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

Qi Zhao

Changjiang University，Jingzhou City, Hebei Province 434000

*Abstract:*Background: Studies have shown that long non-coding RNA (LncRNA) plays a critical role in maintaining genomic instability. As a new tumor marker, the correlation between lncRNA and genomic instability is still worth exploring in bladder cancer.

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 plays an important role in evaluating clinical prognosis. To sum up, this study provides an important research basis and methods for further exploring the role of lncRNA in genomic instability of bladder cancer, and provides a theoretical basis for the identification of bladder cancer biomarkers related to genomic instability.

Keywords: Bladder Cancer; Genome Instability; Long Non-Coding RNAs; Biomarker

Introduction

Bladder cancer is a common malignant tumor of the urinary system, and its morbidity and mortality are on the rise 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 main indicators to predict the progression or recurrence of bladder cancer are tumor grade, lymph node metastasis, p53/pRb signal pathway, receptor tyrosine kinase (Receptortyroinekinase, RTK) pathway, and so on. Moreover, bladder cancer is a rather complex disease with molecular and clinical heterogeneity, as well as its development, progression, and therapeutic response [5]. Therefore, there is an urgent need to find new biomarkers to more accurately evaluate the clinical prognosis of patients with bladder cancer.

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 tumor progression and survival rate [7, 8]. Although the mechanism of genomic instability is not fully understood, there is 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 tumors 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 transcript that is widely defined as larger than 200nt and has little potential to encode a protein [12]. LncRNA plays an important role in different biological processes [13,14], especially the abnormal expression of some lncRNA may affect the proliferation, progression, or metastasis of tumor cells [15, 16]. With the development of bioinformatics technology, it is possible to study the expression profile of lncRNA in tumors 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 tumors, but the function of this lncRNA is still unclear [17,18]. In the past few years, more and more studies have shown that lncRNA also plays a vital role in maintaining genomic instability [19, 20]. For example, Chunzhi Zhang et al. [21] showed that through extensive crosstalk between non-coding RNA and typical DNA damage response (DDR) signal pathway, DDR-induced non-coding RNA expression can provide a precise regulation mechanism of DNA damage response gene expression in time and space. Kirsten M.T et al. [20] found a new LncRNA, MANCR (mitosis-related long non-coding RNA; LINC00704), which is 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 the increase of DNA damage. These shreds of evidence suggest that lncRNA is associated with genomic instability, but the association between this genomic instability-related lncRNA and the clinical prognosis of tumor 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, and to explore the possibility of lncRNA signal as an indicator of genomic instability, to improve the prognostic value of bladder cancer patients.

Materials & Methods

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/). A total of 433 RNA-seq expression data samples were collected, including 19 normal samples and 414 tumor samples, and 412 patients with clinical features and somatic mutations were collected.

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, as shown in figure 1: 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 gnomically 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.

Statistical analysis

We use Euclidean distances and Ward's linkage method for Hierarchical cluster analysis. The relationship between the expression level of genome instability-associated lncRNA 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:

GILncSig(samples)= (Coef value * Expression quantity) lncRNA1+(Coef value * Expression quantity) lncRNA2+………+ (Coef value * Expression quantity) lncRNAn

GILncSig (sample) is the prognostic risk score of patients with Bladder cancer patient. The lncRNAn 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, and the difference of survival rate between the high-risk group and low-risk group was evaluated by the log-rank test. Multivariate Cox regression and stratified analysis were used to evaluate the independence of GILncSig and other key 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 are carried out using Rversion3.6.2.

Functional enrichment analysis

We calculated the Pearson correlation coefficient to measure the correlation between lncRNAs and mRNAs paired expressions, and the first 10 mRNAs were considered to be 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 by using the cluster profile software in R-version 3.6.2 [22].

Results

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 greater 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 (Supplementary Table 1). We used 58 differentially expressed lncRNA sets to perform unsupervised hierarchical clustering analysis on 412 samples of TCGA sets. As shown in figure 2A, 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. The median value of somatic cumulative 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 2B). 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. As shown in figure 2C, 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).

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, which is the top 10 PCG with the strongest correlation 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 2D). 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 (Fig. 2E). Through the analysis of KEGG pathways of LncRNA-related PCG, 12 significantly rich pathways were found, including cytokine-cytokine receptor interaction pathways, calcium signaling pathways, Wnt signaling pathways, and pathways related to carcinogenesis (Fig. 2E).

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 two groups: the training set and the testing set. 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 8 genomic instability-related lncRNAs were significantly associated with the prognosis of bladder cancer patients (P < 0.05; Figure 3A). Also, we hope to screen the LncRNA, with the independent prognostic value from the 8 candidate LncRNA and then conduct multivariate Cox proportional hazard regression analysis on the 8 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; Table 1). 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×expression level of AC078880.3) +(0.0472×expression level of AL355916.1) +(0.0988×expression level of AC037198.1) +(-0.2257×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, while AC078880.3 and LINC02446 may be protective factors, and its high expression is related to prolonged survival.

To test whether the clinical characteristic data of the above training set and the test set are different, we make statistics on the clinical characteristic data of the two groups, including age, gender, grade, stage, and TNM stages. Chi-square test analysis shows that there is no difference between the two groups of clinical characteristics that can be used as a basis for grouping (Supplementary Table 2).

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 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; Figure 3B). 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) (Figure 3 C, D). We classified the patients in the training set according to their scores and 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 (figure 3E). 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 that there were significant differences in somatic mutation pattern and UBQLN4 expression pattern between the high-risk group and low-risk group. As shown in figure 3F, 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; Figure 3F).

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, and there was a significant difference in overall survival. As shown in figure 4A, 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 Figure 1A). The expression of GILncSig, the number of somatic mutation counts, and the expression of UBQLN4 in the test samples are shown in figure 4B. 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; Figure 4C). 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; Figure 4D). The expression of the GILncSig and the distribution of somatic mutation count and UBQLN4 expression in the TCGA samples were illustrated in Figure 4E. The time-dependent ROC curves analysis was applied to the TCGA set, and consistent results were observed as above (AUC = 0.712 (3 years); Supplementary Figure 1B). There was a significant difference in the distribution of somatic mutations between the high-risk group and the low-risk group. (P=1.3e-07, Mann–Whitney U test; Figure 4F).

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, in each age group can 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 young-patient group (log-rank test P<0.001; Figure 5A) as was in the old-patient group (log-rank test P<0.001; Figure 5B). Next, all the patients with bladder cancer were stratified according to the pathological stage, the patients with pathological stage I and II were combined into the early group, and the patients with pathological stage III and IV were combined into the late group. 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; Figure 5C). Similarly, GILncSig could also be used to divide pathological stage III and IV patients into a high-risk group and low-risk group. There was a significant difference in the overall survival rate between the two groups. (log-rank test P<0.001; Figure 5D). 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 group in the female group (log-rank test p < 0.001; Figure 5E), the male group also had a significant difference (log-rank test p < 0.001; Figure 5F). 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 group and the low-risk group in the High-Grade group (logarithmic rank test p < 0.001; figure 5G), 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 (Table 2). In multivariate analysis, in addition to GILncSig, two other clinical factors, age, and stage were observed to be significant.

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) [23] and the 8-lncRNA signature obtained from Lian's research (hereinafter referred to as LianlncSig) [24], and the 14-lncRNA signature obtained from Zhang's research (hereinafter referred to as ZhanglncSig) [25]). Using the same TCGA patient cohort. As shown in Figure5a, the AUC at 3 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) (Figure 6A). Besides, the AUC at 5 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) (Figure 6B). 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.

The GILncSig predicts the results better than 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 (figure 7). 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 [26-29]. It is suggested that GILncSig is also related to the mutation state of TP53, which may be the mutation marker of the TP53 gene.

Discussion

With the increasing incidence of bladder cancer in recent years, researches on bladder cancer tend to be diversified [30-33]. At present, there is still no breakthrough. At the current stage, patients are divided into different treatment groups according to their pathological characteristics, while histopathological features such as staging, grading, and molecular subtypes of bladder tumor are still the most important factors affecting the prognosis of patients [34-38]. However, due to the limitation of traditional clinicopathological features, the clinical prognosis of patients with bladder cancer is still highly heterogeneous [39]. Studies have shown that genomic instability is not only a common feature of most cancers [40-42] but also one of the factors affecting the prognosis of bladder cancer [43]. Genomic instability plays a leading role in the occurrence, development, and recurrence of cancer. 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, quantitative measurement of genomic instability has always been a major challenge [46]. Efforts have been made to identify PCGs and microRNAs associated with genomic instability and to develop genes or miRNA features that predict genomic instability [11,47,48].

Currently, lncRNA is an important part of tumor biology, and their abnormal expression in cancer is associated with disease progression and may serve as prognostic markers for patients [56-58]. Furthermore, studies on the functional mechanism of lncRNA have found that lncRNA is also essential for genomic stability, such as MALAT1 [52] and NEAT1 [53]. Although a great deal of 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 main sources of genomic instability: DNA damage, replication pressure, and chromosome segregation defects. Among them, the regulation of chromosome mitosis is very important to the cell cycle, and the fidelity of chromosome replication and segregation is indispensable to maintain genome stability [54]. EleniPetsalaki et al found that DNA damage proteins play an important role in mitotic cells, while DNA damage response recognizes DNA damage, coordinates cell cycle arrest to repair damaged DNA, or remove affected cells, to prevent genetic changes from being transmitted to the next generation and maintain genomic stability [55]. The Wnt signaling pathway is critically involved in both the development and homeostasis of tissues via the regulation of their endogenous stem cells. Aberrant Wnt signaling has been described as a key player in the initiation of and/or maintenance and development of many cancers, via affecting the behavior of Cancer Stem Cells (CSCs) [56]. 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 tumor mutants and the expression of UBQLN4 in bladder cancer, both of which are important 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, and it has been known in previous genome-wide association analysis [57] 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 [58]. These results suggest that GILncSig can not only predict the prognosis of cancer patients but also be used as an indicator of genomic instability in tumor 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 an important research basis for better evaluation of 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.

Conclusions

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 also provides an important 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.

References

1. John, B.A. and N. Said, Insights from animal models of bladder cancer: recent advances, challenges, and opportunities. Oncotarget, 2017. 8(34): p. 57766-57781.

2. Jazzar, U., et al., Impact of psychiatric illness on decreased survival in elderly patients with bladder cancer in the United States. Cancer, 2018. 124(15): p. 3127-3135.

3. Şanlı, Ö. and Y. Lotan, Alternative therapies in patients with non-muscle invasive bladder cancer. Turk J Urol, 2017. 43(4): p. 414-424.

4. Snell, K.I.E., et al., Exploring the roles of urinary HAI-1, EpCAM & EGFR in bladder cancer prognosis & risk stratification. Oncotarget, 2018. 9(38): p. 25244-25253.

5. Grayson, M., Bladder cancer. Nature, 2017. 551(7679): p. S33.

6. Andor, N., C.C. Maley, and H.P. Ji, Genomic Instability in Cancer: Teetering on the Limit of Tolerance. Cancer Res, 2017. 77(9): p. 2179-2185.

7. Duijf, P.H.G., et al., Mechanisms of Genomic Instability in Breast Cancer. Trends Mol Med, 2019. 25(7): p. 595-611.

8. Ben-David, U., R. Beroukhim, and T.R. Golub, Genomic evolution of cancer models: perils and opportunities. Nat Rev Cancer, 2019. 19(2): p. 97-109.

9. Tam, A.S., et al., Selective defects in gene expression control genome instability in yeast splicing mutants. Mol Biol Cell, 2019. 30(2): p. 191-200.

10. Boysen, G., et al., SPOP mutation leads to genomic instability in prostate cancer. Elife, 2015. 4.

11. Biermann, J., et al., A 17-marker panel for global genomic instability in breast cancer. Genomics, 2020. 112(2): p. 1151-1161.

12. Choudhari, R., et al., Long noncoding RNAs in cancer: From discovery to therapeutic targets. Adv Clin Chem, 2020. 95: p. 105-147.

13. Kopp, F. and J.T. Mendell, Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell, 2018. 172(3): p. 393-407.

14. St Laurent, G., C. Wahlestedt, and P. Kapranov, The Landscape of long noncoding RNA classification. Trends Genet, 2015. 31(5): p. 239-51.

15. Cao, H.L., et al., lncRNA-RMRP promotes proliferation, migration and invasion of bladder cancer via miR-206. Eur Rev Med Pharmacol Sci, 2019. 23(3): p. 1012-1021.

16. Sun, W., N.M. Shen, and S.L. Fu, Involvement of lncRNA-mediated signaling pathway in the development of cervical cancer. Eur Rev Med Pharmacol Sci, 2019. 23(9): p. 3672-3687.

17. Wang, J., et al., LncRNA HOXA-AS2 and its molecular mechanisms in human cancer. Clin Chim Acta, 2018. 485: p. 229-233.

18. Zhao, W., et al., LncRNA HOTAIR influences cell growth, migration, invasion, and apoptosis via the miR-20a-5p/HMGA2 axis in breast cancer. Cancer Med, 2018. 7(3): p. 842-855.

19. Liu, H., Linking lncRNA to genomic stability. Sci China Life Sci, 2016. 59(3): p. 328-9.

20. Tracy, K.M., et al., Mitotically-Associated lncRNA (MANCR) Affects Genomic Stability and Cell Division in Aggressive Breast Cancer. Mol Cancer Res, 2018. 16(4): p. 587-598.

21. Zhang, C. and G. Peng, Non-coding RNAs: an emerging player in DNA damage response. Mutation research. Reviews in mutation research, 2015. 763: p. 202-211.

22. Yu, G., et al., DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics, 2015. 31(4): p. 608-9.

23. Cao, R., et al., Immune-related long non-coding RNA signature identified prognosis and immunotherapeutic efficiency in bladder cancer (BLCA). Cancer cell international, 2020. 20: p. 276.

24. Lian, P., et al., An eight-long non-coding RNA signature as a candidate prognostic biomarker for bladder cancer. Aging, 2019. 11(17): p. 6930-6940.

25. Zhang, X., et al., 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, 2020. 20(Suppl 3): p. 136.

26. Ou, Z., et al., Detection of bladder cancer using urinary cell-free DNA and cellular DNA. Clinical and translational

medicine, 2020. 9(1): p. 4.

27. Ciccarese, C., et al., Tp53 and its potential therapeutic role as a target in bladder cancer. Expert Opin Ther Targets, 2017. 21(4): p. 401-414.

28. Almeida, T.C., et al., Antiproliferative and toxicogenomic effects of resveratrol in bladder cancer cells with different TP53 status. Environmental and molecular mutagenesis, 2019. 60(8): p. 740-751.

29. Lyu, Q., et al., 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, 2020. 27(1): p. 1073274820976665.

30. Patel, V.G., W.K. Oh, and M.D. Galsky, Treatment of muscle-invasive and advanced bladder cancer in 2020. CA: a cancer journal for clinicians, 2020. 70(5): p. 404-423.

31. Guo, C.C. and B. Czerniak, Bladder Cancer in the Genomic Era. Arch Pathol Lab Med, 2019. 143(6): p. 695-704.

32. Hayashi, T., et al., Mutational Landscape and Environmental Effects in Bladder Cancer. Int J Mol Sci, 2020. 21(17).

33. Grossman, H.B., et al., Innovation in Bladder Cancer Immunotherapy. J Immunother, 2016. 39(8): p. 291-7.

34. Humphrey, P.A., et al., The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part B: Prostate and Bladder Tumours. European urology, 2016. 70(1): p. 106-119.

35. Wang, G. and J.K. McKenney, Urinary Bladder Pathology: World Health Organization Classification and American Joint Committee on Cancer Staging Update. Archives of pathology & laboratory medicine, 2019. 143(5): p. 571-577.

36. Alifrangis, C., et al., Molecular and histopathology directed therapy for advanced bladder cancer. Nature reviews. Urology, 2019. 16(8): p. 465-483.

37. Tan, T.Z., et al., Molecular Subtypes of Urothelial Bladder Cancer: Results from a Meta-cohort Analysis of 2411 Tumors. European urology, 2019. 75(3): p. 423-432.

38. Minoli, M., et al., Evolution of Urothelial Bladder Cancer in the Context of Molecular Classifications. International journal of molecular sciences, 2020. 21(16).

39. Matulay, J.T., V.M. Narayan, and A.M. Kamat, Clinical and Genomic Considerations for Variant Histology in Bladder Cancer. Current oncology reports, 2019. 21(3): p. 23.

40. Bartek, J., J. Bartkova, and J. Lukas, DNA damage signalling guards against activated oncogenes and tumour progression. Oncogene, 2007. 26(56): p. 7773-7779.

41. Bakhoum, S.F., et al., Chromosomal instability drives metastasis through a cytosolic DNA response. Nature, 2018. 553(7689): p. 467-472.

42. Petropoulos, M., et al., Replication Licensing Aberrations, Replication Stress, and Genomic Instability. Trends in biochemical sciences, 2019. 44(9): p. 752-764.

43. Vacher, S., et al., Genomic Instability Signature of Palindromic Non-Coding Somatic Mutations in Bladder Cancer. Cancers (Basel), 2020. 12(10).

44. Yang, Y., et al., Complete response to anti-PD-L1 antibody in a metastatic bladder cancer associated with novel MSH4 mutation and microsatellite instability. Journal for immunotherapy of cancer, 2020. 8(1).

45. Biermann, J., et al., Radiation-induced genomic instability in breast carcinomas of the Swedish hemangioma cohort. Genes, chromosomes & cancer, 2019. 58(9): p. 627-635.

46. Ferguson, L.R., et al., Genomic instability in human cancer: Molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition. Seminars in cancer biology, 2015. 35 Suppl.

47. Hypoxic Tumors Share Genomic Instability. Cancer discovery, 2019. 9(3): p. 314.

48. Lee, J.-H., et al., MicroRNA-22 Suppresses DNA Repair and Promotes Genomic Instability through Targeting of MDC1. Cancer research, 2015. 75(7): p. 1298-1310.

49. Bhan, A., M. Soleimani, and S.S. Mandal, Long Noncoding RNA and Cancer: A New Paradigm. Cancer research, 2017. 77(15): p. 3965-3981.

50. Rathinasamy, B. and B.K. Velmurugan, Role of lncRNAs in the cancer development and progression and their regulation by various phytochemicals. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 2018. 102: p. 242-248.

51. Dastmalchi, N., R. Safaralizadeh, and M.M. Nargesi, LncRNAs: Potential Novel Prognostic and Diagnostic Biomarkers in Colorectal Cancer. Current medicinal chemistry, 2020. 27(30): p. 5067-5077.

52. Amodio, N., et al., MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. Journal of hematology & oncology, 2018. 11(1): p. 63.

53. Idogawa, M., et al., Long non-coding RNA NEAT1 is a transcriptional target of p53 and modulates p53-induced transactivation and tumor-suppressor function. International journal of cancer, 2017. 140(12): p. 2785-2791.

54. Ovejero, S., A. Bueno, and M.P. Sacristán, Working on Genomic Stability: From the S-Phase to Mitosis. Genes (Basel), 2020. 11(2).

55. Petsalaki, E. and G. Zachos, DNA damage response proteins regulating mitotic cell division: double agents preserving genome stability. The FEBS journal, 2020. 287(9): p. 1700-1721.

56. Duchartre, Y., Y.M. Kim, and M. Kahn, The Wnt signaling pathway in cancer. Crit Rev Oncol Hematol, 2016. 99: p. 141-9.

57. Gu, J., et al., 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.), 2011. 4(4): p. 514-521.

58. Gago-Dominguez, M., et al., Genetic variations on chromosomes 5p15 and 15q25 and bladder cancer risk: findings from the Los Angeles-Shanghai bladder case-control study. Carcinogenesis, 2011. 32(2): p. 197-202.

Introduction:

Zhao Qi (1995-01) male, Han nationality, Huanggang, Hubei Province, master's degree in reading, Changjiang University, research direction: urinary tumors,mailbox:1365857990@qq.com