



Article

Post-Transcriptional Modifications to miRNAs Undergo Widespread Alterations, Creating a Unique Lung Adenocarcinoma IsomiRome

David E. Cohn¹, Vanessa G. P. Souza^{1,*} , Aisling Forder^{1,†}, Nikita Telkar^{1,2} , Greg L. Stewart¹ and Wan L. Lam¹

¹ Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC V5Z 1L3, Canada; wanlam@bccrc.ca (W.L.L.)

² British Columbia Children's Hospital Research Institute, Vancouver, BC V5Z 4H4, Canada

* Correspondence: vsouza@bccrc.ca

† These authors contributed equally to this work.

Simple Summary: MicroRNAs (miRNAs) play a significant role as epigenetic regulators in cancer. IsomiRs are miRNA molecules that undergo small modifications during miRNA processing, which can affect their stability and their interaction with mRNA targets. While some isomiRs are linked to specific cancers, many, including those in the lung, remain understudied. To address this, small RNA sequencing data from lung adenocarcinoma (LUAD) and adult non-malignant lung (ANL) samples were analyzed to quantify isomiR expression. This analysis identified 16 A-to-I edited isomiRs, 213 5' isomiRs, 128 3' adenylated isomiRs, and 100 3' uridylated isomiRs. A-to-I editing rates correlated with the expression of editing enzymes ADAR and ADARB1, both deregulated in LUAD. LUAD samples had lower A-to-I editing and 3' adenylation rates compared to ANL. Machine learning models based on isomiR data effectively distinguished ANL from stage I/II LUAD samples, suggesting that isomiRs hold potential as cancer biomarkers.

Abstract: Background: MicroRNAs (miRNAs) modulate the expression of oncogenes and tumor suppressor genes, functioning as significant epigenetic regulators in cancer. IsomiRs are miRNA molecules that have undergone small modifications during miRNA processing. These modifications can alter an isomiR's binding stability with mRNA targets, and certain isomiRs have been implicated in the development of specific cancers. Still, the isomiRomes of many tissues, including the lung, have not been characterized; Methods: In this study, we analyzed small RNA sequencing data for three cohorts of lung adenocarcinoma (LUAD) and adult non-malignant lung (ANL) samples. Results: We quantified isomiR expression and found 16 A-to-I edited isomiRs expressed in multiple cohorts, as well as 213 5' isomiRs, 128 3' adenylated isomiRs, and 100 3' uridylated isomiRs. Rates of A-to-I editing at editing hotspots correlated with mRNA expression of the editing enzymes ADAR and ADARB1, which were both observed to be deregulated in LUAD. LUAD samples displayed lower overall rates of A-to-I editing and 3' adenylation than ANL samples. Support vector machines and random forest models were trained on one cohort to distinguish ANL and stage I/II LUAD samples using reads per million (RPM) and frequency data for different types of isomiRs. Models trained on A-to-I editing rates at editing hotspots displayed high accuracy when tested on the other two cohorts and compared favorably to classifiers trained on miRNA expression alone; Conclusions: We have identified isomiRs in the human lung and found that their expression differs between non-malignant and tumor tissues, suggesting they hold potential as cancer biomarkers.

Keywords: isomiR; miRNA; lung adenocarcinoma; ADAR; RNA editing



Citation: Cohn, D.E.; Souza, V.G.P.; Forder, A.; Telkar, N.; Stewart, G.L.; Lam, W.L. Post-Transcriptional Modifications to miRNAs Undergo Widespread Alterations, Creating a Unique Lung Adenocarcinoma IsomiRome. *Cancers* **2024**, *16*, 3322. <https://doi.org/10.3390/cancers16193322>

Received: 2 September 2024

Revised: 24 September 2024

Accepted: 26 September 2024

Published: 28 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

MicroRNAs (miRNAs) are a highly expressed subtype of endogenous small non-coding RNAs that canonically downregulate genes through binding to partially comple-

mentary regions of mRNA and either inhibiting translation or inducing mRNA degradation [1–3]. miRNA dysregulation is a common mechanism through which tumors deregulate critical pathways, and many miRNAs are oncogenes or tumor suppressor genes for certain cancer types [4]. Furthermore, the small size of miRNAs renders them highly stable, and consequently, any miRNAs that are deregulated in tumors are appealing candidate diagnostic biomarkers [5]. These factors have driven interest in the characterization of tumor miRNomes, particularly for cancers such as lung adenocarcinoma (LUAD), which often presents without actionable driver mutations [6].

However, the study of the miRNomes of LUAD and other tumors has been complicated by the fact that most miRNA-coding loci generate not only an unmodified (“canonical”) miRNA but also distinct modified miRNA molecules known as isomiRs [7]. The sequence modifications that distinguish isomiRs arise from a variety of sources, including the cleavage of miRNA precursor molecules by Drosha or Dicer at non-canonical sites, the adenosine deaminase activity of ADAR-family (adenosine deaminase, RNA-specific) proteins, and the activity of terminal nucleotidyltransferase proteins [7] (Figure S1). IsomiRs typically interact with gene-silencing machinery in the same way as canonical miRNAs [8], but their sequence changes mean that they do not necessarily target the same mRNAs [9].

The ubiquity of miRNA deregulation in cancer has motivated research into whether tumors also frequently alter the expression of isomiRs. Several adenosine-to-inosine (A-to-I) edited isomiRs are dysregulated in various forms of cancer, causing downstream changes in gene expression that benefit tumors [10,11]. Different tissues are known to have distinct patterns of isomiR expression [12], which suggests that tumorigenesis in different tissues is likely associated with unique changes to their isomiR transcriptomes (isomiRomes). The characterization of the isomiRomes of specific tumors, such as LUAD, is thus of interest, especially as miRNA deregulation has already been extensively tied to LUAD tumor growth and therapeutic resistance [13–18]. Understanding how the lung isomiRome is altered in LUAD will provide further insight into how LUAD tumors orchestrate and benefit from miRNA deregulation and could enable the improvement of existing miRNA-based biomarker panels through the incorporation of informative isomiRs.

RNA modifications, including A-to-I editing, are associated with a range of lung diseases, particularly lung cancer [19]. Investigations into A-to-I edited isomiRs have utilized data from The Cancer Genome Atlas (TCGA), a major and comprehensive cancer genomics resource [20]. These studies have pinpointed editing sites that distinguish between adult non-malignant lung (ANL) and lung adenocarcinoma (LUAD) tissues and have linked these sites to patient survival [21,22]. Further validation was provided by a TCGA lung-specific study, which identified distinct mRNA targets for edited isomiRs [23]. Additionally, A-to-I edited isomiRs were detected in the British Columbia Cancer Agency (BCCA) lung cohort [24]. These findings were corroborated by a Japanese lung cohort study, which found that reduced editing of miR-99a was associated with patient survival [24].

While these studies have made significant contributions to our understanding of isomiRs in lung cancer, there are areas where further investigation could enhance this understanding. Many of these studies primarily relied on TCGA data, and while TCGA data are comprehensive, validation across additional diverse patient groups is important, given that isomiR expression can vary based on patient characteristics such as race [25]. Moreover, the focus has largely been on A-to-I edited isomiRs, leaving other major types, such as 5' modified, 3' adenylated, and 3' uridylated isomiRs, less explored in the context of lung cancer. Understanding these isomiRs is crucial, as they may have unique functions that could provide deeper insights into disease mechanisms [25]. For example, 5' modified isomiRs may target different mRNAs compared to their canonical counterparts [26], while 3' uridylated and adenylated isomiRs have been shown to influence gene regulation and miRNA stability [27–29].

To expand on earlier discoveries, we conducted a comprehensive study examining the different types of isomiRs across three distinct lung cohorts. We identified differences in isomiR expression between ANL and LUAD tissues, both overall and at the individ-

ual isomiR level. Additionally, we investigated potential causes of these differences by comparing the expression level of key miRNA-modifying genes between ANL and LUAD and analyzing their correlations with isomiR expression. Furthermore, we explored the relationship between isomiR expression and patient outcomes and assessed the potential of isomiRs as biomarkers for stage I/II LUAD.

2. Materials and Methods

2.1. Sample Collection, Sequencing, and Data Processing

Three lung cohorts were used in this study: the British Columbia Cancer Agency (BCCA) internal cohort containing ANL and LUAD samples, The Cancer Genome Atlas LUAD external cohort (TCGA-LUAD, hereafter TCGA) of ANL and LUAD samples, and the Ewha Womans University (EWU) external cohort, which consists of pairs of ANL and LUAD samples. Clinical characteristics are presented in Table 1.

Table 1. Clinical characteristics of patients from the three lung adenocarcinoma (LUAD) cohorts for small RNA sequence analyses.

Characteristic	EWU (<i>n</i> = 48)	TCGA (<i>n</i> = 389)	BCCA (<i>n</i> = 63)
Median Age (Range)	59.5 (37–78)	66 (39–88)	70 (45–86)
Sex			
Male	0 (0%)	173 (44%)	19 (30%)
Female	48 (100%)	216 (56%)	44 (70%)
Smoking History			
Current or Former	7 (15%)	307 (79%)	40 (63%)
Never	41 (85%)	64 (16%)	23 (37%)
Stage			
IA	23 (48%)	102 (26%)	23 (37%)
IB	7 (15%)	103 (26%)	18 (29%)
IIA	5 (10%)	43 (11%)	2 (3%)
IIB	1 (2%)	52 (13%)	11 (17%)
IIIA	12 (25%)	57 (15%)	4 (6%)
IIIB	0 (0%)	6 (2%)	1 (2%)
IV	0 (0%)	17 (4%)	1 (2%)

Legend: BCCA—British Columbia Cancer Agency; EWU—Ewha Womans University; TCGA—The Cancer Genome Atlas; *n*—number.

For mRNA expression data, a total of 76 pairs of ANL and LUAD samples within the BCCA cohort were available for the study [30]. However, fewer LUAD samples (*n* = 63) were sequenced for the small RNA data. The 76 paired ANL and LUAD samples were obtained with informed and written consent from patients at Vancouver General Hospital and with approval from the University of British Columbia/BC Cancer Agency Research Ethics Board (H15-03060), as previously described [31]. After reviewing the histology of the tissue samples, a lung pathologist guided the microdissection process to ensure that ANL and LUAD samples had >80% non-malignant and tumor cell content, respectively. Following this, TRIzol reagent was used to extract total RNA from the frozen sections of samples of the BCCA cohort (Life Technologies, Carlsbad, CA, USA). The small RNA sequencing protocol on the Illumina HiSeq2000 platform (Illumina Inc., San Diego, CA, USA) at Canada's Michael Smith Genome Sciences Centre has been previously described [32]. Small RNA sequencing files were obtained as Binary Alignment Map (BAM) files and converted to unaligned FASTQ files [32]. miRNA expression data for samples in the BCCA cohort were deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GSE175462) [33]. The mRNA expression data for 76 pairs of ANL and LUAD samples in the BCCA cohort were generated using HumanWG-6 Gene Expression BeadChips (Illumina Inc., San Diego, CA, USA), as previously described [30]. Data downloaded from Ensembl Release 104, using Ensembl BioMart [34], were used to convert between probe names and the corresponding Ensembl gene stable IDs or gene names, where necessary [35]. In cases

where a gene had multiple matching probes, the probe corresponding to the highest mean expression level was used.

TCGA miRNA sequencing data were available for 38 ANL samples and 395 LUAD samples, representing 389 LUAD patients after averaging multiple tumor sampling. mRNA sequencing data for the TCGA cohort were accessible for 59 ANL samples and 513 LUAD patients. After averaging multiple tumor samples, 57 paired samples were identified. All data were downloaded as Binary Alignment Map (BAM) files from the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>, accessed on 11 July 2019). The TCGA samples were sequenced on the Illumina HiSeq2500 platform, as previously described [36]. mRNA expression data and patient clinical information, including survival data, were obtained from the University of California Santa Cruz (UCSC) Xena browser (<https://xenabrowser.net/>, accessed on 11 July 2019).

The EWU external cohort consists of 48 pairs of ANL and LUAD samples collected at the Samsung Medical Center. miRNA sequencing data for the EWU cohort were downloaded as FASTQ files from the European Nucleotide Archive (Study Accession: PRJNA434883) (<https://www.ebi.ac.uk/ena/browser/home>, accessed on 12 February 2021) [37]. All EWU samples were sequenced on the Illumina HiSeq2000 platform [37]. Clinical data for this cohort were obtained through the Sequence Read Archive run selector (SRA Accession: SRP133217) (<https://www.ncbi.nlm.nih.gov/sra>, accessed on 12 February 2021). mRNA expression and survival data were not publicly available. Samples with low coverage (<5 million reads) were excluded from all miRNA- and isomiR-level analyses for all cohorts.

2.2. Processing of Small RNA Sequencing Data

The 3' adapter sequences were trimmed from the EWU cohort FASTQ files using the same algorithm that had been applied to the BCCA and TCGA cohorts (described in [36]). The algorithm was modified to accept FASTQ files as input. For all three cohorts, reads with any positions at which the Phred quality score was below 25 were removed using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/, accessed on 1 March 2022). Adapter-trimmed, quality-filtered files were then processed by miRMaster 2.0, which aligned reads to the hg38 assembly and quantified isomiR expression [38]. All default settings were used, apart from the following changes: Protocol = Custom; 3' Adapter = ATCTCGTATGCCGTCTTCTGCTTGT (for the BCCA and TCGA cohorts) or TGGAATTCTCGGGTGCCAAGG (for the EWU cohort); Minimum read length = 15; Sliding window required quality = 1; Alignment tool = Bowtie; Maximum allowed distance for 3' quantification = 2; and Minimum read stack height = 100. Following miRMaster processing, isomiRs with multiple substitutions relative to their corresponding mature miRNA were discarded.

2.3. IsomiR and miRNA Quantification and Nomenclature

For each type of isomiR being studied (5' modified, 3' adenylated, 3' uridylylated, and A-to-I edited), isomiRs of the same miRNA that differed only in non-relevant positions were combined. For example, when quantifying 5' isomiRs, isomiRs that had identical 5' nucleotide additions or deletions were combined, regardless of their 3' variation. Similarly, when quantifying 3' adenylated or uridylylated isomiRs, isomiRs that had the same 3' non-templated addition (NTA) were combined, regardless of any 5' variation. Additionally, for consistency with past studies of miRNA editing, the final two bases of each isomiR were ignored when quantifying edited isomiRs, and edited isomiRs that differed only in those positions were combined [39]. miRNA expression was quantified by combining the reads corresponding to each canonical miRNA with those of all isomiRs that corresponded to the miRNA. All isomiR names begin with the name of their corresponding canonical miRNA. For 5' isomiRs, the miRNA name is then followed by a number indicating the position of the isomiR's 5' end relative to the miRNA's 5' end. A-to-I edited isomiR names have the same format, with an additional number that indicates the position of the edited adenosine

(with 0 corresponding to the isomiR's first nucleotide). 3' adenylated and uridylated isomiR names contain a single number indicating the position of the adenylation (either 1 or 2) relative to the miRNA's 3' end, followed by an "A" or "U".

2.4. Identification of High-Confidence miRNAs and isomiRs

First, isomiRs with substitutions or non-templated additions (NTAs) that matched single nucleotide polymorphisms listed in miRNASNP-v3 (<http://bioinfo.life.hust.edu.cn/miRNASNP/>, accessed on 1 August 2021) [40] or that matched lung cancer somatic mutations listed in COSMIC v94 (<https://cancer.sanger.ac.uk/cosmic>, accessed on 1 August 2021) [41] were discarded. For an isomiR to be considered expressed in a given sample group, it had to meet two criteria, each in $\geq 10\%$ of the group's samples: (a) expression at ≥ 1 RPM and (b) expression at a sufficiently high rate, relative to similar sequences, to distinguish it from a sequencing error. miRNAs were considered expressed if they met criterion (a) only. Criterion (b) was applied using a widely used method first described by Alon et al. for the detection of miRNA editing sites: briefly, the count of an isomiR's reads and the count of otherwise identical reads that lacked the isomiR's characteristic edit/3' NTA/5' variation were input into a binomial cumulative distribution, which used the baseline sequencing error rate to estimate the likelihood that all of the isomiR's reads had arisen from errors when sequencing similar sequences [39]. This sequencing error hypothesis was rejected in a given sample if this likelihood was < 0.05 after Benjamini–Hochberg (BH) correction and there were at least three reads of the isomiR in question. Because small RNA sequencing data are prone to plate-based batch effects [42], the baseline sequencing error rate for substitutions was taken as the greater of the expected rate (Phred 25 = 0.32%) and the observed rate of mismatches of the same type in the same position of a read in samples sequenced on the same plate. For 5', 3' adenylated, and 3' uridylated isomiRs, the error rate was very conservatively set to 2%.

Expressed isomiRs were then discarded if they cross-mapped to other regions of the genome, as this indicated that their sequences may not have arisen through modification of a canonical miRNA transcript. Specifically, an isomiR was discarded if its most highly expressed sequence mapped, using Bowtie [43], to any other genomic site as well as it mapped to the genomic region of its miRNA stem-loop. Since transfer RNAs (tRNAs) receive post-transcriptional 3' additions of CCA [44], isomiRs ending in CCA were also discarded if they cross-mapped to a tRNA-coding genomic region after the removal of those three bases. Finally, an isomiR or miRNA was deemed high-confidence if it was expressed in at least one of the two sample groups (ANL or LUAD) for at least two of the three adult lung cohorts.

2.5. Computation of Sample-Wide miRNA Editing/Adenylation/Uridylation Rates

The frequency of an isomiR was defined as the ratio of its expression to the combined expression of its corresponding canonical miRNA and all of that miRNA's isomiRs. For each sample, a z-score was calculated for each high-confidence isomiR, indicating how its frequency in that sample varied in relation to its frequency in all samples from all three cohorts. The overall rate of a miRNA modification (e.g., 3' adenylation) in a sample was then calculated as the mean z-score for all high-confidence isomiRs of that type in that sample. IsomiRs with an undefined z-score in certain samples, due to those samples having no expression of either the isomiR or the corresponding miRNA, were excluded when the mean z-score for those samples was calculated.

2.6. Determination of Significant Differences between ANL and LUAD Samples

Paired-sample *t*-tests were used to assess whether various quantities (i.e., the expression of miRNA-modifying enzymes or the overall rates of miRNA modifications) in each cohort differed between paired ANL and LUAD samples, with a significance threshold of $p < 0.05$. When determining whether the frequencies of individual isomiRs differed

between ANL and LUAD samples, a false discovery rate (FDR) threshold < 0.05 after BH correction was used instead.

2.7. Analysis of Correlations between Variables

Correlations between the expression of miRNA-modifying enzymes and the rates of miRNA modifications or the frequencies of specific isomiRs were evaluated only in LUAD samples. A correlation was considered significant if Spearman's ρ was different from zero, with a BH-FDR threshold of < 0.05 .

2.8. Survival Analyses

For all survival analyses, LUAD patients in the analyzed cohort were divided into groups based on having above-median or below-median values of the quantity being analyzed. The quantity was deemed to impact outcomes if the difference in overall survival between the two groups was significant, as assessed by a log-rank test ($p < 0.05$).

2.9. Support Vector Machine and Random Forest Classifiers

All classifiers were trained and tested only on ANL and stage I/II LUAD samples, using either miRNA RPM values or either RPM or frequency values for a single type of isomiR. All data were standardized for support vector machine (SVM) classifiers, and 300 trees were grown for each random forest (RF) classifier. isomiRs that had undefined frequencies in any training or testing sample were excluded from frequency-based classifiers.

For classifiers designed to discriminate between ANL and LUAD samples from a single cohort, data for all miRNAs or isomiRs of the chosen type that were expressed in the cohort were used, including those that were not high-confidence. The classification error for SVMs was calculated as the mean classification error of trained models during 10-fold cross-validation (CV), and the classification error for RFs was calculated as the ensemble out-of-bag error.

Classifiers that were designed to test the generalizability of isomiR-based biomarkers were always trained on the largest (TCGA) cohort. Data were used for all miRNAs or isomiRs of the chosen type that were expressed in the TCGA cohort, and that were differentially expressed (for RPM-based classifiers) or differentially frequent (for frequency-based classifiers) between TCGA ANL and TCGA stage I/II LUAD samples. Model hyperparameters were optimized using MATLAB's Bayesopt function. BoxConstraint and KernelScale were optimized for SVMs by minimizing the mean classification error during a 5-fold CV. For the RFs, the minimum leaf size and the percentage of features that were made available to the classifier at each node were optimized, with maximum values of 20 and 70%, respectively, by minimizing the out-of-bag error. Trained, optimized models were tested on the BCCA and EWU cohorts.

2.10. Statistical and Graphical Software

All statistical analysis was performed in MATLAB R2017b (MathWorks, Natick, MA, USA). Except where otherwise specified, all illustrations were also generated in MATLAB.

3. Results

3.1. Characterization of the Lung isomiRome

We found that similar numbers of 5' isomiRs, 3' adenylated isomiRs, and A-to-I edited isomiRs were expressed in each of the three lung cohorts, while there was an elevated number of 3' uridylylated isomiRs in the EWU cohort (Figure 1A). In total, there were 16 A-to-I edited isomiRs, 213 5' isomiRs, 128 3' adenylated isomiRs, 100 3' uridylylated isomiRs, and 654 miRNAs that were expressed in multiple cohorts, and thus deemed "high-confidence" (Figure 1A). Of these 16 edited isomiRs, 13 (81%) had editing sites that were located within their seed sequence (Figure 1B and Table 2). Twelve (75%) of the canonical miRNAs had an uracil immediately 5' of the edited adenine, in line with the known sequence preferences of the editing enzymes ADAR and ADARB1 (Figure 1C and Table 2). IsomiRs with the 3'

NTA of cytosine or guanine were observed in all cohorts and tissue types, but they were far less common than 3' adenylated and uridylylated isomiRs (Figure 1D).

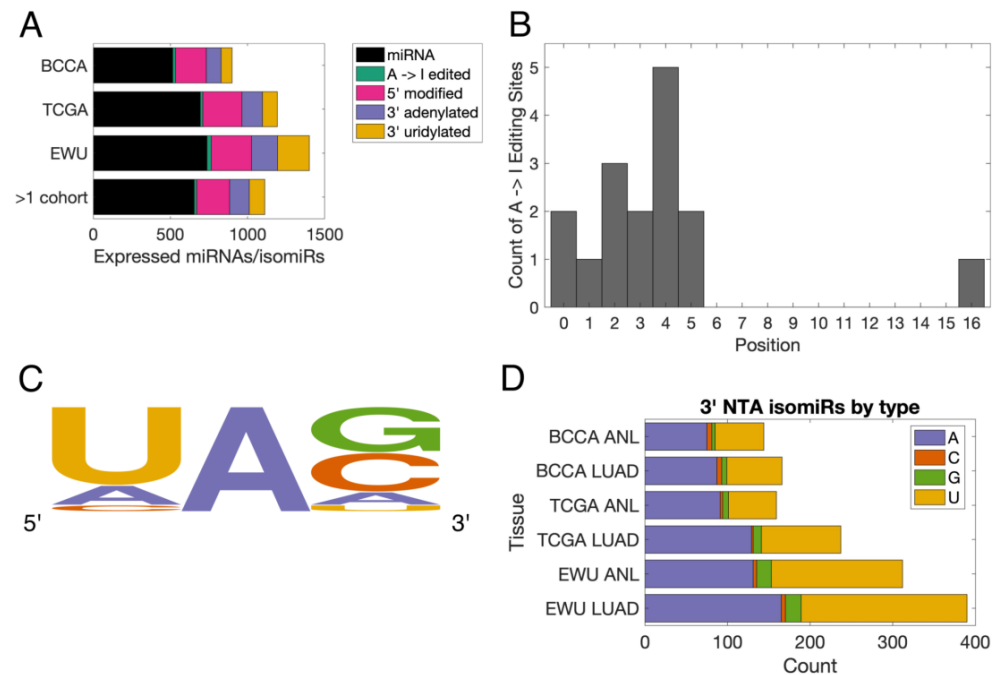


Figure 1. Basic characteristics of the lung isomiRome. (A) Bar chart illustrating the numbers of miRNAs and different types of isomiRs that were expressed in each lung cohort. The bottom row indicates the numbers of high-confidence miRNAs/isomiRs. The bar graph includes isomiRs that are expressed in ANL and LUAD samples in the individual cohorts. (B) Histogram indicating the positions of the 16 high-confidence A-to-I editing sites. Position '0' corresponds to the first nucleotide of a miRNA. (C) Plot of the nucleotides located immediately 5' and 3' of the high-confidence editing sites, with the height of a letter being proportional to its frequency. Generated using WebLogo (v.2.8.2) [45]. (D) Bar chart indicating the numbers of isomiRs with 3' non-templated addition (NTAs) of each nucleotide that was expressed in each sample group. ANL: Adult Non-malignant Lung; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

Table 2. List of high-confidence A-to-I edited isomiRs in the human lung.

miRNA	5' Shift	Edited Position	Adjacent Nucleotides *
hsa-let-7d-3p	0	4	UAC
hsa-miR-151a-3p	−2	2	UAG
hsa-miR-200b-3p	0	4	UAC
hsa-miR-376c-3p	0	5	UAG
hsa-miR-379-5p	0	4	UAG
hsa-miR-381-3p	−1	3	UAC
hsa-miR-381-3p	0	3	UAC
hsa-miR-411-5p	−1	4	UAG
hsa-miR-411-5p	0	4	UAG
hsa-miR-455-5p	0	16	UAC
hsa-miR-4662a-5p	0	2	UAG
hsa-miR-497-5p	0	1	CAG
hsa-miR-589-3p	0	5	AAC
hsa-miR-9903	0	2	UAU
hsa-miR-99a-5p	−1	0	<u>AAA</u>
hsa-miR-99a-5p	0	0	<u>AAA</u>

Legend: * The nucleotides immediately 5' and 3' of the edited adenosine (bolded) are listed. Underlined nucleotides were present only in the pri- or pre-miRNA and not in the mature canonical miRNA. Both the edited and unedited versions of the miRNAs were detected in the sequence reads.

3.2. The Lung isomiRome Is Widely Altered in LUAD

Upon comparing the ANL and LUAD isomiRomes, we found that LUAD samples expressed more isomiRs of all types, except A-to-I edited isomiRs, than ANL samples (Figure 2A). This suggests that the LUAD isomiRome is more diverse than that of ANL, which is consistent with past observations from pan-cancer datasets [46]. To explore differences in the rate of miRNA modifications, we defined the “frequency” of an isomiR to be the ratio of that isomiR’s expression to the total expression of its corresponding canonical miRNA and all of its isomiRs. Using this definition, we found that the frequencies of high-confidence 3’ adenylated isomiRs and A-to-I edited isomiRs were significantly lower in LUAD samples than ANL samples for all three cohorts (Figure 2B). There were no consistent, significant differences in the rates of A-to-I editing, adenylation, or uridylation between groups of tumors of different stages or groups of patients with different smoking histories. However, below-median rates of A-to-I editing in TCGA LUAD samples were near-significantly associated with poorer overall survival ($p = 0.069$), suggesting that the downregulation of miRNA editing in LUAD may have significant biological consequences (Figure 2C).

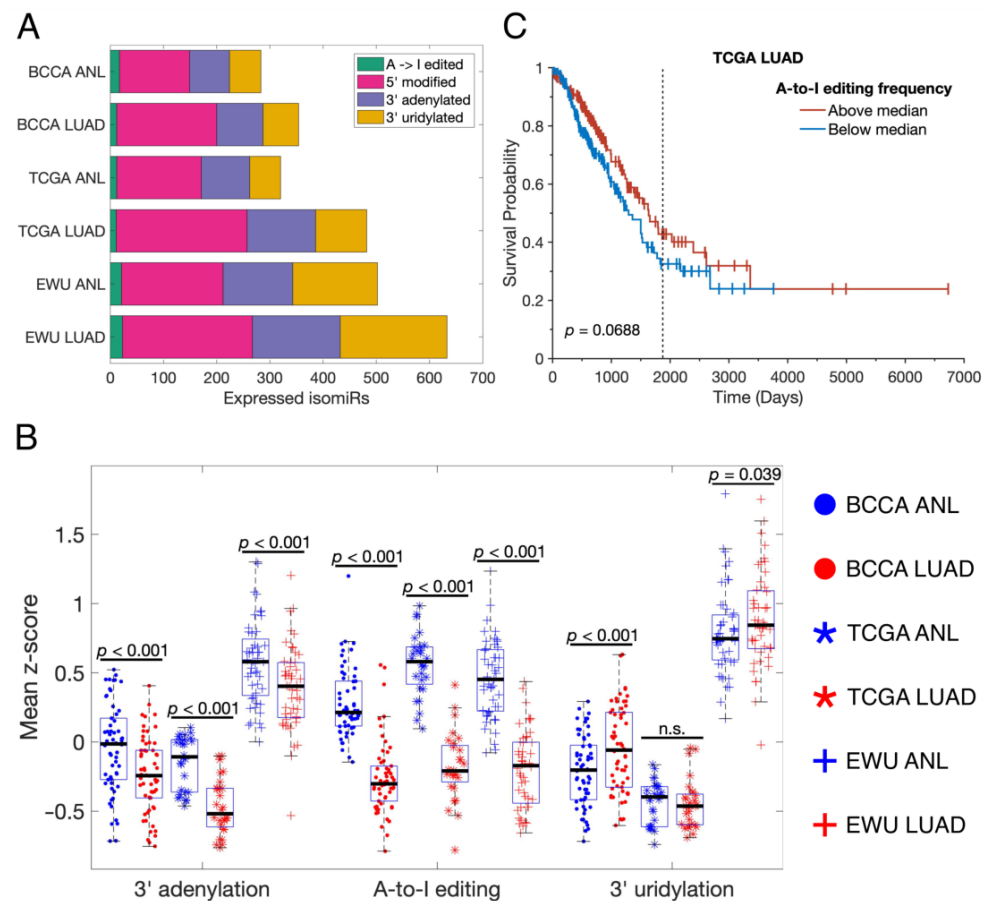


Figure 2. LUAD tumors alter the prevalence and rates of common miRNA modifications. (A) Bar chart showing the numbers of different types of isomiRs that were expressed in each sample group. (B) Scatter-box plots indicating the rates of three major types of miRNA modification in different sample groups, indicated by the legend on the right-hand side. Each point represents a single sample, and only paired ANL and LUAD samples were included. The displayed p -values resulted from paired-sample t -tests. n.s.: not significant. (C) Overall survival curves for TCGA LUAD patients for whom small RNA sequencing and clinical data were available ($n = 379$), stratified by their tumors’ A-to-I miRNA editing rates. The displayed p -value resulted from a log-rank test. The graphical representation was generated using MatSurv [47]. The vertical line indicates the five-year mark. ANL: Adult Non-malignant Lung; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

3.3. miRNA-Modifying Enzymes Are Dysregulated in LUAD

To investigate the causes of these changes to the lung isomiRome in LUAD, we examined the expression of enzymes known to modify miRNAs: ADAR and ADARB1 for A-to-I editing, TENT2 and TENT4B for 3' adenylation, and TENT3A and TENT3B for 3' uridylation. In both the BCCA and TCGA cohorts, ADAR, TENT3A, and TENT3B were more highly expressed in LUAD samples than paired ANL samples, while ADARB1 was less highly expressed in LUAD (Figure 3A,B). Additionally, ADAR expression was significantly positively correlated with the rate of A-to-I editing in both BCCA and TCGA LUAD samples, as was ADARB1 in TCGA LUAD (Figure 3C–F). Interestingly, the above-median expression of ADAR within LUAD samples was significantly associated with poorer overall survival in the TCGA cohort and trended towards significance in the BCCA cohort (Figure 3G,H). This is likely due to ADAR's editing of other RNAs [48], as high rates of miRNA editing were associated with improved outcomes (Figure 2C).

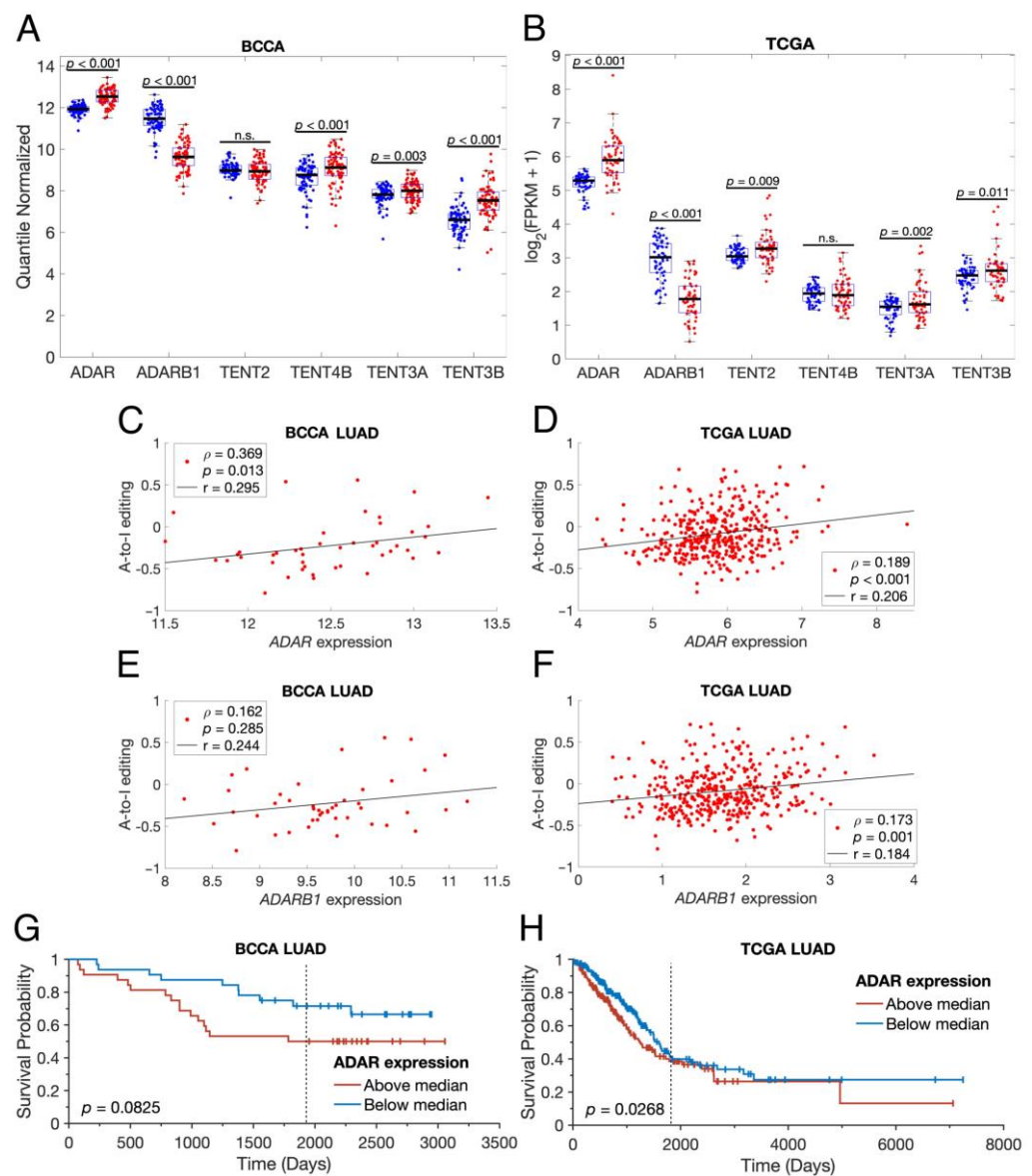


Figure 3. LUAD tumors alter their expression of miRNA-modifying enzymes. (A,B) Scatter-box plots showing the normalized expression of six different miRNA-modifying enzymes in ANL samples (blue points) and paired LUAD samples (red points). The displayed p -values resulted from paired-sample

t-tests. n.s.: not significant, FPKM: fragments per kilobase of transcript per million mapped reads. (C–F) Scatter plots of sample-wide A-to-I editing rates vs. normalized ADAR or ADARB1 expression, where each point represents a single LUAD sample. mRNA expression normalization is identical to that shown in (A,B). Black lines are least-squares lines of best fit. The displayed *p*-values indicate the likelihood that, in an uncorrelated system, Spearman’s ρ would be as far or farther from zero than the displayed ρ . (G,H) Overall survival curves for BCCA LUAD ($n = 64$) and TCGA LUAD ($n = 500$) patients for whom mRNA expression and clinical data were available, stratified by their tumors’ ADAR expression. The displayed *p*-values resulted from log-rank tests. The graphical representations were generated using MatSurv [47]. The vertical line indicates the five-year mark. BCCA: British Columbia Cancer Agency, LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

3.4. Individual miRNAs Are Modified in Distinct Fashions in LUAD

Having examined high-level differences between the ANL and LUAD isomiRNomes, we next investigated how the expression of individual high-confidence isomiRs varied between ANL and LUAD. Eight (50%) of the sixteen A-to-I edited isomiRs had significantly lower frequencies in LUAD in all three cohorts, which is consistent with LUAD’s lower overall editing rate (Figure 4A). Similarly, 36 (28%) of the 3’ adenylated isomiRs had significantly lower frequencies in LUAD in two or more cohorts, as opposed to only 7 (5%) with significantly higher frequencies (Figure 4B). The converse was true for 3’ uridylylated isomiRs (23% with higher frequencies in LUAD; 9% lower) and 5’ isomiRs (19% higher; 11% lower) (Figure 4C,D).

To provide an example of how individual miRNAs are modified differently in LUAD, we decided to look at the isomiRs of hsa-miR-99a-5p in more detail, as that miRNA has high-confidence sites of A-to-I editing, 3’ adenylation, and 3’ uridylation. In all three cohorts, hsa-miR-99a-5p was edited less frequently, adenylated less frequently, and uridylylated more frequently in LUAD than in ANL (Figure 5A–C). As a result, the overall isomiR profile of hsa-miR-99a-5p, which could impact its mRNA targets and stability, was dramatically altered in LUAD. Higher editing of hsa-miR-99a-5p was associated with prolonged survival in the TCGA cohort, suggesting that the edited isomiR may collaborate with the unedited miRNA, which was also associated with prolonged survival and is a known tumor suppressor (Figure 5D,E) [49].

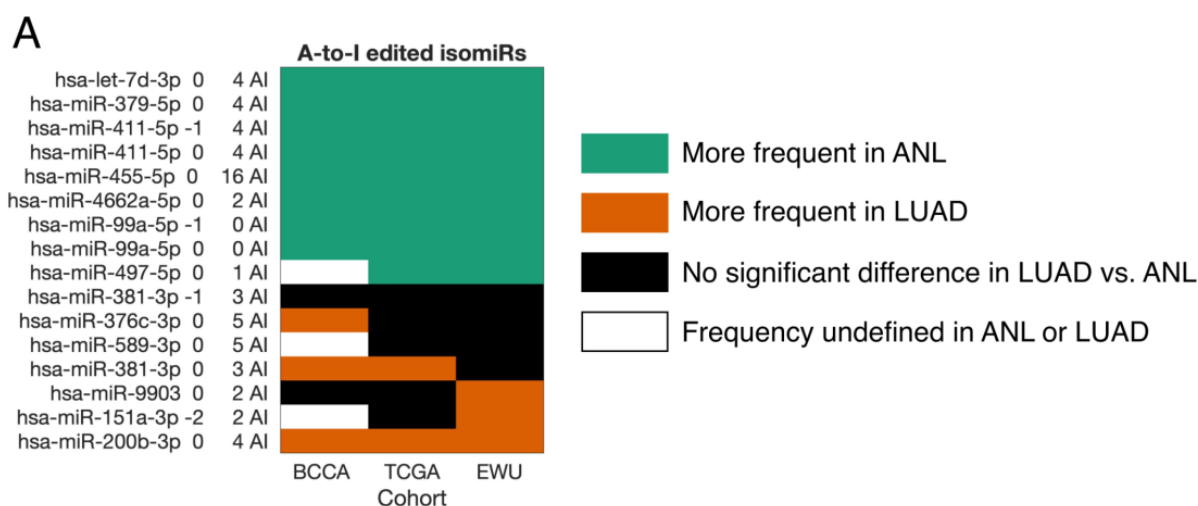


Figure 4. Cont.

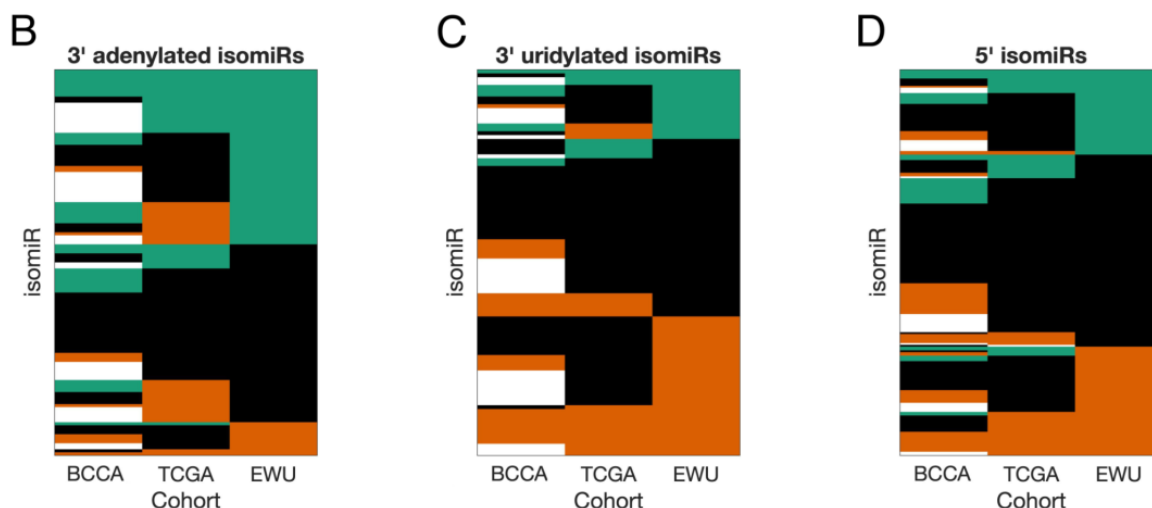


Figure 4. isomiR frequencies are not uniformly altered in LUAD. (A–D) Heatmaps indicating whether the frequencies of individual high-confidence isomiRs (each represented by a row) were significantly altered in LUAD vs. ANL for each of the three adult lung cohorts (represented by columns). Significance was determined by paired-sample *t*-tests with a threshold of a BH-corrected *p*-value < 0.05. White boxes, as indicated in the top-right legend, represent isomiRs whose frequencies were undefined in the ANL or LUAD samples of a particular cohort due to their corresponding miRNAs having zero expression in every sample. ANL: Adult Non-malignant Lung; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

A second interesting example is the A-to-I edited isomiR of hsa-miR-200b-3p, which was the only edited isomiR that was present at a significantly higher frequency in LUAD vs. ANL for all three cohorts (Figure 6A). This suggests that hsa-miR-200b-3p was mainly edited by ADAR, which was also upregulated in LUAD (Figure 6A,B). Indeed, hsa-miR-200b-3p’s editing frequency was significantly correlated with ADAR expression in both BCCA and TCGA LUAD samples (Figure 6B,C). In contrast, the frequency of most other edited isomiRs instead correlated significantly with ADARB1 expression (Table 3). When grouping isomiRs by whether their editing frequency in TCGA LUAD was positively correlated with ADAR, ADARB1, or neither gene, only the ADARB1 group showed a decline in editing in LUAD (Figure 6D). This suggests that the overall decline in editing in LUAD, which is associated with poorer patient outcomes (Figure 2C), is attributable to the decline in ADARB1 expression.

Table 3. Correlations of isomiR editing rates with ADAR/ADARB1 expression.

Edited isomiR	ADAR ρ Value	ADAR <i>p</i> -Value	ADARB1 ρ Value	ADARB1 <i>p</i> -Value	Group *
hsa-let-7d-3p 0 4 AI	0.079	0.124	0.198	9.38×10^{-5}	ADARB1
hsa-miR-151a-3p -2 2 AI	0.267	1.12×10^{-7}	0.067	0.192	ADAR
hsa-miR-200b-3p 0 4 AI	0.452	9.65×10^{-21}	−0.040	0.436	ADAR
hsa-miR-376c-3p 0 5 AI	−0.031	0.547	0.030	0.559	Neither
hsa-miR-379-5p 0 4 AI	−0.030	0.561	0.249	7.63×10^{-7}	ADARB1
hsa-miR-381-3p -1 3 AI	−0.070	0.170	0.006	0.904	Neither
hsa-miR-381-3p 0 3 AI	0.061	0.233	−0.118	0.021	Neither
hsa-miR-411-5p -1 4 AI	−0.047	0.357	0.103	0.044	Neither
hsa-miR-411-5p 0 4 AI	0.008	0.874	0.171	7.97×10^{-4}	ADARB1
hsa-miR-455-5p 0 16 AI	−0.030	0.559	0.268	9.75×10^{-8}	ADARB1
hsa-miR-4662a-5p 0 2 AI	−0.017	0.745	0.224	9.28×10^{-6}	ADARB1
hsa-miR-497-5p 0 1 AI	−0.046	0.374	0.233	3.80×10^{-6}	ADARB1

Table 3. Cont.

Edited isomiR	ADAR ρ Value	ADAR p-Value	ADARB1 ρ Value	ADARB1 p-Value	Group *
hsa-miR-589-3p 0 5 AI	0.319	1.52×10^{-10}	-0.027	0.592	ADAR
hsa-miR-9903 0 2 AI	-0.052	0.343	-0.095	0.081	Neither
hsa-miR-99a-5p -1 0 AI	-0.135	0.008	0.116	0.023	ADARB1
hsa-miR-99a-5p 0 0 AI	-0.081	0.114	0.310	5.48×10^{-10}	ADARB1

Legend: Only TCGA LUAD samples were included in the correlation analysis. The displayed p-values indicate the likelihood that, in an uncorrelated system, Spearman’s ρ would be as far or farther from zero than the displayed ρ. * isomiRs were assigned to the “Neither” group unless they had a positive correlation with ADAR or ADARB1 expression that was significant after BH correction (FDR < 0.05).

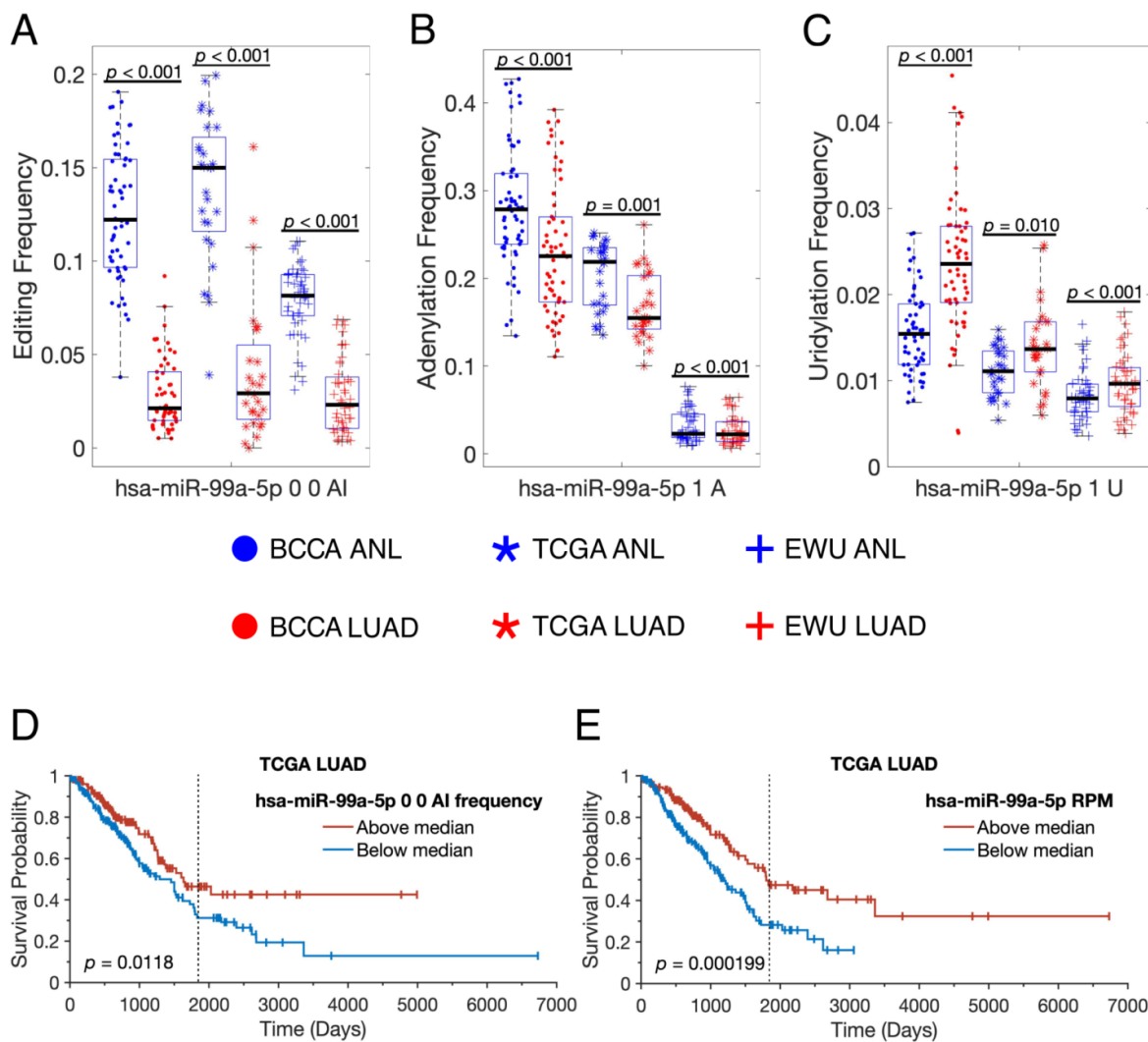


Figure 5. The isomiR profile of hsa-miR-99a-5p is substantially altered in LUAD. (A–C) Scatter-box plots showing the frequencies of high-confidence isomiRs of hsa-miR-99a-5p in different sample groups, which are indicated by the central legend. Each point represents a single sample, and only paired ANL and LUAD samples were included. The displayed p-values resulted from paired-sample t-tests. (D,E) Overall survival curves for TCGA LUAD patients for whom small RNA sequencing and clinical data were available (n = 379), stratified by the frequency with which hsa-miR-99a-5p was edited (D) or by the RPM expression of hsa-miR-99a-5p (E) in their tumors. The displayed p-values resulted from log-rank tests. The graphical representations were generated using MatSurv [47]. The vertical line indicates the five-year mark. ANL: Adult Non-malignant Lung; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

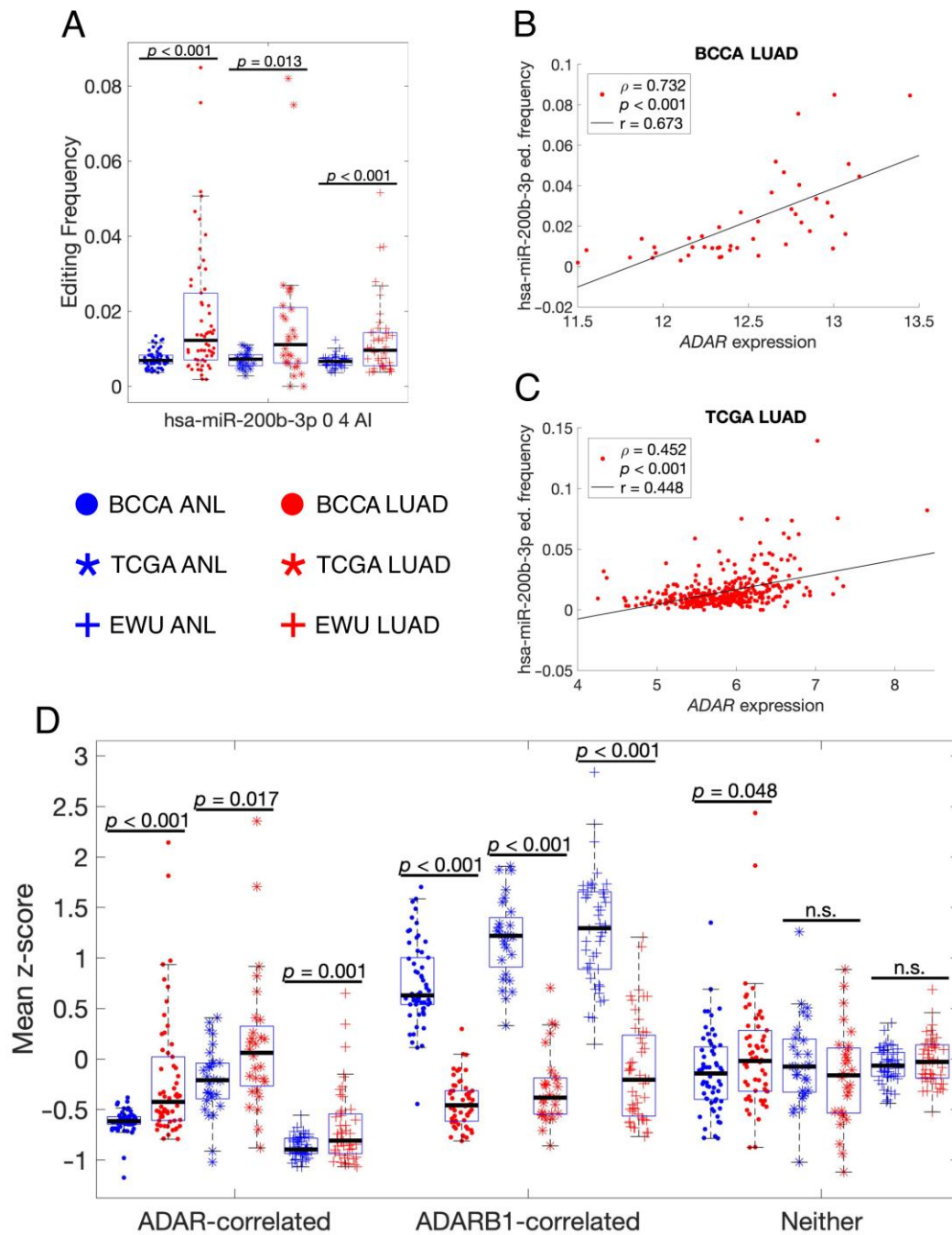


Figure 6. Only ADARB1-correlated edited isomiRs are less frequent in LUAD vs. ANL. **(A)** Scatter-box plots showing the editing frequency of hsa-miR-200b-3p 0 4 AI in different sample groups, which are indicated by the central legend. Each point represents a single sample, and only paired ANL and LUAD samples were included. The displayed p -values resulted from paired-sample t -tests. **(B,C)** Scatter plots of hsa-miR-200b-3p 0 4 AI editing frequencies vs. ADAR expression, where each point represents a single LUAD sample. Black lines are least-squares lines of best fit. The displayed p -values indicate the likelihood that, in an uncorrelated system, Spearman’s ρ would be as far or farther from zero than the displayed ρ . ed.: editing. **(D)** Scatter-box plots showing the rates of A-to-I editing for the three groups of high-confidence edited isomiRs delineated in Table 3. The same paired samples and sample groups were included as for **(A)**, and the displayed p -values resulted from paired-sample t -tests. n.s.: not significant. ANL: Adult Non-malignant Lung; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; LUAD: Lung Adenocarcinoma; TCGA: The Cancer Genome Atlas.

3.5. IsomiR-Based Biomarkers Distinguish ANL and LUAD Samples

Based on the significant differences observed between the ANL and LUAD isomiRomes, we created SVM and RF classifiers to see if isomiRs could differentiate between ANL and LUAD samples from the same cohort. Each classifier was trained on either RPM or frequency data for one type of isomiR or on miRNA RPM data. Out of the 48 isomiR-based classifiers, 44 (92%) were found to be highly accurate (classification error < 5%) in distinguishing ANL and LUAD samples (Table 4).

Table 4. Error rates for single-cohort miRNA and isomiR-based classifiers.

		RPM-Based Classifiers					Frequency-Based Classifiers			
		miRNA	3' U	3' A	A-to-I	5'	3' U	3' A	A-to-I	5'
Mean classification error in 10-fold CV	BCCA SVM	1.56%	1.56%	0.78%	3.13%	2.34%	0.78%	0.00%	4.69%	3.13%
	TCGA SVM	1.45%	1.16%	2.03%	1.45%	0.58%	1.16%	2.62%	1.74%	1.16%
	EWU SVM	2.38%	2.38%	2.38%	2.38%	2.38%	1.19%	3.57%	11.90%	3.57%
	Mean	1.80%	1.70%	1.73%	2.32%	1.77%	1.04%	2.06%	6.11%	2.62%
Out of bag error	BCCA RF	0.78%	0.78%	0.00%	1.56%	0.78%	4.69%	1.56%	3.91%	4.69%
	TCGA RF	1.45%	0.58%	0.87%	1.16%	0.87%	1.45%	2.62%	2.62%	1.74%
	EWU RF	1.19%	0.00%	0.00%	1.19%	0.00%	8.33%	14.29%	3.57%	5.95%
	Mean	1.14%	0.45%	0.29%	1.30%	0.55%	4.82%	6.16%	3.37%	4.13%

Legend: ANL—Adult Non-malignant Lung; BCCA—British Columbia Cancer Agency; EWU—Ewha Womans University; LUAD—Lung Adenocarcinoma; RF—Random Forest; RPM—Reads Per Million; SVM—Support Vector Machine; TCGA—The Cancer Genome Atlas.

To test the generalizability of isomiR-based biomarkers, we trained a second batch of classifiers on either RPM or frequency data for one type of isomiR each, using only TCGA cohort samples, and optimized their hyperparameters through 5-fold cross-validation (CV). Classifiers trained on the RPM values of 5', A-to-I edited, or 3' uridylylated isomiRs were highly accurate when tested on the BCCA and EWU cohorts (mean AUCs of 0.949–0.977) (Table 5 and Figure 7A,B). The classifier trained on the frequency of A-to-I edited isomiRs achieved a similar level of accuracy (mean AUC = 0.972) and compared favorably to classifiers trained on miRNA expression alone (mean AUC = 0.946) (Table 5 and Figure 7C,D). However, the optimal decision threshold for most of these classifiers varied considerably between cohorts (Figure 7A–D).

Table 5. Test cohort AUCs for TCGA-trained miRNA and isomiR-based classifiers.

		RPM-Based Classifiers					Frequency-Based Classifiers			
		miRNA	3' U	3' A	A-to-I	5'	3' U	3' A	A-to-I	5'
Test Cohort AUCs	Features	368	37	41	7	115	17	34	7	36
	BCCA SVM	0.944	0.998	0.962	0.997	0.992	0.663	0.476	0.985	0.716
	EWU SVM	0.951	0.995	0.884	0.934	0.982	0.582	0.752	0.968	0.893
	BCCA RF	0.901	0.891	0.855	0.990	0.894	0.634	0.756	0.981	0.889
	EWU RF	0.988	0.910	0.995	0.988	0.973	0.572	0.704	0.955	0.606
	Mean	0.946	0.949	0.924	0.977	0.960	0.613	0.672	0.972	0.776

Legend: AUC—Area Under the Receiver Operating Characteristic Curve; BCCA—British Columbia Cancer Agency; EWU—Ewha Womans University; LUAD—Lung Adenocarcinoma; RF—Random Forest; RPM—Reads Per Million; SVM—Support Vector Machine; TCGA—The Cancer Genome Atlas.

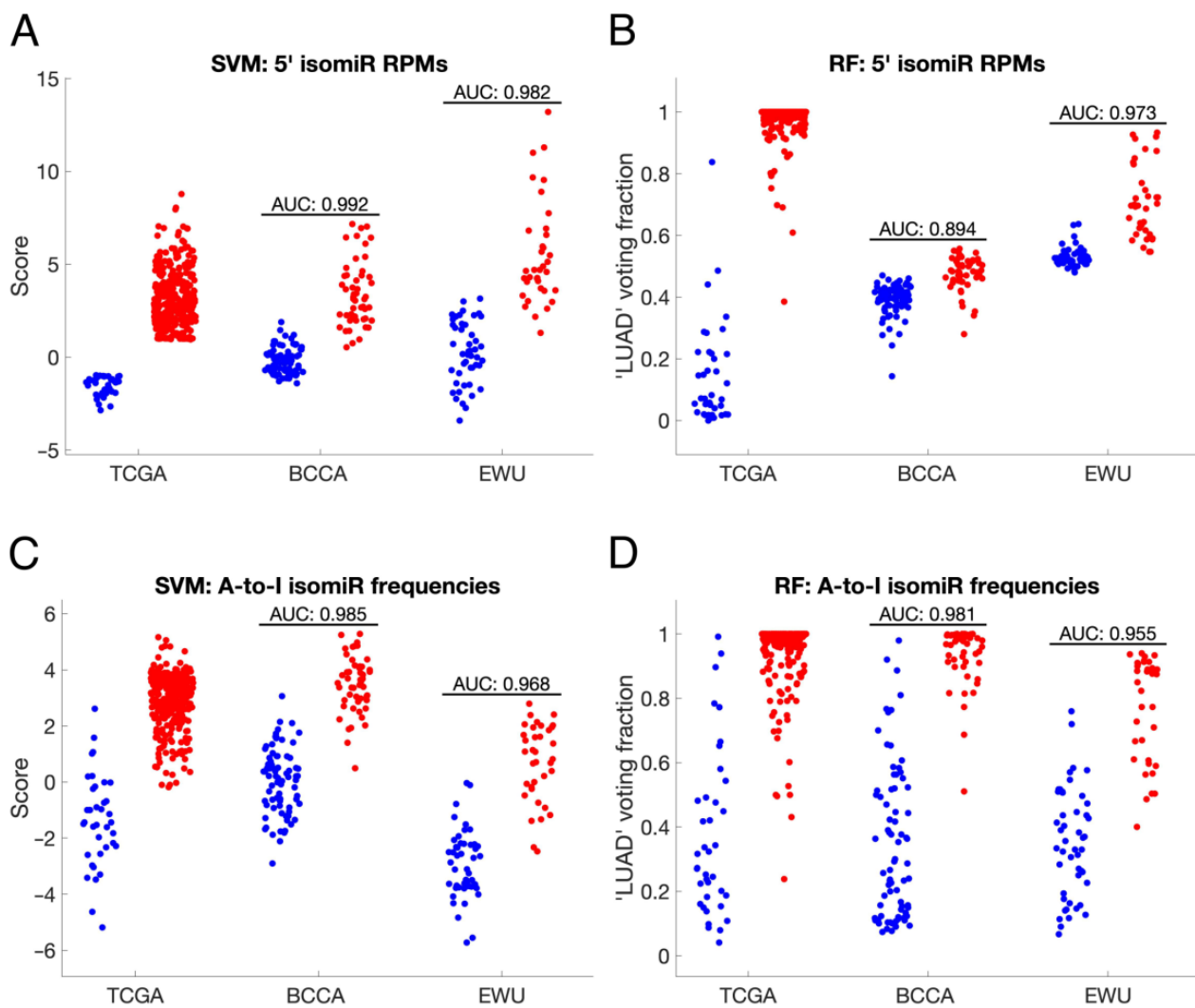


Figure 7. IsomiR-based classifiers accurately distinguish ANL and LUAD samples. (A–D) Plots of SVM scores and RF voting fractions for the training cohort (TCGA) and test cohorts (BCCA and EWU). Blue points represent ANL samples, and red points represent stage I/II LUAD samples. Each classifier was trained on either RPM or frequency data for one type of isomiR, as outlined in each panel’s title. SVM and RF hyperparameters were optimized prior to testing, as described in the methodology. AUC: Area Under the Receiver Operating Characteristic Curve; BCCA: British Columbia Cancer Agency; EWU: Ewha Womans University; RF: Random Forest; RPM: Reads Per Million; SVM: Support Vector Machine; TCGA: The Cancer Genome Atlas.

4. Discussion

In this study, we analyzed the expression of miRNAs and four major types of isomiRs in 163 ANL and 506 LUAD samples from three independent human lung cohorts. To the best of our knowledge, this is the most comprehensive analysis of isomiR expression in the human lung to date and the first study to investigate the expression of multiple types of isomiRs across multiple lung cohorts. The extensive scope of this approach led to the discovery that the expression patterns of these four types of isomiRs are altered in LUAD, making the LUAD isomiRome unique.

We detected the expression of 16 A-to-I edited isomiRs, 213 5' isomiRs, 128 3' adenylylated isomiRs, 100 3' uridylylated isomiRs, and 654 miRNAs in multiple cohorts. In total, isomiRs made up 41% of the unique high-confidence sequences. Not all of these isomiRs will necessarily target different mRNAs than their canonical miRNAs, but this abundance suggests that regulation of the rate of miRNA modifications is a viable mechanism through

which tumors can dysregulate core cellular pathways. Indeed, some of the isomiRs found to be deregulated in this study, such as hsa-miR-200b-3p and the 5' isomiR hsa-miR-140-3p, have been shown to have distinct targets from their canonical miRNAs and to thereby influence cancer cell viability and migration [22,26].

Upon comparing the ANL and LUAD isomiRomes, we found that many isomiRs of each of the four types were expressed at significantly different frequencies, relative to their canonical miRNAs, in LUAD vs. ANL. Decreases in miRNA editing are an established phenomenon in many forms of cancer [21], but we found that the overall rate of 3' adenylation also decreased in LUAD. This change may significantly impact the functional lung miRNome, as 3' adenylation has been shown to diminish miRNA-AGO2 interaction and could also alter a miRNA's 3'-region complementarity with miRNA response elements (MREs) [29]. There was also a trend towards prolonged survival in patients with higher overall miRNA editing rates, which is consistent with the past identification of several edited isomiRs as tumor suppressors [10].

Interestingly, the respective upregulation and downregulation of the ADAR and ADARB1 editing enzymes in LUAD appeared to cause corresponding changes in editing rates: ADARB1-correlated isomiRs, which were a majority, were edited less in LUAD, and the small number of ADAR-correlated isomiRs were more highly edited in LUAD. Yet, despite miRNA editing rates tending to be lower in patients who experienced poor outcomes, low ADAR expression was linked to improved outcomes. This negative impact of ADAR is thought to be due to its editing of other families of RNAs [50]. For example, ADAR can promote LUAD cell migration and invasion by stabilizing the FAK mRNA through A-to-I editing [48]. Future studies should explore the relationship between mRNA expression changes and corresponding protein levels to fully understand their implications for cellular functions and outcomes.

The influence of adenylation and uridylation enzymes on the lung isomiRome was less clear, as rates of adenylation declined in LUAD even though the expression of adenylation enzymes was not consistently altered, and rates of uridylation were not consistently altered even though uridylation enzymes were upregulated in LUAD. One potential explanation for this incongruence is that RNA sequencing takes a snapshot of the transcriptome at a single point in time, and 3' NTAs can significantly alter miRNA stability [28,51]. The observed rate of adenylation/uridylation for miRNAs whose stability is modulated by the NTA will thus vary with the expression and activity of exonucleases, and isomiRs whose NTAs mark them for imminent degradation may go undetected altogether. In addition to this scenario, it is also possible that miRNA adenylation and uridylation in the lung are substantially influenced by enzymes other than the ones investigated here, such as TENT1 and TENT6 [52].

Finally, SVM and RF classifiers trained on either isomiR RPM or frequency values demonstrated high accuracy in distinguishing ANL from stage I/II LUAD samples, even when applied to cohorts different from those they were trained on. Notably, classifiers trained on the RPM or frequency of A-to-I edited isomiRs achieved mean test cohort AUCs of 0.977 and 0.972, respectively, using only seven features each. These classifiers outperformed a previously published statistical model based on miRNA editing frequencies, which achieved a mean AUC of 0.895 upon cross-validation and was not tested in other cohorts [21]. A limitation of most of these classifiers is that the optimal score cutoff varied substantially between cohorts. This variance could be attributed to differences in sample preparation and sequencing methods (e.g., microdissection, Illumina HiSeq2000 vs. 2500) or patient characteristics (e.g., gender, race, and smoking history) across the three cohorts. Ultimately, isomiR-based biomarkers hold promise for the diagnosis of early-stage lung adenocarcinoma if they are measurable from liquid biopsy samples and show consistency across patients. Encouragingly, previous studies have identified isomiRs in serum, with an overlap observed between lung tissue-edited and serum-edited isomiRs [23,53,54].

5. Conclusions

This study provides a comprehensive analysis of isomiR expression across various lung cohorts, offering an extensive overview of multiple isomiR types in the human lung. Our findings suggest that ANL and LUAD tissues have highly diverse isomiRomes. The deregulation of miRNA-modifying enzymes in LUAD contributes to both global and miRNA-specific differences between these isomiRomes, presenting diagnostic opportunities and potentially influencing coding gene regulation in tumors. Future studies may shed light on the function of these isomiRs, as well as the specific mRNAs that they regulate, potentially leading to the discovery of new therapeutic targets.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers16193322/s1>, Figure S1: Biogenesis and functional consequences of selected miRNA modifications.

Author Contributions: D.E.C. and W.L.L. initiated the project. D.E.C. designed, researched, analyzed, and wrote about the topics covered in the article. D.E.C., N.T. and V.G.P.S. designed the figures. A.F., G.L.S., V.G.P.S. and W.L.L. revised the manuscript. W.L.L. is the principal investigator. All authors contributed to data interpretation and manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funds from the Canadian Institutes of Health Research (CIHR, FRN-143345 and 183775), the Terry Fox Foundation, the Lotte and John Hecht Memorial Foundation, and the BC Cancer Foundation.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the University of British Columbia/BC Cancer Agency Research Ethics Board (H15-03060).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets supporting the findings of this study are publicly available as follows: miRNA sequencing data for the TCGA cohort were retrieved from Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>, accessed on 11 July 2019). mRNA expression data and patient clinical information, including survival data, were obtained from the University of California Santa Cruz (UCSC) Xena browser (<https://xenabrowser.net/>, accessed on 11 July 2019). miRNA sequencing data for the Ewha Womans University (EWU) cohort were downloaded from the European Nucleotide Archive (Study Accession: PRJNA434883) (<https://www.ebi.ac.uk/ena/browser/home>, accessed on 12 February 2021). Clinical data for the EWU cohort were acquired through the Sequence Read Archive run selector (SRA Accession: SRP133217) (<https://www.ncbi.nlm.nih.gov/sra>, accessed on 12 February 2021). miRNA expression data for samples in the BCCA cohort were deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO accession: GSE175462).

Acknowledgments: This manuscript is a revised version of Chapter 4 of DE Cohn's published Master's thesis.

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

References

1. Gebert, L.F.R.; MacRae, I.J. Regulation of microRNA Function in Animals. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 21–37. [[CrossRef](#)] [[PubMed](#)]
2. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol.* **2018**, *9*, 402. [[CrossRef](#)]
3. Wilczynska, A.; Bushell, M. The Complexity of miRNA-Mediated Repression. *Cell Death Differ.* **2015**, *22*, 22–33. [[CrossRef](#)]
4. Iqbal, M.A.; Arora, S.; Prakasam, G.; Calin, G.A.; Syed, M.A. MicroRNA in Lung Cancer: Role, Mechanisms, Pathways and Therapeutic Relevance. *Mol. Aspects Med.* **2019**, *70*, 3–20. [[CrossRef](#)] [[PubMed](#)]

5. Jin, X.; Chen, Y.; Chen, H.; Fei, S.; Chen, D.; Cai, X.; Liu, L.; Lin, B.; Su, H.; Zhao, L.; et al. Evaluation of Tumor-Derived Exosomal miRNA as Potential Diagnostic Biomarkers for Early-Stage Non-Small Cell Lung Cancer Using Next-Generation Sequencing. *Clin. Cancer Res.* **2017**, *23*, 5311–5319. [[CrossRef](#)] [[PubMed](#)]
6. Sholl, L.M.; Aisner, D.L.; Varella-Garcia, M.; Berry, L.D.; Dias-Santagata, D.; Wistuba, I.I.; Chen, H.; Fujimoto, J.; Kugler, K.; Franklin, W.A.; et al. Multi-Institutional Oncogenic Driver Mutation Analysis in Lung Adenocarcinoma: The Lung Cancer Mutation Consortium Experience. *J. Thorac. Oncol.* **2015**, *10*, 768–777. [[CrossRef](#)]
7. Tomasello, L.; Distefano, R.; Nigita, G.; Croce, C.M. The MicroRNA Family Gets Wider: The IsomiRs Classification and Role. *Front. Cell Dev. Biol.* **2021**, *9*, 668648. [[CrossRef](#)]
8. Cloonan, N.; Wani, S.; Xu, Q.; Gu, J.; Lea, K.; Heater, S.; Barbacioru, C.; Steptoe, A.L.; Martin, H.C.; Nourbakhsh, E.; et al. MicroRNAs and Their isomiRs Function Cooperatively to Target Common Biological Pathways. *Genome Biol.* **2011**, *12*, R126. [[CrossRef](#)]
9. Tan, G.C.; Chan, E.; Molnar, A.; Sarkar, R.; Alexieva, D.; Isa, I.M.; Robinson, S.; Zhang, S.; Ellis, P.; Langford, C.F.; et al. 5' isomiR Variation Is of Functional and Evolutionary Importance. *Nucleic Acids Res.* **2014**, *42*, 9424–9435. [[CrossRef](#)]
10. Xu, X.; Wang, Y.; Mojumdar, K.; Zhou, Z.; Jeong, K.J.; Mangala, L.S.; Yu, S.; Tsang, Y.H.; Rodriguez-Aguayo, C.; Lu, Y.; et al. A-to-I-Edited miRNA-379-5p Inhibits Cancer Cell Proliferation through CD97-Induced Apoptosis. *J. Clin. Investig.* **2019**, *129*, 5343–5356. [[CrossRef](#)]
11. Shoshan, E.; Mobley, A.K.; Braeuer, R.R.; Kamiya, T.; Huang, L.; Vasquez, M.E.; Salameh, A.; Lee, H.J.; Kim, S.J.; Ivan, C.; et al. Reduced Adenosine-to-Inosine miR-455-5p Editing Promotes Melanoma Growth and Metastasis. *Nat. Cell Biol.* **2015**, *17*, 311–321. [[CrossRef](#)] [[PubMed](#)]
12. Wang, S.; Zheng, Z.; Chen, P.; Wu, M. Tumor Classification and Biomarker Discovery Based on the 5' isomiR Expression Level. *BMC Cancer* **2019**, *19*, 127. [[CrossRef](#)] [[PubMed](#)]
13. Yang, Z.; Fang, S.; Di, Y.; Ying, W.; Tan, Y.; Gu, W. Modulation of NF- κ B/miR-21/PTEN Pathway Sensitizes Non-Small Cell Lung Cancer to Cisplatin. *PLoS ONE* **2015**, *10*, e0121547. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, J.; Wang, J.; Zhao, F.; Liu, Q.; Jiang, K.; Yang, G. MicroRNA-21 (miR-21) Represses Tumor Suppressor PTEN and Promotes Growth and Invasion in Non-Small Cell Lung Cancer (NSCLC). *Clin. Chim. Acta* **2010**, *411*, 846–852. [[CrossRef](#)] [[PubMed](#)]
15. Xia, W.; Chen, Q.; Wang, J.; Mao, Q.; Dong, G.; Shi, R.; Zheng, Y.; Xu, L.; Jiang, F. DNA Methylation Mediated Silencing of microRNA-145 Is a Potential Prognostic Marker in Patients with Lung Adenocarcinoma. *Sci. Rep.* **2015**, *5*, 16901. [[CrossRef](#)]
16. Chen, Z.; Zeng, H.; Guo, Y.; Liu, P.; Pan, H.; Deng, A.; Hu, J. miRNA-145 Inhibits Non-Small Cell Lung Cancer Cell Proliferation by Targeting c-Myc. *J. Exp. Clin. Cancer Res.* **2010**, *29*, 151. [[CrossRef](#)]
17. Grenda, A.; Krawczyk, P.; Błach, J.; Chmielewska, I.; Kubiakowski, T.; Kieszko, S.; Wojas-Krawczyk, K.; Kucharczyk, T.; Jarosz, B.; Paśnik, I.; et al. Tissue MicroRNA Expression as a Predictor of Response to Immunotherapy in NSCLC Patients. *Front. Oncol.* **2020**, *10*, 563613. [[CrossRef](#)]
18. Wang, S.; Su, X.; Bai, H.; Zhao, J.; Duan, J.; An, T.; Zhuo, M.; Wang, Z.; Wu, M.; Li, Z.; et al. Identification of Plasma microRNA Profiles for Primary Resistance to EGFR-TKIs in Advanced Non-Small Cell Lung Cancer (NSCLC) Patients with EGFR Activating Mutation. *J. Hematol. Oncol.* **2015**, *8*, 127. [[CrossRef](#)]
19. Teng, P.-C.; Liang, Y.; Yarmishyn, A.A.; Hsiao, Y.-J.; Lin, T.-Y.; Lin, T.-W.; Teng, Y.-C.; Yang, Y.-P.; Wang, M.-L.; Chien, C.-S.; et al. RNA Modifications and Epigenetics in Modulation of Lung Cancer and Pulmonary Diseases. *Int. J. Mol. Sci.* **2021**, *22*, 10592. [[CrossRef](#)]
20. The Cancer Genome Atlas Program (TCGA)—NCI. Available online: <https://www.cancer.gov/ccg/research/genome-sequencing/tcga> (accessed on 12 August 2024).
21. Pinto, Y.; Buchumenski, I.; Levanon, E.Y.; Eisenberg, E. Human Cancer Tissues Exhibit Reduced A-to-I Editing of miRNAs Coupled with Elevated Editing of Their Targets. *Nucleic Acids Res.* **2018**, *46*, 71–82. [[CrossRef](#)]
22. Wang, Y.; Xu, X.; Yu, S.; Jeong, K.J.; Zhou, Z.; Han, L.; Tsang, Y.H.; Li, J.; Chen, H.; Mangala, L.S.; et al. Systematic Characterization of A-to-I RNA Editing Hotspots in microRNAs across Human Cancers. *Genome Res.* **2017**, *27*, 1112–1125. [[CrossRef](#)] [[PubMed](#)]
23. Nigita, G.; Distefano, R.; Veneziano, D.; Romano, G.; Rahman, M.; Wang, K.; Pass, H.; Croce, C.M.; Acunzo, M.; Nana-Sinkam, P. Tissue and Exosomal miRNA Editing in Non-Small Cell Lung Cancer. *Sci. Rep.* **2018**, *8*, 10222. [[CrossRef](#)] [[PubMed](#)]
24. Maemura, K.; Watanabe, K.; Ando, T.; Hiyama, N.; Sakatani, T.; Amano, Y.; Kage, H.; Nakajima, J.; Yatomi, Y.; Nagase, T.; et al. Altered Editing Level of microRNAs Is a Potential Biomarker in Lung Adenocarcinoma. *Cancer Sci.* **2018**, *109*, 3326–3335. [[CrossRef](#)] [[PubMed](#)]
25. Distefano, R.; Nigita, G.; Le, P.; Romano, G.; Acunzo, M.; Nana-Sinkam, P. Disparities in Lung Cancer: miRNA Isoform Characterization in Lung Adenocarcinoma. *Cancers* **2022**, *14*, 773. [[CrossRef](#)]
26. Salem, O.; Erdem, N.; Jung, J.; Münstermann, E.; Wörner, A.; Wilhelm, H.; Wiemann, S.; Körner, C. The Highly Expressed 5' isomiR of Hsa-miR-140-3p Contributes to the Tumor-Suppressive Effects of miR-140 by Reducing Breast Cancer Proliferation and Migration. *BMC Genom.* **2016**, *17*, 566. [[CrossRef](#)] [[PubMed](#)]
27. D'Ambrogio, A.; Gu, W.; Udagawa, T.; Mello, C.C.; Richter, J.D. Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. *Cell Rep.* **2012**, *2*, 1537–1545. [[CrossRef](#)]

28. Boele, J.; Persson, H.; Shin, J.W.; Ishizu, Y.; Newie, I.S.; Søkilde, R.; Hawkins, S.M.; Coarfa, C.; Ikeda, K.; Takayama, K.; et al. PAPD5-Mediated 3' Adenylation and Subsequent Degradation of miR-21 Is Disrupted in Proliferative Disease. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11467–11472. [[CrossRef](#)] [[PubMed](#)]
29. Burroughs, A.M.; Ando, Y.; de Hoon, M.J.L.; Tomaru, Y.; Nishibu, T.; Ukekawa, R.; Funakoshi, T.; Kurokawa, T.; Suzuki, H.; Hayashizaki, Y.; et al. A Comprehensive Survey of 3' Animal miRNA Modification Events and a Possible Role for 3' Adenylation in Modulating miRNA Targeting Effectiveness. *Genome Res.* **2010**, *20*, 1398–1410. [[CrossRef](#)] [[PubMed](#)]
30. Enfield, K.S.S.; Marshall, E.A.; Anderson, C.; Ng, K.W.; Rahmati, S.; Xu, Z.; Fuller, M.; Milne, K.; Lu, D.; Shi, R.; et al. Epithelial Tumor Suppressor ELF3 Is a Lineage-Specific Amplified Oncogene in Lung Adenocarcinoma. *Nat. Commun.* **2019**, *10*, 5438. [[CrossRef](#)]
31. Vucic, E.A.; Thu, K.L.; Pikor, L.A.; Enfield, K.S.S.; Yee, J.; English, J.C.; MacAulay, C.E.; Lam, S.; Jurisica, I.; Lam, W.L. Smoking Status Impacts microRNA Mediated Prognosis and Lung Adenocarcinoma Biology. *BMC Cancer* **2014**, *14*, 778. [[CrossRef](#)]
32. Cohn, D.E.; Barros-Filho, M.C.; Minatel, B.C.; Pewarchuk, M.E.; Marshall, E.A.; Vucic, E.A.; Sage, A.P.; Telkar, N.; Stewart, G.L.; Jurisica, I.; et al. Reactivation of Multiple Fetal miRNAs in Lung Adenocarcinoma. *Cancers* **2021**, *13*, 2686. [[CrossRef](#)] [[PubMed](#)]
33. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res.* **2002**, *30*, 207–210. [[CrossRef](#)] [[PubMed](#)]
34. Durinck, S.; Moreau, Y.; Kasprzyk, A.; Davis, S.; De Moor, B.; Brazma, A.; Huber, W. BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **2005**, *15*, 3439–3440. [[CrossRef](#)]
35. Howe, K.L.; Achuthan, P.; Allen, J.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Azov, A.G.; Bennett, R.; Bhai, J.; et al. Ensembl 2021. *Nucleic Acids Res.* **2021**, *49*, D884–D891. [[CrossRef](#)]
36. Chu, A.; Robertson, G.; Brooks, D.; Mungall, A.J.; Birol, I.; Coope, R.; Ma, Y.; Jones, S.; Marra, M.A. Large-Scale Profiling of microRNAs for The Cancer Genome Atlas. *Nucleic Acids Res.* **2016**, *44*, e3. [[CrossRef](#)]
37. Yu, N.; Yong, S.; Kim, H.K.; Choi, Y.-L.; Jung, Y.; Kim, D.; Seo, J.; Lee, Y.E.; Baek, D.; Lee, J.; et al. Identification of Tumor Suppressor miRNAs by Integrative miRNA and mRNA Sequencing of Matched Tumor-Normal Samples in Lung Adenocarcinoma. *Mol. Oncol.* **2019**, *13*, 1356–1368. [[CrossRef](#)]
38. Fehlmann, T.; Kern, F.; Laham, O.; Backes, C.; Solomon, J.; Hirsch, P.; Volz, C.; Müller, R.; Keller, A. miRMaster 2.0: Multi-Species Non-Coding RNA Sequencing Analyses at Scale. *Nucleic Acids Res.* **2021**, *49*, W397–W408. [[CrossRef](#)] [[PubMed](#)]
39. Alon, S.; Mor, E.; Vigneault, F.; Church, G.M.; Locatelli, F.; Galeano, F.; Gallo, A.; Shomron, N.; Eisenberg, E. Systematic Identification of Edited microRNAs in the Human Brain. *Genome Res.* **2012**, *22*, 1533–1540. [[CrossRef](#)]
40. Liu, C.-J.; Fu, X.; Xia, M.; Zhang, Q.; Gu, Z.; Guo, A.-Y. miRNASNP-v3: A Comprehensive Database for SNPs and Disease-Related Variations in miRNAs and miRNA Targets. *Nucleic Acids Res.* **2021**, *49*, D1276–D1281. [[CrossRef](#)]
41. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947. [[CrossRef](#)]
42. Ibing, S.; Michels, B.E.; Mosdzien, M.; Meyer, H.R.; Feuerbach, L.; Körner, C. On the Impact of Batch Effect Correction in TCGA isomiR Expression Data. *NAR Cancer* **2021**, *3*, zcab007. [[CrossRef](#)] [[PubMed](#)]
43. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome Biol.* **2009**, *10*, R25. [[CrossRef](#)] [[PubMed](#)]
44. Hou, Y.-M. CCA Addition to tRNA: Implications for tRNA Quality Control. *IUBMB Life* **2010**, *62*, 251–260. [[CrossRef](#)]
45. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. [[CrossRef](#)]
46. McCall, M.N.; Kim, M.-S.; Adil, M.; Patil, A.H.; Lu, Y.; Mitchell, C.J.; Leal-Rojas, P.; Xu, J.; Kumar, M.; Dawson, V.L.; et al. Toward the Human Cellular microRNAome. *Genome Res.* **2017**, *27*, 1769–1781. [[CrossRef](#)] [[PubMed](#)]
47. Creed, J.; Gerke, T.; Berglund, A. MatSurv: Survival Analysis and Visualization in MATLAB. *J. Open Source Softw.* **2020**, *5*, 1830. [[CrossRef](#)]
48. Amin, E.M.; Liu, Y.; Deng, S.; Tan, K.S.; Chudgar, N.; Mayo, M.W.; Sanchez-Vega, F.; Adusumilli, P.S.; Schultz, N.; Jones, D.R. The RNA-Editing Enzyme ADAR Promotes Lung Adenocarcinoma Migration and Invasion by Stabilizing FAK. *Sci. Signal* **2017**, *10*, eaah3941. [[CrossRef](#)]
49. Mizuno, K.; Tanigawa, K.; Nohata, N.; Misono, S.; Okada, R.; Asai, S.; Moriya, S.; Suetsugu, T.; Inoue, H.; Seki, N. FAM64A: A Novel Oncogenic Target of Lung Adenocarcinoma Regulated by Both Strands of miR-99a (miR-99a-5p and miR-99a-3p). *Cells* **2020**, *9*, 2083. [[CrossRef](#)]
50. Chan, T.H.M.; Qamra, A.; Tan, K.T.; Guo, J.; Yang, H.; Qi, L.; Lin, J.S.; Ng, V.H.E.; Song, Y.; Hong, H.; et al. ADAR-Mediated RNA Editing Predicts Progression and Prognosis of Gastric Cancer. *Gastroenterology* **2016**, *151*, 637–650.e10. [[CrossRef](#)]
51. Gutiérrez-Vázquez, C.; Enright, A.J.; Rodríguez-Galán, A.; Pérez-García, A.; Collier, P.; Jones, M.R.; Benes, V.; Mizgerd, J.P.; Mittelbrunn, M.; Ramiro, A.R.; et al. 3' Uridylation Controls Mature microRNA Turnover during CD4 T-Cell Activation. *RNA* **2017**, *23*, 882–891. [[CrossRef](#)]
52. Wyman, S.K.; Knouf, E.C.; Parkin, R.K.; Fritz, B.R.; Lin, D.W.; Dennis, L.M.; Krouse, M.A.; Webster, P.J.; Tewari, M. Post-Transcriptional Generation of miRNA Variants by Multiple Nucleotidyl Transferases Contributes to miRNA Transcriptome Complexity. *Genome Res.* **2011**, *21*, 1450–1461. [[CrossRef](#)] [[PubMed](#)]

53. Koppers-Lalic, D.; Hackenberg, M.; de Menezes, R.; Misovic, B.; Wachalska, M.; Geldof, A.; Zini, N.; de Reijke, T.; Wurdinger, T.; Vis, A.; et al. Non-invasive Prostate Cancer Detection by Measuring miRNA Variants (isomiRs) in Urine Extracellular Vesicles. *Oncotarget* **2016**, *7*, 22566–22578. [[CrossRef](#)] [[PubMed](#)]
54. Koi, Y.; Tsutani, Y.; Nishiyama, Y.; Ueda, D.; Ibuki, Y.; Sasada, S.; Akita, T.; Masumoto, N.; Kadoya, T.; Yamamoto, Y.; et al. Predicting the Presence of Breast Cancer Using Circulating Small RNAs, Including Those in the Extracellular Vesicles. *Cancer Sci.* **2020**, *111*, 2104–2115. [[CrossRef](#)] [[PubMed](#)]

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