

RESEARCH

Open Access



Identification of HMGB2 associated with proliferation, invasion and prognosis in lung adenocarcinoma via weighted gene co-expression network analysis

Xie Qiu^{1†}, Wei Liu^{2†}, Yifan Zheng^{3†}, Kai Zeng⁴, Hao Wang⁵, Haijun Sun^{1*} and Jianhua Dai^{1*}

Abstract

Background: High mobility group protein B2 (HMGB2) is a multifunctional protein that plays various roles in different cellular compartments. Moreover, HMGB2 serves as a potential prognostic biomarker and therapeutic target for lung adenocarcinoma (LUAD).

Methods: In this study, the expression pattern, prognostic implication, and potential role of *HMGB2* in LUAD were evaluated using the integrated bioinformatics analyses based on public available mRNA expression profiles from The Cancer Genome Atlas and Gene Expression Omnibus databases, both at the single-cell level and the tissue level. Further study in the patient-derived samples was conducted to explore the correlation between HMGB2 protein expression levels with tissue specificity, (tumor size-lymph node-metastasis) TNM stage, pathological grade, Ki-67 status, and overall survival. In vitro experiments, such as CCK-8, colony-formation and Transwell assay, were performed with human LUAD cell line A549 to investigate the role of HMGB2 in LUAD progression. Furthermore, xenograft tumor model was generated with A549 in nude mice.

Results: The results showed that the HMGB2 expression was higher in the LUAD samples than in the adjacent normal tissues and was correlated with high degree of malignancy in different public data in this study. Besides, over-expression of HMGB2 promoted A549 cells proliferation and migration while knocking down of HMGB2 suppressed the tumor promoting effect.

Conclusions: Our study indicated that HMGB2 was remarkably highly expressed in LUAD tissues, suggesting that it is a promising diagnostic and therapeutic marker for LUAD in the future.

Keywords: HMGB2, LUAD, Biomarker, Bioinformatics, WG

Introduction

Lung cancer is the most commonly diagnosed cancer worldwide, accounting for the greatest number of deaths among those caused by cancer [1]. The most common histological subtype of lung cancer is lung adenocarcinoma (LUAD), which accounts for more than 40% of the total lung cancer cases [2]. Although LUAD can be diagnosed at an early stage using computerized tomography, many patients with non-small cell lung cancer (NSCLC)

[†]Xie Qiu, Wei Liu and Yifan Zheng contributed equally

*Correspondence: lygthorax@hotmail.com; daijhxw@163.com

¹ Department of Cardiothoracic Surgery, The First People's Hospital of Lianyungang, Lianyungang, People's Republic of China
Full list of author information is available at the end of the article



are diagnosed at an advanced stage. In recent decades, therapies, including surgery, chemotherapy, radiotherapy, molecular-target therapy, and immune therapy, have been used to improve the survival of patients with LUAD [3]. However, the prognosis of patients with advanced-stage LUAD remains grim [4, 5]. The discovery of aberrant gene expression may contribute to the early screening of patients with lung cancer and may increase the percentage of cancer diagnosis in patients at an early stage [6, 7]. Therefore, it is important to identify a novel molecular target to facilitate the early diagnosis and treatment of patients with lung cancer.

The HMGB2 protein, which belongs to the high-mobility group box (HMGB) family, which plays an essential role in transcription, chromatin remodeling, and other processes by binding to single-stranded DNA [8, 9]. *HMGB2* promotes the progression of breast cancer by targeting lactate dehydrogenase B and febrile convulsions 1 proteins [10]. In gastric cancer, *HMGB2* indicates a poor prognosis [11, 12] and is involved in tumor progression along with microRNAs, long non-coding RNAs, and proteins [13–15]. *HMGB2* is also considered as an oncogene in NSCLC and is involved in the chemotherapeutic drug resistance of NSCLC [16–18]. However, the involvement of HMGB2 in LUAD and the proliferation and invasion of LUAD cells have yet to be thoroughly investigated.

Therefore, in this study, a bioinformatics approach was employed to identify the HMGB2 as a potential diagnostic and prognostic marker for LUAD. Subsequently, the expression analysis of the HMGB2 gene in a cohort of patients with was analyzed to investigate the correlation between the *HMGB2* gene expression and the clinical characteristic of patients with LUAD.

Methods

Data acquisition and processing

The gene expression profiles and invasion scores of the single LUAD dataset EXP0068 were downloaded from the CancerSEA database (<http://biocc.hrbmu.edu.cn/CancerSEA/home.jsp>) [19, 20]. Then, the datasets GSE10072 [21], GSE21933 [22], and GSE32863 [23] were downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The TCGA-LUAD dataset was downloaded from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>), and the unit of gene expression was converted from count to transcripts per kilobase million (TPM). The proteomic data of patients with LUAD (PDC000219) was downloaded from The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) (<https://pdc.cancer.gov>) [24]. The signature gene list for the poor survival of LUAD was downloaded from the

MSigDB database [25, 26]. The signature gene list for the invasion was downloaded from the CancerSEA database. Survival data and invasion scores for each of the LUAD samples in TCGA-LUAD, GSE10072, GSE21933, and GSE32863 datasets were quantified using single-sample gene set enrichment analysis (ssGSEA) method with corresponding signature gene patterns using the R package “GSVA”.

Weighted correlation network analysis (WGCNA)

Weighted correlation network analysis (WGCNA) was performed using the R package “WGCNA” with the protein-coding genes (PCGs) expression profiles of a single LUAD cell. In this study, the following parameters were used; soft-threshold $\beta=5$, min module size=30, and threshold to merge similar modules=0.25. Pearson's correlation analysis was performed to identify the features related modules (coefficient > 0.5, $P < 0.05$). Hub genes in the features-related module were identified according to the Pearson's correlation between the gene expressions profiles and feature scores (coefficient > 0.3, $P < 0.05$) and module (coefficient > 0.8, $P < 0.05$).

Protein–protein interaction (PPI) analysis

STRING database (<https://www.string-db.org/>) [27] and Cytoscape software were used to construct the protein–protein interaction (PPI) network according to the instructions provided on their official websites.

Gene ontology (GO) analysis

Gene Ontology (GO) analysis was performed using “ClueGo” in Cytoscape software or Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 with default parameters.

Survival analysis

Kaplan–Meier analysis of patients with LUAD based on the HMGB2 expression was performed using Gene Expression Profiling Interactive Analysis 2.0 (GEPIA 2.0) database (<http://gepia2.cancer-pku.cn/#index>) [28] and PrognScan database (<http://dna00.bio.kyutech.ac.jp/PrognScan/index.html>) [29].

Functional state analysis

Patients with LUAD were divided into two groups based on the median expression of HMGB2, and GSEA analysis was performed on the data from TCGA-LUAD and GSE10072 datasets. The functional state analyses of HMGB2 in the LUAD samples at a single cell level were performed using the CancerSEA database. The TISIDB (<http://cis.hku.hk/TISIDB/index.php>) database [30] was used to evaluate the expression profiles of the *HMGB2*

gene among the different immune subtypes of LUAD (from C1 to C6).

Patient specimens and tissue microarray (TMA) preparation

Formalin-fixed paraffin-embedded tissue blocks, including LUAD and adjacent normal lung tissues, were collected from the The First People's Hospital of Lianyungang. The samples, collected from 2010 to 2015, were obtained from the patients with primary LUAD, who received no chemotherapy or radiotherapy. For the tissue microarray (TMA) construction, all the specimens were re-evaluated and checked by hematoxylin and eosin (HE) staining and the representative areas were selected and prepared into 1.5-mm-thick tissue cores. In this study, a total of 98 LUAD samples with adjacent normal tissues were analyzed. The clinical information of patients, including age, tumor size, Ki-67 status, lymph node status, TNM stage, pathologic grades, and follow-up information for calculating overall survival (OS) rates, were retrieved from the patients' electronic medical records (Additional file 1: Table S1). Clinicopathological classification and staging were performed according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system. The study was reviewed and approved by the Ethics Committee of The First People's Hospital of Lianyungang. The approved number is KY-20190927005. In this study, written informed consent has been obtained from each subject and that all experiments conform to the Declaration of Helsinki.

Immunohistochemistry (IHC)

For immunohistochemistry (IHC) analysis, 3 μ m-thick TMAs slides were dewaxed in xylene and rehydrated in graded ethanol solutions. Antigens were retrieved using the high-pressure heat method with a citrate solution (pH = 6). Then, the slides were incubated with goat serum in a 3% hydrogen peroxide solution for 15 min at room temperature. The samples were then incubated with HMGB2 monoclonal antibody (1:300, Abcam, ab124670) at 4 °C overnight, which was then followed by detection with a universal SP kit (mouse/rabbit streptavidin–biotin detection system, ZSBIO, Cat # SP-9000), following the manufacturer's instructions. The sections were then stained with 3,3-diaminobenzidine (DAB), counterstained with hematoxylin, dehydrated with a graded alcohol series, cleared in xylene, and mounted by neutral resins.

Interpretation and evaluation of IHC results

To analyze IHC expression results, the TMA slides were scanned under an Olympus optical microscope. The HMGB2-positive stains were mainly concentrated in the

nucleus; for expression analysis, both the nuclear-positive staining intensity and percentage of positive cells were graded and multiplied to obtain the overall staining score. Staining intensity was scored on a scale of 0–3 as follows; 0 (negative), 1 (weak), 2 (medium), and 3 (strong). The percentage of positive tumor cells was categorized into five semi-quantitative classes: 0 (\leq 5% positive cells), 1 (6–25% positive cells), 2 (26–50% positive cells), 3 (51–75% positive cells), and 4 ($>$ 76% positive cells). An overall staining score of $>$ 6 was defined as the high HMGB2 expression. The expression score of HMGB2 was analyzed by two independent experienced pathologists.

Cell culture

HMGB2-coding lentivirus vector and HMGB2 knocking-down shRNA vector combined with psPAX and Pmd2.0G were transfected into HEK293T cells. Lentivirus was collected to infect A549 cells after 48 h. All cells were cultured in DMEM medium (Gibco) with 10% fetal bovine serum (FBS) (Gibco) and 1% penicillin–streptomycin (Gibco) at 37 °C with 5% CO₂.

Western blotting

Cells were lysed by pre-cooled RIPA lysis buffer (Sigma) and cell proteins were extracted. The total protein concentration was measured by bicinchoninic acid protein (BCA) assay kit (Sigma) and 10 μ g total protein lysate per sample was separated via SDS-PAGE and transferred to PVDF membranes (Millipore). Membranes were blocked by skim milk for 1 h and then proteins were detected by incubating with primary antibodies HMGB2 (1:300, Abcam, ab124670) and HRP Goat Anti-Rabbit IgG (H + L) (Abcam, Cambridge, UK). Housekeep gene β -Actin (Proteintech, 20,536–1-AP, 1:5000) was used as a loading control. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7738851/-B25>.

Cell growth assay

Cells were plated into 96-well micro-plate (2000 cells/well, 3 parallel wells) at 37 °C and 5% CO₂. Then, the cells were collected at 0 h, 24 h, 36 h, 48 h, 72 h, and cell number was analyzed by using the CCK8 reagent (MCE, HY-K0301) according to manufacturer's instruction. The optical density (450 nm) was used to indicate the number of A549 cells.

Colony-forming assay

Dissociated cells were plated into 6-well plate (200 cells/well, 3 parallel wells). After 3 weeks, colonies were fixed with 4% paraformaldehyde for 30 min and then stained with 0.1% crystal violet solution. Each well was counted for colony-forming under a microscope.

Transwell assay

Appropriately 5×10^4 cells were plated in the top chamber of Transwell (Costar, Cambridge, MA, USA) in serum-free DMEM and DMEM containing 10% FBS was added to the lower chamber. After incubation for 36 h at 37 °C, migrated cells were fixed with 4% paraformaldehyde for 30 min and then stained with 0.1% crystal violet solution. After the non-migrated cells present on the upper surface were removed, each Transwell membrane was photographed and cells were counted.

Cell cycle assay

A549 cells were cultured in 6 wells plate for 12 h. Next, cells were digested with trypsin and suspended into single cell, and fixed with 70% ethanol at 4 °C overnight. Then cells were re-suspended and washed by PBS. 0.2 mg/mL RNase A and 100 µg/mL Propidium in PBS was applied for staining cells for 30 min. The percentages of cells of G0/G1, S, and G2/M were analyzed by counted cell numbers according to red fluorescence emission by Beckman Cytoflex.

Animal experiment

A total of 1×10^6 A549 cells infected with HMGB2-shRNA and Scramble-shRNA were injected into the subcutaneous fat tissue of nude mice aged 6 weeks purchased from Model Animal Research Center of Nanjing University. There were 5 mice in each group and all the operations in accordance with ARRIVE guidelines and animal healthcare of Nanjing Medical University discipline. Tumor volume was calculated using the formula $(\text{length} \times \text{width}^2)/2$. At the end of the experiment, the mice were sacrificed by Carbon dioxide gas anesthesia, and tumor masses were separated and recorded by photograph. The animal procedures were approved by the Nanjing Medical University Health Science Center Institutional Animal Care and Use Committee.

Statistical analysis

Two-tailed t test was utilized to analyze the difference between two groups. Log-rank test was utilized for survival analysis.

Results

Construction of WGCNA and identification of key genes

WGCNA was performed using the PCGs expression profile of a single LUAD dataset after quality control. A total of 6 modules were identified based on the average hierarchical clustering and dynamic tree clipping methods (Fig. 1A). The blue module was the most closely related to the invasion score and poor survival (Fig. 1B).

Similarly, the genes (*HMGB2*, *PTTG1*, *CENPF*, *NUSAP1*, and *TOP2A*) in the blue module, which were closely related to the invasion and poor survival, were identified among the hub genes (Fig. 1C–E).

PPI network and GO analysis of genes in the blue module

The blue module consists of 60 genes, which are listed in Table 1. The PPIs among these 60 genes were identified using the STRING database. A PPI network was constructed using “MCODE” in Cytoscape software, which consisted of 25 genes (*CDC20*, *TOP2A*, *NUSAP1*, *PTTG1*, *RRM2*, *TK1*, *PRC1*, *CENPE*, *CKS2*, *TYMS*, *HMGB2*, *KPNA2*, *CKAP2*, *CENPW*, *UBE2T*, *SMC4*, *TPX2*, *KIAA0101*, *MAD2L1*, *CCNB2*, *NUF2*, *CCNB1*, *BIRC5*, *CKS1B*, and *CDK1*) and illustrated (Fig. 2A). The GO analysis of 25 genes using ClueGo showed that these genes might participate in the segregation of sister chromatid, cyclin-dependent protein serine/threonine kinase activator activity, chromosome condensation, organization of microtubule cytoskeleton involved in mitosis (Fig. 2B).

mRNA expression of HMGB2 was higher in LUAD tissues than in the normal lung tissues

The mRNA expression of the *HMGB2* gene was compared between the LUAD and normal lung tissues using four public datasets. Both the TCGA-LUAD dataset (Fig. 3A) and three GEO datasets (GSE10072, GSE21933, and GSE32863) (Fig. 3B–D) showed that the mRNA expression of the *HMGB2* gene was higher in the LUAD tissues than in that in the normal lung tissues. Besides, the mRNA expression of the *HMGB2* gene was positively correlated with the invasion score of LUAD samples in the TCGA-LUAD, GSE21933, and GSE32863 datasets (Fig. 3E–G).

High mRNA expression of HMGB2 predicted poor prognosis of patients with LUAD

First, using the GEPIA online tool, the patients with LUAD with high *HMGB2* expression were found to exhibit shorter disease-free survival (DFS) than that those with low *HMGB2* expression (Fig. 4A). Next, the PrognScan database demonstrated that the patients with LUAD with high *HMGB2* expression had shorter relapse-free survival (RFS) and OS than that those with low *HMGB2* expression (Fig. 4B). Additionally, it was also found that the mRNA expression of the *HMGB2* gene was positively correlated with the poor survival of LUAD samples in the TCGA-LUAD, GSE10072, GSE21933, and GSE32863 datasets (Fig. 4C–F).

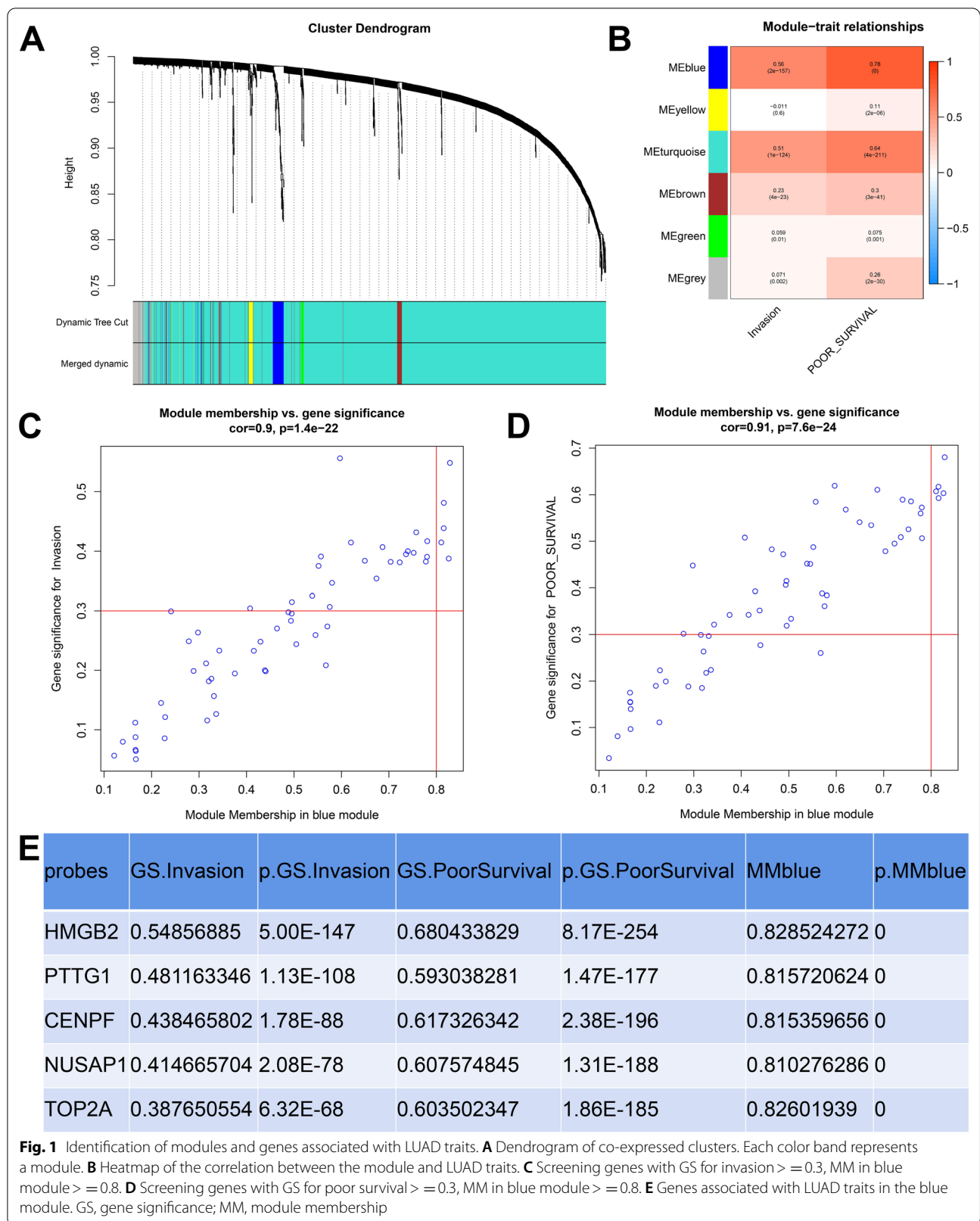


Table 1 List of genes in the blue module

Genes	GS.Invasion	p.GS.Invasion	GS.POOR_SURVIVAL	p.GS.POOR_SURVIVAL	MMblue	p.MMblue
HP1BP3	0.194489465	2.36E-17	0.342049134	2.48E-52	0.375415927	1.72E-63
HMG2	0.248015567	1.54E-27	0.392491809	9.85E-70	0.429403445	1.51E-84
H2AZ1	0.556301554	5.23E-152	0.619374063	5.2E-198	0.596681948	2.81E-180
TXNDC12	0.087574798	0.000152652	0.155133526	1.63E-11	0.166013358	5.41E-13
HMGB2	0.54856885	5E-147	0.680433829	8.17E-254	0.828524272	0
PTTG1	0.481163346	1.13E-108	0.593038281	1.47E-177	0.815720624	0
LMNA	0.085663156	0.000212437	0.111125441	1.5E-06	0.227541962	2.5E-23
CENPF	0.438465802	1.78E-88	0.617326342	2.38E-196	0.815359656	0
MAD2L1	0.43155391	1.81E-85	0.58602481	1.99E-172	0.75751139	0
CDC20	0.416711313	3.05E-79	0.5066494	3.58E-122	0.780645386	0
CNIH4	0.115705602	5.43E-07	0.185395548	6.97E-16	0.317357887	6.58E-45
LBR	0.295566173	6.5E-39	0.414605908	2.2E-78	0.495050243	6.9E-116
RRM2	0.270310779	1.37E-32	0.482672635	1.92E-109	0.464027189	3.2E-100
NUSAP1	0.414665704	2.08E-78	0.607574845	1.31E-188	0.810276286	0
DTYMK	0.126552657	4.17E-08	0.22417539	1.13E-22	0.335943941	1.97E-50
LSM3	0.181587012	2.73E-15	0.263558096	5.23E-31	0.320540308	7.93E-46
ANP32E	0.414532352	2.36E-78	0.568420259	4.41E-160	0.61992395	1.85E-198
SMC4	0.406941975	2.6E-75	0.61080064	3.85E-191	0.686571747	3.45E-260
PRC1	0.400037124	1.3E-72	0.589546434	5.46E-175	0.739849565	5.92878775009496e-323
CCNB1	0.397426228	1.32E-71	0.525596988	5.64E-133	0.752110706	0
BIRC5	0.394914249	1.2E-70	0.508903547	2.02E-123	0.736089122	4.45089906500295e-318
UBE2T	0.391091811	3.3E-69	0.585210459	7.7E-172	0.55671854	2.79E-152
HNRNPH1	0.079826757	0.000559412	0.081351724	0.000436945	0.139160814	1.59E-09
CANX	0.050491299	0.029225736	0.139841541	1.32E-09	0.166869243	4.09E-13
TPX2	0.390565821	5.2E-69	0.572889036	3.81E-163	0.780337799	0
LSM5	0.185672298	6.3E-16	0.217785631	1.85E-21	0.326003446	1.98E-47
H2AZ2	0.232757918	2.31E-24	0.342232892	2.17E-52	0.415424258	1.02E-78
TOP2A	0.387650554	6.32E-68	0.603502347	1.86E-185	0.82601939	0
RAD21	0.199951614	2.85E-18	0.351562846	2.23E-55	0.438644081	1.49E-88
CKAP2	0.383987637	1.41E-66	0.54116512	2.23E-142	0.649091595	1.46E-223
TUBB4B	0.243836908	1.2E-26	0.333798418	8.95E-50	0.504846153	3.53E-121
CCDC34	0.211582621	2.56E-20	0.299505525	5.85E-40	0.314867874	3.38E-44
MSRB2	-0.056602615	0.014495592	-0.034170132	0.1401855	-0.121155745	1.54E-07
NUF2	0.382601836	4.5E-66	0.560125366	1.61E-154	0.777739445	0
HNRNPH3	0.064298362	0.005473179	0.096709601	2.87E-05	0.166721726	4.29E-13
BUB3	0.297794076	1.67E-39	0.471913779	4.72E-104	0.488204343	2.73E-112
PTMS	0.145072376	3.09E-10	0.189829616	1.37E-16	0.220034372	6.98E-22
CKS1B	0.382129621	6.69E-66	0.478582912	2.27E-107	0.703647857	8.68E-279
CCNB2	0.381374568	1.26E-65	0.495160248	6.03E-116	0.722727963	3.24E-301
MZT1	0.283437277	8.55E-36	0.406075716	5.72E-75	0.493581645	4.14E-115
CENPW	0.375372785	1.78E-63	0.487446587	6.75E-112	0.551660618	5.29E-149
CDK1	0.354106101	3.28E-56	0.534635613	2.25E-138	0.67376856	4.57E-247
HMGB3	0.346876287	7.28E-54	0.38390452	1.51E-66	0.580098229	3.46E-168
TUBA1C	0.325429384	2.93E-47	0.451808378	1.76E-94	0.53847834	1.01E-140
ARL6IP1	0.208283525	1E-19	0.260398287	2.77E-30	0.567205804	2.94E-159
UBB	0.19887289	4.34E-18	0.188569008	2.18E-16	0.288429248	4.64E-37
CALM2	0.314837801	3.45E-44	0.318958225	2.28E-45	0.495383249	4.59E-116
SKA2	0.198230093	5.58E-18	0.277240152	2.93E-34	0.440151637	3.21E-89
DCAF7	0.066230214	0.004217493	0.154166518	2.18E-11	0.166071026	5.31E-13

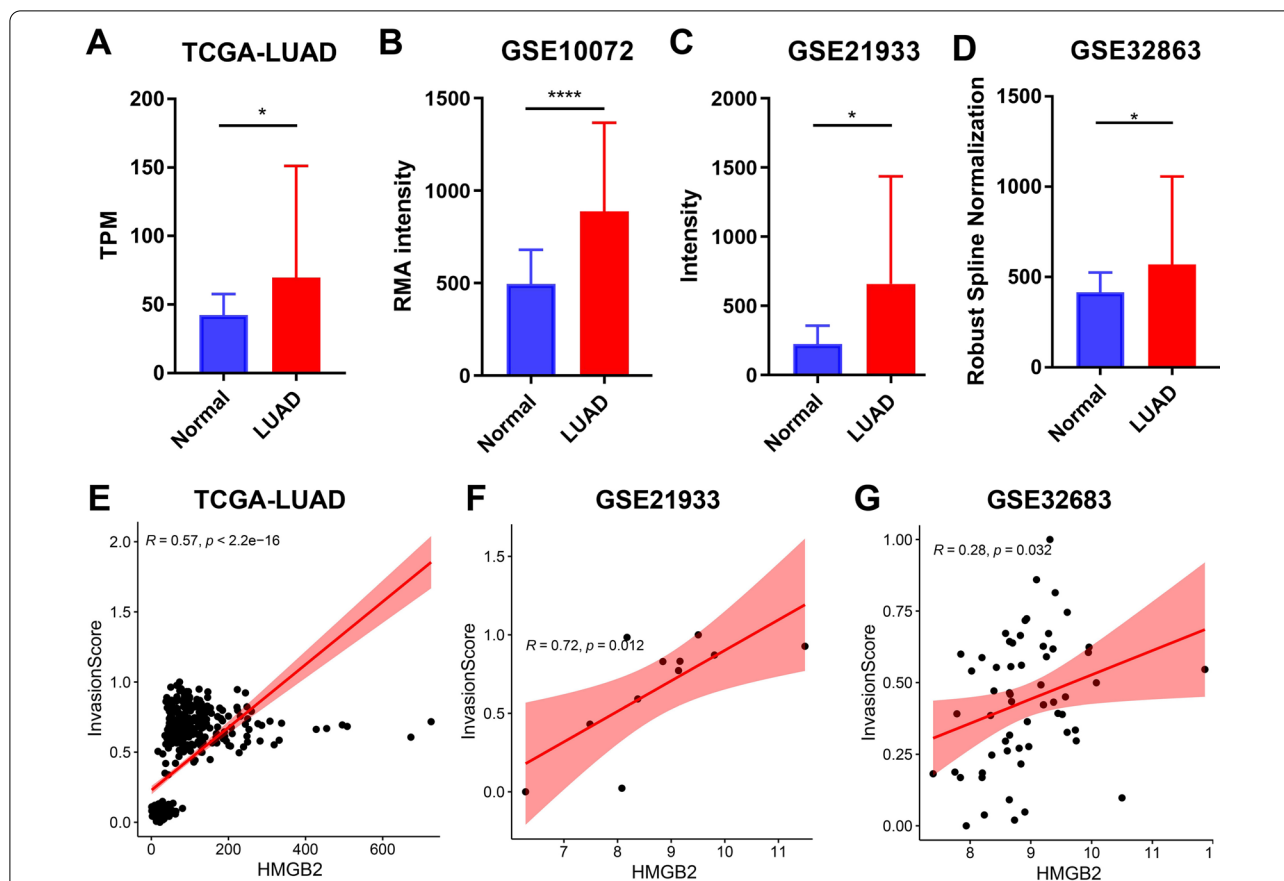


Fig. 3 High expression of *HMGB2* was correlated with invasion in LUAD. **A–D** *HMGB2* mRNA expression was higher in the LUAD than in normal lung samples in **A** TCGA-LUAD, **B** GSE10072, **C** GSE21933, **D** GSE32863 datasets. **E–G** *HMGB2* expression was positively correlated with the invasion score of LUAD samples in **E** TCGA-LUAD, **F** GSE21933, and **G** GSE32863 datasets. $P^* < 0.05$, $P^{**} < 0.01$, $P^{****} < 0.0001$

The expression level of HMGB2 was positively correlated with cell cycle and proliferation in the LUAD tissues

Pearson’s correlation analysis employing the expression patterns of single LUAD cells from two patients with LUAD patients (Pearson’s coefficient > 0.3, $P < 0.05$) revealed that *HMGB2* expression was positively correlated with cell cycle, proliferation, and invasion, as evidenced by the CancerSEA database (Fig. 6A and B). Furthermore, the GSEA analysis, based on the data from TCGA-LUAD and GSE10072 datasets, showed that both the cell cycle and DNA replication pathways were significantly enriched in the LUAD samples with high *HMGB2* expression (Fig. 6C–F).

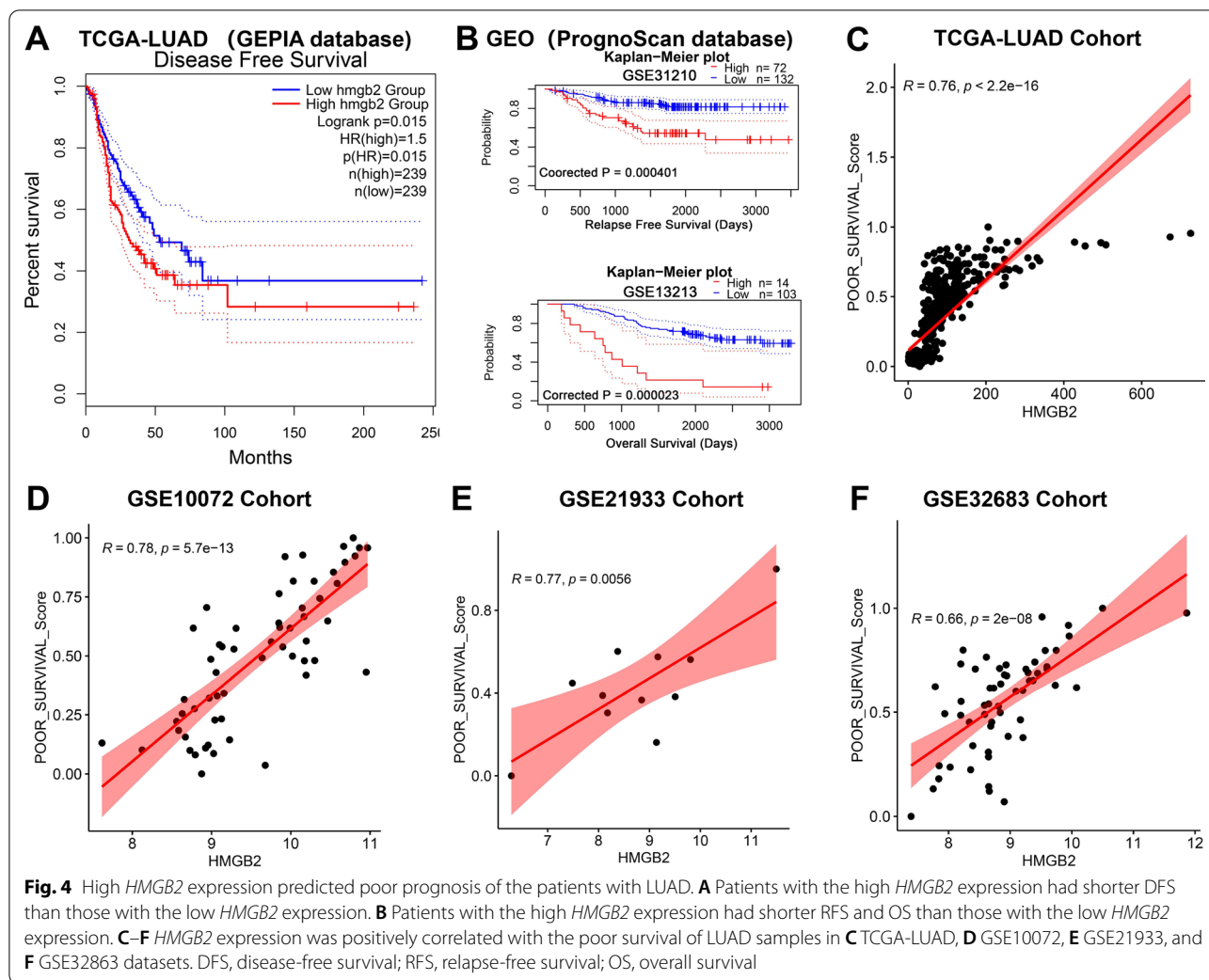
PPI network and GO analysis of HMGB2 and co-expressed genes

HMGB2 was used as the input as a single protein in the STRING database. The PPI network of *HMGB2* and its co-expressed genes consisted of 11 genes (*HMGB2*, *HMGB1*, *HIST1H1A*, *HIST1H1B*, *HIST1H1D*, *H1FO*, *SET*, *APEX1*, *ANP32A*, *NME1*, and *GZMA*), which are shown

in (Fig. 7A). The GO analysis of 11 genes using DAVID showed that these genes might participate in biological processes, such as nucleosome assembly, regulation of mRNA stability, positive regulation of DNA binding, and apoptotic DNA fragment (Fig. 7B); molecular function, such as like poly (A) RNA binding and chromatin DNA binding (Fig. 7C); and cellular component, such as nucleus and nucleoplasm (Fig. 7D).

HMGB2 was negatively related to inflammation in the LUAD tissues

As shown in Fig. 6A, the *HMGB2* expression was negatively correlated with the inflammation score in LUAD patients 3 in EXP0068 cohort at the single-cell level (coefficient = -0.50 and $P < 0.001$). Furthermore, it was found that the C3 immune type (inflammatory) of patients with LUAD had the lowest expression of *HMGB2* as compared to the other immune types (Fig. 8A). Results of Pearson’s correlation analysis showed that the expression of *HMGB2* was negatively correlated with that of chemokines (*CXCL16*, *CX3CR1*, and *CCL14*) (Fig. 8B–D)



and immunostimulatory proteins (*TNFSF13*, *TMEM173*, *IL6R*, and *TNFSF15*) (Fig. 8E–H).

Overexpression of HMGB2 promoted A549 cells proliferation and migration

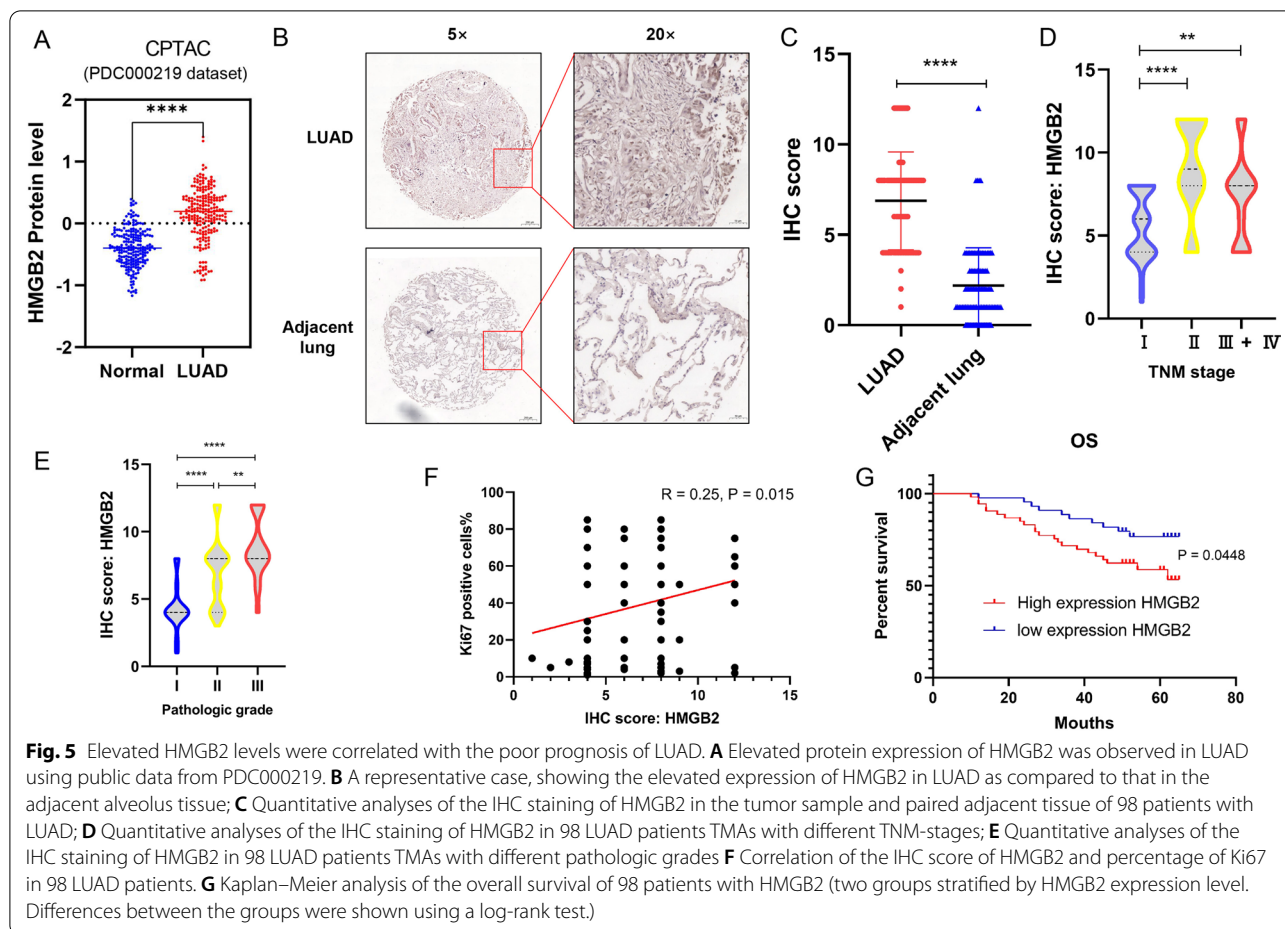
HMGB2 stable expression A549 cell line was constructed and validated as shown in Fig. 9A and Additional file 2; Figure S1. The in vitro experiments’ results showed that over-expression of HMGB2 promoted A549 cells proliferation (Fig. 9B) and colony-formation (Fig. 9C and D). Besides, Transwell assay’ result showed that HMGB2 may promote the ability of A549 cells to migrate (Fig. 9E and F).

Knockdown of HMGB2 inhibited A549 cells proliferation, tumorigenicity and migration

HMGB2 knockdown A549 cell line was constructed and validated as shown in Fig. 10A. The colony assays showed that knockdown of HMGB2 inhibited the colony-formation of A549 cells (Fig. 10B). The in vivo experiments showed that knockdown of HMGB2 inhibited the volumes of the tumors (Fig. 10C). Cell cycle assay showed that the knockdown of HMGB2 increased the number of cells in the G0/G1 phase and decreased the number of cells in the G2/M phase (Fig. 10D). Besides, we observed that knockdown of HMGB2 could inhibited the migration of A549 cells (Fig. 10E).

Discussion

Extensive efforts have been devoted to exploring the promising diagnostic and therapeutic target for improving the prognosis of lung cancer. In this study, we performed WGCNA and ssGSEA analyses with the



expression profiles of patients with LUAD at a single-cell level and identified the *HMGB2* gene was identified as a promising diagnostic and prognostic biomarker for the LUAD (Additional file 2; Figure S2). Furthermore, using an integrated bioinformatics analysis on multiple expression profiles from TCGA and GEO datasets, we found that the *HMGB2* might affect the prognosis of patients with LUAD patients by regulating the proliferation and invasion of LUAD cells.

HMGB2, a member of the family of high mobility group nonhistone chromatin proteins, regulates the processes of transcription, replication, recombination, and DNA repair [31]. *HMGB2* is highly expressed during embryogenesis, however, its expression is limited

in the adult organs and is mainly detected in lymphoid organs and testes. However, previous studies have elaborated on the elevated expression of *HMGB2* in several types of tumor tissues and reported it as an oncogene. In breast cancer, the *HMGB2* is regulated with *ER*, *LDHB*, and *FBP1* to promote the endocrine therapy resistance and tumorigenesis of tumor cells [10, 32, 33]. In gastric cancer, the high expression of *HMGB2* predicts a poor prognosis [11, 34] and its expression is regulated by non-coding RNA, *miRNA-23b-3p*, *miRNA-1297*, *MALAT1*, and *miRNA-873* to promote the proliferation, migration, and invasion of cells [13–15]. In prostate cancer, the early detection of *HMGB2* in prostate tissues using IHC contributed to the early-stage diagnosis of prostate

(See figure on next page.)

Fig. 6 *HMGB2* was correlated with cell cycle and proliferation in LUAD samples. **A** Correlations between the functional states and *HMGB2* expression in patient 3 (LUAD) at the single-cell level. **B** Correlations between the functional states and *HMGB2* expression in patient 4 (LUAD) at the single-cell level. **C, D** GSEA for *HMGB2* using the data from the TCGA-LUAD dataset. **C** Cell cycle and **D** DNA replication pathways were enriched in the high expression group. **E** and **F** GSEA for *HMGB2* using the data from the GSE10072 dataset. **E** Cell cycle and **F** DNA replication pathways were enriched in the high expression group

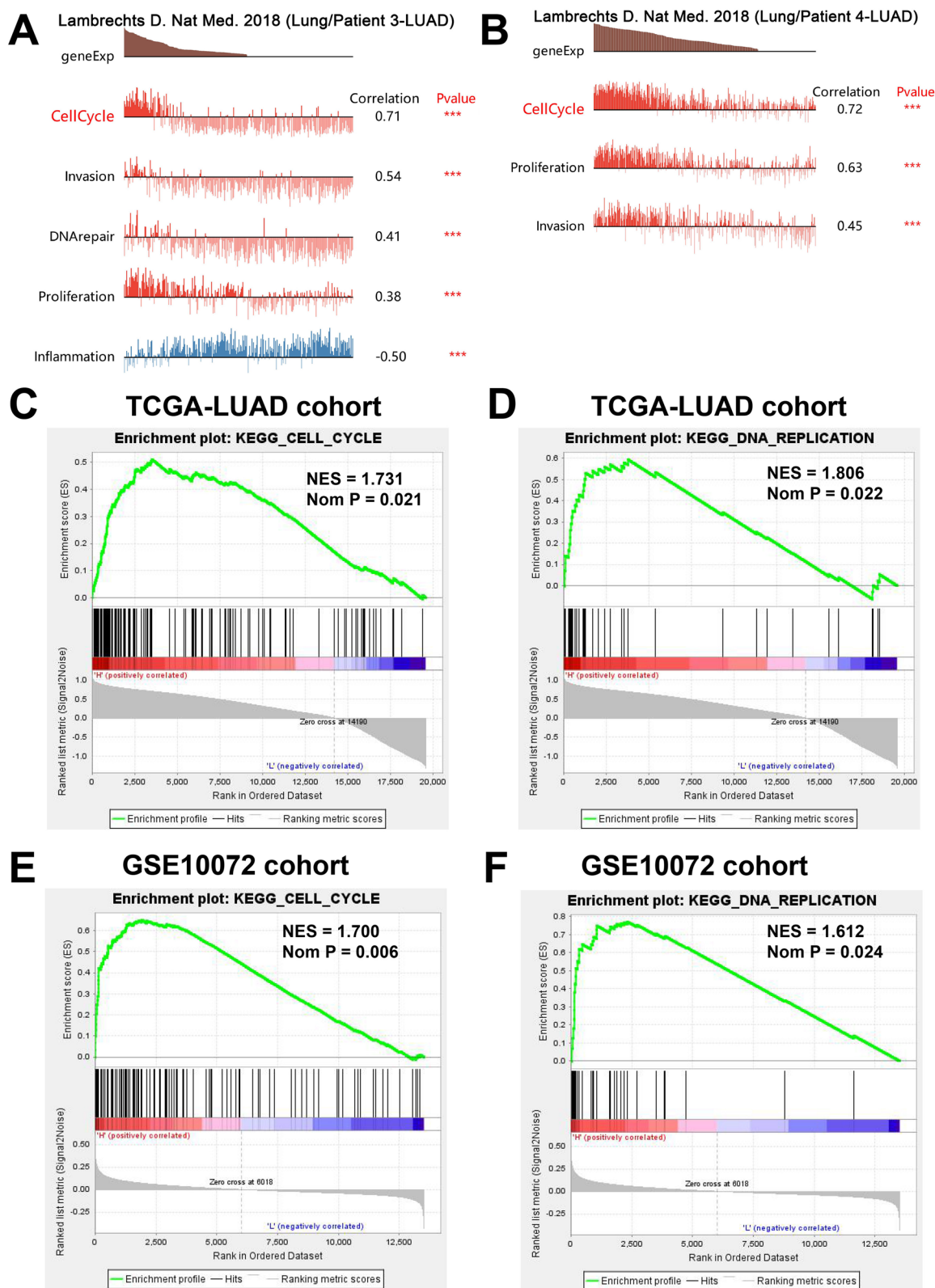
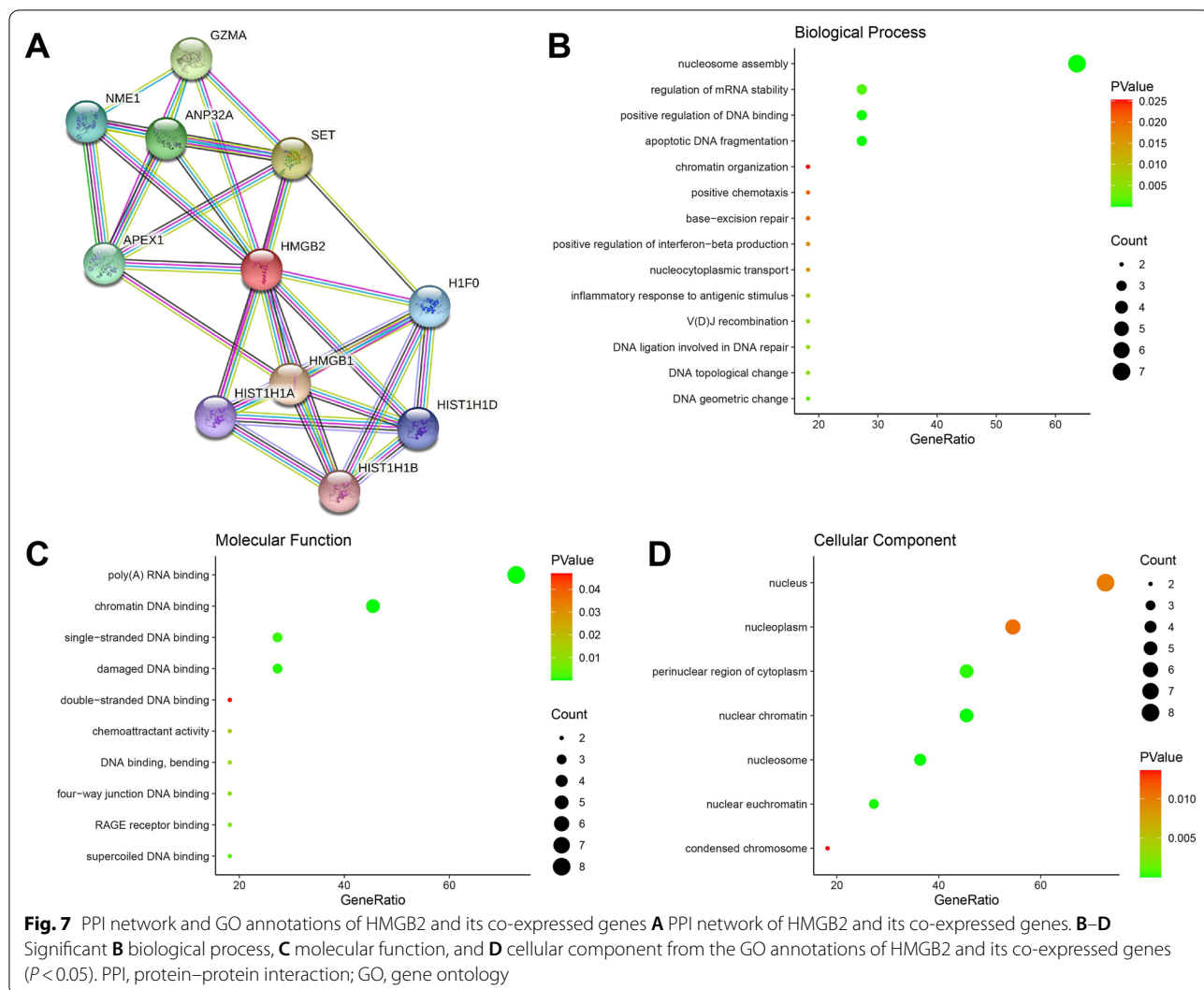


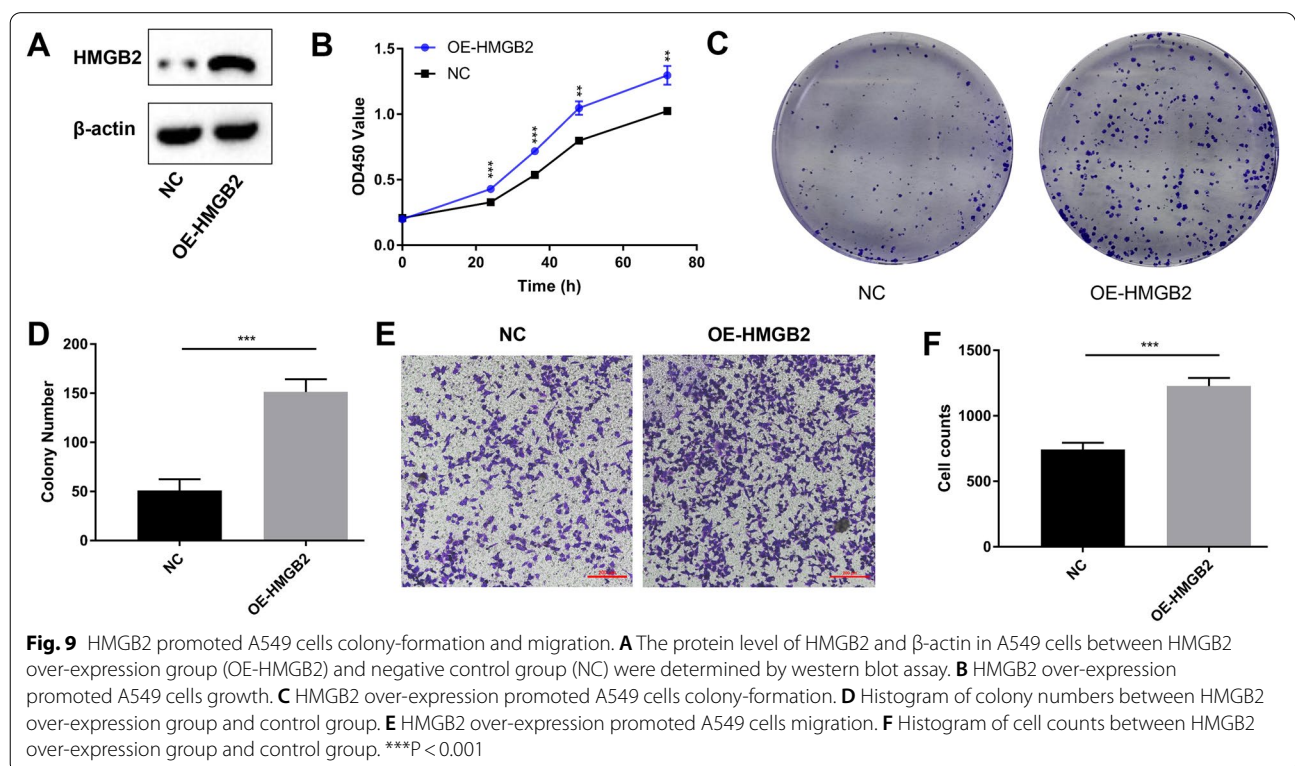
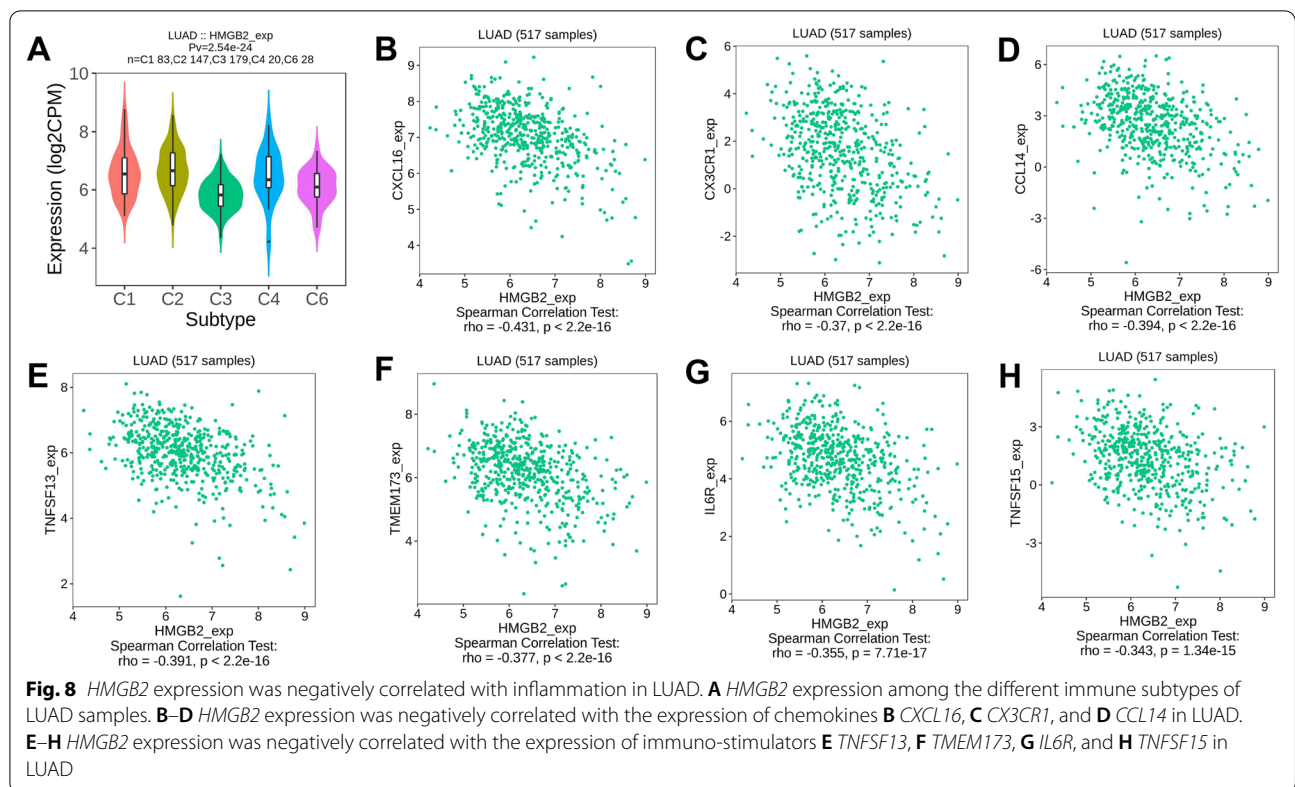
Fig. 6 (See legend on previous page.)

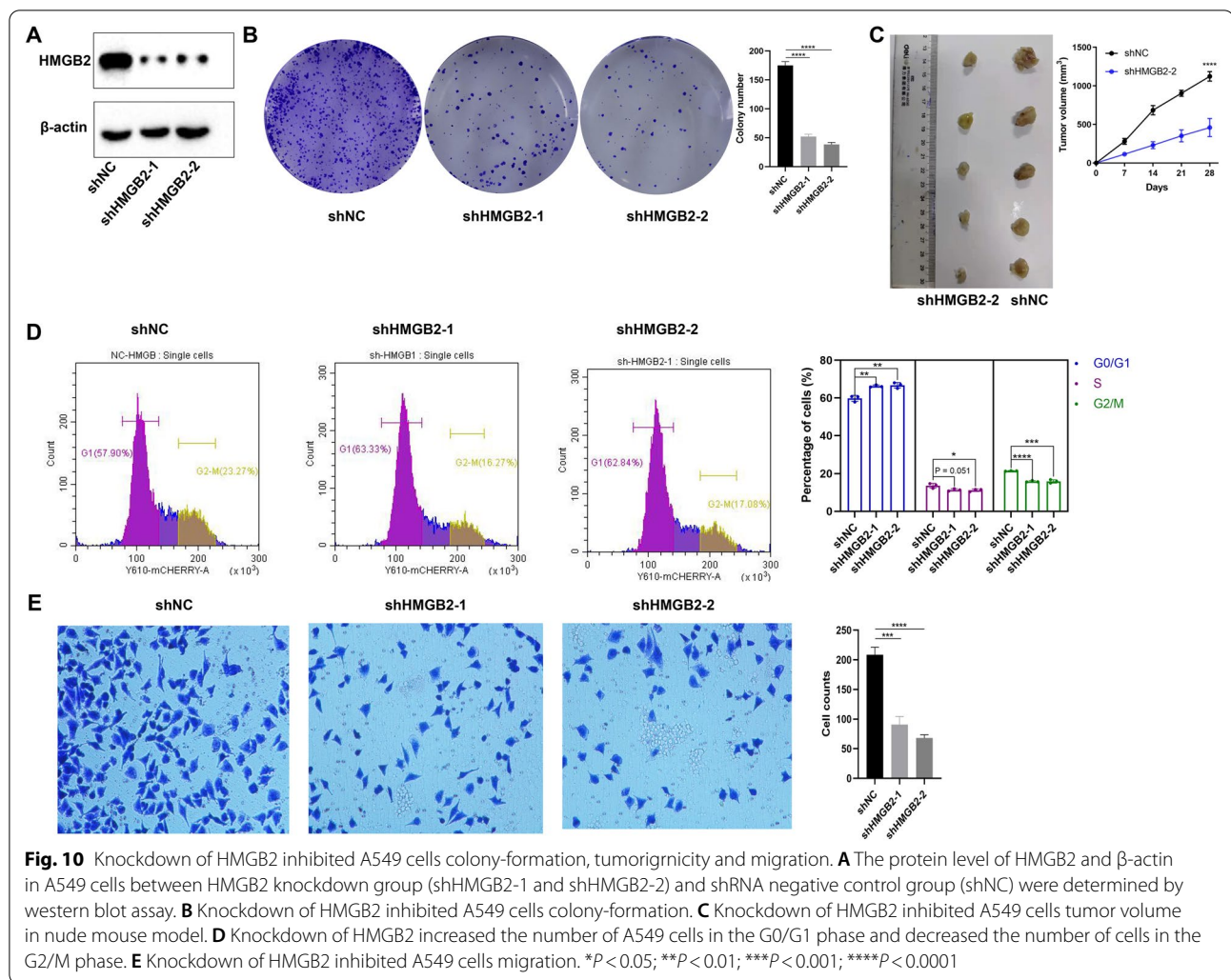


cancer [35]. In cervical cancer, pancreatic cancer, glioma, and ovarian cancer, the HMGB2 was reported as a reliable prognosis predictor. In NSCLC, the high expression of HMGB2 in cancer was associated with the chemotherapy response and poor prognosis [36, 37]. Together with PDIA3, p21, and LINC00184, the HMGB2 can regulate the chemotherapy- and radiotherapy-induced DNA damage to enhance drug resistance [17, 38, 39]. All these results indicated that the HMGB2 might play an important role in the development of tumors, including lung cancer. However, no evidence has demonstrated its effect on tumor growth in NSCLC, especially in LUAD.

In this study, HMGB2 was screened as a promising biomarker for LUAD. Furthermore, the integrated bioinformatics analysis, using the data from public platforms both at tissue and single-cell levels indicated

that the HMGB2 might affect the cell cycle, proliferation, and expression of inflammatory factors in LUAD. For further validation, the correlation of HMGB2 with clinical characteristics was analyzed using TMA staining for HMGB2 on 98 LUAD specimens. The results showed that the expression of HMGB2 was dramatically elevated in tumor cells as compared to that in the normal alveolar cells. Moreover, the HMGB2 expression level was identified to be highly correlated with a poor TNM stage, pathologic grade, and prognosis. Noteworthy, the expression of HMGB2 was positively correlated with that of Ki67 in LUAD. In conclusion, this study demonstrated that HMGB2 is a potential diagnostic and therapeutic indicator for LUAD, suggesting that HMGB2 might be a potential therapeutic target in LUAD.





Conclusions

We propose that HMGB2 may be correlated with proliferation of LUAD cells and it is a promising diagnostic and therapeutic marker for LUAD.

Abbreviations

LUAD: Lung adenocarcinoma; NSCLC: Non-small cell lung cancer; HMGB2: High-mobility group box 2; TCGA: The cancer genome atlas; GEO: Gene expression omnibus; ssGSEA: Single-sample gene set enrichment analysis; WGCNA: Weighted correlation network analysis; PPI: Protein-protein interaction; TMA: Tissue microarray; OS: Overall survival; IHC: Immunohistochemistry.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12890-022-02110-y>.

Additional file 1: Table S1. Clinical informations and IHC score of patients with LUAD.

Additional file 2: Figure S1. Original blot and images. **Figure S2.** HMGB2 expression is not correlated with survival of patients with LUSC.

Acknowledgements

Our research based on public available data from TCGA and GEO. Besides, online databases, such as GEPIA, CancerSEA and TISIDB, provided enormous support for our research. We thank the contributors who provided these databases/resources.

Author contributions

XQ and WL: performed the experiments, wrote the manuscript. YZ: performed the experiments and collected the breast cancer samples. KZ and HW participated in data processing and analysis. JD and HS: supervised the project and provided funds for the whole project. All authors read and approved the final manuscript.

Funding

This research was supported by Youth Elite Project of The First People's Hospital of Lianyungang [QN1903].

Availability of data and materials

The main results of this study are available from public databases, such as CancerSEA database (<http://biocc.hrbmu.edu.cn/CancerSEA/home.jsp>), GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), TCGA database (<https://portal.gdc.cancer.gov/>), MSigDB database (<http://www.gsea-msigdb.org/gsea/index.jsp>), STRING database (<https://www.string-db.org/>), GEPIA2.0 database (<http://gepia2.cancer-pku.cn/#index>), PrognScan database (<http://dna00.bio.kyutech.ac.jp/PrognScan/index.html>), TISIDB database (<http://cis.hku.hk/TISIDB/>),

and CPTAC database (<https://pdc.cancer.gov>). The clinical information and IHC data for human specimens were shown in Additional file 1: Table S1.

Declarations

Ethics approval and consent to participate

All experimental protocols and human specimens collected from The First People's Hospital of Lianyungang were approved by the Ethics Committee of The First People's Hospital of Lianyungang. The approved number is KY-20190927005. All participants provided written informed consent. All methods were carried out in accordance with the Helsinki declaration. All experimental protocols involving animals were approved by the Nanjing Medical University Health Science Center Institutional Animal Care. All experiments involving animals were performed in accordance with relevant guidelines and regulations. All methods involving animals were carried out in accordance with ARRIVE guidelines (<https://arriveguidelines.org>).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Cardiothoracic Surgery, The First People's Hospital of Lianyungang, Lianyungang, People's Republic of China. ²Department of Thoracic Surgery, Haian People's Hospital Affiliated to Nantong University, Haian, People's Republic of China. ³Department of Cardiothoracic Surgery, The Second Affiliated Hospital of Nantong University, Nantong, People's Republic of China. ⁴Department of Thyroid Surgery, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shenzhen 518000, China. ⁵Yancheng TCM Hospital, Nanjing University of Chinese Medicine, Yancheng 224002, China.

Received: 26 April 2022 Accepted: 9 August 2022

Published online: 12 August 2022

References

- Duma N, Santana-Davila R, Molina JR. Non-small cell lung cancer: epidemiology, screening, diagnosis, and treatment. *Mayo Clin Proc*. 2019;94(8):1623–40.
- Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res*. 2016;5(3):288–300.
- Schneider BJ, Ismaila N, Aerts J, Chiles C, Daly ME, Detterbeck FC, et al. Lung cancer surveillance after definitive curative-intent therapy: ASCO guideline. *J Clin Oncol*. 2020;38(7):753–66.
- Uramoto H, Tanaka F. Recurrence after surgery in patients with NSCLC. *Transl Lung Cancer Res*. 2014;3(4):242–9.
- Martinez-Terroba E, Behrens C, de Miguel FJ, Agorreta J, Monso E, Millares L, et al. A novel protein-based prognostic signature improves risk stratification to guide clinical management in early-stage lung adenocarcinoma patients. *J Pathol*. 2018;245(4):421–32.
- Tanoue LT, Tanner NT, Gould MK, Silvestri GA. Lung cancer screening. *Am J Respir Crit Care Med*. 2015;191(1):19–33.
- Goss PE, Strasser-Weippl K, Lee-Bychkovsky BL, Fan L, Li J, Chavarri-Guerra Y, et al. Challenges to effective cancer control in China, India, and Russia. *Lancet Oncol*. 2014;15(5):489–538.
- Shykind BM, Kim J, Sharp PA. Activation of the TFIID-TFIIA complex with HMG-2. *Genes Dev*. 1995;9(11):1354–65.
- Fan Z, Beresford PJ, Zhang D, Lieberman J. HMG2 interacts with the nucleosome assembly protein SET and is a target of the cytotoxic T-lymphocyte protease granzyme A. *Mol Cell Biol*. 2002;22(8):2810–20.
- Fu D, Li J, Wei J, Zhang Z, Luo Y, Tan H, et al. HMGB2 is associated with malignancy and regulates Warburg effect by targeting LDHB and FBP1 in breast cancer. *Cell Commun Signal*. 2018;16(1):8.
- Fang J, Ge X, Xu W, Xie J, Qin Z, Shi L, et al. Bioinformatics analysis of the prognosis and biological significance of HMGB1, HMGB2, and HMGB3 in gastric cancer. *J Cell Physiol*. 2020;235(4):3438–46.
- Cai X, Ding H, Liu Y, Pan G, Li Q, Yang Z, et al. Expression of HMGB2 indicates worse survival of patients and is required for the maintenance of Warburg effect in pancreatic cancer. *Acta Biochim Biophys Sin*. 2017;49(2):119–27.
- Li J, Gao J, Tian W, Li Y, Zhang J. Long non-coding RNA MALAT1 drives gastric cancer progression by regulating HMGB2 modulating the miR-1297. *Cancer Cell Int*. 2017;17:44.
- An Y, Zhang Z, Shang Y, Jiang X, Dong J, Yu P, et al. miR-23b-3p regulates the chemoresistance of gastric cancer cells by targeting ATG12 and HMGB2. *Cell Death Dis*. 2015;6:e1766.
- Yuan F, Zhao ZT, Jia B, Wang YP, Lei W. TSN inhibits cell proliferation, migration, invasion, and EMT through regulating miR-874/HMGB2/beta-catenin pathway in gastric cancer. *Neoplasma*. 2020;67(5):1012–21.
- Li S, Yang J, Xia Y, Fan Q, Yang KP. Long noncoding RNA NEAT1 promotes proliferation and invasion via targeting miR-181a-5p in non-small cell lung cancer. *Oncol Res*. 2018;26(2):289–96.
- Kim HK, Kang MA, Kim MS, Shin YJ, Chi SG, Jeong JH. Transcriptional repression of high-mobility group box 2 by p21 in radiation-induced senescence. *Mol Cells*. 2018;41(4):362–72.
- Zhou G-H, Yi-Yu L, Xie J-L, Gao Z-K, Xiao-Bo W, Yao W-S, Wei-Guang G. Overexpression of miR-758 inhibited proliferation, migration, invasion, and promoted apoptosis of non-small cell lung cancer cells by negatively regulating HMGB. 2019. *Biosci Rep*. <https://doi.org/10.1042/BSR20180855>.
- Yuan H, Yan M, Zhang G, Liu W, Deng C, Liao G, et al. CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res*. 2019;47(D1):D900–8.
- Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med*. 2018;24(8):1277–89.
- Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, Dasgupta A, et al. Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. *PLoS ONE*. 2008;3(2):e1651.
- Lo F-Y, Jer-Wei Chang IS, Chang Y-JC, Hsu H-S, Huang S-FK, Tsai F-Y, Jiang SS, Kanteti R, Nandi S, Salgia R, Wang Y-C. The database of chromosome imbalance regions and genes resided in lung cancer from Asian and Caucasian identified by array-comparative genomic hybridization. *BMC Cancer*. 2012. <https://doi.org/10.1186/1471-2407-12-235>.
- Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. *Genome Res*. 2012;22(7):1197–211.
- Chen Y-J, Roumeliotis TI, Chang Y-H, Chen C-T, Han C-L, Lin M-H, et al. Proteogenomics of non-smoking lung cancer in East Asia delineates molecular signatures of pathogenesis and progression. *Cell*. 2020;182(1):226.
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50.
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605–12.
- Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res*. 2019;47(W1):W556–60.
- Mizuno H, Kitada K, Nakai K, Sarai A. PrognosScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics*. 2009;2:18.
- Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, et al. TISIDB: an integrated repository portal for tumor-immune system interactions. *Bioinformatics*. 2019;35(20):4200–2.
- Thomas JO. HMGB1 and 2: architectural DNA-binding proteins. *Biochem Soc Trans*. 2001;29(Pt 4):395–401.
- Gao KL, Li M, Zhang KP. Imperatorin inhibits the invasion and migration of breast cancer cells by regulating HMGB2. *J Biol Regul Homeost Agents*. 2021;35(1):227–30.
- Redmond AM, Byrne C, Bane FT, Brown GD, Tibbitts P, O'Brien K, et al. Genomic interaction between ER and HMGB2 identifies DDX18 as a

- novel driver of endocrine resistance in breast cancer cells. *Oncogene*. 2015;34(29):3871–80.
34. Cui G, Cai F, Ding Z, Gao L. HMGB2 promotes the malignancy of human gastric cancer and indicates poor survival outcome. *Hum Pathol*. 2019;84:133–41.
 35. Suzuki S, Kato H, Fuji S, Naiki T, Naiki-Ito A, Yamashita Y, et al. Early detection of prostate carcinogens by immunohistochemistry of HMGB2. *J Toxicol Sci*. 2018;43(6):359–67.
 36. Lou N, Zhu T, Qin D, Tian J, Liu J. High-mobility group box 2 reflects exacerbated disease characteristics and poor prognosis in non-small cell lung cancer patients. *Ir J Med Sci*. 2021;191(1):155–62. <https://doi.org/10.1007/s11845-021-02549-8>.
 37. Wang Y, Li XP, Yin JY, Zhang Y, He H, Qian CY, et al. Association of HMGB1 and HMGB2 genetic polymorphisms with lung cancer chemotherapy response. *Clin Exp Pharmacol Physiol*. 2014;41(6):408–15.
 38. Krynetskaia NF, Phadke MS, Jadhav SH, Krynetskiy EY. Chromatin-associated proteins HMGB1/2 and PDIA3 trigger cellular response to chemotherapy-induced DNA damage. *Mol Cancer Ther*. 2009;8(4):864–72.
 39. Wang W, Li L, Zhao L. LINC00184 plays an oncogenic role in non-small cell lung cancer via regulation of the miR-524-5p/HMGB2 axis. *J Cell Mol Med*. 2021;25(21):9927–38. <https://doi.org/10.1111/jcmm.16247>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

