Journal of
Clinical Medicine

MDPI

Article

Diagnostic and Prognostic Potential of MiR-379/656 MicroRNA Cluster in Molecular Subtypes of Breast Cancer

Megha Lal [1,2], Asgar Hussain Ansari [1,2], Anurag Agrawal [1,2] and Arijit Mukhopadhyay [2,3,*]

1 Genomics & Molecular Medicine Unit, CSIR-Institute of Genomics & Integrative Biology, Delhi 110025, India; meghalal110@gmail.com (M.L.); asgar.hussain@igib.in (A.H.A.); a.agrawal@igib.in (A.A.)
2 Academy of Scientific and Innovative Research, Ghaziabad, Uttar Pradesh 201002, India
3 Biomedical Research Centre, Translational Medicine Unit, University of Salford, Manchester M5 4WT, UK
* Correspondence: a.mukhopadhyay@salford.ac.uk; Tel.: +44-0161-295-8129

Abstract: Introduction: Breast cancer is the most frequently diagnosed cancer globally and is one of the most important contributors to cancer-related deaths. Earlier diagnosis is known to reduce mortality, and better biomarkers are needed. MiRNA clusters often co-express and target mRNAs in a coordinated fashion, perturbing entire pathways; they thus merit further exploration for diagnostic or prognostic use. MiR-379/656, at chromosome 14q32, is the second largest miRNA cluster in the human genome and implicated in various malignancies including glioblastoma, melanoma, gastrointestinal tumors and ovarian cancer highlighting its potential importance. In this study, we focus on the diagnostic and prognostic potentials of MiR-379/656 in breast cancer and its molecular subtypes. Materials and Methods: We analyzed miRNA and mRNA next generation sequencing data from 903 primary tumors and 90 normal controls (source: The Cancer Genome Atlas). The differential expression profile between tumor and normal was analyzed using DeSEQ2. Penalized logistic regression modelling (lasso regression) was used to assess the predictive potential of MiR-379/656 expression for tumor and normal samples. The association between MiR-379/656 expression and overall patient survival was studied using Cox Proportional-Hazard Model. The target mRNAs (validated) of MiR-379/656 were annotated via pathway enrichment analysis to understand the biological significance of the cluster in breast cancer. Results: The differential expression analysis for 1390 miRNAs (miRnome) revealed 310 upregulated (22.3%) and 176 downregulated (12.66%) miRNAs in breast cancer patients compared with controls. For MiR-379/656, 32 miRNAs (32/42; 76%) were downregulated. The MiR-379/656 cluster was found to be the most differentially expressed cluster in the human genome ($p < 10^{-30}$). The Basal and Luminal B subtypes showed at least 83% (35/42) of the miRNAs to be downregulated. The binomial model prioritized 15 miRNAs, which distinguished breast cancer patients from controls with 99.15 $\pm$ 0.58% sensitivity and 77.78 $\pm$ 5.24% specificity. Overall, the Basal and Luminal B showed the most effective predictive power with respect to the 15 prioritized miRNAs at MiR-379/656 cluster. The decreased expression of MiR-379/656 was found to be associated with poorer clinical outcome in Basal and Luminal B subtypes, increasing tumor stage and tumor size/extent, and overall patient survival. Pathway enrichment for the validated targets of MiR-379/656 was significant for cancer-related pathways, especially DNA repair, transcriptional regulation by p53 and cell cycle checkpoints (adjusted p-value < 0.05). Conclusions: Genome informatics analysis of high throughput data for MiR-379/656 cluster has shown that a subset of 15 miRNAs from MiR-379/656 cluster can be used for the diagnostic and prognostic purpose of breast cancer and its subtypes—especially in Basal and Luminal B.

Keywords: MiR-379/656; 14q32; breast cancer; prognosis; diagnosis; miRNA; biomarker

Citation: Lal, M.; Ansari, A.H.; Agrawal, A.; Mukhopadhyay, A. Diagnostic and Prognostic Potential of MiR-379/656 MicroRNA Cluster in Molecular Subtypes of Breast Cancer. *J. Clin. Med.* **2021**, *10*, 4071. https://doi.org/10.3390/jcm10184071

Academic Editor: Marco Fiorillo

Received: 8 May 2021
Accepted: 27 August 2021
Published: 9 September 2021

1. Introduction

Breast cancer is characterized by uncontrolled proliferation in the lobules, ducts or in the fatty and the fibrous connective tissue within the breast. WHO administered International Agency for Research on Cancer (IARC) data show that breast cancer leads to the

highest number of cancer cases and deaths in women [1]. This emphasizes the constant need of development of treatment and surveillance strategies for better disease management.

MiRNAs, 20–22 nucleotides in length, are epigenetic regulators that usually function as repressors of gene expression [2]. MiRNAs have been known to play important regulatory roles in disease pathogenesis, particularly in cancer [3–5]. Moreover, miRNAs affect many cancer-related processes including proliferation, cell cycle control, apoptosis, differentiation, migration, metabolism and stress response [6–8]. One of the first reports implicating the involvement of miRNAs in cancer came in early 2000 where miR-15a and miR-16-1 at 13q14 were found to be downregulated or deleted in approximately 68% of the patients suffering from B-cell chronic lymphocytic leukemia [9]. Since then, a number of miRNAs have been discovered to play important regulatory roles in different types of cancer including, breast, colon, gastric, lung, prostate and thyroid [10,11]. Besides, tissue specificity, stability and easy detection in bodily fluids such as blood, serum, and urine provide powerful opportunities for development of miRNA-based biomarkers in cancer diagnostics and therapeutics [12]. For example, a recent report showed that the expression of miR-23a-3p, miR-130a-5p, miR-144-3p, miR-148a-3p and miR-152-3p, in plasma, can be used as biomarkers for early diagnosis of breast cancer [13].

Distribution of miRNAs in the genome is not random, and genomic clusters are frequently seen [14,15]. There are currently 132 such genomic clusters that contain over 20% of known miRNAs [16]. MiRNA clusters are defined as miRNA genes located within 10 Kb of distance on the same chromosome and in the same strand of DNA [17]. MiRNA clusters are frequently co-regulated and co-expressed, targeting multiple mRNA/Protein within the same or similar pathways [18–21]. MiR-17~92 at chromosome 13q31 shows over-expression in different malignancies including breast cancer, lung cancer, B-cell lymphomas and acute lymphoblastic leukemia [22–25]. MiRNA members of MiR-17~92 target mRNAs in a coordinated manner to promote proliferation, increase angiogenesis and sustain cell survival [26–28]. Similarly, MiR-221/222 targets TRPS1, a member of the GATA family of transcriptional repressors in breast cancer. TRPS1 inhibits epithelial to mesenchymal transition (EMT) by directly inhibiting expression of ZEB2 [29]. MiR-99a~125b cluster represses many mRNAs of TP53, Erb and MAPK signaling pathways in multiple myeloma cells [30]. There has been extensive research on miRNAs implicated in cancers; however, studies on large miRNA clusters in cancer are rare. MiR-379/656 on chromosome 14q32 is the second largest miRNA cluster in humans (~45 Kb) and harbors 42 miRNA encoding genes. MiR-379/656 is situated within the imprinted domain, DLK1-DIO3, unique to placental mammal lineage, and shows enriched expression in brain and placenta [31]. The cluster plays an important role in growth and development [32] and is implicated in various malignancies [33–37] including breast cancer [34,38,39]. These reports have been based on a limited number of samples or a limited number of candidate miRNAs. To date, the diagnostic and prognostic potentials of the entire MiR-379/656 cluster as a biomarker in breast cancer and its subtypes remain unknown. We present a comprehensive analysis of miRNA and mRNA next generation sequencing data that addresses the clinical and biological relevance of MiR-379/656 in breast cancer and its molecular subtypes.

2. Materials and Methods

2.1. Genomic Annotation of Clustered miRNAs

The genomic locations of the miRNAs were downloaded from miRBase v18 [40]. MiRNAs within 10 kb of distance, on the same chromosome and on the same strand, were clustered together. We identified 431 miRNAs to be nested within 132 miRNA clusters (Supplementary Material Table S1).

2.2. Downloading and Preprocessing TCGA Data

The miRNA and mRNA expression data of TCGA BRCA and the associated clinical information were downloaded using the R package, TCGAbiolinks [41]. Samples with median tumor purity < 0.6 were filtered out, as per the recommended threshold [42].

There were 1870 miRNAs in total, which were measured across 993 samples. Of which, 480 miRNAs with total read count <10 were removed to improve the accuracy of the analysis (Supplementary Material Table S2). The PVCA module of R package, ExpressionNormalizationWorkflow [43] was used to determine the proportion of variance contributed by different sources of plausible biological (Supplementary Material Table S3) and technical variations (Supplementary Material Table S4). For each biological feature, the contributed variation was <10% (Supplementary Material Figure S1). The residual variation was 86.5%, which might be due to inter-individual differences. In case of technical features, each confounding variables contributed to <10% variation in miRNA expression (Supplementary Material Figure S2). The residual variation was 81.3%, which might be due to inter-individual differences. Survival data and molecular subtype information of TCGA BRCA patient samples were obtained from UCSC Xena browser [44].

2.3. Differential miRNA Expression Analysis in Breast Cancer

The curated miRNA expression matrix consisted of 1390 miRNAs and 993 samples (903 primary tumors and 90 normal controls). The primary tumors were further classified on the basis of their molecular subtypes (PAM50)—Basal (136), Her-2 (59), Luminal A (491) and Luminal B (179). Prior to differential expression analysis, the relationship between sample condition (tumor and normal) and different environmental variables was determined using Fisher's exact test and T-test (Supplementary Material Table S4). It was observed that there was a significant difference between the number of tumor and normal samples sequenced on different Illumina platforms (p-value < 0.01). To determine if this association had an impact on miRNA expression, the data quality was assessed by sample clustering (Supplementary Material Figure S3). The samples formed two distinct clusters with sample condition (tumor, normal). However, with sequencing platform (Illumina Hi-Seq and Illumina GA) the two separate clusters were difficult to define, implying no major contribution to the differential expression. The differential miRNA expression was analyzed using the R package, DESeq2 [45]. The expression values were normalized for sequencing depth and library composition using variance stabilizing transformation and were used for all the subsequent analysis. The significant differences in the expression between different tumor and normal controls were assumed at absolute log2 fold change $\geq$0.6 and adjusted p-value < 0.05. For clustered miRNAs, the significance of differential expression of clustered miRNAs were determined by two-proportions z-test using base R function, "prop.test". The differential expression of the miRNA cluster was compared to a standard defined by the complement of that cluster, i.e., miRNAs not in the cluster.

2.4. Logistic Regression Analysis of MiR-379/656 Expression

The expression matrix was randomly divided into training data (70%; 633 primary tumors and 63 normal samples) and test data (30%; 270 primary tumors and 27 normal samples), using R package, caret [46]. Univariate binomial logistic regression models were built on training data using the base R function; "glm" and miRNAs where the p-value of regression model was <0.05 were removed. Then, lasso regularized logistic regression with 10-fold cross validation was performed on training data with the remaining miRNAs using R package, glmnet [47], to select lambda (λ) parameter with minimum prediction error. Final lasso regularized logistic regression model was built with lambda (λ) where the prediction error was minimum and excluded the miRNA variables with minor contributions (see results for details). The test data were used for subsequent model diagnostics. The area under curve (AUC) of receiver characteristics operating (ROC) curve was used to estimate the accuracy of the model over all possible thresholds. Confusion matrix obtained at optimal threshold was used to compute the sensitivity and specificity of the model. This process of building classification model and its subsequent evaluation were repeated 10 times with different sets of training and test data obtained via resampling. MiRNAs prioritized in 8 out 10 repetitions were further used to build separate regression models with respect to different molecular subtypes of BRCA.

2.5. Cox (Proportional Hazards) Regression Analysis of MiR-379/656

Principal component analysis (PCA) was performed to reduce the dimensions of MiR-379/656 expression using the base R function, "prcomp". The first principal component (PC1), which captured maximum variation in the data, was used to represent the MiR-379/656 meta-expression (Supplementary Material Figure S4). The meta-expression values were then examined for differences among different clinical features of breast cancer (tumor grade, tumor size/extent and molecular subtypes) using Mann–Whitney U-Test. The overlap of patients between tumor grade, tumor size/extent and molecular subtypes was determined using UpSetR plot [48]. No overlap was observed between a particular tumor grade and tumor size/extent with molecular subtype (Supplementary Figures S5 and S6).

For survival analysis, patients' follow-up information was used where overall survival (OS) was defined as the time starting from the date of diagnosis or treatment start until the time of death [49]. The association with patient survival was evaluated using univariate Cox Proportional-Hazard regression analysis. The models were adjusted for all clinical features that showed association with MiR-379/656 meta-expression. Further, Kaplan–Meier curves along with log-rank tests were used to determine statistical differences between the survival of high and low groups, defined by the median expression value as cutoff. The survival analysis was performed using R packages, survival [50,51] and survminer [52].

2.6. Functional Enrichment Analysis of Validated Gene Targets of MiR-379/656

The list of validated gene targets of MiR-379/656 were obtained using miRTarBase 7.0 [53]. To improve the reliability of the analysis, only 901 primary tumor samples with paired miRNA and mRNA expression were considered. Since miRNAs negatively regulate the expression of cognate target genes, anti-regulated miRNA-mRNA pairs were identified using R package, psych [54]. MiRNA-mRNA pairs with spearman correlation coefficient ≤ -0.2 and adjusted p-value < 0.05 were shortlisted to understand the crosstalk among the deregulated pairs of miRNA-mRNA. Interaction networks were plotted using R package, ggraph [55] and tidygraph [56]. The significantly correlated mRNAs of anti-regulated miRNA-mRNA pairs were used for functional enrichment analysis using the R interface of REACTOME pathway database, ReactomePA [57] and R package, clusterProfiler [58]. Pathways with adjusted p-value < 0.05 were considered as significantly enriched.

2.7. Data Processing and Statistical Analysis

Data processing and all statistical analyses mentioned above were performed using programming software, R Foundation For Statistical Computing, Vienna, Austria (version 4.0.3). Statistical significance was assumed if $p < 0.05$. Statistical plots and other data visualizations were generated using the R package, ggplot2, unless specified otherwise [59].

3. Results

3.1. MiR-379/656 Is the Most Significant Differentially Expressed Cluster in Breast Cancer

The differential expression analysis for 1390 miRNAs (miRnome) revealed 310 upregulated (22.3%) and 176 downregulated (12.66%) miRNAs in breast cancer tumor samples compared with normal controls (Figure 1A and Supplementary Material Table S6). In case of MiR-379/656, 32 miRNAs (32/42; 76%) were downregulated (Supplementary Material Figure S7). Similar observations were made when differential expression was studied in 68 paired samples (Supplementary Material Table S6). MiR-379/656 downregulation (confidence intervals: 51% C.I [1]–80% C.I [2]) compared with the 11% downregulation of all miRNAs outside the cluster (1348 miRNAs, used as control) revealed that the proportions were significantly different ($p < 10^{-30}$) (Figure 1B). This indicated that downregulation of the MiR-379/656 cluster in breast cancer was biologically relevant and not attributable to non-specific downregulation of miRNAs.

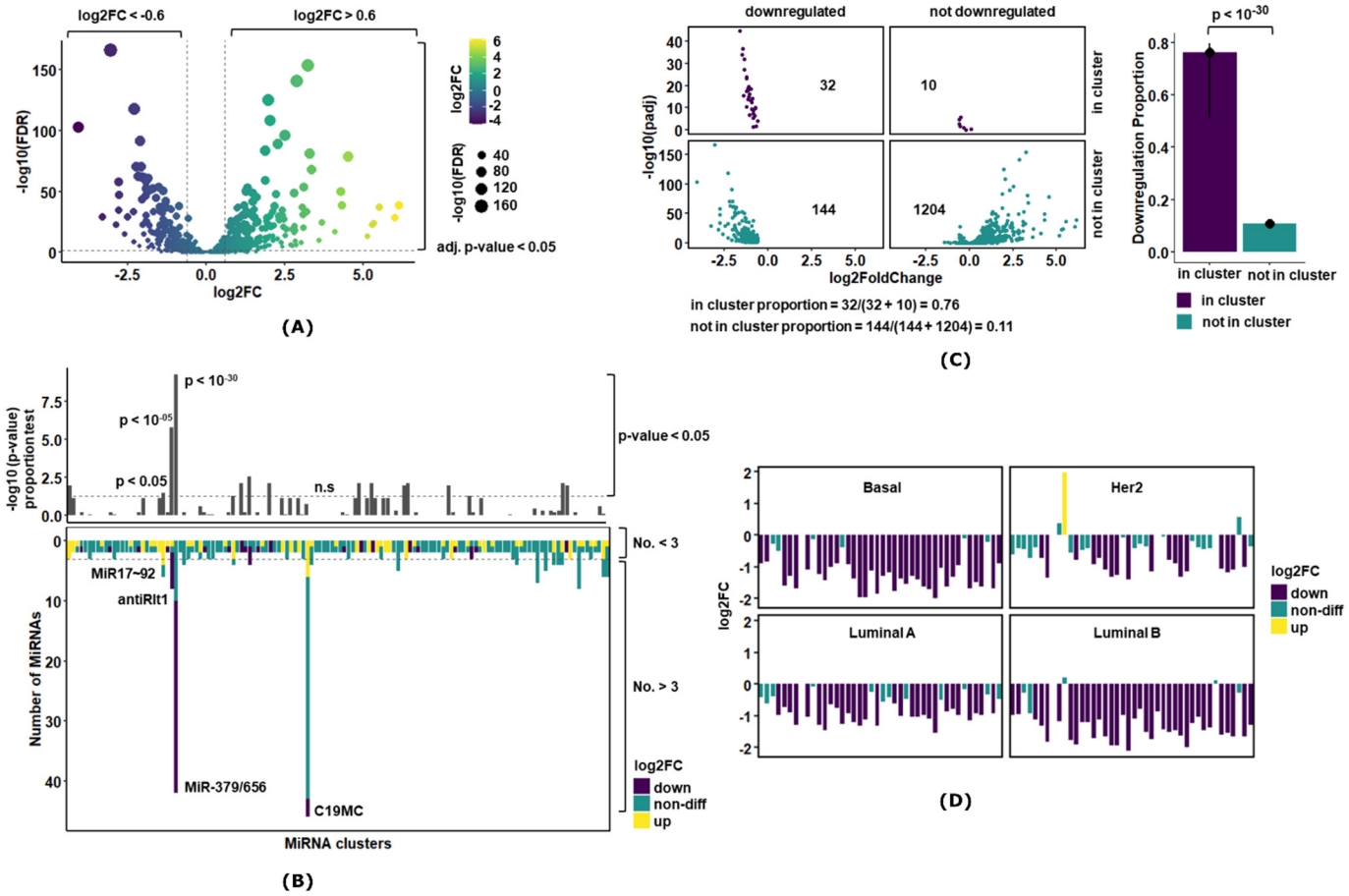


Figure 1. Differential miRNA expression analysis in breast cancer. (**A**) The volcano plot showing miRNA with large fold changes (*x*-axis) along with statistical significance (*y*-axis) in breast cancer. The dotted horizontal line denotes *p*-value of 0.05, and the dotted vertical lines denote absolute log2FC > 0.6. The color represents the biological significance, and size represents statistical significance. The upregulated genes are toward right and downregulated genes are toward the left of the volcano plot. (**B**) The plot represents the 2 × 2 contingency table that compares the proportion of downregulation of miRNAs in the cluster with the standard defined by its complement, i.e., downregulation of miRNAs not in the cluster. The bar-plot represents the downregulation proportions, and the vertical line on the bar represents the 95% C.I. The value above the bars represents the *p*-value of the proportion test. (**C**) The plot represents the proportion test for all 132 clustered miRNAs in the genome. The upper panel shows the *p*-value of the proportion test, and the dotted horizontal line denotes *p*-value of 0.05. The lower panel shows the distribution of differential expression of miRNAs in each cluster where the dotted horizontal represents clusters with 3 or more miRNAs. (**D**) The bar-plot showing miRNA with large fold changes along with statistical significance in molecular subtypes of breast cancer. The downregulated genes are represented by purple color, and upregulated genes toward right are denoted by yellow color.

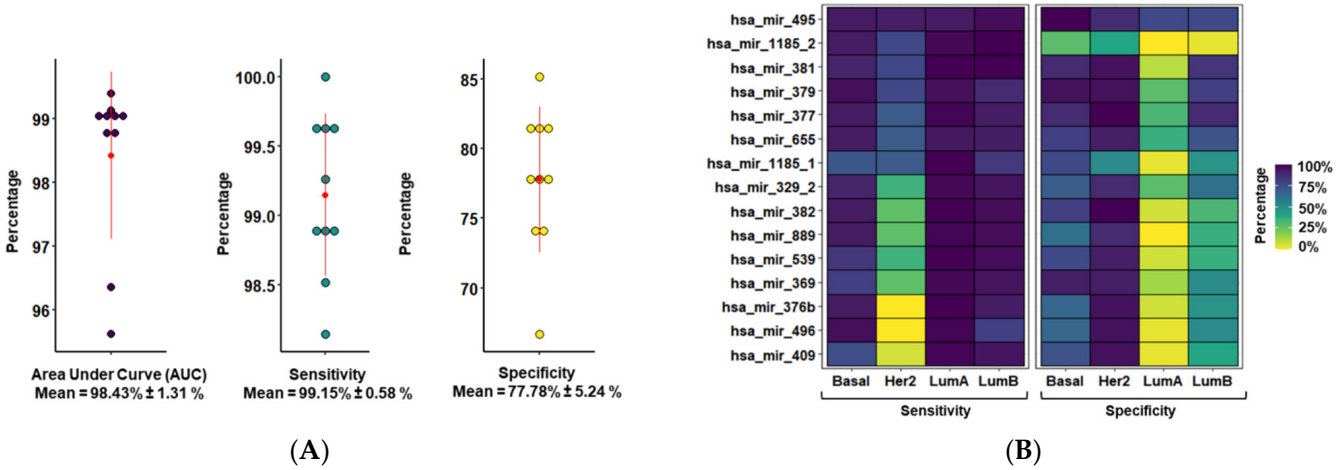
Next, we examined if downregulation of the entire MiR-379/656 cluster was also noticeable in specific molecular subtypes of breast cancer. Differential expression analysis of each subtype against the controls revealed Basal and Luminal B to have at least 83% (35/42) of the miRNAs to be significantly downregulated (Figure 1D; Supplementary Material Table S10). In Her-2 and Luminal A subtypes, the downregulation of miRNAs encoded in miR-379/656 was noted at 43% and 69% respectively. Interestingly, in Her-2 subtype, MiR-323b was upregulated explaining the molecular basis of significant upregulation of MiR-323b with paired samples (Supplementary Material Table S7).

For a comparative and comprehensive analysis of all clustered miRNAs in breast cancer, we compared the expression of MiR-379/656 against all the miRNA clusters (miRBase V18; [40]). We found 14 clusters apart from MiR-379-656 that showed significant over-representation of dysregulated miRNAs (Figure 1C, Supplementary Material Tables S8 and S9). MiR-379/656

was found to be the most significant differentially expressed cluster in the human genome. Amongst others, MiR17~92 on chromosome 13 has been reported as a polycistronic oncomir cluster with coordinated function to promote proliferation, increase angiogenesis and sustain cell survival [60]. The cluster showed 66% upregulation ($p < 0.05$) in our analysis and served as a positive control for the analysis pipeline (Figure 1C, Supplementary Material Tables S8 and S9). Within chromosome 14 and within the imprinted domain DLK1-DIO3, another cluster, anti-Rtl1, also showed 75% downregulation ($p < 10^{-05}$) (Figure 1C, Supplementary Material Tables S8 and S9). Significant downregulation at a genome-wide significance level for both clusters on chromosome 14q suggested disruption of epigenomic features regulating imprinting at DLK1-DIO3 in breast cancer. The largest cluster on the human genome, the imprinted miRNA cluster on chromosome 19, with 46 miRNA encoding genes, showed non-differential expression in breast cancer (Figure 1C, Supplementary Material Tables S8 and S9)—serving as the negative control. Thus, our analysis indicated that the potential epigenetic dysregulation for the MiR-379/656 cluster was specific and relevant for breast cancer pathogenesis and progression. In the subsequent sections, we further analyze the clinical and biological significance of this observation.

3.2. MiR-379/656 Accurately Classifies Tumor and Normal Samples—Especially Basal and Luminal B Subtypes

To understand the clinical relevance of MiR-379/656 downregulation, the prognostic potential of MiR-379/656 for accurate classification of tumor and normal samples was evaluated. First, using univariate logistic regression on training data, miRNAs unlikely to contribute to the pathophysiology were removed ($p > 0.05$; Supplementary Material Table S11). For the remaining miRNAs, most showed an association of lower expression with the disease/tumor (negative β coefficient). This was expected since the cluster was downregulated in breast cancer. After removing the miRNA variables based on the lambda value from 10-fold cross validation (Supplementary Material Table S12), the model prioritized 15 miRNAs in 8 out of 10 repetitions (Supplementary Material Figure S8). The expression level of these 15 miRNAs were able to correctly identify breast cancer patients from controls with 98.43 ± 1.31% accuracy (AUC of ROC; Figure 2A). The sensitivity and specificity to distinguish breast cancer tumors from normal controls was 99.15 ± 0.58% and 77.78 ± 5.24%, respectively.



(**A**) (**B**)

Figure 2. Prognostic potential of MiR-379/656 in breast cancer. (**A**) The plot shows distribution of AUC, sensitivity and specificity of classification models built using different sets of training and test data ($n = 10$). The error bar shows the mean and standard deviation. (**B**) The tile plot showing the sensitivity and specificity measure of 15 miRNAs at MiR-379/656 as determined by univariate binomial logistic regression model across molecular subtypes of breast cancer.

We also evaluated the prognostic potential of the subset of 15 prioritized miRNAs at MiR-379/656 for accurate classification of different molecular subtypes (Supplementary Material Table S13; Figure 2B). The sensitivities of each molecular subtype were 98.70 ± 1.62%, 95.10 ± 4.50%, 91.70 ± 6.24% and 56.5 ± 2.62% for Luminal A, Luminal B, Basal and Her-2, respectively. Whereas the specificity measure of Her-2, Basal, Luminal B and Luminal A was observed at 87.90 ± 1.74%, 77.00 ± 1.73%,54.80 ± 2.36% and 17.80 ± 2.36%, respectively. MiR-495 had the best trade-off between sensitivity (95.48 + 1.81%) and specificity (86.11 ± 10.64%) among all molecular subtypes of BRCA. Overall, the subset of 15 miRNAs at MiR-379/656 predicted tumor outcome with high accuracy—especially for Basal and Luminal B molecular subtypes.

3.3. MiR-379/656 Is Associated with Poor Clinical Outcome in Breast Cancer

We studied the association of 15 prioritized miRNA expression with different clinical features of breast cancer (Figure 3A). The meta-expression of these 15 miRNAs showed decreased expression in Basal and Luminal B subtypes of breast cancer. Further, the expression decreased with increasing tumor stage and tumor size/extent. For survival analysis, the hazard ratio was adjusted for molecular subtype, tumor size and tumor stage to negate the effect of different clinical features on patient survival. The adjusted hazard ratio (HR_{adj}) of most of the 15 miRNAs in breast cancer was <1, implying that decreased expression was associated with worse patient outcome. However, the inter-tumor variability in the miRNA expression profiles revealed a wider spread (95% CI) of the hazard ratio—resulting in marginal significance for most of the candidate miRNAs (Figure 3B). Notable exceptions were miR-487a, miR-889 and miR-379. Kaplan–Meier survival plots revealed significantly worse patient outcomes associated with lower expressions for these miRNAs (miR-487a, miR-379 and miR-494; Figure 3C). These findings implicated that decreased MiR-379/656 expression was associated with poorer clinical outcome.

3.4. MiR-379/656 Target Genes Are Enriched for Cancer-Relevant Pathways

To understand the biological significance of altered expression of the 15 prioritized miRNAs in breast cancer, the experimentally validated mRNA targets were annotated via pathway enrichment analysis. The correlation analysis between these 15 miRNAs and their target mRNAs revealed possible miRNAs-mRNAs regulatory networks (Supplementary Material Table S14). Of the 15 miRNAs, 12 miRNAs were found to be negatively correlated with 103 of its validated targets across 114 interactions, among which MiR-410, MiR-889, MiR-377 and MiR-381 were identified as hub-miRNAs to interact with at least 15 mRNA targets (Figure 4A). Next, the target mRNAs were screened for enrichment of REACTOME signaling pathways. We observed 21 REACTOME signaling pathways to be enriched for the target mRNAs of MiR-379/656 (Supplementary Material Table S15). Cancer-related pathways, especially DNA repair, transcriptional regulation by p53 and cell cycle checkpoints were among the most significantly enriched pathways (Figure 4B). This implied that a cluster-wide downregulation of MiR-379/656 can trigger perturbations of entire mRNA/protein network and regulate oncogenic signaling pathways in a coordinated manner.

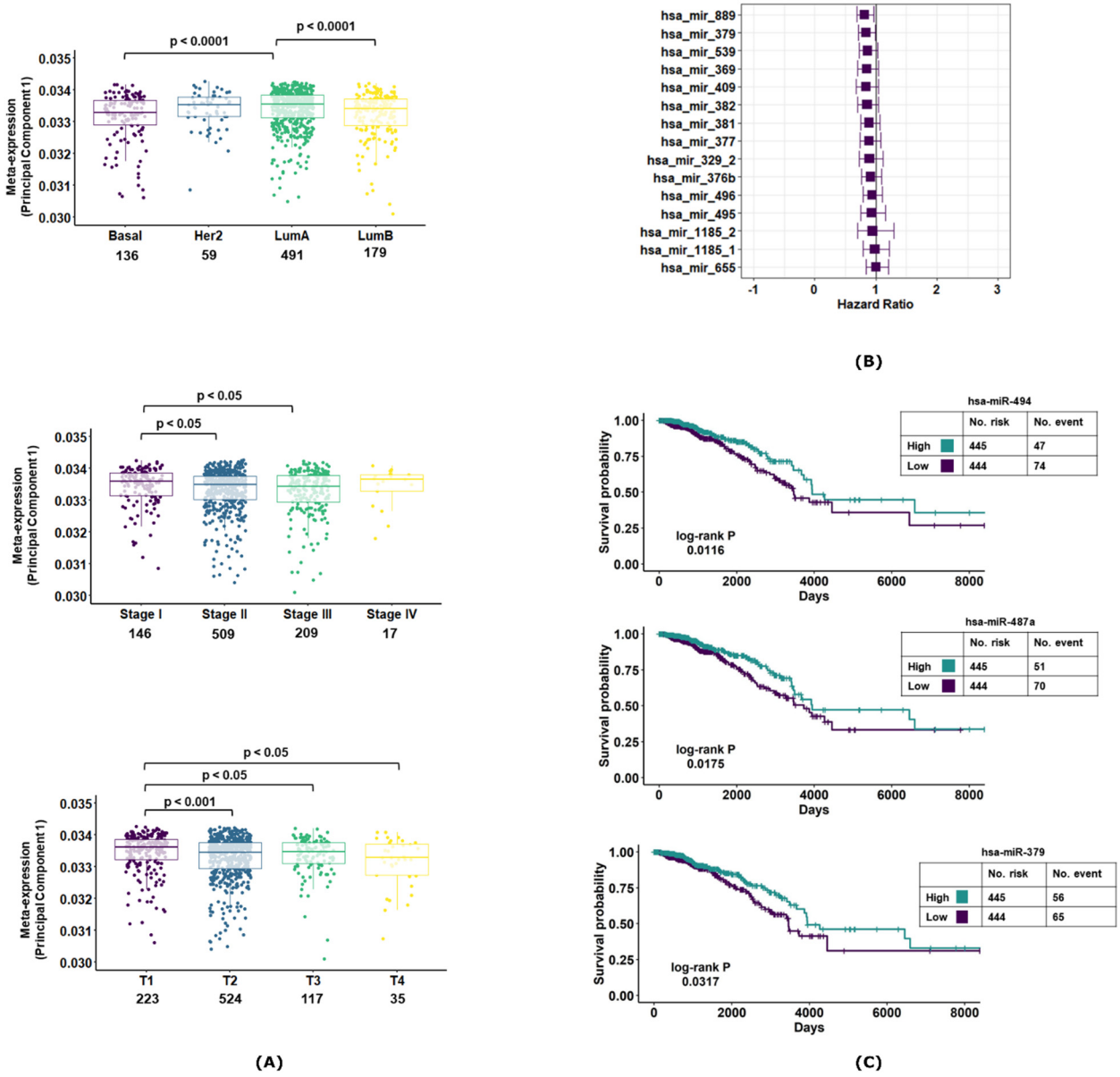


Figure 3. Correlation of MiR-379/656 with clinical features in breast cancer. The boxplot showing the meta-expression of 15 prioritized miRNAs at MiR-379/656 defined by the first principal component across different (**A**) molecular subtypes, tumor stages and tumor size/extent of breast cancer. The value above the boxplot represents the *p*-value of Mann–Whitney U-Test of different group comparisons. (**B**) The forest plot represents the hazard ratio along with 95% C.I. of the subset of 15 miRNAs at MiR-379/656. (**C**) The Kaplan–Meier survival curves of MiR-487a, MiR-379 and MiR-494. The patients in high- and low-expression groups are determined by the median expression value cut-off. The log-rank *p* denotes the *p*-value of the test comparing the differences in the distribution of survival in the two groups.

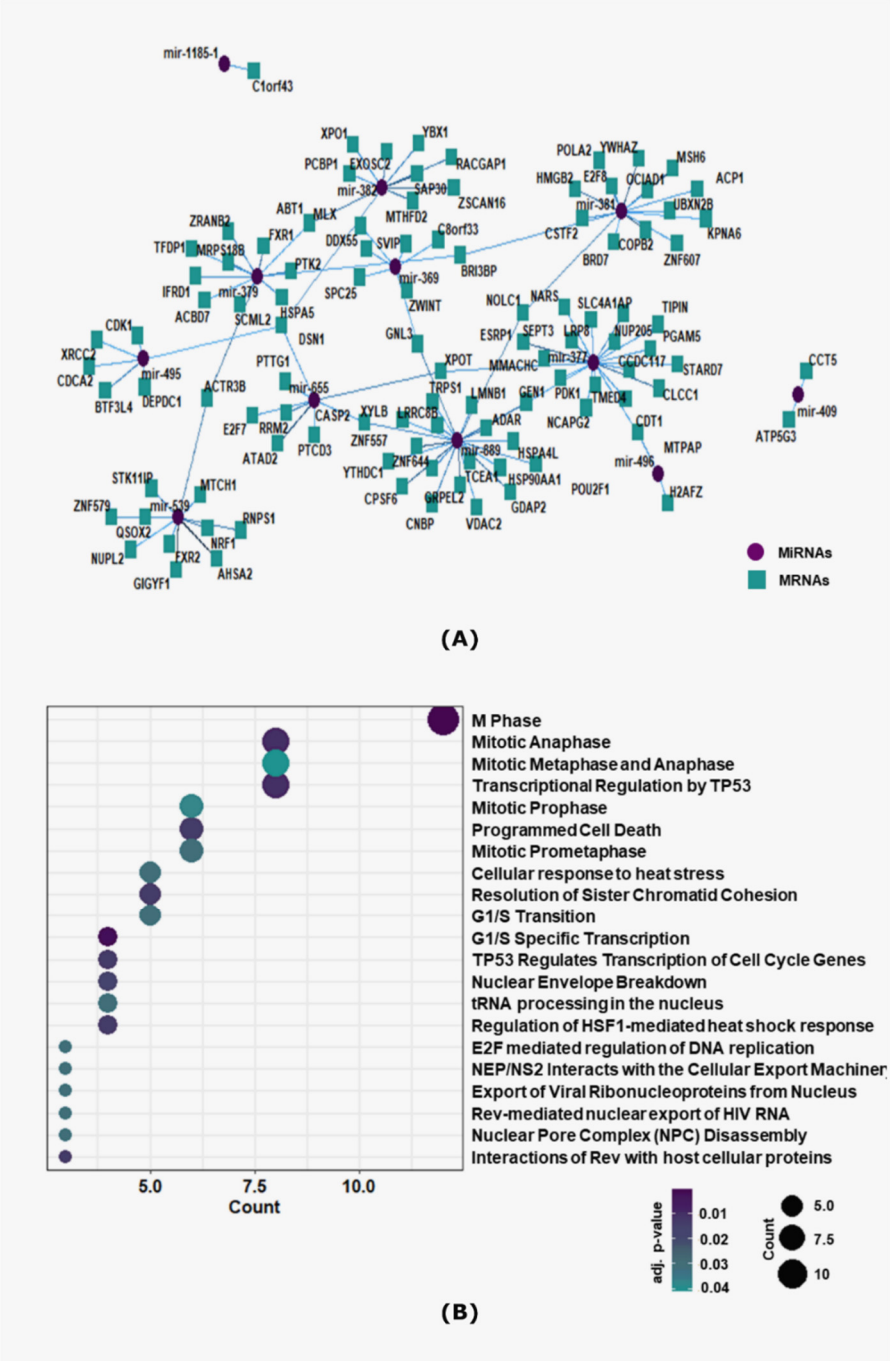


Figure 4. Pathway enrichment of target mRNAs of MiR-379/656 in breast cancer. (**A**) Interaction network plot of MiR-379/656 and negatively correlated target mRNAs. The purple-circled node represents the miRNAs, and the blue-squared nodes represent the mRNAs connected via edges representing interactions. (**B**) The dot-plot showing the top 20 statistically significant REACTOME pathways enriched for anti-regulated target mRNAs. Each point represents a pathway, arranged according to the overlapping count of genes, and the color of the point indicates −log10 adjusted *p*-value.

4. Discussion

MiR-379/656, the second largest miRNA cluster in the human genome, has been implicated in tumor progression or poor clinical outcome suggesting its potential as a cancer biomarker. In a smaller study, we had shown 46% downregulation of the cluster in 80 breast cancer tumor samples compared with controls [34]. This observation was

supported by others including linking the expression of MiR-379/656 to epithelial mesenchymal transition, and stemness [39]. Candidate miRNAs from the cluster, namely, miR-127-5p, miR-544a and miR-655-3p were shown to inhibit metastasis development in an animal model of breast cancer lung colonization [38]. The data we presented in this report are the most comprehensive analysis to date, of the miR-379/656 cluster as a diagnostic and prognostic marker for breast cancer. Our unbiased genome-wide expression analysis of mRNA and miRNA in almost 1000 tumor samples shows that the cluster is profoundly dysregulated, with 76% of the miRNAs significantly downregulated in tumor samples compared with controls. In the entire dataset, there were 68 paired samples (tumor and normal tissue from the same patient), and a separate analysis of that small subset also revealed the same cluster-wide downregulation. Further, we identified a subset of 15 miRNAs that can be used for diagnostic and prognostic purposes of breast cancer and its subtypes—especially for Basal and Luminal B (83% and 86% of the clustered miRNA downregulated, respectively). As expected, molecular subtypes with greater downregulation of MiR-379/656 had a better 'trade-off' between specificity and sensitivity. The expression level of the 15 miRNA panel could diagnose breast cancer samples from controls with 99.15% sensitivity and 77.78% specificity (Figure 2A). Among the four main subtypes, all except Her-2 showed more than 91% sensitivity, while Her-2 showed the highest specificity (87.90%). It is possible that the 15 prioritized miRNAs were not sufficient to capture the molecular characteristics for identification of the HER2 subtype. In addition, the use of multinomial classification instead of binomial logistic regression to classify the molecular subtypes might have yielded different predictions. We chose to use the binomial logistic regression because of the unequal proportion of samples (Basal-136, Her-2-59, Luminal A-491 and Luminal B-179). This would have resulted in a 'class imbalance' in a multinomial model, which would eventually affect the accuracy of the prediction.

We previously reported the expression level of the MiR-379/656 cluster to be a predictor of the tumor grade for brain cancer (Gliomagenesis; [33])—where lower expression of the cluster was correlated with higher tumor grade. In the same study, we also showed that higher miRNA expression was associated with better survival. In this report, we observe the same trend with breast cancer and its subtypes. The decreased expression of the 15 miRNA panel correlated with increasing tumor stage, increasing tumor size/extent and Basal and Luminal B molecular subtypes (Figure 3A). We observed that the decreased miRNA expression was associated with worse patient outcome, albeit with marginal significance. However, a large amount of patient information was lost during follow up (right censoring), which may have impacted the survival analysis. Taken together, this necessitates further investigation using independent datasets. Our study identified MiR-495, MiR-379 and MiR-487a as potential biomarkers for breast cancer. In our analysis, MiR-495 had more than 94% sensitivity and more than 77% specificity in each molecular subtype of breast cancer. MiR-495 is reported to be dysregulated in several cancers, including breast cancer. It modulates the transition of G1 to S phase by regulating the expression of protein Bmi-1 in breast cancer [61]. The long non-coding RNA, SNHG20, regulates HER2 via miR-495 and participates in proliferation, invasion and migration of breast cancer cells [62]. MiR-495 targets STAT-3 to inhibit cell proliferation and migration and to induce apoptosis in breast cancer [63]. The observed downregulation of miR-495 can thus enhance proliferation and migration and reduce apoptosis for breast cancer cells. MiR-379 is reported to have decreased expression in breast cancer patients compared with normal controls [64]. MiR-379 expression showed positive correlation with increasing tumor stage and expression of Cyclin B1. MiR-487a is not directly implicated in breast cancer; however, the miRNA shows deregulation in other cancers [65,66]. These reports support our findings and provide interesting opportunities for the development of miRNA-based biomarkers for early detection of breast cancer.

MicroRNAs are higher order epigenetic regulators. Owing to the imperfect complementarity of the miRNA and the target mRNA sequences, individual miRNAs can bind and regulate multiple mRNAs [67]. Thus, altered expression of one or few miRNAs can

alter the expression of hundreds of mRNAs, and the cascading effect can create havoc in biological systems. As shown in Figure 4B, the subset of 15 miRNAs are expected to perturb pathways relevant for cancer pathophysiology. Cluster-wide downregulation of MiR-379/656 and anti-Rtl1 cluster suggested that epigenomic features responsible for maintenance of imprinting at DLK1-DIO3 might be disrupted in breast cancer. The loss of imprinting in tumors is unexpectedly high, rendering it as one of the most common molecular mechanisms in cancer [68,69]. Corroborative evidence suggests that imprinting defects alter the expression of MiR-379/665 [70–72]. We have previously shown that downregulation of MiR-379/656 in glioblastoma multiforme (GBM) was consistent with hyper methylation of the locus [34]. In lung cancer, hypo methylation of DLK1-DIO3 locus has been correlated with overexpression of miRNAs encoded in the cluster [73,74]. For MiR-379/656 cluster, spanning ~45 Kb in the genome, it will be interesting to explore the possible mechanisms responsible for the said dysregulation in cancers.

This study relies on data gathered from a single source (TCGA), which may influence the findings. Thus, additional data mining and analysis are required to confirm consistency with the TCGA results, and the results are needed to be validated in cell lines and/or patient samples. In future, non-invasive or minimally invasive methods relying on liquid biopsies or circulating tumor cells can be considered for the development of MiR-379/656-based breast cancer diagnostics and/or prognostics. Future studies can determine whether MiR-379/656 and/or the cognate mRNA targets implicated in "cancer-related pathways" can be targeted for reversal of phenotypes.

5. Conclusions

Aberrant expression of miRNAs leads to disease pathogenesis, particularly cancer. This effect may be more pronounced for miRNA clusters. Here, we report a subset of miRNAs at the MiR-379/656 cluster, which can be used for diagnostic and prognostic purposes of breast cancer and its subtypes—especially Basal and Luminal B.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/jcm10184071/s1. Figure S1: Estimation of proportion of variation in miRNA gene expression contributed by known biological features. Information related to smoking history, alcohol history, BMI, height and weight was either missing or not reported. The bar-plot shows the weighted proportion of variance (y-axis) across different confounding variables. Figure S2: Estimation of proportion of variation in miRNA gene expression contributed by known technical variables. Each sample in the TCGA project is assigned a barcode. For example, TCGA-02-0001-01C-01D-0182-01 where "02" refers to the site from where the sample is obtained (tissue source site or TSS); "0001" refers to the serial number participant of the study, "01C" where 01 refers to tumor (sample) and C refers to the third vial, "01D" first portion of DNA sample (portion), "0182" refers to plate number (plate) and "01" refers to the center where sequencing or characterization was performed. In addition to sequencing platform, relevant information related to other possible batch effect (TSS, sample, portion, plate, center) was extracted from the TCGA barcode. Since all of the samples were processed at Canada's Michael Smith Genome Sciences Centre, the center information was excluded from the analysis. The bar-plot shows the weighted proportion of variance (y-axis) across different confounding variables. Figure S3: Data quality assessment by sample clustering. Principal component analysis (PCA) plot, which shows samples clustered along their first two principal component analysis when (A) colored by sample condition and (B) colored by sequencing platform. Data from tumors and controls form two distinct clusters (left panel), whereas data generated in different sequencing platforms do not contribute majorly to the differential analysis (right panel). Figure S4: Scree-plot showing eigenvalues for each individual principal component of 15 prioritized miRNAs. Figure S5: UpSetR plot showing different intersection sets between tumor stages and molecular subtypes. Figure S6: UpSetR plot showing different intersection sets between tumor size/extent and across molecular subtypes. Figure S7: Heat map showing the differential expression of 41 miRNAs at MiR-379/656 in primary tumor and normal controls. Each row represents a miRNA and each column denotes a sample. MiR-300 was not captured by in the current assay. The color gradient indicates the differences in the log2 normalized miRNA expression. The marginal bar plot represents log2 fold change where the dotted vertical line denotes log2FC of −0.6. The

most significant downregulated miRNAs (MiR-329-2, MiR-495, MiR-299, MiR-377, MiR-381 and MiR-379) were highlighted. Figure S8: Recurrence of miRNAs prioritized in lasso regularized classification models in different resampled datasets. The height of the histogram shows the count of recurrence of miRNA, and the color denotes the resampling number. Table S1: List of clustered miRNAs in human genome. Table S2: MiRNA expression across TCGA-BRCA samples. Table S3: Information related to batch effects of TCGA-BRCA samples. Table S4: Clinical information of TCGA-BRCA samples. Table S5: Relationship between sample condition (tumor, normal) and different environmental factors. Fisher's exact test is performed for categorical variables and T-test for continuous variables. Table S6: Differential expression analysis of miRNA expression in breast cancer across 903 primary tumors and 90 normal controls. Table S7: Paired differential expression analysis of MiR-379/656 expression in BRCA. Table S8: Two-proportion z-test comparing the proportion of downregulated clustered miRNAs against proportion of downregulated miRNAs not in cluster. Table S9: Two-proportion z-test comparing the proportion of upregulated clustered miRNAs against proportion of upregulated miRNAs not in cluster. Table S10: Differential expression analysis of miRNA expression in breast cancer across different molecular subtype of breast cancer and 90 normal controls. Table S11: Coefficient of univariate binomial logistic regression model of contributing miRNA variables in breast cancer. Table S12: Coefficient of penalized binomial logistic regression model with lasso regularization. Table S13: Coefficient of univariate binomial logistic regression model of contributing miRNA variables in breast cancer molecular subtypes. Table S14: Correlation coefficient between 15 prioritized miRNAs at MiR-379/656 and their validated targets. Table S15: List of REACTOME pathways enriched for validated targets of 15 prioritized miRNAs at MiR-379/656 sorted by Gene count.

Author Contributions: A.M. and M.L. conceptualized the study. M.L. carried out most of the data curation, investigation and interpretation with critical assistance from A.H.A. M.L. prepared the original draft and the associated figures and tables. A.M. and M.L. reviewed and edited the manuscript. A.M. and A.A. managed the supervision and project administration and provided the necessary funding and infrastructure. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by multiple grants supporting the computational infrastructure provided by the Council of Scientific and Industrial Research, India (CSIR).

Institutional Review Board Statement: The study involved only publicly available data. The overall project was approved by the institutional review board of CSIR-Institute of Genomics & Integrative Biology, Delhi, India.

Data Availability Statement: All the publicly available data used in this manuscript are available to download at https://portal.gdc.cancer.gov/legacy-archive/search/f (accessed on 7 May 2021) and https://xena.ucsc.edu/ (accessed on 7 May 2021).

Acknowledgments: A.H.A. and M.L. acknowledge fellowship from Indian Council of Medical Research (ICMR), India and University of Grants Commission (UGC), India, respectively. We thank Deepanjan Paul for his technical assistance during manuscript preparation. All authors also acknowledge the data resource available with The Cancer Genome Atlas (TCGA).

Conflicts of Interest: The authors declare no conflict of interest. This manuscript has used publicly available data; hence, there are no ethical concerns.

References

1. Global Cancer Observatory. Available online: https://gco.iarc.fr/ (accessed on 7 May 2021).
2. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993**, *75*, 843–854. [CrossRef]
3. Cho, W.C.S. OncomiRs: The discovery and progress of microRNAs in cancers. *Mol. Cancer* **2007**, *6*, 60. [CrossRef]
4. Hammond, S.M. MicroRNAs as oncogenes. *Curr. Opin. Genet. Dev.* **2006**, *16*, 4–9. [CrossRef]
5. Esquela, K.A.; Slack, J. Oncomirs-microRNAs with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269. [CrossRef] [PubMed]
6. Choudhury, Y.; Tay, F.C.; Lam, D.H.; Sandanaraj, E.; Tang, C.; Ang, B.-T.; Wang, S. Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *J. Clin. Investig.* **2012**, *122*, 4059–4076. [CrossRef] [PubMed]
7. Di Leva, G.; Garofalo, M.; Croce, C.M. MicroRNAs in cancer. *Annu. Rev. Pathol.* **2014**, *9*, 287–314. [CrossRef] [PubMed]
8. Giovannetti, E.; Erozenci, A.; Smit, J.; Danesi, R.; Peters, G.J. Molecular mechanisms underlying the role of microRNAs (miRNAs) in anticancer drug resistance and implications for clinical practice. *Crit. Rev. Oncol. Hematol.* **2012**, *81*, 103–122. [CrossRef]

9. Calin, G.A.; Dumitru, C.D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15524–15529. [CrossRef]
10. Cummins, J.M.; Velculescu, V.E. Implications of micro-RNA profiling for cancer diagnosis. *Oncogene* **2006**, *25*, 6220–6227. [CrossRef]
11. Takamizawa, J.; Konishi, H.; Yanagisawa, K.; Tomida, S.; Osada, H.; Endoh, H.; Harano, T.; Yatabe, Y.; Nagino, M.; Nimura, Y.; et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.* **2004**, *64*, 3753–3756. [CrossRef]
12. Manterola, L.; Guruceaga, E.; Perez-Larraya, J.G.; González-Huarriz, M.; Jauregui, P.; Tejada, S.; Diez-Valle, R.; Segura, V.; Samprón, N.; Barrena, C.; et al. A small noncoding RNA signature found in exosomes of GBM patient serum as a diagnostic tool. *Neuro-Oncology* **2014**, *16*, 520–527. [CrossRef]
13. Li, X.; Zou, W.; Wang, Y.; Liao, Z.; Li, L.; Zhai, Y.; Zhang, L.; Gu, S.; Zhao, X. Plasma-based microRNA signatures in early diagnosis of breast cancer. *Mol. Genet. Genom. Med.* **2020**, *8*, e1092. [CrossRef]
14. Marco, A.; Ninova, M.; Ronshaugen, M.; Griffiths-Jones, S. Clusters of microRNAs emerge by new hairpins in existing transcripts. *Nucleic Acids Res.* **2013**, *41*, 7745–7752. [CrossRef] [PubMed]
15. Mohammed, J.; Siepel, A.; Lai, E.C. Diverse modes of evolutionary emergence and flux of conserved microRNA clusters. *RNA* **2014**, *20*, 1850–1863. [CrossRef] [PubMed]
16. Kozomara, A.; Griffiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **2011**, *39*, D152–D157. [CrossRef]
17. Altuvia, Y.; Landgraf, P.; Lithwick, G.; Elefant, N.; Pfeffer, S.; Aravin, A.; Brownstein, M.J.; Tuschl, T.; Margalit, H. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res.* **2005**, *33*, 2697–2706. [CrossRef]
18. Ventura, A.; Young, A.G.; Winslow, M.M.; Lintault, L.; Meissner, A.; Erkeland, S.J.; Newman, J.; Bronson, R.T.; Crowley, D.; Stone, J.R.; et al. Targeted deletion reveals essential and overlapping functions of the miR-17∼92 family of miRNA clusters. *Cell* **2008**, *132*, 875–886. [CrossRef]
19. Kim, Y.-K.; Yu, J.; Han, T.S.; Park, S.-Y.; Namkoong, B.; Kim, D.H.; Hur, K.; Yoo, M.-W.; Lee, H.-J.; Yang, H.-K.; et al. Functional links between clustered microRNAs: Suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. *Nucleic Acids Res.* **2009**, *37*, 1672–1681. [CrossRef] [PubMed]
20. Yuan, X.; Liu, C.; Yang, P.; He, S.; Liao, Q.; Kang, S.; Zhao, Y. Clustered microRNAs' coordination in regulating protein-protein interaction network. *BMC Syst. Biol.* **2009**, *3*, 65. [CrossRef]
21. Wang, J.; Haubrock, M.; Cao, K.-M.; Hua, X.; Zhang, C.-Y.; Wingender, E.; Li, J. Regulatory coordination of clustered microRNAs based on microRNA-transcription factor regulatory network. *BMC Syst. Biol.* **2011**, *5*, 199. [CrossRef]
22. O'Donnell, K.A.; Wentzel, E.A.; Zeller, K.I.; Dang, C.V.; Mendell, J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **2005**, *435*, 839–843. [CrossRef]
23. Tagawa, H.; Seto, M. A microRNA cluster as a target of genomic amplification in malignant lymphoma. *Leukemia* **2005**, *19*, 2013–2016. [CrossRef] [PubMed]
24. Hayashita, Y.; Osada, H.; Tatematsu, Y.; Yamada, H.; Yanagisawa, K.; Tomida, S.; Yatabe, Y.; Kawahara, K.; Sekido, Y.; Takahashi, T. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res.* **2005**, *65*, 9628–9632. [CrossRef]
25. Mavrakis, K.J.; Wolfe, A.L.; Oricchio, E.; Palomero, T.; de Keersmaecker, K.; McJunkin, K.; Zuber, J.; James, T.; Khan, A.A.; Leslie, C.S.; et al. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nat. Cell Biol.* **2010**, *12*, 372–379. [CrossRef] [PubMed]
26. Mestdagh, P.; Boström, A.-K.; Impens, F.; Fredlund, E.; Van Peer, G.; De Antonellis, P.; von Stedingk, K.; Ghesquière, B.; Schulte, S.; Dews, M.; et al. The miR-17-92 microRNA cluster regulates multiple components of the TGF-β pathway in neuroblastoma. *Mol. Cell* **2010**, *40*, 762–773. [CrossRef]
27. Mu, P.; Han, Y.-C.; Betel, D.; Yao, E.; Squatrito, M.; Ogrodowski, P.; de Stanchina, E.; D'Andrea, A.; Sander, C.; Ventura, A. Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. *Genes Dev.* **2009**, *23*, 2806–2811. [CrossRef] [PubMed]
28. Mendell, J.T. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* **2008**, *133*, 217–222. [CrossRef]
29. Stinson, S.; Lackner, M.R.; Adai, A.T.; Yu, N.; Kim, H.-J.; O'Brien, C.; Spoerke, J.; Jhunjhunwala, S.; Boyd, Z.; Januario, T.; et al. miR-221/222 targeting of trichorhinophalangeal 1 (TRPS1) promotes epithelial-to-mesenchymal transition in breast cancer. *Sci. Signal.* **2011**, *4*, pt5. [CrossRef]
30. Feng, M.; Luo, X.; Gu, C.; Li, Y.; Zhu, X.; Fei, J. Systematic analysis of berberine-induced signaling pathway between miRNA clusters and mRNAs and identification of mir-99a∼125b cluster function by seed-targeting inhibitors in multiple myeloma cells. *RNA Biol.* **2015**, *12*, 82–91. [CrossRef]
31. Glazov, E.A.; McWilliam, S.; Barris, W.C.; Dalrymple, B.P. Origin, evolution, and biological role of miRNA cluster in DLK-DIO3 genomic region in placental mammals. *Mol. Biol. Evol.* **2008**, *25*, 939–948. [CrossRef]
32. da Rocha, S.T.; Edwards, C.A.; Ito, M.; Ogata, T.; Ferguson-Smith, A.C. Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends Genet.* **2008**, *24*, 306–316. [CrossRef]

33. Nayak, S.; Aich, M.; Kumar, A.; Sengupta, S.; Bajad, P.; Dhapola, P.; Paul, D.; Narta, K.; Purkrait, S.; Mehani, B.; et al. Novel internal regulators and candidate miRNAs within miR-379/miR-656 miRNA cluster can alter cellular phenotype of human glioblastoma. *Sci. Rep.* **2018**, *8*, 7673. [CrossRef] [PubMed]
34. Laddha, S.V.; Nayak, S.; Paul, D.; Reddy, R.; Sharma, C.; Jha, P.; Hariharan, M.; Agrawal, A.; Chowdhury, S.; Sarkar, C.; et al. Genome-wide analysis reveals downregulation of miR-379/miR-656 cluster in human cancers. *Biol. Direct* **2013**, *8*, 10. [CrossRef]
35. Zehavi, L.; Avraham, R.; Barzilai, A.; Bar-Ilan, D.; Navon, R.; Sidi, Y.; Avni, D.; Leibowitz-Amit, R. Silencing of a large microRNA cluster on human chromosome 14q32 in melanoma: Biological effects of mir-376a and mir-376c on insulin growth factor 1 receptor. *Mol. Cancer* **2012**, *11*, 44. [CrossRef] [PubMed]
36. Haller, F.; von Heydebreck, A.; Zhang, J.D.; Gunawan, B.; Langer, C.; Ramadori, G.; Wiemann, S.; Sahin, O. Localization- and mutation-dependent microRNA (miRNA) expression signatures in gastrointestinal stromal tumours (GISTs), with a cluster of co-expressed miRNAs located at 14q32.31. *J. Pathol.* **2010**, *220*, 71–86. [CrossRef]
37. Zhang, L.; Volinia, S.; Bonome, T.; Calin, G.A.; Greshock, J.; Yang, N.; Liu, C.-G.; Giannakakis, A.; Alexiou, P.; Hasegawa, K.; et al. Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7004–7009. [CrossRef]
38. Uppal, A.; Wightman, S.C.; Mallon, S.; Oshima, G.; Pitroda, S.P.; Zhang, Q.; Huang, X.; Darga, T.E.; Huang, L.; Andrade, J.; et al. 14q32-encoded microRNAs mediate an oligometastatic phenotype. *Oncotarget* **2015**, *6*, 3540–3552. [CrossRef] [PubMed]
39. Cantini, L.; Bertoli, G.; Cava, C.; Dubois, T.; Zinovyev, A.; Caselle, M.; Castiglioni, I.; Barillot, E.; Martignetti, L. Identification of microRNA clusters cooperatively acting on epithelial to mesenchymal transition in triple negative breast cancer. *Nucleic Acids Res.* **2019**, *47*, 2205–2215. [CrossRef]
40. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. miRBase: From microRNA sequences to function. *Nucleic Acids Res.* **2019**, *47*, D155–D162. [CrossRef] [PubMed]
41. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **2016**, *44*, e71. [CrossRef]
42. Mounir, M.; Lucchetta, M.; Silva, T.C.; Olsen, C.; Bontempi, G.; Chen, X.; Noushmehr, H.; Colaprico, A.; Papaleo, E. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* **2019**, *15*, e1006701. [CrossRef]
43. Karthikeyan, M. *Expression Normalization Workflow: Gene Expression Normalization Workflow*, R package version 1.18.0; R Foundation For Statistical Computing: Vienna, Austria, 2021.
44. UCSC Xena. Available online: https://xena.ucsc.edu (accessed on 7 May 2021).
45. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
46. Kuhn, M. Caret: Classification and Regression Training. R Package Version 6.0-86. 2020. Available online: https://CRAN.R-project.org/package=caret (accessed on 7 May 2021).
47. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [CrossRef]
48. Gehlenborg, N. UpSetR: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets. R Package Version 1.4.0. 2019. Available online: https://CRAN.R-project.org/package=UpSetR (accessed on 7 May 2021).
49. NCI Dictionary of Cancer Terms. Available online: https://www.cancer.gov/publications/dictionaries/cancer-terms (accessed on 7 May 2021).
50. Therneau, T. A Package for Survival Analysis in R. R Package Version 3.2-7. 2020. Available online: https://CRAN.R-project.org/package=survival (accessed on 7 May 2021).
51. Therneau, T.M.; Grambsch, P.M. *Modeling Survival Data: Extending the Cox Model*; Springer: New York, NY, USA, 2000; ISBN 0-387-98784-3.
52. Kassambara, A.; Kosinski, M.; Biecek, P. Survminer: Drawing Survival Curves Using "ggplot2". R Package Version 0.4.8. 2020. Available online: https://CRAN.R-project.org/package=survminer (accessed on 7 May 2021).
53. miRTarBase 7.0. Available online: http://mirtarbase.mbc.nctu.edu.tw/ (accessed on 7 May 2021).
54. Revelle, W. *Psych: Procedures for Personality and Psychological Research*; Version=2.0.9; Northwestern University: Evanston, IL, USA, 2020. Available online: https://CRAN.R-project.org/package=psych (accessed on 7 May 2021).
55. Pedersen, T.L. Ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. R Package Version 2.0.3. 2020. Available online: https://CRAN.R-project.org/package=ggraph (accessed on 7 May 2021).
56. Pedersen, T.L. Tidygraph: A Tidy API for Graph Manipulation. R Package Version 1.2.0. 2020. Available online: https://CRAN.R-project.org/package=tidygraph (accessed on 7 May 2021).
57. Yu, G.; He, Q.-Y. ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **2016**, *12*, 477–479. [CrossRef]
58. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS A J. Integr. Biol.* **2012**, *16*, 284–287. [CrossRef] [PubMed]
59. Hadley, W. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
60. Olive, V.; Li, Q.; He, L. mir-17-92: A polycistronic oncomir with pleiotropic functions. *Immunol. Rev.* **2013**, *253*, 158–166. [CrossRef]

61. Wang, L.; Liu, J.-L.; Yu, L.; Liu, X.-X.; Wu, H.-M.; Lei, F.-Y.; Wu, S.; Wang, X. Downregulated miR-495 Inhibits the G1-S Phase Transition by Targeting Bmi-1 in Breast Cancer. *Medicine* **2015**, *94*, e718. [CrossRef] [PubMed]
62. Guan, Y.-X.; Zhang, M.-Z.; Chen, X.-Z.; Zhang, Q.; Liu, S.-Z.; Zhang, Y.-L. Lnc RNA SNHG20 participated in proliferation, invasion, and migration of breast cancer cells via miR-495. *J. Cell. Biochem.* **2018**, *119*, 7971–7981. [CrossRef]
63. Chen, Y.; Luo, D.; Tian, W.; Li, Z.; Zhang, X. Demethylation of miR-495 inhibits cell proliferation, migration and promotes apoptosis by targeting STAT-3 in breast cancer. *Oncol. Rep.* **2017**, *37*, 3581–3589. [CrossRef]
64. Khan, S.; Brougham, C.L.; Ryan, J.; Sahrudin, A.; O'Neill, G.; Wall, D.; Curran, C.; Newell, J.; Kerin, M.J.; Dwyer, R.M. miR-379 regulates cyclin B1 expression and is decreased in breast cancer. *PLoS ONE* **2013**, *8*, e68753. [CrossRef]
65. Ma, J.-B.; Hu, S.-L.; Zang, R.-K.; Su, Y.; Liang, Y.-C.; Wang, Y. MicroRNA-487a promotes proliferation of esophageal cancer cells by inhibiting p62 expression. *Eur. Rev. Med. Pharmacol. Sci.* **2019**, *23*, 1502–1512. [CrossRef] [PubMed]
66. Chang, R.-M.; Xiao, S.; Lei, X.; Yang, H.; Fang, F.; Yang, L.-Y. miRNA-487a Promotes Proliferation and Metastasis in Hepatocellular Carcinoma. *Clin. Cancer Res.* **2017**, *23*, 2593–2604. [CrossRef] [PubMed]
67. Herranz, H.; Cohen, S.M. MicroRNAs and gene regulatory networks: Managing the impact of noise in biological systems. *Genes Dev.* **2010**, *24*, 1339–1344. [CrossRef]
68. Jelinic, P.; Shaw, P. Loss of imprinting and cancer. *J. Pathol.* **2007**, *211*, 261–268. [CrossRef]
69. Leick, M.B.; Shoff, C.J.; Wang, E.C.; Congress, J.L.; Gallicano, G.I. Loss of imprinting of IGF2 and the epigenetic progenitor model of cancer. *Am. J. Stem Cells* **2012**, *1*, 59–74.
70. Kagami, M.; Sekita, Y.; Nishimura, G.; Irie, M.; Kato, F.; Okada, M.; Yamamori, S.; Kishimoto, H.; Nakayama, M.; Tanaka, Y.; et al. Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal upd(14)-like phenotypes. *Nat. Genet.* **2008**, *40*, 237–242. [CrossRef]
71. Stadtfeld, M.; Apostolou, E.; Akutsu, H.; Fukuda, A.; Follett, P.; Natesan, S.; Kono, T.; Shioda, T.; Hochedlinger, K. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **2010**, *465*, 175–181. [CrossRef] [PubMed]
72. Kameswaran, V.; Bramswig, N.C.; McKenna, L.B.; Penn, M.; Schug, J.; Hand, N.J.; Chen, Y.; Choi, I.; Vourekas, A.; Won, K.-J.; et al. Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. *Cell Metab.* **2014**, *19*, 135–145. [CrossRef] [PubMed]
73. Molina-Pinelo, S.; Salinas, A.; Moreno-Mata, N.; Ferrer, I.; Suarez, R.; Andrés-León, E.; Rodríguez-Paredes, M.; Gutekunst, J.; Jantus-Lewintre, E.; Camps, C.; et al. Impact of DLK1-DIO3 imprinted cluster hypomethylation in smoker patients with lung cancer. *Oncotarget* **2018**, *9*, 4395–4410. [CrossRef]
74. González-Vallinas, M.; Rodríguez-Paredes, M.; Albrecht, M.; Sticht, C.; Stichel, D.; Gutekunst, J.; Pitea, A.; Sass, S.; Sánchez-Rivera, F.J.; Lorenzo-Bermejo, J.; et al. Epigenetically Regulated Chromosome 14q32 miRNA Cluster Induces Metastasis and Predicts Poor Prognosis in Lung Adenocarcinoma Patients. *Mol. Cancer Res.* **2018**, *16*, 390–402. [CrossRef]