



Identification and Validation of SNP-Containing Genes With Prognostic Value in Gastric Cancer via Integrated Bioinformatics Analysis

Hui Li^{1†}, Jing Guo^{2†}, Guang Cheng², Yucheng Wei², Shihai Liu³, Yaoyue Qi¹, Gongjun Wang¹, Ruoxi Xiao¹, Weiwei Qi^{2*} and Wensheng Qiu^{2*}

¹ Department of Medicine, Qingdao University, Qingdao, China, ² Department of Oncology, Affiliated Hospital of Qingdao University, Qingdao, China, ³ Central Laboratory, Affiliated Hospital of Qingdao University, Qingdao, China

OPEN ACCESS

Edited by:

Fenglin Liu,
Fudan University, China

Reviewed by:

Amit Gupta,
All India Institute of Medical
Sciences, Rishikesh, India
Bianca Troncarelli Flores,
Portuguese Oncology Institute,
Portugal

*Correspondence:

Weiwei Qi
qwwdz@126.com
Wensheng Qiu
wsqiuqd@163.com

[†]These authors share first authorship

Specialty section:

This article was submitted to
Gastrointestinal Cancers,
a section of the journal
Frontiers in Oncology

Received: 21 May 2020

Accepted: 07 April 2021

Published: 27 April 2021

Citation:

Li H, Guo J, Cheng G, Wei Y, Liu S, Qi Y, Wang G, Xiao R, Qi W and Qiu W (2021) Identification and Validation of SNP-Containing Genes With Prognostic Value in Gastric Cancer via Integrated Bioinformatics Analysis. *Front. Oncol.* 11:564296. doi: 10.3389/fonc.2021.564296

Background: Gastric cancer is one of the most common malignancies worldwide. Although the diagnosis and treatment of this disease have substantially improved in recent years, the five-year survival rate of gastric cancer is still low due to local recurrence and distant metastasis. An in-depth study of the molecular pathogenesis of gastric cancer and related prognostic markers will help improve the quality of life and prognosis of patients with this disease. The purpose of this study was to identify and verify key SNPs in genes with prognostic value for gastric cancer.

Methods: SNP-related data from gastric cancer patients were obtained from The Cancer Genome Atlas (TCGA) database, and the functions and pathways of the mutated genes were analyzed using DAVID software. A protein-protein interaction (PPI) network was constructed using the STRING database and visualized by Cytoscape software, and molecular complex detection (MCODE) was used to screen the PPI network to extract important mutated genes. Ten hub genes were identified using cytoHubba, and the expression levels and the prognostic value of the central genes were determined by UALCAN and Kaplan-Meier Plotter. Finally, quantitative PCR and Western blotting were used to verify the expression of the hub genes in gastric cancer cells.

Results: From the database, 945 genes with mutations in more than 25 samples were identified. The PPI network had 360 nodes and 1616 edges. Finally, cytoHubba identified six key genes (TP53, HRAS, BRCA1, PIK3CA, AKT1, and SMARCA4), and their expression levels were closely related to the survival rate of gastric cancer patients.

Conclusion: Our results indicate that TP53, HRAS, BRCA1, PIK3CA, AKT1, and SMARCA4 may be key genes for the development and prognosis of gastric cancer. Our research provides an important bioinformatics foundation and related theoretical foundation for further exploring the molecular pathogenesis of gastric cancer and evaluating the prognosis of patients.

Keywords: bioinformatics analysis, biomarkers, gastric cancer, prognosis, single nucleotide polymorphisms

INTRODUCTION

Gastric cancer is the fifth most common malignant tumor worldwide and the second leading cause of cancer-related death (1). Although the treatment strategies for gastric cancer have substantially improved in recent years, the mortality rate is still high due to various genetic mutations and abnormal signaling pathways underlying the progression of this disease (2). The occurrence and development of gastric cancer, as a complex disease, involves a series of genetic, epigenetic and phenotypic changes. Gene polymorphisms involved in multiple biological pathways have been identified as potential risk factors for gastric cancer (3). Given the high morbidity and mortality of gastric cancer, identification of its underlying molecular mechanism and genetic characteristics and elucidation of biological indicators for diagnosis and prognosis are essential for the personalized and precise treatment of gastric cancer patients.

Bioinformatics analysis based on high-throughput sequencing is an important method for exploring the molecular mechanism of tumor pathogenesis, identifying biomarkers that can be used for early diagnosis, and discovering therapeutic targets. Single nucleotide polymorphisms (SNPs) are nucleotide polymorphisms that are commonly found in the genome of an organism; among individuals of different species, single nucleotides in the same position of the genomic DNA sequence undergo substitution, insertion or deletion and other mutations, resulting in a single nucleotide change at this site (4). SNPs are generally considered to be the genetic basis of and potential cancer markers that lead to differences in the individual susceptibility to disease. Analysis of SNP-containing genes is important for the early diagnosis and individualized targeted treatment of cancer.

The Cancer Genome Atlas (TCGA) database can be used for high-throughput genomic analysis. To further explore the biological significance of SNP-containing genes in the diagnosis and prognosis of gastric cancer, we downloaded gastric cancer-related SNP data from TCGA database and used bioinformatics analyses, including mutation analysis, function and pathway enrichment analyses, protein-protein interaction (PPI) network analysis and related analyses. Our aim was to explore the mutated genes related to the diagnosis and prognosis of gastric cancer and to provide a scientific theoretical basis for personalized and precise treatment of gastric cancer.

MATERIALS AND METHODS

Data Processing and Analysis

The TCGA data portal was terminated, and all TCGA data were transferred to the newly established genomic data sharing platform (<https://gdc.cancer.gov/>) (5). Since the original data on SNPs in TCGA have not yet been opened to the public, we downloaded the processed SNP-related data of gastric cancer and the original mRNA expression data. The mRNA data were collected from 413 samples (including 32 normal samples and 381 cancer samples). The mutated genes were obtained from the

SNP data of the downloaded gastric cancer samples. The edgeR software package was used to integrate and standardize the downloaded mRNA raw data, and analysis was performed to obtain the differentially expressed genes and their expression levels. The mRNA data provided by TCGA are publicly available, and thus, no approval from the local ethics committee was required.

Functional Enrichment and Pathway Analysis of Mutated Genes

To elucidate the dysfunction caused by these mutated genes, we used the DAVID (<http://www.DAVID.org>) (6) database to perform Gene Ontology (GO) and Kyoto Gene and Genome Encyclopedia (KEGG) analyses on genes from more than 25 mutated samples. As an open source platform, DAVID can be used to determine the association between target molecules. By selecting the GO term and the KEGG pathway and using $P < 0.05$ as the cut-off value, we identified the molecular functions (MFs), biological processes (BPs), cellular components (CCs) and KEGG pathways of the enriched mutated genes.

Construction of the PPI Network of the Mutated Genes and Gene Expression Analysis

The STRING database (<http://string-db.org/>) provides the significant PPIs (7). Cytoscape is used for visual exploration of interactive networks (8). In this study, the STRING database was used to perform PPI network analysis of the selected SNP-containing genes, and then, Cytoscape visualization was used, with a confidence score > 0.4 as the cut-off criterion. The Cytoscape plug-in cytoHubba (9) was used to identify the hub genes by finding the intersections of the first 30 genes from 12 topological analysis methods and then using molecular complex detection (MCODE) to establish the module of the PPI network, with a degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max depth = 100.17 (10).

Kaplan-Meier Survival Curve of the Mutated Genes and Screening of Prognostic Biomarkers

Kaplan-Meier Plotter (<https://kmplot.com/analysis/>) can use gene expression data to assess the survival rate of breast, lung, gastric and ovarian cancer patients. Recurrence-free survival (RFS) and overall survival (OS) data were downloaded from GEO (Affymetrix microarray only), EGA and TCGA. The main purpose of this tool is biomarker evaluation based on meta-analysis (11). Using the Kaplan-Meier chart, we evaluated the effects of the mutated genes on the prognosis of gastric cancer patients and finally identified genes that can be used as prognostic biomarkers for this disease.

Hub Gene Verification Through UALCAN

UALCAN (<http://ualcan.path.uab.edu/>) is a web-based tool that can provide fast and customizable functions based on level 3 RNA-seq and clinical data of 31 cancer types from TCGA database (12). In this study, the UALCAN database was used

to verify the expression of the central genes identified in the module between normal and cancer samples. We chose $P < 0.05$ and fold change > 2 as the threshold.

Analysis of Cancer Genomics Data Through cBioPortal

cBioPortal for Cancer Genomics (<http://cbiportal.org>) provides resources for visualizing and analyzing multidimensional cancer genomics data (13). In this study, based on mutations and changes in the DNA copy number of the four selected subtypes of gastric cancer, we performed an analysis of the genomic changes in pivotal genes.

Cell Culture and Antibodies

AGS, HGC27 and GES-1 cell lines were purchased from the Cell Bank of the Chinese Academy of Sciences and cultured in RPMI-1640 medium containing 10% fetal bovine serum (FBS). FBS and RPMI-1640 were purchased from Gibco (NY, USA). The cells were placed in an incubator at 37°C and a CO₂ concentration of 5%. Antibodies against TP53 (#2527), BRCA1 (#14823), PIK3CA (#4255), AKT1 (#2938), and SMARCA4 (#49360) were purchased from Cell Signaling Technology (Beverly, MA, USA). Antibodies against HRAS (abs137096) and beta actin antibody (abs132001) were obtained from Absin (Shanghai, China). The secondary antibodies used in this study include goat anti-mouse IgG-HRP (abs20001) and goat anti-rabbit IgG-HRP (abs20002), both of which can be obtained from Absin.

qPCR for Detection of the Expression Levels of the Hub Genes

According to the manufacturer's instructions, total RNA was isolated from cells using TRIzol reagent (TaKaRa, Beijing, China) and reverse transcribed into cDNA using PrimeScript RT Master Mix reagent (TaKaRa, Beijing, China). Quantitative real-time PCR (qRT-PCR) was performed using the ABI 7500HT Fast real-time PCR system (Applied Biosystems, California, USA), and then, melting curve analysis was performed. The following cycling conditions were used: 95°C for five minutes, followed by 40 cycles of 95°C for 20 seconds and 60°C for 30 seconds. We used the $2^{-\Delta\Delta Ct}$ method, with GAPDH as an

internal control, to determine the average relative fold change in mRNA expression. The primers are shown in **Table 1**.

Western Blot Analysis

The gastric cancer cells were inoculated into a 6 cm Petri dish, treated for 48 hours, scraped and collected. The cells were dissolved on ice in PMSF-containing RIPA buffer, and then, the mixture was centrifuged at 13,000 x g at 4°C for 5 minutes to remove the cell debris. The supernatant was collected, and the total protein concentration was determined using the BCA protein assay kit. Approximately 20 µg of protein was separated by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The protein was wet transferred to a 0.22 µm polyvinylidene fluoride (PVDF) membrane using a constant current of 300 mA, blocked with 5% skim milk powder in TBST for 2 hours and incubated overnight with the appropriate primary antibody (1:1000). The next day, the membrane was washed 3 times with TBST for 10 minutes. At room temperature, the membrane was incubated with the HRP-conjugated secondary antibody (1:8000) for 2 hours and washed with TBST 3 times for 10 minutes each time. A chemiluminescence kit (Life Technologies, Shanghai, China) was used to observe the bound antibody under the Bio-Rad gel imager infrared imaging system (ChemiDoc XRS +).

Statistical Methods

Data are means \pm standard deviation. Two-tailed unpaired Student's *t* tests were used to assess significance unless stated otherwise. $P < 0.05$ was deemed significant.

RESULTS

Data Processing and Analysis

Using the VarScan method to extract germ/somatic cell mutation data of gastric cancer samples from the second-generation sequencing data of TCGA database as SNP data, we selected 945 genes with mutations in more than 25 samples. Among these genes, 96 genes were mutated in more than 50 samples (**Figure 1**). From TCGA database, 413 samples with gastric cancer gene expression data, including 32 normal tissue samples and 381 cancer tissue samples, were obtained. The patient characteristics are in **Supplementary Table 1**. The edgeR software package was used to analyze the differential expression between the gastric cancer and normal tissue samples (**Figure 2**), with $|\log FC| > 2$ and $P < 0.01$ as the cut-off criteria. We further analyzed the SNP-containing genes and differentially expressed genes in gastric cancer to explore the dysfunction caused by gene mutations and abnormal expression.

Functional Enrichment and Pathway Analysis of the Mutated Genes

To further elucidate the functions of the mutated genes in gastric cancer, we used DAVID online software to perform functional enrichment analysis and pathway analysis on the 945 genes with mutations in more than 25 samples. Functional enrichment

TABLE 1 | The primer of hub genes.

Primer name	Sense	Antisense
TP53	GAGGTTGGCTCTGACTGTACC	TCCGTCCCAG TAGATTACCAC
HRAS	GACGTGCCTGTTGGACATC	CTTCACCGTT TGATCTGCTC
BRCA1	GAAACCGTGCCAAAAGACTTC	CCAAGGTTAGA GAGTTGGACAC
PIK3CA	AGTAGGCAACCGTGAAGAAAAG	GAGGTGAATTG AGGTCCCTAAGA
AKT1	AGCGACGTGGCTATTGTGAAG	GCCATCATTCTT GAGGAGGAAGT
SMARCA4	GAAACAAGACGACTTTGTGACCT	CTTCACGGTTG CCTACTGGT

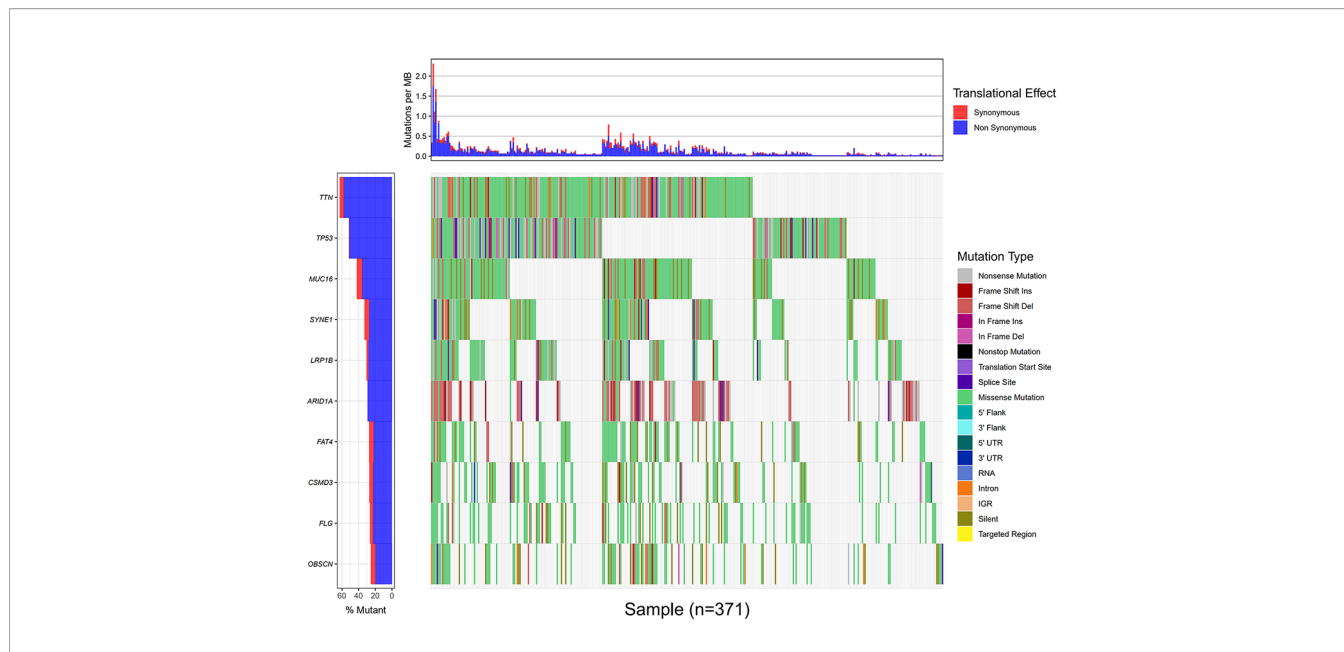


FIGURE 1 | Data processing and analysis. A waterfall map of 10 genes that were mutated in more than 90 samples.

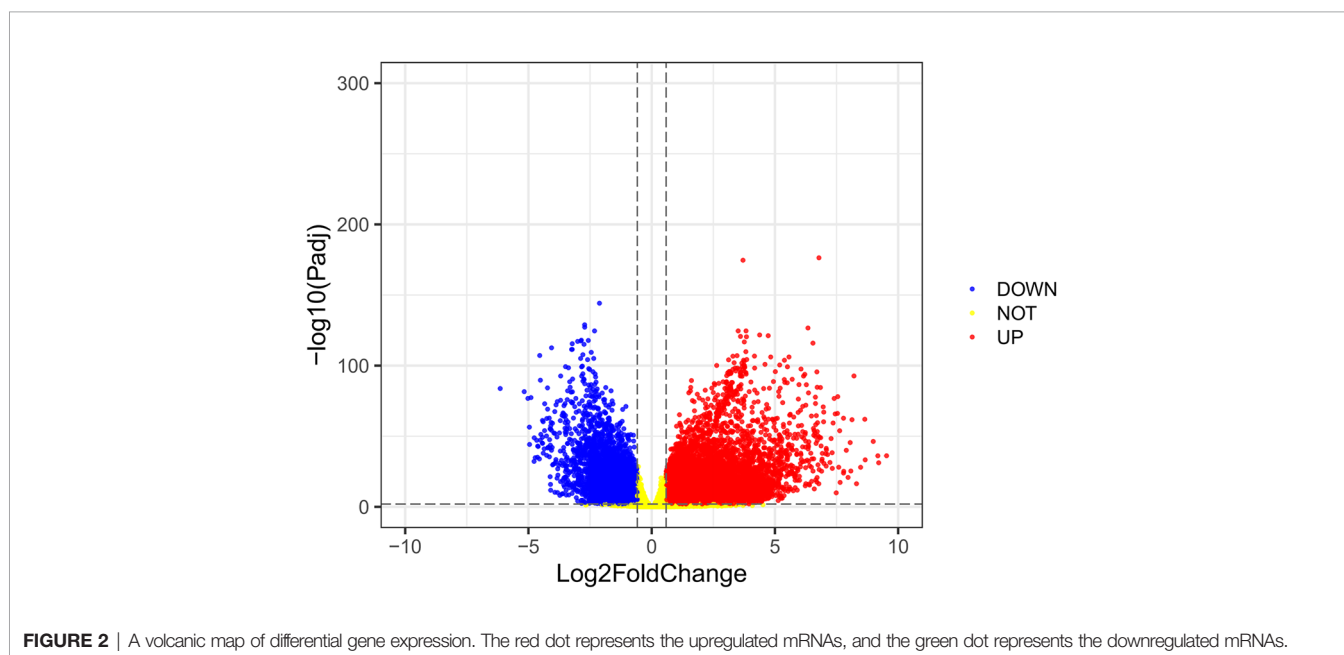


FIGURE 2 | A volcanic map of differential gene expression. The red dot represents the upregulated mRNAs, and the green dot represents the downregulated mRNAs.

analysis showed that in the BP categories, SNP-containing genes were mainly concentrated in homophilic cell adhesion *via* plasma membrane adhesion molecules, membrane depolarization during action potential and neuronal action potential. In the CC categories, mutated genes were mainly categorized in the plasma membrane, proteinaceous extracellular matrix, and voltage-gated sodium channel complex. In the MF categories, these genes were mainly enriched in calcium ion binding, extracellular matrix structural constituent and voltage-gated sodium channel activity (**Table 2**). Pathway enrichment analysis revealed the enrichment of the

SNP-containing genes in many signaling pathways related to cancer, including the phosphatidylinositol 3 kinase (PI3K)/Akt signaling pathway, calcium signaling pathway, and cyclic guanosine monophosphate (cGMP)-PKG signaling pathway (**Figure 3**).

Construction of the PPI Network for the Mutated Genes

To further study the potential relationships between the mutated genes, we used the STRING online database to mine the interactions among these genes. Cytoscape software was used

TABLE 2 | Gene ontology analysis of 945 mutant genes in gastric cancer.

Category	Term	Count	P value
GOTERM_BP_DIRECT	homophilic cell adhesion via plasma membrane adhesion molecules	54	1.13E-35
	negative regulation of transcription from RNA polymerase II promoter	33	0.0075481
	intracellular signal transduction	26	0.0033163
	cell adhesion	23	2.98E-06
	membrane depolarization during action potential	18	8.42E-15
	axon guidance	17	1.11E-05
	heart development	16	2.29E-04
	cell migration	13	0.0085702
	neuronal action potential	12	0.0337519
	microtubule cytoskeleton organization	12	1.60E-09
GOTERM_MF_DIRECT	GO:0005509-calcium ion binding	99	3.414E-26
	GO:0005524-ATP binding	99	1.153E-06
	GO:0008270-zinc ion binding	65	0.0263632
	GO:0003677-DNA binding	42	0.0361054
	GO:0003682-chromatin binding	25	0.0180564
	GO:0044212-transcription regulatory region DNA binding	18	9.495E-06
	GO:0005096-GTPase activator activity	18	0.0093759
	GO:0016887-ATPase activity	16	5.068E-05
	GO:0004842-ubiquitin-protein transferase activity	15	0.0201426
	GO:0004725-protein tyrosine phosphatase activity	14	9.479E-09
GOTERM_CC_DIRECT	integral component of membrane	197	0.0385191
	plasma membrane	133	2.329E-12
	integral component of plasma membrane	54	0.0021151
	proteinaceous extracellular matrix	32	6.187E-12
	cell surface	32	0.0006393
	focal adhesion	31	0.0002001
	dendrite	22	8.302E-07
	neuronal cell body	18	6.336E-05
	voltage-gated potassium channel complex	15	6.437E-07
	postsynaptic density	14	2.216E-07

Top 10 terms were selected according to count and P value <0.05. Count: the number of enriched genes in each term.

to visualize the complex PPI network, which included 360 nodes and 1616 edges (**Figure 4A**). MCODE was used to obtain the important modules from the PPI network, including 25 nodes and 245 edges (**Figure 4B**). Functional and KEGG pathway enrichment analyses showed that the BP categories of important modules mainly included CC organization, biological regulation and cell communication; the CC categories included nucleus, membrane-enclosed lumen and protein-containing complex; and the MF categories mainly included protein-binding, ion binding and transferase activity (**Figure 4C**). KEGG pathway enrichment analysis showed that important module genes were mainly enriched in the FoxO signaling pathway and thyroid hormone signaling pathway (**Figure 5**).

Screening and Survival Analysis of Pivotal Genes

Using the intersection of the first 30 genes in cytoHubba's 12 algorithms, we identified 10 key genes: TP53, EP300, AKT1, HRAS, PTEN, PIK3CA, SMARCA4, CREBBP, BRCA1 and ATM (**Figure 6A**). Metascape tools were used to analyze the pathway and biological process enrichment of hub genes. We observed that key genes are enriched in the PID P53 downstream pathway, apoptosis, regulation cellular response to stress, etc (**Figures 6B, C**). The cBioPortal online platform provided a graphic analysis of the genetic variation of the hub genes. As shown in the figure, 10 key SNP-containing genes all showed a high mutation rate in gastric cancer, with a rate of genome change ranging from 8% to

45% (**Figure 6D**). To determine whether the selected hub genes have clinical correlations, we used Kaplan-Meier curves to analyze the univariate survival of these genes and found that the expression of TP53, HRAS, BRCA1, PIK3CA, AKT1 and SMARCA4 was correlated with prognosis (**Figure 7**). Thus, these genes can be used as prognostic indicators of gastric cancer.

Hub Gene Verification Through UALCAN

UALCAN, an online tool with data from TCGA and GTEx, was used to verify the expression of these key genes in gastric cancer. In this study, according to the RNA sequence data from TCGA database, the mRNA expression levels of 6 genes were compared between the gastric tumor samples and the adjacent normal tissues. These six genes were found to be highly expressed at the transcriptional level in 415 gastric cancer tissues compared with 34 normal tissues (**Figure 8**). We chose $P < 0.05$ and multiple change > 2 as the threshold.

Genomic Changes of the Hub Genes

We used the cBioPortal tool to select 478 samples from TCGA database and explored the genome-specific changes of the hub gene. A summary analysis of cancer types showed that in the gastric cancer data set from TCGA, the proportion of the 6 genes changed from 12.66% to 51.90%, with the lowest to highest levels in mucinous stomach adenocarcinoma, diffuse type stomach adenocarcinoma, stomach adenocarcinoma, and tubular stomach adenocarcinoma (**Figure 9**).

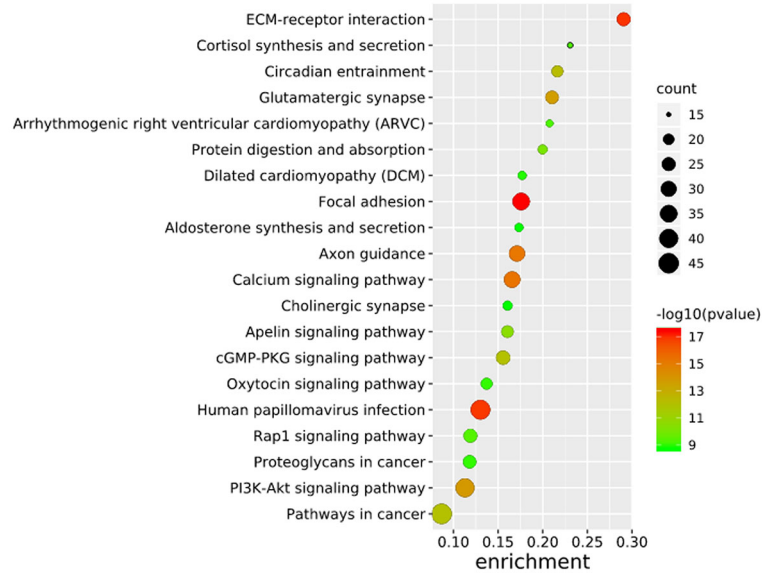


FIGURE 3 | Pathways enrichment map of 945 mutant genes. The top 20 terms with the lowest P value were selected. Count: the number of enriched genes in each term.

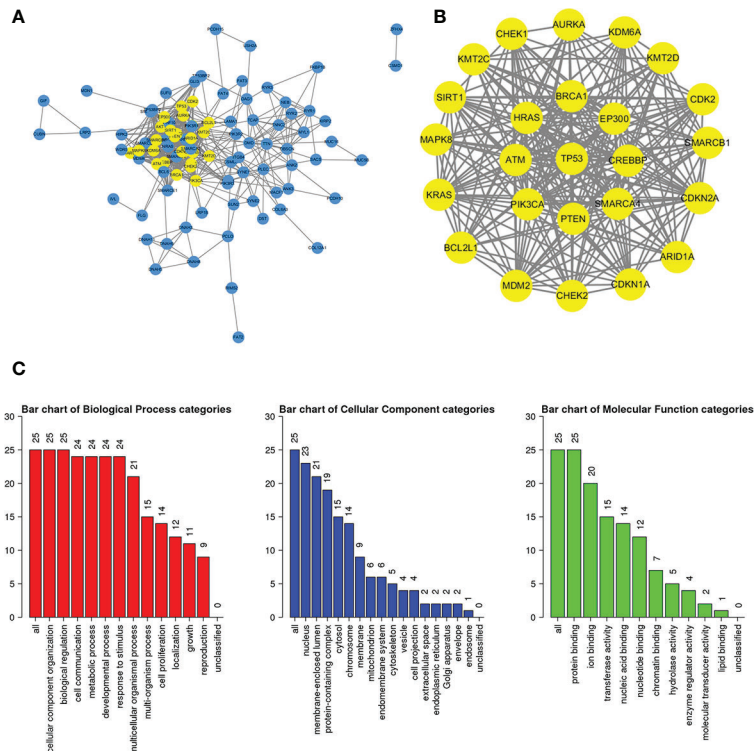


FIGURE 4 | Construction of the PPI network for the mutated genes and analysis of important modules. **(A)** Cytoscape was used to construct a PPI network of 945 mutant genes, including 360 nodes and 1616 edges. **(B)** MCODE was used to obtain an important module with 25 nodes and 245 edges from the PPI network. **(C)** The functional enrichment histogram of important modules. Each biological process, cellular component and molecular function category is represented by a red, blue and green bar, respectively. The height represents the number of IDs in the user list and in the category.

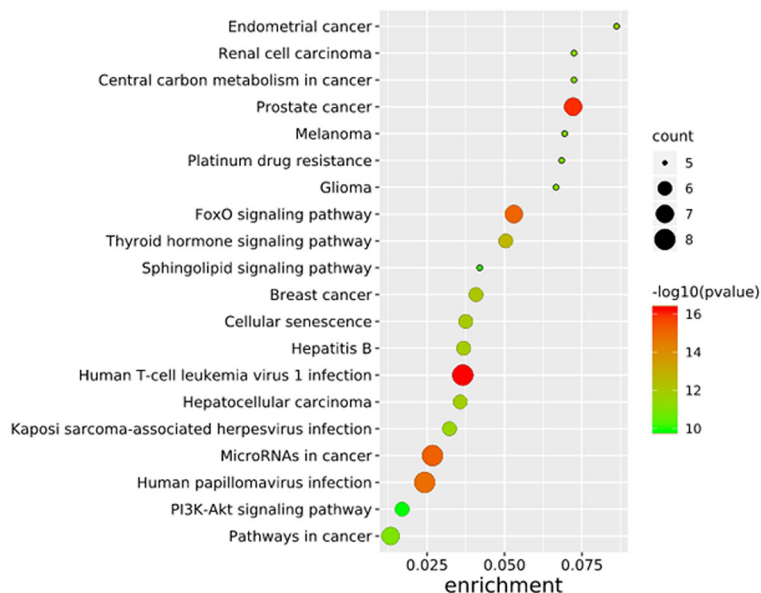


FIGURE 5 | Pathways enrichment map of important modules.

Gene Expression Levels of the Six Genes in Gastric Cancer

qRT-PCR was used to analyze the expression of TP53, HRAS, BRCA1, PIK3CA, AKT1, and SMRACA4 in gastric cancer. The results showed that the expression of the hub genes in AGS and GES-1 cell lines was upregulated compared with that in GES-1 cell lines (Figure 10). Western blot results showed that the expression levels of the six genes in gastric cancer cells were significantly higher than those in normal cells (Figure 11).

DISCUSSION

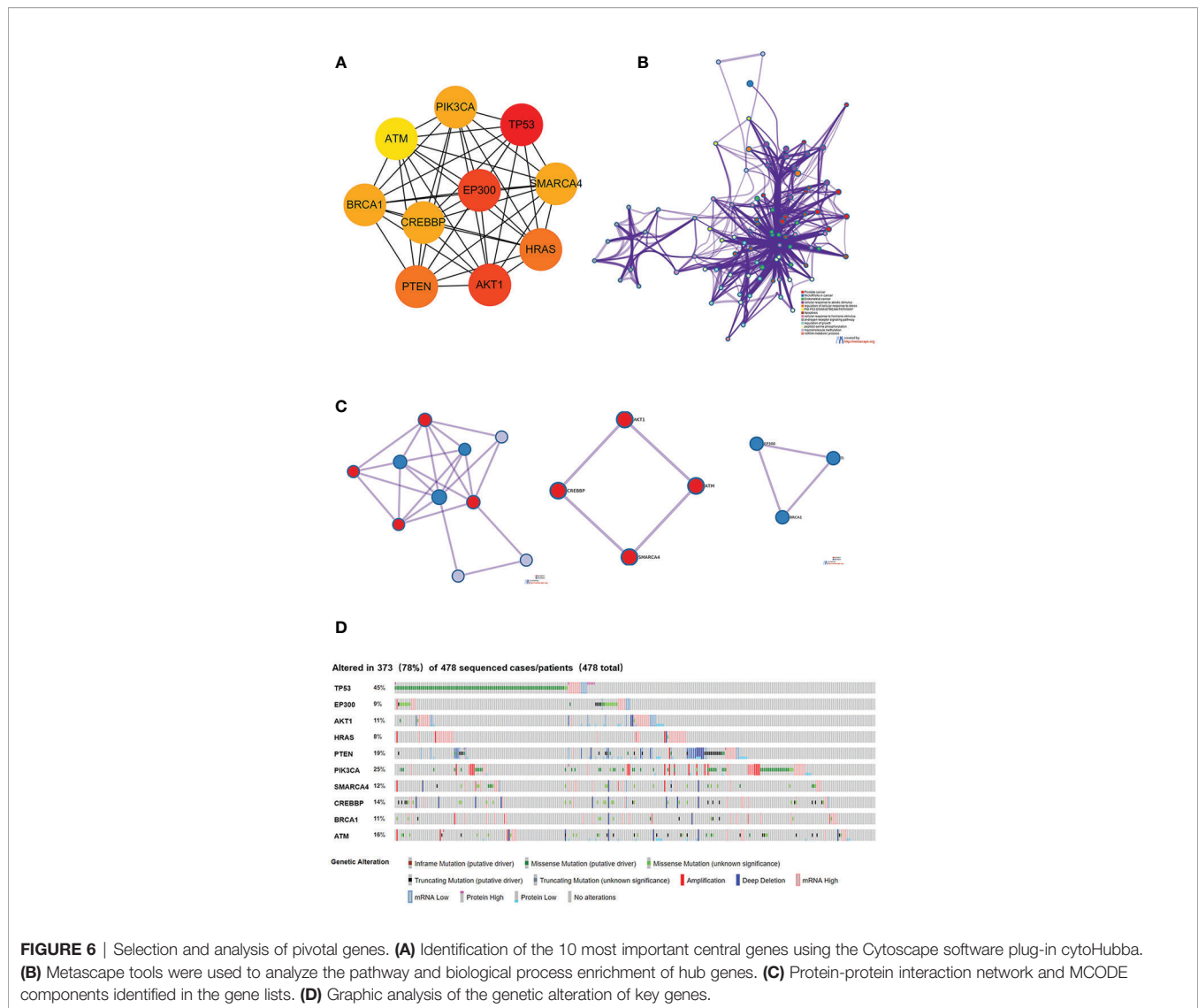
Gastric cancer is a complex disease and the fifth most common malignant tumor worldwide; it is also the third leading cause of cancer-related death (14). To improve the quality of life and prognosis of patients and prolong their survival time, researchers must further clarify the molecular mechanism leading to malignant biological behavior of gastric cancer and identify prognostic markers that affect the development of this disease. According to previous reports, genetic polymorphisms will increase the risk of cancer and are considered to be indicators of poor prognosis in various cancers and potential carcinogenic markers. Therefore, bioinformatics analysis of SNP-containing mutant genes and selection of valuable genes can provide new tools to treat patients and predict prognosis in the clinic.

In this study, we conducted a series of bioinformatics analyses on gastric cancer-related data in TCGA database to screen and identify prognostic biomarkers related to SNP-associated expression. We conducted functional and pathway enrichment analyses of these genes and found that these genes are enriched in

the nucleus and protein complex, mainly regulating multicellular BPs, developmental processes and metabolic processes. Pathway analysis showed that genes mutated in gastric cancer are mainly involved in the PI3K-AKT pathway, CGMP-PKG pathway, calcium signaling pathway, and many other cancer-related pathways. Functional and pathway enrichment analyses revealed the molecular mechanism of SNP-containing genes in the development of cancer.

Six mutated genes, TP53, AKT1, HRAS, PTEN, PIK3CA, SMARCA4 and BRCA1, which are closely related to the occurrence and development of gastric cancer, were screened. Survival analysis showed that the high expression of these six genes was associated with poor prognosis of gastric cancer. Using UALCAN online analysis, we found that the expression of the six genes in gastric cancer tissues was significantly higher than that in normal tissues. Subsequent cell experiments confirmed this result. In addition, we also used cBioPortal tools to study the genomic changes of the key genes in patients with gastric cancer from TCGA database. We found that there were five types of gastric cancer, and tubular gastric adenocarcinoma had the highest frequency of mutations in these genes. The rates of alteration of seven genes ranged from 12.66% to 51.90%.

TP53 (tumor protein p53) is the gene with the highest mutation frequency in gastric cancer (approximately 50%), and it is also the most commonly mutated gene in human cancer. This gene plays an important role in cell cycle arrest, cellular senescence, apoptosis, differentiation and metabolism (15). As a research hotspot in the field of tumor molecular biology, mutations in this gene are related to the poor prognosis of various cancers (16). Most TP53 mutations are missense mutations and gene deletions caused by substitution of single nucleotides, resulting in changes in the TP53 activity.

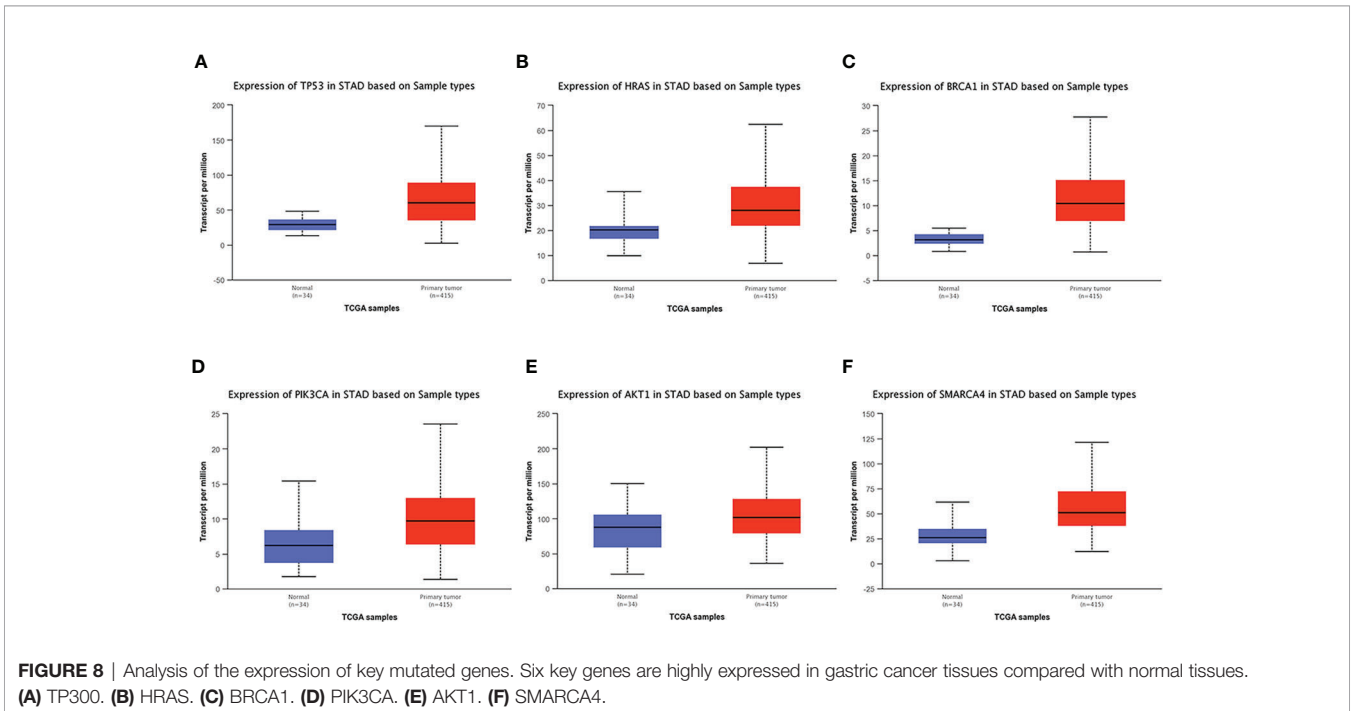
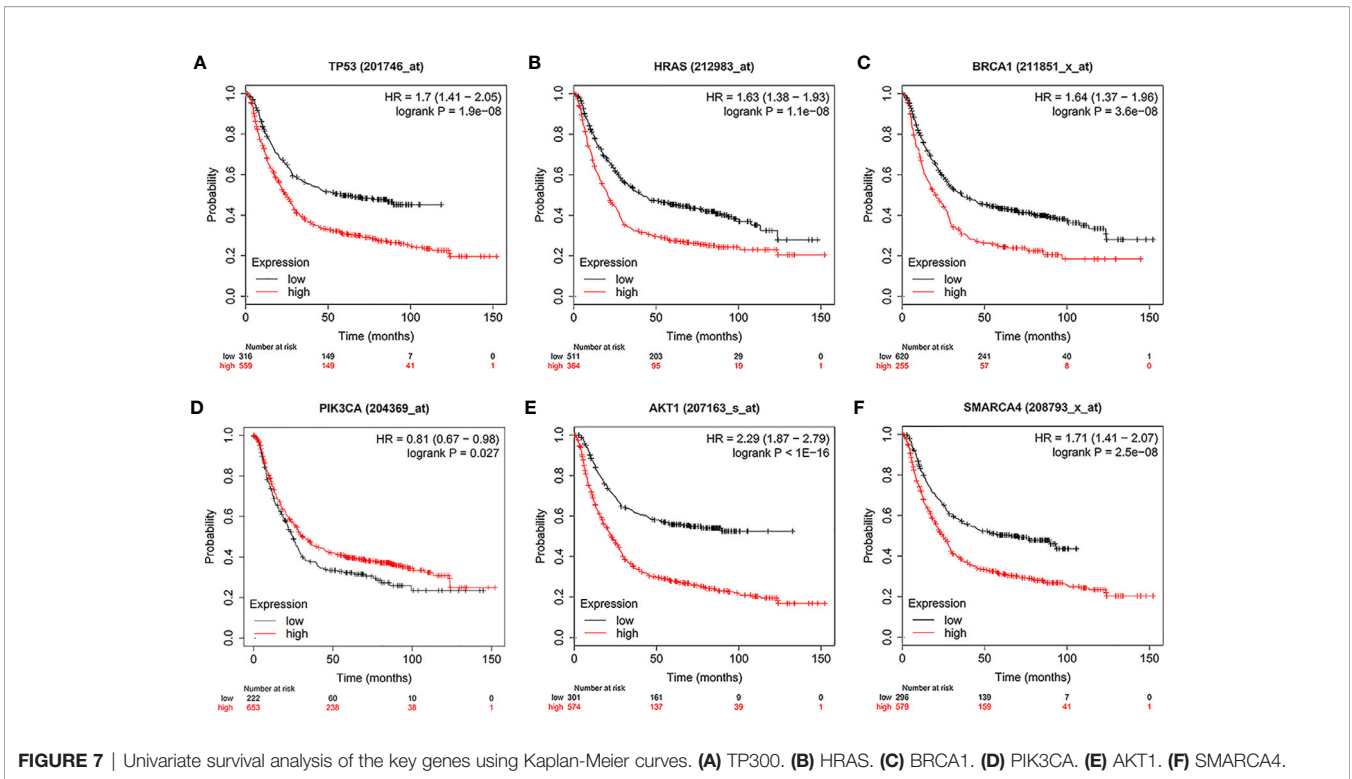


Mutant p53 protein not only loses the antitumor effect of wild-type p53 protein but also increases tumor cell activity, invasion and metastasis and promotes the occurrence and progression of tumors (17). Previous studies have demonstrated the relationship between TP53 and gastric cancer. Ando et al. (18) studied clinical samples from 182 cases of gastric cancer and found that TP53-positive tumors had deeper invasion and more lymph node and liver metastasis than other tumors, and some genes (PICT1, RPL11) were involved in the progression of cancer through TP53 (19). TP53 mutations occur late in gastric carcinogenesis, contributing to the final transition to cancer (20). In addition, Jiang et al. found that TP53 mutation can inhibit tumor immunity in gastric cancer (21). According to the results of our analysis, the degree value of the TP53 gene is the highest among the 10 hub genes, and its increased expression is negatively correlated with the five-year survival rate of gastric cancer, which further confirms the validity of this study. Further exploration of TP53 mutant genotypes will help reveal

the molecular mechanisms underlying the occurrence and development of gastric cancer.

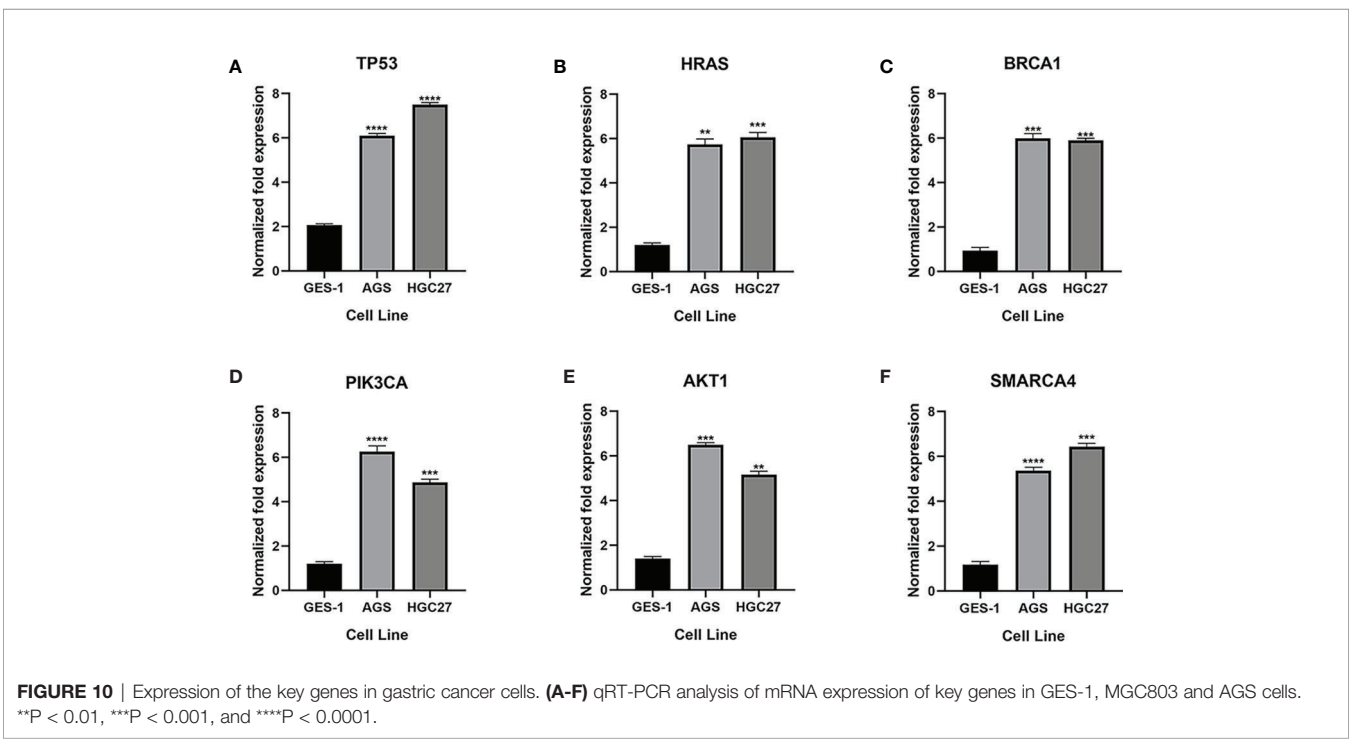
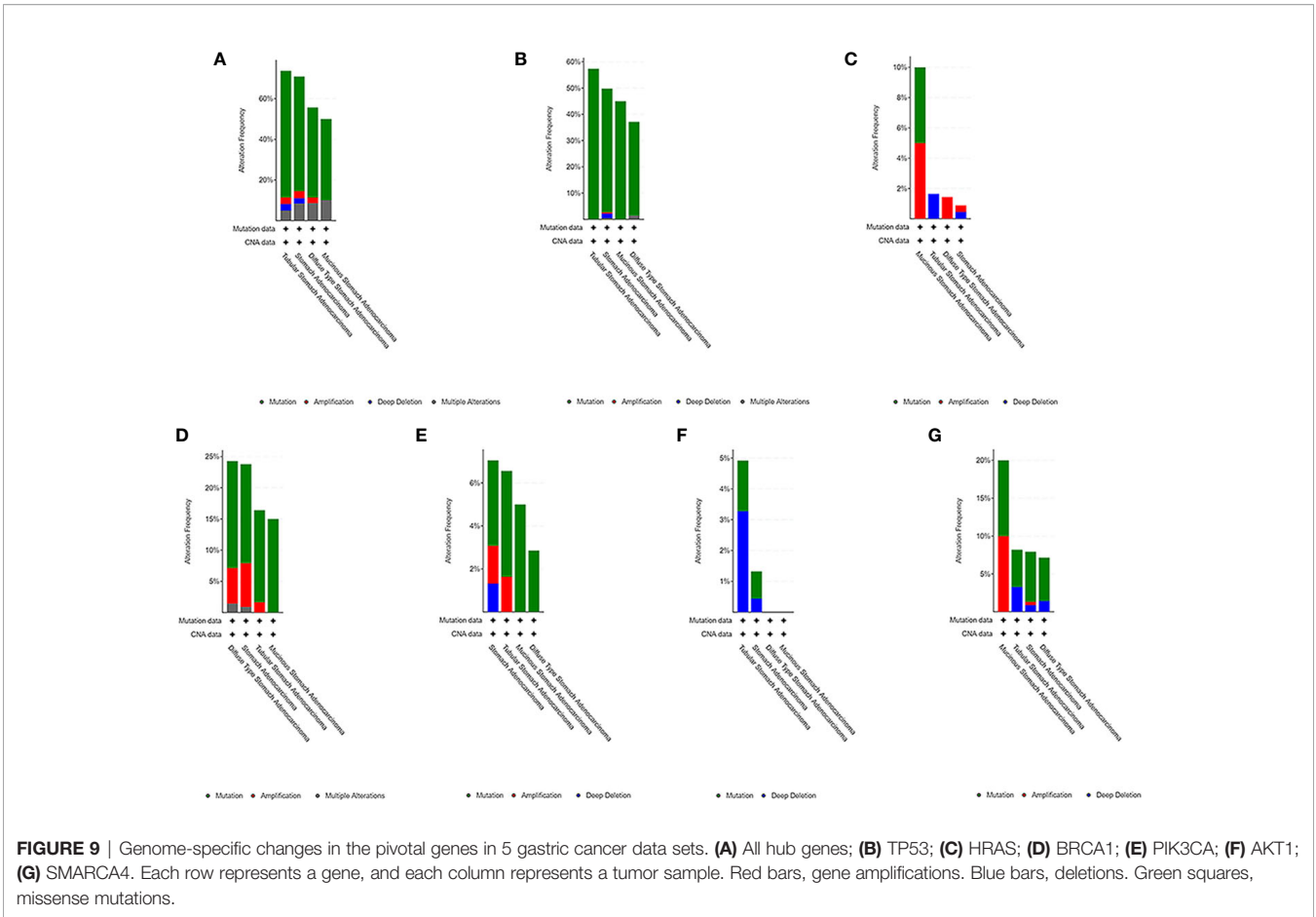
HRAS is a member of the RAS gene family, which participates in the activation of RAS protein signal transduction. The RAS protein is a GDP/GTP-binding protein that mainly regulates proliferation, differentiation and senescence in wild-type cells (22). HRAS functions as an oncogene after activation, and activation commonly occurs through mutations (SNPs, insertions, translocations) and increased expression (23). Diseases related to HRAS include breast cancer, liver cancer, thyroid cancer, and bladder cancer (24–27). Ectopic expression of HRAS was shown to promote the proliferation, migration, invasion, angiogenesis and clone formation of gastric cancer cells (28). Our results suggest that HRAS may play an important role in the diagnosis and treatment of gastric cancer.

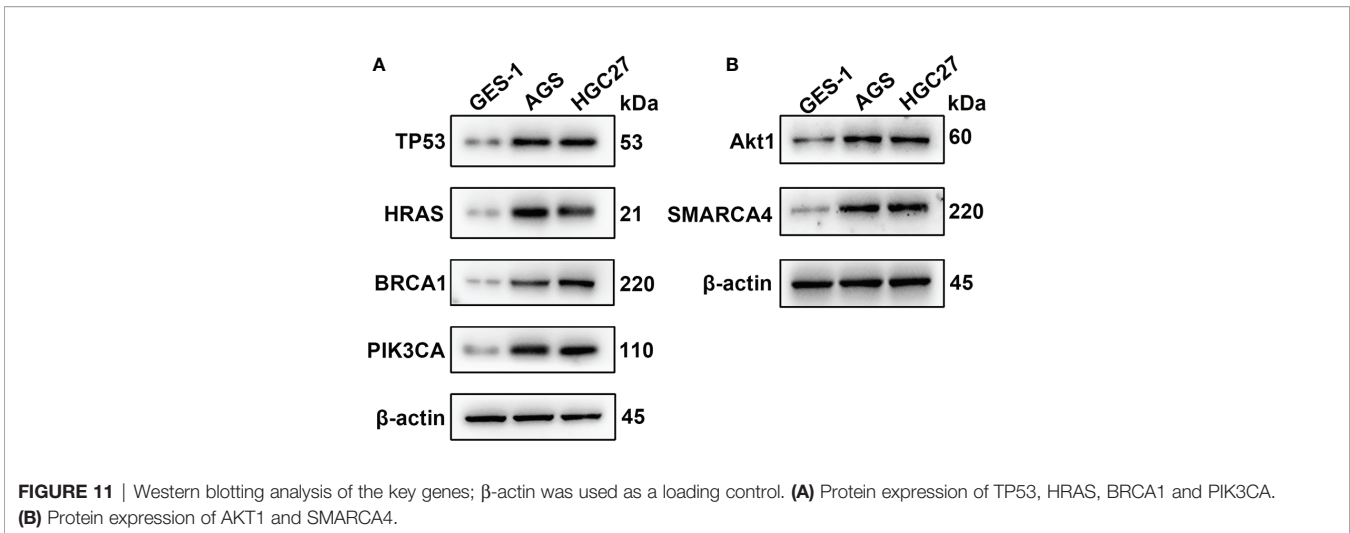
Breast cancer susceptibility gene (BRCA1) belongs a class of tumor suppressor genes with high penetrance that plays an important role in the response to DNA damage (including DNA



double-stranded breaks) (29). The BRCA1 protein can bind to various proteins to regulate gene transcription, maintaining the integrity of the genome. BRCA1 gene mutation leads to DNA replication errors and mutations, which promote abnormal cell proliferation and lead to tumorigenesis (30). Increasing

evidence shows that BRCA1 is widely associated with breast, ovarian and colon cancers (31, 32). Interestingly, the location of BRCA1 in gastric cancer cells was reported to be different, and different expression levels were observed. The expression of BRCA1 in the cytoplasm is downregulated, and the expression





of BRCA1 in the nucleus is upregulated, which is related to the poor prognosis of advanced tumors (33). BRCA1 gene polymorphisms have also been associated with susceptibility to gastric cancer (34, 35). Our study showed that the expression of BRCA1 in gastric cancer was higher than that in normal samples, and high expression was associated with poor prognosis, which indicates that BRCA1 may play contrasting roles in different types of tumors, and the role of BRCA1 in gastric cancer should be further explored.

The PI3K-Akt signal transduction pathway plays an important role in tumorigenesis, development, treatment and prognosis (36). PI3K is a component of the PI3K signaling pathway, plays a key role in the regulation of cell proliferation, survival and adhesion and is often upregulated in human cancer (37). PIK3CA (phosphoinositide-3-kinase, catalytic, alpha gene) encodes the p110 α subunit of PI3K. It plays an important role in tumor cell proliferation, differentiation, transport and metabolism (38). In addition, the PIK3CA pathway regulates angiogenesis and the immune response to cancer (39). PIK3CA mutations have been found in approximately 30% of human cancers (40), including breast cancer, ovarian cancer, colon cancer, and prostate cancer (41–44). In gastric cancer, high PIK3CA protein expression is closely related to tumor invasiveness, tumor phenotype and poor survival of patients (45). PIK3CA mutations were associated with high T stage, poor differentiation and microsatellite instability (46, 47). In our study, correlation analysis showed that the expression of PIK3CA in tumor tissues was significantly higher than that in normal tissues. However, OS analysis indicated the association of better prognosis of gastric cancer with high expression of PIK3CA, suggesting that the role of PIK3CA in gastric cancer is worthy of further exploration.

AKT is a direct downstream target protein of PIK3. Increasing evidence shows that the activation of the AKT protein plays an important biological role in the development of cancer (48). AKT1 is one of the subtypes of AKT. Activated AKT1 phosphorylates many downstream substrates and participates in the regulation of cell growth, metabolism, proliferation, apoptosis and other processes (49). Petrini et al.

found that patients with overexpression of AKT1 in gastric cancer had a poor prognosis, suggesting that AKT1 can be used as a poor prognostic marker for gastric cancer (50). Ghatak et al. found that AKT1 mutation was associated with an abnormal cell cycle in gastric cancer (51). Pathway analysis showed that AKT1 is enriched in PI3K-AKT, MAPK and several other pathways closely related to cancer, which indicates that the gene encoding AKT1 has an important biological function in the development of cancer.

ATP-dependent chromatin remodeling plays an important role in the occurrence and development of cancer, participating in almost all aspects of DNA metabolism, such as transcription, recombination, DNA repair and DNA replication (52). The SWI/SNF complex (BAF complex) was the first discovered mechanism of chromatin remodeling. Proteins encoded by SMARCA4 (also known as BGR1) are members of the SWI/SNF family, have helicase and ATP enzyme activities, and regulate gene transcription by changing the structure of chromatin (53). SMARCA4 is generally considered a tumor suppressor gene (54). However, some recent reports have demonstrated that SMARCA4 plays an important role in cell survival and proliferation in some types of cancer (55, 56). Martinez et al. found that SMARCA4 was highly expressed in 11 kinds of tumor tissues, including gastric cancer tissue, and was related to poor prognosis (57), which reflects the dual role of SMARCA4 in cancer. Previous studies have identified several genes as promising diagnostic and prognostic biomarkers for GC (58). In our study, the expression of SMARCA4 in gastric cancer cells was higher than that in normal cells. OS analysis indicated that high expression of SMARCA4 was linked to poor prognosis, which is consistent with previous studies, and further confirmed our results. When SMARCA4 is mutated in gastric cancer, our data indicated that SMARCA4 does not act as a tumor suppressor, which may be due to the pathological activity of abnormal residual complexes of SWI/SNF.

Finally, we analyzed the protein expression of six key genes. We observed that the expression of TP53, AKT1, HRAS, PTEN, PIK3CA, SMARCA4 and BRCA1 in gastric cancer cells was

higher than that in normal gastric cells. This is consistent with the results of our bioinformatics analysis. In genetic analysis, SNP is widely used as a kind of genetic markers, and some SNP located in genes may directly affect the protein structure or expression level. The six screened genes are highly expressed in gastric cancer cells, suggesting that these SNP mutant genes may play a role as oncogenes in gastric cancer. Based on this finding, a more in-depth study of the mechanism of these genes will help to reveal the role of SNP in the mechanism of cancer.

CONCLUSION

In this study, through bioinformatics and experimental analyses, we found that six SNP-containing genes (TP53, AKT1, HRAS, PTEN, PIK3CA, SMARCA4 and BRCA1) may be key factors in the occurrence and prognosis of gastric cancer and participate in many pathways related to cancer development. Therefore, on this basis, further studies should be performed to detect the polymorphic sites of these genes and explore their corresponding expression levels, which can be used to predict the prognosis of patients. These findings will need to be verified in large-scale clinical studies to determine their accuracy and sensitivity in tumorigenesis and to guide the individualized treatment of patients. However, the focus of this study is to provide new ideas for clinical diagnosis and prognostic evaluation through bioinformatics analysis. Our results provide an important bioinformatics basis and related theoretical basis for guiding follow-up research on gastric cancer.

DATA AVAILABILITY STATEMENT

Publicly available datasets were used in this study. These can be found at The TCGA genomic data sharing platform (<https://gdc.cancer.gov/>).

REFERENCES

- Ang TL, Fock KM. Clinical Epidemiology of Gastric Cancer. *Singapore Med J* (2014) 55(12):621–8. doi: 10.11622/smedj.2014174
- Lazar DC, Taban S, Cornianu M, Faur A, Goldis A. New Advances in Targeted Gastric Cancer Treatment. *World J Gastroenterol* (2016) 22(30):6776–99. doi: 10.3748/wjg.v22.i30.6776
- Peddanna N, Holt S, Verma RS. Genetics of Gastric Cancer. *Anticancer Res* (1995) 15(5B):2055–64.
- Zhang Z, Hao K. Using SAAS-CNV to Detect and Characterize Somatic Copy Number Alterations in Cancer Genomes From Next Generation Sequencing and SNP Array Data. *Methods Mol Biol* (2018) 1833:29–47. doi: 10.1007/978-1-4939-8666-8_2
- Tomczak K, Czerwinska P, Wizniewicz M. The Cancer Genome Atlas (TCGA): An Immeasurable Source of Knowledge. *Contemp Oncol (Pozn)* (2015) 19(1A):A68–77. doi: 10.5114/wo.2014.47136
- Huang da W, Sherman BT, Lempicki RA. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat Protoc* (2009) 4(1):44–57. doi: 10.1038/nprot.2008.211
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. String v10: Protein-Protein Interaction Networks, Integrated Over the Tree of Life. *Nucleic Acids Res* (2015) 43(Database issue):D447–52. doi: 10.1093/nar/gku1003
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* (2003) 13(11):2498–504. doi: 10.1101/gr.1239303
- Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: Identifying Hub Objects and Sub-Networks From Complex Interactome. *BMC Syst Biol* (2014) 8 Suppl 4:S11. doi: 10.1186/1752-0509-8-S4-S11
- Bader GD, Hogue CW. An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinf* (2003) 4:2. doi: 10.1186/1471-2105-4-2
- Lanczky A, Nagy A, Bottai G, Munkacsy G, Szabo A, Santarpia L, et al. miRpower: A Web-Tool to Validate Survival-Associated miRNAs Utilizing Expression Data From 2178 Breast Cancer Patients. *Breast Cancer Res Treat* (2016) 160(3):439–46. doi: 10.1007/s10549-016-4013-7
- Chandrashekar DS, Bachel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. Ualcan: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* (2017) 19(8):649–58. doi: 10.1016/j.neo.2017.05.002
- Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the Cbioportal. *Sci Signal* (2013) 6(269):pl1. doi: 10.1126/scisignal.2004088
- Zhang X, Li M, Chen S, Hu J, Guo Q, Liu R, et al. Endoscopic Screening in Asian Countries Is Associated With Reduced Gastric Cancer Mortality: A

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HL and JG analyzed the data and wrote the manuscript. GC, YW, and SL assisted in editing the manuscript. YQ, GW, and RX contributed to the design of the study. WSQ and WWQ are the corresponding authors of the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by two grants: Natural Science Foundation of China (81602068) and Beijing Xisike Clinical Oncology Research Foundation (Y-HR2018-185).

ACKNOWLEDGMENTS

The authors gratefully acknowledge the data provided by patients and researchers participating in TCGA, GEO, and EGA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.564296/full#supplementary-material>

- Meta-Analysis and Systematic Review. *Gastroenterology* (2018) 155(2):347–54 e9. doi: 10.1053/j.gastro.2018.04.026
15. Soussi T. TP53 Mutations in Human Cancer: Database Reassessment and Prospects for the Next Decade. *Adv Cancer Res* (2011) 110:107–39. doi: 10.1016/B978-0-12-386469-7.00005-0
 16. Li VD, Li KH, Li JT. TP53 Mutations as Potential Prognostic Markers for Specific Cancers: Analysis of Data From The Cancer Genome Atlas and the International Agency for Research on Cancer Tp53 Database. *J Cancer Res Clin Oncol* (2019) 145(3):625–36. doi: 10.1007/s00432-018-2817-z
 17. Wang M, Yang C, Zhang X, Li X. Characterizing Genomic Differences of Human Cancer Stratified by the TP53 Mutation Status. *Mol Genet Genomics* (2018) 293(3):737–46. doi: 10.1007/s00438-018-1416-7
 18. Ando K, Oki E, Zhao Y, Ikawa-Yoshida A, Kitao H, Saeki H, et al. Mortalin is a Prognostic Factor of Gastric Cancer With Normal p53 Function. *Gastric Cancer* (2014) 17(2):255–62. doi: 10.1007/s10120-013-0279-1
 19. Uchi R, Kogo R, Kawahara K, Sudo T, Yokobori T, Eguchi H, et al. PICT1 Regulates TP53 Via RPL11 and is Involved in Gastric Cancer Progression. *Br J Cancer* (2013) 109(8):2199–206. doi: 10.1038/bjc.2013.561
 20. Lin Y, Wu Z, Guo W, Li J. Gene Mutations in Gastric Cancer: A Review of Recent Next-Generation Sequencing Studies. *Tumour Biol* (2015) 36(10):7385–94. doi: 10.1007/s13277-015-4002-1
 21. Jiang Z, Liu Z, Li M, Chen C, Wang X. Immunogenomics Analysis Reveals That TP53 Mutations Inhibit Tumor Immunity in Gastric Cancer. *Transl Oncol* (2018) 11(5):1171–87. doi: 10.1016/j.tranon.2018.07.012
 22. Li S, Balmain A, Counter CM. A Model for RAS Mutation Patterns in Cancers: Finding the Sweet Spot. *Nat Rev Cancer* (2018) 18(12):767–77. doi: 10.1038/s41568-018-0076-6
 23. Chen S, Li F, Xu D, Hou K, Fang W, Li Y. The Function of RAS Mutation in Cancer and Advances in its Drug Research. *Curr Pharm Des* (2019) 25(10):1105–14. doi: 10.2174/1381612825666190506122228
 24. Geyer FC, Li A, Papanastasiou AD, Smith A, Selenica P, Burke KA, et al. Recurrent Hotspot Mutations in HRAS Q61 and PI3K-AKT Pathway Genes as Drivers of Breast Adenomyoepitheliomas. *Nat Commun* (2018) 9(1):1816. doi: 10.1038/s41467-018-04128-5
 25. Pecenkova V, Pajer P, Karafiat V, Kasparova P, Dudlova J, Dvorak M. HRAS, et al. And RON Genes Are Recurrently Activated by Provirus Insertion in Liver Tumors Induced by the Retrovirus Myeloblastosis-Associated Virus 2. *J Virol* (2017) 91(20). doi: 10.1128/JVI.00467-17
 26. Zhang X, Zhang Y. Bladder Cancer and Genetic Mutations. *Cell Biochem Biophys* (2015) 73(1):65–9. doi: 10.1007/s12013-015-0574-z
 27. Untch BR, Dos Anjos V, Garcia-Rendueles MER, Knauf JA, Krishnamoorthy GP, Saqcenca M, et al. Tipifarnib Inhibits Hras-Driven Dedifferentiated Thyroid Cancers. *Cancer Res* (2018) 78(16):4642–57. doi: 10.1158/0008-5472.CAN-17-1925
 28. Wu XY, Liu WT, Wu ZF, Chen C, Liu JY, Wu GN, et al. Identification of HRAS as Cancer-Promoting Gene in Gastric Carcinoma Cell Aggressiveness. *Am J Cancer Res* (2016) 6(9):1935–48.
 29. Jhanwar-Uniyal M. BRCA1 in Cancer, Cell Cycle and Genomic Stability. *Front Biosci* (2003) 8:s1107–17. doi: 10.2741/1131
 30. Somasundaram K. Breast Cancer Gene 1 (BRCA1): Role in Cell Cycle Regulation and DNA Repair—Perhaps Through Transcription. *J Cell Biochem* (2003) 88(6):1084–91. doi: 10.1002/jcb.10469
 31. Paul A, Paul S. The Breast Cancer Susceptibility Genes (BRCA) in Breast and Ovarian Cancers. *Front Biosci (Landmark Ed)* (2014) 19:605–18. doi: 10.2741/4230
 32. Oh M, McBride A, Yun S, Bhattacharjee S, Slack M, Martin JR, et al. BRCA1 and BRCA2 Gene Mutations and Colorectal Cancer Risk: Systematic Review and Meta-Analysis. *J Natl Cancer Inst* (2018) 110(11):1178–89. doi: 10.1093/jnci/djy148
 33. Wang GH, Zhao CM, Huang Y, Wang W, Zhang S, Wang X. BRCA1 and BRCA2 Expression Patterns and Prognostic Significance in Digestive System Cancers. *Hum Pathol* (2018) 71:135–44. doi: 10.1016/j.humpath.2017.10.032
 34. Chen XR, Zhang WZ, Lin XQ, Wang JW. Genetic Instability of BRCA1 Gene At Locus D17S855 is Related to Clinicopathological Behaviors of Gastric Cancer From Chinese Population. *World J Gastroenterol* (2006) 12(26):4246–9. doi: 10.3748/wjg.v12.i26.4246
 35. Shim HJ, Yun JY, Hwang JE, Bae WK, Cho SH, Lee JH, et al. BRCA1 and XRCC1 Polymorphisms Associated With Survival in Advanced Gastric Cancer Treated With Taxane and Cisplatin. *Cancer Sci* (2010) 101(5):1247–54. doi: 10.1111/j.1349-7006.2010.01514.x
 36. Noorolyai S, Shajari N, Baghbani E, Sadreddini S, Baradaran B. The Relation Between PI3K/AKT Signalling Pathway and Cancer. *Gene* (2019) 698:120–8. doi: 10.1016/j.gene.2019.02.076
 37. Mayer IA, Arteaga CL. The PI3K/AKT Pathway as a Target for Cancer Treatment. *Annu Rev Med* (2016) 67:11–28. doi: 10.1146/annurev-med-062913-051343
 38. Liu X, Xu Y, Zhou Q, Chen M, Zhang Y, Liang H, et al. PI3K in Cancer: its Structure, Activation Modes and Role in Shaping Tumor Microenvironment. *Future Oncol* (2018) 14(7):665–74. doi: 10.2217/fon-2017-0588
 39. Sobral-Leite M, Salomon I, Opdam M, Kruger DT, Beelen KJ, van der Noort V, et al. Cancer-Immune Interactions in ER-positive Breast Cancers: PI3K Pathway Alterations and Tumor-Infiltrating Lymphocytes. *Breast Cancer Res* (2019) 21(1):90. doi: 10.1186/s13058-019-1176-2
 40. Stachler MD, Rinehart EM, Garcia E, Lindeman NI. Pik3ca Mutations are Common in Many Tumor Types and are Often Associated With Other Driver Mutations. *Appl Immunohistochem Mol Morphol* (2016) 24(5):313–9. doi: 10.1097/PAI.0000000000000195
 41. Alvarez-Cubero MJ, Martinez-Gonzalez LJ, Robles-Fernandez I, Martinez-Herrera J, Garcia-Rodriguez G, Pascual-Geler M, et al. Somatic Mutations in Prostate Cancer: Closer to Personalized Medicine. *Mol Diagn Ther* (2017) 21(2):167–78. doi: 10.1007/s40291-016-0248-6
 42. Huang D, Sun W, Zhou Y, Li P, Chen F, Chen H, et al. Mutations of Key Driver Genes in Colorectal Cancer Progression and Metastasis. *Cancer Metastasis Rev* (2018) 37(1):173–87. doi: 10.1007/s10555-017-9726-5
 43. Liu S, Tang Y, Yan M, Jiang W. PIK3CA Mutation Sensitizes Breast Cancer Cells to Synergistic Therapy of PI3K Inhibition and AMPK Activation. *Invest New Drugs* (2018) 36(5):763–72. doi: 10.1007/s10637-018-0563-3
 44. Zhang H, Zhou L. Single Nucleotide Polymorphism of PIK3CA and its Interaction With the Environment are Risk Factors for Chinese Han Ovarian Cancer. *Pathol Res Pract* (2019) 215(9):152520. doi: 10.1016/j.prp.2019.152520
 45. Jang SH, Kim KJ, Oh MH, Lee JH, Lee HJ, Cho HD, et al. Clinicopathological Significance of Elevated Pik3ca Expression in Gastric Cancer. *J Gastric Cancer* (2016) 16(2):85–92. doi: 10.5230/jgc.2016.16.2.85
 46. Polom K, Marrelli D, Roviello G, Pascale V, Voglino C, Vindigni C, et al. PIK3CA Mutation in Gastric Cancer and the Role of Microsatellite Instability Status in Mutations of Exons 9 and 20 of the PIK3CA Gene. *Adv Clin Exp Med* (2018) 27(7):963–9. doi: 10.17219/acem/70795
 47. Shen X, Zhao Y, Chen X, Sun H, Liu M, Zhang W, et al. Associations of PIK3CA Mutations With Clinical Features and Prognosis in Gastric Cancer. *Future Oncol* (2019) 15(16):1873–94. doi: 10.2217/fon-2018-0335
 48. Song M, Bode AM, Dong Z, Lee MH. AKT as a Therapeutic Target for Cancer. *Cancer Res* (2019) 79(6):1019–31. doi: 10.1158/0008-5472.CAN-18-2738
 49. LoPiccolo J, Granville CA, Gills JJ, Dennis PA. Targeting Akt in Cancer Therapy. *Anticancer Drugs* (2007) 18(8):861–74. doi: 10.1097/CAD.0b013e3280cc2c6f
 50. Petrini I, Lencioni M, Vasile E, Fornaro L, Belluomini L, Pasquini G, et al. EGFR and AKT1 Overexpression are Mutually Exclusive and Associated With a Poor Survival in Resected Gastric Adenocarcinomas. *Cancer Biomark* (2018) 21(3):731–41. doi: 10.3233/CBM-170865
 51. Ghatak S, Lalnunhlimi S, Lalrohlu F, Pautu JL, Zohmingthanga J, Kunnumakkara AB, et al. Novel AKT1 Mutations Associated With Cell-Cycle Abnormalities in Gastric Carcinoma. *Per Med* (2018) 15(2):79–86. doi: 10.2217/pme-2017-0053
 52. Hargreaves DC, Crabtree GR. ATP-Dependent Chromatin Remodeling: Genetics, Genomics and Mechanisms. *Cell Res* (2011) 21(3):396–420. doi: 10.1038/cr.2011.32
 53. Khavari PA, Peterson CL, Tamkun JW, Mendel DB, Crabtree GR. BRG1 Contains a Conserved Domain of the SWI2/SNF2 Family Necessary for Normal Mitotic Growth and Transcription. *Nature* (1993) 366(6451):170–4. doi: 10.1038/366170a0
 54. Glaros S, Cirrincione GM, Palanca A, Metzger D, Reisman D. Targeted Knockout of BRG1 Potentiates Lung Cancer Development. *Cancer Res* (2008) 68(10):3689–96. doi: 10.1158/0008-5472.CAN-07-6652
 55. Buscarlet M, Krasteva V, Ho L, Simon C, Hebert J, Wilhelm B, et al. Essential Role of BRG, the ATPase Subunit of BAF Chromatin Remodeling Complexes, in Leukemia Maintenance. *Blood* (2014) 123(11):1720–8. doi: 10.1182/blood-2013-02-483495

56. Wu Q, Madany P, Akech J, Dobson JR, Douthwright S, Browne G, et al. The SWI/SNF ATPases Are Required for Triple Negative Breast Cancer Cell Proliferation. *J Cell Physiol* (2015) 230(11):2683–94. doi: 10.1002/jcp.24991
57. Laurette P, Coassolo S, Davidson G, Michel I, Gambi G, Yao W, et al. Chromatin Remodellers Brg1 and Bptf are Required for Normal Gene Expression and Progression of Oncogenic Braf-driven Mouse Melanoma. *Cell Death Differ* (2020) 27(1):29–43. doi: 10.1038/s41418-019-0333-6
58. Wang W, He Y, Zhao Q, Zhao X, Li Z. Identification of Potential Key Genes in Gastric Cancer Using Bioinformatics Analysis. *BioMed Rep* (2020) 12(4):178–92. doi: 10.3892/br.2020.1281

Conflict of Interest: 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 © 2021 Li, Guo, Cheng, Wei, Liu, Qi, Wang, Xiao, Qi and Qiu. 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.