life

MDPI

Article

Machine Learning-Based Identification of Candidate Serum miRNA Features for Pan-Cancer and Cancer Type Classification

Kaiyan Feng [1,†], Yusheng Bao [2,†], Jingxin Ren [2,†], Wei Guo [3], Deling Wang [4], Tao Huang [5,6,*] and Yu-Dong Cai [2,*]

[1] Department of Computer Science, Guangdong AIB Polytechnic College, Guangzhou 510507, China; kyfeng@gdaib.edu.cn
[2] School of Life Sciences, Shanghai University, Shanghai 200444, China; bys@shu.edu.cn (Y.B.); ssdrg@shu.edu.cn (J.R.)
[3] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; gw_1992@sjtu.edu.cn
[4] Department of Radiology, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China; wangdl@sysucc.org.cn
[5] Bio-Med Big Data Center, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[6] CAS Engineering Laboratory for Nutrition, Department of Artificial Intelligence and Digital Health, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China
* Correspondence: huangtao@sibs.ac.cn (T.H.); caiyudong@staff.shu.edu.cn (Y.-D.C.); Tel.: +86-21-54923269 (T.H.); +86-21-66136132 (Y.-D.C.)
† These authors contributed equally to this work.

Abstract

MicroRNA (miRNA) regulation plays a pivotal role in intracellular gene expression. Analysis of miRNA profiles can provide critical insights into disease states. As cancer-associated molecules reported in previous studies, miRNAs may serve as candidate classificatory features for exploratory cancer classification. This research analyzed serum miRNA data from patients with 13 solid cancer types and individuals without cancer. The study comprised two distinct analyses: first, stratifying the dataset into cancer and non-cancer groups to identify miRNAs differentially represented in cancer patients; and second, subdividing the cancer patient data into 13 predefined solid-cancer types to identify candidate miRNA features that discriminate among these cancer types. We employed seven feature-ranking algorithms to evaluate miRNA contributions in both analyses and generate feature lists. Each list was examined using an incremental feature selection method to extract essential miRNAs and build good-performing classification models. Several candidate miRNAs were identified for distinguishing pan-cancer samples from non-cancer ones: miR-4783-3p has been linked to associated with the regulation of endocrine cell differentiation, and miR-663a has been reported in hepatocellular carcinoma and thyroid carcinoma. The analysis also highlighted miRNAs that differentiate solid cancer types, including miR-629-3p, reported to be upregulated in lung and breast cancer, and miR-6087, reported to be downregulated in osteosarcoma and bladder cancer.

Keywords: miRNA; pan-cancer; machine learning; candidate classificatory feature

Check for updates

Academic Editor: Paola Nieri

Received: 16 March 2026
Revised: 14 May 2026
Accepted: 15 May 2026
Published: 20 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Cancer is one of the major diseases worldwide; it is a complex disease that develops largely through genetic alterations in normal cells. These alterations drive uncontrolled cell

proliferation and invasion into other tissues and organs. Early-stage cancers are relatively curable, but when cancer metastasizes [1], mortality increases substantially. Tumor metastasis accounts for most cancer-related deaths. Early detection of cancer, therefore, is of utmost importance. Genomic changes vary significantly among various cancers. Similarly, the expression profiles between primary and metastatic tumors are quite dissimilar with considerable variability [2]. Current medical practice still cannot consistently provide early detection or accurate assessment across cancer types. This presents a strong rationale for developing effective predictive methods. Extensive studies have identified microRNAs (miRNAs) as important cancer-associated markers [3–5]; miRNAs have been reported to show altered expression in tumor cells and have been investigated as candidate therapeutic targets in prior studies [6].

miRNAs, a family of small non-coding RNAs, typically 20–25 nucleotides in length, are important regulators of gene expression in cells. They function by binding to the mRNA of target genes, inhibiting translation or promoting mRNA degradation, thereby regulating gene expression levels [7,8]. Therefore, miRNAs are involved in cell proliferation, apoptosis, and metabolic processes [9]. They may participate in the pathogenesis of many cancer types, partly because miRNAs regulate metastasis-related factors in several cancer types [10,11]. In particular, Hussen et al. have provided a comprehensive framework illustrating how individual miRNAs can serve simultaneously as signatures of cancer progression, tumor staging, and therapeutic response—functioning as oncogenes in some tissue contexts while acting as tumor suppressors in others [11]. This review further emphasized that the diagnostic and prognostic value of circulating miRNAs depends critically on cancer type, disease stage, and the specific biological pathway involved, cautioning against treating any single miRNA as a universally decisive marker. This framework is directly relevant to the present study because the candidate serum miRNAs identified by our machine learning-based pipeline should be interpreted within the broader, context-dependent landscape of miRNA–cancer interactions documented across multiple independent studies. Depending on cancer type, miRNAs can function as either oncogenes or tumor suppressors. For instance, the miR-24 family facilitates tumor growth in oral cancer but serves a tumor-suppressive role in colorectal cancer [12,13]. Our prior studies also suggested that miRNAs participate in mRNA co-expression networks through gene silencing, a mechanism associated with the onset and progression of testicular germ cell tumors [14]. Moreover, we have identified mRNA-miRNA-lncRNA regulatory networks in ovarian cancer and uterine corpus endometrial carcinoma. In these networks, miRNAs interact with lncRNAs through sponge-like mechanisms, attenuating gene-silencing capacity and potentially contributing to cancer initiation and metastasis [15,16]. These findings indicate the context-dependent roles of miRNAs in tumor regulation through diverse pathways. As candidate molecular features, miRNAs may support disease classification, diagnosis and prognostic assessment [10]. The detection of miRNA levels in bodily fluids offers opportunities for early disease diagnosis and prognostic evaluation.

Compared with traditional tissue biopsies, which remain important for tumor characterization, liquid biopsies are notable for their minimal invasiveness and ease of sampling. The repeatability of liquid biopsies enables more effective monitoring of tumor progression, facilitating early cancer diagnosis and treatment response assessment [17,18]. Current technological advances enable precise detection of circulating nucleic acids in serum [19]. Because of the complexity of serum and transcriptomic data, most modern research relies on machine learning analysis methods. Feature selection can partition large datasets into informative subsets; this greatly simplifies the process of biomarker identification [20]. Moreover, a variety of computational and statistical methods have been utilized to identify cancer biomarkers [21], including the analysis of miRNA based features for assessing breast

cancer severity and comprehensive pan-cancer evaluations of genomic profiles [22,23]. Such identification is complicated by the complex biological mechanisms of cancer; studies often identify candidate pan-cancer classificatory features associated with multiple cancer types. Thus, another focus of current research is scoring and ranking candidate features using multi-omics data to further improve the predictive power of the pan-cancer biomarkers [24].

In this study, serum miRNA profiles for various cancer types retrieved from the GEO database were analyzed. The profiles were divided into two datasets. The first dataset contained non-cancer and cancer patients, whereas the second contained only cancer patients, who were further classified into thirteen cancer types. Using seven feature-ranking algorithms, including Least absolute shrinkage and selection operator (LASSO) [25], Light gradient boosting machine (LightGBM) [26], Monte Carlo feature selection (MCFS) [27], Minimum redundancy maximum relevance (mRMR) [28], Random forest (RF_ZL) [29], Categorical boosting (CATboost) [30], and Extreme gradient boosting (XGBoost) [31], we obtained seven feature lists for each dataset. Each feature list was fed into an incremental feature selection (IFS) method [32], using four classification algorithms (Decision tree (DT) [33], K-nearest neighbors (KNN) [34], Random forest (RF) [29], and Support vector machine (SVM) [35]) and Synthetic minority oversampling technique (SMOTE) [36], to build optimal classification models and extract essential features and classification rules. These selected features were considered candidate pan-cancer or cancer type classificatory features. Analysis of these features may clarify serum-based classification patterns and generate hypotheses for subsequent biological validation.

2. Materials and Methods

This study analyzed serum miRNA profiles from cancer patients and non-cancer individuals. Figure 1 outlines the overall workflow of this study. Initially, these data were converted into two patient miRNA expression matrices, corresponding to cancer versus non-cancer classification and cancer type classification. Each matrix was investigated by seven feature-ranking algorithms, resulting in seven feature lists. Subsequently, the IFS method was used to identify the essential miRNA features from each list and build high-performing classification models.

2.1. Dataset

This study used data retrieved from the GEO database with accession ID GSE211692 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE211692, accessed on 20 March 2023) [37], which included serum miRNA profiles from pan-cancer patients with thirteen solid cancer types and non-cancer populations. Two independent analyses were conducted, corresponding to two datasets. First, the patients were divided into one pan-cancer group (9921 patients) and one non-cancer group (6245 patients). Second, the cancer patients were further divided into thirteen groups based on thirteen cancer types, including biliary tract cancer, bladder cancer, bone and soft tissue sarcomas, breast cancer, colorectal cancer, esophageal squamous cell cancer, gastric cancer, hepatocellular cancer, intraparenchymal brain tumors, lung cancer, ovarian cancer, pancreatic cancer, and prostate cancer. Table 1 shows the number of patients in each group. Each patient in the two datasets was represented by the expression levels of 2565 miRNAs. Each dataset was transformed into a two-dimensional matrix (patient-miRNA expression) and fed into a machine-learning-based workflow.

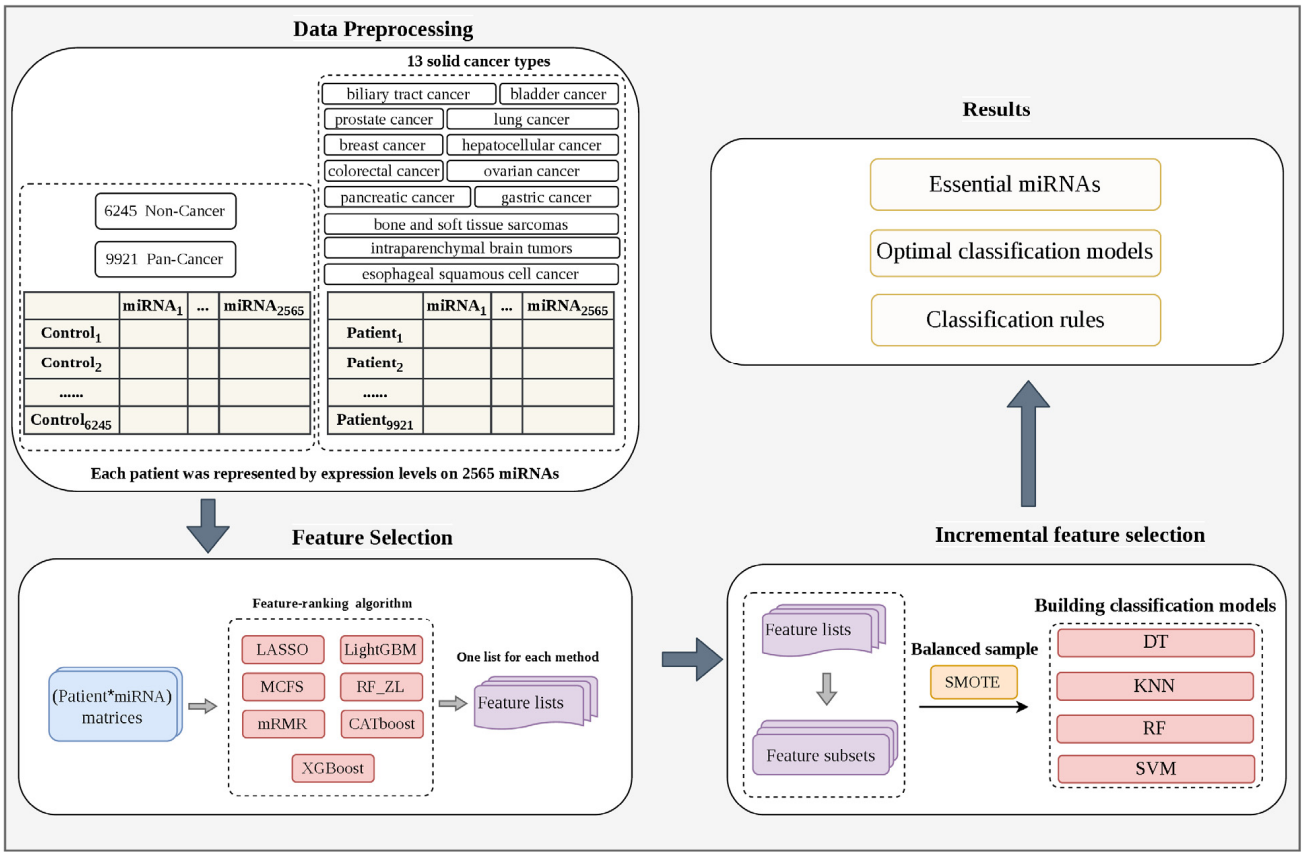


Figure 1. Flow chart of the entire analysis process. Serum miRNA profiles from pan-cancer patients (thirteen solid cancer types) and non-cancer populations were analyzed using multiple machine learning algorithms. Seven feature-ranking algorithms, four classification algorithms, and Synthetic minority oversampling technique (SMOTE) were included. Feature screening was performed using incremental feature selection method, yielding key miRNAs and classification rules, which indicated quantitative miRNA expression patterns for specific populations. Patient*miRNA means the patient-miRNA expression matrices using patients as rows and miRNAs as columns.

Table 1. Number of patients in the two datasets.

Group		Number of Patients
Non-cancer		6245
Pan-cancer	biliary tract cancer	402
	bladder cancer	399
	bone and soft tissue sarcomas	299
	breast cancer	675
	colorectal cancer	1596
	esophageal squamous cell cancer	566
	gastric cancer	1418
	hepatocellular cancer	348
	intraparenchymal brain tumors	241
	lung cancer	1699
	ovarian cancer	400
	pancreatic cancer	851
	prostate cancer	1027
	Total	9921

2.2. Feature-Ranking Algorithms

To analyze the importance of miRNAs in identifying cancer patients from non-cancer populations or classifying solid cancer types, we employed seven feature-ranking algorithms, including LASSO [25], LightGBM [26], MCFS [27], mRMR [28], RF_ZL [29], CATboost [30], and XGBoost [31]. These algorithms have been widely used in marker-screening studies [38–42]. In this study, they were applied to the two complete datasets for ranking miRNAs based on how well they are associated with the target variable. Brief descriptions of these algorithms are provided in File S1. We used the public packages for these seven algorithms and executed them with their default hyperparameters. The sources of the packages are provided in Table S1.

2.3. Incremental Feature Selection

IFS [32] is an approach for selecting important features from a given feature list. First, several feature subsets are constructed such that the top feature in the list constitutes the first feature subset, then the next feature in the list is added to constitute the second feature subset, and so on. Second, one classification model based on a given classification algorithm is constructed on each feature subset, which is evaluated using a cross-validation method [43] and assessed using metrics such as accuracy, F1 measure, or AUC. Finally, the best-performing model is identified as the optimal model, and the features used in this model are extracted as the optimal features. These features generally make essential contributions to classification and should be further analyzed.

2.4. Synthetic Minority Oversampling Technique

SMOTE [36] is an algorithm widely used to address class imbalance in machine learning. It is mainly used when the number of majority-class samples in real-world classification tasks far exceeds the number of minority-class samples. The algorithm operates as follows: first, a sample is randomly selected from the minority class, and its nearest k neighbors in the same class are identified according to the parameter 'k' specified by the user; a new sample is then generated by selecting a new point on the line between the selected sample and its randomly selected neighbor. This new sample is assigned to the minority class to increase its size. The process of generating new samples is repeated until the desired balance ratio is reached between the minority classes and the majority classes.

In this study, SMOTE was used to balance the datasets in two analyses. In the first analysis, cancer patients outnumbered non-cancer patients. SMOTE generated new representations of non-cancer patients until the number of non-cancer patients matched the number of cancer patients. In the second analysis, lung cancer patients constituted the largest class. SMOTE produced new representations of patients in other classes until all classes had equal sizes. Importantly, SMOTE was only applied to the training set in each round of cross-validation. This strategy isolated test sample information from the training procedure, making the results more reliable.

2.5. Classification Algorithms

A total of four classification algorithms were used in this study to construct classification models in the IFS procedure, including DT [33], KNN [44], RF [29], and SVM [45]. Their brief descriptions are available in File S1. In this study, we used the packages of these four classification algorithms implemented in Scikit-learn 1.2.1 [46]. Each package was run using default hyperparameters.

2.6. Cross-Validation Strategy

To ensure the robustness and generalization capability of the classification models in the IFS procedure, we employed a ten-fold cross-validation approach [43,47–51]. The dataset was randomly partitioned into ten equal-sized subsets (folds), maintaining the proportion of samples in each class across all folds. In each iteration, nine folds were used for model training while the remaining fold served as the validation set. This process was repeated ten times, with each fold serving as the validation set exactly once. The final performance metrics were calculated as the average across all ten iterations, providing a more reliable estimate of model performance.

2.7. Performance Evaluation

The purpose of the first analysis was to identify candidate miRNA classificatory features that could distinguish pan-cancer patients from the non-cancer population. This is a binary classification problem. F1 measure [52–60] was used as the key performance metric. Other metrics, such as sensitivity (SN), specificity (SP), accuracy (ACC), precision, and Matthews correlation coefficient (MCC) [61], were also provided for reference. In the second analysis, the samples (patients) were divided into thirteen cancer types. Because the dataset was imbalanced, weights were introduced when evaluating models built on this dataset. The weighted F1 was selected as the key metric and other metrics, including ACC, MCC, and macro F1, were provided for reference.

3. Results

3.1. Results of Feature-Ranking Algorithms

After the serum miRNA profiles were converted into two two-dimensional matrices, each matrix was analyzed by seven feature-ranking algorithms independently, yielding seven feature lists, which are provided in Table S2. Typically, miRNAs (referred to as features in this study) with the highest relevance to cancer were generally positioned at the top of these lists. In the following analysis, we focused on the top features in each list.

3.2. Results of IFS

Each feature list mentioned in Section 3.1 was fed into IFS for subsequent analysis. Generally, essential miRNAs for distinguishing cancer patients from non-cancer individuals and classifying cancer types are usually limited in number. Thus, we only considered the top 1000 features in each list. The inclusion of more features may introduce noise, influencing the reliability of the results. Furthermore, we used a step size of five to construct feature subsets to accelerate the IFS procedure due to our limited computing resources, resulting in 200 subsets from each list. On each subset, four classification models using DT, KNN, RF, and SVM as classification algorithms were constructed and evaluated by ten-fold cross-validation. The cross-validation results were summarized as the metrics mentioned in Section 2.7. In detail, SN, SP, ACC, MCC, precision, and F1 measure were calculated in the first analysis, which are summarized in Table S3, whereas the ACC, MCC, macro F1, and weighted F1 were computed in the second analysis, which are presented in Table S4. The F1 measure was selected as the key metric in the first analysis. Accordingly, IFS curves were plotted to show the dynamics of the model performance as the number of features changed, where F1 measure and number of features were set as the y-axis and x-axis, respectively. The IFS curves for different classification algorithms on seven feature lists are provided in Figures S1 and S2. Likewise, the IFS curves for the second analysis are provided in Figures S3 and S4, where weighted F1 and number of features were set as the y-axis and x-axis, respectively. The results show that the number of features used and predictive ability were not linearly related. There was a significant performance gap

between the classification models constructed using different classification algorithms on different feature lists.

For the first analysis on pan-cancer patients and non-cancer populations, the IFS procedure identified the optimal models on seven feature lists. Specifically, on the feature list yielded by mRMR, the DT model achieved the peak performance with an F1 measure of 0.956, using the top 175 features. The KNN model performed the best on three lists yielded by CATboost, LASSO, and LightGBM. The F1 measures of these models were 0.977 on the feature list yielded by CATboost with the top 50 features, and 0.962 and 0.979 on feature lists yielded by LASSO and LightGBM, respectively, with the top 25 miRNAs each. As for the feature lists yielded by MCFS, RF, and XGBoost, the RF models performed the best. They yielded the F1 measures of 0.962, 0.966, and 0.970 with the top 645 features in the list yielded by MCFS, the top 995 features in the list yielded by RF_ZL, and the top 35 features in the list yielded by XGBoost, respectively. The detailed performance of the above optimal models is listed in Table 2. Specifically, the KNN model using the top 25 features in the list yielded by LightGBM provided the highest F1 measure. This model produced balanced performance, with SN of 0.985, SP of 0.983, and precision of 0.973. It may serve as a useful tool for distinguishing cancer patients from non-cancer individuals in this dataset.

Table 2. Performance of the classification models using the optimal feature subset and the inflection point subset in the pan-cancer and non-cancer comparative analysis.

Feature Ranking Algorithms	Classification Algorithm	Number of Features	SN	SP	Precision	F1 Measure	MCC	ACC
LASSO	KNN	25	0.979	0.965	0.946	0.962	0.938	0.970
LightGBM	KNN	25	0.985	0.983	0.973	0.979	0.965	0.983
MCFS	RF	15 *	0.919	0.992	0.986	0.952	0.924	0.964
	RF	645 **	0.940	0.991	0.985	0.962	0.940	0.971
mRMR	DT	50 *	0.961	0.962	0.941	0.951	0.919	0.962
	DT	175 **	0.966	0.966	0.946	0.956	0.928	0.966
RF_ZL	RF	25 *	0.928	0.985	0.976	0.951	0.923	0.963
	RF	995 **	0.945	0.993	0.989	0.966	0.947	0.975
CATboost	KNN	50	0.984	0.982	0.971	0.977	0.963	0.982
XGBoost	RF	35	0.962	0.987	0.979	0.970	0.952	0.977

Those with * markers are the number of features in the inflection point subset, and those with ** markers are the number of features in the optimal feature subset.

For the second analysis on a cancer patient dataset, the aim was to identify miRNAs that could distinguish thirteen different cancer types. The optimal models on different feature lists were also identified through the IFS procedure and evaluated using weighted F1. This time, the optimal models all used SVM as the classification algorithm across seven feature lists. In particular, they used the top 155, 195, 245, 480, 615, 295, and 240 features in the lists yielded by CATboost, LASSO, LightGBM, MCFS, mRMR, RF_ZL, and XGBoost, respectively, and yielded weighted F1 scores of 0.878, 0.795, 0.884, 0.869, 0.850, 0.876, and 0.832, respectively. The detailed performance of the above optimal models is provided in Table 3. Among these, the SVM model using the top 245 features in the list yielded by LightGBM provided the best performance. The performance of this model on thirteen cancer types, measured by F1 measure, is listed in Table 4. Eight cancer types had F1 measures higher than 0.9, and one cancer type had an F1 measure lower than 0.8, indicating high and balanced performance of the model. These results suggest that this model may be useful for identifying different cancer types.

Table 3. Performance of the classification models using the optimal feature subset and the inflection point subset in thirteen cancer types comparative analysis.

Feature Ranking Algorithms	Classification Algorithm	Number of Features	Weighted F1	MCC	ACC
LASSO	SVM	60 *	0.783	0.766	0.789
	SVM	195 **	0.795	0.778	0.800
LightGBM	SVM	95 *	0.865	0.852	0.867
	SVM	245 **	0.884	0.873	0.886
MCFS	SVM	95 *	0.821	0.805	0.825
	SVM	480 **	0.869	0.857	0.871
mRMR	SVM	115 *	0.821	0.806	0.825
	SVM	615 **	0.850	0.837	0.853
RF_ZL	SVM	115 *	0.859	0.846	0.862
	SVM	295 **	0.876	0.865	0.878
CATboost	SVM	65 *	0.850	0.836	0.853
	SVM	155 **	0.878	0.866	0.880
XGBoost	SVM	140 *	0.821	0.805	0.825
	SVM	240 **	0.832	0.817	0.835

Those with * markers are the number of features in the inflection point subset, and those with ** markers are the number of features in the optimal feature subset.

Table 4. Performance of SVM model using top 245 features in the list yielded by LightGBM on thirteen cancer types.

Cancer Type	F1 Measure	Cancer Type	F1 Measure
biliary tract cancer	0.865	bladder cancer	0.951
bone and soft tissue sarcomas	0.949	breast cancer	0.926
colorectal cancer	0.767	esophageal squamous cell cancer	0.896
gastric cancer	0.884	hepatocellular cancer	0.915
intraparenchymal brain tumors	0.996	lung cancer	0.835
ovarian cancer	0.948	pancreatic cancer	0.933
prostate cancer	0.971	-	-

3.3. Uncovering Biologically Significant Candidate miRNAs

Through the IFS procedure, we obtained several optimal models and features from the two analyses. The optimal features may represent informative serum miRNAs for identifying pan-cancer patients or different cancer types, thereby providing candidate classificatory signals for further evaluation. However, several optimal models used many optimal features, which may include false-positive features. In view of this, we looked for the inflection points in the IFS curves containing the highest F1 measure or weighted F1 yielded by the optimal models using more than 100 features. The inflection point was determined by the threshold-based method. The threshold for F1 measure or weighted F1 was empirically selected according to the trends of the IFS curves. The point first exceeding the threshold was defined as the inflection point. The features corresponding to the inflection point constituted the inflection point subset. The performance of the models using inflection point subsets is listed in Tables 2 and 3. Their performance was slightly

lower than that of the optimal models but required far fewer features. Figure 2 shows the comparison between models using the optimal feature subset and the inflection point subset on the list yielded by LightGBM for the second analysis. The ACC, MCC, and weighted F1 were slightly reduced, whereas the number of features was reduced from 245 to 95. This suggested that the inflection point subset retained the most informative features from the optimal feature set.

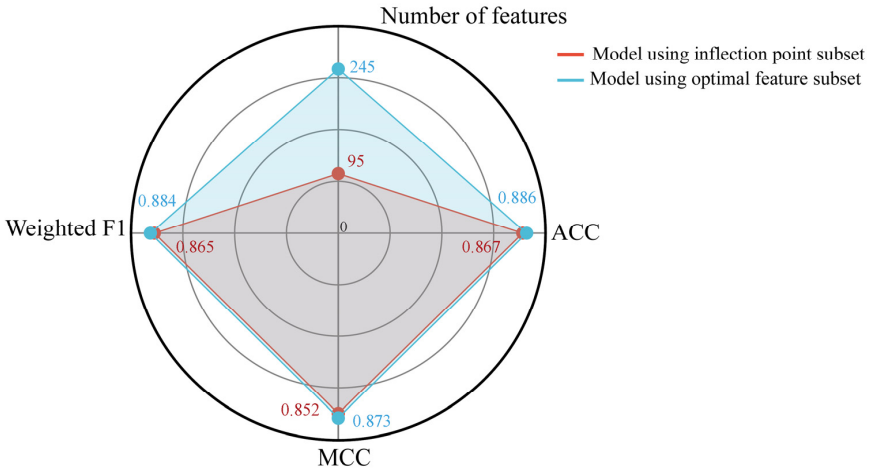


Figure 2. Radar plot to show the performance of the classification models using the optimal feature and inflection point subsets on the feature list yielded by LightGBM in the second analysis on thirteen cancer types. Radar plot showed the three metrics of two models (ACC, MCC, weighted F1) and the number of features used. The model using inflection point subset yielded similar performance but need much fewer features.

To compare the inflection point subsets extracted from different feature lists, an UpSet graph was plotted for each analysis, as shown in Figures 3 and 4. When the optimal feature subset contained fewer than 100 features, we did not identify an inflection point or a corresponding inflection point subset. The optimal feature subset was then treated as the inflection point subset for consistency. The detailed features occurring in different numbers of inflection point subsets are provided in Table S5. It can be observed that several features occurred in multiple inflection point subsets, indicating that the corresponding miRNAs were identified as essential by multiple feature-ranking algorithms. These miRNAs therefore warrant further investigation as candidate classificatory features.

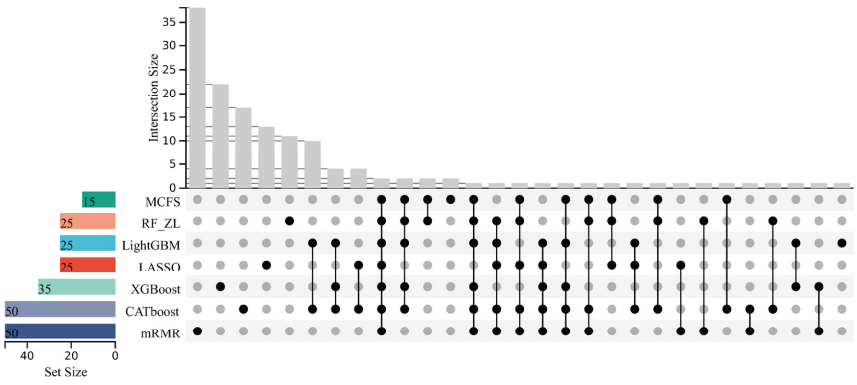


Figure 3. UpSet graph of the inflection point subsets extracted from the lists yielded by LASSO, LightGBM, MCFS, mRMR, RF_ZL, CATboost and XGBoost for the analysis on pan-cancer and non-cancer populations. The graph was plotted based on the number of occurrences of features and the subsets in which they were located. The black dots represented the subsets where the features were located.

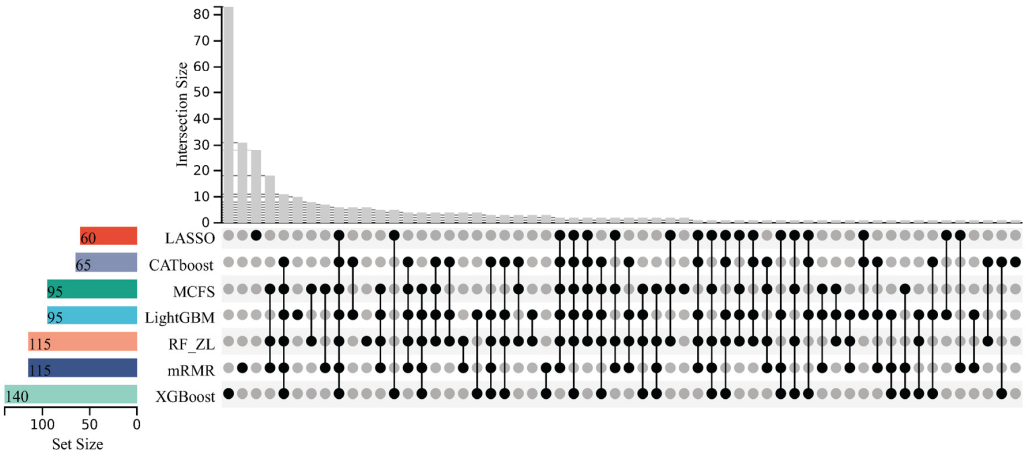


Figure 4. UpSet graph of the inflection point subsets extracted from the lists yielded by LASSO, LightGBM, MCFS, mRMR, RF_ZL, CATboost and XGBoost for the analysis on thirteen cancer types. The graph was plotted based on the number of occurrences of features and the subsets in which they were located. The black dots represented the subsets where the features were located.

3.4. Quantitative Characterization of miRNA Expression Patterns in Different Populations

Among the four classification algorithms used in this study, DT differs from the other three algorithms. It is a white-box algorithm, providing transparent classification procedures. This provides further insight into the essential differences encoded in the two datasets between pan-cancer and non-cancer patients and among patients with different cancer types. We selected the optimal features for DT on each feature list and used all patients to train the tree. A group of classification rules was obtained from the tree and is provided in Tables S6 and S7. Each rule contains several features, their thresholds, and one result (class: non-cancer or pan-cancer in the first analysis and one of thirteen cancer types in the second analysis). Because the features represent miRNAs, each rule indicates a specific miRNA expression pattern for a class. The obtained rules can be useful resources for a deeper investigation into the molecular-level alterations in cancer patients and may help generate hypotheses about associated biological processes that warrant further investigation.

4. Discussion

Using seven feature-ranking algorithms, we identified a set of key miRNAs associated with cancer status and cancer type in this serum dataset. These miRNAs contributed to classification by helping differentiate cancerous from non-cancerous states and distinguish cancer types. This approach nominates candidate serum miRNA features that warrant further investigation as classificatory signals.

To ensure that the biological interpretation of these miRNAs is anchored in the present dataset rather than in prior literature alone, for each identified miRNA we first characterize its behavior within our serum cohort, including the direction and magnitude of change between cancer and non-cancer groups, class-specific distribution across thirteen solid cancer types, and statistical significance after multiple-testing correction, and only then relate these observations to previously reported functions. In the discussion that follows, prior literature is used to contextualize, not to substitute for, the evidence provided by the present study.

4.1. Feature Analysis for Cancer Versus Non-Cancer Identification

The seven feature-ranking algorithms collectively identified 144 key miRNAs (Table S5), of which 137 miRNAs were mapped using the Mienturnet platform and the

miRTarBase database [62]. To focus on miRNA–target interactions, the top 5 miRNAs with the highest node counts were selected for KEGG enrichment analysis, as detailed in Figure S5. As shown in Figure 5, the focal adhesion and adherens junction pathways are significantly enriched. These pathways are involved in cellular motility and migration through the extracellular matrix, and altered pathway activity has been associated with tumor metastasis and drug resistance. Therefore, they provide relevant biological context for interpreting the selected serum miRNA features [63–65]. These observations are consistent with our previous KEGG analyses of cancer gene expression profiles, which also implicated cancer-related pathways [66,67]. Moreover, the enrichment of "MicroRNAs in cancer" is consistent with the classificatory signal observed in our serum data, but should be interpreted as hypothesis-generating rather than as evidence of direct tissue-level mechanism. The enrichment analysis also highlighted signaling pathways related to specific cancer types, providing context for cancer type discrimination. Furthermore, a network containing miRNA–target interactions for these five miRNAs was constructed using the miRNet platform (Figure 6), identifying 1250 gene targets [68]. Further KEGG enrichment analysis revealed that the "Pathways in Cancer" pathway was the most enriched (Table S8), with corresponding genes represented by yellow dots inside the network. Notably, some of these genes overlapped with biomarkers reported in our previous cancer studies, providing context for the current findings [66].

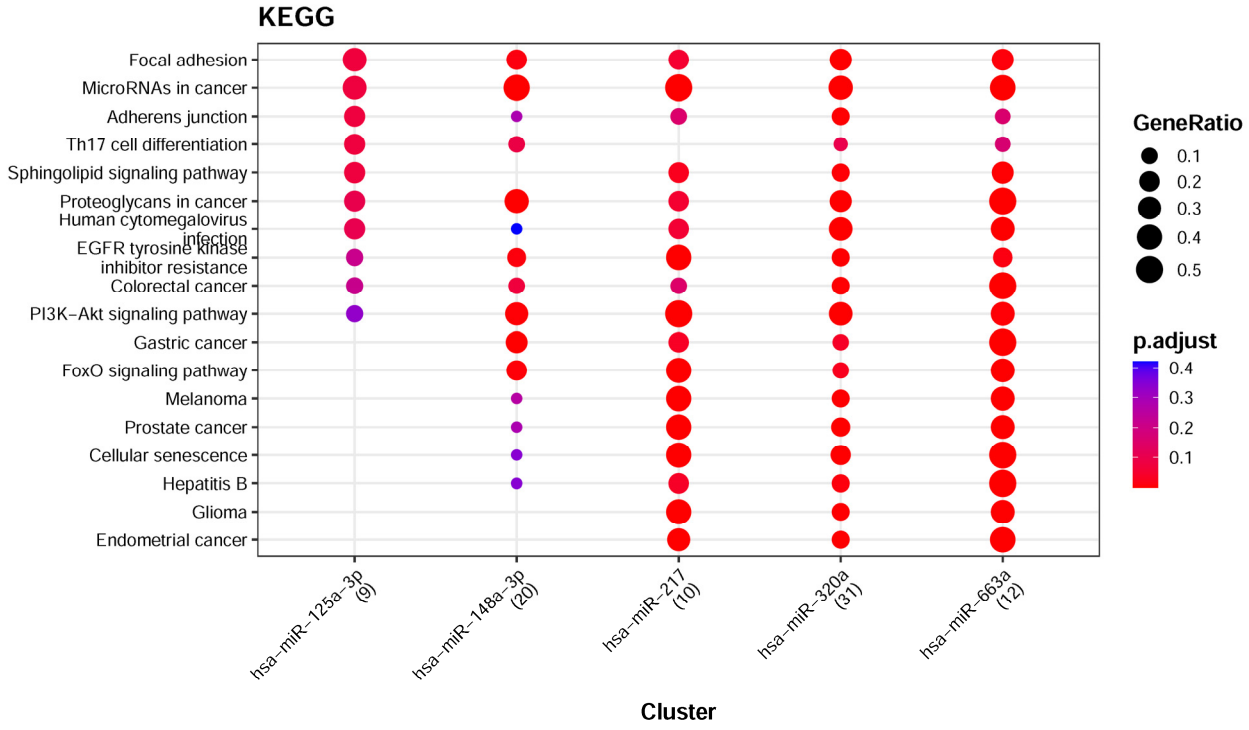


Figure 5. KEGG functional enrichment analysis of the top 5 miRNAs with the highest degree (Pan-Cancer vs. non-Cancer).

To annotate the cancer-related context of the highlighted yellow nodes, we grouped the 59 "Pathways in cancer" genes by functional module. They span the canonical subprograms of KEGG hsa05200: cell cycle and proliferation control (*MYC*, *MAX*, *CDK4/6*, *CDKN1A/1B/2A*, *E2F3*, *CKS1B*); PI3K–AKT and RAS–MAPK signaling downstream of receptor tyrosine kinases (*HRAS*, *KRAS*, *MAPK1*, *GRB2*, *SOS2*, *PIK3CD*, *CRK/CRKL*, *IGF1R*, *MET*, *KITLG*); the Wnt/β-catenin axis (*WNT1/2B/3/7B/10B*, *FZD5*, *DVL3*, *CTNNB1*, *TCF7L2*); TGF-β/SMAD signaling (*TGFB1/2*, *SMAD2*); apoptosis regulation (*TP53*, *BCL2*, *BAX*, *CYCS*, *FAS*); tumor suppression and Hippo/SUMO regulation (*PTEN*, *STK4*, *PIAS4*); inflamma-

tion and cytokine signaling (*IL6*, *STAT3*, *RELA*, *IKBKB*, *TRAF4*); angiogenesis and invasion (*VEGFA*, *HIF1A*, *MMP9*, *LAMA5*, *RHOA*, *RAC1*); and additional cancer-associated effectors including nuclear receptors (*AR*, *PPARD*), the oncogenic tyrosine kinase ABL1, HSP90 chaperones (*HSP90AA1/AB1/B1*), the HIF-regulated glucose transporter SLC2A1 (Warburg effect metabolism), and *PLCG1*. A concise per-gene annotation with the top-5 regulating miRNA(s) is provided in Table S9. These miRNA–target interactions provide pathway-level context for the serum miRNA features.

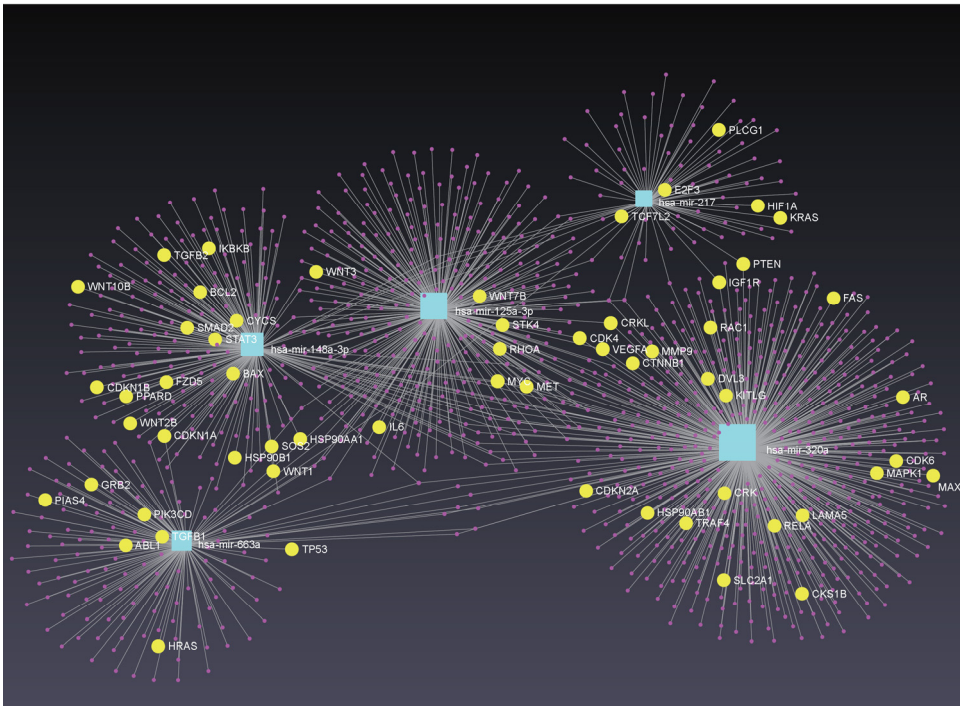


Figure 6. miRNA–Target interactions network of the top 5 miRNAs and KEGG functional enrichment analysis (Pan-Cancer vs. non-Cancer). The yellow dots represent the genes related to Pathways in cancer. A concise functional annotation of these 59 genes, together with the top 5 miRNAs that regulate each of them, is provided in Table S9.

Before relating these selected features to the existing literature, we first summarize their behavior within our own serum cohort (Table S10; class-specific distributions are shown in Figure S6). All features highlighted in this section showed significant differential expression between pan-cancer and non-cancer samples by two-sided Wilcoxon rank-sum test with Benjamini–Hochberg correction ($q < 0.005$ in every case). Specifically, miR-4783-3p (log2FC = +2.63; median log2 expression increased from 5.34 in non-cancer to 8.54 in pan-cancer; selected by 7/7 ranking methods), miR-663a (log2FC = +1.66; 10.51 to 12.30; 7/7), miR-5100 (log2FC = +3.08; 10.22 to 13.73; 5/7), miR-1307-3p (log2FC = +2.52; 5.47 to 8.42; 5/7), miR-8073 (log2FC = +2.32; 5.69 to 8.27; 4/7), miR-1469 (log2FC = +1.27; 10.58 to 11.91; 2/7) and miR-151a-5p (log2FC = +4.54; −2.95 to 3.06; 1/7) were markedly upregulated in pan-cancer serum relative to non-cancer controls, whereas miR-6784-5p (log2FC = −1.27; 12.64 to 11.26; 6/7) and miR-4456 (log2FC = −0.45; 2.19 to 0.34; 3/7) were downregulated. These quantitative observations, drawn from the present cohort, serve as the primary basis for the candidate classificatory feature claims made below; the cited literature is intended to place these findings in biological context, not to substitute for them.

Within the 144 identified key features, the highest-ranked features contributed most to discriminating cancerous from non-cancerous states, leading us to select certain features for in-depth analysis grounded in the existing literature. miR-4783-3p, recognized by all seven feature-ranking algorithms, was retained as a candidate classificatory feature in

this serum dataset. Previous research has shown that it targets the *INSM1/IA-1* gene [69], which is linked to insulinoma and is crucial for endocrine cell differentiation, marking miR-4783-3p as a potential indicator of endocrine differentiation in various tumors [70,71]. Moreover, miR-4783-3p has been acknowledged for its high diagnostic accuracy as a biomarker for ovarian cancer [69]. Similarly, miR-663a, identified by all seven algorithms, has been reported across several cancer types; prior tissue-based studies described its candidate function as a tumor suppressor through TGF-β1 regulation in liver [72] and thyroid cancer [73] cells. The regulatory elements of TGF-β signal transduction, including the pan-cancer biomarker Smad protein we previously discovered, are prevalent in numerous cancers [74]. miR-663a also inhibits the proliferation and invasion of glioblastoma cells by targeting *PIK3CD* and acts as a tumor suppressor in gastric and lung cancers [75–77]. Thus, miR-4783-3p and miR-663a are highlighted by our analysis as key candidate classificatory features warranting further validation.

For instance, the presence of miR-6784-5p in six feature subsets reflects its dysregulation in several cancers such as ovarian cancer [78], breast cancer [79], and prostate cancer [80]. It is therefore a candidate classificatory feature for further evaluation [81]. miR-1307-3p appears in five feature subsets, suggesting strong relevance to multiple cancers. It targets the DAB2-interacting protein, enhancing proliferation, invasion, and migration of liver cancer cells [82]. Additionally, it influences breast cancer proliferation by targeting the tumor suppressor gene *SMYD4* (SET and MYND domain containing 4) and contributes to gastric and colon cancer development [83–85].

miR-8073, found in four feature subsets, is associated with potential targets including *FOXM1*, *CCND1*, *MBD3*, *KLK10*, and *CASP2*, which are crucial in cell cycle regulation and proliferation [86–89]. In prior tissue-based studies, this miRNA has been reported to act as a candidate tumor suppressor associated with reduced target protein levels and inhibition of tumor growth.

miR-4456, identified in three subsets, shows downregulation in osteosarcoma and acts as a regulatory molecule in metastasis, exhibiting tumor-suppressive properties [90]. miR-1469, present in two subsets, targets *STAT5A* to regulate apoptosis in lung cancer cells [91]. miR-151a-5p, appearing in a single subset, facilitates the epithelial–mesenchymal transition in colorectal cancer cells, similar to the role of *CTBP2*, a colon cancer biomarker we previously identified [92], consistent with the reported role of EMT in colon cancer described in prior tissue-based studies.

4.2. Feature Analysis for Cancer Type Classification

Seven feature-ranking algorithms identified a total of 307 key miRNAs (Table S5), of which 297 miRNAs were mapped using the Mienturnet platform and the miRTarBase database. To further investigate miRNA-target gene relationships, the top 5 miRNAs were selected for detailed enrichment analysis, as presented in Figure S7. The subsequent KEGG pathway analysis, shown in Figure 7, highlighted several processes that are closely related to the development and progression of cancer. Of particular interest was the involvement of the PI3K-Akt signaling pathway in tumor cell proliferation [93], while the process of Efferocytosis has been noted not only for the engulfment of apoptotic cells but also for modulating the tumor microenvironment to foster immune suppression and promote tumor invasion [94].

Pathways such as MicroRNAs in cancer, Prostate cancer, and Breast cancer, which were highlighted in the enrichment analysis, indicate that miRNAs are useful in the discrimination of cancer types. The application of Disease Ontology enrichment analysis provided additional detail on the association of the miRNAs with different types of cancers. The major disease classes were cancers or tumors; as shown in Figure 8, these classes

significantly matched our samples and are consistent with the cancer-type-related signal captured by the selected serum miRNA features within this dataset.

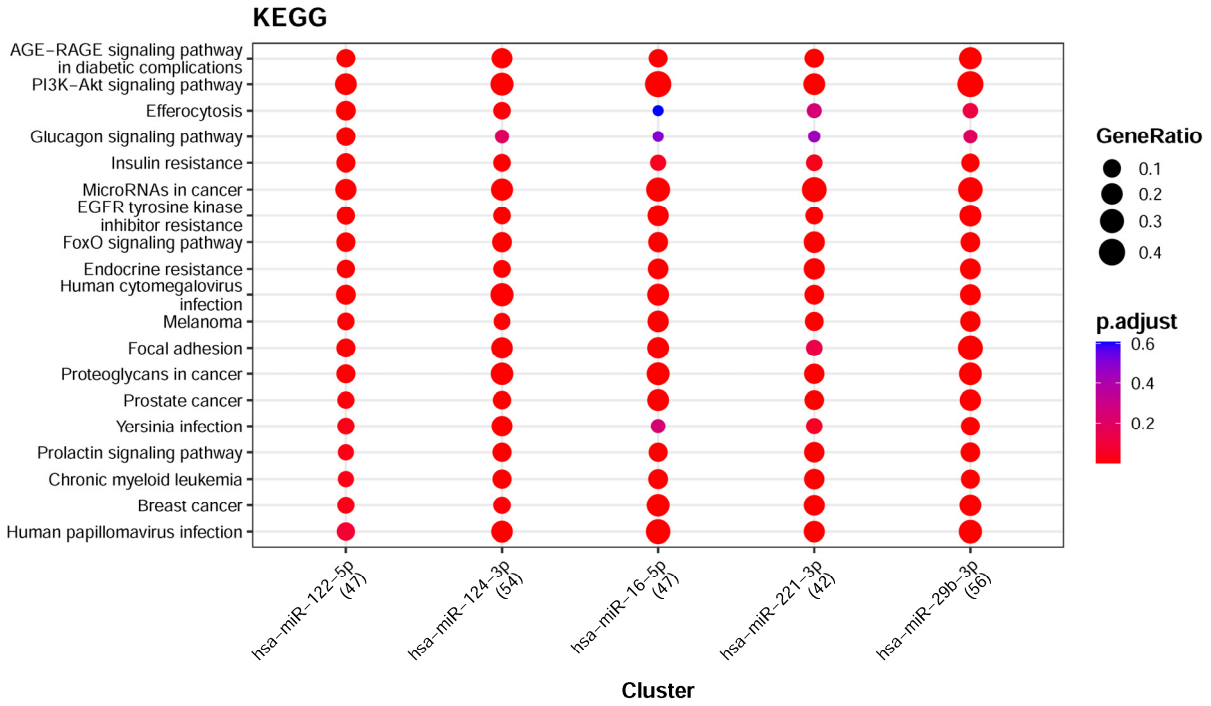


Figure 7. KEGG functional enrichment analysis of the top 5 miRNAs with the highest degree (Within Pan-Cancer).

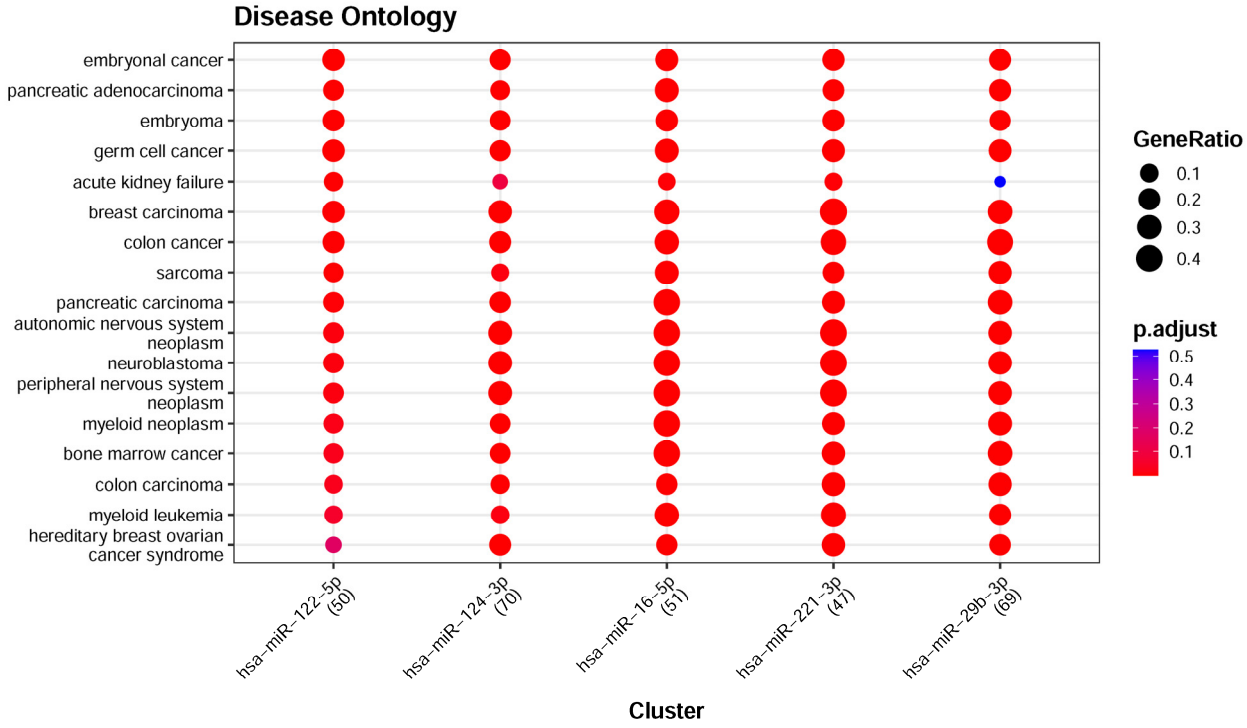


Figure 8. Disease ontology enrichment analysis of the top 5 miRNAs.

Meanwhile, the miRNA–target interaction network for the top 5 features was constructed (Figure 9). KEGG pathway analysis indicated that "Pathways in cancer" was the most significantly enriched pathway (Table S11). Notably, several related target genes

overlapped with previously identified pan-cancer biomarkers, consistent with previously reported observations.

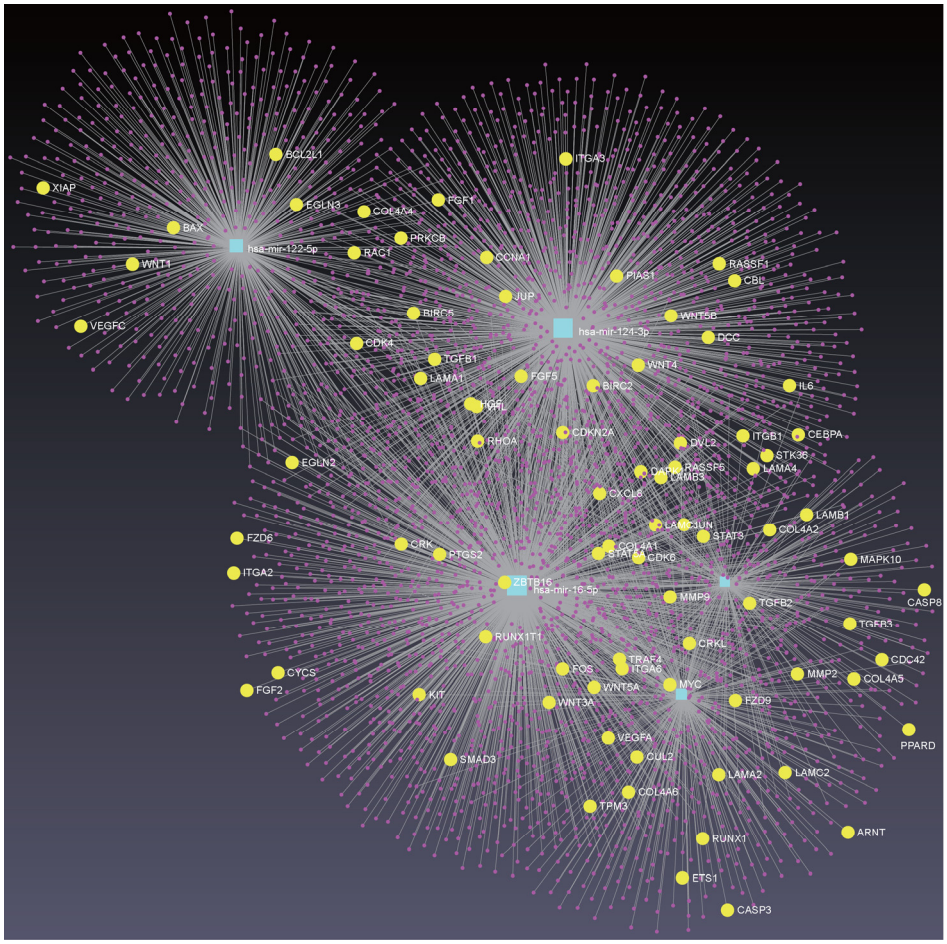


Figure 9. miRNA–Target interaction network of the top 5 miRNAs and KEGG functional enrichment analysis (Within Pan-Cancer).

To examine whether multiple miRNAs in this top-5 set share common targets, we quantified, for each of the 3825 unique genes in the Figure 9 network, how many of the five miRNAs regulate it. Fifty-seven genes are cooperatively targeted by three or more of the five miRNAs as shown in Table S12: one by all five, five by four, and 51 by three. The single gene regulated by all five miRNAs is *AKT3*, one of the three AKT isoforms and a central effector of PI3K–AKT survival and proliferation signaling, whose amplification or overexpression has been reported in ovarian, breast, and glioblastoma cancers and associated with resistance to targeted therapy [95]. Five further genes are each targeted by four of the five miRNAs. *CDK6* is a cyclin-D partner that phosphorylates Rb to drive the G1–S transition and is deregulated in lymphomas, gliomas, melanoma, and breast cancer; it is the clinical target of palbociclib, ribociclib and abemaciclib [96]. *PSMB5* encodes the catalytic chymotrypsin-like β5 subunit of the 20S proteasome, the direct target of bortezomib and carfilzomib in multiple myeloma, and its upregulation or mutation is a recognized mechanism of proteasome inhibitor resistance [97]. *DDX6* is a DEAD-box RNA helicase that regulates mRNA decay and translational repression in P-bodies and is oncogenic in colorectal, hepatocellular and head-and-neck cancers [98]. *GALNT3* (polypeptide N-acetylgalactosaminyltransferase 3) catalyzes initial mucin-type O-glycosylation of Ser/Thr residues and, when aberrantly expressed, contributes to pancreatic, ovarian, and gastric tumor invasion and poor prognosis [99]. *CALU* encodes calumenin, an ER Ca2+-

binding chaperone that is upregulated in lung adenocarcinoma and hepatocellular carcinoma, where it has been linked to migration, chemoresistance and epithelial–mesenchymal transition [100]. The 51 genes targeted by three of the five miRNAs include canonical cancer regulators such as *TP53*, *MYCN*, *CCND2*, *PIK3R1*, *MKI67*, *HMGA1* and *SP1*; EMT and invasion/MMP-regulatory factors (*ZEB2*, *RHOA*, *RECK*, *TIMP3*); basement membrane components (*COL4A1/A2*, *LAMC1*); inflammation and immune modulators (*REL/NF-κB*, *ENTPD1/CD39*); the RNAi effector AGO2; and ER stress/chaperone genes (*CALR*, *HMOX1*, *SERPINH1*). Collectively, the cooperatively regulated targets converge on PI3K–AKT signaling, cell cycle control, and ECM/invasion programs. A full annotation of all 57 shared targets, with their regulating miRNA(s) and cancer relevance notes, is provided in Table S12. As with the enrichment results above, this shared-target signal should be interpreted as hypothesis-generating rather than as direct evidence of tissue-level mechanism.

For the cancer-type classification analysis, we again first describe what is directly observable in our own dataset before turning to previously reported functions (Table S11; cross-cancer distributions and the corresponding heatmap are shown in Figure S8). All features highlighted in this section showed highly significant differential expression across the 13 solid tumor types and between pan-cancer and non-cancer samples (BH-FDR $q < 0.005$ in every case). In aggregate pan-cancer vs. non-cancer comparisons, miR-629-3p (log2FC = +0.21; selected by 7/7 ranking methods in the cancer type analysis), miR-6087 (log2FC = +0.37; 6/7), miR-422a (log2FC = +0.10; 6/7), miR-221-3p (log2FC = +2.61; 5/7), miR-29b-3p (log2FC = +4.07; 4/7), miR-4726-5p (log2FC = +0.29; 3/7) and miR-6165 (log2FC = +0.13; 2/7) were upregulated, while miR-133b (log2FC = −0.32; 1/7) was modestly downregulated. More importantly for cancer type discrimination, several of these features—notably miR-629-3p, miR-221-3p and miR-29b-3p—displayed strongly class-specific distribution patterns across the 13 tumor types (Figure S8), which is the feature-level property that makes them informative for type classification rather than for simple detection of cancer status. The biological discussion that follows should therefore be read as placing these cohort-level observations in the context of prior molecular studies.

Within the 307 identified key features, the highest-ranked features contributed most to discriminating cancer types, leading us to select certain features for in-depth analysis grounded in the existing literature. miR-629-3p, identified in seven feature subsets, emerges as a candidate classificatory feature; tissue-based studies have previously reported its involvement in various cancer types. It has been implicated in cancer progression; for instance, in breast cancer, miR-629-3p targets the leukemia inhibitory factor receptor (*LIFR*), facilitating lung metastasis [101]. In lung cancer, miR-629-3p is upregulated and targets the surfactant protein C (*SFTPC*) to promote tumor cell proliferation [102]. In the case of prostate cancer, it targets large tumor suppressor 2 (*LATS2*) to drive tumor cell development [103], while in the case of pancreatic cancer, it targets FOX3, resulting in the latter's enhanced tumor-promoting effect [104]. These reports of increased miR-629-3p expression in lung, breast, prostate, and pancreatic cancer tissues support its candidacy as a classificatory feature for these cancer types.

miR-6087, appearing in six feature subsets, is linked to osteosarcoma. Its downregulation, mediated by exosomes from human bone marrow mesenchymal stem cells (hBMSCs), impedes osteosarcoma growth [105]. It has also been identified as a specific candidate classificatory feature for bladder cancer [106]. In both osteosarcoma and bladder cancer, miR-6087's expression is diminished, reflecting its tumor-suppressive function and marking it as a candidate classificatory feature.

miR-422a, present in six feature subsets, has been implicated across several cancers. It targets *MAPK1* and *AKT1* in colorectal tumors, and proline-rich protein 2 in breast cancer stem cells, inhibiting tumor development [107,108]. In osteosarcoma, miR-422a directly

targets *TGFβ2* to regulate downstream transcriptional factors [109] or targets *BCL2L2* and *KRAS*, acting as a tumor suppressor by inhibiting cell proliferation and migration [110]. Furthermore, it is associated with increased chemotherapy sensitivity in gastric cancer [111]. The downregulation of miR-422a in rectal cancer, osteosarcoma, breast cancer, and gastric cancer tissues highlights its significance as a candidate classificatory feature.

miR-221-3p, expressed in five subsets, promotes proliferation and inhibits apoptosis in pancreatic cancer cells [112], and impedes ovarian cancer progression and migration by targeting *ARF4* [113]. It has also been reported as a diagnostic and prognostic marker for colorectal and gastric cancers [114,115]. Its involvement in the regulatory network of hepatocellular carcinoma is also substantial [116]. Therefore, miR-221-3p has been reported to be highly expressed in these cancers and is considered a critical candidate classificatory feature for pancreatic, ovarian, colorectal, liver, and gastric cancers.

miR-29b-3p, expressed in four subsets, has been described as a crucial regulatory molecule in the process of epithelial-to-mesenchymal transition in bladder and colorectal cancers [117,118]. Moreover, it plays a significant role in the treatment of breast and ovarian cancers [118,119], and it has been described as an important diagnostic marker in prostate cancer [120]. Accordingly, given its reported downregulation in these cancers, miR-29b-3p may serve as a candidate classificatory feature for these conditions.

miR-4726-5p, expressed in three subsets, acts as a tumor suppressor in hepatocellular carcinoma [121]; it was documented to be downregulated in that cancer type and may thus be considered a candidate feature for that carcinoma. miR-6165 is expressed in two subsets and induces proliferation and migration in hepatocellular carcinoma, as well as in breast cancer [122,123], and its upregulation in these tissues marks it as a candidate feature for these cancers. Finally, miR-133b appeared in one subset and is reported as a tumor suppressor in esophageal cancer [124], whose downregulated expression marks its potential as a candidate classificatory feature for the same type of cancer.

In summary, supported by multiple research studies, the selected features represent promising candidate signatures that warrant further validation in independent cohorts.

4.3. Cancer Versus Non-Cancer Classification Rule Analysis

Among the features identified by the seven classifiers, miR-663a and miR-5100 are especially strong candidate classificatory features in this dataset. As described above, miR-663a expression is upregulated in a wide range of cancers, suggesting that its candidate classificatory potential is broad. Similarly, miR-5100 expression is also upregulated in tumor cells, and overexpression of miR-5100 has been shown to increase the proliferation of lung cancer cells and downregulate Rab6 expression, leading to the growth and metastasis of tumors [125]. Conversely, miR-5100, overexpressed in pancreatic cancer, suppresses cell proliferation, migration, and invasion [126]. It is also identified as a critical regulator of apoptosis in gastric cancer [127] and a detection marker in prostate cancer [128]. While miR-4730 was selected in six classifier subsets and has previously been described in pancreatic cancer literature as a candidate molecular feature, it was reported to inhibit tumor growth in liver cancer [129]. miR-1290 is present in five classifier sets and has been widely described as a candidate classificatory feature in several types of cancers, with its possible application in early diagnosis [130]. Similarly, miR-4456 has also appeared in five classification rule sets, given its association with metastasis suppression in osteosarcoma [90].

In summary, this literature review indicates that the features identified by our machine learning approach nominate candidate miRNA classificatory features for further external validation.

4.4. Cancer Type Classification Rule Analysis

miR-1343-3p is consistently present in all 13 pan-cancer classification rules. Therefore, it may act as a candidate classificatory feature in those cancers. Previous studies have identified that miR-1343-3p targets the *TEAD4* gene, thereby regulating its expression and acting as a tumor suppressor in gastric and colorectal cancers [131,132]. In pancreatic and ovarian cancers, miR-1343-3p was inhibited by different lncRNAs, which promote tumor development [133,134]. In brief, miR-1343-3p exhibited striking variations among all 13 cancer types. Among these, established roles in six cancer types have been supported by multiple studies, while the remaining seven cancer types remained unexplored. This finding suggests its potential utility as a candidate classificatory feature.

Our findings suggest that miR-629-3p was selected as a candidate classificatory feature for cholangiocarcinoma, hepatocellular carcinoma, and pancreatic cancer, based on our classification criteria. While the involvement of miR-629-3p in pancreatic and liver cancers has been documented, its role and underlying mechanisms in cholangiocarcinoma warrant further exploration. miR-122-5p was predominantly selected as a candidate classificatory feature for bladder cancer, where urothelial carcinoma-associated 1 (*UCA1*), a marker specific to bladder cancer, is notably increased in bladder tumor cells. Acting as a tumor suppressor [135], miR-122-5p is antagonized by *UCA1*, which sponges miR-122-5p to facilitate tumorigenesis [136], providing biological context for the appearance of miR-122-5p as a candidate classificatory feature in this dataset. miR-187-5p is identified predominantly as a candidate classificatory feature for osteosarcoma and ovarian cancer; in osteosarcoma, it suppresses the *S100A4* gene, inhibiting cell proliferation and metastasis [137], while in ovarian cancer, it targets *DAB2* (homolog-2), curbing epithelial–mesenchymal transition and tumor growth [138]. miR-1225-3p, distinctively associated with a subset of breast cancers, may act as a tumor suppressor, given its reduced expression in tumors. This miRNA is diminished by circ_0000518, which indirectly elevates *SOX4*, thereby enhancing breast cancer cell proliferation and migration [139]. In colorectal cancer, miR-6746-5p, significantly overexpressed, appears to promote metastasis and is important in modulating drug response [140]. For esophageal cancer, miR-3648, which is markedly increased, targets *APC2*, playing a pivotal role in cell proliferation and prognosis [141,142]. miR-221-3p, associated mainly with hepatocellular carcinoma in this analysis, has been reported in prior tissue-based studies to promote tumor growth and invasion through suppression of inhibitors such as *SOCS3* [116,143]. Similarly, miR-328-5p, an oncogene in liver cancer, targets *DAB2IP* to enhance malignancy in hepatocellular carcinoma [144]. miR-4763-3p is mainly a classificatory feature of glioma, bone and soft tissue sarcoma, esophageal cancer, and prostate cancer. Its relationship with these cancers has not been validated, but the expression of this miRNA is upregulated in glioma patients and can be used to distinguish diffuse glioma from non-cancerous conditions [145]. miR-195-5p, a pancreatic cancer candidate classificatory feature, is suppressed in tumors, potentially acting as a tumor suppressor. This suppression is attributed to the B-RAF oncogenic kinase mutation that activates non-coding RNAs, inhibiting miR-195-5p, which is also regulated by various non-coding RNAs in pancreatic tumor progression [146,147]. Lastly, miR-1290, a marker for prostate cancer, is reduced in tumors; its secretion by cancer-associated fibroblasts promotes tumor growth and metastasis [148].

In conclusion, our research has identified candidate classificatory features that exhibit strong correlations with specific cancer types, supporting the internal utility of our classification model and rules as a starting point for further external validation in independent cohorts.

4.5. Limitations

Several limitations should be acknowledged. First, all performance metrics were obtained by ten-fold cross-validation on a single serum miRNA dataset without independent external or prospective validation; therefore, internal cross-validation, even when carefully performed, should not be regarded as a substitute for true external prospective validation or as evidence of clinically deployable performance. Accordingly, the selected miRNAs should be interpreted as candidate classificatory features and not as validated clinical biomarkers. The absence of detailed clinical information for the non-cancer cohort also means that benign, inflammatory, or other systemic conditions in some controls could not be fully assessed, which may influence estimates of classificatory specificity. Second, although SMOTE mitigated class imbalance, synthetic oversampling cannot fully substitute for real minority-class samples, and performance on rare cancer types should be interpreted with caution. Third, the biological interpretation of the selected miRNAs relies on prior literature to contextualize our data-driven observations; no direct functional validation was performed. Fourth, circulating miRNA levels are sensitive to pre-analytical and biological factors—including hemolysis, freeze–thaw cycles, storage, fasting, age, sex, stage, treatment, and batch effects—and cohort composition is inherently uneven across tumor types (e.g., breast and ovarian samples are obligately female, prostate male, and sarcoma patients substantially younger). Although all samples originated from a single biobank profiled on a uniform 3D-Gene platform with standard normalization, and Matsuzaki reported no significant association of diagnostic performance with age or sex, sample-level metadata on hemolysis, storage, freeze–thaw history, treatment, and batch were unavailable to us as secondary users. None of the miRNAs highlighted in Sections 4.1 and 4.2 belongs to well-known hemolysis-sensitive species (e.g., miR-451a, miR-486-5p, miR-92a), making hemolysis-driven confounding less likely to explain the selected features; nevertheless, we cannot formally exclude that part of the classificatory signal reflects demographic or technical structure. Independent validation in prospective, multi-center cohorts with harmonized pre-analytical protocols, matched clinical metadata, and dedicated functional assays will therefore be essential before any candidate miRNA identified here can be considered a validated clinical biomarker.

5. Conclusions

Serum miRNA profiles from patients with thirteen solid cancer types (pan-cancer) and non-cancer populations were analyzed using a machine learning approach. The study consisted of two independent analyses: (1) to differentiate between pan-cancer and non-cancer populations, and (2) to classify samples among thirteen predefined solid cancer types within this dataset. For the first analysis, the optimal model showed high internal cross-validation performance for binary classification, and decision-tree-derived rules revealed quantitative miRNA expression thresholds associated with the cancer versus non-cancer distinction. For the second analysis, internal cross-validation-based discrimination among the thirteen cancer types was achieved, and class-specific candidate miRNAs were identified. We derived quantitative classification rules that help summarize serum miRNA expression patterns. It should be noted that this closed-set classification task does not equate to cancer-of-unknown-primary diagnosis or metastatic tumor origin tracing, which would require fundamentally different study designs. These results nominate candidate serum miRNA classificatory features for further investigation rather than validated clinical biomarkers; all findings are based on internal cross-validation of a single retrospective dataset and should be considered exploratory and hypothesis-generating. Direct mechanistic and therapeutic relevance at the tissue level was not addressed by the present analysis

and would require independent validation in prospective multi-center cohorts, alongside dedicated tissue-based and functional studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/life16050850/s1, Figure S1: IFS curves for the first analysis on pan-cancer and non-cancer populations comparison using the feature lists yielded by four feature-ranking algorithms. *y*-axis is the F1 measure metric and *x*-axis is the number of features. A. IFS curves based on the feature list yielded by LASSO. B. IFS curves based on the feature list yielded by LightGBM. C. IFS curves based on the feature list yielded by MCFS. D. IFS curves based on the feature list yielded by mRMR. Figure S2: IFS curves for the first analysis on pan-cancer and non-cancer populations comparison using the feature lists yielded by three feature-ranking algorithms. *y*-axis is the F1 measure metric and *x*-axis is the number of features. A. IFS curves based on the feature list yielded by RF_ZL. B. IFS curves based on the feature list yielded by CATboost. C. IFS curves based on the feature list yielded by XGBoost. Figure S3: IFS curves for the second analysis on patients with thirteen different cancer types using the feature lists yielded by four feature-ranking algorithms. *y*-axis is the Weighted F1 metric and *x*-axis is the number of features. A. IFS curves based on the feature list yielded by LASSO. B. IFS curves based on the feature list yielded by LightGBM. C. IFS curves based on the feature list yielded by MCFS. D. IFS curves based on the feature list yielded by mRMR. Figure S4: IFS curves for the second analysis on patients with thirteen different cancer types using the feature lists yielded by three feature-ranking algorithms. *y*-axis is the Weighted F1 metric and *x*-axis is the number of features. A. IFS curves based on the feature list yielded by RF_ZL. B. IFS curves based on the feature list yielded by CATboost. C. IFS curves based on the feature list yielded by XGBoost. Figure S5: Network degree plots of microRNA-genes network based on MIENTURNET platform and miRTarBase database (Pan-Cancer vs. non-Cancer). Figure S6: Violin plots showing the serum expression distributions of the nine miRNAs in Pan-Cancer versus Non-Cancer samples. Figure S7: Network degree plots of microRNA-genes network based on MIENTURNET platform and miRTarBase database (Within Pan-Cancer). Figure S8: Heatmap showing the class-specific expression distributions of the eight miRNAs across the thirteen solid cancer types. File S1: Description on feature-ranking and classification algorithms. Table S1. Details of machine learning algorithms used in this study. Table S2: Feature lists yielded by LASSO, LightGBM, MCFS, mRMR, RF_ZL, CATboost, XGBoost for two analyses. Table S3: Performance of the classification models constructed in the IFS procedure for the first analysis on pan-cancer vs. non-cancer. Table S4: Performance of the classification models constructed in the IFS procedure for the second analysis on thirteen cancer types. Table S5: Intersection of inflection point subsets identified from feature lists yielded by LASSO, LightGBM, MCFS, mRMR, RF_ZL, CATboost, XGBoost. The features that appear in 7, 6, 5, 4, 3, 2, and 1 inflection point subset are shown. Table S6: Classification rules generated by the optimal decision tree models on seven feature lists for the first analysis on pan-cancer vs. non-cancer. Table S7: Classification rules generated by the optimal decision tree models on seven feature lists for the second analysis on thirteen cancer types. Table S8: KEGG enrichment analysis results of miRNA-target network based on the miRNet platform for the first analysis on pan-cancer vs. non-cancer. Table S9: Description of genes related to the Pathways in Cancer and their distribution in functional sub-modules. Table S10: Expression statistics and ranking-algorithm selection for the miRNAs we selected. For each miRNA, the table reports mean ± SD and median [IQR] of log2-transformed serum expression in Non-Cancer and Pan-Cancer groups, log2 fold change (Pan-Cancer minus Non-Cancer), Wilcoxon rank-sum test *p* value with Benjamini–Hochberg correction (q), and the number of the seven ranking methods that selected the feature. Table S11: KEGG enrichment analysis results of miRNA-target network based on the miRNet platform for the second analysis on thirteen cancer types. Table S12: Shared miRNA target genes and their functions (within Pan-Cancer).

Author Contributions: Conceptualization, T.H. and Y.-D.C.; methodology, K.F., Y.B. and J.R.; validation, T.H.; formal analysis, K.F., Y.B., J.R., W.G. and D.W.; data curation, T.H.; writing—original draft preparation, K.F., Y.B. and J.R.; writing—review and editing, Y.-D.C.; funding acquisition, T.H. and Y.-D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Major Project of Guangzhou National Laboratory (GZNL2024A01003), the Fund of the Key Laboratory of Tissue Microenvironment and Tumor of Chinese Academy of Sciences (202002), Shandong Provincial Natural Science Foundation (ZR2022MC072).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data analyzed in this study is available in GEO database with accession ID GSE211692 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE211692, accessed on 20 March 2023). The analyzed results are contained within the article or Supplementary Material.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

miRNA	MicroRNA
CATboost	Categorical boosting
LASSO	Least absolute shrinkage and selection operator
LightGBM	Light gradient boosting machine
MCFS	Monte Carlo feature selection
mRMR	Minimum redundancy maximum relevance
XGBoost	Extreme gradient boosting
SMOTE	Synthetic minority oversampling technique
IFS	Incremental feature selection
RF	Random forest
KNN	k-nearest neighbor
DT	Decision tree
SVM	Support vector machine
LIFR	Leukemia inhibitory factor receptor
SFTPC	Surfactant protein C
LATS2	Large tumor suppressor 2
hBMSC	Human bone marrow mesenchymal stem cell

References

1. Roy, P.S.; Saikia, B.J. Cancer and cure: A critical analysis. *Indian J. Cancer* **2016**, *53*, 441–442. [CrossRef]
2. Robinson, D.R.; Wu, Y.-M.; Lonigro, R.J.; Vats, P.; Cobain, E.; Everett, J.; Cao, X.; Rabban, E.; Kumar-Sinha, C.; Raymond, V.; et al. Integrative clinical genomics of metastatic cancer. *Nature* **2017**, *548*, 297–303. [CrossRef] [PubMed]
3. Srinivas, P.R.; Kramer, B.S.; Srivastava, S. Trends in biomarker research for cancer detection. *Lancet Oncol.* **2001**, *2*, 698–704. [CrossRef] [PubMed]
4. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529. [CrossRef] [PubMed]
5. Lujambio, A.; Lowe, S.W. The microcosmos of cancer. *Nature* **2012**, *482*, 347–355. [CrossRef]
6. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; et al. Microrna expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838. [CrossRef] [PubMed]
7. Hill, M.; Tran, N. Mirna interplay: Mechanisms and consequences in cancer. *Dis. Model. Mech.* **2021**, *14*, dmm047662. [CrossRef]
8. Bartel, D.P. Micrornas: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [PubMed]
9. Saliminejad, K.; Khorram Khorshid, H.R.; Soleymani Fard, S.; Ghaffari, S.H. An overview of microrna: Biology, functions, therapeutics, and analysis methods. *J. Cell Physiol.* **2019**, *234*, 5451–5465. [PubMed]
10. He, B.; Zhao, Z.; Cai, Q.; Zhang, Y.; Zhang, P.; Shi, S.; Xie, H.; Peng, X.; Yin, W.; Tao, Y.; et al. Mirna-based biomarkers, therapies, and resistance in cancer. *Int. J. Biol. Sci.* **2020**, *16*, 2628–2647.
11. Hussen, B.M.; Hidayat, H.J.; Salihi, A.; Sabir, D.K.; Taheri, M.; Ghafouri-Fard, S. Microrna: A signature for cancer progression. *Biomed. Pharmacother.* **2021**, *138*, 111528. [CrossRef] [PubMed]

12. Lin, S.C.; Liu, C.J.; Lin, J.A.; Chiang, W.F.; Hung, P.S.; Chang, K.W. Mir-24 up-regulation in oral carcinoma: Positive association from clinical and in vitro analysis. *Oral Oncol.* **2010**, *46*, 204–208. [CrossRef]
13. Gao, Y.; Liu, Y.; Du, L.; Li, J.; Qu, A.; Zhang, X.; Wang, L.; Wang, C. Down-regulation of mir-24-3p in colorectal cancer is associated with malignant behavior. *Med. Oncol.* **2015**, *32*, 362. [CrossRef]
14. Liu, F.; Dong, H.; Mei, Z.; Huang, T. Investigation of mirna and mrna co-expression network in ependymoma. *Front. Bioeng. Biotechnol.* **2020**, *8*, 177. [CrossRef]
15. Zhou, Y.; Zheng, X.; Xu, B.; Hu, W.; Huang, T.; Jiang, J. The identification and analysis of mrna-lncrna-mirna cliques from the integrative network of ovarian cancer. *Front. Genet.* **2019**, *10*, 751. [CrossRef]
16. Liu, C.; Zhang, Y.H.; Deng, Q.; Li, Y.; Huang, T.; Zhou, S.; Cai, Y.D. Cancer-related triplets of mrna-lncrna-mirna revealed by integrative network in uterine corpus endometrial carcinoma. *Biomed. Res. Int.* **2017**, *2017*, 3859582. [CrossRef]
17. Crowley, E.; Di Nicolantonio, F.; Loupakis, F.; Bardelli, A. Liquid biopsy: Monitoring cancer-genetics in the blood. *Nat. Rev. Clin. Oncol.* **2013**, *10*, 472–484. [CrossRef]
18. Siravegna, G.; Marsoni, S.; Siena, S.; Bardelli, A. Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 531–548. [CrossRef]
19. Nikanjam, M.; Kato, S.; Kurzrock, R. Liquid biopsy: Current technology and clinical applications. *J. Hematol. Oncol.* **2022**, *15*, 131. [CrossRef] [PubMed]
20. Grewal, J.K.; Tessier-Cloutier, B.; Jones, M.; Gakkhar, S.; Ma, Y.; Moore, R.; Mungall, A.J.; Zhao, Y.; Taylor, M.D.; Gelmon, K.; et al. Application of a neural network whole transcriptome–based pan-cancer method for diagnosis of primary and metastatic cancers. *JAMA Netw. Open* **2019**, *2*, e192597. [CrossRef] [PubMed]
21. Singh, A.; Shannon, C.P.; Gautier, B.; Rohart, F.; Vacher, M.; Tebbutt, S.J.; Lê Cao, K.-A. Diablo: An integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **2019**, *35*, 3055–3062. [CrossRef]
22. Mar-Aguilar, F.; Mendoza-Ramírez, J.A.; Malagón-Santiago, I.; Espino-Silva, P.K.; Santuario-Facio, S.K.; Ruiz-Flores, P.; Rodríguez-Padilla, C.; Reséndez-Pérez, D. Serum circulating microrna profiling for identification of potential breast cancer biomarkers. *Dis. Markers* **2013**, *34*, 163–169. [CrossRef]
23. Li, Y.; Kang, K.; Krahn, J.M.; Croutwater, N.; Lee, K.; Umbach, D.M.; Li, L. A comprehensive genomic pan-cancer classification using the cancer genome atlas gene expression data. *BMC Genom.* **2017**, *18*, 508. [CrossRef] [PubMed]
24. Zhao, N.; Guo, M.; Wang, K.; Zhang, C.; Liu, X. Identification of pan-cancer prognostic biomarkers through integration of multi-omics data. *Front. Bioeng. Biotechnol.* **2020**, *8*, 268. [CrossRef] [PubMed]
25. Ranstam, J.; Cook, J.A. Lasso regression. *Br. J. Surg.* **2018**, *105*, 1348. [CrossRef]
26. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*; Curran Associates Inc.: Long Beach, CA, USA, 2017; pp. 3149–3157.
27. Dramiński, M.; Koronacki, J. Rmcfs: An r package for monte carlo feature selection and interdependency discovery. *J. Stat. Softw.* **2018**, *85*, 1–28. [CrossRef]
28. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef] [PubMed]
29. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
30. Dorogush, A.V.; Ershov, V.; Gulin, A. Catboost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363. [CrossRef]
31. Chen, T.; Guestrin, C. Xgboost: A Scalable Tree Boosting System. In *Proceedings of the The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016*; Association for Computing Machinery: San Francisco, CA, USA, 2016; pp. 785–794.
32. Liu, H.; Setiono, R. Incremental feature selection. *Appl. Intell.* **1998**, *9*, 217–230. [CrossRef]
33. Safavian, S.R.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man. Cybern.* **1991**, *21*, 660–674. [CrossRef]
34. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. Knn Model-Based Approach in Classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, Proceedings of the OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, 3–7 November 2003*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 986–996.
35. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
36. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. Smote: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
37. Matsuzaki, J.; Kato, K.; Oono, K.; Tsuchiya, N.; Sudo, K.; Shimomura, A.; Tamura, K.; Shiino, S.; Kinoshita, T.; Daiko, H.; et al. Prediction of tissue-of-origin of early stage cancers using serum mirnomes. *JNCI Cancer Spectr.* **2023**, *7*, pkac080. [CrossRef]

38. Li, H.; Huang, F.; Liao, H.; Li, Z.; Feng, K.; Huang, T.; Cai, Y.D. Identification of COVID-19-specific immune markers using a machine learning method. *Front. Mol. Biosci.* **2022**, *9*, 952626. [CrossRef]
39. Li, Z.; Guo, W.; Ding, S.; Chen, L.; Feng, K.; Huang, T.; Cai, Y.D. Identifying key microrna signatures for neurodegenerative diseases with machine learning methods. *Front. Genet.* **2022**, *13*, 880997. [CrossRef]
40. Lu, J.; Meng, M.; Zhou, X.; Ding, S.; Feng, K.; Zeng, Z.; Huang, T.; Cai, Y.D. Identification of COVID-19 severity biomarkers based on feature selection on single-cell rna-seq data of cd8(+) t cells. *Front. Genet.* **2022**, *13*, 1053772. [CrossRef]
41. Ren, J.; Zhou, X.; Huang, K.; Chen, L.; Guo, W.; Feng, K.; Huang, T.; Cai, Y.D. Identification of key genes associated with persistent immune changes and secondary immune activation responses induced by influenza vaccination after COVID-19 recovery by machine learning methods. *Comput. Biol. Med.* **2023**, *169*, 107883. [CrossRef] [PubMed]
42. Ma, Q.L.; Huang, F.M.; Guo, W.; Feng, K.Y.; Huang, T.; Cai, Y.D. Machine learning classification of time since bnt162b2 COVID-19 vaccination based on array-measured antibody activity. *Life* **2023**, *13*, 1304. [CrossRef]
43. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the International Joint Conference on Artificial Intelligence, Montreal, QC, Canada, 20–25 August 1995*; Lawrence Erlbaum Associates Ltd.: Montreal, QC, Canada, 1995; pp. 1137–1145.
44. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
45. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
46. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
47. Liao, H.; Ma, Q.; Chen, L.; Guo, W.; Feng, K.; Bao, Y.; Zhang, Y.; Shen, W.; Huang, T.; Cai, Y.-D. Machine learning analysis of cd4+ t cell gene expression in diverse diseases: Insights from cancer, metabolic, respiratory, and digestive disorders. *Cancer Genet.* **2025**, *290–291*, 56–60. [CrossRef]
48. Chen, L.; Lu, Y.; Xu, J.; Zhou, B. Prediction of drug's anatomical therapeutic chemical (atc) code by constructing biological profiles of atc codes. *BMC Bioinform.* **2025**, *26*, 86. [CrossRef]
49. Yuan, F.; Huang, F.; Cao, X.; Zhang, Y.-H.; Feng, K.; Bao, Y.; Huang, T.; Cai, Y.-D. Integrative multi-omics machine learning reveals novel driver genes associations in lung adenocarcinoma. *Biochim. Biophys. Acta (BBA) Proteins Proteom.* **2026**, *1874*, 141113. [CrossRef]
50. Ren, J.; Liao, H.; Chen, L.; Bao, Y.; Guo, W.; Feng, K.; Huang, T.; Cai, Y.-D. Identification of gene signatures associated with multisystem inflammatory syndrome in children after SARS-CoV-2 infection. *Curr. Bioinform.* **2025**. [CrossRef]
51. Chen, L.; Yang, J.; Zhou, B.; Cai, Y.-D. Plysptm-hgnn: Predicting lysine ptm sites of proteins using hybrid graph neural networks. *BMC Bioinform.* **2026**, *27*, 32. [CrossRef] [PubMed]
52. Powers, D. Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation. *J. Mach. Learn. Technol.* **2011**, *2*, 37–63.
53. Chen, L.; Chen, Y.; Zhou, B. Hclamcmi: Prediction of circrna-mirna interactions based on hypergraph contrastive learning and an attention mechanism. *J. Chem. Inf. Model.* **2025**, *65*, 12099–12115. [CrossRef]
54. Ren, J.; Gao, Q.; Zhou, X.; Feng, K.; Guo, W.; Huang, T.; Cai, Y.-D. Identification of gene signatures differentiating cancer from normal tissues across histological classifications of gastric adenocarcinoma via machine learning methods. *Biochem. Genet.* **2026**. [CrossRef] [PubMed]
55. Ma, Q.; Ren, J.; Chen, L.; Guo, W.; Feng, K.; Zhang, Y.; Shen, W.; Huang, T.; Cai, Y.-D. Identifying transcriptional signatures of leukocytes in tissue and blood for multicancer diagnosis by using machine learning methods. *Cancer Genet.* **2026**, *302–303*, 13–26. [CrossRef]
56. Chen, L.; Gu, J.; Zhou, B. Pmislocmf: Predicting mirna subcellular localizations by incorporating multi-source features of mirnas. *Brief. Bioinform.* **2024**, *25*, bbae386. [CrossRef]
57. Chen, L.; Xu, L.; Zhou, B.; Chen, Y. Anticannet: A graph convolution and chemical llm framework for predicting anti-cancer small molecules. *Curr. Bioinform.* **2026**.
58. Chen, L.; Hu, J.; Zhou, B. Predicting circrna subcellular localization by fusing circrna sequence and network information. *Sci. Rep.* **2026**, *16*, 12775. [CrossRef] [PubMed]
59. Chen, L.; Xun, X.; Zhou, B. Root-associated protein prediction using a protein large language model and hypergraph convolutional networks. *Sci. Rep.* **2026**, *16*, 4876. [CrossRef] [PubMed]
60. Bao, Y.; Zhou, X.; Feng, K.; Guo, W.; Huang, T.; Cai, Y.-D. Convergent autoencoder and machine learning methodologies for SARS-CoV-2 essential host factor identification based on interferon pathway loss-of-function perturb-seq data mining. *Curr. Bioinform.* **2026**.
61. Matthews, B.W. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochim. Biophys. Acta* **1975**, *405*, 442–451. [CrossRef]
62. Licursi, V.; Conte, F.; Fiscon, G.; Paci, P. Mienturnet: An interactive web tool for microrna-target enrichment and network-based analysis. *BMC Bioinform.* **2019**, *20*, 545. [CrossRef] [PubMed]

63. Boissan, M.; De Wever, O.; Lizarraga, F.; Wendum, D.; Poincloux, R.; Chignard, N.; Desbois-Mouthon, C.; Dufour, S.; Nawrocki-Raby, B.; Birembaut, P.; et al. Implication of metastasis suppressor nm23-h1 in maintaining adherens junctions and limiting the invasive potential of human cancer cells. *Cancer Res.* **2010**, *70*, 7710–7722. [CrossRef]
64. Eke, I.; Cordes, N. Focal adhesion signaling and therapy resistance in cancer. *Semin. Cancer Biol.* **2015**, *31*, 65–75. [CrossRef]
65. Tilghman, R.W.; Parsons, J.T. Focal adhesion kinase as a regulator of cell tension in the progression of cancer. *Semin. Cancer Biol.* **2008**, *18*, 45–52. [CrossRef]
66. Ding, S.; Li, H.; Zhang, Y.H.; Zhou, X.; Feng, K.; Li, Z.; Chen, L.; Huang, T.; Cai, Y.D. Identification of pan-cancer biomarkers based on the gene expression profiles of cancer cell lines. *Front. Cell Dev. Biol.* **2021**, *9*, 781285. [CrossRef]
67. Li, J.; Xu, Q.; Wu, M.; Huang, T.; Wang, Y. Pan-cancer classification based on self-normalizing neural networks and feature selection. *Front. Bioeng. Biotechnol.* **2020**, *8*, 766. [CrossRef]
68. Chang, L.; Zhou, G.; Soufan, O.; Xia, J. Mirnet 2.0: Network-based visual analytics for mirna functional analysis and systems biology. *Nucleic Acids Res.* **2020**, *48*, W244–W251. [CrossRef]
69. Hamidi, F.; Gilani, N.; Belaghi, R.A.; Sarbakhsh, P.; Edgünlü, T.; Santaguida, P. Exploration of potential mirna biomarkers and prediction for ovarian cancer using artificial intelligence. *Front. Genet.* **2021**, *12*, 724785. [CrossRef]
70. Gierl, M.S.; Karoulias, N.; Wende, H.; Strehle, M.; Birchmeier, C. The zinc-finger factor insm1 (ia-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. *Genes Dev.* **2006**, *20*, 2465–2478. [CrossRef]
71. Juhlin, C.C.; Zedenius, J.; Höög, A. Clinical routine application of the second-generation neuroendocrine markers isl1, insm1, and secretagogin in neuroendocrine neoplasia: Staining outcomes and potential clues for determining tumor origin. *Endocr. Pathol.* **2020**, *31*, 401–410. [CrossRef]
72. Zhang, C.; Chen, B.; Jiao, A.; Li, F.; Sun, N.; Zhang, G.; Zhang, J. Mir-663a inhibits tumor growth and invasion by regulating tgf-β1 in hepatocellular carcinoma. *BMC Cancer* **2018**, *18*, 1179. [CrossRef]
73. Wang, Z.; Zhang, H.; Zhang, P.; Dong, W.; He, L. Microrna-663 suppresses cell invasion and migration by targeting transforming growth factor beta 1 in papillary thyroid carcinoma. *Tumor Biol.* **2016**, *37*, 7633–7644. [CrossRef]
74. Lu, J.; Li, J.; Ren, J.; Ding, S.; Zeng, Z.; Huang, T.; Cai, Y.D. Functional and embedding feature analysis for pan-cancer classification. *Front. Oncol.* **2022**, *12*, 979336. [CrossRef]
75. Shi, Y.; Chen, C.; Zhang, X.; Liu, Q.; Xu, J.-L.; Zhang, H.-R.; Yao, X.-H.; Jiang, T.; He, Z.-C.; Ren, Y.; et al. Primate-specific mir-663 functions as a tumor suppressor by targeting pik3cd and predicts the prognosis of human glioblastoma. *Clin. Cancer Res.* **2014**, *20*, 1803–1813. [CrossRef]
76. Pan, J.; Hu, H.; Zhou, Z.; Sun, L.; Peng, L.; Yu, L.; Sun, L.; Liu, J.; Yang, Z.; Ran, Y. Tumor-suppressive mir-663 gene induces mitotic catastrophe growth arrest in human gastric cancer cells. *Oncol. Rep.* **2010**, *24*, 105–112. [CrossRef]
77. Liu, Z.Y.; Zhang, G.L.; Wang, M.M.; Xiong, Y.N.; Cui, H.Q. Microrna-663 targets tgfb1 and regulates lung cancer proliferation. *Asian Pac. J. Cancer Prev.* **2011**, *12*, 2819–2823.
78. Yaghoobi, H.; Babaei, E.; Hussen, B.M.; Emami, A. Ebst: An evolutionary multi-objective optimization based tool for discovering potential biomarkers in ovarian cancer. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2021**, *18*, 2384–2393. [CrossRef]
79. Li, Z.; Chen, Z.; Hu, G.; Zhang, Y.; Feng, Y.; Jiang, Y.; Wang, J. Profiling and integrated analysis of differentially expressed circrnas as novel biomarkers for breast cancer. *J. Cell Physiol.* **2020**, *235*, 7945–7959. [CrossRef]
80. Liu, H.-P.; Lai, H.-M.; Guo, Z. Prostate cancer early diagnosis: Circulating microrna pairs potentially beyond single micrornas upon 1231 serum samples. *Brief. Bioinform.* **2021**, *22*, bbaa111. [CrossRef]
81. Chen, J.W.; Dhahbi, J. Identification of four serum mirnas as potential markers to screen for thirteen cancer types. *PLoS ONE* **2022**, *17*, e0269554. [CrossRef]
82. Chen, S.; Wang, L.; Yao, B.; Liu, Q.; Guo, C. Mir-1307-3p promotes tumor growth and metastasis of hepatocellular carcinoma by repressing dab2 interacting protein. *Biomed. Pharmacother.* **2019**, *117*, 109055. [CrossRef]
83. Han, S.; Zou, H.; Lee, J.W.; Han, J.; Kim, H.C.; Cheol, J.J.; Kim, L.S.; Kim, H. Mir-1307-3p stimulates breast cancer development and progression by targeting smyd4. *J. Cancer* **2019**, *10*, 441–448. [CrossRef]
84. Guo, Z.; Zhang, Y.; Xu, W.; Zhang, X.; Jiang, J. Engineered exosome-mediated delivery of circdido1 inhibits gastric cancer progression via regulation of mir-1307-3p/socs2 axis. *J. Transl. Med.* **2022**, *20*, 326. [CrossRef]
85. Chen, Q.; Mao, Y.; Meng, F.; Wang, L.; Zhang, H.; Wang, W.; Hua, D. Rs7911488 modified the efficacy of capecitabine-based therapy in colon cancer through altering mir-1307-3p and tyms expression. *Oncotarget* **2017**, *8*, 74312–74319. [CrossRef]
86. Duan, N.; Hu, X.; Yang, X.; Cheng, H.; Zhang, W. Microrna-370 directly targets foxm1 to inhibit cell growth and metastasis in osteosarcoma cells. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 10250–10260.
87. Zhang, L.; Zheng, Y.; Sun, Y.; Zhang, Y.; Yan, J.; Chen, Z.; Jiang, H. Mir-134-mbd3 axis regulates the induction of pluripotency. *J. Cell Mol. Med.* **2016**, *20*, 1150–1158. [CrossRef]
88. White, N.M.; Chow, T.F.; Mejia-Guerrero, S.; Diamandis, M.; Rofael, Y.; Faragalla, H.; Mankaruous, M.; Gabril, M.; Girgis, A.; Yousef, G.M. Three dysregulated mirnas control kallikrein 10 expression and cell proliferation in ovarian cancer. *Br. J. Cancer* **2010**, *102*, 1244–1253. [CrossRef]

89. Guha, M.; Xia, F.; Raskett, C.M.; Altieri, D.C. Caspase 2-mediated tumor suppression involves survivin gene silencing. *Oncogene* **2010**, *29*, 1280–1292. [CrossRef] [PubMed]
90. Si, Z.; Zhou, S.; Shen, Z.; Luan, F.; Yan, J. Knockdown of lncrna fezf1-as1 inhibits metastasis of osteosarcoma cells by mir-4456/galnt10. *Res. Sq.* **2020**. [CrossRef]
91. Xu, C.; Zhang, L.; Li, H.; Liu, Z.; Duan, L.; Lu, C. Mirna-1469 promotes lung cancer cells apoptosis through targeting stat5a. *Am. J. Cancer Res.* **2015**, *5*, 1180–1189.
92. Zhang, N.; Wang, M.; Zhang, P.; Huang, T. Classification of cancers based on copy number variation landscapes. *Biochim. Biophys. Acta* **2016**, *1860*, 2750–2755. [CrossRef]
93. Luo, J.; Manning, B.D.; Cantley, L.C. Targeting the pi3k-akt pathway in human cancer: Rationale and promise. *Cancer Cell* **2003**, *4*, 257–262. [CrossRef]
94. Qiu, H.; Shao, Z.; Wen, X.; Liu, Z.; Chen, Z.; Qu, D.; Ding, X.; Zhang, L. Efferocytosis: An accomplice of cancer immune escape. *Biomed. Pharmacother.* **2023**, *167*, 115540. [CrossRef]
95. Turner, K.M.; Sun, Y.; Ji, P.; Granberg, K.J.; Bernard, B.; Hu, L.; Cogdell, D.E.; Zhou, X.; Yli-Harja, O.; Nykter, M. Genomically amplified akt3 activates DNA repair pathway and promotes glioma progression. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3421–3426. [CrossRef] [PubMed]
96. Álvarez-Fernández, M.; Malumbres, M. Mechanisms of sensitivity and resistance to cdk4/6 inhibition. *Cancer Cell* **2020**, *37*, 514–529. [CrossRef]
97. Oerlemans, R.; Franke, N.E.; Assaraf, Y.G.; Cloos, J.; van Zantwijk, I.; Berkers, C.R.; Scheffer, G.L.; Debipersad, K.; Vojtekova, K.; Lemos, C. Molecular basis of bortezomib resistance: Proteasome subunit β5 (psmb5) gene mutation and overexpression of psmb5 protein. *Blood J. Am. Soc. Hematol.* **2008**, *112*, 2489–2499. [CrossRef]
98. Taniguchi, K.; Iwatsuki, A.; Sugito, N.; Shinohara, H.; Kuranaga, Y.; Oshikawa, Y.; Tajirika, T.; Futamura, M.; Yoshida, K.; Uchiyama, K. Oncogene rna helicase ddx6 promotes the process of c-myc expression in gastric cancer cells. *Mol. Carcinog.* **2018**, *57*, 579–589. [CrossRef]
99. Chugh, S.; Meza, J.; Sheinin, Y.M.; Ponnusamy, M.P.; Batra, S.K. Loss of n-acetylgalactosaminyltransferase 3 in poorly differentiated pancreatic cancer: Augmented aggressiveness and aberrant erbb family glycosylation. *Br. J. Cancer* **2016**, *114*, 1376–1386. [CrossRef]
100. Li, Y.; Sun, S.; Zhang, H.; Jing, Y.; Ji, X.; Wan, Q.; Liu, Y. Calu promotes lung adenocarcinoma progression by enhancing cell proliferation, migration and invasion. *Respir. Res.* **2024**, *25*, 267. [CrossRef] [PubMed]
101. Wang, J.; Song, C.; Tang, H.; Zhang, C.; Tang, J.; Li, X.; Chen, B.; Xie, X. Mir-629-3p may serve as a novel biomarker and potential therapeutic target for lung metastases of triple-negative breast cancer. *Breast Cancer Res.* **2017**, *19*, 72. [CrossRef] [PubMed]
102. Li, B.; Meng, Y.-Q.; Li, Z.; Yin, C.; Lin, J.-P.; Zhu, D.-J.; Zhang, S.-B. Mir-629-3p-induced downregulation of sftpc promotes cell proliferation and predicts poor survival in lung adenocarcinoma. *Artif. Cells Nanomed. Biotechnol.* **2019**, *47*, 3286–3296. [CrossRef] [PubMed]
103. Li, Y.; Zeng, S.; Cao, L. Mir-629 repressed lats2 expression and promoted the proliferation of prostate cancer cells. *Horm. Metab. Res.* **2023**, *55*, 573–579. [CrossRef]
104. Yan, H.; Li, Q.; Wu, J.; Hu, W.; Jiang, J.; Shi, L.; Yang, X.; Zhu, D.; Ji, M.; Wu, C. Mir-629 promotes human pancreatic cancer progression by targeting foxo3. *Cell Death Dis.* **2017**, *8*, e3154. [CrossRef]
105. Yang, X.; Zhou, Y.; Lai, Q.; Huang, S.; Liu, X.; Hu, B.; Dai, M.; Zhang, B. Mirna-6087 delivered by exosomes derived from human bone marrow mesenchymal stem cells, inhibition of mir-6087 in exosomes can significantly inhibit the proliferation, invasion and migration of tumor cells. *Res. Sq.* **2020**. [CrossRef]
106. Usuba, W.; Urabe, F.; Yamamoto, Y.; Matsuzaki, J.; Sasaki, H.; Ichikawa, M.; Takizawa, S.; Aoki, Y.; Niida, S.; Kato, K.; et al. Circulating mirna panels for specific and early detection in bladder cancer. *Cancer Sci.* **2019**, *110*, 408–419. [CrossRef]
107. Wei, W.-T.; Nian, X.-X.; Wang, S.-Y.; Jiao, H.-L.; Wang, Y.-X.; Xiao, Z.-Y.; Yang, R.-W.; Ding, Y.-Q.; Ye, Y.-P.; Liao, W.-T. Mir-422a inhibits cell proliferation in colorectal cancer by targeting akt1 and mapk1. *Cancer Cell Int.* **2017**, *17*, 91. [CrossRef] [PubMed]
108. Zou, Y.; Chen, Y.; Yao, S.; Deng, G.; Liu, D.; Yuan, X.; Liu, S.; Rao, J.; Xiong, H.; Yuan, X.; et al. Mir-422a weakened breast cancer stem cells properties by targeting plp2. *Cancer Biol. Ther.* **2018**, *19*, 436–444. [CrossRef]
109. Liu, M.; Xiusheng, H.; Xiao, X.; Wang, Y. Overexpression of mir-422a inhibits cell proliferation and invasion, and enhances chemosensitivity in osteosarcoma cells. *Oncol. Rep.* **2016**, *36*, 3371–3378. [CrossRef]
110. Zhang, H.; He, Q.-Y.; Wang, G.-C.; Tong, D.-K.; Wang, R.-K.; Ding, W.-B.; Li, C.; Wei, Q.; Ding, C.; Liu, P.-Z.; et al. Mir-422a inhibits osteosarcoma proliferation by targeting bcl2l2 and kras. *Biosci. Rep.* **2018**, *38*, BSR20170339. [CrossRef]
111. Zhou, Z.; Lin, Z.; He, Y.; Pang, X.; Wang, Y.; Ponnusamy, M.; Ao, X.; Shan, P.; Tariq, M.A.; Li, P.; et al. The long noncoding rna d63785 regulates chemotherapy sensitivity in human gastric cancer by targeting mir-422a. *Mol. Ther. Nucleic Acids* **2018**, *12*, 405–419. [CrossRef]
112. Li, F.; Xu, J.W.; Wang, L.; Liu, H.; Yan, Y.; Hu, S.Y. Microrna-221-3p is up-regulated and serves as a potential biomarker in pancreatic cancer. *Artif. Cells Nanomed. Biotechnol.* **2018**, *46*, 482–487. [CrossRef]

113. Wu, Q.; Ren, X.; Zhang, Y.; Fu, X.; Li, Y.; Peng, Y.; Xiao, Q.; Li, T.; Ouyang, C.; Hu, Y.; et al. Mir-221-3p targets arf4 and inhibits the proliferation and migration of epithelial ovarian cancer cells. *Biochem. Biophys. Res. Commun.* **2018**, *497*, 1162–1170. [CrossRef]
114. Tao, K.; Yang, J.; Guo, Z.; Hu, Y.; Sheng, H.; Gao, H.; Yu, H. Prognostic value of mir-221-3p, mir-342-3p and mir-491-5p expression in colon cancer. *Am. J. Transl. Res.* **2014**, *6*, 391–401. [PubMed]
115. Zhang, Y.; Huang, H.; Zhang, Y.; Liao, N. Combined detection of serum mir-221-3p and mir-122-5p expression in diagnosis and prognosis of gastric cancer. *J. Gastric Cancer* **2019**, *19*, 315–328. [CrossRef]
116. Li, H.; Zhang, B.; Ding, M.; Lu, S.; Zhou, H.; Sun, D.; Wu, G.; Gan, X. C1qtnf1-as1 regulates the occurrence and development of hepatocellular carcinoma by regulating mir-221-3p/socs3. *Hepatol. Int.* **2019**, *13*, 277–292. [CrossRef] [PubMed]
117. Luo, M.; Li, Z.; Wang, W.; Zeng, Y.; Liu, Z.; Qiu, J. Long non-coding rna h19 increases bladder cancer metastasis by associating with ezh2 and inhibiting e-cadherin expression. *Cancer Lett.* **2013**, *333*, 213–221. [CrossRef] [PubMed]
118. Ding, D.; Li, C.; Zhao, T.; Li, D.; Yang, L.; Zhang, B. Lncrna h19/mir-29b-3p/pgrn axis promoted epithelial-mesenchymal transition of colorectal cancer cells by acting on wnt signaling. *Mol. Cells* **2018**, *41*, 423–435.
119. Gu, L.; Li, Q.; Liu, H.; Lu, X.; Zhu, M. Long noncoding rna tug1 promotes autophagy-associated paclitaxel resistance by sponging mir-29b-3p in ovarian cancer cells. *OncoTargets Ther.* **2020**, *13*, 2007–2019. [CrossRef] [PubMed]
120. Worst, T.S.; Previti, C.; Nitschke, K.; Diessl, N.; Gross, J.C.; Hoffmann, L.; Frey, L.; Thomas, V.; Kahlert, C.; Bieback, K.; et al. Mir-10a-5p and mir-29b-3p as extracellular vesicle-associated prostate cancer detection markers. *Cancers* **2020**, *12*, 43. [CrossRef]
121. Tang, Q.; Li, X.; Chen, Y.; Long, S.; Yu, Y.; Sheng, H.; Wang, S.; Han, L.; Wu, W. Solamargine inhibits the growth of hepatocellular carcinoma and enhances the anticancer effect of sorafenib by regulating hottip-tug1/mir-4726-5p/muc1 pathway. *Mol. Carcinog.* **2022**, *61*, 417–432. [CrossRef]
122. Shen, D.; Zhao, H.; Zeng, P.; Ge, M.; Shrestha, S.; Zhao, W. Circular rna circ_0001459 accelerates hepatocellular carcinoma progression via the mir-6165/igf1r axis. *Ann. N. Y. Acad. Sci.* **2022**, *1512*, 46–60. [CrossRef] [PubMed]
123. Ebrahimi, S.O.; Reiisi, S. Mir-6165 dysregulation in breast cancer and its effect on cell proliferation and migration. *ISMJ* **2021**, *24*, 439–453. [CrossRef]
124. Kano, M.; Seki, N.; Kikkawa, N.; Fujimura, L.; Hoshino, I.; Akutsu, Y.; Chiyomaru, T.; Enokida, H.; Nakagawa, M.; Matsubara, H. Mir-145, mir-133a and mir-133b: Tumor-suppressive mirnas target fscn1 in esophageal squamous cell carcinoma. *Int. J. Cancer* **2010**, *127*, 2804–2814. [CrossRef]
125. Huang, H.; Jiang, Y.; Wang, Y.; Chen, T.; Yang, L.; He, H.; Lin, Z.; Liu, T.; Yang, T.; Kamp, D.W.; et al. Mir-5100 promotes tumor growth in lung cancer by targeting rab6. *Cancer Lett.* **2015**, *362*, 15–24. [CrossRef] [PubMed]
126. Chijiiwa, Y.; Moriyama, T.; Ohuchida, K.; Nabae, T.; Ohtsuka, T.; Miyasaka, Y.; Fujita, H.; Maeyama, R.; Manabe, T.; Abe, A.; et al. Overexpression of microrna-5100 decreases the aggressive phenotype of pancreatic cancer cells by targeting podxl. *Int. J. Oncol.* **2016**, *48*, 1688–1700. [CrossRef]
127. Zhang, H.; Wang, J.; Wang, Y.; Li, J.; Zhao, L.; Zhang, T.; Liao, X. Long non-coding lef1-as1 sponge mir-5100 regulates apoptosis and autophagy in gastric cancer cells via the mir-5100/dek/ampk-mtor axis. *Int. J. Mol. Sci.* **2022**, *23*, 4787. [CrossRef]
128. Mello-Grand, M.; Bruno, A.; Sacchetto, L.; Cristoni, S.; Gregnanin, I.; Dematteis, A.; Zitella, A.; Gontero, P.; Peraldo-Neia, C.; Ricotta, R.; et al. Two novel ceramide-like molecules and mir-5100 levels as biomarkers improve prediction of prostate cancer in gray-zone psa. *Front. Oncol.* **2021**, *11*, 769158. [CrossRef]
129. Yang, W.; Ju, H.Y.; Tian, X.F. Hsa-mir-4730 as a new and potential diagnostic and prognostic indicators for pancreatic cancer. *Eur. Rev. Med. Pharmacol. Sci.* **2020**, *24*, 8801–8811.
130. Kalhori, M.R.; Soleimani, M.; Arefian, E.; Alizadeh, A.M.; Mansouri, K.; Echeverria, J. The potential role of mir-1290 in cancer progression, diagnosis, prognosis, and treatment: An oncomir or onco-suppressor microrna? *J. Cell Biochem.* **2022**, *123*, 506–531. [CrossRef]
131. Wang, L.; Bo, X.; Yi, X.; Xiao, X.; Zheng, Q.; Ma, L.; Li, B. Exosome-transferred linc01559 promotes the progression of gastric cancer via pi3k/akt signaling pathway. *Cell Death Dis.* **2020**, *11*, 723. [CrossRef]
132. Bhat, I.P.; Rather, T.B.; Bhat, G.A.; Maqbool, I.; Akhtar, K.; Rashid, G.; Parray, F.Q.; Besina, S.; Mudassar, S. Tead4 nuclear localization and regulation by mir-4269 and mir-1343-3p in colorectal carcinoma. *Pathol.-Res. Pract.* **2022**, *231*, 153791. [CrossRef]
133. Chen, X.; Wang, J.; Xie, F.; Mou, T.; Zhong, P.; Hua, H.; Liu, P.; Yang, Q. Long noncoding rna linc01559 promotes pancreatic cancer progression by acting as a competing endogenous rna of mir-1343-3p to upregulate raf1 expression. *Aging* **2020**, *12*, 14452–14466. [CrossRef]
134. Li, Y.; Zhao, Z.; Sun, D.; Li, Y. Novel long noncoding rna linc02323 promotes cell growth and migration of ovarian cancer via tgf-β receptor 1 by mir-1343-3p. *J. Clin. Lab. Anal.* **2021**, *35*, e23651. [CrossRef]
135. Wang, X.S.; Zhang, Z.; Wang, H.C.; Cai, J.L.; Xu, Q.W.; Li, M.Q.; Chen, Y.C.; Qian, X.P.; Lu, T.J.; Yu, L.Z.; et al. Rapid identification of uca1 as a very sensitive and specific unique marker for human bladder carcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2006**, *12*, 4851–4858. [CrossRef]
136. Zhou, Y.; Meng, X.; Chen, S.; Li, W.; Li, D.; Singer, R.; Gu, W. Imp1 regulates uca1-mediated cell invasion through facilitating uca1 decay and decreasing the sponge effect of uca1 for mir-122-5p. *Breast Cancer Res.* **2018**, *20*, 32. [CrossRef]

137. Xiao, Y.; Zhao, Q.; Du, B.; Chen, H.-y.; Zhou, D.-Z. Microrna-187 inhibits growth and metastasis of osteosarcoma by downregulating s100a4. *Cancer Investig.* **2018**, *36*, 1–9. [CrossRef]
138. Chao, A.; Lin, C.Y.; Lee, Y.S.; Tsai, C.L.; Wei, P.C.; Hsueh, S.; Wu, T.I.; Tsai, C.N.; Wang, C.J.; Chao, A.S.; et al. Regulation of ovarian cancer progression by microrna-187 through targeting disabled homolog-2. *Oncogene* **2012**, *31*, 764–775. [CrossRef]
139. Shujuan, K.; Zhongxin, L.; Jingfang, M.; Zhili, C.; Wei, W.; Liu, Q.; Li, Y. Circular rna circ_0000518 promotes breast cancer progression through the microrna-1225-3p/sry-box transcription factor 4 pathway. *Bioengineered* **2022**, *13*, 2611–2622. [CrossRef]
140. Tubita, V.; Segui-Barber, J.; Lozano, J.J.; Banon-Maneus, E.; Rovira, J.; Cucchiari, D.; Moya-Rull, D.; Oppenheimer, F.; Del Portillo, H.; Campistol, J.M.; et al. Effect of immunosuppression in mirnas from extracellular vesicles of colorectal cancer and their influence on the pre-metastatic niche. *Sci. Rep.* **2019**, *9*, 11177. [CrossRef]
141. Rashid, F.; Awan, H.M.; Shah, A.; Chen, L.; Shan, G. Induction of mir-3648 upon er stress and its regulatory role in cell proliferation. *Int. J. Mol. Sci.* **2017**, *18*, 1375. [CrossRef]
142. Zhang, D.; Yin, H.; Bauer, T.L.; Rogers, M.P.; Velotta, J.B.; Morgan, C.T.; Du, W.; Xu, P.; Qian, X. Development of a novel mir-3648-related gene signature as a prognostic biomarker in esophageal adenocarcinoma. *Ann. Transl. Med.* **2021**, *9*, 1702. [CrossRef]
143. Dong, Y.; Zhang, N.; Zhao, S.; Chen, X.; Li, F.; Tao, X. Mir-221-3p and mir-15b-5p promote cell proliferation and invasion by targeting axin2 in liver cancer. *Oncol. Lett.* **2019**, *18*, 6491–6500. [CrossRef] [PubMed]
144. Liu, Z.; Yu, Y.; Huang, Z.; Kong, Y.; Hu, X.; Xiao, W.; Quan, J.; Fan, X. Circrna-5692 inhibits the progression of hepatocellular carcinoma by sponging mir-328-5p to enhance dab2ip expression. *Cell Death Dis.* **2019**, *10*, 900. [CrossRef]
145. Ohno, M.; Matsuzaki, J.; Kawauchi, J.; Aoki, Y.; Miura, J.; Takizawa, S.; Kato, K.; Sakamoto, H.; Matsushita, Y.; Takahashi, M.; et al. Assessment of the diagnostic utility of serum microrna classification in patients with diffuse glioma. *JAMA Netw. Open* **2019**, *2*, e1916953. [CrossRef]
146. Gao, K.-F.; Zhao, Y.-F.; Liao, W.-J.; Xu, G.-L.; Zhang, J.-D. Cers6-as1 promotes cell proliferation and represses cell apoptosis in pancreatic cancer via mir-195-5p/wipi2 axis. *Kaohsiung J. Med. Sci.* **2022**, *38*, 542–553. [CrossRef]
147. Zhou, W.-Y.; Zhang, M.-M.; Liu, C.; Kang, Y.; Wang, J.-O.; Yang, X.-H. Long noncoding rna linc00473 drives the progression of pancreatic cancer via upregulating programmed death-ligand 1 by sponging microrna-195-5p. *J. Cell Physiol.* **2019**, *234*, 23176–23189. [CrossRef] [PubMed]
148. Wang, S.; Du, P.; Cao, Y.; Ma, J.; Yang, X.; Yu, Z.; Yang, Y. Cancer associated fibroblasts secreted exosomal mir-1290 contributes to prostate cancer cell growth and metastasis via targeting gsk3β. *Cell Death Discov.* **2022**, *8*, 371. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.