Identification and Elucidation of the Protective isomiRs in Lung Cancer Patient Prognosis

Fu-Mei Hsieh, Su-Ting Lai, Ming-Fong Wu and Chen-Ching Lin*

Institute of Biomedical Informatics, National Yang Ming Chiao Tung University, Taipei, Taiwan

MicroRNAs (miRNAs) are approximately 20–22 nucleotides in length, which are well known to participate in the post-transcriptional modification. The mature miRNAs were observed to be varied on 5′ or 3′ that raise another term—the isoforms of mature miRNAs (isomiRs), which have been proven not the artifacts and discussed widely recently. In our research, we focused on studying the 5′ isomiRs in lung adenocarcinoma (LUAD) in The Cancer Genome Atlas (TCGA). We characterized 75 isomiRs significantly associated with better prognosis and 43 isomiRs with poor prognosis. The 75 protective isomiRs can successfully distinguish tumors from normal samples and are expressed differently between patients of early and late stages. We also found that most of the protective isomiRs tend to be with downstream shift and upregulated compared with those with upstream shift, implying that a possible selection occurs during cancer development. Among these protective isomiRs, we observed a highly positive and significant correlation, as well as in harmful isomiRs, suggesting cooperation within the group. However, between protective and harmful, there is no such a concordance but conversely more negative correlation, suggesting the possible antagonistic effect between protective and harmful isomiRs. We also identified that two isomiRs miR-181a-3p|-3 and miR-181a-3p|2, respectively, belong to the harmful and protective groups, suggesting a bidirectional regulation of their originated archetype—miR-181a-3p. Additionally, we found that the protective isomiRs of miR-21-5p, which is an oncomiR, may be evolved as the tumor suppressors through producing isomiRs to hinder metastasis. In summary, these results displayed the characteristics of the protective isomiRs and their potential for developing the treatment of lung cancer.

Keywords: miRNA, isomiR, lung adenocarcinoma, survival analysis, miR-181a-3p, miR-21-5p

INTRODUCTION

MicroRNAs (miRNAs) are a class of small non-coding RNAs (ncRNA) of approximately 22 nts in length encoded in the genome of plants and animals, which have been found to regulate gene expression in the post-transcriptional level (Lee et al., 1993; Reinhart et al., 2000). MiRNAs were found to bind on the 3′ untranslated region (UTR) of target mRNA, which was called miRNA response element (MRE), and usually repress the target mRNA expression (Lai, 2002). Typically, miRNA is first transcribed to primary miRNA (pri-miRNA) by RNA polymerase II and folded to a hairpin structure. Then, the Drosha-DGCR8 complex binds on the pri-miRNA to prune the stem

and produce miRNA precursor (pre-miRNA), which is approximately 60 nts. The pre-miRNAs are further exported into the cytoplasm. In the cytoplasm, the hairpin loop of pre-miRNA is trimmed by Dicer to generate a paired miRNA duplex, each approximately 22 nts. Either strand of the paired miRNA duplex can further bind to the Argonaute protein (AGO) (Olina et al., 2018), form a silencing complex (RISC), and then regulate, usually repress, the target gene expression (Bartel, 2018).

"Lung cancer has been noticed as one of the leading causes of death in cancers in the United States and around the world" (Dela Cruz et al., 2011). Besides, lung cancer has a poor prognosis and thus is worth noticing its progression in treatment, biomarkers, etc. The most common type is non-small cell lung cancer (NSCLC), which takes 85% of lung cancer. NSCLC can be further classified into three subtypes: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. In this study, we mainly focused on lung adenocarcinoma (LUAD), which comprises 40% of all lung cancer (Zappa and Mousa, 2016). MiRNAs are involved in many mechanisms in cells, such as immune responses (Sun et al., 2013; Xu et al., 2014) and tumor progression. They can be tumor suppressors and oncogenes (Esquela-Kerscher and Slack, 2006). More importantly, because of their ubiquitous characteristic, miRNAs are widely detected and found to contribute to cancer treatment (Berindan-Neagoe et al., 2014; Shah et al., 2016). Recent studies have provided insight into the role of miRNAs in tumorigenesis and predicted miRNAs as biomarkers in lung cancer (Azizi et al., 2021; Zhong et al., 2021). Additionally, the functions of miRNA in lung cancer have been widely reported (Iqbal et al., 2019; Wu et al., 2019).

With the development of sequencing technology, the mature miRNAs were observed to be varied on 5′ or 3′ that raise the isoforms of mature miRNAs (isomiRs) (Morin et al., 2008; Guo et al., 2016). Notably, these isomiRs were indicated to be mainly generated from imprecise cleavage or shifting by Drosha and Dicer rather than the degradation during sample preparation for sequencing (Lee et al., 2010; Neilsen et al., 2012). Additionally, some isomiRs can be generated by non-template nucleotide addition, such as adenylation at the 3′-end (Lu et al., 2009). Accordingly, the identification of isomiRs extends the functional significance and variance of miRNA regulation (Mercey et al., 2017). IsomiRs can generally be categorized into three types: 3′ isomiR, 5′ isomiR, and polymorphic isomiR. MiRNA can interact with its mRNA target by Watson–Crick pairing typically through the strong connection of "seed region," which is located on 2-7mer (or 8mer) of 5′ miRNA and guided mRNA targeting (Bartel, 2009; Tan et al., 2014). Therefore, the generation of 5′ isomiRs could alternate the seed region and affect the downstream regulation and functional roles of miRNAs. Recently, the isomiRs have been discovered to show different expression levels among disease subtypes, gender differences (Loher et al., 2014; Guo et al., 2016). The existence of isomiRs may also help to distinguish the cancer type (Telonis et al., 2017). Also, isomiRs may play the dominant miRNA in certain tissues or even become the canonical miRNAs through evolution (Tan et al., 2014). The miRNA regulation relies on miRNA–mRNA complementary binding. The presence or even domination of isomiR indicates the concern of nucleotides alternation in miRNA regulation,

therefore elevating the urgency of clarifying the variance of miRNA regulation derived from isomiRs.

MATERIALS AND METHODS

Collection and Preprocessing of miRNA and mRNA Expression Profiles

The miRNA expression profiles "isoforms.quantification.txt" were obtained from The Cancer Genome Atlas (TCGA)-GDC portal[1], which recorded the coordinate information and expression level miRNAs and their isomiRs. We initially downloaded 567 LUAD samples. For data cleaning, 87 were removed as they contain "annotation.txt," which may be problematic. Also, three duplicated files from the same participant (TCGA-44-6775) and two "Recurrent Solid Tumor" files were skipped. Finally, 475 LUAD miRNA expression files, which contain 442 primary solid tumors and 33 solid tissue normal, were kept. Meanwhile, the RNA-seq data "htseq.counts" were also downloaded. After data cleaning as the same criteria with miRNA expression profiles, 480 RNA-Seq samples (46 normal, 434 tumor) remained. Finally, we kept a total of 444 RNA-Seq samples (14 normal, 430 tumor) that match the isomiR expression profiles for the analysis of miRNA-regulated target genes.
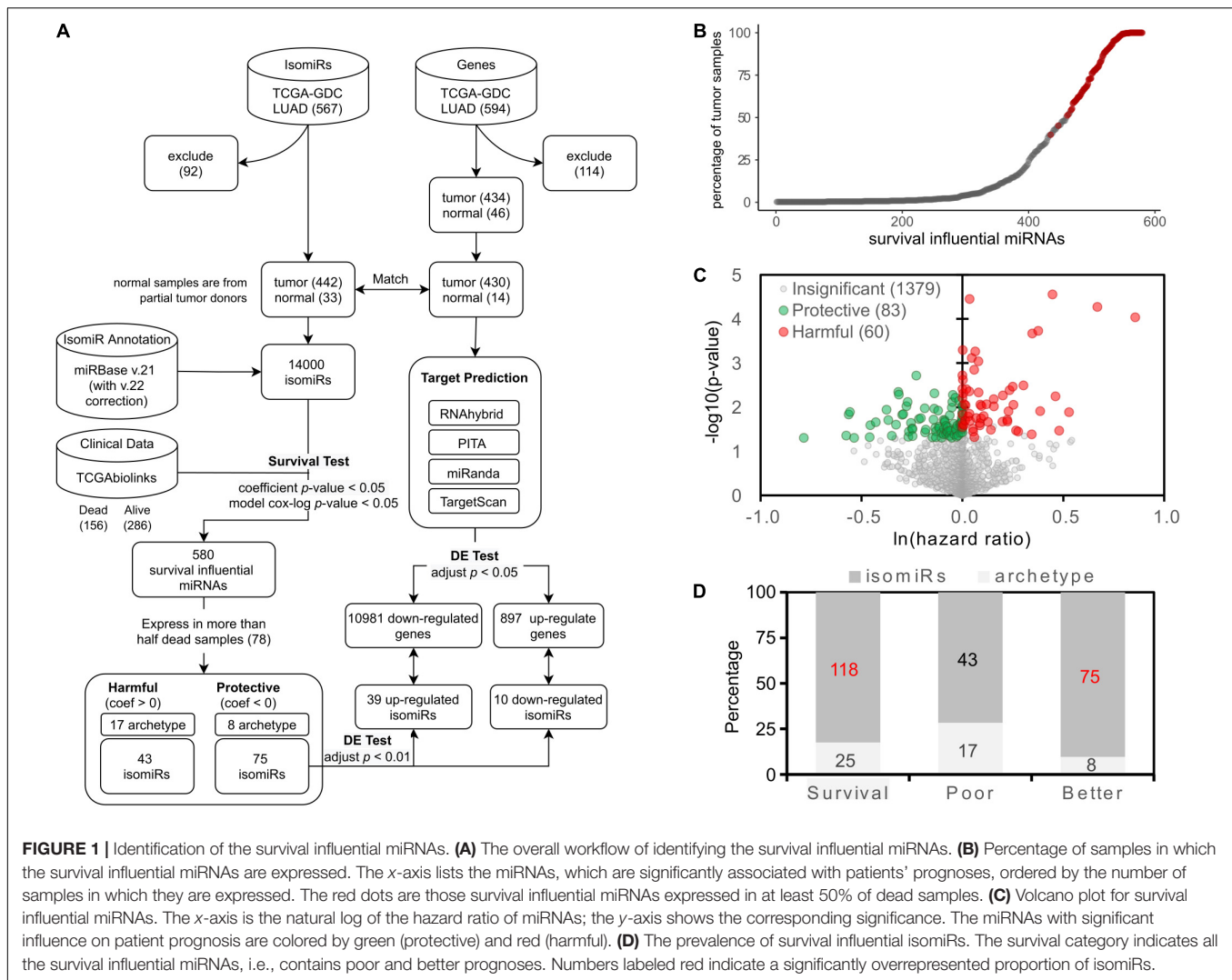
The Nomenclature of isomiRs

In our research, the isomiRs are defined by the relative 5′ start site of archetype miRNA. Since the miRNA annotation of TCGA data was according to the miRBase V21, and the archetype miRNA coordinate "hsa.gff3" was initially downloaded from miRBase (V21 from http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2014; Kozomara et al., 2019). The nomenclature of isomiRs was defined by subtracting the 5′ start site coordinates of the archetype miRNA, which are recorded in "hsa.gff3," and each corresponding isomiR, which is recorded in "isoforms.quantification.txt," followed by previous studies (Telonis et al., 2015, 2017). For example, "MIMAT0000062|-2" and "MIMAT0000062|4" are isomiRs of the archetype "MIMAT0000062" with respectively two upstream and four downstream nucleotides shifting of the archetype. Additionally, we denoted the archetypes as the miRNA with zero shift; for example, the archetype of MIMAT0000062 is named as MIMAT0000062|0. Positive or negative strands and location on chromosomes were all considered in different subtraction cases. Herein, we use miRNAs to represent the set of archetype miRNA and isomiRs. Furthermore, the miRBase has been updated to V22 (Kozomara et al., 2019); we then updated the annotation of miRNAs and isomiRs in TCGA dataset for the analysis in this study. The overall workflow of this study is depicted in **Figure 1A**.

Identification of Survival Influential and Differentially Expressed isomiRs

In this study, we applied the Cox regression model to discover survival influential miRNAs in LUAD. The Cox regression model

---

[1] https://portal.gdc.cancer.gov/

FIGURE 1 | Identification of the survival influential miRNAs. **(A)** The overall workflow of identifying the survival influential miRNAs. **(B)** Percentage of samples in which the survival influential miRNAs are expressed. The *x*-axis lists the miRNAs, which are significantly associated with patients' prognoses, ordered by the number of samples in which they are expressed. The red dots are those survival influential miRNAs expressed in at least 50% of dead samples. **(C)** Volcano plot for survival influential miRNAs. The *x*-axis is the natural log of the hazard ratio of miRNAs; the *y*-axis shows the corresponding significance. The miRNAs with significant influence on patient prognosis are colored by green (protective) and red (harmful). **(D)** The prevalence of survival influential isomiRs. The survival category indicates all the survival influential miRNAs, i.e., contains poor and better prognoses. Numbers labeled red indicate a significantly overrepresented proportion of isomiRs.

was performed using the coxph function of the Survival R package (Therneau, 2015). The clinical data of LUAD patients were retrieved by *TCGAbiolinks* (v.2.18.0) (Mounir et al., 2019) in *Bioconductor* (v.3.12) release. We then performed the multivariate Cox regression model—the covariates are the expression level of the tested miRNA and the confounding clinical factors—to assess the influence of a miRNA on patient survival. Then, incorporated confounding factors are "age_at_index," "cigarettes_per_day," and "ajcc_pathologic_t." In this study, we denoted the miRNAs with positive and negative coefficients in the model as harmful and protective to patient survival, respectively. The significance of the coefficient β is determined by comparing with the null model, which hypothesizes that the changes of tested miRNA expression level do not affect patient survival, and tested by a likelihood ratio test and the Wald test (Andersen and Gill, 1982; Therneau et al., 2000). The Wald test examines whether the observed regression coefficient statistically differs from zero and reported a z-score (standard score) and a *p*-value estimated by z-score for the observed regression coefficient. The miRNAs with the coefficient of *p*-value < 0.05 were recognized as the

survival influential miRNAs, that is, they have a significant impact on patient survival. Consequently, 580 miRNAs were identified as survival influential miRNAs. However, we observed that approximately 68% identified survival influential miRNAs were expressed in less than 20% of patients (**Figure 1B**). In addition, previous studies (Concato et al., 1995; Peduzzi et al., 1995; Harrell et al., 1996) suggested that 10-20 events per predictor variable (EPV) could be appropriate for the construction of Cox regression model. In our Cox regression model, we used four predictor variables, which are expression level of the tested miRNA, and three confounding factors; thus, the proper number of events (dead samples) could be ranged from 30 to 80. Accordingly, to preserve enough coverage of patients in which the tested miRNA expressed and met the proper EPV, we decided to keep the survival influential miRNAs expressed in more than 78 dead samples, which also can cover approximately 50% samples (**Figure 1B**, red circles). Finally, 143 survival influential miRNAs are used in the following analysis.

Next, we used the *limma* package (v.3.46.0) (Ritchie et al., 2015) and *edgeR* (v.3.32.1) (McCarthy et al., 2012) to identify the

differentially expressed (DE) isomiRs and their mRNA targets. To be consistent with the survival analysis, we used 1,522 isomiRs, which are expressed in more than 50% dead tumor samples and cover at least 41% samples, to perform differential expression analysis. The isomiR expression profiles were normalized by the Upper Quartile method and modeled with *limFit* function and *eBayes* function.

To speculate the possible functions of these isomiRs, we used the RNA-seq data of the LUAD samples as target gene expression profiles. The 3′-UTR sequences of mRNAs were also retrieved to predict the potential isomiRs' binding targets. The RNA-seq data were first normalized with TMM, a method for estimating relative RNA production levels (Robinson and Oshlack, 2010), and using *voom* function and *eBayes* function for modeling. We found 8,557 DE genes (adjusted *p*-value = 0.01), with 3,484 upregulated and 5,073 downregulated.

Prediction of isomiR-regulated Target Genes

We mainly use R (v4.0.3) and its packages in this research. The *Biopython* (Cock et al., 2009), which is a Python package, was used to get sequences of isomiRs and the transcript sequences. Since we only focused on the 5′ variation of isomiRs, the sequence length to where 3′ should stop was defined by the longest one for each isomiR in the whole LUAD dataset.

With the whole gene expression (444 samples) and isomiRs expression (475 samples), there are four tools used for isomiRs' target prediction, which are: *miRanda* (v.3.3a) (Enright et al., 2003; John et al., 2004), *RNAhybrid* (v.2.1.2) (Rehmsmeier et al., 2004), *PITA* (version 6, Aug-31-08) (Kertesz et al., 2007), and TargetScan (v.7.2) (Agarwal et al., 2015); the predicted isomiR-target pairs numbers were 106,335,587, 106,439,448, 9,231,767, and 77,795,459, respectively. These algorithms follow different principles, such as the seed match, sequence complementary, the secondary structure, and free energy, thus none of them could predict the isomiR-target interactions comprehensively. Thus, we set a stringent cut-off for pairs presented in three of the four prediction tools, where only 13,091,963 pairs were kept. Finally, combined with the DE isomiR, DE gene results, and survival analysis, the protective isomiR-target pairs were classified to 10,981 (39 upregulated isomiR vs. downregulated gene) and 897 (10 dowregulated isomiR vs. upregulated gene).

Functional Analysis of isomiRs in the Protective and Harmful Groups

The isomiR-regulated network for the functional test was selected by Spearman's correlation, estimated by *stats:cor.test*, and visualized by *psych:cor.plo*t. Because there are many zero counts in sample-wise dimension, we compute every isomiR–isomiR correlation by intersecting non-zero value. Thus, not every correlation has the same sample number. Then, the correlation was selected by a positive coefficient (rho > 0) and top 5% by ranking both *p*-value and rho; also, only the nodes with degree > 1 were selected. For analyzing biological processes of Gene Ontology, we used *clusterProfiler:enrichGO* function (v.3.18.1), where except for the qvalueCutoff (set as 1), the

other parameters were set as default. The discovered functional modules were further summarized by REVIGO (Supek et al., 2011) algorithm with a similarity ≥ 0.9 that is calculated from the Resnik algorithm (Resnik, 1999) and visualized by CirGO package (Kuznetsova et al., 2019).

RESULTS

Identification of Survival Influential isomiRs in Lung Cancer

In this study, we performed the multivariate Cox regression model to identify survival influential miRNAs. The workflow is depicted in **Figure 1A**. To meet the appropriate EPV (Concato et al., 1995; Peduzzi et al., 1995; Harrell et al., 1996) and preserve the sufficient coverage of patients in which the tested miRNA was expressed (**Figure 1B**), only the identified survival influential miRNAs expressed in more than 50% dead patients (78) were kept for the following analysis. Consequently, 143 miRNAs are identified—60 and 83 miRNAs are harmful and protective, respectively (**Figure 1C**). Notably, 118 out of 143 survival influential miRNAs are isoforms (isomiRs) (**Figure 1D**). This proportion (83%) is significantly overrepresented (*p*-value = $3.21*10^{-5}$, Fisher's exact test). In other words, the archetype miRNAs are significantly underrepresented in these 143 survival influential miRNAs. Moreover, isomiRs associated with better prognosis of patients are significantly enriched in these 118 survival influential isomiRs (75/118, *p*-value = $8.39*10^{-7}$, Fisher's exact test). However, both archetype miRNAs and isomiRs are neither significantly overrepresented nor underrepresented in miRNAs associated with poor prognosis of patients. These results suggested that the survival influential isomiRs might tend to be protective and, thus, play roles as tumor suppressors and putative therapeutics for patients with lung cancer.

Characteristics of the Identified Protective isomiRs

To investigate the characteristics of these protective isomiRs, we studied their expression profiles in LUAD patients. **Figure 2A** shows the fold change and adj. *p*-value of these 75 protective isomiRs. Among the 75 protective isomiRs, 39 and 10 isomiRs are significantly up- and downregulated, respectively. The proportion of upregulated isomiRs (52%) is overrepresented, but the significance is moderate (right-tailed *p*-value = 0.09, Fisher's exact test), and the proportion of downregulated isomiRs (13%) is close to randomly expected (right-tailed *p*-value = 0.44, Fisher's exact test). This observation suggested that the identified protective isomiRs might tend to be upregulated in lung cancer. Additionally, we observed that the 75 protective isomiRs can separate tumor from normal samples significantly in the principal component analysis (PCA) (**Figure 2B**), but the expression profiles of completely 14,000 isomiRs and 1,522 selected isomiRs cannot (**Supplementary Figure 1**). Besides, using the principal components (PC1 and PC2) of 75 protective isomiRs with logistic regression model
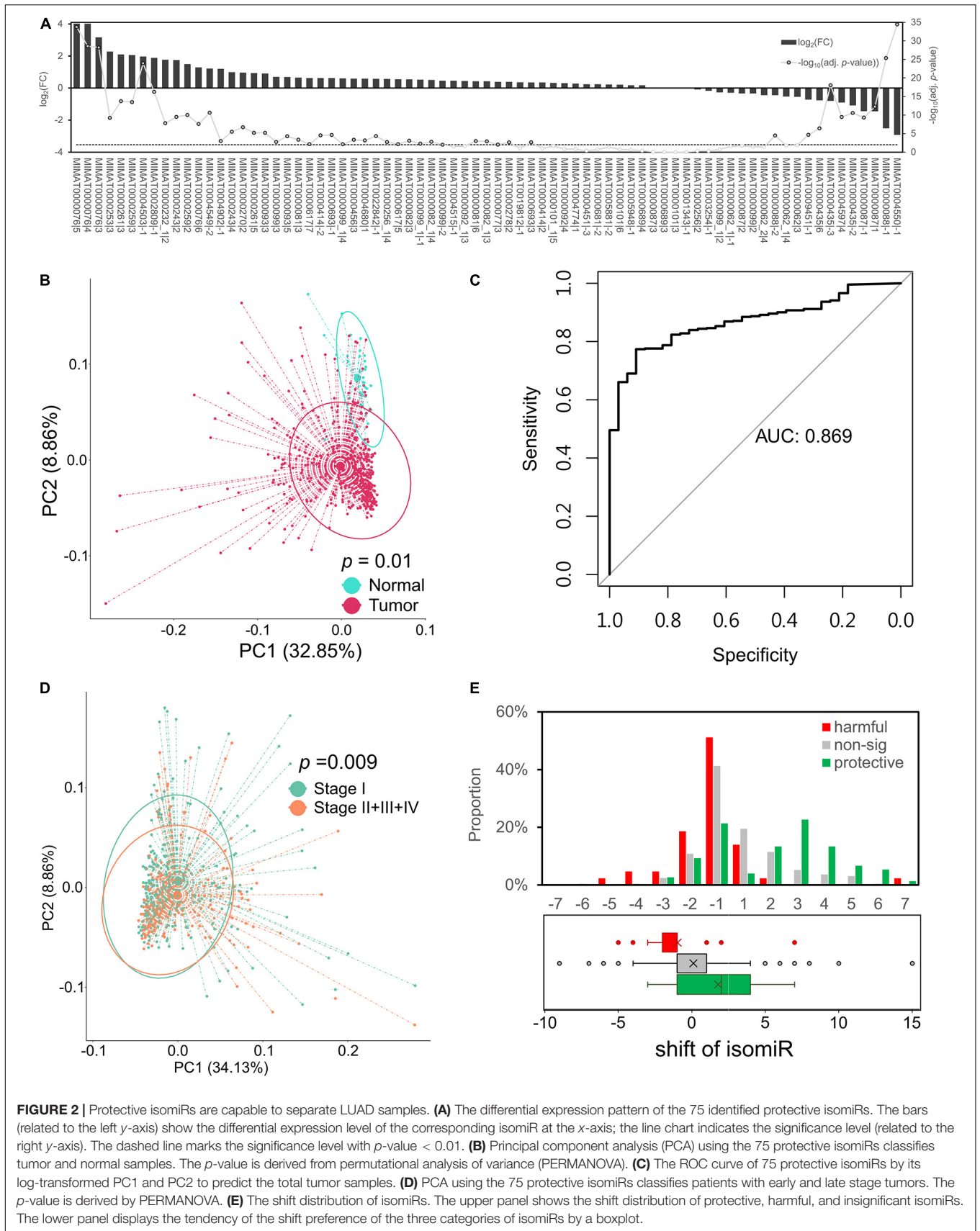
FIGURE 2 | Protective isomiRs are capable to separate LUAD samples. **(A)** The differential expression pattern of the 75 identified protective isomiRs. The bars (related to the left y-axis) show the differential expression level of the corresponding isomiR at the x-axis; the line chart indicates the significance level (related to the right y-axis). The dashed line marks the significance level with p-value < 0.01. **(B)** Principal component analysis (PCA) using the 75 protective isomiRs classifies tumor and normal samples. The p-value is derived from permutational analysis of variance (PERMANOVA). **(C)** The ROC curve of 75 protective isomiRs by its log-transformed PC1 and PC2 to predict the total tumor samples. **(D)** PCA using the 75 protective isomiRs classifies patients with early and late stage tumors. The p-value is derived by PERMANOVA. **(E)** The shift distribution of isomiRs. The upper panel shows the shift distribution of protective, harmful, and insignificant isomiRs. The lower panel displays the tendency of the shift preference of the three categories of isomiRs by a boxplot.

(by *glm* function in R) to predict the samples, the area under curve (AUC) reaches 0.86 (**Figure 2C**). These results suggested that the 75 protective isomiRs can classify the LUAD samples. The expression profiles of the 75 protective isomiRs are also distinguishable between patients of the early (I) and late (II + III + IV) stage (**Figure 2D**). The main difference between early and late stages is metastasis or not. Therefore, these 75 protective isomiRs might be involved in the regulation of tumor metastasis. Among these 75 protective isomiRs, 50 are positively shifted (downstream relative to archetype), and the remaining 25 are negatively shifted (upstream relative to archetype). Interestingly, we found that the protective isomiRs tend to possess positive shifts (p-value = $8.39*10^{-5}$, Fisher's exact test) and the harmful isomiRs negative shifts (p-value = $4.38*10^{-4}$, Fisher's exact test) (**Figure 2E**). Moreover, we observed that the protective isomiRs with downstream shift tend to be more upregulated than those with upstream shift. The average fold changes of upstream and downstream isomiRs are 1.56 and 1.01, respectively (p-value = 0.0564, Wilcoxon rank sum test, **Supplementary Figure 2**). Accordingly, these above results imply that the protective isomiRs with downstream shift might be selected to express against tumorigenesis during cancer development.
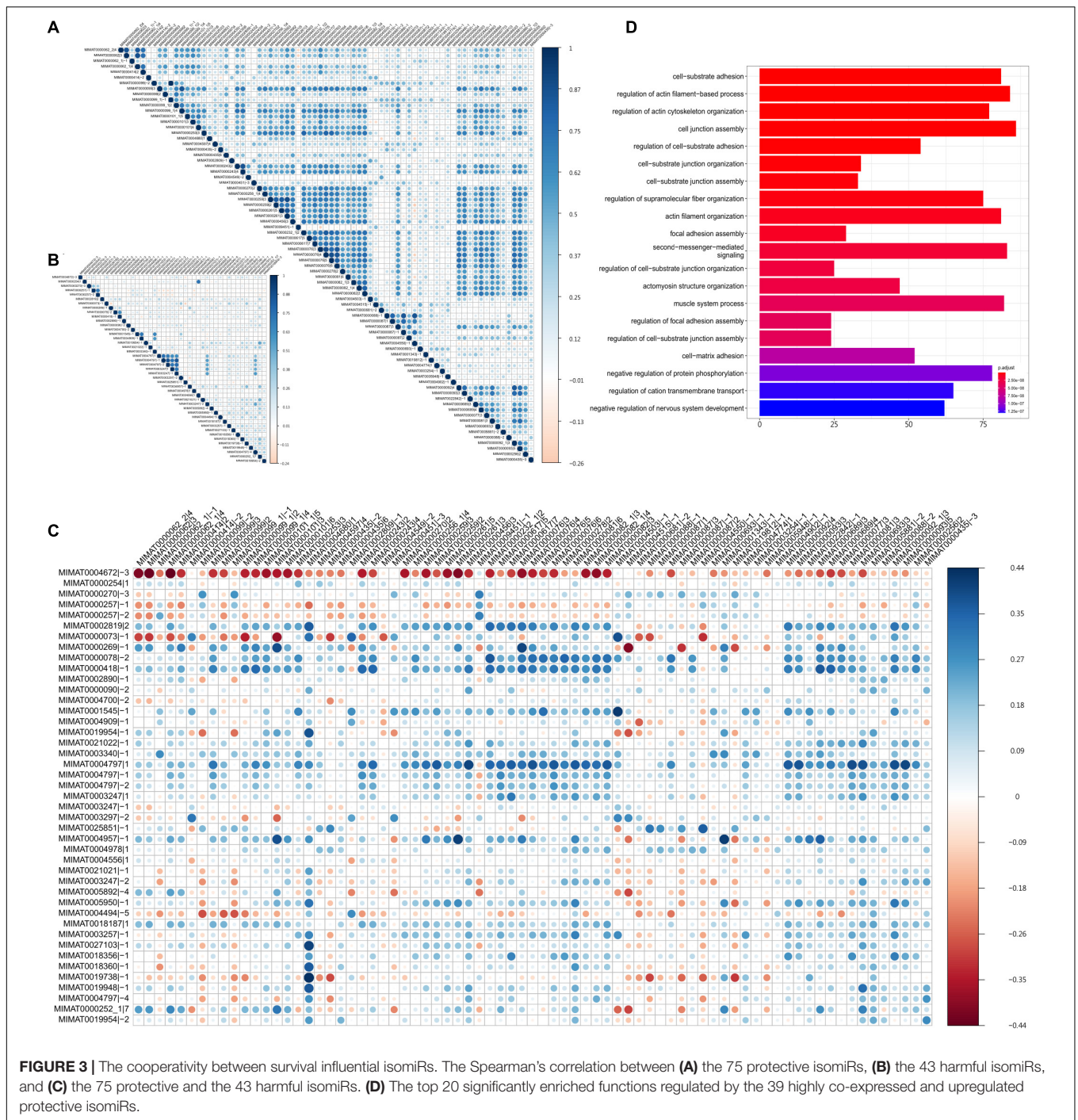
Positive Correlation Between the Protective isomiRs

In this study, we observed the prevalent positive correlation of expression profiles between the protective or harmful isomiRs—within-group (**Figures 3A,B**). Additionally, among the protective isomiRs, approximately 80% of pairs are significantly and positively correlated (z-score of Spearman's correlation coefficient > 2). This proportion is significantly overrepresented (p-value < 10^{-271}, Fisher's exact test). Among the harmful isomiRs, approximately 44% of pairs are significantly and positively correlated. This proportion is overrepresented, but the significance is moderate (p-value = 0.16, Fisher's exact test). These observations suggest the possible cooperativity between isomiRs with a similar influence on patient prognosis. Furthermore, the cooperativity between protective isomiRs might be stronger than that between harmful isomiRs. On the other hand, we found that the significant anti-correlations among either protective or harmful isomiRs are significantly underrepresented (protective: 2%, p-value = $5.68*10^{-40}$; harmful: 3%, p-value = $1.1*10^{-6}$, Fisher's exact). This observation further suggests that the isomiRs with similar influence on patient prognosis might not tend to be against each other.

On the other hand, we did not observe the concordance expression correlation between protective and harmful isomiRs (**Figure 3C**), suggesting that the cooperativities between protective and harmful isomiRs might not be prevalent. Furthermore, we found that the significantly negative correlations of expression profiles between protective and harmful isomiRs are significantly overrepresented (p-value = $1.38*10^{-5}$, Fisher's exact test), demonstrating the potential antagonistic relationship between protective and harmful isomiRs.

To further investigate the cooperative regulation of activated protective isomiRs, we focused on significantly 39 upregulated protective isomiRs. The functional enrichment analysis for the significantly downregulated target genes of these 39 protective isomiRs is shown in **Figure 3D**. We found that these target genes are significantly overrepresented in the cell migration-related functions, suggesting their roles in tumor metastasis. This observation further recapitulates the protective isomiRs' characteristic that can classify patients with early stage tumors from late stage very well. More importantly, the above results imply that the identified upregulated protective isomiRs might repress the genes involved in tumor metastasis to benefit patient prognosis.

The Bidirectional Regulation of Patient Prognosis

The interaction between isomiRs and their corresponding archetype miRNAs has been observed in lung cancer (Chan et al., 2013; Liang et al., 2020). More specifically, the isomiRs may disturb the classical regulatory network of their corresponding archetype miRNAs, even cause conflicting drug responses in cancers. In this study, we noticed, among the identified survival influential isomiRs, that MIMAT0000270|2 (hsa-miR-181a-3p|2) and MIMAT0000270|-3 (hsa-miR-181a-3p|-3) are originated from the same archetype—MIMAT0000270 (hsa-miR-181a-3p), but hsa-miR-181a-3p|2 is protective and hsa-miR-181a-3p|-3 is harmful. Interestingly, a previous study has observed that the hsa-miR-181a could function as an oncomiR or tumor suppressor in acute myeloid leukemia (Qiang et al., 2020). These observations suggest that hsa-miR-181a-3p could possess bidirectional regulation to patient prognosis. Moreover, the proportion of overlapped targets between these two isomiRs is small (0.02, Jaccard Index), emphasizing that these two isomiRs might regulate distinct downstream pathways. Additionally, these two isomiRs are upregulated in tumors (**Figure 4A**), suggesting that their regulatory downstream pathways might be repressed. We then performed functional enrichment analysis on their downregulated target genes separately (**Figures 4B,C**). We observed that the downregulated target genes of the protective hsa-miR-181a-3p|2 were significantly enriched in muscle cell or myotube functions (**Figure 4B**); the downregulated target genes of the harmful hsa-miR-181a-3p|-3 tend to be involved in fat cell differentiation and positive regulation of neuronal differentiation (**Figure 4C**). Previous studies have reported that muscle loss could be a significant predictor of patient mortality (Shiroyama et al., 2019; Rosa-Caldwell et al., 2020), supporting the protective role of hsa-miR-181a-3p|2 in lung cancer. In addition, we observed that the downregulated target genes of these two isomiRs are significantly enriched in the epithelial–mesenchymal transition (EMT)-associated functions (**Figures 4B,C**, marked by red). EMT is a reversible cell transition process, where cells could progressively lose the polarity, cell–cell junction, fixation, and finally turn into the mesenchymal morphology (Dongre and Weinberg, 2019). Evidence has shown that EMT was an epigenetic process, independent of DNA sequence alterations (Tam and Weinberg, 2013), and could

FIGURE 3 | The cooperativity between survival influential isomiRs. The Spearman's correlation between **(A)** the 75 protective isomiRs, **(B)** the 43 harmful isomiRs, and **(C)** the 75 protective and the 43 harmful isomiRs. **(D)** The top 20 significantly enriched functions regulated by the 39 highly co-expressed and upregulated protective isomiRs.

be significant during neoplasia, contributing to the malignant progression (Mani et al., 2008; Singh and Settleman, 2010), as well as in lung cancer (O'Leary et al., 2018). For example, the harmful isomiR—hsa-miR-181a-3p|-3—may regulate "regulation of actin cytoskeleton organization" and "protein localization to plasma membrane," pointing out that the dysregulation of its targeted genes might lose the function in fixation. On the other hand, the protective isomiR—hsa-miR-181a-3p|2—may also take part in EMT. The hsa-miR-181a-3p|2 may repress the function of

the glucocorticoid metabolic process that could promote tumor cell invasion and lung metastasis (Shi et al., 2019). The other functions, such as "regulation of establishment of cell polarity" and "skeletal muscle fiber development," may also indicate the role. The above observations suggest the potential bidirectional regulation of the hsa-miR-181a-3p. That is, these two isomiRs are derived from the same arm of pre-miRNA but demonstrate the opposite way in regulating tumor metastasis, and further patient prognosis.
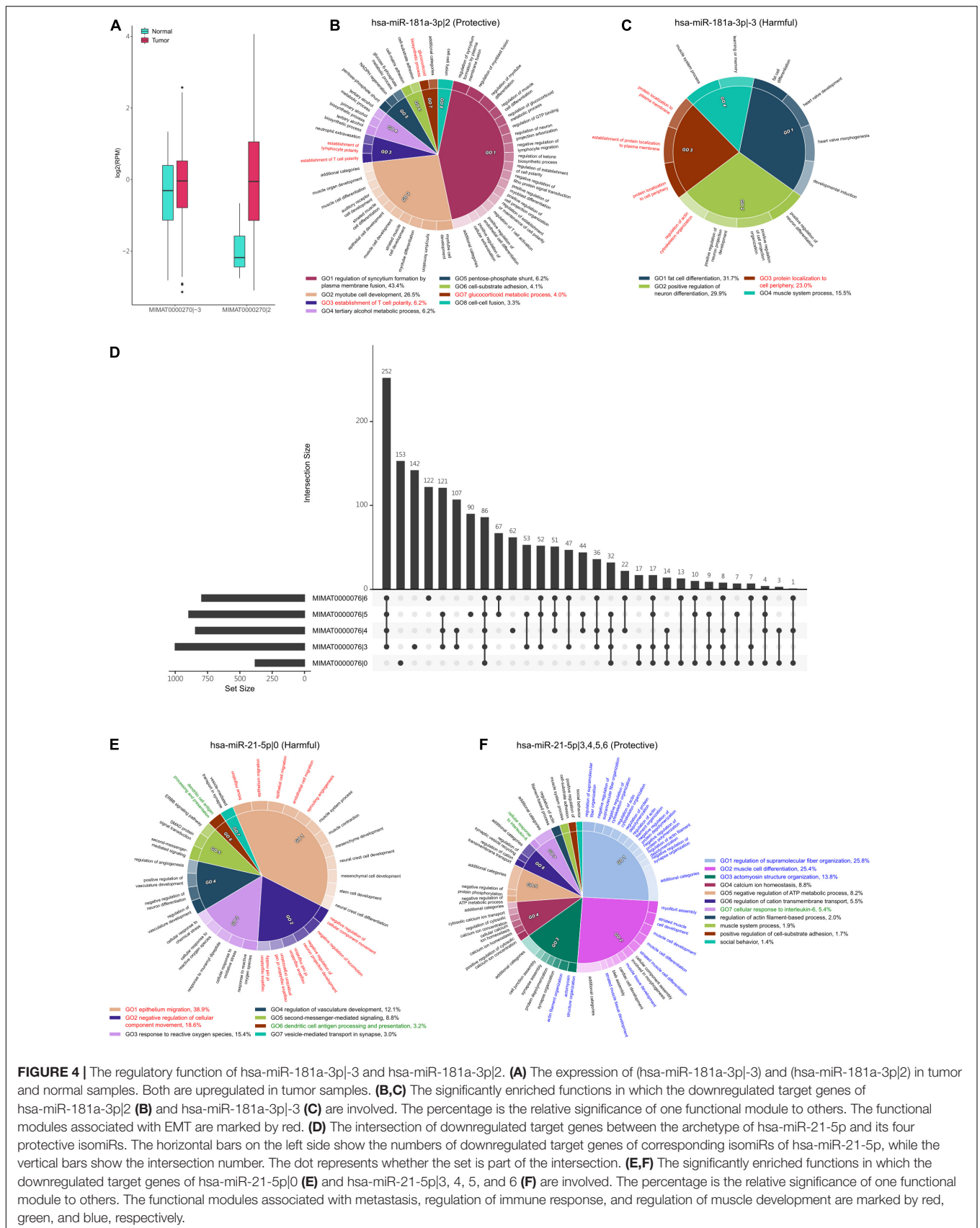
FIGURE 4 | The regulatory function of hsa-miR-181a-3p|-3 and hsa-miR-181a-3p|2. **(A)** The expression of (hsa-miR-181a-3p|-3) and (hsa-miR-181a-3p|2) in tumor and normal samples. Both are upregulated in tumor samples. **(B,C)** The significantly enriched functions in which the downregulated target genes of hsa-miR-181a-3p|2 **(B)** and hsa-miR-181a-3p|-3 **(C)** are involved. The percentage is the relative significance of one functional module to others. The functional modules associated with EMT are marked by red. **(D)** The intersection of downregulated target genes between the archetype of hsa-miR-21-5p and its four protective isomiRs. The horizontal bars on the left side show the numbers of downregulated target genes of corresponding isomiRs of hsa-miR-21-5p, while the vertical bars show the intersection number. The dot represents whether the set is part of the intersection. **(E,F)** The significantly enriched functions in which the downregulated target genes of hsa-miR-21-5p|0 **(E)** and hsa-miR-21-5p|3, 4, 5, and 6 **(F)** are involved. The percentage is the relative significance of one functional module to others. The functional modules associated with metastasis, regulation of immune response, and regulation of muscle development are marked by red, green, and blue, respectively.

In addition, miR-21 was found to be an oncogene in several cancers (Bica-Pop et al., 2018). In NSCLC, it was upregulated in patients (Yu et al., 2010; Yang et al., 2015), which also resembles our study. We observed that the archetype hsa-miR-21-5p (MIMAT0000076|0) is harmful to patients' prognoses. However, the hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6 are protective. Moreover, these four isomiRs are the top four upregulated isomiRs in the identified protective miRNAs. Interestingly, we found that the protective isomiRs, hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6, have the highest intersection of target genes, while the second-highest target gene set was exclusively presented in the archetype hsa-miR-21-5p|0 (**Figure 4D**). This observation suggests the possible cooperativity between these four isomiRs and divergent downstream regulation from the archetype—hsa-miR-21-5p. As expected, we observed that the archetype hsa-miR-21-5p|0 is largely involved in regulating tumor metastasis-associated functions (**Figure 4E**, marked as red). More specifically, the downregulated target genes of hsa-miR-21-5p|0 are significantly enriched in the negative regulation of cell migration and the negative regulation of cell motility. This observation suggests that the hsa-miR-21-5p|0 may promote tumor metastasis by repressing the negative regulation of cell migration or motility. Moreover, the hsa-miR-21-5p|0 might repress the function of dendritic cell antigen processing and presentation to help tumor progression (**Figure 4E**, marked as green). Interestingly, the intersection of downregulated target genes between the four isomiRs, hsa-miR-21-5p|3, 4, 5, and 6, is largely significantly enriched in the function of regulation of muscle development (**Figure 4F**, marked as blue). This observation shows the protective roles of these four isomiRs that might co-regulate the muscle development to control muscle loss and prolong patient survival (Shiroyama et al., 2019; Rosa-Caldwell et al., 2020). In addition, we found that these four isomiRs regulate the cellular response to interleukin-6 (IL-6) (**Figure 4F**, marked as green). Previous studies have demonstrated the IL-6 can promote lung cancer metastasis (Gross et al., 2018; Jiang et al., 2018; Liu et al., 2020). Accordingly, these four isomiRs might prevent the patients from tumor metastasis by repressing the function of IL-6. Briefly, these above results provide a possible insight in bidirectional regulation between archetype miRNA and its isomiRs: these four isomiRs might be against the regulation of archetype hsa-miR-21-5p to repress tumor metastasis.

DISCUSSION

Herein, we performed a standard pipeline to identify the survival influential miRNAs to patient prognosis. We found that the isomiRs are significantly enriched in the identified survival influential miRNAs, especially for the isomiRs associated with the better prognosis of patients, that is, they are protective. These protective isomiRs might be potential therapeutics for lung cancer. Previous studies have reported that hsa-miR-21-5p was upregulated consistently across cancers (Ferracin et al., 2015) and recognized as oncomiR in lung cancer

(Li and Wu, 2018; Yuan et al., 2018). Interestingly, in this study, the isomiRs originated from hsa-miR-21-5p are upregulated and associated with a better prognosis of patients: they might be tumor suppressors. IsomiRs possess different seed regions from the archetypes, suggesting that the downstream regulation could vary from an archetype to its isomiRs. Accordingly, the possible scenario is that these four isomiRs (hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6) might evolve tumor suppressor function during cancer development. Additionally, we noticed that the protective isomiRs tend to be provided with downstream shift. Moreover, these protective isomiRs with downstream shift are more likely to be upregulated than those with upstream shift. These results suggest that the isomiRs with downstream shift might be selected during carcinogenesis to prolong patient mortality. However, further experiments are required to validate this hypothesis.

On the other hand, we observed strong cooperative interactions within isomiRs with the same prognostic effect to patients, but antagonistic interactions between protective and harmful isomiRs. Furthermore, this cooperativity can be observed from the isomiRs originated from the same precursor and shared a large proportion of regulatory target genes. For example, hsa-miR-181b-5p|−1 and −2 are both harmful isomiRs, originate from hsa-miR-181b-5p, and share significantly overrepresented target genes (Jaccard Index = 0.45; p-value < 0.001, Fisher's exact test); hsa-miR-182-5p|2 and 3 are both protective isomiRs, originate from hsa-miR-182-5p, and share significantly overrepresented target genes (Jaccard Index = 0.4; p-value < 0.001, Fisher's exact test). On the other hand, the antagonistic interaction can also be observed from the small shared proportion of target genes between isomiRs, for example, hsa-miR-181a-3p|2 and hsa-miR-181a-3p|-3. This observation might further imply the synergistic effect between the isomiRs exerting a similar prognostic influence on patients. That is, the combination of multiple survival influential isomiRs could benefit or worsen patient prognosis more.

In addition, we identified that hsa-miR-21-5p|0 and its four isomiRs—hsa-miR-21-5p|3, 4, 5, and 6—could play critical roles in regulating tumor metastasis. More importantly, these four isomiRs are associated with a better prognosis of patients and might be against the harmful archetype: hsa-miR-21-5p. Among the downregulated target genes of these four protective isomiRs, *HER2*, *AKT*, and *DDR2* have been observed to be aberrantly expressed in NSCLC (Tsakonas and Ekman, 2018). Moreover, *DDR2* (Discoidin domain receptor tyrosine kinase 2) plays an important role in cellular connectivity, survival, migration, and cell proliferation (Bargal et al., 2009; Fathi et al., 2018), which contributed to therapeutic target in squamous cell lung cancer (Hammerman et al., 2011). Interestingly, we found that *DDR2* is the common downregulated target gene of hsa-miR-21-5p|3, 4, and 5, but not the archetype. This exclusive regulation of oncogene might support the proposed scenario, and these four isomiRs are against the regulation of archetype to repress cancer metastasis.

In summary, we identified survival influential miRNAs for lung cancer patients. We noticed that isomiRs are more associated with patients' mortality, especially for a better

prognosis than archetype miRNAs. Moreover, the expression profiles of the protective isomiR set can be a good predictor for classifying tumor and normal samples, as well as early and late stage patients. On the other hand, we observed the cooperativity between isomiRs with similar prognostic effects to patients and possible antagonistic interaction between protective and harmful isomiRs. In the end, we gave two examples—hsa-miR-181-3p and hsa-miR-21-5p—to demonstrate that the production of isomiRs could exert a distinct downstream regulatory effect, even opposite phenotypic change: better and poor prognoses.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The LUAD isoform miRNA expression profile as well as transcriptome expression profile ("isoforms.quantification.txt", "htseq.counts") for this study can be found in the TCGA-GDC porta (https://portal.gdc.cancer.gov/). The isomiR nomenclature was followed by miRBase (http://www.mirbase.org/), where the coordinate archetype miRNA ("hsa.gff3") downloaded from.

AUTHOR CONTRIBUTIONS

C-CL conceived, designed, and coordinated the study and revised the manuscript. F-MH, S-TL, M-FW, and C-CL implemented the computational method and carried out the analysis. F-MH and C-CL drafted the manuscript. All authors read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.702695/full#supplementary-material

REFERENCES

Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4:e05005. doi: 10.7554/eLife.05005

Andersen, P. K., and Gill, R. D. (1982). Cox's Regression Model for Counting Processes: a Large Sample Study. *Ann. Statist.* 10, 1100–1120. doi: 10.1214/aos/1176345976

Azizi, M. I. H. N., Othman, I., and Naidu, R. (2021). The Role of MicroRNAs in Lung Cancer Metabolism. *Cancers* 13:1716. doi: 10.3390/cancers13071716

Bargal, R., Cormier-Daire, V., Ben-Neriah, Z., Le Merrer, M., Sosna, J., Melki, J., et al. (2009). Mutations in DDR2 gene cause SMED with short limbs and abnormal calcifications. *Am. J. Hum. Genet.* 84, 80–84. doi: 10.1016/j.ajhg.2008.12.004

Bartel, D. P. (2009). MicroRNAs: target Recognition and Regulatory Functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Bartel, D. P. (2018). Metazoan MicroRNAs. *Cell* 173, 20–51. doi: 10.1016/j.cell.2018.03.006

Berindan-Neagoe, I., Monroig Pdel, C., Pasculli, B., and Calin, G. A. (2014). MicroRNAome genome: a treasure for cancer diagnosis and therapy. *CA Cancer J. Clin.* 64, 311–336. doi: 10.3322/caac.21244

Bica-Pop, C., Cojocneanu-Petric, R., Magdo, L., Raduly, L., Gulei, D., and Berindan-Neagoe, I. (2018). Overview upon miR-21 in lung cancer: focus on NSCLC. *Cell. Mol. Life Sci.* 75, 3539–3551. doi: 10.1007/s00018-018-2877-x

Chan, Y. T., Lin, Y. C., Lin, R. J., Kuo, H. H., Thang, W. C., Chiu, K. P., et al. (2013). Concordant and discordant regulation of target genes by miR-31 and its isoforms. *PLoS One* 8:e58169. doi: 10.1371/journal.pone.0058169

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163

Concato, J., Peduzzi, P., Holford, T. R., and Feinstein, A. R. (1995). Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J. Clin. Epidemiol.* 48, 1495–1501. doi: 10.1016/0895-4356(95)00510-2

Dela Cruz, C. S., Tanoue, L. T., and Matthay, R. A. (2011). Lung cancer: epidemiology, etiology, and prevention. *Clin. Chest Med.* 32, 605–644. doi: 10.1016/j.ccm.2011.09.001

Dongre, A., and Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 20, 69–84. doi: 10.1038/s41580-018-0080-4

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in Drosophila. *Genome Biol.* 5:R1. doi: 10.1186/gb-2003-5-1-r1

Esquela-Kerscher, A., and Slack, F. J. (2006). Oncomirs — microRNAs with a role in cancer. *Nat. Rev. Cancer* 6, 259–269. doi: 10.1038/nrc1840

Fathi, Z., Mousavi, S. A. J., Roudi, R., and Ghazi, F. (2018). Distribution of KRAS, DDR2, and TP53 gene mutations in lung cancer: an analysis of Iranian patients. *PLoS One* 13:e0200633. doi: 10.1371/journal.pone.0200633

Ferracin, M., Lupini, L., Salamon, I., Saccenti, E., Zanzi, M. V., Rocchi, A., et al. (2015). Absolute quantification of cell-free microRNAs in cancer patients. *Oncotarget* 6, 14545–14555. doi: 10.18632/oncotarget.3859

Gross, A. C., Cam, H., Phelps, D. A., Saraf, A. J., Bid, H. K., Cam, M., et al. (2018). IL-6 and CXCL8 mediate osteosarcoma-lung interactions critical to metastasis. *JCI Insight* 3:e99791. doi: 10.1172/jci.insight.99791

Guo, L., Liang, T., Yu, J., and Zou, Q. (2016). A Comprehensive Analysis of miRNA/isomiR Expression with Gender Difference. *PLoS One* 11:e0154955. doi: 10.1371/journal.pone.0154955

Hammerman, P. S., Sos, M. L., Ramos, A. H., Xu, C., Dutt, A., Zhou, W., et al. (2011). Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer Discov.* 1, 78–89. doi: 10.1158/2159-8274.CD-11-0005

Harrell, F. E. Jr., Lee, K. L., and Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* 15, 361–387. doi: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Iqbal, M. A., Arora, S., Prakasam, G., Calin, G. A., and Syed, M. A. (2019). MicroRNA in lung cancer: role, mechanisms, pathways and therapeutic relevance. *Mol. Aspects Med.* 70, 3–20. doi: 10.1016/j.mam.2018.07.003

Jiang, M., Wang, Y., Zhang, H., Ji, Y., Zhao, P., Sun, R., et al. (2018). IL-37 inhibits invasion and metastasis in non-small cell lung cancer by suppressing the IL-6/STAT3 signaling pathway. *Thorac. Cancer* 9, 621–629. doi: 10.1111/1759-7714.12628

John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004). Human MicroRNA targets. *PLoS Biol.* 2:e363. doi: 10.1371/journal.pbio.0020363

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39, 1278–1284. doi: 10.1038/ng2135

Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucl. Acids Res.* 47, D155–D162. doi: 10.1093/nar/gky1141

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucl. Acids Res.* 42, D68–D73. doi: 10.1093/nar/gkt1181

Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O., and Filipovska, A. (2019). CirGO: an alternative circular way of visualising gene ontology terms. *BMC Bioinform.* 20:84. doi: 10.1186/s12859-019-2671-2

Lai, E. C. (2002). Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* 30, 363–364. doi: 10.1038/ng865

Lee, L. W., Zhang, S., Etheridge, A., Ma, L., Martin, D., Galas, D., et al. (2010). Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA* 16, 2170–2180. doi: 10.1261/rna.2225110

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854. doi: 10.1016/0092-8674(93)90529-y

Li, X., and Wu, X. (2018). MiR-21-5p promotes the progression of non-small-cell lung cancer by regulating the expression of SMAD7. *Onco. Targets Ther.* 11, 8445–8454. doi: 10.2147/ott.S172393

Liang, T., Han, L., and Guo, L. (2020). Rewired functional regulatory networks among miRNA isoforms (isomiRs) from let-7 and miR-10 gene families in cancer. *Comput. Struct. Biotechnol. J.* 18, 1238–1248. doi: 10.1016/j.csbj.2020.05.001

Liu, W., Wang, H., Bai, F., Ding, L., Huang, Y., Lu, C., et al. (2020). IL-6 promotes metastasis of non-small-cell lung cancer by up-regulating TIM-4 via NF-kappaB. *Cell Prolif.* 53:e12776. doi: 10.1111/cpr.12776

Loher, P., Londin, E. R., and Rigoutsos, I. (2014). IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies. *Oncotarget* 5, 8790–8802. doi: 10.18632/oncotarget.2405

Lu, S., Sun, Y.-H., and Chiang, V. L. (2009). Adenylation of plant miRNAs. *Nucl. Acids Res.* 37, 1878–1885. doi: 10.1093/nar/gkp031

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715. doi: 10.1016/j.cell.2008.03.027

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi: 10.1093/nar/gks042

Mercey, O., Popa, A., Cavard, A., Paquet, A., Chevalier, B., Pons, N., et al. (2017). Characterizing isomiR variants within the microRNA-34/449 family. *FEBS Lett.* 591, 693–705. doi: 10.1002/1873-3468.12595

Morin, R. D., O'Connor, M. D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A. L., et al. (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 18, 610–621. doi: 10.1101/gr.7179508

Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., et al. (2019). New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* 15:e1006701. doi: 10.1371/journal.pcbi.1006701

Neilsen, C. T., Goodall, G. J., and Bracken, C. P. (2012). IsomiRs–the overlooked repertoire in the dynamic microRNAome. *Trends Genet.* 28, 544–549. doi: 10.1016/j.tig.2012.07.005

O'Leary, K., Shia, A., and Schmid, P. (2018). Epigenetic Regulation of EMT in Non-Small Cell Lung Cancer. *Curr. Cancer Drug Targets* 18, 89–96. doi: 10.2174/1568009617666170203162556

Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A., and Esyunina, D. M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry* 83, 483–497. doi: 10.1134/s0006297918050024

Peduzzi, P., Concato, J., Feinstein, A. R., and Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* 48, 1503–1510. doi: 10.1016/0895-4356(95)00048-8

Qiang, P., Pan, Q., Fang, C., Fozza, C., Song, K., Dai, Y., et al. (2020). MicroRNA-181a-3p as a diagnostic and prognostic biomarker for acute myeloid leukemia. *Mediterr. J. Hematol. Infect. Dis.* 12:e2020012. doi: 10.4084/MJHID.2020.012

Rehmsmeier, M., Steffen, P., Hochsmann, M., and Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *Rna* 10, 1507–1517. doi: 10.1261/rna.5248604

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 403, 901–906. doi: 10.1038/35002607

Resnik, P. (1999). Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language. *J. Artif. Int. Res.* 11, 95–130. doi: 10.1613/jair.514

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:R25. doi: 10.1186/gb-2010-11-3-r25

Rosa-Caldwell, M. E., Fix, D. K., Washington, T. A., and Greene, N. P. (2020). Muscle alterations in the development and progression of cancer-induced muscle atrophy: a review. *J. Appl. Physiol.* 128, 25–41. doi: 10.1152/japplphysiol.00622.2019

Shah, M. Y., Ferrajoli, A., Sood, A. K., Lopez-Berestein, G., and Calin, G. A. (2016). microRNA Therapeutics in Cancer - An Emerging Concept. *Ebiomedicine* 12, 34–42. doi: 10.1016/j.ebiom.2016.09.017

Shi, W., Wang, D., Yuan, X., Liu, Y., Guo, X., Li, J., et al. (2019). Glucocorticoid receptor-IRS-1 axis controls EMT and the metastasis of breast cancers. *J. Mol. Cell Biol.* 11, 1042–1055. doi: 10.1093/jmcb/mjz001

Shiroyama, T., Nagatomo, I., Koyama, S., Hirata, H., Nishida, S., Miyake, K., et al. (2019). Impact of sarcopenia in patients with advanced non-small cell lung cancer treated with PD-1 inhibitors: a preliminary retrospective study. *Sci. Rep.* 9:2447. doi: 10.1038/s41598-019-39120-6

Singh, A., and Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* 29, 4741–4751. doi: 10.1038/onc.2010.215

Sun, Y. M., Lin, K. Y., and Chen, Y. Q. (2013). Diverse functions of miR-125 family in different cell contexts. *J. Hematol. Oncol.* 6:6. doi: 10.1186/1756-8722-6-6

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800

Tam, W. L., and Weinberg, R. A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.* 19, 1438–1449. doi: 10.1038/nm.3336

Tan, G. C., Chan, E., Molnar, A., Sarkar, R., Alexieva, D., Isa, I. M., et al. (2014). 5′ isomiR variation is of functional and evolutionary importance. *Nucleic Acids Res.* 42, 9424–9435. doi: 10.1093/nar/gku656

Telonis, A. G., Loher, P., Jing, Y., Londin, E., and Rigoutsos, I. (2015). Beyond the one-locus-one-miRNA paradigm: microRNA isoforms enable deeper insights into breast cancer heterogeneity. *Nucleic Acids Res.* 43, 9158–9175. doi: 10.1093/nar/gkv922

Telonis, A. G., Magee, R., Loher, P., Chervoneva, I., Londin, E., and Rigoutsos, I. (2017). Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* 45, 2973–2985. doi: 10.1093/nar/gkx082

Therneau, T. (2015). *A Package for Survival Analysis in S. version 2*. 38. Available online at: https://CRAN.R-project.org/package=survival (accessed November 17, 2020).

Therneau, T. M., Grambsch, and Patricia, M. (2000). *Modeling Survival Data: extending the Cox Model*. New York: Springer-Verlag.

Tsakonas, G., and Ekman, S. (2018). Oncogene-addicted non-small cell lung cancer and immunotherapy. *J. Thorac. Dis.* 10, S1547–S1555. doi: 10.21037/jtd.2018.01.82

Wu, K.-L., Tsai, Y.-M., Lien, C.-T., Kuo, P.-L., Hung, J. L., and Jen, Y. (2019). The Roles of MicroRNA in Lung Cancer. *Int. J. Mol. Sci.* 20:1611. doi: 10.3390/ijms20071611

Xu, L. L., Shi, C. M., Xu, G. F., Chen, L., Zhu, L. L., Zhu, L., et al. (2014). TNF-α, IL-6, and leptin increase the expression of miR-378, an adipogenesis-related microRNA in human adipocytes. *Cell Biochem. Biophys.* 70, 771–776. doi: 10.1007/s12013-014-9980-x

Yang, J. S., Li, B. J., Lu, H. W., Chen, Y., Lu, C., Zhu, R. X., et al. (2015). Serum miR-152, miR-148a, miR-148b, and miR-21 as novel biomarkers in non-small

cell lung cancer screening. *Tumour Biol.* 36, 3035–3042. doi: 10.1007/s13277-014-2938-1

Yu, L., Todd, N. W., Xing, L., Xie, Y., Zhang, H., Liu, Z., et al. (2010). Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. *Int. J. Cancer* 127, 2870–2878. doi: 10.1002/ijc.25289

Yuan, Y., Xu, X. Y., Zheng, H. G., and Hua, B. J. (2018). Elevated miR-21 is associated with poor prognosis in non-small cell lung cancer: a systematic review and meta-analysis. *Eur Rev Med Pharmacol. Sci.* 22, 4166–4180. doi: 10.26355/eurrev_201807_15410

Zappa, C., and Mousa, S. A. (2016). Non-small cell lung cancer: current treatment and future advances. *Transl. Lung Cancer Res.* 5, 288–300. doi: 10.21037/tlcr.2016.06.07

Zhong, S., Golpon, H., Zardo, P., and Borlak, J. (2021). miRNAs in lung cancer. A systematic review identifies predictive and prognostic miRNA candidates for precision medicine in lung cancer. *Transl. Res.* 230, 164–196. doi: 10.1016/j.trsl.2020.11.012