

Model Construction and Identification of Genome Instability-Associated LncRNA Signature Markers in Bladder Cancer

Jun Li, Hui Pan, Xiaowu Pi, Ying Xiong*

The First Affiliated Hospital of Yangtze University, Jingzhou, China

Email: *31350266@qq.com

How to cite this paper: Li, J., Pan, H., Pi, X.W. and Xiong, Y. (2022) Model Construction and Identification of Genome Instability-Associated LncRNA Signature Markers in Bladder Cancer. *Journal of Biosciences and Medicines*, 10, 46-63. <https://doi.org/10.4236/jbm.2022.101006>

Received: December 9, 2021

Accepted: January 14, 2022

Published: January 17, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Background: Studies have shown that long non-coding RNA (LncRNA) plays a critical role in maintaining genomic instability. The correlation between LncRNA and genomic instability is still worth exploring in bladder cancer as a new tumour marker. **Methods:** Therefore, combined with the LncRNA expression profile and somatic mutation profile of bladder cancer, we established a computing framework of LncRNA related to genomic instability and identified 58 new LncRNA related to genomic instability. Next, we identified a LncRNA signature (GILncSig), based on these 58 new genes, which divided patients into high-risk and low-risk groups. The clinical prognosis was significantly different and was further verified in an independent cohort of patients. **Results:** We confirmed that GILncSig is related to the genomic mutation rate of bladder cancer, suggesting that GILncSig can be used as an indicator of genomic instability. The results show that GILncSig has prognostic value independent of age, sex, grade, and stage and is vital in evaluating clinical prognosis. To sum up, this study provides a vital research basis and methods for further exploring the role of LncRNA in the genomic instability of bladder cancer and provides a theoretical basis for the identification of bladder cancer biomarkers related to genomic instability.

Keywords

Bladder Cancer, LncRNA, Genomic Instability, Biomarkers, Bioinformatics

1. Introduction

Bladder cancer is a common malignant tumour of the urinary system. The American Cancer Society estimates 76,960 new cases and 16,390 bladder cancer deaths in the US with a male: female ratio of 3:1 [1]. These figures indicate that

bladder cancer morbidity and mortality have risen in recent years [1] [2]. Early bladder cancer has a good prognosis, but it is easy to relapse and develop into muscle-infiltrating bladder cancer (muscle-infiltrating bladder cancer, MIBC), and MIBC has a strong ability to invade and metastasize, the prognosis is poor, and the 5-year survival rate is less than 50% [3] [4]. At present, the leading indicators to predict the progression or recurrence of bladder cancer are tumour grade, lymph node metastasis, p53/pRb signal pathway, receptor tyrosine kinase (Receptortyrosinekinase, RTK) pathway, and so on. Moreover, bladder cancer is a rather complex disease with molecular and clinical heterogeneity and its development, progression, and therapeutic response [5]. Therefore, there is an urgent need to find new biomarkers to evaluate the clinical prognosis of patients with bladder cancer more accurately.

It is reported that genomic instability is one of the prominent features of cancer [6]. More importantly, some studies have confirmed that genomic instability is an important prognostic factor, and the degree of genomic instability is related to tumour progression and survival rate [7] [8]. Although the mechanism of genomic instability is not fully understood, there has been evidence in recent years that some molecular signatures can be used as indicators of genomic instability [8] [9]. For example, GuntherBoysen *et al.* [10] revealed that the most common mutant gene, SPOP in primary prostate cancer, is related to genomic instability. SPOP mutations partly promote the development of prostate tumours through genomic instability, suggesting that SPOP mutations may enhance the response to DNA damage treatment. JanaBiermann *et al.* [11] identified 17 markers related to genomic stability and found that these markers have a significant impact on the clinical prognosis of breast cancer. A long non-coding RNA (LncRNA) is a widely defined transcript as more significant than 200 nt and has little potential to encode a protein [12]. LncRNA plays a vital role in different biological processes [13] [14], especially the abnormal expression of some lncRNA may affect the proliferation, progression, or metastasis of tumour cells [15] [16]. With the development of bioinformatics technology, some scholars have found a series of potentially available biomarkers of urinary tumours through bioinformatics [17]. It is possible to use bioinformatics to study the expression profile of lncRNA in tumours and normal tissues on a large scale, and it has been found that there are a large number of abnormal expressions of lncRNA in various tumours. However, the function of this lncRNA is still unclear [18] [19]. In the past few years, more and more studies have shown that lncRNA also plays a vital role in maintaining genomic instability [20] [21]. For example, Chunzhi Zhang *et al.* [22] showed that through extensive crosstalk between non-coding RNA and typical DNA damage response (DDR) signal pathway, DDR-induced non-coding RNA expression could provide a precise regulation mechanism of DNA damage response gene expression in time and space. Kirsten M.T *et al.* [21] found a new lncRNA, MANCR (mitosis-related long non-coding RNA; LINC00704), up-regulated in breast cancer samples and cells. The deletion of the MANCR

gene in triple-negative breast cancer cells significantly decreased cell proliferation and vitality, accompanied by increased DNA damage. These shreds of evidence suggest that lncRNA is associated with genomic instability. However, the association between this genomic instability-related lncRNA and the clinical prognosis of tumour patients is not clear.

In this study, the lncRNA expression profile and somatic mutation profile in the bladder cancer genome were combined to establish the calculation framework of lncRNA related to genomic instability, explore the possibility of lncRNA signal as an indicator of genomic instability, and improve the prognostic value of bladder cancer patients.

2. Materials & Methods

2.1. Data Collection

The clinical characteristics, RNA-seq expression data, and somatic mutation information of patients with bladder cancer were collected from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>). 433 RNA-seq expression data samples were collected, including 19 standard samples and 414 tumour samples, and 412 patients with clinical features and somatic mutations were collected.

2.2. Identification of lncRNAs Associated with Genomic Instability

To identify the lncRNA associated with genomic instability, we established a computational framework for genomic instability-related lncRNAs, which combines the lncRNA expression profile and somatic mutation profile in the bladder cancer genome. First, the cumulative number of somatic mutations for each patient was computed; Second, patients were ranked in decreasing order of the cumulative number of somatic mutations; Then, the top 25% of patients were defined as genomic unstable (GU)-like group, and the last 25% were defined genomically stable (GS)-like group; Next, The lncRNAs expression profiles of GU group and GS group were compared by microarray significance analysis (SAM) method; Finally, differentially expressed lncRNAs (fold change > 1 or < -1 and false discovery rate (FDR) adjusted $P < 0.05$) were defined as genome instability-associated lncRNAs.

2.3. Statistical Analysis

We use Euclidean distances and Ward's linkage method for Hierarchical cluster analysis. The relationship between genome instability-associated lncRNA expression level and overall survival was evaluated by univariate and multivariate Cox proportional hazard regression analysis. According to the coefficients of multiple regression analysis and the expression levels of prognostic genome instability-associated lncRNA, we constructed a genome instability-associated lncRNA signature (GILncSig) to predict the results as follows:

$$\text{GILncSig (samples)} = (\text{Coef value} * \text{Expression quantity}) \text{ lncRNA1} + (\text{Coef}$$

value * Expression quantity) lncRNA2 + ... + (Coef value * Expression quantity) lncRNAn.

GILncSig (sample) is the prognostic risk score of patients with Bladder cancer. The lncRNA represents the nth prognostic lncRNA, and Expression quantity is the expression level of lncRNA for the patient. The Coef value represents the contribution of lncRNA to the prognostic risk score obtained from the regression coefficient of multivariate Cox analysis. Taking the median score of the patients in the training set as the risk cutoff to classify patients into the high-risk group with high GILncSig or low-risk group with low GILncSig.

The survival rate and median survival time of each prognostic risk group were calculated by the Kaplan-Meier method. The log-rank test evaluated the difference in survival rate between the high-risk and low-risk groups. Multivariate Cox regression and stratified analysis were used to evaluate the independence of GILncSig and other critical clinical factors. Hazard ratio (HR) and 95% confidence interval (CI) are calculated by Cox analysis. The performance of the GILncSig is also evaluated by the time-dependent receiver operating characteristic (ROC) curve. All statistical analyses were carried out using Rversion3.6.2.

2.4. Functional Enrichment Analysis

We calculated the Pearson correlation coefficient to measure the correlation between lncRNAs and mRNAs paired expressions, and the first ten mRNAs were considered co-expressed lncRNA-related partners. To predict the potential function of lncRNAs, we conducted functional enrichment analysis of co-expressed lncRNA-related mRNA partners to identify significantly enriched Gene Ontology (GO) terms and the Kyoto Encyclopedia of Gene and Genome (KEGG) pathway. The function enrichment analysis was carried out using the cluster profile software in R-version 3.6.2 [23].

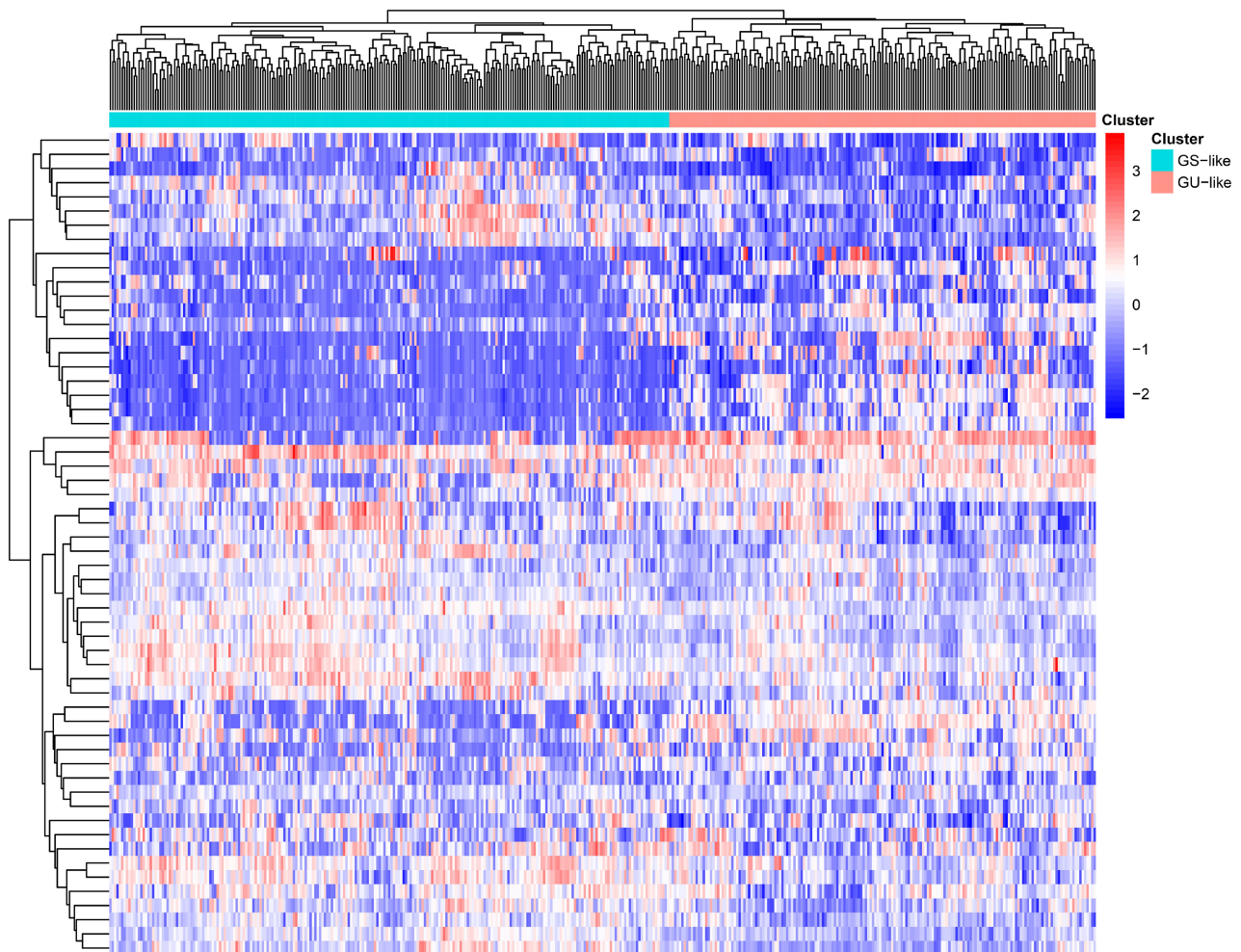
3. Results

3.1. Identification of Genomic Instability-Related lncRNAs in Bladder Cancer Patients

To identify determine the lncRNA, associated with genomic instability, the cumulative number of somatic mutations in each patient was calculated and sorted in descending order. According to the cumulative number of somatic mutations, the first 25% (nude 105) and the last 25% (nude 105) of patients were divided into GU-like groups and GS-like groups. Then the lncRNA expression profiles of the Gu-like group and GS-like group were compared to find out the significant difference of lncRNA. Using the SAM method, we found that a total of 58 lncRNA were considered to be significantly differentially expressed, their folding changes were more outstanding than or less than -1, and their FDR-adjusted P values were less than 0.05. Among them, 35 lncRNA were up-regulated and 23 down-regulated in the GU-like group. We used 58 differentially expressed lncRNA sets to perform unsupervised hierarchical clustering analysis on 412

samples of TCGA sets. As shown in **Figure 1(a)**, all 412 samples were divided into two groups according to 58 differentially expressed lncRNA expression levels. There was a significant difference in the pattern of somatic mutation between the two groups. The group with high accumulative somatic cell mutation was called the GU-like group, and the other group was the GS-like group. The median value of cumulative somatic mutations in the GU-like group was significantly higher than that in the GS-like group ($P = 3.2E-06$, Mann-Whitney U test; **Figure 1(b)**). Next, we compared the expression levels of the UBQLN4 gene (a newly discovered driver of genomic instability) in the GS-like group and GU-like group. The expression of UBQLN4 in the Gu-like group was significantly higher than that in the GS-like group ($P = 0.0021$, Mann-Whitney U test; **Figure 1(c)**).

To determine whether the potential functions and pathways involved in 58 lncRNA are related to genomic instability, we conducted functional enrichment analysis to predict potential functions. We first measured the expression correlation between 58 differentially expressed lncRNA and protein-coding gene (PCG) and obtained the lncRNA-related PCG, the top 10 PCG with the strongest correlation



(a)

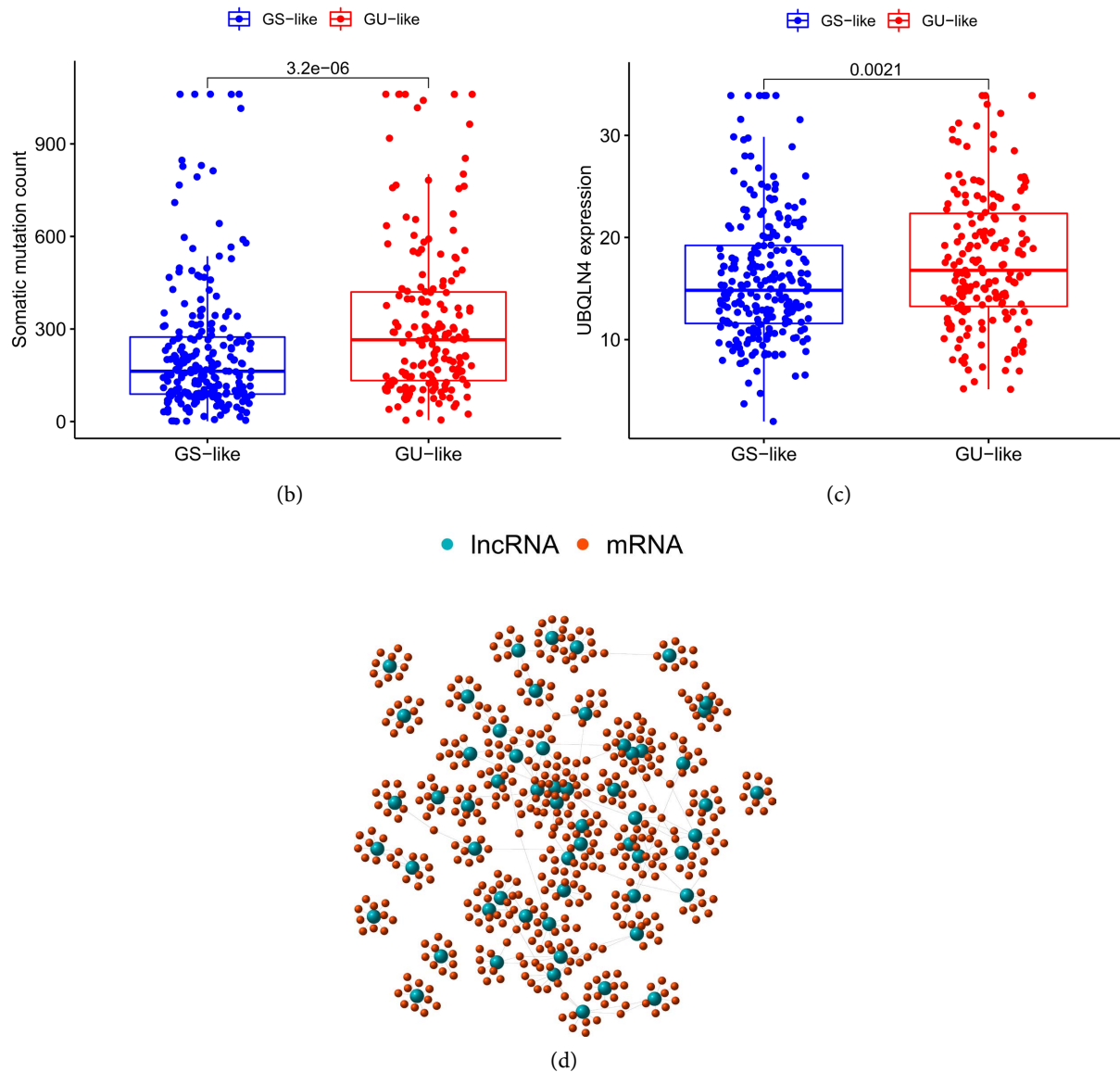


Figure 1. (a) All 412 samples were divided into two groups according to the expression levels of 58 differentially expressed lncRNA. There was a significant difference in the pattern of somatic mutation between the two groups. The group with high accumulative somatic cell mutation was called the GU-like group, and the other group was the GS-like group; (b) The median value of somatic cumulative mutations in the GU-like group was significantly higher than that in the GS-like group ($P = 3.2 \times 10^{-6}$, Mann-Whitney U test; 1(b)). (c) As shown in **Figure 1(c)**, the expression of UBQLN4 in the Gu-like group was significantly higher than that in the GS-like group ($P = 0.0021$, Mann-Whitney U test). (d) lncRNAs-mRNA co-expression network;

with each lncRNAs. A lncRNAs-mRNA co-expression network is constructed, in which the nodes are lncRNA and mRNAs, if they are related to each other, then connect lncRNA and mRNAs; (**Figure 1(d)**). GO analysis of lncRNA-related PCG showed that mRNAs in the network were closely related to the formation and development of chromosomes and cellular genes, including the regulation of mitosis, chromosome segregation, and mitosis. Through the analysis of KEGG pathways of lncRNA-related PCG, 12 significantly rich pathways were found, including cytokine-cytokine receptor interaction pathways, calcium sig-

nalling pathways, Wnt signalling pathways, and pathways related to carcinogenesis.

3.2. Development of LncRNAs Signature Derived from Genomic Instability for the Training Set Result Prediction

To further explore the clinical value of lncRNA associated with these candidates' genomic instability, 412 bladder cancer patients from the TCGA project were divided into the training and testing sets. To screen lncRNA, related to prognosis, univariate Cox proportional hazard regression was used to analyze the relationship between the expression level of 58 genomic instability-related lncRNA and OS in the training set. It was found that eight genomic instability-related lncRNAs were significantly associated with the prognosis of bladder cancer patients ($P < 0.05$; **Figure 2**). Also, we hope to screen the lncRNA with the independent prognostic value from the eight candidate lncRNA and then conduct multivariate Cox proportional hazard regression analysis on the eight candidates lncRNA. Finally, four of the eight candidates, lncRNA (AC078880.3, AL355916.1, AC037198.1, and LINC02446), were identified as independent prognostic indicators because they still had prognostic significance in multivariate COX ($P < 0.05$). Then, according to the coefficient of multivariate Cox analysis and the expression levels of four independent lncRNA related to prognostic instability, a genomic instability-related lncRNA signal (GILncSig) was constructed to evaluate the prognostic risk of bladder cancer patients. GILncSig score = $(-0.5860 \times \text{expression level of AC078880.3}) + (0.0472 \times \text{expression level of AL355916.1}) + (0.0988 \times \text{expression level of AC037198.1}) + (-0.2257 \times \text{expression level of LINC02446})$. In GILncSig, lncRNAAL355916.1 and LINC037198.1 may be risk factors, suggesting that its high expression is related to poor prognosis. At the same time, AC078880.3 and LINC02446 may be protective factors, and their high expression is related to prolonged survival.

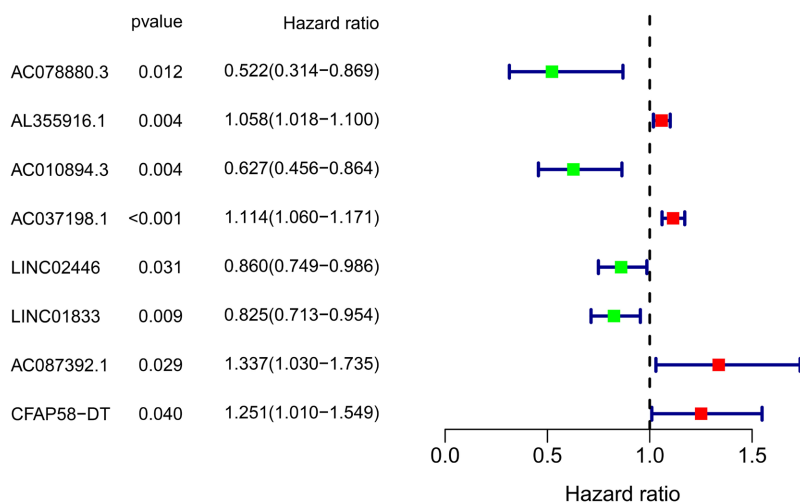


Figure 2. The eight genomic instability-related lncRNAs are significantly associated with the prognosis of bladder cancer patients ($P < 0.05$).

To test whether the characteristic clinical data of the above training set and the test set are different, we make statistics on the characteristic clinical data of the two groups, including age, gender, grade, stage, and TNM stages. Chi-square test analysis shows no difference between the two groups of clinical characteristics that can be used as a basis for grouping.

GILncSig was used to obtain the risk score of each patient in the training set, and then these patients were divided into different predictive groups with the median risk score as the threshold. The group with a patient score equal to or above the threshold is called the high-risk group, and the other group is called the low-risk group. Kaplan-Meier analysis showed that the survival outcome of patients in the low-risk group was significantly better than that in the high-risk group ($p < 0.001$, log-rank test). The time-dependent ROC curve of GILncSig was analyzed, and the area under the curve (AUC) was 0.740 (3 years) and 0.801 (5 years). We classified the patients in the training set according to their scores. We observed that the expression level of GILncSig, the number of somatic mutations, and the expression level of UBQLN4 changed with the increase of the score. In the high score group, the expression of risk gene AL355916.1LINC037198.1 was up-regulated, while the expression of protective genes, AC078880.3 and LINC02446 were down-regulated. On the contrary, the GILncSig showed the opposite expression pattern in patients with low scores. Comparative analysis showed significant differences in somatic mutation pattern and UBQLN4 expression pattern between the high-risk and low-risk groups. The number of somatic mutations in the high-risk group was significantly higher than that in the low-risk group. (Median somatic mutation counts 300 versus 150, $P = 8.6e-05$, Mann-Whitney U test).

3.3. Independent Validation of GILncSig in the Bladder Cancer Data Set Based on RNA-Seq Platform

To test the effectiveness of GILncSig, the prognostic performance of GILncSig was then tested in patients with an independent TCGA test set. When the test set used the same GILncSig and risk cutoff as the training set, the patients in the test set were divided into the high-risk group and low-risk group. There was a significant difference in overall survival. The overall survival rate of patients in the high-risk group was much worse than that in the low-risk group ($P = 0.001$, log-rank test). Based on the analysis of the time-dependent ROC curves of GILncSig in the test set, the AUC value is 0.684 (3 years). There was a significant difference in the mode of somatic mutation between the high-risk group and low-risk group ($P = 0.00024$, Mann-Whitney U test). The prognosis of GILncSig in the TCGA set was similar to that mentioned above. Patients in the TCGA set were divided into a high-risk group and a low-risk group, and the median survival time of patients in the high-risk group was shorter than that in the low-risk group ($P < 0.001$, log-rank test). The time-dependent ROC curves analysis was applied to the TCGA set. Consistent results were observed above (AUC = 0.712 (3 years)); There was a significant difference in the distribution of somatic muta-

tions between the high-risk and low-risk groups. ($P = 1.3e-07$, Mann-Whitney U test).

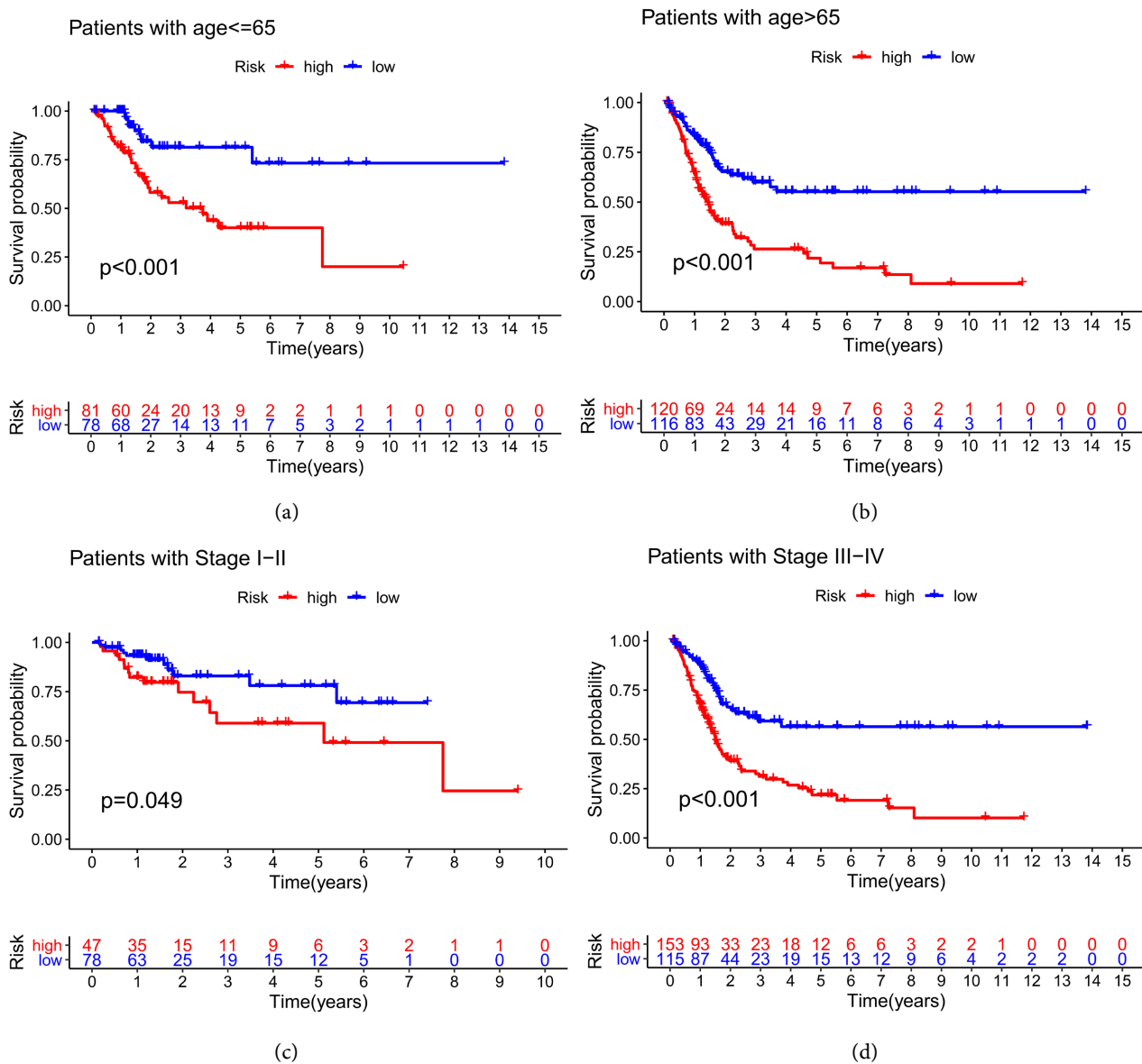
3.4. Independence of the GILncSig from Other Clinical Factors

We conducted a stratified analysis of GILncSig to determine whether GILncSig had prognostic value independent of age, Gender, Grade, and Stage. Patients in the TCGA set were stratified into a young-patient and an old-patient group according to age (age = 65). Patients who use GILncSig can be further divided into high-risk or low-risk groups in each age group. There was a significant difference in overall survival between the high-risk and low-risk groups in the young-patient group (log-rank test $P < 0.001$; **Figure 3(a)**) as was in the old-patient group (log-rank test $P < 0.001$; **Figure 3(b)**). Next, all the patients with bladder cancer were stratified according to the pathological stage, the patients with pathological stages I and II were combined into the early group, and the patients with pathological stages III and IV were combined into the late group. The GILncSig could classify patients with pathologic stage I or II into high-risk and low-risk groups, and there is a significant difference in overall survival rate between the two groups (log-rank test $P = 0.049$; **Figure 3(c)**). Similarly, GILncSig could also be used to divide pathological stage III and IV patients into a high-risk group and a low-risk group. There was a significant difference in the overall survival rate between the two groups (log-rank test $P < 0.001$; **Figure 3(d)**). According to the sex of the patients, the patients in the TCGA set were divided into female groups and male groups. Patients in the GILncSig group, the female group, and the male group can be further divided into high-risk or low-risk groups. There was a significant difference in overall survival rate between the high-risk and low-risk groups in the female group (log-rank test $p < 0.001$; **Figure 3(e)**), the male group also had a significant difference (log-rank test $p < 0.001$; **Figure 3(f)**). According to the grade of the patients, the TCGA set was divided into a High-Grade and a Low-Grade group. Patients using each grade group of GILncSig could be further divided into high-risk or low-risk groups. There was a significant difference in overall survival between the high-risk and low-risk groups in the High-Grade group (logarithmic rank test $p < 0.001$; **Figure 3(g)**), but there was no significant difference in the Low-Grade group. Also, to evaluate whether the prognostic value of GILncSig is independent of common clinical variables, we performed multivariate Cox regression analysis on age, sex, grade, stage, and GILncSig-based prognostic risk score model. The results of multivariate analysis showed that after adjusting for age, sex, grade, and stage, GILncSig was significantly correlated with the overall survival rate of each group. In multivariate analysis, in addition to GILncSig, two other clinical factors, age, and stage were significant.

3.5. Comparison of Survival Prediction Performance between the GILncSig and Existing LncRNA Related Signatures

We further compared the prediction performance of GILncSig with the three

recently published lncRNA signatures: The 5-lncRNA signature obtained from Cao's research (hereinafter referred to as CaoLncSig) [24] and the 8-lncRNA signature obtained from Lian's research (hereinafter referred to as LianLncSig) [25], and the 14-lncRNA signature obtained from Zhang's research (hereinafter referred to as ZhangLncSig) [26]). They are using the same TCGA patient cohort. The AUC at three years of OS for the GILncSig is 0.710, which is significantly higher than that of CaoLncSig (AUC = 0.657) and LianLncSig (AUC = 0.590) and ZhangLncSig (AUC = 0.548). Besides, the AUC at five years of OS for the GILncSig is 0.801, which is significantly higher than that of CaoLncSig (AUC = 0.645) and LianLncSig (AUC = 0.630) and ZhangLncSig (AUC = 0.600). In addition, the number of lncRNA in our GILncSig is relatively moderate. These results show that GILncSig has better prognostic performance than the three recently published lncRNA signatures in predicting survival.



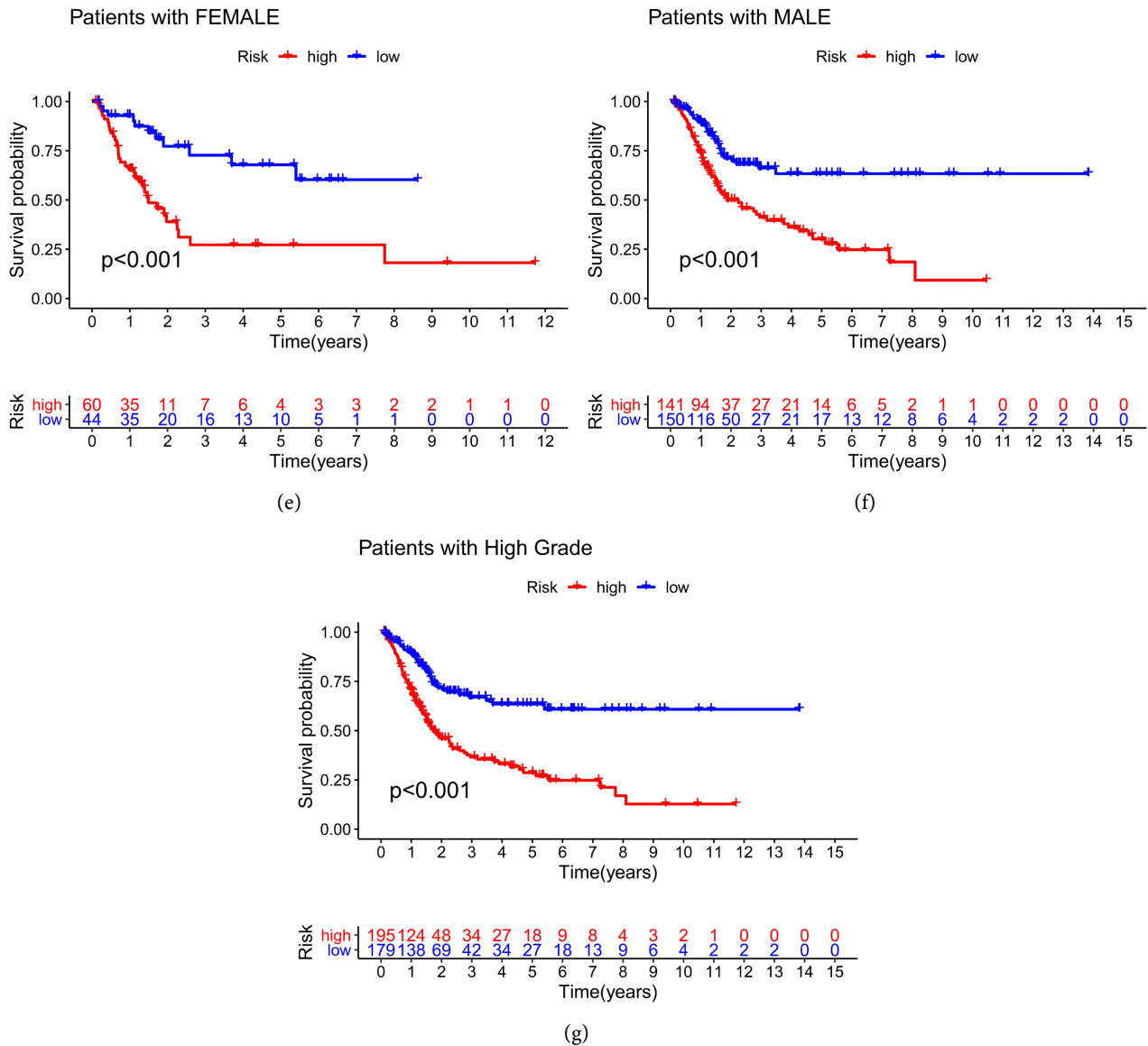


Figure 3. There was a significant difference in overall survival between the high-risk and low-risk groups in the young-patient group (log-rank test $P < 0.001$; 3(a)) as was in the old-patient group (log-rank test $P < 0.001$; 3(b)). The GILncSig could classify the patients with pathologic stage I or II into a high-risk group and low-risk group, and there is a significant difference in overall survival rate between the two groups (log-rank test $P = 0.049$; 3(c)). Similarly, GILncSig could also be used to divide pathological stage III and IV patients into a high-risk group and a low-risk group. There was a significant difference in the overall survival rate between the two groups. (log-rank test $P < 0.001$; 3(d)). There was a significant difference in overall survival rate between the high-risk and low-risk groups in the female group (log-rank test $p < 0.001$; 3(e)), the male group also had a significant difference (log-rank test $p < 0.001$; 3(f)). Patients using each grade group of GILncSig could be further divided into high-risk or low-risk groups. There was a significant difference in overall survival between the high-risk group and the low-risk group in the High-Grade group (logarithmic rank test $p < 0.001$; 3(g)), but there was no significant difference in the Low-Grade group.

3.6. The GILncSig Predicts the Results Better Than the TP53 Mutation Status

Further analysis showed that in the training set, testing set, and TCGA set, the proportion of patients with TP53 mutations in the high-risk group was significantly higher than that in the low-risk group. In the training set, 55% of patients

in the high-risk group possessed TP53 mutations, significantly higher than 42% of patients in the low-risk group (chi-square test $P = 0.012$). Similar results were found in the testing set and the TCGA set. In the testing set, 55% of patients in the high-risk group possessed TP53 mutations, significantly higher than 40% of patients in the low-risk group (chi-square test $P = 0.050$). In the TCGA set, 54% of patients in the high-risk group possessed TP53 mutations, significantly higher than 42% of patients in the low-risk group (chi-square test $P = 0.022$). It is well known that TP53 maintains genomic stability, and TP53 mutation is associated with poor survival and can be used as an independent prognostic marker for bladder cancer [27] [28] [29] [30]. It is suggested that GILncSig is also related to the mutation state of TP53, which may be the mutation marker of the TP53 gene.

4. Discussion

With the increasing incidence of bladder cancer in recent years, research on bladder cancer is diversified [31] [32] [33] [34]. At present, there is still no breakthrough. At the current stage, patients are divided into different treatment groups according to their pathological characteristics. At the same time, histopathological features such as staging, grading, and molecular subtypes of bladder tumours are still the most critical factors affecting the prognosis of patients [35] [36] [37] [38] [39]. However, due to the limitation of standard clinicopathological features, the clinical prognosis of patients with bladder cancer is still highly heterogeneous [40]. Studies have shown that genomic instability is a common feature of most cancers [41] [42] [43] and one of the factors affecting the prognosis of bladder cancer [44]. Genomic instability plays a leading role in cancer occurrence, development, and recurrence. Therefore, the category and degree of genomic instability have important diagnostic and prognostic significance. There is growing evidence that abnormal transcription or epigenetic changes of genes *in vivo* or *in vitro* can lead to genomic instability. However, the quantitative measurement of genomic instability has always been a significant challenge [45]. Efforts have been made to identify PCGs and microRNAs associated with genomic instability and develop genes or miRNA features that predict genomic instability [46] [47].

Currently, lncRNA is an essential part of tumour biology, and their abnormal expression in cancer is associated with disease progression and may serve as prognostic markers for patients [48] [49] [50]. Furthermore, studies on the functional mechanism of lncRNA have found that lncRNA is also essential for genomic stability, such as MALAT1 [51] and NEAT1 [52]. Although much evidence has confirmed that lncRNA is associated with genomic instability, the construction of such models and the systematic exploration of their clinical value in bladder cancer is still lacking. Therefore, we have developed a computational framework for identifying genomic instability-related lncRNA that combines bladder cancer gene expression profiles and mutant phenotypes. Then, we identified 58 new lncRNA associated with genomic instability. Based on the functional analysis of the genes co-expressed with 58 lncRNAs associated with

genomic instability, our observation showed that the genes co-expressed with the 58 lncRNAs were enriched in the regulation of mitosis. There are three primary sources of genomic instability: DNA damage, replication pressure, and chromosome segregation defects. The regulation of chromosome mitosis is essential to the cell cycle, and the fidelity of chromosome replication and segregation is indispensable to maintaining genome stability [53]. EleniPetsalaki *et al.* found that DNA damage proteins play an essential role in mitotic cells. In contrast, DNA damage response recognizes DNA damage, coordinates cell cycle arrest to repair damaged DNA, removes affected cells, prevents genetic changes from being transmitted to the next generation, and maintains genomic stability [54]. The Wnt signalling pathway is critically involved in the development and homeostasis of tissues via regulating their endogenous stem cells. Aberrant Wnt signalling has been described as a critical player in the initiation of and maintenance, and development of many cancers via affecting the behaviour of Cancer Stem Cells (CSCs) [55]. We further investigated whether genomic instability-associated lncRNA could predict clinical outcomes and generated lncRNA signatures consisting of four genomic instability-associated lncRNA (AC078880.3, AL355916.1, AC037198.1 and LINC02446). The GILncSig divided patients into two risk groups, and there was a significant difference in survival rate in the training set, which was verified in an independent test set. Furthermore, the GILncSig is significantly associated with the phenotype of tumour mutants and the expression of UBQLN4 in bladder cancer, both of which are essential indicators of genomic instability. After a careful literature search, we found that the biological functions of this four lncRNA in GILncSig have not been reported. However, we found that lncRNA; AL355916.1 is located on the 14q21 chromosome. It has been known in previous genome-wide association analyses [56] that this region may be associated with a reduced risk of bladder cancer. Another lncRNARP11-358L4.1 is located in the 15q25 region of the chromosome, which is reported to be a susceptible region for bladder cancer risk [57]. These results suggest that GILncSig can predict the prognosis of cancer patients and be used as an indicator of genomic instability in tumour patients. Also, the TP53 mutation rate in the high-risk group of GILncSig was significantly higher than that in the low-risk group, suggesting that the mutation state of GILncSig was consistent with that of TP53.

Our study has constructed new biological markers of bladder cancer and provided a vital research basis for better evaluating genomic instability and prognosis in patients with bladder cancer. However, there are still some limitations that need to be further studied in the future. Although GILncSig has been validated in TCGA datasets, more independent datasets and clinical trials are needed to verify GILncSig, to ensure its reliability and repeatability.

5. Conclusion

Our study proposes a computational framework based on the mutation rate of bladder cancer mutants to identify lncRNA related to genomic instability. This

framework provides a theoretical basis for our later study of lncRNA related to genomic instability of bladder cancer and provides a meaningful way and method for exploring the biological function of lncRNA. By combining bladder cancer gene expression profile, somatic mutation profile, and clinical information, we identified lncRNA signal derived from genomic instability as an independent prognostic marker for risk subgroup stratification of bladder cancer patients, which was successfully verified in an independent cohort of patients. Through further prospective verification, GILncSig may play an important role in genomic instability and personalized decision-making in patients with bladder cancer.

Ethics Approval and Consent to Participate

The data used in our study were obtained from public databases TCGA, therefore, ethical approval was not required.

Acknowledgements

We would like to thank the teachers of the Research Center of the First Affiliated Hospital of Yangtze University and Yangtze University School of Medicine for their help and the patients and researchers who participated in TCGA for providing the data.

Funding

This study was supported by the Youth Talents of the Hubei Provincial Health Council (Grant No. WJ2021Q014, WJ2017Q041).

Conflicts of Interest

The authors have declared that they have no competing interests. The sources that funded this study played no role in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

References

- [1] John, B.A. and Said, N. (2017) Insights from Animal Models of Bladder Cancer: Recent Advances, Challenges, and Opportunities. *Oncotarget*, **8**, 57766-57781. <https://doi.org/10.18632/oncotarget.17714>
- [2] Jazzar, U., *et al.* (2018) Impact of Psychiatric Illness on Decreased Survival in Elderly Patients with Bladder Cancer in the United States. *Cancer*, **124**, 3127-3135. <https://doi.org/10.1002/cncr.31404>
- [3] Şanlı, Ö. and Lotan, Y. (2017) Alternative Therapies in Patients with Non-Muscle Invasive Bladder Cancer. *Turkish Journal of Urology*, **43**, 414-424. <https://doi.org/10.5152/tud.2017.64624>
- [4] Snell, K.I.E., *et al.* (2018) Exploring the Roles of Urinary HAI-1, EpCAM & EGFR in Bladder Cancer Prognosis & Risk Stratification. *Oncotarget*, **9**, 25244-25253. <https://doi.org/10.18632/oncotarget.25397>
- [5] Grayson, M. (2017) Bladder Cancer. *Nature*, **551**, S33.

- <https://doi.org/10.1038/551S33a>
- [6] Andor, N., Maley, C.C. and Ji, H.P. (2017) Genomic Instability in Cancer: Teetering on the Limit of Tolerance. *Cancer Research*, **77**, 2179-2185. <https://doi.org/10.1158/0008-5472.CAN-16-1553>
- [7] Duijf, P.H.G., et al. (2019) Mechanisms of Genomic Instability in Breast Cancer. *Trends in Molecular Medicine*, **25**, 595-611. <https://doi.org/10.1016/j.molmed.2019.04.004>
- [8] Ben-David, U., Beroukhim, R. and Golub, T.R. (2019) Genomic Evolution of Cancer Models: Perils and Opportunities. *Nature Reviews Cancer*, **19**, 97-109. <https://doi.org/10.1038/s41568-018-0095-3>
- [9] Tam, A.S., et al. (2019) Selective Defects in Gene Expression Control Genome Instability in Yeast Splicing Mutants. *Molecular Biology of the Cell*, **30**, 191-200. <https://doi.org/10.1091/mbc.E18-07-0439>
- [10] Boysen, G., et al. (2015) SPOP Mutation Leads to Genomic Instability in Prostate Cancer. *Elife*, **4**, e09207. <https://doi.org/10.7554/eLife.09207>
- [11] Biermann, J., et al. (2020) A 17-Marker Panel for Global Genomic Instability in Breast Cancer. *Genomics*, **112**, 1151-1161. <https://doi.org/10.1016/j.ygeno.2019.06.029>
- [12] Choudhari, R., et al. (2020) Long Noncoding RNAs in Cancer: From Discovery to Therapeutic Targets. *Advances in Clinical Chemistry*, **95**, 105-147. <https://doi.org/10.1016/bs.acc.2019.08.003>
- [13] Kopp, F. and Mendell, J.T. (2018) Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell*, **172**, 393-407. <https://doi.org/10.1016/j.cell.2018.01.011>
- [14] St Laurent, G., Wahlestedt, C. and Kapranov, P. (2015) The Landscape of Long Noncoding RNA Classification. *Trends in Genetics*, **31**, 239-251. <https://doi.org/10.1016/j.tig.2015.03.007>
- [15] Cao, H.L., et al. (2019) lncRNA-RMRP Promotes Proliferation, Migration and Invasion of Bladder Cancer via miR-206. *European Review for Medical and Pharmacological Sciences*, **23**, 1012-1021.
- [16] Sun, W., Shen, N.M. and Fu, S.L. (2019) Involvement of lncRNA-Mediated Signaling Pathway in the Development of Cervical Cancer. *European Review for Medical and Pharmacological Sciences*, **23**, 3672-3687.
- [17] Zhao, Q., Cheng, Y. and Xiong, Y. (2021) LTF Regulates the Immune Microenvironment of Prostate Cancer through JAK/STAT3 Pathway. *Frontiers in Oncology*, **11**, Article ID: 692117. <https://doi.org/10.3389/fonc.2021.692117>
- [18] Wang, J., et al. (2018) lncRNA HOXA-AS2 and Its Molecular Mechanisms in Human Cancer. *Clinica Chimica Acta*, **485**, 229-233. <https://doi.org/10.1016/j.cca.2018.07.004>
- [19] Zhao, W., et al. (2018) lncRNA HOTAIR Influences Cell Growth, Migration, Invasion, and Apoptosis via the miR-20a-5p/HMGA2 Axis in Breast Cancer. *Cancer Medicine*, **7**, 842-855. <https://doi.org/10.1002/cam4.1353>
- [20] Liu, H. (2016) Linking lncRNA to Genomic Stability. *Science China Life Sciences*, **59**, 328-329. <https://doi.org/10.1007/s11427-016-5009-6>
- [21] Tracy, K.M., et al. (2018) Mitotically-Associated lncRNA (MANCR) Affects Genomic Stability and Cell Division in Aggressive Breast Cancer. *Molecular Cancer Research*, **16**, 587-598. <https://doi.org/10.1158/1541-7786.MCR-17-0548>
- [22] Zhang, C. and Peng, G. (2015) Non-Coding RNAs: An Emerging Player in DNA

- Damage Response. *Mutation Research. Reviews in Mutation Research*, **763**, 202-211. <https://doi.org/10.1016/j.mrrev.2014.11.003>
- [23] Yu, G., et al. (2015) DOSE: An R/Bioconductor Package for Disease Ontology Semantic and Enrichment Analysis. *Bioinformatics*, **31**, 608-609. <https://doi.org/10.1093/bioinformatics/btu684>
- [24] Cao, R., et al. (2020) Immune-Related Long Non-Coding RNA Signature Identified Prognosis and Immunotherapeutic Efficiency in Bladder Cancer (BLCA). *Cancer Cell International*, **20**, 276. <https://doi.org/10.1186/s12935-020-01362-0>
- [25] Lian, P., et al. (2019) An Eight-Long Non-Coding RNA Signature as a Candidate Prognostic Biomarker for Bladder Cancer. *Aging*, **11**, 6930-6940. <https://doi.org/10.18632/aging.102225>
- [26] Zhang, X., et al. (2020) A Prognostic Index Based on a Fourteen Long Non-Coding RNA Signature to Predict the Recurrence-Free Survival for Muscle-Invasive Bladder Cancer Patients. *BMC Medical Informatics and Decision Making*, **20**, Article No. 136. <https://doi.org/10.1186/s12911-020-1115-2>
- [27] Ou, Z., et al. (2020) Detection of Bladder Cancer Using Urinary Cell-Free DNA and Cellular DNA. *Clinical and Translational Medicine*, **9**, 4. <https://doi.org/10.1186/s40169-020-0257-2>
- [28] Ciccarese, C., et al. (2017) Tp53 and Its Potential Therapeutic Role as a Target in Bladder Cancer. *Expert Opinion on Therapeutic Targets*, **21**, 401-414. <https://doi.org/10.1080/14728222.2017.1297798>
- [29] Almeida, T.C., et al. (2019) Antiproliferative and Toxicogenomic Effects of Resveratrol in Bladder Cancer Cells with Different TP53 Status. *Environmental and Molecular Mutagenesis*, **60**, 740-751. <https://doi.org/10.1002/em.22297>
- [30] Lyu, Q., et al. (2020) Alterations in TP53 Are a Potential Biomarker of Bladder Cancer Patients Who Benefit from Immune Checkpoint Inhibition. *Cancer Control Journal of the Moffitt Cancer Center*, **27**, Paper No. 1073274820976665. <https://doi.org/10.1177/1073274820976665>
- [31] Patel, V.G., Oh, W.K. and Galsky, M.D. (2020) Treatment of Muscle-Invasive and Advanced Bladder Cancer in 2020. *CA: A Cancer Journal for Clinicians*, **70**, 404-423. <https://doi.org/10.3322/caac.21631>
- [32] Guo, C.C. and Czerniak, B. (2019) Bladder Cancer in the Genomic Era. *Archives of Pathology & Laboratory Medicine*, **143**, 695-704. <https://doi.org/10.5858/arpa.2018-0329-RA>
- [33] Hayashi, T., et al. (2020) Mutational Landscape and Environmental Effects in Bladder Cancer. *International Journal of Molecular Sciences*, **21**, E6072. <https://doi.org/10.3390/ijms21176072>
- [34] Grossman, H.B., et al. (2016) Innovation in Bladder Cancer Immunotherapy. *Journal of Immunotherapy*, **39**, 291-297. <https://doi.org/10.1097/CJI.0000000000000130>
- [35] Humphrey, P.A., et al. (2016) The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs Part B: Prostate and Bladder Tumours. *European Urology*, **70**, 106-119. <https://doi.org/10.1016/j.eururo.2016.02.028>
- [36] Wang, G. and McKenney, J.K. (2019) Urinary Bladder Pathology: World Health Organization Classification and American Joint Committee on Cancer Staging Update. *Archives of Pathology & Laboratory Medicine*, **143**, 571-577. <https://doi.org/10.5858/arpa.2017-0539-RA>
- [37] Alifrangis, C., et al. (2019) Molecular and Histopathology Directed Therapy for Advanced Bladder Cancer. *Nature Reviews. Urology*, **16**, 465-483.

- <https://doi.org/10.1038/s41585-019-0208-0>
- [38] Tan, T.Z., et al. (2019) Molecular Subtypes of Urothelial Bladder Cancer: Results from a Meta-Cohort Analysis of 2411 Tumors. *European Urology*, **75**, 423-432. <https://doi.org/10.1016/j.eururo.2018.08.027>
- [39] Minoli, M., et al. (2020) Evolution of Urothelial Bladder Cancer in the Context of Molecular Classifications. *International Journal of Molecular Sciences*, **21**, 5670. <https://doi.org/10.3390/ijms21165670>
- [40] Matulay, J.T., Narayan, V.M. and Kamat, A.M. (2019) Clinical and Genomic Considerations for Variant Histology in Bladder Cancer. *Current Oncology Reports*, **21**, 23. <https://doi.org/10.1007/s11912-019-0772-8>
- [41] Bartek, J., Bartkova, J. and Lukas, J. (2007) DNA Damage Signalling Guards against Activated Oncogenes and Tumour Progression. *Oncogene*, **26**, 7773-7779. <https://doi.org/10.1038/sj.onc.1210881>
- [42] Bakhoun, S.F., et al. (2018) Chromosomal Instability Drives Metastasis through a Cytosolic DNA Response. *Nature*, **553**, 467-472. <https://doi.org/10.1038/nature25432>
- [43] Petropoulos, M., et al. (2019) Replication Licensing Aberrations, Replication Stress, and Genomic Instability. *Trends in Biochemical Sciences*, **44**, 752-764. <https://doi.org/10.1016/j.tibs.2019.03.011>
- [44] Vacher, S., et al. (2020) Genomic Instability Signature of Palindromic Non-Coding Somatic Mutations in Bladder Cancer. *Cancers (Basel)*, **12**, 2882. <https://doi.org/10.3390/cancers12102882>
- [45] Ferguson, L.R., et al. (2015) Genomic Instability in Human Cancer: Molecular Insights and Opportunities for Therapeutic Attack and Prevention through Diet and Nutrition. *Seminars in Cancer Biology*, **35**, S5-S24. <https://doi.org/10.1016/j.semcancer.2015.03.005>
- [46] (2019) Hypoxic Tumors Share Genomic Instability. *Cancer Discovery*, **9**, 314. <https://doi.org/10.1158/2159-8290.CD-NB2019-012>
- [47] Lee, J.-H., et al. (2015) MicroRNA-22 Suppresses DNA Repair and Promotes Genomic Instability through Targeting of MDC1. *Cancer Research*, **75**, 1298-1310. <https://doi.org/10.1158/0008-5472.CAN-14-2783>
- [48] Bhan, A., Soleimani, M. and Mandal, S.S. (2017) Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Research*, **77**, 3965-3981. <https://doi.org/10.1158/0008-5472.CAN-16-2634>
- [49] Rathinasamy, B. and Velmurugan, B.K. (2018) Role of lncRNAs in the Cancer Development and Progression and Their Regulation by Various Phytochemicals. *Bio-medicine & Pharmacotherapy*, **102**, 242-248. <https://doi.org/10.1016/j.biopha.2018.03.077>
- [50] Dastmalchi, N., Safaralizadeh, R. and Nargesi, M.M. (2020) LncRNAs: Potential Novel Prognostic and Diagnostic Biomarkers in Colorectal Cancer. *Current Medicinal Chemistry*, **27**, 5067-5077. <https://doi.org/10.2174/0929867326666190227230024>
- [51] Amodio, N., et al. (2018) MALAT1: A Druggable Long Non-Coding RNA for Targeted Anti-Cancer Approaches. *Journal of Hematology & Oncology*, **11**, Article No. 63. <https://doi.org/10.1186/s13045-018-0606-4>
- [52] Idogawa, M., et al. (2017) Long Non-Coding RNA NEAT1 Is a Transcriptional Target of p53 and Modulates p53-Induced Transactivation and Tumor-Suppressor Function. *International Journal of Cancer*, **140**, 2785-2791.

-
- <https://doi.org/10.1002/ijc.30689>
- [53] Ovejero, S., Bueno, A. and Sacristán, M.P. (2020) Working on Genomic Stability: From the S-Phase to Mitosis. *Genes (Basel)*, **11**, 225.
<https://doi.org/10.3390/genes11020225>
- [54] Petsalaki, E. and Zachos, G. (2020) DNA Damage Response Proteins Regulating Mitotic Cell Division: Double Agents Preserving Genome Stability. *The FEBS Journal*, **287**, 1700-1721. <https://doi.org/10.1111/febs.15240>
- [55] Duchartre, Y., Kim, Y.M. and Kahn, M. (2016) The Wnt Signaling Pathway in Cancer. *Critical Reviews in Oncology/Hematology*, **99**, 141-149.
<https://doi.org/10.1016/j.critrevonc.2015.12.005>
- [56] Gu, J., *et al.* (2011) A Genome-Wide Association Study Identifies a Locus on Chromosome 14q21 as a Predictor of Leukocyte Telomere Length and as a Marker of Susceptibility for Bladder Cancer. *Cancer Prevention Research (Philadelphia, Pa.)*, **4**, 514-521. <https://doi.org/10.1158/1940-6207.CAPR-11-0063>
- [57] Gago-Dominguez, M., *et al.* (2011) Genetic Variations on Chromosomes 5p15 and 15q25 and Bladder Cancer Risk: Findings from the Los Angeles-Shanghai Bladder Case-Control Study. *Carcinogenesis*, **32**, 197-202.
<https://doi.org/10.1093/carcin/bgq233>