


Multi-Omics analysis identifies a lncRNA-related prognostic signature to predict bladder cancer recurrence

Zhipeng Xu^{a,*}, Hui Chen^{b,*}, Jin Sun^{c,*}, Weipu Mao^{a,*}, Shuqiu Chen^a, and Ming Chen ^{a,d}

^aDepartment of Urology, Affiliated Zhongda Hospital of Southeast University, Nanjing, China; ^bDepartment of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China; ^cDepartment of Urology, Xuyi People's Hospital, Huaian, China; ^dDepartment of Urology, Zhongda Hospital Lishui Branch, Nanjing, China

ABSTRACT

Bladder cancer (BLCA) is one of the most common cancers worldwide with high recurrence rate. Hence, we intended to establish a recurrence-related long non-coding RNA (lncRNA) model of BLCA as a potential biomarker based on multi-omics analysis. Multi-omics data including copy number variation (CNV) data, mutation annotation files, RNA expression profiles and clinical data of The Cancer Genome Atlas (TCGA) BLCA cohort (303 cases) and GSE31684 (93 cases) were downloaded from public database. With multi-omics analysis, twenty lncRNAs were identified as the candidates related with BLCA recurrence, CNVs and mutations in training set. Ten-lncRNA signature were established using least absolute shrinkage and selection operation (LASSO) and Cox regression. Then, various survival analysis was used to assess the power of lncRNA model in predicting BLCA recurrence. The results showed that the recurrence-free survival time of high-risk group was significantly shorter than that of low-risk group in training and testing sets, and the predictive value of ten-lncRNA signature was robust and independent of other clinical variables. Gene Set Enrichment Analysis (GSEA) showed this signature were associated with immune disorders, indicating this signature may be involved in tumor immunology. After compared with the other reported lncRNA signatures, ten-lncRNA signature was validated as a superior prognostic model in predicting the recurrence of BLCA. The effectiveness of the model was also evaluated in bladder cancer samples via qRT-PCR. Thus, the novel ten-lncRNA signature, constructed based on multi-omics data, had robust prognostic power in predicting the recurrence of BLCA and potential clinical implications as biomarkers.

ARTICLE HISTORY

Received 20 October 2021
Accepted 22 October 2021

KEYWORDS

Multi-omics data; copy number variation; mutation; bladder cancer recurrence; long non-coding rnas¹

1. Introduction


Bladder cancer (BLCA) ranks tenth among cancers in incidence, with an estimated 549,000 new cases and 200,000 deaths worldwide in 2018 [1]. Transitional cell carcinomas account for over 90% in BLCA [2]. Approximately 70% of patients with transitional cell carcinomas were diagnosed as non-muscle-invasive bladder cancer (NMIBC), the others were diagnosed as muscle-invasive bladder cancer (MIBC) [2,3]. Despite the development in multimodal treatment, both NMIBC and MIBC have the high rates of recurrence [3–6]. NMIBC presents high risk of recurrence ranging from 31% to 78% at five years, but are generally not life threatening [2,7]. For MIBC, recurrence-free survival (RFS) and 5-year overall survival (OS) rates

after radical cystectomy are 68.0% and 57.7%, respectively [5]. The high risk of recurrence and progression contributes to high mortality in MIBC [4,8,9]. Thus, it is clinically significant to screen for key biomarkers to assess the possibility of recurrence in patients with BLCA.

Long non-coding RNAs (lncRNAs), a major class of non-coding RNAs, are RNA transcripts with more than 200 base pairs [10]. Recent studies have also validated the roles of lncRNA in BLCA. Zhan et al. have reported that significantly upregulated expression of lncRNA *SOX2OT* was closely related with stemness phenotype in BLCA [11]. Wang et al. have verified that upregulation of *BLACAT2* made contributions to the BLCA lymphatic metastasis by upregulating the expression of VEGF-C in

CONTACT Ming Chen  mingchenseu@126.com; Shuqiu Chen  chenshuqiuseu@163.com  Department of Urology, Affiliated Zhongda Hospital of Southeast University, Nanjing, China

*These authors contributed equally to this work.

 Supplemental data for this article can be accessed [here](#)

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

epigenetic mechanisms [12]. *LNMT1* can promote lymphatic metastasis through epigenetically activating CCL2-dependent macrophage recruitment in BLCA [13]. These findings demonstrated that aberrant expression of lncRNAs was involved in BLCA initiation and progression, which represented their potential as diagnostic and prognostic biomarkers. Thus, many researchers have focused on predicting the prognosis of patients with BLCA basing on lncRNA profiles. For instance, Joep J et al. revealed that lncRNAs profiling could provide additional information for BLCA subtyping, which contributed to precision patient management [14]. Du et al. [15] constructed an epithelial-mesenchymal transition (EMT)-associated lncRNA signature to predict the prognosis of BLCA patients. Besides, Mao et al. [16] developed a ten-lncRNA signature to predict the outcome and immune status of BLCA. Additionally, Gao et al. [17] identified a six-lncRNA signature as a robust prognostic marker in BLCA through COX regression analysis. However, limited data on tumor etiology was taken into consideration when only transcriptome data was comprehensively analyzed with clinical features. Recently, some studies have paid attention to distinguish cancer patients with different clinical outcomes via multi-omics data [18–21]. Manikandan et al. [22] identified that amplified P4HA1 gene was related to hypoxia in breast cancer via multi-omics analysis. Zhao et al. [23] provided a novel insight into molecular subtypes for lung adenocarcinoma basing on multi-omics data (genomics, epigenomics, and transcriptomics). Chaudhary et al. [24] evaluated prognosis features of hepatocellular carcinoma patients via multi-omics data, which showed robustness in several external cohorts. However, multi-omics analysis has not been performed in constructing the lncRNA-associated prognostic model in patients with BLCA.

The previous studies have indicated the critical roles of lncRNAs in BLCA, and prompted the potential value of multi-omics analysis on constructing prognostic model. Hence, we assumed that multi-omics analysis may contribute to construct a robust lncRNA predictive model for BLCA recurrence, and intended to construct a novel lncRNA signature on the basis of multi-omics

data, including transcriptome data, clinical data, copy number variation (CNV) data and mutation annotation data in the cohort of BLCA patients from The Cancer Genome Atlas (TCGA) database and GSE31684 dataset. We successfully constructed a ten-lncRNA signature for RFS of BLCA patients by a contemporary clinical-practical statistical method, least absolute shrinkage and selection operation (LASSO), and COX regression model. Our results validated that the ten-lncRNA signature could act as an independent prognostic predictor of BLCA recurrence. Besides, the novel lncRNA signature may be associated with tumor immunology. We present the following article in accordance with the TRIPOD reporting checklist.

2. Materials and methods

2.1. Pre-processing of lncRNA-associated information and clinical data from public databases

The RNA expression profiles, clinical data and genomic copy number variation data in the TCGA database were obtained via the UCSC cancer browser (<https://xenabrowser.net/datapages/>). And the mutation annotation file (MAF) was extracted via GDC client. After downloading the ‘fragments per kilobase of transcript per million fragments sequenced’ (FPKM) data of RNA-Seq from TCGA database, we obtained the lncRNA expression profiles by cross-referring to ensemble ID of lncRNAs from GENCODE project [25]. Then, ‘transcripts per million’ (TPM) values were calculated according to the FPKM values. Finally, the TPM values were normalized by Z-score.

In addition, we downloaded the clinical data and transcriptome profiles of Series GSE31684 from GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) [26]. The microarray raw data were downloaded from GEO database. We obtained the lncRNA probes from the manufacturer’s website (<http://www.affymetrix.com>), and mapped the probe sequences to the human genome (hg19) without mismatch. Next, 5076 probes fell into lncRNAs through re-annotation. Eventually, the

expression data was normalized via the quantile-normalization approach [27,28].

Herein, we concentrated on BLCA recurrence. Hence, we excluded the BLCA samples whose RFS time was unknown, not described or less than 30 days. We randomly selected three-quarters of samples from the BLCA cohort in TCGA database as the training dataset. The main reason for selecting three-quarters of samples is to include more samples in the training dataset in order to obtain more stable results. To avoid the selection bias, we performed the re-randomizations for 1000 times, and calculated the AUC distribution of each dataset, mainly ranging from 0.65 to 0.76 (Supplemental figure 1A). Eventually, 303 cases from TCGA database and 93 cases from Series GSE31684 in GEO database were included in this study. We randomly selected 227 cases from all 303 cases in the TCGA database as the training datasets ($n = 227$). The first testing datasets consisted of all 303 patients in the TCGA database. The second testing datasets were the 93 samples from Series GSE31684 in GEO database. We constructed the lncRNA-related prognosis model on the basis of the training dataset. The prognosis model was validated in the testing datasets. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

2.2. Gene CNV analysis

As for the copy number variation data obtained from TCGA database, the genomic regions with somatic copy number alterations (SCNA) were determined by GISTIC 2.0 [29]. The GISTIC peaks of amplification or deletion with p -value < 0.05 were regarded as significance. Next, in order to identify lncRNAs with CNV in BLCA, we mapped the genomic regions to GISTIC peaks.

2.3. Identification of lncRNAs associated with significantly mutated genes (SMGs)

Based on the MAF files from TCGA database, we identified SMGs with q -value < 0.05 using MutSig 2.0 algorithms [30]. In order to identify the SMGs-related lncRNAs, we labeled the samples in the training datasets with the mutation or non-mutation of each SMG. Abnormal expression of

lncRNAs associated with each SMG was determined by rank sum test with p -value < 0.01 .

2.4. Identification of BLCA recurrence-associated lncRNAs

Univariate COX regression analysis was performed to screen out the lncRNAs associated with RFS. So as to further assess which lncRNA could act as the dependent variable factor, multivariable COX regression analysis was conducted. R software and bioconductor were utilized for analysis. P -value < 0.05 was considered as significant.

2.5. Construction of the lncRNA signature

On the basis of univariate COX regression analysis, gene CNV analysis and identification of SMGs-related lncRNAs, the lncRNAs associated with CNV, gene mutation and RFS time in BLCA were screened out as candidate genes for developing the lncRNA signature. Next, we intended to further select the candidate genes and construct a prognostic model with high accuracy by utilizing LASSO model [31]. The ‘glmnet’ package and R software were applied to perform LASSO regression algorithm.

After further selecting the candidate genes by LASSO regression algorithm, multivariable Cox regression analysis was performed on reserved candidate genes. Those lncRNAs with the lowest Akaike information criteria (AIC) value were screened out as the final candidate genes [32]. Next, the following formula was utilized to compute the risk score:

$$\text{Risk Score} = \sum_{i=1}^n (\text{Coefficient}_{\text{lncRNA}_i} \times \text{Expression}_{\text{lncRNA}_i})$$

On the basis of training dataset, we performed ROC (receiver-operating characteristic) curves analysis with the AUC (area under curve) at five years of RFS. Next, the Youden’s index was calculated as the optimal cutoff point to distinguish BLCA cases into high or low recurrence risk sets [33]. Kaplan-Meier RFS curve analysis and the log-rank test were applied in comparing RFS in two sets. The survfit and survdif function of ‘survival’ packages and R software were applied to perform Kaplan-Meier RFS curve analysis and the log-rank test.

2.6. Validation of the prognosis power of lncRNA signature

We validated the signature in the validation datasets, namely the entire TCGA datasets and GSE31684 datasets. We utilized the Youden's index to distinguish the cases in validation datasets into high and low recurrence risk sets. Next, we compare the RFS in high and low recurrence risk sets via Kaplan-Meier RFS analysis, log-rank test and further stratified analysis. In order to assess the prognostic power of lncRNA signature, we performed univariate and multivariate COX regression analysis in training dataset and testing datasets.

2.7. Functional enrichment analysis and tumor immune microenvironment characteristics

Gene Set Enrichment Analysis (GSEA) was performed based on the training dataset [34,35]. The RNA expression profiles were used as the input file and labeled with lncRNA signature-based risk score. The CIBERSORT method was performed to evaluate the relative abundance of immune cell profiling [36]. To evaluate the components in each sample, we applied the ESTIMATE algorithm to calculate the score of stromal, immune and ESTIMATE score in each sample [37]. In addition, the correlation was further analyzed with Spearman Rank correlation.

2.8. Clinical samples and cell lines

The 33 cases of BLCA specimens used in our study were collected from the Shanghai Tenth People's Hospital of Tongji University, China. The 33 patients underwent radical cystectomy from May 2014 to June 2014. The 33 cases were staged based on the 7th AJCC staging system. This study was approved by the Ethics committee of the Shanghai Tenth People's Hospital of Tongji University (SHSY-IEC-4.1/19-120/01). Prior informed consent was obtained from all of the patients. The bladder transitional cell carcinoma cell lines (UM-UC-3 and T24) were cultured in DMEM (Gibco, USA) containing 10% fetal bovine serum (FBS; Gibco, USA).

2.9. Transfection and qRT-PCR

In order to knockdown the expression of AGAP2-AS1, small interfering RNA (siRNA) was purchased from GenePharma (Shanghai, China). The siRNAs were transfected into cells using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. We isolated the total RNA from cell lines and tissues using TRIzol (Takara, Japan) based on the manufacturer's instructions. lncRNA reverse transcription were conducted with a New Poly(A) Tailing Kit (ThermoFisher Scientific) and PrimeScript RT Master Mix Kit (RR036A, TaKaRa), respectively. qRT-PCR was performed using a Universal SYBR Green Master Mix (4,913,914,001, Roche) with a 7500 Real-Time PCR System (Applied Biosystems, USA). We normalized the relative lncRNA expression levels to GAPDH, respectively. The sequences of siRNAs and primers in this study were listed in Supplemental table 6.

2.10. Cellular proliferation assay and transwell assay

For the CCK-8 assay (CCK-8, Dojindo, Kumamoto, Japan), we first seeded cells in triplicate in a 96-well plate at a density of 2000 cells per well. At the indicated time points, we added 10 μ L CCK-8 solution to each well. After a 2 h incubation, the absorbance was determined using a microplate reader. For the colony formation assay, we seeded cells in six-well plates at a density of 500 cells per well and cultured the plate for 2 weeks. Subsequently, we fixed the cells in 75% ethanol and stained them with crystal violet. Colonies were observed and counted under a light microscope. Cell migration was analyzed using Transwell chambers (Corning, USA). Cells were cultured in serum-free DMEM in upper chamber to inhibit cell proliferation.

3. Results

In this study, we performed the comprehensive analysis of transcriptome data, clinical data, CNV data and mutation annotation data in the cohort of BLCA patients from TCGA database and

GSE31684 dataset. We constructed a ten-lncRNA signature for RFS of BLCA patients via LASSO and Cox regression model. Next, our results validated that the ten-lncRNA signature could act as a robust and independent prognostic predictor of BLCA recurrence in the training and testing datasets. Further analysis indicated that the novel lncRNA signature may be associated with tumor immunology. In addition, we verified that ten-lncRNA signature constructed via multi-omics analysis had better performance in predicting BLCA recurrence than the two reported lncRNA signature models.

3.1. Screening of BLCA relapse-associated lncRNAs

We pre-processed a series of information from the public database, including extraction of the RNA expression profiles, array re-annotation, exclusion of the samples with unclear RFS state, and randomly assigning cases into the training or validation cohorts as mentioned in the Materials and Methods. Next, we overviewed the clinical data of the training dataset and testing datasets, and displayed relevant clinical information including age, relapse state, gender, TNM stage, subtype, grade and tumor stage for each dataset in supplemental table 1. Next, univariate Cox regression analysis was conducted to identify those lncRNAs related to RFS in training dataset. 38 lncRNAs were chosen as the candidates for the subsequent study, and the results of univariate Cox regression analysis were shown in table 1 and supplemental table 2.

3.2. Analysis of gene CNV in BLCA

In past decades, more and more evidences demonstrated that genomic alterations could contribute to aberrant expression of lncRNA in various cancer [38–40]. Here, we identified the genes with significant genomic amplification or deletion in BLCA. Significant amplifications were shown in Figure 1(a). We documented genes with significant amplifications in supplemental table 3, such as significant amplification of *LINC00709* on segment 10p14 (q-value = $1.23e-19$), significant amplification of *LOC101929622* on segment 8p11.23 (q-value = $2.05e-12$) and significant

amplification of *LINC01195* on segment 16p13.2 (q-value = $7.49e-07$). Eventually, 717 genes were identified as amplified genes. The significant deletions in BLCA were shown in Figure 1(b). The genes significantly deleted on each fragment were recorded in supplemental table 4, such as significant deletion of *LOC100286922* on segment 2q37.1 (q-value = $3.05e-21$), significant deletion of *LOC101929066* on segment 8p21.3 (q-value = $2.96e-32$) and significant deletion of *LINC00208* on segment 8p23.2 (q-value = $3.28e-31$). We identified 875 deleted genes in total.

3.3. Mining of lncRNAs associated with significantly mutated genes (SMGs)

We identified significant mutations according to TCGA mutation annotation data via Mutsig 2.0, and obtained 32 genes with significant mutations. Based on the TCGA training dataset, the distribution of synonymous mutations, missense mutations, frame insertions or deletions, frame movements, nonsense mutations, splice sites and other nonsynonymous mutations in the 32 genes were analyzed and shown (Figure 1(c)). The upper graph showed the amount of nonsynonymous and synonymous mutations of the 32 genes in every case. And the right histogram showed the number of clinical samples with mutations in each gene among the 32 genes. It has been reported that some of the 32 genes were closely associated with tumor initiation and progression, such as *CDKN1A*, *CDKN2A*, *ELF3*, *HRAS*, *PIK3CA*, *RB1* and so on. Then, we intended to identify the lncRNAs related to gene mutation in the 32 genes. We used the mutation state in each gene as a label and analyzed the difference between the expression of each lncRNA in the mutant and non-mutant sets via rank-sum test. lncRNAs with p-value < 0.01 were considered to be significantly associated with the mutation in some gene. As a result, we identified 2665 lncRNAs whose expression was related to gene mutation (supplemental table 5).

3.4. Establishment of the relapse-associated lncRNA signature

According to above results, we found that 20 lncRNAs were related to genomic CNVs and gene mutation among 38 lncRNAs associated

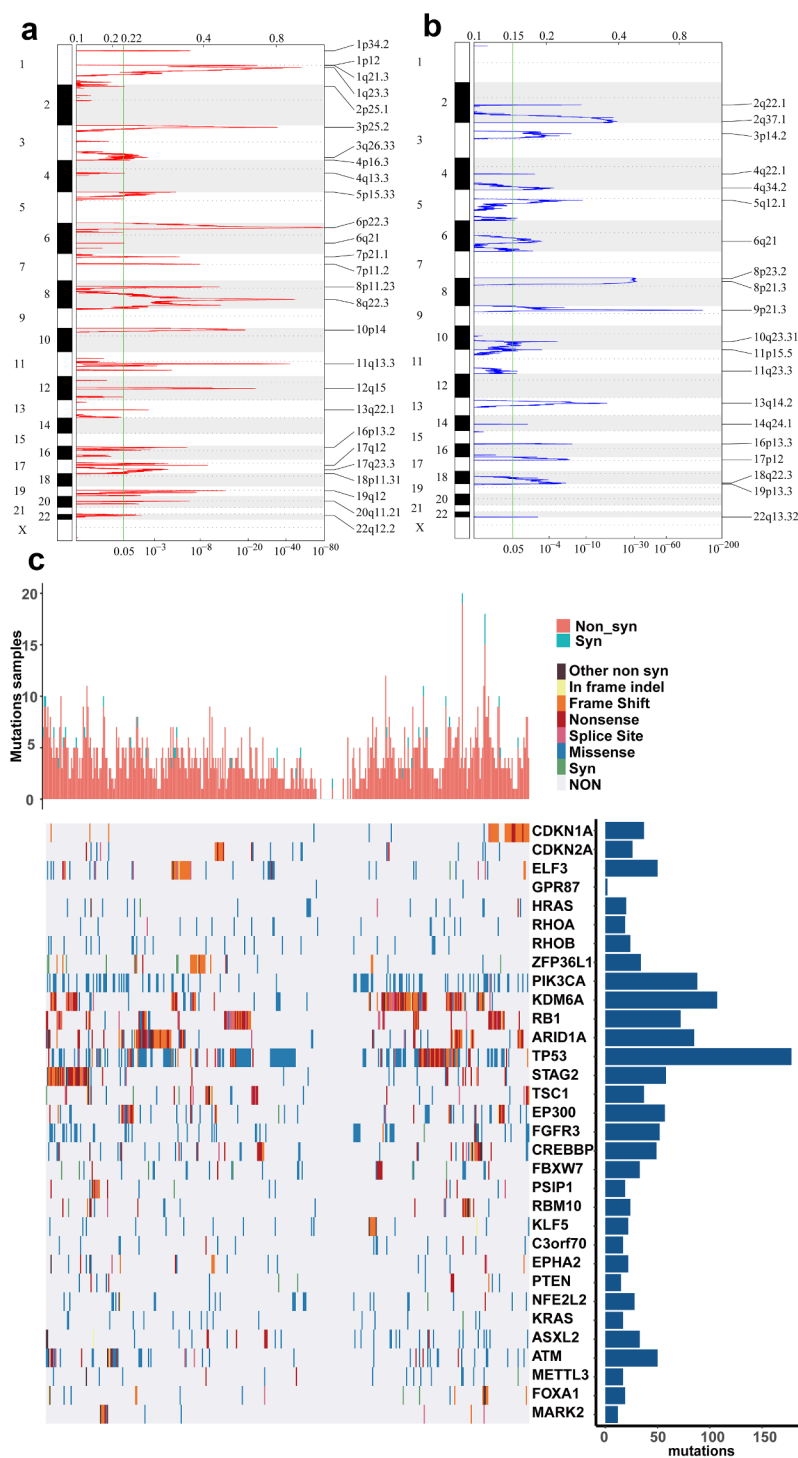


Figure 1. CNVs and mutations analysis of genome loci in BLCA. (a) The significantly amplified fragments (red) in BLCA genome were shown. (b) The significant deleted fragments (blue) in BLCA were shown. The rows are arranged according to the genome loci. (c) Distribution of mutations in 32 genes with significant mutation frequencies on basis of training dataset. Upper bar graph showed the total number of non-synonymous and synonymous mutations of 32 genes in each patient. And right histogram showed the number of clinical samples with mutations in each gene among the 32 genes.

with recurrence (Figure 2(a)). In order to further select the candidate genes and construct a prognostic model while maintaining high accuracy, LASSO model was applied in developing

a predictive signature with these screened 20 lncRNAs. LASSO method could build a penalty function to construct a refined model. LASSO evaluations of the coefficients of variables could effectively shrink

coefficients and set some coefficients to zero. Hence, LASSO regression model retained the advantages of subset shrinkage and was an approach for biased estimation in processing the multi-collinearity data, which could select variables while estimating the model parameters and handle the multi-collinearity better. LASSO model performed the analysis of the change trajectory of each variable. More coefficients of independent variable approached zero with the

lambda increasing gradually (Figure 2(b)). Then, we intended to construct the model with 3-fold cross-validation and analyzed the confidence interval under each lambda. We found that the model was optimal at $\lambda = 0.03046723$ and selected the corresponding 16 lncRNAs as the candidate genes (Figure 2(c)).

Afterward, multivariate COX regression was performed on the 16 lncRNAs, and 10 lncRNAs with AIC: 841.77 (the lowest AIC value) were

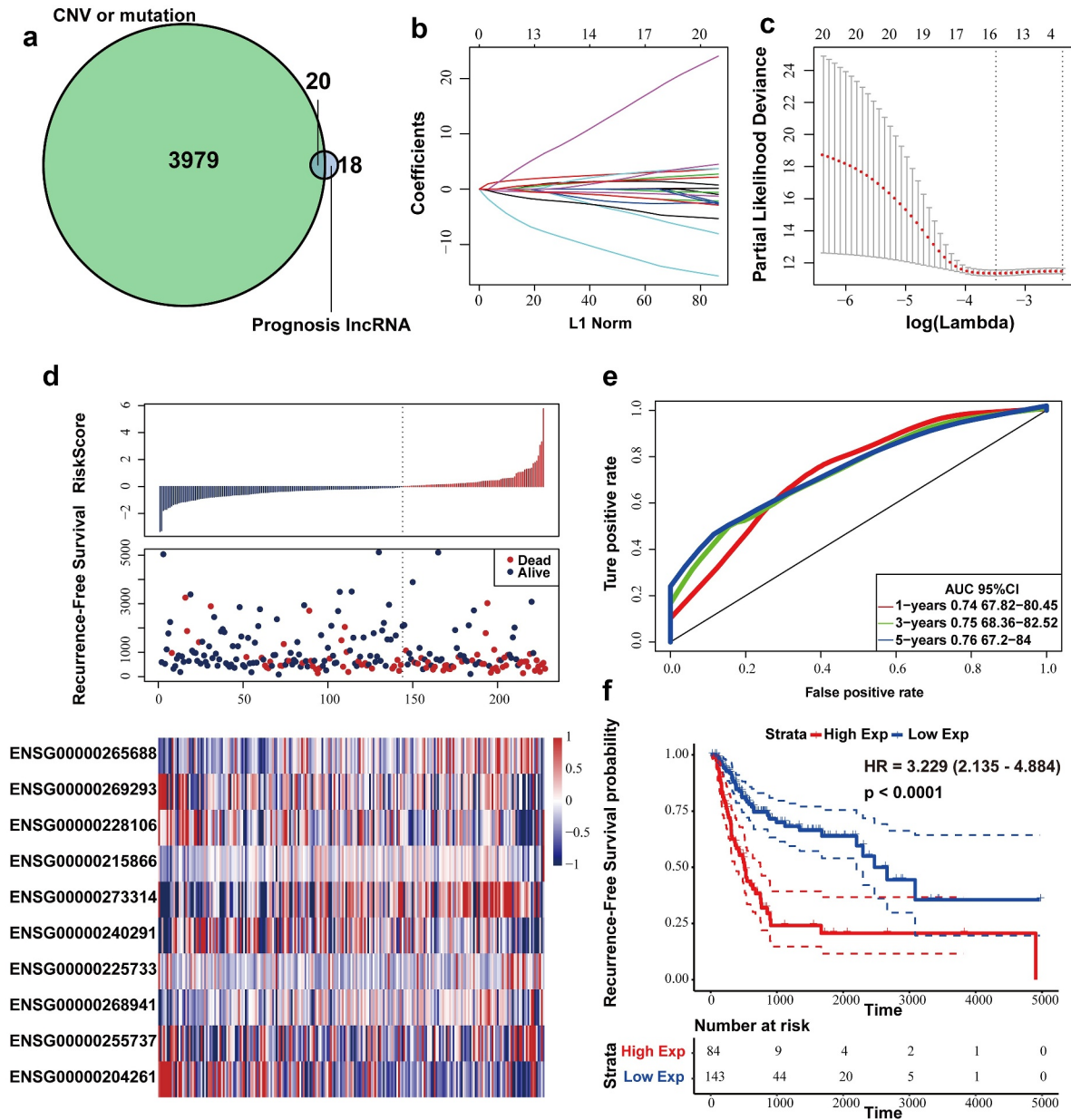


Figure 2. Construction and evaluation of the lncRNA signature based on multi-omics data. (a) Venn diagram about lncRNAs associated with genomic CNVs, gene mutation or BLCA recurrence. (b) The change trajectory of every independent variable. Horizontal axis represents the log value of independent variable lambda, and vertical axis represents the coefficient of independent variable. (c) Confidence intervals under each lambda. (d) Distribution of risk score, RFS and lncRNA expression of each case. (e) ROC curve analysis base on ten-lncRNA signature. (f) Kaplan-Meier RFS curve analysis.

eventually selected as candidate genes whose details was displayed in Table 2. We computed risk score for each patient via the following formula:

$$\begin{aligned} \text{RiskScore} = & 7.021 * \exp \text{LINC01711} + 5.153 * \\ & \exp \text{MAFG-DT} - 12.12 * \exp \text{ZSCAN16-AS1} \\ & + 22.9 * \exp \text{AC005229.4} + 0.4837 * \exp \text{FGD5-} \\ & \text{AS1} + 4.255 * \exp \text{AGAP2-AS1} + 8.5 * \\ & \exp \text{LINC01356} - 9.161 * \exp \text{AL392172.1} \\ & - 10.86 * \exp \text{AL450384.2} - 5.485 * \exp \text{PSMB8-AS1} \end{aligned}$$

In the Cox regression analysis, *LINC01711*, *MAFG-DT*, *AC005229.4*, *FGD5-AS1*, *AGAP2-AS1* and *LINC01356* had positive coefficients, indicating that upregulation of these 6 lncRNAs was related to shorter RFS time. However, *ZSCAN16-AS1*, *AL392172.1*, *AL450384.2* and *PSMB8-AS1* with negative coefficients were considered as beneficial prognostic factors in BLCA.

The RFS status, risk score and expression of 10 lncRNAs in each patient from the training datasets were shown in Figure 2(d). According to the 5-year RFS prediction AUC in the training datasets, the Youden's index (-0.4533178) was calculated as the optimal cutoff point to classify the samples into high (n = 84) or low (n = 143)-risk sets (Figure 2(e)), and Kaplan-Meier RFS analysis was performed (log-rank test p-value < 0.0001, HR = 3.229) (Figure 2(f)).

3.5. Validating prognostic power of the relapse-associated lncRNA signature

Firstly, we verified ten-lncRNA signature in the entire TCGA dataset. We utilized the Youden's index, calculated based on training dataset, to distinguish the cases in the entire TCGA datasets into high (n = 105) and low (n = 196)-risk sets. We analyzed RFS, risk score and 10 lncRNAs' expression level of each case in the entire TCGA dataset (Figure 3(a)). AUC for ten-lncRNA signature was 0.73 at the RFS in the fifth year (Figure 3(b)). As shown in Figure 3(c), we compared RFS time of high and low-risk sets (log-rank test p-value < 0.0001, HR = 2.887).

Afterward, we used GSE31684 dataset as the other testing dataset. According to the same cutoff

point (Youden's index), the cases in GSE31684 dataset were divided into high (n = 30) and low (n = 63)-risk sets. We analyzed RFS, risk score and 10 lncRNAs' expression level in the GSE31684 dataset (Figure 3(d)), and AUC for RFS in the first, third, and fifth year was 0.71, 0.65, and 0.67, respectively (Figure 3(e)). Kaplan-Meier RFS curve was utilized to compare RFS of high and low-risk sets (log-rank p value = 0.0023, HR = 2.587) (Figure 3(f)). These results suggested that the patients with higher risk score had shorter RFS time and higher recurrence rates in the testing datasets.

3.6. Evaluating whether ten-lncRNA signature had robust prognostic power in BLCA

Firstly, stratified analysis was conducted to assess the relapse-predictive power of ten-lncRNA signature at different age, tumor stages or subtypes. All 303 cases in entire TCGA dataset were divided into younger (n = 130) and elderly (n = 173) datasets at the age (65 years old). As shown in Figures 4(a,b), ten-lncRNA signature could effectively distinguish each dataset into high and low relapse-risk sets. Next, all 303 cases were re-stratified into three different datasets according to the tumor stage (stage II, n = 108; stage III, n = 102; stage IV, n = 89). Ten-lncRNA signature could classify the tumor stage III or IV dataset into high and low relapse-risk sets via medium risk score (log-rank test p-value = 0.0017, Stage III; log-rank test p-value = 0.0049, Stage IV) (Figures 4(c,d)). However, ten-lncRNA signature could not distinguish these patients in stage II into different groups with different RFS (log-rank test p-value = 0.27, Stage II) (Supplemental Figure 1B). Finally, we stratified all 303 cases into non-papillary (n = 197) and papillary (n = 102) datasets based on subtypes. Ten-lncRNA signature could classify each dataset into high and low relapse-risk sets with different RFS (log-rank test p-value = 0.00025, Figure 4(e); log-rank test p-value = 0.019, Figure 4(f)).

To validate whether ten-lncRNA signature was an independent predictive factor for BLCA recurrence, we performed Cox regression analysis in the training and testing datasets. As shown in Table 3,

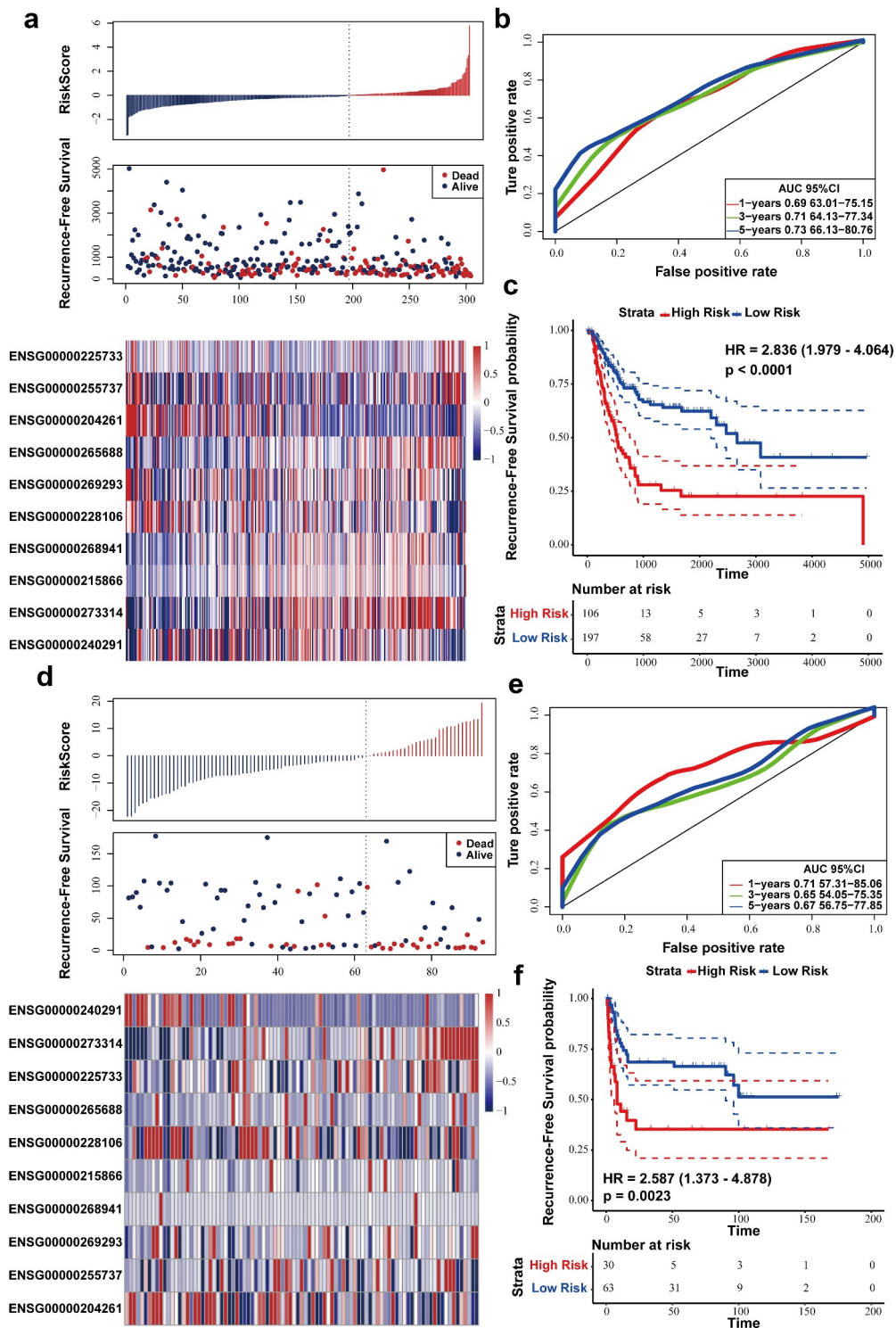


Figure 3. Evaluating the prognostic power of ten-lncRNA signature in the testing datasets. (a) Distribution of risk score, RFS and lncRNA expression of each case in the entire TCGA dataset. (b) ROC curve analysis of ten-lncRNA signature in the entire TCGA dataset. (c) Kaplan-Meier RFS curve analysis of high and low relapse-risk sets in the entire TCGA dataset. (d) Distribution of risk score, RFS and lncRNA expression of every patient in GSE31684 dataset. (e) ROC curve analysis of ten-lncRNA signature in GSE31684 dataset. (f) Kaplan-Meier RFS curve of high and low relapse-risk groups according to ten-lncRNA signature in GSE31684 dataset.

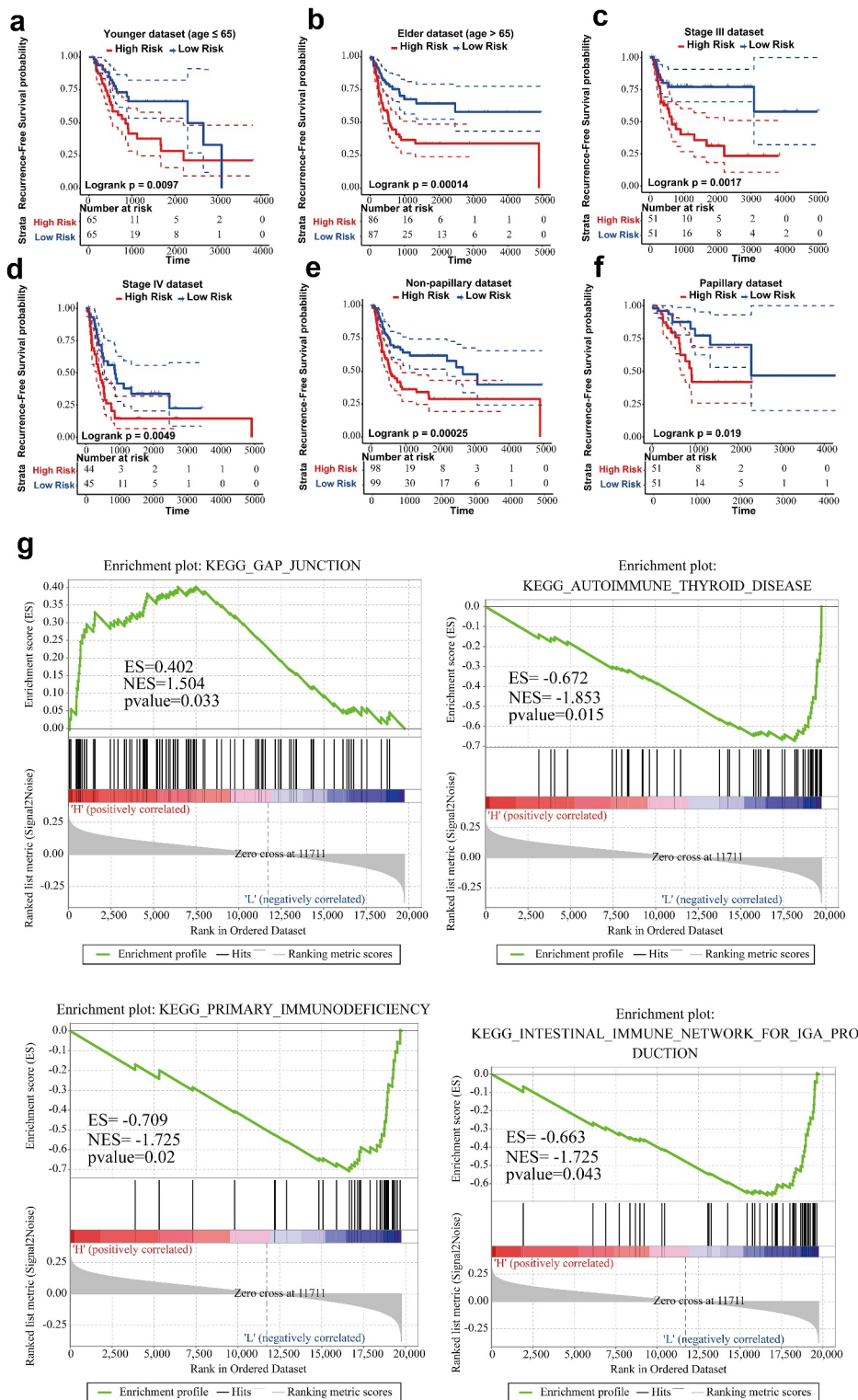


Figure 4. Stratified analysis on the basis of age, stage or subtype. (a,b) Kaplan-Meier RFS curve analysis in the younger or elderly dataset. (c,d) Kaplan-Meier RFS curve analysis in Stage III or Stage IV cohorts. (e,f) Kaplan-Meier RFS curve analysis in the non-papillary or papillary dataset. (g) Results of GSEA analysis in the TCGA training dataset.

we analyzed the association between RFS and clinical variables including ten-lncRNA signature.

In training dataset, univariate Cox regression analysis indicated that pathologic N stage,

pathologic M stage, tumor stage and risk score were related to RFS of BLCA patients. And we found that risk score and pathologic N stage were related to RFS in the multivariate Cox

regression analysis. In the entire TCGA dataset, univariate Cox regression analysis revealed that risk score, N stage, M stage and tumor stage were significantly associated with RFS of BLCA patients. And risk score and pathologic N stage were significantly associated with RFS in multivariate Cox regression analysis. In the GSE31684 dataset, we found that risk score was related to RFS in both univariate and multivariate Cox regression analysis. Taken together, ten-lncRNA signature has independent prognostic power for RFS prediction in patients with BLCA patients.

3.7. Pathway enrichment analysis and tumor immune microenvironment characteristics

In consideration of robust prognostic power of ten-lncRNA signature for BLCA recurrence, we supposed that these ten lncRNAs could take part in the progression of BLCA. We performed the Gene Set Enrichment Analysis (GSEA) among cohorts in the TCGA training datasets with the gene set named 'c2.cp.kegg.v6.0.symbols'. The RNA expression profiles were used as input files and labeled with risk score of the ten-lncRNA signature. The pathways with significantly enrichment were shown in Table 4. As shown in Figure 4(g), GSEA revealed that some significantly enriched KEGG pathways were related to the tumorigenesis, tumor progression and immune disorders.

Considering the potential relationship between ten-lncRNA signature and immune disorders revealed in GSEA, we performed the analysis on the tumor microenvironment and the infiltration of immune cells. The CIBERSORT algorithm was performed to evaluate the abundance of diverse immune cells. As shown in Figure 5(a), the infiltration of M2 macrophage was positively correlated with the risk score based on the ten-lncRNA signature ($R = 0.303$, P -value < 0.01). Next, the association between tumor microenvironment and the ten-lncRNA signature was assessed via the ESTIMATE algorithm. The results indicated that the risk score had the significant and weak correlation with the immune score, which was consistent with the worse prognosis in the patients with higher risk score (Figure 5(b)).

3.8. Comparing ten-lncRNA signature with reported lncRNA signatures in BLCA

By searching for literature about lncRNA signatures, we chose two models associated with recurrence in BLCA: four-lncRNA signature (PMID: 28,060,759) [41] and six-lncRNA signature (PMID: 31,338,862) [17]. We recalculated risk scores of each patient in training dataset according to lncRNAs in the two selected models. Next, we utilized ROC curve analysis to classify cases into high and low-risk sets by Youden's index. The results suggested that the AUC for RFS in the fifth year was 0.69 for the six-lncRNA signature (p -value = 0.00041) (Figure 6(a,b)) and the AUC for RFS in the third year was 0.60 for the four-lncRNA signature (p -value = 0.012) (Figure 6(c, d)). On the other hand, the AUC of 3-year and 5-year RFS prediction for the ten-lncRNA signature was 0.75 and 0.76, respectively (Figure 3(e)). By comparing the results of four-lncRNA signature, six-lncRNA signature and ten-lncRNA signature, we confirmed that the ten-lncRNA signature developed in this study had better performance in predicting BLCA recurrence.

3.9. Survival analysis of the ten-lncRNA signature in BLCA samples

We determined the expression of the ten lncRNAs used to construct the relapse-associated lncRNA signature in the collected 33 BLCA samples. The clinicopathological characteristics of 33 patients was displayed in supplemental table 7. Based on the formula mentioned previously, we calculated the risk score for each sample. As shown in Figure 7(a), the survival analysis indicated that patients with higher risk score intended to have shorter RFS time, with marginal significance (P value = 0.047). In addition, we performed the survival analysis on the ten lncRNAs in the 33 samples. The results indicated that higher expression of AGAP2-AS1 and LINC01711 was significantly associated with higher possibility of BLCA recurrence (AGAP2-AS1, P value = 0.017; LINC01711, P value = 0.046) (Figure 7(b) and Supplemental figure 2).

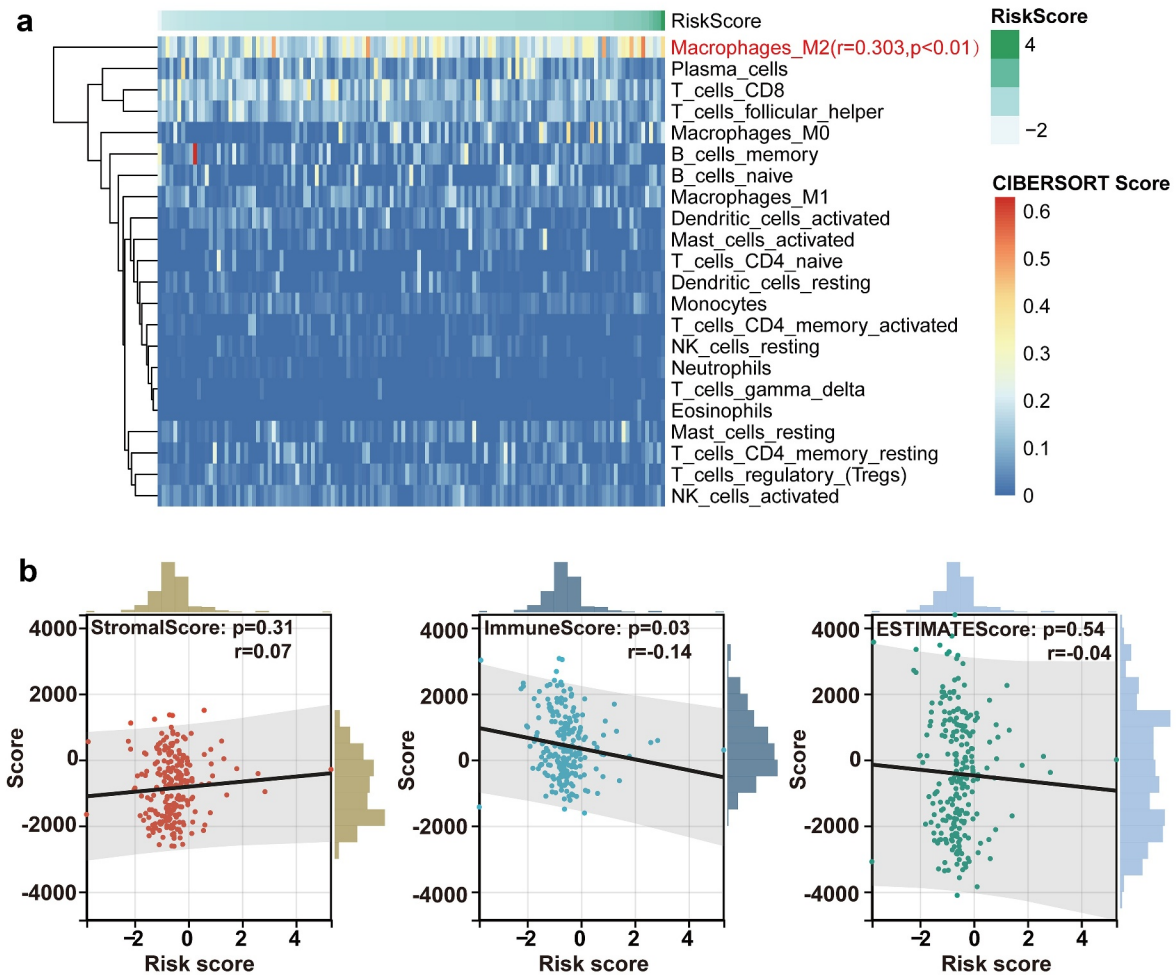


Figure 5. The analysis on the tumor immune microenvironment characteristics. (a) Evaluation of the correlation between the abundance of diverse immune cells and the ten-lncRNA signature via the CIBERSORT algorithm. (b) Evaluating the correlation between tumor microenvironment and the ten-lncRNA signature via the ESTIMATE algorithm.

3.10. AGAP2-AS1 knockdown suppresses cell proliferation and migration in BLCA cells

On the basis of survival analysis, we chose AGAP2-AS1 for further experimental study, because AGAP2-AS1 was most significantly associated with BLCA recurrence (P value = 0.017). LncRNA qRT-PCR was applied to evaluate the knockdown of AGAP2-AS1 (Figure 7(c)). The CCK-8 assay revealed that downregulation of AGAP2-AS1 expression presented a lower growth rate than the negative control in UM-UC-3 and T24 cells (Figure 7(d)). The colony formation assay further confirmed that AGAP2-AS1 knockdown could significantly inhibit BLCA cell proliferation (Figures 7(e,f)). On the other hand, the Transwell assay indicated that AGAP2-AS1

knockdown significantly inhibited cell migration compared with the control (Figures 7(g,h)). Results above suggested that AGAP2-AS1 knockdown could inhibit cell proliferation and migration in BLCA cells.

4. Discussion

Due to the development of high throughput RNA sequencing, expression pattern of lncRNAs was uncovered in diverse cancers [42,43]. Many studies have documented that expression pattern of some lncRNAs were specific in some cancer, even in different stage of some tumor [44–50]. In the recent years, a great number of studies have highlighted that dysregulation of lncRNAs was

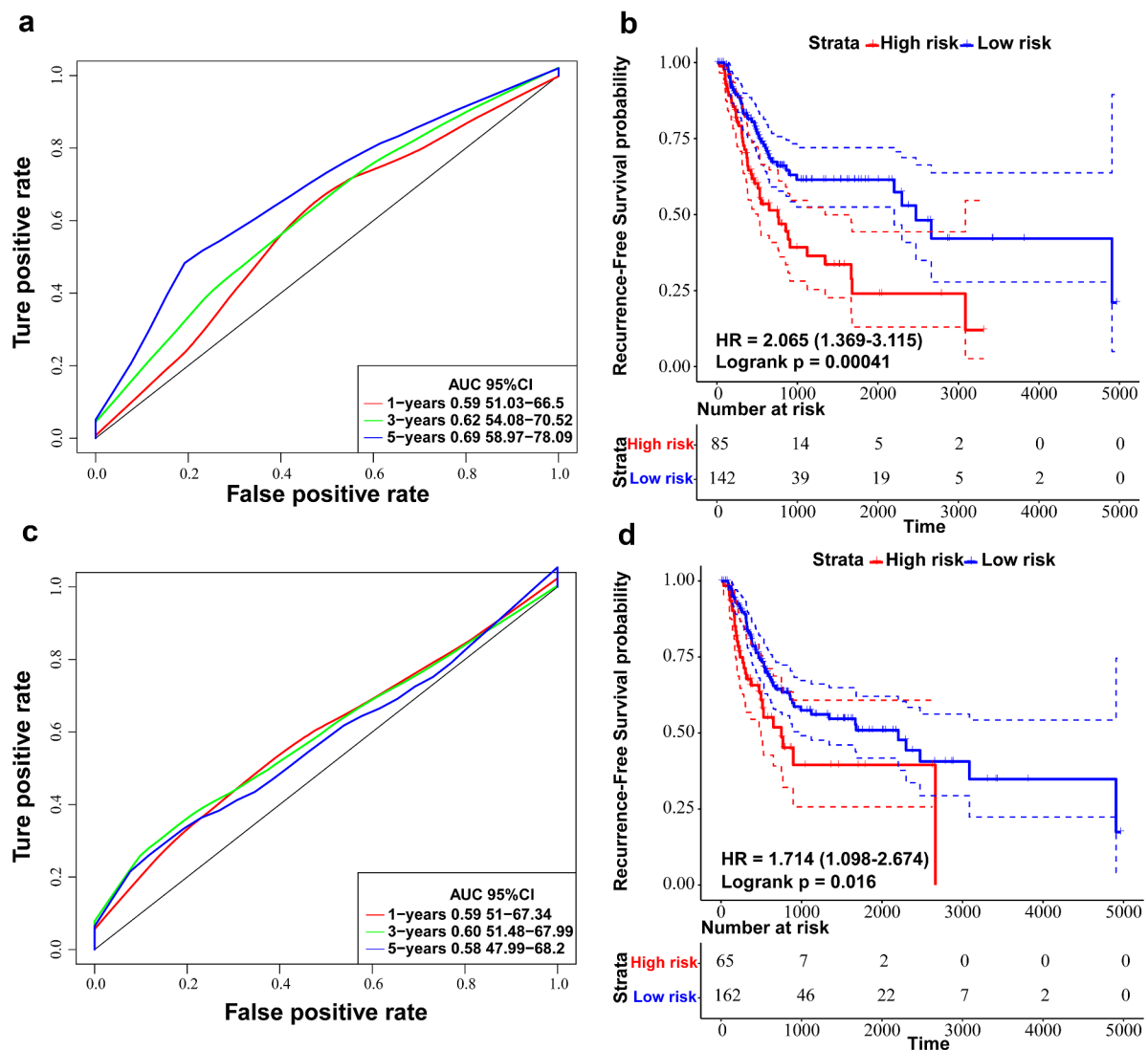


Figure 6. Comparing ten-lncRNA signature with two reported lncRNA signatures in BLCA. (a) ROC curve analysis and Kaplan-Meier RFS curve analysis on the reported 6-lncRNA signature in the TCGA training dataset. (b) ROC curve analysis and Kaplan-Meier RFS curve analysis on the reported 4-lncRNA signature in the TCGA training dataset.

involved in cancer progression [51–54]. These evidences indicated that lncRNAs have the potential to act as biomarkers for prognostic prediction in human cancers. Hence, it is necessary to identify aberrant expression pattern of lncRNAs and reveal their possible roles in BLCA development and recurrence.

Herein, we performed multi-omics analysis of transcriptome, genomic CNV, mutation annotation and clinical data of BLCA in TCGA database, in order to find the lncRNAs whose aberrant expression was associated with BLCA recurrence. Our study uncovered that 38 lncRNAs were significantly related to BLCA

recurrence through univariate Cox regression analysis. Then, CNV analysis revealed that 1592 genes had significant amplification or deletion in their genome loci and some genomic alterations contributed to the dysregulation of lncRNAs expression in BLCA. In addition, gene mutations analysis showed that there were a total of 32 genes with significant mutation, including some genes closely related to tumor initiation and progression. Further analysis indicated that expression pattern of 2665 lncRNAs was associated with these genes' mutations. On the basis of our analysis mentioned above, we found that 20 lncRNAs were associated with gene mutation and CNV

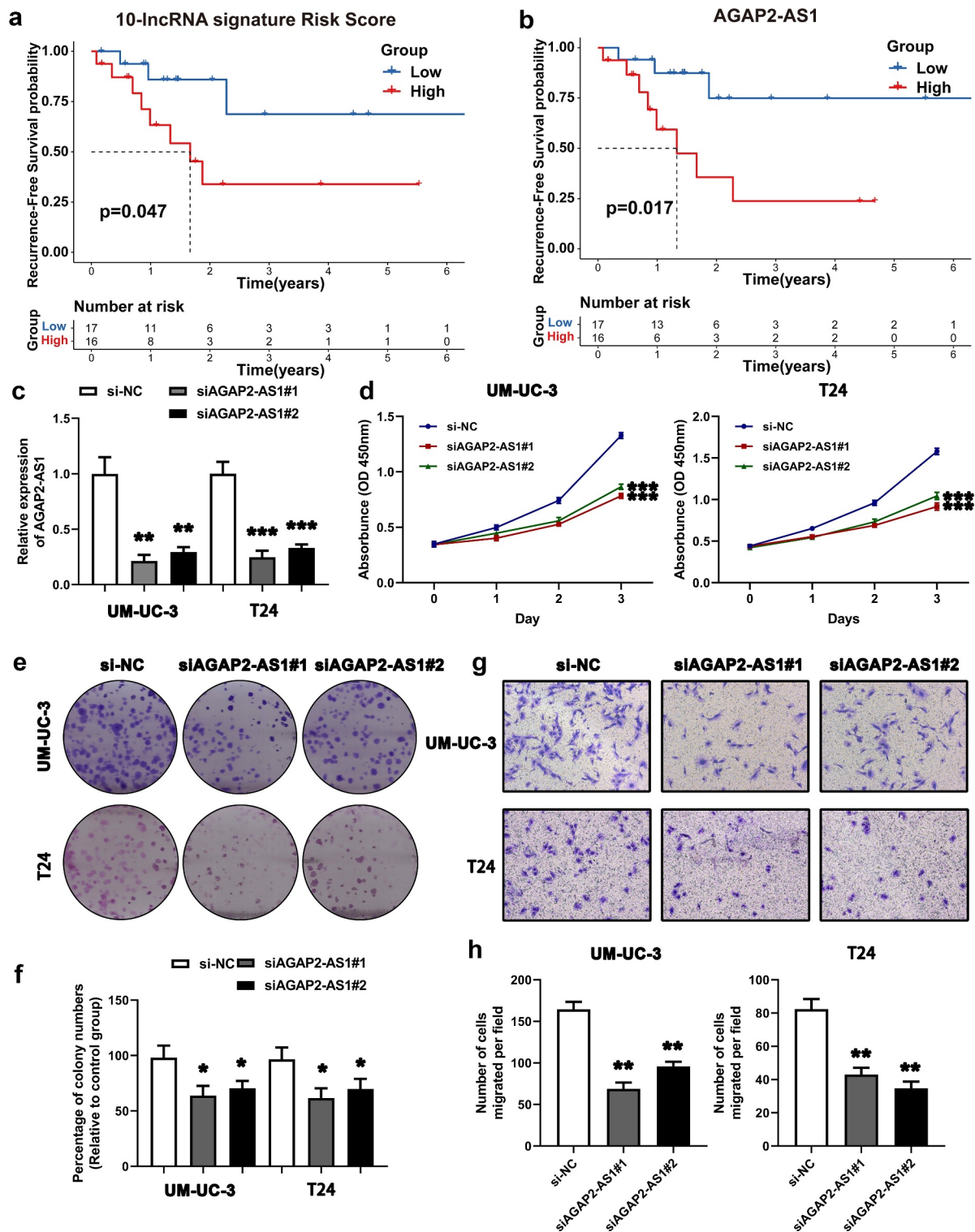


Figure 7. Survival analysis of the ten-lncRNA signature in BLCA samples and experimental study on the biofunction of AGAP2-AS1 on the BLCA cells. (a) Kaplan-Meier RFS curve analysis on the ten-lncRNA signature in the 33 BLCA samples. (b) Kaplan-Meier RFS curve analysis on the expression of AGAP2-AS1 in the 33 BLCA samples. (c) Evaluating the efficiency of AGAP2-AS1 knockdown via qRT-PCR. (d) CCK-8 assay on the effect of AGAP2-AS1 knockdown on cell proliferation. The OD value among different groups was found to be significantly different by two-way ANOVA. * $p < 0.05$, ** $p < 0.01$; *** $p < 0.001$. The data are expressed as the mean \pm SD. (e, f) Colony formation assay on the BLCA cells transfected with siRNA. (g, h) Transwell assay was used to evaluate the migration of BLCA cells transfected with siRNA.

among 38 lncRNAs associated with BLCA recurrence. We further selected the 16 lncRNAs as candidates from the 20 lncRNAs using the LASSO model. Multivariate COX regression analysis eventually selected out 10 lncRNAs to develop a recurrence-associated lncRNA signature in BLCA.

Moreover, validation of the recurrence-associated lncRNA signature indicated that ten-lncRNA signature had predictive power for the recurrence of BLCA. We conducted Cox regression analysis in training dataset and testing datasets, and the analysis verified that ten-lncRNA signature was an independent prognostic factor for BLCA recurrence. Stratified analysis indicated that ten-lncRNA signature could effectively classify cases into high and low recurrence-risk sets in different subgroups. Taken together, ten-lncRNA signature based on multi-omics analysis could act as a robust and independent predictor for BLCA recurrence.

Furthermore, we performed the GSEA to reveal the potential pathways related to ten-lncRNA signature. The results of GSEA contained several significantly enriched BLCA-associated pathways, such as ‘gap junction’, ‘primary immunodeficiency’, ‘intestinal immune network for IgA production’, ‘autoimmune thyroid disease’ and ‘tight junction’. We supposed that ten-lncRNA signature may contribute to tumor immunity whose dysregulations could play a critical role in BLCA relapse.

By comparing ten-lncRNA signature with the reported four-lncRNA signature and six-lncRNA signature, we validated that ten-lncRNA signature had better performance in predicting BLCA recurrence than the two reported lncRNA signature models, which suggested that our approach of multi-omics analysis on transcriptome data, genomic CNV data, mutation annotation data and clinical data may be superior in constructing the prognostic signature.

Above all, the ten-lncRNA signature had robust predictive power, which was an independent prognostic factor for BLCA relapse. Hence, ten-lncRNA signature could have potential implications as prognostic markers for BLCA recurrence. On the other hand, the approach we utilized for developing biomarkers may contribute to studying

cancer-associated RNA expression profiles in the future. Due to the limited samples collected from patients, the survival analysis based on the 33 BLCA patients revealed that only AGAP2-AS1 among the ten lncRNAs was associated with RFS in BLCA. Herein, we chose AGAP2-AS1 for further functional experiments, and the results revealed that AGAP2-AS1 knockdown could inhibit the cell proliferation and migration in BLCA cells for the first time. However, further investigation should be performed to validate biological functions and potential mechanisms of ten lncRNAs in BLCA.

5. Conclusion

The novel ten-lncRNA signature, constructed based on multi-omics data, had robust prognostic power in predicting the recurrence of BLCA and potential clinical implications as biomarkers for personalized management of BLCA.

Research highlights

Multi-omics analysis on CNV, mutation annotation, RNA expression and clinical data.

Constructing a novel and robust lncRNA signature to predict BLCA recurrence.

The lncRNA signature may be associated with tumor immunology.

AGAP2-AS1 knockdown could inhibit cell proliferation and migration in BLCA cells.

Acknowledgements

We acknowledge TCGA and GEO database for providing their platforms and contributors for uploading their meaningful datasets. The authors wish to thank Dr. Wang Keyi from Shanghai Tenth People’s Hospital of Tongji University for his help in collecting clinical samples and relevant experimental validation.

Disclosure statement

No potential conflict of interest was reported by the author (s).

Funding

This work was supported by the National Natural Science Foundation of China (81872089, 81672551, 81670632,

81871157, 82070773, 82102831); Jiangsu Provincial Key Research and Development Program (BE2019751), Natural Science Foundation of Jiangsu Province (BK20201271) and The National Key Research and Development Program of China (SQ2017YFSF090096).

Data availability statement

Original data were uploaded to a recognized data repository (URL: <https://zenodo.org/record/5153530#.YZtEZU5ByUk>; DOI: 10.5281/zenodo.5153530)

Author statement

Zhipeng Xu: Conceptualization, Methodology, Software, Writing - original draft preparation. Hui Chen: Data curation, Formal analysis, Writing - original draft preparation. Jin Sun: Validation, Writing - review and editing. Weipu Mao: Software, Visualization. Shuqiu Chen: Funding acquisition, Supervision. Ming Chen: Conceptualization, Funding acquisition, Supervision.

ORCID

Ming Chen  <http://orcid.org/0000-0002-3572-6886>

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(6):394–424.
- [2] Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. *Lancet.* 2009;374(9685):239–249.
- [3] Kamat AM, Hahn NM, Efstathiou JA, et al. Bladder cancer. *Lancet.* 2016;388(10061):2796–2810.
- [4] Del Bene G, Calabro F, Giannarelli D, et al. Neoadjuvant vs. Adjuvant chemotherapy in muscle invasive bladder cancer (MIBC): analysis from the RISC database. *Front Oncol.* 2018;8:463.
- [5] Hautmann RE, Gschwend JE, de Petriconi RC, et al. Cystectomy for transitional cell carcinoma of the bladder: results of a surgery only series in the neobladder era. *J Urol.* 2006;176(2):486–492. discussion 91–2.
- [6] Wosnitzer MS, Hruby GW, Murphy AM, et al. A comparison of the outcomes of neoadjuvant and adjuvant chemotherapy for clinical T2-T4aN0-N2M0 bladder cancer. *Cancer.* 2012;118(2):358–364.
- [7] Sylvester RJ, van der Meijden AP, Oosterlinck W, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur Urol.* 2006;49(3):466–5. discussion 75–7. DOI:10.1016/j.eururo.2005.12.031.
- [8] Patel VG, Oh WK, Galsky MD. Treatment of muscle-invasive and advanced bladder cancer in 2020. *CA Cancer J Clin.* 2020;70(5):404–423.
- [9] Dobruch J, Daneshmand S, Fisch M, et al. Gender and bladder cancer: a collaborative review of etiology, biology, and outcomes. *Eur Urol.* 2016;69(2):300–310.
- [10] Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. *J Clin Invest.* 2016;126(8):2775–2782.
- [11] Zhan Y, Chen Z, He S, et al. Long non-coding RNA SOX2OT promotes the stemness phenotype of bladder cancer cells by modulating SOX2. *Mol Cancer.* 2020;19(1):25.
- [12] He W, Zhong G, Jiang N, et al. Long noncoding RNA BLACAT2 promotes bladder cancer-associated lymphangiogenesis and lymphatic metastasis. *J Clin Invest.* 2018;128(2):861–875.
- [13] Chen C, He W, Huang J, et al. LNMAT1 promotes lymphatic metastasis of bladder cancer via CCL2 dependent macrophage recruitment. *Nat Commun.* 2018;9(1):3826.
- [14] de Jong JJ, Liu Y, Robertson AG, et al. Long non-coding RNAs identify a subset of luminal muscle-invasive bladder cancer patients with favorable prognosis. *Genome Med.* 2019;11(1):60.
- [15] Du Y, Wang B, Jiang X, et al. Identification and validation of a stromal EMT related lncRNA signature as a potential marker to predict bladder cancer outcome. *Front Oncol.* 2021;11:620674.
- [16] Mao X, Chen S, Li G. Identification of a ten-long noncoding RNA signature for predicting the survival and immune status of patients with bladder urothelial carcinoma based on the GEO database: a superior machine learning model. *Aging (Albany NY).* 2021;13(5):6957–6981.
- [17] Gao X, Zhang S, Chen Y, et al. Development of a novel six-long noncoding RNA signature predicting survival of patients with bladder urothelial carcinoma. *J Cell Biochem.* 2019;120(12):19796–19809.
- [18] Zhang S, Zhang J, An Y, et al. Multi-omics approaches identify SF3B3 and SIRT3 as candidate autophagic regulators and druggable targets in invasive breast carcinoma. *Acta Pharm Sin B.* 2021;11(5):1227–1245.
- [19] Gout AM, Arunachalam S, Finkelstein DB, et al. Data-driven approaches to advance research and clinical care for pediatric cancer. *Biochim Biophys Acta Rev Cancer.* 2021;1876(1):188571.
- [20] Lee D, Park Y, Kim S. Towards multi-omics characterization of tumor heterogeneity: a comprehensive review of statistical and machine learning approaches. *Brief Bioinform.* 2021;22:3.
- [21] Lewis JE, Kemp ML. Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat Commun.* 2021;12(1):2700.
- [22] Murugesan M, Premkumar K. Systemic multi-omics analysis reveals amplified P4HA1 gene associated with

- prognostic and hypoxic regulation in breast cancer. *Front Genet.* 2021;12:632626.
- [23] Zhao Y, Gao Y, Xu X, et al. Multi-omics analysis of genomics, epigenomics and transcriptomics for molecular subtypes and core genes for lung adenocarcinoma. *BMC Cancer.* 2021;21(1):257.
- [24] Chaudhary K, Poirion OB, Lu L, et al. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res.* 2018;24(6):1248–1259.
- [25] Harrow J, Denoeud F, Frankish A, et al. GENCODE: producing a reference annotation for ENCODE. *Genome Biol.* 2006;7(Suppl 1):S4 1–9.
- [26] Riester M, Taylor JM, Feifer A, et al. Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer. *Clin Cancer Res.* 2012;18(5):1323–1333.
- [27] Hicks SC, Irizarry RA. Quantro: a data-driven approach to guide the choice of an appropriate normalization method. *Genome Biol.* 2015;16:117.
- [28] Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249–264.
- [29] Mermel CH, Schumacher SE, Hill B, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 2011;12(4):R41.
- [30] Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499(7457):214–218.
- [31] Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385–395.
- [32] Liu G, Locascio JJ, Corvol JC, et al. Prediction of cognition in Parkinson's disease with a clinical-genetic score: a longitudinal analysis of nine cohorts. *Lancet Neurol.* 2017;16(8):620–629.
- [33] Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J.* 2005;47(4):458–472.
- [34] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550.
- [35] Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet.* 2003;34(3):267–273.
- [36] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–457.
- [37] Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013;4:2612.
- [38] Raposo CJ, McElroy KA, Fuchs SM. The Epithelial adhesin 1 tandem repeat region mediates protein display through multiple mechanisms. *FEMS Yeast Res.* 2020;20:3.
- [39] Luo WL, Luo MX, He RZ, et al. Reveals the pan-cancer landscape of bone morphogenetic proteins. *Med Sci Monit.* 2020;26:e920943.
- [40] Seal DB, Das V, Goswami S, et al. Estimating gene expression from DNA methylation and copy number variation: a deep learning regression model for multi-omics integration. *Genomics.* 2020;112(4):2833–2841.
- [41] Bao Z, Zhang W, Dong D. A potential prognostic lncRNA signature for predicting survival in patients with bladder urothelial carcinoma. *Oncotarget.* 2017;8(6):10485–10497.
- [42] Barik GK, Sahay O, Behera A, et al. Keep your eyes peeled for long noncoding RNAs: explaining their boundless role in cancer metastasis, drug resistance, and clinical application. *Biochim Biophys Acta Rev Cancer.* 2021;1876(2):188612.
- [43] Ahmad S, Abbas M, Ullah MF, et al. Long non-coding RNAs regulated NF-kappaB signaling in cancer metastasis: micromanaging by not so small non-coding RNAs. *Semin Cancer Biol.* 2021. DOI:10.1016/j.semcancer.2021.07.015.
- [44] Kong FE, Tang YQ, Gong YF, et al. Identification of prognostic claudins signature in hepatocellular carcinoma from a hepatocyte differentiation model. *Hepatol Int.* 2020;14(4):521–533.
- [45] Bo H, Zhu F, Liu Z, et al. Integrated analysis of high-throughput sequencing data reveals the key role of LINC00467 in the invasion and metastasis of testicular germ cell tumors. *Cell Death Discov.* 2021;7(1):206.
- [46] Unfried JP, Marin-Baquero M, Rivera-Calzada A, et al. Long noncoding RNA NIHCOLE promotes ligation efficiency of DNA double-strand breaks in hepatocellular carcinoma. *Cancer Res.* 2021;81(19):4910–4925.
- [47] Li RH, Tian T, Ge QW, et al. A phosphatidic acid-binding lncRNA SNHG9 facilitates LATS1 liquid-liquid phase separation to promote oncogenic YAP signaling. *Cell Res.* 2021;31(10):1088–1105.
- [48] Huang Z, Zhou JK, Peng Y, et al. The role of long noncoding RNAs in hepatocellular carcinoma. *Mol Cancer.* 2020;19(1):77.
- [49] Cheng J, Meng J, Zhu L, et al. Exosomal noncoding RNAs in Glioma: biological functions and potential clinical applications. *Mol Cancer.* 2020;19(1):66.
- [50] Cao Y, Tian T, Li W, et al. Long non-coding RNA in bladder cancer. *Clin Chim Acta.* 2020;503:113–121.
- [51] Liu Y, Zhang P, Wu Q, et al. Long non-coding RNA NR2F1-AS1 induces breast cancer lung metastatic dormancy by regulating NR2F1 and DeltaNp63. *Nat Commun.* 2021;12(1):5232.

- [52] Kikuchi Y, Tokita S, Hiramata T, et al. CD8+ T-cell immune surveillance against a tumor antigen encoded by the oncogenic long non-coding RNA, PVT1. *Cancer Immunol Res.* 2021. DOI:10.1158/2326-6066.CIR-20-0964.
- [53] Ma L, Xu A, Kang L, et al. LSD1-Demethylated LINC01134 confers oxaliplatin resistance via SP1-induced p62 transcription in hepatocellular carcinoma. *Hepatology* 2021. DOI:10.1002/hep.32079.
- [54] Zhang MX, Zhang LZ, Fu LM, et al. Positive feedback regulation of lncRNA PVT1 and HIF2alpha contributes to clear cell renal cell carcinoma tumorigenesis and metastasis. *Oncogene.* 2021;40(37):5639–5650.