

Role of Rho guanine nucleotide exchange factors in non-small cell lung cancer

Rui-Jie Zeng^{a,b,c,#}, Wei-Jie Xie^{a,b,#}, Chun-Wen Zheng^{a,b,#}, Wan-Xian Chen^{a,b}, Si-Meng Wang^{a,b}, Zheng Li^{a,b}, Chi-Bin Cheng^{a,b}, Hai-Ying Zou^{a,b}, Li-Yan Xu^{a,d}, and En-Min Li^{a,b}

^aDepartment of Biochemistry and Molecular Biology, Shantou University Medical College, Shantou China; ^bThe Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area, Shantou University Medical College, Shantou China; ^cDepartment of Gastroenterology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou China; ^dInstitute of Oncologic Pathology, Shantou University Medical College, Shantou China

ABSTRACT

Conventionally, Rho guanine nucleotide exchange factors (GEFs) are known activators of Rho guanine triphosphatases (GTPases) that promote tumorigenesis. However, the role of Rho GEFs in non-small cell lung cancer (NSCLC) remains largely unknown. Through the screening of 81 Rho GEFs for their expression profiles and correlations with survival, four of them were identified with strong significance for predicting the prognosis of NSCLC patients. The four Rho GEFs, namely ABR, PREX1, DOCK2 and DOCK4, were downregulated in NSCLC tissues compared to normal tissues. The downregulation of ABR, PREX1, DOCK2 and DOCK4, which can be attributed to promoter methylation, is correlated with poor prognosis. The underexpression of the four key Rho GEFs might be related to the upregulation of MYC signaling and DNA repair pathways, leading to carcinogenesis and poor prognosis. Moreover, overexpression of ABR was shown to have a tumor-suppressive effect in PC9 and H1703 cells. In conclusion, the data reveal the unprecedented role of ABR as tumor suppressor in NSCLC. The previously unnoticed functions of Rho GEFs in NSCLC will inspire researchers to investigate the distinct roles of Rho GEFs in cancers, in order to provide critical strategies in clinical practice.

ARTICLE HISTORY

Received 22 July 2021
Revised 12 October 2021
Accepted 13 October 2021

KEYWORDS

Rho GEFs; ABR; PREX1;
DOCK2; DOCK4; NSCLC

1. Introduction

Lung cancer is one of the most common malignancies and a leading cause of cancer deaths (>1 million annually) worldwide [1,2]. Despite therapeutic advances, the prognosis of lung cancer remains poor, and more than half of the patients diagnosed with lung cancer die within one year [3]. Non-small cell lung cancer (NSCLC) constitutes approximately 85% of all lung cancers, of which lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are the most common subtypes [3,4]. A comprehensive understanding of genetic alterations and associated mechanisms in the development of NSCLC is required to more effectively predict the prognosis of patients and identify druggable targets in cancer therapeutics.

Rho guanosine triphosphatases (GTPases) are essential molecular switches involved in the regulation of numerous downstream pathways in various types of cancer [5,6]. The cycling between guanosine

triphosphate (GTP)-bound (active) state and guanosine diphosphate (GDP)-bound (inactive) state contributes to the activation or inactivation of downstream effectors [7]. In the 'on' state (GTP-bound) of Rho GTPases, target proteins are recognized and subsequently a response is generated until the Rho GTPases are turned to 'off' state (GDP-bound) [5]. Ras-related C3 botulinum toxin substrate 1 (RAC1), Ras homolog family member A (RHOA) and cell division control protein 42 homolog (CDC42), which are the most important and extensively studied members of Rho GTPases, have been identified to regulate actin cytoskeleton reorganization, migration, and metastasis and promote the development of lung cancer [8–11].

Rho GTPases activity is principally regulated by three types of regulatory proteins, including guanine nucleotide exchange factors (GEFs) for activation, GTPase-activating proteins (GAPs) for

CONTACT Li-Yan Xu  lyxu@stu.edu.cn  Institute of Oncologic Pathology, Shantou University Medical College, Shantou 515041, China; En-Min Li  lyxu@stu.edu.cn  The Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area Shantou University Medical College, Shantou 515041, China

[#]These authors contributed equally to this work.

 Supplemental data for this article can be accessed [here](#)

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

inactivation and GDP-dissociation inhibitors (GDIs) for GDP-bound form stabilization [12]. GEFs catalyze GDP release and subsequently assist in its binding to GTP, converting inactive Rho GTPases to their active states for further functioning [13]. As activators of Rho GTPases, Rho GEFs have attracted the attention of researchers in recent years. Rho GEFs can be divided into two distinct families: the diffuse B-cell lymphoma (DBL) family and the dedicator of cytokinesis (DOCK) family [14]. There are 70 members of the DBL family GEFs and 11 members of the DOCK family GEFs [14,15]. The altered expression or mutation of Rho GEFs has been identified in human cancers [16]. However, compared to the roles of Rho GTPase family members, the roles of Rho GEF family members in NSCLC remain largely unclear. Consequently, exploring the altered expression and relevant mechanisms of Rho GEFs leading to NSCLC development is significant.

Herein, we hypothesize that Rho GEFs can exert functions in addition to Rho GTPase activation in NSCLC. The aim of our study was to analyze the expression of Rho GEFs and their functions in NSCLC. Certain Rho GEFs will be selected for their observably altered expression and significant predictive value for prognosis in NSCLC patients. Moreover, methylation and mutation profiles were analyzed to interpret the observed altered expression of the selected members. The potential mechanisms related to the selected Rho GTPases in NSCLC development were investigated. *In vitro* experiments were performed to confirm the role of the selected Rho GTPase in NSCLC.

2. Materials and methods

2.1 Phylogenetic and protein-protein interaction (PPI) analysis

The protein sequences of human DBL family Rho GEFs (70 members) and human DOCK family Rho GEFs (11 members) were obtained from UniProt (<https://www.uniprot.org/>). The phylogenetic tree was constructed by MEGA7.0 software using neighbor-joining method with default parameters and 1000 bootstrap replicates [17]. The final tree containing the information of the domain organization was visualized by the online

tool Evolview [18]. PPI networks were created by Search Tool for the Retrieval of Interacting Genes (STRING) (correlation score > 0.4) [19].

2.2 Oncomine database analysis

Oncomine (<https://www.oncomine.org/>) is an online mining platform by which the expression analyses on comparing the transcriptome data in most types of cancer with their corresponding normal tissues can be performed [20,21]. The mRNA expression levels of genes encoding Rho GEFs in NSCLC tissues, represented by LUAD and LUSC, as well as normal lung tissues were analyzed by Oncomine. In this study, P -value < 0.05, fold change > 2.0 and top 5% gene rank were selected as the thresholds. The resulting data were input into and visualized by Microsoft Office Excel 16.0 software (Microsoft Corporation, Redmond, CA).

2.3 Gene Expression Profiling Interactive Analysis (GEPIA) database analysis

GEPIA (<http://gepia.cancer-pku.cn/>) is an online tool that contains differential expression analysis between tumors and normal tissues based on The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) data [22]. To date, TCGA has produced RNA-Seq data including 9736 tumor samples across 33 cancer types, as well as data containing 726 adjacent normal tissues. The GTEx project contains RNA-Seq data for more than 8000 normal samples from unrelated donors. GEPIA integrates the information from cancer genomics big data for end users. The GEPIA database was used to validate the transcriptional profiles of Rho GEFs in patients with NSCLC.

2.4 UALCAN database analysis

UALCAN (<http://ualcan.path.uab.edu/>) is a web server that uses TCGA RNA-seq and clinical data from 31 cancer types [23]. This database provides a platform for identifying candidate biomarkers specific for tumor sub-groups. The UALCAN database was utilized to analyze the mRNA expression of normal tissues and NSCLC specimens from different sub-groups based on nodal metastasis status and tumor stages.

2.5 Kaplan–Meier Plotter database analysis

Kaplan–Meier Plotter (<https://kmplot.com/>) is an online database capable of assessing the effects of 54,000 genes on survival in 21 types of cancer. This system includes gene chip and RNA-seq data from the Gene Expression Omnibus (GEO), European Genome-Phenome Archive (EGA), and TCGA database. In this database, data of lung cancer are available [24]. The patient specimens were divided into high expression and low expression groups using the JetSet best probe set and auto-selected best cutoff [25]. Outlier arrays were excluded to control the array quality [26]. Overall survival (OS) and first progression (FP) were analyzed in patients with NSCLC. Log-rank *P*-value and hazard ratio (HR) with 95% confidence interval (CI) were calculated and displayed on the web server. The Kaplan–Meier Plotter database was used to evaluate the prognostic value of genes encoding Rho GEFs in NSCLC.

2.6 MEXPRESS database analysis

MEXPRESS (<https://mexpress.be/>) is an online data visualization tool for the visualization and integration of TCGA expression, DNA methylation and clinical data [27,28]. The MEXPRESS database was employed to investigate the promoter methylation status of selected genes in NSCLC specimens compared to normal tissues, and only CpG sites with statistically significant results were reported.

2.7 cBioPortal for Cancer Genomics (cBioPortal) database analysis

cBioPortal (<http://cbioportal.org>) is a web resource for the exploration, visualization and analysis of cancer genomics data [29]. The datasets selected in our study were LUAD (TCGA, Firehose Legacy, containing 580 samples) and LUSC (TCGA, Firehose Legacy, containing 503 samples). The platform for methylation sequencing information was the Illumina Human Methylation 450 (HM450). The correlation between the expression of the selected genes and DNA methylation in LUAD and LUSC was determined by cBioPortal. Moreover, the mutations of the selected genes and their relationship with gene expression were investigated using cBioPortal.

2.8 Gene Set Enrichment Analysis (GSEA)

GSEA was used to determine whether a defined gene set was expressed with significant differences under two different biological conditions [30]. The TCGA data regarding LUAD (*n* = 479) and LUSC (*n* = 501) patients were downloaded from the Genomic Data Commons (GDC) Data Portal website (<https://portal.gdc.cancer.gov/>). Subsequently, the patients were classified into two groups (high vs. low expression) for each selected gene, and the cutoffs were determined as medians. GSEA was conducted on the mRNA expression data of the selected genes using GSEA 4.0.3 software (<http://software.broadinstitute.org/gsea/>). A list of protein-coding genes of hallmark gene sets (H) from the Molecular Signatures Database (MSigDB) (<https://www.gsea-msigdb.org/gsea/msigdb/>) was obtained for GSEA analysis [31], with 1,000 permutations to calculate the normalized enrichment score (NES) and *P* values. Gene sets with a false discovery rate (FDR) of < 0.25 and *P* < 0.05 were regarded as significantly enriched gene sets.

2.9 Cell lines and cell culture

293 T cells, adenocarcinoma cell PC9 and squamous cell H1703 in human NSCLC were used in this study to verify the tumor-suppressive effect of ABR. The 293 T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with high glucose (Hyclone; GE Healthcare Life Sciences, Logan, UT, USA). PC9 and H1703 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium (Gibco; Thermo Fisher Scientific, Inc., Waltham, MA, USA). All media were supplemented with 10% fetal bovine serum (FBS Gibco; Thermo Fisher Scientific, Inc., Waltham, MA, USA), 100 U/mL penicillin and 100 mg/L streptomycin. All cell lines were incubated in a 5% CO₂, 95% humidity incubator at 37°C.

2.10 Plasmid and stable-overexpression cell lines construction

The full-length sequences of ABR CF1: CCGCTCG AGGCCACCATGGAGCCGCTCAGCCACCGG and ABR CR1:GGAATTCCACGTCGGTGGAGA AGTACAG were amplified from 293 T cells with

primers (NM_021962.5). Then ABR sequences were connected to pcDNA3.1 with a flag tag and subcloned into the pBOBI-IRES-EGFP-T2A-Puro vector. After the virus was packaged in 293 T cells using Lipofectamine 3000 (Life Invitrogen, USA), ABR and an empty vector were transfected into the target cell line lung cancer cells for overexpression. According to the fluorescence expression, the cells were screened with puromycin (1 $\mu\text{g}/\text{mL}$) for 4 to 6 days.

2.11 Protein extraction and Western blotting

Total protein was extracted from all cell lines with RIPA buffer including 100X Thermo Cocktail Protease inhibitor (Thermo Fisher Scientific, Inc., Waltham, MA, USA). The cell lysates were subjected to SDS-PAGE and transferred to a polyvinylidene fluoride (PVDF) membrane. The PVDF membrane was blocked with 5% fat-free milk and incubated with the primary antibody directed against ABR (Abcam, UK) at room temperature for 1 h. Subsequently, the PVDF membrane was washed with TBS-Tween 20 (TBST) for three times, incubated with a horseradish peroxidase (HRP) goat anti-Rabbit (IgG) secondary antibody (Abcam, UK) at room temperature for 1 h, and washed again with TBST for three times. The PVDF membrane was imaged using FluorChem M (ProteinSimple, CA, USA). After stripping, the PVDF membrane was re-probed with β -actin antibody (Proteintech Group, USA) and the HRP-conjugated Affinipure Goat Anti-Mouse IgG (H + L) (Proteintech Group, CA, USA) as the secondary antibody.

2.12 Cell proliferation assay

The cells were seeded into 96 well plates at a density of 10,000 cells per well. Cell viability was measured at 2h, 24 h, 48 h and 72 h after inoculation. The absorbance of 490 nm was measured at 37°C for 1 h after incubation with Cell Titer 96[®] Aqueous One Solution Cell Proliferation Assay (MTS) (Promega, WI, USA). Results are representative of three independent experiments.

2.13 Colony formation assay

Cells were seeded at a density of 1,000 cells per well in 6-well plates and cultured at 37°C for colony formation. After 7 to 14 days, the cells were fixed in

anhydrous formaldehyde for 15 min and stained with 0.1% (w/v) crystal violet at room temperature for 15 min. The colonies were photographed using a Fluor Chem R Imaging system (Protein Simple) and counted by Image J software. Results are representative of three independent experiments.

2.14 Cell migration assay

After starvation in serum-free medium for 12 h, the cells were seeded into a Transwell chamber at a density of 50,000 cells per well. The cells were cultured in serum-free or 10% FBS medium in a chamber. The cells were cultured in a 37°C and 5% CO₂ incubator for 24 h. Finally, the cells were fixed in anhydrous formaldehyde for 15 min, stained with 0.1% (w/v) crystal violet for 15 min, and rinsed with PBS. After that, the cells were placed under a 200X microscope, photographed for 10 fields, and counted by Image J software. Results are representative of three independent experiments.

3. Results

Rho GEFs are conventionally known as activators of Rho GTPases, which promote carcinogenesis. To date, there is a lack of comprehensive understanding of the role of Rho GEFs in NSCLC. The aim of our study was to provide a more comprehensive understanding of Rho GEFs. We hypothesized that Rho GEFs can exert functions in addition to Rho GTPase activation in NSCLC. Our bioinformatic analysis and in vitro experiments confirmed our hypothesis for the Rho GEF ABR.

3.1 Structure and domains of Rho GEF family proteins

The human Rho GEFs are comprised of 81 members, which can be mainly divided into two families, namely DOCK family (Figure 1a) and DBL family (Figure 1b). The structures of Rho GEFs are various, but conserved domains have also been identified. Proteins of the DBL Rho GEFs were defined by the presence of conserved Dbl-homology (DH) domains. The DBL family of Rho GEFs contains 70 members with 72 DH domains, among which TRIO and KALIRIN possess two DH domains, and the other members all contain one DH domain, which is responsible for stimulating

guanine nucleotide exchange. In addition to the DH domain, most DBL Rho GEFs contain a pleckstrin-homology (PH) domain following the DH domain. The PH domains are able to bind to phospholipids and are essential to properly mediate the localization of DBL Rho GEFs to their action sites via lipid binding. Proteins of the DOCK Rho GEFs were classified by the presence of DOCK-homology region 1 (DHR-1) and DOCK-homology region 2 (DHR-2) domains. Similar to the DH and PH domains, DHR-2 domains are responsible for guanine nucleotide exchange, and DHR-1 domains interact with phospholipids, which aid in targeting DOCK Rho GEFs to the plasma membrane. Apart from the DH, PH, DHR-1, and DHR-2 domains, other domains of Rho GEFs, such as Src homology 2 (SH2) and Src homology 3 (SH3) domains, are important in mediating protein-protein interactions, protein-lipid interactions, and messenger binding. In order to highlight the highly conserved domains of Rho GEFs, the details of the other domains displayed in [Figure 1](#) are illustrated in [Supplementary Figure S1](#). The PPIs of the Rho GEFs are demonstrated in [Supplementary Fig. S2](#),

3.2 The expression levels of Rho GEFs are altered in lung cancer.

Oncomine database analysis was performed to investigate the altered mRNA expression in lung cancer. There are 43 genes encoding DBL GEFs ([Supplementary Figure S3A](#)) and 7 genes ([Supplementary Figure S3B](#)) encoding DOCK GEFs showing remarkable alterations of mRNA expression levels in lung cancer compared to normal tissues. For the DBL family of Rho GEFs, 24 genes were downregulated in lung cancer, whereas 13 genes were upregulated, and the other 6 genes showed contradictory results of aberrant expression in different datasets. For the DOCK family of Rho GEFs, 6 genes were underexpressed in lung cancer, while only 1 gene was identified to be overexpressed. Moreover, GEPIA database analysis was used to verify the altered mRNA expression of genes encoding Rho GEFs in NSCLC. Due to the fact that gene expression patterns can vary in LUAD and LUSC, the following analyses were performed independently in LUAD and LUSC [32]. Most of the genes encoding DBL Rho GEFs were downregulated in LUAD ([Figure 2a](#)) and LUSC ([Figure 2b](#)). For all of the DOCK Rho GEFs except

DOCK7, the median expression of genes was also decreased in LUAD ([Figure 2a](#)) and LUSC ([Figure 2b](#)) compared to normal samples. Furthermore, the UALCAN and Kaplan Meier Plotter database analyses for all genes encoding Rho GEFs were performed (data not shown). The genes showing contradictory effects within or across different databases were excluded. Eventually, ABR and PREX1 in the DBL family of Rho GEFs, as well as DOCK2 and DOCK4 in the DOCK family of Rho GEFs, were selected for this study.

3.3 ABR, PREX1, DOCK2 and DOCK4 are downregulated in NSCLC.

In order to obtain an overview of the expression patterns of ABR in various tumor types and in the corresponding normal tissues, a body map was generated by GEPIA ([Figure 3a](#)). Compared to normal lung tissue, the expression of ABR is lower in lung cancer. To further validate the downregulation of ABR in lung cancer, specific datasets from Oncomine analysis are illustrated ([Figure 3a](#)). From the Garber Lung dataset, the expression of ABR in LUSC was markedly lower than that in normal lung tissue ($P < 0.001$, fold change -2.435), while the datasets of LUAD with statistical significance are lacking in the Oncomine database. Therefore, the GEPIA database was used to confirm our results. From the GEPIA database analysis of both LUAD and LUSC, ABR was significantly downregulated in tumors ($P < 0.05$). The downregulation of PREX1 ([Figure 3b](#)), DOCK2 ([Figure 3c](#)), and DOCK4 ([Figure 3d](#)) was also validated by Oncomine and GEPIA database analysis. These results imply that ABR, PREX1, DOCK2, and DOCK4 can function as tumor suppressors in NSCLC, while more specific data on tumor subgroups and the corresponding survival data are needed for confirmation. Consequently, the expression levels of the selected genes in the tumor subgroups divided by nodal metastasis status or individual stages were investigated.

From the UALCAN database analysis, ABR ([Supplementary Figure 4a](#)), PREX1 ([Supplementary Figure 4c](#)), DOCK2 ([Supplementary Figure 4e](#)), and DOCK4 ([Supplementary Figure 4g](#)) were significantly downregulated in each nodal metastasis status (N0, N1, N2, and N3) of LUAD and LUSC when compared to normal tissues. In addition, the expression of ABR ([Supplementary Figure 4b](#)), PREX1 ([Supplementary](#)

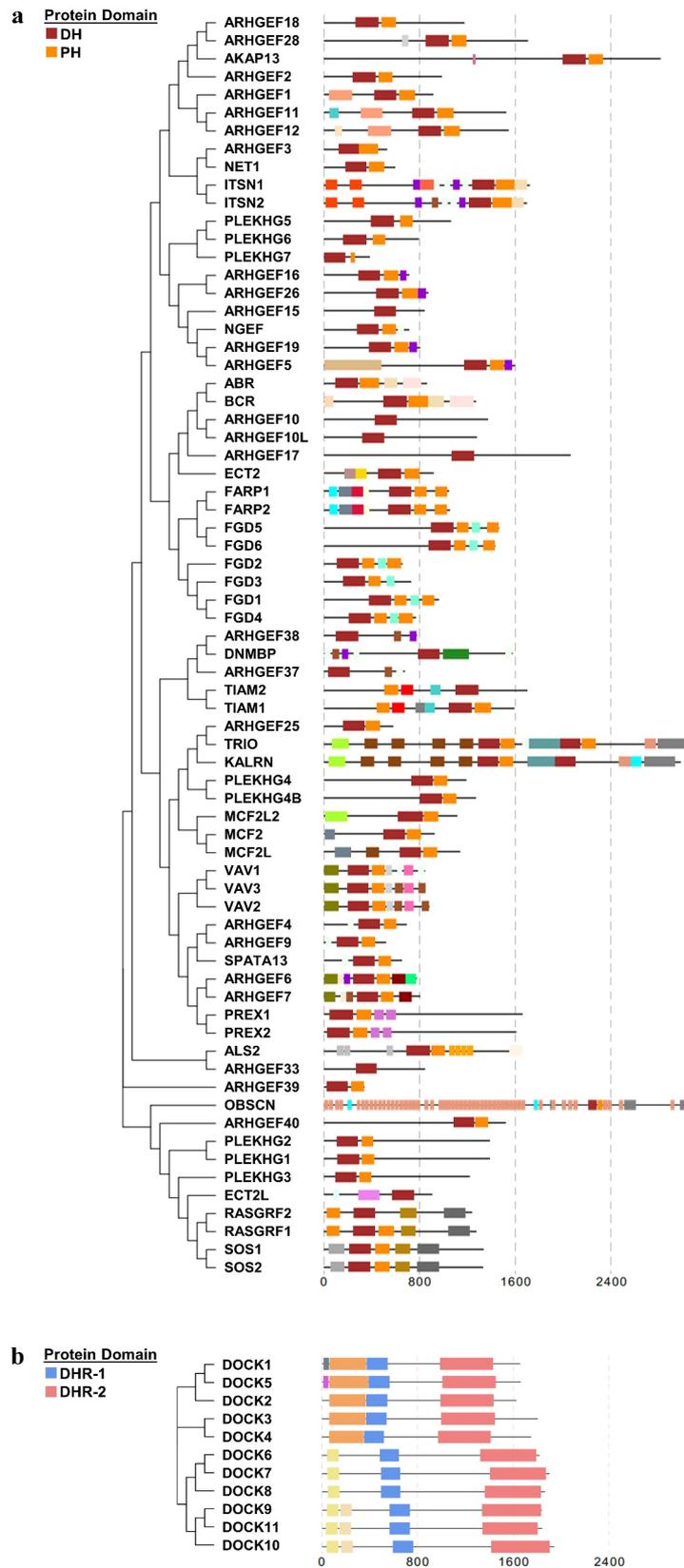


Figure 1. The human DBL family and DOCK family of Rho GEFs. (a) The 70 members of DBL family Rho GEFs and (b) 11 members of DOCK family Rho GEFs are arranged in phylogenetic tree by MEGA7.0 software, using the protein sequences obtained from UniProt. The protein domain architecture is visualized by Evolview. Each color represents one distinct structural domain. The DH, PH domains of DBL Rho GEF family and DHR-1, DHR2 domains of DOCK Rho GEF family with the corresponding colors are highlighted in the left panel. All proteins are visualized in the same scale except OBSCN. For other structural domains, please refer to FigureS1. DH: Dbl homology; PH: Pleckstrin homology; DHR-1: DOCK homology region 1; DHR-2: DOCK homology region.

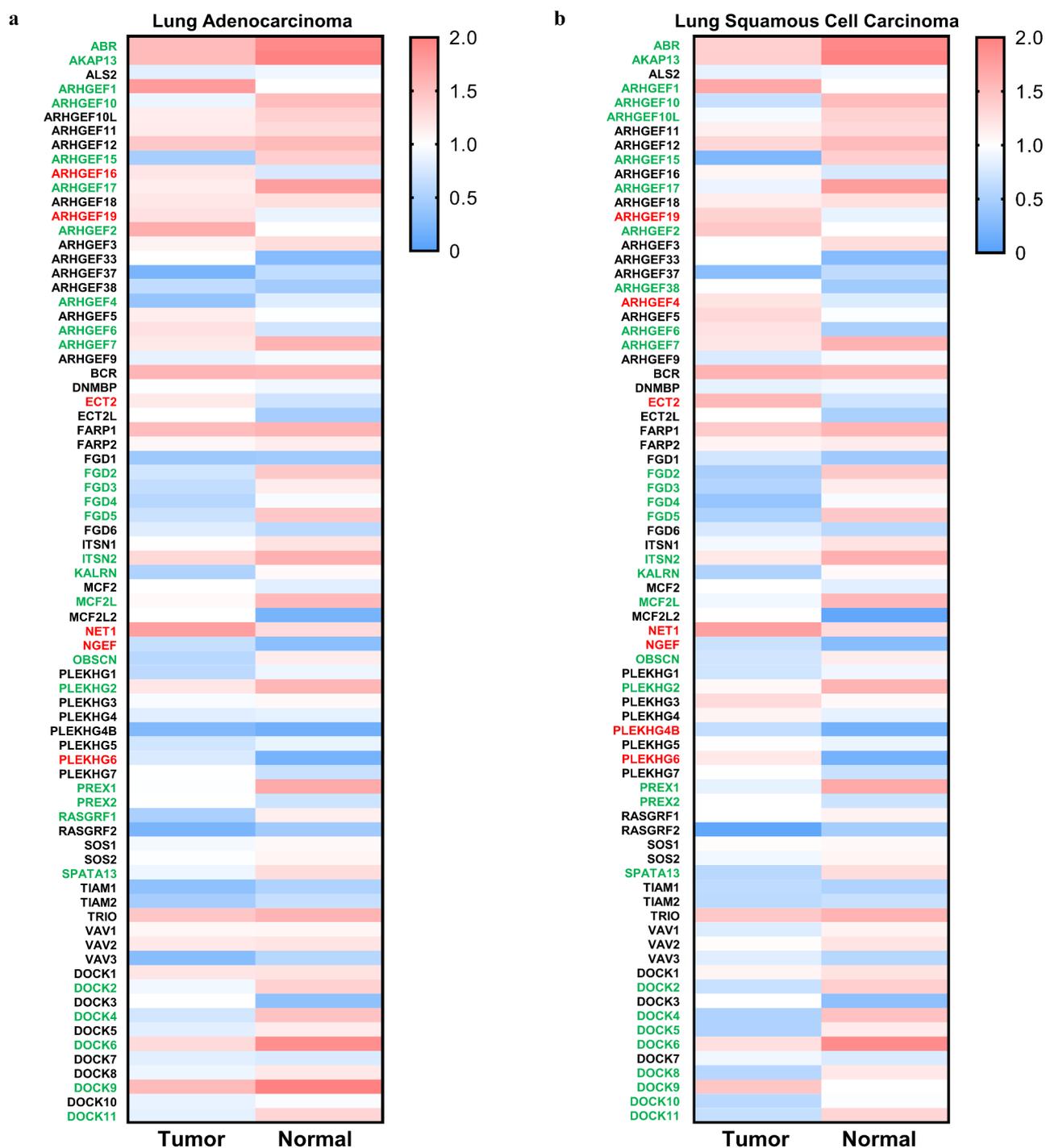


Figure 2. Expression of Rho GEFs in NSCLC. The expression levels of Rho GEFs in NSCLC, represented by LUSC and LUAD, are analyzed by GEPIA database. (a) The median expression of Rho GEFs in LUAD compared to normal tissues. (b) The median expression of Rho GEFs in LUSC compared to normal tissues. The transcripts per million (TPM) of genes are normalized by $\log_{10}(\text{TPM})$. Gene name with red (higher expression compared to normal tissue) and green (lower expression compared to normal tissue) color indicates $P < 0.05$ for their expressions in tumors compared to normal tissues.

Figure 4d), DOCK2 (Supplementary figure 4f), and DOCK4 (Supplementary Figure 4h) were observably lower in each stage (stage 1, stage 2, stage 3 and stage 4) of LUAD and LUSC compared to normal samples.

Although the median expression of the selected genes generally decreased as nodal metastasis status or tumor stage advanced, the majority of these data showed no statistical significance. Taken together,

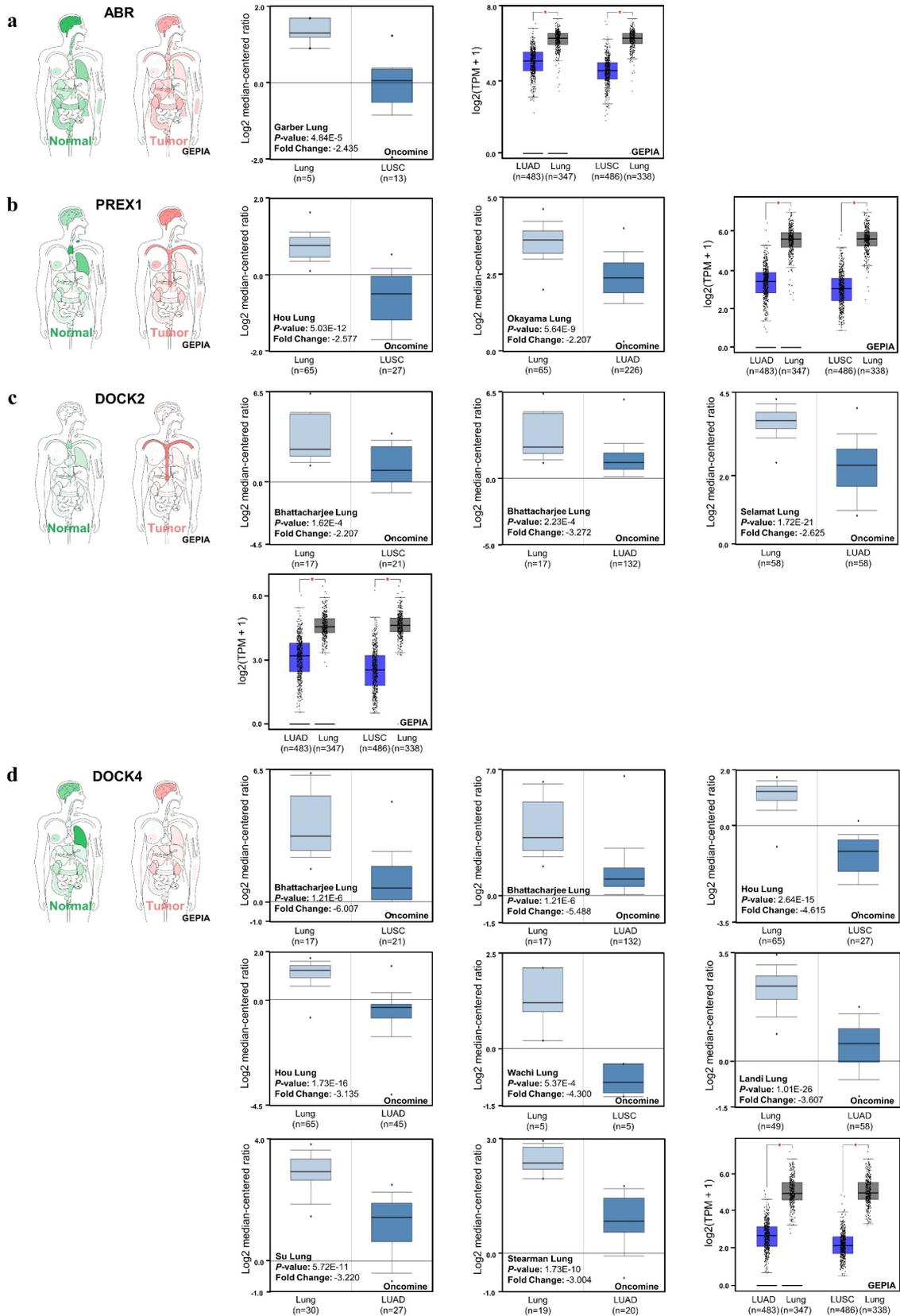


Figure 3. Expression of ABR, PREX1, DOCK2 and DOCK4 in NSCLC. (a) The body map of various organs (GEPIA), the cohort of Garber Lung (Oncomine), and the LUSC and LUAD patient specimens with normal tissues for the expression of ABR (GEPIA) are illustrated. (b) The body map (GEPIA), the cohorts of Hou Lung and Okayama Lung (Oncomine), and the LUSC and LUAD patient samples with normal tissues for PREX1 expression (GEPIA) are demonstrated. (c) The body map including different organs (GEPIA),

the cohorts of Bhattacharjee Lung and Selamat Lung (Oncomine), and the LUSC and LUAD patient samples with normal tissues for DOCK2 expression (GEPIA) are shown. (d) The body map (GEPIA), the cohorts of Bhattacharjee Lung, Hou Lung, Wachi Lung, Landi Lung, Su Lung and Stearman Lung (Oncomine), and the LUSC and LUAD patient samples with normal tissues for DOCK4 expression (GEPIA) are illustrated. Red: tumor tissue. Green: normal tissue. The deeper color in the body map represents higher expression.* $P < 0.05$.

these consistent results confirm that ABR, PREX1, DOCK2, and DOCK4 are notably downregulated in NSCLC, and their prognostic values were subsequently investigated.

3.4 Downregulations of ABR, PREX1, DOCK2 and DOCK4 are prognostic factors for NSCLC.

Using the Kaplan–Meier Plotter database, survival analysis was performed to determine the prognostic value of the selected genes in NSCLC patients. With regard to overall survival, ABR (Figure 4a), PREX1 (Figure 4c), DOCK2 (Figure 4e), and DOCK4 (Figure 4g) all demonstrated significant values in the prognosis of LUAD and LUSC patients, and a lower expression of these genes predicted poor OS in patients. For LUAD patients, the expression of DOCK4 was more significantly correlated with OS (HR = 0.33; 95% CI: 0.24–0.44; $P = 1.4E-14$) compared to that of ABR, PREX1, and DOCK2. In LUSC patients, DOCK4 expression was also more markedly correlated with OS (HR = 0.64; 95% CI: 0.49–0.82; $P = 0.00055$) than ABR, PREX1, and DOCK2 expression. However, for the prognosis of tumor progression, ABR showed no significance in both LUAD and LUSC patients (Figure 4b). In addition, lower DOCK4 expression was associated with earlier FP in LUAD patients, but not in LUSC patients (Figure 4h). The downregulation of PREX1 (Figure 4d) and DOCK2 (figure 4f) were significantly correlated with earlier tumor progression in LUSC and LUAD patients. Among the four Rho GEFs, DOCK4 predicted the progression in LUAD patients with a higher significance (HR = 0.42; 95% CI: 0.29–0.62; $P = 5.7E-06$), while DOCK2 more observably predicted LUSC progression (HR = 0.48; 95% CI: 0.29–0.8; $P = 0.0041$). Collectively, through the analysis using Oncomine, GEPIA, UALCAN, and Kaplan–Meier Plotter databases, consistent results highlight the downregulation of ABR, PREX1, DOCK2, and DOCK4 in NSCLC, and they notably predict the poor prognosis of NSCLC patients.

3.5 Promoter methylation levels contribute to the aberrant expression of ABR, PREX1, DOCK2, and DOCK4 in NSCLC

DNA methylation in promoter regions is important for modulating gene expression by inducing stable epigenetic inhibition of gene expression. Therefore, whether methylation is correlated with the expression of ABR, PREX1, DOCK2, and DOCK4 was examined. cBioPortal database analysis revealed that the promoter methylation levels of ABR ($P = 6.91E-20$, $r = -0.41$ for LUAD; $P = 7.31E-8$, $r = -0.28$ for LUSC) (Figure 5a), PREX1 ($P = 1.786E-5$, $r = -0.20$ for LUAD; $P = 2.689E-3$, $r = -0.16$ for LUSC) (Figure 5b), DOCK2 ($P = 5.12E-10$, $r = -0.29$ for LUAD; $P = 4.785E-5$, $r = -0.21$ for LUSC) (Figure 5c) and DOCK4 ($P = 1.638E-6$, $r = -0.22$ for LUAD; $P = 1.609E-5$, $r = -0.22$ for LUSC) (Figure 5d) were negatively correlated with mRNA expression in NSCLC. MEXPRESS database revealed a large number of probes that were significantly associated with the expressions of ABR, PREX1, DOCK2 and DOCK4 (Supplementary FigureS5). The above results suggest that the hypermethylation of ABR, PREX1, DOCK2, and DOCK4 is not only identified in NSCLC in comparison to normal tissues, but is also negatively correlated with their mRNA expression in NSCLC.

3.6 Analysis of ABR, PREX1, DOCK2, and DOCK4 mutations in NSCLC.

Genetic alterations of ABR, PREX1, DOCK2, and DOCK4 were studied by the cBioPortal database. The proportion and distribution of TCGA samples with genetic alterations in LUAD (Supplementary FigureS6A) and LUSC (Supplementary FigureS6B) are shown, and DOCK2 possesses the highest proportion of genetic alterations in both LUAD (13%) and LUSC (10%) among the four selected Rho GEFs. The locations of the mutations are further

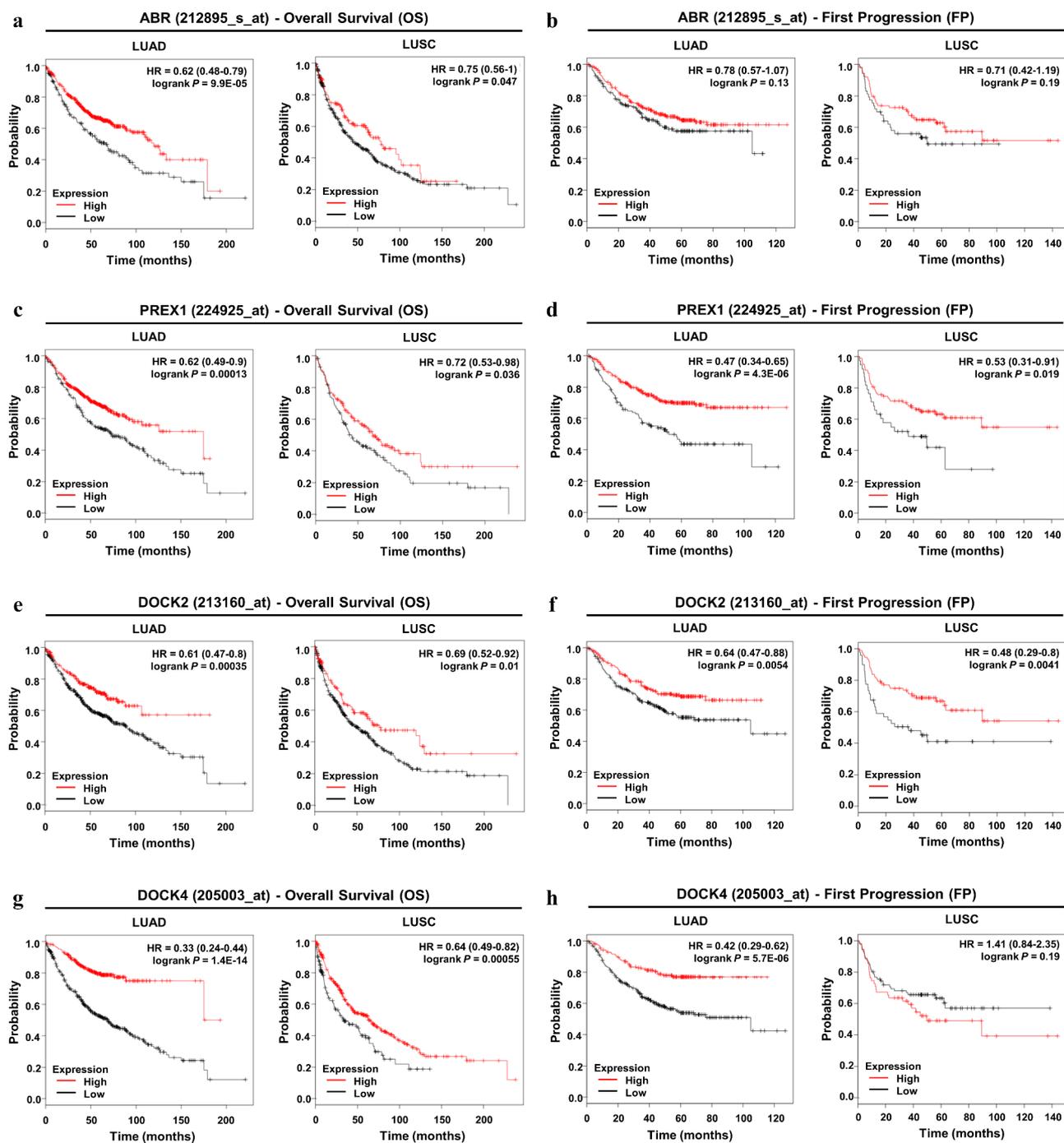


Figure 4. Prognostic value of ABR, PREX1, DOCK2 and DOCK4 expression in NSCLC. (A, C, E, G) Survival curves referring to OS are plotted for LUAD and LUSC patients. **(B, D, F, H)** Survival curves with regard to FP are generated for patients with LUAD and LUSC. Log-rank P values and HRs with 95% CIs are displayed. FP: first progression; HR: hazard ratio; OS: overall survival.

illustrated in Supplementary FigureS6C and Supplementary FigureS6D. The majority of the mutations in the four selected Rho GEFs were missense mutations, and several mutations were located in the regions encoding the conserved

domains of Rho GEFs. However, the mutations of the selected genes have no statistically significant impact on the mRNA expression in LUAD (Supplementary FigureS6E) and LUSC (Supplementary FigureS6F).

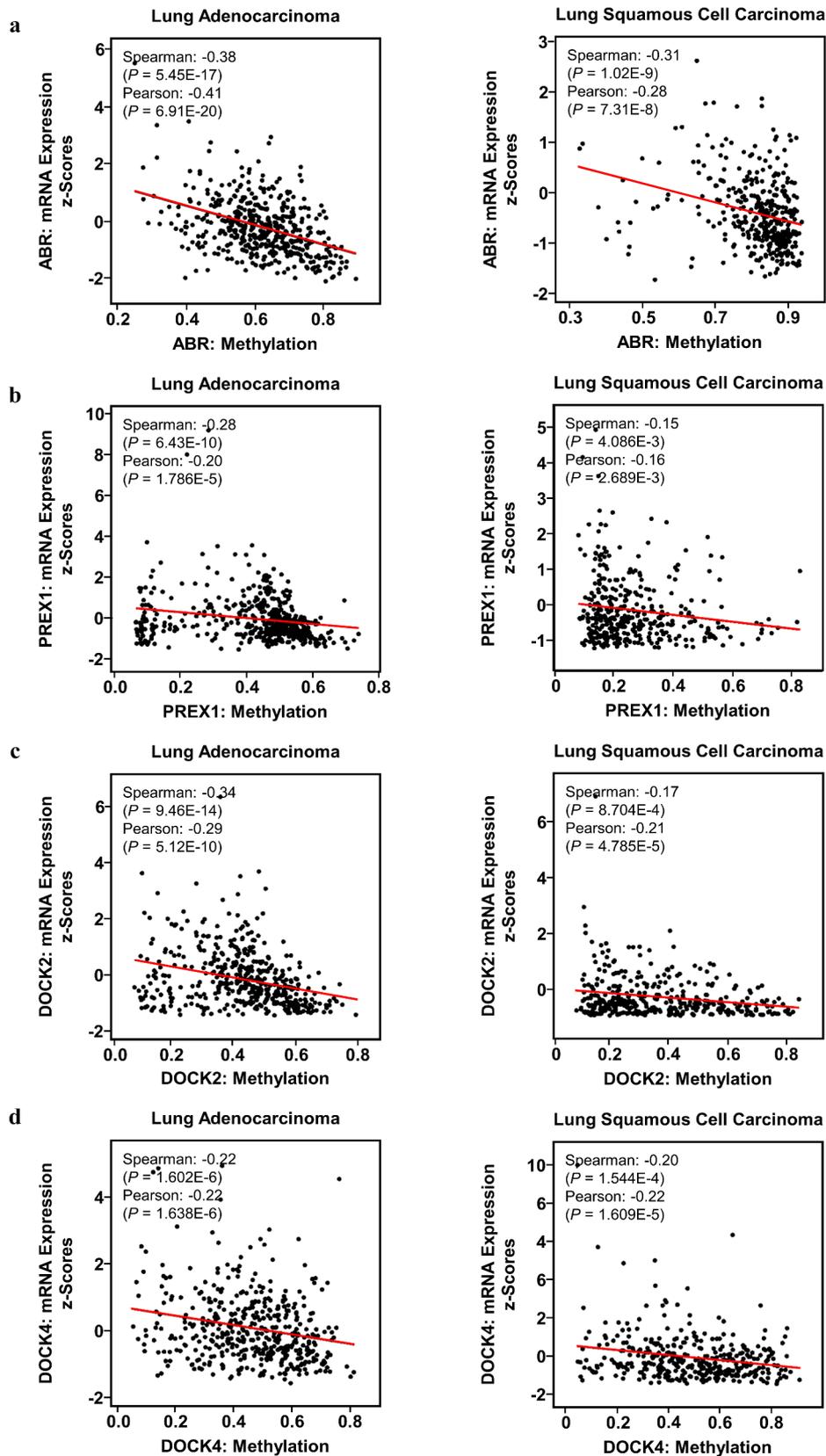


Figure 5. Impact of DNA methylation on ABR, PREX1, DOCK2 and DOCK4 expression in NSCLC. Visualization of TCGA data for the promoter region methylation and mRNA expression of (a) ABR, (b) PREX1, (c) DOCK2 and (d) DOCK4 in LUAD and LUSC, compared to those in normal tissues using MEXPRESS. The correlation of (e) ABR, (f) PREX1, (g) DOCK2 and (h) DOCK4 expression with promoter methylation is analyzed by cBioPortal database. Spearman's correlation analysis and Pearson's correlation analysis are performed in the database. The regression line is used to illustrate the correlation trend.

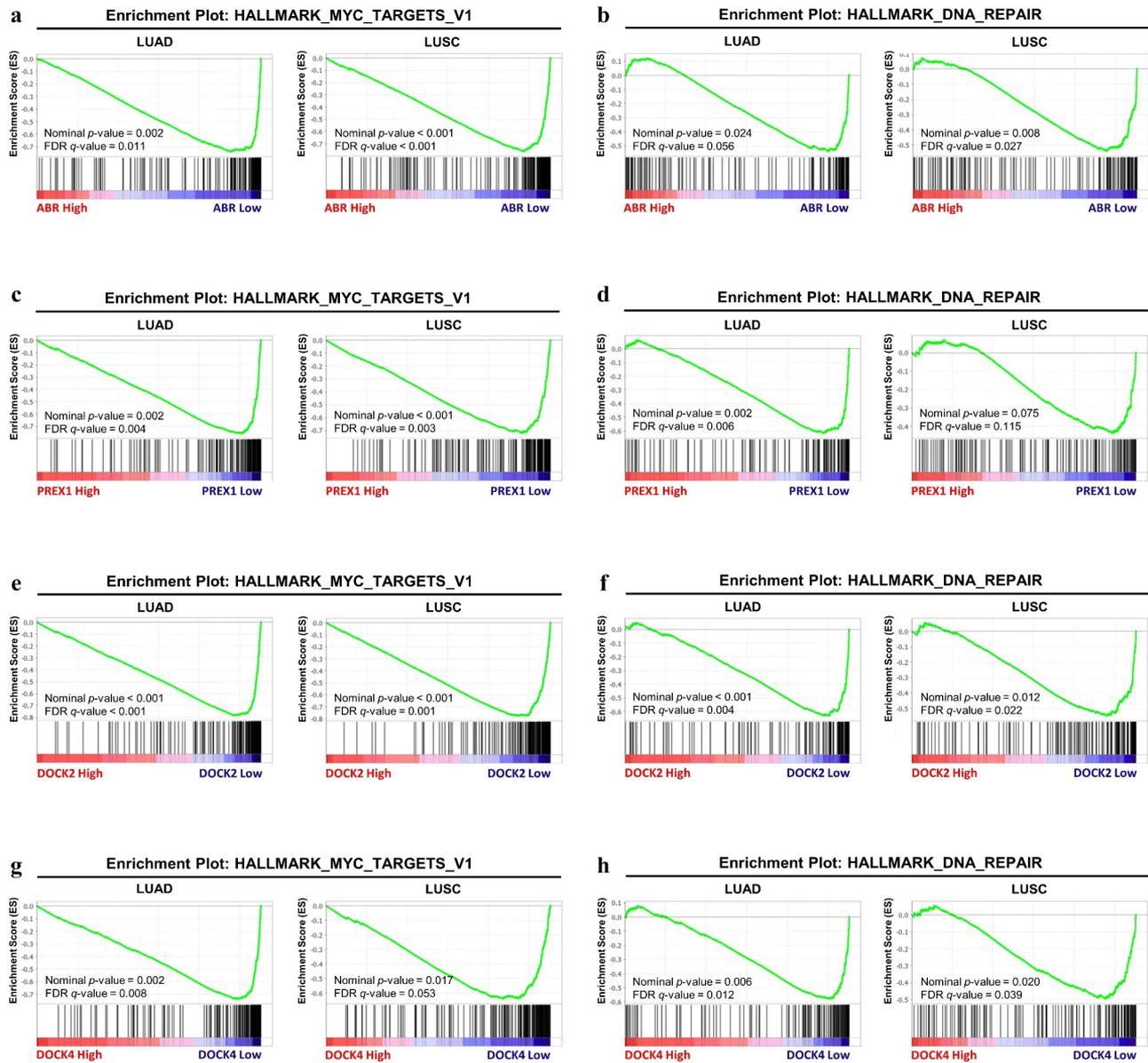


Figure 6. GSEA analysis between the high-expression group and low-expression group in ABR, PREX1, DOCK2 and DOCK4 for NSCLC. The correlation between the enrichment of MYC signaling gene set and (a) ABR, (c) PREX1, (e) DOCK2 and (g) DOCK4 expression in LUAD and LUSC. The association between the expression of (b) ABR, (d) PREX1, (f) DOCK2 and (h) DOCK4 expression with DNA repair gene set enrichment. The barcode plot demonstrates gene positions in each gene set. The horizontal bar indicates positive (red) and negative correlation (blue) with gene expression. FDR: false discovery rate.

3.7 GSEA identifies potential mechanisms by which ABR, PREX1, DOCK2, and DOCK4 regulate NSCLC development and progression

Based on the data from multiple analyses, the lower expression of ABR, PREX1, DOCK2, and DOCK4 can lead to NSCLC development and progression, and therefore, GSEA by TCGA data was employed to explain the underlying mechanisms. The expression matrix of NSCLC patients is

divided into high-expression and low-expression groups based on the median expression levels of ABR, PREX1, DOCK2, and DOCK4, respectively. GSEA revealed that, in LUAD and LUSC, patients with lower expression of ABR (nominal $P = 0.002$, FDR = 0.011 for LUAD; nominal $P < 0.001$, FDR < 0.001 for LUSC) (Figure 6a), PREX1 (nominal $P = 0.002$, FDR = 0.004 for LUAD; nominal $P < 0.001$, FDR = 0.003 for LUSC) (Figure 6c),

Table 1. Gene Set Enrichment Analysis (GSEA) of gene sets significantly enriched in LUAD and LUSC.

Tumor Types	Genes Encoding GEFs	Enriched Gene Sets	Size	ES	NES	Nominal p -value	FDR q -value	FWER p -value
LUAD	ABR	HALLMARK_MYC_TARGETS_V1	199	-0.73912	-2.03656	0.002016	0.011191	0.027
		HALLMARK_DNA_REPAIR	150	-0.53791	-1.77365	0.023762	0.056209	0.191
	PREX1	HALLMARK_MYC_TARGETS_V1	199	-0.75718	-2.15914	0.001957	0.00364	0.007
		HALLMARK_DNA_REPAIR	150	-0.61628	-2.06687	0.00198	0.00618	0.022
	DOCK2	HALLMARK_MYC_TARGETS_V1	199	-0.7828	-2.27298	0	7.28E-04	0.002
		HALLMARK_DNA_REPAIR	150	-0.63067	-2.13108	0	0.004392	0.016
	DOCK4	HALLMARK_MYC_TARGETS_V1	199	-0.73671	-2.02493	0.001938	0.008172	0.032
		HALLMARK_DNA_REPAIR	150	-0.57858	-1.94703	0.005725	0.012478	0.056
LUSC	ABR	HALLMARK_MYC_TARGETS_V1	199	-0.7608	-2.25999	0	7.39E-04	0.003
		HALLMARK_DNA_REPAIR	150	-0.54108	-1.88272	0.008403	0.02707	0.116
	PREX1	HALLMARK_MYC_TARGETS_V1	199	-0.72201	-2.17652	0	0.003185	0.009
		HALLMARK_DNA_REPAIR	150	-0.43784	-1.51756	0.075397	0.115496	0.491
	DOCK2	HALLMARK_MYC_TARGETS_V1	199	-0.77524	-2.279	0	0.001378	0.002
		HALLMARK_DNA_REPAIR	150	-0.54779	-1.92649	0.012097	0.022362	0.091
	DOCK4	HALLMARK_MYC_TARGETS_V1	199	-0.63924	-1.9451	0.016736	0.05325	0.077
		HALLMARK_DNA_REPAIR	150	-0.49595	-1.81377	0.02004	0.039442	0.183

ES: enrichment score; NES: normalized enrichment score; FDR: false discovery rate; FWER: family-wise error rate.

DOCK2 (nominal $P < 0.001$, FDR < 0.001 for LUAD; nominal $P < 0.001$, FDR = 0.001 for LUSC) (Figure 6e) and DOCK4 (nominal $P = 0.002$, FDR = 0.008 for LUAD; nominal $P = 0.017$, FDR = 0.053 for LUSC) (Figure 6g) have a higher expression of genes associated with MYC signaling, which are essential for the development of lung cancer [33–35]. Moreover, our results also revealed that the DNA repair pathway was significantly enriched under the downregulation of ABR (nominal $P = 0.024$, FDR = 0.056) (Figure 6b), PREX1 (nominal $P = 0.002$, FDR = 0.006) (Figure 6d), DOCK2 (nominal $P < 0.001$, FDR = 0.004) (figure 6f), and DOCK4 (nominal $P = 0.006$, FDR = 0.012) in LUAD (Figure 6h). Similarly, the DNA repair pathway enrichment was identified under ABR (nominal $P = 0.008$, FDR = 0.027) (Figure 6b), DOCK2 (nominal $P = 0.012$, FDR = 0.022) (figure 6f), and DOCK4 (nominal $P = 0.020$, FDR = 0.039) (Figure 6h) downregulation in LUSC. However, the DNA repair pathway was not enriched in the PREX1-low-expression group in LUSC (nominal $P = 0.075$, FDR = 0.115) (Figure 6d). The details of the data are summarized in Table 1.

3.8 ABR shows tumor suppression effect in proliferation, migration, and cloning ability in PC9 and H1703 cells

In addition to analyzing the tumor-suppressive effects of ABR, PREX1, DOCK2, and DOCK4 through bioinformatics, we further verified the

role of ABR in proliferation, migration, and cloning ability in PC9 and H1703 cells experimentally. Cell proliferation experiments of PC9 and H1703 cell lines with overexpression of NC and ABR1 showed that overexpression of ABR inhibited cell proliferation in both cell types (Figure 7a). The migration capacity of ABR overexpression was measured in PC9 and H1703 cells, which showed that ABR presented a tumor-suppressive effect when compared to the NC group (Figure 7b). Cloning ability was also tested and analyzed, and ABR overexpression showed a suppressive effect in the number of clones in both PC9 and H1703 cells (Figure 7c). The above results suggest that the overexpression of ABR, as a representative of Rho GEFs, is prone to exert a tumor-suppressive effect in NSCLC.

4. Discussions

Rho GEFs are known for their critical roles as molecular switches in activating Rho GTPases, and therefore function as regulators of various diseases not limited to cancer [5,6]. ABR, as indicated by its name, was first identified as a breakpoint cluster region (BCR)-related protein that shares high homology with BCR in humans (68% amino acid identity) [36]. ABR maintains the normal reactivity of the innate immune system, and the altered function of ABR can lead to the development of leukemia [37,38]. PREX1 was first discovered in the cytosol of neutrophils [39]. PREX1

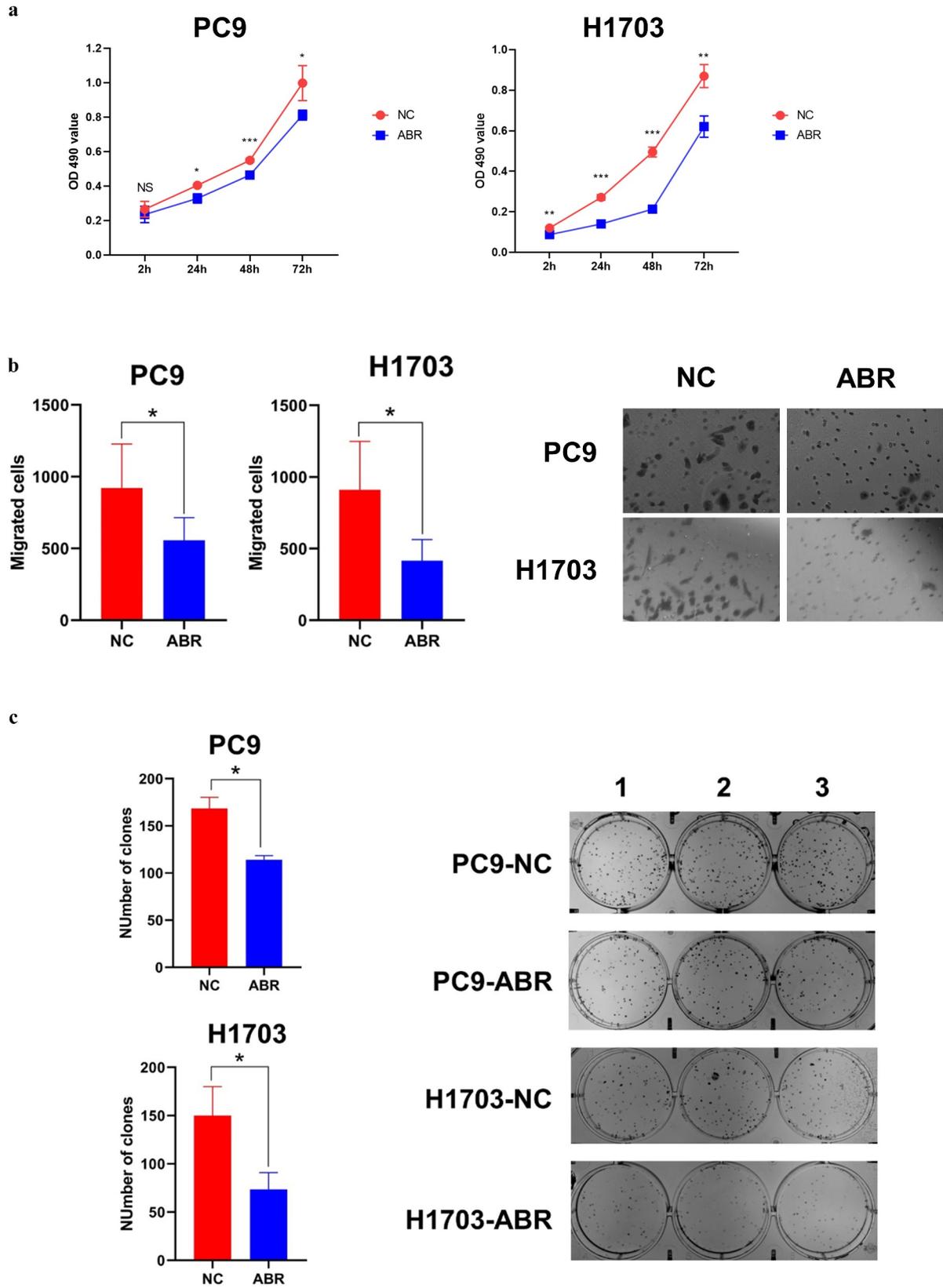


Figure 7. ABR shows tumor suppression effect in proliferation, migration and cloning ability in PC9 and H1703 cells. The cell proliferation experiment of PC9 and H1703 cell lines with overexpression of ABR (a). The migration capacity of PC9 and H1703 cells under ABR overexpression is shown, and the right panel illustrates tumor cells that invaded the chamber (b). The cloning formation ability of PC9 and H1703 cells under ABR overexpression, and the right panel demonstrates the colonies formed in each well (c). NC: control group with transfection by empty vector. ABR: ABR overexpressed group. NS: non-significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

is important in regulating reactive oxygen species (ROS) production, migration, and chemotaxis of neutrophils [40,41]. Elevated expression of PREX1 has been associated with the development of melanoma, prostate cancer, and breast cancer [42–44]. DOCK2 expression was initially deemed to be restricted to hematopoietic cells [45]. Although it is predominantly expressed in lymphocytes and hematopoietic tissues, recent research has revealed the tumor-promoting role of DOCK2 in lymphoma and colorectal cancer [46–48]. Distinct from the identification of ABR, PREX1 and DOCK2 in non-cancer cells, DOCK4 identification was initially reported in osteosarcoma cells, in which DOCK4 was deleted during tumor progression [49]. However, DOCK4 has also been reported to promote breast cancer development and is associated with bone metastasis [50,51]. To date, the roles of ABR, PREX1, DOCK2, and DOCK4 in lung cancer are largely unknown, and the underlying mechanisms remain to be explored.

Through the integration of data from multiple databases, our study demonstrated that ABR, PREX1, DOCK2, and DOCK4 are downregulated in NSCLC. In addition, NSCLC subgroups with higher nodal metastasis levels or cancer stages generally have lower median expression of ABR, PREX1, DOCK2, and DOCK4, whereas statistical significance for the comparison among these subgroups remains to be verified with a larger number of samples. Notably, the downregulation of ABR, PREX1, DOCK2, and DOCK4 are associated with poor overall survival of NSCLC patients, suggesting that these four novel Rho GEFs can serve as promising biomarkers for predicting the OS of patients. For the prognostic value for disease progression, only PREX1 and DOCK2 were statistically significant in both LUAD and LUSC patients. Although DOCK4 is not a candidate marker for predicting the disease progression in LUSC patients, its higher expression is markedly correlated with delayed tumor progression in LUAD patients. Hence, ABR, PREX1, DOCK2, and DOCK4 downregulation are characteristic of NSCLC, and they are promising for predicting the prognosis of NSCLC patients.

Gene expression and repression within cancer cells can be controlled by the epigenetic mechanisms of DNA methylation, which is the interaction

between genes and phenotypes without causing mutations in the DNA sequence [52,53]. DNA methylation involves the covalent addition of methyl groups to the C-5 position of cytosine rings, especially in a CpG dinucleotide [54]. In mammals, approximately 70% of the promoters are rich in unmethylated CpG [55]. Hypermethylation of CpG sites in promoters is typically associated with gene silencing at the transcriptional levels [56]. In lung cancer, DNA hypermethylation of tumor suppressors represents a hallmark and an early event in tumorigenesis [57]. Our study revealed that NSCLC samples from TCGA database contain higher methylation levels in the promoters of ABR, PREX1, DOCK2 and DOCK4 genes compared to the normal samples. Moreover, in NSCLC specimens, elevated levels of promoter methylation were markedly associated with lower expression of ABR, PREX1, DOCK2, and DOCK4. Thus, DNA hypermethylation might contribute to the downregulation of ABR, PREX1, DOCK2, and DOCK4 in NSCLC, and the methylation profiles of the four key Rho GEFs may be novel biomarkers for lung cancer screening.

In addition to epigenetic alterations, gene mutations can also affect gene expression levels [58]. The majority of mutations for ABR, PREX1, DOCK2, and DOCK4 in NSCLC are missense and truncating mutations, which means a change of a single amino acid into another and a change in the DNA that can shorten the protein, respectively. Our results illustrate that the expression of ABR, PREX1, DOCK2, and DOCK4 in NSCLC are not correlated with mutations. It should be noted that loss-of-function or gain-of-function mutations can lead to potential inhibitory or tumorigenic effects [59]. From our analysis, a few mutations occurred in the RHOGEF (DH), PH, DOCK-C2 (DHR-1) and DHR-2 domains, indicating that the critical functions of Rho GEF can be altered by the mutations, and this is yet to be confirmed.

Mechanistically, the downregulation of ABR, PREX1, DOCK2 and DOCK4 might upregulate MYC signaling and DNA repair pathways, as identified by GSEA using TCGA data. The MYC oncogene encodes a transcription factor that triggers gene expression in cancer cells [60]. MYC

signaling is implicated in the pathogenesis of most human cancers, and its deregulation is correlated with poor survival of patients [61]. MYC activation is associated with many features of cancer, including protein synthesis, proliferation and altered cellular pathways [61,62]. MYC upregulation is detected in >40% of NSCLC cases and is related to the loss of cell differentiation and tumor progression [63,64]. Therefore, ABR, PREX1, DOCK2, and DOCK4 downregulation could lead to the development of NSCLC via upregulation of the MYC protein expression and its downstream targets.

In clinical practice, chemotherapy and radiotherapy are the gold standards for the treatment of patients with lung cancer, and they have largely prolonged the survival of patients [65]. The therapeutic effects of platinum-based chemotherapeutic drugs and ionizing radiation are mediated by DNA damage [66–68]. Nevertheless, enhanced DNA repair mechanisms counteract the therapeutic benefits to patients, thereby leading to poor patient survival. Hence, the downregulation of ABR, PREX1, DOCK2, and DOCK4 promotes cancer development and leads to a poor prognosis by activating MYC and DNA repair signaling pathways in NSCLC, and further *in vitro* and *in vivo* studies are necessary to further confirm their association. Moreover, agents targeting DNA damage repair mechanisms have shown promise in NSCLC clinical models [67]. These four key Rho GEFs are also promising as biomarkers for predicting the response to DNA-damage regimens.

In spite of the conventional perspective that Rho GEFs are activators of Rho GTPases, based on comprehensive bioinformatic analysis and *in vitro* validation, our study unexpectedly but reasonably revealed that the Rho GEF ABR is a tumor suppressor in NSCLC. In addition to the above findings, it should be noted that the following evidence supports our results. ABR contains a Rho GAP domain in addition to its Rho GEF domain, functioning as a dual Rho GEF/GAP [69]. In addition, the prominent expression of ABR suggests its correlation with lymphocyte infiltration via cytokine secretion in NSCLC tissues. Tumor-infiltrating lymphocytes can attack and eliminate tumor cells, and therefore contribute to a better prognosis in patients with NSCLC [70,71].

For future studies, researchers should not simply regard all Rho GEFs as tumor promoters because of their involvement in activating Rho GTPases. Instead, the functions of Rho GEFs, other than the activation of Rho GEFs in cancers, should be noted.

5. Conclusion

In conclusion, our study identified that ABR, PREX1, DOCK2 and DOCK4 can serve as promising biomarkers for predicting the prognosis of NSCLC patients. The downregulation of ABR, PREX1, DOCK2 and DOCK4 in NSCLC can be induced by promoter methylation, and their methylation profiles might be potential indicators for lung cancer screening. The downregulation of the four key genes might enhance the tumorigenic MYC signaling and the DNA repair pathway. Our *in vitro* studies confirmed the role of ABR as a tumor suppressor in NSCLC cells. Future studies on methylation, regulation patterns and specific mechanisms are required.

Acknowledgements

The authors are grateful to Taylor & Francis Editing Services for English language editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Natural Science Foundation of China-Guangdong Joint Fund (U1601229) and the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds, pdjh2019a0182).

Author contributions

R-JZ, W-JX, C-WZ, L-YX, and E-ML conceptualized the study; R-JZ, W-JX, C-WZ, W-XC, S-MW, L-YX, and E-ML developed the methodology; W-JX carried out the experiments; R-JZ, C-WZ, W-XC, S-MW, LZ, C-BC, L-YX, and E-ML carried out investigation; R-JZ, W-JX, C-WZ and W-XC wrote the original draft of the manuscript; S-MW, LZ, C-BC, L-YX, and E-ML reviewed and edited the manuscript;

W-JX, L-YX, and E-ML acquired funding; L-YX, and E-ML provided resources; and L-YX and E-ML supervised the study.

Abbreviations

ABR, Active breakpoint cluster region-related; BCR, Breakpoint cluster region; CDC42, Cell division control protein 42 homolog; CI, Confidence interval; DH, Dbl-homology; DOCK, Dedicator of cytokinesis; DBL, Diffuse B-cell lymphoma; DHR, DOCK-homology region; EGA, European Genome-Phenome Archive; FDR, False discovery rate; FP, First progression; GDI, GDP-dissociation inhibitor; GEO, Gene Expression Omnibus; GEPIA, Gene Expression Profiling Interactive Analysis; GSEA, Gene Set Enrichment Analysis; GDC, Genomic Data Commons; GTEx, Genotype-Tissue Expression; GAP, GTPase-activating protein; GEF, Guanine nucleotide exchange factor; GDP, Guanosine diphosphate; GTPases, Guanosine triphosphatases; GTP, Guanosine triphosphate; HR, Hazard ratio; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; MSigDB, Molecular Signatures Database; NES, Normalized enrichment score; OS, Overall survival; PREX1, Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 1; PH, Pleckstrin-homology; RHOA, Ras homolog family member A; RAC1, Ras-related C3 botulinum toxin substrate 1; SH, Src homology; TCGA, The Cancer Genome Atlas.

References

- [1] de Groot PM, Wu CC, Carter BW, et al. The epidemiology of lung cancer. *Transl Lung Cancer Res.* **2018**;7(3):220–233.
- [2] Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Health.* **2019**;85. DOI:10.5334/aogh.2352
- [3] Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res.* **2016**;5(3):288–300.
- [4] Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature.* **2018**;553(7689):446–454.
- [5] Etienne-Manneville S, Hall A. Rho gtpases in cell biology. *Nature.* **2002**;420(6916):629–635.
- [6] Haga RB, Ridley AJ. Rho gtpases: Regulation and roles in cancer cell biology. *Small GTPases.* **2016**;7(4):207–221.
- [7] Kumawat A, Chakrabarty S, Kulkarni K. Nucleotide dependent switching in rho gtpase: Conformational heterogeneity and competing molecular interactions. *Sci Rep.* **2017**;7(1):1–11.
- [8] Zou T, Mao X, Yin J, et al. Emerging roles of RAC1 in treating lung cancer patients. *Clin Genet.* **2017**;91(4):520–528.
- [9] Konstantinidou G, Ramadori G, Torti F, et al. Rhoa-fak is a required signaling axis for the maintenance of kras-driven lung adenocarcinomas. *Cancer Discov.* **2013**;3(4):444–457.
- [10] Xiao X-H, Lv L-C, Duan J, et al. Regulating cdc42 and its signaling pathways in cancer: Small molecules and microrna as new treatment candidates. *Molecules.* **2018**;23(4):787.
- [11] Porter AP, Papaioannou A, Malliri A. Deregulation of rho gtpases in cancer. *Small GTPases.* **2016**;7(3):123–138.
- [12] Sosa MS, Lopez-Haber C, Yang C, et al. Identification of the rac-gef p-rex1 as an essential mediator of erbb signaling in breast cancer. *Mol Cell.* **2010**;40(6):877–892.
- [13] Cardama GA, González N, Maggio J, et al. Rho gtpases as therapeutic targets in cancer. *Int J Oncol.* **2017**;51(4):1025–1034.
- [14] Lawson CD, Ridley AJ. Rho gtpase signaling complexes in cell migration and invasion. *J Cell Biol.* **2018**;217(2):447–457.
- [15] Cook DR, Rossman KL, Der CJ. Rho guanine nucleotide exchange factors: Regulators of rho gtpase activity in development and disease. *Oncogene.* **2014**;33(31):4021–4035.
- [16] Lazer G, Katzav S. Guanine nucleotide exchange factors for rho gtpases: Good therapeutic targets for cancer therapy? *Cell Signal.* **2011**;23(6):969–979.
- [17] Kumar S, Stecher G, Tamura K. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* **2016**;33(7):1870–1874.
- [18] He Z, Zhang H, Gao S, et al. Evolvew v2: An online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* **2016**;44(W1):W236–W241.
- [19] Szklarczyk D, Gable AL, Lyon D, et al. String v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**;47(D1):D607–d613.
- [20] Rhodes DR, Kalyana-Sundaram S, Mahavisno V, et al. OncoPrint 3.0: Genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia (New York, NY).* **2007**;9(2):166.
- [21] Rhodes DR, Yu J, Shanker K, et al. OncoPrint: a cancer microarray database and integrated data-mining platform. *Neoplasia (New York, NY).* **2004**;6(1):1.
- [22] Tang Z, Li C, Kang B, et al. Gepia: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **2017**;45(W1):W98–W102.
- [23] Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. Ualcan: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia.* **2017**;19(8):649–658.
- [24] Györfy B, Surowiak P, Budczies J, et al. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One.* **2013**;8(12):e82241–e82241.
- [25] Li Q, Birkbak NJ, Györfy B, et al. Jcset: Selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics.* **2011**;12(1):474–474.

- [26] Györfy B, Benke Z, Lánckzy A, et al. Recurrenceonline: An online analysis tool to determine breast cancer recurrence and hormone receptor status using microarray data. *Breast Cancer Res Treat.* 2012;132(3):1025–1034.
- [27] Koch A, De Meyer T, Jeschke J, et al. Mexpress: Visualizing expression, DNA methylation and clinical tcga data. *BMC Genomics.* 2015;16(1):636.
- [28] Koch A, Jeschke J, Van Criekinge W, et al. Mexpress update 2019. *Nucleic Acids Res.* 2019;47(W1):W561–W565.
- [29] Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cbiportal. *Sci Signal.* 2013;6(269):pl1–pl1.
- [30] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Nat Acad Sci.* 2005;102(43):15545–15550.
- [31] Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (msigdb) 3.0. *Bioinformatics.* 2011;27(12):1739–1740.
- [32] Faruki H, Mayhew GM, Serody JS, et al. Lung adenocarcinoma and squamous cell carcinoma gene expression subtypes demonstrate significant differences in tumor immune landscape. *J Thorac Oncol.* 2017;12(6):943–953.
- [33] Brägelmann J, Böhm S, Guthrie MR, et al. Family matters: How myc family oncogenes impact small cell lung cancer. *Cell Cycle.* 2017;16(16):1489–1498.
- [34] Soucek L, Whitfield JR, Sodik NM, et al. Inhibition of myc family proteins eradicates kras-driven lung cancer in mice. *Genes Dev.* 2013;27(5):504–513.
- [35] Yang S, Liu Y, Li M-Y, et al. Foxp3 promotes tumor growth and metastasis by activating wnt/ β -catenin signaling pathway and emt in non-small cell lung cancer. *Mol Cancer.* 2017;16(1):124–124. DOI:10.1186/s12943-017-0700-1
- [36] Tan EC, Leung T, Manser E, et al. The human active breakpoint cluster region-related gene encodes a brain protein with homology to guanine nucleotide exchange proteins and gtpase-activating proteins. *J Biol Chem.* 1993;268(36):27291–27298.
- [37] Cunnick JM, Schmidhuber S, Chen G, et al. Bcr and abr cooperate in negatively regulating acute inflammatory responses. *Mol Cell Biol.* 2009;29(21):5742–5750.
- [38] Namasu CY, Katzerke C, Bräuer-Hartmann D, et al. Abr, a novel inducer of transcription factor *c/ebp α* , contributes to myeloid differentiation and is a favorable prognostic factor in acute myeloid leukemia. *Oncotarget.* 2017;8(61):103626.
- [39] Welch HC, Coadwell WJ, Ellson CD, et al. P-rax1, a ptdins (3, 4, 5) p3-and g $\beta\gamma$ -regulated guanine-nucleotide exchange factor for rac. *Cell.* 2002;108(6):809–821.
- [40] Welch HCE, Condliffe AM, Milne LJ, et al. P-rax1 regulates neutrophil function. *Curr Biol.* 2005;15(20):1867–1873.
- [41] Cervantes-Villagrana RD, Adame-García SR, García-Jiménez I, et al. G $\beta\gamma$ signaling to the chemotactic effector p-rax1 and mammalian cell migration is directly regulated by g α q and ga13 proteins. *J Biol Chem.* 2019;294(2):531–546.
- [42] Lindsay CR, Lawn S, Campbell AD, et al. P-rax1 is required for efficient melanoblast migration and melanoma metastasis. *Nat Commun.* 2011;2(1):1–9.
- [43] Qin J, Xie Y, Wang B, et al. Upregulation of pip3-dependent rac exchanger 1 (p-rax1) promotes prostate cancer metastasis. *Oncogene.* 2009;28(16):1853–1863.
- [44] Dillon LM, Bean JR, Yang W, et al. P-rax1 creates a positive feedback loop to activate growth factor receptor, pi3k/akt and mek/erk signaling in breast cancer. *Oncogene.* 2015;34(30):3968–3976.
- [45] Nishihara H, Kobayashi S, Hashimoto Y, et al. Non-adherent cell-specific expression of dock2, a member of the human cdm-family proteins. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research.* 1999;1452(2):179–187.
- [46] Gadea G, Blangy A. Dock-family exchange factors in cell migration and disease. *Eur J Cell Biol.* 2014;93(10–12):466–477.
- [47] Miao S, Zhang RY, Wang W, et al. Overexpression of dedicator of cytokinesis 2 correlates with good prognosis in colorectal cancer associated with more prominent CD8 + lymphocytes infiltration: a colorectal cancer analysis. *J Cell Biochem.* 2018;119(11):8962–8970.
- [48] Wang L, Nishihara H, Kimura T, et al. Dock2 regulates cell proliferation through rac and erk activation in b cell lymphoma. *Biochem Biophys Res Commun.* 2010;395(1):111–115.
- [49] Yajnik V, Paulding C, Sordella R, et al. Dock4, a gtpase activator, is disrupted during tumorigenesis. *Cell.* 2003;112(5):673–684.
- [50] Kobayashi M, Harada K, Negishi M, et al. Dock4 forms a complex with sh3yl1 and regulates cancer cell migration. *Cell Signal.* 2014;26(5):1082–1088.
- [51] Westbrook JA, Wood SL, Cairns DA, et al. Identification and validation of dock4 as a potential biomarker for risk of bone metastasis development in patients with early breast cancer. *J Pathol.* 2019;247(3):381–391.
- [52] Pfeifer GP. Defining driver DNA methylation changes in human cancer. *Int J Mol Sci.* 2018;19(4):1166.
- [53] Li L, Li C, Mao H, et al. Epigenetic inactivation of the cpg demethylase tet1 as a DNA methylation feedback loop in human cancers. *Sci Rep.* 2016;6(1):26591–26591. DOI:10.1038/srep26591
- [54] Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer.* 2011;2(6):607–617.
- [55] Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of cpg dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A.* 2006;103(5):1412–1417.

- [56] Belinsky SA. Silencing of genes by promoter hypermethylation: Key event in rodent and human lung cancer. *Carcinogenesis*. 2005;26(9):1481–1487.
- [57] Langevin SM, Kratzke RA, Kelsey KT. Epigenetics of lung cancer. *Transl Res*. 2015;165(1):74–90.
- [58] Zhang Z, Miteva MA, Wang L, et al. Analyzing effects of naturally occurring missense mutations. *Comput Math Methods Med*. 2012;2012:805827–805827.
- [59] Kim KH, Roberts CWM. Targeting ezh2 in cancer. *Nat Med*. 2016;22(2):128–134.
- [60] Stine ZE, Walton ZE, Altman BJ, et al. Myc, metabolism, and cancer. *Cancer Discov*. 2015;5(10):1024–1039.
- [61] Chen H, Liu H, Qing G. Targeting oncogenic myc as a strategy for cancer treatment. *Signal Transduct Target Ther*. 2018;3(1):5–5.
- [62] Gabay M, Li Y, Felsher DW. Myc activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb Perspect Med*. 2014;4(6):a014241.
- [63] Lorenz J, Friedberg T, Paulus R, et al. Oncogene overexpression in non-small-cell lung cancer tissue prevalence and clinicopathological significance. *Clin Investig*. 1994;72(2):156–163.
- [64] Mustachio LM, Roszik J, Farria AT, et al. Repression of gcn5 expression or activity attenuates c-myc expression in non-small cell lung cancer. *Am J Cancer Res*. 2019;9(8):1830–1845.
- [65] Lemjabbar-Alaoui H, Hassan OU, Yang YW, et al. Lung cancer: Biology and treatment options. *Biochim Biophys Acta*. 2015;1856(2):189–210.
- [66] Woods D, Turchi JJ. Chemotherapy induced DNA damage response: convergence of drugs and pathways. *Cancer Biol Ther*. 2013;14(5):379–389.
- [67] Sen T, Gay CM, Byers LA. Targeting DNA damage repair in small cell lung cancer and the biomarker landscape. *Transl Lung Cancer Res*. 2018;7(1):50–68.
- [68] Butkiewicz D, Drosik A, Suwiński R, et al. Influence of DNA repair gene polymorphisms on prognosis in inoperable non-small cell lung cancer patients treated with radiotherapy and platinum-based chemotherapy. *Int J Cancer*. 2012;131(7):E1100–E1108.
- [69] Vaughan EM, Miller AL, Yu H-YE, et al. Control of local rho gtpase crosstalk by abr. *Curr Biol*. 2011;21:270–277.
- [70] Yu P, Fu Y-X. Tumor-infiltrating t lymphocytes: friends or foes? *Lab Invest*. 2006;86(3):231–245.
- [71] Geng Y, Shao Y, He W, et al. Prognostic role of tumor-infiltrating lymphocytes in lung cancer: a meta-analysis. *Cell Physiol Biochem*. 2015;37(4):1560–1571.