Research Article

Identification of a New Prediction Model for Bladder Cancer Related to Immune Functions and Chemotherapy Using Gene Sets of Biological Processes

Huidong Zhou,1,2 Guo Huang,3,4 Tao Chen,2 Jianglei Zhang,1 and Jun Ouyang1

1Department of Urology, The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu, China
2Department of Urology, The Fourth Hospital of Changsha, Changsha, Hunan, China
3Hengyang Medical College, University of South China, Hengyang, Hunan, China
4Hunan Province Key Laboratory of Tumor Cellular & Molecular Pathology, Cancer Research Institute, Hengyang Medical School, University of South China, Hengyang, Hunan, China

Correspondence should be addressed to Jun Ouyang; ouyangjun99@sina.com

Received 15 May 2022; Revised 20 September 2022; Accepted 29 September 2022; Published 30 October 2022

Academic Editor: Hassan Dastsooz

Copyright © 2022 Huidong Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Biological processes serve crucial functions in the initiation and development of cancer. Therefore, we constructed and validated a model for bladder cancer (BLCA) with good predictive power for immunity, prognosis, and therapy. Methods. Using the expression of the gene sets based on biological processes, BLCA patients were divided into three clusters by consensus cluster analysis. By performing LASSO regression analysis twice, key genes were selected, and the biological processes-related genes’ (BPRG) score was calculated. Differences in immune infiltration, tumor microenvironment, tumor mutation burden, immunotherapy, and sensitivity towards chemotherapy were analyzed between two groups divided by BPRG score. Results. Good accuracy was observed for the three clusters. They showed different prognoses and levels of immune cell infiltration. The selected key genes were mainly enriched in immune-related pathways. The high-BPRG score group was related to poor prognosis, higher immune cell infiltration, interstitial scores, and increased tumor mutation. Moreover, the effects of immunotherapy were good, and those of chemotherapy were poor. Conclusion. Overall, key genes may be involved in various complex immune regulation processes. Therefore, the quantification and verification of the BPRG score are expected to facilitate the understanding of the immunosuppressive microenvironment in BLCA and guide the choice of chemotherapeutic drugs and immunotherapeutic regimens and help predict the prognoses of patients with BLCA.

1. Introduction

Bladder cancer (BLCA) is a common malignancy worldwide. The Centers for Disease Control and Prevention in the United States estimated 83,730 new cases of BLCA in 2021 [1]. Once the tumor progresses to the local advanced or the metastasis stage, the effects of traditional treatment methods, including surgery and chemotherapy, are not favorable [2]. Immune checkpoint inhibitor (ICI) therapy, such as CTL-associated protein 4 (CTLA4), programmed death-ligand 1 (PD-L1), and programmed death 1 (PD1), has brought new hope for the treatment of advanced BLCA [3]. Although BLCA is immunogenic cancer with characteristics of high tumor mutation burden (TMB) and heterogeneity [4], the efficacy of ICI treatment in many patients is insignificant. Several factors including the tumor microenvironment (TME) influence the effectiveness of ICI treatment. Coupled with high recurrence rates and risk of metastases in BLCA, only a few biomarkers can well predict these patients’ prognosis and therapeutic effectiveness [5]. Therefore, the identification and
quantification of molecular maps that affect the choice of drug treatment regimens are crucial. With recent advancements in the understanding of the biological behavior of BLCA and the progress in high throughput sequencing technology, the identification of new indicators to predict the prognosis of patients with BLCA is possible [6].

Many biological processes (BP) are linked to and altered during the initiation and development of cancer [7]. Therefore, determining BPs involved and quantifying the expressions of relevant genes in the study of cancer cell behavior are crucial. In addition, linking treatment sensitivity of some key components of BPs may also facilitate personalized medication guidance [8]. Paterno et al. show that DNA damage responses may increase breast cancer risk and radiosensitivity amongst these patients. Aberrant DNA repair BPs is associated with malignancy in breast cancer patients [9]. Similarly, the process of epithelial-mesenchymal transition (EMT) is tightly linked to invasion and metastasis in BLCA [10].

Herein, we aimed to establish a scoring model using biological process-related genes (BPRG) for prognostic prediction and guidance of clinical treatment for BLCA. First, in the TCGA-BLCA cohort, we calculated the expression of 7658 gene sets related to BPs by GSVA and divided BLCA patients into three clusters. Different clusters showed differential processes and immune infiltration levels. Subsequently, key genes related to BPs were selected, and a BPRG score model was constructed by the LASSO-Cox method. After classification, the model was found to be closely linked to immune responses and chemotherapy. In conclusion, our results demonstrated the potential relationship of BPRG with prognosis, immune microenvironment, immunotherapy, and chemotherapy in patients with BLCA might be beneficial for developing personalized therapy and assessment strategies.

2. Materials and Methods

2.1. Raw Data Extraction and Processing. The clinical characteristics and transcriptome of BLCA patients were obtained from the GEO and The Cancer Genome Atlas (TCGA) databases, respectively (GSE13507 and TCGA-BLCA) [11]. The RNAseq and somatic mutations data were downloaded in the FPKM format from TCGA. The average value was taken when multiple Ensembl IDs were mapped to the same gene. All sequencing data downloaded from the GEO database were processed by log quadratic transformation, background adjustment, and normalization. Finally, the data of 565 BLCA patients were retained for further analysis after excluding those without survival information.

2.2. Gene Set Variation Analysis (GSVA). GSVA is a nonparametric unsupervised method [12] to assess the enrichment in sequenced gene sets. The gene expression matrix was transformed into an expression matrix of gene sets to evaluate the differences in enrichment among samples. Herein, the “GSVA” R package and gene sets of BPs in the MSigDB database were employed.

2.3. Tumor Subtype Determination. Based on the expression profile of the BP gene sets, consensus cluster analysis was conducted using the “CancerSubtypes” R package to identify the characteristics of different gene sets [13] and classify BLCA patients. Kaplan-Meier analysis was performed to compare differences in the overall survival of patients among clusters. Pearson’s chi-squared test was employed to analyze the correlation between clinical factors and clusters; the results were visualized using the “Complex Heatmap” package.

2.4. Screening Prognosis-Related Gene Set. The “Limma” R package was used to screen the differential gene sets among the three clusters [14]. Those relevant to survival were selected by univariate Cox and KM survival analyses (a total of 66). Five prognosis-related gene sets were screened by LASSO regression using the “glmnet” package in R [15]. The derived coefficients were utilized to obtain the risk score, and the patients were classified into high- and low-risk groups using the median value of the risk coefficient. The differences in OS between the groups were evaluated by the KM survival curve, and significance was assessed using the log-rank test. The same score calculations, cut-off values, and analysis methods were used to verify the GSE13507 cohort.

2.5. Construction of the Predictive Model. 112 genes were extracted from the prognosis-related gene sets, and 27 key survival-related genes were screened by univariate Cox and KM survival analyses. All BLCA patients were randomly classified into test (n = 282) and training (n = 283) groups in the ratio of 1:1, following which the Lasso-Cox regression algorithm was executed. The “glmnet” R package was used to reduce the risk of overfitting, and the model was established by 10x cross-validation. Key genes were selected following the multivariate Cox analysis, and BPRG scores were established (BPRG score = Σ(Expi * coefi)), where coefi and Expi correspondingly indicate each gene’s coefficient and expression. The patients were classified into low and high-BPRG score groups based on the median score. The KM survival analysis was performed, and a receiver operating characteristic (ROC) curve was plotted; the findings were verified in the validation set.

2.6. Establishment and Verification of the Nomograph Scoring System. Clinical characteristics and BPRG score were utilized to construct the predictive nomogram using the “rms” R package. In the nomograph scoring system, each variable for each sample was matched to a score, and the total score was calculated by adding the addition of the scores of all variables. Nomograms were evaluated using time-dependent ROC curves for 1-, 3-, and 5-year survival. The calibration chart for the nomogram was used to describe the consistency between the predicted 1-, 3-, and 5-year survival events and the obtained results.

2.7. Assessment of Immune Status between Different Groups. The ESTIMATE algorithm was used to quantify the abundances of 22 immune cell infiltrates among the three clusters and between the high- and low-BPRG score groups to assess the proportion of tumor-infiltrating immune cells (TIICs) [16, 17]. The correlation of immune functions and interstitial scores with the TME in different groups was compared. Moreover, the benefits of immunotherapy between the high-
RNAseq data of BLCA patients from TCGA (n = 430)

Excluded tumor tissues with incomplete clinical information (N = 30)

NMF algorithm for subtype of BLCA

Survival analysis

Overlapped gene sets of three subtype (n = 505)

Univariate COX regression analysis

Lasso COX regression analysis

Validation cohorts GSE13507 n = 145

5 prognosis-related gene sets

Immune infiltration

GSEA

Immune infiltration

Tumor burden mutation

Tumor microenvironment

ICIs response

Chemo drugs response

Prognostic nomogram

Type

Normal

Tumor

Figure 1: The flowchart of the study design and the heat map of TCGA data. (a) Flowchart of the critical steps in the study. (b) Enrichment score for gene sets of biological processes between cancers and paracancerous tissues in TCGA-BLCA.
Figure 2: Classification of TCGA-BLCA cohort based on the biological processes enrichment score. (a) BLCA patients were grouped into three clusters according to the consensus clustering matrix for $K = 3$. (b) The PCA scatter plot indicating that three clusters were well distinguished. (c) NMF matrix heat map of each cluster. (d) Kaplan-Meier curves showing three gene clusters that are significantly related to overall survival in BLCA patients (log-rank, $p < 0.001$). (e) Identification of clustering value using silhouette width plots which shows a good accuracy.
Figure 3: Continued.
and low-BPRG score groups were assessed by using the tumor immune dysfunction and exclusion (TIDE) algorithm [18]. The TIDE score was calculated using the TIDE website after normalizing the expression data.

2.8. Gene Mutation and Drug Sensitivity Analyses. The mutation annotation format (MAF) in the TCGA database was generated using the "maftools" R package to determine differences in somatic mutations among BLCA patients between the high- and low- BPRG score groups [19]. We also calculated the TMB score for each BLCA patient in both groups, visualized the results using KM curves, and examined the significance of log-rank tests. The R package, pRRo-

2.9. Statistical Analysis. Independent t-tests were employed to compare the continuous variables between the two groups. The χ² test was used for analyzing classified data. TIDE scores between the groups were compared by the Wilcoxon test. KM survival curves were plotted following univariate survival analysis, and the log-rank test was employed to test the significance of the results. The Cox regression model was employed for univariate and
Figure 4: Continued.
multivariate survival analyses. Two-tailed \( P < 0.05 \) was regarded as statistically significant. All statistical data were analyzed using the packages in R (https://www.r-project).

3. Results

3.1. Changes in BP Gene Sets between BLCA and Adjacent Normal Tissues. To interpret the changes in the enrichment of BP gene sets between BLCA and adjacent normal tissues, first, we downloaded the BP gene set (c5.go.bp.v7.5.1.symbols, including 7658 gene sets) from the GSEA database, following which the BLCA transcriptome and clinical data from TCGA and GEO databases were extracted. TCGA cohort comprised 411 tumor tissues and 19 adjacent normal tissues, while GSE13507 included 165 tumor tissues and 58 adjacent tissues. The flow of this study and the results of the GSVA of the TCGA-BLCA cohort are shown in Figures 1(a) and 1(b).

3.2. Determination of BLCA Clusters and their Correlation with Clinical Features. TCGA-BLCA was subtyped using the "CancerSubtype" package in R software. The NMF algorithm yielded \( K = 3 \) as the optimal cluster number, that is, the cohort could be divided into three clusters according to the GSVA-related score of each sample (Figures 2(a)–2(c)). A total of 30 patients were excluded due to the lack of clinical data. In the follow-up study, prognostic differences across the three clusters were analyzed using clinical data (Figure 2(d)). In particular, C1 showed a significantly good prognosis, while C2 exhibited the worst prognosis. The contour width in the figure was 0.97 (Figure 2(e)), which indicated a good match between the BLCA samples and the corresponding clusters.

Subsequently, we analyzed the clinicopathological differences between three clusters (Figure 3(a)) and found that TNM stage, pathological stage, and histological grade varied in different clusters. In general, as compared to C2 and C3 patients, C1 patients were in the good to medium clinical stage, pathological grade, and TNM stage, and these differences were statistically significant. We further analyzed the differences in TIICs and immune function among different clusters. Twenty-two immune cell types and immune functions had the highest scores in type C3, while the lowest scores were obtained in C1 (Figures 3(b), 3(c)).

3.3. Screening Prognostic Gene Sets. We calculated the differential gene sets between the three clusters (adj. \( p < 0.05 \) and \( \text{fold change} > 1.69 \)), then took their intersection and combined the gene sets with the clinical data for univariate Cox regression and KM survival analyses. 112 differential gene sets were obtained. Next, these gene sets were analyzed by LASSO regression and cross-validated (Figures 4(a)–4(c)). Finally, 5 target gene sets were determined according to the risk coefficients as follows: regulation of neuron projection regeneration, regulation of cardiocyte differentiation, GOBP convergent extension, regulation of cell fate specification, and relaxation of cardiac muscle. According to the risk score, TCGA-BLCA patients were classified into two groups for survival analysis (Figure 4(d)). The OS survival rate of patients in the low-risk group was higher (\( p < 0.001 \)). The above risk score calculation formula and cut-off value were utilized in GSE13507 to verify the risk score’s stability and

Figure 4: Construction and evaluation of the prognosis-related gene sets based on BP. (a) The common gene sets among differential gene sets between three clusters. (b, c) LASSO regression identified 5 prognosis-related gene sets. (d, e) KM curves of overall survival across different risk groups in the training set and validation set (log-rank, \( p < 0.001 \)).
Figure 5: Continued.
accuracy, which illustrated that the screening method was meaningful. (Figure 4(e)).

3.4. Construction of the Prognostic Model Using Key Genes. A total of 112 genes were extracted from the target 5 gene set. After survival and LASSO regression analyses and cross-validation, 5 key genes (PTN, SPP1, KREMEN1, EGFR, and LBX2) were obtained (Figure 5(a–d)). The accuracy of the ROC curve for prognostic prediction was verified. The AUC of 1-, 3-, and 5-year OS survival rates was 0.720, 0.679 and 0.669, respectively (Figure 5(e)).

3.5. Establishing the Nomogram. We compared the differences in clinical characteristics between the low- and high-BPRG score BLCA patients to further supplement the clinical applicability of the prognosis model (Figure 6(a)). The results suggested that patients with advanced TNM stage, pathological stage, histological grade, and older age were more frequent in the high-BPRG score group. Univariate Cox regression analysis indicated that gender (p < 0.006), age (p < 0.001), TNM stage (p < 0.001), grade (p < 0.001), and BPRG score (p < 0.001) were factors for adverse OS. After multivariate Cox regression, the BPRG score was confirmed as an independent prognostic factor (Figures 6(b) and 6(c)). To facilitate the clinical application of the prognostic prediction model, an individualized prediction model was constructed, including age, TNM stage, and pathological grade, and correction curves for 1-, 3-, and 5 years were plotted for 328 patients. Each variable was standardized with a score between 0 and 100; these scores were added to obtain a total score for each BLCA patient to predict the probability of survival at one, three, and five years. The C-index of this nomogram reached 0.742 (95% CI: 0.702–0.781). Thus, the prognosis of TCGA-BLCA patients predicted by the nomogram model was in agreement with the actual prognosis (Figures 6(d) and 6(e)).

3.6. Differences in Molecular Characteristics between High- and Low-BPRG Groups. GSEA was used to identify enriched gene sets among different groups. Interestingly, activation of immune response, cell adhesion mediated by integrin, chemokine signaling pathway, cytokine receptor interaction, and other pathways related to immune response were enriched in the high-BPRG score group (Figure 7(a)).

Next, gene mutations were analyzed to understand the molecular differences between the two groups. The waterfall diagram shows the top 20 mutated genes in the high score group and the low score group. TP53 exhibited a high mutation frequency in both groups, especially in the high score group (53%) (Figures 7(b) and 7(c)). Although no significant differences were found between the two groups, owing to the fundamental clinical prognostic value of TMB, the TMB of each BLCA patient was calculated and classified into a low TMB group or high TMB group with the optimal cut-off value of 10.158. The survival time of patients with high TMB was significantly longer than that of patients with low TMB. Given the predictive value of TMB and prognostic score for BLCA, we conducted a subgroup survival analysis using the TMB. The overall survival rate was the lowest in the high score and low TMB group (Figures 7(d–f)).

3.7. Evaluation of Immune-Related Characteristics. Since the immune-related pathways were mainly enriched in the high-BPRG score group, the ESTIMATE and ssGSEA scores were used to compare immune cell infiltration between the two groups. Most immune cells, including memory CD4 and CD8T cells, macrophages, activated CD4 and CD8T cells, and resting natural killer (NK) cells, were abundant in the high score group, while CD56 bright natural killer cells
Figure 6: Continued.
and monocytes were abundant in the low score group (Figure 8(a)). In addition, the differences in immune functions between the two groups were calculated (Figure 8(b)). These results were similar to those of immune infiltration levels. In addition, the ESTIMATE algorithm was employed to calculate stromal and immune scores for each patient; significant differences were obtained between the two groups ($P < 0.001$) (Figure 8(c)). We also analyzed the correlation between 5 key genes and immune cells. (Figure 8(d)).

3.8. Correlation between Immunotherapy and Chemotherapy in Different Groups. Subsequently, to examine the role of different scores in antitumor therapy patients with BLCA, we evaluated the correlation between the two groups of patients with immunotherapy and chemotherapy. First, the expression of immune checkpoint-related genes between the two groups was compared (Figure 9(a)), whereby the expression in the high score groups was generally high. The TIDE algorithm was analyzed to predict the immunotherapy responders in BLCA patients. The results showed that the percentage of immunotherapy responders was more efficient in the high-BPRG score group (20%) compared with the low score group (11%) (chi-squared test, $P = 0.004$; Figures 9(b) and 9(c)), which was consistent with the results of immune checkpoint-related gene expression. Since BLCA patients are sensitive to chemotherapy, the IC50 value was measured for each BLCA patient according to the prophetic algorithm. The IC50 of gemcitabine ($P = 0.0008$), docetaxel ($P = 6.1e-05$), and cisplatin ($P = 0.0004$) decreased significantly (Figures 9(d–f)). These findings indicated that BLCA patients in the low score group were more sensitive to chemotherapy.

4. Discussion

The development and progression of tumors are extremely complex and can be attributed to several genetic abnormalities rather than just one or a few genes. Therefore, gene set-
Enriched in high BPRG score group

(a)

Figure 7: Continued.
Altered in 231 (94.29%) of 245 samples

Altered in 141 (94%) of 150 samples

Figure 7: Continued.
based analysis is attracting increasing attention. Lymphatic metastasis or muscle invasion of tumor cells can affect the treatment outcomes in BLCA and lead to a poor prognosis [21]. In recent years, with further understanding of the relationship of the immune system with the development and occurrence of tumors, immunotherapy is a sought-after treatment strategy for BLCA [22]. However, because BLCA is a highly heterogeneous tumor, these patients usually show...
**Figure 8: Continued.**

(a) 
Expression values for various cell types across different BPRGscore levels.

(b) 
Distribution of TME score, Immune score, and ESTIMATE score across BPRGscore levels.

(c) 
Further analysis of cell type expression and BPRGscore relationship.
different prognoses. Therefore, several studies have focused on identifying new biomarkers to guide individual immunotherapy and predict prognosis.

As machine learning technology proved good accuracy in classification and identification [23, 24]. Herein, after GSVA for the BP gene sets, we classified the patients in the TCGA-BLCA cohort according to the levels of gene set expression. As shown in Figure 2, the classification based on the result of GSVA was relatively accurate, and the average contour area reached 0.97. In our study, three clusters showed significant differences in prognosis, immune infiltration levels, and functions. Due to this method being based on gene sets expression, although the accuracy was high, the applicability may not be substantial. Thus, after extracting the target genes from different gene sets, LASSO-Cox regression was performed twice to analyze the dimensionally reduced target genes, and finally, 5 key genes were obtained (PTN, SPP1, KREMEN1, EGFR, and LBX2). According to the calculated cut-off value, the patients were classified into low- and high-BPRG score groups. In the test group, survival time in the low score group was longer. Similar results were obtained in the validation group. The higher the score, the worse the clinical features and the shorter the expected survival time. Thus, we constructed the nomogram and plotted the correction curve, which could predict the patients’ OS more accurately and was more conducive to clinical application.

Some previous studies have reported these 5 key genes, while others have also shown their close relationship with cancer. Pleiotroin (PTN) is a protein-coding gene. The protein encoded by this gene is a secretory heparin-binding growth factor. PTN promotes nerve invasion in pancreatic cancer [25]. Overexpression of SPP1 is related to a poor prognosis of melanoma. Silencing SPP1 inhibits melanoma cells’ migration, invasion, and proliferation [26]. Somatic mutations in KREMEN1 found in human cancer can affect proapoptotic activity and support tumor suppressive functions [27]. EGFR is often expressed at high levels in different cancers, and its expression positively correlates with tumor progression and poor prognosis [28]. EGFR is also activated in basal cells, including those in BLCA; its overexpression is also associated with poor prognosis [29]. LBX2 serves carcinogenic functions in LUAD and is involved in tumor migration, proliferation, and invasion throughout EMT processes [30].

Interestingly, after GSEA of key genes in the high score group, the pathways related to cancer promotion and immune activation were enriched in the high score group, including activation of immune response, cell adhesion mediated by integrin, chemokine signaling pathway, and cytokine receptor interaction. Integrins, chemokines, and cytokines are also closely related to cancer. According to Cooper and Giancotti, integrin signaling can drive several tumor cell functions, including initiation, metastasis
Figure 9: Evaluation of the influence of BPRG score on immunotherapeutic and chemotherapeutics efficacies for BLCA patients. (a) The expression of immune checkpoints in different BPRG score groups. (b) Sankey chart displays the distribution of responder and BPRG score. (c) Comparison of immunotherapy responder among two BPRG score groups. (d-f) Drug sensitivity analysis for gemcitabine, docetaxel, and cisplatin in BLCA patients belonging to high- and low-BPRG score groups. *p < 0.05; **p < 0.01; ***p < 0.001.
reactivation, and epithelial plasticity, along with resistance to oncogene and targeted immune therapies [31]. Bagati et al. show that integrin αvβ6 can promote immune escape in breast cancer [32]. In BLCA, Liu et al. show that integrin β can promote drug resistance and growth in these cells through the Y-box binding protein 1-dependent signaling pathway [33]. Chemokines can participate in the generation and recruitment of immune cells, which help in forming a microenvironment that promotes tumor progression [34]. Considering that the key genes were closely related to immune functions, immune responses are important for the occurrence and treatment of BLCA [35]. We then compared the differences in TICs, immune-related functions, interstitial cell infiltration levels, immune checkpoints, and immunotherapy between the groups. In the high score group, the abundances of immune cells including CD8 cells and macrophages were high, and the related immune functions were enriched. Moreover, the interstitial score of the high score group was also higher. Stromal and immune cells are the main components of TME. High immune and stromal scores are associated with poor prognosis [36]. This suggested that the high evaluation group may show a poor prognosis.

The patients in the high score group also showed increased expression of immune checkpoint-related genes, including PD1, PDL1, and CTLA4 which suggested that the high score group may benefit more from immunotherapy [37]. The immune checkpoints serve regulatory functions in immune responses by inhibiting the initiation of immune cell responses and immune monitoring [38, 39]. TIDE is a computational framework that simulates the two main mechanisms of tumor immune escape. It predicts the results of immunotherapy [18]. Subsequently, we verified our hypothesis using the TIDE algorithm, and the results were consistent with previous findings for immune cell infiltration and expression of immune checkpoint-related genes. Unexpectedly, the high rating group showed a lower survival rate. We further evaluated the differences in somatic mutation data between the two groups. The three genes with the highest mutation rates were TTN, KMT2D, and TP53 in BLCA, similar to Song et al. [40]. Our calculations showed that the high TMB group had a longer survival time, which is also in line with the conclusion of Samstein et al. [41]. Moreover, high TMB cannot predict the immune checkpoint blockade responses in all cancer types [42]. A high level of immune checkpoint gene expression may stimulate the microenvironment of immunosuppression and lead to the immune escape of tumor cells [43]. Therefore, tumor immunity is a complex and detailed mechanism, and one-sided predictions cannot be made using unilateral factors. Therefore, we speculated that the high score group did not show a good prognosis due to the activated immune pathway, and high expression of immune cells were eliminated by the presence of other cells in the immune microenvironment, such as stromal cells. Simultaneously, high expression of immune checkpoints also exerts synergistic regulatory effects. Next, the TCIA database was used for analyzing the chemosensitivity between the two groups. Overall, the chemosensitivity in the low score group was significantly higher than that in the high score group, especially in treatment using gemcitabine, docetaxel, and cisplatin. This finding suggested that chemotherapeutic drugs may be more effective in the low score group, and thus, the score is a potential predictor of chemotherapeutic benefits in BLCA patients.

However, there are some limitations to this study. First, this was a retrospective study based on publicly available databases. There may be some biases in patient selection and relevant clinical information. Further research is needed to verify our conclusions. Second, the established model did not have a good predictive value for some immunotherapy patients, which may be related to the highly heterogeneous feature of BLCA. Therefore, more extensive research is needed, including single-cell sequencing analysis to more accurately explain the specific changes in the TME. Finally, these findings need to be verified through in vivo and in vitro experiments in the future.

5. Conclusions

In conclusion, we systematically analyzed the relevant information on key genes screened based on BPs and found significant differences in clinical characteristics, immune microenvironment, immunotherapy, and chemotherapy between the low and high score patients. The findings demonstrated the clinical significance of comprehensively evaluating the BPRG score for each BLCA patient, which is expected to facilitate the formulation of individual therapy strategies by oncologists.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Huidong Zhou organized and wrote the manuscript. Huidong Zhou and Guo Huang carried out the data analysis. Janglei Zhang and Tao Chen participated in the study design and data collection. Jun Ouyang and Huidong Zhou revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This research was supported by the Suzhou Science Research Foundation (201900180034).

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," CA: a Cancer Journal for Clinicians, vol. 70, no. 1, pp. 7–30, 2020.
[2] H. von der Maase, L. Sengelov, J. T. Roberts et al., "Long-Term survival results of a randomized trial comparing gemcitabine
plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer," Journal of Clinical Oncology, vol. 23, no. 21, pp. 4602–4608, 2005.

[3] J. E. Rosenberg, J. Hoffman-Censits, T. Powles et al., “Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial," Lancet, vol. 387, no. 10031, pp. 1909–1920, 2016.

[4] C. Kandoth, M. D. McLellan, F. Vandin et al., “Mutational landscape and significance across 12 major cancer types," Nature, vol. 502, no. 7471, pp. 333–339, 2013.

[5] M. Nishino, N. H. Ramaiya, H. Hatabu, and F. S. Hodi, “Inferring immune-checkpoint blockade: response evaluation and biomarker development," Nature Reviews. Clinical Oncology, vol. 14, no. 11, pp. 655–668, 2017.

[6] L. Tran, J.-F. Xiao, N. Agarwal, J. E. Dux, and D. Theodorescu, "Advances in bladder cancer biology and therapy," Nature Reviews. Cancer, vol. 21, no. 2, pp. 104–121, 2021.

[7] T. Song, S. Cao, S. Tao, S. Liang, W. du, and Y. Liang, "A novel unsupervised algorithm for biological process-based analysis on cancer," Scientific Reports, vol. 7, no. 1, pp. 4671, 2017.

[8] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," Cell, vol. 144, no. 5, pp. 646–674, 2011.

[9] C. Patrano, S. Sterpone, A. Testa, and R. Cozzi, "Polymorphisms in base excision repair genes: breast cancer risk and individual radiosensitivity," World J Clin Oncol., vol. 5, no. 5, pp. 874–882, 2014.

[10] H. Wang, Y. Mei, C. Luo et al., “Single-cell analyses reveal mechanisms of cancer stem cell maintenance and epithelial-mesenchymal transition in recurrent bladder cancer," Clinical Cancer Research, vol. 27, no. 22, pp. 6265–6278, 2021.

[11] W.-J. Kim, E.-J. Kim, S.-K. Kim et al., “Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer," Molecular Cancer, vol. 9, no. 1, p. 3, 2010.

[12] S. Hänzelmann, R. Castelo, and J. Guinney, “GSVA: gene set variation analysis for microarray and RNA-seq data," BMC Bioinformatics, vol. 14, no. 1, p. 7, 2013.

[13] M. D. Wilkerson and D. N. Hayes, “ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," Bioinformatics, vol. 26, no. 12, pp. 1572-1573, 2010.

[14] M. E. Ritchie, B. Shipman, D. Wu et al., “Lima powers differential expression analyses for RNA-sequencing and microarray studies," Nucleic Acids Research, vol. 43, no. 7, p. e47, 2015.

[15] I. S. Lossos, D. K. Czerwinski, A. A. Alizadeh et al., "Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes," The New England Journal of Medicine, vol. 350, no. 18, pp. 1828–1837, 2004.

[16] G. Bindea, B. Mlecnik, M. Tosolini et al., “Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer," Immunity, vol. 39, no. 4, pp. 782–795, 2013.

[17] K. Yoshihara, M. Shahmoradgoli, E. Martinez et al., “Inferring tumour purity and stromal and immune cell admixture from expression data," Nature Communications, vol. 4, no. 1, p. 2612, 2013.

[18] P. Jiang, S. Gu, D. Pan et al., “Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," Nature Medicine, vol. 24, no. 10, pp. 1550–1558, 2018.

[19] A. Mayakonda, D.-C. Lin, Y. Assenov, C. Plass, and H. P. Koeffler, “Maftools: efficient and comprehensive analysis of somatic variants in cancer," Genome Research, vol. 28, no. 11, pp. 1747–1756, 2018.

[20] P. Geelhofer, N. Cox, and R. S. Huang, "pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels," Plos One, vol. 9, no. 9, article e107468, 2014.

[21] B. L. Jacobs, C. T. Lee, and J. E. Montie, "Bladder cancer in 2010: how far have we come?," CA: a Cancer Journal for Clinicians, vol. 60, no. 4, pp. 244–272, 2010.

[22] E. Feld, J. Harton, N. J. Meropol et al., “Effectiveness of first-line immune checkpoint blockade versus carboplatin-based chemotherapy for metastatic urothelial cancer," European Urology, vol. 76, no. 4, pp. 524–532, 2019.

[23] B. Yang, W. Bao, and J. Wang, "Active disease-related compound identification based on capsule network," Briefings in Bioinformatics, vol. 23, no. 1, article bbab462, 2022.

[24] W. Bao, Q. Cui, B. Chen, and B. Yang, "Phage_UniR_LGBM: phage virion proteins classification with UniRep features and LightGBM model," Computational and Mathematical Methods in Medicine, vol. 2022, Article ID 9470683, 8 pages, 2022.

[25] J. Yao, X.-F. Hu, X.-S. Feng, and S. G. Gao, "Pleiotrosphin promotes perineural invasion in pancreatic cancer," World Journal of Gastroenterology, vol. 19, no. 39, pp. 6555–6558, 2013.

[26] G. Deng, F. Zeng, J. Su et al., "BET inhibitor suppresses melanoma progression via the noncanonical NF-kB/SPPI pathway," Theranostics, vol. 10, no. 25, pp. 11428–11443, 2020.

[27] F. Caurer, I. Sumia, and A. Pierani, "Kvremen1 and Dickkopf1 control cell survival in a Wnt-independent manner," Cell Death and Differentiation, vol. 23, no. 2, pp. 323–332, 2016.

[28] S. Talukdar, L. Emdad, S. K. Das et al., "EGFR: an essential receptor tyrosine kinase-regulator of cancer stem cells," Advances in Cancer Research, vol. 147, pp. 161–188, 2020.

[29] R. I. Nicholson, J. M. Gee, and M. E. Harper, "EGFR and cancer prognosis," European Journal of Cancer, vol. 37, Supplement 4, pp. 9–15, 2001.

[30] J. Hu, Y. Bai, Q. Zhang, M. Li, R. Yin, and L. Xu, "Identification of LBX2 as a novel causal gene of lung adenocarcinoma," Thorac Cancer., vol. 11, no. 8, pp. 2137–2145, 2020.

[31] J. Cooper and F. G. Giancotti, "Integrin signaling in cancer: mechanotransduction, stemness, epithelial plasticity, and therapeutic resistance," Cancer Cell, vol. 35, no. 3, pp. 347–367, 2019.

[32] A. Bagati, S. Kumar, P. Jiang et al., "Integrin αvβ6-TGFβ-SOX4 pathway drives immune evasion in triple-negative breast cancer," Cancer Cell, vol. 39, no. 1, pp. 54–67.e9, 2021.

[33] S. Liu, L. Chen, H. Zhao, Q. Li, R. Hu, and H. Wang, "Integrin β8 facilitates tumor growth and drug resistance through a Y-box binding protein 1-dependent signaling pathway in bladder cancer," Cancer Science, vol. 111, no. 7, pp. 2423–2430, 2020.

[34] A. J. Ozga, M. T. Chow, and A. D. Luster, "Chemokines and the immune response to cancer," Immunity, vol. 54, no. 5, pp. 859–874, 2021.

[35] A. K. Schneider, M. F. Chevalier, and L. Derré, "The multifaceted immune regulation of bladder cancer," Nature Reviews. Urology, vol. 16, no. 10, pp. 613–630, 2019.

[36] B. A. McGregor, A.-K. A. Lalani, W. Xie et al., "Activity of cabozantinib after immune checkpoint blockade in metastatic..."
clear-cell renal cell carcinoma,” *European Journal of Cancer*, vol. 135, pp. 203–210, 2020.

[37] R. Cao, B. Ma, G. Wang, Y. Xiong, Y. Tian, and L. Yuan, “Identification of autophagy-related genes signature predicts chemotherapeutic and immunotherapeutic efficiency in bladder cancer (BLCA),” *Journal of Cellular and Molecular Medicine*, vol. 25, no. 12, pp. 5417–5433, 2021.

[38] W. C. M. Dempke, K. Fenchel, P. Uciechowski, and S. P. Dale, “Second- and third-generation drugs for immuno-oncology treatment--the more the better?,” *European Journal of Cancer*, vol. 74, pp. 55–72, 2017.

[39] I. Le Mercier, J. L. Lines, and R. J. Noelle, “Beyond CTLA-4 and PD-1, the generation Z of negative checkpoint regulators,” *Frontiers in Immunology*, vol. 6, p. 418, 2015.

[40] X. Song, S. Xin, Y. Zhang et al., “Identification and quantification of iron metabolism landscape on therapy and prognosis in bladder cancer,” *Frontiers in Cell and Development Biology*, vol. 10, article 810272, 2022.

[41] R. M. Samstein, C.-H. Lee, A. N. Shoushtari et al., “Tumor mutational load predicts survival after immunotherapy across multiple cancer types,” *Nature Genetics*, vol. 51, no. 2, pp. 202–206, 2019.

[42] D. J. McGrail, P. G. Pilić, N. U. Rashid et al., “High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types,” *Annals of Oncology*, vol. 32, no. 5, pp. 661–672, 2021.

[43] G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, and R. D. Schreiber, “Cancer immunoediting: from immunosurveillance to tumor escape,” *Nature Immunology*, vol. 3, no. 11, pp. 991–998, 2002.