Analysis of mRNA expression differences in bladder cancer metastasis based on TCGA datasets

**Sha Liu[1,2], Jiazhong Shi[1], Yuting Liu[1],
Liwei Wang[3], Jingqi Zhang[3], Yaqin Huang[1],
Zhiwen Chen[3] and Jin Yang[1]** 🆔

Abstract

Objective: To investigate the metastatic mechanism of muscle invasive bladder cancer (MIBC), which accounts for approximately 30% of all bladder cancer cases, and is a considerable medical problem with high metastatic and mortality rates.

Methods: The mRNA levels of patients with metastatic MIBC and nonmetastatic MIBC from The Cancer Genome Atlas dataset were compared. An integrated bioinformatics analysis was performed of the differentially expressed genes (DEGs), and analyses of Gene Ontology, Kyoto Encyclopaedia of Genes and Genomes pathway, protein-protein interaction, and survival were performed to investigate differences between metastatic and nonmetastatic MIBC.

Results: Data from 264 patients were included (131 with, and 133 without, metastasis). A total of 385 significantly DEGs were identified, including 209 upregulated genes and 176 downregulated genes. Based on results using the STRING database and the MCODE plugin of Cytoscape software, two clusters were obtained. Moreover, two genes were identified that may be valuable for prognostic analysis: Keratin 38, type I (*KRT38*) and Histone cluster 1, H3f (*HIST1H3F*).

Conclusion: The *KRT38* and *HIST1H3F* genes may be important in metastasis of MIBC.

[3]Department of Urology, the First Affiliated Hospital of the Third Military Medical University, Chongqing, China

Corresponding author:
Jin Yang, Department of Cell Biology, Third Military Medical University, 30 Gaotanyanzheng Street, Shapingba District, Chongqing, China.
Email: jinyangtmmu@sina.com

[1]Department of Cell Biology, Third Military Medical University, Chongqing, China
[2]Department of Urology, Chinese People's Armed Police Force Tibet Corps Hospital, Lhasa, Tibet, China

Introduction

Bladder cancer is the tenth most prevalent type of malignancy globally; an estimated 549 000 people develop the disease, and 200 000 people die each year. The incidence of bladder cancer is higher in men than in women (with an incidence of 9.6/100 000 and a mortality of 3.2/100 000 in men).[1] Although the number of cases diagnosed as non-muscle invasive bladder cancer (NMIBC) is 2–4 times higher than that of muscle invasive bladder cancer (MIBC),[2,3] MIBC is more prone to recurrence or progression, and is the leading cause of death from bladder cancer. The low survival rate in patients with MIBC is attributed to metastases to the local pelvic lymph nodes and other organs at high risk.[2] Once metastases occur, the five-year survival rate of patients with MIBC is approximately 6%, and effective treatments are lacking.[2,4] MIBC metastasis is the primary cause of death,[5] and exploring the mechanism of MIBC metastasis is vital.

Several key genes, such as lymph node metastasis associated transcript 1 (*LNMAT1*),[6] long intergenic non-protein coding RNA 958 (*LINC00958*, also known as *BLACAT2*),[7] Rho GDP dissociation inhibitor beta (*ARHGDIB*, also known as *RhoGDIβ*),[8] and the long non-coding RNA FOXD2 adjacent opposite strand RNA 1 (*FOXD2-AS1*),[9] are reported to be related to MIBC metastasis. Additionally, individual signalling pathways, such as the phosphoinositide 3-kinase–AKT serine/threonine kinase 1– mechanistic target of rapamycin kinase and receptor tyrosine kinase–RAS– extracellular signal-regulated kinase pathways, which contribute to the progression of MIBC, have recently attracted much attention.[2] It is worth noting that overexpression and mutations in fibroblast growth factor receptor 3 (*FGFR3*) and each member of the ERBB family are associated with bladder cancer.[10] Although some progress has been made in investigating MIBC, little therapeutic progress has occurred in the past two decades, and molecular mechanisms underlying the development and metastasis of MIBC remain unclear.

Patients with distant metastasis of bladder cancer are not suitable for surgery; therefore, cancer specimens with organ metastasis are difficult to obtain. Due to the accessibility of The Cancer Genome Atlas (TCGA) research network, determination of the pathogenic and metastatic mechanisms of MIBC and screening for novel biomarkers with prognostic value can be performed using these data. The aim of the present study was to increase knowledge of the metastatic mechanism of MIBC by analysing sequencing and corresponding clinical data from patients with MIBC downloaded from the TCGA database (https://www.cancer.gov/aboutnci/organization/ccg/research/structural-genomics/tcga). Gene-level data from nonmetastatic and metastatic MIBC samples were analysed, and differentially expressed genes (DEGs) were identified. Subsequently, Gene Ontology (GO) term, Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway enrichment, protein-protein interaction (PPI), and survival analyses were performed to identify crucial genes and pathways that are closely related to MIBC.

Materials and methods

Data acquisition and analysis

Bladder tumour transcriptome sequencing data were downloaded from the TCGA database with HTSeq-counts as the workflow type, transcriptome profiling as the data category, and gene expression quantification as the data type. The cohort included bladder cancer specimens and adjacent nontumor specimens. After downloading, the data were merged into a gene profiling

matrix, the Ensembl gene (ENSG) numbers in the matrix file were transformed into gene symbols, and the individual genotypes were annotated corresponding to the Homo_sapiens.GRH.38.95.gtf file (downloaded from the official Ensembl website: https://www.ensembl.org/index.html?redirect = no). Then, mRNA expression data were extracted from the transcriptome sequencing data for subsequent analysis.

Documents that contained corresponding clinical information were also downloaded from the TCGA. Demographic and clinical data, such as patient age, sex, ethnicity, survival time and neoplasm tumour-node-metastasis (TNM) stage, were extracted. Samples with incomplete clinical information were excluded. Bladder cancer metastases were defined as lymph node metastases (N1) and distant metastases (M1), and some patients have both lymph node metastases and distant metastases. The samples were divided into a metastasis group and a non-metastasis group.

The current study was approved by the Ethics Committee of the Third Military Medical University. Informed patient consent was not required as the results shown are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Identification of DEGs

RNA-seq data of nonmetastatic and metastatic MIBC tissues were analysed via R software, version 3.5.2 (originally developed at the University of Auckland, New Zealand: cran.r-project.org) using the edgeR software package, version 3.22.5,[11] to identify DEGs. In the present study, the DEGs were acquired by setting the following thresholds: log2-based fold change | log2FC|>1 and *P* value < 0.01. The R package ggplot2, version 3.1.0 (cran.r-project.org) was used to map the volcano plot of all DEGs, and the R package pheatmap,

version 1.0.12 (cran.r-project.org) was utilized to generate a heat map with the top 50 genes, using the largest logFC value.

GO and KEGG pathway enrichment analyses of DEGs

The DEGs were divided into up- and down-regulated DEGs with the aim of further investigating their characteristics. Next, all DEGs were analysed using GO and KEGG pathway enrichment methods. The Bioconductor R packages clusterProfiler, version 3.10.1 and org.Hs.eg.db, version 3.6.0 (https://www.bioconductor.org/packages/release/) were used to implement this process. A *P* value <0.05 was considered statistically significant.

PPI network analysis

A PPI network was constructed using the STRING online database,[12] to explore the interactions among the encoded proteins of the DEGs. Disconnected nodes in the network were disregarded, and the required interaction score was set at a high confidence level (value = 0.7). After downloading information on the interactions among the DEGs, the data were imported into Cytoscape software (https://cytoscape.org/) and the network topology was analysed using the Cytoscape MCODE plugin to determine the hub module and candidate hub genes of the PPI network.

Identification of prognostic genes

Associations between gene expression and survival were analysed using the R package survival, version 2.43-3 (https://cran.r-project.org/src/contrib/Archive/survival/). Each individual hub gene in both cluster 1 and cluster 2 was subjected to survival analysis. Then, the potential prognostic genes with a *P* value < 0.05 were screened. Associations between survival and key genes were

verified using the UALCAN interactive web portal.[13]

Results

Sample grouping and processing

The cohort obtained from downloading bladder tumour TCGA transcriptome sequencing data included a total of 433 tumour samples from 408 patients (414 bladder cancer specimens and 19 adjacent nontumour specimens). TCGA documents that contained corresponding clinical information were also downloaded for a total of 412 cases. Following exclusion of samples with incomplete clinical information, patients were divided into those with or without metastases. All samples in the non-metastatic group had expression data, while two samples in the metastatic group did not. Finally, a total of 264 bladder tumour mRNA samples were included for analysis: 133 in the nonmetastatic group and 131 in the metastatic group (Table 1).

Identification of DEGs

To thoroughly investigate differences between the metastatic and nonmetastatic groups, 385 significantly DEGs (209 upregulated and 176 downregulated DEGs) were identified (Figure 1). To more intuitively present the global expression changes in the DEGs, the top 50 genes with the largest fold change in expression were selected in both groups and a heatmap was created (Figure 2).

GO and KEGG enrichment analyses of DEGs

Up- and downregulated DEGs were analysed using GO (Figure 3) and KEGG (Figure 4) pathway enrichment methods to determine the characteristics of the identified DEGs. The cellular components (CCs)

Table 1. Clinical features of 264 patients with muscle invasive bladder cancer, divided into those with or without metastasis.

| | Cancer status | |
| | Metastasis | Non-metastasis |
Clinical feature	($n = 131$)	($n = 133$)
Survival time, days	698	889
Survival status		
Alive	49	93
Dead	82	40
Age, years		
<60	19	36
≥60	112	97
Sex		
Female	30	27
Male	98	106
Subgroup		
Non-papillary	99	68
Papillary	30	63
T stage		
T1	3	3
T2	39	67
T3	60	54
T4	28	7
N stage		
N0	0	133
N1+	129	0
M stage		
M0	44	133
M1	10	0
Clinical Stage		
I	0	1
II	0	64
III	0	65
IV	130	2
Grade		
High	129	113
Low	0	19

Data presented as *n* days or *n* patient prevalence.
T, tumour; N, node; M, metastasis.

of the 209 upregulated DEGs were principally enriched in the DNA packaging complex, nucleosome, nuclear nucleosome, protein-DNA complex, intermediate filament, and intermediate filament cytoskeleton, and their molecular functions (MFs) mainly included hormone activity, pattern
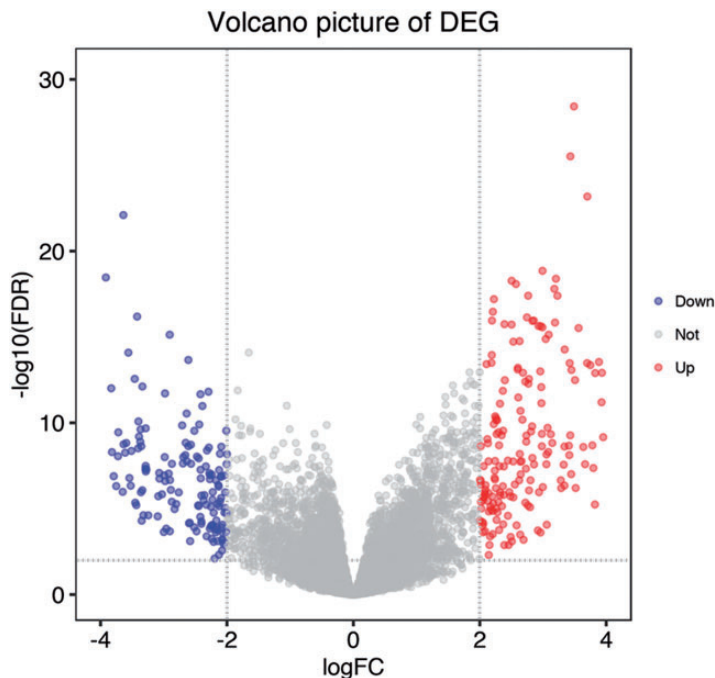
Figure 1. Volcano plot of the 385 DEGs. The transverse axis represents the logFC, while the vertical axis represents the logFDR. Red, upregulated genes with a logFC ≥ 2; blue, downregulated genes with a logFC ≤ -2; DEG, differentially expressed gene; FDR, false discovery rate; FC, fold change.

binding and polysaccharide binding. These genes mainly participate in the biological processes (BPs) of epidermal cell differentiation, epidermis development, keratinization, keratinocyte differentiation, and nucleosome assembly (Figure 3a). KEGG analysis demonstrated that these upregulated DEGs are principally related to individual pathways such as systemic lupus erythematosus, alcoholism, the oestrogen signalling pathway, salivary secretion, and viral carcinogenesis (Figure 4a). The 176 downregulated DEGs were mainly involved in ion channel complexes, transmembrane transporter complexes, transporter complexes, cation channel complexes, the neuronal cell body and other CCs. The corresponding MFs included passive transmembrane transporter activity, channel activity, substrate-specific channel activity and ion-gated channel activity. These

genes mainly participate in the BPs that regulate membrane potential (Figure 3b). In terms of KEGG analysis, the downregulated DEGs are mainly involved in pathways such as the neuroactive ligand-receptor interaction, pancreatic secretion, and nicotine addiction (Figure 4B).

PPI network analysis

A PPI network with 175 nodes and 357 edges was constructed to further investigate interactions among the proteins encoded by the DEGs (Figure 5a). A total of 175 nodes were further analysed by employing the MCODE plugin to detect hub modules. The top two significant modules were determined, and these two clusters contained multiple histone family and keratin family genes (Figure 5b). The top 30 hub genes
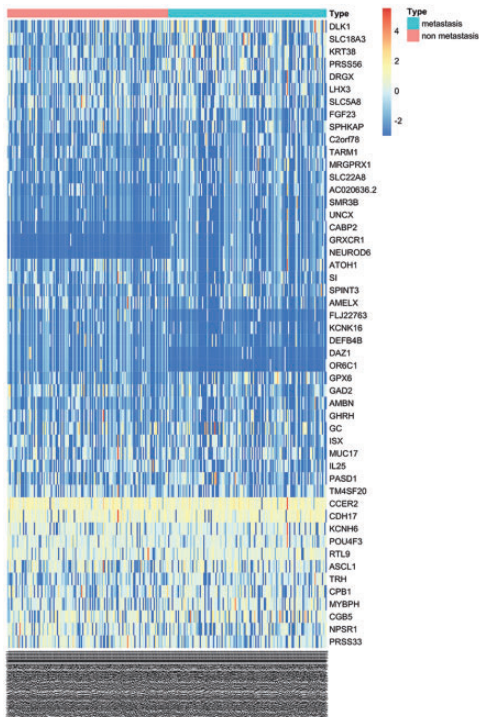
Figure 2. Heatmap of the top 50 differentially expressed genes with the largest logFC value. Samples were subdivided and ordered into metastasis or non-metastasis groups. The colour of each dot represents the expression of the gene in the sample. The deeper the red colour, the higher the expression; the deeper the blue colour, the lower the expression. FC, fold change.

that may play vital roles in the metastasis of MIBC are displayed (Figure 5c).

Survival analysis of key genes

Survival analysis was conducted with 21 candidate genes identified in the top two significant modules, to identify genes that may affect survival rate in patients with bladder cancer. Two candidate genes, Keratin 38, type I (*KRT38*) and Histone cluster 1, H3f (*HIST1H3F*), were revealed to exert significant effects on the overall survival of patients with bladder cancer (Figure 6). Patients whose tissues had high

HIST1H3F expression levels had significantly longer overall survival than patients whose tissues had low *HIST1H3F* expression levels ($P < 0.05$). Moreover, patients with low *KRT38* expression levels had a good prognosis compared with those with high expression ($P < 0.05$).

Discussion

The present study demonstrated that *KRT38* and *HIST1H3F* may be two novel prognostic biomarkers of MIBC and may exert an important function in the process of MIBC metastasis. In addition, the present research provides new information on the mechanism of bladder cancer metastasis.

Survival analysis revealed that patients with low expression of the *KRT38* gene have higher survival rates. To the best of the authors' knowledge, the relationship between *KRT38* and bladder cancer prognosis has not been reported previously. Few reports have described the relationship between this gene and diseases in *Homo sapiens*, most likely due to the application of sequencing technology, which allows the identification of valuable molecules that are not detectable by microarray technology.

As a member of the largest subgroup of intermediate filaments, keratin is closely related to the cellular cytoskeleton,[14] and keratin proteins are recognized to indicate the differentiation status of tumour cells and act as markers of prognosis in patients with cancer.[15,16] Previous studies have revealed that keratin acts as a prognostic marker in colorectal cancer, oral squamous cell carcinoma (SCC), breast cancer, and bladder cancer.[14,17–19]

Recently, an unsupervised hierarchical clustering analysis was performed based on MIBC gene expression profiling data to define expression pattern subtypes, and a subtype called 'basal' was verified by researchers.[17,20–23] The basal subtype is
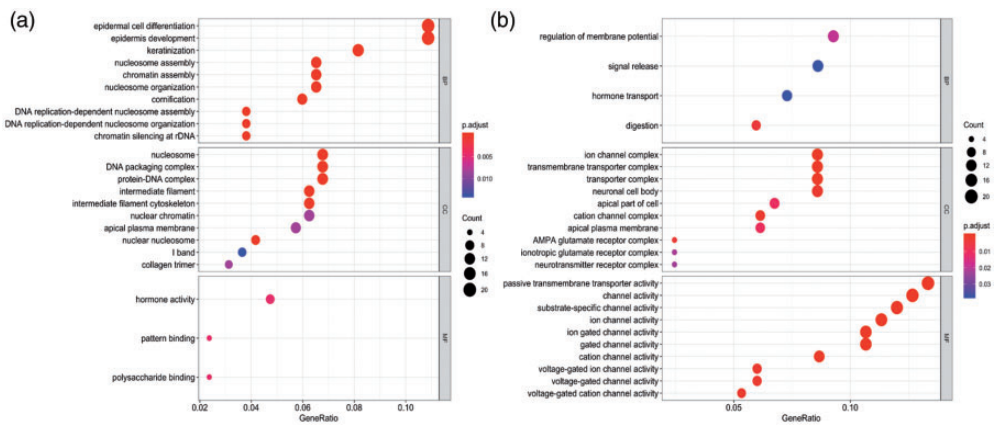
Figure 3. Dot plots showing Gene Ontology analysis results of: (a) upregulated DEGs with a logFC ≥ 2; and (b) downregulated DEGs with a logFC ≥ 2. The colour of each dot represents the FDR of each term involved in the analysis. The size of each dot represents the gene counts of this term involved in the analysis. DEGs, differentially expressed genes; FDR, false discovery rate; BP, biological process; CC, cellular component; MF, molecular function.
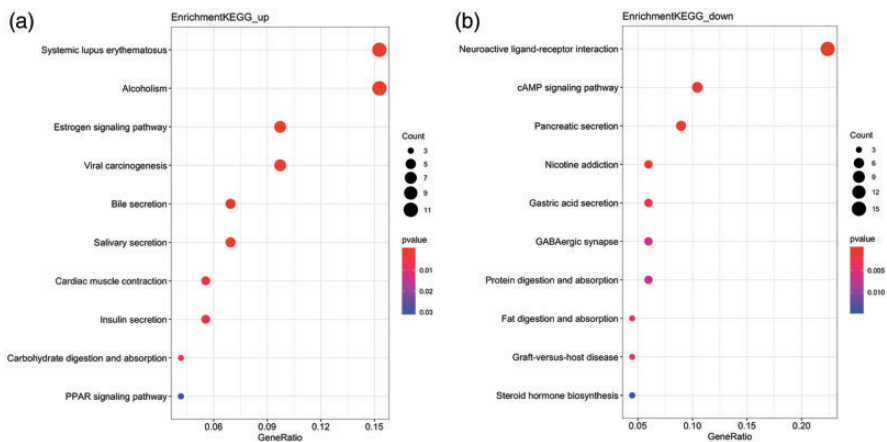

Figure 4. Dot plots showing Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis results of: (a) upregulated differentially expressed genes (DEGs); and (b) downregulated DEGs. The colour of each dot represents the *P* value of each term involved in the analysis. The size of each dot represents the gene counts of this term involved in the analysis.

characterized by the level of keratins, including keratin 5 (*KRT5*), keratin 14 (*KRT14*), keratin 6A (*KRT6A* or *KRT6C*) and keratin 6B (*KRT6B*). Similar results were obtained in a study by Kim et al;[2] in that the basal subtype accounted for 29% of invasive bladder cancers, and the levels of *KRT14*, *KRT5*, *KRT6A/B*, and keratin

16 (*KRT16*) were increased in these bladder cancer tissues. The studies mentioned above suggest that abnormally expressed keratins are important in the development of bladder cancer and that *KRT38* may be a promising prognostic biomarker.

In the present study, almost half of the hub genes (11/30) obtained from the PPI
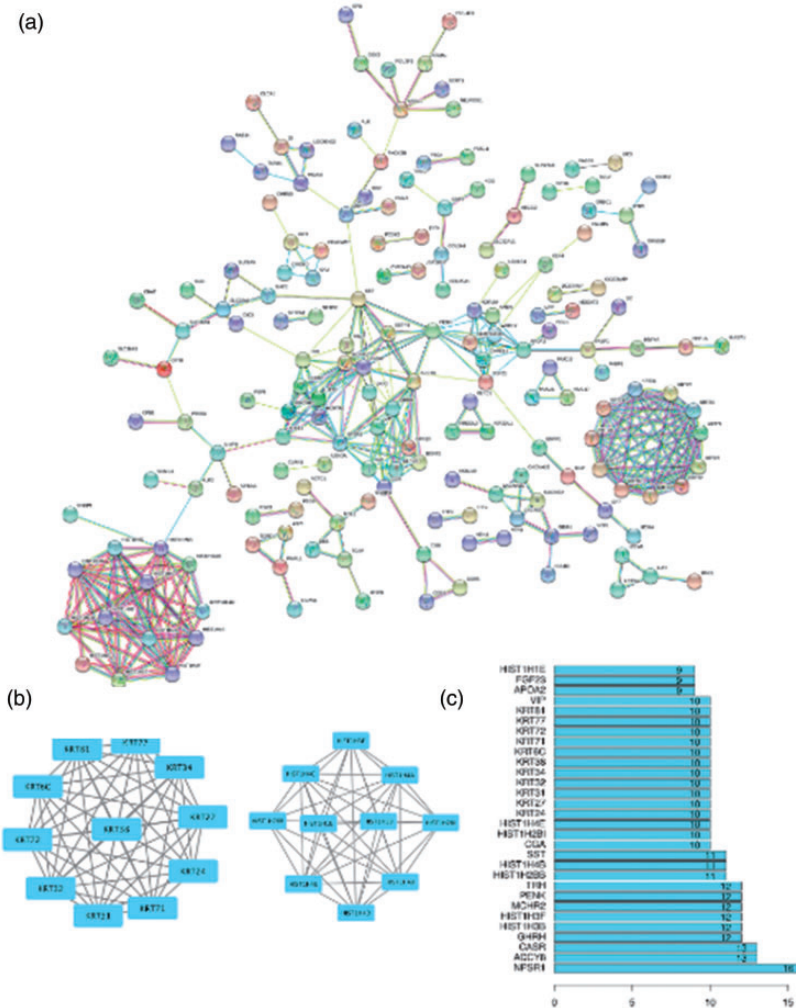
Figure 5. Protein-protein interaction (PPI) network analysis results, showing: (a) PPI network of differentially expressed genes; (b) module analysis of the PPI network; and (c) top thirty hub genes identified from the PPI network.

network (Figure 5c showing the top thirty genes) are members of the keratin family. Previous studies have shown that cancer cells can migrate and invade the body after epithelial-mesenchymal transition. Realignment occurs in the cytoskeleton and epithelial biomarkers, and e-cadherin and keratin are disrupted in this process.[24–26] Moreover, Joosse et al.[27] reported that changes in keratin levels occurred during the lymph node metastasis of primary breast cancer. The results of the present study are consistent with those of previous studies and suggest that changes in keratin expression may be associated with bladder tumour metastasis. However, further exploration and experimental validation are needed.

The HIST1H3F protein, a replication-dependent histone, has been suggested as
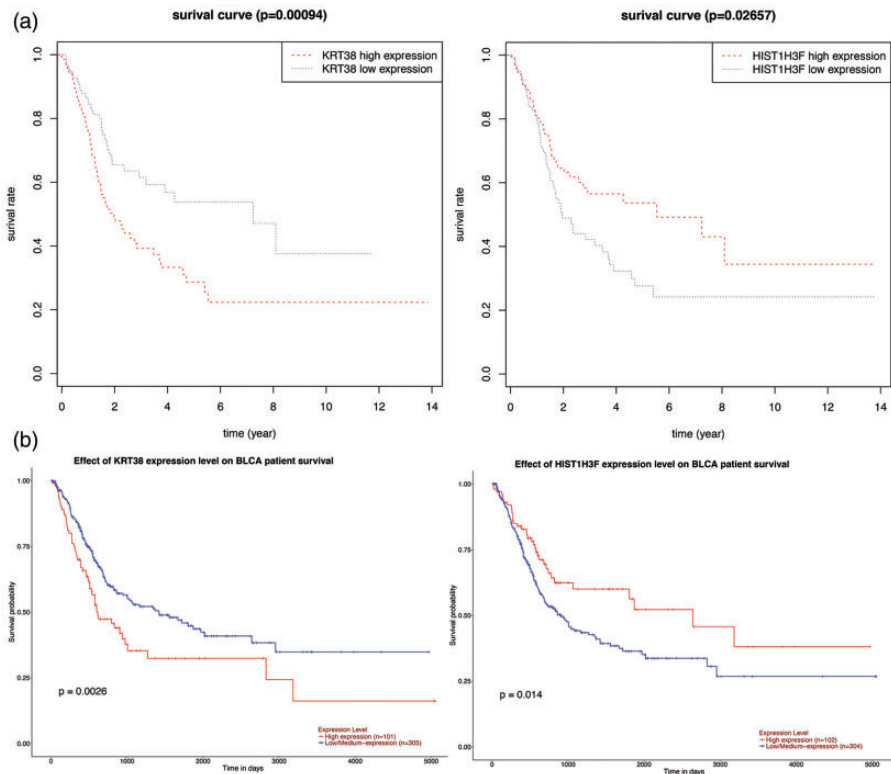
Figure 6. Survival curves showing: (a) relationship between survival and key genes in the present study population; and (b) validation of relationship between survival and key genes using UALCAN website tools. Patients whose tissues had high Histone cluster 1, H3f (*HIST1H3F*) expression levels showed significantly longer overall survival than patients whose tissues had low *HIST1H3F* expression levels. Moreover, patients with low Keratin 38, type I (*KRT38*) levels had a better prognosis than those with high expression levels ($P < 0.05$).

a potential biomarker for various cancers, and in SCCs of the larynx, the combination of HIST1H3F with other molecules is used to predict patient prognosis.[28] However, the present report is the first to indicate that HIST1H3F may be used as a prognostic marker in patients with MIBC. Histones are indispensable structural proteins related to DNA in eukaryotic cells and are divided into five major families, H1/H5, H2A, H2B, H3, and H4, of which, H1/H5 are considered linker histones, while the others are considered core histones.[29] Among the hub genes involved in the present study, except for *HIST1H1E*, which encodes a

linker histone, the remaining genes (*HIST1H4A-E*, *HIST1H2BB*, *HIST1H2 BI*, *HIST1H3F*, and *HIST1H3B*) encode core histones.[29] Core histones have two major functions, compacting DNA strands and chromatin regulation, which are mainly related to histone posttranslational modifications. Complex mechanisms, such as covalent histone modifications and DNA methylation, are used to dynamically regulate the chromatin structure.[30] Similarly, in the present study, the main GO terms of cluster 1 were associated with protein-DNA complex assembly, the nucleosome and chromatin assembly. Considering the

functions of histones in previous studies, it is concluded that the aberrant level of histones may reflect the decreased control of cell cycle progression, which is a typical mechanism of rapid tumour proliferation.

The present study is based on the TCGA sequencing database, and a total of 264 samples (metastasis: non-metastasis, 131: 133) were selected. Because the quantity of samples was balanced, DEGs could be precisely determined. Moreover, routine methods of high-throughput data analyses, including GO and KEGG enrichment analyses, were also used.

The KEGG analysis revealed that upregulated DEGs in MIBC are mainly related to pathways such as systemic lupus erythematosus, alcoholism, the oestrogen signalling pathway, salivary secretion, and viral carcinogenesis. Interestingly, one factor that increases the morbidity and mortality of infectious diseases and cancers is alcohol consumption.[31] Natural killer (NK) cells have been revealed to exert a significant inhibitory effect on tumour metastasis to draining lymph nodes, while alcohol consumption damages NK cells or inhibits NK cell production. The decreased number of NK cells in lymph nodes has been related to increased melanoma metastasis to draining lymph nodes.[32] Based on the present findings, it may be concluded that similar mechanisms also play a significant role in the metastasis of bladder cancer. Additionally, a recent study showed that the loss of oestrogen receptor (ER)α was positively associated with the grade and invasiveness of tumours.[33] Furthermore, the ERβ level was upregulated in high-grade/invasive tumours and related to a poor prognosis. The results of the current study support the results of previous studies from a different perspective.

The downregulated DEGs in MIBC samples were found to be mainly involved in pathways such as the neuroactive ligand-receptor interaction, pancreatic secretion, nicotine addiction, and cAMP pathway. Fang et al.[34] demonstrated that neuroactive ligand-receptor interactions were related to bladder cancer progression at the Ta-T1/ T1-T2 stages. Cigarette smoking is another known risk factor for bladder cancer,[35] and DEGs were specifically enriched in nicotine addiction pathways. In addition, a previous study showed that activation of the cAMP pathway could inhibit bladder cancer cell invasion by targeting microtubule associated protein 4-dependent microtubule dynamics.[36] The reasons why DEGs in the present study were enriched in several unusual pathways, such as systemic lupus erythematosus and pancreatic secretion pathways, remain unknown; however, one possible explanation is that the enrichment methods that were used are based on over-representation analysis (ORA) methods. Although ORA methods provide robust and reliable results, they have some limitations in revealing the molecular mechanism involved.[37] Another possible explanation is the present authors limited background knowledge of bladder cancer.

In summary, using bioinformatics analysis, crucial genes and pathways were identified that are closely correlated with the prognosis and metastasis of MIBC. These results may provide novel and promising prognostic biomarkers for patients with MIBC, and more detailed information for exploring the molecular mechanism of bladder tumour metastasis. Nevertheless, further molecular biology experiments are needed to verify these findings.

ORCID iD

Jin Yang ⓘ https://orcid.org/0000-0001-6014-3482

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68: 394–424.

2. Kim J, Akbani R, Creighton CJ, et al. Invasive bladder cancer: genomic insights and therapeutic promise. *Clin Cancer Res* 2015; 21: 4514–4524.

3. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016; 66: 7–30.

4. Edge SB and Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* 2010; 17: 1471–1474.

5. Alonso DF, Ripoll GV, Garona J, et al. Metastasis: recent discoveries and novel perioperative treatment strategies with particular interest in the hemostatic compound desmopressin. *Curr Pharm Biotechnol* 2011; 12: 1974–1980.

6. Chen C, He W, Huang J, et al. LNMAT1 promotes lymphatic metastasis of bladder cancer via CCL2 dependent macrophage recruitment. *Nat Commun* 2018; 9: 3826.

7. He W, Zhong G, Jiang N, et al. Long non-coding RNA BLACAT2 promotes bladder cancer–associated lymphangiogenesis and lymphatic metastasis. *J Clin Invest* 2018; 128: 861–875.

8. Huang H, Jin H, Zhao H, et al. RhoGDIβ promotes Sp1/MMP-2 expression and bladder cancer invasion through perturbing miR-200c-targeted JNK2 protein translation. *Mol Oncol* 2017; 11: 1579–1594.

9. Su F, He W, Chen C, et al. The long non-coding RNA FOXD2-AS1 promotes bladder cancer progression and recurrence through a positive feedback loop with Akt and E2F1. *Cell Death Dis* 2018; 9: 233.

10. Mazzola CR, Siddiqui KM, Billia M, et al. Dovitinib: rationale, preclinical and early clinical data in urothelial carcinoma of the bladder. *Expert Opin Investig Drugs* 2014; 23: 1553–1562.

11. McCarthy DJ, Chen Y and Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012; 40: 4288–4297.

12. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; 47: D607–D613.

13. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 2017; 19: 649–658.

14. Majumdar D, Tiernan JP, Lobo AJ, et al. Keratins in colorectal epithelial function and disease. *Int J Exp Pathol* 2012; 93: 305–318.

15. Barak V, Goike H, Panaretakis KW, et al. Clinical utility of cytokeratins as tumor markers. *Clin Biochem* 2004; 37: 529–540.

16. Ho PL, Kurtova A and Chan KS. Normal and neoplastic urothelial stem cells: getting to the root of the problem. *Nat Rev Urol* 2012; 9: 583–594.

17. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 2014; 507: 315–322.

18. Coelho BA, Peterle GT, Santos M, et al. Keratins 17 and 19 expression as prognostic markers in oral squamous cell carcinoma. *Genet Mol Res* 2015; 14: 15123–15132.

19. Merkin RD, Vanner EA, Romeiser JL, et al. Keratin 17 is overexpressed and predicts poor survival in estrogen receptor–negative/human epidermal growth factor receptor-2–negative breast cancer. *Hum Pathol* 2017; 62: 23–32.

20. Choi W, Porten S, Kim S, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell* 2014; 25: 152–165.

21. Damrauer JS, Hoadley KA, Chism DD, et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. *Proc Natl Acad Sci U S A* 2014; 111: 3110–3115.

22. Sjödahl G, Lauss M, Lövgren K, et al. A molecular taxonomy for urothelial carcinoma. *Clin Cancer Res* 2012; 18: 3377–3386.

23. Volkmer JP, Sahoo D, Chin RK, et al. Three differentiation states risk-stratify bladder cancer into distinct subtypes. *Proc Natl Acad Sci U S A* 2012; 109: 2078–2083.

24. Radisky DC. Epithelial-mesenchymal transition. *J Cell Sci* 2005; 118: 4325–4326.

25. Thiery JP and Sleeman JP. Complex networks orchestrate epithelial–mesenchymal transitions. *Nat Rev Mol Cell Biol* 2006; 7: 131–142.

26. Willipinski-Stapelfeldt B, Riethdorf S, Assmann V, et al. Changes in cytoskeletal protein composition indicative of an epithelial-mesenchymal transition in human micrometastatic and primary breast carcinoma cells. *Clin Cancer Res* 2005; 11: 8006–8014.

27. Joosse SA, Hannemann J, Spötter J, et al. Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells. *Clin Cancer Res* 2012; 18: 993–1003.

28. Mirisola V, Mora R, Esposito AI, et al. A prognostic multigene classifier for squamous cell carcinomas of the larynx. *Cancer Lett* 2011; 307: 37–46.

29. Bhasin M, Reinherz EL and Reche PA. Recognition and classification of histones using support vector machine. *J Comput Biol* 2006; 13: 102–112.

30. Wang GG, Allis CD and Chi P. Chromatin remodeling and cancer, part I: covalent histone modifications. *Trends Mol Med* 2007; 13: 363–372.

31. Ben-Eliyahu S, Page GG, Yirmiya R, et al. Acute alcohol intoxication suppresses natural killer cell activity and promotes tumor metastasis. *Nat Med* 1996; 2: 457–460.

32. Zhang H, Zhu Z and Meadows GG. Chronic alcohol consumption decreases the percentage and number of NK cells in the peripheral lymph nodes and exacerbates B16BL6 melanoma metastasis into the draining lymph nodes. *Cell Immunol* 2011; 266: 172–179.

33. Miyamoto H, Yao JL, Chaux A, et al. Expression of androgen and oestrogen receptors and its prognostic significance in urothelial neoplasm of the urinary bladder. *BJU Int* 2012; 109: 1716–1726.

34. Fang ZQ, Zang WD, Chen R, et al. Gene expression profile and enrichment pathways in different stages of bladder cancer. *Genet Mol Res* 2013; 12: 1479–1489.

35. Marsit CJ, Karagas MR, Danaee H, et al. Carcinogen exposure and gene promoter hypermethylation in bladder cancer. *Carcinogenesis* 2006; 27: 112–116.

36. Ou Y, Zheng X, Gao Y, et al. Activation of cyclic AMP/PKA pathway inhibits bladder cancer cell invasion by targeting MAP4-dependent microtubule dynamics. *Urol Oncol* 2014; 32: 47.e21–47.e28.

37. Wang X, Yin T, Li B, et al. Progress in gene functional enrichment analysis. *Scientia Sinica Vitae* 2016; 46: 363–373 [In Chinese, English abstract].