EBioMedicine
Published by THE LANCET

Research paper

miRNA arm switching identifies novel tumour biomarkers

Liang Chen [a], Huiyan Sun [b], Changliang Wang [a], Yang Yang [a], Menglei Zhang [a], Garry Wong [a,*]

[a] Cancer Centre, Centre of Reproduction, Development and Aging, Faculty of Health Sciences, University of Macau, Taipa, Macau S.A.R., China
[b] Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun 130012, China

ARTICLE INFO

ABSTRACT

Background: microRNAs have been reported to play critical roles in cancer and to have potential as diagnostic biomarkers. During miRNA biogenesis, one strand of the miRNA hairpin precursor is preferentially selected as a functionally mature miRNA, while the other strand is typically degraded. Arm switching occurs when the strand preference is changed. This preference can be different and can change dynamically depending upon the species, tissue types, or development stages. Due to recent advances in next-generation sequencing methods, arm switching has been observed in a variety of cancers.
Methods: A tumour miRNA-Seq dataset was collected from The Cancer Genome Atlas (TCGA). The support vector machine (SVM) method combined with 5-fold cross validation was applied to select the best combination of arm-switched miRNA tumour markers. Survival analysis was also applied to identify patient survival associated miRNA markers.
Findings: We observed 51 arm-switched miRNAs and of these, 7 were associated with patient survival. Twenty-three 1-combination arm switching miRNAs with excellent diagnostic value were identified. Interestingly, ovarian cancer showed a significant difference in arm switching pattern compared with 32 other cancers.
Interpretation: These results suggest that arm switching miRNAs could be used as potential biomarkers for various cancers.
Fund: This work was partially supported by the National Natural Science Foundation of China (no. 61472158, 61572227), and University of Macau Faculty of Health Sciences (MYRG2016-00101-FHS).

1. Introduction

miRNAs are single-stranded endogenous small non-coding RNAs that comprise 1–3% of the human genome [1]. miRNAs have been found to play a crucial role in maintaining and dynamically altering the levels of mRNAs. One of the well-known functions of miRNA is the post-transcriptional regulation of gene expression [2]. miRNAs are initially transcribed as primary transcripts (pri-miRNAs) by RNA polymerase II (Pol II) and are subsequently cleaved to 65–70 nucleotide precursors (pre-miRNAs) by RNA endonuclease Drosha [3,4]. pre-miRNAs occur as thermodynamically stable hairpin structures that are exported to the cytoplasm via Exportin-5 and processed to a short double-stranded RNA duplex by RNA endonuclease Dicer. miRNA duplexes are then loaded into particular types of AGO proteins. Following AGO loading, one strand is degraded [2,3]. The other strand preferentially remains and forms a mature RNA-induced silencing complex (RISC) that binds to the complementary mRNA and cleaves the

transcript products. This action prevents the translation of mRNAs into proteins [5,6]. Canonical miRNA biogenesis is shown in Fig. 1A.

In the short double-stranded RNA duplex, the strand with presumable biological activity is named the guide strand, while the discarded strand is termed the passenger strand (or miRNA star) [7]. Fig. 1B shows the models of AGO sorting and strand selection. The central mismatching and bulge on the miRNA duplex can promote sorting of the duplex to the appropriate type of AGO protein in *Drosophila* [7]. In contrast, the four human AGO proteins (AGO1–4) are associated with miRNA duplexes without any clear preference [4]. Strand selection is not random, and some selection factors have been revealed. Duplex end stability is one determinant factor, where the strand whose 5′ end has less structure is favoured to be chosen as the guide strand [8,9]. The 5′ terminal nucleotide effects the choice of guide strand and different AGO proteins prefer different nucleotides. For example, AGO1 tends to select the strand with U at the 5′ end as a guide strand [9–11]; while the pocket within AGO2 prefers a pA or pU 5′ terminus [12].

The asymmetry of miRNA strand selection does not necessarily imply that the passenger strand is non-functional. Many studies have demonstrated that the passenger strand can also constitute a mature RISC and interact with target genes [13–15]. The 5′ and 3′ side of the

* Corresponding author.
  E-mail address: GarryGWong@umac.mo (G. Wong).

pre-miRNA hairpin structure define the two arms of miRNA. The miR-5p and miR-3p represent 5′ mature miRNA and 3′ mature miRNA, respectively, as shown in Fig. 1C. In different tissues, developmental stages and species, the biological system may select different arms as the dominant expressed one [16–18]. The phenomenon of alternative strand selection has been termed 'arm switching' [19]. The molecular mechanisms that underlie miRNA strand selection and that lead to arm switching are not entirely clear. Previous studies have suggested that Dicer cleavage governs arm switching [20–22]. Alternative Drosha processing which changes the thermodynamic stability of the miRNA duplex ends can also explain an arm switching event [23]. Both arms could be co-accumulated as miRNA pairs in some tissues while being subjected to arm selection in other tissues [18]. Arm switching can frequently occur (~11%) in insects between orthologous miRNAs [24]. Other studies have shown that two arms from the same precursor tend to have different targeting properties [25]. In brief, arm switching provides a basic mechanism to evolve the function of miRNA and may be a fundamental mechanism of miRNA diversification [14,19,26–28].

The importance of miRNAs in cancer has been demonstrated via direct interaction with oncogenes or tumour suppressor genes [29]. Differentially expressed miRNAs have been proposed as biomarkers for cancers [30,31]. Circulating miRNA could also be a potential biomarker for cancer diagnosis and prognosis [32]. Moreover, some cancer-related studies have integrated miR-5p and miR-3p together [33–35]. For example,

miR-193a-5p and miR-193a-3p were found significantly increased in breast cancer and revealed that strand expression preference may be a means of modulating miRNA function [36]. miR-5p and miR-3p differential expression between tumour tissue and tumour-adjacent tissue has been reported in gastric cancer [37]. miR-5p and miR-3p have been found that co-express and cross-target in colon cancer cells [38]. Moreover, dysregulation of miR-5p/miR-3p specific miRNAs were found in gastric cancer in a TCGA data set [39]. Furthermore, a study of specific differences in the function of an arm-switched miRNA in a lung cancer has been performed [40]. Recently, a database for small RNA sequencing in human cancer research can visualize expression of arm switching miRNAs has been published [41]. Although the mechanism of arm switching is not fully understood, some studies have linked arm switching to cancers. Nonetheless, there are still no miRNA biomarkers based on arm switching to our knowledge. In this study, we investigated the arm switching pattern in different tumours and demonstrated that some pre-miRNAs tend to select different arms in tumour and tumour-adjacent tissues. We also correlated arm switching events in pan-cancers with patient survival. Our results support the notion that miRNA arm switching could be a novel biomarker for different cancers.

2. Materials and methods

In order to discover novel tumour biomarkers based on miRNA arm switching, we implemented a workflow as shown in Fig. 2. The overall workflow contains four parts: arm switching miRNA identification, enrichment analysis, survival analysis, and biomarker discovery. The arm switching miRNA identification module marks strand selection preference between tumour and tumour-adjacent samples for each pre-miRNA initially and then identifies arm-switched miRNAs. The function analysis module annotates different arms of pre-miRNA and reveals the function for each arm for each biomarker candidate. The survival analysis module is used to identify arm-switched miRNAs which are associated with patient survival. The biomarker discovery module applies SVM to find combination biomarkers of arm-switched miRNAs.

2.1. Data sources

Thirty three cancer types including 11158 datasets were collected from the public database TCGA (http://cancergenome.nih.gov/). A summary table of the data sets that includes the cancer types are shown in Supplementary Table 1. The sample types we collected were tumour tissue and tumour-adjacent tissue (Normal tissue). Processed miRNA-Seq data were directly extracted from tab-delimited (.txt) data files in level 3. The survival information was obtained from the clinical files and incomplete samples were excluded from the survival analysis. Basic information of miRNAs was extracted from miRBase (Release 21, http://www.mirbase.org/) [42]. The annotation of TCGA miRNA-Seq level 3 data is based on miRBase v16, so there are only 1046 pre-miRNAs of which 318 pre-miRNAs are annotated with two arms. The experiment validated targets were extracted from miRTarBase 7.0 [43] and TarBase 8.0 [44].

2.2. Arm switching miRNA identification

To evaluate the arm switching patterns, we calculated a miR-5p ratio for each pre-miRNA in each sample originally developed by Hu et al. [45]. We defined an average miR-5p ratio arm switch criteria for multi-samples in this study as:

$$\text{Average Ratio}_{miR\text{-}5p} = \frac{1}{N} \sum_{i=1}^{N} \frac{miR\text{-}5p_i}{miR\text{-}5p_i + miR\text{-}3p_i}$$

miR-5p (or miR-3p) represents the expression RPM value of miR-5p (or miR-3p). N is the number of samples in each tumour. We determine the dominant arm by averaging the miR-5p ratio. The ratio was adopted
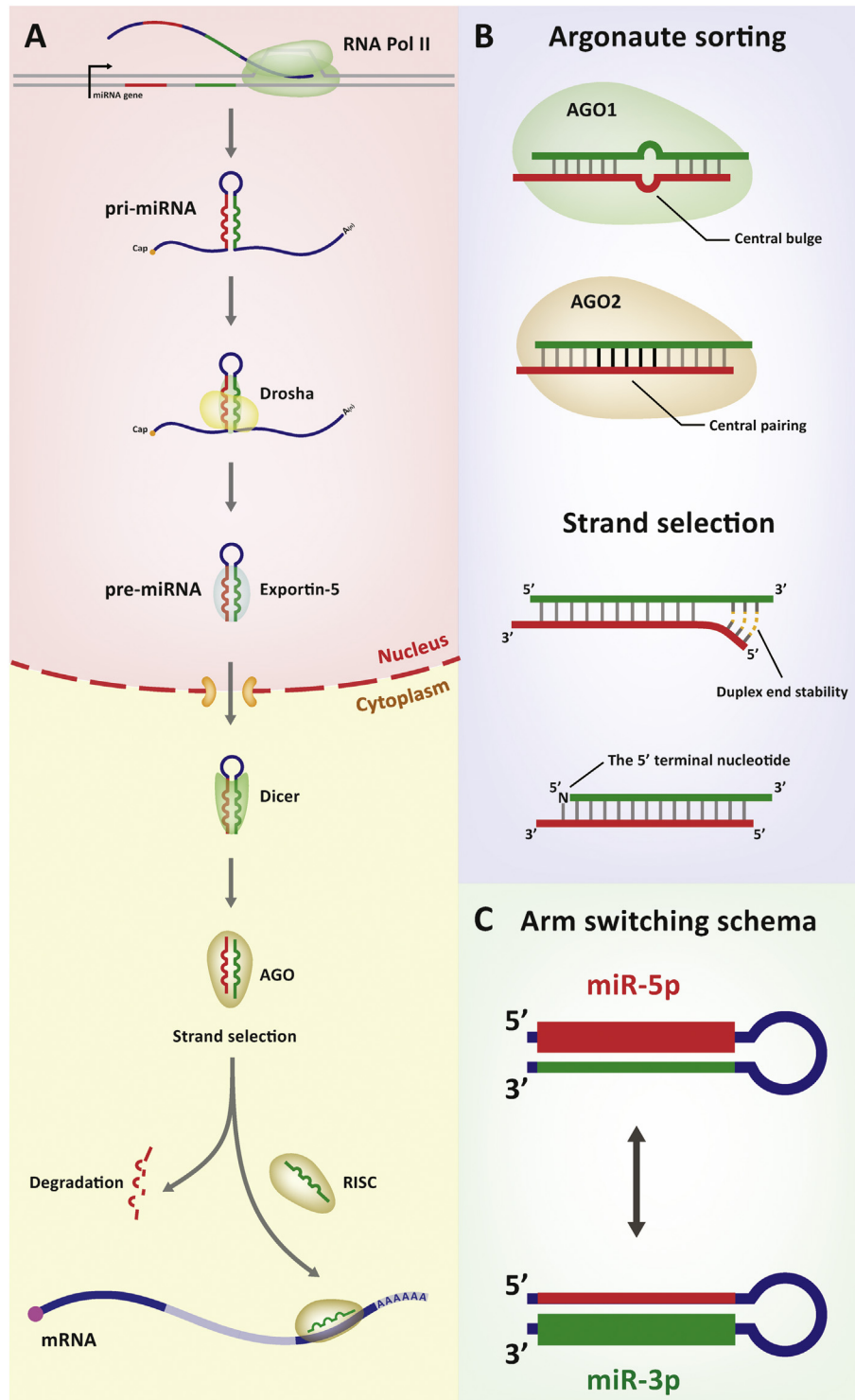
Fig. 1. Canonical miRNA biogenesis and arm switching schema. (A) Biogenesis of miRNA is shown from transcription of the primary miRNA to targeting of mRNA via the RISC complex. (B) Argonaute (AGO) sorting and selection of the two strands of miRNA duplex are shown. The two arms are labelled with red and green, respectively. (C) Arm switching schema is shown. The blue line represents the pre-miRNA also known as the hairpin. The red and green lines represent the locations of the miR-5p and miR-3p arms, respectively. The thickness of the red or green line indicates the expression level.

as a measure of the arm selection, where $1 \geq$ ratio > 0.5 indicates miR-5p dominant expression, and $0.5 >$ ratio ≥ 0 indicates miR-3p dominant expression.

In normal or tumour samples, we labelled each pre-miRNA by the above calculation. We then defined an arm switching event between normal and tumour samples by finding a transformation between miR-5p and miR-3p dominant expression.

2.3. Enrichment analysis

The experimentally validated targets were extracted from miRTarBase 7.0 [43] and TarBase 8.0 [44]. Bioconductor R package clusterProfiler [46] and GOSemSim [47] were applied to perform enrichment analysis and to measure the similarity of GO terms, respectively. We chose the Wang method to measure the semantic similarity of GO terms [48].
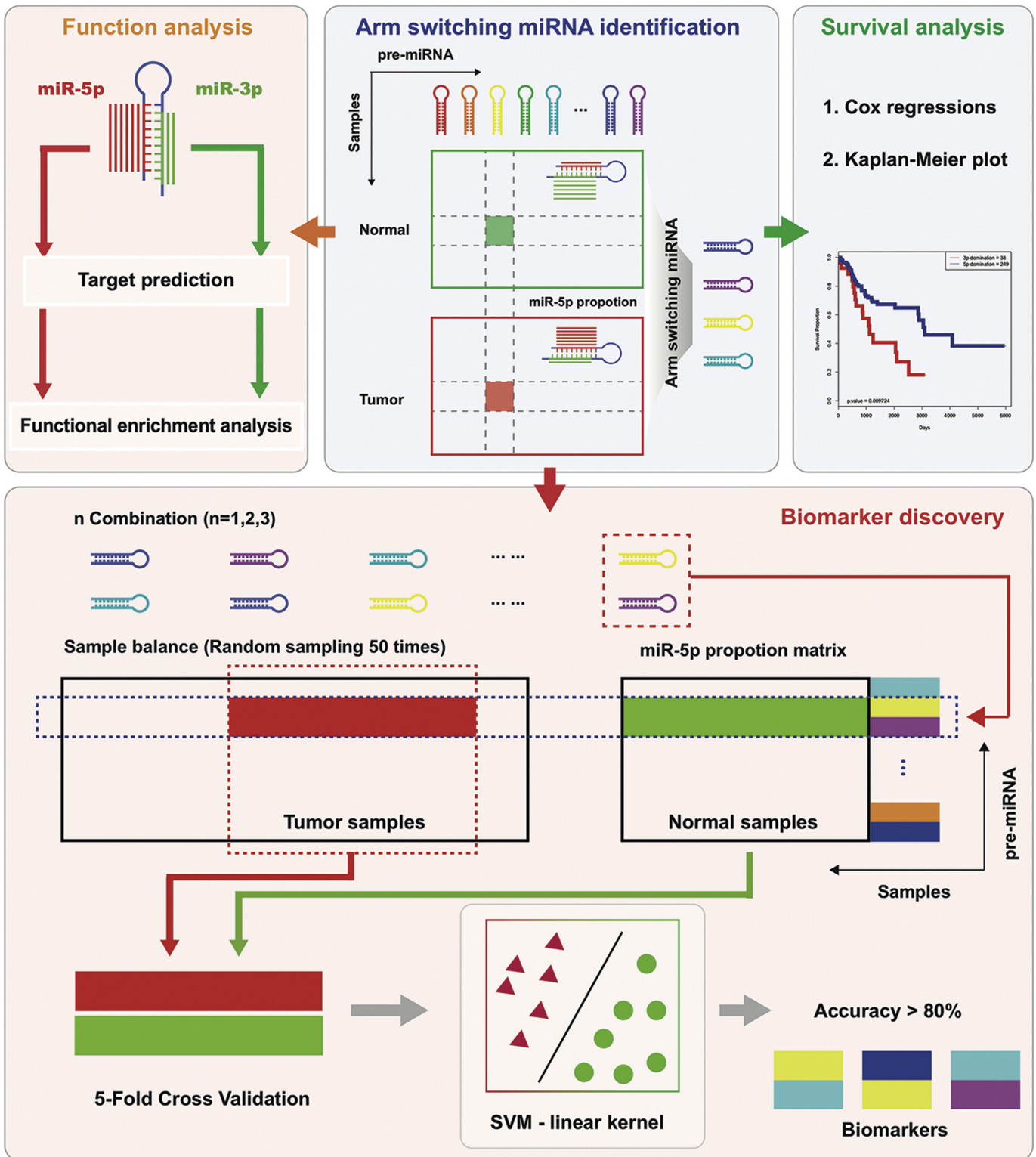
Fig. 2. Workflow of this study. The study is divided into four parts: arm switching miRNA identification, functional enrichment analysis, survival analysis, and biomarker discovery. Data sources and tools used in each part are provided in the methods.

2.4. Survival analysis

Multivariate Cox regressions were performed with the coxph function from the R survival library to test statistical significance [49].

2.5. Biomarker discovery

miRNA expression profiles can be used for classifying tumours [50], and up- or down- regulated miRNAs could be a biomarker for diagnosis [51]. Briefly, in most studies, miRNA signatures are the expression of

miRNA. In this study, we applied arm switching events in place of expression levels to discover the differences in arm switching signatures between tumour and normal tissues.

The workflow of biomarker discovery is shown in Fig. 2. In the TCGA dataset, the number of tumour samples was always more than normal samples and it was necessary to balance the sample number before training. We sampled the same number of tumour samples as normal samples for each cancer. We randomly sampled 50 times to make the negative and positive training sets to have equal size, and the average classification accuracy is based on each time. For each k-combination miRNA, we applied SVM with linear kernel and combined a 5-fold cross validation method to select the best combination of tumour markers [52]. The average classification accuracy of each k-combination miRNA should be above 80% to meet the conditions of selection.

3. Results

3.1. Arm switching events identified between tumour and tumour-adjacent

Based on the workflow described in methods (Fig. 2), we searched for arm switching miRNAs between tumour and tumour-adjacent tissues. We identified 51 miRNAs that displayed arm switching in 16 different tumours: Colon adenocarcinoma (COAD), Kidney renal papillary cell carcinoma (KIRP), Uterine corpus endometrial carcinoma (UCEC), Kidney chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), Esophageal carcinoma (ESCA), Prostate adenocarcinoma (PRAD), Bladder urothelial carcinoma (BLCA), Head and neck squamous cell carcinoma (HNSC), Lung adenocarcinoma (LUAD), Lung squamous cell carcinoma (LUSC), Cholangiocarcinoma (CHOL), Liver hepatocellular carcinoma (LIHC), Breast invasive carcinoma (BRCA), Thyroid carcinoma (THCA), and Stomach adenocarcinoma (STAD) with a criteria that the number of the tumour-adjacent samples is >8. Fig. 3 shows 51 arm switching miRNAs detected from 16 tumours vs. tumour-adjacent tissues (normal tissues). Results of basic statistical analysis of these arm switching miRNAs are in Supplementary Table 2.

We found some miRNAs that could frequently arm switch. They were hsa-mir-30e (in 10 tumours), hsa-mir-17 (in 9 tumours),

hsa-mir-671 (in 8 tumours), hsa-mir-106b (in 7 tumours), hsa-mir-181c (in 7 tumours), hsa-mir-556 (in 7 tumours), hsa-mir-324 (in 6 tumours), and hsa-mir-1228 (in 5 tumours). While most arm-switching events were consistent in arm preference (miR-5p to miR-3p or vice versa) in tumour tissues, hsa-mir-106b and hsa-mir-130b showed different arm preference depending upon the tumour (miR-5p dominant in some, miR-3p in others). COAD showed the most number of arm-switched miRNAs.

3.2. Functional enrichment analysis of paired miR-3p and miR-5p

In order to explore the arm switching effect of the 51 pre-miRNAs, we performed GO and KEGG enrichment analysis for miR-3p and miR-5p arms, respectively. For this analysis, we used the experimentally validated targets of 51 paired miRNAs extracted from miRTarBase 7.0 [43] and TarBase 8.0 [44]. The results of these 51 arm switching miRNAs in heat map format are shown in Fig. 4. The full enrichment analysis results in text format are shown in Supplementary Table 3. We found that for most pre-miRNAs, the respective targets of the two individual arms are associated with different terms, and similarity of terms shows different functions. While the arm switching event of some miRNAs do not affect the function, such as mir-30e and mir-17.

3.3. Novel tumour biomarkers based on miRNA arm switching

We applied SVM with 5-fold cross validation method to exhaustively search n ($n = 1, 2, 3$) combination of 51 arm switching miRNAs for 16 cancers (tumour-adjacent samples number ≥ 8). One-, 2- and 3-combination biomarkers with average accuracy are listed in Supplementary Table 4 whose average classification power accuracy is >80%. From Supplementary Table 4, we found 23 unique arm switching miRNAs from 8 cancers (BLCA, CHOL, COAD, KICH, LIHC, LUAD, LUSC, and UCEC). Boxplots of expression from the 2 arms of the 23 potential biomarkers that reached statistical significance are shown in Fig. 5.

Among the best biomarkers, hsa-mir-29b-2 selects miR-29-3p arm as dominantly expressed in tumour and the miR-5p ratio of hsa-mir-29b-2 can classify COAD tumour and normal with 100%
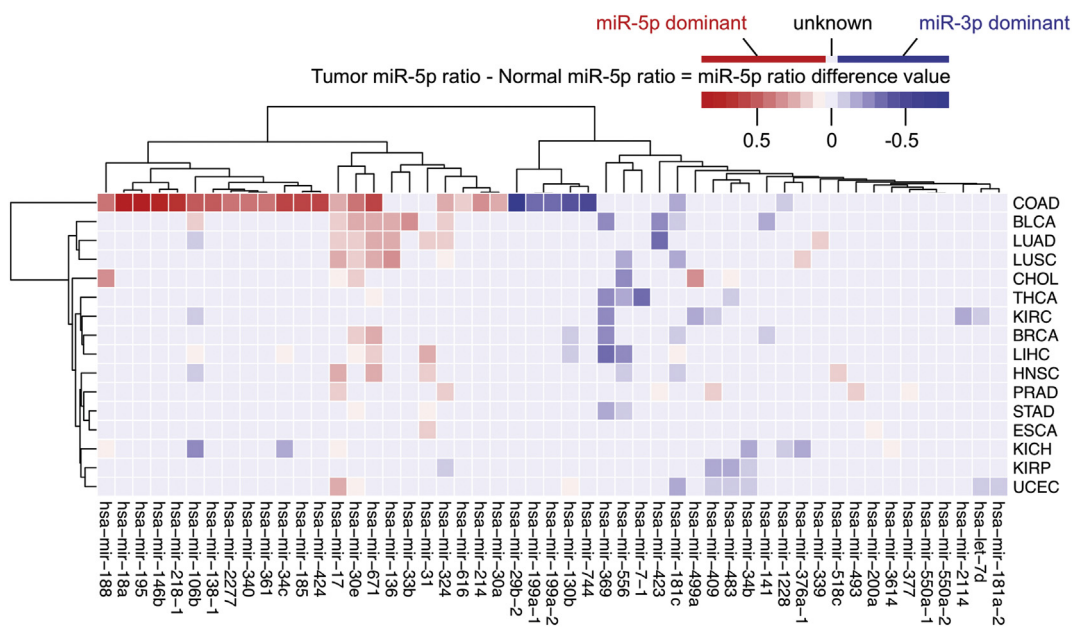
Fig. 3. Heat map of arm switching miRNA in different tumours. The value in the heat map was defined as the difference between tumour miR-5p ratio and tumour-adjacent (normal) miR-5p ratio. Red indicates the miR-5p dominantly expressed in a certain tumour compared to its adjacent tissue. Blue indicates the miR-3p dominantly expressed and light grey indicates unknown (not expressed or not switched). Rows indicate cancer types and columns indicate arm-switched miRNAs.

accuracy. As a tumour promoter, miR-29-3p mediates epithelial-mesenchymal transition (EMT) and promotes metastasis in breast cancer and colon cancer [53]. The miR-5p ratio of hsa-mir-136 could classify tumour and normal in LUAD (accuracy: 0.83), LUSC (accuracy: 0.92), and BLCA (accuracy: 0.81). The miR-5p ratio of hsa-mir-30e could classify tumour and normal in BLCA (accuracy: 0.83), CHOL (accuracy: 0.93), COAD (accuracy: 0.96), LIHC (accuracy: 0.81), and LUSC (accuracy: 0.82). miR-136-5p and miR-30e-5p are dominantly expressed in tumour tissue (see Fig. 3).

3.4. Arm switching events identified in pan-cancer

We also investigated arm switching between 33 different tumour types. Fig. 6 shows a landscape of arm switching events in pan-cancer. In Fig. 6, Ovarian serous cystadenocarcinoma (OV) is significantly different from other cancers in arm switching patterns. In most miRNAs, we could observe three patterns: miR-5p preference, miR-3p preference, and not expressed in tumours compared to others. When pre-miRNA arm preference changes, it usually does not change in many different cancers, which suggests arm switching may be a specific event for each cancer and not a general cancer mechanism.

We tested the association of the identified 51 miRNAs based on miR-5p expression ratio of pre-miRNA with patient survival. We found 7 miRNAs that were highly associated with patient survival. The Kaplan-Meier plots of survival association with the 7 biomarkers are shown in Fig. 7. Ten associations were identified where arm-switched miRNAs were associated with increased survival. Some of the miRNAs were associated with increased survival in more than one cancer. For example, hsa-mir-181c was associated with BRCA, LIHC, and LUSC. Conversely, 3 cancers (KIRP, LIHC, and UCEC) had more than one arm-switched miRNA associated with it. The strand dominance did not favour miR-5p or miR-3p as some miRNAs were represented by both strands. While those miRNAs identified were significant ($p < .05$, Cox regression), some of the differences in survival were dramatic such as hsa-mir-34b (KIRP) and hsa-mir-181a-2 (UCEC).

4. Discussion

Cancer is a complex disease that involves multiple factors that play a role in its development and progression. miRNA is one of these key factors, and its dysregulated expression has been altered in different cancer types and during several stages of cancer. In the present study, we applied and validated miRNA arm switching expression as a basis for
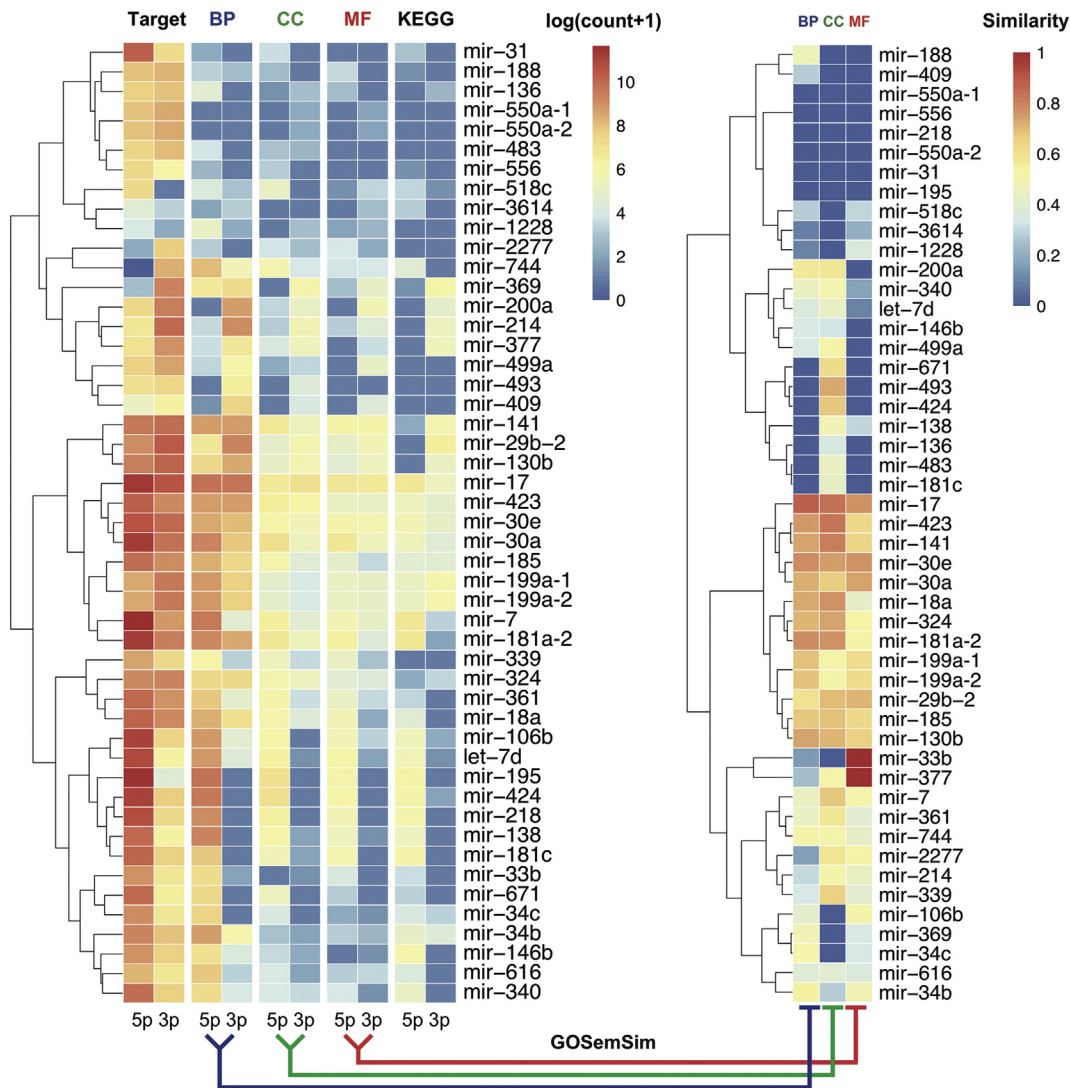

Fig. 4. Heat map of functional comparison of arm-switched miRNAs. The left heat map shows the target gene count, GO term count (BP, biological process; CC, cellular compartment; MF, molecular function), and KEGG pathway count. The value in the heat map is log (count +1). Rows are pre-miRNAs and columns are miR-5p and miR-3p arms, respectively. The right heat map shows the semantic similarity between enriched terms of the two arms. The detailed results of enrichment analysis of arm switching miRNAs are provided in Supplementary Table 3.
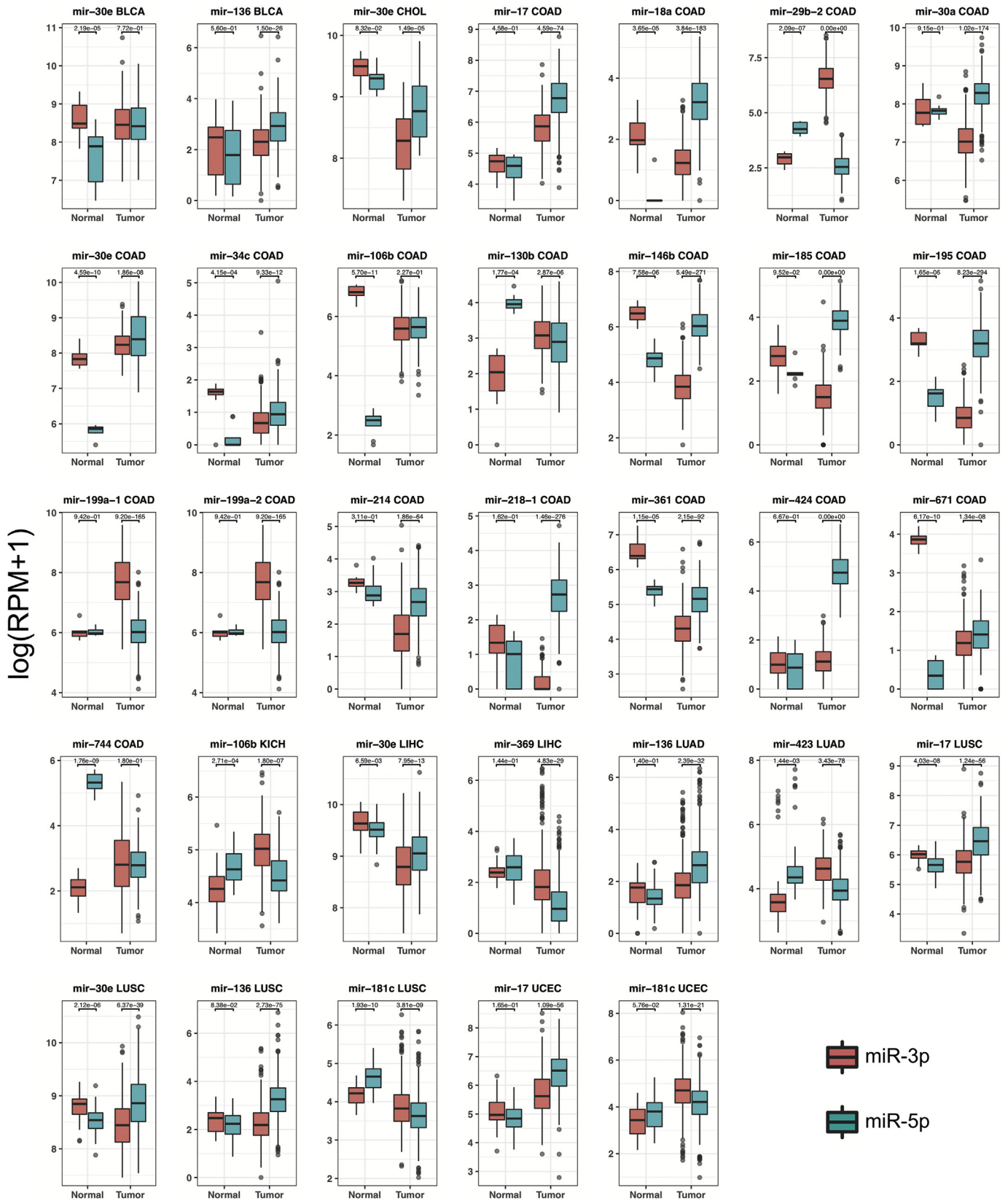
Fig. 5. Boxplot of expression of two arms in arm switching miRNA biomarkers. The boxplot illustrates the expression distribution of two arms from the pre-miRNA in different sample types (tumour or tumour-adjacent), using the log of RPM plus 1. The products from different arms are distinguished by colour (red, miR-3p and blue, miR-5p). The significance level value is shown above the box plots (*t*-test, *p*-value). The tissue types are marked below the box plots (Normal (tumour-adjacent) and tumour). The y-axis represents the expression level of reads log (RPM + 1).
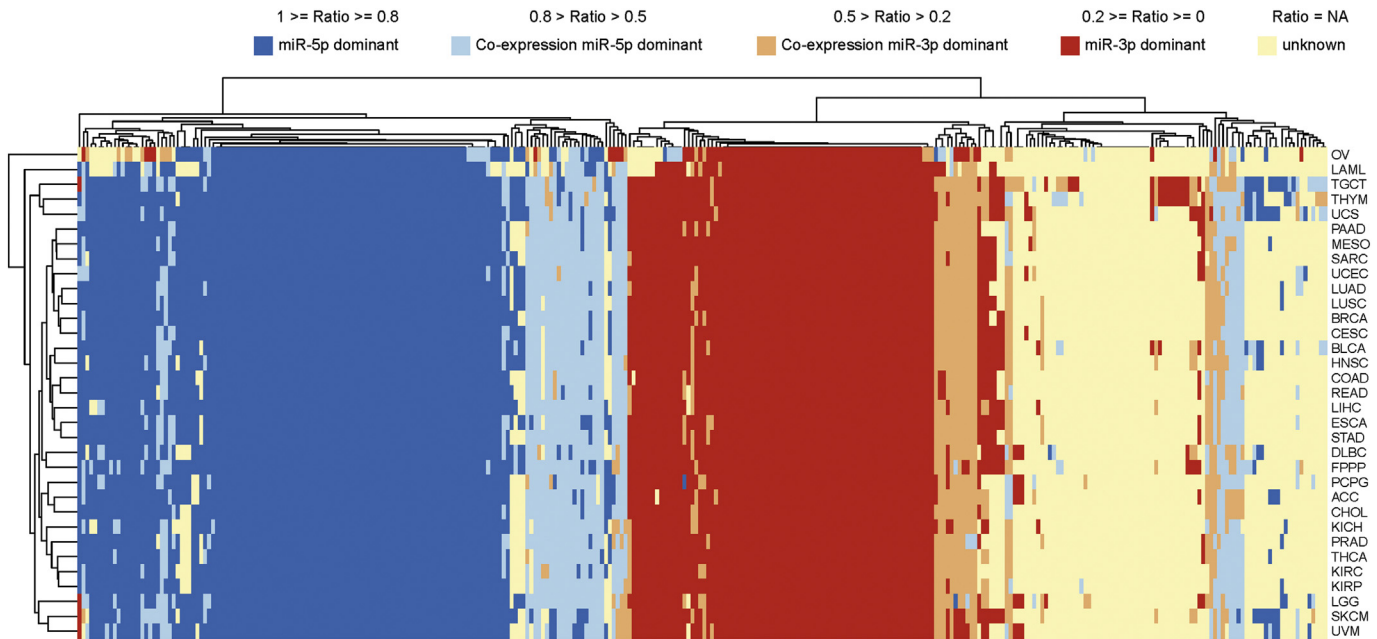
Fig. 6. Heat map of 318 pre-miRNA arm selection preference events in pan-cancer. Blue represents miR-5p is the dominant arm and red indicates miR-3p is dominant. The ratio was adopted as a measure of the arm selection, where 1 ≥ ratio ≥ 0.8 indicates miR-5p dominant expression, 0.8 > ratio > 0.5 indicates co-expression miR-5p dominant expression, 0.5 > ratio > 0.2 indicates co-expression miR-3p dominant expression, and 0.2 ≥ ratio ≥ 0 indicates miR-3p dominant expression. Arm switching degree under the detection threshold or when data is unavailable are filled by light orange.

biomarker discovery. We identified 51 pre-miRNAs that displayed arm switching in different cancer types. All arms from 51 pre-miRNAs have been associated with cancer from previous studies. A summary of these studies are shown in Supplementary Table 5. We observed that the miR-5p and miR-3p were treated separately in those studies and were found dysregulated in different tumours. Some were reported as arm switching miRNAs or analysed as miR-3p and miR-5p together by previous studies. For example, in 2015, Pundhir et al. found mir-30e and mir-30a could arm switch in 9 different tissue cell lines based on ENCODE datasets [54]. Earlier studies have found that two arms of mir-17 could participate in the same biological process. In 2013, Yang et al. found that miR-17-5p and miR-17-3p could target TIMP3 at the same time to promote the growth and invasion of prostate cancer [55]. Shan et al. found that miR-17-5p and miR-17-3p induced liver cancer by targeting PTEN and GalNT7 [56]. Impressively, some arm switching miRNAs from the 51 pre-miRNAs could be circulating miRNA biomarkers for cancers. They were serum miRNAs: miR-369-5p, miR-30e-5p, and miR-423-3p and plasma miRNAs: miR-671-3p, miR-483-3p, and miR-744-5p.

While the TCGA database has proven to be a rich resource for biomarker studies, our results would be strengthened if we could validate some of our findings in independent data sets. Taking a sampling of datasets from the GEO database, we were able to confirm arm switching of mir-324, mir-181c, mir-556, mir-195, mir-424, mir-185, and mir-483 in the same direction (miR-5p dominant to miR-3p dominant or vice versa) and in same or similar tumour types. The GEO dataset, platform, tumour type and number of samples and t-test values are provided in Supplementary Table 6. GO and KEGG analysis of their experimentally verified targets are provided in Supplementary Fig. 1. As we did not exhaustively search the GEO database and we confirmed in different microarray and sequencing platforms, there may be more arm-switched miRNAs that could be validated.

We also identified 7 arm switching miRNAs associated with patient survival and 23 high accuracy 1-combination arm switching miRNAs that classify tumour and normal and many 2- and 3-

combination biomarkers. Based on the enrichment analysis, we were not able to determine whether the two arms of miRNA work cooperatively or not. The arm switching pattern generated from 33 tumours denote that arm switching does not frequently occur, and most miRNAs tend to conservatively choose one arm as the dominantly expressed one.

Some studies analysed miR-3p and miR-5p forms together. Dual strands of hsa-mir-150 and hsa-mir-145 act as tumour-suppressors in prostate cancer and lung cancer respectively [57,58]. In previous studies, miR-195-5p tends to suppress the development of the tumour. In colon cancer, miR-195-5p promotes cell apoptosis and suppresses tumour cell invasion and metastasis by downregulating the Bcl-2 gene [59]. In liver cancer, miR-195-5p targets PHF19 and CBX4 to suppress the development of the tumour [60,61]. Based on Supplementary Table 3, we found that miR-195-5p has a well-defined biological function, while miR-195-3p had no enriched pathways and non-significant GO terms. miR-17-5p and miR-17-3p can enrich the same KEGG pathways related to cancer, such as Transcriptional misregulation in cancer (hsa05202), Colorectal cancer (hsa05210), Glioma (hsa05214), Bladder cancer (hsa05219), Chronic myeloid leukemia (hsa05220), Hepatocellular carcinoma (hsa05225), and Gastric cancer (hsa05226).

The expression levels of miRNAs may be small, and this limits sensitivity. Nonetheless, our study demonstrates the concept of using miRNA arm-switching in cancer tissues to observe dynamic changes in miRNA at the molecular level. As sequencing technologies advance, and costs for sequence determination decrease rapidly, the feasibility of sequencing tissues for biomarkers is becoming greater. Our studies indicate here specific miRNA arms that should be interrogated as markers when tissue is available. Finally, while the function of RNA arm switching events remains to be elucidated, our results provide additional motivation for its investigation. Such knowledge would be useful not only in enlightening basic molecular biology mechanisms, but also in gaining knowledge in pathophysiology of cancer.
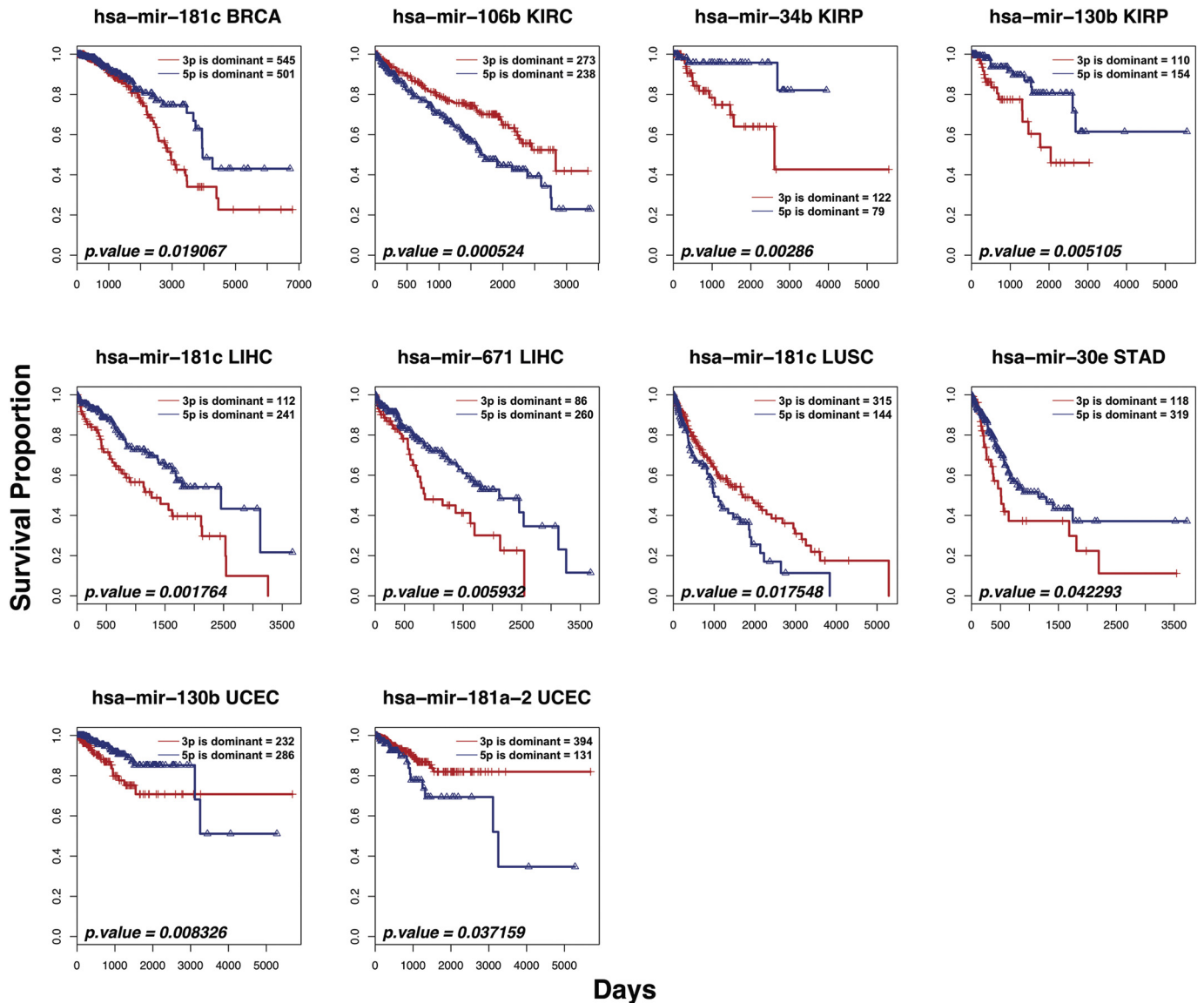
Fig. 7. Kaplan-Meier plot of survival associated biomarkers. Seven different pre-miRNAs biomarkers in 7 tumours are shown. The survival proportion and survival in days are indicated. Blue and red represent dominant expression of miR-5p or miR-3p, respectively. Number of samples and statistical significance from Cox proportional hazards regression are indicated inside the plots. Abbreviations for the tumour types are: BRCA, Breast invasive carcinoma; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LIHC, Liver hepatocellular carcinoma; LUSC, Lung squamous cell carcinoma; STAD, Stomach adenocarcinoma; UCEC, Uterine corpus endometrial carcinoma.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2018.11.003.

Declarations of interests

The authors have no conflicts of interest.

Author contributions

Garry Wong and Liang Chen designed research and wrote the manuscript; Liang Chen and Huiyan Sun performed the analysis; Changliang Wang, Yang Yang, and Menglei Zhang performed data collection and pre-processing of data. All authors approved the manuscript.

References

[1] Zhao Y, Srivastava D. A developmental view of microRNA function. Trends Biochem Sci 2007;32(4):189–97.
[2] Bartel DP. Metazoan MicroRNAs. Cell 2018;173(1):20–51.
[3] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004; 116(2):281–97.

[4] Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 2014;15 (8):509–24.

[5] Pillai RS. MicroRNA function: multiple mechanisms for a tiny RNA? RNA (New York, NY) 2005;11(12):1753–61.

[6] Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 2010;11(9):597–610.

[7] Okamura K, Liu N, Lai EC. Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. Mol Cell 2009;36(3):431–44.

[8] Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell 2003;115(2):209–16.

[9] Noland CL, Doudna JA. Multiple sensors ensure guide strand selection in human RNAi pathways. RNA (New York, NY) 2013;19(5):639–48.

[10] Hu HY, Yan Z, Xu Y, et al. Sequence features associated with microRNA strand selection in humans and flies. BMC Genomics 2009;10:413.

[11] Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD. Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA (New York, NY) 2010;16(1):43–56.

[12] Suzuki HI, Katsura A, Yasuda T, et al. Small-RNA asymmetry is directly driven by mammalian Argonautes. Nat Struct Mol Biol 2015;22(7):512–21.

[13] Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res 2007;17(12):1850–64.

[14] Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, Lai EC. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. Nat Struct Mol Biol 2008;15(4):354–63.

[15] Guo L, Lu Z. The fate of miRNA* strand through evolutionary analysis: implication for degradation as merely carrier strand or potential regulatory molecule? PLoS One 2010;5(6):e11387.

[16] Liu N, Okamura K, Tyler DM, Phillips MD, Chung WJ, Lai EC. The evolution and functional diversification of animal microRNA genes. Cell Res 2008;18(10):985–96.

[17] Chiang HR, Schoenfeld LW, Ruby JG, et al. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev 2010;24(10): 992–1009.

[18] Ro S, Park C, Young D, Sanders KM, Yan W. Tissue-dependent paired expression of miRNAs. Nucleic Acids Res 2007;35(17):5944–53.

[19] Griffiths-Jones S, Hui JH, Marco A, Ronshaugen M. MicroRNA evolution by arm switching. EMBO Rep 2011;12(2):172–7.

[20] Dudoit S, Fridlyand J. A prediction-based resampling method for estimating the number of clusters in a dataset. Genome Biol 2002;3(7) RESEARCH0036.

[21] Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell 2003;115(2):199–208.

[22] Wilson RC, Tambe A, Kidwell MA, Noland CL, Schneider CP, Doudna JA. Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis. Mol Cell 2015;57(3):397–407.

[23] Wu H, Ye C, Ramirez D, Manjunath N. Alternative processing of primary microRNA transcripts by Drosha generates 5′ end variation of mature microRNA. PLoS One 2009;4(10):e7566.

[24] Marco A, Hui JH, Ronshaugen M, Griffiths-Jones S. Functional shifts in insect microRNA evolution. Genome Biol Evol 2010;2:686–96.

[25] Marco A, Macpherson JI, Ronshaugen M, Griffiths-Jones S. MicroRNAs from the same precursor have different targeting properties. Silence 2012;3(1):8.

[26] Guo L, Zhang H, Zhao Y, Yang S, Chen F. Selected isomiR expression profiles via arm switching? Gene 2014;533(1):149–55.

[27] de Wit E, Linsen SE, Cuppen E, Berezikov E. Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Res 2009;19(11):2064–74.

[28] Berezikov E. Evolution of microRNA diversity and regulation in animals. Nat Rev Genet 2011;12(12):846.

[29] Hayes J, Peruzzi PP, Lawler S. MicroRNAs in cancer: biomarkers, functions and therapy. Trends Mol Med 2014;20(8):460–9.

[30] Liu R, Chen X, Du Y, et al. Serum microRNA expression profile as a biomarker in the diagnosis and prognosis of pancreatic cancer. Clin Chem 2012;58(3):610–8.

[31] Lee EJ, Gusev Y, Jiang J, et al. Expression profiling identifies microRNA signature in pancreatic cancer. Int J Cancer 2007;120(5):1046–54.

[32] Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 2008;105(30): 10513–8.

[33] Wu X, Zeng R, Wu S, Zhong J, Yang L, Xu J. Comprehensive expression analysis of miRNA in breast cancer at the miRNA and isomiR levels. Gene 2015;557(2): 195–200.

[34] Zhang H, Yang S, Guo L, Zhao Y, Shao F, Chen F. Comparisons of isomiR patterns and classification performance using the rank-based MANOVA and 10-fold cross-validation. Gene 2015;569(1):21–6.

[35] Mitra R, Lin CC, Eischen CM, Bandyopadhyay S, Zhao Z. Concordant dysregulation of miR-5p and miR-3p arms of the same precursor microRNA may be a mechanism in inducing cell proliferation and tumorigenesis: a lung cancer study. RNA (New York, NY) 2015;21(6):1055–65.

[36] Tsai KW, Leung CM, Lo YH, et al. Arm selection preference of MicroRNA-193a varies in breast cancer. Sci Rep 2016;6:28176.

[37] Li SC, Liao YL, Ho MR, Tsai KW, Lai CH, Lin WC. miRNA arm selection and isomiR distribution in gastric cancer. BMC Genomics 2012(13 Suppl 1):S13.

[38] Choo KB, Soon YL, Nguyen PN, Hiew MS, Huang CJ. MicroRNA-5p and -3p co-expression and cross-targeting in colon cancer cells. J Biomed Sci 2014;21:95.

[39] Kuo WT, Su MW, Lee YL, et al. Bioinformatic interrogation of 5p-arm and 3p-arm specific miRNA expression using TCGA datasets. J Clin Med 2015;4(9):1798–814.

[40] Lin MH, Chen YZ, Lee MY, et al. Comprehensive identification of microRNA arm selection preference in lung cancer: miR-324-5p and -3p serve oncogenic functions in lung cancer. Oncol Lett 2018;15(6):9818–26.

[41] Chung IF, Chang SJ, Chen CY, et al. YM500v3: a database for small RNA sequencing in human cancer research. Nucleic Acids Res 2017;45(D1) D925-D31.

[42] Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 2014;42(Database issue):D68–73.

[43] Chou CH, Shrestha S, Yang CD, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res 2018;46(D1): D296–302.

[44] Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res 2018;46(D1) D239-D45.

[45] Hu W, Wang T, Yue E, Zheng S, Xu JH. Flexible microRNA arm selection in rice. Biochem Biophys Res Commun 2014;447(3):526–30.

[46] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics 2012;16(5):284–7.

[47] Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics 2010;26(7): 976–8.

[48] Wang JZ, Du Z, Payattakool R, Yu PS, Chen CF. A new method to measure the semantic similarity of GO terms. Bioinformatics 2007;23(10):1274–81.

[49] Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. Springer Science & Business Media; 2013.

[50] Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. Nature 2005;435(7043):834–8.

[51] Andorfer CA, Necela BM, Thompson EA, Perez EA. MicroRNA signatures: clinical biomarkers for the diagnosis and treatment of breast cancer. Trends Mol Med 2011;17 (6):313–9.

[52] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. , 1995Montreal, Canada: Ijcai; 1995; 1137–45.

[53] Jiang H, Zhang G, Wu JH, Jiang CP. Diverse roles of miR-29 in cancer (review). Oncol Rep 2014;31(4):1509–16.

[54] Pundhir S, Gorodkin J. Differential and coherent processing patterns from small RNAs. Sci Rep 2015;5:12062.

[55] Yang X, Du WW, Li H, et al. Both mature miR-17-5p and passenger strand miR-17-3p target TIMP3 and induce prostate tumor growth and invasion. Nucleic Acids Res 2013;41(21):9688–704.

[56] Shan SW, Fang L, Shatseva T, et al. Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. J Cell Sci 2013;126(Pt 6):1517–30.

[57] Okato A, Arai T, Kojima S, et al. Dual strands of pre-miR150 (miR1505p and miR1503p) act as antitumor miRNAs targeting SPOCK1 in naive and castration-resistant prostate cancer. Int J Oncol 2017;51(1):245–56.

[58] Mataki H, Seki N, Mizuno K, et al. Dual-strand tumor-suppressor microRNA-145 (miR-145-5p and miR-145-3p) coordinately targeted MTDH in lung squamous cell carcinoma. Oncotarget 2016;7(44):72084–98.

[59] Yang B, Tan Z, Song Y. Study on the molecular regulatory mechanism of MicroRNA-195 in the invasion and metastasis of colorectal carcinoma. Int J Clin Exp Med 2015;8 (3):3793–800.

[60] Zheng C, Li J, Wang Q, et al. MicroRNA-195 functions as a tumor suppressor by inhibiting CBX4 in hepatocellular carcinoma. Oncol Rep 2015;33(3):1115–22.

[61] Xu H, Hu YW, Zhao JY, et al. MicroRNA-195-5p acts as an anti-oncogene by targeting PHF19 in hepatocellular carcinoma. Oncol Rep 2015;34(1):175–82.