Biochemistry and Biophysics Reports 40 (2024) 101851

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

A pan-cancer analysis of the oncogenic function of HMGB1 in human tumors

Hui-min Yang [a,1], Xiang-ning Zhao [b,1], Xiao-ling Li [c], Xi Wang [c], Yu Pu [c], Dong-kai Wei [c], Zhe Li [d,*]

[a] Shanghai Singlera Medical Laboratory, 1rst Floor, No. 20 Building, 500 Furonghua Road, Pudong New District, Shanghai, China
[b] Department of Surgical Oncology, Shanghai Mengchao Cancer Hospital, 118 Qianyang Road, Jiading District, Shanghai, China
[c] Suzhou Func Biotech Inc, Suzhou, Jiangsu, China
[d] Department of Breast Surgery, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, 528 Zhangheng Road, Pudong New District, Shanghai, China

ARTICLE INFO

Keywords:
Cancer
HMGB1
Mutation
Prognostic
Immune cell infiltration

ABSTRACT

Background: Although high mobility group box protein 1 (*HMGB1*) has been researched in relation to cancer in many investigations, a thorough investigation of its role in pan-cancer has yet to be conducted. With the objective of bridging this gap, we delved into the functions of *HMGB1* in various tumors.

Methods: This investigation employed The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases to examine *HMGB1* gene expression differences and correlation with survival across various human tumors. Then, genetic alterations of *HMGB1* were analyzed by tool cBioPortal, and immune cell infiltration was assessed. Finally, we gathered clinial samples from 95 patients with various types of solid tumor and performed somatic mutation analysis using panel sequencing. This further highlighted the role of HMGB1 in different solid tumors.

Results: There was a notable elevation of *HMGB1* gene expression in tumor tissues as opposed to non-cancerous tissues across the bulk of tumor types. Elevated *HMGB1* gene expression had a connection with shorter overall survival, progression-free survival, and disease-free survival in specific tumor types. Genetic alterations of *HMGB1* suggested that the amplifications and mutations of *HMGB1* may impact the prognosis of breast cancer (BRCA) and liver hepatocellular carcinoma (LIHC). Both BRCA and mesothelioma (MESO) displayed a connection between the infiltration of cancer-associated fibroblasts (CAFs) and *HMGB1* gene expression. Moreover, *HMGB1* co-expression analysis revealed its association with genes involved in RNA splicing, mRNA processing, and modulation of mRNA metabolic processes. Additionally, a pathway analysis by use of the Kyoto Encyclopedia of Genes and Genomes (KEGG) unveiled that *HMGB1* was implicated in the pathogenic mechanisms of "Hepatitis B," "Viral Carcinogenesis," and "Hepatocellular Carcinoma." Based on somatic mutation analysis of 95 patients with different solid tumors, we found that the frequency of HMGB1 mutations was higher in Liver cancer patients compared to other solid tumors. This finding is consistent with our in-silico study results. Additionally, we discovered that the frequency of HMGB1 mutations ranked among the top 20 mutated genes in the 95 patients' data, indicating that HMGB1 plays an important role in the development and prognosis of various solid tumors.

Conclusion: This pan-cancer study of *HMGB1* underscores its potential as a signature marker and target for the management of various tumor types.

1. Introduction

Cancer is a complex disease marked by abnormal cell growth. Treating it is challenging due to its diverse nature, resistance to conventional therapies, and the heavy burden of side effects on patients. However, new strategies like personalized medicine, immunotherapy,

* Corresponding author.
E-mail address: lizhezsh@126.com (Z. Li).
[1] Hui-min Yang and Xiang-ning Zhao share the co-first authorship.

https://doi.org/10.1016/j.bbrep.2024.101851
Received 9 May 2024; Received in revised form 30 September 2024; Accepted 16 October 2024

targeted drug therapies, and gene editing show promise in tackling these issues. Despite the obstacles, ongoing research brings hope for better treatment outcomes and improved patient survival [1]. Considering the intricacies of tumorigenesis, a pan-cancer investigation of oncogene expression and its connection to possible molecular mechanisms and patient outcomes is imperative [2]. The use of The Cancer Genome Atlas (TCGA) and The Gene Expression Omnibus (GEO), which present extensive functional genomics data from various tumors, enables relevant pan-cancer investigations [3,4].

High mobility group box protein 1 (*HMGB1*) is a non-histone multifunctional nuclear protein falling under the damage-associated molecular pattern family, which performs critical functions in cancers of the kidney, lung, pancreas cancer, stomach, colon, liver, breast, prostate, and cervix, as well as blood system tumors and melanoma [5,6]. *HMGB1* plays a role in diverse cellular mechanisms, including DNA repair, genome stability, recombination, autophagy, necrosis, and apoptosis, and enhances tumor cell proliferation, invasion, and metastasis while improving the immune response, inducing tumor cell apoptosis, and inhibiting tumor growth [6–10]. Studies have also found that *HMGB1* mRNA expression is raised in gastrointestinal cancer tissues relative to adjacent differentiated cancerous tissues and pre-cancer lesions. Moreover, *HMGB1* mRNA expression in non-cancer tissues around advanced or poorly differentiated cancers is also enhanced. However, regular tissue expression around well-differentiated gastric cancer does not increase [11].

Recent studies have shown that HMGB1 interacts with autophagy, a cellular self-degradation process that controls the quantity and quality of proteins, organelles, and other cellular components. HMGB1 can affect autophagy in several ways. For instance, it can directly interact with autophagy-related proteins like Beclin 1, thereby regulating the initiation and progression of autophagy [12]. Secondly, HMGB1 can modulate autophagy by activating signaling pathways such as RAGE and TLR4 [13]. Lastly, the release of HMGB1 itself is also regulated by autophagy [14]. In summary, the intricate relationship between HMGB1 and autophagy significantly influences cellular survival and death. Further investigation into the molecular mechanisms underlying these interactions could reveal new therapeutic targets for related diseases.

This study features a novel pan-cancer investigation of *HMGB1* that employs the TCGA and GEO databases. Our comprehensive analysis investigated gene expression and alterations, survival outcomes, immune cell infiltration (ICI), and relevant cellular pathways, with the aim of elucidating the possible molecular mechanisms that are hypothesized to underlie the pathogenesis or clinical manifestation of various types of cancer, in which *HMGB1* may play a role.

2. Materials and methods

2.1. Gene expression assessment

A comprehensive assessment of *HMGB1* gene expression in various tumors was implemented by use of multiple web-based resources. Specifically, we utilized the "Gene_DE" module of Tumor Immune Estimation Resource version 2 (TIMER2 (http://timer.cistrome.org/), to assess the discrepancies in *HMGB1* gene expression between non-cancerous and tumor tissues. Regarding tumors that lacked typical or had limited normal tissues as controls in the TIMER2 database, the "Expression Analysis-Box Plots" module of the Gene Expression Profiling Interactive Analysis, version 2 (GEPIA2, http://gepia2.cancer-pku.cn/#analysis) was utilized [15]. By use of the Genotype-Tissue Expression (GTEx) database, we obtained box plots depicting the expression discrepancy between tumor and non-cancerous tissues, with thresholds of log2 fold change (FC) = 1 and P = 0.01, while also ensuring "Match TCGA normal and GTEx data." Further, we explored *HMGB1* gene expression levels in stage I–IV of all tumors in the TCGA dataset by selecting the "Pathological Stage Plot" module of GEPIA2, generating violin plots using gene expression data transformed by the log2 [transcripts per million (TPM) +1]. Finally, the Clinical Proteomic Tumor Analysis Consortium (CPTAC) dataset was examined by use of the UALCAN (https://ualcan.path.uab.edu/analysis-prot.html), an interactive web tool presenting cancer omics information [16].

2.2. Survival prognosis assessment

With the aim of ascertaining the interrelation between survival outcomes and *HMGB1* gene expression in TCGA tumors, the "Survival Analysis" module of GEPIA2 was employed [11]. This investigation provided us with Kaplan-Meier (KM) curves reflecting overall survival (OS) and disease-free survival (DFS), along with a map indicating the significance of survival outcomes. We used 50 % cutoff-high and 50 % cutoff-low as the criteria to distinguish between high and low expression cohorts. To delve deeper into the interrelation between survival and *HMGB1* gene expression across distinct cancer types, we turned to the KM plotter (http://kmplot.com/analysis/) [17]. This tool enables the assessment of the impact of 54k genes (mRNA, miRNA, and protein) on survival in 21 types of cancer. Our analysis of the "Pan-cancer RNA-seq" module allowed us to ascertain the connection between survival and *HMGB1* gene expression in distinct cancer types. The databases used in this analysis include GEO, EGA, and TCGA.

2.3. Gene variation assessment

To assess gene variations, we accessed the cBioPortal (https://www.cbioportal.org/) [18,19]. Subsequent to entering *HMGB1* into the "Quick Search" module, we accessed the "Cancer Type Summary" module, which allowed us to view the gene mutation type, frequency of alternation, and copy number alterations (CNAs) across all tumors in the TCGA dataset. Furthermore, the "Mutation" module was employed to explore the mutated site within the protein or three-dimensional (3D) structure. Additionally, we utilized the "Comparison" module to assess the difference in OS, DFS, and progression-free survival (PFS) in the *HMGB1* mutated group versus the non-mutated group in colorectal cancer (CRC). We analyzed the resulting data were analyzed by use of KM plots and determined the P-value computed *via* the log-rank test.

2.4. ICI analysis

With the objective of delving into the interrelation between cancer-associated fibroblasts (CAFs) and *HMGB1* gene expression across all tumors in the TCGA dataset, we employed the "Immune" module of TIMER2 and made an assessment of the ICI degree by Microenvironment Cell Populations-counter (MCP-counter), Tracking of Indels by Decomposition (TIDE), xCell, and Estimating the Proportions of Immune and Cancer Cells (EPIC) algorithms, and computed partial correlation (PC) and P-values by use of a Spearman's rank correlation analysis adjusted for purity. The heat map and scatter plot were used for representing the results.

2.5. Enrichment analysis of genes associated with HMGB1

We search for proteins that bind to *HMGB1* by querying the STRING database (https://string-db.org/) and inputting "*HMGB1*" as the protein and "Homo sapiens" as the organism. The search parameters used for the STRING database included a minimum score threshold of "low confidence (0.150)", with network edges considered as "evidence." Additionally, a maximum of 50 interactors were displayed in the primary shell, with "experiments" being the preferred interaction sources. This process yielded 50 proteins that bind to *HMGB1*, according to experimental evidence.

By use of the "Similar Gene Detection" module of GEPIA2, the identified 100 most highly correlated genes with *HMGB1* were analyzed in healthy tissues and all tumors in the TCGA dataset. With the aim of examining the interrelation between *HMGB1* and the top five chosen

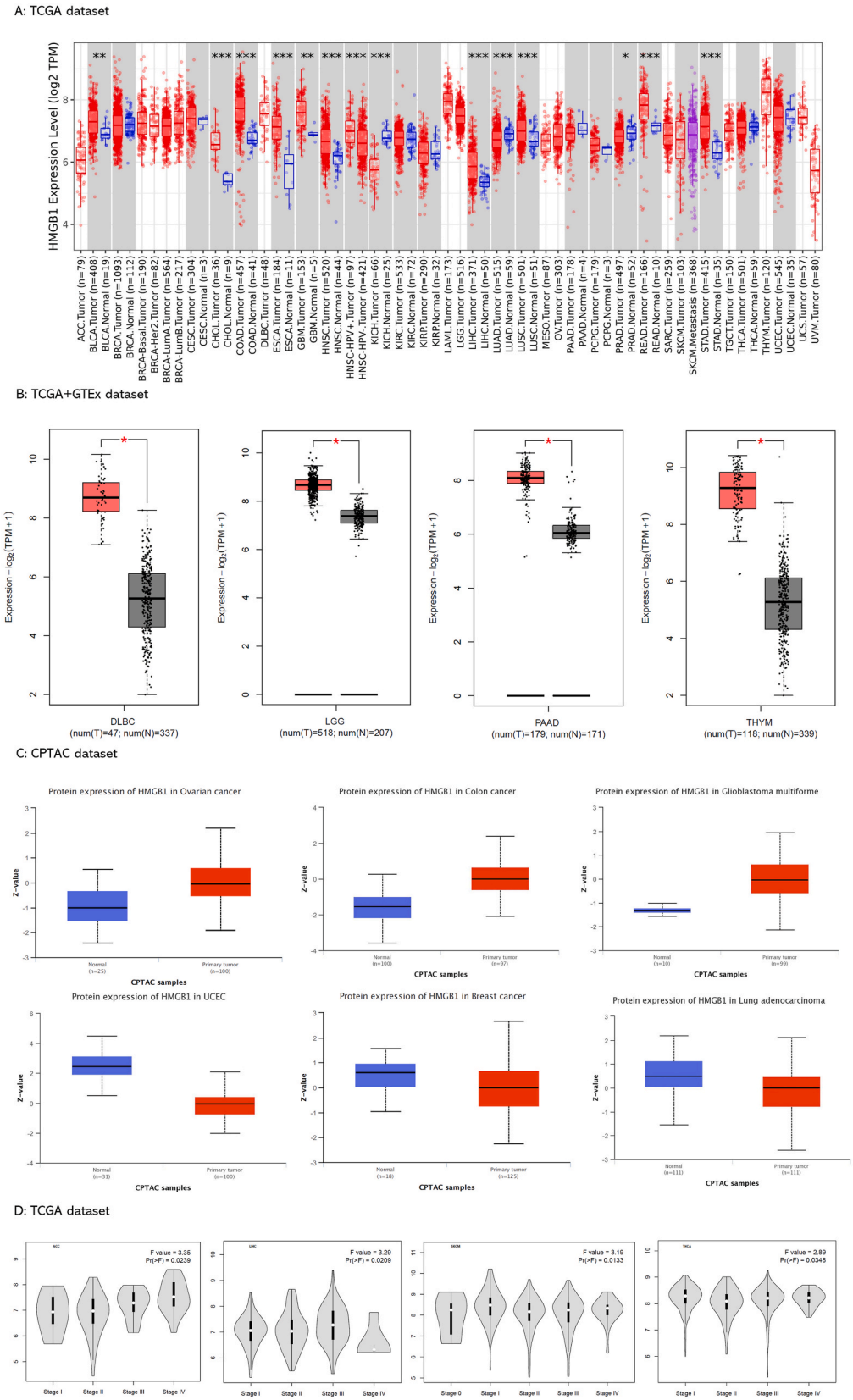

Fig. 1. *HMGB1* gene expression levels in various tumors and pathological stages. (**A**) *HMGB1* gene expression in various cancer types or subtypes. *P < 0.05; **P < 0.01; ***P < 0.001. (**B**) Comparison results of *HMGB1* gene expression levels between DLBC, LGG, PAAD, and THYM tissues in the TCGA dataset and the matching non-cancerous tissues in the GTEx dataset. *P < 0.05. (**C**) Comparison results of the total protein expression levels of *HMGB1* between non-cancerous tissues and primary OV, COAD, GBM, UCEC, BRCA, and LUAD tissues in the CPTAC dataset. ***P < 0.001. (**D**) Analysis data of*HMGB1* gene expression levels in stage I–IV of ACC, LIHC, SKCM, and THCA in the TCGA dataset, with log2 (TPM+1) used as the log scale.

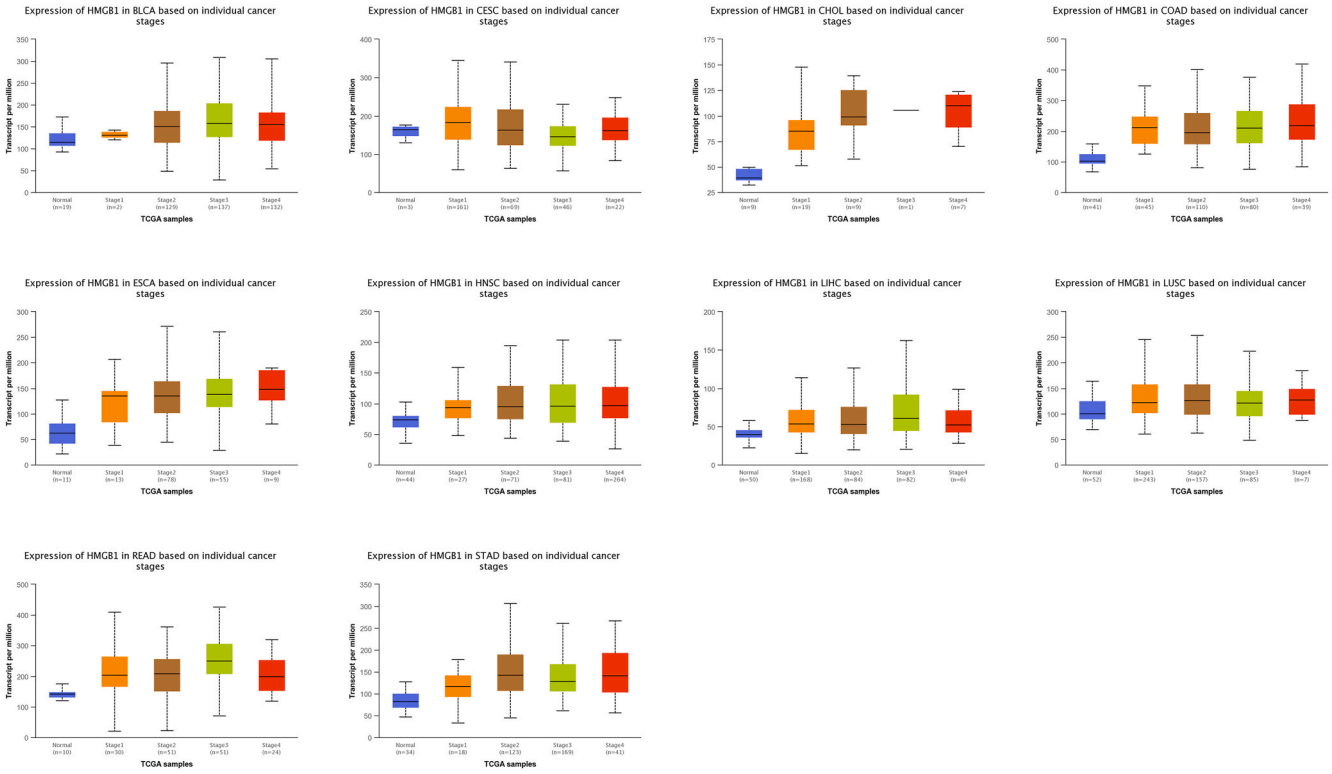

Fig. 2. Association between various pathological stages and *HMGB1* gene expression. The UALCAN database is used to assess *HMGB1* gene expression in stage I–IV of various cancer types in the TCGA dataset. The X-axis represents pathological stages and the corresponding number of samples, while the Y-axis represents TPM. N, normal; S, stage. A P-value marked in red indicates statistical significance between the two groups. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

genes, we implemented a gene-to-gene correlation analysis by use of the "Correlation Analysis" module of GEPIA2, which utilized Pearson correlation coefficients (R). Dot plots were generated based on log2 TPM, providing both P-values and R. Also, we employed TIMER2's "Gene_corr" module to obtain heatmap data, including PC and P-values acquired from Spearman's rank correlation analysis adjusted for purity, for the chosen genes. Additionally, we performed a comparative analysis of the genes that interact with HMGB1 and those that interact with it using Jenn [20], an interactive tool for creating Venn diagrams.

We then merged the two datasets and conducted a pathway analysis by use of the Kyoto Encyclopedia of Genes and Genomes (KEGG). Regarding functional annotation chart data, we uploaded gene lists with the chosen identifier ("OFFICIAL_GENE_SYMBOL") and organism ("Homo sapiens") to the Database for Annotation, Visualization, and Integrated Discovery (DAVID). and visualized enriched pathways by use of "tidyr" and "ggplot2″ packages in R [R-3.6.3, 64-bit] (https://www.r-project.org/). The "clusterProfiler" package in R was employed for Gene Ontology (GO) enrichment analysis. Molecular function, cellular component, and biological process data were displayed as cnetplots by use of the cnetplot function (colorEdge = T, circular = F, node label = T). A two-tailed significance level of $P < 0.05$ was indicative of statistical significance.

2.6. Patient samples

We collected tissue samples from 95 cancer patients, including those with liver, lung, gastric, colorectal, and other types of cancer. Genomic DNA was extracted from FFPE sections using commercial kits (QIAamp DNA FFPE Tissue Kit; QIAGEN) following the manufactures.

2.7. Library preparation and sequencing

For targeted sequencing, genomic DNA was quantified using the Qubit 3.0 with the dsDNA HS Assay Kit (ThermoFisher Scientific). Library preparations were carried out with the KAPA Hyperplus Kit (KAPA Biosystems). For the targeted panel, customized 225 genes PANEL probes (Integrated DNA Technologies) were used for hybridization enrichment. The capture reaction was performed using the hybridization and wash kit (Integrated DNA Technologies) according to the manufacture. Captured libraries were amplified on beads with Illumina p5 (5′ AAT GAT ACG GCG ACC ACC GA 3′) and p7 primers (5′ CAA GCA GAA GAC GGC ATA CGA GAT 3') using KAPA HiFi HotStart ReadyMix (KAPA Biosystems), then followed by purification with Agencourt AMPure XP beads (Beckman). The library fragment size was determined using the Bioanalyzer 4200 (Agilent Technologies), then sequenced on the MGI2000 system (BGI).

2.8. Sequence analysis

Sequencing data (FASTQ) was first trimmed to remove sequencing adapters, then aligned to the human genome hg19 using BWA. Somatic mutation detection was performed using GATK4 with default filtering conditions, resulting in VCF files containing mutation information. The VCF files were annotated using ANNOVAR. Using conda v22.9.0, VEP v93.2, and vcf2maf v1.6.16, the VCF files were converted to MAF files and merged. The R package maftools is used to read MAF files, while the dplyr package is employed for data processing. Gene mutation results from multiple samples are combined, and samples without gene mutations are filtered out. A waterfall plot is then created using the ggplot2 package.

3. Results

3.1. Gene expression assessment data

The GEPIA2 database analysis revealed a striking difference in *HMGB1* gene expression between 14 types of TCGA cancers and their corresponding normal controls (Fig. 1A). Notably, *HMGB1* gene expression was higher in eleven cancers, namely bladder urothelial carcinoma (BLCA), cholangiocarcinoma (CHOL), liver hepatocellular carcinoma (LIHC), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), glioblastoma multiforme (GBM), stomach adenocarcinoma (STAD), lung squamous cell carcinoma (LUSC), HNSC-HPV, and rectum adenocarcinoma (READ), as opposed to that in non-cancerous tissues. On the contrary, *HMGB1* gene expression was uncovered to be lower in lung adenocarcinoma (LUAD), kidney chromophobe (KICH), and prostate cancer (PRAD). To gain further insight into the variations in *HMGB1* gene expression between non-cancerous and tumor tissues, we employed the GTEx dataset for cancers that lacked non-cancerous tissues as controls in the TIMER2 database. There was a noticeable increase in *HMGB1* gene expression in low-grade glioma (LGG), diffuse large B-cell lymphoma (DLBC), thymoma (THYM), and pancreatic adenocarcinoma (PAAD) than in non-cancerous tissues ($P < 0.05$) (Fig. 1B). At length, these findings provided evidence that *HMGB1* gene expression rose dramatically in most tumors.

As uncovered by the protein data of *HMGB1* from the CPTAC dataset, primary COAD, ovarian cancer (OV), and GBM tissues displayed upregulated expression levels, while breast cancer (BRCA), uterine corpus endometrial carcinoma (UCEC), and LUAD tissues exhibited opposite results, as opposed to their respective non-cancerous tissues (Fig. 1C).

Additionally, we utilized the "Pathological Stage Diagram" module of GEPIA2 to clarify whether the pathological stage of diverse cancer types and *HMGB1* gene expression are interrelated. The findings demonstrated a marked link between *HMGB1* gene expression and the

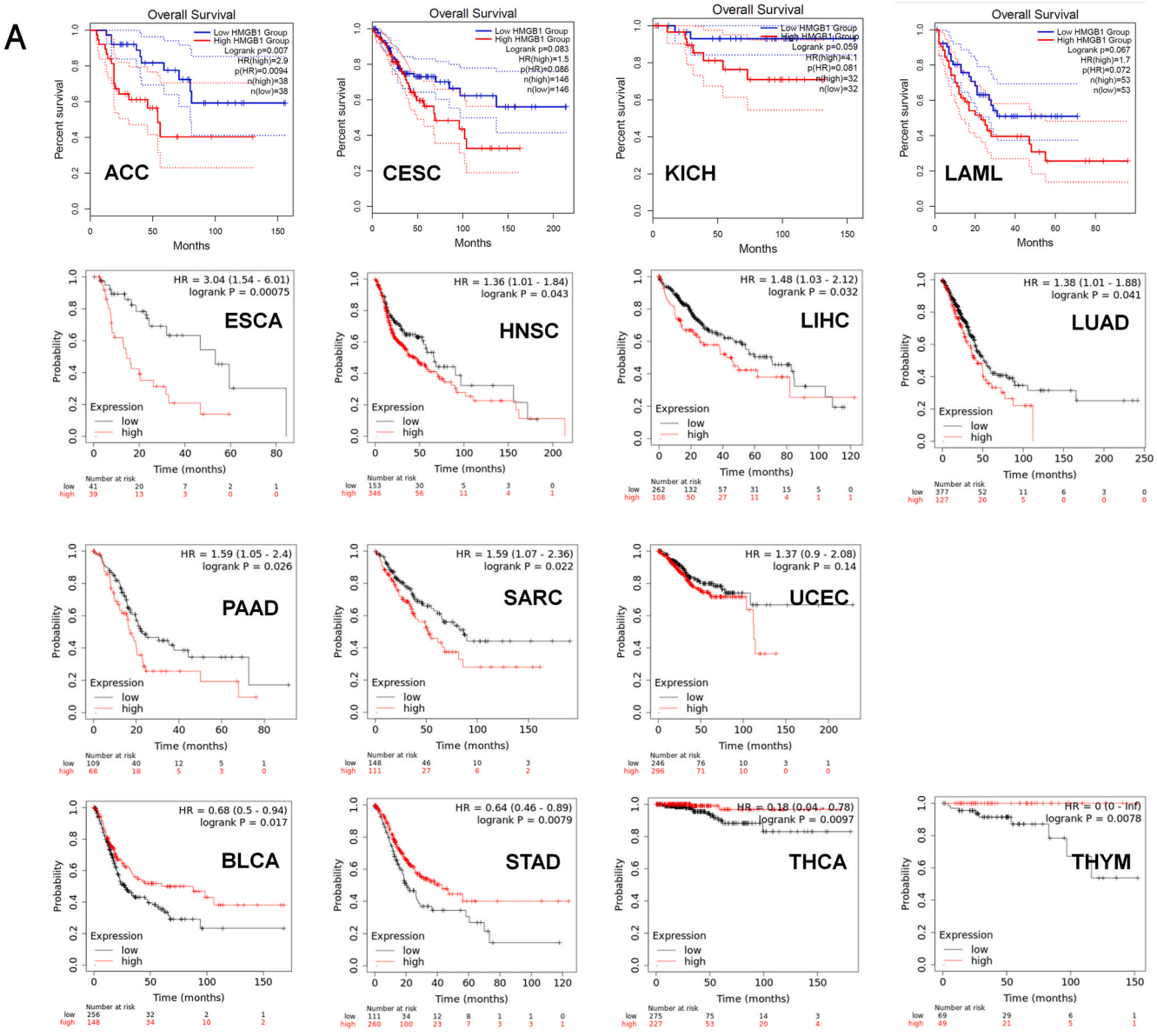

Fig. 3. Connection between survival outcomes of cancer patients in the TCGA database and *HMGB1* gene expression. (**A**) The OS of various cancers in TCGA is assessed by use of the GEPIA2 tool to determine *HMGB1* gene expression. (**B**) The DFS of various cancers in TCGA is assessed by use of the GEPIA2 tool to determine *HMGB1* gene expression. OS, overall survival; DFS, disease-free survival; TCGA, The Cancer Genome Atlas; GEPIA2, Gene Expression Profiling Interactive Analysis, version 2.

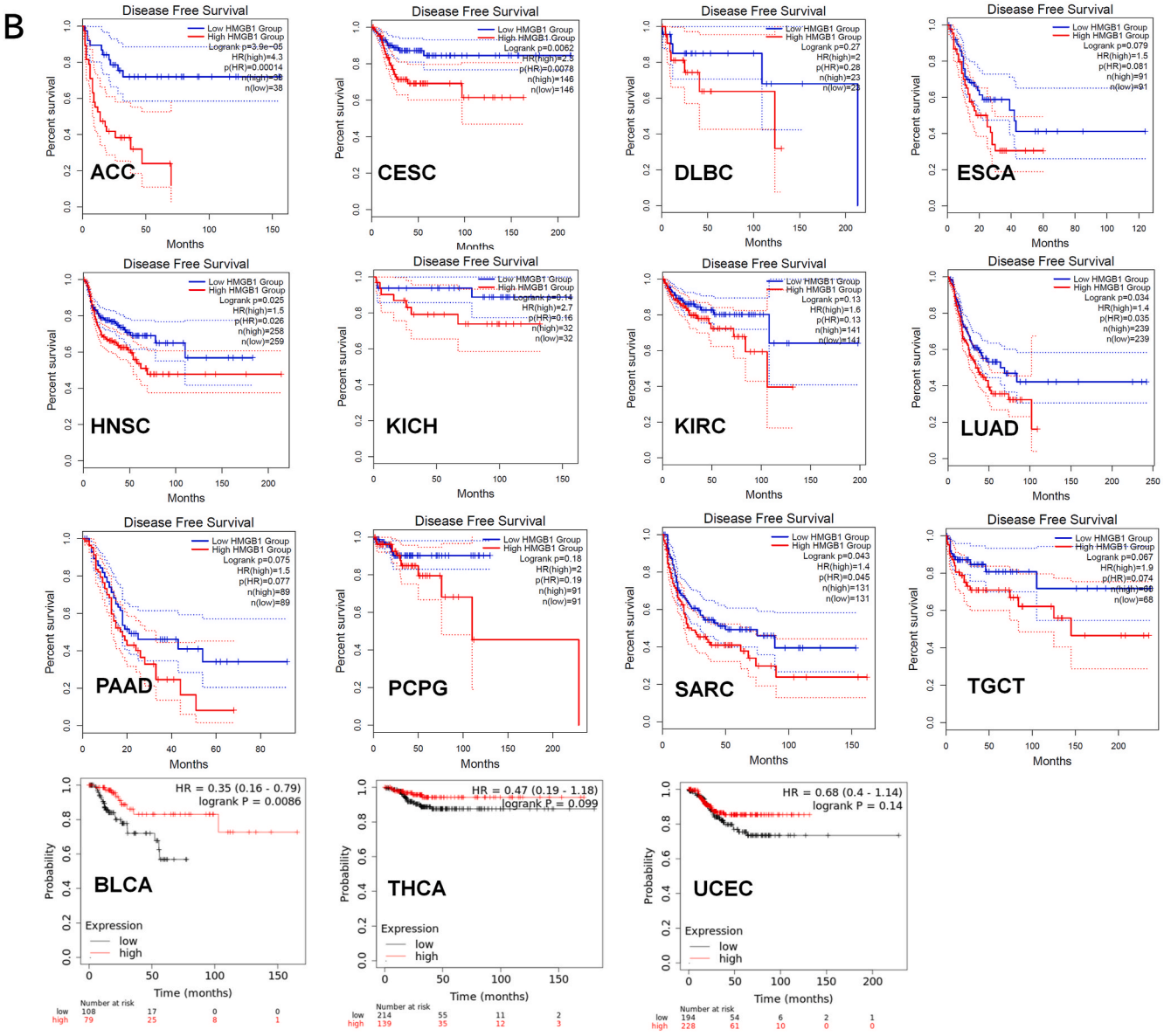

Fig. 3. (*continued*).

pathological stage of thyroid carcinoma (THCA), adrenocortical carcinoma (ACC), LIHC, and skin cutaneous melanoma (SKCM) (Fig. 1D). However, this correlation was not observed in other cancer types.

We proceeded to examine the interrelation between *HMGB1* gene expression and the pathological stage of individuals with diverse TCGA cancers. Our findings demonstrated a noteworthy elevation of *HMGB1* gene expression during the early stages of BLCA, CHOL, LIHC, COAD, LUSC, ESCA, cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), HNSC, READ, and STAD (Fig. 2). Nonetheless, we observed no pronounced differences between the advanced and early stages of cancer, implying *HMGB1* might participate in the onset of cancer rather than its progression.

3.2. Survival analysis data

HMGB1 gene expression was used to categorize individuals with cancer into high- and low-expression groups. In the TCGA and GEO datasets, the interrelation between *HMGB1* gene expression and patient outcomes in various tumor types was examined. Our investigation demonstrated that high *HMGB1* gene expression exhibited a correlation with shorter OS in ACC, CESE, KICH, and LAML (P = 0.007, 0.083, 0.059, and 0.067, respectively) within the TCGA project using GEPIA2. Similarly, survival analysis utilizing the KM plotter unveiled that shorter OS in ESCA, HNSC, LIHC, LUAD, PAAD, SARC, and UCEC (P = 0.00075, 0.043, 0.032, 0.041, 0.026, 0.022, and 0.14, respectively) within the TCGA project could be attributed to high *HMGB1* gene expression. Conversely, low *HMGB1* gene expression had a connection to shorter OS in BLCA, STAD, THCA, and THYM (P = 0.017, 0.0079, 0.0097, and 0.0078, respectively) (Fig. 3A).

Based on the data presented, high *HMGB1* gene expression showed striking relevance to poor prognosis across several tumor types, indicating the gene's potential as a pan-cancer prognostic marker.

Additionally, a striking association was uncovered between high *HMGB1* gene expression and shorter DFS in ACC, CESC, DLBC, ESCA, HNSC, ESCA, KICH, KIRC, LUAD, PAAD, PCPG, SARC, and TGCT patients (P = 3.9E-5, 0.0062, 0.27, 0.079, 0.079, 0.025, 0.14, 0.13, 0.034, 0.075, 0.18, 0.043, and 0.067, respectively). Conversely, low *HMGB1* gene expression was tied to shorter DFS in BLCA, THCA, and UCEC (P = 0.0086, 0.0086, and 0.14, respectively) (Fig. 3B).

Meanwhile, we observed that *HMGB1* gene expression had a variable

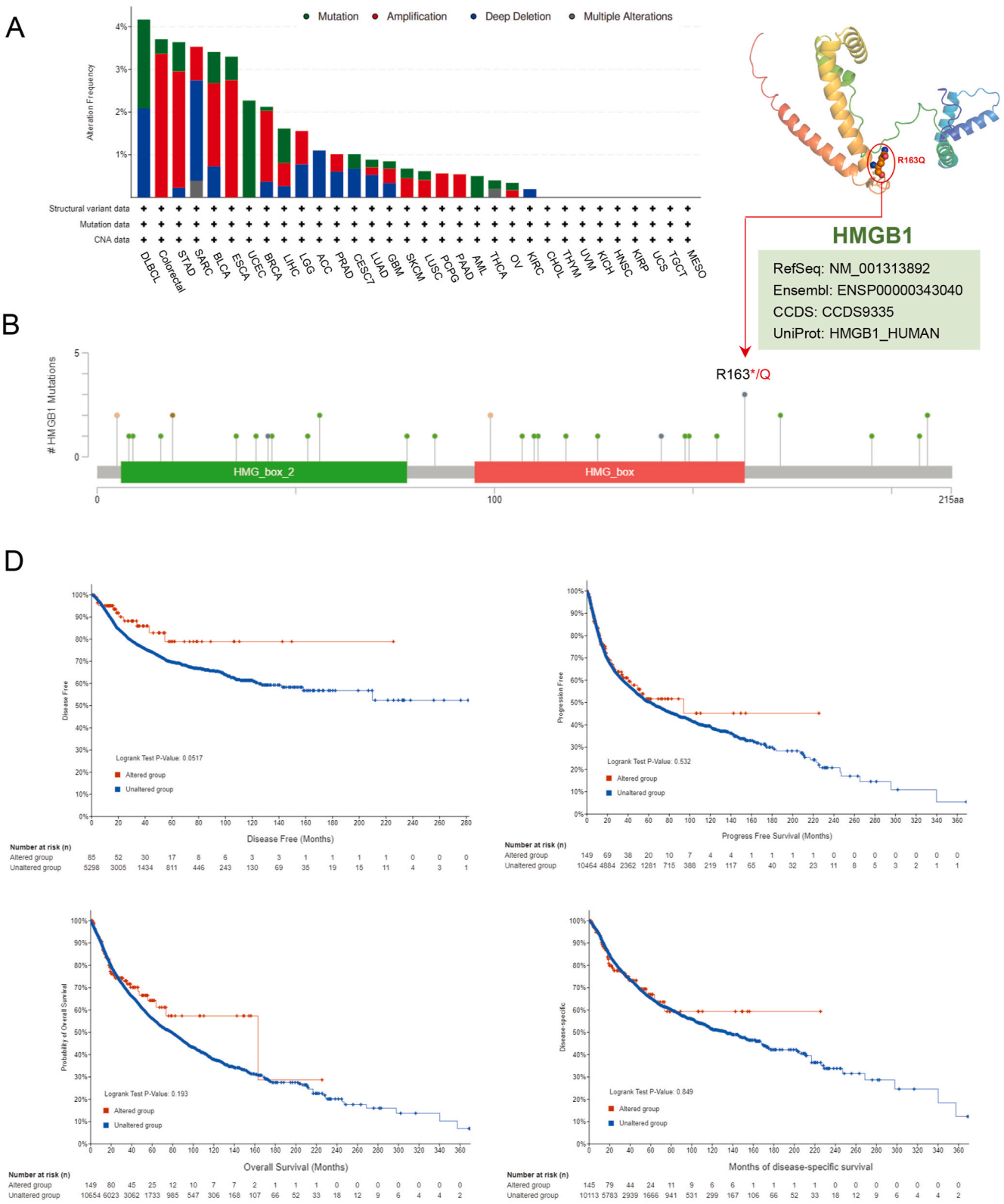

Fig. 4. Mutation characteristics of *HMGB1* across various tumors in the TCGA database. (**A**) The frequency and type of *HMGB1* mutations in TCGA cancers are explored by use of the cBioPortal tool. (**B**) *HMGB1* mutation sites in TCGA cancers are presented. (**C**) The R163Q mutation site is visualized in the 3D structure of *HMGB1*. (**D**) The possible interrelation between *HMGB1* mutation status and DFS, PFS, OS, and disease-specific survival of COAD is analyzed by use of the cBioPortal tool.

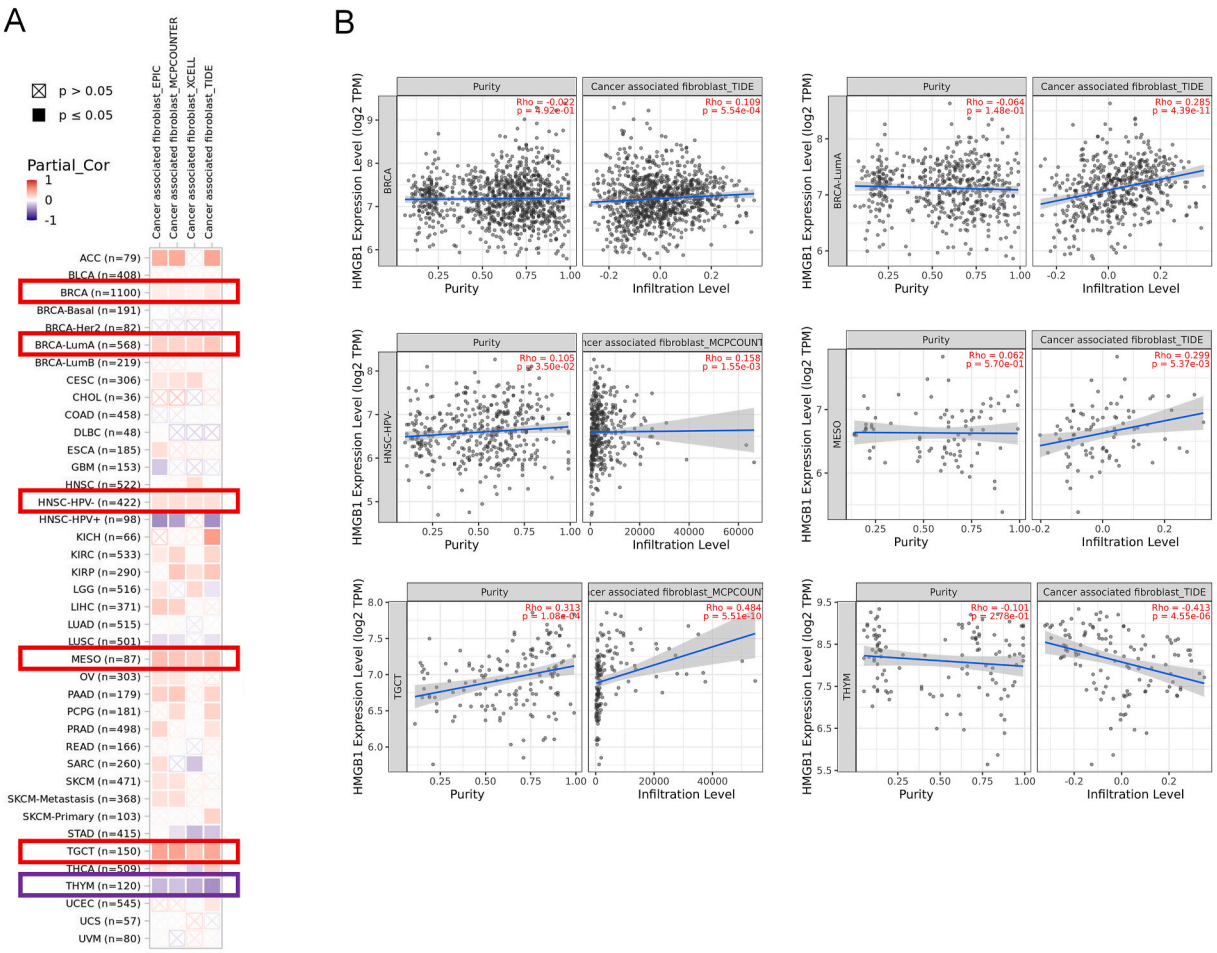

Fig. 5. Interrelation between CAF infiltration and *HMGB1* gene expression. (**A**) The infiltration degree of CAFs is analyzed using the XCELL, MCPCOUNT, TIDE, and EPIC algorithms for exploring the interrelation between *HMGB1* gene expression and CAF infiltration. (**B**) The interrelation between *HMGB1* gene expression and CAF infiltration across HNSC-HPV, MESO, BRCA, BRCA-LUMA, and TGCT.

association with the prognosis of distinct tumors. In most cancers, elevated *HMGB1* gene expression displayed relevance to unfavorable OS and DFS outcomes. However, in BLCA and THCA, low *HMGB1* gene expression had a connection to poor OS and DFS outcomes.

3.3. Genetic alteration landscape of HMGB1 in distinct cancer types

We examined *HMGB1* genetic alterations in distinct tumor subtypes from TCGA cohorts. The frequency of *HMGB1* genetic alterations was uncovered to be the highest in DLBC tumor samples (>4 %) (Fig. 4A). In this cancer subtype, mutation and deep deletions accounted for almost half of the variation. Moreover, the primary type of genetic alterations in COAD, STAD, and ESCA was found to be amplification, with an alteration frequency exceeding 2 %.

Fig. 4B provides a detailed summary of the types, sites, and frequency of *HMGB1* genetic alterations. Our analysis uncovered that the predominant alterations were nonsense and missense mutations in *HMGB1*. As an instance, a nonsense mutation (R163*) was detected in the HMG domain of one GBM case and another UCEC case, while a missense mutation (R163Q) was identified in the HMG domain of a single UCEC case. According to Fig. 4C, the R163Q mutation was in the 3D structure of *HMGB1*.

Additionally, we utilized the "Comparison/Survival" module for investigating the prospective interrelation between *HMGB1* genetic alterations and survival outcomes of COAD patients. The data in Fig. 4D indicated no significant difference in predicting COAD cases with *HMGB1* alterations compared with cases without *HMGB1* alterations.

3.4. ICI analysis data

Tumor-infiltrating immune cells (TIICs) are vital in the tumor microenvironment (TME), as they have a strong association with tumor growth, progression, and spread [21,22]. According to reports, CAFs can regulate the function of TIICs in the stroma of the TME [23,24]. Hence, we employed such algorithms as TIMER, XCELL, MCPCOUNT, TIDE, and EPIC to determine the potential correlations between *HMGB1* gene expression and CAFs or other immune cells across TCGA cancer species. Our analysis revealed that the degree of ICI had a positive connection to *HMGB1* gene expression in HNSC-HPV, MESO, BRCA, BRCA-LUMA, and TGCT. Conversely, in THYM, we observed a negative correlation. Our findings are presented in Fig. 5 as a heatmap and scatter plot.

3.5. Enrichment analysis data of HMGB1-related partners

With the objective of elucidating the biological mechanism developed by *HMGB1* towards cancer formation, we attempted to ascertain candidate proteins that bind to *HMGB1* for subsequent analysis of their interactions with this *HMGB1*. By use of the STRING online tool, we were able to isolate 50 *HMGB1*-binding proteins with confirmed scientific evidence. These proteins are presented in Fig. 6A as an interaction network diagram. Next, the GEPIA2 platform was employed to amalgamate the entire expression data of TCGA tumors and a list of the 100 most highly correlated genes with *HMGB1* was acquired. Our analysis uncovered that *HMGB1* gene expression had a positive connection to the expressions of Heterogeneous Nuclear Ribonucleoproteins A2/B1

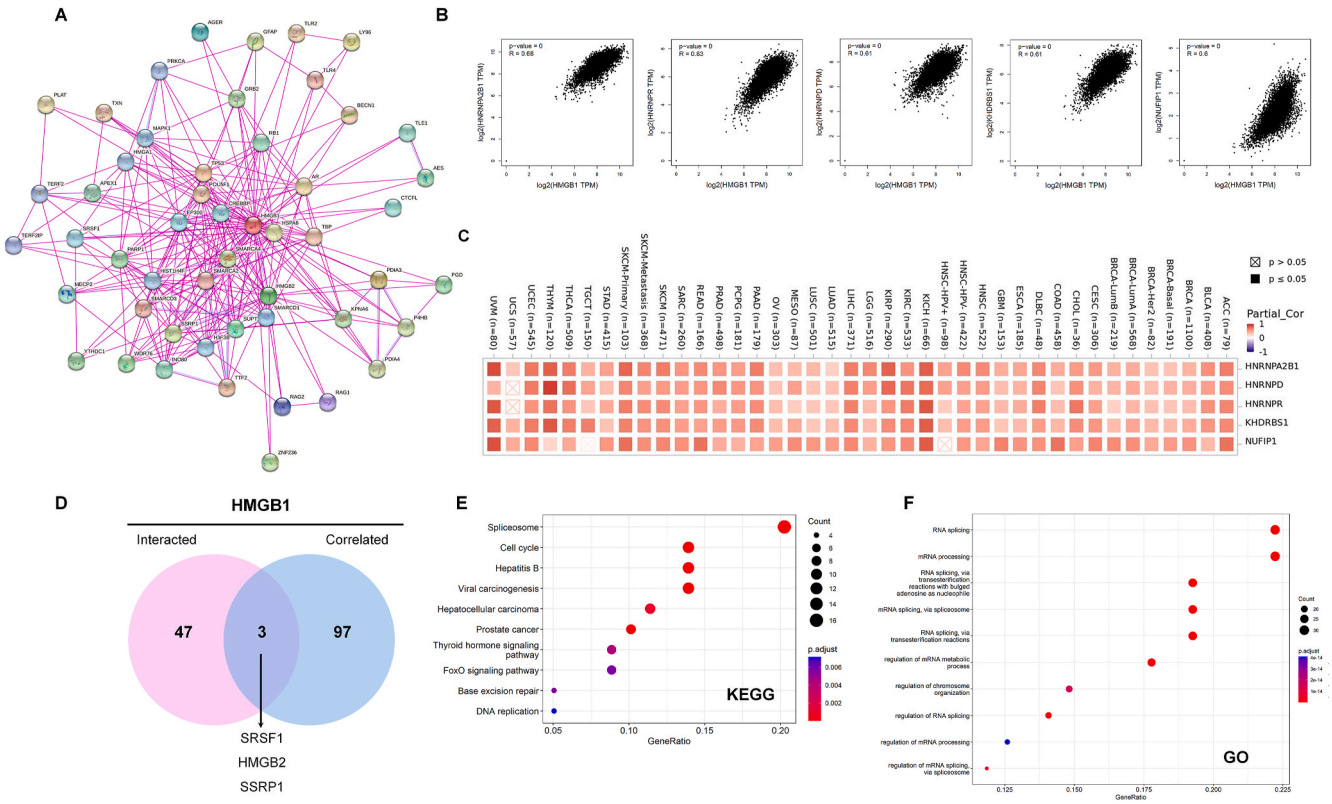

Fig. 6. Enrichment analysis results of genes associated with *HMGB1*. (**A**) The STRING tool is utilized to identify proteins that bind to *HMGB1*, according to experimental evidence. (**B**) The GEPIA2 tool is employed to analyze the 100 most highly correlated genes with *HMGB1*in TCGA datasets and assess the interrelation between *HMGB1* and chosen targeting genes, including HNRNPA2B1, HNRNPR, HNRNPD, KHDRBS1, and NUFIP1. (**C**) The heatmap displays corresponding data in exact cancer types. (**D**) The genes that bind to *HMGB1* and interact with it are subjected to an intersection assessment. (**E**) KEGG pathway analysis is implemented based on the genes that bind to *HMGB1* and interact with it. (**F**) The cnetplot function is used to display molecular function data in the GO analysis.

(HNRNPA2B1), Heterogeneous Nuclear Ribonucleoprotein R (HNRNPR), Heterogeneous Nuclear Ribonucleoprotein D (HNRNPD), KH RNA Binding Domain Containing, Signal Transduction Associated 1 (KHDRBS1) (R = 0.68, 0.63, 0.61, and 0.61, respectively), and Nuclear FMRP Interacting Protein 1 (NUFIP1) genes (R = 0.6) (all $P < 0.001$), as depicted in Fig. 6B. Additionally, the heatmap data for the majority of tumor types examined uncovered *HMGB1* had a positive connection with the five mentioned genes, as depicted in Fig. 6C. Through analysis of the intersection of the above two groups, SRSF1, HMGB2, and SSRP1 were identified as overlapping genes (Fig. 6D).

Two datasets were integrated, and KEGG and GO enrichment analyses were implemented to investigate potential pathways involved in tumor pathogenesis. The KEGG data showed that *HMGB1* might be involved in the "Spliceosome," "Cell Cycle," and "Hepatitis B″ pathways (Fig. 6E). Further analysis using GO enrichment unveiled that the majority of the genes identified had a connection to pathways associated with RNA metabolism, including RNA splicing (through transesterification reactions that involve a bulged adenosine as a nucleophile or by means of a spliceosome complex), mRNA processing, regulation of mRNA metabolic processes, and other related biological processes (Fig. 6F).

3.6. *HMGB1* genetic alterations and liver cancer

The above results consistently indicated that HMGB1 protein was highly expressed in HCC compared with normal tissues and correlated with the clinical prognosis of patients. Therefore, we analyzed systematically 95 primary human tumors by applying next-generation sequencing to characterize their somatic mutations, and the "maftools" R package was implemented to analyze and visualize the landscape of mutation profiles. The different genetic mutations in each sample were shown in the waterfall plot, with various color annotations representing different mutation types (Fig. 7A, Supplementary file 1). We detected mutations in nearly 84.21 % (80/95) of tumor samples, with the highest frequency of HMGB1 genetic alterations found in HCC tumor samples (>58 %, 7/12). Missense mutations accounted for the highest fraction (Fig. 7B), while single-nucleotide variant (SNVs) and insertions were more frequent than deletions and others (Fig. 7B),and the most common type of variant was C > T (Fig. 7C). Additionally, considering the total number of mutations and counting multiple hits alone, we recalculated the top 20 mutated genes (Fig. 7D), which showed slight differences from the previous list (Fig. 7A). In both comparisons, missense mutations accounted for the largest fraction, consistent with the results in Fig. 7A.

4. Discussion

In 1973, Ernest Johns, Graham Goodwin and colleagues extracted a set of nonhistone proteins from calf thymus chromatin [25,26]. Subsequently, these proteins were named "high mobility group" (HMG) proteins because of its rapid migration in polyacrylamide gel electrophoresis [27]. The HMGB family members are known to be implicated in various pathological progress, including inflammation [28], immune system disorders [29], and carcinogenesis [30]. A growing body of research has reported their essential role in various cancers, such as HCC [31], CRC [30], and lung cancer [32]. *HMGB1* is essential in the cell membrane, nucleus, cytoplasm, and extracellular space [33]. *HMGB1* regulates gene transcription within the nucleus by organizing DNA and nucleosomes [34–36]. Moreover, *HMGB1* exhibits high expression in the cytoplasm of the colon, liver, stomach, and brain

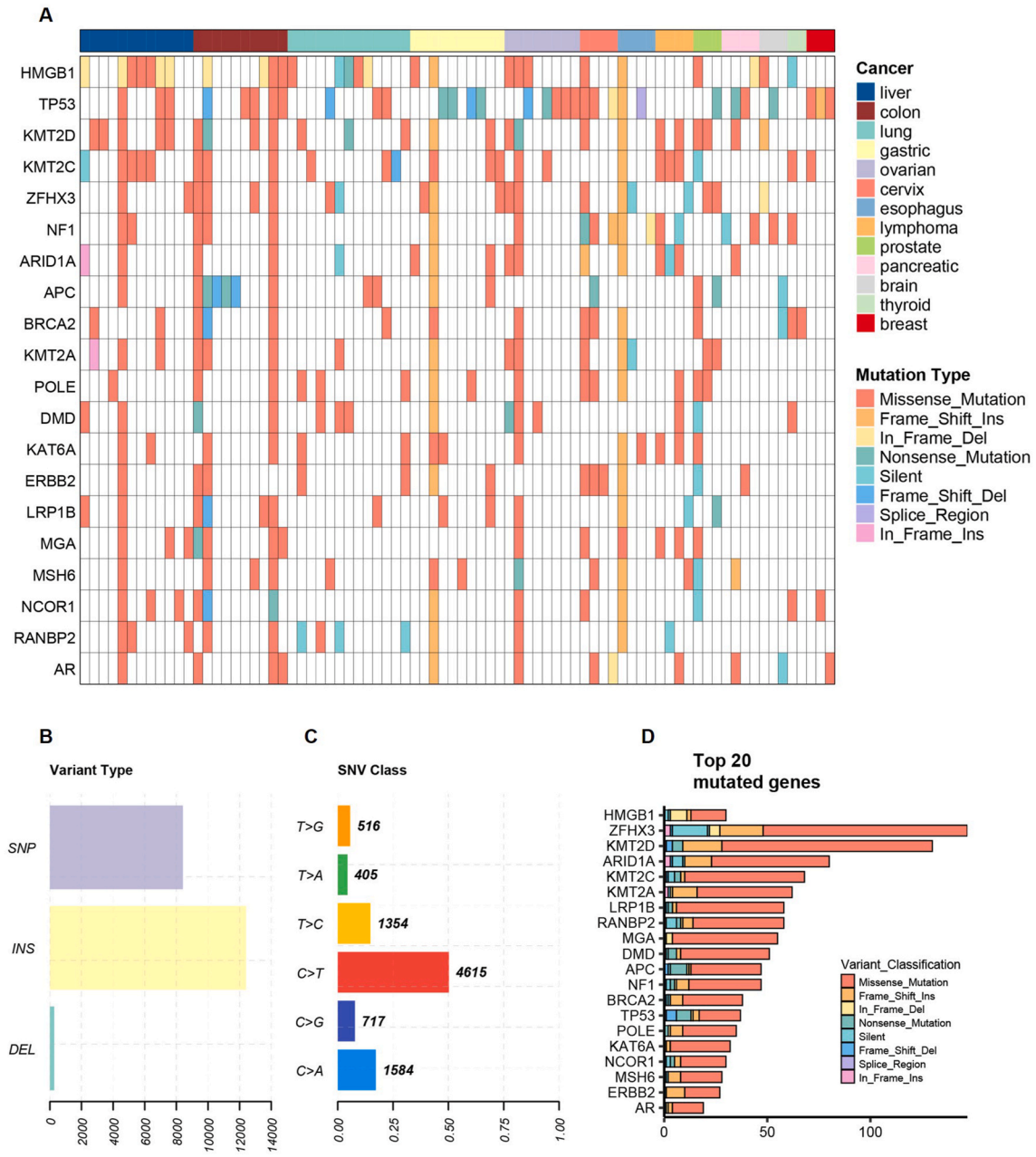

Fig. 7. Enrichment analysis results of genes associated with *HMGB1*. Landscape of mutation profiles in tumor samples. **(A)** Mutation information of each gene in each sample was shown in the waterfall plot, with various color annotations to distinguish different mutation types. **(B–C)** According to different classification categories, missense mutation, SNP, INS and C > T mutation accounted for a larger proportion. **(D)** Top 20 mutated genes in tumor samples. SNP, single-nucleotide polymorphism; SNV, single-nucleotide variant. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

tissues and co-localizes with enzyme proteins [33,37], indicating its vital role in regulating cytoplasmic autophagy. Hui Guan et al. pointed out that HMGB1 was not only differentially expressed in various tumors but also significantly correlated with both clinical prognosis and the immune microenvironment, it is predicted that HMGB1 may be a potential target for tumor therapy [38]. Although pan-cancer analysis of *HMGB1* have been provided in several studies, they focused on the role of HMGB1 in different tumors, the common molecular mechanisms underlying *HMGB1* function across different contexts are still not fully understood. We also note that several studies analyzed genetic variation of HMGB1 in different tumors in TCGA database, there is a lack of profiling in Asian or Chinese populations. Our objective was to comprehensively examine *HMGB1* gene expression, gene changes, ICI, and genetic enrichment in various tumors by combining data from TCGA, CPTAC, and GEO databases. In addition, we focused on the role of HMGB1 in liver cancer, and found that *HMGB1* genetic alterations was uncovered to be highest in liver cancer.

First, our results revealed that *HMGB1* exhibits high expression in most cancers, including BLCA, LIHC, CHOL, COAD, ESCA, and THYM., with low expression in a few cancers, such as UCEC and BRCA. *HMGB1* expression had distinct prognostic values in diverse cancer types. For instance, in THYM, the GEPIA2 tool was adopted to study how high *HMGB1* gene expression was statistically interrelated with OS ($P = 0.046$) and found that the higher the *HMGB1* gene expression, the better the OS of THYM. We further evaluated ICI in THYM expressing *HMGB1* using various algorithms and uncovered a significant inverse correlation

between ICI and CAFs (Fig. 4A).

HMGB1 is a conserved protein binding to chromatin and interacting with DNA. Apart from its well-established role in systemic inflammation as an extracellular mediator, it also exerts an early intermediary impact in hepatic ischemia-reperfusion injury [39]. The involvement of *HMGB1* in human malignancies, such as liver cancer, BRCA, prostate cancer, colon cancer, and malignant mesothelioma, is related to unfavorable outcomes [7,40–42]. Under hypoxic conditions, HCC cells can release *HMGB1*, which in turn facilitates tumor invasion and metastasis by producing various inflammatory mediators [43]. *In vitro* studies have shown that hepatitis B virus (HBV) X protein can enhance *HMGB1* secretion by HCC cells, with *HMGB1* gene expression levels in both the nucleus and cytoplasm positively correlated with the HBV load. As such, *HMGB1* holds promise as a candidate predictive biomarker and target for HBV-induced HCC therapy [44].

Enrichment analysis was employed in this study to construct an interaction network of the *HMGB1* protein, and the top 5 genes with the most significant relationship were identified. Our results revealed that *HMGB1* gene expression level had positive correlations with HNRNPA2B1, HNRNPD, HNRHPR, KHDRBS1, and NUFIP1 ($R = 0.68$, 0.63, 0.61, 0.61, and 0.6, respectively). We integrated *HMGB1*-binding protein and *HMGB1* information in all tumors and implemented KEGG and GO enrichment analyses. Our findings unveiled that *HMGB1* is probably implicated in the pathogenesis of "Hepatitis B," "Viral Carcinogenesis," "Hepatocellular Carcinoma", “Prostate cancer”, “Thyroid hormone signaling pathway” et al.

The above results consistently indicated that the expression level of HMGB1 was higher in the tumor tissue than that of normal tissue of Hepatocellular carcinoma (HCC) patients and was related to clinical prognosis. HCC is the most common type of primary liver cancer, and is usually precipitated by chronic viral infections (hepatitis B and C), non-alcoholic steatohepatitis, heavy alcohol use, and other factors which may lead to chronic inflammation and cirrhosis of the liver [45]. Therefore, we focused on the role of HMGB1 in liver cancer. By analyzing the somatic mutation profiles, we found that HMGB1 genetic alterations were higher in liver cancer (58.33 %, 7/12) compared to other tumors, such as colorectal (40.0 %, 4/6), lung (33.33 %, 5/15), ovarian (30.00 %, 3/10), brain (25.00 %, 1/4), gastric (20 %, 2/10), esophagus (20 %, 1/5), prostate (20 %, 1/5), cervix (20 %, 1/5), pancreatic (20 %, 1/5), lymphoma (20 %, 1/5), thyroid (20 %, 1/5), breast (0.00 %, 0/4) (Supplementary file 2). However, it cannot be ruled out that there may be some preferences due to the limited clinical sample size in this enrollment. We also found that SNP and INS, which could explain a significant proportion of the heritable risk of cancer [46], is the main variant type in cancer. In addition, HMGB1 was found in the top 20 mutated genes which means HMGB1 was also frequently identified in solid tumors.

In conclusion, this study presents a thorough investigation of *HMGB1* in various cancer types, elucidating the statistical associations of *HMGB1* gene expression with clinical outcomes, gene mutations, and ICI. By using clinical tumor samples, our study elucidates the involvement of *HMGB1* in tumorigenesis and provides valuable insights into developing anti-tumor strategies that target *HMGB1*. Therefore, experimental and clinical investigations are warranted to establish the prospective use of *HMGB1* in prognosis prediction and cancer treatment. Nonetheless, it is important to acknowledge the limitations inherent in our study. These limitations include a constrained sample size and an absence of comprehensive exploration into the relationship between HMGB1 and liver cancer. Future research should aim to elucidate these mechanisms in greater detail.

CRediT authorship contribution statement

Hui-min Yang: Writing – original draft, Methodology, Data curation, Conceptualization. **Xiang-ning Zhao:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Xiao-ling Li:** Writing – original draft, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Xi Wang:** Writing – original draft, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Yu Pu:** Software, Methodology, Formal analysis. **Dong-kai Wei:** Writing – review & editing, Software, Methodology, Investigation, Conceptualization. **Zhe Li:** Writing – review & editing, Conceptualization.

Ethics approval and consent to participate

The present investigation was conducted following the criteria for good clinical practice (GCP). Informed consent from the patients for the collection of their clinical data was obtained following the Declaration of Helsinki. The protocol was approved by the IRB of Shu guang Hospital affiliated with Shanghai University of TCM (2020-903-112-01).

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed in this study are included in this article and additional files.

Funding

No specific funding was received for this work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2024.101851.

Data availability

Data will be made available on request.

References

[1] B. Liu, H. Zhou, L. Tan, et al., Exploring treatment options in cancer: tumor treatment strategies, Signal Transduct. Targeted Ther. 9 (1) (2024) 175.
[2] X. Cui, X. Zhang, M. Liu, et al., A pan-cancer analysis of the oncogenic role of staphylococcal nuclease domain-containing protein 1 (SND1) in human tumors, Genomics 112 (6) (2020) 3958–3967.
[3] H. Xiao, K. Wang, D. Li, et al., Evaluation of FGFR1 as a diagnostic biomarker for ovarian cancer using TCGA and GEO datasets, PeerJ 9 (2021) e10817.
[4] A. Blum, P. Wang, J.C. Zenklusen, SnapShot: TCGAanalyzed tumors, Cell 173 (2018) 530.
[5] P. Scaffidi, T. Misteli, M.E. Bianchi, Release of chromatin protein HMGB1 by necrotic cells triggers inflammation, Nature 418 (6894) (2002) 191–195.
[6] H.X. Yan, H.P. Wu, H.L. Zhang, et al., p53 promotes inflammation-associated hepatocarcinogenesis by inducing HMGB1 release, Journal of hepatology 59 (4) (2013) 762–768.
[7] R. Kang, Q. Zhang, H.J. Zeh, et al., HMGB1 in cancer: good, bad, or both? HMGB1 in cancer, Clin. Cancer Res. 19 (15) (2013) 4046–4057.
[8] D. Tang, R. Kang, I.I.I.H.J. Zeh, et al., High-mobility group box 1 and cancer, Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1799 (1–2) (2010) 131–140.
[9] G.P. Sims, D.C. Rowe, S.T. Rietdijk, et al., HMGB1 and RAGE in inflammation and cancer, Annu. Rev. Immunol. 28 (2009) 367–388.

[10] U. Andersson, K.J. Tracey, HMGB1 is a therapeutic target for sterile inflammation and infection, Annu. Rev. Immunol. 29 (2011) 139.
[11] Y.Y. Xiang, D.Y. Wang, M. Tanaka, et al., Expression of high-mobility group-l mRNA in human gastrointestinal adenocarcinoma and corresponding non-cancerous mucosa, Int. J. Cancer 74 (1997) 1–6.
[12] R. Kang, K.M. Livesey, I.I.I.H.J. Zeh, et al., HMGB1: a novel Beclin 1-binding protein active in autophagy, Autophagy 6 (8) (2010) 1209–1211.
[13] D. Tang, R. Kang, K.M. Livesey, et al., Endogenous HMGB1 regulates autophagy, JCB (J. Cell Biol.) 190 (5) (2010) 881–892.
[14] K. Yang, F. Cao, W. Wang, et al., The relationship between HMGB1 and autophagy in the pathogenesis of diabetes and its complications, Front. Endocrinol. 14 (2023) 1141516.
[15] Z. Tang, B. Kang, C. Li, et al., GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis, Nucleic acids research 47 (W1) (2019) W556–W560.
[16] F. Chen, D.S. Chandrashekar, S. Varambally, et al., Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers, Nat. Commun. 10 (1) (2019) 1–15.
[17] Á. Nagy, G. Munkácsy, B. Győrffy, Pancancer survival analysis of cancer hallmark genes, Sci. Rep. 11 (1) (2021) 1–10.
[18] J. Gao, B.A. Aksoy, U. Dogrusoz, et al., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, Sci. Signal. 6 (269) (2013) pl1, pl1.
[19] E. Cerami, J. Gao, U. Dogrusoz, et al., The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data, Cancer Discov. 2 (5) (2012) 401–404.
[20] P. Bardou, J. Mariette, F. Escudié, et al., jvenn: an interactive Venn diagram viewer, BMC Bioinf. 15 (1) (2014) 1–7.
[21] S.X. Ge, D. Jung, R. Yao, ShinyGO: a graphical gene-set enrichment tool for animals and plants, Bioinformatics 36 (8) (2020) 2628–2629.
[22] A. Steven, B. Seliger, The role of immune escape and immune cell infiltration in breast cancer, Breast Care 13 (1) (2018) 16–21.
[23] W.H. Fridman, J. Galon, M.C. Dieu-Nosjean, et al., Immune infiltration in human cancer: prognostic significance and disease control, Cancer Immunol. Immunother. (2010) 1–24.
[24] M.Q. Kwa, K.M. Herum, C. Brakebusch, Cancer-associated fibroblasts: how do they contribute to metastasis? Clin. Exp. Metastasis 36 (2) (2019) 71–86.
[25] G.H. Goodwin, E.W. Johns, Isolation and characterisation of two calf-thymus chromatin non-histone proteins with high contents of acidic and basic amino acids, Eur. J. Biochem. 40 (1) (1973) 215–219.
[26] J. Xue, J.S. Suarez, M. Minaai, et al., HMGB1 as a therapeutic target in disease, J. Cell. Physiol. 236 (5) (2021) 3406–3419.
[27] R. Chen, R. Kang, D. Tang, The mechanism of HMGB1 secretion and release, Experimental & molecular medicine 54 (2) (2022) 91–102.
[28] HMG-1 as a late mediator of endotoxin lethality in mice, Science 285 (5425) (1999) 248–251.
[29] M.E. Bianchi, M.P. Crippa, A.A. Manfredi, et al., High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair, Immunol. Rev. 280 (1) (2017) 74–82.
[30] K.J. Cheng, M.A. Alshawsh, Mohamed E.H. Mejia, et al., HMGB1: an overview of its versatile roles in the pathogenesis of colorectal cancer, Cell. Oncol. 43 (2) (2020) 177–193.
[31] L. Wang, X.N. Xie, X.H. Song, et al., Upregulation of miR-200b inhibits hepatocellular carcinoma cell proliferation and migration by targeting HMGB3 protein, Technol. Cancer Res. Treat. 17 (2018) 1533033818806475.
[32] L. Wu, L. Yang, The function and mechanism of HMGB1 in lung cancer and its potential therapeutic implications, Oncol. Lett. 15 (5) (2018) 6799–6805.
[33] R. Kang, R. Chen, Q. Zhang, et al., HMGB1 in health and disease, Mol Aspects Med 40 (2014) 1–116.
[34] C. Bonne-Andrea, F. Harper, J. Sobczak, et al., The role of HMG1 protein in nucleosome assembly and in chromatin replication, Adv. Exp. Med. Biol. 179 (1984) 479–488.
[35] R. Kang, D. Tang, N.E. Schapiro, et al., The HMGB1/RAGE inflammatory pathway promotes pancreatic tumor growth by regulating mitochondrial bioenergetics, Oncogene 33 (5) (2014) 567–577.
[36] J. Singh, G.H. Dixon, High mobility group proteins 1 and 2 function as general class II transcription factors, Biochemistry 29 (26) (1990) 6295–6302.
[37] H. Lee, N. Shin, M. Song, et al., Analysis of nuclear high mobility group box 1 (HMGB1)-binding proteins in colon cancer cells: clustering with proteins involved in secretion and extranuclear function, J. Proteome Res. 9 (9) (2010) 4661–4670.
[38] H. Guan, M. Zhong, K. Ma, et al., The comprehensive role of high mobility group box 1 (HMGB1) protein in different tumors: a pan-cancer analysis, J. Inflamm. Res. (2023) 617–637.
[39] A. Tsung, R. Sahai, H. Tanaka, et al., The nuclear factor HMGB1 mediates hepatic injury after murine liver ischemia-reperfusion, The Journal of experimental medicine 201 (7) (2005) 1135–1143.
[40] S. Jube, Z.S. Rivera, M.E. Bianchi, et al., Cancer cell secretion of the DAMP protein HMGB1 supports progression in malignant MesotheliomaHMGB1 secretion supports mesothelioma growth, Cancer Res. 72 (13) (2012) 3290–3301.
[41] Y.D. Dong, L. Cui, C.H. Peng, et al., Expression and clinical significance of HMGB1 in human liver cancer: knockdown inhibits tumor growth and metastasis in vitro and in vivo, Oncol. Rep. 29 (1) (2013) 87–94.
[42] D. Wu, Y. Ding, S. Wang, Q. Zhang, L. Liu, Increased expression of high mobility group box 1(HMGB1) is associated with progression and poor prognosis in human nasopharyngeal carcinoma, J. Pathol. 216 (2008) 167–175.
[43] W. Yan, Y. Chang, X. Liang, et al., High-mobility group box 1 activates caspase-1 and promotes hepatocellular carcinoma invasiveness and metastases, Hepatology 55 (6) (2012) 1863–1875.
[44] S. Chen, Z. Dong, P. Yang, et al., Hepatitis B virus X protein stimulates high mobility group box 1 secretion and enhances hepatocellular carcinoma metastasis, Cancer letters 394 (2017) 22–32.
[45] K. Sankar, J. Gong, A. Osipov, et al., Recent advances in the management of hepatocellular carcinoma, Clin. Mol. Hepatol. 30 (1) (2024) 1.
[46] V. Fanfani, L. Citi, A.L. Harris, et al., The landscape of the heritable cancer genome, Cancer Res. 81 (10) (2021) canres.3348.