



Identification of Hub Genes Associated With Progression and Prognosis in Patients With Bladder Cancer

Xin Yan^{1,2†}, Xiao-Ping Liu^{1†}, Zi-Xin Guo¹, Tong-Zu Liu¹ and Sheng Li^{1,2,3*}

¹ Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan, China, ² Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, China, ³ Human Genetics Resource Preservation Center of Hubei Province, Wuhan, China

OPEN ACCESS

Edited by:

Xing Chen,
China University of Mining
and Technology, China

Reviewed by:

Raul Isea,
Fundación Instituto de Estudios
Avanzados (IDEA), Venezuela
Zhi-Ping Liu,
Shandong University, China

*Correspondence:

Sheng Li
lisheng-znyy@whu.edu.cn

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Genetics

Received: 30 November 2018

Accepted: 15 April 2019

Published: 07 May 2019

Citation:

Yan X, Liu X-P, Guo Z-X, Liu T-Z
and Li S (2019) Identification of Hub
Genes Associated With Progression
and Prognosis in Patients With
Bladder Cancer.
Front. Genet. 10:408.
doi: 10.3389/fgene.2019.00408

Given that most bladder cancers (BCs) are diagnosed in advanced stages with poor prognosis, this study aims to find novel biomarkers associated with the progression and prognosis in patients with BC. 1,779 differentially expressed genes (DEGs) between BC samples and normal bladder tissues were identified in total. Then, 24 DEGs were regarded as candidate hub genes by constructing a protein-protein interaction (PPI) network and a random forest model. Among them, six genes (BUB1B, CCNB1, CDK1, ISG15, KIF15, and RAD54L) were eventually identified by using five analysis methods (one-way Analysis of Variance analysis, spearman correlation analysis, distance correlation analysis, receiver operating characteristic curve, and expression values comparison), which were correlated with the progression and prognosis of BC. Moreover, the validation of hub genes was conducted based on GSE13507, Oncomine, and CBioPortal. Results of univariate Cox regression analysis showed that the expression levels of all the hub genes were influence features of overall survival (OS) and cancer specific survival (CSS) based on GSE13507, and we further established a six-gene signature based on the expression levels of the six genes and their Cox regression coefficients. This signature showed good potential for clinical application suggested by survival analysis (OS: Hazard Ratio = 0.484, 95%CI: 0.298–0.786; $P = 0.0034$; CSS: Hazard Ratio = 0.244, 95%CI: 0.121–0.493, $P < 0.0001$) and decision curve analysis. In conclusion, our study indicates that six hub genes have great predictive value for the prognosis and progression of BC and may contribute to the exploration of further basic and clinical research of BC.

Keywords: bladder cancer, hub genes, survival analysis, enrichment analysis, prognosis, bioinformatics

INTRODUCTION

Bladder cancer (BC) is the most common malignancy of the urinary system (Ferlay et al., 2010). According to recent research, the incidence of BC is growing worldwide (Ebrahimi et al., 2019). There were 437,442 new cases of BC in 2016 (Ebrahimi et al., 2019). BC ranks the first in urinary malignancies among males (Chen et al., 2016). Therefore, early diagnosis, postoperative monitoring, prognosis evaluation and more effective individualized treatment of BC are extremely important. At present, cystoscopy and biopsy are still the gold standard for diagnosing BC (Emerson and Cheng, 2005). Cystoscopy is an invasive examination,

which is not accepted by most patients (Harris et al., 1997). Urinary cytology examination is the main method for diagnosing BC and postoperative follow-up (Matsuzaki et al., 2016). The specificity of urinary cytology examinations was high, ranging from 90 to 96%, while the sensitivity of diagnosis for low grade and early stage BC was very low (Kiyoshima et al., 2016). Biomarkers are the frontiers and hotspots in the screening and diagnosis of BC (Kiyoshima et al., 2016).

In recent years, a few biomarkers have been used for the diagnosis of BC. For instance, Nuclear Matrix Protein-22 (NMP-22), a protein component found in the spindle of mitotic cells, could ensure chromosome segregation when cells underwent mitosis. A small amount of NMP-22 could be found in the urine of normal people, and BC might exist when its content exceeded the threshold (Miakhil et al., 2013). However, the sensitivity and specificity of NMP-22 were reported controversially in literatures (47–90%) (Tilki et al., 2011). Given that most BCs were diagnosed with advanced stages, the prognosis of patients with BC remained extremely poor. Therefore, it is of top priority to develop novel and specific prognostic markers for patients with BC.

In the field of machine learning and data mining, random forest is an ensemble learning method for classification, regression and other tasks (Rigatti, 2017). Random forest realizes its function by constructing a multitude of decision trees at training time (Svetnik et al., 2003). Then, it further outputs the class that is the mode of the classification or mean prediction (regression) of the individual trees (Svetnik et al., 2003). This method has a lot of advantages, one of these is that it gives estimates of what variables are important in the classification (Svetnik et al., 2003), which interests us most. It means that we can construct a random forest model to narrow down the number of genes (variables) according to the importance of variables.

Previous studies only identified differentially expressed genes (DEGs) based on one database and there was a lack of validation when using other databases (Jia et al., 2015). Although some studies had validated their results through reverse transcription-quantitative polymerase chain reaction (RT-qPCR) and western blotting, researchers only identified hub genes by constructing a co-expression network (Zhang et al., 2017) or a protein-protein interaction (PPI) network (Bi et al., 2015). In this study, we not only constructed a PPI network of DEGs to pick out hub genes, but also innovatively constructed a random forest model to further narrow down the number of hub genes in the PPI network by using these genes as features and their expression levels as feature values. Furthermore, unlike other studies, we used many different methods to identify hub genes associated with the progression and prognosis of patients with BC among these candidate hub genes and further validated them based on other databases.

MATERIALS AND METHODS

BC Gene Expression Studies

To screen DEGs, we downloaded TCGA-BLCA data including 413 BCs with clinical information and 19 normal bladder samples

from The Cancer Genome Atlas (TCGA) database¹. Another independent dataset GSE13507 (Kim et al., 2010; Lee et al., 2010) was downloaded from the gene expression omnibus (GEO) database² to verify our results. GSE13507, an expression profile based on Agilent GPL6102 platform (Illumina Human-6 v2.0 Expression Beadchip), included 165 primary BC samples, 23 recurrent non-muscle invasive tumor tissues, 10 normal bladder mucosae and 58 bladder mucosae surrounding cancer.

Data Processing and DEG Screening

The research process of this study was showed in **Figure 1**. For GSE13507, we used “affy” in R for normalization and log₂ transformation by using “rma” algorithm (Gautier et al., 2004). After weeding out samples with incomplete clinical information, a total of 427 samples (408 BCs and 19 normal samples) were used to select DEGs by using package “edgeR” (Robinson et al., 2010) in R. We considered genes as DEG when they reached the standard: Adjust *P* value < 0.05, and |log₂ fold change (FC)| ≥ 2 (Sun et al., 2017; Li et al., 2018).

Functional Enrichment Analysis

We performed Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for DEGs to find out their lurking functions by using R package “clusterProfiler” (Yu et al., 2012). In this study, we only showed the results of biological process (BP) and KEGG. Gene sets at *P* < 0.05 were considered to be significantly enriched (Li et al., 2018).

Candidate Hub Gene Identification

Firstly, by means of the Search Tool for the Retrieval of Interacting Genes (STRING) (Szklarczyk et al., 2015), we built the PPI network of DEGs. Parameters setting: Network scoring: degree cutoff = 2; Cluster finding: node score cutoff = 0.2, k-core = 2, and max. depth = 100 (Sun et al., 2017). In this study, we calculated the degree of genes by network analyzer (a tool in Cytoscape, Shannon et al., 2003). After that, genes with degree ≥ 50 were considered to be hub genes in the PPI work. Secondly, in order to pick out the most important factors associated with the progression among them, we further constructed a random forest model of hub genes in the PPI network by using package “randomForest” (Liaw and Wiener, 2002) in R. After that, genes which reached the standards (both of MeanDecreaseAccuracy and MeanDecreaseGini ranked top 50) (Svetnik et al., 2003) were considered as candidate hub genes.

Hub Gene Identification

In this study, five different methods were used to identify hub genes among candidate hub genes using GEO dataset GSE13507. The one-way ANOVA test and spearman correlation analysis were performed using SPSS (version 21.0). We used R package “ggplot2” (Wickham, 2015) to visualize the results. Meanwhile, we used R package “energy” (Rizzo and Szekely, 2016) to complete the distance correlation analysis to overcome

¹<https://genome-cancer.ucsc.edu/>

²<https://www.ncbi.nlm.nih.gov/geo/>

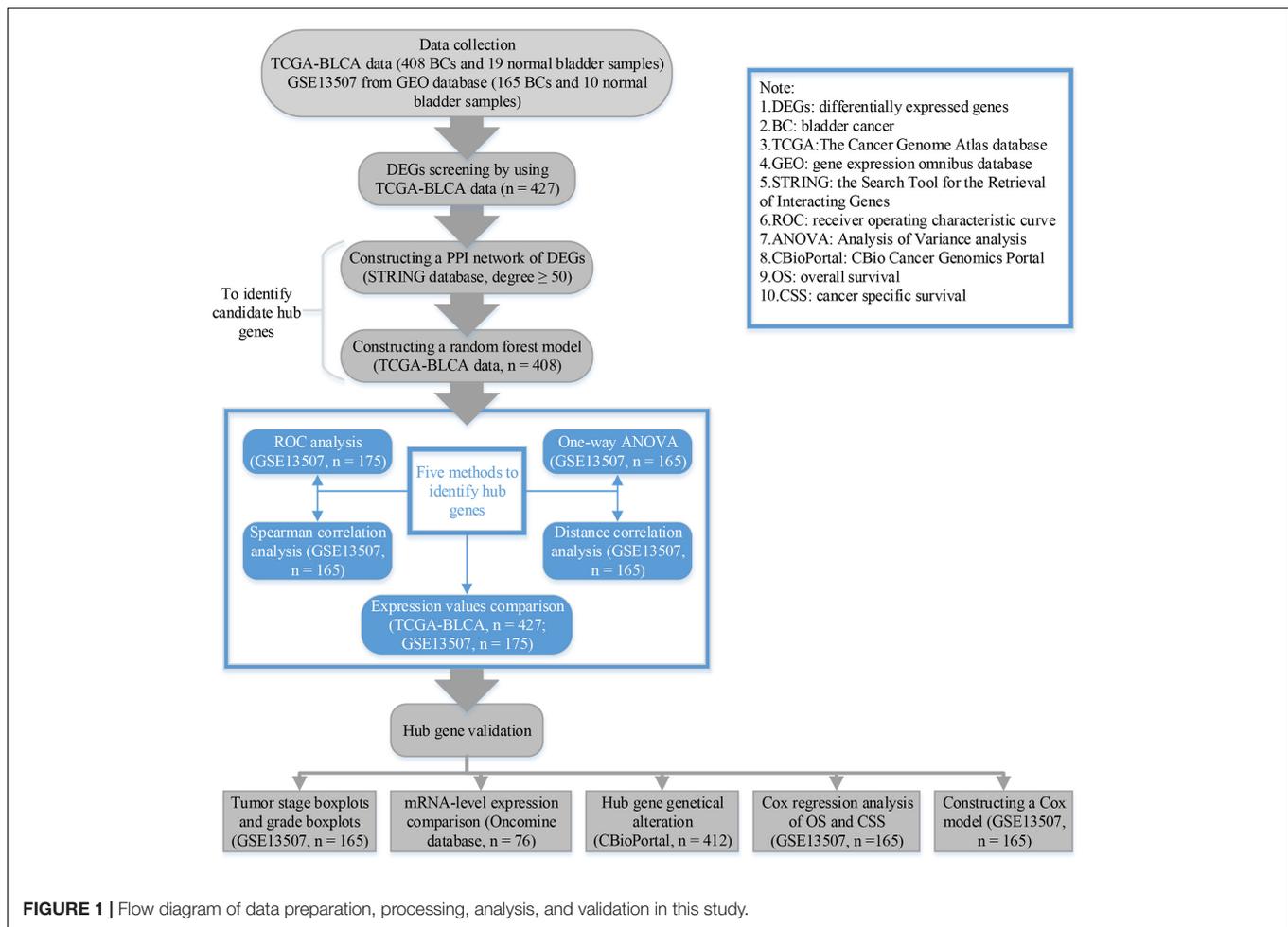


FIGURE 1 | Flow diagram of data preparation, processing, analysis, and validation in this study.

the weaknesses of spearman correlation. All of the three analyses were performed to explore the correlation between gene expression levels and tumor grade to pick out genes associated with tumor progression based on GSE13507. Moreover, by means of R package “plotROC” (Sachs, 2015), Receiver operating characteristic (ROC) analysis was performed. In GSE13507, we worked out the AUC to distinguish BC samples from normal tissues. After that, we compared the expression levels of candidate hub genes in BC and normal bladder tissues using GSE13507 ($n = 175$) and TCGA-BLCA data ($n = 427$). The boxplots were drawn using R package “ggstatsplot” (Patil and Powell, 2018) and GEPIA (Tang et al., 2017) (Gene Expression Profiling Interactive Analysis). Genes which satisfied the standard ($P < 0.05$ in all analysis and $AUC \geq 0.85$) (Yuan et al., 2018) were considered to be hub genes. An upset plot was also performed using R package “UpSetR” (Conway et al., 2017) to overlap genes in these five analysis methods.

Hub Gene Validation and Genetical Alteration

Based on GSE13507, T stage (Ta, T1, T2, T3, and T4) boxplots and tumor grade (low and high) boxplots were performed

using “ggstatsplot.” A one-way ANOVA test was performed to measure the statistical significance in stage boxplots. In addition, we validated the mRNA-level expression of hub genes based on Oncomine³ (Rhodes et al., 2004). In the present study, 28 tumors and 48 normal bladder samples from Sanchez-Carbayo Bladder 2 were included (Sanchez-Carbayo et al., 2006). We used an unpaired *t*-test to measure the statistical significance in grade boxplots and mRNA-level validation. Visualization and analysis of cancer genomic datasets can be realized by using the CBio Cancer Genomics Portal⁴ (Cerami et al., 2012; Gao et al., 2013). In the present study, CBioPortal was used for exploring genetic alterations of hub genes and relationships between genes and drugs (the data of CbioPortal originated from TCGA, 412 samples in total were included in this step).

Univariate and Multivariate Cox Regression Analysis

In the present study, we included the hub gene expression values and other important clinical information (gender, age, tumor grade, T stage, N stage, and M stage) into univariate Cox analysis

³<http://www.oncomine.org/>

⁴<http://www.cbioportal.org/>

of overall survival (OS) by using dataset GSE13507. After that, we constructed a Cox model by combining the regression coefficient (beta) with gene expression values. The Cox model and factors with p -Value < 0.05 in univariate Cox analysis were included for multivariate Cox analysis. To do this, we could determine if the prediction of hub genes was independent from other clinical features. We used R package “forestplot” (Aut and Aut, 2016) for visualization.

Survival Analysis and Decision Curve Analysis (DCA)

For investigating the influence of the Cox model on the OS and cancer specific survival (CSS) of BC patients, we calculated the risk score of every sample in GSE13507. BCs were divided into two groups (high risk and low risk) according to the median risk score of GSE13507. Kaplan–Meier survival analysis was conducted with the information from GSE13507 by using R package “survival” (Therneau, 2015). This package also generated a Kaplan–Meier survival curve. In addition, DCA was used to further validate the diagnostic value of this model and hub genes.

RESULTS

Identification of DEGs Between BC and Normal Bladder Tissues

In total, 1,779 DEGs were identified under the threshold of adjust P -value < 0.05 and $|\log_2FC| \geq 2$. Among them, 908 genes were up-regulated and 871 genes were down-regulated

(Supplementary Figure S1). The adjust P -value and \log_2FC of DEGs were available in Supplementary Table S1.

GO Biological Processes and KEGG Analysis

For primary comprehensions of these DEGs, GO and KEGG pathway analyses were performed. According to GO biological processes analysis, the up-regulated DEGs were enriched in 102 BPs. The top 10 were DNA packaging, nucleosome assembly, chromatin assembly, nucleosome organization, DNA conformation change, chromatin assembly or disassembly, DNA replication-dependent nucleosome assembly, DNA replication-dependent nucleosome organization, chromatin silencing at rDNA, and chromatin silencing (Figure 2A). As for the down-regulated DEGs, they were enriched in 326 BPs in total. The top 10 were muscle system process, muscle contraction, regulation of blood circulation, muscle organ development, regulation of heart contraction, heart contraction, heart process, regulation of muscle system process, striated muscle contraction, and regulation of muscle contraction (Figure 2B). While in KEGG pathway analysis, the up-regulated DEGs were enriched in only 5 KEGG pathways including systemic lupus erythematosus, alcoholism, viral carcinogenesis, cell cycle, and neuroactive ligand-receptor interaction (Figure 2C). Meanwhile, the down-regulated DEGs were enriched in 33 KEGG pathways totally. The top 10 were dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), calcium signaling pathway, arrhythmic right ventricular cardiomyopathy (ARVC), vascular smooth muscle contraction, cGMP-PKG signaling pathway, adrenergic signaling in cardiomyocytes, regulation of muscle contraction, and insulin secretion (Figure 2D).

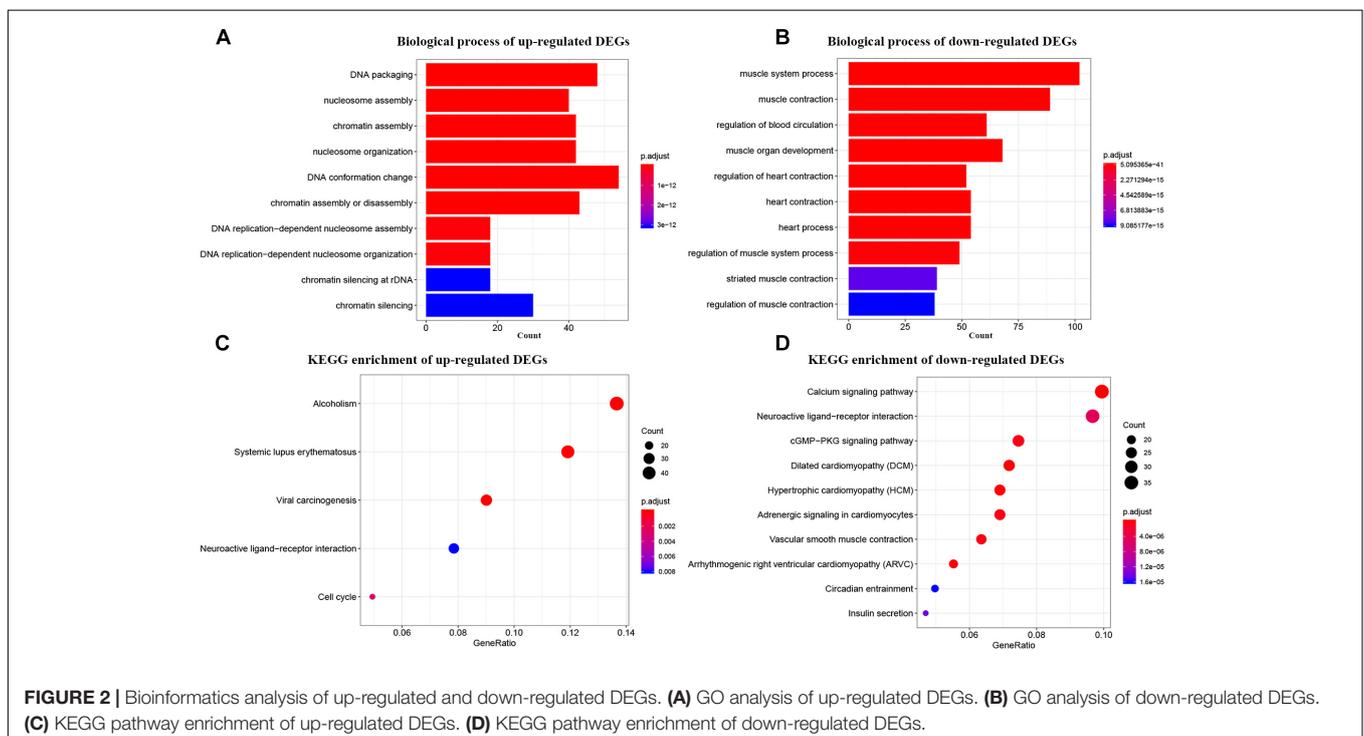


FIGURE 2 | Bioinformatics analysis of up-regulated and down-regulated DEGs. **(A)** GO analysis of up-regulated DEGs. **(B)** GO analysis of down-regulated DEGs. **(C)** KEGG pathway enrichment of up-regulated DEGs. **(D)** KEGG pathway enrichment of down-regulated DEGs.

neuroactive ligand-receptor interaction, insulin secretion, and circadian entrainment (Figure 2D).

Candidate Hub Gene Identification

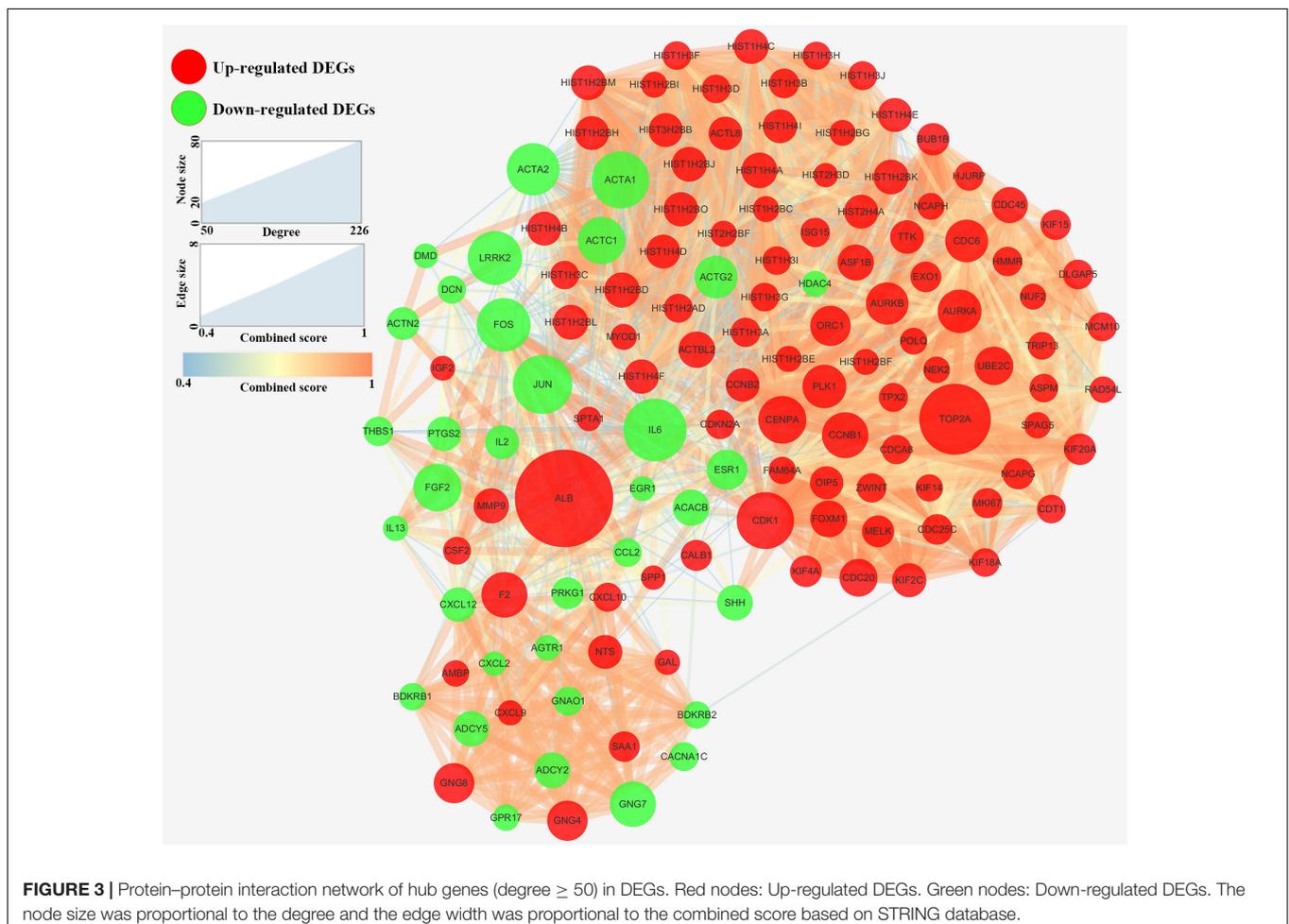
At the beginning, we constructed a PPI network of DEGs (Supplementary Figure S2). 134 genes with degree ≥ 50 were considered as hub genes in the PPI network (Figure 3). In order to narrow down the number of hub genes in the PPI network, we also constructed a random forest model. According to the results, 24 genes were common in genes with MeanDecreaseAccuracy ranked in the top 50 and genes with MeanDecreaseGini ranked in the top 50 (Supplementary Table S2). Genes with MeanDecreaseAccuracy ranked in the top 30 and genes with MeanDecreaseGini ranked in the top 30 were shown in Supplementary Figure S3. We regarded these 24 genes as candidate hub genes for further analysis.

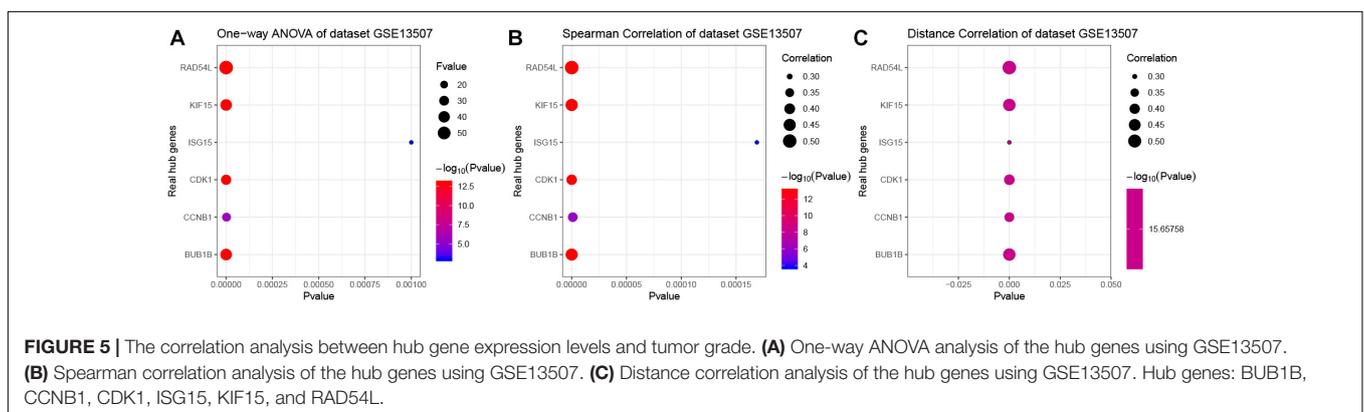
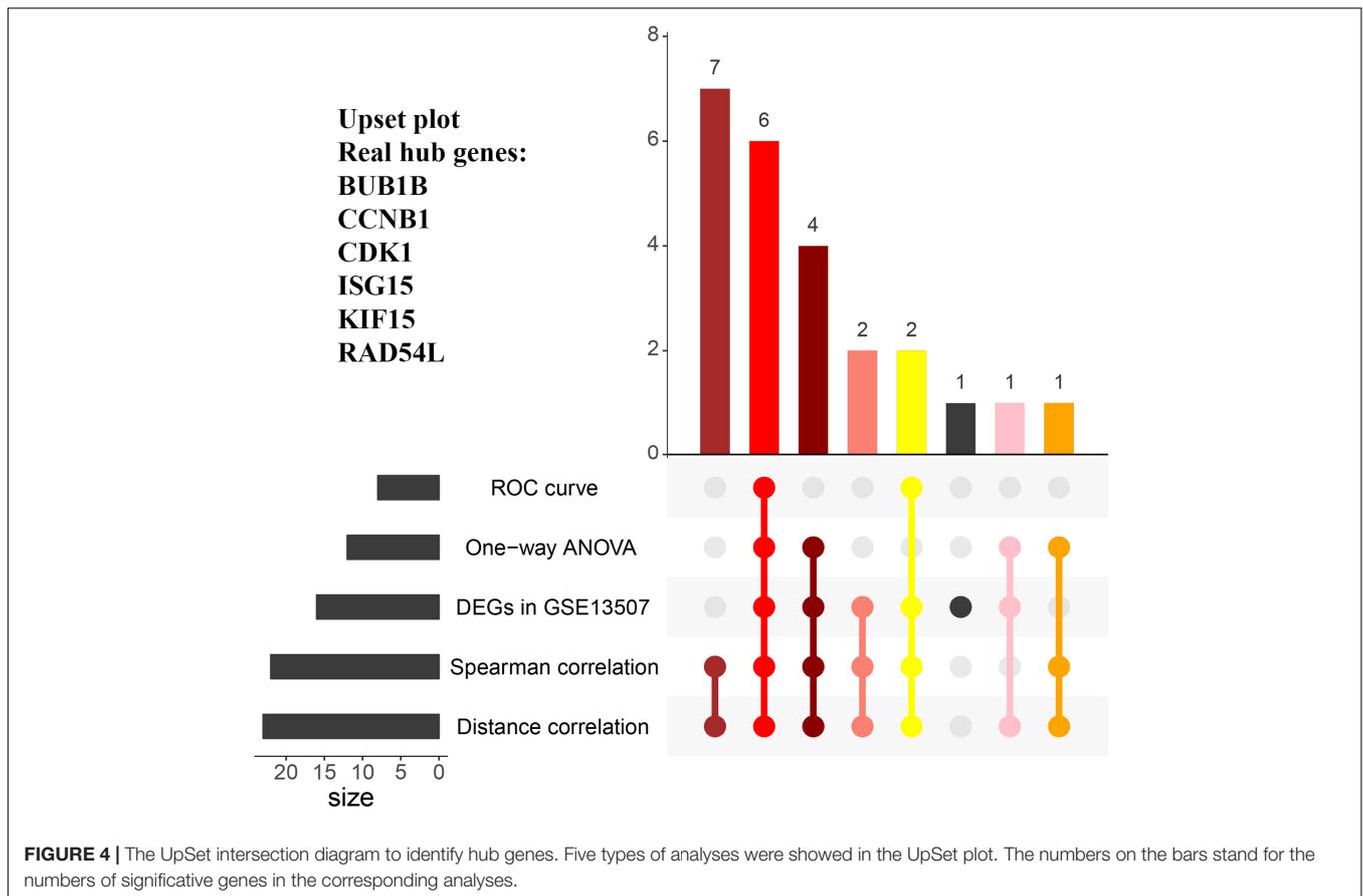
Hub Gene Identification

When exploring the correlation between gene expression levels and tumor grade, 12 genes were identified by one-way ANOVA, 22 genes were picked out by the spearman correlation analysis, 24 genes were screened by the distance correlation analysis ($P < 0.05$, Supplementary Table S3). Then, we performed

ROC curve analysis, eight genes which reached the standard of $AUC \geq 0.85$ were finally screened out (Supplementary Table S3). In addition, 16 genes were differentially expressed in tumor tissues compared with normal tissues based on GSE13507.

In the end, six genes [BUB1B (BUB1 mitotic checkpoint threonine kinase B), CCNB1 (cyclin B1), CDK1 (cyclin-dependent kinase 1), ISG15 (Interferon-stimulated gene 15 kDa protein), KIF15 (Kinesin family member15), and RAD54L (RAD54 like)] were identified as hub genes, because they showed significant P -value in all the five analysis procedures (Figure 4). As shown in Figure 5, the six hub genes were all significantly associated with the progression of BC. Among them, RAD54L might be the biggest factor affecting tumor grade (one-way ANOVA: 56.778, $P < 0.001$; spearman correlation: 0.504, $P < 0.001$; distance correlation: 0.522, $P < 0.001$). The results of ROC curve indicated that RAD54L could distinguish BC samples from normal tissues best, among all the hub genes (BUB1B: $AUC = 0.934$; CCNB1: $AUC = 0.884$; CDK1: $AUC = 0.869$; ISG15: $AUC = 0.908$; KIF15: $AUC = 0.888$; and RAD54L: $AUC = 0.951$, Figure 6). Moreover, based on GSE13507 and TCGA-BC data, the expression levels of these six genes were significantly higher in tumor tissues (Supplementary Figure S4).

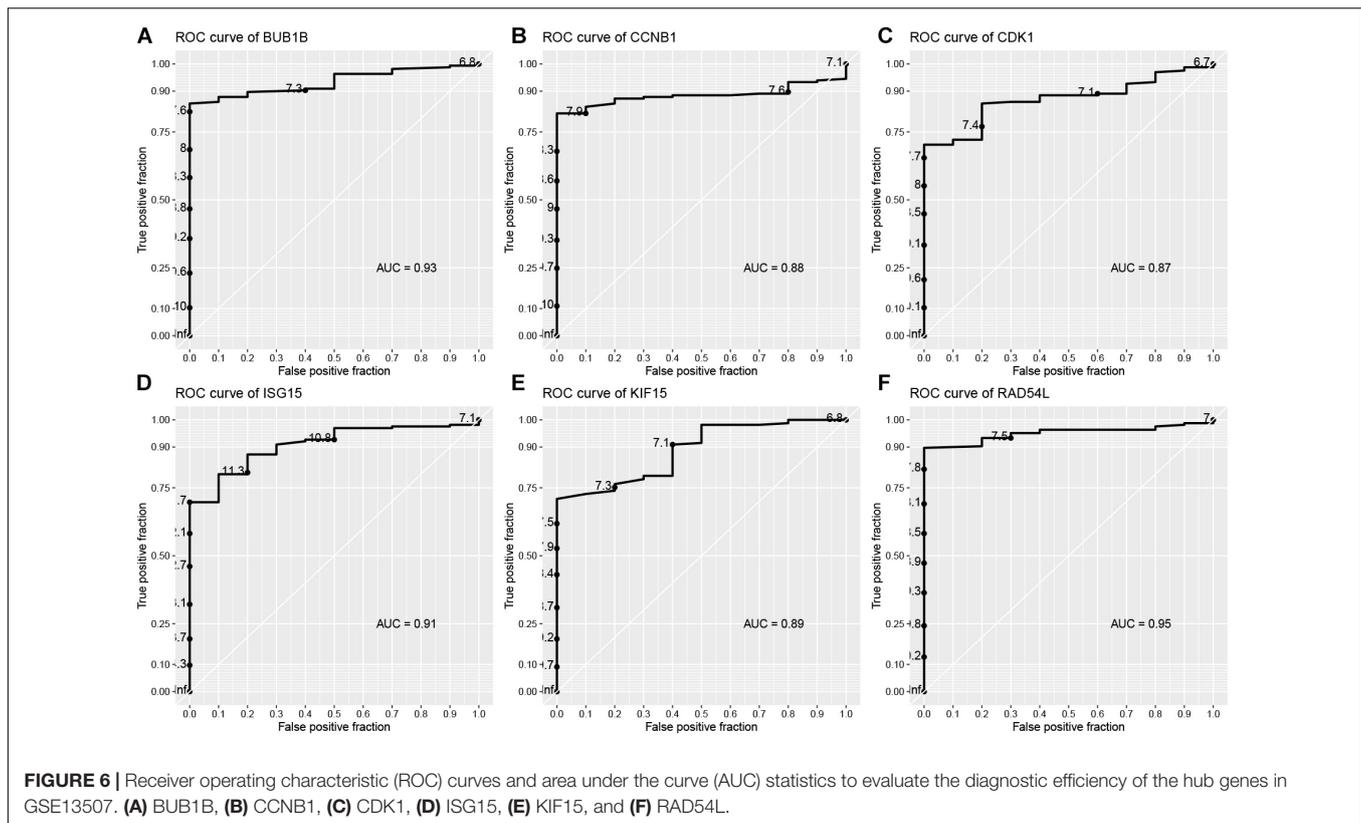




Validation and Genetical Alteration of Hub Genes

Based on GSE13507, the tumor grade and stage boxplots of six hub genes were shown in **Supplementary Figures S5A,B**. The results suggested that all the hub genes were significantly associated with the grade and T stage of tumor. In addition, mRNA expression levels were all significantly higher in tumor tissues compared with those in normal tissues (BUB1B: $t = -8.109$, $P < 0.001$; CCNB1: $t = -9.942$, $P < 0.001$; CDK1: $t = -9.784$, $P < 0.001$; ISG15: $t = -4.008$, $P < 0.001$; KIF15: $t = -3.781$, $P < 0.001$; and RAD54L: $t = -2.944$,

$P = 0.005$; **Supplementary Figure S5C**), which was suggested by the Oncomine database. These results made the six hub genes we screened out reliable. As for genetical alteration, six hub genes were altered in 123 (30%) of 412 patients (**Figure 7B**). As shown in **Figure 7A**, KIF15 altered most (10%) and the main type was mRNA upregulation. A network contained 46 genes (4 hub genes and 42 most variant genes) was showed in **Figure 7C**. As for the relationship between anticancer drugs and hub genes, we found CCNB1 and CDK1 were the targets of cancer drugs. But there was no drug targeting to the rest four hub genes, which might be novel therapeutic targets for patients with BC.



Cox Regression Analysis of OS and CSS Among Patients With BC

Univariate Cox regression analysis showed that almost all the factors we included were influence features of OS and CSS, except gender (Figures 8A,B). Following this we established a six-gene signature, with the risk scores calculated based on the expression levels of the six genes and Cox regression coefficients as follows: risk score of OS = $BUB1B \times 0.347 + CCNB1 \times 0.324 + CDK1 \times 0.218 + ISG15 \times 0.264 + KIF15 \times 0.268 + RAD54L \times 0.386$, risk score of CSS = $BUB1B \times 0.741 + CCNB1 \times 0.507 + CDK1 \times 0.445 + ISG15 \times 0.432 + KIF15 \times 0.508 + RAD54L \times 0.784$. After that, we included the six-gene signature into the multivariate Cox analysis. The results showed that even being adjusted by other factors, risk scores of the six-gene signature were still relevant to OS and CSS among patients with BC (Figures 8C,D).

Survival Analysis of Real Hub Genes and DCA

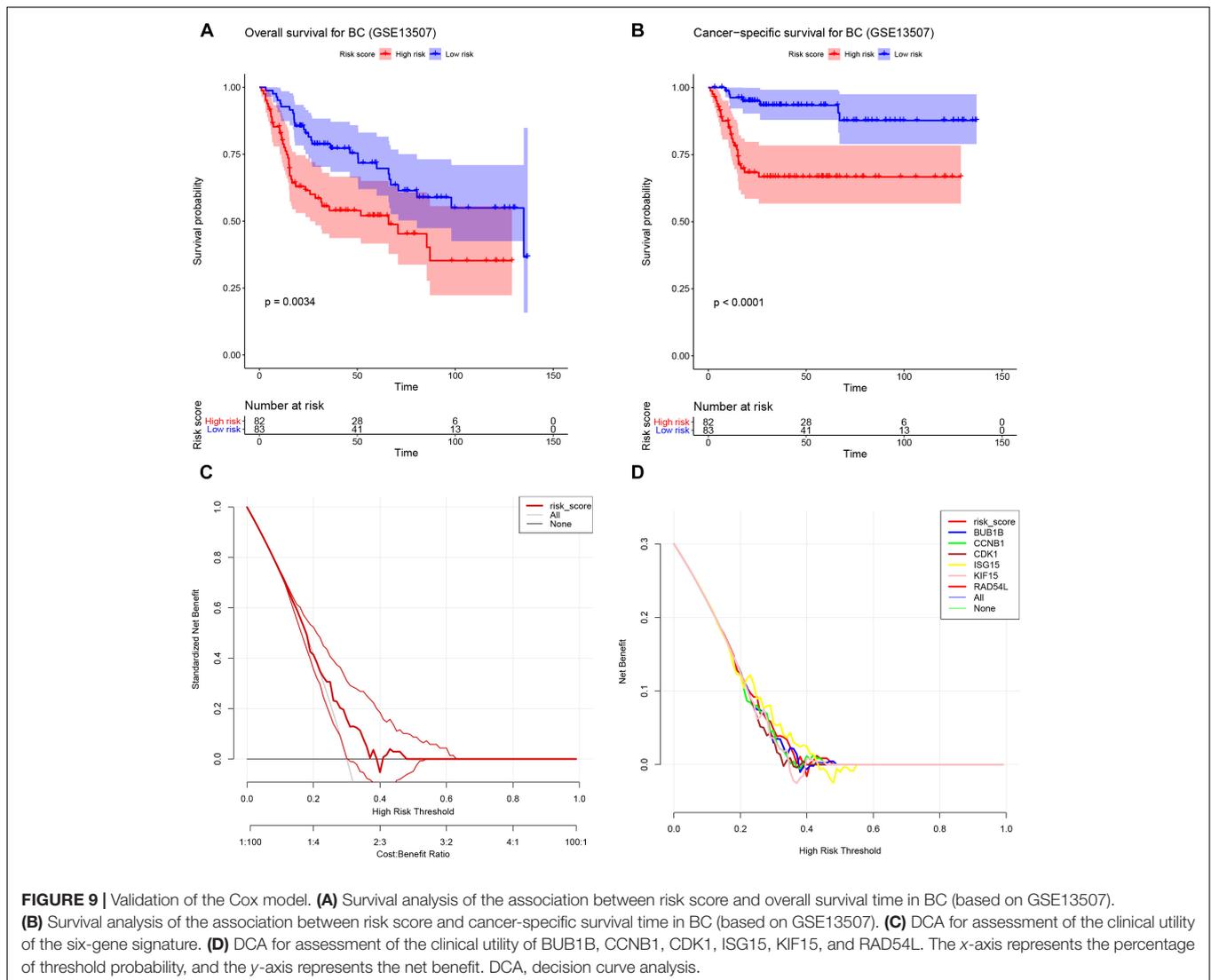
One hundred sixty-five BC patients' prognostic information was obtained from GSE13507. The result suggested that the high-risk group (Hazard Ratio = 0.484, 95%CI of ratio: 0.298–0.786, $P = 0.0034$) had worse OS for patients with BC (Figure 9A). As for CSS analysis, the high-risk group (Hazard Ratio = 0.244, 95%CI of ratio: 0.121–0.493, $P < 0.0001$) was obviously associated with poor CSS for patients with BC (Figure 9B), as shown in Figure 9C, except when then value of Threshold Probability

(Pt) = 0.40; this signature showed high potential, because the Pt (the red thick line in the figure) ensured better net benefits compared with all (gray line in the figure), or none, of the options (black line in the figure). But we could not distinguish this signature from the single gene models because it did not ensure better net benefits compared with others (Figure 9D).

DISCUSSION

Bladder cancer, which is among the leading causes of cancer death globally, can occur at any age (Ebrahimi et al., 2019). Therefore, there is a pressing need for sensitive and novel biomarkers of BC.

After determining our goals, we identified 1,779 DEGs in total. These DEGs were made up of 908 up-regulated genes and 871 down-regulated genes. According to the GO biological processes analysis, the up-regulated DEGs majorly participated in DNA packaging, nucleosome assembly, chromatin assembly, nucleosome organization, DNA conformation change, chromatin assembly or disassembly, DNA replication-dependent nucleosome assembly, DNA replication-dependent nucleosome organization, chromatin silencing at rDNA, and chromatin silencing. As for KEGG analysis, the up-regulated DEGs were relevant to systemic lupus erythematosus, alcoholism, viral carcinogenesis, cell cycle, and neuroactive ligand-receptor interaction; meanwhile down-regulated DEGs were obviously relevant to DCM, HCM, calcium signaling pathway, ARVC, and vascular smooth muscle contraction. On the basis of degree of



that the hub genes we identified were associated with tumor progression, we performed Spearman correlation analysis, one-way ANOVA, and distance correlation analysis by regarding these genes and tumor grade as variables based on GSE13507. In addition, based on TCGA-BLCA data and GSE13507, we compared the expression levels of candidate hub genes in BC and normal bladder tissues to pick out genes highly expressed in BC compared with those in normal tissues. Genes highly expressed in BC might be associated with the happening of BC. By these five analyses, six hub genes (BUB1B, CCNB1, CDK1, ISG15, KIF15, and RAD54L) related to the progression and poor prognosis of BC were finally identified. This meant we tried our best to ensure that the hub genes we screened were disease genes.

In order to validate the six hub genes, we firstly performed stage and grade boxplots based on GSE13507. The results suggested that the high expressions of hub genes were associated with the malignant degree and progression of BC. Secondly, the mRNA expression levels of hub genes were all significantly

higher in tumor tissues compared with those in normal tissues, which demonstrated that the six genes played important roles in the occurrence and progression of BC. Furthermore, in order to validate the prognosis value of the hub genes, we brought 6 factors and the hub genes expression values into Cox regression analysis among BC patients based on GSE13507. The results suggested that expression levels of all the six hub genes were associated with OS and CSS among patients with BC. We then established a six-gene signature and the risk scores were calculated combining the expression levels of the six genes and Cox regression coefficients. The Cox multivariate analysis showed that risk score, N stage, and M stage were relevant to OS and CSS among patients with BC. In order to verify the prognosis value of this six-gene signature we performed survival analysis and the results showed that the high-risk group was obviously associated with poor OS and CSS for patients with BC.

With the development of bioinformatics and high-throughput sequencing, studies indicated that small molecules might have

a beneficial or detrimental effect against diseases (Qin et al., 2017). This made it possible to regard genes as novel therapeutic targets (Hansen et al., 2019; Song et al., 2019). So that we used CBioPortal to explore the relationship between hub genes and drugs aiming at finding new targets for anticancer drugs in this study. We found that CCNB1 and CDK1 were already the targets for anticancer drugs, which meant the remaining genes (BUB1B, ISG15, KIF15, and RAD54L) might become potent drug targets. CCNB1 showed higher expression in most tumor cells compared with normal cells (Fang et al., 2014). This caused deficiencies in MPF (maturation promoting factor) phosphorylation regulation mechanism (Zhang H. et al., 2018). In order to carry on the anti-tumor treatment, medicines inhibited the function of MPF through targeting CCNB1 to prevent cell mitosis (Egloff et al., 2006). Some recent studies also thought CCNB1 was a drug target. Freitas et al. (2016) thought hierridin B was a potential anticancer compound that targeted CCNB1. A study in breast cancer confirmed that targeting CCNB1 was useful in BRCA1-associated breast cancer therapy (Choi et al., 2018). In clinical, Resveratrol (Joe et al., 2002), quercetin (Choi et al., 2001), and genistein (Wiseman et al., 2007) were the representative targeted drugs. As for CDK1, it was a co-chaperone of CCNB1 (Wang et al., 2014). CDK1 could form complexes with cyclin B1 (Wang et al., 2014). CDK1/cyclin B1 complexes not only played an important role in cell division, but also increased the activity of strong mitochondria (Wang et al., 2014). According to the results of recent research, targeting CDK1 could overcome apoptotic resistance in patients with colorectal cancer (Zhang P. et al., 2018). In clinical there existed many targeted drugs for CDK1, most of which belonged to CDK inhibitors. The famous drugs were flavopiridol (Prithviraj et al., 2013) and palbociclib (Todd et al., 2015). Compared with the first-generation inhibitors, palbociclib had better selectivity, higher therapeutic index, and less side effects (Mariaule and Belmont, 2014). This suggested that improving the selectivity of CDK inhibitors was the key to improving the treatment index. For a deeper and better understanding of the remaining four genes, a literature review was carried out. BUB1B, also known as BUBR1, was an important functional protein of mitotic detection point (Pinto et al., 2014). The changes of BUB1B expression played an important role in tumorigenesis and progression (Myrie et al., 2000). ISG15 was up-regulated in the uterus, corpus luteum and liver during early pregnancy in animals as reported (Zhang L. et al., 2018). KIF15, a member of kinesin superfamily protein, could promote cell mitotic and participate in cellular material transportation (Woehlke and Schliwa, 2000). RAD54L, encoded by gene RAD54L, was shown to play an important role in homologous recombination related repair or DNA double-strand breaks (Rencic et al., 1996). In summary, we found that almost all the six hub genes were associated with cell cycle and mitosis.

Some limitations of this study also should be discussed. Firstly, lack of experiments (*in vivo* and *in vitro* validation) might be one limitation of our study. Secondly, according to the results, the six hub genes were all up-regulated in BC, but the mechanism of up-regulation was not clear. There might need more evidences to find

out the biological basis. Therefore, further molecular biological experiments are needed to confirm the function of these hub genes and how they perform their roles in the progression of BC.

CONCLUSION

In conclusion, by using a series of bioinformatics and retrospective analyses, the present study identified six hub genes (BUB1B, CCNB1, CDK1, ISG15, KIF15, and RAD54L), which were significantly associated with progression and prognosis of BC. These hub genes were all up-regulated in BC and four of them might be novel drug targets. Further and more in-depth study is necessary to confirm the results of the study. In any case, our study could provide some strong basis for gene targeting therapy of BC in the future and the six hub genes might be potential and novel target genes for BC.

AUTHOR CONTRIBUTIONS

SL, X-PL, and XY reviewed the relevant literature and drafted the manuscript. XY, X-PL, Z-XG, and T-ZL conducted all the statistical analyses. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (81802541), the Zhongnan Hospital of Wuhan University Science, Technology and Innovation Seed Fund (znp2017050), and the 351 Talent Project of Wuhan University (LuoJia Young Scholars: SL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00408/full#supplementary-material>

FIGURE S1 | Volcano plot visualizing DEGs in TCGA-BC data.

FIGURE S2 | The whole Protein-protein interaction (PPI) network of DEGs. Red nodes: Up-regulated DEGs. Green nodes: Down-regulated DEGs.

FIGURE S3 | A random forest plot: genes with MeanDecreaseAccuracy ranked top 30 and genes with MeanDecreaseGini ranked top 30.

FIGURE S4 | Expression levels of hub genes in BC and normal tissues based on GSE13507 (A) and GEPIA (B).

FIGURE S5 | Validation of the six hub genes. (A) Grade plot of the hub genes using GSE13507. (B) Stage plot of the hub genes using GSE13507. (C) mRNA expression levels of hub genes suggested by Oncomine database.

TABLE S1 | Differentially expressed genes (DEGs) between BC samples and normal bladder tissues.

TABLE S2 | Common genes in MeanDecreaseAccuracy and MeanDecreaseGini ranked top 50.

TABLE S3 | One-way ANOVA analysis, spearman correlation analysis, distance correlation analysis, and AUC of candidate hub genes.

REFERENCES

- Aut, M. G., and Aut, T. L. (2016). *forestplot: Advanced Forest Plot Using 'grid' Graphics*. Available at: <http://CRAN.R-project.org/package=forestplot> (accessed September 1, 2017).
- Bi, D., Ning, H., Liu, S., Que, X., and Ding, K. (2015). Gene expression patterns combined with network analysis identify hub genes associated with bladder cancer. *Comput. Biol. Chem.* 56, 71–83. doi: 10.1016/j.compbiolchem.2015.04.001
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2, 401–404. doi: 10.1158/2159-8290.CD-12-0095
- Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer statistics in China, 2015. *CA Cancer J. Clin.* 66, 115–132. doi: 10.3322/caac.21338
- Choi, E. K., Lim, J. A., Kim, J. K., Jang, M. S., Kim, S. E., Baek, H. J., et al. (2018). Cyclin B1 stability is increased by interaction with BRCA1, and its overexpression suppresses the progression of BRCA1-associated mammary tumors. *Exp. Mol. Med.* 50:136. doi: 10.1038/s12276-018-0169-z
- Choi, J. A., Kim, J. Y., Lee, J. Y., Kang, C. M., Kwon, H. J., Yoo, Y. D., et al. (2001). Induction of cell cycle arrest and apoptosis in human breast cancer cells by quercetin. *Int. J. Oncol.* 19, 837–844. doi: 10.3892/ijo.19.4.837
- Conway, J. R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33, 2938–2940. doi: 10.1093/bioinformatics/btx364
- Ebrahimi, H., Amini, E., Pishgar, F., Moghaddam, S. S., Nabavizadeh, B., Rostamabadi, Y., et al. (2019). Global, regional, and national burden of bladder cancer, 1990–2016: results from the global burden of disease study 2016. *J. Urol.* doi: 10.1097/JU.000000000000025 [Epub ahead of print].
- Egloff, A. M., Vella, L. A., and Finn, O. J. (2006). Cyclin B1 and other cyclins as tumor antigens in immunosurveillance and immunotherapy of cancer. *Cancer Res.* 66, 6–9. doi: 10.1158/0008-5472.CAN-05-3389
- Emerson, R. E., and Cheng, L. (2005). Immunohistochemical markers in the evaluation of tumors of the urinary bladder: a review. *Anal. Quant. Cytol. Histol.* 27:301. doi: 10.1159/000325904
- Fang, Y., Yu, H., Liang, X., Xu, J., and Cai, X. (2014). Chk1-induced CCNB1 overexpression promotes cell proliferation and tumor growth in human colorectal cancer. *Cancer Biol. Ther.* 15, 1268–1279. doi: 10.4161/cbt.29691
- Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* 127, 2893–2917. doi: 10.1002/ijc.25516
- Freitas, S., Martins, R., Costa, M., Leao, P. N., Vitorino, R., Vasconcelos, V., et al. (2016). Hierridin B isolated from a marine cyanobacterium alters vdaC1, mitochondrial activity, and cell cycle genes on HT-29 colon adenocarcinoma cells. *Mar. Drugs* 14:E158. doi: 10.3390/md14090158
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioportal. *Sci. Signal.* 6:11. doi: 10.1126/scisignal.2004088
- Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy - analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315. doi: 10.1093/bioinformatics/btg405
- Hansen, A. L., Mukai, K., Schopfer, F. J., Taguchi, T., and Holm, C. K. (2019). STING palmitoylation as a therapeutic target. *Cell. Mol. Immunol.* 16, 236–241. doi: 10.1038/s41423-019-0205-5
- Harris, R. L., Cundiff, G. W., Theofrastous, J. P., Yoon, H., Bump, R. C., and Addison, W. A. (1997). The value of intraoperative cystoscopy in urogynecologic and reconstructive pelvic surgery. *Am. J. Obst. Gynecol.* 177, 1369–1371.
- Jia, Z., Ai, X., Sun, F., Zang, T., Guan, Y., and Gao, F. (2015). Identification of new hub genes associated with bladder carcinoma via bioinformatics analysis. *Tumori* 101, 117–122. doi: 10.5301/tj.5000196
- Joe, A. K., Liu, H., Suzui, M., Vural, M. E., Xiao, D., and Weinstein, I. B. (2002). Resveratrol induces growth inhibition, S-phase arrest, apoptosis, and changes in biomarker expression in several human cancer cell lines. *Clin. Cancer Res.* 8, 893–903. doi: 10.1159/000048589
- Kim, W. J., Kim, E. J., Kim, S. K., Kim, Y. J., Ha, Y. S., Jeong, P., et al. (2010). Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Mol. Cancer* 9:3. doi: 10.1186/1476-4598-9-3
- Kiyoshima, K., Akitake, M., Shiota, M., Takeuchi, A., Takahashi, R., Inokuchi, J., et al. (2016). Prognostic significance of preoperative urine cytology in low-grade non-muscle-invasive bladder cancer. *Anticancer Res.* 36, 799–802.
- Lee, J. S., Leem, S. H., Lee, S. Y., Kim, S. C., Park, E. S., Kim, S. B., et al. (2010). Expression signature of E2F1 and its associated genes predict superficial to invasive progression of bladder tumors. *J. Clin. Oncol.* 28, 2660–2667. doi: 10.1200/JCO.2009.25.0977
- Li, M. X., Jin, L. T., Wang, T. J., Feng, Y. J., Pan, C. P., Zhao, D. M., et al. (2018). Identification of potential core genes in triple negative breast cancer using bioinformatics analysis. *Onco. Targets Ther.* 11, 4105–4112. doi: 10.2147/OTT.S166567
- Liaw, A., and Wiener, M. (2002). Classification and regression by random forest. *R. News* 2, 18–22.
- Mariaule, G., and Belmont, P. (2014). Cyclin-dependent kinase inhibitors as marketed anticancer drugs: where are we now? A short survey. *Molecules* 19, 14366–14382. doi: 10.3390/molecules190914366
- Matsuzaki, K., Fujita, K., Ujike, T., Uemura, M., Nonomura, N., and Miyagawa, Y. (2016). Is it necessary to carry out intraoperative retrograde upper urinary tract cytology examination in bladder cancer patients with normal upper urinary tract appearance and suspicious or positive voided urine cytology? *Int. J. Urol.* 23, 623–624. doi: 10.1111/iju.13100
- Miakhil, I., Parker, S. G., Kommu, S. S., and Nethercliffe, J. (2013). A review of molecular biomarkers for bladder cancer. *Int. J. Med. Biomed. Res.* 2, 186–194. doi: 10.14194/ijmbr.235
- Myrie, K. A., Percy, M. J., Azim, J. N., Neeley, C. K., and Petty, E. M. (2000). Mutation and expression analysis of human BUB1 and BUB1B in aneuploid breast cancer cell lines. *Cancer Lett.* 152, 193–199. doi: 10.1016/S0304-3835(00)00340-2
- Patil, I., and Powell, C. (2018). *ggstatsplot: 'ggplot2' Based Plots with Statistical Details*. Available at: <https://CRAN.R-project.org/package=ggstatsplot> (accessed March 17, 2019).
- Pinto, M., Vieira, J., Ribeiro, F. R., Soares, M. J., Henrique, R., Oliveira, J., et al. (2014). Overexpression of the mitotic checkpoint genes BUB1 and BUBR1 is associated with genomic complexity in clear cell kidney carcinomas. *Cell. Oncol.* 30, 389–395. doi: 10.3233/CLO-2008-0439
- Prithviraj, B., Simmons, G. L., and Steven, G. (2013). Cyclin-dependent kinase inhibitor therapy for hematologic malignancies. *Expert Opin. Inv. Drug.* 22, 723–738. doi: 10.1517/13543784.2013.789859
- Qin, H., Zhao, A., and Fu, X. (2017). Small molecules for reprogramming and transdifferentiation. *Cell. Mol. Life Sci.* 74, 3553–3575. doi: 10.1007/s00018-017-2586-x
- Rencic, A., Gordon, J., Otte, J., Curtis, M., Kovatich, A., Zoltick, P., et al. (1996). Detection of JC virus DNA sequence and expression of the viral oncoprotein, tumor antigen, in brain of immunocompetent patient with oligoastrocytoma. *Proc. Natl. Acad. Sci. U.S.A.* 93, 7352–7357.
- Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., et al. (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 6, 1–6. doi: 10.1016/S1476-5586(04)80047-2
- Rigatti, S. J. (2017). Random forest. *J. Insur. Med.* 47, 31–39. doi: 10.17849/insm-47-01-31-39.1
- Rizzo, M. L., and Szekely, G. J. (2016). *energy: E-Statistics: Multivariate Inference Via the Energy of Data*. Available at: <https://cran.r-project.org/web/packages/energy/index.html> (accessed August 11, 2008).
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Sachs, M. C. (2015). plotROC: a tool for plotting ROC curves. *J. Stat. Softw. Code Snipp.* 79, 1–19. doi: 10.18637/jss.v079.c02
- Sanchez-Carbayo, M., Socci, N. D., Lozano, J., Saint, F., and Cordon-Cardo, C. (2006). Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *J. Clin. Oncol.* 24, 778–789. doi: 10.1200/JCO.2005.03.2375
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of

- biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Song, M., Bode, A. M., Dong, Z., and Lee, M. H. (2019). AKT as a therapeutic target for cancer. *Cancer Res.* 79, 1019–1031. doi: 10.1158/0008-5472.CAN-18-2738
- Sun, C., Yuan, Q., Wu, D., Meng, X., and Wang, B. (2017). Identification of core genes and outcome in gastric cancer using bioinformatics analysis. *Oncotarget* 8, 70271–70280. doi: 10.18632/oncotarget.20082
- Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., and Feuston, B. P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* 43, 1947–1958. doi: 10.1021/ci034160g
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huertacepas, J., et al. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003
- Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx247
- Therneau, T. M. (2015). Survival: survival analysis. *Technometrics* 46, 111–112.
- Tilki, D., Burger, M., Dalbagni, G., Grossman, H. B., Hakenberg, O. W., Palou, J., et al. (2011). Urine markers for detection and surveillance of non-muscle-invasive bladder cancer. *Eur. Urol.* 60, 484–492. doi: 10.1016/j.eururo.2011.05.053
- Todd, V. A., Chris, B., Arndt, K. T., and Abraham, R. T. (2015). Molecular pathways: targeting the cyclin D-CDK4/6 axis for cancer treatment. *Clin. Cancer Res.* 21, 2905–2910. doi: 10.1158/1078-0432.CCR-14-0816
- Wang, Z., Fan, M., Candas, D., Zhang, T. Q., Qin, L., Eldridge, A., et al. (2014). Cyclin B1/Cdk1 coordinates mitochondrial respiration for cell-cycle G2/M progression. *Dev. Cell* 29, 217–232. doi: 10.1016/j.devcel.2014.03.012
- Wickham, H. (2015). ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. doi: 10.1002/wics.147
- Wiseman, D. A., Werner, S. R., and Crowell, P. L. (2007). Cell cycle arrest by the isoprenoids perillyl alcohol, geraniol, and farnesol is mediated by p21(Cip1) and p27(Kip1) in human pancreatic adenocarcinoma cells. *J. Pharmacol. Exp. Ther.* 320, 1163–1170. doi: 10.1124/jpet.106.111666
- Woehlke, G., and Schliwa, M. (2000). Walking on two heads: the many talents of kinesin. *Nat. Rev. Mol. Cell Bio.* 1, 50–58. doi: 10.1038/35036069
- 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
- Yuan, L., Qian, G., Chen, L., Wu, C. L., Dan, H. C., Xiao, Y., et al. (2018). Co-expression network analysis of biomarkers for adrenocortical carcinoma. *Front. Genet.* 9:328. doi: 10.3389/fgene.2018.00328
- Zhang, D. Q., Zhou, C. K., Chen, S. Z., Yang, Y., and Shi, B. K. (2017). Identification of hub genes and pathways associated with bladder cancer based on co-expression network analysis. *Oncol. Lett.* 14, 1115–1122. doi: 10.3892/ol.2017.6267
- Zhang, H., Zhang, X., Li, X., Meng, W. B., Bai, Z. T., Rui, S. Z., et al. (2018). Effect of CCNB1 silencing on cell cycle, senescence, and apoptosis through the p53 signaling pathway in pancreatic cancer. *J. Cell. Physiol.* 234, 619–631. doi: 10.1002/jcp.26816
- Zhang, L., Xue, J., Wang, Q., Lv, W., Mi, H., Liu, Y., et al. (2018). Changes in expression of ISG15, progesterone receptor and progesterone-induced blocking factor in ovine thymus during early pregnancy. *Theriogenology* 121, 153–159. doi: 10.1016/j.theriogenology.2018.08.018
- Zhang, P., Kawakami, H., Liu, W., Zeng, X., Strebhardt, K., Tao, K., et al. (2018). Targeting CDK1 and MEK/ERK overcomes apoptotic resistance in BRAF-mutant human colorectal cancer. *Mol. Cancer Res.* 16, 378–389. doi: 10.1158/1541-7786.MCR-17-0404
- Zhao, X., and Liu, Z. P. (2019). Analysis of topological parameters of complex disease genes reveals the importance of location in a biomolecular network. *Genes* 10:E143. doi: 10.3390/genes10020143

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Yan, Liu, Guo, Liu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.