Abstract. Non-small cell lung cancer (NSCLC) is a leading cause of cancer-associated mortality worldwide. NSCLC includes lung adenocarcinoma and lung squamous cell carcinoma. Tumor metastasis is the major cause of mortality of patients with NSCLC. However, the mechanisms underlying NSCLC metastasis remain largely elusive. The present study focused on exploring the roles of RBAK in NSCLC. The present study demonstrated that RB-associated KRAB zinc finger (RBAK) was upregulated in NSCLC compared with non-tumorous tissues by analyzing Gene Expression Omnibus (GEO) datasets and The Cancer Genome Atlas (TCGA) dataset. High expression of RBAK was associated with poor disease-free survival of patients with NSCLC by analyzing TCGA dataset. Furthermore, an RBAK-mediated protein-protein interaction network was constructed to reveal the potential underlying mechanisms by which RBAK drives NSCLC progression. The authors found that RBAK was involved in regulating a number of transcription factors, including androgen receptor, forkhead box A1, tumor protein 53, and E2F transcription factor 1, 2 and 4, suggesting that RBAK may have a role in regulating gene transcription. GO and KEGG enrichment analyses of the genes co-expressed with RBAK revealed that RBAK is involved in regulating a number of biological functions, including the Wnt signaling pathway, mRNA splicing, protein polyubiquitination, cell-cell adhesion and focal adhesion. Transwell and wound healing assays demonstrated that knockdown of RBAK suppressed NSCLC cell migration and invasion. The present study enhances the current understanding of the important roles of RBAK in NSCLC metastasis and may provide useful information for the development of novel treatment approaches.

Introduction

Non-small cell lung cancer (NSCLC) is a leading cause of cancer-associated mortality worldwide, and the incidence and mortality rates of NSCLC have significantly increased (1,2). NSCLC includes lung adenocarcinoma (LUAD) and lung squamous cell carcinoma. Tumor metastasis is recognized as the major cause of mortality of patients with NSCLC (3). Previous studies have identified a number of genes that are involved in regulating NSCLC metastasis, including twist family bHLH transcription factor 1 (4,5), patched 1 (PTCH1) (6) and transforming growth factor β-induced long non-coding RNA (7). Elucidation of the molecular mechanisms underlying NSCLC metastasis is urgently required, as this will assist with the identification of novel therapeutic targets for NSCLC.

A previous study has reported that RB-associated KRAB zinc finger (RBAK) interacts with tumor suppressor retinoblastoma 1 (8). However, the roles of RBAK in human diseases remain largely elusive. A number of studies have indicated that RBAK is involved in cancer progression (9). In prostate cancer (PCa), RBAK interacts with androgen receptor (AR) and is involved in the regulation of the cell cycle (10). Wan et al (9) revealed that knockdown of RBAK inhibits PCa growth by inducing cell apoptosis. A high expression level of RBAK has been associated with a shorter survival time for patients with PCa following radical prostatectomy (9). However, the prognostic value and molecular functions of RBAK in NSCLC have remained elusive.

The present study identified that RBAK is upregulated in NSCLC. Furthermore, the potential effects of RBAK on NSCLC cell migration and invasion were investigated. The present study may provide useful information that promotes the understanding of the important roles of RBAK in NSCLC metastasis.
Materials and methods

Analysis of public datasets. In the present study, three Gene Expression Omnibus (GEO) datasets, including GSE19188 (11), GSE19804 (12) and GSE18842 (13), were downloaded from the National Cancer for Biotechnology Information database for analysis of RBAK expression in NSCLC (https://www.ncbi.nlm.nih.gov/geo/). The dataset GSE19188 includes 91 NSCLC samples and 65 normal lung samples. Furthermore, GSE19804 contains 60 paired NSCLC and normal samples, and GSE18842 includes 46 NSCLC samples and 45 control samples. The raw data of the mRNA expression profiles were downloaded and analyzed using R language software. Background correction, quartile data normalization and probe summarization were applied to the original data. The limma tool in Bioconductor software (http://www.bioconductor.org/) was used to identify differentially expressed genes (DEGs) between the NSCLC and normal samples. The significance of each DEG was calculated by a classical t-test and was presented as a P-value. In addition, The Cancer Genome Atlas (TCGA) lung adenocarcinoma (LUAD) dataset, with 59 normal and 517 LUAD samples, which included clinical features such as survival time, were downloaded from cBioPortal (http://www.cbioportal.org/).

Bioinformatics analysis. The Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics tool (https://david.ncifcrf.gov/) was used to perform gene ontology (GO) and Kyoto Encyclopedia of genes and genomes (KEGG) pathway enrichment analyses. The present study used Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org/) to construct the PPI network for target genes (minimum required interaction score >0.4). In addition, Cytoscape software version 3.6.0 (http://www.cytoscape.org/download-platforms.html) was used for visualization of the PPI networks.

Cell culture. The human cell lines A549, H1299, 95D and H1975 were purchased from the Cell Bank of the Chinese Academy of Sciences. The NSCLC cell lines were cultured in RPMI 1640 medium containing 10% fetal bovine serum (FBS), 1% sodium pyruvate and 1% penicillin/streptomycin (all Gibco; Thermo Fisher Scientific, Inc.) at 37°C in a humidified atmosphere containing 5% CO₂.

Lentiviral constructs and transfection. The RBAK short hairpin (sh)RNA (5'-CCGGGCTGCGTGATCAATAGAAA CTCGAGTTTTAGATACGCAAGCAGCTTTTT-3') was purchased from Shanghai GeneChem Co., Ltd. Recombinant lentiviral plasmids carrying RBAK shRNA were constructed according to the manufacturer's protocol. The empty recombinant lentiviral plasmids were used as negative controls. Reverse transcription-quantitative (RT-q)PCR was performed to detect the knockdown efficiency of RBAK shRNA.

RT-qPCR. Total RNA was extracted from the cells using TRIzol (Invitrogen; Thermo Fisher Scientific, Inc.). Subsequently, 0.5-1.0 μg RNA was reverse-transcribed to complementary (c)DNA using the RevertAid First Strand cDNA Synthesis kit (Takara Biotechnology Co., Ltd.). qPCR was performed with the iQSYBR Green Supermix using the ABI Prism 7900 platform (Applied Biosystems; Thermo Fisher Scientific, Inc.). The PCR conditions were as follows: Initial denaturation at 95°C for 10 sec, followed by 40 cycles of 95°C for 5 sec, 57°C for 15 sec and 72°C for 20 sec. GAPDH was used as an internal control. The relative expression levels of the target genes were calculated using the 2^(-ΔΔCq) method (14). The following primers were used for qPCR: GAPDH forward, 5'-TGACTCTACACAGCGCACACCA-3' and reverse, 5'-CACCCCTGTTGTGCTAGC AAA-3'; RBAK forward, 5'-TTACAGGGATGTGAAGTT GGA-3' and reverse, 5'-CTCCTTGAGGGGTTTTTGCT-3'.

Wound healing assay. A total of 5,000 A549 cells were seeded into a 6-well plate. A circular hole was generated 16 h after serum starvation of the cells using an Oris™ plate. Serum starvation medium was RPMI 1640 containing 1% sodium pyruvate and 1% penicillin/streptomycin (all Gibco; Thermo Fisher Scientific, Inc.). Images were captured after 0, 24 and 48 h using a microscope (CK40-F200; Olympus).

Transwell assay. Transwell assays were performed using 8-mm pore size Transwell® plates (Corning Inc.), according to the manufacturer's protocol. In brief, 50,000 A549 cells were seeded into the upper chamber of the Transwell plate. The upper wells were filled with serum-free RPMI 1640 and the bottom wells were filled with RPMI 1640 containing 10% FBS. After 72 h of incubation, the number of migrated cells was counted and images were captured. Cells were stained with 0.1% crystal violet for 20 min at room temperature. The invasion assay was performed using the above protocol with the Transwell plates pre-coated with Matrigel® basement membrane matrix (BD Biosciences).

Statistical analysis. Statistical analyses were performed using SPSS software (version 16.0.0; SPSS, Inc., Chicago, IL, USA). All experiments were performed three times. Depending on the comparison, a Student's t-test or Mann-Whitney U-test was performed to assess the statistical significance between two groups. Kaplan Meier and Cox regression analyses were applied to measure disease-free survival (DFS), and log-rank test was conducted to determine the survival difference based on RBAK expression. P<0.05 was considered to indicate a statistically significant difference.

Results

RBAK is overexpressed in NSCLC. The functional roles and expression pattern of RBAK in human cancer, particularly NSCLC, have remained largely elusive. The present study first analyzed RBAK expression in NSCLC using three GEO datasets, namely GSE19188, GSE19804 and GSE18842. As presented in Fig. 1, it was observed that RBAK was significantly overexpressed in the NSCLC samples compared to normal samples (Fig. 1A-C). Subsequently, TCGA LUAD dataset with 59 normal and 517 LUAD samples was analyzed. It was identified that patients with LUAD had a higher RBAK expression level compared with that in the normal controls (Fig. 1D).

High RBAK expression is associated with poor survival in patients with NSCLC. Subsequently, Kaplan-Meier analysis
was performed to investigate the prognostic value of RBAK in NSCLC. The cut-off finder online system (http://molpath.charite.de/cutoff/) was used to identify the cut-off value to stratify all NSCLC samples into RBAK-high and RBAK-low groups. It was revealed that the DFS rate was lower for patients with LUAD or lung squamous cell carcinoma in the RBAK-high expression group than for those in the RBAK-low expression group (Fig. 2). These results suggest that RBAK may serve as a biomarker for NSCLC.

Bioinformatics analysis of RBAK in NSCLC. Considering that the function of RBAK in NSCLC has remained elusive, the present study performed a bioinformatics analysis of its roles and interactions. First, an RBAK-mediated protein-protein interaction network was constructed using the STRING online system (https://string-db.org/cgi/network.pl). A total of 57 single nodes and 576 edges were included in the RBAK-mediated network. Figure 2. High RBAK expression is associated with shorter disease-free survival for patients with LUAD. High RBAK expression is associated with poor disease-free survival in LUAD. RBAK, RB-associated KRAB zinc finger; TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma.
protein-protein interaction network (Fig. 3). It was identified that the network included a number of transcription factors, including AR, forkhead box (FOX)A1, tumor protein (TP)53, and E2F transcription factor 1, 2 and 4, suggesting that RBAK may have a role in regulating gene transcription. Therefore, the present study performed a co-expression analysis for RBAK to reveal its potential roles in NSCLC.

Subsequently, a functional analysis of RBAK in NSCLC was performed based on its co-expressed genes. Gene ontology enrichment analysis demonstrated that RBAK is involved in regulating the Wnt signaling pathway, mRNA splicing, protein polyubiquitination, NF-kB-inducing kinase (NIK)/NF-kB signaling, regulation of mRNA stability, histone H3-K4 methylation, macroautophagy, DNA repair, chromatin remodeling, Hippo signaling, cell-cell adhesion, G2/M transition of mitotic cell cycle and focal adhesion assembly (Fig. 4A). Furthermore, Kyoto Encyclopedia of Genes and Genomes analysis revealed that RBAK is significantly associated with various types of human cancer, including NSCLC, in addition to the cell cycle, thyroid hormone signaling pathway, tight junction and FOXO signaling pathways (Fig. 4B). Metastasis is a key cause of NSCLC-associated mortality. According to the aforementioned analysis, RBAK is involved in regulating cell-cell adhesion and focal adhesion, suggesting that RBAK may affect the progression of metastasis in NSCLC.

Knockdown of RBAK inhibits cell migration in NSCLC. By detecting the endogenous expression of RBAK in the human A549, H1299, 95D and H1975 cell lines, it was identified that RBAK expression is higher in A549 cells. To further validate the effect of RBAK on cell metastasis, the present study performed a loss-of-function assay with NSCLC A549 cells. As presented in Fig. 5, a wound-healing assay was used to investigate the role of RBAK in the regulation of NSCLC cell migration. Images of the area void of cells were

Figure 3. Construction of an RBAK-mediated protein-protein interaction network. An RBAK-mediated protein-protein interaction network in non-small-cell lung cancer was generated, which included 57 single nodes and 576 edges. RBAK is shown in red, genes that were positively co-expressed with RBAK are denoted as green nodes and genes that were negatively co-expressed with RBAK are denoted as blue nodes. RBAK, RB-associated KRAB zinc finger.

Figure 4. Bioinformatics analysis of RBAK in NSCLC. (A) Gene Ontology and (B) Kyoto Encyclopedia of Genes and Genomes analyses were performed to predict potential functions of RBAK in NSCLC using its co-expressing genes. NSCLC, non-small-cell lung cancer; RBAK, RB-associated KRAB zinc finger. NIK, NF-kB-inducing kinase.
obtained at 0, 24 and 48 h. After 24 h, ~25% of the wound area was repaired by migrating cells in the control group, while only 11% of the wound area was repaired by cells in the RBAK-knockdown group (Fig. 5B and C). After 48 h, ~30% of the wound area was repaired by migrating cells in the control group, while only 12% of the wound area was repaired in the RBAK-knockdown group (Fig. 5B and C). The transfection efficiency is presented in Fig. 5A. The knockdown efficiency is presented in Fig. 5A and the transfection efficiency is presented in Fig. S1.

Furthermore, the present study investigated the roles of RBAK in the regulation of NSCLC cell migration by using a Transwell assay. As illustrated in Fig. 6, it was revealed that RBAK knockdown significantly decreased the number of migrating A549 cells compared with that in the negative control group (Fig. 6A and B).

Knockdown of RBAK suppresses the invasive ability of NSCLC cells. In addition, the roles of RBAK in the regulation of NSCLC cell invasion were investigated. At 72 h post-transfection, the ability of A549 cells to invade through Matrigel was assessed. It was identified that knockdown of RBAK significantly inhibited A549 cell invasion. RBAK knockdown decreased the number of invading A549 cells by 88% compared with that in the negative control group (Fig. 7A and B).

Discussion

Tumor metastasis is a major cause of mortality in patients with NSCLC (3). The mechanisms underlying NSCLC metastasis remain to be further investigated. The present study was the first, to the best of our knowledge, to identify RBAK as a
novel regulator of metastasis in NSCLC. RBAK was observed to be overexpressed in NSCLC and high expression was associated with shorter DFS times in patients with NSCLC. Bioinformatics analysis demonstrated that RBAK is involved in regulating various metastasis-associated biological functions. These roles were experimentally validated, and in vitro assays confirmed that knockdown of RBAK suppressed NSCLC cell migration and invasion.

A limited number of studies have indicated that RBAK is involved in regulating cancer progression. Higher expression levels of RBAK have been associated with a shorter survival time for patients with PCa (9). However, to the best of our knowledge, the expression pattern of RBAK in NSCLC has remained largely elusive. In the present study, analysis of public datasets revealed that the expression levels of RBAK in NSCLC were higher compared with those in normal samples. Furthermore, a high RBAK expression level was associated with poorer DFS of NSCLC patients. These results suggest that RBAK may serve as a novel prognostic marker for NSCLC.

RBAK has been identified to interact with AR and promote cell proliferation and the cell cycle in PCa (15). In the present study, a Bioinformatics analysis was performed to investigate the potential roles of RBAK in NSCLC. RBAK-mediated protein-protein interaction networks were constructed, which revealed that RBAK interacts with various key regulators in NSCLC, including AKT1 (16), AKT2 (16), AKT3 (17), E2F1 (18), E2F2 (19), E2F3 (20), RB1 (21), cyclin D 1 (21), cyclin-dependent kinase 4 (22), epidermal growth factor receptor (23) and TP53 (24). The present study indicated that RBAK may be involved in regulating NSCLC progression and metastasis by influencing cell-cell adhesion and focal adhesion. Tumor metastasis is the major cause of mortality for patients with NSCLC. In previous studies, a number of metastasis regulators have been identified in NSCLC. For instance, Wan et al (6) reported that the 3'-untranslated region of PTCH1 may promote NSCLC metastasis by activating the Wnt signaling pathway and mRNA splicing. These pathways have crucial roles in NSCLC progression. Wnt signaling regulators have widely been observed to be dysregulated in NSCLC. Wnt signaling may regulate NSCLC cell proliferation (27), apoptosis (28) and metastasis (29). Furthermore, the Wnt pathway may enhance resistance to chemotherapy and radiotherapy (30). However, the effect of RBAK on these pathways requires further validation. In addition, in vivo experimental validation of the functions of RBAK remains to be performed in future studies. Furthermore, the upstream regulators of RBAK still require to be investigated. Future studies focusing on these points may promote the understanding regarding the important roles of RBAK in NSCLC progression.

In conclusion, the present study demonstrated that RBAK is upregulated in NSCLC vs. non-tumorous tissues samples and may serve as a novel prognostic marker. Furthermore, the present study was the first, to the best of our knowledge, to reveal that RBAK is a novel regulator of cell metastasis. The present study may serve as a basis for the development of novel potential therapeutic approaches for NSCLC.

Acknowledgements
Not applicable.

Funding
No funding was received.

Availability of data and materials
The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Authors’ contributions
Conception and design: BH and GY. Development of methodology and in vitro assays: BH, BW and HW. Analysis and interpretation of data: BH, CZ, YW and LF. Writing, review and/or revision of the manuscript: BH and GY. All authors read and approved the final manuscript.

Ethics approval and informed consent
Not applicable.

Patients’ consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.


