Pan-cancer analysis of BLM as a prognostic and immunological marker in human tumors

yifu hou (✉ 296724305@qq.com)
Sichuan Provincial People's Hospital: Sichuan Academy of Medical Sciences and Sichuan People's Hospital

Chuan LAN
Affiliated Hospital of North Sichuan Medical College

Dongme Ren
Affiliated Hospital of North Sichuan Medical College

Xi Wang
Sichuan Provincial People's Hospital: Sichuan Academy of Medical Sciences and Sichuan People's Hospital

Research

Keywords: Bloom syndrome helicase (BLM), DNA, TCGA, cancer genome atlas, Gene expression omnibus, prostate cancer (PC)

DOI: https://doi.org/10.21203/rs.3.rs-837352/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Pan-cancer analysis of BLM as a prognostic and immunological marker in human tumors

Authors
Yifu HOU1,2,3, Chuan LAN1, Dongmei Ren1, Xi WANG2
1Department of Hepatobiliary Surgery and Center of Severe Acute Pancreatitis, The Affiliated Hospital of North Sichuan Medical College, Nanchong 637100, China
2Department of Organ Transplantation, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu 610072, China
3Chinese Academy of Sciences Sichuan Translational Medicine Research Hospital, Chengdu 610072, China

*Correspondence:
Corresponding Author: Yifu HOU.
houyifu0726@foxmail.com
Email: Chuan LAN: hxlanchuan@163.com
Dongmei Ren: 838257310@qq.com
Xi WANG:1171118458@qq.com

Abstract
Background: Bloom syndrome helicase (BLM) is critical for DNA replication, repair, transcription, telomere maintenance and other cellular metabolic processes. Studies reported that BLM is closely related to the development of some types of tumors, however there is no pan-cancer analysis of BLM.

Results: BLM was over-expressed in almost all types of cancers and associated with overall survival and disease-free survival in multiple cancers. We found a stage-specific expression of BLM among them. A correlation of BLM expression with the estimated infiltration value of cancer-associated fibroblasts, endothelial cells, immunostimulator, and immunoinhibitor was discovered in part types of cancers. Finally, we found that BLM was associated with cell cycle, p53 signaling pathway, and other mechanisms.

Conclusion: Our study provided a relatively comprehensive understanding of BLM as an oncogene and predictor of prognosis in various tumor types.

Methods: We analyzed data from TCGA (The cancer genome atlas), GEO (Gene expression omnibus), and GTEx databases by using bioinformatic tools, to evaluate the expression difference, prognostic value, genetic alterations, immune infiltrates, and biological functions of BLM in 33 types of cancers.
Introduction

Genomic instability and mutations is one of the hallmarks of cancer which derived result from impairment of DNA damage repair[1]. Bloom syndrome helicase (BLM) is one of the important members of the RecQ family of DNA helicases, which is critical for DNA replication, repair, transcription, telomere maintenance and other cellular metabolic processes.[2,3,4] A rare disorder named Bloom’s syndrome is caused by germ-line mutation in the BLM gene, which is characterised by cancer predisposition, growth retardation, immunodeficiency, sunlight hypersensitivity and impaired fertility.[5] Recently, a few studies found that BLM is highly expressed in various kinds of tumors and partly correlated with poor prognosis. In summary with that, Kaul et al. concluded that BLM plays a role on the dual functions as a tumor suppressor or as a pro-oncogene.[6]

On one hand, the tumor suppressive function of BLM has also been identified by many studies. Studies using mouse models demonstrated that haploinsufficiency of BLM led to an early onset of lymphomas and colorectal cancer. [7,8] In mice, BLM acts as a factor essential for maintaining genomic stability and is involved in the prevention or reduction of tumor development. [9] Similar to mice models, patients with Bloom’s syndrome also develop a spectrum of cancers at a very early age, of which leukemia and lymphomas are the most common malignancies followed by colorectal cancer. Thus, the lack of functional BLM may hamper its ability to regulate the expression of these proto-oncogenes, causing the promotion of tumorigenesis.[10,11] On the other hand, BLM mRNA is reported to be overexpressed in almost all types of cancer tissues as compared with normal tissues.[12] It has been demonstrated that BLM promotes cell proliferation and lead poor prognosis of prostate cancer (PC). [13,14] Otherwise, Alzahrani et al found that BLM overexpression was also related to poor survival in lung and gastric cancer patients.[15] Moreover, BLM mRNA and protein levels are found to be overexpressed in patients with colorectal cancer. [16,17] Furthermore, accumulative evidences suggest that a significant BLM mRNA over-expression along with high BLM cytoplasmic localization has been associated with poor breast cancer-specific survival.[18,19,20]

Given the crucial role of BLM in tumorigenesis, we conducted a pan-cancer analysis of BLM expression with patient prognoses based on The Cancer Genome Atlas (TCGA) datasets, GEO datasets, and Genotype-Tissue Expression (GTEx) database. In addition, we explored the correlations between BLM expression and DNA methylation, immune infiltration, and genetic mutation by bioinformatic tools, such as tumor immune estimation resource, version 2 (TIMER2), GEPIA, cBioportal and so on. Our findings indicated statistical correlations of BLM gene expression with clinical prognosis, immune infiltration, and genetic mutation, which suggests that BLM is a potential prognostic biomarker.
Results

Gene Expression Analysis Data

We applied the TIMER2 tool to analyze the expression status of BLM across all cancer types of TCGA. As shown in Fig. 1A, the expression level of BLM in the tumor tissues of BLCA (Bladder urothelial carcinoma), BRCA, CESC (Cervical squamous cell carcinoma and endocervical adenocarcinoma), CHOL (Cholangiocarcinoma), COAD (Colon adenocarcinoma), ESCA (Esophageal carcinoma), HNSC (Head and neck squamous cell carcinoma), KIRC (Kidney renal clear cell carcinoma), KIRP (Kidney renal papillary cell carcinoma), GBM (Glioblastoma multiforme), LIHC (Liver hepatocellular carcinoma), LUAD, LUSC (Lung squamous cell carcinoma), PCPG (Pheochromocytoma and Paraganglioma), PRAD (Prostate adenocarcinoma), READ (Rectum adenocarcinoma) (P<0.001), SKCM (Skin Cutaneous Melanoma), STAD (Stomach adenocarcinoma) (P<0.01), THCA (Thyroid carcinoma), and UCEC (P<0.05) is higher than the corresponding control tissues, respectively.

Next, we further evaluated the expression difference of BLM between the normal tissues and tumor tissues after including the normal tissue of the GTEx dataset. As shown in Fig. 1B, except for the types of tumors mentioned above, the expression level of BLM is up-regulated in tumor tissue of ACC (Adrenocortical carcinoma), DLBC (Lymphoid neoplasm diffuse large B-cell lymphoma), LAML (Acute myeloid leukemia), LGG (Brain lower grade glioma), OV, PAAD (Pancreatic adenocarcinoma), TGCT (Testicular germ cell tumors), THYM (Thymoma), and UCS (Uterine carcinosarcoma). However, we did not observe a significant difference for the remnant tumors, such as KICH (Kidney chromophobe) and SARC (Sarcoma).

Next, we analyzed the total protein expression of BLM by using the CPTAC dataset, which showed that BLM is highly expressed in the tumor tissues of clear cell RCC (p<0.001), LUAD (p<0.05), and UCEC (p<0.01) than in normal tissues, whereas BLM is lower expressed in the tumor tissues of OV (p<0.05) (Fig. 1C).

We also used GEPIA2 approach to analyze the correlations between BLM expression and the pathological stages of cancers, which showed that a stage-specific expression of BLM in ACC, BRCA, KICH, KIRC, KIRP, LIHC, LUAD, and THCA (Fig. 1D, all P<0.05) but not others (Fig. S1).

Survival analysis data

In order to evaluate the prognostic value of BLM, we classified the cancer cases into high-expression and low-expression groups according to the expression levels of BLM and investigated the correlation of BLM expression with the prognosis of patients with various types of tumors based on the datasets of TCGA and GEO. As shown in Fig. 2A, high expression of BLM was associated with poor prognosis of OS for cancers of ACC (P<0.001), KIRC (P=0.0045), KIRP (P=0.0017), LGG (P<0.001), LUAD (P=0.039), MESO (P<0.001), PAAD (P=0.044), SARC (P=0.031), and SKCM (P=0.016). In contrast, the high
expression of BLM was associated with favor prognosis of OS for CESC (P=0.021), LUSC (P=0.018), and OV (P=0.029). With regard to DFS, the data (Fig. 2B) showed a correlation between high BLM expression and poor prognosis for the cancers including ACC (P=0.0039), KIRP (P=0.0064), LGG (P=0.001), LIHC (P=0.0048), MESO (Mesothelioma) (P=0.038), PAAD (P=0.0062), and PRAD (P<0.001). Using univariate Cox regression analysis, we found that increased BLM expression was a risk factor for ACC, CHOL, KICH, KIRC, LGG, MESO, and PAAD. In contrast, decreased BLM expression was a risk factor for CESC and THYM (Fig. 3).

**DNA methylation analysis data**

By using the UALCAN approach, we investigated the potential association between BLM DNA methylation and the pathogenesis of different tumors in the TCGA project. Compared with normal tissues, we observed an increased methylation level of BLM in tumor tissues of PAAD (P<0.001), LUSC (P<0.001), SARC (P<0.001), ESCA (P<0.01), and KIRC (P<0.001), but a reduced methylation level for TGCT (P<0.01), UCEC (P<0.01), and BLCA (P<0.001) (Fig. 4).

**Genetic alteration analysis data**

Using cBioportal tool, we observed the genetic alteration status of BLM in different tumor samples of the TCGA cohorts. As shown in Fig. 3A, the highest alteration frequency of BLM (almost 8%) appears for patients with undifferentiated stomach adenocarcinoma with “mutation” as the only type. The “amplification” type of CNA was the primary type in the cases of esophagogastric adenocarcinoma, which show an alteration frequency of ~5% (Fig. 4A). Besides, we found that all sarcoma cases with genetic alteration (>2% frequency) had amplification of BLM (Fig. 5A).

The types, sites and case number of the BLM genetic alteration are further presented in Fig. 4B. We found that missense mutation of BLM was the main type of genetic alteration, and N515Mfs*16 alteration was detected in 3 cases of colorectal adenocarcinoma, 2 cases of STAD and 1 case of UCEC (Fig. 5B).

Additionally, we analyzed the potential association between genetic alteration of BLM and the clinical survival prognosis of cases with different types of cancer. As shown in Fig. 5C, it indicated that cases with altered BLM showed better prognosis in disease-free (P=0.0318) survival, but not overall (P=0.383) and disease-specific (P=0.639) and progression-free (P=0.141) survival when compared with cases without BLM alteration.

**Immune Infiltration Analysis Data**

Tumor-infiltrating immune cells, as pivotal components of the tumor microenvironment, were tightly associated with the initiation, progression or metastasis of cancer [22, 23]. Cancer-associated fibroblasts in the stroma of the tumor microenvironment were reported to participate in modulating the function of various tumor-infiltrating immune cells [24, 25]. Thus, we applied the TIMER, CIBERSORT, CIBERSORT-ABS, TIDE, XCELL,
MCPCOUNTER, QUANTISEQ and EPIC algorithms to explore the correlation between the infiltration level of cancer-associated fibroblasts and endothelial cells and BLM expression in multiple tumor types of TCGA. Interestingly, we discovered a negative correlation of BLM expression and the estimated infiltration value of cancer-associated fibroblasts for the BRCA, BRCA-Her2, LUSC, STAD, TGCT, and THYM, whereas BLM expression was positively correlated with that in KIRP (Fig. 6). Meanwhile, we found a negative correlation of BLM expression and the endothelial cell infiltration for the BRCA, COAD, KIRC, LUAD, STAD, TGCT, THCA, and THYM, while a positive correlation was found between the endothelial cells and BLM expression in the tumors of GBM (Figure S2).

By using TISIDB, we also explored the association of BLM expression with immune-related chemokines, immunostimulator, and immunoinhibitor. The analysis of immune-related chemokines revealed that BLM was associated with most of the chemokines in KIRC, TGCT, and THCA. Among all tumor types, CCL14 and CX3CL1 were the most common chemokines which correlated with BLM expression (Fig. S3A). With regard to immunostimulator and immunoinhibitor, we found that BLM was positively correlated with most of them in KIRC, TGCT, and THCA, whereas BLM was negatively correlated with that in GBM and UCS (Fig. S3B-S3C).

**Enrichment analysis of BLM-related partners**

To further distinguish the molecular mechanism of the BLM gene in tumorigenesis, we attempted to filter out the targeting BLM-binding proteins and the BLM expression-correlated genes for pathway enrichment analyses. By using STRING tool, we obtained a total of 50 BLM-binding proteins, which were originated from experimental evidence. The interaction network of these proteins were shown at Fig. 7A. Meanwhile, we used the GEPIA2 tool to combine all tumor expression data of TCGA and acquired the top 100 genes that correlated with BLM expression. An intersection analysis by Jvenn tool of the above two groups showed 11 common genes, including BRCA1 (Breast cancer type 1 susceptibility protein), CHAF1A (Chromatin assembly factor 1 subunit A), CHEK1 (Serine/threonine-protein kinase Chk1), ERCC6L (DNA excision repair protein ERCC-6-like), EXO1 (Exonuclease 1), FANCA (Fanconi anemia group A protein), FANCD2 (Fanconi anemia group D2 protein), FEN1 (Flap endonuclease 1), RAD51 (DNA repair protein RAD51 homolog 1), RAD54L (DNA repair and recombination protein RAD54-like), and TIPIN (TIMELESS-interacting protein) (Fig. 7B). As shown in Fig. 7C, the BLM expression was positively correlated with that of BRCA1 (R=0.52), CHAF1A (R=0.49), CHEK1 (R=0.51), ERCC6L (R=0.48), EXO1 (R=0.52), FANCA (R=0.48), FANCD2 (R=0.55), FEN1 (R=0.48), RAD51 (R=0.52), RAD54L (R=0.52), and TIPIN (R=0.55) genes (all P<0.001). The Fig. 7D showed the corresponding heatmap data of a positive correlation between BLM and the above 11 genes in the majority of detailed cancer types.

Finally, we combined the two datasets to perform KEGG and GO enrichment analyses. The KEGG data of Fig. 7E suggest that “cell cycle”, “mismatch repair”, “cellular senescence”, “p53 signaling pathway”, “microRNAs in cancer”, and “platinum drug resistance” might be involved in the effect of BLM on tumor pathogenesis. In addition, the
GO enrichment analysis indicated that most of these genes are linked to “DNA replication” in the BP category, chromosomal region in the CC category, and ATPase activity in the MF category (Fig. 7F-7H).

Discussion

Dysregulation of DNA damage repair plays a fundamental role in carcinogenesis. BLM has important roles in the initiation and regulation of homologous recombination repair of DNA double-strand breaks. [2,3,4]. Therefore, it is reasonable to extrapolate that an optimal level of BLM is necessary to maintain genome stability. Both high and low levels or loss of BLM may lead to genomic instability and may eventually promote tumorigenesis. [15,6] Although recent evidence shows BLM mRNA to be over expressed in various cancers including colon, breast and prostate when compared with the normal samples. [16,18,13] However, the roles of BLM in human pan-cancer are not well understood. In this study, we identified the relationship of the BLM in multiple cancer tumorigenesis models via a pan-cancer analysis of TCGA, CPTAC, TIMER2, and GTEx databases.

According to the TCGA and GTEx dataset, the gene expression analysis showed that significantly higher expression level of BLM occurred almost in all types of malignances except KICH and SARC, as compared with corresponding normal tissues. Moreover, a stage-specific expression of BLM were found in ACC, BRCA, KICH, KIRC, KIRP, LIHC, LUAD, and THCA, which indicated role of BLM in tumorigenesis in various types of cancer. Although BLM was highly expressed in most tumors, the survival prognosis analysis data suggested distinct conclusions for different tumors. BLM was over-expressed and associated with worse OS in ACC, KIRC, KIRP, LGG, LUAD, MESO, PAAD and SKCM. In contrast, the high expression of BLM was associated with favor OS in CESC, LUSC and OV. However, BLM expression were not necessarily associated with prognosis in other types of cancer. Along with that, the results of cox regression showed that BLM expression was independent risk factor for OS in ACC, CESC, KIRC, LGG, MESO, and PAAD.

By now, several researches have testified the impact of BLM in several kinds of cancers. Ruan et al found that over-expressed BLM can either promote cell proliferation of PC3 cell lines in vitro study, or promote tumor growth by directly interacting with EZH2 in vitro and vivo study[14]. Otherwise, they also validated that BLM was highly expressed in tumor tissue when compared with normal and hyperplasia tissues, which was accordingly to the result of our pan-cancer analysis. Another study investigated the mRNA and protein expression of BLM in a huge amount of patients with breast cancer, which showed that a high expression of BLM was correlated with poor breast cancer specific survival. However, the relationship between BLM expression and overall survival was not displayed[18]. In addition, the role of BLM gene in cholangiocarcinoma was studied by Du et al[21]. This study showed that the expression of BLM was much higher in tumor tissue and cells than in non-tumor tissue and cells by using Immunohistochemical staining and qRT-PCR. After BLM was knockdown by siRNA, the proliferation and migration ability of CCA cells were
inhibited, while the cell cycle was arrested. This is highly consistent with our current research.

Gene mutations play an important role in the pathogenesis of some cancers [22]. In this study, the highest alteration frequency of BLM (almost 8%) appears for patients with undifferentiated stomach adenocarcinoma with “mutation” as the only type. We found that missense mutation of BLM was the main type of genetic alteration. These missense mutations have been shown to abolish the ATPase and DNA binding activity with some of them losing ATP binding activity, thus rendering the BLM protein catalytically inactive [23,24]. In literatures, BLM mutation that confers low-to-moderate penetrance risk for developing CRC [25,7] Our study indicated that cases with altered BLM showed better prognosis in disease-free survival. Specific gene mutations may predict patient prognosis and treatment response. However, the precise biological mechanism is not yet fully investigated.

The tumor microenvironment (TME) is a tumor promoting setting that tumor cells use to evade immune surveillance. The presence of the TME significantly influences therapeutic response and clinical outcome. Previous studies have identified various components that participate in the formation of the TME, including cancer-associated fibroblasts (CAF), endothelial cells, and so on [26,27,28,29] Surprisingly, we discovered a negative correlation of BLM expression and the estimated infiltration value of CAF and endothelial cell for a plethora of cancer. This result demonstrated the tumor suppressive function of BLM once again. CAF and endothelial cells serve pro-tumorigenic roles in the TME via secretion of various growth factors, cytokines, and chemokines and via degradation of the extracellular matrix [30,31]. This association between BLM and TME might be another reason for the prognostic implications of BLM in various cancers.

We characterized the function of differentially expressed BLM via GO enrichment analysis and KEGG pathway enrichment analysis. We found that differentially expressed BLM was mainly associated with regulation of cell cycle, mismatch repair, cellular senescence, p53 signaling pathway, microRNAs in cancer, and platinum drug resistance. The forementioned studies of prostate cancer and cholangiocarcinoma both proved that BLM was associated with p53 signaling pathway and cell cycle. This, in turn, confirms our pan-cancer analysis. Whether BLM participated in the mentioned pathways in other types of cancer needs further investigation.

Conclusion

Regarding genomic instability and mutations is one of the hallmarks of cancer which derived result from impairment of DNA damage repair, there has been increasingly more attention for BLM because of its critical role on DNA replication, repair, transcription, telomere maintenance. By using bioinformatic analysis, our study provided a relatively comprehensive understanding of BLM as an oncogene and predictor of prognosis in various tumor types.

Materials and methods
Gene Expression Analysis
In this study, we used TIMER2 (tumor immune estimation resource, version 2, http://timer.cistrome.org/) to analyze the expression differences of BLM between tumor types and adjacent normal tissues. For partial types of tumors without or only with limited normal tissue [e.g., TCGA-DLBC (Lymphoid neoplasm diffuse large B-cell lymphoma), TCGA-PAAD (Pancreatic adenocarcinoma), etc.], we analyzed the BLM expression based on combining the data for normal tissues from the GTEx (Genotype-Tissue Expression) database with data from The Cancer Genome Atlas (TCGA). All expression data were normalized via log2 conversion. The GEPIA2 (Gene Expression Profiling Interactive Analysis, version 2) tool (http://gepia2.cancer-pku.cn/#analysis) was used to analyze the BLM expression in different pathological stages of all TCGA cancers. The log2 [TPM (Transcripts per million) +1] transformed expression data were applied for the box or violin plots. Additionally, we used UALCAN tool (http://ualcan.path.uab.edu/analysis prot.html) to conduct the protein expression analysis of the CPTAC (Clinical proteomic tumor analysis consortium) dataset based on cancer Omics data(15). Expression level of the total protein of BLM has been compared between primary and normal tissues from available datasets of five tumors, which included BRCA (breast cancer), OV (ovarian cancer), KIRC (Kidney renal clear cell carcinoma), LUAD (Lung adenocarcinoma), and UCEC (Uterine corpus endometrial carcinoma).

Survival Prognosis Analysis
We used GEPIA2 tool to obtain the OS (Overall Survival) and DFS (Disease-Free Survival) significance map data and survival plots of BLM across all TCGA tumors. We used cutoff-high (50%) and cutoff low (50%) values as the expression thresholds to split the high-expression and low-expression cohorts (16). The log-rank test was used in the hypothesis testing. For survival analysis, the HR and p value were calculated using univariate Cox regression analysis.

DNA methylation analysis data
We also used the UALCAN tool to investigate the potential association between BLM DNA methylation and the pathogenesis of different tumors in the TCGA project.

Genetic Alteration Analysis
We used cBioPortal tool (https://www.cbioportal.org/) to obtain the data of alteration frequency, mutation type, mutated site information, and CNA (Copy number alteration) across all TCGA tumors. Survival data, including the overall, disease-free, progression-free, and disease-free survival differences were compared for all the TCGA cancer types, with or without BLM genetic alteration. Kaplan-Meier plots with log-rank P-value were generated as well.
Immune Infiltration Analysis

TIMER2 tool was also used to investigate the relationship between BLM expression and immune infiltrates across all TCGA tumors. Cancer-associated fibroblast and endothelial cell were selected for detailed analysis. The TIMER, TIDE, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, MCPCOUNTER and EPIC algorithms were applied for estimations. The data were visualized as a heatmap and a scatter plot. Additionally, we used TISIDB website (http://cis.hku.hk/TISIDB/index.php) for the analysis of immune-related chemokines, immunostimulator, and immunoinhibitor. The P-values and partial correlation (cor) values were obtained via the purity-adjusted Spearman’s rank correlation test.

BLM-Related Gene Enrichment Analysis

We firstly used STRING website (https://string-db.org/) for the analysis of protein-protein interaction network. The main parameters were: minimum required interaction score (“Low confidence (0.150)”), meaning of network edges (“evidence”), max number of interactors to show (“no more than 50 interactors” in 1st shell) and active interaction sources (“experiments”). Secondly, we used GEPIA2 tool to obtain the top 100 BLM-correlated genes based on the datasets of all TCGA tumor and normal tissues. Then we used Jvenn, an interactive Venn diagram viewer [19], to conduct an intersection analysis to compare the BLM-binding and interacted genes. After we obtained the overlapping molecules from Jvenn, we conducted a pairwise gene-gene Pearson correlation analysis between BLM and the selected genes by using TIMER2 tool. The P-values and the correlation coefficient (R) were calculated and are indicated in the corresponding figure panels. Heatmap representation of the expression profile for the selected genes contains the partial correlation (cor) and P-value in the purity adjusted Spearman’s rank correlation test.

We merged and filtered the two sets of data to perform KEGG (Kyoto encyclopedia of genes and genomes) pathway analysis. The enriched pathways were visualized as cnetplots with the “tidy” and “ggplot2” R packages. In addition, we applied the “clusterProfiler” R package to conduct GO (Gene ontology) enrichment analysis. The data for BP (Biological process), CC (Cellular component), and MF (Molecular function) were visualized as bubble charts. The R language software [R-3.6.3, 64-bit] (https://www.r-project.org/) was used in this analysis. Two-tailed P<0.05 was considered to be statistically significant (17).

FIGURE LEGENDS

Fig. 1. Expression level of BLM gene in different tumors and pathological stages. (A) The expression status of the BLM gene in various types of cancers or specific cancer subtypes based on TCGA database. ** P<0.01; *** P<0.001. (B) The expression status of the BLM gene in different
cancers based on TCGA and GTEx databases. **P<0.01; ***P<0.001. (C) The expression level of BLM total protein between normal tissue and primary tissue of KIRC, LUAD, OV, and UCEC based on the CPTAC dataset. *P<0.05; ***P<0.001. (D) The expression levels of the BLM gene in varied main pathological stages (stage I, stage II, stage III, and stage IV) of ACC, BLCA, KICH, KIRC, KIRP, LIHC, LUAD, and THCA based on the TCGA data. Log2 (TPM+1) was applied for log-scale.

**Fig. 2. Correlation between BLM gene expression and survival prognosis of different cancers in TCGA.** Overall survival (A) and disease-free survival (B) analyses of different cancers in TCGA by BLM gene expression. The survival map and Kaplan-Meier curves with positive results are shown.

**Fig. 3. Association between BLM gene expression and the overall survival of cancer patients.** The forest map shows the results of Univariate Cox Regression analysis for OS in various types of tumors.

**Fig. 4. Pan-cancer analysis of DNA methylation of BLM.** The positive results of correlation between DNA methylation and mRNA expression of BLM in TCGA pan-cancer are shown.

**Fig. 5. Mutation feature of BLM in different tumors of TCGA.** We analyzed the mutation features of BLM for the TCGA tumors by using the cBioPortal tool. The alteration frequency with mutation type (A) and mutation site (B) are shown. The potential correlation between mutation status and overall, disease-specific, disease-free and progression-free survival (C) by using the cBioPortal tool are displayed.

**Fig. 6. Correlation analysis between BLM expression and immune infiltration of cancer-associated fibroblasts.** Different algorithms explored the potential correlations of (A) expression of the BLM gene and (B) infiltration of cancer-associated fibroblasts across all types of cancer in TCGA.

**Fig. 7. BLM-related gene enrichment analysis.** (A) The available experimentally determined BLM-binding proteins by using the STRING tool. (B) An intersection analysis of the BLM-binding and top 100 BLM-correlated genes in TCGA projects using the GEPIA2 was analyzed, including BRCA1, CHAF1A, CHEK1, ERCC6L, EXO1, FANCA, FANCD2, FEN1, RAD51, RAD54L, and TIPIN. (C) The corresponding heatmap data in the detailed cancer types are displayed. (D) The expression correlation between BLM and selected targeting genes. (E) The cnetplot for KEGG pathway analysis was performed based on the BLM-binding and interacted genes. The bubble charts for the biological process (BP) (F), cellular component (CC) (G), and molecular function (MF) (H) in GO analysis are also shown.

**Acknowledgments**

We would like to acknowledge the Oncomine, GEPIA, Kaplan-Meier Plotter, THPA, TCGA, cBioPortal, STRING, GeneMANIA, and DAVID databases, which
were used free of charge. Meanwhile, we would like to thank Editage (www.editage.cn) for English language editing.

**Author Contributions**

Chuan LAN and Yifu HOU designed the study and analyze the bioinformatics. Dongmei REN, xi WANG prepared the figures and tables, also typeset the references. Chuan LAN drafted the manuscript and Yifu HOU revised it critically for important intellectual content. All authors read and approved the final manuscript.

**Funding**

None.

**Availability of data and materials**

The raw data of this study are derived from the Oncomine, GEPIA, Kaplan-Meier Plotter, THPA, TCGA, cBioPortal, STRING, GeneMANIA, and DAVID databases, which are publicly available databases.

**Ethics approval and consent to participate**

Not necessary.

**Consent for publication**

Not applicable.

**Conflict of Interest**

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.
References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646-74. [Hallmarks of cancer: the next generation - PubMed (nih.gov)]

2. Sharma S, Doherty KM, Brosh RM Jr. Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability. Biochem J. 2006 Sep 15;398(3):319-37. [Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability - PubMed (nih.gov)]

3. Croteau DL, Popuri V, Opresko PL, Bohr VA. Human RecQ helicases in DNA repair, recombination, and replication. Annu Rev Biochem. 2014;83:519-52. [Human RecQ helicases in DNA repair, recombination, and replication - PubMed (nih.gov)]

4. Brosh RM Jr. DNA helicases involved in DNA repair and their roles in cancer. Nat Rev Cancer. 2013 Aug;13(8):542-58. [DNA helicases involved in DNA repair and their roles in cancer - PubMed (nih.gov)]

5. Cheok CF, Bachrati CZ, Chan KL, Ralf C, Wu L, Hickson ID. Roles of the Bloom's syndrome helicase in the maintenance of genome stability. Biochem Soc Trans. 2005 Dec;33(Pt 6):1456-9. [Roles of the Bloom's syndrome helicase in the maintenance of genome stability - PubMed (nih.gov)]

6. Kaur E, Agrawal R, Sengupta S. Functions of BLM Helicase in Cells: Is It Acting Like a Double-Edged Sword? Front Genet. 2021 Mar;12:634789. [Functions of BLM Helicase in Cells: Is It Acting Like a Double-Edged Sword? - PubMed (nih.gov)]

7. de Voer RM, Hahn MM, Mensenkamp AR, Hoischen A, Gilissen C, Henkes A, Spruijt L, van Zelst-Stams WA, Kets CM, Verwiel ET, Nagtegaal ID, Schackert HK,
van Kessel AG, Hoogerbrugge N, Ligtenberg MJ, Kuiper RP. Deleterious Germline BLM Mutations and the Risk for Early-onset Colorectal Cancer. Sci Rep. 2015 Sep 11;5:14060.Deleterious Germline BLM Mutations and the Risk for Early-onset Colorectal Cancer - PubMed (nih.gov)

8. Ronkin MA, Dushanova GA, Chistiakova MA. Sostoianie mozgovogo krovoobrashcheniia u bol'nykh khronicheskimi nespetsificheskimi zabolevaniiami legkikh po dannym reoentselografii [Cerebrovascular circulation in chronic nonspecific lung diseases according to rheoencephalographic data]. Sov Med. 1975 Dec;(12):36-40.

9. McDaniel LD, Chester N, Watson M, Borowsky AD, Leder P, Schultz RA. Chromosome instability and tumor predisposition inversely correlate with BLM protein levels. DNA Repair (Amst). 2003 Dec 9;2(12):1387-404.Chromosome instability and tumor predisposition inversely correlate with BLM protein levels - PubMed (nih.gov)

10. German J, Sanz MM, Ciocci S, Ye TZ, Ellis NA. Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. Hum Mutat. 2007 Aug;28(8):743-53. doi: 10.1002/humu.20501.Syndrome-causing mutations of the BLM gene in persons in the Bloom’s Syndrome Registry - PubMed (nih.gov)

11. Cunniff C, Bassetti JA, Ellis NA. Bloom's Syndrome: Clinical Spectrum, Molecular Pathogenesis, and Cancer Predisposition. Mol Syndromol. 2017 Jan;8(1):4-23.Bloom’s Syndrome: Clinical Spectrum, Molecular Pathogenesis, and Cancer Predisposition - PubMed (nih.gov)

12. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, Varambally S. UALCAN: A Portal for
Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia.

2017 Aug;19(8):649-658.  
[UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses – PubMed (nih.gov)]

13.Ledet EM, Antonarakis ES, Isaacs WB, Lotan TL, Pritchard C, Sartor AO. 
Germline BLM mutations and metastatic prostate cancer. Prostate. 2020 Feb;80(2):235-237. [Germline BLM mutations and metastatic prostate cancer – PubMed (nih.gov)]

14.Ruan Y, Xu H, Ji X, Zhao J. BLM interaction with EZH2 regulates MDM2 expression and is a poor prognostic biomarker for prostate cancer. Am J Cancer Res. 2021 Apr 15;11(4):1347-1368. [BLM interaction with EZH2 regulates MDM2 expression and is a poor prognostic biomarker for prostate cancer – PubMed (nih.gov)]

15.Alzahrani FA, Ahmed F, Sharma M, Rehan M, Mahfuz M, Baeshen MN, Hawsawi Y, Almatrafi A, Alsagaby SA, Kamal MA, Warsi MK, Choudhry H, Jamal MS. Investigating the pathogenic SNPs in BLM helicase and their biological consequences by computational approach. Sci Rep. 2020 Jul 23;10(1):12377. [Investigating the pathogenic SNPs in BLM helicase and their biological consequences by computational approach – PubMed (nih.gov)]

16.Lao VV, Welcsh P, Luo Y, Carter KT, Dzieciatkowski S, Dintzis S, Meza J, Sarvetnick NE, Monnat RJ Jr, Loeb LA, Grady WM. Altered RECQ Helicase Expression in Sporadic Primary Colorectal Cancers. Transl Oncol. 2013 Aug 1;6(4):458-69. [Altered RECQ Helicase Expression in Sporadic Primary Colorectal Cancers – PubMed (nih.gov)]

17.Votino C, Laudanna C, Parcesepe P, Giordano G, Remo A, Manfrin E, Pancione M. Aberrant BLM cytoplasmic expression associates with DNA
damage stress and hypersensitivity to DNA-damaging agents in colorectal cancer. J Gastroenterol. 2017 Mar;52(3):327-340.

Aberrant BLM cytoplasmic expression associates with DNA damage stress and hypersensitivity to DNA-damaging agents in colorectal cancer - PubMed (nih.gov)

18. Arora A, Abdel-Fatah TM, Agarwal D, Doherty R, Moseley PM, Aleskandarany MA, Green AR, Ball G, Alshareeda AT, Rakha EA, Chan SY, Ellis IO, Madhusudan S. Transcriptomic and Protein Expression Analysis Reveals Clinicopathological Significance of Bloom Syndrome Helicase (BLM) in Breast Cancer. Mol Cancer Ther. 2015 Apr;14(4):1057-65. Transcriptomic and Protein Expression Analysis Reveals Clinicopathological Significance of Bloom Syndrome Helicase (BLM) in Breast Cancer - PubMed (nih.gov)

19. Zhu X, Chen H, Yang Y, Xu C, Zhou J, Zhou J, Chen Y. Distinct prognosis of mRNA expression of the five RecQ DNA-helicase family members - RECQL, BLM, WRN, RECQL4, and RECQL5 - in patients with breast cancer. Cancer Manag Res. 2018 Dec 5;10:6649-6668. Distinct prognosis of mRNA expression of the five RecQ DNA-helicase family members - RECQL, BLM, WRN, RECQL4, and RECQL5 - in patients with breast cancer - PubMed (nih.gov)

20. Zhang W, Yang S, Liu J, Bao L, Lu H, Li H, Pan W, Jiao Y, He Z, Liu J. Screening antiproliferative drug for breast cancer from bisbenzylisoquinoline alkaloid tetrandrine and fangchinoline derivatives by targeting BLM helicase. BMC Cancer. 2019 Oct 28;19(1):1009. Screening antiproliferative drug for breast cancer from bisbenzylisoquinoline alkaloid tetrandrine and fangchinoline derivatives by targeting BLM helicase - PubMed (nih.gov)

21. Du X, Zhang C, Yin C, Wang W, Yan X, Xie D, Zheng X, Zheng Q, Li M, Song Z. High BLM Expression Predicts Poor Clinical Outcome and Contributes to Malignant Progression in Human Cholangiocarcinoma. Front Oncol. 2021 Mar
High BLM Expression Predicts Poor Clinical Outcome and Contributes to Malignant Progression in Human Cholangiocarcinoma - PubMed (nih.gov)

22. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A, Lopez-Bigas N. A compendium of mutational cancer driver genes. Nat Rev Cancer. 2020 Oct;20(10):555-572. A compendium of mutational cancer driver genes - PubMed (nih.gov)

23. Rong SB, Väliaho J, Vihinen M. Structural basis of Bloom syndrome (BS) causing mutations in the BLM helicase domain. Mol Med. 2000 Mar;6(3):155-64. Structural basis of Bloom syndrome (BS) causing mutations in the BLM helicase domain - PubMed (nih.gov)

24. Guo RB, Rigolet P, Ren H, Zhang B, Zhang XD, Dou SX, Wang PY, Amor-Gueret M, Xi XG. Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. Nucleic Acids Res. 2007;35(18):6297-310. Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein - PubMed (nih.gov)

25. Sokolenko AP, Iyevleva AG, Preobrazhenskaya EV, Mitiushkina NV, Abyshcheva SN, Suspitsin EN, Kuligina ESh, Gorodnova TV, Pfeifer W, Togo AV, Turkevich EA, Ivantsov AO, Voskresenskiy DV, Dolmatov GD, Bit-Sava EM, Matsko DE, Semiglazov VF, Fichtner I, Larionov AA, Kuznetsov SG, Antoniou AC, Imyanitov EN. High prevalence and breast cancer predisposing role of the BLM c.1642 C>T (Q548X) mutation in Russia. Int J Cancer. 2012 Jun
15;130(12):2867-73. High prevalence and breast cancer predisposing role of the BLM c.1642 C>T (Q548X) mutation in Russia - PubMed (nih.gov)

26. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med. 2013 Nov;19(11):1423-37. Microenvironmental regulation of tumor progression and metastasis - PubMed (nih.gov)

27. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. Science. 2015 Apr 3;348(6230):74-80. T cell exclusion, immune privilege, and the tumor microenvironment - PubMed (nih.gov)

28. Ridge SM, Sullivan FJ, Glynn SA. Mesenchymal stem cells: key players in cancer progression. Mol Cancer. 2017 Feb 1;16(1):31. Mesenchymal stem cells: key players in cancer progression - PubMed (nih.gov)

29. Kalluri R. The biology and function of fibroblasts in cancer. Nat Rev Cancer. 2016 Aug 23;16(9):582-98. The biology and function of fibroblasts in cancer - PubMed (nih.gov)

30. Liao Z, Tan ZW, Zhu P, Tan NS. Cancer-associated fibroblasts in tumor microenvironment - Accomplices in tumor malignancy. Cell Immunol. 2019 Sep;343:103729. Cancer-associated fibroblasts in tumor microenvironment - Accomplices in tumor malignancy - PubMed (nih.gov)

31. Sobierajska K, Ciszewski WM, Sacewicz-Hofman I, Niewiarowska J. Endothelial Cells in the Tumor Microenvironment. Adv Exp Med Biol. 2020;1234:71-86. Endothelial Cells in the Tumor Microenvironment - PubMed (nih.gov)
**Figure 1**

Expression level of BLM gene in different tumors and pathological stages. (A) The expression status of the BLM gene in various types of cancers or specific cancer subtypes based on TCGA database. **P<0.01; *** P<0.001. (B) The expression status of the BLM gene in different cancers based on TCGA and GTEx datasets.
GTEx databases. ** P<0.01; *** P<0.001. (C) The expression level of BLM total protein between normal tissue and primary tissue of KIRC, LUAD, OV, and UCEC based on the CPTAC dataset. * P<0.05; *** P<0.001. (D) The expression levels of the BLM gene in varied main pathological stages (stage I, stage II, stage III, and stage IV) of ACC, BLCA, KICH, KIRC, KIRP, LIHC, LUAD, and THCA based on the TCGA data. Log2 (TPM+1) was applied for log-scale.
Correlation between BLM gene expression and survival prognosis of different cancers in TCGA. Overall survival (A) and disease-free survival (B) analyses of different cancers in TCGA by BLM gene expression. The survival map and Kaplan-Meier curves with positive results are shown.

| Tumor type | HR (95% CI)   | P value |
|------------|---------------|---------|
| ACC        | 4.916(1.988-12.156) | <0.001  |
| BLCA       | 0.980(0.732-1.312)    | 0.892   |
| BRCA       | 1.081(0.785-1.487)    | 0.633   |
| CESC       | 0.604(0.376-0.971)    | 0.038   |
| CHOL       | 3.253(1.119-9.456)    | 0.03    |
| COAD       | 0.888(0.602-1.309)    | 0.548   |
| DLBC       | 0.414(0.083-2.062)    | 0.282   |
| ESCA       | 0.992(0.606-1.624)    | 0.976   |
| GBM        | 0.933(0.665-1.308)    | 0.686   |
| HNSC       | 1.117(0.855-1.458)    | 0.417   |
| KICH       | 8.214(1.027-15.896)   | 0.047   |
| KIRC       | 1.563(1.155-2.116)    | 0.004   |
| LAML       | 0.956(0.628-1.456)    | 0.834   |
| LGG        | 2.026(1.421-2.886)    | <0.001  |
| LUAD       | 1.213(0.910-1.616)    | 0.187   |
| LUSC       | 0.812(0.620-1.064)    | 0.132   |
| MESO       | 2.590(1.588-4.223)    | <0.001  |
| OV         | 0.900(0.695-1.166)    | 0.424   |
| PAAD       | 1.706(1.125-2.587)    | 0.012   |
| PCPGG      | 2.512(0.499-12.648)   | 0.264   |
| PRAD       | 1.590(0.442-5.713)    | 0.478   |
| READ       | 0.522(0.238-1.146)    | 0.105   |
| SARC       | 1.451(0.975-2.161)    | 0.106   |
| SKCM       | 1.250(0.954-1.638)    | 0.106   |
| STAD       | 0.793(0.571-1.100)    | 0.165   |
| THCA       | 1.073(0.398-2.893)    | 0.89    |
| THYM       | 0.177(0.036-0.879)    | 0.034   |
| UCEC       | 1.253(0.830-1.991)    | 0.283   |
| UCS        | 0.938(0.477-1.844)    | 0.854   |
| UVM        | 1.279(0.557-2.939)    | 0.562   |

Figure 3

Association between BLM gene expression and the overall survival of cancer patients. The forest map shows the results of Univariate Cox Regression analysis for OS in various types of tumors.
Figure 4

Pan-cancer analysis of DNA methylation of BLM. The positive results of correlation between DNA methylation and mRNA expression of BLM in TCGA pan-cancer are shown.
Figure 5

Mutation feature of BLM in different tumors of TCGA. We analyzed the mutation features of BLM for the TCGA tumors by using the cBioPortal tool. The alteration frequency with mutation type (A) and mutation site (B) are shown. The potential correlation between mutation status and overall, disease-specific, disease-free and progression-free survival (C) by using the cBioPortal tool are displayed.
Correlation analysis between BLM expression and immune infiltration of cancer-associated fibroblasts. Different algorithms explored the potential correlations of (A) expression of the BLM gene and (B) infiltration of cancer-associated fibroblasts across all types of cancer in TCGA.
Figure 7

BLM-related gene enrichment analysis. (A) The available experimentally determined BLM-binding proteins by using the STRING tool. (B) An intersection analysis of the BLM-binding and top 100 BLM-correlated genes in TCGA projects using the GEPIA2 was analyzed, including BRCA1, CHAF1A, CHEK1, ERCC6L, EXO1, FANCA, FANCD2, FEN1, RAD51, RAD54L, and TIPIN. (C) The corresponding heatmap data in the detailed cancer types are displayed. (D) The expression correlation between BLM and selected targeting
genes. (E) The cnetplot for KEGG pathway analysis was performed based on the BLM-binding and interacted genes. The bubble charts for the biological process (BP) (F), cellular component (CC) (G), and molecular function (MF) (H) in GO analysis are also shown.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.pdf