Identification and Quantification of Iron Metabolism Landscape on Therapy and Prognosis in Bladder Cancer

Xiaodong Song†, Sheng Xin†, Yucong Zhang2, Jiaquan Mao1, Chen Duan1, Kai Cui1, Liang Chen1, Fan Li1, Zheng Liu1, Tao Wang1, Jihong Liu1, Xiaming Liu1* and Wen Song1*

1Department of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, 2Department of Geriatric, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

The morbidity of bladder cancer (BLCA) is high and has gradually elevated in recent years. BLCA is also characterized by high recurrence and high invasiveness. Due to the drug resistance and lack of effective prognostic indicators, the prognosis of patients with BLCA is greatly affected. Iron metabolism is considered to be a pivot of tumor occurrence, progression, and tumor microenvironment (TME) in tumors, but there is little research in BLCA. Herein, we used univariate COX regression analysis to screen 95 prognosis-related iron metabolism-related genes (IMRGs) according to transcription RNA sequencing and prognosis information of the Cancer Genome Atlas (TCGA) database. TCGA-BLCA cohort was clustered into four distinct iron metabolism patterns (C1, C2, C3, and C4) by the non-negative matrix factorization (NMF) algorithm. Survival analysis showed that C1 and C3 patterns had a better prognosis. Gene set variant analysis (GSVA) revealed that C2 and C4 patterns were mostly enriched in carcinogenic and immune activation pathways. ESTIMATE and single sample gene set enrichment analysis (ssGSEA) also confirmed the level of immune cell infiltration in C2 and C4 patterns was significantly elevated. Moreover, the immune checkpoint genes in C2 and C4 patterns were observably overexpressed. Studies on somatic mutations showed that the tumor mutation burden (TMB) of C1 and C4 patterns was the lowest. Chemotherapy response assessment revealed that C2 pattern was the most sensitive to chemotherapy, while C3 pattern was the most insensitive. Then we established the IMRG prognosis signature (IMRGscore) by the least absolute shrinkage and selection operator (LASSO), including 13 IMRGs (TCIRG1, CTSE, ATP6V0A1, CYP2C8, RNF19A, CYP4Z1, YPEL5, PLOD1, BMP6, CAST, SCD, IFNG, and ASIC3). We confirmed IMRGscore could be utilized as an independent prognostic indicator. Therefore, validation and quantification of iron metabolism landscapes will help us comprehend the formation of the BLCA immunosuppressive microenvironment, guide the selection of chemotherapeutic drugs and immunotherapy, and predict the prognosis of patients.

Keywords: bladder cancer, iron metabolism, tumor microenvironment, prognostic signature, nomogram, bioinformactics
INTRODUCTION

Bladder cancer (BLCA) is one of the most familiar malignant tumors in the urinary system, with about 81400 new cases and 17900 deaths in the United States in 2020 (Siegel et al., 2020). Approximately 75% of BLCA was found to be non-muscle invasive bladder cancer (NMIBC), which was characterized by a high recurrence rate (45% 5-year recurrence rate) (Berdik, 2017; Babjuk et al., 2019). Transurethral resection of bladder tumor (TURBT), chemotherapy, BCG vaccine, radiotherapy, and radical cystectomy are the main treatments for BLCA patients (Berdik, 2017). Chemotherapy and immunotherapy are also important strategies for conservative treatment of BLCA (Yin et al., 2016; Rouanne et al., 2018). However, some patients are not sensitive to these drug therapies. Due to the high recurrence rate, high metastatic risk, and patient’s dissatisfaction with the treatment effect, it is of great significance to identify and quantify some molecular landscapes that have impacts on the choice of drug treatment, and explore a novel indicator that predicts the prognosis of BLCA patients.

Iron is a vital trace element for cell proliferation and growth in the human body (Torti and Torti, 2013). In cancer, the absorbability, effusion, storage, and regulation of iron are entirely disturbed, which indicates that the reprogramming of iron metabolism would induce the dysregulation in tumor cells division and survival (Andrews, 2008; Manz et al., 2016; Wang et al., 2018; Jung et al., 2019). Iron plays a dual role in cancer (Thévenod, 2018). Epidemiological investigations revealed that excess iron was a hazard factor of carcinogenesis (Stevens et al., 1994; Wu et al., 2004; Fonseca-Nunes et al., 2014). The accumulation of iron supports tumor worsening in proliferation, metabolism, and metastasis (Torti et al., 2018). Cancer cells exhibit a phenotype search for iron through disordering regulation of iron-binding proteins (Dufès et al., 2013; Bialasek et al., 2019). On the other hand, iron dependence of cancer cells affects many cell death modes, including ferroptosis, a form of iron-dependent cell death (Mou et al., 2019; Battaglia et al., 2020). Inducing ferroptosis of cancer cells has become a new hotspot in the research and development of cancer treatment (Hassannia et al., 2019; Xu et al., 2019).

There has been little research on iron metabolism in BLCA. The study was conducted to confirm whether iron metabolism had an effect on the molecular microenvironment of BLCA, as well as its ability to predict the clinical prognosis. We first clustered the TCGA-BLCA cohort into different iron metabolic patterns on the basis of the expression of iron metabolism-related genes (IMRGs). Then the survival prognosis, GSVA analysis, tumor immune microenvironment (TIME), somatic mutations chemotherapy, and immunotherapy response among different patterns were analyzed. Eventually, we established a prognostic signature associated with iron metabolism and confirmed that it is an effective independent predictor in BLCA patients.

Table 1: Characteristics of patients included in the study.

| Variable                  | TCGA-BLCA cohort (n = 400) | GSE13507 cohort (n = 165) |
|---------------------------|----------------------------|---------------------------|
| Age                       |                            |                           |
| ≤70                       | 228 (57.00)                | 109 (66.06)               |
| >70                       | 172 (43.00)                | 56 (33.94)                |
| Gender                    |                            |                           |
| MALE                      | 296 (74.00)                | 135 (81.82)               |
| FEMALE                    | 104 (26.00)                | 30 (18.18)                |
| T stage                   |                            |                           |
| TX                        | 1 (0.25)                   | 0                         |
| T0                        | 1 (0.25)                   | 0                         |
| Ta                        | 0                         | 24 (14.55)                |
| T1                        | 3 (0.75)                   | 80 (48.48)                |
| T2                        | 117 (29.25)                | 31 (18.79)                |
| T3                        | 190 (47.5)                 | 19 (11.52)                |
| T4                        | 57 (14.25)                 | 11 (6.67)                 |
| N stage                   |                            |                           |
| NK                        | 36 (9.00)                  | 1 (0.61)                  |
| N0                        | 233 (58.25)                | 149 (90.30)               |
| N1                        | 44 (11.00)                 | 8 (4.85)                  |
| N2                        | 75 (18.75)                 | 6 (3.64)                  |
| N3                        | 7 (1.75)                   | 1 (0.61)                  |
| Unknown                   | 5 (1.25)                   | 0                         |
| Pathologic stage          |                            |                           |
| Stage 0                   | 0                         | 23 (13.94)                |
| Stage I                   | 2 (0.50)                   | 80 (48.48)                |
| Stage II                  | 127 (31.75)                | 26 (15.76)                |
| Stage III                 | 138 (34.50)                | 29 (17.58)                |
| Stage IV                  | 131 (32.75)                | 7 (4.24)                  |
| Unknown                   | 2 (0.50)                   | 0                         |
| Histologic grade          |                            |                           |
| Low grade                 | 20 (5.00)                  | 105 (63.64)               |
| High grade                | 377 (94.25)                | 60 (36.36)                |
| unknown                   | 3 (0.75)                   | 0                         |

MATERIALS AND METHODS

Retrieval of Iron Metabolism-Related Genes

A set of IMRGs was sorted from multiple gene sets from Molecular Signatures Database (MsigDB) (http://www.gsea-msigdb.org/gsea/msigdb/index.jsp), including GOMF_IRON_ION_BINDING, GOBP_IRON_ION_TRANSPORT, GOBP_RESPONSE_TO_IRON_ION, GOBP_IRON_ION_METABOLISM, GOBP_IRON_IMPORT INTO_CELL, GOBP_IRON_ION_IMPORT ACROSS_PLASMA_MEMBRANE, GOMF_2IRON_2SULFUR_CLUSTER_BINDING, GOFP_4IRON_4SULFUR_CLUSTER_BINDING, GOBP_IRON_COORDINATION_ENTITY_TRANSPORT, GOBP_CELLULAR_IRON_METABOLISM, GOBPHEME_METABOLIC_PROCESS, HEME_BIOSYNTHETIC_PROCESS, MODULE_540, HALLMARK_HEME_METABOLISM and REACTOMEIRON_UPTAKE_ABD_TRANSPORT. After removing the duplicate genes from all gene sets, a total of 515 IMRGs were retrieved.
TABLE 2 | Prognosis-related IMRGs selected by univariate COX regression analysis.

| Gene     | HR   | z     | p-value |
|----------|------|-------|---------|
| NFE2     | 1.25892 | -3.0568 | 0.00224 |
| ALKBH2   | 0.71241 | -2.4545 | 0.01411 |
| ALKBH3   | 1.3569 | 2.1987 | 0.0129 |
| ALOX5    | 0.85739 | -2.5977 | 0.02367 |
| ASCC3    | 0.68739 | -2.2624 | 0.00239 |
| ATP5F1A2 | 0.71955 | -2.1818 | 0.02913 |
| ATP6V0A1 | 1.81787 | 3.6315 | 0.00028 |
| ATP6V0D1 | 1.45986 | 2.7342 | 0.00625 |
| ATP6V1A1 | 1.37125 | 2.0984 | 0.03587 |
| ATP6V1C2 | 1.28480 | 2.2647 | 0.02253 |
| BMP6     | 1.30778 | 2.9345 | 0.00350 |
| CAST     | 1.33555 | 2.8034 | 0.00334 |
| CD201    | 1.25977 | 2.3200 | 0.02034 |
| CDAO3    | 0.64021 | -2.3562 | 0.01851 |
| CIR1     | 0.52578 | -3.1615 | 0.0157 |
| CISD1    | 1.28174 | 2.0195 | 0.04343 |
| CLTC     | 1.52718 | 3.0834 | 0.02025 |
| CROCCP2  | 0.66104 | -2.9905 | 0.00006 |
| CTSE     | 0.84792 | -3.9943 | 0.03831 |
| CUL1     | 1.46334 | 2.1223 | 0.00329 |
| CYBRD1   | 1.13803 | 2.1286 | 0.01697 |
| CYP19A1  | 1.82857 | 3.2558 | 0.00089 |
| CYP1B1   | 1.12555 | 2.3874 | 0.01018 |
| CYP26B1  | 1.22486 | 2.8893 | 0.00991 |
| CYP27B1  | 0.76528 | -2.0634 | 0.02173 |
| CYP2C8   | 0.48055 | -3.2681 | 0.00243 |
| CYP2D6   | 0.53783 | -2.5791 | 0.00216 |
| CYP2D7   | 0.39076 | -3.3122 | 0.00093 |
| CYP2F1   | 1.47375 | 2.8389 | 0.02169 |
| CYP2R1   | 0.68076 | -2.2516 | 0.01018 |
| CYP2W1   | 1.14759 | 2.3205 | 0.03832 |
| CYP3A5   | 0.88535 | -2.1530 | 0.01534 |
| CYP4A22  | 0.01075 | -2.3088 | 0.00022 |
| CYP4F12  | 0.85576 | -3.1193 | 0.00215 |
| CYP4F8   | 0.89174 | -2.6547 | 0.00208 |
| CYP4Z1   | 0.51083 | -3.0875 | 0.04189 |
| CYP4Z2P  | 0.71692 | -3.1295 | 0.07688 |
| CYPS1A1  | 1.54962 | 3.0743 | 0.00049 |
| CYPTB1   | 1.25543 | 3.0436 | 0.02001 |
| DNM2     | 0.72877 | -2.6658 | 0.01018 |
| ENDO1    | 1.36254 | 3.4875 | 0.02169 |
| EPO     | 0.8764 | -2.3261 | 0.00034 |
| FA2H     | 0.86757 | -2.3261 | 0.01688 |
| FECH     | 1.29593 | 2.1268 | 0.01481 |
| FTO      | 1.44207 | 2.1268 | 0.00110 |
| G6PD     | 1.18156 | 2.4243 | 0.02667 |
| GCLM     | 1.19098 | 2.5395 | 0.04924 |
| HJV      | 0.00736 | -2.2164 | 0.01388 |
| HRG      | 0.28674 | -1.9665 | 0.00304 |
| IFLG     | 0.72938 | -2.4803 | 0.01371 |
| ISCU     | 0.69714 | -2.3199 | 0.00034 |
| LAMP2    | 1.22151 | 2.0780 | 0.01699 |
| MBAT2    | 1.31363 | 3.5807 | 0.01934 |
| MKRN1    | 0.66446 | -2.3893 | 0.02424 |
| MYC      | 1.14183 | 2.4370 | 0.02823 |
| NARF     | 0.73477 | -2.3388 | 0.00972 |
| NDUFV2   | 0.71090 | -2.3388 | 0.01416 |
| NFE2     | 1.25892 | 2.0291 | 0.00768 |
| NUBPL    | 1.71372 | 2.1941 | 0.02424 |
| OGFOD1   | 1.56046 | 2.0567 | 0.02823 |
| P3H1     | 1.27676 | 2.5856 | 0.01699 |
| P3H3     | 1.16777 | 2.4552 | 0.00972 |

**HR**, hazard ratio.

Acquisition and Process of Original Data

Transcript RNA sequencing, clinical data, and somatic mutation data for patients with BLCA were obtained from the Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). The cohort included 411 BLCA tissues and 19 normal bladder tissues. The TCGA-BLCA level 3 RNA-seq data was downloaded as fragments per kilobase of transcript per million mapped reads (FPKM), and when multiple Ensembl IDs were mapped to a single gene symbol in the RNA sequencing data, gene expression is annotated in an average expression. The GSE13507 dataset was analyzed as an external validation cohort. The gene expression profile and clinical information for the microarray dataset came from the Gene Expression Synthesis (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). All sequencing data were processed with log two transformation, background adjustment, normalization, final summarization through the “Affy” package in R. The clinical information of all BLCA patients included in this study is shown in Tables 1, 2.

NMF Clustering for Iron Metabolism Patterns

We matched IMRGs’ RNAseq data and overall survival (OS) information of the TCGA-BLCA dataset. A univariate Cox
regression analysis was carried out to determine the prognosis-related IMRGs in BLCA, with a screening criterion of \( p < 0.05 \). Non-negative matrix factorization (NMF) is applied to determine distinct iron metabolism-related patterns with the help of the “NMF” R package. NMF algorithm decomposes the original matrix into two non-negative matrices to identify the potential features in the gene expression profile (Brunet et al., 2004). Repeat the deposition and aggregate the results to obtain consistent clustering. According to the cophenetic coefficient, contour, and sample size, \( k = 4 \) was determined as the best cluster number.

**GSVA**

Gene set variant analysis (GSVA), is a nonparametric, unsupervised algorithm. GSVA transforms the isolate gene expression matrix to an expression matrix of particular gene sets as features. The algorithm is implemented based on the “GSVA” R package. The difference of expression matrix after transformation was analyzed by the “limma” package to find the difference of enriched functions among different iron metabolic patterns.

**Evaluation of Tumor Immune Microenvironment**

In order to assess the TIME status of BLCA, we used single sample gene set enrichment analysis (ssGSEA), ESTIMATE, and CIBERSORT in R. ssGSEA investigated congenital and adaptive immune cells as well as a variety of immune-related functions. The Normalized Enrichment Score (NES) was to embody the relative amount of each TIME infiltration unit in patients. ESTIMATE predicted the level of infiltrating matrix and immune cells by calculating stromal and immune scores, comprehensively obtained the ESTIMATE score for evaluating the TIME. We also evaluated the relative fraction of 22 tumor-infiltrating immune cells (TIICs), including B cells, T cells, natural killer (NK) cells, macrophages, dendritic cells (DCs), eosinophils, neutrophils, and so on in each cancer sample with CIBERSORT algorithm. \( p < .05 \) was the threshold of a credible sample for estimating the proportion of immune cells.

**Evaluation of Immunotherapy and Chemotherapy on Iron Metabolism Patterns**

Tumor Immune Dysfunction and Exclusion (TIDE) score is a computational framework developed based on the analysis and modeling of characteristic genes for T cell exclusion and T cell dysfunction in immunosuppression at high levels of cytotoxic T lymphocytes (CTL) (Jiang et al., 2018). We applied four indicators to predict the efficacy of immunotherapy, including exclusion score, dysfunction score, microsatellite instability (MSI), and TIDE. The chemotherapeutic response of BLCA patients was evaluated by Genomics of Drug Sensitivity in Cancer (GDSC) (https://www.cancerRxgene.org). Six chemotherapeutic drugs in BLCA treatment were chosen, including Gemcitabine, Cisplatin, Docetaxel, Mitomycin-C, Doxorubicin, and Paclitaxel. The ridge regression algorithm was used to calculate the half-maximal inhibitory concentration (IC50), and satisfactory prediction accuracy was obtained through 10 times cross-validation (Geeleher et al., 2014). The calculation process was completed by the “pRRophetic” R package.

**Construction and Validation of IMRG Prognostic Signature**

According to the prognosis-related IMRGs in the univariate Cox regression model, the “glmnet” package in R performed the least absolute shrinkage and selection operator (LASSO) to identify important prognostic IMRGs and select one standard error (SE) above the minimum criteria. The multivariate Cox regression model made it more optimized and practical. Finally, the IMRGscore formula was obtained:

\[
Risk\text{ }score = (\exp\text{ }Gene1 \times \text{coefficient } gene1) + (\exp\text{ }Gene2 \times \text{coefficient } gene2) + \ldots + (\exp\text{ }GeneN \times \text{coefficient } GeneN)\\
\]

On the basis of the optimal cut-off of IMRGscore obtained by the “surv_cutpoint” function in R, we divided BLCA patients into high-risk and low-risk groups. With the help of Kaplan-Meier analysis (“survival” package) and receiver operating characteristic (ROC) curve (“timeROC” package), the predictive ability of the prognostic signature was evaluated. The diagnostic accuracy was estimated by the area under the curve (AUC). The same IMRGscore calculation formula, cut-off value and, analysis methods were applied in the GSE13507 cohort to validate the signature.

**Establishment and Evaluation of the Nomogram**

Nomogram is an intuitive clinical prognosis prediction model integrating a variety of clinicopathological features related to prognosis. We established a nomogram model to provide a more accurate prediction of prognosis for clinical patients based on IMRGscore and clinical pathological characteristics. First, univariate Cox regression analysis was utilized to assess the predicted values of variables. Then further determined the coefficient via multivariate Cox regression analysis. The “rms” R package then established a nomogram for predicting the operating system. Concordance index (C-index) and calibration analysis were applied to estimate the accuracy and consistency. Finally, the clinical application value of the nomogram is evaluated using Decision Curve Analysis (DCA). These analyses were performed with “survival” package.

**Statistical Analysis**

All statistical analyses were completed with R software (version 4.0.4) in this study. Before establishing and verifying the prognostic signature, the batch differences between the TCGA dataset and GEO datasets were removed through the “sva” package. Wilcoxon rank-sum test was to verify the significance of the difference in two groups. When comparing more than two
groups, the Kruskal Wallis test was selected to verify the difference. Set $p$-value < .05 as a statistically significant standard.

RESULTS

To describe our research intuitively and systematically, we showed the research process in Figure 1.

Characterization of Iron Metabolism Patterns in BLCA

Through the univariate COX regression analysis ($p < .05$) of the TCGA-BLCA patients with integrated survival information and cancer tissue expression profile, 95 IMRGs were selected as prognosis-related genes (Table 1). Then we clustered the TCGA-BLCA cohort by NMF algorithm based on these genes. According to cophenetic coefficients, we decided $k = 4$ as the best cluster number (Figure 2A). Figure 2B was the NMF matrix heatmap when $k = 4$, including C1 subtype 89 cases, C2 subtype 141 cases, C3 subtype 91 cases, and C4 subtype 79 cases. Kaplan-Meier survival curves showed that the prognosis of patients in C1 and C3 patterns was better than that of C1 and C2 patterns ($p = .020$) (Figure 2C).

Figure 2D shows the expression of prognosis-related IMRGs in iron metabolic patterns. We also analyzed the clinicopathological differences among distinct iron metabolism patterns (Figure 2E). It was found that the proportion of TNM stages, pathologic stage, histologic grade, OS, DSS, and PFI events was dissimilar among patterns, and the incidence of advanced clinicopathological results in C2 and C4 patterns tended to increase.

Through GSVA analysis, we obtained the rich-concentration pathways among iron metabolism patterns (Figure 2F). We found that C4 pattern was positively related to multiple stromal, carcinogenic, and immune activation related pathways, including TGF-$\beta$ signaling pathway, WNT signaling pathway, MAPK signaling pathway, JAK-STAT signaling pathway, T cell receptor signaling pathway, chemokine signaling pathway, B cell receptor signaling pathway, cytokine-cytokine receptor interaction, NOD-like receptor signaling pathway, TOLL-like receptor signaling pathway, and so on. C2 pattern showed a similar trend to C4 pattern, but C2 was also significantly expressed in a variety of DNA damage repair related pathways. The correlation score of most carcinogenic and immune activation-related pathways was reduced in C1 and C3 patterns. While C3 pattern was also found to exhibit an enriched trend in DNA damage repair related pathways.

Tumor Immune Microenvironment of Iron Metabolism Patterns

In order to investigate whether there are differences in TIME among distinct iron metabolism patterns, we used ESTIMATE and ssGSEA scores for evaluation. ESTIMATE showed that there
were significant differences in the stromal score ($p < .001$), immune score ($p < .001$), and ESTIMATE score ($p < .001$) among the three patterns, of which C4 was the highest, C1 and C3 were the lowest (Figure 3A). Then we analyzed the infiltration differences in immune cells among iron metabolism patterns. The ssGSEA score suggested that the infiltration of all 22 TIICs in iron metabolism patterns was significantly different, among which, the ssGSEA score of TIICs in C1 and C3 patterns was lower and C2 and C4 patterns were higher (Figure 3B). And the enrichment trend of immune-related functions in iron metabolism patterns was similar to that of immune cell infiltration (Figure 3C). Additionally, the expression levels of major histocompatibility complex (MHC) molecules, costimulatory molecules, and adhesion molecules roundly decreased in C1 and C3 patterns (Figure 3D).

**Tumor Somatic Mutation in Iron Metabolism Patterns**

The tumorigenesis frequently occurs after the accumulation of gene mutations (Martincorena and Campbell, 2015). It is also reported that tumor mutation burden (TMB) can be used as a potential prognostic indicator for BLCA (Chan et al., 2019). Consequently, we used the "maftools" R package to show the distribution of somatic mutations and the differences of TMB in various iron metabolism patterns. Through the simple nucleotide variation information of TCGA-BLCA, the mutation spectrum and TMB of each sample was obtained. In BLCA samples, the 20 genes with the highest mutation rate were TP53, TTN, KMT2D, MUC16, ARID1A, KDM6A, PIK3CA, SYNE1, RB1, HMCN1, FGFR3, RYR2, KMT2C, MACF1, EP300, FLG, FAT4, STAG2, ATM and OBSCN (Figures 4A–D). C2 pattern had the highest mutation rate of TP53, while the mutation of TTN and KMT2D mostly happened in C3 pattern. The mutation rates of these three genes in C1 and C4 patterns were significantly reduced. Most gene mutations were missense-mutation. In patients with BLCA, high TMB indicated a better prognosis (hazard ratio [HR] = .65 (.48–.88), $p = .005$) (Figure 4E). Additionally, we found that the TMB of C2 and C3 patterns was significantly upper than that of C1 and C4 patterns (Figure 4F).

**Evaluation of Immunotherapy in Iron Metabolism Patterns**

Lately, immune checkpoint inhibitors (ICIs) have gradually become the second-line treatment for advanced BLCA. Therefore, we analyzed the expression of some immune checkpoints (PDCD1 (PD-1), CD274 (PD-L1), PDCD1LG2 (PD-L2), LAG3, TIGIT, IDO1, and CTLA4) among different iron metabolism patterns to predict the efficacy of immunotherapy (Figure 5A). The expression

---

**FIGURE 2** | Non-negative matrix factorization clustering of iron metabolism patterns in TCGA-BLCA cohort. (A) Cophenetic coefficients of NMF algorithm clustering number from 2–10 (B) NMF matrix heatmap when k = 4. (C) Kaplan-Meier survival analysis of iron metabolism patterns. (D) Expression Heatmap of prognosis-related IMRGs. (E) Correlation between iron metabolism pattern and clinicopathological characteristics. (F) Heatmap of GSVA analysis among iron metabolism patterns. *$p < .05$; **$p < .01$; ***$p < .001$. 

---

**FIGURE 3** | (A) Cophenetic coefficients of NMF algorithm clustering number from 2–10 (B) NMF matrix heatmap when k = 4. (C) Kaplan-Meier survival analysis of iron metabolism patterns. (D) Expression Heatmap of prognosis-related IMRGs. (E) Correlation between iron metabolism pattern and clinicopathological characteristics. (F) Heatmap of GSVA analysis among iron metabolism patterns. *$p < .05$; **$p < .01$; ***$p < .001$.
of all immune checkpoints in C2 and C4 patterns was signally higher than that in the other two patterns. This suggested that C2 and C4 might be more suitable for ICIs treatment. However, the high expression level of immune checkpoints may be related to the formation of the immunosuppressive microenvironment (Dunn et al., 2002). This conclusion was confirmed in Figures 5B,C. C4 immune exclusion score was observably higher than other iron metabolism patterns. The immune dysfunction score of C4 pattern also increased signally, while C3 was the lowest. Moreover, we used the TIDE algorithm to evaluate ICIs response, in which the MSI of C1 and the TIDE of C4 were the highest (Figures 5D,E).

Chemotherapeutics Drugs Response in Iron Metabolism Patterns
Chemotherapeutics drugs are widely used in the treatment of BLCA, including intravesical instillation and systemic chemotherapy. Consequently, we evaluated the IC50 values of
six commonly used chemotherapeutic drugs (Gemcitabine, Cisplatin, Docetaxel, Mitomycin-C, Doxorubicin, and Paclitaxel) in each sample (Figures 6A–F). Among the six drugs, C2 patterns showed the lowest IC50 value. In Gemcitabine, Cisplatin, Docetaxel, and Paclitaxel, the IC50 value of C4 pattern was also lower than that of C1 and C3 patterns. As for C3 pattern, the IC50 value in Gemcitabine, Doxorubicin, and Paclitaxel was higher than that of C1 pattern. The above results strongly indicated that C2 was the most sensitive to chemotherapeutic drugs, C4 was the second, while C3 was more insensitive to chemotherapeutic drugs.

Establishment of the IMRG Prognostic Signature in the TCGA-BLCA Cohort

We selected 400 BLCA patients with explicit, non-zero OS and survival status to establish an IMRG signature from the TCGA database. Then used the LASSO Cox regression model to calculate, and selected an SE higher than the minimum standard to further screen the prognostic genes. Finally, through the multivariate COX regression analysis, we obtained the optimal prognostic signature containing 13 IMRGs, including TCIRG1, CTSE, ATP6V0A1, CYP2C8, RNF19A, CYP4Z1, YPEL5, PLOD1, BMP6, CAST, SCD, IFNG, and ASIC3 (Figures 7A,B). And constructed a formula to evaluate the IMRGscore of each patient: IMRGscore = −(.18775 × TCIRG1 expression) − (.073 × CTSE expression) + (.33856 × ATP6V0A1 expression) − (.37089 × CYP2C8 expression) − (.30306 × RNF19A expression) − (.27636 × CYP4Z1 expression) − (.35016 × YPEL5 expression) + (.17559 × PLOD1 expression) + (.25065 × BMP6 expression) + (.23398 × CAST expression) + (.13313 × SCD expression) − (.52087 × IFNG expression) − (.57726 × ASIC3 expression). And according to the optimal cut-off value (cut point = 1.78265), samples were decomposed into low- and high-risk groups. Kaplan-Meier survival analysis showed that the OS of the low-risk group was longer (hazard ratio [HR] = 4.49 (3.29–6.14), \( p < .001 \)) (Figure 7C). And the AUCs for the 1-, 3-, and 5-year OS survival rates were .741, .772, and .755, respectively (Figure 7D). The risk score distribution, survival status graph, and expression profile heatmap were shown in Figures 7E–G. The proportion of patient deaths was observably positively correlated with the IMRGscore. The expression of ATP6V0A1, PLOD1, BMP6, CAST, and SCD were up-regulated in the high-risk group, while TCIRG1, CTSE, CYP2C8, RNF19A, CYP4Z1, YPEL5, IFNG, and ASIC3 were down-regulated.
FIGURE 5 | Evaluation of immunotherapeutic therapy in iron metabolism patterns. (A) The expression of immune checkpoints in iron metabolism patterns. The comparisons of exclusion score (B), dysfunction score (C), MSI (D) and TIDE (E) among iron metabolism patterns. *p < 0.05; **p < 0.01; ***p < 0.001.

FIGURE 6 | Evaluation of chemotherapy in iron metabolism patterns. The comparisons in IC50 value of Gemcitabine (A), Cisplatin (B), Docetaxel (C), Mitomycin-C (D), Doxorubicin (E), and Paclitaxel (F) among iron metabolism patterns.
Confirmation of the IMRG Signature in the GSE13507 Cohort

As the test set, 165 BLCA samples in the GSE13507 cohort were grouped using the same IMRGscore calculation formula and cut-off value of the train set to validate the applicability and stability of the IMRG signature. Consistent with the above conclusion, patients in the low-risk group had a better OS (hazard ratio [HR] = 2.65 (1.52–4.60), p = .001) in the GSE13507 cohort (Figure 8A). The AUCs for the 1-, 3-, and 5-year OS survival rates were .753, .630, and .552, respectively (Figure 8B). The conclusions of the risk score distribution, survival status graph, and expression profile heatmap were consistent with the training set (Figures 8C–E).

Clinical Relevance of the IMRG Signature

To further supplement the clinical application value of the IMRG prognostic signature, we integrated the significant differences in IMRGscore among distinct subgroups of BLCA patients with clinicopathological characteristics (Figure 9A). Heatmap indicated that the advanced TNM stages, pathologic stage, histologic grade, aging, and worse OS, DSS, and PFI events had an elevated trend in IMRGscore.

Since the significant correlation between signature and clinicopathological stage, we determined whether the IMRGscore was a clinically independent predictor of BLCA patients (Figures 9B,C). Univariate Cox regression analysis showed that advanced pathologic stage (p < .001), aging (p = .005), and higher IMRGscore (p < .001) were unfavorable factors for OS. After performing the multivariate Cox regression analysis, we confirmed that the IMRGscore was an independent prognostic parameter.

Establishment of a Nomogram Based on the IMRG Signature

According to the above result from univariate Cox regression analysis of the TCGA-BLCA cohort, we established a nomogram model containing pathologic stage, age, and IMRGscore (Figure 10A). After removing the patients without complete information and the subgroups of variables with few samples, a total of 362 patients were included. We standardized each variable with a score from 0 to 100 and summed the scores to obtain the total number of points for each BLCA patient. The predicted 1-, 3-, and 5-year survival probabilities of each patient were standardized according to the relationship between the positions along with the prognosis and total points axes. The C-index reached 0.694 (95% CI: 0.653–0.735). Figures 10B–D suggested that the nomogram model predicted that the prognosis results of TCGA-BLCA patients would fit well with the actual prognosis results. Besides, DCA curves revealed that the signature provided patients with a stable and significant net benefit in BLCA patients (Figures 10E–G). Then based on the nomogram model, we built the Kaplan-Meier survival curve and the time-dependent ROC curves. In the TCGA-BLCA cohort, we divided samples into high- and low-risk groups with the optimal cut-off value (cut point = −.122532). Patients in the high-risk group showed a significantly poor OS (hazard ratio [HR] = 4.22 (2.92–6.10), p < .001, Figure 10H). The AUCs for the 1-, 3-, and 5-year OS survival rates were .764, .769, and .760, respectively (Figure 10I). Additionally, we verified the nomogram model in the GSE13507 cohort. The cut-off value was consistent with the train set. Kaplan-Meier survival analysis showed that patients in the low-risk group had a better OS than those in the high-risk group (hazard ratio [HR] = 6.47 (2.89–14.49), p < .001, Figure 10J).

FIGURE 7 | Construction of IMRG signature and prognosis analysis based on the training set. (A,B) LASSO regression identified 13 IMRGs. (C) Kaplan-Meier survival analysis between IMRGscore-defined groups. (D) Time-dependent ROC curves of IMRG signature. (E) IMRGscore distribution. (F) Survival status map. (G) IMRGs expression profiles heatmap.
and the AUCs were .896, .906, and .915 (Figure 10K).

GSEA
To further comprehend the effect of IMRGs expression on the biological characteristics of BLCA, we carried on GSEA analysis in IMRGscore-defined groups (Figures 11A,B). The Kyoto Encyclopedia of Genes and Genomes (KEGG) results revealed that in the high-risk group, the main enrichment pathways were ECM receptor interaction, regulation of actin cytoskeleton, MAPK signaling pathway, WNT signaling pathway, pathways in cancer. While the low-risk group is mainly concentrated in allograft rejection, asthma, primary immunodeficiency, and so on. Furthermore, Figures 11C,D showed the enrichment of the high- and low-risk groups in the Gene Ontology biological process (GOBP). We found that the low-risk group was enriched in multiple immune functions, such as activation of immune response, adaptive immune response, B cell-mediated immunity, and so on.

Correlation Between Tumor Immune Microenvironment and IMRG Signature
As we mentioned before, iron metabolism is closely related to TIME. Based on the CIBERSORT algorithm, we calculated the proportion of 22 TIICs in each TCGA-BLCA sample (Figure 12A). After selecting samples with significant immune cell fraction results ($p < .05$), 195 samples were included in the difference analysis, including 141 cases in the low-risk group and 54 cases in the high-risk group. Then the difference in the proportion of TIICs between the IMRGscore-defined groups was shown in Figure 12B. It was found that the fraction of CD8 T cells, activated CD4 memory T cells, follicular helper T cells (TFH) and regulatory T cells (Treg) in the low-risk group was significantly higher. In contrast, the fraction of M0 macrophages was upper in the high-risk group. Furthermore, high levels of CD8 T cells ($p = .004$), activated CD4 memory T cells ($p = .013$) and TFH ($p = .041$) were observably associated with better OS (Figures 12C–E), while increased M0 macrophages ($p = .035$) indicated a poor OS (Figure 12F).

DISCUSSION
Bladder cancer is a heterogeneous malignancy. Patients with BLCA generally show different prognoses because of the molecular discrepancy (Knowles and Hurst, 2015; Guo and Czerniak, 2019). At present, it is generally believed that clinical or pathological stages were insufficient to predict the
The prognosis of patients with BLCA (Konety, 2006; Rosenberg et al., 2013) is crucial, making it essential to develop a more accurate and efficient model to predict survival outcomes. Recent studies have linked iron involvement to cancer appearance and progression. Iron metabolism has been shown to play a significant role in various stages of carcinogenesis. Murata M. demonstrated that iron released from damaged transferrin could mediate the Fenton reaction and generate ROS, contributing to carcinogenic processes (Murata, 2018). A cross-sectional study revealed that the serum iron concentration in patients with BLCA was lower than that in the control group (Mazdak et al., 2010). However, studies exploring the role of iron metabolism in BLCA treatment and prognosis are scarce.

In this study, patients in the TCGA-BLCA cohort were divided into four iron metabolism patterns based on the expression of prognosis-related IMRGs. Survival analysis indicated that patterns C1 and C3 had a better prognosis. Our findings also showed that patterns C2 and C4 had higher enrichment scores in various pathways, including NOD-like receptors (NLRs), Toll-like receptors (TLRs), and Notch signaling. These pathways are crucial in coordinating the tumor immune microenvironment and promoting angiogenesis, cancer stem cells, and chemotherapy resistance, thus enhancing tumor risk (Liu et al., 2019). TLRs are transmembrane pattern recognition receptors that detect and defend microbial pathogens through the innate immune response (Brennan and Gilmore, 2018). The activation of the Notch signaling pathway is observed in multiple components of the tumor microenvironment (TME), such as angiogenesis, tumor stem cell maintenance, immune infiltration, or therapeutic resistance (Meurette and Mehlen, 2018). C2 and C4 iron metabolism patterns were significantly infiltrated by immune cells.
FIGURE 10 | Nomogram model based on clinicopathological characteristics and IMRGscore. (A) Nomogram for predicting the probability of OS over 1, 3, and 5 years. (B–D) Calibration curves for predicting the fitness of the nomogram model in 1, 3, and 5 years. (E–G) DCA curves based on three predictors of 1, 3, and 5 years. (H) Kaplan-Meier analysis of nomogram model in the TCGA-BLCA cohort. (I) Time-dependent ROC curves of nomogram model in the TCGA-BLCA cohort. (J) Kaplan-Meier analysis of nomogram model in the GSE13507 cohort. (K) Time-dependent ROC curves of nomogram model in the GSE13507 cohort.
and the expression of MHC genes was highly elevated. These characteristics are consistent with an immune-inflamed phenotype. On the other hand, C1 and C3 are suspected to be immune-desert phenotypes due to lack of immune infiltration and antigen presentation. Additionally, we found that multiple immune checkpoints (PD-1, PD-L1, PD-L2, LAG3, TIGIT, IDO1, and CTLA4) were highly expressed in C2 and C4 patterns, which might indicate that immunotherapy has a better effect on them. Furthermore, studies have shown that the high-level expression of immune checkpoint genes might be a stimulative to the immunosuppressive microenvironment and led to the immune escape of tumor cells (Dunn et al., 2002). The TIDE algorithm confirmed this conclusion. Therefore, we speculate that the reason why C2 and C4 patterns do not show a better prognosis of immunoinflammatory phenotype is that the antitumor effect based on the activated immune pathway and high infiltration level of T cells were eliminated by the formation of the immunosuppressive microenvironment. The above evidence proved that iron metabolism is of great significance in shaping various TME landscapes in BLCA.

Somatic mutation is not only the driving factor of cancer, but also the guiding basis for diagnosis and treatment. The three genes with the highest mutation rate in BLCA were TP53, TTN, and KMT2D. The mutation rate of TP53 in C2 iron metabolism pattern was the highest, while in C1 pattern was the lowest. In C2 and C3 patterns, the incidence of TTN and KMT2D mutations was significantly higher. Detection of TP53 mutation was conducive to estimating the high risk of early lesions (Olivier et al., 2010). Single TTN gene mutation often indicated high TMB (Oh et al., 2020). The mutation of epigenetic regulator KMT2D was a biomarker of poor prognosis in some cancers (Ferrero et al., 2020). Additionally, we found that high TMB suggested that patients with BLCA had a better prognosis through the TCGA database. Consistently, the TMB of C1 and C4 patterns was observably decreased. These results show that iron metabolism has a complex interaction with somatic mutation.

**FIGURE 11** Correlation between IMRG signature and biological functions. KEGG results of the high-risk group (A) and low-risk group (B). GOBP results of the high-risk group (C) and low-risk group (D).
Chemotherapy is still one of the main treatments for BLCA. According to the latest guidelines of the European Association of Urology (EAU), all muscle-invasive bladder cancer (MIBC) patients with physical conditions can apply platinum-based neoadjuvant chemotherapy before operation (Witjes et al., 2021). This study investigated the efficacy of six common chemotherapeutic drugs on iron metabolism mode, including Gemcitabine, Cisplatin, Docetaxel, Mitomycin-C, Doxorubicin, and Paclitaxel. The results showed that C2 pattern was the most sensitive to these chemotherapeutic drugs, while C3 pattern was the most insensitive. This provided a reference basis for the selection of clinical chemotherapy drugs.

Due to the effect of iron metabolism on the tumorigenesis and progression of cancer, it is of great importance to establish an IMRG signature for predicting the prognosis of patients with BLCA. Herein, we applied stepwise regression analysis to compose a clinical prognostic signature for BLCA patients with 13 IMRGs (TCIRG1, CTSE, ATP6V0A1, CYP2C8, RNF19A, CYP4Z1, YPEL5, PLOD1, BMP6, CAST, SCD, IFNG, and ASIC3). A test set was utilized to confirm its accuracy and stability. Moreover, the IMRGscore was elevated in multiple advanced clinicopathological stages. And multivariate Cox regression analysis verified that IMRGscore was an independent prognostic index of BLCA patients. Finally, we combined some clinicopathological features, including pathologic stage, age, and IMRGscore to construct a nomogram that accurately predicted the prognosis of patients with BLCA. The accuracy and clinical contributions were verified by calibration analysis and DCA.

Several studies have shown that these 13 IMRGs are closely related to cancers, and some of these genes have been confirmed about BLCA. TCIRG1, one of the V-ATPase subunits, is abnormally overexpressed in patients with recurrent hepatocellular carcinoma, and enhances the ability of metastasis by regulating the growth, death, and epithelial to mesenchymal transition of cancer cells (Yang et al., 2018). A retrospective study suggests that CTSE can be used as an independent prognostic marker for NMIBC, so as to guide the treatment of patients (Lin et al., 2001). ATP6V0A1 was demonstrated that it could enhance the fusion of autophagosomes and lysosomes, up-regulate autophagy volume accumulation, and finally induce autophagic cell death (Hsin et al., 2012). CYP2C8 can be regulated by GAS5/miR-382-3p in hepatocellular carcinoma and play an anticancer role (Li and Chen, 2020). RNF19A was confirmed to be overexpressed in non-small cell lung cancer, which plays a carcinogenic role by destroying the function of p53 (Cheng et al., 2021). CYP4Z1 was found to be highly expressed in BLCA and positively correlated with the progression of histologic grade and pathologic stage (Al-Saraireh et al., 2021). YPEL5 was found to be a marker for the early detection of cancer.
to be inhibited by METTL3-m6A (N6-methyladenosine)-YTHDF2 axis in colorectal cancer, promoting the growth and metastasis of tumor (Zhou et al., 2021). The expression of PLOD1 was an independent prognostic factor in BLCA patients, and downregulated by inhibitor could significantly reduce the invasiveness of BLCA cells (Yamada et al., 2019). The expression of BMP6, a key endogenous regulator of iron metabolism, was affected by Med19, which could promote bone metastasis and invasiveness of bladder cancer (Andriopoulos Jr et al., 2009; Wen et al., 2013). Public clinical data also confirm that BMP6 is a prognostic marker for bladder cancer (Yuen et al., 2012). Calpastatin (CAST) is involved in many important physiological processes, including cell cycle, ECM, cancer cell proliferation, metastasis, and apoptosis (Nian and Ma, 2021). SCD can protect cancer cells from oxidative stress and ferroptosis through mediated lipogenesis in prostate cancer with over-activation of PI3K-AKT-mTOR signaling (Yi et al., 2020). SCD has also been shown to reduce proliferation and invasion of BLCA cells when inhibited (Piao et al., 2019). Iron metabolism can affect innate immune response by affecting IFNG mediated immune response pathway in macrophages (Nairz et al., 2014). And IFNG was demonstrated to inhibit the activity of bladder cancer stem cells (Qiu et al., 2020). ASIC3 has an H+ gating function, which promotes the acid-induced epithelial-mesenchymal transition in pancreatic cancer cells (Zhu et al., 2017).

GSEA analysis showed that several cancer-related pathways were enriched in the high-risk group. The unique biochemical and biophysical properties of ECM when it is dysregulated are the key drivers of cancer progression (Walker et al., 2018). The MAPK signaling pathway is considered to be related to cell proliferation, differentiation, migration, aging, and apoptosis (Sun et al., 2015). The Wnt signaling pathway is an important driving factor to maintain tissue development and homeostasis. Abnormal Wnt signaling will cause the occurrence and progression of many cancers by affecting cancer stem cells (Duchartre et al., 2016). Furthermore, CIBERSORT analysis indicated that several TIICs with differential infiltration levels had significant correlations between the prognosis of BLCA patients, and played a regulatory role in the balance of iron metabolism. After being activated by immunotherapy, CD8 T cells can enhance ferroptosis-specific lipid peroxidation and increase ferroptosis in tumor cells, which contributes to the anti-tumor effect (Wang et al., 2019). Macrophages are the regulatory hub of iron metabolism. Macrophages phagocytize and degrade aging and damaged erythrocytes to recover iron, and also have the ability to release iron. The accumulation of M2 macrophages is often associated with poor prognosis, and M2 macrophages possess iron release characteristics (Recalcati et al., 2010). Therefore, the accumulation of M2 macrophages may further aggravate the disorder of iron metabolism.

Our research has obvious advantages in the study of iron metabolism characteristics and the clinical application of BLCA. First, we divided BLCA patients into distinct iron metabolism landscapes to further confirm the relationship between iron metabolism and TME and somatic mutations. At the same time, we also verified that iron metabolism landscapes had guiding significance for chemotherapy drugs and immunotherapy selection. Next, we established the IMRG prognostic signature and proved IMRGscore was an independent prognostic factor for BLCA patients, and it also had the predictive ability for clinicopathological characteristics. At present, our research is still insufficient. First, it is a retrospective study. Deviations in variables such as clinicopathological characteristics of patients most likely exist. Second, our prognostic signature depends on gene expression and does not take into account the effects of gene mutation, methylation, or other factors on the prognosis of BLCA patients. Finally, the prognostic signature can be incorporated into large sample prospective studies to further verify its clinical value.

DATA AVAILABILITY STATEMENT

Raw data for this study were generated at the TCGA database with the cancer type of BLCA. The datasets used and/or analyzed during the current study are available from the GEO database (GSE13507). Derived data supporting the findings are available from the corresponding author (WS) on reasonable request.

AUTHOR CONTRIBUTIONS

XL and WS designed and guided the work. XS and SX participated in data collecting, data processing, figures preparation, and manuscript writing. YZ, JM, and CD contributed to statistical analysis. KC, LC, and FL contributed to manuscript writing and article publishment. ZL and TW contributed to manuscript draft writing. JL revised the manuscript critically. All authors provided critical advice for the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Number: 82072838), Tongji Outstanding Young Researcher Funding (Grant number: 2020YQ13), Huazhong University of Science and Technology (Grant Number: 2019kfyXJKC06), Natural Science Foundation of Hubei Province (Grant Number: ZRMS2020002466).

ACKNOWLEDGMENTS

The authors thank the members of the Urology Department of Tongji Hospital of Huazhong University of science and technology for their valuable support and useful discussion. We also thank TCGA, GEO, and MSigDB databases for the availability of the data.
