

## Prognosis related miRNAs, DNA methylation, and epigenetic interactions in lung adenocarcinoma

H. LIU<sup>1</sup>, H. ZHAO<sup>2,\*</sup>

<sup>1</sup>Cardiovascular Medicine Center, the Affiliated Hospital of Yan'an University, Yan'an, Shaanxi 716000, China; <sup>2</sup>Department of Thoracic Surgery, Shangluo Central Hospital, Shangluo, Shaanxi 726000, China

\*Correspondence: zhaohuaqian@aliyun.com

Received October 29, 2018 / Accepted January 16, 2019

Our study aimed to identify prognosis related epigenetic interactions of DNA methylation-miRNA-gene in lung adenocarcinoma. The RNA-seq, DNA methylation and miRNA-seq data of squamous cell cancer samples were downloaded from TCGA. The DNA methylation-miRNA-gene interactions were collected via Illumina methylation platform and miRTarBase database. Linear regression model was utilized for the identification of epigenetic interactions. The epigenetic interactions related to prognosis were selected via Kaplan-Meier analysis. Genes in the interactions were used for pathway enrichment. Differentially expressed genes (DEGs) between high methylation level/high miRNA expression level (H/H) and low methylation level/low miRNA expression level (L/L) samples were screened. The correlations of epigenetic interactions with clinical features were also explored. Total of 454 lung adenocarcinoma patient samples were collected. The 1063 interactions were comprised of 1083 DNA methylation probes, 271 miRNAs and 528 genes, including cg14146378-hsa-mir-205-ARID1B, cg15375596-hsa-mir-1275-IGF1R, cg26691953-hsa-mir-195-CCNT1, etc. A total of 95 epigenetic interactions were significantly associated with prognosis. Among all the identified DEGs, low-expressed RASSF4, ZNF704, TFDP1, PLXNB2, TMC04, ZNF878, ARID1B and high-expressed ZNF704, ZNF451, THOP1, IGF1R were related with poor prognosis, while low-expressed LDHB, ARID2, PRKCSH, HDAC4, NIPA1, RABAC1, TRIM28 and high-expressed FAM160B1, DNAAF3, CCNT1, ADAP1, ZFPM1, CCL11 were related with good prognosis. Fifteen epigenetic interactions were significantly related with clinical features. Gene expression and N-glycan trimming in the ER and Calnexin/Calreticulin cycle were two significant enriched pathways. Interactions of cg14146378-hsa-mir-205-ARID1B and cg15375596-hsa-mir-1275-IGF1R may be used as prognosis indicators in lung adenocarcinoma.

*Key words: epigenetic interaction, RNA-seq, DNA methylation, lung adenocarcinoma, prognosis*

Lung cancer is a malignant lung tumor characterized by uncontrolled cell growth in lung tissues. Lung cancer ranks as the most common cause of cancer-related death in men and second most common one in women, with 1.6 million deaths among 1.8 million patients worldwide [1, 2]. The two main types of lung cancer are small cell lung carcinoma (SCLC) and non-small cell carcinoma (NSCLC). NSCLC comprises approximately 85% of all diagnoses lung malignancies, with squamous cell carcinoma and adenocarcinoma as the two pathological types [3, 4]. Lung adenocarcinoma is the most common histologic type of NSCLC [5]. Moreover, more than half of all lung adenocarcinoma cases present with advanced metastatic disease and associated with a 5-year survival rate of 4% [6]. Though small molecule tyrosine kinase inhibitors and immunotherapy could led to obvious survival benefits in selected patients, the overall cure and survival rates for

NSCLC particularly in metastatic disease remain low [7]. Therefore, continued research into molecular mechanisms is required to expand the clinical benefit to improve prognosis in patients with lung adenocarcinoma.

It is of great importance to identify the molecular regulators during the occurrence and development of cancers. Epigenetic alterations are remarkably associated with the development of cancer, and it is reported that DNA methylation plays critical roles in lung tumorigenesis and may potentially be proposed as a diagnostic biomarker [8]. MiRNAs can also be used as treatment target and prognosis indicators for cancer [9, 10]. Previous study showed there are synergistic regulatory mechanisms between DNA methylation and miRNA within the epigenome of cancer associated genes, the interactions of which are related with survival [11]. Shiva-kumar et al. found 11 significant epigenetic interactions

between DNA methylation and miRNA, which target E2F3, CCND1, UTP6, CDADC1, SLC35E3, METRNL, TPCN2, NACC2, VGLL4, and PTEN, and they were related with survival of patients with bladder cancer [12]. He et al. found that methylation-regulated miR-149 modulates chemoresistance by targeting GlcNAc N-deacetylase/N-sulfotransferase-1 in human breast cancer [13]. Pal et al. revealed that the key developmental miRNAs are regulated by global changes in histone modification, thus linking the mammary epigenome with genome-wide changes in the expression of genes and miRNAs [14]. However, the influence of the DNA methylation and miRNAs to transcriptional expression changes in lung adenocarcinoma remains unclear.

In this study, DNA methylation, miRNA-seq, and mRNA-seq data of lung adenocarcinoma samples were downloaded from The Cancer Genome Atlas (TCGA). These data were merged for integrated analysis of the prognosis related DNA methylation-miRNA-gene interactions.

## Materials and methods

**Data resources and preprocessing.** The DNA methylation, miRNA expression and gene expression data of lung adenocarcinoma samples with available clinical information were downloaded from TCGA (<https://cancergenome.nih.gov/>) on 10 Marth, 2018. These three types of data were based on Infinium HM450 BeadChip platform, IL/Lumina HiSeq miRNA-Seq platform, and IL/Lumina HiSeq RNA-Seq platform, respectively. Clinical informations of these samples were also available (Table 1).

The miRNA and mRNA expression profiles were normalized by the standard quantile normalization method using preprocessCore Version 1.40.0 in R3.4.1 [15]. DNA methylation probes with “NA” information were deleted. After standardization, all data were log<sub>2</sub> transformed for further analysis.

**Target genes of DNA methylation and miRNAs.** DNA methylation data were mapped to the relative genes using the recorded DNA methylation and gene pairs in Illumina

methylation platform (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>) [16, 17]. Target genes of miRNAs were obtained from the recorded pairs in miRTarBase database (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>) [18].

**Epigenetic interactions.** Linear regression model was utilized for identification of epigenetic (DNA methylation-miRNA-gene) interactions. Likelihood ratio test (LRT) was conducted between full model (gene expression=miRNA + methylation + miRNA × methylation) and reduced model (gene expression=miRNA + methylation) [19]. The epigenetic interactions with Bonferroni adjusted p-value <0.05 were selected as significantly correlated epigenetic interactions.

**Pathway enrichment.** Genes in the significantly correlated epigenetic interactions were used for pathway enrichment by over-representation analysis based on Consensus-PathDB database (<http://cpdb.molgen.mpg.de/>) [20, 21]. The parameter q-value <0.05 was the cut-off threshold.

**Survival analysis.** In order to investigate the connections between epigenetic interactions and patient prognosis, samples were firstly divided into 9 subgroups according to the DNA methylation levels and miRNA expression levels. The 9 subgroups included low methylation level and low miRNA expression level (L/L), low methylation level and high miRNA expression level (L/H), low methylation level and median miRNA expression level (L/M), median methylation level and low miRNA expression level (M/L), median methylation level and high miRNA expression level (M/H), median methylation level and median miRNA expression level (M/M), high methylation level and low miRNA expression level (H/L), and high methylation level and high miRNA expression level (H/H), and high methylation level and median miRNA expression level (H/M). The survival comparison was performed between H/H and L/L subgroups using the Kaplan-Meier analysis. The p<0.05 was considered as significant.

**Differentially expressed genes (DEGs).** The genes in the significantly correlated epigenetic interactions that were differentially expressed between H/H and L/L subgroups were considered as the prognosis-related genes. The p<0.05 obtained from Student’s T-test was the cut-off threshold. The DEGs among the other four subgroups (L/L, L/H, H/L and H/H) were also identified using analysis of variance (ANOVA), and p<0.05 was set as the threshold.

**Correlation analysis of epigenetic interactions with clinical features.** The correlation analysis of epigenetic interactions with clinical features were also conducted. The samples were divided into different groups according to the clinical features, including the pathologic stage (stage I, stage IA, and stage IB samples were merged as stage I samples; stage II, stage IIA, and stage IIB samples were merged as stage II samples; stage III, stage IIIA, stage IIIB, stage IV were merged as ≥stage III samples), and measure of response (complete response, partial response, and stable disease samples were merged as benefit group; clinical progressive disease samples were recognized as non-benefit group). The correla-

**Table 1. Clinical information of lung adenocarcinoma samples in TCGA database.**

Clinical variables	Values (N=454)
Sex (Male/Female)	210/244
Age (Mean±SD)	65.68±10.03
Stage (I/II/III/IV/NA)	246/110/73/20/5
Smoking status (Smoker/Non-smoker/NA)	177/266/11
Smoking history (1/2/3/4/5/NA)	66/104/118/145/4/17
Number pack years smoked (Mean±SD)	40.89±27.11
Location in lung parenchyma (Central/Peripheral/NA)	53/113/288
Residual tumor (R0/R1/R2/RX/NA)	292/11/1/12/128
Drug response (Benefit/No-benefit/NA)	180/68/206

NA, data unavailable.

tions of epigenetic interactions with pathologic stage, race, icd\_o\_3\_site, tobacco smoking history, residual tumor were analyzed via Chi square test, the correlations with measure of response, location in lung parenchyma were analyzed via Fisher test, and correlations with number pack years smoked was analyzed via T test.

## Results

**Data resources and target genes.** The DNA methylation, miRNA expression and gene expression data of tissue samples from 454 patients with lung adenocarcinoma were collected. After preprocessing, the expression data of 395,945 DNA methylation probes, 632 miRNAs, and 31 035 genes were collected for further analysis.

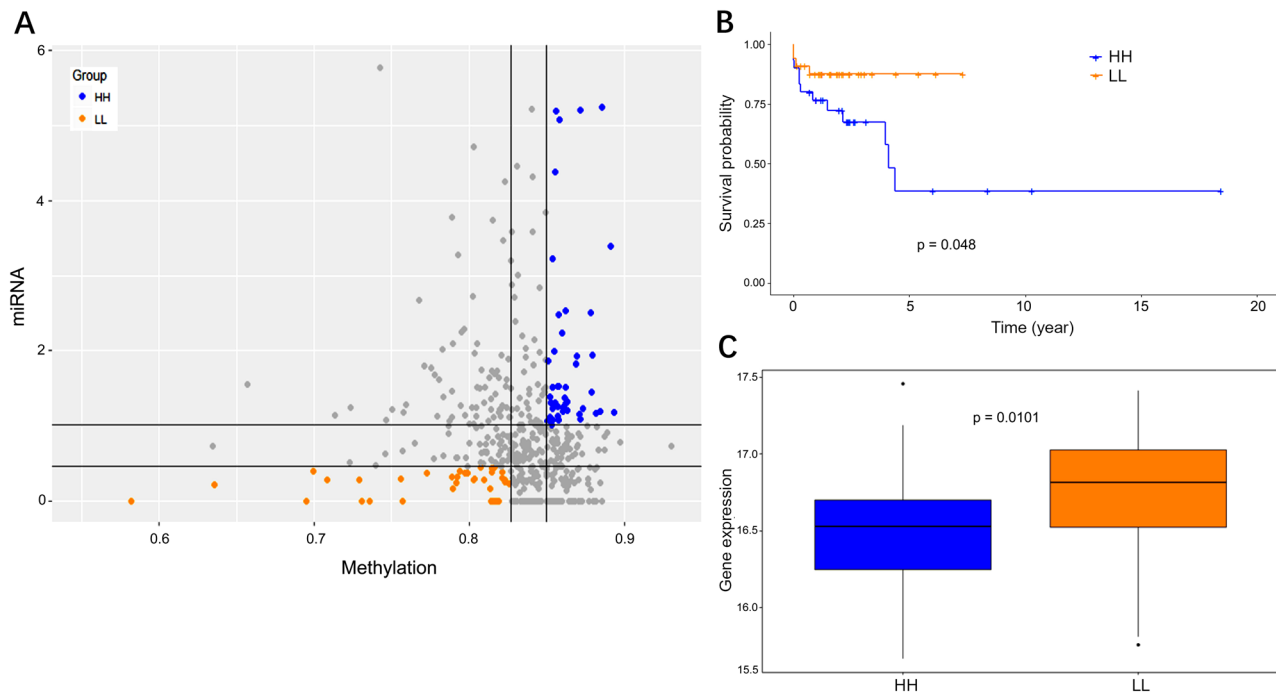
**Epigenetic interactions and pathway enrichment.** A total of 1063 epigenetic interactions between miRNAs and DNA methylations were found to be significantly corrected with the gene expressions (Bonferroni adjusted  $p$ -value  $< 0.05$ ), which involved with 1,083 DNA methylation probes, 271 miRNAs, and 528 genes. The interactions included cg14146378-hsa-mir-205-ARID1B, cg15375596-hsa-miR-1275-IGF1R, cg26691953-hsa-mir-195-CCNT1, cg06126698-hsa-mir-183-ZFPM1, cg11753018-hsa-mir-47b-RASSF4, cg02703822-hsa-mir-942-ZNF704, cg04364540-hsa-mir-133b-FAM160B1, etc.

The top 20 significant pathways were listed in Table 2, such as Gene Expression ( $q$ -value  $= 7.83 \times 10^{-5}$ ), N-glycan

trimming in the ER and Calnexin/Calreticulin cycle ( $q$ -value  $= 5.52 \times 10^{-4}$ ), Signaling by TGF-beta Receptor Complex ( $q$ -value  $= 5.89 \times 10^{-4}$ ), Generic Transcription Pathway ( $q$ -value  $= 6.24 \times 10^{-4}$ ), and Josephin domain DUBs ( $q$ -value  $= 6.45 \times 10^{-4}$ ) (Table 2).

**Prognosis related epigenetic interactions.** A total of 95 epigenetic interactions were found to be significantly associated with prognosis ( $p < 0.05$ ). For example, the samples were divided into 9 subgroups based on the methylation level and miRNA expression level (Figure 1A) for the interaction of cg14146378-hsa-mir-205-ARID1B. The prognosis between H/H group samples and L/L group samples was significantly different ( $p = 0.048$ ) (Figure 1B). The result showed that high cg14146378 methylation level and high hsa-mir-205 expression samples had a worse prognosis. The ARID1B expression level in H/H group is obviously lower than that in L/L group ( $p = 0.0101$ ) (Figure 1C). There were 76 common target gene among the 95 prognosis-related epigenetic interactions, such as PTEN, ERBB2, MAP3K2, ARID1B, CREB1, TRIM28, TFPP1, LAST1, IGF1R, and TRAF.

**Differentially expressed genes.** The identified DEGs involved 25 epigenetic interactions by comparing the expression levels of the 76 common target genes mentioned in H/H and L/L samples. The results showed that RASSF4, ZNF704, LDHB, TFDP1, ARID2, PLXNB2, PRKCSH, TRIM28, HDAC4, TMCO4, NIPA1, ZNF 728, RABAC1, and ARID1B expression levels in H/H groups were significantly lower than that in L/L group ( $p$ -value  $< 0.05$ ), while



**Figure 1.** Epigenetic interacting pairs divided subgroups and survival comparison. A) the stratification of samples via methylation levels and miRNA expression levels; B) Kaplan-Meier curve; C) gene expression level in high methylation level / high miRNA expression level (H/H) and low methylation level / low miRNA expression level (L/L) sample group.

**Table 2. The top 20 significant over-representation pathways of genes in the DNA-methylation-miRNA-gene interacting groups.**

Pathways	q-value
Gene Expression	$7.83 \times 10^{-5}$
Disease	$5.52 \times 10^{-4}$
N-glycan trimming in the ER and Calnexin/Calreticulin cycle	$5.52 \times 10^{-4}$
Signaling by TGF-beta Receptor Complex	$5.89 \times 10^{-4}$
Generic Transcription Pathway	$6.24 \times 10^{-4}$
Josephin domain DUBs	$6.45 \times 10^{-4}$
Signaling by ERBB2	$6.45 \times 10^{-4}$
Signaling by PDGF	$6.45 \times 10^{-4}$
Downregulation of ERBB2 signaling	$6.45 \times 10^{-4}$
PIP3 activates AKT signaling	$6.45 \times 10^{-4}$
Transcriptional Regulation by TP53	$6.45 \times 10^{-4}$
PI3K/AKT activation	$6.92 \times 10^{-4}$
GAB1 signalosome	$6.92 \times 10^{-4}$
Oncogene Induced Senescence	$7.65 \times 10^{-4}$
Downstream signal transduction	$1.06 \times 10^{-3}$
Axon guidance	$1.14 \times 10^{-3}$
Signaling by FGFR	$1.14 \times 10^{-3}$
Diseases of signal transduction	$1.14 \times 10^{-3}$
EGFR downregulation	$1.14 \times 10^{-3}$
Signaling by FGFR2	$1.14 \times 10^{-3}$

FAM160B1, DNAAF3, ZNF451, CCNT1, THOP1, IGF1R, ADAP1, ZFPM1, and CCL11 expression levels in H/H group were significantly higher than that in L/L group (p-value <0.05, Table 3). Moreover, low-expressed RASSF4, ZNF704, TFDPI1, PLXNB2, TMC04, ZNF878, ARID1B and high-expressed ZNF704, ZNF451, THOP1, IGF1R were related with poor prognosis. Low-expressed LDHB, ARID2, PRKCSH, HDAC4, NIPA1, RABAC1, TRIM28 and high-expressed FAM160B1, DNAAF3, CCNT1, ADAP1, ZFPM1, CCL11 were related with good prognosis.

**Correlation of epigenetic interactions with clinical features.** There were 15 epigenetic interactions that were significantly related to clinical features among the 25 epigenetic interactions contained DEGs (Table 4). The cg15375596-has-miR-1275-IGF1R was correlated with pathologic stage, six epigenetic interactions (cg11753018-hsa-mir-147b-RASSF4, cg02703822-hsa-mir-942-ZNF704, cg25139110-hsa-mir-30e-TFDPI1, cg07276957-hsa-mir-374a-ARID2, cg22103585-hsa-mir-335-CCNT1, and cg08130793-hsa-mir-125a-HDAC4) were related to tobacco smoking history, two epigenetic interactions (cg12944146-hsa-mir-615-PLXNB2 and cg05663122-hsa-mir-342-TRIM28) were related to number pack years smoked, two epigenetic interactions

**Table 3. Differentially expressed genes between H/H and L/L group samples.**

Gene	Methylation	miRNA	LRT corrected p-value	Survival analysis p-value	T test p-value
RASSF4	cg11753018	hsa-mir-147b	$4.58 \times 10^{-6}$	$5.33 \times 10^{-3}$	$7.04 \times 10^{-3}$
ZNF704	cg02703822	hsa-mir-942	$2.99 \times 10^{-2}$	$5.59 \times 10^{-3}$	$2.46 \times 10^{-3}$
FAM160B1	cg04364540	hsa-mir-133b	$2.98 \times 10^{-3}$	$6.49 \times 10^{-3}$	$6.69 \times 10^{-3}$
LDHB	cg02659794	hsa-mir-197	$3.16 \times 10^{-4}$	$7.09 \times 10^{-3}$	$3.88 \times 10^{-3}$
TFDPI1	cg25139110	hsa-mir-30e	$5.00 \times 10^{-3}$	$8.82 \times 10^{-3}$	$3.29 \times 10^{-2}$
ARID2	cg07276957	hsa-mir-374a	$5.31 \times 10^{-3}$	$1.09 \times 10^{-2}$	$4.91 \times 10^{-2}$
PLXNB2	cg12944146	hsa-mir-615	$4.22 \times 10^{-2}$	$1.12 \times 10^{-2}$	$3.29 \times 10^{-2}$
ZNF704	cg02703822	hsa-mir-23b	$2.42 \times 10^{-2}$	$1.68 \times 10^{-2}$	$4.48 \times 10^{-2}$
DNAAF3	cg18326578	hsa-mir-34c	$4.29 \times 10^{-3}$	$1.77 \times 10^{-2}$	$3.51 \times 10^{-6}$
ZNF451	cg09757620	hsa-mir-6808	$3.13 \times 10^{-2}$	$1.94 \times 10^{-2}$	$7.39 \times 10^{-5}$
PRKCSH	cg05358168	hsa-mir-361	$1.27 \times 10^{-4}$	$2.06 \times 10^{-2}$	$5.94 \times 10^{-3}$
CCNT1	cg22103585	hsa-mir-335	$2.09 \times 10^{-3}$	$2.17 \times 10^{-2}$	$3.22 \times 10^{-8}$
HDAC4	cg08130793	hsa-mir-125a	$5.94 \times 10^{-4}$	$2.36 \times 10^{-2}$	$1.34 \times 10^{-2}$
THOP1	cg03651904	hsa-mir-484	$1.61 \times 10^{-2}$	$2.42 \times 10^{-2}$	$7.10 \times 10^{-4}$
PLXNB2	cg23041250	hsa-mir-615	$1.12 \times 10^{-2}$	$3.27 \times 10^{-2}$	$4.04 \times 10^{-2}$
IGF1R	cg15375596	hsa-mir-1275	$1.33 \times 10^{-2}$	$3.31 \times 10^{-2}$	$4.64 \times 10^{-3}$
ADAP1	cg27572072	hsa-mir-3614	$8.14 \times 10^{-5}$	$3.44 \times 10^{-2}$	$1.21 \times 10^{-3}$
TMC04	cg27230749	hsa-mir-615	$1.63 \times 10^{-4}$	$3.70 \times 10^{-2}$	$1.06 \times 10^{-2}$
ZFPM1	cg04464062	hsa-mir-29c	$5.83 \times 10^{-3}$	$3.72 \times 10^{-2}$	$6.96 \times 10^{-4}$
NIPA1	cg03382910	hsa-mir-1468	$2.43 \times 10^{-6}$	$3.91 \times 10^{-2}$	$3.70 \times 10^{-4}$
ZNF878	cg26626525	hsa-mir-550a-3	$3.42 \times 10^{-2}$	$4.39 \times 10^{-2}$	$1.40 \times 10^{-4}$
RABAC1	cg00471664	hsa-mir-150	$1.76 \times 10^{-4}$	$4.43 \times 10^{-2}$	$6.76 \times 10^{-3}$
ARID1B	cg14146378	hsa-mir-543	$1.46 \times 10^{-3}$	$4.77 \times 10^{-2}$	$1.01 \times 10^{-2}$
TRIM28	cg05663122	hsa-mir-342	$4.58 \times 10^{-4}$	$4.84 \times 10^{-2}$	$3.91 \times 10^{-2}$
CCL11	cg08499840	hsa-mir-766	$3.63 \times 10^{-4}$	$4.91 \times 10^{-2}$	$4.72 \times 10^{-2}$

\*LRT, likelihood ratio test.

**Table 4. Significant epigenetic interacting pairs with clinical features.**

Gene	Methylation	miRNA	Clinical Feature	p-value
IGF1R	cg15375596	hsa-mir-1275	stage event pathologic stage	$3.15 \times 10^{-2}$
RASSF4	cg11753018	hsa-mir-147b	tobacco smoking history	$7.44 \times 10^{-4}$
ZNF704	cg02703822	hsa-mir-942	tobacco smoking history	$1.57 \times 10^{-2}$
TFDP1	cg25139110	hsa-mir-30e	tobacco smoking history	$8.05 \times 10^{-3}$
ARID2	cg07276957	hsa-mir-374a	tobacco smoking history	$4.29 \times 10^{-3}$
CCNT1	cg22103585	hsa-mir-335	tobacco smoking history	$4.07 \times 10^{-2}$
HDAC4	cg08130793	hsa-mir-125a	tobacco smoking history	$3.77 \times 10^{-4}$
PLXNB2	cg12944146	hsa-mir-615	number pack years smoked	$4.35 \times 10^{-2}$
TRIM28	cg05663122	hsa-mir-342	number pack years smoked	$4.62 \times 10^{-2}$
RASSF4	cg11753018	hsa-mir-147b	location in lung parenchyma	$1.83 \times 10^{-2}$
FAM160B1	cg04364540	hsa-mir-133b	location in lung parenchyma	$1.32 \times 10^{-2}$
LDHB	cg02659794	hsa-mir-197	measure of response	$4.58 \times 10^{-2}$
TFDP1	cg25139110	hsa-mir-30e	measure of response	$1.07 \times 10^{-3}$
PLXNB2	cg12944146	hsa-mir-615	measure of response	$2.38 \times 10^{-2}$
PLXNB2	cg23041250	hsa-mir-615	measure of response	$8.12 \times 10^{-4}$

(cg11753018-hsa-mir-147b-RASSF4 and cg04364540-hsa-mir-133b-FAM160B1) were related to location in lung parenchyma, and four epigenetic interactions (cg02659794-hsa-mir-197-LDHB, cg25139110-hsa-mir-30e-TFDP1, cg12944146-hsa-mir-615-PLXNB2, and cg23041250-hsa-mir-615-PLXNB2) were related to measure of response.

## Discussion

Based on the public RNA-seq, DNA methylation and miRNA-seq data in TCGA, the DNA methylation-miRNA-gene interacting network in lung adenocarcinoma samples was established. Afterwards, the epigenetic interactions of DNA methylation-miRNA-gene were screened via linear regression analysis. According to the DNA methylation levels and miRNA expression levels, the samples were divided into 9 subgroups for survival analysis. A total of 95 prognosis related epigenetic interactions were collected, such as cg14146378-hsa-mir-205-ARID1B and cg15375596-hsa-mir-1275-IGF1R. Among the 95 prognosis-related epigenetic interactions, 25 interactions showed significantly differential gene expression levels between L/L and H/H samples. Moreover, low-expressed RASSF4, ZNF704, TFDP1, PLXNB2, TMC04, ZNF878, ARID1B and high-expressed ZNF704, ZNF451, THOP1, IGF1R were related with poor prognosis. At the same time, low-expressed LDHB, ARID2, PRKCSH, HDAC4, NIPA1, RABAC1, TRIM28 and high-expressed FAM160B1, DNAAF3, CCNT1, ADAP1, ZFPM1, CCL11 were related with good prognosis.

ARID1B encodes AT-rich interactive domain containing protein 1B, also known as BAF250B. It is a member of the mammalian SWI/SNF/Sucrose Non-Fermentable (SWI/SNF) chromatin remodeling complex [22]. ARID1B gene mutation can cause syndromic and non-syndromic intellectual disability, and it is also related with childhood neuro-

blastoma [23, 24]. ARID1B might play tumor suppressor role in pancreatic cancer since it exhibited significantly reduced expressions in pancreatic cancer samples, especially in samples from advanced-stage tumors when compared to normal pancreas [25]. In our study, ARID1B was found to be downregulated in the H/H samples with worse prognosis which is consistent with this study. However, it has also been reported that high expression of ARID1B expression is associated with poor outcomes of bladder urothelial carcinoma and low ARID1B expression may act as a predictive index for the benefit from adjuvant chemotherapy [26]. This contradiction might be due to the different kinds of tumors. It is reported that ARID1B may cooperate with ARID1A to promote cancer formation in a unique functional dependence [27]. Overexpressed miRNA-205 is identified in many cancers, including lung cancer, breast cancer, prostate cancer, colorectal cancer [28]. miR-205 is involved in the tumorigenesis of non-small cell lung cancer through modulating PTEN signaling pathway [29]. Low expression of miR-205 functions as a prognostic marker in early stage non-small cell lung cancer [30]. In our study, ARID1B was found to be downregulated in the H/H samples with good prognosis. Therefore, it could be inferred that the epigenetic interaction of cg14146378-hsa-mir-205-ARID1B might help the good clinical outcomes by inhibiting the expressions of ARID1B.

Moreover, 15 epigenetic interactions were found significantly related to clinical features among the 25 epigenetic interactions contained DEGs in our study. The cg15375596-hsa-mir-1275-IGF1R was correlated with pathologic stage. The type 1 insulin-like growth factor receptor (IGF1R) is a transmembrane heterotetrameric protein involved in promoting oncogenic transformation, cancer cell growth, survival, and metabolism [31, 32]. According to the study of 184 surgically treated NSCLC patients, high IGF1R and

EGFR co-expression was associated with poor survival [33]. IGF1R gene and protein expressions were higher in squamous cell carcinomas, and a positive IGF1R expression was significantly associated with grade III differentiation [34, 35]. A meta-analysis of 17 independent studies concluded that IGF1R positive expression could act as an unfavorable factor for disease-free survival in NSCLC patients, and IGF1R expression was associated with smoking status and tumor size [36]. The expression of miR-1275 was variable in oral squamous cell carcinoma, with high levels associated to regional lymph node invasion [37]. It has been revealed that miR-1275 exerted its anti-tumor role in lung cancer by regulating FOXC1 expression [38]. In our study, IGF1R was overexpressed in the H/H samples and correlated with poor survival. Therefore, it could be speculated that DNA methylation cg15375596 influenced the miR-1275 expression followed by inducing abnormal IGF1R expression which may be used to predict the prognosis of lung adenocarcinoma.

## Conclusions

Genetic changes including DNA methylation, miRNA and gene abnormality exert important roles in tumorigenesis. The identified epigenetic interactions related to prognosis may help the understanding of cancers. Interactions of cg14146378-hsa-mir-205-ARID1B and cg15375596-hsa-miR-1275-IGF1R may be used as prognosis indicators in lung adenocarcinoma. However, the findings obtained from bioinformatics analysis should be used cautiously before further validations are available.

## References

- [1] FORMAN D, FERLAY J. The global and regional burden of cancer, pp 16–53. In: BW Stewart, CP Wild (Eds.). World Cancer Report 2014. IARC, Lyon 2014, p. 632. ISBN 978-92-832-0443-5
- [2] BRAMBILLA E, TRAVIS WD. Lung cancer, pp 350–361. In: BW Stewart, CP Wild (Eds.). World Cancer Report 2014. IARC, Lyon 2014, p. 632. ISBN 978-92-832-0443-5
- [3] BREMNES RM, BUSUND LT, KILVÆR TL, ANDERSEN S, RICHARDSEN E et al. The Role of Tumor-Infiltrating Lymphocytes in Development, Progression, and Prognosis of Non-Small Cell Lung Cancer. *J Thorac Oncol* 2016; 11: 789–800. <https://doi.org/10.1016/j.jtho.2016.01.015>
- [4] CAMPBELL JD, ALEXANDROV A, KIM J, WALA J, BERGER AH et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* 2016; 48: 607–616. <https://doi.org/10.1038/ng.3564>
- [5] YANG SM, CHEN LW, WANG HJ, CHEN LR, LOR KL et al. Extraction of radiomic values from lung adenocarcinoma with near-pure subtypes in the International Association for the Study of Lung Cancer/the American Thoracic Society/the European Respiratory Society (IASLC/ATS/ERS) classification. *Lung Cancer* 2018; 119: 56–63. <https://doi.org/10.1016/j.lungcan.2018.03.004>
- [6] SIEGEL RL, MILLER KD, JEMAL A. Cancer statistics, 2017. *CA Cancer J Clin* 2017; 67: 7–30. <https://doi.org/10.3322/caac.21387>
- [7] HERBST RS, MORGENZTERN D, BOSHOFF C. The biology and management of non-small cell lung cancer. *Nature* 2018; 553: 446–454. <https://doi.org/10.1038/nature25183>
- [8] SHI YX, WANG Y, LI X, ZHANG W, ZHOU HH et al. Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer. *BMC Genomics* 2017; 18: 901. <https://doi.org/10.1186/s12864-017-4223-3>
- [9] GAO X, WU Y, YU W, LI H. Identification of a seven-miRNA signature as prognostic biomarker for lung squamous cell carcinoma. *Oncotarget* 2016; 7: 81670–81679. <https://doi.org/10.18632/oncotarget.13164>
- [10] WU C, XU B, ZHOU Y, JI M, ZHANG D et al. Correlation between serum IL-1beta and miR-144-3p as well as their prognostic values in LUAD and LUSC patients. *Oncotarget* 2016; 7: 85876–85887. <https://doi.org/10.18632/oncotarget.13042>
- [11] HE DX, GU F, GAO F, HAO JJ, GONG D et al. Genome-wide profiles of methylation, microRNAs, and gene expression in chemoresistant breast cancer. *Sci Rep* 2016; 6: 24706. <https://doi.org/10.1038/srep24706>
- [12] SHIVAKUMAR M, LEE Y, BANG L, GARG T, SOHN KA et al. Identification of epigenetic interactions between miRNA and DNA methylation associated with gene expression as potential prognostic markers in bladder cancer. *BMC Medical Genomics* 2017; 10: 65–75. <https://doi.org/10.1186/s12920-017-0269-y>
- [13] HE DX, GU XT, LI YR, JIANG L, JIN J et al. Methylation-regulated miR-149 modulates chemoresistance by targeting GlcNAc N-deacetylase/N-sulfotransferase-1 in human breast cancer. *FEBS J* 2014; 281: 4718–4730. <https://doi.org/10.1111/febs.13012>
- [14] PAL B, CHEN Y, BERT A, HU Y, SHERIDAN JM et al. Integration of microRNA signatures of distinct mammary epithelial cell types with their gene expression and epigenetic portraits. *Breast Cancer Res* 2015; 17: 85. <https://doi.org/10.1186/s13058-015-0585-0>
- [15] BOLSTAD BM, IRIZARRY RA, ASTRAND M, SPEED TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003; 19: 185–193.
- [16] BIBIKOVA M, LE J, BARNES B, SAEDINIA-MELNYK S, ZHOU L et al. Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics* 2009; 1: 177–200. <https://doi.org/10.2217/epi.09.14>
- [17] BIBIKOVA M, BARNES B, TSAN C, HO V, KLOTZLE B et al. High density DNA methylation array with single CpG site resolution. *Genomics* 2011; 98: 288–295. <https://doi.org/10.1016/j.ygeno.2011.07.007>
- [18] CHOU CH, SHRESTHA S, YANG CD, CHANG NW, LIN YL et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018; 46: D296–D302. <https://doi.org/10.1093/nar/gkx1067>
- [19] WILKS SS. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Ann Math Statist* 1938; 9: 60–62. <https://doi.org/10.1214/aoms/1177732360>

- [20] KAMBUROV A, STELZL U, LEHRACH H, HERWIG R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res* 2013; 41: D793–D800. <https://doi.org/10.1093/nar/gks105>
- [21] KAMBUROV A, WIERLING C, LEHRACH H, HERWIG R. ConsensusPathDB--a database for integrating human interaction networks. *Nucleic Acids Res* 2009; 37: D623–D628. <https://doi.org/10.1093/nar/gkn698>
- [22] RONAN JL, WU W, CRABTREE GR. From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet* 2013; 14: 347–359. <https://doi.org/10.1038/nrg3413>
- [23] SANTEN GW, ATEN E, SUN Y, ALMOMANI R, GILISEN C et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet* 2012; 44: 379–380. <https://doi.org/10.1038/ng.2217>
- [24] SAUSEN M, LEARY RJ, JONES S, WU J, REYNOLDS CP et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat Genet* 2013; 45: 12–17. <https://doi.org/10.1038/ng.2493>
- [25] KHURSHEED M, KOLLA JN, KOTAPALLI V, GUPTA N, GOWRISHANKAR S et al. ARID1B, a member of the human SWI/SNF chromatin remodeling complex, exhibits tumour-suppressor activities in pancreatic cancer cell lines. *Br J Cancer* 2013; 108: 2056–2062. <https://doi.org/10.1038/bjc.2013.200>
- [26] WANG B, XIE H, MA C, ZHANG G, GAN H et al. Expression of ARID1B Is Associated with Poor Outcomes and Predicts the Benefit from Adjuvant Chemotherapy in Bladder Urothelial Carcinoma. *J Cancer* 2017; 8: 3490–3497. <https://doi.org/10.7150/jca.19109>
- [27] HELMING KC, WANG X, WILSON BG, VAZQUEZ F, HASWELL JR et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med* 2014; 20: 251–254. <https://doi.org/10.1038/nm.3480>
- [28] JIANG J, LEE EJ, GUSEV Y, SCHMITTGEN TD. Real-time expression profiling of microRNA precursors in human cancer cell lines. *Nucleic Acids Res* 2005; 33: 5394–5403. <https://doi.org/10.1093/nar/gki863>
- [29] Lei, L., Y. Huang, and W. Gong, miR-205 promotes the growth, metastasis and chemoresistance of NSCLC cells by targeting PTEN. *Oncology reports*, 2013. 30(6): p. 2897–2902
- [30] PARK KS, MOON YW, RAFFELD M, LEE DH, WANG Y et al. High cripto-1 and low miR-205 expression levels as prognostic markers in early stage non-small cell lung cancer. *Lung Cancer* 2018; 116: 38–45. <https://doi.org/10.1016/j.lungcan.2017.12.010>
- [31] MORGILLO F, WOO JK, KIM ES, HONG WK, LEE HY. Heterodimerization of insulin-like growth factor receptor/epidermal growth factor receptor and induction of survivin expression counteract the antitumor action of erlotinib. *Cancer Res* 2006; 66: 10100–10111. <https://doi.org/10.1158/0008-5472.CAN-06-1684>
- [32] FRASCA F, PANDINI G, SCIACCA L, PEZZINO V, SQUATRITO S et al. The role of insulin receptors and IGF-I receptors in cancer and other diseases. *Arch Physiol Biochem* 2008; 114: 23–37. <https://doi.org/10.1080/13813450801969715>
- [33] GATELY K, FORDE L, CUFFE S, CUMMINS R, KAY EW et al. High coexpression of both EGFR and IGF1R correlates with poor patient prognosis in resected non-small-cell lung cancer. *Clin Lung Cancer* 2014; 15: 58–66. <https://doi.org/10.1016/j.clcc.2013.08.005>
- [34] CAPPUZZO F, TALLINI G, FINOCCHIARO G, WILSON RS, LIGORIO C et al. Insulin-like growth factor receptor 1 (IGF1R) expression and survival in surgically resected non-small-cell lung cancer (NSCLC) patients. *Ann Oncol* 2010; 21: 562–567. <https://doi.org/10.1093/annonc/mdp357>
- [35] DZIADZIUSZKO R, MERRICK DT, WITTA SE, MENDOZA AD, SZOSTAKIEWICZ B et al. Insulin-like growth factor receptor 1 (IGF1R) gene copy number is associated with survival in operable non-small-cell lung cancer: a comparison between IGF1R fluorescent in situ hybridization, protein expression, and mRNA expression. *J Clin Oncol* 2010; 28: 2174–2180. <https://doi.org/10.1200/JCO.2009.24.6611>
- [36] ZHAO S, QIU Z, HE J, LI L, LI W. Insulin-like growth factor receptor 1 (IGF1R) expression and survival in non-small cell lung cancer patients: a meta-analysis. *Int J Clin Exp Pathol* 2014; 7: 6694–6704.
- [37] MANIKANDAN M, DEVA MAGENDHRA RAO AK, ARUNKUMAR G, MANICKAVASAGAM M, RAJKUMAR KS et al. Oral squamous cell carcinoma: microRNA expression profiling and integrative analyses for elucidation of tumorigenesis mechanism. *Mol Cancer* 2016; 15: 28. <https://doi.org/10.1186/s12943-016-0512-8>
- [38] MA X, YANG X, BAO W, LI S, LIANG S et al. Circular RNA circMAN2B2 facilitates lung cancer cell proliferation and invasion via miR-1275/FOXK1 axis. *Biochem Biophys Res Commun* 2018; 498: 1009–1015. <https://doi.org/10.1016/j.bbrc.2018.03.105>