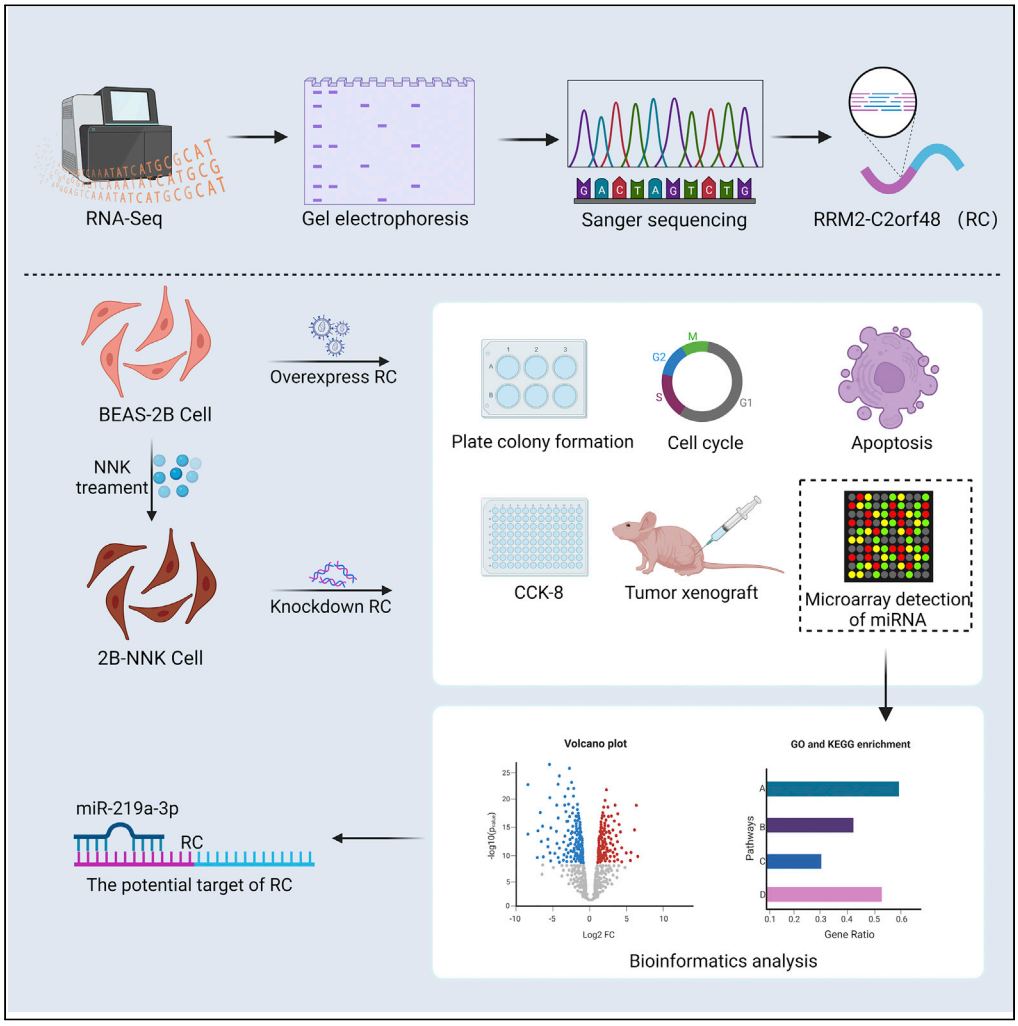
iScience

CellPress
OPEN ACCESS

Article

Chimeric RNA *RRM2-C2orf48* plays an oncogenic role in the development of NNK-induced lung cancer



Jiazhen Zhou, Xinchao Guan, Enwu Xu, Jiaxin Zhou, Rui Xiong, Qiaoyuan Yang

qiaoyuan_yang@gzhmu.edu.cn

Highlights

RRM2-C2orf48 had higher expression levels in lung cancer patients and cell lines

RRM2-C2orf48 expression was upregulated by NNK exposure

RRM2-C2orf48 played an oncogenic role in lung cancer

MiR-219a-2-3p may be a potential target of *RRM2-C2orf48*

Zhou et al., iScience *26*, 105708
January 20, 2023
https://doi.org/10.1016/j.isci.2022.105708

iScience

CellPress
OPEN ACCESS

Article

Chimeric RNA *RRM2-C2orf48* plays an oncogenic role in the development of NNK-induced lung cancer

Jiazhen Zhou,[1,3] Xinchao Guan,[1,3] Enwu Xu,[2] Jiaxin Zhou,[1] Rui Xiong,[1] and Qiaoyuan Yang[1,4,*]

SUMMARY

Chimeric RNAs have been used as biomarkers and therapeutic targets for multiple types of cancers. However, less attention has been paid to their mechanism of action in neoplasia. Here, we reported that high-expressed chimeric RNA *RRM2-C2orf48* was found in malignantly transformed BEAS-2B cells induced by 4-(methyl nitrosamine)-1-(3-pyridinyl)-1-butanone (NNK) in 74 lung cancer patients and several lung cancer cell lines. The expression level of *RRM2-C2orf48* was significantly correlated with lymph node metastasis, distant metastasis, tumor-lymph node-metastasis (TNM) stage, and smoking. Overexpressing *RRM2-C2orf48* promoted cell growth and accelerated the process of NNK-induced lung cancer. *RRM2-C2orf48* knockdown inhibited the growth of *RRM2-C2orf48*-overexpressing BEAS-2B cells. Finally, we identified *miR-219a-2-3p* as a potential target of *RRM2-C2orf48 in lung cancer*. In summary, chimeric RNA *RRM2-C2orf48* accelerated the process of NNK-induced lung cancer, and *miR-219a-2-3p* may be involved in this process.

INTRODUCTION

Lung cancer remains the leading cause of cancer-related deaths globally.[1] It is the leading cause of cancer-related deaths in men and the second most common cause of cancer-related deaths in women. The number of new lung cancer cases and deaths are increasing each year. According to the Global Cancer Observatory, the number of new lung cancer cases in 2020 was more than 2.2 million, which accounted for 11.4% of all cancer cases, and the number of deaths due to lung cancer was more than 1.79 million, which accounted for 18.0% of all cancer-related deaths.[2] The number of new lung cancer cases and related deaths have increased since 2018, in which 2.09 million new cases and 1.76 million deaths were reported.[3]

The risk factors for lung cancer are sex, age, socioeconomic status, race, and geography, but tobacco smoking remains the predominant risk factor.[4,5] There are more than 6,000 different chemicals in tobacco smoke, including nicotine, aromatic amines, N-nitrosamines, aldehydes, and polycyclic aromatic hydrocarbons.[6] One of the biggest concerns is the presence of nicotine as the nitrification of nicotine and alkaloids produces a tobacco-specific nitrosamine, 4-(methyl nitrosamine)-1-(3-pyridinyl)-1-butanone (NNK), which accounts for the largest proportion of tobacco-specific nitrosamines.[7] In recent years, an increasing number of studies have focused on NNK. These studies have found that NNK has a variety of adverse effects on the human body, including DNA damage, oxidative stress, and phosphorylation. Of the tobacco-related carcinogens, NNK is the most potent risk factor for lung cancer progression and poor outcomes.[8] It promotes lung tumorigenesis[9] and increases the invasive ability of lung cancer cells.[10,11] NNK also affects biodegradation and metabolism, and it binds with DNA and hemoglobin in animals, resulting in a high level of carcinogenicity.[12] Although NNK has been found to induce lung cancer through various mechanisms, its oncogenic mechanisms have not been thoroughly elucidated.

Recent studies have identified chimeric RNAs as new potential biomarkers for lung cancer diagnosis and prognosis. A chimeric RNA is an RNA molecule comprising two or more RNA sequences from different loci that are not typically found in the same gene.[13] Chimeric RNAs are new RNA molecules, with or without encoding ability, that often play important carcinogenic roles.[14] Chimeric RNAs were first detected in the 1960s and are thought to be generated solely by the fusion of genes generated by chromosomal rearrangements.[15] For example, some chimeric RNAs are typically present in cancer tissue. These include *BCR-ABL1* in chronic myeloid leukemia[16] and *KIF13B-RET* and *CD74-NRG1* in lung cancer.[17,18] These canonical chimeric RNAs are well characterized and known to be expressed in a cancer-specific pattern.[13,19] Some

[1]The Institute for Chemical Carcinogenesis, School of Public Health, Guangzhou Medical University, Xinzao, Panyu District, Guangzhou 511436, China

[2]Department of Thoracic Surgery, General Hospital of Southern Theater Command, PLA, Guangzhou 510010, PR China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: qiaoyuan_yang@gzhmu.edu.cn

https://doi.org/10.1016/j.isci.2022.105708

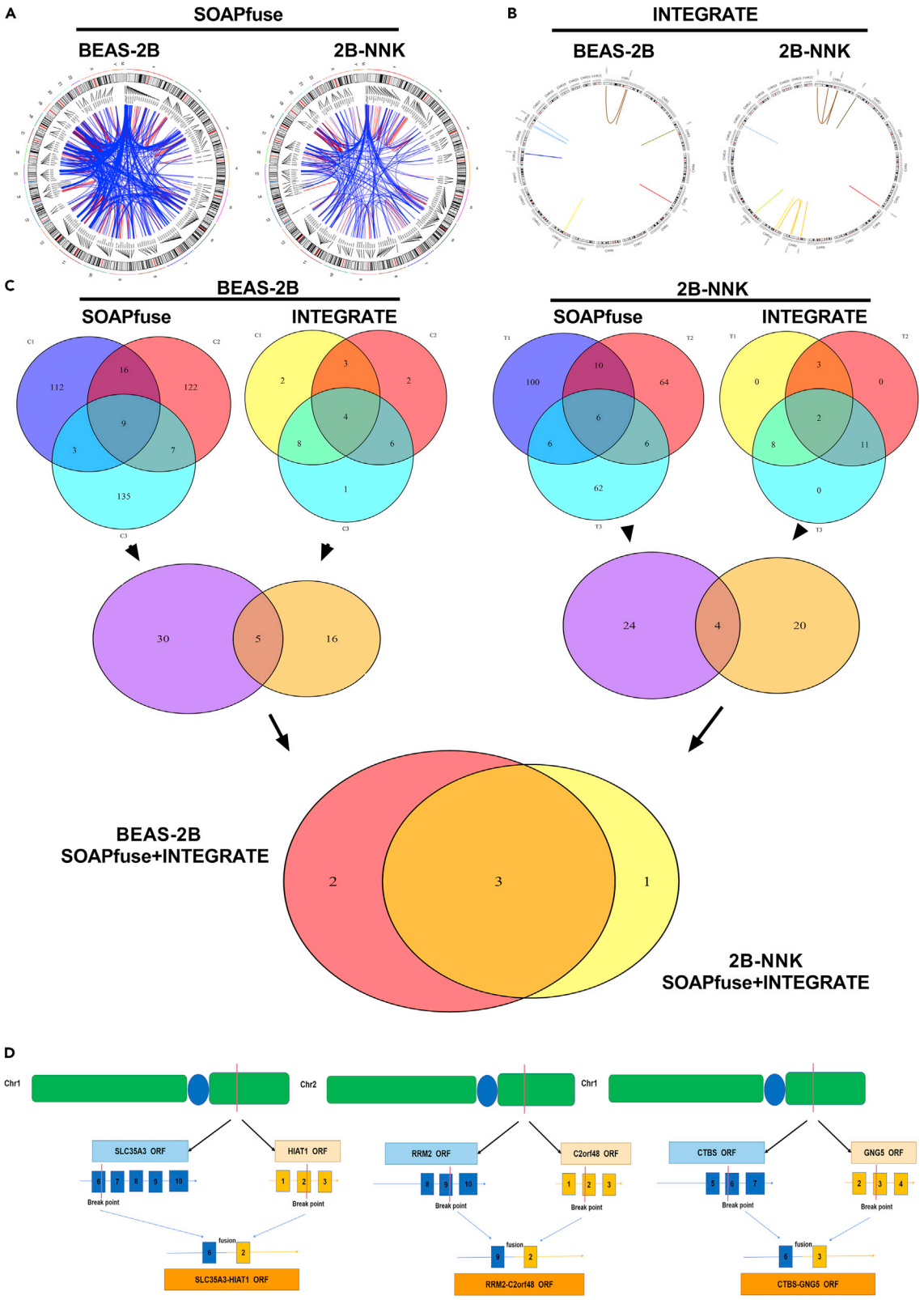

A
SOAPfuse
BEAS-2B
2B-NNK
B
INTEGRATE
BEAS-2B
2B-NNK
C
BEAS-2B
SOAPfuse
INTEGRATE
2B-NNK
SOAPfuse
INTEGRATE
112
16
122
9
3
7
135
2
3
2
4
8
6
1
100
10
64
6
6
6
62
0
3
0
2
8
11
0
30
5
16
24
4
20
BEAS-2B
SOAPfuse+INTEGRATE
2
3
1
2B-NNK
SOAPfuse+INTEGRATE
D
Chr1
Chr2
Chr1
SLC35A3 ORF
HIAT1 ORF
RRM2 ORF
C2orf48 ORF
CTBS ORF
GNG5 ORF
Break point
fusion
SLC35A3-HIAT1 ORF
RRM2-C2orf48 ORF
CTBS-GNG5 ORF

Figure 1. Chimeric RNAs identification

(A and B) Total chimeric RNAs detected in BEAS-2B and 2B-NNK cells using (A) SOAPfuse and (B) INTEGRATE.

(C) An integrated analysis performed using SOAPfuse and INTEGRATE to achieve a higher repeated detection rate. Three chimeric RNAs were selected for further study.

(D) The formation of chimeric RNAs.

chimeric RNAs have been used as targets of clinical therapy in cancer. For example, NTRK-fusion-positive lipofibromatosis-like neural tumors have been successfully treated with an NTRK-fusion inhibitor and *EML4-ALK*-fusion-positive lung cancer is treated with crizotinib.[20,21]

Chimeric RNAs have been found to be generated from *trans*-splicing and *cis*-splicing between adjacent genes (*cis*-SAGe), without genomic aberrations.[22] Recent studies have suggested that chimeric RNAs generated by *cis*-SAGe play a potentially important role in normal biological functions in humans.[23] There have been few studies of *cis*-SAGe chimeric RNAs because only a small number of *cis*-spliced RNA chimeras have been experimentally confirmed in mammalian cells.[24] Despite this, the recurrent *cis*-SAGe chimeric RNAs that have been experimentally confirmed have been found to play important roles in cell growth regulation and are potentially valuable as diagnostic and prognostic cancer biomarkers.[25,26] For example, two *cis*-SAGe chimeric RNAs, *BCL2L2-PABPN1* and *CHFR-GOLGA3*, have been found to be expressed at higher levels in bladder cancer tissue compared to adjacent normal tissue. These two fusions are mainly enriched in the cell nuclear fraction, suggesting a potential long noncoding RNA (lncRNA) role.[27] Similarly, the *cis*-SAGe *SLC45A3-ELK4* is highly expressed in prostate cancer, and its expression level correlates with prostate cancer progression.[28] *SLC45A3-ELK4* functions as a lncRNA, controlling cell proliferation in prostate cancer through its transcript, as it does not generate a translated protein.[29] Moreover, *D2HGDH-GAL3ST2* is more frequently seen in prostate cancer samples than in normal tissue samples. Silencing this chimera results in a marked reduction in cell proliferation rate and cell motility.[30] High levels of *RRM2-C2orf48* expression may be a useful predictor of metastatic potency in nasopharyngeal carcinoma patients, which has potential implications for nasopharyngeal carcinoma diagnosis and therapy.[31] Furthermore, *RRM2-C2orf48* promotes the proliferation of colorectal cancer cells and has a different expression pattern and function than its parental genes.[32] Many previous studies have demonstrated that *cis*-SAGe forms a new RNA molecule that functions like an mRNA or an lncRNA and that these fusion RNAs are usually highly expressed in cancer.[33] However, there are no reported studies of the effect of smoking on these chimeric RNAs.

In this study, we identified three chimeric RNAs, *SLC35A3-HIAT1*, *RRM2-C2orf48*, and *CTBS-GNG5*, using RNA sequencing of BEAS-2B cells malignantly transformed by NNK treatment. Furthermore, we demonstrated that *RRM2-C2orf48* promoted cell growth and was related to the development of NNK-induced lung cancer. This is the first study to demonstrate that the chimeric RNA *RRM2-C2orf48* is involved in the induction of lung cancer by the environmental carcinogen NNK.

RESULTS

Recurrent chimeric RNA identification in BEAS-2B cells malignantly transformed by NNK treatment

To identify chimeric RNAs that may contribute to NNK-induced lung cancer, normal human bronchial epithelial BEAS-2B cells that had been malignantly transformed by treatment with NNK (2B-NNK) and normal BEAS-2B cells were subjected to RNA sequencing. To reduce the false-positive rate for chimeric RNA identification, we used two chimeric RNA identification software tools, SOAPfuse and INTEGRATE, to perform an integrated analysis. As shown in Figures 1A and 1B, SOAPfuse identified 407 chimeric RNA candidates in BEAS-2B cells and 254 in 2B-NNK cells, while INTEGRATE identified 26 chimeric RNA candidates in BEAS-2B cells and 24 in 2B-NNK cells. The chimeric RNAs were detected in at least two samples, and those detected in both BEAS-2B and 2B-NNK cells were further investigated (Figure 1C). Finally, three chimeric RNAs, *SLC35A3-HIAT1*, *RRM2-C2orf48*, and *CTBS-GNG5*, met the criteria. As shown in Figure 1D, the upstream and downstream regions of each chimeric RNA were positioned on the same chromosome, and the distance between the two genes was found to be less than 3 Mb, indicating that these chimeric RNAs may have formed via *cis*-splicing. Previous studies have shown that *RRM2-C2orf48*, *SLC35A3-HIAT1*, and *CTBS-GNG5* form by *cis*-splicing.[31,32,34] Moreover, a detailed analysis of the fusion site of the chimeric RNAs showed that *SLC35A3-HIAT1* comprised the sixth exon of the sense strand of *SLC35A3* and the second exon of the sense strand of *HIAT1*, *RRM2-C2orf48* comprised the ninth exon of

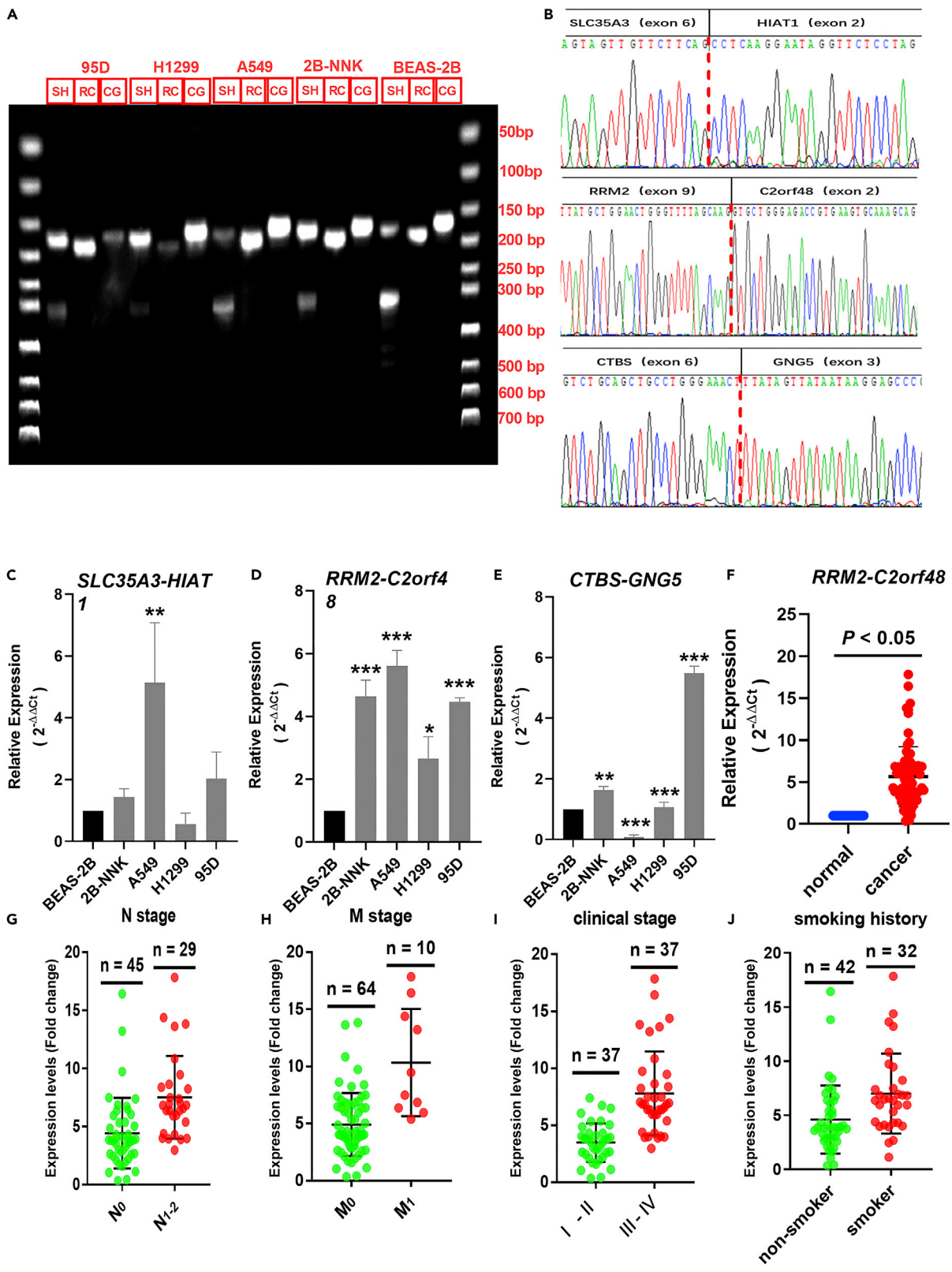

A
95D
H1299
A549
2B-NNK
BEAS-2B
SH RC CG SH RC CG SH RC CG SH RC CG SH RC CG
50bp
100bp
150 bp
200 bp
250 bp
300 bp
400 bp
500 bp
600 bp
700 bp
B
SLC35A3 (exon 6)
HIAT1 (exon 2)
RRM2 (exon 9)
C2orf48 (exon 2)
CTBS (exon 6)
GNG5 (exon 3)
C
SLC35A3-HIAT1
D
RRM2-C2orf48
E
CTBS-GNG5
F
RRM2-C2orf48
P < 0.05
Relative Expression (2-ΔΔCt)
BEAS-2B
2B-NNK
A549
H1299
95D
normal
cancer
G
N stage
H
M stage
I
clinical stage
J
smoking history
Expression levels (Fold change)
n = 45
n = 29
n = 64
n = 10
n = 37
n = 37
n = 42
n = 32
N0
N1-2
M0
M1
I - II
III - IV
non-smoker
smoker

Figure 2. Validation of chimeric RNAs and their expression in lung cancer cells and *RRM2-C2orf48* expression in lung cancer tissues with different clinicopathological characteristics

(A) Representative agarose gel image of RT-PCR products of chimeric RNAs from lung cancer cell lines and the normal human bronchial epithelial cell line BEAS-2B.

(B–E) Representative image of Sanger sequencing results of chimeric RNA RT-PCR products. The expression levels of chimeric RNAs, (C) *SLC35A3-HIAT1*, (D) *RRM2-C2orf48*, and (E) *CTBS-GNG5* in lung cancer cell lines.

(F–J) The expression level of *RRM2-C2orf48* in 74 paired lung cancer tissues and noncancerous lung tissues. Nonparametric assessment of the correlation between *RRM2-C2orf48* expression level and (G) lymph node metastasis, (H) distant metastasis, (I) clinical stage, and (J) smoking in 74 patients with lung cancer. Data are represented as mean ± SEM from at least three independent experiments. Student's *t* test was used for data analysis. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

the sense strand of *RRM2* and the second exon of the sense strand of *C2orf48*, and *CTBS-GNG5* comprised the sixth exon of the antisense strand of *CTBS* and the third exon of the antisense strand of *GNG5*.

Chimeric RNA validation in lung cancer cell lines and tissues

RT-PCR and Sanger sequencing were used to validate the chimeric RNAs. Primers annealing to the parental genes and flanking the fusion junction site were designed. *RRM2-C2orf48*, *SLC35A3-HIAT1*, and *CTBS-GNG5* showed PCR products of the expected size, but *SLC35A3-HIAT1* showed an additional nonspecific product in the RT-PCR assay (Figure 2A). The PCR products were confirmed as the corresponding chimeric RNAs by Sanger sequencing (Figure 2B). To examine the expression levels of these chimeric RNAs in lung cancer, qRT-PCR was performed in four human lung cancer cell lines (2B-NNK, A549, H1299, and 95D) and a normal human bronchial epithelial cell line (BEAS-2B). *RRM2-C2orf48* showed higher expression levels in all four lung cancer cell lines than in BEAS-2B cells. *CTBS-GNG5* showed higher expression levels in 2B-NNK, H1299, and 95D cell lines but a lower expression level in A549 cells than in BEAS-2B cells, whereas *SLC35A3-HIAT1* was expressed at a significantly higher level in A549 cells compared with BEAS-2B cells (all $p < 0.05$, Figures 2C, 2D, and 2E). *RRM2-C2orf48* was found to be most preferentially expressed in lung cancer cell lines.

Thus, we used qRT-PCR to further examine whether *RRM2-C2orf48* was expressed at high levels in lung cancer patients. Unanimously, *RRM2-C2orf48* expression was significantly upregulated in lung cancer tissues from the 74 patients compared with its expression levels in corresponding noncancerous lung tissues ($p < 0.05$, Figure 2F). Meanwhile, higher *RRM2-C2orf48* expression levels were observed in patients with lymph node metastasis (N stage, $p < 0.05$, Figure 2G), with distant metastasis (M stage, $p < 0.05$, Figure 2H), with an advanced tumor-lymph node-metastasis (TNM) stage ($p < 0.05$, Figure 2I), and in those who smoked ($p < 0.05$, Figure 2J) than in patients without N stage, without M stage, without advanced TNM stage, or who did not smoke, respectively. Thus, an increased expression level of *RRM2-C2orf48* was correlated with N stage, M stage, advanced TNM stage, and smoking. However, no significant relationship was found between *RRM2-C2orf48* expression level and age, sex, lung cancer subtype, or the largest tumor dimension ($p > 0.05$, Table 1).

RRM2-C2orf48 was upregulated, and the malignant transformation of BEAS-2B cells was accelerated by NNK treatment

Based on the observations described above, the correlation between smoking and *RRM2-C2orf48* required further investigation. Therefore, we studied the effect of NNK, the main component and most potent carcinogen in cigarette smoke, on the expression level of *RRM2-C2orf48* in BEAS-2B cells. Moreover, colony formation assays were performed to determine whether *RRM2-C2orf48* overexpression accelerated NNK-induced carcinogenesis.

A Cell Counting Kit-8 (CCK-8) proliferation assay was used to assess the effect of NNK on the viability of BEAS-2B cells. A significant reduction in the viability of BEAS-2B cells was seen after 24 h of NNK treatment. After 24 h of treatment with 50, 100, 200, and 400 mg/L NNK, cell viability was reduced by 15.87, 19.23, 26.68, and 52.40%, respectively, indicating a dose-response relationship between NNK and cell viability (Figure 3A). Next, we examined the expression level of *RRM2-C2orf48* in BEAS-2B cells by qRT-PCR after 0, 50, 100, and 200 mg/L NNK treatment. The 400 mg/L dose was not used as it markedly suppressed cell proliferation. We found that *RRM2-C2orf48* expression levels increased after NNK treatment, with the 200 mg/L dose significantly inducing the upregulation of *RRM2-C2orf48* expression (Figure 3B). We then used soft agarose colony formation assays to examine the effect of *RRM2-C2orf48* overexpression on the NNK-induced malignant transformation of BEAS-2B cells. First, we constructed a stable

Table 1. Nonparametric assessment of the correlations between clinicopathological factors and *RRM2-C2orf48* expression levels in 74 patients with lung cancer

Clinical Characteristic	Number of patients (n = 74)	RRM2-C2orf48 expression Mean ± SD	P-value
Age (years)			0.897
≥60	38	5.47 ± 3.43	
<60	36	5.82 ± 3.76	
Gender			0.292
Male	48	5.95 ± 3.75	
Female	26	5.06 ± 3.21	
T stage			0.881
T_{1-2}	4	6.59 ± 5.48	
T_{3-4}	70	5.58 ± 3.49	
N stage			<0.001
N_0	44	4.42 ± 3.04	
N_{1-2}	30	7.53 ± 3.56	
M stage			<0.001
M_0	64	4.91 ± 2.76	
M_1	10	10.34 ± 4.70	
TNM stage			<0.001
I–II	37	3.49 ± 1.69	
III–IV	37	7.79 ± 3.67	
Histological subtype			0.96
LUSC	29	5.52 ± 2.84	
LUAD	40	4.96 ± 2.96	
SCLC	4	13.13 ± 6.33	
LCLC	1	6.37 ± 0	
Largest tumor dimension (cm)			0.604
>3.0	38	5.75 ± 3.48	
≤3.0	36	5.52 ± 3.71	
Smoking history			0.001
Smoker	32	7.00 ± 3.70	
Nonsmoker	42	4.60 ± 3.14	

RRM2-C2orf48-overexpressing model using BEAS-2B cells transfected with a lentivirus vector carrying the *RRM2-C2orf48* sequence (BEAS-2B-RC cells) and a negative control cell model using BEAS-2B cells transfected with an empty lentivirus vector (BEAS-2B-V cells, Figure 3C). Western blotting and qRT-PCR analyses confirmed the successful construction of BEAS-2B-RC cells and that the overexpression of *RRM2-C2orf48* did not affect the expression of its parent genes, *RRM2* and *C2orf48* (Figures 3D–3G). Based on the CCK-8 assay results, we selected 100 mg/L as the dose for the multiple NNK exposure experiment due to its slight suppression of the viability of BEAS-2B cells. We set the observation endpoint of the multiple NNK exposure experiment at the time when a significant increase in colony formation rate was observed in the NNK treatment groups. As expected, we observed a significant increase in colony formation rate in NNK-treated BEAS-2B, BEAS-2B-V, and BEAS-2B-RC cells compared with the corresponding control cells (Figures 3H and 3I). The tumor size of nude mice was measured every 3 days starting on the seventh day after injection; we found that the tumor volume of the BEAS-2B-RC-NNK group was significantly bigger than that of the BEAS-2B-NNK group ($p < 0.05$, Figure 3J). At the observation endpoint, the final NNK exposure doses were 1.0, 1.2, and 1.2 mg for BEAS-2B-RC, BEAS-2B-V, and BEAS-2B cells, respectively (Figure 3K). These results indicated that *RRM2-C2orf48* overexpression may accelerate the malignant transformation of BEAS-2B Cells.

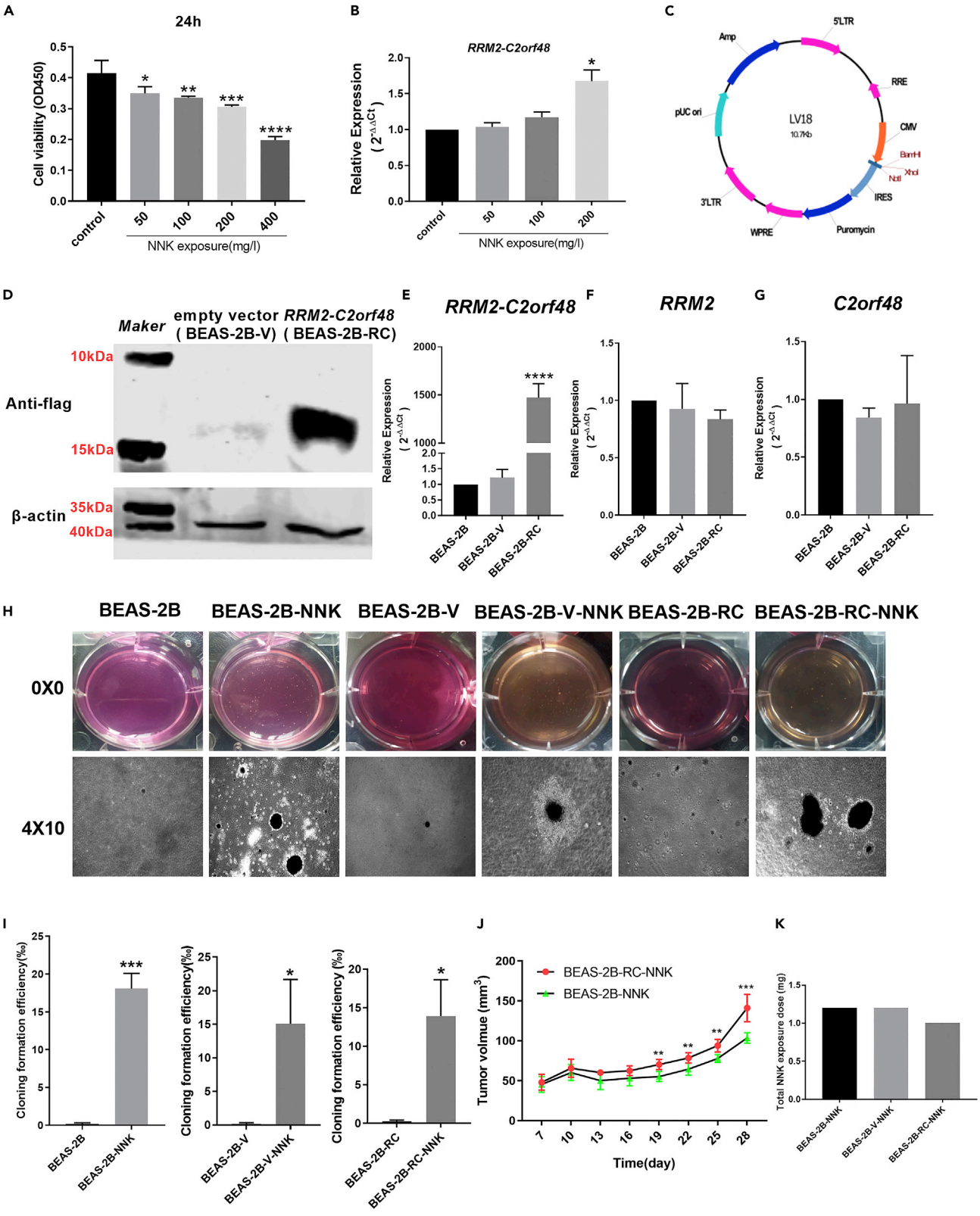

A
24h
Cell viability (OD450)
control
50
100
200
400
NNK exposure(mg/l)
B
RRM2-C2orf48
Relative Expression
control
50
100
200
NNK exposure(mg/l)
C
5'LTR
Amp
RRE
pUC ori
LV18
10.7Kb
CMV
BamHI
XhoI
NotI
IRES
3'LTR
WPRE
Puromycin
D
Maker
empty vector (BEAS-2B-V)
RRM2-C2orf48 (BEAS-2B-RC)
10kDa
Anti-flag
15kDa
β-actin
35kDa
40kDa
E
RRM2-C2orf48
Relative Expression
BEAS-2B
BEAS-2B-V
BEAS-2B-RC
F
RRM2
G
C2orf48
H
BEAS-2B
BEAS-2B-NNK
BEAS-2B-V
BEAS-2B-V-NNK
BEAS-2B-RC
BEAS-2B-RC-NNK
0X0
4X10
I
Cloning formation efficiency(‰)
Cloning formation efficiency (%)
J
BEAS-2B-RC-NNK
BEAS-2B-NNK
Tumor volmue (mm3)
Time(day)
K
Total NNK exposure dose (mg)

Figure 3. Correlation between NNK and *RRM2-C2orf48*

(A) CCK-8 proliferation assays were used to examine the viability of BEAS-2B cells after 24 h of NNK treatment.

(B) The chimeric RNA *RRM2-C2orf48* was detected by qRT-PCR after treating BEAS-2B cells with NNK.

(C) Lentivirus vector model.

(D–G) Western blotting analysis. Quantification of the expression levels of (E) *RRM2-C2orf48*, (F) *RRM2*, and (G) *C2orf48*.

(H) Representative image of soft agarose colony formation assay of BEAS-2B, BEAS-2B-V, and BEAS-2B-RC after NNK treatment.

(I) Quantification of soft agarose colony formation results.

(J) NNK dose used in soft agarose colony formation assays.

(K) Tumor growth volume in the two groups over 28 days; tumor volume is expressed as mm^3. Data are represented as mean ± SEM from at least three independent experiments. Student's *t* test was used for data analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

RRM2-C2orf48 overexpression promoted cell growth in BEAS-2B cells

CCK-8, plate colony formation, and flow cytometry assays were performed on stably overexpressing *RRM2-C2orf48* cells to investigate the function of *RRM2-C2orf48*. The results of CCK-8 assays showed that, compared with BEAS-2B and BEAS-2B-V cells, BEAS-2B-RC cells had higher optical density values, indicating that *RRM2-C2orf48* overexpression promoted cell proliferation ($p < 0.05$, Figure 4A). As seen in Figures 4B and 4C, colony formation efficiency was higher for BEAS-2B-RC cells (50.50%) than BEAS-2B and BEAS-2B-V cells (33.66% and 34.33%, respectively). In addition, *RRM2-C2orf48* overexpression increased the colony formation rate ($p < 0.05$, Figures 4B and 4C). The percentage of apoptotic BEAS-2B and BEAS-2B-V cells was higher than the percentage of apoptotic BEAS-2B-RC cells ($p < 0.05$, Figures 4D and 4E). To investigate the mechanism whereby *RRM2-C2orf48* affects the cell cycle, we analyzed the cell cycle progression after cell cycle synchronization at the G0/G1 and S + G2/M phases. Our results demonstrated that *RRM2-C2orf48* overexpression promoted G0/G1 phase entry to the S + G2/M phase and inhibited S + G2/M exit into the G0/G1 phase (both $p < 0.05$, Figures 4F and 4G), which indicated that *RRM2-C2orf48* overexpression also promoted cell cycle progression.

RRM2-C2orf48 knockdown inhibited the growth of *RRM2-C2orf48*-overexpressing BEAS-2B cells

To further determine whether the effect of NNK on cell proliferation was regulated by *RRM2-C2orf48*, we designed a small interfering RNA (siRNA) targeting the fusion junction site (siRC), as previously described.[32] siRC effectively silenced the chimeric RNA *RRM2-C2orf48* but did not affect its parent genes *RRM2* and *C2orf48* (Figures 5B–5D and 6A–6C). In contrast to the CCK-8 assay results from the *RRM2-C2orf48* overexpression experiments, an siRNA targeting *RRM2-C2orf48* suppressed the proliferation of control BEAS-2B-RC cells and NNK-treated BEAS-2B cells (both $p < 0.05$, Figures 5E and 6D). Plate colony formation assays showed that an siRNA taergeting *RRM2-C2orf48* decreased the colony formation rate when compared with the corresponding control cells (both $p < 0.05$, Figures 5F, 5G, 6E, and 6F). Furthermore, *RRM2-C2orf48* knockdown markedly increased apoptosis ($p < 0.05$, Figures 5H, 5I, 6G, and 6H). Meanwhile, *RRM2-C2orf48* deficiency induced cell-cycle arrest in control BEAS-2B-RC cells and NNK-treated BEAS-2B cells ($p < 0.05$, Figures 5J, 5K, 6I, and 6J). Collectively, these results suggest that *RRM2-C2orf48* regulates cell growth and may be involved in lung cancer development.

RRM2-C2orf48 regulated cell growth via *miR-219a-2-3p*

To explore the downstream mechanism of *RRM2-C2orf48*, transcriptome sequencing analysis was performed for BEAS-2B-V and BEAS-2B-RC cells. Based on the cut-off criteria of a p value <0.05 and a fold change >1.5, 49 upregulated microRNAs (miRNAs) and 77 downregulated miRNAs were detected (Figure 7A).

First, the 10 miRNAs with the greatest fold change in expression level were validated by qRT-PCR. The qRT-PCR results showed that *miR-219a-2-3p* levels significantly increased when *RRM2-C2orf48* was overexpressed via a lentivirus vector and its expression levels decreased after silencing *RRM2-C2orf48* via siRNA ($p < 0.05$, Figures 7B and 7C). However, *miR-9-3p* and *miR-323a-3p* were not affected ($p > 0.05$, Figure 7B), and the remaining miRNAs with low expression levels could not be detected. Next, two databases (miRDB and TargetScan) were used for *miR-219a-2-3p* target gene prediction, and the predicted target genes were used in gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. *miR-219a-2-3p* was found to have 362 predicted target genes via the miRDB database and 198 predicted target genes via the TargetScan database. Thus, 484 predicted target genes were used for the KEGG pathway and GO enrichment analyses (Figure 7D). The KEGG and

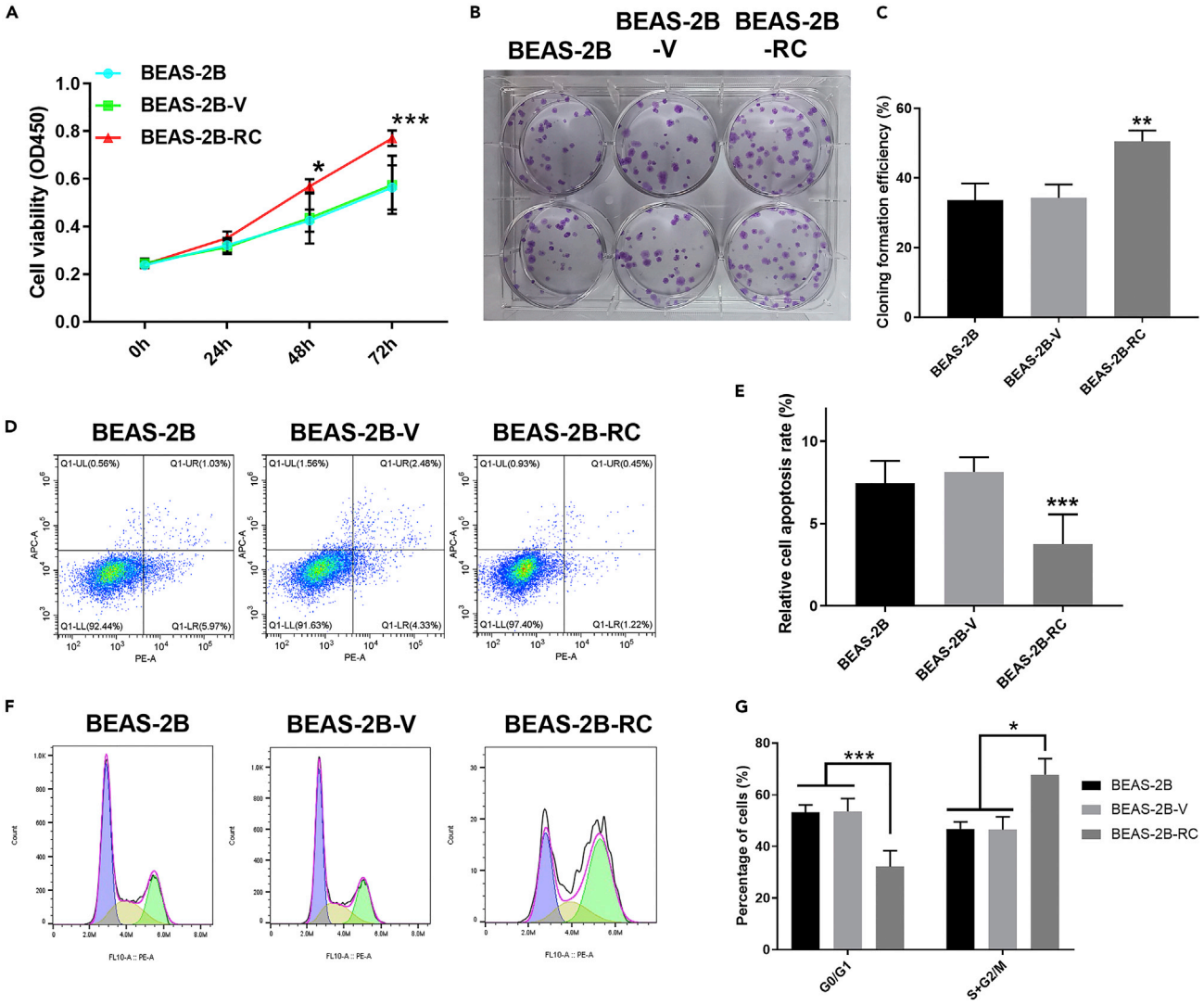


Figure 4. Effect of overexpressing *RRM2-C2orf48* in BEAS-2B cells

(A) CCK-8 assay.

(B) Representive image of plate colony formation assay.

(C) Quantification of plate colony formation assay results.

(D) Representive image of apoptosis analysis via flow cytometry.

(E) Quantification of apoptosis.

(F) Representive image of cell cycle analysis via flow cytometry.

(G) Quantification of cell cycle analysis. Data are represented as mean ± SEM from at least three independent experiments. One-way ANOVA followed by Dunnett's t test were used for data analysis. *p < 0.05, * *p < 0.01, ***p < 0.001, ****p < 0.0001.

GO enrichment analyses showed that the oncogenic effects of *miR-219a-2-3p* may, at least in part, be mediated by the regulation of the cancer-related pathways, p53 signaling pathway and proteoglycans in cancer (Figure 7E), and the biological process, cell growth (Figure 7F). These results further suggest that *RRM2-C2orf48* may be involved in cell growth by regulating *miR-219a-2-3p*.

DISCUSSION

In this study, we identified potential chimeric RNAs by an RNA-sequencing analysis of BEAS-2B and 2B-NNK cells. To reduce the false positive rate of chimeric RNA identification, both SOAPfuse and INTEGRATE were used to identify chimeric RNAs. The advantage of SOAPfuse is that it has a higher

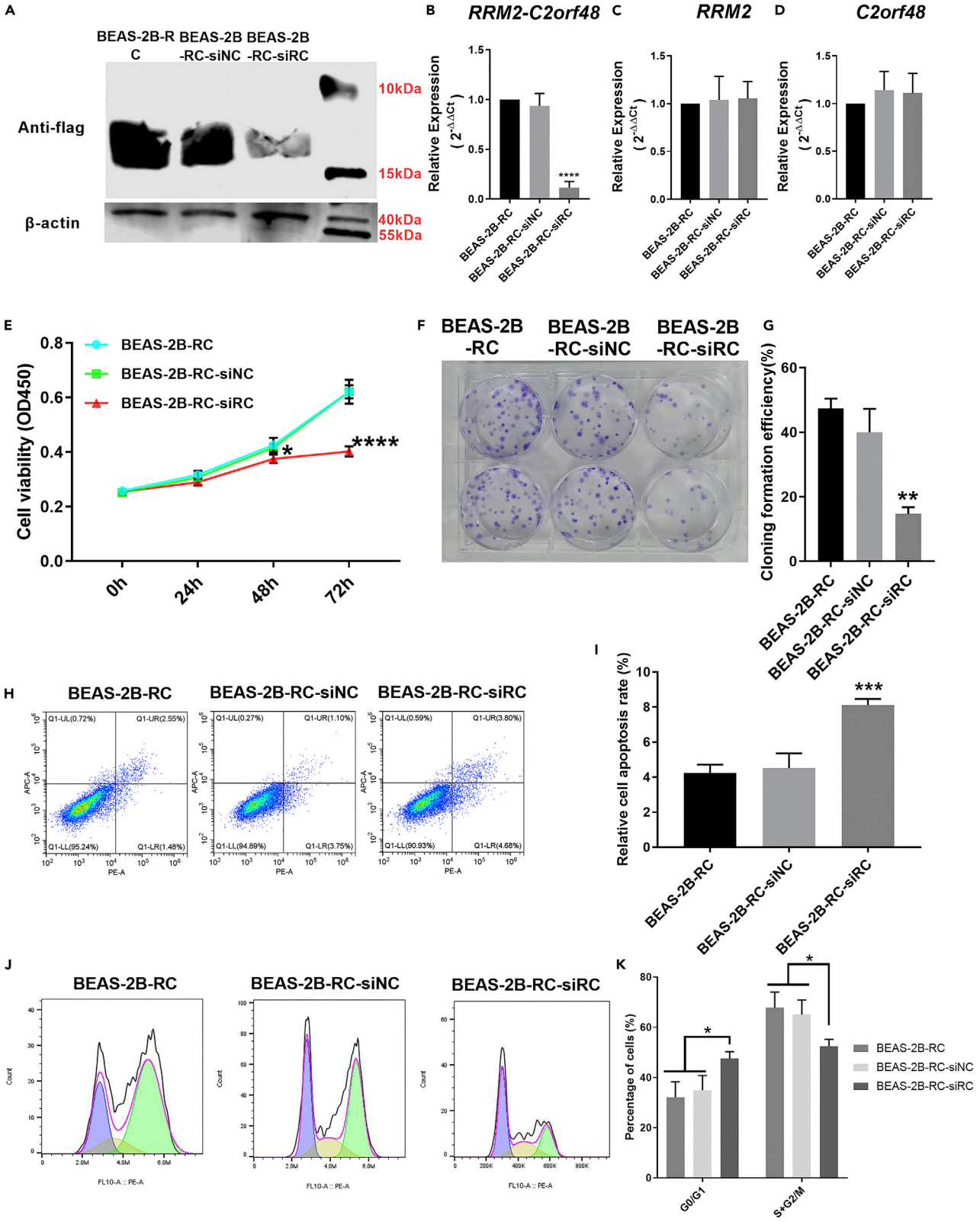

A
BEAS-2B-RC
BEAS-2B-RC-siNC
BEAS-2B-RC-siRC
Anti-flag
β-actin
10kDa
15kDa
40kDa
55kDa
B
RRM2-C2orf48
C
RRM2
D
C2orf48
Relative Expression
BEAS-2B-RC
BEAS-2B-RC-siNC
BEAS-2B-RC-siRC
E
Cell viability (OD450)
0h
24h
48h
72h
F
G
Cloning formation efficiency(%)
H
Q1-UL(0.72%)
Q1-UR(2.55%)
Q1-LL(95.24%)
Q1-LR(1.48%)
Q1-UL(0.27%)
Q1-UR(1.10%)
Q1-LL(94.89%)
Q1-LR(3.75%)
Q1-UL(0.59%)
Q1-UR(3.80%)
Q1-LL(90.93%)
Q1-LR(4.68%)
APC-A
PE-A
I
Relative cell apoptosis rate (%)
J
Count
FL10-A :: PE-A
K
Percentage of cells (%)
G0/G1
S+G2/M

Figure 5. Effect of silencing *RRM2-C2orf48* in BEAS-2B cells overexpressing *RRM2-C2orf48*

(A–D) Western blotting analysis. Quantification of the expression levels of (B) *RRM2-C2orf48*, (C) *RRM2*, and (D) *C2orf48*.

(E) CCK-8 assay results.

(F) Representive image of plate colony formation assay.

(G) Quantification of plate colony formation assay.

(H) Representive image of apoptosis analysis via flow cytometry.

(I) Quantification of apoptosis.

(J) Representive image of cell cycle analysis via flow cytometry.

(K) Quantification of cell cycle analysis. Data are represented as mean ± SEM from at least three independent experiments. One-way ANOVA followed by Dunnett's t test were used for data analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

detection efficiency and consumes fewer computing resources,[35] whereas INTEGRATE has the advantage of greater accuracy in fusion RNA identification.[36]

NNK is viewed as the main carcinogen in tobacco smoke as it induces lung cancer through various mechanisms, such as producing DNA adducts and inducing DNA mutations, DNA methylation, and immunosuppression.[37] Recently, chimeric RNAs were identified as novel lung cancer biomarkers, but their role in mediating the effects of smoking has not been studied. We addressed the question of whether smoking regulates chimeric RNAs and whether chimeric RNAs play a role in the process of smoking-induced lung cancer.

Chimeric RNAs were discovered many years ago, but the majority of studies in the field have focused on RNA chimeras produced from fusion genes and less attention has been given to those that are produced from *cis*-SAGe.[38] Furthermore, there is no customized software for specifically identifying *cis*-SAGe chimeric RNAs or *trans*-spliced chimeric RNAs. Thus, we used a specific screening method based on the characteristics of *cis*-SAGe chimeric RNAs and focused on chimeric RNAs derived from *cis*-SAGe to study their role in smoking-induced lung cancer. We identified three chimeric RNAs, *RRM2-C2orf48*, *CTBS-GNG5*, and *SLC35A3-HIAT1*, and confirmed them by RT-PCR and Sanger sequencing. *RRM2-C2orf48* and *CTBS-GNG5* were found to be produced from *cis*-SAGe. Although *SLC35A3-HIAT1* could not be experimentally validated, it had similar characteristics to *cis*-SAGe chimeric RNAs, indicating that the screening method has great potential for identifying *cis*-SAGe chimeric RNAs.

High expression levels of *RRM2-C2orf48* have been detected in nasopharyngeal carcinoma and colorectal cancer cells.[31,32] Consistently, *RRM2-C2orf48* was found to be highly expressed in 2B-NNK cells and other lung cancer cell lines, indicating that *RRM2-C2orf48* is widely expressed in nasopharyngeal, colorectal, and lung cancer cells. Moreover, our analysis demonstrated that *RRM2-C2orf48* was enriched in lung cancer tissues, whereas it was expressed at low levels in noncancerous lung tissue. Meanwhile, the *RRM2-C2orf48* expression level was found to be associated with clinical staging, N staging, M staging, and smoking status, based on a large subset of patients with lung cancer. Therefore, we confirmed the presence of *RRM2-C2orf48* in lung cancer cells and tissues. Moreover, based on the results presented above, we propose that *RRM2-C2orf48* may be the critical functional molecule involved in smoking-induced lung cancer. However, the relationship between the tobacco-specific carcinogen NNK and *RRM2-C2orf48* requires further investigation.

It is widely known that tobacco smoking is one of the most important carcinogenic exposures, and continuing smokers experience up to a 25% lifetime risk of developing a smoking-related cancer, particularly lung cancer.[39] We investigated whether the tobacco-specific carcinogen NNK regulated *RRM2-C2orf48* and the role of *RRM2-C2orf48* in NNK-induced lung cancer. We observed that both acute and chronic NNK exposure increased the expression level of *RRM2-C2orf48*, while overexpressing *RRM2-C2orf48* may have accelerated the processes of the NNK-induced malignant transformation of the noncancerous immortalized bronchial epithelial cell line BEAS-2B, potentially explaining the reason for the abnormally high expression level of *RRM2-C2orf48* in 2B-NNK cells compared with BEAS-2B cells. To the best of our knowledge, our study is the first to report a link between smoking-induced lung cancer and chimeric RNAs and sheds light on a novel function of chimeric RNAs in the etiology of lung cancer.

The most important characteristics of malignant tumors are uncontrolled cell proliferation and infinite tumor growth.[40] Previous studies have shown that NNK regulates numerous oncogenes to enhance cell proliferation, migration, and invasion and thus induce lung cancer.[41] Combined with our data, it is reasonable

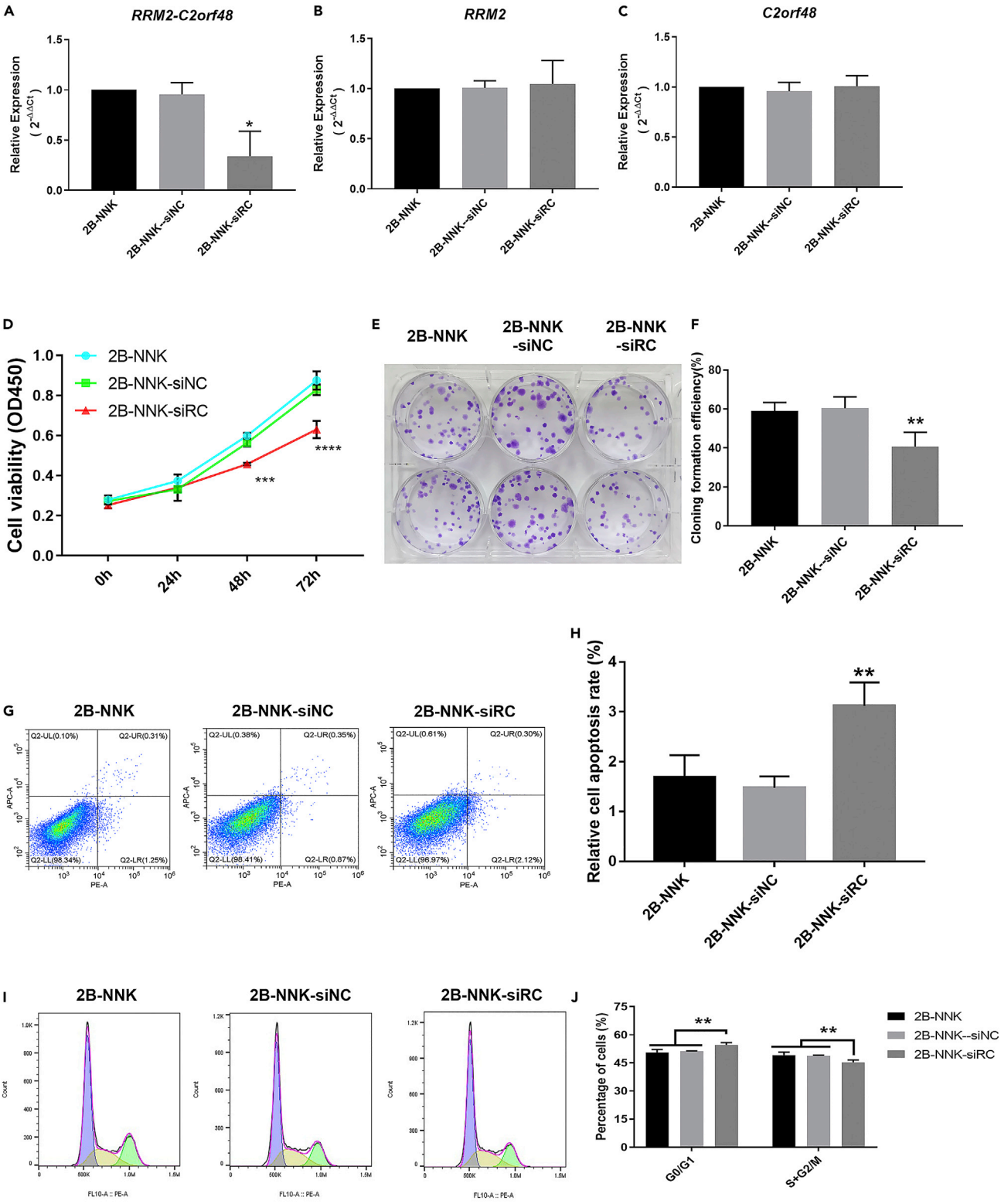

A
RRM2-C2orf48
Relative Expression (2^-ΔΔCt)
2B-NNK
2B-NNK-siNC
2B-NNK-siRC
B
RRM2
C
C2orf48
D
Cell viability (OD450)
0h
24h
48h
72h
E
F
Cloning formation efficiency(%)
G
Q2-UL(0.10%)
Q2-UR(0.31%)
Q2-LL(98.34%)
Q2-LR(1.25%)
Q2-UL(0.38%)
Q2-UR(0.35%)
Q2-LL(98.41%)
Q2-LR(0.87%)
Q2-UL(0.61%)
Q2-UR(0.30%)
Q2-LL(96.97%)
Q2-LR(2.12%)
APC-A
PE-A
H
Relative cell apoptosis rate (%)
I
Count
FL10-A :: PE-A
J
Percentage of cells (%)
G0/G1
S+G2/M
2B-NNK
2B-NNK--siNC
2B-NNK-siRC

Figure 6. Effect of silencing *RRM2-C2orf48* in 2B-NNK cells

(A–C) Quantification of the expression level of (A) *RRM2-C2orf48*, (B) *RRM2*, and (C) *C2orf48*.

(D) CCK-8 assay analysis.

(E) Representive image of plate colony formation assay.

(F) Quantification of plate colony formation assay.

(G) Representive image of apoptosis analysis via flow cytometry.

(H) Quantification of apoptosis.

(I) Representive image of cell cycle analysis via flow cytometry.

(J) Quantification of cell cycle analysis. Data are represented as mean ± SEM from at least three independent experiments. One-way ANOVA followed by Dunnett's t test were used for data analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

to speculate that NNK regulates the chimeric RNA *RRM2-C2orf48* to induce lung cancer. *cis*-SAGe chimeric RNAs with high expression levels in cancer usually play a role in cell growth. For example, silencing *D2HGDH-GAL3ST2* results in a marked decrease in cell growth,[30] *TSNAX-DISC1* promotes cell cycle progression,[42] and *DUS4L-BCAP29* regulates gastric cancer cell proliferation.[43] *RRM2-C2orf48* has been shown to promote colorectal cancer cell growth and development, but it has not been studied in lung cancer. We show that overexpressing *RRM2-C2orf48* enhanced BEAS-2B cell proliferation, whereas silencing *RRM2-C2orf48* resulted in the suppression of BEAS-2B-RC and 2B-NNK cell proliferation. These results further suggest that *RRM2-C2orf48* enhances cell proliferation and thus may further accelerate the development of lung cancer.

Until now, the mechanism responsible for the regulation of chimeric RNAs has remained unknown, but it has usually been correlated with the regulatory mechanism of the parental genes because they are homologous sequences.[33] However, we found that the expression level of *RRM2-C2orf48* was not correlated with the expression levels of *RRM2* or *C2orf48* when *RRM2-C2orf48* was overexpressed or silenced. These findings suggest that *RRM2-C2orf48* has a unique regulatory pathway different from its parental genes *RRM2* and *C2orf48*. Previous studies have demonstrated that the products of chimeric RNAs may be regulated by noncoding RNAs such as circRNAs and miRNAs and that some chimeric RNAs have functions similar to lncRNAs.[29,44,45] Therefore, there may be a close relationship between noncoding RNAs and chimeric RNAs, but further investigation is required to confirm this. In this study, we performed an RNA-sequencing analysis of normal bronchial epithelial BEAS-2B cells with *RRM2-C2orf48* overexpression. We found that *miR-219a-2-3p* was upregulated when *RRM2-C2orf48* was overexpressed and was downregulated when *RRM2-C2orf48* was silenced. In addition, *miR-219a-2-3p* has been found to regulate the growth of pituitary adenoma cells and thyroid cancer cells.[46,47] Bioinformatic analysis was used to predict the target genes of *miR-219a-2-3p*, and they were found to be mainly enriched in the "cell growth" pathway. Together, these results suggested that the *RRM2-C2orf48*/*miR-219a-2-3p* pathway may play an important role in the regulation of cell growth.

Previous studies have found that there is a certain connection between chimeric RNA and noncoding RNA. For example, circRNA can regulate the production of chimeric protein formed by the translation of chimeric RNA to play the function of oncogene (Panni et al. 2020). Fujun Qin et al. found that chimeric RNA *SLC45A3-ELK4* regulates cell growth in tumor cells through its transcript rather than its translated protein,[29] which suggests that chimeric RNA may not only have the general properties of protein-coding genes but also have functional properties similar to lncRNA; thus we speculate that fusion gene may affect miRNA. According to the supposed, we preliminarily explored the relationship between chimeric RNA and miRNA and found that *miR-219a-2-3p* was associated with cell growth pathways which may interact with *RRM2-C2orf48* that causes malignant transformation of cells together probably.

In summary, our results demonstrated that NNK induces lung cancer presumably through the chimeric RNA, *RRM2-C2orf48*, which regulates cell growth. Moreover, *miR-219a-2-3p* may be a target of *RRM2-C2orf48* in this process. These results suggest that the chimeric RNA *RRM2-C2orf48* may be a therapeutic target for lung cancer.

Limitations of the study

Because researchers found that the chimeric RNA is also important for normal cell growth, our results only suggested that *RRM2-C2orf48* plays an oncogenic role in the process of NNK-induced lung cancer, which means it may not be an oncogene in other carcinogenic process. In this study, we found that *miR-219a-2-3p* may be a target of *RRM2-C2orf48*. However, we did not further demonstrate their relationship, whether direct or indirect. Further study needs to prove the regulation mechanism of *RRM2-C2orf48* to *miR-219a-2-3p*.

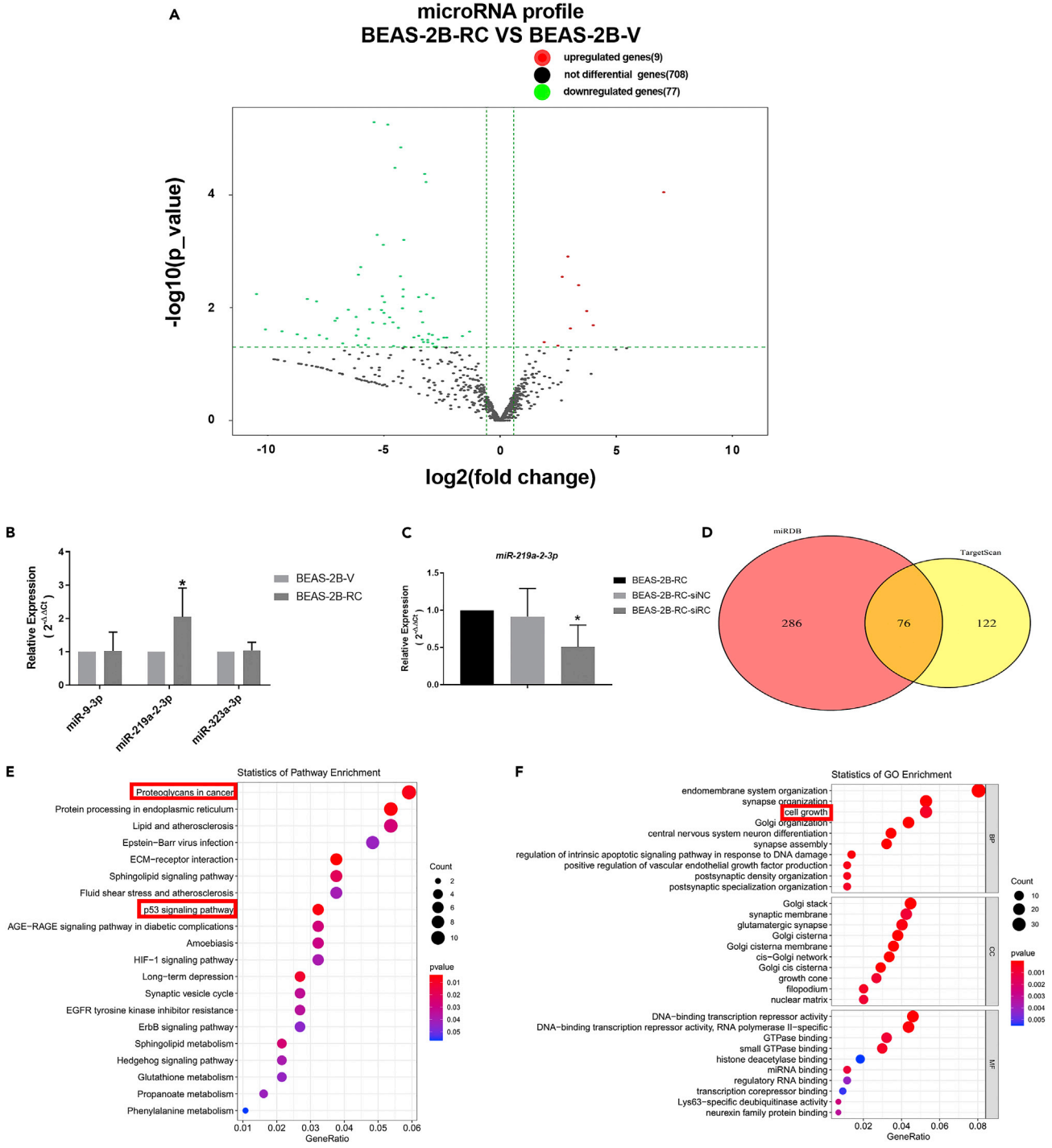


Figure 7. Regulatory effect of *RRM2-C2orf48* on microRNAs

(A) Volcano plot of BEAS-2B-RC cells compared with BEAS-2B-V cells.

(B) Quantification of miRNA expression level in cells overexpressing *RRM2-C2orf48*.

(C) Quantification of *miR-219a-2-3p* expression level after silencing *RRM2-C2orf48*.

(D) The target genes of *miR-219a-2-3p* used for the KEGG pathway and GO enrichment analysis.

(E and F) KEGG pathway enrichment analysis and (F) GO enrichment analysis of predicted target genes of *miR-219a-2-3p*. Data are represented as mean ± SEM from at least three independent experiments. Student's *t* test and one-way ANOVA followed by Dunnett's t test were used for data analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant Number: 81773385 to Y.Q.) and Natural Science Foundation of Guangdong Province (Grant Number: 2019A1515011298 to Y.Q.).

AUTHOR CONTRIBUTIONS

Jiazhen Zhou and Qiaoyuan Yang performed the study scheme and experimental design. Jiazhen Zhou and Xinchao Guan conducted experimental operation, data collection, and statistical analysis. Qiaoyuan Yang, Jiazhen Zhou, Xinchao Guan, Jiaxin Zhou, and Rui Xiong performed the complete writing, revision, and review of the paper. Enwu Xu provided technical and material support. All authors read and approved the final paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 2, 2022
Revised: October 24, 2022
Accepted: November 29, 2022
Published: January 20, 2023

REFERENCES

1. (2018). Cancer Fact Sheets, Lung Cancer. https://gco.iarc.fr/today/fact-sheets-cancers.

2. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *71*, 209–249. https://doi.org/10.3322/caac.21660.

3. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *68*, 394–424. https://doi.org/10.3322/caac.21492.

4. Bade, B.C., and Dela Cruz, C.S. (2020). Lung cancer 2020: epidemiology, etiology, and prevention. Clin. Chest Med. *41*, 1–24. https://doi.org/10.1016/j.ccm.2019.10.001.

5. Malhotra, J., Malvezzi, M., Negri, E., La Vecchia, C., and Boffetta, P. (2016). Risk factors for lung cancer worldwide. Eur. Respir. J. *48*, 889–902. https://doi.org/10.1183/13993003.00359-2016.

6. Rodgman, A., and Perfetti, T.A. The Chemical Components of Tobacco and

Tobacco Smoke, Second Edition (Taylor and Francis)

7. Hoffmann, D., and Hecht, S.S. (1985). Nicotine-derived N-nitrosamines and tobacco-related cancer: current status and future directions. Cancer Res. *45*, 935–944.

8. Park, J.E., Jang, Y.L., and Jang, C.Y. (2017). The tobacco carcinogen NNK disturbs mitotic chromosome alignment by interrupting p53 targeting to the centrosome. Toxicol. Lett. *281*, 110–118. https://doi.org/10.1016/j.toxlet.2017.09.019.

9. Boo, H.J., Min, H.Y., Jang, H.J., Yun, H.J., Smith, J.K., Jin, Q., Lee, H.J., Liu, D., Kweon, H.S., Behrens, C., et al. (2016). The tobacco-specific carcinogen-operated calcium channel promotes lung tumorigenesis via IGF2 exocytosis in lung epithelial cells. Nat. Commun. *7*, 12961. https://doi.org/10.1038/ncomms12961.

10. Kovi, R.C., Johnson, C.S., Balbo, S., Hecht, S.S., and O'Sullivan, M.G. (2018). Metastasis to the F344 rat pancreas from lung cancer induced by 4-(Methylnitrosamino)- 1-(3-pyridyl)-1-butanone and enantiomers of its metabolite 4-(Methylnitrosamino)-1-(3-pyridyl)- 1-butanol, constituents of tobacco products. Toxicol. Pathol. *46*, 184–192. https://doi.org/10.1177/0192623317751573.

11. Wang, Y., Shi, L., Li, J., Wang, H., and Yang, H. (2020). Involvement of twist in NNK exposure-promoted lung cancer cell migration and invasion. Toxicol. Vitro *63*, 104740. https://doi.org/10.1016/j.tiv.2019.104740.

12. Yalcin, E., and de la Monte, S. (2016). Tobacco nitrosamines as culprits in disease: mechanisms reviewed. J. Physiol. Biochem. *72*, 107–120. https://doi.org/10.1007/s13105-016-0465-9.

13. Wu, H., Singh, S., Shi, X., Xie, Z., Lin, E., Li, X., and Li, H. (2020). Functional heritage: the evolution of chimeric RNA into a gene. RNA Biol. *17*, 125–134. https://doi.org/10.1080/15476286.2019.1670038.

14. Chwalenia, K., Facemire, L., and Li, H. (2017). Chimeric RNAs in cancer and normal physiology. Wiley Interdiscip. Rev. RNA *8*. https://doi.org/10.1002/wrna.1427.

15. Guo, M.B., Xiao, Z.D., Dai, Z.M., Zhu, L., Lei, H., Diao, L.T., and Xiong, Y.Y. (2020). The landscape of long noncoding RNA-involved and tumor-specific fusions across various cancers. Nucleic Acids Res. *48*, 12618–12631. https://doi.org/10.1093/nar/gkaa1119.

16. Jabbour, E., and Kantarjian, H. (2018). Chronic myeloid leukemia: 2018 update on diagnosis, therapy and monitoring. Am. J. Hematol. *93*, 442–459. https://doi.org/10.1002/ajh.25011.

17. Murayama, T., Nakaoku, T., Enari, M., Nishimura, T., Tominaga, K., Nakata, A., Tojo, A., Sugano, S., Kohno, T., and Gotoh, N. (2016). Oncogenic fusion gene CD74-NRG1 confers cancer stem cell-like properties in lung cancer through a IGF2 autocrine/paracrine circuit. Cancer Res. *76*, 974–983. https://doi.org/10.1158/0008-5472.Can-15-2135.

18. Xia, D., Le, L.P., Iafrate, A.J., and Lennerz, J. (2017). KIF13B-NRG1 gene fusion and KRAS amplification in a case of natural progression of lung cancer. Int. J. Surg. Pathol. *25*, 238–240. https://doi.org/10.1177/1066896917693092.

19. Chwalenia, K., Qin, F.J., Singh, S., and Li, H. (2019). A cell-based splicing reporter system to identify regulators of cis-splicing between adjacent genes. Nucleic Acids Res. *47*, e24. https://doi.org/10.1093/nar/gky1288.

20. Lee, M.S., Kim, R.N., I, H., Oh, D.Y., Song, J.Y., Noh, K.W., Kim, Y.J., Yang, J.W., Lira, M.E., Lee, C.H., et al. (2016). Identification of a novel partner gene, KIAA1217, fused to RET: functional characterization and inhibitor sensitivity of two isoforms in lung adenocarcinoma. Oncotarget *7*, 36101–36114. https://doi.org/10.18632/oncotarget.9137.

21. Pekova, B., Sykorova, V., Mastnikova, K., Vaclavikova, E., Moravcova, J., Vlcek, P., Lastuvka, P., Taudy, M., Katra, R., Bavor, P., et al. (2021). NTRK fusion genes in thyroid carcinomas: clinicopathological characteristics and their impacts on prognosis. Cancers *13*. https://doi.org/10.3390/cancers13081932.

22. Chen, C., Haddox, S., Tang, Y., Qin, F., and Li, H. (2021). Landscape of chimeric RNAs in non-cancerous cells. Genes *12*. https://doi.org/10.3390/genes12040466.

23. Qin, F., Song, Y., Zhang, Y., Facemire, L., Frierson, H., and Li, H. (2016). Role of CTCF in regulating SLC45A3-ELK4 chimeric RNA. PLoS One *11*, e0150382. https://doi.org/10.1371/journal.pone.0150382.

24. Jia, Y., Xie, Z., and Li, H. (2016). Intergenically spliced chimeric RNAs in cancer. Trends Cancer *2*, 475–484. https://doi.org/10.1016/j.trecan.2016.07.006.

25. Wu, H., Li, X., and Li, H. (2019). Gene fusions and chimeric RNAs, and their implications in cancer. Genes Dis *6*, 385–390. https://doi.org/10.1016/j.gendis.2019.08.002.

26. Plebani, R., Oliver, G.R., Trerotola, M., Guerra, E., Cantanelli, P., Apicella, L., Emerson, A., Albiero, A., Harkin, P.D., Kennedy, R.D., and Alberti, S. (2012). Long-range transcriptome sequencing reveals cancer cell growth regulatory chimeric mRNA. Neoplasia *14*, 1087–1096. https://doi.org/10.1593/neo.121342.

27. Zhu, D.J., Singh, S., Chen, X., Zheng, Z.S., Huan, J., Lin, T.X., and Li, H. (2019). The landscape of chimeric RNAs in bladder urothelial carcinoma (vol 110, pg 50, 2019). Int. J. Biochem. Cell Biol. *111*, 11. https://doi.org/10.1016/j.biocel.2019.03.011.

28. Zhang, Y.M., Gong, M., Yuan, H.L., Park, H.G., Frierson, H.F., and Li, H. (2012). Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. Cancer Discov. *2*, 598–607. https://doi.org/10.1158/2159-8290.Cd-12-0042.

29. Qin, F.J., Zhang, Y.M., Liu, J., and Li, H. (2017). SLC45A3-ELK4 functions as a long non-coding chimeric RNA. Cancer Lett. *404*, 53–61. https://doi.org/10.1016/j.canlet.2017.07.007.

30. Qin, F., Song, Z., Chang, M., Song, Y., Frierson, H., and Li, H. (2016). Recurrent cis-SAGe chimeric RNA, D2HGDH-GAL3ST2, in prostate cancer. Cancer Lett. *380*, 39–46. https://doi.org/10.1016/j.canlet.2016.06.013.

31. Han, P., Chen, R.H., Wang, F., Zeng, J.Y., Yu, S.T., Xu, L.H., Cai, Q., Liang, F.Y., Xia, T.L., Lin, Z.R., et al. (2017). Novel chimeric transcript RRM2-c2orf48 promotes metastasis in nasopharyngeal carcinoma. Cell Death Dis. *8*, e3047. https://doi.org/10.1038/cddis.2017.402.

32. Wu, H., Singh, S., Xie, Z.Q., Li, X.R., and Li, H. (2020). Landscape characterization of chimeric RNAs in colorectal cancer. Cancer Lett. *489*, 56–65. https://doi.org/10.1016/j.canlet.2020.05.037.

33. Shi, X., Singh, S., Lin, E., and Li, H. (2021). Chimeric RNAs in cancer. Adv. Clin. Chem. *100*, 1–35. https://doi.org/10.1016/bs.acc.2020.04.001.

34. Babiceanu, M., Qin, F., Xie, Z., Jia, Y., Lopez, K., Janus, N., Facemire, L., Kumar, S., Pang, Y., Qi, Y., et al. (2016). Recurrent chimeric fusion RNAs in non-cancer tissues and cells. Nucleic Acids Res. *44*, 2859–2872. https://doi.org/10.1093/nar/gkw032.

35. Jia, W., Qiu, K., He, M., Song, P., Zhou, Q., Zhou, F., Yu, Y., Zhu, D., Nickerson, M.L., Wan, S., et al. (2013). SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. Genome Biol. *14*, R12. https://doi.org/10.1186/gb-2013-14-2-r12.

36. Zhang, J., White, N.M., Schmidt, H.K., Fulton, R.S., Tomlinson, C., Warren, W.C., Wilson, R.K., and Maher, C.A. (2016). INTEGRATE: gene fusion discovery using whole genome and transcriptome data. Genome Res. *26*, 108–118. https://doi.org/10.1101/gr.186114.114.

37. Ge, G.Z., Xu, T.R., and Chen, C. (2015). Tobacco carcinogen NNK-induced lung cancer animal models and associated carcinogenic mechanisms. Acta Biochim. Biophys. Sin. *47*, 477–487. https://doi.org/10.1093/abbs/gmv041.

38. Zhou, J.H., Liao, J.S., Zheng, X., and Shen, H. (2012). Chimeric RNAs as potential biomarkers for tumor diagnosis. BMB Rep *45*, 133–140. https://doi.org/10.5483/BMBRep.2012.45.3.133.

39. Fasanelli, F., Baglietto, L., Ponzi, E., Guida, F., Campanella, G., Johansson, M., Grankvist, K., Johansson, M., Assumma, M.B., Naccarati, A., et al. (2015). Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. Nat. Commun. *6*, 10192. https://doi.org/10.1038/ncomms10192.

40. Wang, J., Tang, C., Yang, C., Zheng, Q., and Hou, Y.C. (2019). Tropomyosin-1 functions as a tumor suppressor with respect to cell proliferation, angiogenesis and metastasis in renal cell carcinoma. J. Cancer *10*, 2220–2228. https://doi.org/10.7150/jca.28261.

41. Li, M.Y., Liu, L.Z., Li, W.D., Ng, C.S.H., Liu, Y., Kong, A.W.Y., Zhao, Z.L., Wang, S.S., Qi, H.L., Jia, H., et al. (2019). Ambient fine particulate matter inhibits 15-lipoxygenases to promote lung carcinogenesis. J. Exp. Clin. Cancer Res. *38*, 359. https://doi.org/10.1186/s13046-019-1380-z.

42. Li, N., Zheng, J., Li, H., Deng, J., Hu, M., Wu, H., Li, W., Li, F., Lan, X., Lu, J., and Zhou, Y. (2014). Identification of chimeric TSNAX-DISC1 resulting from intergenic splicing in endometrial carcinoma through high-throughput RNA sequencing. Carcinogenesis *35*, 2687–2697. https://doi.org/10.1093/carcin/bgu201.

43. Kim, H.P., Cho, G.A., Han, S.W., Shin, J.Y., Jeong, E.G., Song, S.H., Lee, W.C., Lee, K.H., Bang, D., Seo, J.S., et al. (2014). Novel fusion transcripts in human gastric cancer revealed by transcriptome analysis. Oncogene *33*, 5434–5441. https://doi.org/10.1038/onc.2013.490.

44. Huang, W., Fang, K., Chen, T.Q., Zeng, Z.C., Sun, Y.M., Han, C., Sun, L.Y., Chen, Z.H., Yang, Q.Q., Pan, Q., et al. (2019). circRNA circAF4 functions as an oncogene to regulate MLL-AF4 fusion protein expression and inhibit MLL leukemia progression. J. Hematol. Oncol. *12*, 103. https://doi.org/10.1186/s13045-019-0800-z.

45. Xie, Z., Tang, Y., Su, X., Cao, J., Zhang, Y., and Li, H. (2019). PAX3-FOXO1 escapes miR-495 regulation during muscle differentiation. RNA Biol. *16*, 144–153. https://doi.org/10.1080/15476286.2018.1564464.

46. Wang, Y., Zhao, J., Zhang, C., Wang, P., Huang, C., and Peng, H. (2020). MiR-219a-2-3p suppresses cell proliferation and promotes apoptosis by targeting MDM2/p53 in pituitary adenomas cells. Biosci. Biotechnol. Biochem. *84*, 911–918. https://doi.org/10.1080/09168451.2020.1715780.

47. Yang, C., Zhang, S., Chang, X., Huang, Y., Cui, D., and Liu, Z. (2021). MicroRNA-219a-2-3p modulates the proliferation of thyroid cancer cells via the HPSE/cyclin D1 pathway. Exp. Ther. Med. *21*, 659. https://doi.org/10.3892/etm.2021.10091.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-flag antibody	Abcam	Cat#ab205606
Mouse monoclonal anti-β-actin antibody	Abcam	Cat#ab8226
goat anti-rabbit IgG secondary antibody	LI-COR	RRID AB_2651127
goat anti-mouse IgG secondary antibody	LI-COR	RRID AB_2687825
Bacterial and virus strains		
pMSCV-BamHI-hoI-*RRM2-C2orf48*-flag	GenePharma	Y644
Biological samples		
Lung cancer patient's lung cancer tissue	General Hospital of Southern Theater Command	N/A
Chemicals, peptides, and recombinant proteins		
4-(methyl nitrosamine)-1-(3-pyridinyl)-1-butanone (NNK)	Sigma-Aldrich	Cat#78013
Critical commercial assays		
SuperScript™ IV One-Step RT-PCR System	Invitrogen	Cat#12594100
E.Z.N.A.® Gel Extraction Kit	Omega	Cat#D2501
GoScript™ Reverse Transcription System Kit	Promega	Cat#A5000
Bulge-Loop miRNA qRT-PCR Starter Kit	Ribobio	Cat#C10211-1
Cell Counting Kit-8	Dojindo Laboratories	Cat#341-08001
Annexin V-PE/7-AAD Apoptosis Detection Kit	KeyGen Biotech	Cat#KGA1018
Cell Cycle Detection Kit	KeyGen Biotech	Cat#KGA511
Deposited data		
Raw and analysis data	This paper	GEO: GSE217991
Original western blot images	This paper	Mendeley Data, https://doi.org/10.17632/dsbt8mt5sd.1
Experimental models: Cell lines		
Human: 95D cells	This paper	N/A
Human: A549 cells	ATCC	Cat#bio-69092
Human: H1299 cells	ATCC	Cat#bio-73106
Human: BEAS-2B cells	ATCC	Cat#bio-68091
Human: BEAS-2B-NNK cells	This paper	N/A
Experimental models: Organisms/strains		
Mice: BALB/c nude mice	Guangdong Medical laboratory animal center	N/A
Oligonucleotides		
Primer for *SLC35A3-HIAT1*	This paper	N/A
Primer for *RRM2-C2orf48*	This paper	N/A
Primer for *CTBS-GNG5*	This paper	N/A
Primer for *GAPDH*	This paper	N/A
siRNA argeting sequence: *RRM2-C2orf48* #1: TTAGCAAGGTGCTGGGAGA	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
SOAPfuse	Jia et al.[35]	https://sourceforge.net/projects/soapfuse/
INTEGRATE	Zhang et al.[36]	https://sourceforge.net/projects/integrate/
R script (version 4.1.1)	Bell Laboratories	https://www.r-project.org/
Prism software (version 8.0.2; GraphPad Software Inc)	Dr. Harvey Motulsky	https://www.graphpad.com/demos/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagent and resource may be directed to and will be fulfilled by the lead contact, PhD. Qiaoyuan Yang (qiaoyuan_yang@gzhmu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Single-cell RNA-seq data have been deposited at GEO: GSE217991 and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Original western blot images have been deposited at Mendeley data: https://doi.org/10.17632/dsbt8mt5sd.1 and are publicly available as of the date of publication. The DOI is listed in the key resources table. Microscopy data reported in this paper will be shared by the lead contact upon request. No original code was written for this analysis. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The lung cancer cell lines 95D, A549, H1299 and the human bronchial epithelial cell line BEAS-2B were obtained from American Type Culture Collection (Manassas, VA, USA).

BEAS-2B cells were treated with NNK (Sigma-Aldrich, Saint Louis, MO, USA) at a concentration of 100 mg/L. After 24 h of exposure, the culture medium was discarded, and the cells were washed twice with PBS (PBS) and grown in fresh complete culture medium. When the cells were 80% confluent, the cells were passaged. When the cells were 70% confluent, the above NNK exposure steps were repeated and the cells were passaged to the sixth generation. The NNK-induced malignant transformation of the BEAS-2B cells was detected using the soft-agar clone formation rates. And these cells were designated as 2B-NNK cells.

BEAS-2B and 2B-NNK cells were grown in bronchial epithelial basal medium (BEBM; Clonetics, Basel, Switzerland) supplemented with SingleQuots (Clonetics). 95D and H1299 cells were grown in RPMI-1640 medium (HyClone, Logan, UT, USA) with 10% fetal bovine serum (FBS; Thermo Fisher Scientific, Waltham, MA, USA). A549 cells were grown in F-12K medium (HyClone) supplemented with 10% FBS. All cells were incubated at 37°C in a humidified incubator containing 5% CO2.

METHOD DETAILS

NNK treatment conditions

NNK (Sigma-Aldrich) was dissolved in dimethyl sulfoxide (DMSO) to a concentration of 1 g/mL. For NNK exposure experiments, BEAS-2B cells were cultured in 6-well plates to a final concentration of 0 (0.1% DMSO), 50, 100, 200, or 400 mg/L NNK for 24 h. For the multiple NNK exposure experiments, cells were cultured in 6-well plates in 100 mg/L NNK for 24 h every 4 days. The medium was changed or cells were sub-cultured as necessary.

Patients and tissue samples

Informed consent was obtained from all patients and institutional approval was granted by the Review Board of the General Hospital of Southern Theater Command and the Guangzhou Medical University prior to commencing this study. Lung cancer tissues with paired adjacent non-tumor tissues were obtained from 81 patients who had undergone surgical resection between August 2013 and June 2019 at the General Hospital of Southern Theater Command. Patients were included if they had undergone initial radical resection of lung cancer without adjuvant chemotherapy. All tissues samples were verified by pathological examination. Due to the lack of detected expression of the chimeric RNA *RRM2-C2orf48* in some samples, data from 74 lung cancer patients were finally obtained. The clinicopathological characteristics of the 74 patients, including sex, age, smoking history, daily smoking consumption, histology, tumor-lymph node-metastasis (TNM) stage, tumor size, lymph node stage and distant metastasis status are summarized in Table 1. Patients were classified as a "smoker" (smoking every day for at least 6 months) if they had any of the following statuses: current smoker; ever smoked but ceased smoking <3 years ago or second-hand smoker (patients passively exposed to smoking every week) for at least 6 months. Otherwise, the patients were classified as "non-smokers".

RNA extraction and PCR assays

Total RNA was extracted from cells using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. RNA concentration was determined using a Multiskan Spectrum microplate reader (BioTek, Winooski, VT, USA). A PCR approach was used to verify the presence of the chimeric RNAs *SLC35A3-HIAT1*, *RRM2-C2orf48* and *CTBS-GNG5* using the SuperScript™ IV One-Step RT-PCR System (Invitrogen, Life Technologies, Grand Island, NY, USA). PCR amplification was performed in a total volume 50 μL, consisting of 25 μL of 2× Platinum™ SuperFi™ RT-PCR Master Mix, 2.5 μL of each primer (10 μM), 0.5 μL of SuperScript™ IV RT Mix, template RNA based on the RNA concentration and nuclease-free water. The PCR cycling parameters were as follows: denaturation at 95°C for 2 min and 40 cycles of denaturation at 95°C for 15 s, annealing at 60°C for 1 min and extension at 72°C for 1 min. The PCR primer sequences were listed below:

SLC35A3-HIAT1 Sense 5'-TGATGGAGAACTGGTATCAAAGAA-3'.

Anti-sense 5'-GCTGTCAATAGTCCCCAAGC-3'.

RRM2-C2orf48 Sense 5'-AGCTCATTGGGATGAATTGC-3'.

Anti-sense 5'-TTGGTCTCCATCTCCCTCAC-3'.

CTBS-GNG5 Sense 5'-TGGGATAAAGATCAGCGGGC-3'.

Anti-sense 5'-GACTTTCTGGGGTCTGAAGGG-3'.

GAPDH Sense 5'-ACAGTCAGCCGCATCTTCTT-3'.

Anti-sense 5'-GACTCCGACCTTCACCTTCC-3'.

Gel electrophoresis

PCR products were separated by gel electrophoresis. Twenty microliters of each PCR product were electrophoresed at 100 V for 1 h through a 2% agarose gel in 1× Tris-borate- EDTA buffer. Ten microliters of a 50bp DNA marker ladder were included on both sides of each sample. After electrophoresis, the agarose gels were visualized and photographed using a GelView 6000Pro instrument (Guangzhou Biolight Biotechnology Co., Ltd., Guangzhou, China).

Purification of PCR products and Sanger sequencing

After electrophoresis, agarose gel segments containing the PCR products were harvested under UV light and purified using the E.Z.N.A.® Gel Extraction Kit (Omega, Norcross, GA, USA). After purification and extraction using an activated adsorption column, the PCR products were dissolved in sterile water and incubated at 50°C. The concentration of the purified PCR products was determined using a Multiskan Spectrum

microplate reader (BioTek). Qualified samples were sent for Sanger sequencing (Sangon Biotech, Guangzhou, China).

Reverse transcription and quantitative PCR assays

cDNA was prepared from total RNA by reverse transcription (RT) using a GoScript™ Reverse Transcription System Kit (Promega, Madison, WI, USA). RT was performed using the following steps: an initial denaturation step at 70°C for 5 min, 4°C for 5 min, annealing at 25°C for 5 min, extension at 70°C for 1 h and holding at 4°C. An Applied Biosystems™ 7500 Fast Dx Real-Time PCR Instrument (Thermo Fisher Scientific) together with GoTaq qPCR Master Mix (Promega) were used for quantitative (q) PCR to measure gene expression levels. qPCR was performed with the following cycling parameters: denaturation at 95°C for 2 min and 40 cycles of denaturation at 95°C for 15 s, annealing at 60°C for 1 min and extension at 72°C for 1 min. We used *GAPDH* as an endogenous control. Relative gene expression levels were determined using the $2^{-\Delta\Delta Ct}$ method.

miRNA qRT-PCR was performed using the Bulge-Loop miRNA qRT-PCR Starter Kit (Ribobio, Guangzhou, China) following the manufacturer's instructions. We used *U6* as an endogenous control for miRNA qRT-PCR. Relative miRNA expression levels were determined using the $2^{-\Delta\Delta Ct}$ method.

Cell viability assays

Cell proliferation was measured using Cell Counting Kit-8 (CCK-8; Dojindo Laboratories, Kumamoto, JPN) following the manufacturer's instructions. Briefly, cells were seeded into 96-well plates with approximately 2 000 cells in each well, and were then incubated for 0, 24, 48 or 72 h. After incubation, the cells were washed with PBS (PBS), 10 μL of the CCK-8 reagent was added to each well and the cells were incubated for another 1 h at 37°C. The absorbance at 450 nm was then measured at various time points using a Multiskan Spectrum microplate reader (BioTek).

Lentivirus vector and transfection

The human *RRM2-C2orf48* chimeric RNA was cloned into the pMSCV vector using BamHI (or BglI) and XhoI restriction sites (GenePharma, Guangzhou, China). The 213 bp *RRM2-C2orf48* sequence was as follows:

5'-ATGGAGTTCCTCACTGAGGCCTTGCCTGTGAAGCTCATTGGGATGAATTGCAC.

TCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGGTTTTAGCAAGGTGC
TGGGAGACCGTGAAGTGCAAAGCAGATGGAGTCCAGGCCCTCGGGGAGACAGCACGCCAGTGAGG
GAGATGGAGACCAATCACCCACCCAGTGTGTGA-3'.

BEAS-2B cell lines were infected with a lentivirus vector containing the human chimeric RNA *RRM2-C2orf48* to form stably transfected cell lines. The empty vector was used as a negative control. The cells were selected and maintained in puromycin (4 μg/mL).

Western blotting

Total proteins were extracted from the cells using radioimmunoprecipitation buffer (Beyotime, Jiangsu, China) on ice for 30 min. The cells were then centrifuged at 4°C for 20 min at 14,000 rpm. Protein concentrations were measured using a bicinchoninic acid protein assay kit (Beyotime). Total proteins were separated using sodium dodecyl sulfate–polyacrylamide gel electrophoresis, transferred to polyvinylidene fluoride membranes (Millipore, Burlington, MA, USA), blocked with buffer supplemented with 5% non-fat milk and 0.1% Tween 20 and then incubated with a primary anti-flag antibody (Abcam, Cambridge, UK) overnight at 4°C. An anti-β-actin antibody (Abcam) was used as a protein loading control. The membranes were then incubated with the secondary antibodies IRDye® 800CW goat anti-rabbit IgG and IRDye® 800CW goat anti-mouse IgG (LI-COR, Lincoln, NE, USA) for 1 h at room temperature in the dark. After the membranes were washed in Tris-buffered saline with 0.1% Tween 20, the labeled protein bands were visualized using an LI-COR Odyssey Infrared Imaging System.

RNAi

A small interfering RNA (siRNA) was used to knock down *RRM2-C2orf48*. The double-stranded custom siRNA molecule was synthesized by Ribobio. Cells were seeded in 6-well plates and siRNA transfection

was performed once the cells reached 50-60% confluency. siRNA transfection was performed using Lipofectamine® 2000 Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Total RNA was extracted after 48 h of transfection and the efficiency of siRNA-based interference was determined by qRT-PCR and western blotting. The siRNA target sequences were as follows:

RRM2-C2orf48 5'-TTAGCAAGGTGCTGGGAGA-3'.

Negative control siRNA 5'-CGTACGCGGAATACTTCGA-3'.

Plate colony formation assay

Cells were harvested, seeded in 6-well plates at 100 cells per well and maintained in BEBM medium containing 5% FBS (replaced every 3 days). Fourteen days later, the cells were fixed with paraformaldehyde and stained with crystal violet. Finally, the number of stained colonies was counted. Colony formation efficiency was determined using the following formula: colony formation efficiency (%) = (number of clones/number of inoculated cells) × 100%.

Soft agarose colony formation assay

Each well of the six-well plates was pre-coated with 1.5 mL of 0.6% agar in complete medium containing 5% FBS. The cells were harvested, suspended in 0.35% agar in complete medium containing 5% FBS and then plated in triplicate on the pretreated 6-well plate (1.5 mL of medium with 5 000 cells per well). The cells were then incubated at 37°C in a humidified incubator containing 5% CO_2 for 4 weeks. The resulting colonies were photographed and colonies larger than 0.05 mm were counted.

Tumor Xenograft assay

BEAS-2B and BEAS-2B-RC cells treated with multiple exposure of NNK at a concentration of 100 mg/L were harvested for tumorigenesis experiments in nude mice. Then 2 500 000 cells were suspended in 150 μL of a 1:1 mixture of Matrigel® Basement Membrane Matrix High Concentration (Corning, NY, U.S.) and complete medium to form a suspension. Twelve healthy BALB/c nude mice (6 males and 6 females) were obtained from Guangdong Medical Laboratory Animal Center (Guangzhou, China) which randomly divided into two groups (n = 6 per group; 3 males and 3 females). The suspension was injected into the right lower groin of each nude mouse. The tumor volume in the left groin was continuously observed and measured from the seventh day after injection of the mixture until the 28th day after injection. The increase in tumor volume was assessed by measuring the tumor volume with the formula: volume (mm^3) = (D^2 × d) × 0.5^2, where D and d were the longest and shortest tumor diameters in the nude mice respectively.

Flow cytometric analysis

Apoptosis was assessed using the Annexin V-PE/7-AAD Apoptosis Detection Kit (KeyGen Biotech, Nanjing, China) according to the manufacturer's instructions. Briefly, cells that had been harvested with trypsin, without EDTA, were washed twice with cold PBS and then resuspended in Binding Buffer. Fifty microliters of the cell suspension (100,000 cells) was then transferred to a flow cytometry tube, and 5 μL of 7-aminoactinomycin was added. After incubation for 5-15 min in the dark, 450 μL of Binding Buffer and 1 μL of annexin V-phycoerythrin were added to each tube. After incubation for 5-15 min in the dark, flow cytometric analysis was performed using a CytoFLEX Cytometer (Beckman-Coulter, Brea, CA, USA).

Each phase of the cell cycle was analyzed by flow cytometry using a Cell Cycle Detection Kit (KeyGen Biotech). Briefly, cells were collected and fixed with 75% ethanol overnight at 4°C. The cells were then washed in PBS and stained with propidium iodide/RNase staining buffer for 30-60 min at room temperature in the dark. Flow cytometric analysis was performed using a CytoFLEX Cytometer (Beckman-Coulter).

Pathway enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using R software (version 4.1.1). The package "org.Hs.eg.db" was used to identify *trans* genes and the package "clusterprofiler" was used to perform enrichment analysis using the GO and KEGG databases. Statistical significance was considered as $p < 0.05$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using Prism software (GraphPad Prism software version 8.0.2). Differences between groups were analyzed by Student's *t* test or one-way ANOVA followed by Dunnett's t test. The results of quantitative data in histograms, line charts and individual data points were presented as mean ± SEM Statistical significance was considered as $p < 0.05$. The meaning of asterisks number were *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Ethics approval

All patients' informed consents were obtained before this study and the experimental protocol was approved by the Ethics Committee for Human Experiments of the General Hospital of Southern Theater Command and Guangzhou medical university (Permit protocol: 2017-532-18).