MDPI

Article

Integrated Tissue and Blood miRNA Expression Profiles Identify Novel Biomarkers for Accurate Non-Invasive Diagnosis of Breast Cancer: Preliminary Results and Future Clinical Implications

Fei Su [1,†], Ziyu Gao [1,2,†], Yueyang Liu [1], Guiqin Zhou [3], Ying Cui [1], Chao Deng [1,2], Yuyu Liu [1,2], Yihao Zhang [2], Xiaoyan Ma [2], Yongxia Wang [2], Lili Guan [4], Yafang Zhang [2,*] and Baoquan Liu [2,5,*]

1 College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China
2 Department of Anatomy, Harbin Medical University, Harbin 150081, China
3 Department of Immunology, Harbin Medical University, Harbin 150081, China
4 Department of Information Management, Shanghai Lixin University of Accounting and Finance, Shanghai 200000, China
5 Department of Modern Medicine and Pharmacy, University of Tibetan Medicine, Lhasa 850000, China
* Correspondence: 100549@hrbmu.edu.cn (B.L.); yafangzhang2008@aliyun.com (Y.Z.)
† These authors contributed equally to this work.

Abstract: We aimed to identify miRNAs that were closely related to breast cancer (BRCA). By integrating several methods including significance analysis of microarrays, fold change, Pearson's correlation analysis, *t* test, and receiver operating characteristic analysis, we developed a decision-tree-based scoring algorithm, called Optimized Scoring Mechanism for Primary Synergy MicroRNAs (O-PSM). Five synergy miRNAs (hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p) were identified using O-PSM, which were used to distinguish normal samples from pathological ones, and showed good results in blood data and in multiple sets of tissue data. These five miRNAs showed accurate categorization efficiency in BRCA typing and staging and had better categorization efficiency than experimentally verified miRNAs. In the Protein-Protein Interaction (PPI) network, the target genes of hsa-miR-342-5p have the most regulatory relationships, which regulate carcinogenesis proliferation and metastasis by regulating Glycosaminoglycan biosynthesis and the Rap1 signaling pathway. Moreover, hsa-miR-342-5p showed potential clinical application in survival analysis. We also used O-PSM to generate an R package uploaded on github (SuFei-lab/OPSM accessed on 22 October 2021). We believe that miRNAs included in O-PSM could have clinical implications for diagnosis, prognostic stratification and treatment of BRCA, proposing potential significant biomarkers that could be utilized to design personalized treatment plans in BRCA patients in the future.

Keywords: tissue; blood; breast cancer; primary synergy miRNA; diagnostic marker

1. Introduction

Breast cancer (BRCA) is the most common cancer worldwide [1–5], accounting for 11.7% of new cancer cases, which has risen rapidly [6,7]. To diagnose BRCA in its early stages, reduce its mortality, and reduce treatment-related harm to patients [8,9], a current challenge is to identify noninvasive and accurate biological markers to be used as indicators for early screening and diagnosis of BRCA.

MicroRNAs (miRNAs) are a class of non-coding single-stranded RNA molecules with a length of approximately 22 nucleic acids [1,10–12], which can be found in tissues and blood [13–15]. For patients with different cancers, tumor-specific or related changes have been found in the free nucleic acids in blood circulation [16,17]. MiRNAs constitute short non-coding RNAs of post-transcriptional regulatory genes [18–20] and play an important

role in development [21] and in normal physiological activities, and they can act as carcinogens [22,23] or tumor inhibitory regulators [24,25]. An increasing number of studies have recently reported close associations between miRNAs and cancers [25,26]. According to REMARK's suggestion [27], we aimed to identify miRNAs closely related to BRCA as clinical diagnostic markers [1,28–30].

The majority of previous studies have focused on tissues [31,32] or blood [33,34] to search for cancer-related markers, but we hope to identify markers that have significantly combined properties in both tissues and blood, to increase diagnostic accuracy. Blood-derived markers [35–39] are not stable and may show opposite trends in different studies. Tissue-derived markers perform more consistently, but tissue sampling can be physically damaging to patients. Therefore, we hope to find markers of tissue and blood coexistence for BRCA identification and diagnosis. Furthermore, we searched for "tissue-blood shared miRNAs as cancer biomarkers" regarding BRCA or other carcinomas, but there were only a few studies that focus on marker miRNAs that work in both tissues and blood [40–44]. Thereby, we hope to fill the gap of miRNAs shared by tissues and blood as tumor biomarkers.

In this study, we integrated many methods—Significance Analysis of Microarrays (SAM), fold change (FC), Pearson's correlation analysis, *t* test, receiver operating characteristic (ROC), and decision trees—to develop a new method, called Optimized Scoring Mechanism for Primary Synergy MicroRNAs (O-PSM), that can identify disease-related miRNAs that a play major role in BRCA. This algorithm was performed for each feature selection set, first filtering key miRNAs using SAM, FC, Pearson's correlation analysis, and *t* test, then constructing decision trees to obtain synergistic miRNAs, and in each tree, further selecting miRNA sets to ensure they lead to maximum purity at each branch of the nodes. This binary tree was split until it stopped growing; thus, the miRNAs in the tree represent a set of feature combinations from the root node to the leaf nodes that contribute synergistically to the classification. The combined miRNAs classify the samples sequentially according to the hierarchical structure of the tree and jointly decide to identify the diagnosis of BRCA.

We used O-PSM to process and analyze BRCA data obtained from The Cancer Genome Atlas (TCGA) and tested the classification performance in tissues and blood using data obtained from the Gene Expression Omnibus (GEO), to identify five BRCA-related miRNAs: hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p, which exhibited satisfactory classification efficiencies in both tissues and blood. The highest single area under the curve (AUC) of these five miRNAs was the AUC of hsa-miR-139-5p, which was equal to 0.9941 with a CI of 98.44–100%, and combined AUC of 0.9975 with a CI of 99.36–100%. In independent sets, the highest single AUC was 0.8793 (CI: 78.69–97.16%) from hsa-miR-486-5p, and the combined AUC was 0.9105 (CI: 99.36–100%) in serum, with 0.9714 (CI: 93.4–100%) from hsa-miR-139-5p and 1.0000 (CI: 99.36~100%) in the tissue independent set 1, 0.9714 (CI: 93.4–100%), from hsa-miR-139-5p and 1.0000 (CI: 99.36–100%) in tissue independent set 2.

We also assessed the categorization efficiency of the five screened miRNAs for BRCA staging and typing, the latter being the one better characterized, surprisingly, by the panel. In addition, we compared hsa-miR-125a-5p [45,46] and hsa-miR-146b-5p [45,47], which were experimentally confirmed in tissues and blood from previous reports. The results showed that miRNAs identified by O-PSM had better categorization efficiency. We also analyzed the target genes of five miRNAs and obtained Protein-Protein Interaction (PPI) networks and functional enrichment of target genes by STRING and Metascape databases. In addition, survival analysis showed that high expression of hsa-miR-342-5p had a great impact on patient prognosis. Our results showed that the combined use of the panel of identified miRNAs may be effective for non-invasive diagnosis and prognostic stratification of patients, in order to design a novel biomarkers-based approach for BRCA management in clinical practice.

2. Materials and Methods

2.1. Data

The TCGA BRCA RNA-seq data were obtained from the UCSC Xena database (https://xenabrowser.net/ (accessed on 1 October 2020)), which included data from 748 disease patients, 187 healthy patients, and a total of 2253 miRNAs. The prognostic data and phenotype information were also downloaded from the UCSC Xena database. There were three independent test sets in the GEO (https://www.ncbi.nlm.nih.gov/geo/ (accessed on 1 October 2020)) for the classification effectiveness test; two of them involved tissue data, and one involved blood data. The tissue independent set 1 was from GSE42128 [48] platform 3 (GPL15018 Agilent-031181 Unrestricted_Human_miRNA_V16.0_Microarray 030840), with 1205 miRNAs, which had data from 28 cancer patients and 20 para-carcinoma patients. The tissue independent set 2 was from GSE57897 [49] (GPL18722→Homo sapiens microRNA array), with 1849 miRNAs, which had data from 422 cancer patients and 31 healthy controls. The blood data were from GSE42128 platform 2 (GPL16224→Exiqon LNA RT-PCR Human panels (1 and 2)), with 274 miRNAs, which had preoperative serum data from 32 cancer patients and 22 healthy controls (Figure 1A).

We used GSE81002 [50–52] (GPL10656 Agilent-029297 Human miRNA Microarray v14 Rev.2 (miRNA ID version)) and GSE97811 [53] (GPL21263 3D-Gene Human miRNA V21_1.0.0, microarray data) from GEO to validate the classification effect of the screened miRNAs for BRCA typing and staging. GSE81002 (GPL10656), which has 128 miRNAs overlapping with TCGA BRCA data, and all 5 marker miRNAs in it, left a total of 425 samples. We removed 50 samples without pam50 subtype; then, there were 45 basal-like samples, 44 normal-like samples, 155 Luminal A samples, 89 Luminal B samples, and 42 HER-2 (+) samples, 375 samples in total. GSE97811, which has 2222 miRNAs overlap with TCGA BRCA data, and all 5 marker miRNAs in it, consisted of 61 samples, including 28 stage 1 samples, 28 stage 2 samples, and 5 stage 3 samples.

2.2. Random Sampling

Random sampling was performed using BRCA data downloaded from TCGA. The diseased and healthy data were divided into four parts for random combination; three were training sets and one was a test set. To enrich our research samples and generate more accurate results, two-thirds of the training set was divided into a feature selection set, and one-third was the feature evaluation set. One sampling produced 12 feature selection sets, and we sampled 100 times at random to produce 1200 feature selection sets. We finally obtained 1200 feature selection sets ($i1k1 - i12k100$), 1200 feature evaluation sets ($i1k1 - i12k100$), and 1200 internal test sets ($i1k1 - i12k100$), to use in subsequent data analyses (Figure 1B). I is the number of permutations of random combinations, k is the number of random samples.

2.3. Differential Expression Analyses

The SAM method (R Package: siggenes) was then used to select differentially expressed miRNAs from 1200 features selection sets, and characteristic miRNAs with $p < 0.05$ were identified in each set. FC was performed on the characteristic miRNAs obtained from SAM in each set, and the threshold was set as $\log_2(|FC\ value|) > \log_2(1.2)$, to identify the miRNAs with significant differential expressions (Figure 1B).

2.4. Clustering

To further obtain co-expression relationships between the miRNAs identified by SAM and FC, these miRNAs were clustered using Pearson's correlation analysis with correlation coefficient (r) > 0.6, to obtain the sets of co-expressed characteristic miRNAs (Figure 1B).
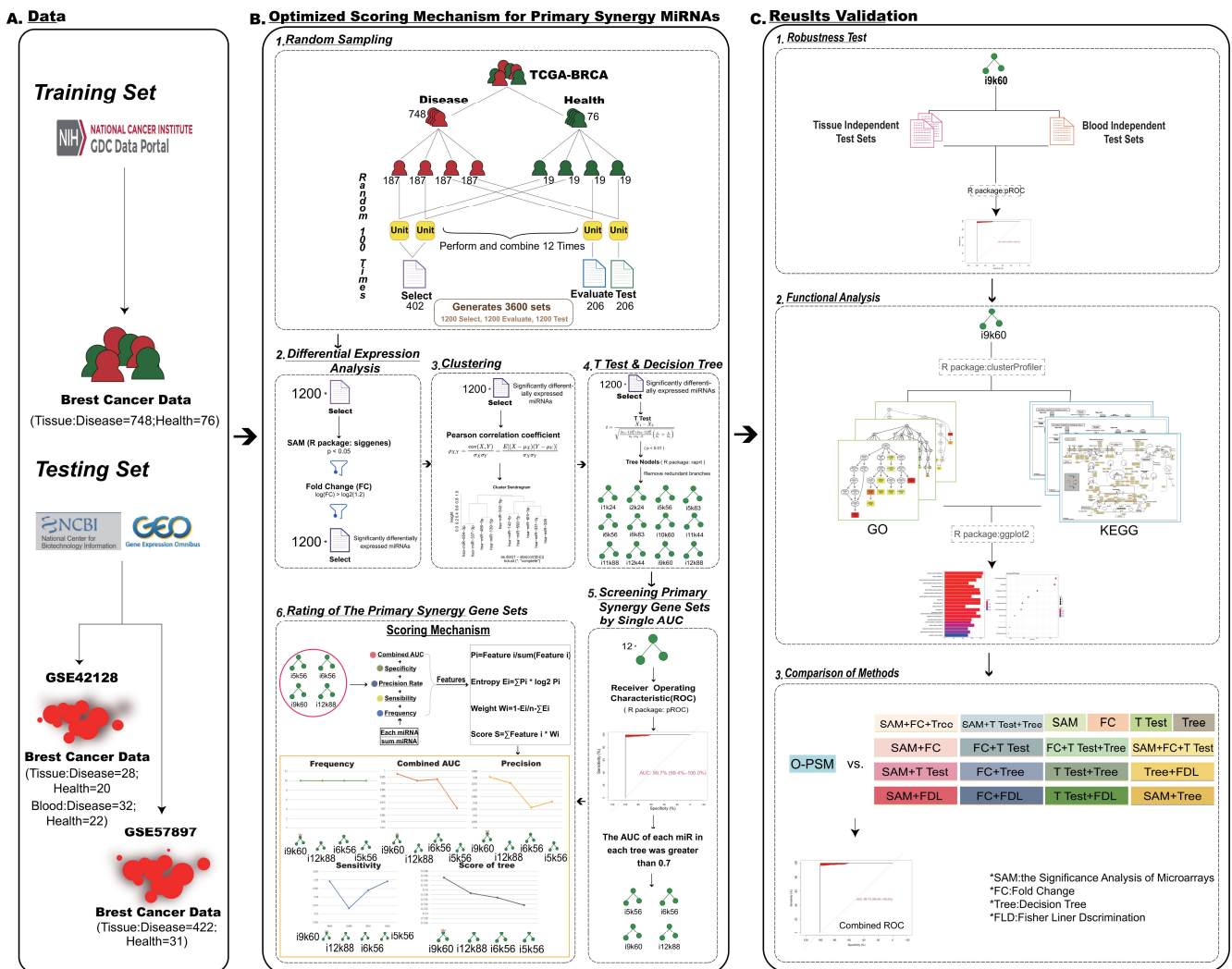
Figure 1. Flow Chart. (**A**) Data acquisition. (**B**) The set of microRNAs (miRNAs) with major synergistic effect was screened by the synthesis algorithm Optimized Scoring Mechanism for Primary Synergy MicroRNAs (O-PSM). 1. Four-fold cross sampling. 2. Differential expression analysis. 3. Clustering was performed using the pearson correlation coefficient. 4. The significant decision tree models were obtained by *t* test and decision tree. 5. Screening primary synergy gene sets by single Area Under Curve (AUC) of each miRNA > 0.7. 6. After rating the primary synergy gene sets, keep the highest scoring marker set. (**C**) Results validation.

2.5. The T Test and Decision Tree

The unpaired *t* test ($p < 0.01$) was performed on the clustering results, and the characteristic miRNAs from each set identified by this test were used for the decision trees (R package: rpart). To identify more significant sets, we also removed trees with redundant branches (Figure 1B).

2.6. Screening Primary Synergy miRNA Sets by Single AUCs

We used ROC analyses to select miRNA sets. The AUC of the trees were calculated using the R package, pROC. AUC value, accuracy and sensitivity were obtained from each set. To select more representative primary synergy miRNAs, the single AUC value of each miRNA in the feature evaluation set was greater than 0.7 during standard screening (Figure 1B).

2.7. Ratings of the Primary Synergy miRNA Sets

The combined AUC value, specificity, accuracy, and sensitivity of each miRNA in the remaining feature evaluation sets and the frequency of occurrence of each miRNA was combined to score the sets. The entropy weighting method was used to transform these indices into scores of the primary synergy miRNA sets (Figure 1B).

$$P_{xy} = m_{xy} \ / \ \left(\sum\nolimits_x m_{xy} \right) \quad (1)$$

x is for each feature set, y is the indices, $y \in (1, 2, 3, \dots, n)$. m_{xy} is set as the AUC value, specificity, accuracy, sensitivity and frequency of miRNA in the feature set.

$$E_y = - \sum\nolimits_x P_{xy} \cdot log\left(P_{xy}\right) \quad (2)$$

E_x is the entropy.

$$W_y = \left(1 - E_y\right) \ / \ \left(5 - \sum E_y\right) \quad (3)$$

W_x is the weight of entropy. Then, we assigned entropy weight to each set for subsequent scoring.

$$S_x = \sum\nolimits_y m_{xy} \cdot W_y \quad (4)$$

The scores (S_x) of each miRNA set were calculated and compared. The set with the highest score was identified as the one with significant synergistic effects.

We selected the decision tree method, not only because it showed the synergistic effect of miRNAs, but also because it provided us with the primary and secondary relationships between them. The root node was the node with the most important significance in the whole tree.

2.8. O-PSM for Primary Synergy miRNAs

By integrating the series of methods including random sampling, SAM, FC, Pearson's correlation analysis, *t* test, ROC and decision tree, we proposed the decision tree-based scoring algorithm, which involves the aforementioned steps (Figure 1A,B).

2.9. Comparison of Methods

We compared O-PSM with a range of methods, including SAM, FC, *t* test, decision tree and their combinations, for a total of 19 methods, including the use of Fisher Liner Discrimination and random forest to further determine the advantages of O-PSM (Supplementary Table S7). Additionally, we also evaluated the robustness of O-PSM by randomly permutating the data class labels. When the number of miRNAs screened by the compared methods exceeded 5, we randomly selected 5 miRNAs for comparison.

2.10. Functional Annotation of miRNAs

We obtained a total of 2752 target genes for 5 miRNAs from the miRTarBase database (https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/php/index.php (accessed on 26 March 2021)). Target genes that were significantly negatively correlated with markers were screened by Pearson's correlation coefficient. Using the negatively correlated target genes, we enriched Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for differentially high expressed miRNAs (hsa-miR-342-5p and hsa-miR-331-3p) and differentially low expressed miRNAs (hsa-miR-139-5p, hsa-miR-486-3p and hsa-miR-654-3p) using the R package "clusterProfiler", respectively, to obtain relevant functional annotations.

2.11. Construction of PPI Network and In-Depth Network Analysis

We obtained the target genes of these 5 miRNAs by querying miRTarBase (https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/php/index.php (accessed on 26 March 2021)), with 354 of them in the STRING database (https://string-db.org (accessed

on 28 March 2021)). We used these 354 target genes to build a PPI network and analyze their mechanisms. In order to obtain more information and make the PPI network clearer, we utilized Cytoscape software to visualize it.

Molecular Complex Detection (MCODE) in Cytoscape was used to explore hub modules of PPI network, with the threshold degree cut-off = 2, node score cut-off = 0.2, k-core = 2, and max depth = 100.

Finally, we obtained the functional annotation and the pathway enrichment of those target genes in Metascape (https://metascape.org (accessed on 28 March 2021)).

2.12. Survival Analysis

Using TCGA BRCA prognostic information, of which there were 741 samples, we first analyzed the effect of different clinical characteristics on patient survival by univariate cox regression. Then, we collected factors that were associated statistically significantly with the survival status of BRCA patients. We then combined remarkable factors in a multivariate cox regression analysis to obtain independent factors that could influent the prognosis of patients. Analysis was performed using the "survival" package and "survminer" package in R. Then, we plotted Kaplan-Meier (K-M) curve by R package "ggsurvplot".

3. Results

3.1. Identification of miRNAs with Primary Synergy Using O-PSM of Tissues

By performing O-PSM, 1200 sets were obtained, trees were generated in each feature selection set, and trees with redundant branches were subsequently removed. To determine the efficiency of discrimination, ROCs in the feature evaluation sets were calculated for the remaining 12 miRNA sets without redundant branches of the trees involving i1k24, i2k24, i5k56, i5k83, i6k56, i6k83, i9k60, i10k60, i11k44, i11k88, i12k44, and i12k88 (Figure 2A–L, Supplementary Table S1).

To identify the miRNA sets with high quality, the sets with each single AUC value greater than 0.7 were identified. The combined ROC, specifically, accuracy of each miRNA set, and frequency of occurrence of each miRNA in all miRNA sets were used to score using O-PSM. Finally, the highest ranked set was identified as i9k60, with a score of 0.7452.

There were 41 miRNAs in the raw data of i9k60. SAM selected 30 miRNAs with $p < 0.05$. After FC screening (|FC value| > 1.2), 10 miRNAs remained. Using clustering, eight co-expression sets were identified. All sets had t test p values less than 0.01. After pruning the decision tree, we had five miRNAs (hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p) that had major synergistic effects on BRCA. The hsa-miR-139-5p was the miRNA that had the highest single AUC (0.9941), and it was also the root node of the tree, which confirmed the importance of hsa-miR-139-5p.

3.2. Significant Categorization of Tissues

In the internal evaluated sets, the combined AUC of those miRNAs reached 0.9975, exceeding the AUC of five single miRNAs, which showed that the combined effect was better than the individual effects (Supplementary Table S2 and Figure 3A–F). Hsa-miR-139-5p with the highest frequency was also the miRNA with the highest individual AUC. It was also the root node of the decision tree (Figure 2N), highlighting its significance among the five miRNAs.

In the internal test sets, the results were the same as in the internal test set. The single AUC of hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p are 0.9904, 0.7084, 0.7974, 0.9294 and 0.8224, respectively, and the combined AUC is 0.9930 (Supplementary Table S3 and Supplementary Figure S1).

In independent validation set 1, the results were verified once (Supplementary Table S4 and Supplementary Figure S2), indicating the accurate classification performance of the miRNAs. The single AUC of hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p are 0.9714, 0.5625, 0.6446, 0.8982 and 0.9125, the combined AUC is 1.0000.
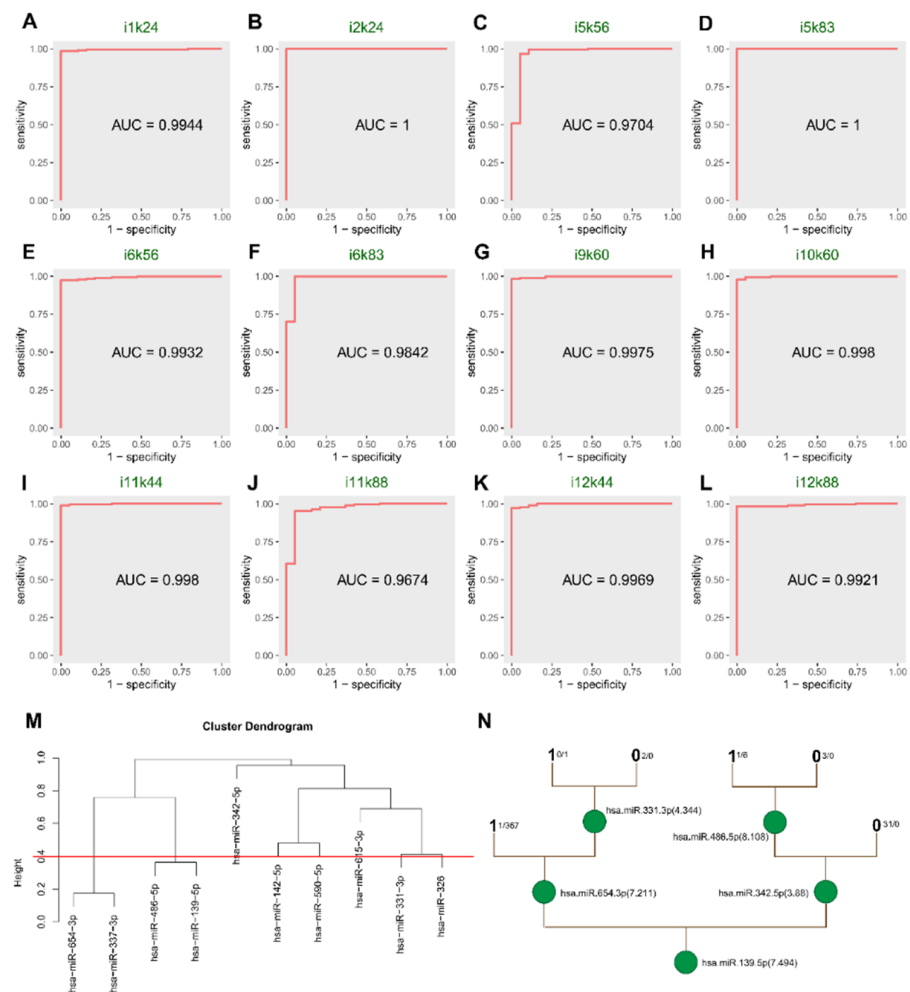
Figure 2. ROC of 12 trees in evaluated sets and cluster and tree plots of i9k60. (**A–L**) The ROC of 12 decision trees without redundant branches in the evaluated sets. (**M**) The 10 differentially expressed in i9k60 miRNAs were divided into 8 co-expression sets by Pearson's correlation coefficient analysis. (**N**) Unpaired *t* test ($p < 0.01$) was performed on the clustering results, the characteristic miRNAs in i9k60 screened by *t* test were used for the decision tree (R package: rpart). Through pruning the tree, we ended up with 5 miRNAs: hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p.

To further test the combined effects of these five miRNAs, we tested them in tissue independent set 2. The AUC results of hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p in GSE57897 showed that the combined effect reached 0.7181, exceeding each single AUC value, which showed the primary synergy of these five miRNAs again (Supplementary Table S5 and Supplementary Figure S3).

3.3. Significant Categorization of Serum

The five miRNAs that showed primary synergy also achieved satisfactory results in the independent test sets of serum (GSE42128 platform 2), with 0.9105 for the combined AUC, indicating that the joint use of five miRNAs also had significant classification performance in serum (Supplementary Table S6 and Figure 3G–L).

Both in tissues and blood (serum), the individual ROC was not as high as the combined value, indicating that the combined effect of these miRNAs was better than that of individual miRNAs. The results showed the reliable efficiency of O-PSM and synergy of these five miRNAs.
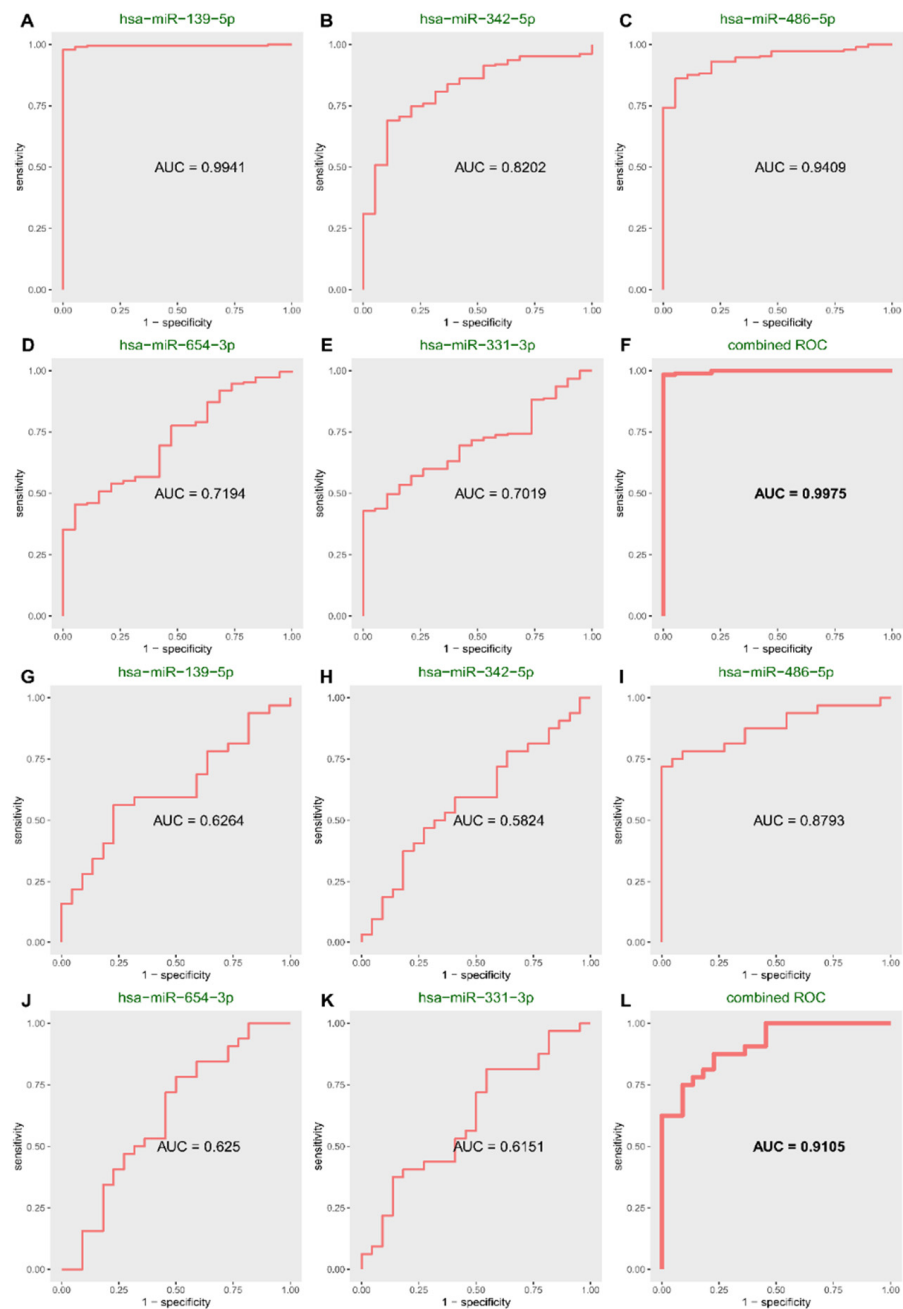
Figure 3. The single ROC of each miRNA and their combined ROC in the evaluated set and blood independent test set of i9k60. (**A–E**) Single ROC of each miRNA in evaluated set of i9k60. (**F**) Combined ROC of the 5 miRNAs in evaluated set of i9k60. (**G–K**) Single ROC of each miRNA of i9k60 in blood independent test set. (**L**) Combined ROC of the 5 miRNAs of i9k60 in blood independent test set.

3.4. Categorization Performed Remarkably Well in Typing and Staging of BRCA

For BRCA typing, GSE81002 was used to remove samples without subtype information. When samples of the subtype basal-like were cases, other subtypes were classified as controls. The same was true for other subtypes. The i9K60 miRNAs collection showed accurate categorization efficiency in basal-like BRCA and normal-like BRCA (Table 1).

Next, we evaluated the classification efficacy of i9k60 miRNAs for different stages of BRCA. Similar to typing, when stage 1 BRCA samples were used as case group, samples at other stages were used as control. The same was performed for other stages. The analysis revealed that i9k60 miRNAs had better classification efficacy when in stage 2. The effect

was not significant in other stages, and we speculated that this is due to the small amount of data (Table 2).

Table 1. The Area Under Curve (AUC) and Confidence Interval (CI) value of i9k60 in typing (GSE81002).

Subtype	Combined AUC	CI
Basal-like	0.840	78.94~89.15%
Normal-like	0.876	83.05~92.13%
Luminal A	0.680	62.66~73.39%
Luminal B	0.686	62.30~75.00%
HER-2 (+)	0.686	60.21~76.95%

Table 2. The AUC and CI value of i9k60 in Staging (GSE97811).

Stage	Combined AUC	CI
1	0.657	51.78~79.70%
2	0.723	59.08~85.48%
3	0.615	62.12~85.03%

3.5. Comparison with miRNAs That Had Been Experimentally Confirmed

We compared the selected miRNA collections with those that had been experimentally verified to be closely related to BRCA, to further determine the collection abilities of i9k60 miRNAs. The miRNAs that met the requirements were in previous reports, showing that hsa-miR-125a-5p and hsa-miR-146b-5p simultaneously existed in tissues and blood. ROC was used to compare our sets of data with these previously reported data. The results showed that the categorization efficiency of hsa-miR-125a-5p and hsa-miR-146b-5p differed greatly from the miRNAs screened by us (Figure 4). For the miRNA set screened by O-PSM, not only was the AUC value of these five miRNAs alone better than hsa-miR-125a-5p and hsa-miR-146b-5p, but also the combined ROCs of the internal data set and the three independent validation sets were 0.9975, 1.0000, 0.7181, and 0.9105, respectively, which were all higher than hsa-miR-125a-5p and hsa-miR-146b-5p. The novel marker miRNAs we screened not only achieved good classification results in several data sets, but they also obtained great results when compared with the two experimentally validated BRCA-associated miRNAs. Therefore, we believed that the miRNAs screened by the O-PSM method are promising as markers for BRCA identification and diagnosis.

3.6. Comparison of Methods

To test whether O-PSM had advantages over other methods, we compared it with a series of methods, including SAM, FC, *t* test, decision tree, and their combinations, 18 methods in total, including using Fisher Liner Discrimination to further identify the advantages of O-PSM (Supplementary Table S7). When the numbers of miRNAs screened by the comparison methods were more than five, we randomly selected five miRNAs for comparison. All the comparison results showed that there was a difference between these methods and O-PSM. The characteristic miRNAs screened by O-PSM had the highest combined AUC in TCGA-BRCA (0.9975), and the combined AUC in tissue independent validation set 1 tied with the AUC obtained by other partial methods and achieved better results in both tissue independent validation set 2 and blood independent validation set. Together, O-PSM showed good classification performance and could identify disease-related miRNAs with primary synergies. We also permuted the TCGA BRCA data to more accurately confirm the efficacy of these five miRNAs as clinical diagnostic markers for BRCA. However, the permuted data could not pass O-PSM, further validating the stringent screening conditions of O-PSM.
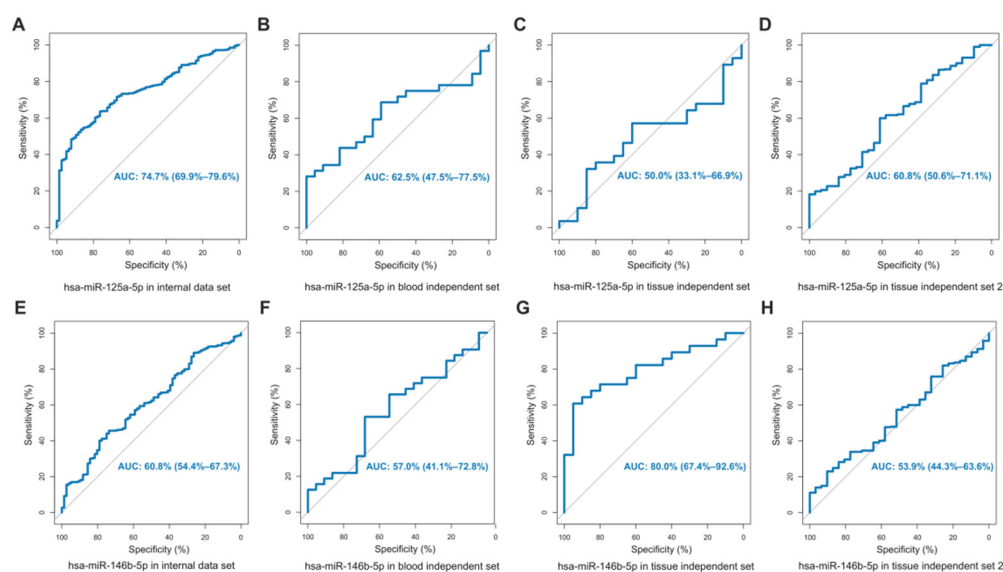
Figure 4. ROC of hsa-miR-125-5p and hsa-miR-146b-5p. (**A–D**) AUC for hsa-miR-125-5p was detected in TCGA-BRCA, blood independent set (GSE42128 platform2), tissue independent set 1 (GSE42128 platform3) and tissue independent set 2 (GSE57897). (**E–H**) AUC for hsa-miR-146b-5p was detected in TCGA-BRCA, blood independent set (GSE42128 platform2), tissue independent set 1 (GSE42128 platform3) and tissue independent Set 2 (GSE57897).

3.7. Functional Annotation

The development, progression and metastasis are highly complex processes that involve multiple biological functions and pathways. We queried the target genes of marker miRNAs in miRTarBase, screened the target genes negatively co-expressed with markers by Pearson's correlation coefficient, and annotated miRNAs functionally according to the significant negatively co-expressed target genes. Then, we performed GO and KEGG functional annotation by R package "clusterProfiler" for differentially high expressed miRNAs (hsa-miR-342-5p and hsa-miR-331-3p) and differentially low expressed miRNAs (hsa-miR-139-5p, hsa-miR-486-3p and hsa-miR-654-3p). We found that four out of five marker miRNAs participate in multiple biological procedures and pathways that may affect BRCA progression by regulating target genes (Figure 5).

Based on GO (Figure 5A,B) and KEGG (Figure 5C,D), we identified the upregulated miRNAs involved in the Rap1 signaling pathway, focal adhesion, Ras signaling pathway, MAPK signaling pathway and proteoglycans in cancer (Figure 5E), which were closely related to cell adhesion and other related functions. Adhesion is associated with cancer metastasis [54]. Cell migration is central to numerous physiological processes, including embryonic development, immune surveillance and wound healing, and dysregulation of migration is critical for cancer propagation. Among the target genes, *EGFR* is epidermal growth factor receptor [55–57], which is frequently expressed at high levels in different forms of cancer, and its expression is often positively correlated with cancer progression and poor prognosis [58–63]. *FLT4* is FMS-like tyrosine kinase 4 [64], *ID1* is a DNA-binding protein inhibitor [65], *MET* is proto-oncogene tyrosine protein kinase Met [66], and *PRKCA* is the classical protein kinase C alpha type [67–69], which are receptor tyrosine kinases (RTKs). RTKs are cell surface receptors with specific structural and biological characteristics that react to environmental clues by activating proper signaling cascades in cancer cells. *PGF* is placental growth factor [70], which belongs to growth factors (GF). In cancer cells, *PGF* mediates a range of pro-metastatic cellular events, including engaging endothelial cells to build blood supply, enhancing invasiveness and cell movement [71]. The upregulated markers reduce the expression of target genes, leading to dysregulation of processes such as signal transduction, cellular processes, and human diseases, and inhibit functions such as cell adhesion, which may promote tumor metastasis (Figure 5F).
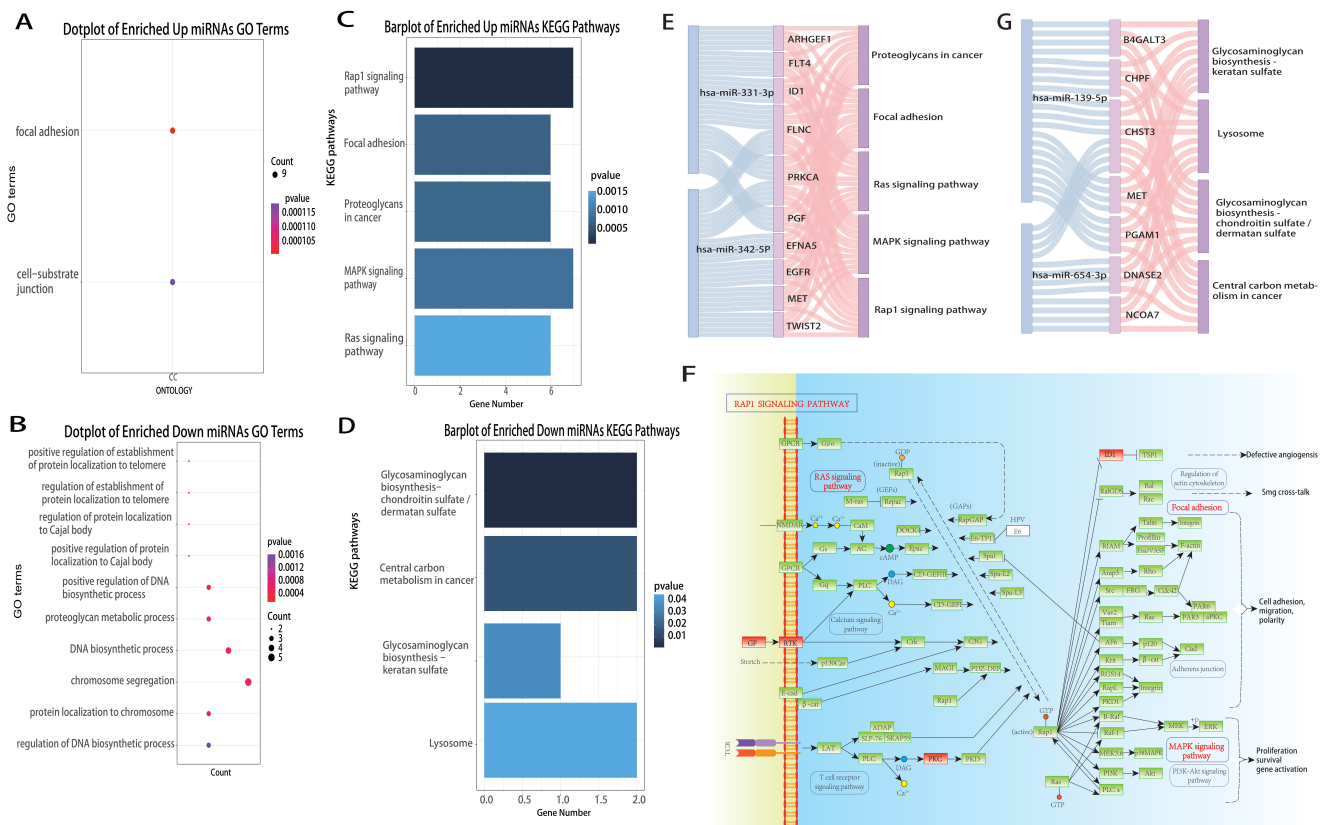
Figure 5. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations of hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, hsa-miR-654-3p. (**A**,**B**) Through R package "clusterProfiler", we obtained the GO annotation of these miRNAs. (**C**,**D**) Through R package "clusterProfiler", we obtained the KEGG enrichment analysis of them. (**E**) Sankey plot of upregulated miRNAs—target genes—pathways. (**F**) KEGG pathways plot of target genes of upregulated miRNAs. (**G**) Sankey plot of downregulated—target genes—pathways.

The target genes of downregulated miRNAs are mainly involved in glycosaminoglycan biosynthesis [72,73] and lysosome [74] and central carbon metabolism in cancer [75]; and these functions are tightly related to angiogenesis, cell growth, cell proliferation and apoptosis (Figure 5G). We assumed that marker miRNAs downregulate expression, reduce the repressive effect on target genes, enable the overactivation of related pathways and indirectly promote tumor development.

3.8. PPI Network and Mechanism Analysis of Target Genes

With the STRING database, we obtained the PPI network of target genes (Figure 6A), in the whole network, among them, the target genes that possessed the most interactions are targets of hsa-miR-342-5p, which have 65 relationships. The first three tightly connected hub interworking subnets are filtered by MCODE (Figure 6B–D).

Then, we obtained the functional annotations of all target genes of five marker miR-NAs in Metascape, and the results (Figure 6E) showed that they are highly related to the phosphatidylinositol phosphate biosynthetic process with the *p* value of 9.886×10^{-8}. Phosphatidylinositol 3-phosphate is a lipid that regulates membrane dynamics protein sorting and cell signaling [76]. They are also related to regulation of the purine nucleotide metabolic process ($p = 1.028 \times 10^{-6}$). Purines are basic components of nucleotides in cell proliferation; thus, impaired purine metabolism is associated with the progression of cancer, membrane trafficking ($p = 1.999 \times 10^{-6}$), which is a focal point for targeting cancer [77], negative regulation of the protein modification process ($p = 3.963 \times 10^{-6}$), vesicle organization ($p = 5.970 \times 10^{-6}$), PID MTOR 4PATHWAY ($p = 1.629 \times 10^{-5}$), neg-

ative regulation of apoptotic signaling pathway ($p = 1.897 \times 10^{-5}$), malignant pleural mesothelioma ($p = 2.421 \times 10^{-5}$), adipogenesis ($p = 2.489 \times 10^{-5}$) and brain development ($p = 2.630 \times 10^{-5}$).

Through the PPI network mining modules, we found that the first three modules all contain the target genes of hsa-miR-342-5p; thus, we think that the differential expression of hsa-miR-342-5p may have a more important effect on BRCA. Therefore, we analyzed the effect of hsa-miR-342-5p on the prognosis of BRCA patients.

First, we determined the factors that would affect the prognosis of BRCA patients by univariate cox regression. The analysis revealed that hsa-miR-342-5p as well as primary tumor stage (T stage), regional lymph nodes stage (N stage) and distant metastasis (M stage), which are also known as the TNM stage, were associated statistically significantly with the survival status of BRCA patients (Figure 6D). The effect of gender, race and radiation therapy on patient prognosis was not remarkable. We then combined hsa-miR-342-5p with the TNM stage in a multivariate cox regression analysis, and the results showed that hsa-miR-342-5p remarkably influenced patient survival despite the influence of TNM stage (Figure 6D). Thus, we concluded that hsa-miR-342-5p could influence the prognosis of BRCA patients significantly as an independent factor. By plotting the K-M curve, we could see that when hsa-miR-342-5p was highly expressed in patients, there was better prognosis of patients (Figure 6E).

We then analyzed the expression of target genes of hsa-miR-342-5p, which significantly enriches pathways, and could see that their target genes were remarkably low expressed in disease and relatively high expressed in normal samples, consistent with our findings from functional and survival analyses (Figure 6F).
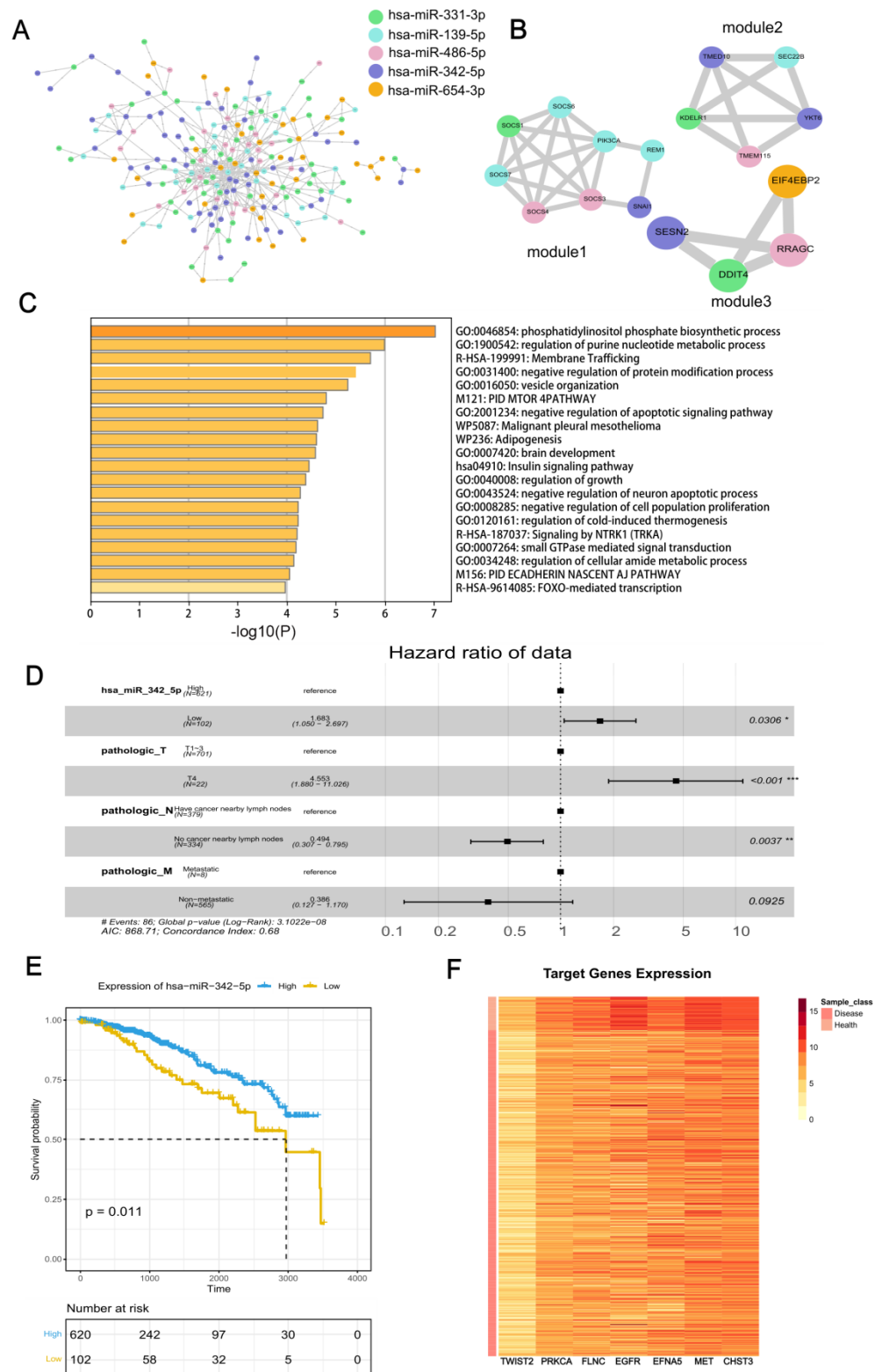
Figure 6. Protein-Protein Interaction (PPI) network and mechanism of the target genes and survival analysis of miRNAs. (**A**) PPI network of target genes obtained from STRING database. (**B**) The top three clusters filtered by MCODE. (**C**) Functional annotation of target genes by Metascape. (**D**) Multivariate cox regression analysis of factors that influence prognosis of BRCA patients. '*': $p < 0.05$, '**': $p < 0.01$, '***': $p < 0.001$. (**E**) Prognostic impact of high and low expression of hsa-miR-342-5p on TCGA BRCA patients. (**F**) Heat map of significantly differentially expressed hsa-miR-342-5p target genes in BRCA patients and normal samples.

4. Discussion

Identifying non-invasive markers of tumors has been a pressing issue. In urologic tumors, miRNAs have been shown to be useful biomarkers, but a large body of research data is not yet available for clinical practice [78]. However, it is now clear to us that different cancers can have very different clinical presentations in different patients. The use of biomarkers can be a very promising strategy. Epigenetics-based biomarkers such as dysregulated DNA methylations, deregulated expression of chromatin structure proteins and miRNAs or nt-RNAs or lncRNAs could have a high impact on clinical practice in oncology. Nevertheless, the transfer from laboratory to clinical practice remains slow. This is why translational research, with clinical implications, is the future of oncology research.

Avan et al. had compiled the potential value of tissue and circulating miRNAs for prognostic and therapeutic applications in BRCA in 2018 [79], and they concluded that miRNAs are promising for early detection of BRCA, predicting prognosis, and monitoring patient response to treatment based on preclinical and clinical investigations of tissue-specific miRNAs and circulating miRNAs. However, the performance of circulating miRNAs has been inconsistent and may lead to conflicting conclusions in different studies. Tissue-specific miRNAs are relatively stable, but biopsy can cause some damage to the body. We hope to mine BRCA markers in miRNAs common to tissue and blood, so that marker miRNAs can have the advantages of both: non-invasive and stable.

We proposed O-PSM to identify more efficient and accurate non-invasive disease-related miRNAs. Five miRNAs (hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p) were identified from BRCA tissues and blood samples of patients, which played a major role in this disease. These five miRNAs showed precise classification performance in BRCA tissues and serum, showed appropriate significance for the clinical diagnosis of BRCA, and strongly supported the screening ability of O-PSM. We placed the entire O-PSM process into an R package stored at https://github.com/SuFei-lab/OPSM.git (accessed on 22 October 2021). Verification in tissues and blood confirmed that the combined effect of these five miRNAs provided a novel biomarker system for diagnosis of BRCA. Furthermore, the study suggested the possibility of using non-invasive methodologies to achieve an accurate diagnostic and prognostic definition of BRCA cases.

We verified the categorization efficacy of screened miRNAs in terms of BRCA typing [50–52] and staging [53], indicating that the collection of miRNAs with major synergistic effects obtained by O-PSM was closely related to BRCA. We also compared with experimentally verified miRNAs (hsa-miR-125a-5p [45,46] and hsa-miR-146b-5p [45,47]) related to BRCA. By comparing the miRNAs screened by O-PSM with those reported in the literature, it was determined that miRNAs obtained by O-PSM had better categorization efficiencies. Thus, we concluded that the fresh markers screened by O-PSM has good robustness.

We conducted preliminary survival analysis of hsa-miR-342-5p, which showed that the marker is statistically significantly related with prognosis of BRCA patients ($p < 0.05$), and the higher the marker expression, the better the survival outcome.

In addition to bioinformatic analysis, we also searched the literature in the hope of finding the relationship between marker miRNAs and anti-cancer drugs and exploring the mechanisms by which markers act as potential drug targets to influence patient treatment outcomes and prognosis. The search revealed that hsa-miR-342-5p regulates the expression of genes involved in tamoxifen-mediated apoptosis and cell cycle progression in tumor cells. Restoration of hsa-miR-342-5p expression may represent a novel therapeutic approach to sensitize and inhibit the growth of tamoxifen-refractory breast tumors [80]. Aside from BRCA therapeutics, we also retrieved marker miRNAs that play a role in other cancer treatments. Deng et al. found that hsa-miR-342-5p may act as a tumor suppressor in osteosarcoma (OS) by targeting *Wnt7b* to inhibit the effects on OS cells viability, migration, invasion, sensitivity to Doxorubicin and apoptosis [81]. Tang et al. found that hsa-miR-139-5p increased apoptosis and inhibited cisplatin (DDP), induced non-small cell lung cancer (NSCLC) cell proliferation in vitro by regulating the PI3K/AKT/caspase-3 signaling pathway, and sensitized NSCLC cells to DDP by targeting *HOXB2*. Modulation

of hsa-miR-139-5p expression reversed DDP resistance and increased chemosensitivity of therapeutic NSCLC [82]. Furthermore, Fentanyl can inhibit the viability and invasion of NSCLC cells by inducing hsa-miR-331-3p and reducing *HDAC5* [83]. Hsa-miR-486-3p is an important mediator in regulating sorafenib resistance by targeting *FGFR4* and *EGFR*, thus providing a potential target for HCC treatment [84]. Allicin upregulates hsa-miR-486-3p and enhances the sensitivity of TMZ in glioblastoma. Allicin may be used as adjuvant chemotherapy for TMZ to improve patient prognosis, while hsa-miR-486-3p may be a potential target for glioblastoma treatment to improve outcomes [85]. Hsa-miR-654-3p enhances DDP sensitivity in Ovarian Cancer (OVC) cells by downregulating *QPRT* expression, and inhibition of hsa-miR-654-3p reverses the inhibitory effect of *QPRT*-targeted short interfering RNA on OVC cell proliferation and chemoresistance [86]. This shows that marker miRNAs participate in the treatment of diverse cancers, further demonstrating that our screened miRNAs may serve as potential drug targets for cancer treatment. We speculate that the drugs mentioned in our review may also have some implications for BRCA treatment.

Our study has a number of limitations: first, because TCGA BRCA data and GEO data (GSE42128) had little in common, it might have had an impact on the results. Nonetheless, because of the good screening capability of O-PSM, this method can also be used for other diseases to provide theoretical support for the development of better methods of diagnose, stratification and personalized treatment plans for patients. For example, it has recently been shown that tumor-specific immuno-profiling based on biomarkers can be used in bladder cancer patients treatable with immune checkpoints inhibitors [87]. Second, we also performed permutation on TCGA BRCA data to more accurately confirm the efficacies of these five miRNAs as BRCA clinical diagnosis markers. However, the data after random perturbation could not pass O-PSM, further validating the strict screening conditions of O-PSM. Third, we analyzed the prognostic impact of only one marker (hsa-miR-342-5p) and may have overlooked information about the prognostic impact of other markers on BRCA patients. Fourth, there are only a few studies focused on tissue and blood miRNAs in BRCA; thus, the research studies in comparison with previously reported biomarkers were limited, which proves that this research is of great significance; therefore, further study is necessary.

In the future, we hope that more scientists can further explore the miRNAs (hsa-miR-139-5p, hsa-miR-331-3p, hsa-miR-342-5p, hsa-miR-486-5p, and hsa-miR-654-3p) screened by O-PSM. Despite the limitations of our study, we believe that our results can be clinically significant in BRCA cases. Moreover, we hope to be able to extend O-PSM to more diseases and organs, to find effective and reliable biomarkers, and to contribute to advancements in the field of accurate non-invasive cancer diagnosis and prognostic stratification of cancer patients. The final goal is to be able to design tumor-specific personalized treatment plans for cancer patients in the next future.

References

1. De Smet, E.G.; Mestdagh, P.; Vandesompele, J.; Brusselle, G.G.; Bracke, K.R. Non-coding RNAs in the pathogenesis of COPD. *Thorax* **2015**, *70*, 782–791. [CrossRef] [PubMed]
2. Harbeck, N.; Gnant, M. Breast cancer. *Lancet* **2017**, *389*, 1134–1150. [CrossRef]
3. Fan, L.; Strasser-Weippl, K.; Li, J.J.; St Louis, J.; Finkelstein, D.M.; Yu, K.D.; Chen, W.Q.; Shao, Z.M.; Goss, P.E. Breast cancer in China. *Lancet Oncol.* **2014**, *15*, e279–e289. [CrossRef]
4. Brook, N.; Brook, E.; Dharmarajan, A.; Dass, C.R.; Chan, A. Breast cancer bone metastases: Pathogenesis and therapeutic targets. *Int. J. Biochem. Cell Biol.* **2018**, *96*, 63–78. [CrossRef] [PubMed]
5. Duijf, P.H.G.; Nanayakkara, D.; Nones, K.; Srihari, S.; Kalimutho, M.; Khanna, K.K. Mechanisms of genomic instability in breast cancer. *Trends Mol. Med.* **2019**, *25*, 595–611. [CrossRef] [PubMed]
6. Merino Bonilla, J.A.; Torres Tabanera, M.; Ros Mendoza, L.H. Breast cancer in the 21st century: From early detection to new therapies. *Cáncer Mama Siglo XXI Detección Precoz A Los Nuevos Tratamientos. Radiol.* **2017**, *59*, 368–379. [CrossRef]
7. Gomez, S.L.; Von Behren, J.; McKinley, M.; Clarke, C.A.; Shariff-Marco, S.; Cheng, I.; Reynolds, P.; Glaser, S.L. Breast cancer in Asian Americans in California, 1988–2013: Increasing incidence trends and recent data on breast cancer subtypes. *Breast Cancer Res. Treat.* **2017**, *164*, 139–147. [CrossRef]
8. Berry, D.A.; Cronin, K.A.; Plevritis, S.K.; Fryback, D.G.; Clarke, L.; Zelen, M.; Mandelblatt, J.S.; Yakovlev, A.Y.; Habbema, J.D.; Feuer, E.J. Effect of screening and adjuvant therapy on mortality from breast cancer. *N. Engl J. Med.* **2005**, *353*, 1784–1792. [CrossRef]
9. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: An independent review. *Lancet* **2012**, *380*, 1778–1786. [CrossRef]
10. Acunzo, M.; Romano, G.; Wernicke, D.; Croce, C.M. MicroRNA and cancer–a brief overview. *Adv. Biol. Regul.* **2015**, *57*, 1–9. [CrossRef]
11. Chen, C.; Tan, R.; Wong, L.; Fekete, R.; Halsey, J. Quantitation of microRNAs by real-time RT-qPCR. *Methods Mol. Biol.* **2011**, *687*, 113–134. [CrossRef] [PubMed]
12. Mellis, D.; Caporali, A. MicroRNA-based therapeutics in cardiovascular disease: Screening and delivery to the target. *Biochem. Soc. Trans.* **2018**, *46*, 11–21. [CrossRef] [PubMed]
13. Turchinovich, A.; Weiz, L.; Langheinz, A.; Burwinkel, B. Characterization of extracellular circulating microRNA. *Nucleic Acids Res.* **2011**, *39*, 7223–7233. [CrossRef] [PubMed]
14. Fricke, A.; Cimniak, A.F.V.; Ullrich, P.V.; Becherer, C.; Bickert, C.; Pfeifer, D.; Heinz, J.; Stark, G.B.; Bannasch, H.; Braig, D.; et al. Whole blood miRNA expression analysis reveals miR-3613-3p as a potential biomarker for dedifferentiated liposarcoma. *Cancer Biomark.* **2018**, *22*, 199–207. [CrossRef] [PubMed]
15. Lagos-Quintana, M.; Rauhut, R.; Yalcin, A.; Meyer, J.; Lendeckel, W.; Tuschl, T. Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* **2002**, *12*, 735–739. [CrossRef]
16. Osumi, H.; Shinozaki, E.; Yamaguchi, K.; Zembutsu, H. Clinical utility of circulating tumor DNA for colorectal cancer. *Cancer Sci.* **2019**, *110*, 1148–1155. [CrossRef] [PubMed]

17. Castillo, J.; Bernard, V.; San Lucas, F.A.; Allenson, K.; Capello, M.; Kim, D.U.; Gascoyne, P.; Mulu, F.C.; Stephens, B.M.; Huang, J.; et al. Surfaceome profiling enables isolation of cancer-specific exosomal cargo in liquid biopsies from pancreatic cancer patients. *Ann. Oncol.* **2018**, *29*, 223–229. [CrossRef] [PubMed]

18. Ko, N.Y.; Chen, L.R.; Chen, K.H. The role of micro RNA and long-non-coding RNA in osteoporosis. *Int. J. Mol. Sci.* **2020**, *21*, 4886. [CrossRef]

19. Huang, H.Y.; Lin, Y.C.; Li, J.; Huang, K.Y.; Shrestha, S.; Hong, H.C.; Tang, Y.; Chen, Y.G.; Jin, C.N.; Yu, Y.; et al. miRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* **2020**, *48*, D148–D154. [CrossRef]

20. Suzuki, H.I. MicroRNA control of TGF-β signaling. *Int. J. Mol. Sci.* **2018**, *19*, 1901. [CrossRef]

21. Zhang, Y.; Wang, Z.; Gemeinhart, R.A. Progress in microRNA delivery. *J. Control. Release* **2013**, *172*, 962–974. [CrossRef] [PubMed]

22. Humphries, B.; Wang, Z.; Yang, C. The role of microRNAs in metal carcinogen-induced cell malignant transformation and tumorigenesis. *Food Chem. Toxicol.* **2016**, *98*, 58–65. [CrossRef] [PubMed]

23. Kunz, M.; Göttlich, C.; Walles, T.; Nietzer, S.; Dandekar, G.; Dandekar, T. MicroRNA-21 versus microRNA-34: Lung cancer promoting and inhibitory microRNAs analysed in silico and in vitro and their clinical impact. *Tumour Biol.* **2017**, *39*, 1010428317706430. [CrossRef] [PubMed]

24. Xiao, H. MiR-7-5p suppresses tumor metastasis of non-small cell lung cancer by targeting NOVA2. *Cell Mol. Biol. Lett.* **2019**, *24*, 60. [CrossRef] [PubMed]

25. Bertoli, G.; Cava, C.; Castiglioni, I. MicroRNAs: New biomarkers for diagnosis, prognosis, therapy prediction and therapeutic tools for breast cancer. *Theranostics* **2015**, *5*, 1122–1143. [CrossRef]

26. O'Bryan, S.; Dong, S.; Mathis, J.M.; Alahari, S.K. The roles of oncogenic miRNAs and their therapeutic importance in breast cancer. *Eur. J. Cancer* **2017**, *72*, 1–11. [CrossRef]

27. Sauerbrei, W.; Taube, S.E.; McShane, L.M.; Cavenagh, M.M.; Altman, D.G. Reporting recommendations for tumor marker prognostic studies (REMARK): An abridged explanation and elaboration. *J. Natl. Cancer Inst.* **2018**, *110*, 803–811. [CrossRef]

28. Yang, Y.; Weng, W.; Peng, J.; Hong, L.; Yang, L.; Toiyama, Y.; Gao, R.; Liu, M.; Yin, M.; Pan, C.; et al. Fusobacterium nucleatum increases proliferation of colorectal cancer cells and tumor development in mice by activating toll-like receptor 4 signaling to nuclear factor-κB, and up-regulating expression of MicroRNA-21. *Gastroenterology* **2017**, *152*, 851–866.e24. [CrossRef]

29. Yu, T.; Ma, P.; Wu, D.; Shu, Y.; Gao, W. Functions and mechanisms of microRNA-31 in human cancers. *Biomed. Pharmacother.* **2018**, *108*, 1162–1169. [CrossRef]

30. Mulrane, L.; McGee, S.F.; Gallagher, W.M.; O'Connor, D.P. miRNA dysregulation in breast cancer. *Cancer Res.* **2013**, *73*, 6554–6562. [CrossRef]

31. Ma, X.; Guo, J.; Liu, K.; Chen, L.; Liu, D.; Dong, S.; Xia, J.; Long, Q.; Yue, Y.; Zhao, P.; et al. Identification of a distinct luminal subgroup diagnosing and stratifying early stage prostate cancer by tissue-based single-cell RNA sequencing. *Mol. Cancer* **2020**, *19*, 147. [CrossRef] [PubMed]

32. Gurel, B.; Ali, T.Z.; Montgomery, E.A.; Begum, S.; Hicks, J.; Goggins, M.; Eberhart, C.G.; Clark, D.P.; Bieberich, C.J.; Epstein, J.I.; et al. NKX3.1 as a marker of prostatic origin in metastatic tumors. *Am. J. Surg. Pathol.* **2010**, *34*, 1097–1105. [CrossRef] [PubMed]

33. Praharaj, P.P.; Bhutia, S.K.; Nagrath, S.; Bitting, R.L.; Deep, G. Circulating tumor cell-derived organoids: Current challenges and promises in medical research and precision medicine. *Biochim. Biophys. Acta Rev. Cancer* **2018**, *1869*, 117–127. [CrossRef] [PubMed]

34. Vaidyanathan, R.; Soon, R.H.; Zhang, P.; Jiang, K.; Lim, C.T. Cancer diagnosis: From tumor to liquid biopsy and beyond. *Lab Chip* **2018**, *19*, 11–34. [CrossRef] [PubMed]

35. Shan, X.; Zhang, L.; Zhu, D.X.; Zhou, X.; Zhang, H.; Liu, Q.X.; Tang, J.W.; Wen, W.; Wang, T.S.; Zhu, W.; et al. Serum microRNA expression profiling revealing potential diagnostic biomarkers for lung adenocarcinoma. *Chin. Med. J.* **2020**, *133*, 2532–2542. [CrossRef]

36. Zhang, Y.; Roth, J.A.; Yu, H.; Ye, Y.; Xie, K.; Zhao, H.; Chang, D.W.; Huang, M.; Li, H.; Qu, J.; et al. A 5-microRNA signature identified from serum microRNA profiling predicts survival in patients with advanced stage non-small cell lung cancer. *Carcinogenesis* **2019**, *40*, 643–650. [CrossRef]

37. Kondrotienė, A.; Daukša, A.; Pamedytytė, D.; Kazokaitė, M.; Žvirblienė, A.; Daukšienė, D.; Simanavičienė, V.; Klimaitė, R.; Golubickaitė, I.; Stakaitis, R.; et al. Plasma-derived miRNA-222 as a candidate marker for papillary thyroid cancer. *Int. J. Mol. Sci.* **2020**, *21*, 6445. [CrossRef]

38. Shan, X.; Zhang, H.; Zhang, L.; Zhou, X.; Wang, T.; Zhang, J.; Shu, Y.; Zhu, W.; Wen, W.; Liu, P. Identification of four plasma microRNAs as potential biomarkers in the diagnosis of male lung squamous cell carcinoma patients in China. *Cancer Med.* **2018**, *7*, 2370–2381. [CrossRef]

39. Dittmar, R.L.; Liu, S.; Tai, M.C.; Rajapakshe, K.; Huang, Y.; Longton, G.; DeCapite, C.; Hurd, M.W.; Paris, P.L.; Kirkwood, K.S.; et al. Plasma miRNA biomarkers in limited volume samples for detection of early-stage pancreatic cancer. *Cancer Prev. Res.* **2021**, *14*, 729–740. [CrossRef]

40. Lou, W.; Ding, B.; Xu, L.; Fan, W. Construction of potential glioblastoma multiforme-related miRNA-mRNA regulatory network. *Front. Mol. Neurosci.* **2019**, *12*, 66. [CrossRef]

41. Cheng, Y.; Su, Y.; Wang, S.; Liu, Y.; Jin, L.; Wan, Q.; Liu, Y.; Li, C.; Sang, X.; Yang, L.; et al. Identification of circRNA-lncRNA-miRNA-mRNA competitive endogenous RNA network as novel prognostic markers for acute myeloid leukemia. *Genes* **2020**, *11*, 868. [CrossRef] [PubMed]

42. Lu, Z.; Gao, Y. Screening differentially expressed genes between endometriosis and ovarian cancer to find new biomarkers for endometriosis. *Ann. Med.* **2021**, *53*, 1377–1389. [CrossRef] [PubMed]

43. Wang, F.; Xu, X.; Zhang, N.; Chen, Z. Identification and integrated analysis of hepatocellular carcinoma-related circular RNA signature. *Ann. Transl. Med.* **2020**, *8*, 294. [CrossRef] [PubMed]

44. Zhang, L.; Tao, H.; Li, J.; Zhang, E.; Liang, H.; Zhang, B. Comprehensive analysis of the competing endogenous circRNA-lncRNA-miRNA-mRNA network and identification of a novel potential biomarker for hepatocellular carcinoma. *Aging* **2021**, *13*, 15990–16008. [CrossRef] [PubMed]

45. Elango, R.; Alsaleh, K.A.; Vishnubalaji, R.; Manikandan, M.; Ali, A.M.; Abd El-Aziz, N.; Altheyab, A.; Al-Rikabi, A.; Alfayez, M.; Aldahmash, A.; et al. MicroRNA expression profiling on paired primary and lymph node metastatic breast cancer revealed distinct microRNA profile associated With LNM. *Front. Oncol.* **2020**, *10*, 756. [CrossRef] [PubMed]

46. Fu, S.W.; Chen, L.; Man, Y.G. miRNA biomarkers in breast cancer detection and management. *J. Cancer* **2011**, *2*, 116–122. [CrossRef]

47. Li, M.; Zou, X.; Xia, T.; Wang, T.; Liu, P.; Zhou, X.; Wang, S.; Zhu, W. A five-miRNA panel in plasma was identified for breast cancer diagnosis. *Cancer Med.* **2019**, *8*, 7006–7017. [CrossRef]

48. Chan, M.; Liaw, C.S.; Ji, S.M.; Tan, H.H.; Wong, C.Y.; Thike, A.A.; Tan, P.H.; Ho, G.H.; Lee, A.S. Identification of circulating microRNA signatures for breast cancer detection. *Clin. Cancer Res.* **2013**, *19*, 4477–4487. [CrossRef]

49. Zhao, X.G.; Hu, J.Y.; Tang, J.; Yi, W.; Zhang, M.Y.; Deng, R.; Mai, S.J.; Weng, N.Q.; Wang, R.Q.; Liu, J.; et al. miR-665 expression predicts poor survival and promotes tumor metastasis by targeting NR4A3 in breast cancer. *Cell Death Dis.* **2019**, *10*, 479. [CrossRef]

50. Aure, M.R.; Vitelli, V.; Jernström, S.; Kumar, S.; Krohn, M.; Due, E.U.; Haukaas, T.H.; Leivonen, S.K.; Vollan, H.K.; Lüders, T.; et al. Integrative clustering reveals a novel split in the luminal A subtype of breast cancer with impact on outcome. *Breast Cancer Res.* **2017**, *19*, 44. [CrossRef]

51. Knutsen, E.; Lellahi, S.M.; Aure, M.R.; Nord, S.; Fismen, S.; Larsen, K.B.; Gabriel, M.T.; Hedberg, A.; Bjørklund, S.S.; Bofin, A.M.; et al. The expression of the long NEAT1_2 isoform is associated with human epidermal growth factor receptor 2-positive breast cancers. *Sci. Rep.* **2020**, *10*, 1277. [CrossRef] [PubMed]

52. Pladsen, A.V.; Nilsen, G.; Rueda, O.M.; Aure, M.R.; Borgan, Ø.; Liestøl, K.; Vitelli, V.; Frigessi, A.; Langerød, A.; Mathelier, A.; et al. DNA copy number motifs are strong and independent predictors of survival in breast cancer. *Commun. Biol.* **2020**, *3*, 153. [CrossRef] [PubMed]

53. Hironaka-Mitsuhashi, A.; Matsuzaki, J.; Takahashi, R.U.; Yoshida, M.; Nezu, Y.; Yamamoto, Y.; Shiino, S.; Kinoshita, T.; Ushijima, T.; Hiraoka, N.; et al. A tissue microRNA signature that predicts the prognosis of breast cancer in young women. *PLoS ONE* **2017**, *12*, e0187638. [CrossRef] [PubMed]

54. Shen, J.; Cao, B.; Wang, Y.; Ma, C.; Zeng, Z.; Liu, L.; Li, X.; Tao, D.; Gong, J.; Xie, D. Hippo component YAP promotes focal adhesion and tumour aggressiveness via transcriptionally activating THBS1/FAK signalling in breast cancer. *J. Exp. Clin. Cancer Res.* **2018**, *37*, 175. [CrossRef] [PubMed]

55. Shao, H.; Cheng, H.Y.; Cook, R.G.; Tweardy, D.J. Identification and characterization of signal transducer and activator of transcription 3 recruitment sites within the epidermal growth factor receptor. *Cancer Res.* **2003**, *63*, 3923–3930.

56. Habib, A.A.; Chatterjee, S.; Park, S.K.; Ratan, R.R.; Lefebvre, S.; Vartanian, T. The epidermal growth factor receptor engages receptor interacting protein and nuclear factor-kappa B (NF-kappa B)-inducing kinase to activate NF-kappa B. Identification of a novel receptor-tyrosine kinase signalosome. *J. Biol. Chem.* **2001**, *276*, 8865–8874. [CrossRef]

57. Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Yarden, Y.; Libermann, T.A.; Schlessinger, J.; et al. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* **1984**, *309*, 418–425. [CrossRef]

58. Maemondo, M.; Inoue, A.; Kobayashi, K.; Sugawara, S.; Oizumi, S.; Isobe, H.; Gemma, A.; Harada, M.; Yoshizawa, H.; Kinoshita, I.; et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N. Engl. J. Med.* **2010**, *362*, 2380–2388. [CrossRef]

59. Talukdar, S.; Emdad, L.; Das, S.K.; Fisher, P.B. EGFR: An essential receptor tyrosine kinase-regulator of cancer stem cells. *Adv. Cancer Res.* **2020**, *147*, 161–188. [CrossRef]

60. Zhang, Z.; Stiegler, A.L.; Boggon, T.J.; Kobayashi, S.; Halmos, B. EGFR-mutated lung cancer: A paradigm of molecular oncology. *Oncotarget* **2010**, *1*, 497–514. [CrossRef]

61. Riely, G.J.; Pao, W.; Pham, D.; Li, A.R.; Rizvi, N.; Venkatraman, E.S.; Zakowski, M.F.; Kris, M.G.; Ladanyi, M.; Miller, V.A. Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin. Cancer Res.* **2006**, *12*, 839–844. [CrossRef] [PubMed]

62. Shigematsu, H.; Lin, L.; Takahashi, T.; Nomura, M.; Suzuki, M.; Wistuba, I.I.; Fong, K.M.; Lee, H.; Toyooka, S.; Shimizu, N.; et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J. Natl. Cancer Inst.* **2005**, *97*, 339–346. [CrossRef] [PubMed]

63. Pao, W.; Chmielecki, J. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat. Rev. Cancer* **2010**, *10*, 760–774. [CrossRef] [PubMed]

64. Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D. Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene. *Genomics* **1992**, *13*, 475–478. [CrossRef]

65. Ruzinova, M.B.; Benezra, R. Id proteins in development, cell cycle and cancer. *Trends Cell Biol.* **2003**, *13*, 410–418. [CrossRef]
66. Park, M.; Dean, M.; Kaul, K.; Braun, M.J.; Gonda, M.A.; Vande Woude, G. Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6379–6383. [CrossRef]
67. Finkenzeller, G.; Marmé, D.; Hug, H. Sequence of human protein kinase C α. *Nucleic Acids Res.* **1990**, *18*, 2183. [CrossRef]
68. Kolch, W.; Heidecker, G.; Kochs, G.; Hummel, R.; Vahidi, H.; Mischak, H.; Finkenzeller, G.; Marmé, D.; Rapp, U.R. Protein kinase C α activates RAF-1 by direct phosphorylation. *Nature* **1993**, *364*, 249–252. [CrossRef]
69. Darieva, Z.; Han, N.; Warwood, S.; Doris, K.S.; Morgan, B.A.; Sharrocks, A.D. Protein kinase C regulates late cell cycle-dependent gene expression. *Mol. Cell Biol.* **2012**, *32*, 4651–4661. [CrossRef]
70. Maglione, D.; Guerriero, V.; Viglietto, G.; Delli-Bovi, P.; Persico, M.G. Isolation of a human placenta cDNA coding for a protein related to the vascular permeability factor. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9267–9271. [CrossRef]
71. Taylor, A.P.; Rodriguez, M.; Adams, K.; Goldenberg, D.M.; Blumenthal, R.D. Altered tumor vessel maturation and proliferation in placenta growth factor-producing tumors: Potential relationship to post-therapy tumor angiogenesis and recurrence. *Int. J. Cancer* **2003**, *105*, 158–164. [CrossRef] [PubMed]
72. Kim, B.T.; Kitagawa, H.; Tamura, J.; Saito, T.; Kusche-Gullberg, M.; Lindahl, U.; Sugahara, K. Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode α 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/heparin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7176–7181. [CrossRef] [PubMed]
73. Funderburgh, J.L. Keratan sulfate: Structure, biosynthesis, and function. *Glycobiology* **2000**, *10*, 951–958. [CrossRef] [PubMed]
74. Eskelinen, E.L.; Tanaka, Y.; Saftig, P. At the acidic edge: Emerging functions for lysosomal membrane proteins. *Trends Cell Biol.* **2003**, *13*, 137–145. [CrossRef]
75. Soga, T. Cancer metabolism: Key players in metabolic reprogramming. *Cancer Sci.* **2013**, *104*, 275–281. [CrossRef]
76. Schink, K.O.; Raiborg, C.; Stenmark, H. Phosphatidylinositol 3-phosphate, a lipid that regulates membrane dynamics, protein sorting and cell signalling. *Bioessays* **2013**, *35*, 900–912. [CrossRef]
77. Jin, H.; Tang, Y.; Yang, L.; Peng, X.; Li, B.; Fan, Q.; Wei, S.; Yang, S.; Li, X.; Wu, B.; et al. Rab GTPases: Central coordinators of membrane trafficking in cancer. *Front. Cell Dev. Biol.* **2021**, *9*, 648384. [CrossRef]
78. Mancini, M.; Zazzara, M.; Zattoni, F. Stem cells, biomarkers and genetic profiling: Approaching future challenges in urology. *Urologia* **2016**, *83*, 4–13. [CrossRef]
79. Bahrami, A.; Aledavood, A.; Anvari, K.; Hassanian, S.M.; Maftouh, M.; Yaghobzade, A.; Salarzaee, O.; ShahidSales, S.; Avan, A. The prognostic and therapeutic application of microRNAs in breast cancer: Tissue and circulating microRNAs. *J. Cell Physiol.* **2018**, *233*, 774–786. [CrossRef]
80. Cittelly, D.M.; Das, P.M.; Spoelstra, N.S.; Edgerton, S.M.; Richer, J.K.; Thor, A.D.; Jones, F.E. Downregulation of miR-342 is associated with tamoxifen resistant breast tumors. *Mol. Cancer* **2010**, *9*, 317. [CrossRef]
81. Liu, Q.; Wang, Z.; Zhou, X.; Tang, M.; Tan, W.; Sun, T.; Deng, Y. miR-342-5p inhibits osteosarcoma cell growth, migration, invasion, and sensitivity to doxorubicin through targeting Wnt7b. *Cell Cycle* **2019**, *18*, 3325–3336. [CrossRef] [PubMed]
82. Du, H.; Bao, Y.; Liu, C.; Zhong, A.; Niu, Y.; Tang, X. miR-139-5p enhances cisplatin sensitivity in non-small cell lung cancer cells by inhibiting cell proliferation and promoting apoptosis via the targeting of Homeobox protein *HOX*-B2. *Mol. Med. Rep.* **2021**, *23*, 104. [CrossRef] [PubMed]
83. Gong, S.; Ying, L.; Fan, Y.; Sun, Z. Fentanyl inhibits lung cancer viability and invasion via upregulation of miR-331-3p and repression of HDAC5. *Onco Targets Ther.* **2020**, *13*, 13131–13141. [CrossRef] [PubMed]
84. Ji, L.; Lin, Z.; Wan, Z.; Xia, S.; Jiang, S.; Cen, D.; Cai, L.; Xu, J.; Cai, X. miR-486-3p mediates hepatocellular carcinoma sorafenib resistance by targeting FGFR4 and EGFR. *Cell Death Dis.* **2020**, *11*, 250. [CrossRef]
85. Wu, H.; Li, X.; Zhang, T.; Zhang, G.; Chen, J.; Chen, L.; He, M.; Hao, B.; Wang, C. Overexpression miR-486-3p promoted by allicin enhances temozolomide sensitivity in glioblastoma via targeting MGMT. *Neuromol. Med.* **2020**, *22*, 359–369. [CrossRef]
86. Niu, Y.C.; Tong, J.; Shi, X.F.; Zhang, T. MicroRNA-654-3p enhances cisplatin sensitivity by targeting QPRT and inhibiting the PI3K/AKT signaling pathway in ovarian cancer cells. *Exp. Ther. Med.* **2020**, *20*, 1467–1479. [CrossRef]
87. Mancini, M.; Righetto, M.; Noessner, E. Checkpoint inhibition in bladder cancer: Clinical expectations, current evidence, and proposal of future strategies based on a tumor-specific immunobiological approach. *Cancers* **2021**, *13*, 6016. [CrossRef]