu et al., 2020; Shen et al., 2025).

It is worth noting that piRNA IDs that are consecutive appear to be similar sequences. For example, piR-55152 (chr 8:56,073,866-56,073,905) and piR-55151 (chr 8:56,073,867-56,073,905) which were identified as significant in the TCGA LUSC univariate Cox analysis, and were both associated with OS individually, map to the same locations and have the same sequence other than piR-55152, which is one nucleotide longer than piR-55151. In this study we left them as separate IDs to match the deposition in the RNACentral database. However, functional validation via methods such as qRT-PCR, piRNA mimic or inhibitor studies, luciferase reporter assays, and localization staining, may provide more clarity on whether they are distinct entities or redundant naming. Such approaches offer orthogonal evidence supporting their biological relevance beyond potential technical or computational biases such as algorithmic assumptions, batch effects, reference database constraints, and alignment challenges (Fehlmann et al., 2021).

Although sparse in the literature, some oncofetal piRNAs we identified have been deposited in piRBase, the largest piRNA database, with evidence of their expression in lung cell lines such as H522 cells (a LUAD cell line), and in human bronchial epithelial (HBE) cells (Mei et al., 2015; Wang et al., 2022). Among the piRNAs identified in the TCGA-LUAD cohort, piR-34804 (piR-hsa-26940) and piR-61135 (piR-hsa-25313) were detected in the H522 cell line, with piR-34804 meeting the author's expression cut-off of 20 reads or more. In the LUSC cohort, piR-33404 (piR-hsa-23555) and piR-34342 (piR-hsa-26492) also met the expression cut-off defined by the authors in H522 cells. piR-34342 exhibited particularly high read counts in H522 cells (227 reads) and was expressed in several other

NSCLC and HBE cell lines. Although these piRNAs were not central to the authors findings, evidence of their presence in cell lines warrants further studies to determine their functional roles and significance in lung biology.

5 Conclusion

Our study provides novel insights into the role of piRNAs in lung cancer, particularly oncofetal piRNAs uniquely expressed in tumor tissues and associated with patient survival. These findings suggest that piRNAs may serve as potential biomarkers for lung cancer prognosis, with the possibility of further elucidating their functional roles in tumorigenesis. Given the similarities between fetal development and cancer, further investigation into piRNA biogenesis and their regulation in tumor tissues may uncover valuable therapeutic targets for lung cancer.

Data availability statement

The data presented in the study were retrieved from previously published sources. The fetal lung and BCCA LUAD data presented in the study were retrieved from the Gene Expression Omnibus (GEO) repository, accession number GSE175462. The TCGA LUAD and LUSC data were retrieved from the Genomic Data Commons (GDC) Portal. The original contributions presented in the study are included in the article and Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by BC Cancer and University of British Columbia Research Ethics Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

MP: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing. VS: Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review and editing. DC: Conceptualization, Data curation, Formal Analysis, Methodology, Writing – review and editing. NT: Writing – review and editing. GS: Conceptualization, Writing – review and editing. KB: Writing – review and editing. PR: Conceptualization, Funding acquisition, Writing – review and editing. VM: Conceptualization, Supervision, Writing – review and editing. WR: Conceptualization, Supervision, Writing – review and editing. WL: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by funds from the Canadian Institutes of Health Research (CIHR, FRN-143345 and 183775), the Terry Fox Foundation, the Lotte and John Hecht Memorial Foundation, and the BC Cancer Foundation. V.P.G.S. was supported by the BC Cancer Rising Stars Fellowship.

Acknowledgments

We acknowledge Brenda Minatel and Adam Sage for their contributions during the early stages of this project and their previous work on miRNAs, which laid the foundation for this study. Additionally, we recognize that part of the results presented here are based on data generated by the TCGA Research Network (<https://www.cancer.gov/tcga>). All figures were generated using BioRender (<https://www.biorender.com/>).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Bartos, M., Siegl, F., Kopkova, A., Radova, L., Oppelt, J., Vecera, M., et al. (2021). Small RNA sequencing identifies PIWI-interacting RNAs deregulated in Glioblastoma-piR-9491 and piR-12488 reduce tumor cell colonies *in vitro*. *Front. Oncol.* 11, 707017. doi:10.3389/fonc.2021.707017
- Becker-Santos, D. D., Thu, K. L., English, J. C., Pikor, L. A., Martinez, V. D., Zhang, M., et al. (2016). Developmental transcription factor NFIB is a putative target of oncofetal miRNAs and is associated with tumour aggressiveness in lung adenocarcinoma. *J. Pathol.* 240, 161–172. doi:10.1002/path.4765
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170
- Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511, 543–550. doi:10.1038/nature13385
- Cao, J., Zhang, Z., Zhou, L., Luo, M., Li, L., Li, B., et al. (2023). Oncofetal reprogramming in tumor development and progression: novel insights into cancer therapy. *MedComm* (2020), 4–e427. doi:10.1002/mco.2427
- Cassim, S., Chepul, L., Keenan, R., Kidd, J., Firth, M., and Lawrenson, R. (2019). Patient and carer perceived barriers to early presentation and diagnosis of lung cancer: a systematic review. *BMC Cancer* 19, 25. doi:10.1186/s12885-018-5169-9
- Chen, D., Zhang, L.-Q., Huang, J.-F., Liu, K., Chuai, Z.-R., Yang, Z., et al. (2014). BRAF mutations in patients with non-small cell lung cancer: a systematic review and meta-analysis. *PLoS One* 9, e101354. doi:10.1371/journal.pone.0101354
- Chen, T., Dai, X., Dai, J., Ding, C., Zhang, Z., Lin, Z., et al. (2020). AFP promotes HCC progression by suppressing the HuR-mediated Fas/FADD apoptotic pathway. *Cell Death Dis.* 11, 822. doi:10.1038/s41419-020-03030-7
- Cheung, W. K. C., and Nguyen, D. X. (2015). Lineage factors and differentiation states in lung cancer progression. *Oncogene* 34, 5771–5780. doi:10.1038/ncr.2015.85
- Cohn, D. E., Barros-Filho, M. C., Minatel, B. C., Pewarchuk, M. E., Marshall, E. A., Vucic, E. A., et al. (2021). Reactivation of multiple fetal miRNAs in lung adenocarcinoma. *Cancers (Basel)* 13, 2686. doi:10.3390/cancers13112686
- Conway, J. R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33, 2938–2940. doi:10.1093/bioinformatics/btx364
- Coulouarn, C., Derambure, C., Lefebvre, G., Daveau, R., Hiron, M., Scotte, M., et al. (2005). Global gene repression in hepatocellular carcinoma and fetal liver, and suppression of dudulin-2 mRNA as a possible marker for the cirrhosis-to-tumor transition. *J. Hepatol.* 42, 860–869. doi:10.1016/j.jhep.2005.01.027
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. doi:10.1093/gigascience/giab008
- Das, V., Kalita, J., and Pal, M. (2017). Predictive and prognostic biomarkers in colorectal cancer: a systematic review of recent advances and challenges. *Biomed. Pharmacother.* 87, 8–19. doi:10.1016/j.biopha.2016.12.064
- Deng, X., Liao, T., Xie, J., Kang, D., He, Y., Sun, Y., et al. (2024). The burgeoning importance of PIWI-Interacting RNAs in cancer progression. *Sci. China Life Sci.* 67, 653–662. doi:10.1007/s11427-023-2491-7
- Du, X., Shao, Y., Qin, H.-F., Tai, Y.-H., and Gao, H.-J. (2018). ALK-rearrangement in non-small-cell lung cancer (NSCLC). *Thorac. Cancer* 9, 423–430. doi:10.1111/1759-7714.12613
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. doi:10.1093/nar/30.1.207
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Waite Hill, OH: Society for Industrial and Applied Mathematics. doi:10.1137/1.9781611970319
- Fehlmann, T., Backes, C., Kahraman, M., Haas, J., Ludwig, N., Posch, A. E., et al. (2017). Web-based NGS data analysis using miRMaster: a large-scale meta-analysis of human miRNAs. *Nucleic Acids Res.* 45, 8731–8744. doi:10.1093/nar/gkx595
- Fehlmann, T., Kern, F., Laham, O., Backes, C., Solomon, J., Hirsch, P., et al. (2021). miRMaster 2.0: multi-species non-coding RNA sequencing analyses at scale. *Nucleic Acids Res.* 49, W397–W408. doi:10.1093/nar/gkab268
- Feng, L., Wang, J., Cao, B., Zhang, Y., Wu, B., Di, X., et al. (2014). Gene expression profiling in human lung development: an abundant resource for lung adenocarcinoma prognosis. *PLoS One* 9, e105639. doi:10.1371/journal.pone.0105639
- Feng, Y., Li, Y., Dai, W., Mo, S., Li, Q., and Cai, S. (2018). Clinicopathologic features and prognostic factors in alpha-fetoprotein-producing colorectal cancer: analysis of 78 cases. *Cell Physiol. Biochem.* 51, 2052–2064. doi:10.1159/000495824
- Fu, X.-D. (2014). Non-coding RNA: a new frontier in regulatory biology. *Natl. Sci. Rev.* 1, 190–204. doi:10.1093/nsr/nwu008

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1611805/full#supplementary-material>

- Goldman, M. J., Craft, B., Hastie, M., Repčeka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and interpreting cancer genomics data via the xena platform. *Nat. Biotechnol.* 38, 675–678. doi:10.1038/s41587-020-0546-8
- Haibe-Kains, B., Desmedt, C., Sotiriou, C., and Bontempi, G. (2008). A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatics* 24, 2200–2208. doi:10.1093/bioinformatics/btn374
- Halic, M., and Moazed, D. (2009). Transposon silencing by piRNAs. *Cell* 138, 1058–1060. doi:10.1016/j.cell.2009.08.030
- Hassan, K. A., Chen, G., Kalemkerian, G. P., Wicha, M. S., and Beer, D. G. (2009). An embryonic stem cell-like signature identifies poorly differentiated lung adenocarcinoma but not squamous cell carcinoma. *Clin. Cancer Res.* 15, 6386–6390. doi:10.1158/1078-0432.CCR-09-1105
- Heagerty, P. J., and Chaudhuri, S. (2022). survivalROC: time-dependent ROC curve estimation from censored survival data. Available online at: <https://CRAN.R-project.org/package=survivalROC>.
- Hendriks, L. E. L., Remon, J., Faivre-Finn, C., Garassino, M. C., Heymach, J. V., Kerr, K. M., et al. (2024). Non-small-cell lung cancer. *Nat. Rev. Dis. Prim.* 10, 71. doi:10.1038/s41572-024-00551-9
- Herbst, R. S., Morgensztern, D., and Boshoff, C. (2018). The biology and management of non-small cell lung cancer. *Nature* 553, 446–454. doi:10.1038/nature25183
- Hu, M., and Shivdasani, R. A. (2005). Overlapping gene expression in fetal mouse intestine development and human colorectal cancer. *Cancer Res.* 65, 8715–8722. doi:10.1158/0008-5472.CAN-05-0700
- Huang, J.-F., Jiang, H.-Y., Cai, H., Liu, Y., Zhu, Y.-Q., Lin, S.-S., et al. (2019). Genome-wide screening identifies oncofetal lncRNA Ptn-dt promoting the proliferation of hepatocellular carcinoma cells by regulating the ptn receptor. *Oncogene* 38, 3428–3445. doi:10.1038/s41388-018-0643-z
- Iwasaki, Y. W., Siomi, M. C., and Siomi, H. (2015). PIWI-interacting RNA: its biogenesis and functions. *Annu. Rev. Biochem.* 84, 405–433. doi:10.1146/annurev-biochem-060614-034258
- Kolde, R. (2019). *Pheatmap: pretty heatmaps*. 1.0.12. doi:10.32614/CRAN.package.pheatmap
- Li, D., Luo, Y., Gao, Y., Yang, Y., Wang, Y., Xu, Y., et al. (2016). piR-651 promotes tumor formation in non-small cell lung carcinoma through the upregulation of cyclin D1 and CDK4. *Int. J. Mol. Med.* 38, 927–936. doi:10.3892/ijmm.2016.2671
- Lu, C., Bera, K., Wang, X., Prasanna, P., Xu, J., Janowczyk, A., et al. (2020). A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicenter, retrospective study. *Lancet Digit. Health* 2, e594–e606. doi:10.1016/s2589-7500(20)30225-9
- Ma, Y., Zhang, P., Wang, F., Yang, J., Yang, Z., and Qin, H. (2010). The relationship between early embryo development and tumorigenesis. *J. Cell Mol. Med.* 14, 2697–2701. doi:10.1111/j.1582-4934.2010.01191.x
- Martinez, V. D., Vucic, E. A., Thu, K. L., Hubaux, R., Enfield, K. S. S., Pikor, L. A., et al. (2015). Unique somatic and malignant expression patterns implicate PIWI-interacting RNAs in cancer-type specific biology. *Sci. Rep.* 5, 10423. doi:10.1038/srep10423
- Martinez, V. D., Cohn, D. E., Telkar, N., Minatel, B. C., Pewarchuk, M. E., Marshall, E. A., et al. (2021a). Profiling the small non-coding RNA transcriptome of the human placenta. *Sci. Data* 8, 166. doi:10.1038/s41597-021-00948-1
- Martinez, V. D., Sage, A. P., Minatel, B. C., Marshall, E. A., Price, E. M., Becker-Santos, D. D., et al. (2021b). Human placental piwi-interacting RNA transcriptome is characterized by expression from the DLK1-DIO3 imprinted region. *Sci. Rep.* 11, 14981. doi:10.1038/s41598-021-93885-3
- Mei, Y., Clark, D., and Mao, L. (2013). Novel dimensions of piRNAs in cancer. *Cancer Lett.* 336, 46–52. doi:10.1016/j.canlet.2013.04.008
- Mei, Y., Wang, Y., Kumari, P., Shetty, A. C., Clark, D., Gable, T., et al. (2015). A piRNA-like small RNA interacts with and modulates p-ERM proteins in human somatic cells. *Nat. Commun.* 6, 7316. doi:10.1038/ncomms8316
- Mizejewski, G. J. (2002). Biological role of alpha-fetoprotein in cancer: prospects for anticancer therapy. *Expert Rev. Anticancer Ther.* 2, 709–735. doi:10.1586/14737140.2.6.709
- National Cancer institute (2019). The cancer genome atlas program (TCGA) - NCI. Available online at: <https://www.cancer.gov/ccg/research/genome-sequencing/tcga> (Accessed July 11, 2019).
- Navarro, A., Tejero, R., Viñolas, N., Cordeiro, A., Marrades, R. M., Fuster, D., et al. (2015). The significance of PIWI family expression in human lung embryogenesis and non-small cell lung cancer. *Oncotarget* 6, 31544–31556. doi:10.18632/oncotarget.3003
- Peng, L., Song, L., Liu, C., Lv, X., Li, X., Jie, J., et al. (2016). piR-55490 inhibits the growth of lung carcinoma by suppressing mTOR signaling. *Tumour Biol.* 37, 2749–2756. doi:10.1007/s13277-015-4056-0
- R Core Team (2013). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>.
- Reeves, M. E., Firek, M., Jliedi, A., and Amaar, Y. G. (2017). Identification and characterization of RASSF1C piRNA target genes in lung cancer cells. *Oncotarget* 8, 34268–34282. doi:10.18632/oncotarget.15965
- Riquelme, I., Pérez-Moreno, P., Letelier, P., Brebi, P., and Roa, J. C. (2021). The emerging role of PIWI-interacting RNAs (piRNAs) in gastrointestinal cancers: an updated perspective. *Cancers (Basel)* 14, 202. doi:10.3390/cancers14010202
- Roovers, E. F., Rosenkranz, D., Mahdipour, M., Han, C.-T., He, N., Chuva de Sousa Lopes, S. M., et al. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Rep.* 10, 2069–2082. doi:10.1016/j.celrep.2015.02.062
- RStudio Team (2020). *RStudio: integrated development environment for R*. Boston, MA: RStudio, PBC. Available online at: <http://www.rstudio.com/>.
- Schröder, M. S., Culhane, A. C., Quackenbush, J., and Haibe-Kains, B. (2011). Survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 27, 3206–3208. doi:10.1093/bioinformatics/btr511
- Sharma, A., Blériot, C., Currenti, J., and Ginhoux, F. (2022). Oncofetal reprogramming in tumour development and progression. *Nat. Rev. Cancer* 22, 593–602. doi:10.1038/s41568-022-00497-8
- Shen, Y., Chen, J.-Q., and Li, X.-P. (2025). Differences between lung adenocarcinoma and lung squamous cell carcinoma: driver genes, therapeutic targets, and clinical efficacy. *Genes Dis.* 12, 101374. doi:10.1016/j.gendis.2024.101374
- Shi, S., Yang, Z.-Z., Liu, S., Yang, F., and Lin, H. (2020). PIWIL1 promotes gastric cancer via a piRNA-independent mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 117, 22390–22401. doi:10.1073/pnas.2008724117
- Siegel, R. L., Giaquinto, A. N., and Jemal, A. (2024). Cancer statistics, 2024. *CA Cancer J. Clin.* 74, 12–49. doi:10.3322/caac.21820
- Tan, A. C., and Tan, D. S. W. (2022). Targeted therapies for lung cancer patients with oncogenic driver molecular alterations. *J. Clin. Oncol.* 40, 611–625. doi:10.1200/JCO.21.01626
- Teixeira, F. K., Okuniewska, M., Malone, C. D., Coux, R.-X., Rio, D. C., and Lehmann, R. (2017). piRNA-mediated regulation of transposon alternative splicing in the soma and germ line. *Nature* 552, 268–272. doi:10.1038/nature25018
- The Cancer Genome Atlas Research Network, Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., et al. (2013). The cancer Genome Atlas pan-cancer analysis project. *Nat. Genet.* 45, 1113–1120. doi:10.1038/ng.2764
- Therneau, T. M. (2024). A package for survival analysis in R. Available online at: <https://CRAN.R-project.org/package=survival>.
- Therneau, T. M., and Grambsch, P. M. (2000). *Modeling survival data: extending the cox model*. New York: Springer.
- Wang, J., Shi, Y., Zhou, H., Zhang, P., Song, T., Ying, Z., et al. (2022). piRBase: integrating piRNA annotation in all aspects. *Nucleic Acids Res.* 50, D265–D272. doi:10.1093/nar/gkab1012
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag. Available online at: <https://ggplot2.tidyverse.org>.
- Zens, P., Bello, C., Scherz, A., Koenigsdorf, J., Pöllinger, A., Schmid, R. A., et al. (2021). A prognostic score for non-small cell lung cancer resected after neoadjuvant therapy in comparison with the tumor-node-metastases classification and major pathological response. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc.* 34, 1333–1344. doi:10.1038/s41379-021-00777-y
- Zhang, Q., Zhu, Y., Cao, X., Tan, W., Yu, J., Lu, Y., et al. (2023). The epigenetic regulatory mechanism of PIWI/piRNAs in human cancers. *Mol. Cancer* 22, 45. doi:10.1186/s12943-023-01749-3