Research Paper

Identification of microRNA differentially expressed in three subtypes of non-small cell lung cancer and in silico functional analysis

Yanjun Hu[1,*], Luqing Wang[2,*], Jingxian Gu[3], Kai Qu[3] and Yunxia Wang[4]

[1]Department of Clinical Laboratory, Liaocheng People's Hospital, Taishan Medical College, Liaocheng 252000, Shandong Province, China

[2]Department of Nuclear Medicine, Radioimmunology Room, Liaocheng People's Hospital, Taishan Medical College, Liaocheng 252000, Shandong Province, China

[3]Department of Hepatobiliary Surgery, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China

[4]Department of Intensive Care Unit, Liaocheng Infectious Diseases Hospital, Liaocheng 252000, Shandong Province, China

[*]These authors have contributed equally to this work

Correspondence to: Yunxia Wang, **email:** wyx9039@163.com

Kai Qu, **email:** joanne8601@163.com

ABSTRACT

Emerging studies demonstrated that miRNAs played fundamental roles in lung cancer. In this study, we attempted to explore the clinical significance of the miRNA signature in different histological subtypes of non-small cell lung cancer (NSCLC). Three miRNome profiling datasets (GSE19945, GSE25508 and GSE51853) containing lung squamous cell carcinoma (SCC), lung adenocarcinoma (ADC) and large cell lung cancer (LCLC) samples were obtained for bioinformatics and survival analysis. Moreover, pathway enrichment and coexpression network were performed to explore underlying molecular mechanism. MicroRNA-375 (miR-375), miR-203 and miR-205 were identified as differentially expressed miRNAs (DEmiRNAs) which distinguished SCC from other NSCLC subtypes. Pathway enrichment analysis suggested that Hippo signaling pathway was combinatorically affected by above mentioned three miRNAs. Coexpression analysis of three miRNAs and the Hippo signaling pathway related genes were conducted based on another dataset, GSE51852. Four hub genes (*TP63*, *RERE*, *TJP1* and *YWHAE*) were identified as the candidate targets of three miRNAs, and three of them (*TP63*, *TJP1* and *YWHAE*) were validated to be downregulated by miR-203 and miR-375, respectively. Finally, survival analysis further suggested the prognostic value of three-miRNA signature in SCC patients. Taken together, our study compared the miRNA profiles among three histological subtypes of NSCLC, and suggested that a three-miRNA signature might be potential diagnostic and prognostic biomarkers for SCC patients.

INTRODUCTION

Lung cancer remains the leading cause of cancer-related death both worldwide and in China [1, 2]. Non-small cell lung cancer (NSCLC) including

three main histological subtypes, lung squamous cell carcinoma (SCC), lung adenocarcinoma (ADC) and large cell lung cancer (LCLC), accounts for over 80% cases of lung cancer [3]. Increasing studies demonstrated that therapeutic effect varied in different histological

subtypes of NSCLC, suggesting that early confirmation of histology is of critical importance [4-6]. Nevertheless, there are emerging researches into the biological features of different histological subtypes of NSCLC, but the fundamental molecular mechanisms remain elusive. For instance, smoking is a much greater risk of SCC than of ADC [7]. Thus, a clear understanding of the differences from morphological to genetic levels between them will definitely have a positive effect on precise therapy.

MicroRNAs (miRNAs), comprising 19-23 nucleotides or so, are a class of endogenously expressed RNAs belonging to non-coding RNA family [8]. In the last decade, miRNAs have attracted so much attention as biomarkers with great potential in the early diagnosis, and clinical outcome prediction of multiple malignancies [9-11]. In NSCLC, emerging evidence suggested miRNAs as potential diagnostic and prognostic biomarkers [12-14]. Several miRNAs were reported to be capable of distinguishing different subtypes of NSCLC [15, 16].

Therefore, exploration of miRNA-based biomarkers provides opportunities for investigate the molecular changes among different histological subtypes of NSCLC.

Recently, high-throughput platforms have been utilized as a fundamental method dealing with massive genetic data. Data from high-throughput platforms including microarrays and next-generation sequence, if analyzed by bioinformatics methods, can provide amounts of useful information for screening cancer biomarkers and therapeutic targets [17]. Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), hosted by the National Center for Biotechnology Information (NCBI), serves as a public genetic expression profile repository for a wide range of high-throughput experimental data. Therefore, in the present study, we adopted high-throughput data from GEO database for bioinformatics analysis to identify the differentially diagnostic miRNA signature between three subtypes of NSCLC.
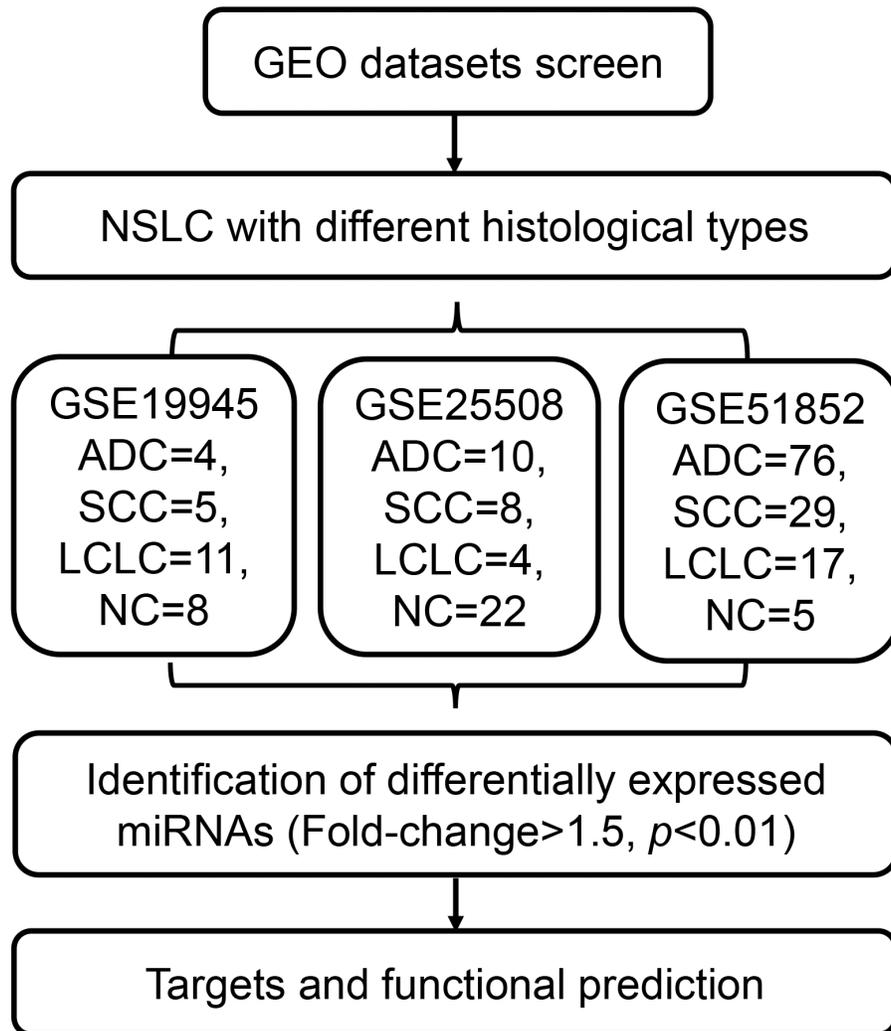

Figure 1: Schematic of workflow depicting the eligible datasets and analysis procedure.

RESULTS

Identification of differentially expressed miRNAs (DEmiRNAs) in NSCLC

The workflow of this study was presented in Figure 1. Firstly, we compared the miRNA expression profiles in three miRNA datasets (GSE19945, GSE25508 and GSE51852). DEmiRNAs between cancer and normal tissue samples were identified in each dataset. In GSE19945, 39 up-regulated and 52 down-regulated miRNAs were obtained (Figure 2A). Based on GSE51853, a total of 53 dysregulated miRNAs were identified including 14 up-regulated and 31 down-regulated miRNAs (Figure 2B). In GSE25508, there were 29 up-regulated and 7 down-regulated miRNAs (Figure 2C). When comparing the list of dysregulated miRNAs among three datasets, six upregulated miRNAs (miR-130b, miR-210, miR-183, miR-21, miR-200b and miR-96) and five down-regulated miRNAs (miR-30a, miR-638, miR-486, miR-30d and miR-145) (Figure 2D).

Identification of DEmiRNAs in different histological subtypes of NSCLC

We further compared miRNA expression profiles between cancer and normal tissue samples in different histological subtypes of NSCLC (ADC, SCC and LCLC) of each dataset. A Venn diagram was conducted to examine the overlap of resulting DEmiRNAs lists from three datasets in each histological type of NSCLC. In ADC, there were 12 up- and 11 down-regulated miRNAs listed in more than two datasets (Figure 3A). Robust rank aggregation (RRA) analysis showed miR-210 and miR-130b was significantly up-regulated with a Bonferroni-corrected p-value less than 0.05 (Figure 3B), but no downregulated miRNA reached statistical significance after Bonferroni correction (Figure 3C). Similarly, a total of 11 up- and 26 down-regulated miRNAs commonly listed in SCC datasets (Figure 3D). Integrated meta-analysis confirmed that miR-205, miR-210 and miR-130b was up-regulated, and miR-486 and miR-30a were down-regulated (Figure 3E and 3F). In LCLC, there were 14 up- and 27 down-regulated miRNAs were selected (Figure 3G), and miR-30a*, miR-30a, miR-497, miR-144 and miR-145 were identified as commonly down-regulated miRNAs in LCLC (Figure 3H and 3I).

MiR-203, miR-205 and miR-375 distinguished SCC from other NSCLC subtypes

We also explored the DEmiRNAs between SCC and other two histological subtypes of NSCLC, LCLC and ADC. It was found that a total of 6 miRNAs were
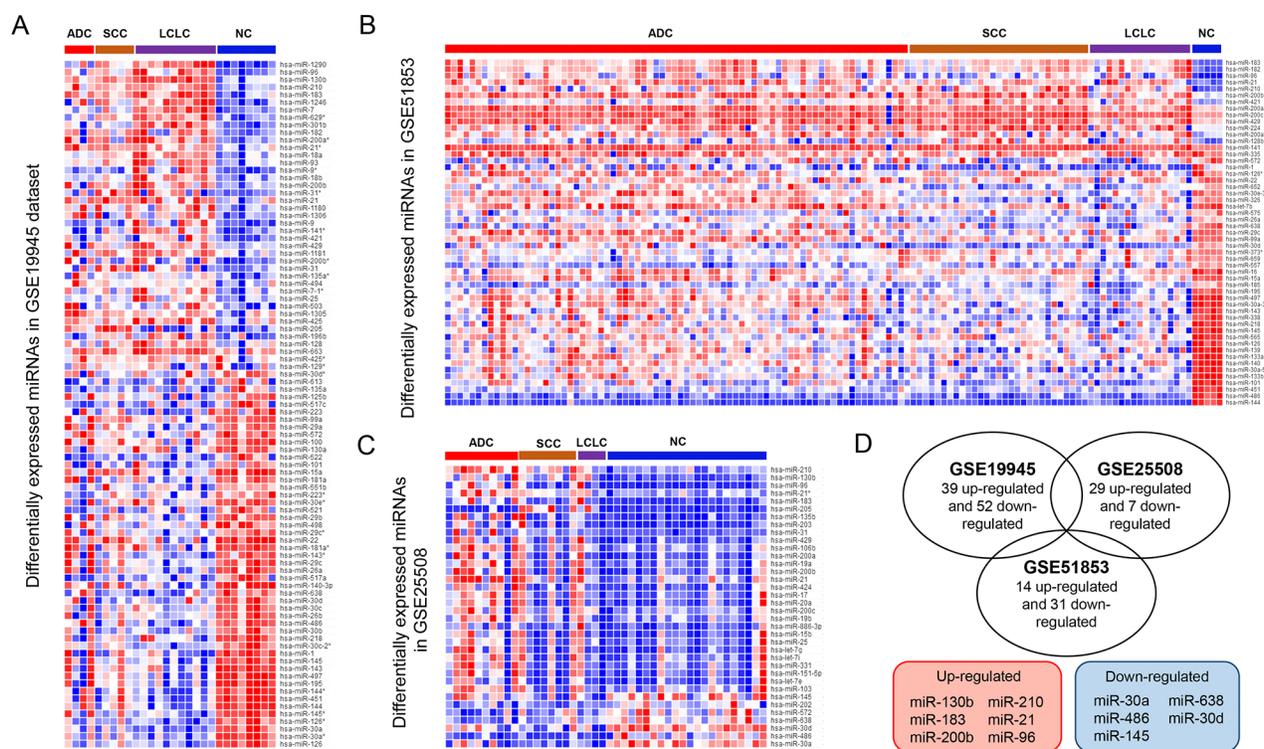

Figure 2: Identification of common DEmiRNAs in NSCLC. Heat map of the significantly dysregulated miRNAs between the cancer and normal tissue samples in GSE19945 **(A)**, GSE51852 **(B)** and GSE25508 **(C)**. **(D)** (*Upper*) Venn diagram between the up-regulated and down-regulated miRNAs from each dataset. (*Lower*) the lists of the common dysregulated miRNAs. Red: up-regulation; blue: down-regulation.

significantly dysregulated from the comparison of SCC versus LCLC (Figure 4A). The corresponding RRA approach revealed that miR-205, miR-375 and miR-203 were the commonly dysregulated miRNAs (Figure 4B). When comparing SCC with ADC, four significantly dysregulated miRNAs were selected (Figure 4C). The comprehensive meta-analysis showed that miR-375 and miR-203 were significantly dysregulated (Figure 4D). Although miR-205 was not significant, it was the third dysregulated miRNA collectively. In addition, we did not obtain positive result from comparison between ADC and

LCLC (data not shown). Interestingly, miR-375, miR-203 and miR-205 were listed in both SCC versus ADC and SCC versus LCLC. Therefore, miR-375, miR-203 and miR-205 were selected for further analysis.

Function prediction of miR-203, miR-205 and miR-375 in SCC

Target genes of miR-203, miR-205 and miR-375 were predicted by TargetScan and mapped in Cytoscape (Figure 5A). Kyoto Encyclopedia of Genes and Genomes
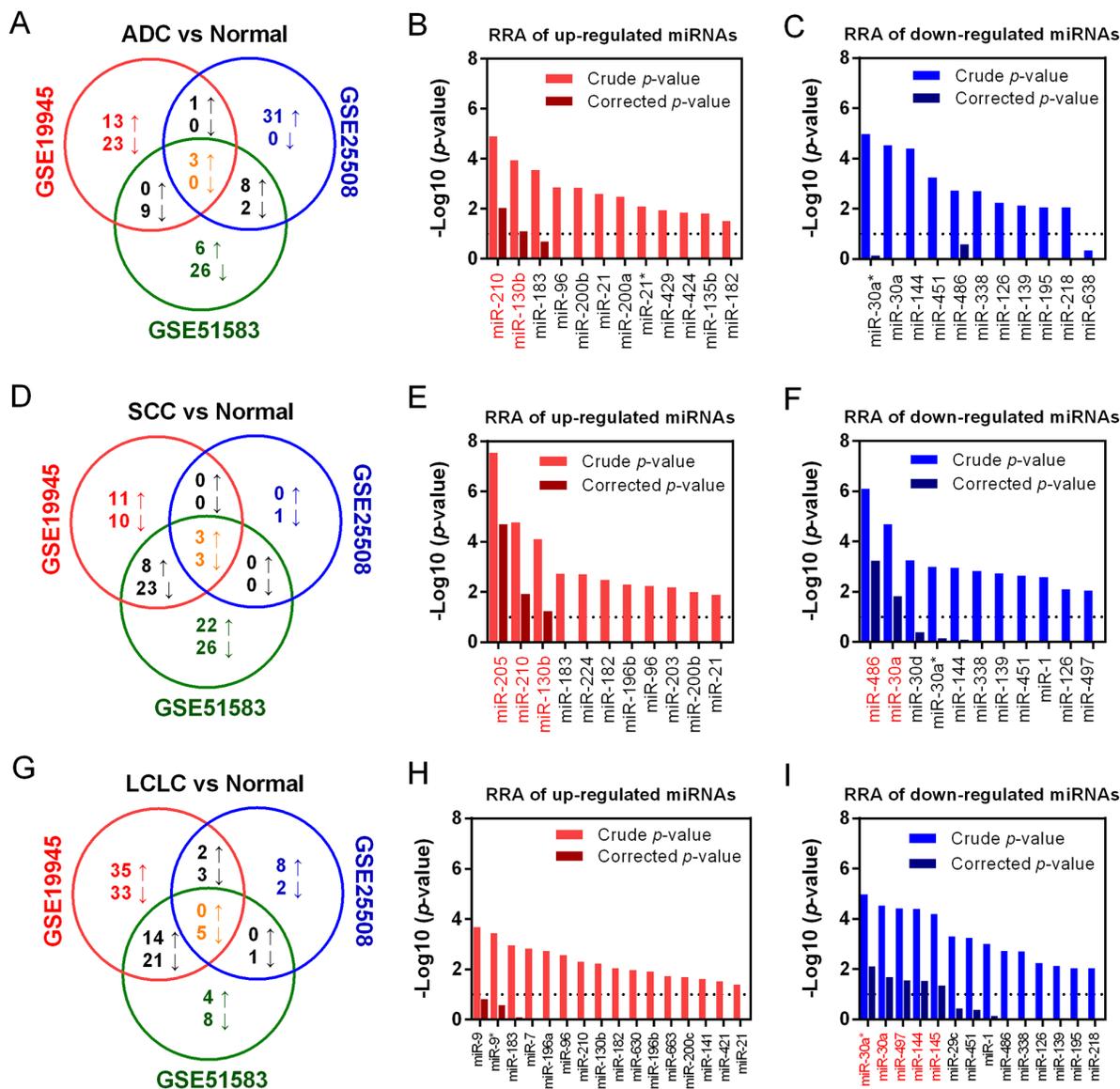

Figure 3: Identification of DEmiRNAs in three histological subtype of NSCLC. (A) Overlapped DEmiRNAs in ADC were presented as Venn diagram. RRA analysis of up-regulated **(B)** and down-regulated miRNAs **(C)** to identify common DEmiRNAs in ADC. **(D)** Overlapped DEmiRNAs in SCC were presented. RRA analysis of up-regulated **(E)** and down-regulated miRNAs **(F)** in SCC were listed. **(G)** Overlapped DEmiRNAs in LCLC were presented. RRA analysis of up-regulated **(H)** and down-regulated miRNAs **(I)** in LCLC were listed. For the RRA column diagram, red: crude p-value (up-regulation); Dark red: corrected p-value (up-regulation). Blue: crude p-value (down-regulation); Dark blue: corrected p-value (down-regulation).

(KEGG) pathway enrichments of target genes were carried out, respectively (Figure 5B-5D). "Pathway in cancer (KEGG_05200)", "Focal adhesion (KEGG_04510)" and "Adherens junction (KEGG_04520)" were commonly top pathways by function prediction of three miRNAs. Moreover, the combinatorial effects of three miRNAs in pathways were predicted by the DIANA tools (http://www.microrna.gr/miRPathv2). Hippo signaling pathway was the most enriched KEGG pathway (Figure 5E). Recently, increasing evidence suggested that Hippo pathway, which was demonstrated to regulate cell proliferation, differentiation and migration, played a critical role in lung cancer, particularly in NSCLC [23-25]. We therefore employed another GEO dataset, GSE51852, to further analysis the expression correlations between three miRNAs and genes in Hippo signaling pathway (Figure 6A). Pearson correlation was calculated and yielded three sets of genes that were significantly correlated with miR-375, miR-203 and miR-205 (Figure 6B). We then conducted interaction network based on three miRNAs and their targets which reached statistical significance. From the comprehensive network, 4 genes (*TP63*, *RERE*, *TJP1* and *YWHAE*) were found to be interacted with all three miRNAs, and 6 genes (*YAP1*, *AMOT*, *WWC1*, *PTPN14*, *TSHZ2* and *STK4*) were interacted with two miRNAs (Figure 6C). In addition, we also validated the regulationship between three miRNAs and above Hippo signaling genes. Our data revealed that *TP63* and *TJP1* were downregulated by miR-203 (Figure 7A and 7B), meanwhile, *YWHAE* was downregulated by miR-375 (Figure 7C).

Prognostic values of miR-203, miR-205 and miR-375 in SCC

Kaplan-Meier curve was conducted to analyze the association between three miRNAs (miR-203, miR-205 and miR-375) expression and overall survival in SCC
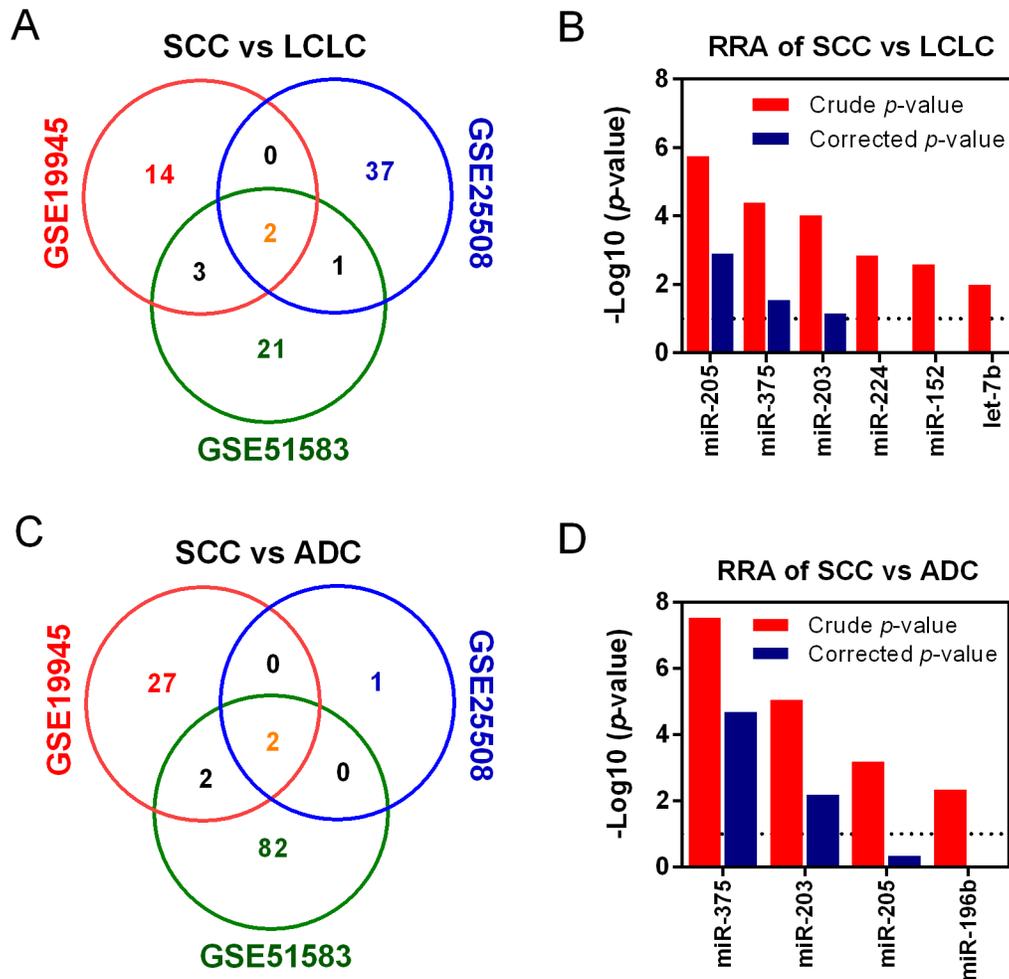

Figure 4: Identification of DEmiRNAs among histological subtypes of NSCLC. Venn diagram of overlapped dysregulated miRNAs of SCC vs. LCLC **(A)** and SCC vs. ADC **(C)**. Common DEmiRNAs of SCC vs. LCLC **(B)** and SCC vs. ADC **(D)** were identified using RRA method. For the RRA column diagram, red: rude p-value; blue: corrected p-value.
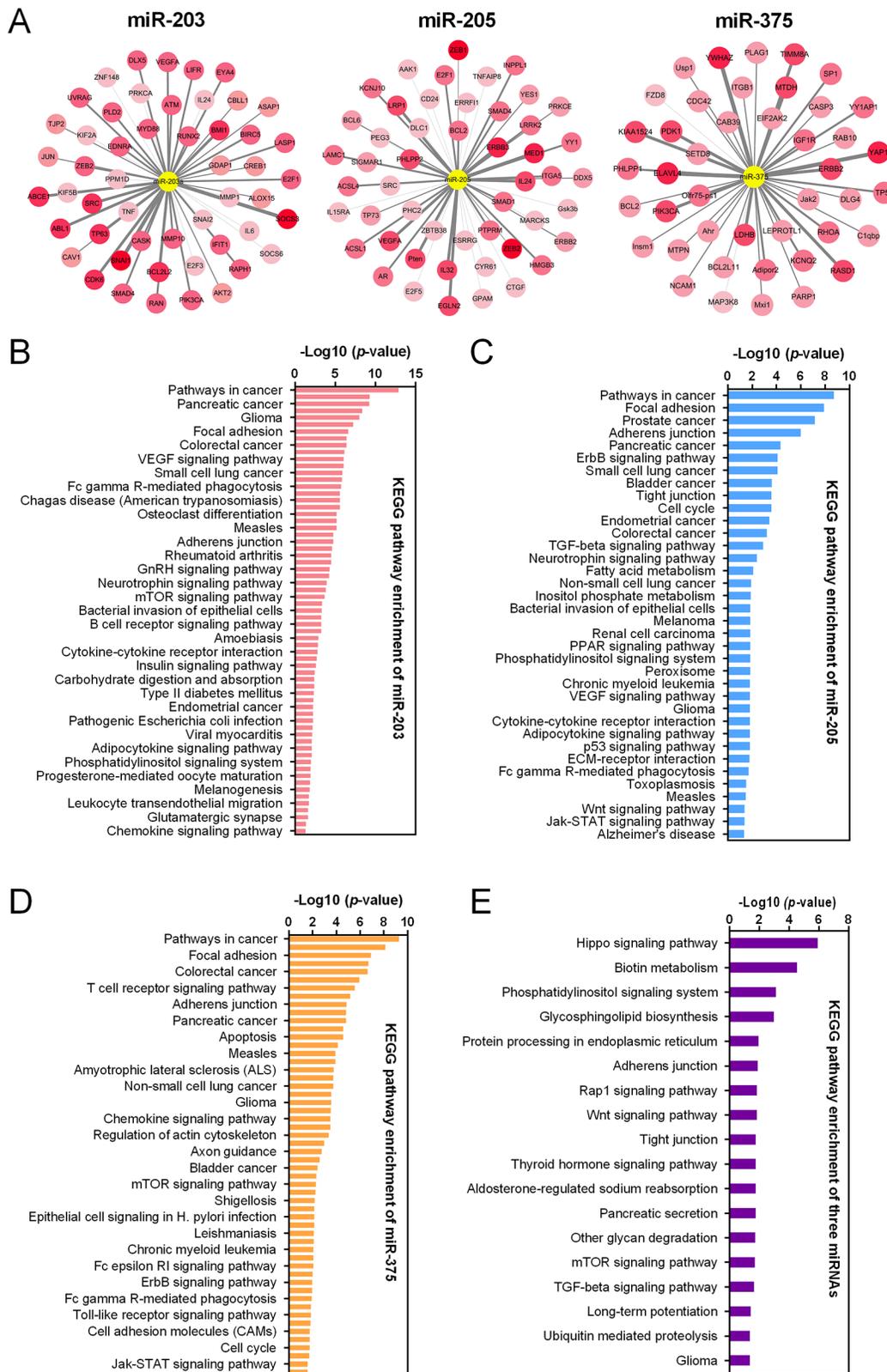
Figure 5: MiRNA-gene network and KEGG pathway enrichment. (A) The interaction between three miRNAs (miR-203, miR-205 and miR-375) and of their predicted target genes. The central yellow circles represented miRNAs. The color of the circles with the names of the target genes on it varies according to their cumulative weighted context++ score. KEGG pathway enrichments of predicted target genes of miR-203 **(B)**, miR-205 **(C)** and miR-375 **(D)**, were presented, respectively. **(E)** The combinatorial effects of three miRNAs in KEGG pathways were predicted.
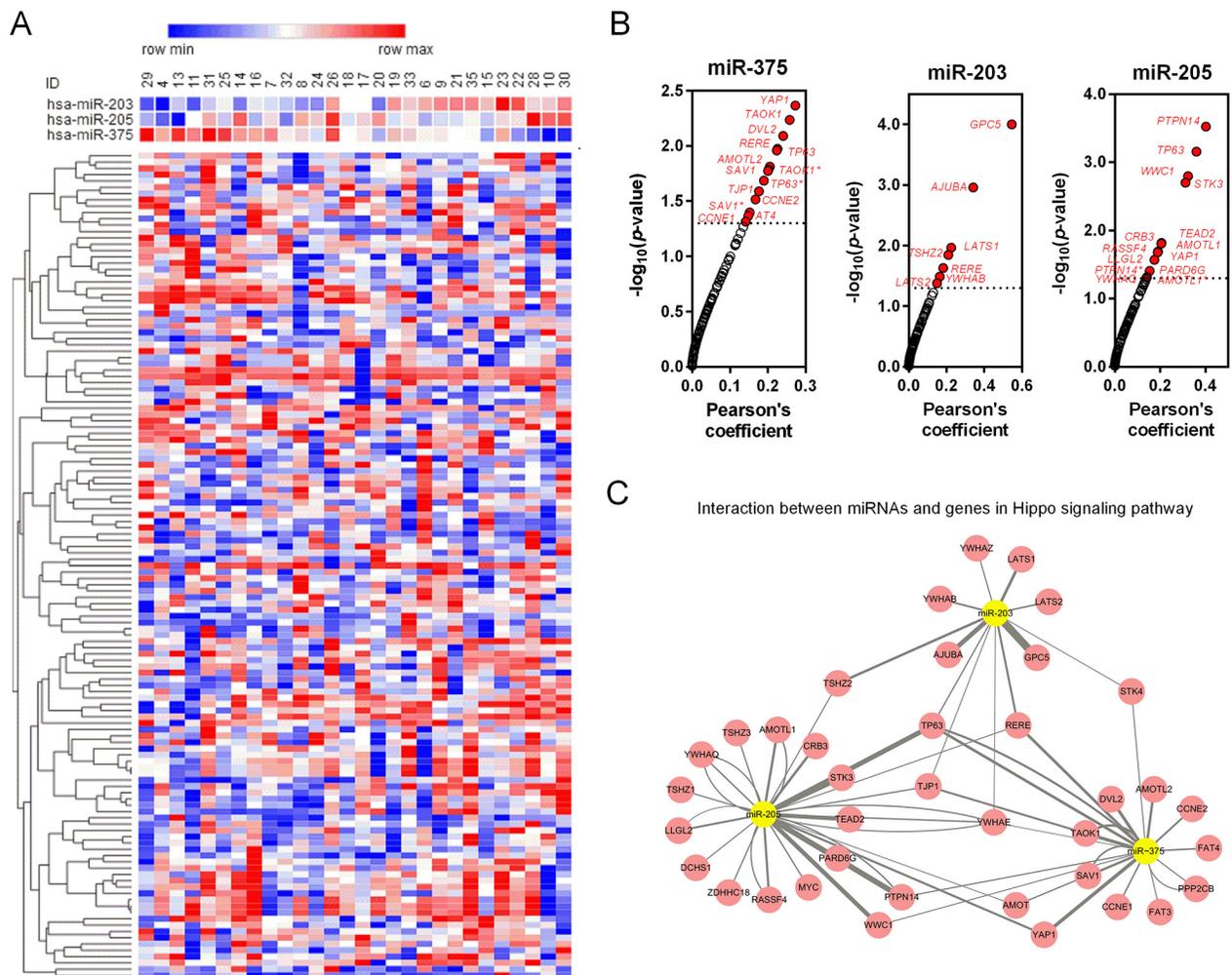
Figure 6: Coexpression analysis of three miRNAs and Hippo signaling related genes. (A) The expression profile of 3 miRNAs (miR-203, miR-205 and miR-375) and 90 Hippo signaling related genes was presented as heatmap. **(B)** Pearson correlation for each miRNA and 90 Hippo signaling related genes. The genes which considered to have statistically significant relationship with the miRNA were marked red. **(C)** The interaction network of the significant couples derived from Pearson correction.
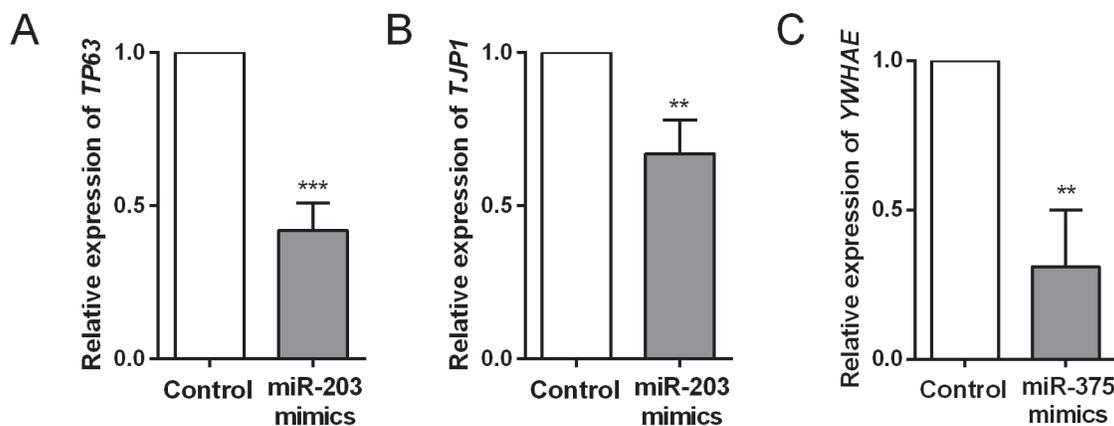

Figure 7: Comparison of target gene expression levels between control and miRNA mimics -transfected HCC cells. (A) Downregulation of *TP63* by miR-203 mimics transfection; **(B)** Downregulation of *TJP1* by miR-203 mimics transfection; **(C)** Downregulation of *YWHAE* by miR-375 mimics transfection.

patients. To reduce inter-lab variability, prognostic data were obtained from two independent datasets based on sequencing (TCGA-LUSC dataset, n=134) and microarray platforms (GSE16025 dataset, n=61), respectively. By using the SurvMicro web tool, we compared the overall survival between the SCC patients with high risk and low risk groups, according to their expression levels of three miRNAs. As shown in Figure 8, combination of miR-203, miR-205 and miR-375 was significantly associated with overall survival of SCC in both two datasets, with Log-Rank *p*-values being 0.0032 and 0.0286, respectively. The Cox regression model suggested that SCC patients in high risk groups showed shorter overall survival time, compared with those in low risk groups. The corresponding Hazard ratio (HR) values were 2.57 (95% confidence interval (CI), 1.37-4.84) and 2.15 (95%CI, 1.08-4.25) in TCGA and GSE16025 datasets, respectively, suggesting that three-miRNA signature was related to the overall survival in SCC patients.

DISCUSSION

Lung cancer is initiated by various genetic alterations, and investigation of cancer genome landscapes have a profound effect on clinical practice [26]. Recently, high-throughput methods have been widely used to screen thousands of genes simultaneously which makes genomic study much more efficiently [27]. Bioinformatic analysis based on high-throughput data will definitely provide more evidence for understanding molecular mechanisms of carcinogenesis. Here, we employed three miRNA expression profiles (GSE19945, GSE25508 and GSE51853) to search for the key miRNAs as candidate clinical biomarkers in different histological subtypes of NSCLC. After the comprehensive analysis, miR-375, miR-203

and miR-205 were found to distinguish SCC from other subtypes of NSCLC. Moreover, function prediction and coexpression analysis suggested that above three miRNAs might affect cancer progression and Hippo signaling pathway. Finally, the prognostic values of three-miRNA signature were confirmed by two independent datasets.

Recent studies showed the clinical features varied among different NSCLC subtypes [29, 30]. SCC, accounting for some 30% of lung cancer cases, is more closely associated with tobacco smoking than ADC/LCLC [26, 28, 29]. Most SCCs are centrally located and tend to originate from proximal bronchi whereas ADC/LCLC are often peripheral [30]. ADC arises most frequently from bronchial mucosal glands and are highly vascular [31]. And the least common subtype, LCLC was highly malignant, one subtype of which has neuroendocrine properties [30]. ADC/LCLC was easier and earlier to metastasize than SCC, the latter always grows slowly and are more likely to remain in the central area of lung [32, 33]. Furthermore, the responses to targeted therapy are also different in patients with different histological subtypes [34, 35]. Therefore, exploration the underlying molecular changes among different histological subtypes might provide evidence for understanding mechanisms of NSCLC. In this study, we compared the miRNA expression profiles of SCC/ADC/LCLC samples with normal lung samples using three GEO datasets. From ADC vs. normal, there were two miRNAs significantly up-regulated (miR-210 and miR-130b) and none down-regulated. And from SCC vs. normal, 3 miRNAs (miR-205, miR-210 and miR-130b) were significantly up-regulated and 2 (miR-486 and miR-30a) were significantly down-regulated. As for LCLC vs. normal, 5 (miR-30a*, miR-30a, miR-497, miR-144 and miR-145) were significantly down-regulated and none up-regulated. Moreover, we compared different subtypes
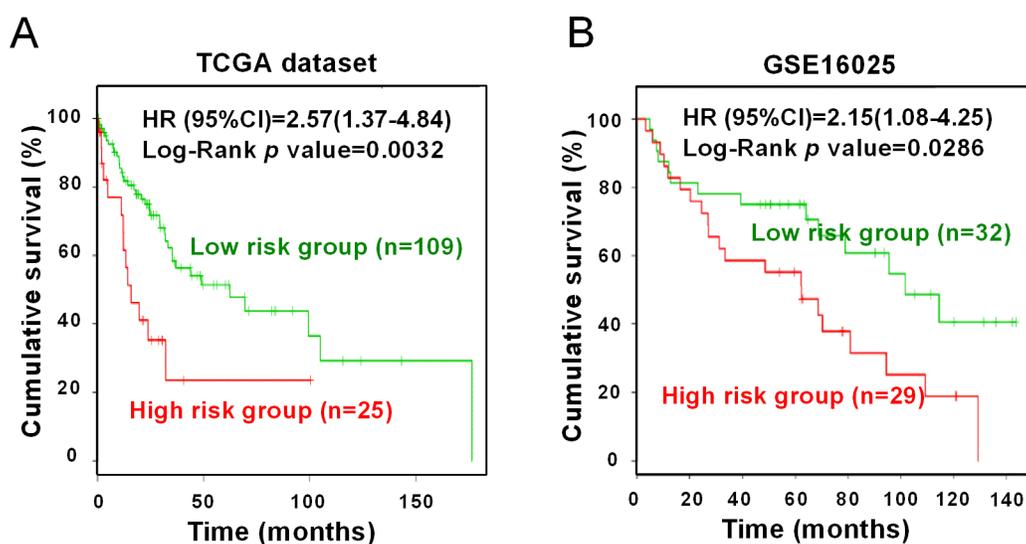

Figure 8: Survival analysis of three miRNAs. Kaplan-Meier survival curves were conducted to compare overall survival of SCC patients in high or low risk groups. **(A)** TCGA; **(B)** GSE16025. Green: low risk group. Red: high risk group.

of NSCLC (SCC vs. ADC, SCC vs. LCLC and ADC vs LCLC) using the similar method. Interestingly, we found miR-203, miR-205 and miR-375 distinguished SCC from other NSCLC subtypes, consisting with previous studies [36-38]. Moreover, we also confirmed that the signature of three-above mentioned miRNAs was also associated with overall survival of SCC patients. Canueto et al. reported the prognostic values of miR-205 and miR-203 in cutaneous squamous cell carcinoma [39]. Additionally, miR-375 was also found to be negatively associated with clinical outcomes in NSCLC [40, 41]. Therefore, our above findings provide evidence for investigating mechanisms and exploring therapeutic target for NSCLC, especially for SCC patients.

In order to further explore the function of three identified miRNAs, pathway enrichment analysis of their target genes was performed. Hippo signaling pathway was the most predicted pathway affected by miR-203, miR-205 and miR-375. Hippo signaling pathway was reported to be closely related to the occurrence of multiple malignancies, including NSCLC [42-44]. Therefore, we further performed coexpression analysis of three miRNAs and Hippo signaling pathway related genes. Four hub genes (*TP63*, *RERE*, *TJP1* and *YWHAE*) were identified as the pivotal genes that might be affected by three miRNAs. Furthermore, our validation data confirmed the association between miRNAs and three Hippo signaling pathway related genes (*TP63*, *TJP1* and *YWHAE*). *TP63* (tumor protein p63), a member of p53 family, was considered as one of "squamous markers" [45], and its copy number variation was reported to be associated with SCC [46]. *TJP1* (tight junction protein 1), also known as *ZO-1* (Zonula Occludens-1), participated in cell-cell adhesion and is reported to be associated with skin SCC [47]. Recent studies suggested that *TJP1* might be a useful prognostic predictor of NSCLC [48, 49]. *YWHAE*, a gene encoding 14-3-3 protein, was found to involve in carcinogenesis and progression of SCC [50]. Taken together, both our findings and previous reports suggested the critical roles of three miRNAs (miR-203, miR-205 and miR-375) in regulating Hippo signaling pathway. Therefore, investigation the underlying interaction between above mentioned miRNA and genes will be of particular interest to study the molecular mechanisms of NSCLC.

In conclusion, we identified differentially expressed miRNAs and their possible target genes involved in different histological subtypes of NSCLC. The novel findings in the present study were that miR-203, miR-205 and miR-375 could distinguish SCC from other subtypes of NSCLC, and were associated with overall survival in SCC patients. Hippo signaling pathway might be affected by above-mentioned three miRNAs. Our results provide evidence to better understand the molecular changes underlying distinctive biological properties of different NSCLC subtypes.

MATERIALS AND METHODS

Microarray datasets

The miRNA datasets were searched from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). The search strategy was ("lung" OR "pulmonary" OR "respiratory" OR "bronchi" [Mesh]) AND ("cancer" OR "carcinoma" OR "tumor" OR neoplas* OR malignan* OR "squamous Cell Carcinoma" OR "adenocarcinoma" OR "large cell lung cancer" [Mesh]) AND (MicroRNA OR miRNA [Mesh]). Finally, three datasets (GSE19945, GSE25508 and GSE51853) were selected for further analysis. These three datasets contained the histological information of NSLC (ADC, SCC and LCLC).

GSE19945 was conducted through GPL9948 platform (Agilent Human 0.6K miRNA Microarray G4471) and consisted of 4 ADC, 5 SCC, 11 LCLC and 8 normal lung tissue samples. Meanwhile, GSE51853, based on GPL6480 platform (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F), consisted of 76 ADC, 29 SCC, 17 LCLC and 5 normal lung tissue samples [18]. GSE25508 was conducted by GPL7731 platform (Agilent-019118 Human miRNA Microarray 2.0 G4470B) to explore miRNA expression panel with or without asbestos exposure [19]. From GSE25508 dataset, miRNA profiling data of 10 ADC, 8 SCC, 4 LCLC and 22 non-cancer lung tissue samples were extracted for analysis.

Identification of DEmiRNAs

DEmiRNAs between different groups (cancer vs normal, or comparison among different histological subtypes) were screened based on three GSE datasets, respectively. Only fold change (FC)>=1.5 and *p*-value for *t*-test <0.05 was considered statistically significant. The heatmap of DEmiRNAs was conducted by a web-based tool, Morpheus (https://software.broadinstitute.org/morpheus/). Common DEmiRNAs among three GSE datasets was conducted by Venny 2.1.0 (http://bioinfogp.cnb.csic.es/tools/venny/). Moreover, we carried out a comprehensive meta-analysis of the expression of the miRNA screened above through a RRA method to identify DEmiRNAs [17]. This approach requires ranked miRNA lists and the total probe number of each miRNome profiling study. MiRNAs were re-ranked by corrected *p*-value.

Prediction of target genes

Target genes of the DEmiRNAs was predicted by a web tool, TargetScanHuman 7.1 (http://www.targetscan.org/vert_71/). The cumulative weighted context++ score <-0.4 (-0.2 for miR-375) was adopted as the cut-off value. These genes were intended for pathway enrichment. GSE51852,

another subseries of GSE51855, was a gene expression profile based on the same sample as GSE51853. The interaction between the DEmiRNAs and the significantly genes was mapped in Cytoscape [20].

Pathway enrichment analysis

Gene Ontology (GO) and KEGG pathway enrichment was performed by Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david-d.ncifcrf.gov/summary.jsp), a web tool for gene functional annotation [21]. Identifier of official gene symbol and homo sapiens were selected.

Cell culture and transfection

The HCC cell line, SMMC-7721, used in this study were obtained from the Liver Cancer Institute, Fudan University (Shanghai, China), and maintained in DMEM containing 10% fetal bovine serum at 37°C with 5% CO_2. The miRNA mimics were synthesized by GenePharma (Shanghai, China). The transfections of miR-375 and miR-203 mimics were performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the procedure recommended by the manufacturer.

Reverse transcription-polymerase chain reaction (RT-PCR)

Total RNA was isolated from the different cell groups using the RNAfast200 Total RNA Extract Kit (Fastgene, Shanghai, China), and 2 μg RNA was reverse transcribed to cDNA by the RevertAid™ First Strand cDNA Synthesis Kit (Fermentas, MBI, Lithuania). Quantitative RT-PCR was carried out using the SYBR® PrimeScriptTM miRNA RT-PCR Kit and SYBR® Premix Ex TaqTM (TaKaRa Biotechnology, Dalian, China). The Fold change of gene expression was calculated based on the threshold cycle (Ct) as relative to β-actin using a 2-Δ(ΔCt) method. The sequences of PCR primers for target genes are listed as followings. *TP63*: forward primer, 5'-GGACCAGCAGATTCAGAACGG-3'; reverse primer, 5'-AGGACACG TCGAAACTGTGC-3'; *TJP1*: forward primer, 5'-CAACATACAGTGACGCTTCACA-3'; reverse primer, 5'- CACTATTGACGTTTCCCCACTC-3'; *YWHAE*: forward primer, 5'- GATTCGGGAATATC GGCAAATGG-3'; reverse primer, 5'-GCTGGAATGAG GTGTTTG TCC-3'. All primer pairs were synthesized by TaKaRa.

Survival analysis

The prognostic value of DEmiRNAs was evaluated by SurvMicro (http://bioinformatica.mty.itesm.mx:8080/Biomatec/Survmicro.jsp) [22]. Kaplan-Meier curves were conducted based on TCGA-LUSC and GSE16025 datasets. MiRNA profile of TCGA-LUSC dataset was sequenced using Illumina Genoma Analyzer IIx Equipment. GSE16025 dataset contained 61 SCC samples and 10 matched normal lung samples and profiled on MirVana miRNA Bioarrays (version 2, Ambio). HR and 95% CI of the association between miRNAs and overall survival was estimated using a Cox proportional hazards analysis. Statistical assessment was performed using the Log Rank test.

Statistical analysis

Statistical analyses were performed using STATA software, version 12.0. The paired student *t*-test was used to compare the differences in gene expression between control and miRNA mimics -transfected HCC cells. Pearson correlation was applied to evaluate the expression correlation between the DEmiRNAs and the selected genes. *p*-value less than 0.05 was considered statistically different.

Abbreviations

ADC: lung adenocarcinoma; CI: confidence interval; Ct: threshold cycle; DAVID: Database for Annotation, Visualization and Integrated Discovery; DEmiRNA: differentially expressed miRNA; FC: fold change; GO: Gene Ontology; GEO: Gene Expression Omnibus; HR: hazard ratio; KEGG: Kyoto Encyclopedia of Genes and Genomes; LCLC: large cell lung cancer; NSCLC: non-small cell lung cancer; miRNA: microRNAs; NCBI: National Center for Biotechnology Information; RRA: robust rank aggregation; SCC: lung squamous cell carcinoma.

Authors contributions

Kai Qu and Yunxia Wang: Designed the research and wrote the paper; Yanjun Hu and Luqing Wang: Collected and analyzed data; Kai Qu: Constructed figures; Jingxian Gu and Kai Qu: Drafted and revised the manuscript.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors declare no competing financial interests.

FUNDING

REFERENCES

1. Torre LA, Siegel RL and Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016; 893:1-19.

2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin. 2016; 66:115-132.

3. Capelletto E and Novello S. Emerging new agents for the management of patients with non-small cell lung cancer. Drugs. 2012; 72:37-52.

4. Peng G, Zinner RG, Wang Y, Treat J, Monberg M, Obasaju CK, Herbst RS, Novello S and Scagliotti GV. Comparison of patient outcomes stratified by histology among pemetrexed (P)-treated patients (pts) with stage IIIB/IV non-small cell lung cancer (NSCLC) in two phase II trials. Journal of Applied Mathematics & Mechanics. 2008; 76:636-645.

5. Scagliotti GV, Parikh P, Von PJ, Biesma B, Vansteenkiste J, Manegold C, Serwatowski P, Gatzemeier U, Digumarti R and Zukin M. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. Journal of Clinical Oncology. 2008; 26:3543-3551.

6. Scagliotti G, Hanna N, F, Sugarman K, Blatter J, Peterson P, Simms L and Shepherd F. The differential efficacy of pemetrexed according to NSCLC histology: a review of two Phase III studies. Oncologist. 2009; 14:253-263.

7. Sobue T, Yamamoto S, Hara M, Sasazuki S, Sasaki S and Tsugane S. Cigarette smoking and subsequent risk of lung cancer by histologic type in middle-aged Japanese men and women: the JPHC study. International Journal of Cancer Journal International Du Cancer. 2002; 99:245-251.

8. Dong H, Lei J, Ding L, Wen Y, Ju H and Zhang X. MicroRNA: function, detection, and bioanalysis. Chemical Reviews. 2013; 113:6207.

9. Li J, Tan S, Kooger R, Zhang C and Zhang Y. MicroRNAs as novel biological targets for detection and regulation. Chemical Society Reviews. 2014; 43:506-517.

10. Boeri M, Pastorino U and Sozzi G. Role of microRNAs in lung cancer: microRNA signatures in cancer prognosis. Cancer Journal. 2012; 18:268-274.

11. Ling H, Krassnig L, Bullock MD and Pichler M. MicroRNAs in Testicular Cancer Diagnosis and Prognosis. Urologic Clinics of North America. 2016; 43:127-134.

12. Zhan B, Lu D, Luo P and Wang B. Prognostic Value of Expression of MicroRNAs in Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. Clin Lab. 2016; 62:2203-2211.

13. Berghmans T, Ameye L, Willems L, Paesmans M, Mascaux C, Lafitte JJ, Meert AP, Scherpereel A, Cortot AB, Cstoth I, Dernies T, Toussaint L, Leclercq N and Sculier JP. Identification of microRNA-based signatures for response and survival for non-small cell lung cancer treated with cisplatin-vinorelbine A ELCWP prospective study. Lung Cancer. 2013; 82:340-345.

14. Li X, Shi Y, Yin Z, Xue X and Zhou B. An eight-miRNA signature as a potential biomarker for predicting survival in lung adenocarcinoma. J Transl Med. 2014; 12:159.

15. Patnaik S, Mallick R, Kannisto E, Sharma R, Bshara W, Yendamuri S and Dhillon SS. MiR-205 and MiR-375 microRNA assays to distinguish squamous cell carcinoma from adenocarcinoma in lung cancer biopsies. J Thorac Oncol. 2015; 10:446-453.

16. Huang W, Hu J, Yang DW, Fan XT, Jin Y, Hou YY, Wang JP, Yuan YF, Tan YS, Zhu XZ, Bai CX, Wu Y, Zhu HG and Lu SH. Two microRNA panels to discriminate three subtypes of lung carcinoma in bronchial brushing specimens. Am J Respir Crit Care Med. 2012; 186:1160-1167.

17. Võsa U, Kolde R, Vilo J, Metspalu A and Annilo T. Comprehensive meta-analysis of microRNA expression using a robust rank aggregation approach. Methods in Molecular Biology. 2014; 1182:361-373.

18. Arima C, Kajino T, Tamada Y, Imoto S, Shimada Y, Nakatochi M, Suzuki M, Isomura H, Yatabe Y, Yamaguchi T, Yanagisawa K, Miyano S and Takahashi T. Lung adenocarcinoma subtypes definable by lung development-related miRNA expression profiles in association with clinicopathologic features. Carcinogenesis. 2014; 35:2224-2231.

19. Nymark P, Guled M, Borze I, Faisal A, Lahti L, Salmenkivi K, Kettunen E, Anttila S and Knuutila S. Integrative analysis of microRNA, mRNA and aCGH data reveals asbestos- and histology-related changes in lung cancer. Genes Chromosomes Cancer. 2011; 50:585-597.

20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. 2003.

21. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA. DAVID: Database for Annotation, visualization, and Integrated Discovery. Genome Biology. 2003; 4:P3.

22. Aguirregamboa R and Trevino V. SurvMicro: assessment of miRNA-based prognostic signatures for cancer clinical outcomes by multivariate survival analysis. Bioinformatics. 2014; 30:1630.

23. Meng Z, Moroishi T and Guan KL. Mechanisms of Hippo pathway regulation. Genes Dev. 2016; 30:1-17.

24. Yeung B, Yu J and Yang X. Roles of the Hippo pathway in lung development and tumorigenesis. Int J Cancer. 2016; 138:533-539.

25. Das S and Teoh SL. The emerging role of the Hippo pathway in lung cancers: Clinical implications. Curr Drug Targets. 2016.

26. Cersosimo RJ. Lung cancer: a review. American Journal of Health-System Pharmacy. 2002; 59:611-642.

27. Kulasingam V, Pavlou MP and Diamandis EP. Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. Nature Reviews Cancer. 2010; 10:371-378.

28. Pesch B, Kendzia B, Gustavsson P, Jöckel KH, Johnen G, Pohlabeln H, Olsson A, Ahrens W, Gross IM and Brüske I. Cigarette smoking and lung cancer--relative risk estimates for the major histological types from a pooled analysis of case-control studies. International Journal of Cancer. 2010; 131:1210–1219.

29. Okamoto T, Suzuki Y, Fujishita T, Kitahara H, Shimamatsu S, Kohno M, Morodomi Y, Kawano D and Maehara Y. The prognostic impact of the amount of tobacco smoking in non-small cell lung cancer--differences between adenocarcinoma and squamous cell carcinoma. Lung Cancer. 2014; 85:125-130.

30. FR H, Spreafico A, Novello S, Wood MD, Simms L and Papotti M. The prognostic and predictive role of histology in advanced non-small cell lung cancer: a literature review. Journal of Thoracic Oncology. 2008; 3:1468-1481.

31. Travis WD, Garg K, Franklin WA, Wistuba II, Sabloff B, Noguchi M, Kakinuma R, Zakowski M, Ginsberg M and Padera R. Bronchioloalveolar carcinoma and lung adenocarcinoma: the clinical importance and research relevance of the 2004 World Health Organization pathologic criteria. Journal of Thoracic Oncology Official Publication of the International Association for the Study of Lung Cancer. 2006; 1:S13-19.

32. Usui S, Minami Y, Shiozawa T, Iyama S, Satomi K, Sakashita S, Sato Y and Noguchi M. Differences in the prognostic implications of vascular invasion between lung adenocarcinoma and squamous cell carcinoma. Lung Cancer. 2013; 82:407-412.

33. D'Amico TA. Angiogenesis in non-small cell lung cancer. Seminars in Thoracic & Cardiovascular Surgery. 2004; 16:13-18.

34. Besse B, Ropert S and Soria JC. Targeted therapies in lung cancer. Annals of Oncology Official Journal of the European Society for Medical Oncology. 2007; 18:ix135-142.

35. Chirieac LR and Dacic S. Targeted Therapies in Lung Cancer. Surgical Pathology Clinics. 2010; 3:71-82.

36. Landi MT, Zhao Y, Rotunno M, Koshiol J, Liu H, Bergen AW, Rubagotti M, Goldstein AM, Linnoila I and Marincola FM. MicroRNA expression differentiates histology and predicts survival of lung cancer. Clinical Cancer Research. 2010; 16:430-441.

37. Bishop JA, Benjamin H, Cholakh H, Chajut A, Clark DP and Westra WH. Accurate classification of non-small cell lung carcinoma using a novel microRNA-based approach. Clinical Cancer Research An Official Journal of the American Association for Cancer Research. 2010; 16:610-619.

38. Zhang YK, Zhu WY, He JY, Chen DD, Huang YY, Le HB and Liu XG. miRNAs expression profiling to distinguish lung squamous-cell carcinoma from adenocarcinoma subtypes. Journal of cancer research and clinical oncology. 2012; 138:1641-1650.

39. Canueto J, Cardenoso-Alvarez E, Garcia-Hernandez JL, Galindo-Villardon P, Vicente-Galindo P, Vicente-Villardon JL, Alonso-Lopez D, De Las Rivas J, Valero J, Moyano-Saez E, Fernandez-Lopez E, Mao JH, Castellanos-Martin A, Roman-Curto C and Perez-Losada J. miR-203 and miR-205 expression patterns identify subgroups of prognosis in cutaneous squamous cell carcinoma. The British journal of dermatology. 2016.

40. Yu H, Jiang L, Sun C, Li Guo L, Lin M, Huang J and Zhu L. Decreased circulating miR-375: a potential biomarker for patients with non-small-cell lung cancer. Gene. 2014; 534:60-65.

41. Li Y, Jiang Q, Xia N, Yang H and Hu C. Decreased expression of microRNA-375 in nonsmall cell lung cancer and its clinical significance. J Int Med Res. 2012; 40:1662-1669.

42. Zhou Z, Hao Y, Liu N, Raptis L, Tsao MS and Yang X. TAZ is a novel oncogene in non-small cell lung cancer. Oncogene. 2011; 30:2181-2186.

43. Wang Y, Dong QQ, Li Z, Wang E and Qiu X. Overexpression of yes-associated protein contributes to progression and poor prognosis of non-small-cell lung cancer. Cancer Science. 2010; 101:1279-1285.

44. Cui ZL, Han FF, Peng XH, Chen X, Luan CY, Han RC, Xu WG and Guo XJ. YES-associated protein 1 promotes adenocarcinoma growth and metastasis through activation of the receptor tyrosine kinase Axl. Int J Immunopathol Pharmacol. 2012; 25:989-1001.

45. Rekhtman N, Ang DC, Sima CS, Travis WD and Moreira AL. Immunohistochemical algorithm for differentiation of lung adenocarcinoma and squamous cell carcinoma based on large series of whole-tissue sections with validation in small specimens. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2011; 24:1348-1359.

46. Yang Z, Zhuan B, Yan Y, Jiang S and Wang T. Integrated analyses of copy number variations and gene differential expression in lung squamous-cell carcinoma. Biological research. 2015; 48:47.

47. Morita K, Tsukita S and Miyachi Y. Tight junction-associated proteins (occludin, ZO-1, claudin-1, claudin-4) in squamous cell carcinoma and Bowen's disease. The British journal of dermatology. 2004; 151:328-334.

48. Wang Y, Zhou J, Zeng F, Huang Y, Zhou S and Liu X. [Expression and clinical significance of ZO-1 in patients with

non-small cell lung cancer]. Zhongguo Fei Ai Za Zhi. 2011; 14:146-150.

49. Ni S, Xu L, Huang J, Feng J, Zhu H, Wang G and Wang X. Increased ZO-1 expression predicts valuable prognosis in non-small cell lung cancer. Int J Clin Exp Pathol. 2013; 6:2887-2895.

50. Sun N, Wu Y, Huang B, Liu Q, Dong Y, Ding J and Liu Y. Decreased expression of 14-3-3 sigma, an early event of malignant transformation of respiratory epithelium, also facilitates progression of squamous cell lung cancer. Thoracic cancer. 2015; 6:715-721.