Check for updates

Comprehensive Analyses of Mutation-Derived Long-Chain Noncoding RNA Signatures of Genome Instability in Kidney Renal Papillary Cell Carcinoma

Jian Li[1†], Shimei Wei[1†], Yan Zhang[2], Shuangshuang Lu[1], Xiaoxu Zhang[1], Qiong Wang[1], Jiawei Yan[1], Sanju Yang[3], Liying Chen[3], Yunguang Liu[1]* and Zhijing Huang[1]*

[1]Department of Pediatrics, Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China, [2]Department of Pediatrics, Shanxi Children's Hospital, Taiyuan, China, [3]Graduate School of Youjiang Medical University for Nationalities, Baise, China

Background: The role of long-chain noncoding RNA (lncRNA) in genomic instability has been demonstrated to be increasingly importance. Therefore, in this study, lncRNAs associated with genomic instability were identified and kidney renal papillary cell carcinoma (KIRP)-associated predictive features were analysed to classify high-risk patients and improve individualised treatment.

Methods: The training ($n = 142$) and test ($n = 144$) sets were created using raw RNA-seq and patient's clinical data of KIRP obtained from The Cancer Genome Atlas (TCGA).There are 27 long-chain noncoding RNAs (lncRNAs) that are connected with genomic instability, these lncRNAs were identified using the 'limma' R package based on the numbers of somatic mutations and lncRNA expression profiles acquired from KIRP TCGA cohort. Furthermore, Cox regression analysis was carried out to develop a genome instability-derived lncRNA-based gene signature (GILncSig), whose prognostic value was confirmed in the test cohort as well as across the entire KIRP TCGA dataset.

Results: A GILncSig derived from three lncRNAs (BOLA3-AS1, AC004870, and LINC00839), which were related with poor KIRP survival, was identified, which was split up into high- and low-risk groups. Additionally, the GILncSig was found to be an independent prognostic predictive index in KIRP using univariate and multivariate Cox analysis. Furthermore, the prognostic significance and characteristics of GilncSig were confirmed in the training test and TCGA sets. GilncSig also showed better predictive performance than other prognostic lncRNA features.

Conclusion: The function of lncRNAs in genomic instability and the genetic diversity of KIRP were elucidated in this work. Moreover, three lncRNAs were screened for prediction of the outcome of KIRP survival and novel insights into identifying cancer biomarkers related to genomic instability were discussed.

Keywords: genomic instability, KIRP, lncRNAs, mutator phenotype, prognostic

INTRODUCTION

Kidney renal papillary cell carcinoma (KIRP) accounts for approximately 15% of all renal cell carcinoma (RCC) cases, which is the second most common type of RCC after clear cell RCC (Znaor et al., 2015). KIRP, a renal parenchymal malignancy, includes two subtypes (type 1 and type 2), with type 2 having a worse prognosis. The cause of KIRP remains unclear, the loss or mutation of the von Hippel Lindau (VHL) gene is observed in roughly 85% of KIRP tumours, which is one of the inevitable initial steps in the development of KIRP. It has been reported that VHL mutation through Akt/GSK-3 β Signalling pathway mediated SALL4 overexpression to promote the occurrence and vascularization of clear cell renal cell carcinoma (Sun J. et al., 2020). In addition, chromosome abnormality is another important factor in the occurrence and development of KIRP, such as Chromosome 3p Deletion (Jonasch et al., 2021) and Xp11.2 translocation (Dong et al., 2021). At present, clinicopathological features are considered the main diagnostic criteria for KIRP; however, the prediction results are inaccurate owing to the inconsistent criteria.(Sukov et al., 2012). Recently, various studies have focussed on developing potential targets that could benefit KIRP treatment, such as foretinib; however, these medications are only effective against type 1 KIRP and not to the aggressive type 2 KIRP (Choueiri et al., 2013; Choueiri et al., 2014). Currently, a standard treatment method for KIRP is not available, and the mortality of patients with KIRP is still very high. Therefore, identifying new biomarkers and formulating effective diagnosis and treatment strategies, which can improve the survival rate of patients with KIRP, is crucial.

Genomic instability induced by imperfect mismatch repair systems is characterised by extensive mutations throughout the genome, especially in highly repetitive microsatellite regions. This suggests that genomic instability can be used as a predictor of illness outcome, with the association of mutation accumulation to tumour advancement and survival (Suzuki et al., 2003; Ottini et al., 2006). Moreover, the underlying mechanism of genomic instability requires further elucidation. A study reports the potential use of molecular features to quantitatively measure genomic instability (Tam et al., 2019). For example, a study found 10 miRNA markers related to genomic instability that were associated with ovarian cancer (OV) prognosis (Zeng et al., 2018; Zhang et al., 2019). In the past decade, high-throughput genome sequencing technology has promoted the discovery of various prognostic factors, such as long-chain noncoding RNA (lncRNA). lncRNA is a common noncoding RNA that has a length of no less than 2000 nucleotides and is involved in the development of cancer [8]. Its abnormal expression or behaviour impacts cancer development (Chi et al., 2019). For example, lncRNA *Arlnc1* promotes androgen receptor-regulated prostate cancer progression (Zhang et al., 2018), and *Malat-1* promotes tumour epithelial–mesenchymal transformation (EMT) and further promotes the progression of RCC (Zhang et al., 2015). Additionally, numerous studies have published that lncRNAs are involved in cell survival, proliferation, migration and genomic stability (Moran et al., 2012; Iyer et al., 2015; Yao et al., 2019; Zhou et al., 2019). Some genomic instability-related lncRNA indicators have recently been found to predict prognosis in cancer patients, such as lung adenocarcinoma (Peng et al., 2021) and colon cancer (Yin et al., 2021). However, the potential biological process and genomic instability related to the clinical significance of lncRNA in KIRP remain unclear.

In this study, we used The Cancer Genome Atlas (TCGA) to develop a computational framework derived from the mutational hypothesis of lncRNA expression profile and somatic mutation profile in tumour genomes, which can be used as prognostic risk models, were identified to improve patient stratification and promote personalised treatment decision-making.

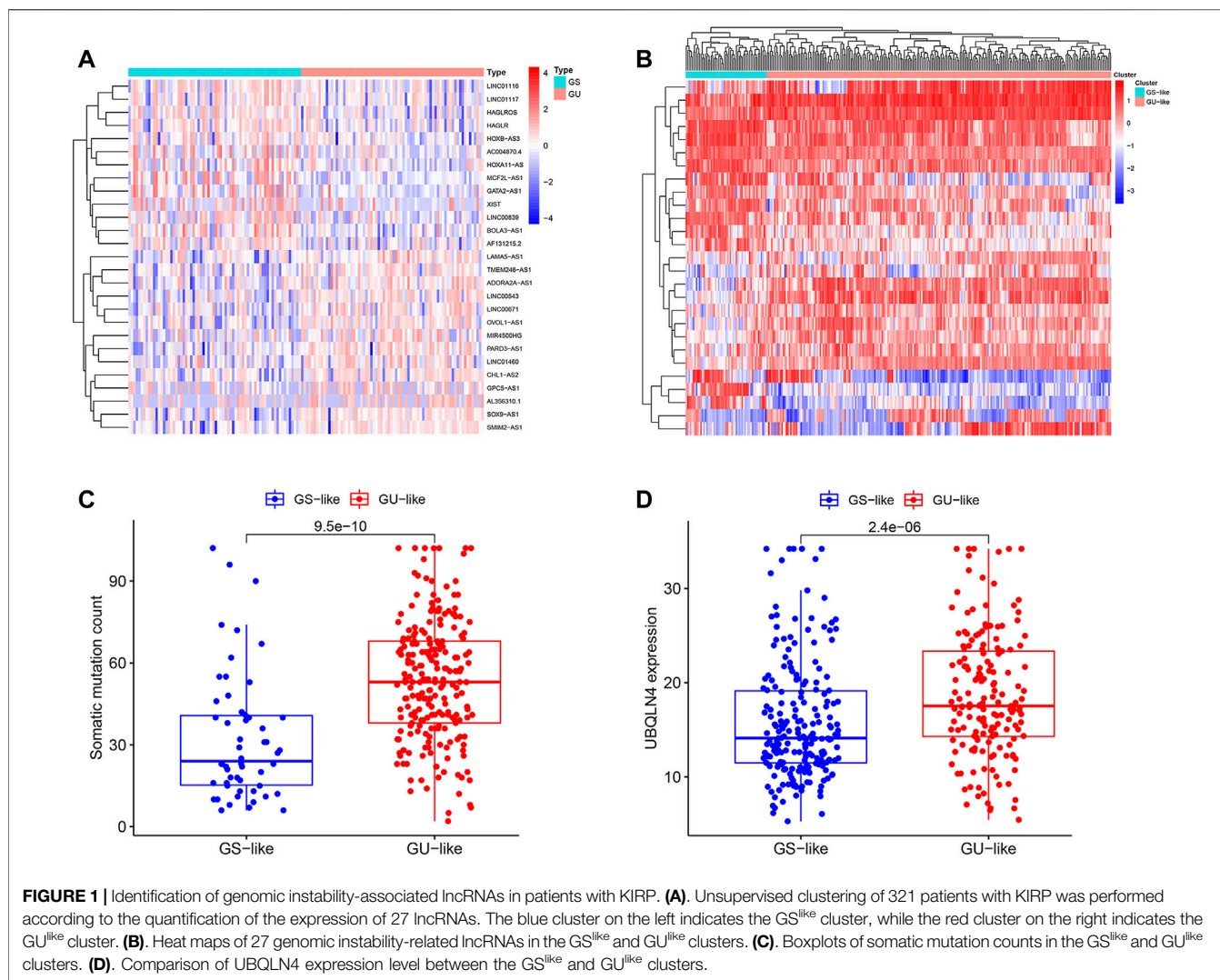# MATERIALS AND METHODS

Data Collection

Clinical characteristics, RNA-seq and somatic mutation data of patients with KIRP were downloaded from TCGA database, including the expression profiles of paired lncRNA and mRNA, somatic mutations, patient's survival information and clinicopathological characteristics. The training set comprised TCGA dataset, which included 321 patients who were divided at random into the training (144 patients) and test groups (142 patients). These groups were carried out in order to find lncRNA signatures that could be used for prognostic index for KIRP and construct prognostic risk models. Additionally, another independent KIRP cancer validation set (GSE3494) was downloaded from the GEO database, consisting of a large number of participants and Corresponding clinicopathological characteristics.

Screening of lncRANs Related to Genomic Instability

Somatic mutation information and matched lncRNA expression profiles were obtained from the KIRP-TCGA dataset in order to identify lncRNAs that are associated with genomic instability. The total number of somatic mutations in each patient was calculated using Perl. The subjects were sorted in descending order according to the number of somatic mutations. The genomic unstable (GU) group consisted of the top 25% of samples with the greatest mutation frequency, while the bottom 25% of the samples with the lowest mutation frequency was considered as the genomic stable (GS) group. The expression patterns of these two groups were compared to identify lncRNAs that are linked to genomic instability using Wilcoxon test analysis with the package of 'limma'. $p < 0.05$ and $|\log2 (FC)| > 1$ were determined to the screening thresholds.

Analyses of Hierarchical Clustering and Construction of a Co-expression Network

Depending on the levels of expression of the selected genome instability-related lncRNAs, a hierarchical clustering analysis was done using 'limma' (Liu et al., 2021) and 'sparcl' (Yin et al., 2021) packages to separate the patients into two groups: GU-like (GU^{like}) and GS-like (GS^{like}). The top 10 mRNAs that were highly linked with each lncRNA were chosen as target genes using the Pearson correlation analysis to construct the lncRNA and mRNA co-expression network.

FIGURE 1 | Identification of genomic instability-associated lncRNAs in patients with KIRP. **(A)**. Unsupervised clustering of 321 patients with KIRP was performed according to the quantification of the expression of 27 lncRNAs. The blue cluster on the left indicates the GS[like] cluster, while the red cluster on the right indicates the GU[like] cluster. **(B)**. Heat maps of 27 genomic instability-related lncRNAs in the GS[like] and GU[like] clusters. **(C)**. Boxplots of somatic mutation counts in the GS[like] and GU[like] clusters. **(D)**. Comparison of UBQLN4 expression level between the GS[like] and GU[like] clusters.

Establishment of a Prognostic Signature

Survival analysis was used to assess the role of genomic instability-associated lncRNAs prognostic using the 'survival' R package. Further, univariate and multivariate Cox regression analyses were performed to investigate the relationship between lncRNA expression and clinical prognosis. Based on the expression levels of the genome instability-related prognostic lncRNAs and their associated coefficient, the genomic instability-associated lncRNA prognostic index was constructed as follows:

$$\text{GILncSig}\,(\text{riskscore}) = \sum = 1 \text{ncoef}\,(\text{lncRNAi}) * \text{expr}\,(\text{lncRNAi})$$

GILncSig is a prognostic risk score for patients with KIRP. KIRP patients were divided into the high-risk (risk[hi]) and low-risk (risk[lo]) groups in accordance with the median value of GILncSig. The Kaplan–Meier method was carried out to calculate the prognostic value of the GILncSig based on the middle of GILncSig scores, and multivariate Cox regression and hierarchical analyses were performed to determine the independence of GILncSig against other key clinicopathological features. Additionally, the receiver operating characteristic (ROC) curve (AUC) was performed to examine the prediction accuracy of GILncSig.

Functional and Pathway Enrichment Analyses

The enrichment analysis of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were carried out on the mRNA in the co-expression network using the 'clusterProfiler' package (Yu et al., 2012) with the threshold of $p < 0.05$.

RESULTS

Identification of Genomic Instability-Associated lncRNAs in KIRP Patients

Based on the cumulative somatic mutation count per patient, the top 25% ($n = 61$) and the bottom 25% ($n = 61$) of the KIRP subjects were assigned to the GU[like] and GS[like] groups. A total of

FIGURE 2 | Functional annotations of genomic instability-related lncRNAS in patients with KIRP. **(A)**.Co-expression network of mRNAs and lncRNAs associated with genomic instability. Gene Ontology **(B)** and Kyoto Encyclopedia of Genes and Genomes enrichment analysis **(C)** of the lncRNA co-expressed mRNAs.
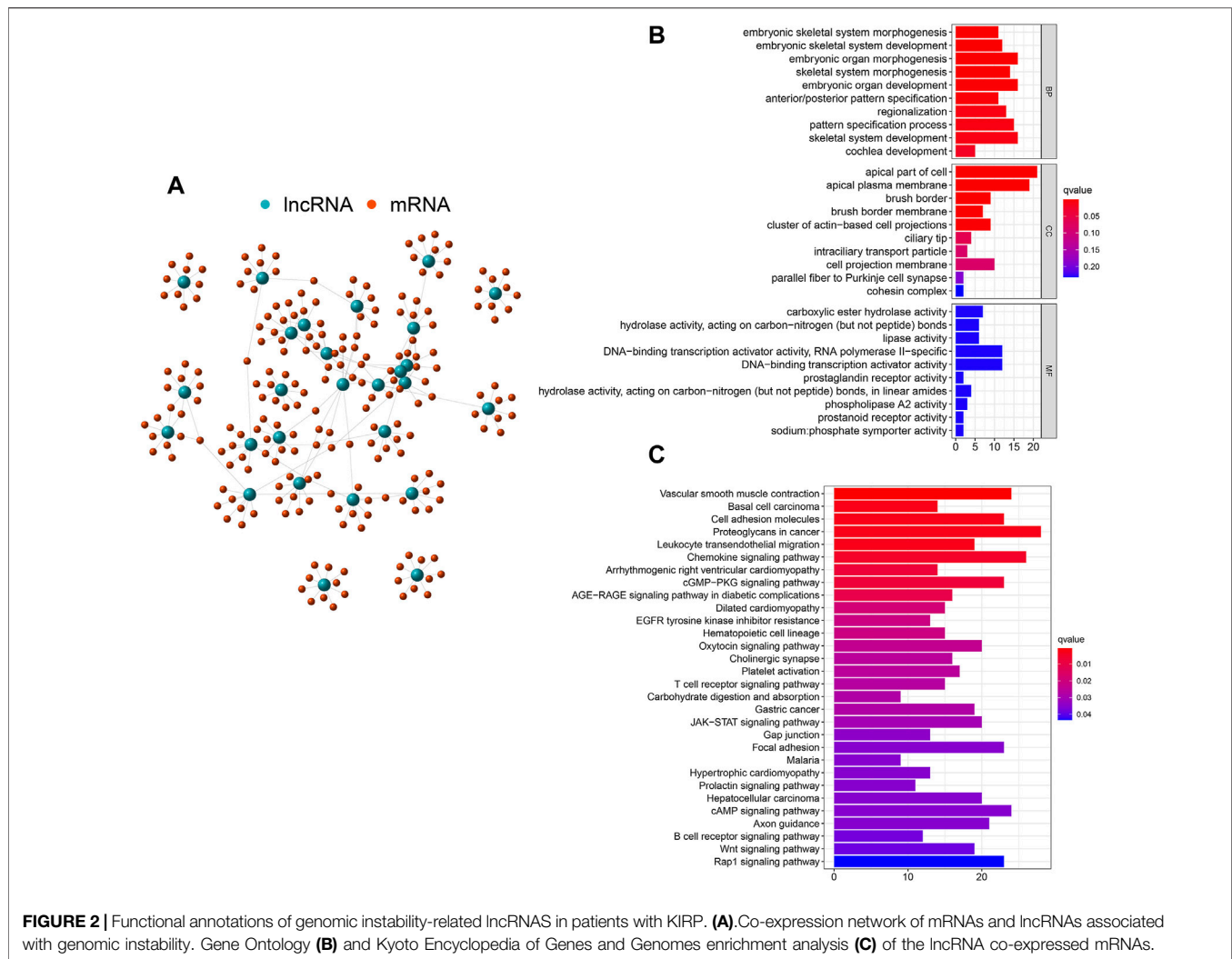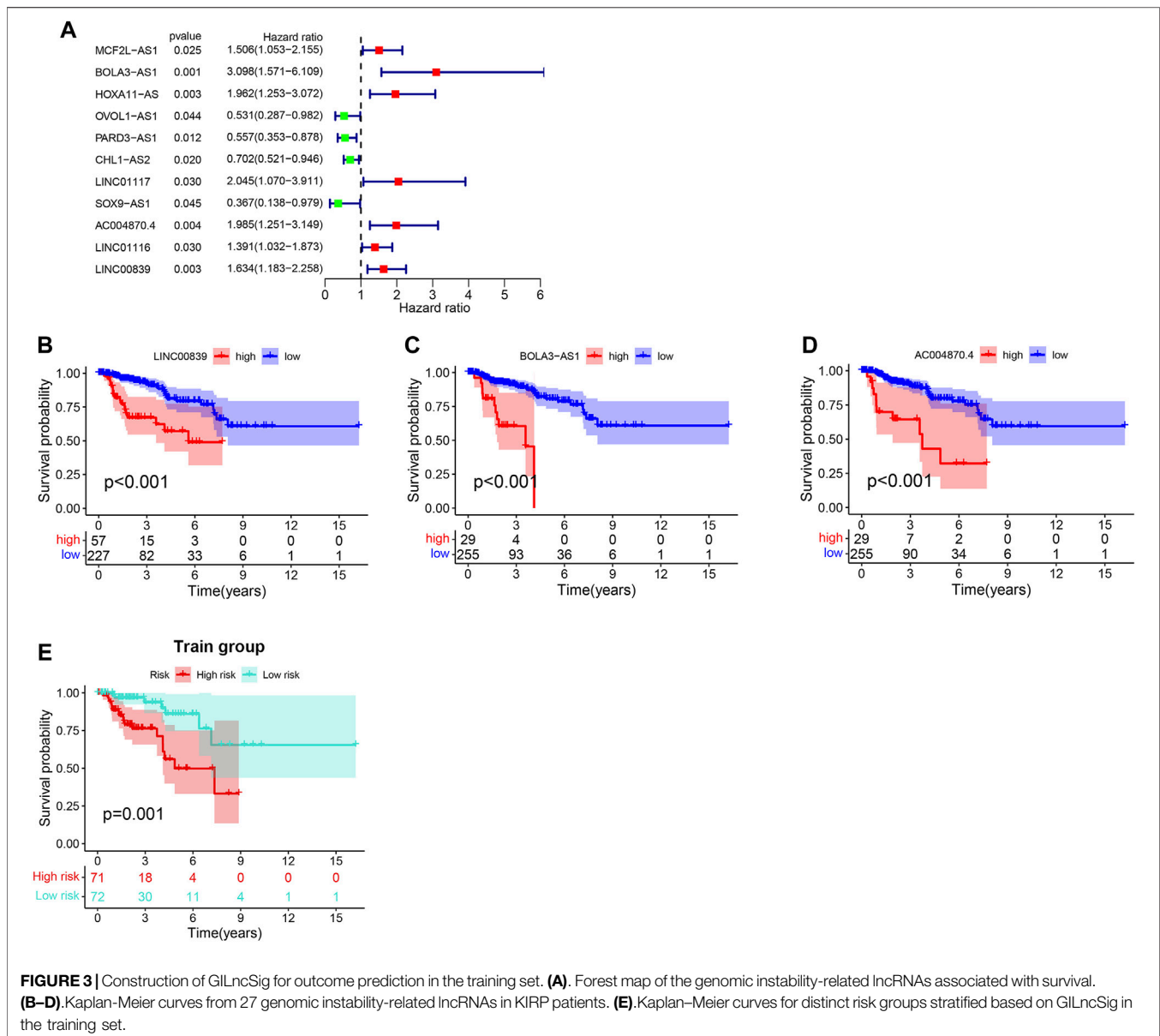
27 differentially expressed lncRNAs were determined at $p < 0.05$ and $|\log 2 (FC)| > 1$ by comparing differences in lncRNA expression, which revealed 14 up-regulated and 13 down-regulated lncRNAs in the GU$^{\text{like}}$ group (**Supplementary Table S1**). These lncRNAs that were up-regulated and down-regulated were chosen to visualize a heat map (**Figure 1A**). The expression level of 155 genomic instability-related lncRNAs was used to divide the subjects into the GU$^{\text{like}}$ or GS$^{\text{like}}$ clusters (**Figure 1B**). Moreover, compared with GS$^{\text{like}}$ clusters, the GU$^{\text{like}}$ cluster had a higher frequency of somatic mutations (**Figure 1C**). Besides, the UBQLN4 expression in GU$^{\text{like}}$ cluster group is remarkably higher than in the GS$^{\text{like}}$ clusters (**Figure 1D**). The biological functions of 146 lncRNAs related target genes were analysed using the GO and KEGG pathway enrichment analyses. The target genes were chosen from the top 10 mRNAs that were highly associated with each lncRNA. Subsequently, the lncRNA and mRNA co-expression network was built based on the correlation between the lncRNAs and mRNAs (**Figure 2A**). GO analysis revealed that (**Figure 2B**) the biological processes of the identified mRNAs were mainly enriched in embryonic organ morphogenesis, embryonic organ development and skeletal system

development, with the cell components mainly concentrated in the apical cell, apical plasma membrane and brush border. The molecular functions were mainly concentrated during skeletal system development, gastric acid secretion regulation and proximal/distal pattern formation. KEGG pathway analysis of the most identified target genes was concentrated in proteoglycans for cancer, chemokine signalling pathway and cAMP signalling pathway (**Figure 2C**).

Development of GILncSig as a Prognostic Index in the Training Set

The training ($n = 144$) and test sets ($n = 142$) were created by randomly dividing all samples received from the TCGA database. In the training set, 14 prognoses-related lncRNAs were screened from 27 genomic instability-related lncRNAs using Univariate Cox analysis, and a forest plot was drawn (**Figure 3A**). Additionally, the survival curve also revealed that the selected genes were linked to a poor prognosis in patients with KIRP, except ADORA2A-AS1, GPC5-AS1, OVOL1-AS1, RARD3-AS1, CHL1-AS2, and SOX9-AS1 (**Supplementary Figure S1**,

FIGURE 3 | Construction of GILncSig for outcome prediction in the training set. **(A)**. Forest map of the genomic instability-related lncRNAs associated with survival. **(B–D)**.Kaplan-Meier curves from 27 genomic instability-related lncRNAs in KIRP patients. **(E)**.Kaplan–Meier curves for distinct risk groups stratified based on GILncSig in the training set.

Figure 3B). The univariate and multivariate Cox proportional hazards regression analysis revealed three lncRNAs that were identified as high risk factors (HR > 1) and independent prognostic lncRNAs (**Supplementary Table S2**, **Figures 3A–D**). A GILncSig was created to assess the prognostic risk of individuals with KIRP based on the expression level of the three independent prognostic genomic instability-associated lncRNAs using multivariate Cox analysis as follows: LINC00839 expression level × 0.339 + BOLA3-AS1 expression level × 0.734 AC004870 expression level × −0.687. It was consistent with the previous analysis results, of the GILncSig, the coefficient of these three lncRNAs was positive, indicating that these lncRNAs may be a risky factor because their high expression was associated with a poor prognosis. When the median risk score for the training
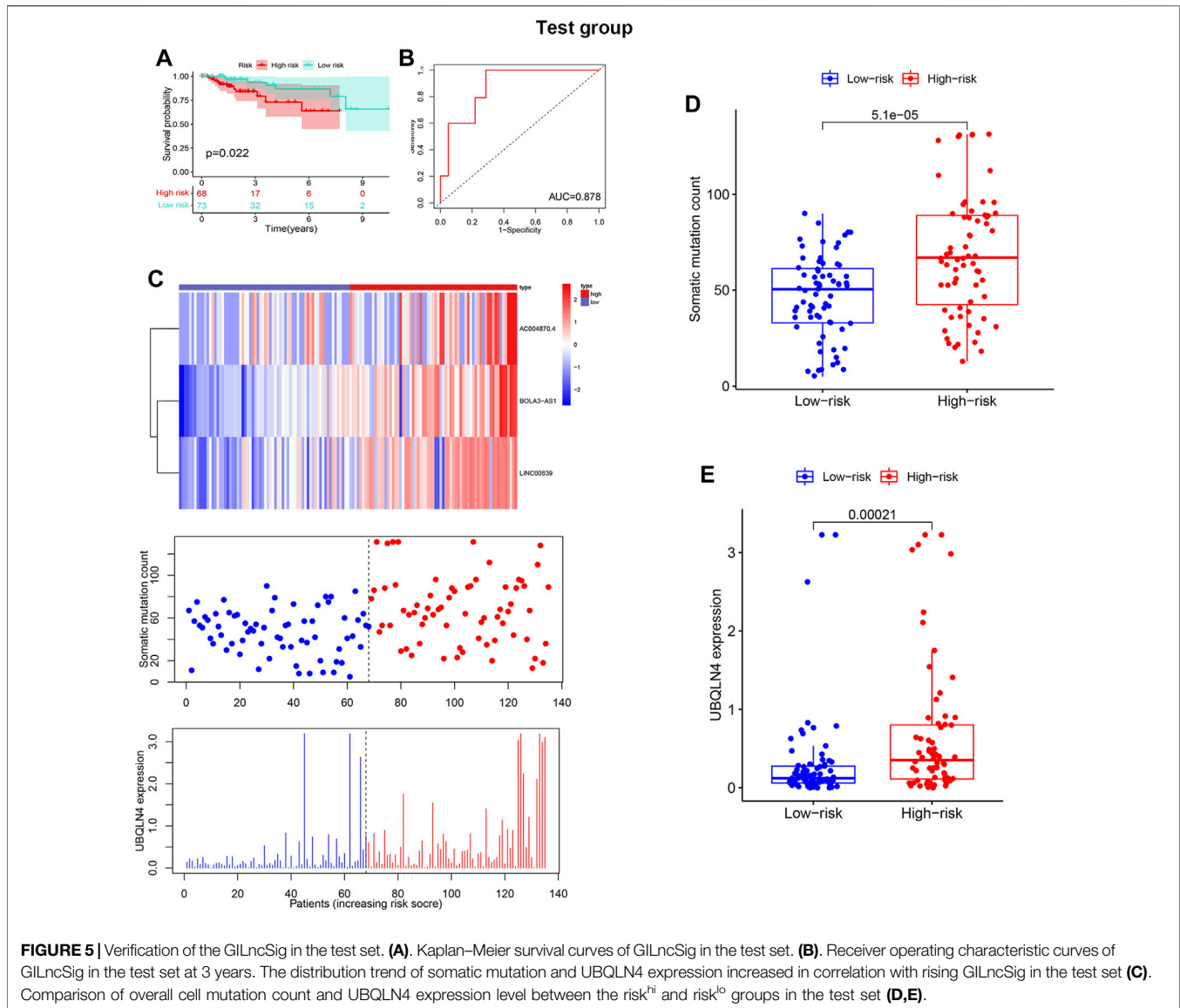
concentration was utilized as the dividing point, we found that the KIRP patients with risk high group had worse survival rate (**Figure 3E**). The AUC of GILncSig was 0.816 according to the ROC curve (**Figure 4A**), which further confirmed its role as a prognostic marker of KIRP. Furthermore, the patients were ranked in ascending order according to their risk score for observing GILncSig trends, somatic mutation numbers and UBQLN4 expression value. Patients with risk[hi] scores showed an increasing trend in somatic mutation count and UBQLN4 expression level (**Figure 4B**). Moreover, compared with the risk[lo] group, the high-risk group had more numbers of somatic mutations (**Figure 4C**). Similarly, the risk[hi] group showed higher expression levels of UBQLN4 than the risk[lo] group (**Figure 4D**).

FIGURE 4 | Independent validation of GILncSig in the training set with RNA-seq platform. **(A)**. Time receiver operating characteristic curve analysis of GILncSig in the training set at 3 years. **(B)**. lncRNA expression patterns, somatic mutation distribution and UBQLN4 expression based on the increasing trend of GILncSig score. Additionally, total somatic mutation distribution with increasing GILncSig score is displayed. **(C)**. Comparison of the total number of somatic mutations in different GILncSig groups. **(D)**. The expression level of UBQLN4 in various GILncSig groups.

Validation of the GILncSig in a Distinct Dataset

The test and TCGA sets were used to verify the prognostic value of GILncSig. The test set, which included 142 patients, showed that the survival rate of the high-risk group ($n = 85$) was much lower than that of the risk[lo] group ($n = 91$) (**Figure 5A**), with an AUC of 0.878 (**Figure 5B**). Additionally, the test set showed that the expression of GILncSig and the expression level of UBQLN4, as well as somatic mutation number were similar to the training set (**Figure 5C**). The test group showed substantial difference in the number of somatic mutations count between the risk[hi] and risk[lo] groups (**Figure 5D**); in addition, UBQLN44 expression in

the high-risk group was higher than that in the risk[lo] group (**Figure 5E**). Moreover, GILncSig in the TCGA dataset showed similar results. The overall survival rate in the risk[hi] group was significantly lower than that in the risk[lo] group ($p < 0.001$, **Figure 6A**), with an AUC of 0.833 for the TCGA set (**Figure 6B**). The TCGA set also showed similar GILncSig expression, somatic mutation count and UBQLN4 expression (**Figure 6C**). In the TCGA dataset, higher somatic mutation numbers were observed in the high-risk group (**Figure 6D**). It was revealed that there was a statistically significant difference in UBQLN4 expression between the two risk groups (**Figure 6E**).

FIGURE 5 | Verification of the GILncSig in the test set. **(A)**. Kaplan–Meier survival curves of GILncSig in the test set. **(B)**. Receiver operating characteristic curves of GILncSig in the test set at 3 years. The distribution trend of somatic mutation and UBQLN4 expression increased in correlation with rising GILncSig in the test set **(C)**. Comparison of overall cell mutation count and UBQLN4 expression level between the risk[hi] and risk[lo] groups in the test set **(D,E)**.
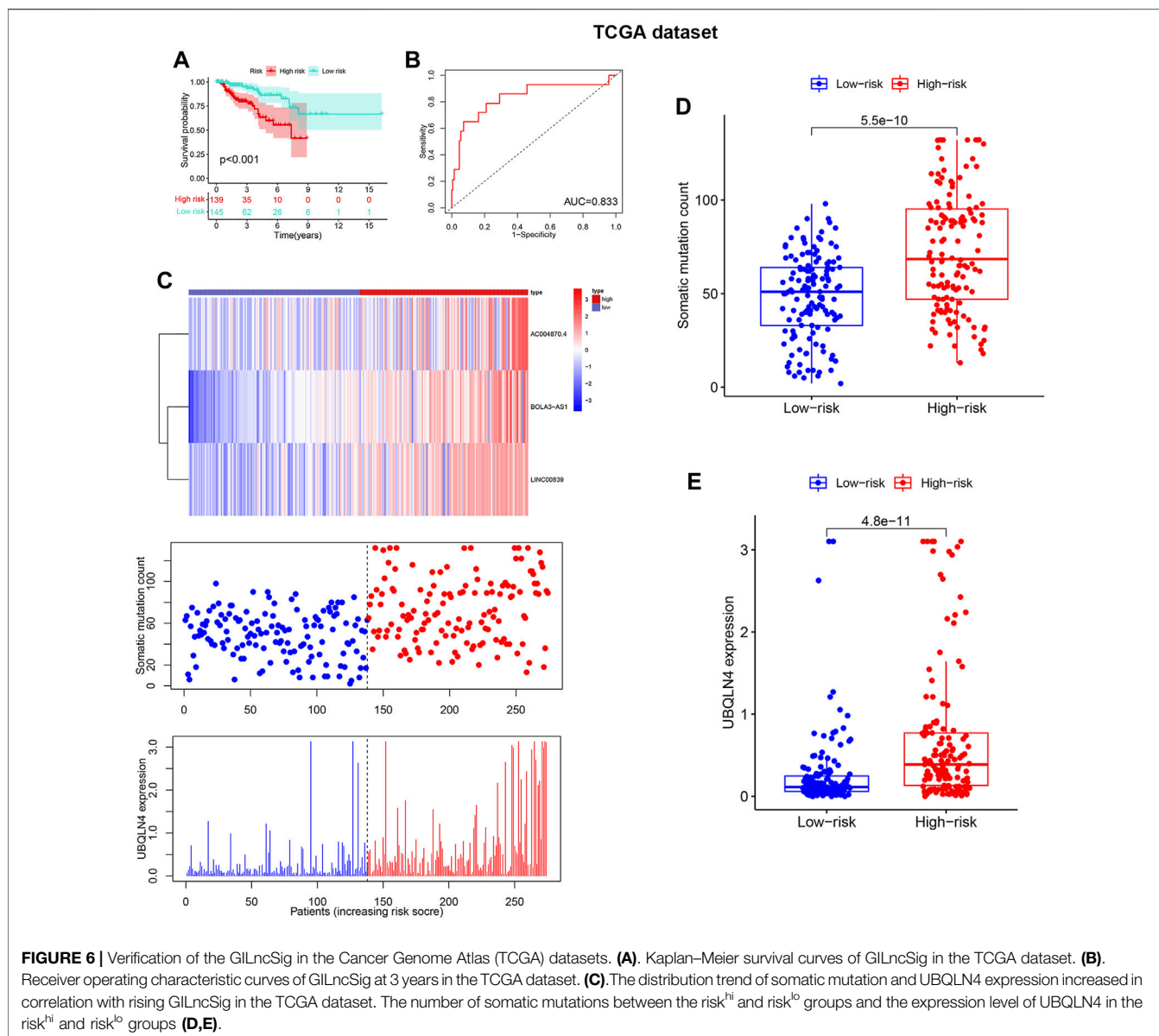
GILncSig is Prognostically Independent of Other Clinicopathological Factors

Univariate Cox and multivariate Cox analysis on the age, gender, pathological stage malignancy and prognostic risk score of GILncSig was performed. GILncSig was found to be an independent prognostic predictor in patients with KIRP after controlling for age, gender, and clinical stage (**Supplementary Figure S2A,B**).

GILncSig's Performance is Compared to Those of Other lncRNA-Related Signatures and Molecular Biomarkers

The predictive performance of GILncSig was compared with two recently published lncRNA characteristics. A 3-lncRNA signature was obtained from a study conducted by Wang (LilncSig is the name given to this signature), and an 8-lncRNA signature was obtained from a work conducted by Sun (referred to as SunlncSig). The AUC of the 3-year overall survival (AUC = 0.828) of GILncSig was higher than that of Wanglncsig (AUC = 0.680) and Sunlncsig (AUC = 0.716) (**Figure 7A**), indicating that GILncSig has a higher predictability when compared to the other two previously reported lncRNA signatures. Besides, the predictive value of GILncSig was further compared with two KIRP related molecular biomarkers. A 3-m6A signature (m6ASig) was built form a research by Wang (Sun Z. et al., 2020), and a 4 tumor microenvironment signature (TMESig) was constructed form a study by Su (Luo et al., 2021). The results showed that the AUC of the 3-year overall survival of GILncSig was higher than that of TMESig (AUC = 0.787), in addition, it has the same value as m6ASig (AUC = 0.828) (**Figure 7B**).These data suggest that GILncSig is a promising biomarker in KIRP.
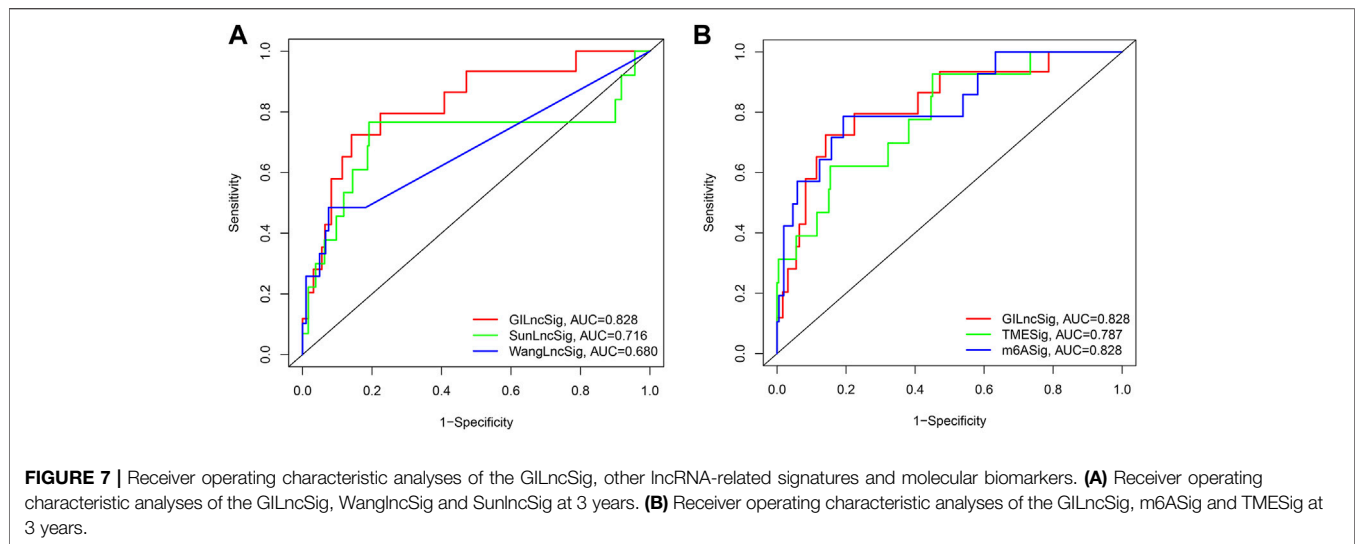
FIGURE 6 | Verification of the GILncSig in the Cancer Genome Atlas (TCGA) datasets. **(A)**. Kaplan–Meier survival curves of GILncSig in the TCGA dataset. **(B)**. Receiver operating characteristic curves of GILncSig at 3 years in the TCGA dataset. **(C)**.The distribution trend of somatic mutation and UBQLN4 expression increased in correlation with rising GILncSig in the TCGA dataset. The number of somatic mutations between the risk[hi] and risk[lo] groups and the expression level of UBQLN4 in the risk[hi] and risk[lo] groups **(D,E)**.

DISCUSSION

KIRP remains a clinical challenge owing to its high histological heterogeneity, poor prognosis and limited treatment options. The majority of the research on KIRP has been on a few well-known cancer-related genes, for example, hsa-mir-3199 and hsa-mir-1293 were found to be novel prognostic biomarkers of KIRP (Luo et al., 2017). Additionally, several mRNAs were discovered to be predictive of the survival probability of KIRP patients (Gao et al., 2019), such as CCNB2, IGF2BP3, KIF18A, PTTG1, and BUB1. However, the incidence and mortality rates of KIRP have been annually increasing; however, the underlying mechanism of its carcinogenesis remains unclear. Therefore, identifying new reliable molecular features to predict the survival rate of patients with KIRP is crucial. Technologies such as high-throughput and transcriptome sequencing have largely promoted cancer research in the past decade. This study discovered a new signature based on three lncRNAs that are related with genomic instability that can be used to predict the prognosis of patients with KIRP, which was previously unknown, as well as the prognostic valve of these signature was confirmed using data from TCGA. The prognostic signature developed exhibited good performance in distinguishing high-risk patients with KIRP that had a poor prognosis, which may contribute to their clinical treatment.

lncRNA, an important noncoding RNA, is a critical player in cell homeostasis, cell proliferation, migration, and genome stability, among other processes (Huarte, 2015). The lncRNA GClnc1 is linked to the malignant behaviour of bladder cancer (Zhuang et al., 2019). Previous research has found a number of long noncoding RNAs (lncRNAs) that are linked with the progression and prognosis of RCC. For instance, HCP5

FIGURE 7 | Receiver operating characteristic analyses of the GILncSig, other lncRNA-related signatures and molecular biomarkers. **(A)** Receiver operating characteristic analyses of the GILncSig, WanglncSig and SunlncSig at 3 years. **(B)** Receiver operating characteristic analyses of the GILncSig, m6ASig and TMESig at 3 years.

inhibits the growth and metastasis of RCC cells by regulating the mir-214-3p/mapk1 axis (Hao et al., 2020). Moreover, lncRNAs are closely associated with the incidence of RCC and have been identified as a new prognostic marker in recent years (Bhan et al., 2017; Wang et al., 2019). For example, Xia et al. developed a result prediction model based on nine lncRNAs with redox-related functions (Qi-Dong et al., 2020). Additionally, Zhang et al. identified five lncRNAs associated with the overall survival of RCC (Zhang et al., 2021). LncRNAs, in addition, are strongly associated with genomic stability and have been shown to enhance the onset, progression, and metastasis of RCC (Lee et al., 2016; Zhai et al., 2017). A recent study reported six genomic instability-associated lncRNAs (LINC02678, HOXA10-AS, RHOXF1-AS1, AC010789.1, LINC01150, and TGFB2-AS1) that were associated with gastric cancer prognosis and somatic mutations (Sun et al., 2021). However, studies reporting on lncRNAs associated with genomic instability in KIRP are limited.

3,840 cancer-associated lncRNAs were retrieved from the TCGA database and 27 lncRNAs related to genomic instability were found in this study to further investigate the potential impacts of genomic instability on KIRP. Cox regression analysis showed that 11 genomic instability-associated lncRNAs were related to the survival of patients with KIRP. Furthermore, The GILncSIg was constructed using three mutation-derived lncRNAs, which can be used to predict the outcome of KIRP patients. Among these lncRNAs with predictive characteristics related to genome instability, BOLA3-AS1, LINC00839, and AC004870.4 were considered the risk factors. BOLA3-AS1 is a different transcript of BOLA3 (Fagerberg et al., 2014), which is associated with high-risk myelodysplastic syndrome and plays a crucial function in the formation of blood cells, such as platelets, erythrocytes and bone marrow cells (Szikszai et al., 2020). Furthermore, the expression of BOLA3-AS1 has been reported to be up-regulated in gastric cancer cell lines and is related to the poor prognosis of gastric

cancer (Zhang et al., 2021). Studies have shown that LINC00839 targets the mir-338-3p/GLUT1 axis to promote the cell proliferation, migration, invasion and glycolysis of neuroblastoma cells, resulting in the poor survival rate of patients with neuroblastoma (Sahu et al., 2018; Yang et al., 2021). In addition, LINC00389 promotes the proliferation, migration, and invasion and inhibits the apoptosis of liver and lung cancer cells (Meng et al., 2021; Yu et al., 2021). However, as far as we are aware at this time, AC004870.4 has been published for the first time in this study. Therefore, well-designed investigations should be carried out in the future to explore potential functions and mechanisms of AC004870.4 in cancer, which may reveal their potential as therapeutic targets for patients with KIRP.

Furthermore, we analysed the characteristics of GILncSIg based on the clinical information of KIRP. The survival curve revealed that patients in the high-risk group had a lower overall survival rate than those in the low-risk group. Notably, the results of time-dependent ROC curve analysis showed that the AUC value of GILncSig in the training and test sets exceeded 0.60, indicating the accuracy of the signature. These results were verified using the GEO dataset, which showed similar results. Moreover, the time-dependent ROC curve revealed that the signature developed in this study is an independent prognostic factor when compared with other clinicopathological features (including age, sex, and tumour stage). Other prognostic features of KIRP have been previously reported. However, in terms of ROC analysis, the signature established in this study has been proven to be more accurate in predicting survival outcomes than the signatures reported in previous studies. Furthermore, we assessed whether genomic instability was related to GILncSIg and found that the tumour mutation phenotype and UBEQL4 expression level were important indicators of genomic instability. UBEQL4 is a novel regulator of cancer genomic instability, which can lead to genomic instability to inhibit homologous recombination-mediated DSB repair (HRR) activity (Jachimowicz et al.,

2019). The results revealed that the somatic mutation rate and expression level of UBEQL4 in the high-risk group were significantly higher than those in the low-risk group. This finding indicated that GILncSIg was closely related to genomic instability and further demonstrated the key role of lncRNAs in maintaining genomic instability and the importance of examining the underlying mechanisms of lncRNAs in epigenetics. Although this study provides promising insights into a better genomic instability assessment and KIRP prognosis, it had some limitations. First, although GILncSig was verified using the TCGA and GEO datasets, it will be necessary to use more independent datasets in order to further verify GILncSig in order to confirm its robustness and reproducibility. Second, it is necessary to conduct biological experiments on these predicted lncRNAs *in vitro* and *in vivo*. In conclusion, this study proposes a computational framework derived from a hypothesis based on mutations to identify lncRNAs related to genomic instability, which introduces novel outlooks and resources for further studying the role of lncRNAs in genomic instability. Based on lncRNA expression profile, somatic mutation profile and clinical data of KIRP patients, a prognostic marker of lncRNA derived from genomic instability was constructed to predict the prognosis of people suffering from KIRP. GILncSig was found to be a potential prognostic index that was independent of other usual clinicopathological factors, according to further investigation, thereby introducing a new target for KIRP treatment.

REFERENCES

Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Res.* 77, 3965–3981. doi:10.1158/0008-5472.can-16-2634

Chi, Y., Wang, D., Wang, J., Yu, W., and Yang, J. (2019). Long Non-Coding RNA in the Pathogenesis of Cancers. *Cells* 8 (9), 1015. doi:10.3390/cells8091015

Choueiri, T. K., Pal, S. K., McDermott, D. F., Morrissey, S., Ferguson, K. C., Holland, J., et al. (2014). A Phase I Study of Cabozantinib (XL184) in Patients with Renal Cell Cancer. *Ann. Oncol.* 25, 1603–1608. doi:10.1093/annonc/mdu184

Choueiri, T. K., Vaishampayan, U., Rosenberg, J. E., Logan, T. F., Harzstark, A. L., Bukowski, R. M., et al. (2013). Phase II and Biomarker Study of the Dual MET/VEGFR2 Inhibitor Foretinib in Patients with Papillary Renal Cell Carcinoma. *J. Clin. Oncol.* 31, 181–186. doi:10.1200/jco.2012.43.3383

Dong, J., Xu, W., Ji, Z., and Pan, B. (2021). Xp11.2 Translocation Renal Cell Carcinoma: Clinical Characteristics and Potential Prognostic Predictors. *Dis. Markers* 2021, 5647933. doi:10.1155/2021/5647933

Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., et al. (2014). Analysis of the Human Tissue-specific Expression by Genome-Wide Integration of Transcriptomics and Antibody-Based Proteomics. *Mol. Cel Proteomics* 13, 397–406. doi:10.1074/mcp.m113.035600

Gao, Z., Zhang, D., Duan, Y., Yan, L., Fan, Y., Fang, Z., et al. (2019). A Five-Gene Signature Predicts Overall Survival of Patients with Papillary Renal Cell Carcinoma. *PLoS One* 14, e0211491. doi:10.1371/journal.pone.0211491

Hao, J.-F., Chen, P., Li, H.-Y., Li, Y.-J., and Zhang, Y.-L. (2020). Effects of LncRNA HCP5/miR-214-3p/MAPK1 Molecular Network on Renal Cell Carcinoma Cells. *Cancer Manag. Res.* 12, 13347–13356. doi:10.2147/cmar.s274426

Huarte, M. (2015). The Emerging Role of lncRNAs in Cancer. *Nat. Med.* 21, 1253–1261. doi:10.1038/nm.3981

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The Landscape of Long Noncoding RNAs in the Human Transcriptome. *Nat. Genet.* 47, 199–208. doi:10.1038/ng.3192

Jachimowicz, R. D., Beleggia, F., Isensee, J., Velpula, B. B., Goergens, J., Bustos, M. A., et al. (2019). UBQLN4 Represses Homologous Recombination and Is Overexpressed in Aggressive Tumors. *Cell* 176, 505–519. doi:10.1016/j.cell.2018.11.024

Jonasch, E., Walker, C. L., and Rathmell, W. K. (2021). Clear Cell Renal Cell Carcinoma Ontogeny and Mechanisms of Lethality. *Nat. Rev. Nephrol.* 17, 245–261. doi:10.1038/s41581-020-00359-2

Lee, S., Kopp, F., Chang, T.-C., Sataluri, A., Chen, B., Sivakumar, S., et al. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* 164, 69–80. doi:10.1016/j.cell.2015.12.017

Liu, S., Wang, Z., Zhu, R., Wang, F., Cheng, Y., and Liu, Y. (2021). Three Differential Expression Analysis Methods for RNA Sequencing: Limma, EdgeR, DESeq2. *J. Vis. Exp.* 175, e62528. doi:10.3791/62528

Luo, L., Zhou, H., and Su, H. (2021). Identification of 4-Genes Model in Papillary Renal Cell Tumor Microenvironment Based on Comprehensive Analysis. *BMC Cancer* 21, 553. doi:10.1186/s12885-021-08319-0

Luo, W., Wang, L., Luo, M. H., Huang, Y. Z., Yang, H., Zhou, Y., et al. (2017). hsa-mir-3199-2 and Hsa-mir-1293 as Novel Prognostic Biomarkers of Papillary Renal Cell Carcinoma by COX Ratio Risk Regression Model Screening. *J. Cel. Biochem.* 118, 3488–3494. doi:10.1002/jcb.26008

Meng, B., Han, F., Gao, B., Zhuang, H., Zhang, X. Z., Wang, Y. J., et al. (2021). Effects of LINC00839 Targeting miR-3666 on Proliferation, Migration and Invasion of Hepatocellular Carcinoma Cells. *Zhonghua Zhong Liu Za Zhi* 43, 1148–1155. doi:10.3760/cma.j.cn112152-20200222-00116

Moran, V. A., Perera, R. J., and Khalil, A. M. (2012). Emerging Functional and Mechanistic Paradigms of Mammalian Long Non-coding RNAs. *Nucleic Acids Res.* 40, 6391–6400. doi:10.1093/nar/gks296

Ottini, L., Falchetti, M., Lupi, R., Rizzolo, P., Agnese, V., Colucci, G., et al. (2006). Patterns of Genomic Instability in Gastric Cancer: Clinical Implications and

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Materials**.

AUTHOR CONTRIBUTIONS

JL and SW contributed equally to this study. YL, ZH designed this work. JL and SW wrote the manuscript. YZ performed the bioinformatics analysis. SL, XZ, QW, JY, LC, and SW performed the data review. All authors have read and approved the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.874673/full#supplementary-material

Perspectives. *Ann. Oncol.* 17 (Suppl. 7), vii97–vii102. doi:10.1093/annonc/mdl960

Peng, B., Li, H., Na, R., Lu, T., Li, Y., Zhao, J., et al. (2021). Identification of a Novel Prognostic Signature of Genome Instability-Related LncRNAs in Early Stage Lung Adenocarcinoma. *Front. Cel Dev. Biol.* 9, 706454. doi:10.3389/fcell.2021.706454

Qi-Dong, X., Yang, X., Lu, J. L., Liu, C. Q., Sun, J. X., Li, C., et al. (2020). Development and Validation of a Nine-Redox-Related Long Noncoding RNA Signature in Renal Clear Cell Carcinoma. *Oxid Med. Cel Longev* 2020, 6634247. doi:10.1155/2020/6634247

Sahu, D., Ho, S. Y., Juan, H. F., and Huang, H. C. (2018). High-Risk, Expression-Based Prognostic Long Noncoding RNA Signature in Neuroblastoma. *JNCI Cancer Spectr.* 2, pky015. doi:10.1093/jncics/pky015

Sukov, W. R., Lohse, C. M., Leibovich, B. C., Thompson, R. H., and Cheville, J. C. (2012). Clinical and Pathological Features Associated with Prognosis in Patients with Papillary Renal Cell Carcinoma. *J. Urol.* 187, 54–59. doi:10.1016/j.juro.2011.09.053

Sun, J., Jiang, Q., Chen, H., Zhang, Q., Zhao, J., Li, H., et al. (2021). Genomic Instability-Associated lncRNA Signature Predicts Prognosis and Distinct Immune Landscape in Gastric Cancer. *Ann. Transl Med.* 9, 1326. doi:10.21037/atm-21-3569

Sun, J., Tang, Q., Gao, Y., Zhang, W., Zhao, Z., Yang, F., et al. (2020a). VHL Mutation-Mediated SALL4 Overexpression Promotes Tumorigenesis and Vascularization of clear Cell Renal Cell Carcinoma via Akt/GSK-3β Signaling. *J. Exp. Clin. Cancer Res.* 39, 104. doi:10.1186/s13046-020-01609-8

Sun, Z., Jing, C., Xiao, C., Li, T., and Wang, Y. (2020b). Prognostic Risk Signature Based on the Expression of Three m6A RNA Methylation Regulatory Genes in Kidney Renal Papillary Cell Carcinoma. *Aging* 12, 22078–22094. doi:10.18632/aging.104053

Suzuki, K., Ohnami, S., Tanabe, C., Sasaki, H., Yasuda, J., Katai, H., et al. (2003). The Genomic Damage Estimated by Arbitrarily Primed PCR DNA Fingerprinting Is Useful for the Prognosis of Gastric Cancer. *Gastroenterology* 125, 1330–1340. doi:10.1016/j.gastro.2003.07.006

Szikszai, K., Krejcik, Z., Klema, J., Loudova, N., Hrustincova, A., Belickova, M., et al. (2020). LncRNA Profiling Reveals that the Deregulation of H19, WT1-AS, TCL6, and LEF1-AS1 Is Associated with Higher-Risk Myelodysplastic Syndrome. *Cancers (Basel)* 12 (10), 2726. doi:10.3390/cancers12102726

Tam, A. S., Sihota, T. S., Milbury, K. L., Zhang, A., Mathew, V., and Stirling, P. C. (2019). Selective Defects in Gene Expression Control Genome Instability in Yeast Splicing Mutants. *Mol. Biol. Cel* 30, 191–200. doi:10.1091/mbc.e18-07-0439

Wang, L., Cho, K. B., Li, Y., Tao, G., Xie, Z., and Guo, B. (2019). Long Noncoding RNA (lncRNA)-Mediated Competing Endogenous RNA Networks Provide Novel Potential Biomarkers and Therapeutic Targets for Colorectal Cancer. *Int. J. Mol. Sci.* 20 (22), 5758. doi:10.3390/ijms20225758

Yang, L., Pei, L., and Yi, J. (2021). LINC00839 Regulates Proliferation, Migration, Invasion, Apoptosis and Glycolysis in Neuroblastoma Cells through miR-338-3p/GLUT1 Axis. *Neuropsychiatr. Dis. Treat.* 17, 2027–2040. doi:10.2147/ndt.s309467

Yao, R.-W., Wang, Y., and Chen, L.-L. (2019). Cellular Functions of Long Noncoding RNAs. *Nat. Cel Biol* 21, 542–551. doi:10.1038/s41556-019-0311-8

Yin, T., Zhao, D., and Yao, S. (2021). Identification of a Genome Instability-Associated LncRNA Signature for Prognosis Prediction in Colon Cancer. *Front. Genet.* 12, 679150. doi:10.3389/fgene.2021.679150

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). ClusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* 16, 284–287. doi:10.1089/omi.2011.0118

Yu, X., Jiang, Y., Hu, X., and Ge, X. (2021). LINC00839/miR-519d-3p/JMJD6 Axis Modulated Cell Viability, Apoptosis, Migration and Invasiveness of Lung Cancer Cells. *Folia Histochem. Cytobiol* 59 (4), 271–281. doi:10.5603/fhc.a2021.0022

Zeng, X., Liu, L., Lü, L., and Zou, Q. (2018). Prediction of Potential Disease-Associated microRNAs Using Structural Perturbation Method. *Bioinformatics* 34, 2425–2432. doi:10.1093/bioinformatics/bty112

Zhai, W., Sun, Y., Guo, C., Hu, G., Wang, M., Zheng, J., et al. (2017). LncRNA-SARCC Suppresses Renal Cell Carcinoma (RCC) Progression via Altering the Androgen Receptor(AR)/miRNA-143-3p Signals. *Cell Death Differ* 24, 1502–1517. doi:10.1038/cdd.2017.74

Zhang, H.-M., Yang, F.-Q., Chen, S.-J., Che, J., and Zheng, J.-H. (2015). Upregulation of Long Non-Coding RNA MALAT1 Correlates with Tumor Progression and Poor Prognosis in Clear Cell Renal Cell Carcinoma. *Tumor Biol.* 36, 2947–2955. doi:10.1007/s13277-014-2925-6

Zhang, X., Zou, Q., Rodriguez-Paton, A., and Zeng, X. (2019). Meta-Path Methods for Prioritizing Candidate Disease miRNAs. *IEEE/ACM Trans. Comput. Biol. Bioinf.* 16, 283–291. doi:10.1109/tcbb.2017.2776280

Zhang, Y., Pitchiaya, S., Cieślik, M., Niknafs, Y. S., Tien, J. C.-Y., Hosono, Y., et al. (2018). Analysis of the Androgen Receptor-Regulated lncRNA Landscape Identifies a Role for ARLNC1 in Prostate Cancer Progression. *Nat. Genet.* 50, 814–824. doi:10.1038/s41588-018-0120-1

Zhang, Z., Tang, Y., Liu, Y., Zhuang, H., Lin, E., Xie, L., et al. (2021). A Novel Immune-Related lncRNA-Based Model for Survival Prediction in Clear Cell Renal Cell Carcinoma. *J. Immunol. Res.* 2021, 9921466. doi:10.1155/2021/9921466

Zhou, M., Zhao, H., Wang, X., Sun, J., and Su, J. (2019). Analysis of Long Noncoding RNAs Highlights Region-Specific Altered Expression Patterns and Diagnostic Roles in Alzheimer's Disease. *Brief Bioinform* 20, 598–608. doi:10.1093/bib/bby021

Zhuang, C., Ma, Q., Zhuang, C., Ye, J., Zhang, F., and Gui, Y. (2019). LncRNA GClnc1 Promotes Proliferation and Invasion of Bladder Cancer through Activation of MYC. *FASEB j.* 33, 11045–11059. doi:10.1096/fj.201900078rr

Znaor, A., Lortet-Tieulent, J., Laversanne, M., Jemal, A., and Bray, F. (2015). International Variations and Trends in Renal Cell Carcinoma Incidence and Mortality. *Eur. Urol.* 67, 519–530. doi:10.1016/j.eururo.2014.10.002

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.