Heliyon 10 (2024) e37046

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Helicobacter pylori infection facilitates cell migration and potentially impact clinical outcomes in gastric cancer

Ling Ou [a,1], Hengrui Liu [b,e,1], Chang Peng [a], Yuanjing Zou [a], Junwei Jia [c], Hui Li [c], Zhong Feng [a,c,**], Guimin Zhang [d,***], Meicun Yao [a,*]

[a] School of Pharmaceutical Sciences (Shenzhen), Sun Yat-sen University, Shenzhen, 518107, China
[b] Cancer Institute, Jinan University, Guangzhou, China
[c] International Pharmaceutical Engineering Lab of Shandong Province, Feixian, 273400, Shandong, China
[d] Lunan Pharmaceutical Group Co., Ltd, Linyi, 276000, Shandong, China
[e] Tianjin Yinuo Biomedical Co., Ltd, Tianjin, China

ARTICLE INFO

Keywords:
Helicobacter pylori
Gastric cancer
Cell infection
Bioinformatics

ABSTRACT

Gastric cancer is a significant health concern worldwide. *Helicobacter pylori* (HP) infection is associated with gastric cancer risk, but differences between HP-infected and HP-free gastric cancer have not been studied sufficiently. The objective of this study was to investigate the effects of HP infection on the viability and migration of gastric cancer cells and identify potential underlying genetic mechanisms as well as their clinical relevance. Cell counting kit-8, lactate dehydrogenase, wound healing, and transwell assay were applied in the infection model of multiple clones of HP and multiple gastric cancer cell lines. Genes related to HP infection were identified using bioinformatics analysis and subsequently validated using real-time quantitative PCR. The association of these genes with immunity and drug sensitivity of gastric cancer was analyzed. Results showed that HP has no significant impact on viability but increases the migration of gastric cancer cells. We identified 1405 HP-upregulated genes, with their enriched terms relating to cell migration, drug, and immunity. Among these genes, the 82 genes associated with survival showed a significant impact on gastric cancer in consensus clustering and LASSO prognostic model. The top 10 hub HP-associated genes were further identified, and 7 of them were validated in HP-infected cells using real-time quantitative PCR, including ERBB4, DNER, BRINP2, KCTD16, MAPK4, THPO, and VSTM2L. The overexpression experiment showed that KCTD16 medicated the effect of HP on gastric cancer migration. Our findings suggest that HP infection may enhance the migratory potential of gastric cancer cells and these genes might be associated with immunity and drug sensitivity of gastric cancer. In human subjects with gastric cancer, HP presence in tumors may affect migration, immunity, and drug sensitivity.

* Corresponding author. School of Pharmaceutical Sciences (Shenzhen), Sun Yat-sen University, Shenzhen, 518107, China.
** Corresponding author. International Pharmaceutical Engineering Lab, Shandong, 273400, China.
*** Corresponding author. Lunan Pharmaceutical Group Co., Ltd, Linyi, 276000, Shandong, China.
E-mail addresses: ouling5@mail2.sysu.edu.cn (L. Ou), lh@yinuobiomedical.cn (H. Liu), pengch56@mail2.sysu.edu.cn (C. Peng), zouyj6@mail2.sysu.edu.cn (Y. Zou), jiajunwei98@163.com (J. Jia), 1207561886@qq.com (H. Li), fengzhong22@163.com (Z. Feng), lunanzhangguimin@163.com (G. Zhang), lssymc@mail.sysu.edu.cn (M. Yao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e37046
Received 25 April 2024; Received in revised form 23 August 2024; Accepted 27 August 2024
Available online 28 August 2024

Abbreviations

Helicobacter pylori HP
The Cancer Genome Atlas TCGA
Stomach adenocarcinoma STAD
Gene Ontology GO
Kyoto Encyclopedia of Genes and Genomes KEGG
gene set enrichment analysis GSEA
Kaplan-Meier KM
one-class logistic regression OCLR
Density of Maximum Neighborhood Component DMNC
Immune checkpoint blockade ICB
Tumor Immune Dysfunction and Exclusion TIDE
Least absolute shrinkage and selection operator LASSO
Real-time Quantitative-PCR RT-QPCR

1. Introduction

Gastric cancer is a common type of tumor and is currently the fourth leading cause of cancer-related death [1–10]. For early-stage gastric cancer, endoscopic resection is the major clinical therapy [11]. Late-stage gastric cancer results in much worse survival although surgical treatment combined with chemotherapy brings a better prognosis for patients [11]. Targeted therapies have been developed for tumors in the stomach [12]. However, due to the possibility of drug resistance in gastric cancer and the heterogeneity of patients, it is necessary to identify additional targets for therapy and markers for prognosis. *Helicobacter pylori* (HP) is a group 1 carcinogen categorized by the World Health Organization [13,14] and may increase the risk of developing gastric cancer [15]. In developing countries, up to 90 % of the population can be infected with HP by age 10, although lower in developed countries [16]. Over 80 % of cases of gastric cancer may be attributable to the bacterium [17,18]. This underscores the need for effective public health strategies to reduce the prevalence of HP infection [16].

HP colonizes the human stomach by penetrating the mucus layer, adhering to the gastric mucosa, and neutralizing stomach acid with urease. Key factors for its persistence include urease production, adaptability to the stomach environment, immune evasion, genetic diversity, and biofilm formation. HP infection can lead to an increase in gastrin secretion, a hormone that stimulates the proliferation of gastric mucosal cells, which raises the risk of polyp formation. Additionally, HP infection can trigger an abnormal immune response, causing infiltration of inflammatory cells and the release of cytokines. This further exacerbates damage to the gastric mucosa and promotes the development of polyps. The successful eradication of HP for the first time has been demonstrated to decrease hyperplastic polyps in patients and, in turn, reduce the incidence of gastric cancer [19].

The key virulence factors of HP that facilitate its tumorigenic effects include cytotoxin-associated gene A (CagA), vacuolating cytotoxin A (VacA), and urease. CagA can disrupt cell signaling pathways and promote inflammation and cell transformation [20]. VacA causes vacuolation and damage to gastric epithelial cells [14]. Urease helps HP survive in the acidic stomach environment and contributes to the inflammatory response. In particular, HP infection has been seen to reduce the expression of the tumor suppressor protein TP53, leading to an increased risk of cancer occurrence [21]. HP infection promotes the epithelial-to-mesenchymal transition (EMT) process [22] and induces RAS protein activators like 2 (RASAL2), an oncogene. HP strains carrying the cag-pathogenicity island (cag-PAI) activate the Type IV secretion system [23]. This system enables the delivery of bacterial cytokines into gastric epithelial cells, thereby driving the progression of EMT through phenotypic modifications within these cells [24]. The expression of RASAL2 is induced by HP mediated by the nuclear factor-kappaB (NF-kappaB) and leads to the upregulation of beta-catenin transcriptional activity [25]. Tight junction proteins (TJPs) display an epithelial defense mechanism, TJP1 is unregulated in HP-positive gastric cancer, and after HP eradication, its higher expression was reversed [26]. However, so far, many other genetic targets of HP toward gastric cancer and mechanisms underlying the impact of HP infection on gastric cancer remain largely blurred.

In the overall design of this study, we initially conducted experiments to investigate the functional effects of HP infection on gastric cancer cells. Subsequently, we employed bioinformatic analyses to identify key genes involved in HP infection. Herein, we used TCGA data to analyze the differentially expressed genes in HP-positive gastric cancer, in order to elucidate the potential effects of HP on gastric tumors. These genes were then further validated through experimental testing. Finally, we used bioinformatic analysis to explore the clinical implications of these validated genes in gastric cancer. Our findings provide valuable insight into the influence of HP on gastric cancer.

2. Methods

2.1. Cell culture and helicobacter pylori (HP) infection

Helicobacter pylori strain ATCC 43504 and strain ATCC 700392 (American Type Culture Collection, Manassas, VA) were cultured on 5 % sheep blood agar under a tri-gas incubator with 5 % O_2, 10 % CO_2, and 85 % N_2 at 37 °C for approximately 48 h as a previous study

[27]. SGC-7901 and HGC27 human gastric cancer cells (ATCC, Manassas, VA) were cultured in RPMI 1640 medium containing 10 % fetal bovine serum under 5 % CO_2 at 37 °C. SGC-7901 and HGC27 cells were seeded with 200,000 cells in a 25 cm cell culture Petri dish. Then SGC-7901 and HGC27 cells were infected with *HP* strain ATCC 43504 or strain ATCC 700392 at a multiplicity of infection of 100:1 (MOI = 100:1) for 24 h respectively for RNA extraction. For observation of HP easily, before infection, HP strain ATCC 43504 or strain ATCC 700392 was marked with FITC at 37 °C for 1 h, then washed with Phosphate-buffered saline (PBS) and centrifuged at 2400 g for 5 min, repeated three times, and finally, marked-FITC HP strain ATCC 43504 or strain ATCC 700392 strain was collected and suspended in PBS for cell infection. FITC labeling serves as a rapid and intuitive method to ascertain whether HP has successfully invaded host cells. By quantifying fluorescence intensity, it becomes feasible to quantitatively analyze the number of HP strains, thereby assessing the severity of infection.

2.2. Cell viability

Cell viability was conducted using Cell Counting Kit-8 (CCK-8, Beyotime, China). Cells were plated in a 96-well plate and then the next day cells were infected with HP strain ATCC 43504 or strain ATCC 700392 at a multiplicity of infection (MOI) of 50:1 or 100:1 for 24 h respectively. After infection, 10 μL CCK-8 was incubated for 2 h, and the OD 450 nm was conducted in a microplate reader (BMG CLARIOstar, Germany). The control of the OD 450 nm was set as 100 % viability.

2.3. Lactate dehydrogenase (LDH) assay

The LDH assay measures the release of lactate dehydrogenase in cell culture. LDH is a cytosolic enzyme present in many cells. When cells are damaged, the cell membrane becomes compromised, allowing LDH to leak out into the surrounding medium. An increase in LDH activity in the medium indicates cell damage. After 24 h infection with HP strain ATCC 43504 or strain ATCC 700392, the supernatant of SGC-7901 cells was collected and centrifuged at 3000 rpm for 10 min. The LDH assay was conducted using a lactate dehydrogenase assay kit (Nanjing Jiancheng Bioengineering Institute, China) according to the manufacturer's protocol. The OD value of 450 nm was detected by a microplate reader (BMG CLARIOstar, Germany).

2.4. Wound healing assay

Cells were plated in a 24-well and after growing at confluence 100 %, cells were scratched using 10 μL pipette tips to create a "wound". Then cells were washed with PBS twice and cultured in serum-free medium. The images were captured for 0 h. The following cells were infected with marked-FITC HP strain ATCC 43504 or marked-FITC strain ATCC 700392 or a clinical HP strain ICDC11101 at MOI with 50:1 or 100:1. After 24 h, cells were photographed using IncuCyte® S3 (Sartorius, Germany).

2.5. Transwell assay

A total of 20,000 HGC27 cells were seeded into the upper chamber of a transwell that had been coated with Matrigel, and this was done using 200 μL of blank medium. Following this, the lower chamber was filled with either 500 μL of medium containing 10 % fetal bovine serum. Then HP at a multiplicity of infection (MOI) ratio of 100:1 was added into the upper chamber. After an incubation period of 24 h, the transwell chambers were fixed with paraformaldehyde for 15 min. The cells were then stained with crystal violet, after which they were washed with phosphate-buffered saline (PBS). Finally, the chamber was photographed.

2.6. TCGA data collection

RNA sequencing transcriptome data and clinical data of HP-positive and negative gastric cancer were downloaded from TCGA [28,29], following the guidelines and policies. The HP infection information was calculated with PathSeq [30] by the TCGA research network. There are 163 cases included in the analysis of HP differential expression analysis which included HP infection information in the data set downloaded. The baseline clinical information was provided in S-Table 1.

2.7. Differentially expressed gene analysis

HP-upregulated genes in gastric cancer were identified by differentially expressed gene analysis with a cut-off (log2 (fold-change) = 1 and *P*-value = 0.05) using the Limma package.

2.8. Functional enrichment analysis

The GO biological and KEGG pathway enrichment analyses were performed using R with clusterProfiler package. The org.Hs.eg.db package was run for annotation; the GOplot package was used to calculate Z-score. The GSEA analysis was conducted with the clusterProfiler package, c2.cp.v7.2.symbols.gmt as a reference, and MSigDB Collections as a gene set database [31,32]. The cut-off False discovery rate (FDR) and p.adjust were set at 0.25 and 0.05 respectively.

2.9. Survival analysis

The survival association was analyzed using the univariate Cox regression analysis [33]. Kaplan-Meier (KM) plot and log-rank analysis [34] were utilized to display the prognostic curve of representative genes.

2.10. The identification of hub genes

The protein-protein interaction network was established using STRING [35] with an interaction score of >0.15. The top 10 hub nodes in the network were calculated with the Density of Maximum Neighborhood Component (DMNC) algorithm using the Hubba [36] in Cytoscape [37].

2.11. Consensus clustering and LASSO model

All the samples were subtyped by consensus clustering using the survival-associated HP-up-regulated gene set. The ConsensusClusterPlus [38] package was used for clustering. The one-class logistic regression (OCLR) machine learning algorithm [39,40] was applied to stemness analysis. The M6A and Ferroptosis-related genes are based on previous studies [41,42]. The LASSO (least absolute shrinkage and selection operator) model was constructed using the R package glmnet [43].

2.12. Real-time Quantitative-PCR (RT-QPCR)

The cells were washed with PBS twice and lysed with RNA extract buffer for RNA extraction using RNeasy Mini Kit (Qiagen, Germany) followed by instructions. RNA was reversed to cDNA using PrimeScript™ RT Master Mix (Perfect Real Time) (Takara, Japan). 100 ng cDNA acted as DNA template, designed forward and reverse primers, ROX dye II and TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (Takara, Japan) was used to perform Real-time qPCR using Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific, USA). The running procedure was 95 °C 30 s for 1 cycle, then 95 °C 5 s, 60 °C 34 s for 40 cycles and the dissociation stage was 95 °C 15 s, 60 °C 1min, 95 °C 15s, and 60 °C 15 s. These sequences of primers are listed in Table 1 as follows. Gene expressions were calculated using the $2^{(-\triangle\triangle Ct)}$ method. GAPDH was used as a relative reference gene.

2.13. Immune analysis

We used single-sample Gene Set Enrichment Analysis (ssGSEA) to assess the levels of immune cell infiltration based on integrated immune gene sets from previous studies. This analysis was based on the R GSVA package version 1.34.0 [44]. Immune checkpoint blockade (ICB) responses of subtypes were predicted using the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm [45]. All the samples in the TCGA-STAD were used in this analysis and the data were normalized to log2 (TPM+0.01) prior to the analysis.

2.14. Drug analysis

Correlation between HP-associated genes and predicted drug IC50 in gastric cancer. The IC50 of drugs of each sample was predicted using models trained by IC50 and expression profiles of cancer cell lines from Genomics of Drug Sensitivity in Cancer (GDSC). The prediction process was implemented by the R package pRRophetic [46]. The "pRRophetic" package in R was employed to estimate the IC50 values of various anticancer drugs across a range of cell lines. The analysis was based on genomic profiles, utilizing ridge regression to model the relationship between drug response and genomic alterations. For data preparation, we integrated gene expression profiles sourced from the Cancer Genome Project (CGP). The "pRRophetic" package facilitated the implementation of the predictive model, with parameters meticulously adjusted to optimize the estimation. This methodological approach allowed for a precise estimation of drug sensitivity, providing valuable insights into the potential efficacy of anticancer drugs based on the molecular characteristics of individual cell lines. All the samples in the TCGA-STAD were used in this analysis and the data were normalized to

Table 1
Gene sequences for Real-time qPCR.

Gene Name	Forward sequences	Reverse sequences
GAPDH	AGGTCGGAGTCAACGGATTTG	TGTAAACCATGTAGTTGAGGTCA
DNER	AAGGCTATGAAGGTCCCAACT	CTGAGAGCGAGGCAGGATTT
ERBB4	GTCCAGCCCAGCGATTCTC	AGAGCCACTAACACGTAGCCT
BRINP2	CAGTGTGGCACTCGGTTTAGA	AAGTCAGCATACTCCTGAGCG
KCTD16	ACTCGCCATTCCACATTGATAAG	TGGAGTCCTTGGCTAGATCATTA
MAPK4	CGGTGTCAATGGTTTGGTGC	GACGATGTTGTCGTGGTCCA
THPO	AACTGCAAGGCTAACGCTGT	GACATGGGAGTCACGAAGCA
VSTM2L	CAGAGACACCCCATGACATGA	GCTCCGTACATACCACCACTG
F5	TCCAGGCCGAGAATACACCTA	CGATTTGCTTGTCAAACGTCTTC
DAO	AATCTCGGGCTACAACCTCTT	TCAGCTTCCGAAATCCCAGAA
NKAIN2	ATCAGTGGGCACCTATCCTG	GAGGACTAGCCAGACAGCATA

log2 (TPM+0.01) prior to the analysis.

2.15. Expression regulation analysis

Copy number variation (CNV) data were collected from the NCI Genomic Data Commons. CNV data were processed with GISTICS2.0 [47]. In the GISTICS, −2 represents deep deletion; −1 represents shallow deletion; 0 represents diploid; 1 represents gain; 2 represents amplification. The correlation analysis was based on the above four types of CNV values. This method has been used in a previous study [48]. The methylation data were downloaded from the TCGA database. Because one gene might have multiple methylation sites, we performed correlation analysis and only presented the site most negatively correlated with gene expression. The miRNA regulation data was collected from multiple databases, including targetscan and miRanda predicted as well as experimentally verified data from papers, TarBase, miRTarBase, and mir2disease. The regulation was confirmed by analyzing the correlation between mRNA and miRNA expression. The correlation was calculated in all TCGA samples not only gastric cancer because the STAD cohort only was insufficient to provide firm results. In the network, the connection between a miRNA node and a gene node represent miRNA regulation of a gene. The edge width is defined by the absolute value of the correlation coefficient between the miRNA and the gene. The visNetwork R package was employed to generate a network and the networkD3 package was used to generate the interactive network in the supplementary materials.

2.16. Statistical significance

A significance level of $P < 0.05$ was used as the cutoff for statistical significance. Spearman correlation analyses were conducted for all correlation analyses. For intergroup comparisons, the stats R package was utilized to analyze the differences between groups, with the software automatically selecting the appropriate analysis method to calculate significance. Regarding the assumption of normality on numerical variables, the choice of display and statistical test methods depended on whether a variable was deemed to follow a normal distribution. If a variable was determined to be normally distributed (for conducting t-tests), the corresponding variable could be selected, and the program would automatically return the relevant options. The summary format for the numerical variable would then be displayed as "Mean ± Standard Deviation." Similarly, for categorical variables, the choice of statistical test method was influenced. If a variable was identified to require a chi-square test, the corresponding variable could be selected, and the program would return the appropriate options.

3. Results

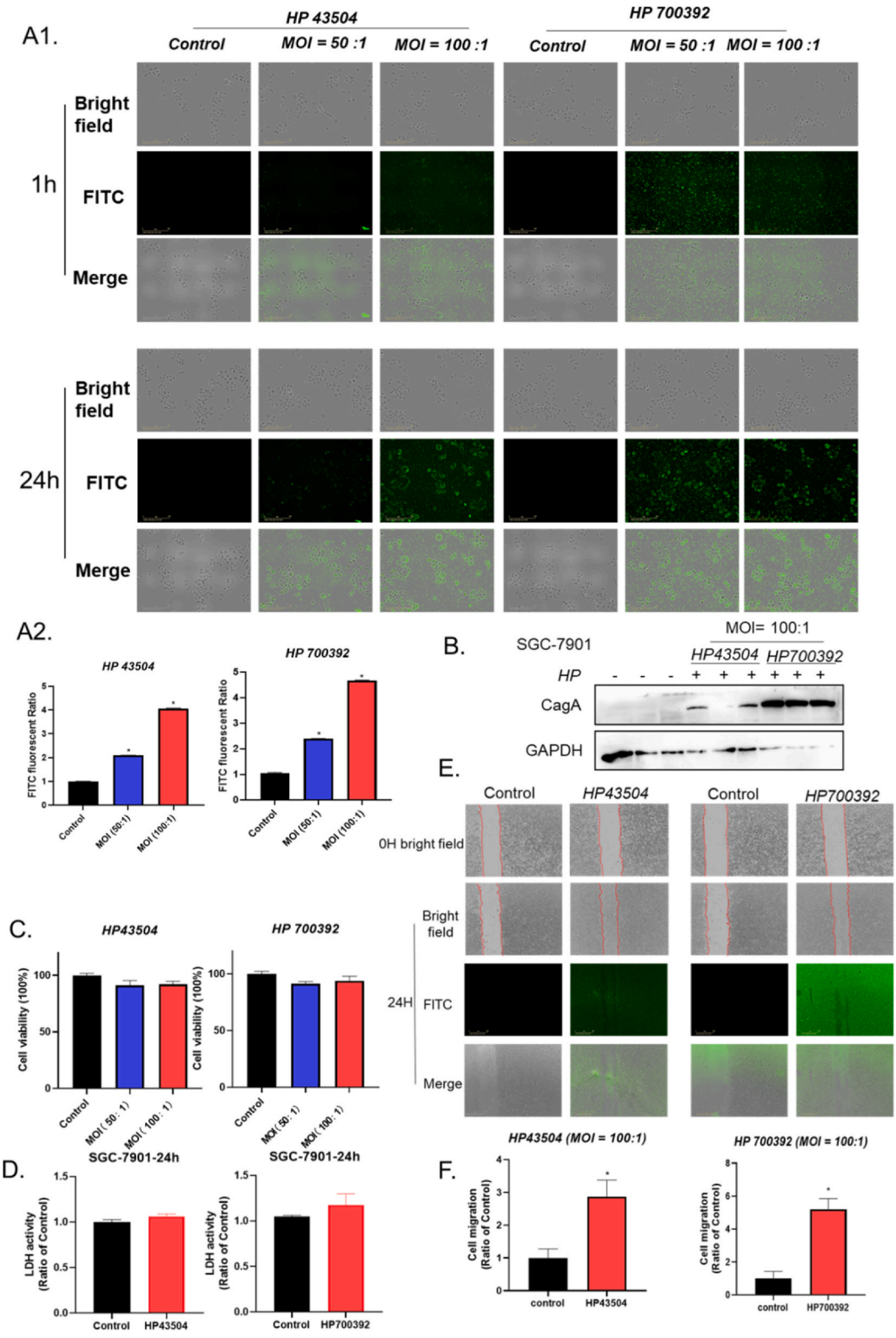
3.1. Effect of HP on human gastric cancer cells

To determine the potential effect of HP on human gastric tumor cells, we employed an HP infection gastric tumor cell model and detected the impact of HP infection on cell viability and migration. SGC-7901 human gastric cancer cells were infected with two HP strains respectively, HP ATCC 43504 or HP ATCC 700392, at a multiplicity of infection (MOI) of 50:1 or 100:1 for 24 h. HP was dyed and marked by FITC for 1 h before infection of SGC-7901 cells. We observed cell morphology at 1 h and 24 h and also tested the viability, Lactate dehydrogenase assay (LDH), and migration. Both MOI of 50:1 and 100:1 infected the cells successfully with different infection rates, as MOI of 100:1 had a higher GFP intensity than MOI of 50:1 (Fig. 1A1&A2). To confirm the infection of HP, we determined the cytotoxin-associated gene A (Cag A) expression using western-blot (Fig. 1B). The objective of using Western blot analysis for detecting CagA expression in HP-infected cells is to specifically identify and quantify CagA protein, a critical virulence factor of HP. This method also verifies infection status since CagA expression serves as a biological marker for HP infection. Utilizing antibodies that target CagA, this technique confirms its presence in infected cells and allows for a semi-quantitative assessment of its expression levels. This is crucial for understanding the correlation between HP infection and CagA expression.

Functional assay results showed that neither of the HP stains significantly altered the cell viability and LDH release of gastric cancer cells (Fig. 1C and D). However, the wound healing migration assay revealed that both of the HP stains significantly facilitated the migration of cells (Fig. 1E and F). In Fig. 1, panel A&B shows that the HP is successfully infected with gastric cancer cells. Panel C&D rejected the hypothesis that HP affects cell proliferation. Although HP has been reported to promote cell proliferation [49], in our study, for these specific cell lines and our experimental condition (medium, culture time, infection approach), the difference in proliferation (generally, because we did not directly test proliferation, what we have are overall viability and overall cell death) was not significant. In panel E&F, the main finding is that HP affects cell migration. All in all, our study using an HP infection gastric tumor cell model demonstrated that HP infection did not significantly affect cell viability or cytotoxicity in SGC-7901 human gastric cancer cells. However, it did promote cell migration as observed in the wound healing migration assay. These findings suggest that HP infection may play a role in enhancing the migratory potential of gastric cancer cells, highlighting its potential contribution to tumor metastasis and underscoring the importance of considering HP infection in the management and treatment of gastric cancer.

3.2. Enrichment of HP-upregulated genes in gastric cancer

To clarify the effect of HP infection on gastric cancer, TCGA STAD data were collected and gene expressions in gastric cancer tissues with or without HP infection were distinguished to analyze the role of HP in gastric cancer. This study started with the differentially expressed gene analysis to identify the higher expressed genes in HP-positive samples (n = 18) compared with the HP-negative samples



(caption on next page)

Fig. 1. Impact of HP on human gastric cancer cells. SGC-7901 human gastric cancer cells were infected with HP ATCC43504 or HP ATCC700392 at a multiplicity of infection (MOI) of 50:1 or 100:1 for 1 h or 24 h. **A1.** The cell morphology of SGC-7901 was observed (200×) after HP infection. HP was dyed and marked by FITC for 1 h before infection of SGC-7901 cells. **A2.** The FITC fluorescent ratio of SGC-7901 cells after HP infection for 24 h. **B.** Cytotoxin-associated gene A (Cag A) expression was measured by Western blot. **C.** The cell viability of SGC-7901 cells was assayed by CCK-8 assay after HP infection for 24 h. **D.** The Lactate dehydrogenase assay (LDH) activity was detected after HP infection for 24 h. **E.** Representative images of cell migration of SGC-7901 cells (40×) after HP infection. **F.** Cell migration ratio was measured. $^{*}P < 0.05$.

(n = 145). The analysis identified 1405 genes that were significantly higher expressed in HP-positive samples (Fig. 2A). To investigate the functions and biological processes of these genes involved in gastric cancer, we conducted GO/KEGG pathway enrichment and Gene Set Enrichment Analysis (GSEA). GO/KEGG enrichment analysis facilitates the understanding of the biological roles of genes associated with gastric cancer through multiple avenues. It categorizes genes into functional groups and pathways, supports hypothesis generation regarding disease mechanisms, identifies potential drug targets, enables comparative studies, integrates with other datasets, and aids in predictive modeling. Together, these functions enhance our understanding and treatment of the disease. As shown in the representative enriched terms in Fig. 2B, the GO/KEGG pathway enrichment showed that identified 1405 genes were enriched in multiple terms that involved cancer migration, proliferation, drug, and immunity, such as cell-cell adhesion terms, growth factor activity, drug catabolic and transmembrane transportation terms, humoral immune response, etc. The complete results of the enrichment analysis are provided in the supplementary file "supplementary table for enrichment analysis". On the other hand, the GSEA results only identified two pathways that the HP-upregulated genes significantly involved, including "Reactome keratinization" and "Reactome formation of the cornified envelope". As keratinization helps maintain the mechanical stability of individual cells and epithelial tissues, HP might affect cell adhesion by regulating keratinization in gastric cancer, thereby impacting cell migration (Fig. 2C). In conclusion, this study provides insights into the potential mechanisms by which HP infection influences gene expression and affects biological processes in gastric cancer. The findings suggest that HP may contribute to the aggressive behavior of gastric cancer through modulation of cell adhesion, migration, proliferation, drug response, and immune processes. Further research is needed to validate these findings and unravel the precise molecular mechanisms underlying HP-induced effects in gastric cancer.

3.3. Identification of HP-associated genes

To identify survival-associated HP-upregulated genes, we conducted a univariate Cox regression analysis to determine the survival association of the HP-upregulated genes. Results revealed that among 1405 HP-upregulated genes, 82 genes were significantly risky factors. In this survival analysis, we separated TCGA patients into two groups based on the expression level of each gene, using the median expression level as the cutoff. We then applied Cox regression analysis to assess survival differences and calculated the Hazard Ratio (HR) for each gene. The top survival-critical genes included SYT14, THPO, and PCDHA11. The top 60 genes associated with worse overall survival are shown in Fig. 4A. The changes in these genes inferred that HP affects survival. The survival-associated HP-upregulated genes might contribute to the potential risk of HP in gastric cancer, hence the gastric cancer subtypes based on these genes can be used to investigate the comprehensive influence of HP infection on gastric cancer. Based on the consensus cumulative distribution function (CDF) plotting, the optimized cluster number (K) was 4 (S-Fig. 1A and 1B). By the Non-Negative Matrix Factorization (NMF) method, patients were clustered into four subtypes (S-Fig. 1C). The heatmap and PCA plotting of the subtypes also demonstrated the distinction of four subtypes (S-Fig. 1D and 1E). The objective of clustering analysis was not to identify HP-associated subgroups, as the differentiation between HP-positive and HP-negative groups has already been established based on the identified differential expression of these genes. Instead, the primary aim of clustering was to treat the 82 genes as a collective entity and explore the overall impact they exert. By examining the subtype differences, we gain a comprehensive understanding of the potential consequences of HP infection on these genes and their functional implications. Results revealed that patients in different subtypes had different overall survival (Fig. 3A) as well as progression-free survival, disease-specific survival, and disease-free survival (S-Fig. 2). We also analyzed the association between the subtypes and clinical characteristics, including gender, ethnicity, TNM staging, grade, and radiotherapy. However, we did not find any significant association (S-Fig. 3). C1 has a relatively higher percentage of HP-positive samples, hence, might associate with HP infection (Fig. 3B). Moreover, we analyzed the mutation rate of a typical tumor inhibitor, TP53. Results indicated that the mutation rates in different subgroups were significantly different (Fig. 3C). TP53 mutation has been previously used in molecular subtyping gastric cancer [50]. Here, we compared our subtype with the previous subtyping system using a Sanky plot. Although the p53 significantly enriched in the C3 HP cluster, the difference is not dramatic. In the Sanky plot, the molecular subtype was not correlated with C3 (Fig. 3D). The molecular subtype was based on Asian Cancer Research Group subtypes [51]. This demonstrated that our cluster generated a unique subtyping system apart from some current molecular subtypes. The survival analysis suggested the HP cluster can be used for prognosis. Hence, it can provide prognostic value addition to the current molecular subtypes. Furthermore, we also calculate the mRNAsi score for the evaluation of cancer stemness in different subtypes. Results showed that there were significant differences between the subtypes (Fig. 3E). Another analysis we did was to predict the drug response of three FDA-approved gastric cancer drugs of patients in different subtypes, including Mitomycin, Doxorubicin, and Docetaxel. Results predicted that IC50 of Mitomycin were significantly different (Fig. 3F) while the IC50s of Doxorubicin and Docetaxel were not (S-Fig. 4). This analysis suggested that HP infection might affect the sensitivities of cancer toward some gastric cancer drugs. Recent studies also suggested that M6A methylation and ferroptosis were critical for cancers, thus, this study analyzed the levels of M6A-related and ferroptosis-related genes across subtypes. Results showed that many of the M6A-related and

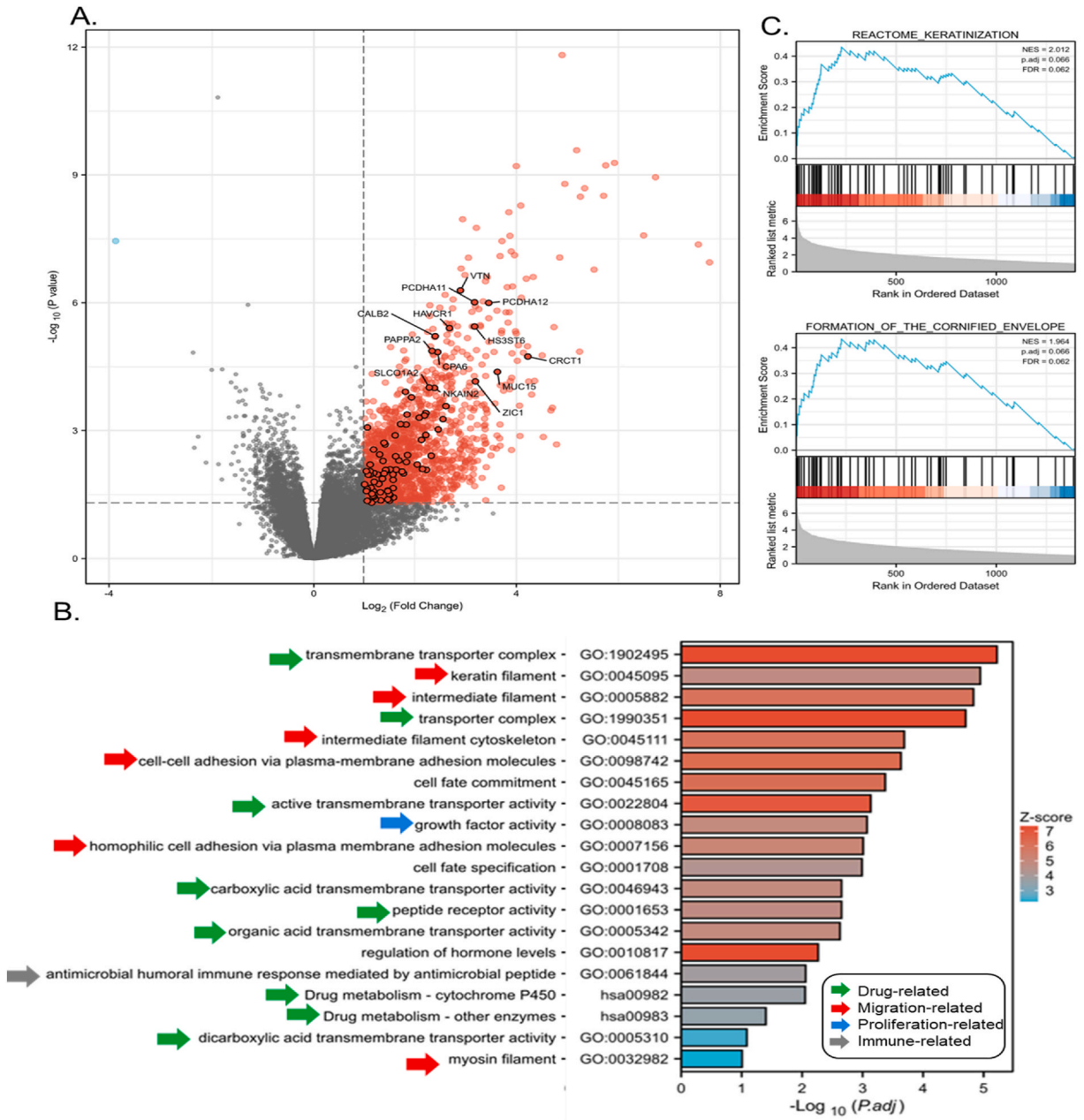
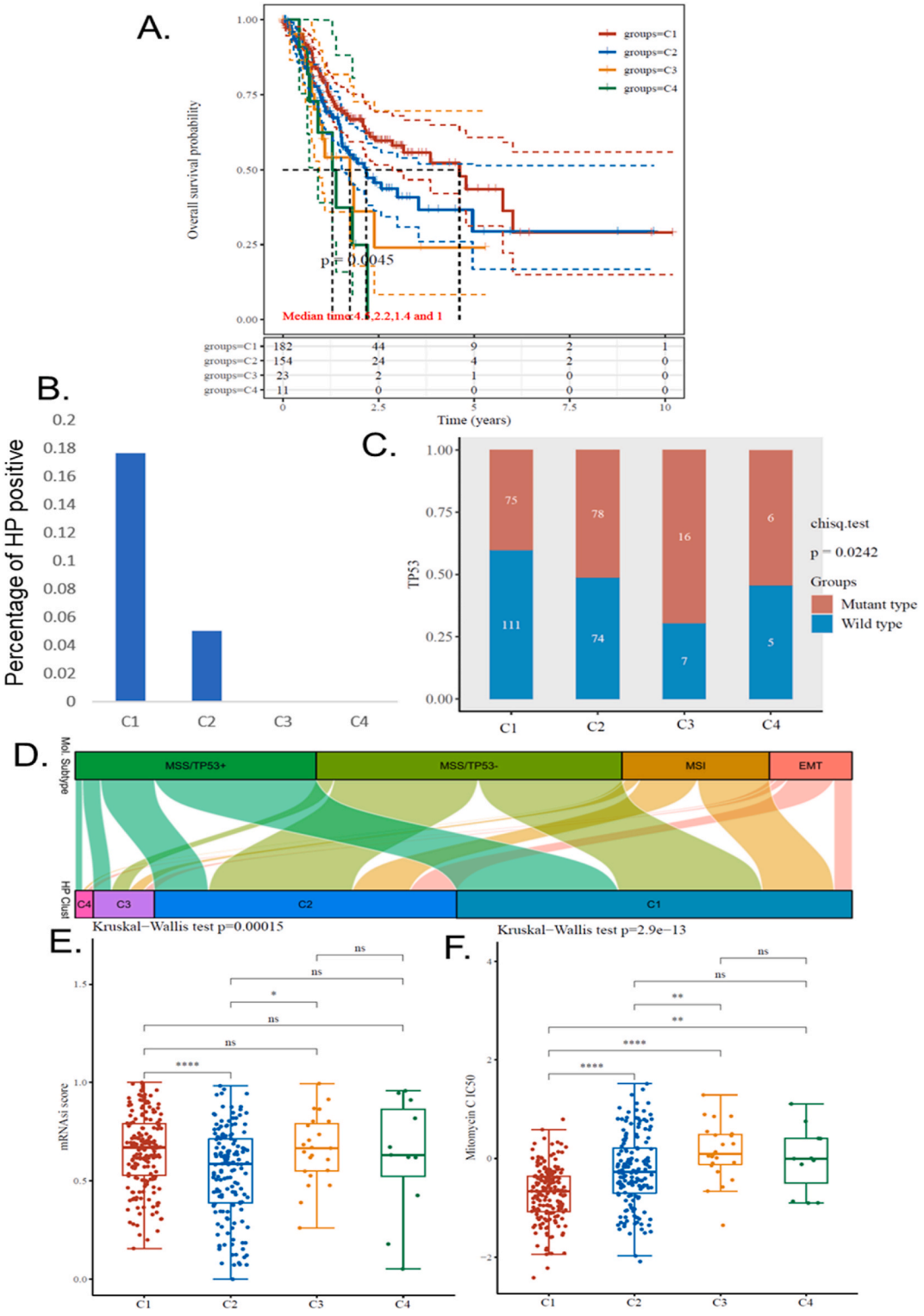


Fig. 2. Enrichment of HP-upregulated genes in gastric cancer. TCGA STAD cohort was analyzed. **A.** Volcano plot of the differentially expressed genes between HP-positive and HP-negative gastric cancer tissues. The x-axis represents the fold change of the gene in HP-positive gastric cancer compared to HP-negative cancer. The y-axis represents the p-value of the difference. Genes associated with worse overall survival were highlighted with black borders. **B.** Representative GO and KEGG enrichment results of HP-upregulated genes. The length of the bar represents the adjusted p-value of the enrichment analysis and the colors of the bars represent the activation Z-score. The activation Z-score makes predictions about HP by using information about the direction of gene regulation. Complete results were provided in the supplementary file "supplementary table for enrichment analysis". **C.** GSEA enrichment of HP-upregulated genes. GSEA enrichment was conducted and the cancer-related pathways with significance were plotted in the figure. The peak point of the blue plot on the top of each panel represents the enrichment score, which tells how over- or under-expressed these genes are with respect to the ranked list. The second part of the graph (middle with red and blue) shows where the rest of the genes related to the pathway or feature are located in the ranking. The third part of the figure (bottom with grey part) shows how this metric is distributed along the list.



(caption on next page)

Fig. 3. The difference among gastric cancer subtypes based on HP cluster. The patients in TCGA-STAD cohorts were clustered into 4 subtypes using the Non-Negative Matrix Factorization (NMF) method based on the survival-associated HP-upregulated gens in gastric cancer. **A.** Overall survival KM plot of the subtypes. **B.** Percentage of HP positive samples in HP cluster subtypes. **C.** TP53 mutation rate of the HP cluster subtypes. **D.** Sanky plot comparing HP clusters and molecular subtypes of gastric cancer. **E.** The mRNAsi score of the HP cluster subtypes. The mRNAsi score, also known as the mRNA-based stemness index, is a computational score that quantifies the degree of stemness or pluripotency in cancer cells based on their mRNA expression profiles. **F.** Predicted Mitomycin sensitivity of the HP cluster subtypes. The IC50 of drugs of each sample was predicted using models trained by IC50 and expression profiles of cancer cell lines from Genomics of Drug Sensitivity in Cancer (GDSC).

ferroptosis-related genes were significantly different in different subtypes, inferring that HP might affect M6A and ferroptosis (S-Fig. 5).

Taken together, our analysis of survival-associated HP-upregulated genes identified a set of genes that were significantly associated with worse overall survival in gastric cancer. Furthermore, the clustering of patients based on these genes revealed distinct subtypes, which showed differences in survival, TP53 mutation rates, stemness characteristics, and predicted drug responses. These findings highlight the potential impact of HP infection on gastric cancer prognosis and provide insights into the molecular subtyping and personalized treatment strategies for patients with HP-associated gastric cancer.

3.4. Protein-protein interaction (PPI) network and hub genes

To display the interconnection of survival-critical HP-upregulated genes, we established a PPI network. The construction of a PPI network for survival-critical HP-upregulated genes elucidates their interrelationships within the regulatory framework, containing 82 nodes and 265 edges. This network not only aids in identifying hub genes that are central and likely crucial to the gene set but also shows that these genes are functionally interconnected. The density of the network is indicative of the gene interactions: a high density suggests numerous interactions and a tightly connected network, potentially vital for cancer development, whereas a low density indicates fewer interactions and possible involvement in disparate pathways. Nodes with high connectivity are pivotal and may serve as key therapeutic targets due to their central roles in the cancer process. Conversely, nodes with low connectivity, while potentially less critical to overall cancer development, might have more specialized functions. Although the PPI network is broadly based and not specific to gastric cancer or cancer in general, it provides insights into the interrelationships among a list of genes. Nodes with high connectivity (hubs) are central and could play critical roles in the cancer process, potentially serving as key therapeutic targets. We calculated the top 10 hub genes in the network using the DMNC algorithm, including ERBB4, DNER, F5, MAPK4, KCTD16, THPO, DAO, BRINP2, NKAIN2, and VSTM2L (Fig. 4B). We further plot the KM survival curves for these 10 genes and the HR and p-value of the best cut-off are shown in Fig. 4C. To develop the potential of the HP-upregulated gene set, we constructed a prognostic model. In this study, we utilized LASSO regression to select the genes included and the coefficients in the prognostic model from the 10 hub genes identified above (S-Fig. 6A and B). 6 genes were selected to optimize the prognostic model, including ERBB4, F5, MAPK4, KCTD16, THPO, and VSTM2L with coefficients of 0.1345, 0.0645, 0.0219, 0.0334, 0.0877, and 0.0332 respectively (S-Fig. 6C). The log-rank analysis demonstrated that the risk score generated by the model can predict the survival of gastric cancer patients (S-Fig. 6D and E). Collectively, our analysis of the survival-critical HP-upregulated genes revealed a functionally associated protein-protein interaction network, highlighting the interconnection of these genes in gastric cancer. The identification of hub genes such as ERBB4, F5, and THPO, along with the development of a prognostic model using LASSO regression, demonstrates the potential of these genes as prognostic markers for gastric cancer patients. These findings provide valuable insights into the molecular mechanisms and potential therapeutic targets associated with HP-upregulated genes in gastric cancer, with implications for personalized treatment approaches.

3.5. Expression of HP-associated genes

To validate the HP up-regulation of the 10 hub genes, we conducted an RT-QPCR assay to test their alteration in gastric cancer cells after 24 h of HP infection. F5, DAO, and NKAIN2 were not detected in the RT-QPCR assay, while the other 7 genes were detected. Results revealed that both of the HP upregulated gene expressions of *ERBB4*, *DNER*, *BRINP2*, *KCTD16*, *MAPK4*, *THPO*, and *VSTM2L* in SGC-7901 human gastric cancer cells respectively (Fig. 4D and E). (We would refer to *ERBB4*, *DNER*, *BRINP2*, *KCTD16*, *MAPK4*, *THPO*, and *VSTM2L* as "HP-associated genes" in the rest of the paper.) Among these genes, KCTD16 was the most remarkable differential gene caused by *HP* with a fold change of almost 10. However, these genes were not necessarily expressed higher in gastric cancer compared to normal tissues. Instead, unpaired comparisons using TCGA and GTEx data suggested that VSTM2L, MAPK4, ERBB4, and DNER were expressed lower in the tumor, while THOP was the only gene expressed higher in the tumor (S-Fig. 7A). The paired comparisons using TCGA cancer and para-cancer tissue data further confirmed that MAPK4 and DNER were expressed significantly lower in cancer (S-Fig. 7B). A comparison of HP-positive and HP-negative with normal control was also provided in S-Fig. 7C for reference. Notice that some of these genes are expressed higher in normal than tumor tissue, we believe the genes that are expressed higher in normal than tumor tissue might be cancer inhibitors. They are up-regulated by HP because of some negative feedback of cancer-promoting pathways. One possibility is that these genes are associated with gastric intestinal metaplasia more than gastric adenocarcinoma, but further studies are required to figure out this issue. All in all, our RT-QPCR assay confirmed the upregulation of ERBB4, DNER, BRINP2, KCTD16, MAPK4, THPO, and VSTM2L genes in gastric cancer cells upon HP infection, with KCTD16 showing the most significant upregulation. However, the expression levels of these genes in gastric cancer compared to normal tissues varied, with some genes being downregulated in tumor samples. These findings highlight the complex gene expression patterns associated with HP infection and gastric cancer, emphasizing the need for subsequent investigation to understand their medical significance and potential

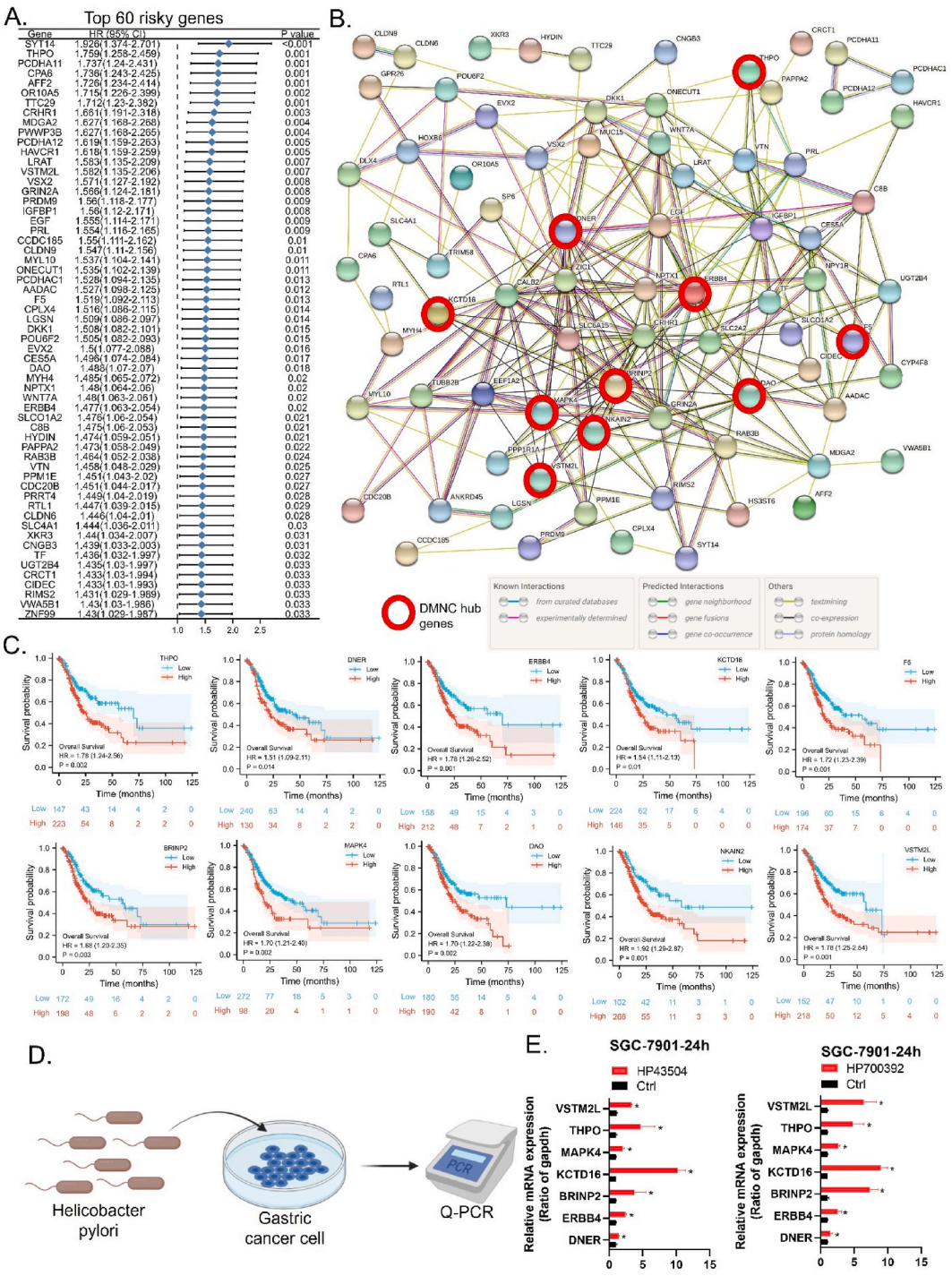


Fig. 4. Identification of the critical HP-associated genes in gastric cancer. **A.** Survival association of the HP-upregulated genes. TCGA-STAD cohort was used to analyze the survival association. The association between HP-associated genes and the overall survival of gastric cancer patients were analyzed using Cox regression. The hazard ratio and p-value were presented in a forest plot. The top 60 genes associated with worse overall survival were shown. **B.** Protein-protein interaction network of survival-associated genes in gastric cancer. The top 10 hub genes were marked by red nods. As described in the legend, this network includes data from curated databases, experiments, predicted interactions, etc. **C.** Survival KM plot of the top 10 hub genes in gastric cancer patients. TCGA STAD data were used. **D.** The schematic diagram of the validation of HP-upregulated gene expression. **E.** Validation of HP-upregulated genes in HP-infected SGC-7901 human gastric cancer cells. SGC-7901 human gastric cancer cells were infected with HP ATCC43504 or HP ATCC700392 at a multiplicity of infection (MOI) of 100:1 for 24 h. The gene expressions were detected using Real-time qPCR. GAPDH was used as a relative reference. $^{*}P < 0.05$.

implications for targeted therapies.

3.6. Clinical association of HP-associated genes in gastric cancer

We believed that the potential role of the HP-associated genes in the gastric tumor might mediate the impact of HP on cancer, hence, we investigated the clinical association of these 7 genes in gastric cancer patients using the TCGA-STAD cohort. First, we compared the clinical difference between high gene-expressing samples and low gene-expressing samples (cutoff: 50 % vs 50 %) in the T stage, N stage, M stage, pathologic stage, primary therapy outcome, gender, race, age, histological type, residual tumor, histologic grade, anatomic neoplasm subdivision, reflux history, anti-reflux treatment, and Barrett's esophagus. As shown in Table 2, age was associated with all these genes. VSTM2L was associated with N stage pathologic stage and primary therapy outcome. THPO was associated with T stage, race, histologic grade, and Barrett's esophagus. THPO was also associated with histological type and histologic grade. ERBB4 was associated with pathologic stage and race. Detailed results were provided in the supplementary file "Supplementary Table for the clinical association of HP-associated genes". Age is the only continuous variable and we provided additional analysis for age with the *t*-test or Wilcoxon method (depending on the homogeneity of variance test) which was shown at the bottom of the supplementary table. In summary, we observed significant clinical associations between the HP-associated genes (ERBB4, DNER, BRINP2, KCTD16, MAPK4, THPO, and VSTM2L) and various clinicopathological factors in gastric cancer patients, including age, tumor stage, race, histological type, and primary therapy outcome. These findings suggest that these genes may play a role in gastric cancer progression and could potentially serve as prognostic or therapeutic targets, highlighting their medical significance and the need for further investigation into their functional implications.

3.7. Correlation analysis of HP-associated genes in gastric cancer

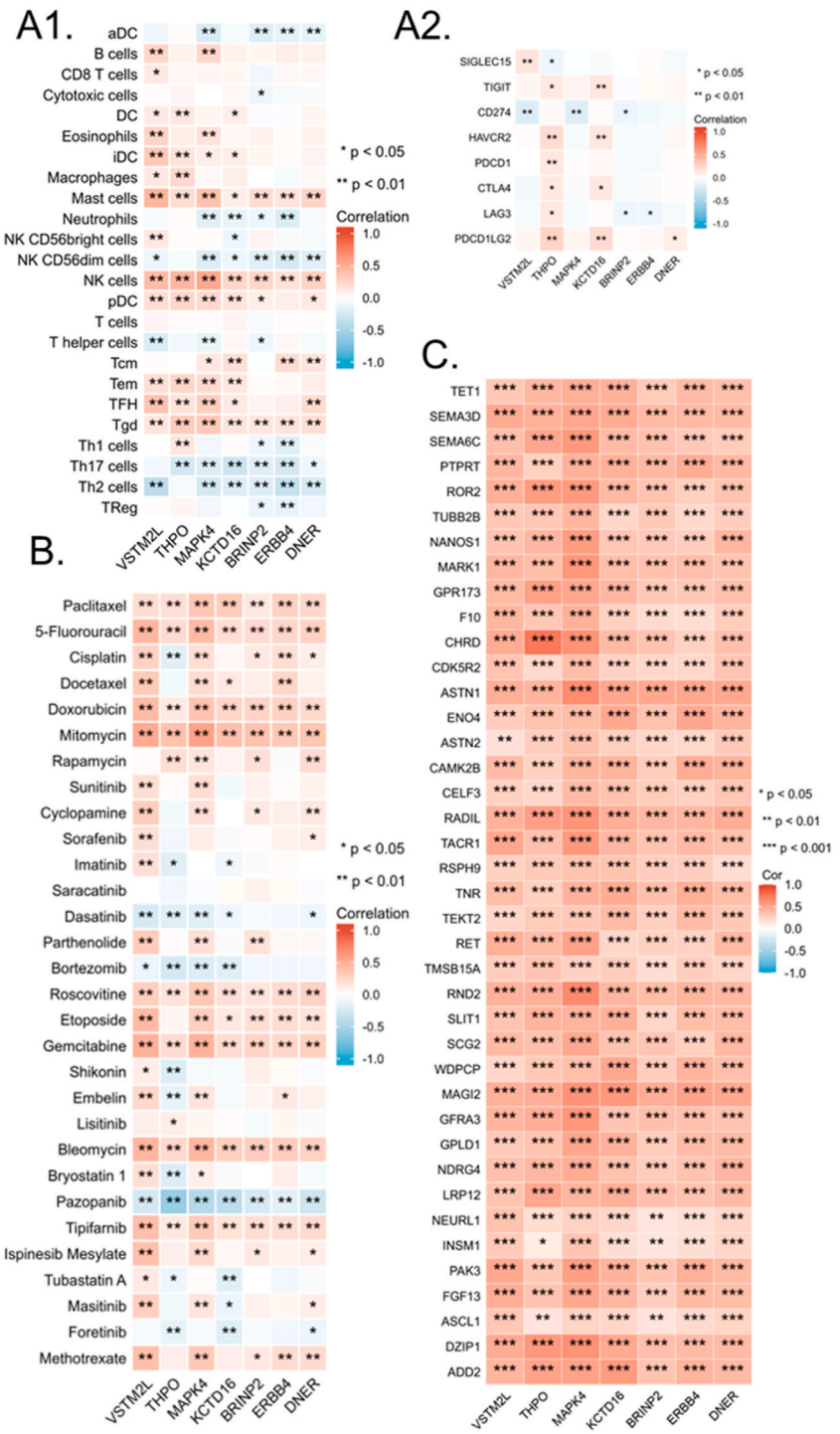
Our enrichment analysis inferred that HP might affect gastric cancer immunity; therefore, we explored the immune association of the HP-associated genes in gastric cancer. The recent advances in genomic sequencing and bioinformatics have allowed for the high-throughput analysis and interpretation of complex disease-related datasets, which is an ideal approach to quantifying the tumor-infiltrating immune cells of various cancers. Analyzing HP-negative and HP-positive subjects separately to see if the HP-positive samples are more inflamed overall is a sensible idea. However, given the low case number of the HP-positive group, the correlation is hardly significant. Hence, we analyzed the correlation for all samples.

Single-sample Gene Set Enrichment Analysis (ssGSEA) is an extension of Gene Set Enrichment Analysis (GSEA) that computes separate enrichment scores for each pairing of a sample and gene set. This transformation of a single sample's gene expression profile to a gene set enrichment profile enables researchers to characterize the tumor-infiltrating immune cells of the tumor microenvironment rather than through immunohistochemistry and flow cytometry. ssGSEA thus permits the definition of immune cell-related gene sets, which can be used to quantify the density of tumor-infiltrating immune cells. We calculated the correlation between HP-associated genes and predicted immune cell infiltration levels in gastric cancer. The correlations indicated that the HP-associated genes were significantly correlated with multiple immune cells. For example, mast cells, NK cells, and Tgd cells were positively correlated with all 7 genes (Fig. 5A1). Notably, Th17 and Treg cells were apparently less abundant in HP + cases than HP- cases, which is unusual because HP infection is known to elicit both a Th17 response and Tregs [52,53]. In this study, we conducted an analysis to investigate the relationship between core HP genes and immune infiltration. Our findings revealed a significant negative correlation between these 7 genes and Th17 and Treg cells. However, it is important to note that our study's correlational nature does not provide direct evidence of a decrease in Th17 and Treg cells following HP infection. Experimental validation would be required to establish causality. Nonetheless, these findings are intriguing and suggest a potential association between HP infection and alterations in

Table 2
Clinical association of HP-associated genes in gastric cancer (TCGA-STAD data).

Gene	VSTM2L	THPO	MAPK4	KCTD16	BRINP2	ERBB4	DNER
T stage	ns	*	ns	ns	ns	ns	ns
N stage	*	ns	ns	ns	ns	ns	ns
M stage	ns	ns	ns	ns	ns	ns	ns
Pathologic stage	*	ns	ns	ns	ns	*	ns
Primary therapy outcome	*	ns	ns	ns	ns	ns	ns
Gender	ns	ns	ns	ns	ns	ns	ns
Race	ns	*	ns	ns	ns	*	ns
Age	*	*	*	*	*	*	*
Histological type	ns	ns	ns	ns	*	ns	ns
Residual tumor	ns	ns	ns	ns	ns	ns	ns
Histologic grade	ns	*	ns	ns	*	ns	ns
Anatomic neoplasm subdivision	ns	ns	ns	ns	ns	ns	ns
Reflux history	ns	ns	ns	ns	ns	ns	ns
Antireflux treatment	ns	ns	ns	ns	ns	ns	ns
Barretts esophagus	ns	*	ns	ns	ns	ns	ns

Note: ns: not significant.
*$p < 0.05$.



(caption on next page)

Fig. 5. Correlation analysis of HP-associated genes in gastric cancer. These analyses were the results of analysis using the TCGA-STAD cohort. **A1**. Correlation between HP-associated genes and predicted immune cell infiltration levels in gastric cancer. The immune cell infiltration levels were calculated using the ssGSEA algorithm. The immune cell infiltration levels of each bulk sample were calculated respectively and the correlation between the immune cell infiltration levels and the levels of HP-associated genes were calculated respectively. **A2**. Correlation between HP-associated genes and immune checkpoint levels in gastric cancer. **B**. Correlation between HP-associated genes and predicted drug IC50 in gastric cancer. The IC50 of drugs of each sample was predicted using models trained by IC50 and expression profiles of cancer cell lines from Genomics of Drug Sensitivity in Cancer (GDSC). Then the correlation between the IC50 of drugs and the levels of HP-associated genes were calculated respectively. **C**. Correlation between HP-associated genes and migration-associated genes. The migration-associated gene list was downloaded from GSEA Human Gene Set "GOBP_CELL_MOTILITY". Results of the top 40 migration-associated genes were plotted. *P < 0.05, **P < 0.01, ***P < 0.001.

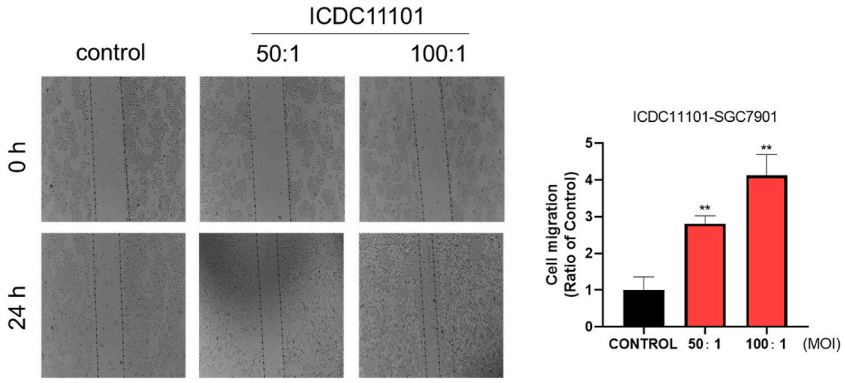


Fig. 6. Validation of the impact of HP on gastric cancer cell migration using a clinical HP strain. SGC7901 cells were scratched, washed with PBS, and initially photographed at 0 h. Subsequently, they were infected with the clinical HP strain ICDC11101 at an MOI of 100:1. After 24 h of infection, the cells were photographed again to assess changes over the 24 h.

immune cell populations. Future studies involving peripheral blood analysis from patients could provide further insights into the complex interplay between HP and immune cells.

To explore the value of these genes in immune therapies as well as the potential impact of HP infection on immune therapies, we calculated the correlation between HP-associated genes and immune checkpoint levels in gastric cancer. Results revealed that only a few immune checkpoints were associated with these genes. Among these results, THPO had the most significant correlation with the immune checkpoints (Fig. 5A2). These inferred that these genes might be impactful in some types of immune therapy.

We noticed that our enrichment analysis also inferred that HP might impact drug sensitivity, hence, we calculated the correlation between HP-associated genes and predicted drug IC50 in gastric cancer. The analysis combines several computational and statistical tools to predict in vivo drug response, using models trained on cell line data as described in a previous paper [46]. To do this, we used published data from the Genomics of Drug Sensitivity in Cancer (GDSC), which included baseline gene expression microarray data and sensitivity to drugs in a panel of almost a large number of cell lines. By building a statistical model from these data, we capture a significant amount of the variability in drug response in patients. Then the predicted drug sensitivity of the sample in the TCGA-STAD cohort was calculated using these models. In this study, we calculated drugs of several FDA-approved medicines used for gastric cancer chemotherapy, such as 5-Fluorouracil, Cisplatin, Docetaxel, Doxorubicin, and Mitomycin, as well as a series of potential gastric cancer drugs suggested such as Gemcitabine, Shikonin and so on. The detailed information of these drugs is shown in S-Table 2. The TCGA-STAD dataset is all treatment-naïve samples. The IC50 was predicted using a model trained by multiple cancer cell lines of different types of tumors, thus, this analysis only provides some hints for gastric cancer. Results revealed that the levels of these genes were significantly correlated with multiple gastric cancer drugs respectively. Remarkably, these 7 genes significantly correlate to most of the gastric cancer first-line drugs available (Fig. 5B). These results strongly suggested that HP infection might alter the sensitivity of gastric cancer to chemotherapy.

Last but not least, we have demonstrated that HP infection enhances the mobility of gastric cancer cells with experimental evidence, here we calculated the correlation between HP-associated genes and migration-associated genes in gastric cancer tissue to identify potential targets that mediated the effects of the HP-associated genes on gastric cancer migration. The migration-associated gene list was downloaded from GSEA Human Gene Set GOBP_CELL_MOTILITY (GO:0048870). Results revealed that the 7 HP-associated genes were positively correlated to multiple migration-associated genes (Fig. 5C). Therefore, these identified migration-associated genes might mediate the effect of HP on gastric cancer migration. Our enrichment analysis only suggests that HP can affect migration. While the correlation analysis shows it is these genes that mediate the impact of HP on cell migration. Further studies are required to prove this conclusion. Overall, our analysis revealed that the HP-associated genes in gastric cancer are significantly correlated with immune cell infiltration, suggesting their potential role in modulating the tumor microenvironment and influencing immune therapy outcomes. Additionally, these genes showed significant correlations with immune checkpoints, indicating their relevance to immune checkpoint blockade therapies. Furthermore, the correlation between these genes and drug sensitivity suggests

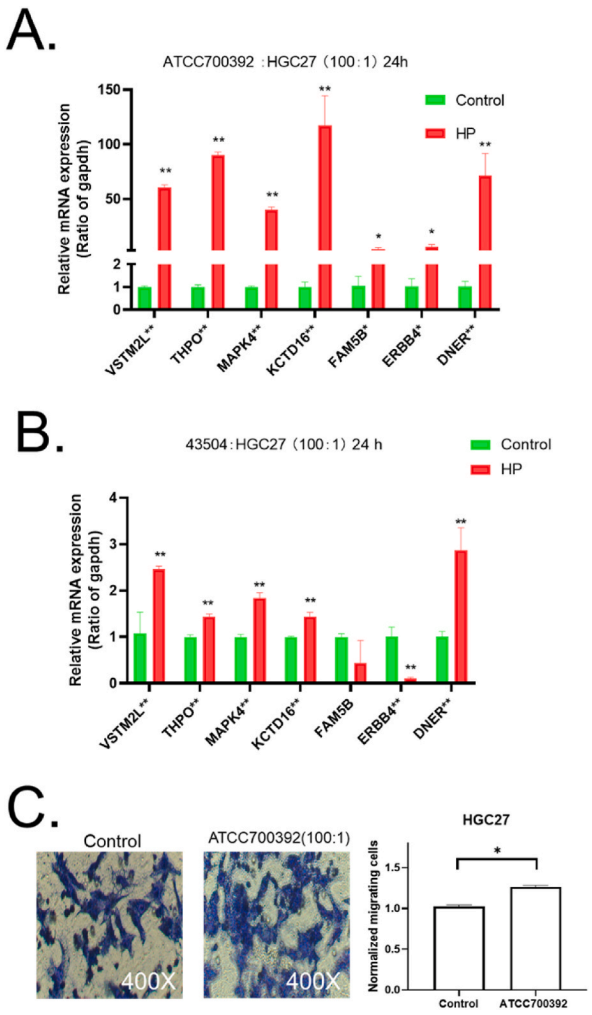


Fig. 7. Validation of cell line specificity and additional functional assay. **A**. The mRNA expression of the seven genes in HGC27 after ATCC 700392 infection. **B**. The mRNA expression of the seven genes in HGC27 after ATCC 43504 infection. **C**. Cell migration transwell assay. The transwell migration assay was conducted using HGC27 cells seeded in a 24-well transwell plate. The bacterium ATCC 700392 was applied at an MOI of 100:1 in the upper chamber, and migration was assessed over 24 h.

that HP infection may alter the response of gastric cancer cells to chemotherapy. Finally, the positive correlation between the HP-associated genes and migration-associated genes highlights potential targets that mediate the effects of HP on gastric cancer cell migration, emphasizing their significance in tumor progression and metastasis. These findings provide valuable insights into the interplay between HP infection, tumor biology, and therapeutic strategies in gastric cancer. However, these analyses are associative studies and do not imply causal effects for these genes. Thus, they support the identification of genes that can serve as clinical biomarkers to predict cancer progression, metastasis, and chemosensitivity. Additionally, it identifies potential candidates for future research.

3.8. Expression regulation of HP-associated genes in gastric cancer

Furthermore, we analyzed copy number variation (CNV), methylation, and miRNA to explore mechanisms underlying the expression regulation of HP-associated genes in gastric cancer. The heterozygous and homozygous amplification and deletion profiles were plotted (S-Fig. 8A and B). The GISTIC score generated by GISTIC2.0 indicates the copy number of a gene: 2 indicates a deep loss, possibly a homozygous deletion; −1 indicates a shallow loss, possibly a heterozygous deletion; 0 is indicative of a diploid state; 1 suggests a low-level gain (a few extra copies, often broad), possibly a heterozygous amplification; and 2 indicates a high-level amplification (more copies, often focal), possibly a homozygous amplification. Results showed that none of the 7 HP-associated genes has expression levels that significantly correlate with the corresponding CNV (S-Fig. 8C). Although there is a relatively low number of cases, we provided CNV numbers for HP-positive patients in S-Fig. 8D. Given that these genes are pivotal in HP pathology and may mediate the effects of HP on gastric cancer cells, abnormalities in their copy number could alter their expression levels and,

consequently, HP's impact. For instance, a score of '-2′ might result in non-expression of the gene (absence of gene copies), reducing HP-mediated impact if the gene is involved in such mediation. In addition, DNA methylation of these genes in gastric cancer was also analyzed. Because there are multiple DNA methylation locations within a given transcriptional sequence, the correlation was calculated using the data from the location that had the highest DNA methylation level. Although the expression might not be regulated by CNV, the correlation of DNA methylation and mRNA expression showed that the expression of 6 of these HP-associated genes significantly correlated with the methylation level, including ERBB4, DNER, BRINP2, KCTD16, MAPK4, and THPO (VSTM2L did not have sufficient data for analysis) (S-Fig. 9A). These results indicated that the methylation of these genes regulates their expression. Further, a survival analysis of gene methylation revealed that the methylation level of these genes is correlated with disease-specific survival, with the exception of KCTD16. Additionally, the methylation level of these genes is correlated with overall survival, except for KCTD16 and DNER (S-Fig. 9B). However, the DNA methyltransferases did not show a significant difference between HP-positive and negative gastric cancer (S-Fig. 9C). The comparison in methylation level between HP infection positive and negative revealed that there were only a few significances, such as cg18991165 of KCTD16 was higher methylation in HP positive compared to HP negative. Although without significance, generally, the average methylation level of THPO is higher in HP positive compared to HP negative. For most of the methylation locations, the average methylation level of BRINP2 is lower in HP positive compared to HP negative (S-Fig. 9D). These trends might potentially infer the subtle effects of HP on the methylation of these genes. Moreover, we also constructed a miRNA-gene regulation network based on database data and gene-miRNA expression correlation (S-Fig. 10). An interactive version of the network is provided in the supplementary file "Supplementary miRNA network of HP-associated genes". This network visualizes the potential interactions between genes and miRNAs, providing a comprehensive profile of the potential pathway by which miRNAs regulate genes. Color and node size were used to cluster the nodes, with node size proportional to degree, and edge widths determined by the absolute value of the correlation coefficient. This provides valuable information for further study into the role of miRNAs in HP-cancer regulation. To sum up, our analysis of copy number variation (CNV), DNA methylation, and miRNA regulation shed light on the regulatory mechanisms underlying the expression of HP-associated genes in gastric cancer. While CNV did not significantly correlate with gene expression, DNA methylation emerged as a crucial factor regulating the expression of several genes. Moreover, the survival analysis highlighted the prognostic significance of DNA methylation in these genes. The miRNA-gene regulation network provided further insights into the potential pathways through which miRNAs may influence the expression of HP-associated genes. These

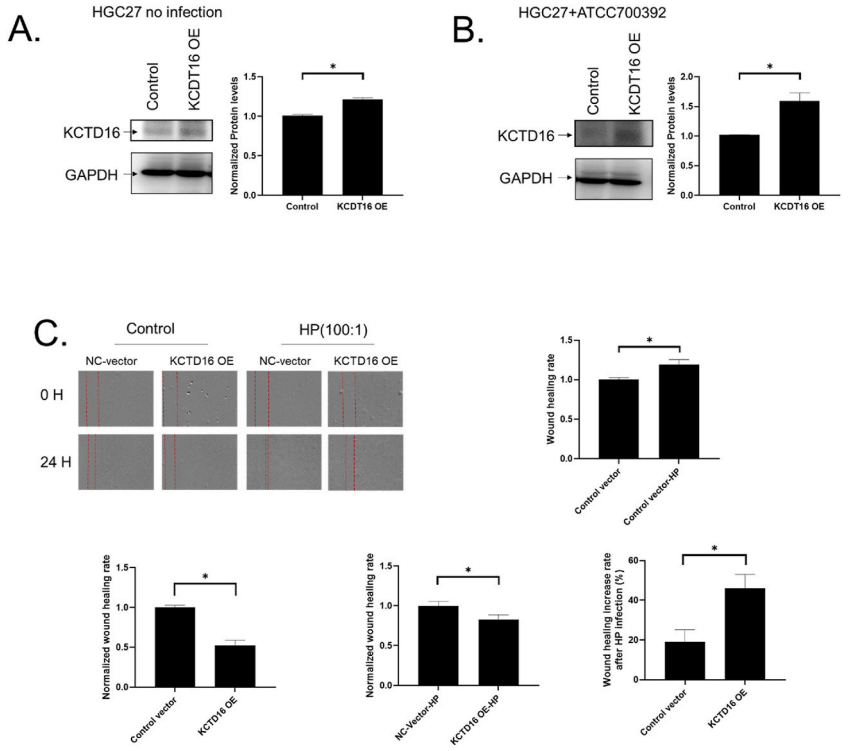


Fig. 8. Mechanistic studies underlying the HP infection of gastric cancer cells. **A**. The KCTD16 expression of HGC27 cells in non-HP-infecting gastric cancer cells with KCTD16 overexpression. **B**. The KCTD16 expression of HGC27 cells in HP-infecting gastric cancer cells with KCTD16 overexpression. HGC27 cells were transfected with control vectors or KCTD16 plasmid for 48 h, then ATCC700392 at MOI of 100: 1 was added for 24 h. **C**. The normalized wound healing rate of HGC27 was analyzed after overexpressed KCTD16 and infection with or without HP ATCC700392. HGC27 cells were transfected with either a control vector or a KCTD16 plasmid for 48 h. After transfection, the cells were scratched, washed with PBS, and initially photographed at 0 h. Subsequently, they were exposed to ATCC 700392 at an MOI of 100:1 for 24 h, after which they were photographed again at the 24-h mark.

findings enhance our understanding of the molecular mechanisms involved in HP-related gastric cancer and offer potential targets for further research and therapeutic interventions.

3.9. Validation of the impact of HP on gastric cancer cell migration using a clinical HP strain

To validate the impact of HP infection on gastric cancer cell migration, we used the clinical HP strain ICDC11101 and conducted a wound healing assay. The data indicate that infection with the clinical HP strain ICDC11101 in the SGC7901 gastric cancer cell line enhanced the migration of these cells (Fig. 6).

3.10. Validation of cell line specificity and additional functional assay

To further elucidate the impact of HP on gastric cancer, we carried out a series of experiments using the HGC27 cell line, another model of gastric cancer. In these experiments, HP strains ATCC 700392 and ATCC 43504 induced the upregulation of seven genes (VSTM2L, THPO, MAPK4, KCTD16, FAM5B, ERBB4, and DNER) at a multiplicity of infection (MOI) of 100:1 after 24 h of incubation. Notably, ATCC 700392 exhibited a significant effect on the upregulation of the KCTD16 gene, whereas ATCC 43504 showed a pronounced increase in the expression of DNER. Additionally, we performed a transwell assay, which revealed that ATCC 700392 significantly promotes the invasive capabilities of HGC27 cells (Fig. 7).

3.11. Mechanistic studies underlying the HP infection of gastric cancer cells

Based on the RT-QPCR results, KCTD16 was identified as one of the most prominently affected genes by HP infection. To delve deeper into the mechanism by which HP influences gastric cancer cells, we conducted experiments to investigate the role of KCTD16 in the context of HP's impact on gastric cancer cells. We overexpressed the KCTD16 protein in the cells and incubated them for 48 h before subjecting them to HP infection for an additional 24 h. We then performed a scratch assay to assess cell migration. Figures A and B were used to detect the overexpression of KCTD16 via Western blot analysis. Figure C presents the scratch assay results, which indicate that overexpression of KCTD16 inhibits the wound healing rate. Moreover, after HP infection, the suppressed healing rate was restored, suggesting that KCTD16 overexpression can counteract the migration promotion caused by HP infection. Further research suggests that there is a certain relationship between KCTD16 and HP. To assess cell migration, we performed a wound-healing assay. The negative control vector did not mitigate the effects of HP, validating the functionality of our transfection model. The results indicated that overexpression of KCTD16 inhibited wound healing in both the negative vector group and the HP-infected group. In contrast, wound healing in the negative control group increased by about 20 % following HP infection. While, in the KCTD16 overexpression group, wound healing increased by about 45 % after HP infection. Therefore, KCTD16 appears to counteract the migration-promoting effects of HP infection (Fig. 8).

4. Discussion

In the past few decades, the advent of high-throughput technologies and the application of various statistical tools have enabled remarkable advancements in cancer research. However, very few system-level analyses have taken into account the cancer stages, which are known to be integral to prognosis and therapy. To elucidate the dynamics of the network structure across normal and four tumor-stage phenotypes, a previous study [54] analyzed 276 samples of primary tumor tissues. This analysis revealed a drastic difference between the structure of the normal network and that of the tumor network. The analysis of connectivity dynamics showed that hub genes present in the normal network but not in the tumor networks played an important role in tumorigenesis, while those unique to a tumor network were enriched in specific biological terms. Moreover, the previous study [54] discovered three clusters of genes that had specific dynamic features across the five phenotypes and were enriched in stage-specific biological terms. Integrating the results from the expression analysis and the connectivity analysis indicated that the stages of tumor should be taken into account and a system-level analysis serves as a complement and refinement of the traditional expression analysis.

This study applied bioinformatic analysis based on open databases. TCGA aims to help better understand the mechanism underlying cancer development [28]. TCGA STAD Ranse cohort provided gene expression with HP infection data. One of the limitations is that the differential expression gene analysis was subjected to a relatively small case number in the HP-positive gastric cancer group. We need to emphasize that the HP-positive data were assessed using the Path-Seq method and in this specific study, only 11 % of samples were assessed, which was a bit lower than it should be. The Path-Seq method is just one of the approaches for assessing HP status and may lack sensitivity and/or specificity, hence, our results are just a starting point and need to be confirmed using better methods for HP detection in future studies. Nevertheless, in our study, although there might be systemic errors during the sample collection and data analysis, our study performed experiments to validate the identified target genes. As the complexity of the effects of HP on gastric cancer, a "dimensional reduction" process for the altered genes is required. Analyzing the subtypes based on the HP up-regulation gene set is a practicable and low-cost strategy to investigate the comprehensive impacts of HP on gastric cancer. Therefore, this study provided an overview of the impacts of HP on gastric cancer.

This study aimed to investigate the effect of HP infection on gastric cancer by analyzing gene expression data from gastric cancer tissues with or without HP infection. The GO/KEGG pathway enrichment analysis revealed that the 1405 upregulated genes in HP-positive samples were enriched in various terms related to cancer migration, proliferation, drug metabolism, and immunity. On the other hand, the GSEA results identified only two pathways significantly associated with the HP-upregulated genes: "Reactome

keratinization" and "Reactome formation of the cornified envelope." This suggests that HP might influence cell adhesion by regulating keratinization in gastric cancer, potentially impacting cell migration. A previous study suggested that HP affects cancer migration by suppressing the activation of the JAK1/STAT3 signaling pathway [55]. Additionally, the findings suggest that a gene called FTO contributes to the development of Hp-induced gastric cancer, playing a role in promoting the proliferation and migration of gastric cancer cells [56]. However, interpreting enrichment results in the context of heterogeneous diseases like gastric cancer can be particularly challenging due to the complex nature of the disease. Gastric cancer exhibits significant biological heterogeneity, with variations in genetic mutations, molecular subtypes, and pathological features across different patients. This diversity can complicate enrichment analyses because pathways or genes significant in one subtype might be irrelevant in others, making it difficult to draw generalized conclusions. Moreover, achieving a large enough sample size to represent all subtypes equally is often challenging, which affects statistical power and the reliability of the findings. Additionally, multiple testing and correction methods can sometimes be overly conservative or not fully appropriate for such diverse datasets, leading to potential misinterpretation of the results. These factors necessitate careful consideration and sophisticated analytical strategies to accurately interpret enrichment results in gastric cancer research.

HP infection appears to have a significant impact on gene expression in gastric cancer tissues, with 1405 genes being significantly upregulated in HP-positive samples compared to HP-negative samples. This suggests that HP infection may play a role in modulating gene expression patterns in gastric cancer. The enrichment analysis indicates that the upregulated genes in HP-positive samples are involved in multiple processes associated with cancer progression, including cell adhesion, growth factor activity, drug metabolism, and immune response. These findings suggest that HP infection may contribute to the aggressive behavior of gastric cancer by promoting cell migration, and proliferation, and altering immune processes. Altered immune processes play a crucial role in the progression and treatment of gastric cancer. In many patients, the ability of the immune system to recognize and eliminate cancer cells is compromised, which can facilitate tumor growth and metastasis. For instance, the tumor microenvironment often features immunosuppressive elements such as regulatory T cells and myeloid-derived suppressor cells, which inhibit the activity of effector immune cells capable of attacking the tumor. This suppression not only aids in the cancer evasion of immune surveillance but also creates barriers to conventional treatments like chemotherapy. However, this understanding of immune system interactions has spurred the development of immunotherapy strategies targeting these mechanisms. Treatments such as checkpoint inhibitors, which rejuvenate the immune system to fight cancer by blocking inhibitory signals to T cells, have shown promise in improving outcomes for gastric cancer patients. Thus, the manipulation of immune processes, once thoroughly understood, offers a powerful avenue for enhancing the efficacy of treatments and potentially altering the course of the disease. In this study, we find that genes related to HP are correlated with both immune cell activity and drug sensitivity, presenting promising avenues for future research in gastric cancer treatment. Notably, most of these genes show a positive correlation with the IC50 values of various drugs, suggesting that HP may decrease the sensitivity of gastric cancer cells to these treatments. Additionally, HP-related genes are also positively correlated with the presence of certain immune cells, such as natural killer (NK) cells, indicating that HP infection might enhance NK cell infiltration into the tumor microenvironment. These findings underscore the complex interplay between HP infection, immune response, and therapeutic resistance in gastric cancer.

The GSEA results highlight the potential involvement of keratinization in the mechanism of HP-induced effects on gastric cancer. Keratinization is known to play a role in maintaining mechanical stability in epithelial tissues. Therefore, it is plausible that HP infection may affect cell adhesion and migration by influencing keratinization processes in gastric cancer. The transmembrane protein is the “sensor” of cells toward the microenvironment where HP exerts impacts on cells. In vitro studies have reported HP presenting microenvironment-induced epithelial-mesenchymal transition (EMT) of gastric cancer cells [57]. The EMT decreases cell-cell adhesion and facilitates cell migration and invasion [58]. Hence, these studies were in line with our enrichment results that HP might modulate cell-cell adhesion and might affect cell migration and invasion. In addition, the EMT is essential for the initiation of metastasis in cancer progression [58], which might account for the potentially worse prognosis of HP in gastric cancer. In the cell model, we failed to observe the impact of HP infection on cell viability and LDH release, but the infection enhanced the mobility of cancer cells in the wound healing assay. These migration results further supported the fact that HP might influence cancer cell migration.

Clinical data suggested that HP infection results in higher mortality in gastric cancer [59] and the eradication of HP reduced the death of gastric cancer patients [60]. Different tests for HP have been developed [61]. Our analysis revealed that some up-regulated genes were associated with worse survival in gastric cancer patients and they might mediate the effects of HP on gastric cancer patient survival. The subtypes based on these genes were not associated with other clinical characteristics, indicating that HP infections are independent of gender and ethnicity. However, although 82 HP-upregulated genes were identified to be survival-associated, the subtypes were not associated with the TNM staging. We suggested that the effect of HP might be not as strong as to alter the tumor size or the metastasis of gastric cancer. However, the limited number of pathways identified by GSEA compared to the GO/KEGG pathway enrichment analysis suggests that HP infection might have a more nuanced and specific effect on certain biological processes. Further studies are needed to explore these pathways and validate their functional relevance in gastric cancer with HP infection. The provided supplementary table for enrichment analysis can serve as a valuable resource for researchers interested in investigating the specific genes and pathways involved in HP-associated gastric cancer.

Although HP might not be the major risk factor for gastric cancer, our analysis revealed that it can potentially affect the therapy of gastric cancer. On one hand, the subtypes showed differences in TP53 mutation and stemness, indicating that therapy targeting the TP53 pathway or stemness-related pathways can be affected by HP. Other potential therapeutic targets such as M6A and ferroptosis pathways were also found to be potentially affected by HP. The difference in the sensitivity of Mitomycin is an example demonstrating the potential effects of HP on drug treatment. The presence of HP infection in gastric cancer patients significantly impacts drug sensitivity, with profound clinical implications. HP can modify the gastric microenvironment, potentially affecting chemotherapy

efficacy. Changes in gastric acidity and alterations in drug metabolism within the stomach can influence the bioavailability and effectiveness of chemotherapy agents. This variation may necessitate adjustments in dosing or drug selection to optimize treatment efficacy. Additionally, the role of HP in promoting inflammation and possibly affecting DNA repair mechanisms can lead to differential responses to treatment, further complicating standard therapeutic approaches. Consequently, understanding the HP status of gastric cancer patients could be crucial for tailoring more effective and personalized treatment strategies. On the other hand, our study suggested that HP infection might affect immune therapy via the seven HP-associated genes identified. A previous study suggested that HP infection might impact the effectiveness of PD-1/PD-L1 blockade therapy [62]. Yet, further validation is required in the future.

We bioinformatically identified and experimentally validated seven HP-associated genes, including ERBB4, DNER, BRINP2, KCTD16, MAPK4, THPO, and VSTM2L. Most of these genes have been less studied in gastric cancer. These seven genes have not been studied in other infection models. ERBB4 is a member of the ERBB family of receptor tyrosine kinases, which includes the epidermal growth factor receptor, ERBB2, and ERBB3 [63,64]. ERBB4 has been reported to improve gastric cancer cell growth [65]. Delta/Notch-like epidermal growth factor-related receptor, short as DNER, acts as a prognostic biomarker and may be a potential therapeutic target of GC patients and inhibition of DNER diminishes gastric cancer growth and metastasis [66,67]. BMP/retinoic acid inducible neural specific 2, also named BRINP2, is involved in the cell cycle and plays a positive regulation of neuron differentiation [68]. According to the Alliance of Genome Resources [69], KCTD16 encodes a potassium channel tetramerization domain-containing protein which was involved in the G protein-coupled receptor signaling pathway. MAPK4 encodes the mitogen-activated protein kinase 4. This gene has been reported to play roles in multiple cancer types, such as prostate cancer [70], breast cancer [71,72], cervical cancer [73], ovarian cancer [74], etc. Thrombopoietin, called THPO, is a humoral growth factor that is necessary for megakaryocyte proliferation and maturation, as well as for thrombopoiesis, which was reported as a prognostic biomarker for gastric cancer [75]. VSTM2L, whose full name is V-set and transmembrane domain containing 2 like, is involved in the negative regulation of neuron apoptotic process [76]. Recent research revealed that VSTM2L is a new player involved in anoikis resistance and is a promising biomarker for ovarian cancer metastasis and prognosis [77].

Our analysis revealed that the regulation analysis revealed that their expression is associated with their methylation level. In a previous paper, the differential m6A methylation of DNER in human gingival tissue was revealed by m6A microarray and transcriptomic analysis, and it was shown to be different in periodontitis compared to periodontal health [78]. Methylation of the MAPK4 gene promoter is an innovative epigenetic biomarker for predicting recurrence risk in patients with thymic epithelial tumors [79]. VSTM2L has been previously reported to be associated with the CpG island methylator phenotype [80]. It is unclear whether the expression of these genes is upstream or downstream of methylation, or both. In the clustering analysis, we conducted the M6A methylation pathway gene analysis and found some differences. However, the analysis to compare the methylation of the methylation of these genes in HP positive and negative revealed that, although there are some trends, it is hard to tell if they are impacted by HP. We constructed a novel machine learning prognostic model for clinical applications. Interestingly, these genes potentially affect the immune microenvironment and drug sensitivity of gastric cancer. We proposed that these genes mediated the impact of HP on gastric cancer.

A previous study also suggested that THPO facilitated gastric cancer cell growth, invasion, and migration [75]. As THPO was up-regulated after HP infection, we believe it accounts for the increase in cell migration. Surprisingly, the M stage was not significantly associated with any of these HP-associated genes, we suggested that the low number of HP cases in the cohorts might limit the reliability of the association analysis. Moreover, we also identified potential targets for the effect of THPO as well as the other HP-associated genes on cell migration of gastric cancer. Nevertheless, many of our analyses were mostly based on computational algorithms and correlation analysis, thus, the causation between the gene expression and functions required experimental validation. In addition, HP is a bacterium that has adapted unique mechanisms to survive and persist in the acidic environment of the human stomach. It evades the stomach's acid by neutralizing its immediate surroundings and attaching to the stomach lining, while also interfering with local immune responses [81]. HP has been reported to be associated with cancer immunity [82], which is in line with our results that these key genes were correlated with immune cell infiltration levels in gastric cancer. Data also revealed the possibility of HP impacting cancer immune therapy [83], which requires more studies in the future. Previous studies also suggest that HP is associated with alcohol consumption and cigarette smoking [84].

TCGA has been instrumental in advancing our understanding of the molecular underpinnings of cancer through its comprehensive genomic and molecular characterizations across numerous cancer types. It has been widely used for cancer studies [85–96]. However, TCGA does have limitations. One significant issue is the lack of longitudinal data, as most samples in TCGA represent a single snapshot in time, typically collected from primary, untreated tumors. This limits insights into tumor evolution, the impacts of treatment, and mechanisms of resistance or recurrence. Furthermore, the diversity of patient demographics in TCGA may not fully represent global populations, potentially affecting the generalizability of findings across different ethnicities and geographic regions. Another limitation is the variability in the quality and types of data collected, which can affect the reproducibility and robustness of certain analyses. Finally, while TCGA provides a wealth of genomic data, it often lacks detailed environmental, lifestyle, and clinical intervention data, which are crucial for understanding the multifactorial nature of cancer etiology and treatment responses.

In conclusion, HP infection promotes gastric cancer migration by regulating the expression of key genes associated with migration, immunity, and drug sensitivity. The exact mechanisms by which HP affects these gene expressions and consequently the biological behavior of gastric cancer remain incompletely understood. Further research is crucial to delineate these interactions, which could unveil new therapeutic targets or lead to more personalized treatment protocols. Understanding the impact of HP on gene expression could also help in predicting treatment responses and tailoring interventions that mitigate its effects on cancer progression. Moreover, as research advances, it may reveal whether eradication of HP before cancer development could alter the course of the disease or improve the efficacy of existing therapies, providing a dual strategy of treatment and prevention in gastric cancer management. These

insights could be pivotal in shaping future research agendas and clinical strategies, potentially improving outcomes for patients with this complex disease.

Ethics statement

Approval of the research protocol by an Institutional Reviewer Board.
N/A.

Informed consent

N/A.

Registry and the registration no. of the study/trial

N/A.

Animal studies

N/A.

CRediT authorship contribution statement

Ling Ou: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Hengrui Liu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation, Conceptualization. **Chang Peng:** Writing – review & editing, Validation, Formal analysis, Data curation. **Yuanjing Zou:** Writing – review & editing, Validation, Formal analysis, Data curation. **Junwei Jia:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Hui Li:** Writing – review & editing, Validation, Methodology, Formal analysis. **Zhong Feng:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Guimin Zhang:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Meicun Yao:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Meicun Yao reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 81973552). The authors thank the support of Weifen Chen, Zongxiong Liu, Bryan Liu, and Yaqi Yang.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e37046.

References

[1] R.L. Siegel, K.D. Miller, H.E. Fuchs, A. Jemal, Cancer statistics, CA: a cancer journal for clinicians 72 (2022) 7–33, 2022.
[2] C. Xia, X. Dong, H. Li, M. Cao, D. Sun, S. He, F. Yang, X. Yan, S. Zhang, N. Li, Cancer statistics in China and United States, 2022: profiles, trends, and determinants, Chin. Med. J. 135 (2022) 584–590.
[3] A.N. Giaquinto, K.D. Miller, K.Y. Tossas, R.A. Winn, A. Jemal, R.L. Siegel, Cancer statistics for African American/black people 2022, CA: a cancer journal for clinicians 72 (2022) 202–229.
[4] Kang M.J., Jung K.W., Bang S.H., Choi S.H., Park E.H., Yun E.H., Kim H.J., Kong H.J., Im J.S., Seo H.G., Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2020,Cancer Res Treat 55, 2023, 385–399.
[5] Canadian Cancer Statistics: a 2022 special report on cancer prevalence, Health Promot Chronic Dis Prev Can 43 (2023) 49.
[6] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA Cancer, J Clin 73 (2023) 17–48.
[7] W. Ju, R. Zheng, S. Zhang, H. Zeng, K. Sun, S. Wang, R. Chen, L. Li, W. Wei, J. He, Cancer statistics in Chinese older people, 2022: current burden, time trends, and comparisons with the US, Japan, and the Republic of Korea, Sci. China Life Sci (2023) 1079–1091.

[8] D. Xiang, S. Hu, T. Mai, X. Zhang, L. Zhang, S. Wang, K. Jin, J. Huang, Worldwide cancer statistics of adults over 75 years old in 2019: a systematic analysis of the global burden of disease study 2019, BMC Publ. Health 22 (2022) 1979.
[9] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D.M. Parkin, M. Pineros, A. Znaor, F. Bray, Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods, Int. J. Cancer 144 (2019) 1941–1953.
[10] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J Clin 71 (2021) 209–249.
[11] F.M. Johnston, M. Beckman, Updates on management of gastric cancer, Curr. Oncol. Rep. 21 (2019) 67.
[12] Z. Song, Y. Wu, J. Yang, D. Yang, X. Fang, Progress in the treatment of advanced gastric cancer, Tumour biology 39 (2017) 1010428317714626.
[13] S.E. Crowe, Helicobacter pylori infection, N. Engl. J. Med. 380 (2019) 1158–1165.
[14] M.T.S. Al-Ouqaili, R.A. Hussein, Y.H. Majeed, F. Al-Marzooq, Study of vacuolating cytotoxin A (vacA) genotypes of ulcerogenic and non-ulcerogenic strains of Helicobacter pylori and its association with gastric disease, Saudi J. Biol. Sci. 30 (2023) 103867.
[15] R. Vogelmann, M.R. Amieva, The role of bacterial pathogens in cancer, Curr. Opin. Microbiol. 10 (2007) 76–81.
[16] Z.W. Aziz, S.H. Saleem, H.A. Al-Nuaimy, Helicobacter pylori in gastric biopsy: a histochemical and immunohistochemical assessment, Ann. Coll. Med. Mosul. 41 (2020) 139–147.
[17] M. Plummer, S. Franceschi, J. Vignat, D. Forman, C. de Martel, Global burden of gastric cancer attributable to Helicobacter pylori, Int. J. Cancer 136 (2015) 487–490.
[18] C. de Martel, D. Georges, F. Bray, J. Ferlay, G.M. Clifford, Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis, Lancet Glob Health 8 (2020) e180–e190.
[19] Y.S. Cho, S.Y. Nam, H.S. Moon, T.H. Kim, S.E. Kim, J.T. Jung, Helicobacter pylori eradication reduces risk for recurrence of gastric hyperplastic polyp after endoscopic resection, Korean J Intern Med 38 (2023) 167–175.
[20] J. Wang, R. Deng, S. Chen, S. Deng, Q. Hu, B. Xu, J. Li, Z. He, M. Peng, S. Lei, T. Ma, Z. Chen, H. Zhu, C. Zuo, Helicobacter pylori CagA promotes immune evasion of gastric cancer by upregulating PD-L1 level in exosomes, iScience 26 (2023) 108414.
[21] M.A. Abu-Lubad, G.F. Helaly, W.J. Haddadin, D.A.K. Jarajreh, A.A. Aqel, M.A. Al-Zeer, Loss of p53 expression in gastric epithelial cells of Helicobacter pylori-infected Jordanian patients, Int J Microbiol 2022 (2022) 7779770.
[22] S. Bacon, L. Seeneevassen, A. Fratacci, F. Rose, C. Tiffon, E. Sifre, M.M. Haykal, M.M. Moubarak, A. Ducournau, L. Bruhl, S. Claverol, C. Tokarski, A. R. Gouloumi, I.S. Pateras, T. Daubon, P. Lehours, C. Varon, O.C.B. Martin, Nrf2 downregulation contributes to epithelial-to-mesenchymal transition in Helicobacter pylori-infected cells, Cancers 14 (2022) 4316.
[23] T. Kawahara, Y. Kuwano, S. Teshima-Kondo, T. Sugiyama, T. Kawai, T. Nikawa, K. Kishi, K. Rokutan, Helicobacter pylori lipopolysaccharide from type I, but not type II strains, stimulates apoptosis of cultured gastric mucosal cells, J. Med. Invest. 48 (2001) 167–174.
[24] N. Kim, Chemoprevention of gastric cancer by Helicobacter pylori eradication and its underlying mechanism, J. Gastroenterol. Hepatol. 34 (2019) 1287–1295.
[25] L. Cao, S. Zhu, H. Lu, M. Soutto, N. Bhat, Z. Chen, D. Peng, J. Lin, J. Lu, P. Li, C. Zheng, C. Huang, W. El-Rifai, Helicobacter pylori-induced RASAL2 through activation of nuclear factor-kappaB promotes gastric tumorigenesis via beta-catenin signaling Axis, Gastroenterology 162 (2022) 1716–1731 e1717.
[26] S. Choi, N. Kim, J.H. Park, R.H. Nam, C.H. Song, H.S. Lee, Effect of Helicobacter pylori infection and its eradication on the expression of tight junction proteins in the gastric epithelium in relation to gastric carcinogenesis, Helicobacter 27 (2022) e12929.
[27] R.A. Hussein, M.T.S. Al-Ouqaili, Y.H. Majeed, Detection of clarithromycin resistance and 23SrRNA point mutations in clinical isolates of Helicobacter pylori isolates: phenotypic and molecular methods, Saudi J. Biol. Sci. 29 (2022) 513–520.
[28] K. Tomczak, P. Czerwińska, M. Wiznerowicz, The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge, Contemporary oncology (Poznan, Poland) 19 (2015) A68–A77.
[29] Comprehensive molecular characterization of gastric adenocarcinoma, Nature 513 (2014) 202–209.
[30] A.D. Kostic, A.I. Ojesina, C.S. Pedamallu, J. Jung, R.G. Verhaak, G. Getz, M. Meyerson, PathSeq: software to identify or discover microbes by deep sequencing of human tissue, Nat. Biotechnol. 29 (2011) 393–396.
[31] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (2012) 284–287.
[32] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 15545–15550.
[33] V.S. Stel, F.W. Dekker, G. Tripepi, C. Zoccali, K.J. Jager, Survival analysis II: Cox regression, Nephron Clin. Pract. 119 (2011) c255–c260.
[34] N. Grafféo, F. Castell, A. Belot, R. Giorgi, A log-rank-type test to compare net survival distributions, Biometrics 72 (2016) 760–769.
[35] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (2021) D605–d612.
[36] C.Y. Lin, C.H. Chin, H.H. Wu, S.H. Chen, C.W. Ho, M.T. Ko, Hubba: hub objects analyzer–a framework of interactome hubs identification for network biology, Nucleic Acids Res. 36 (2008) W438–W443.
[37] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.
[38] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (2010) 1572–1573.
[39] H. Lian, Y.P. Han, Y.C. Zhang, Y. Zhao, S. Yan, Q.F. Li, B.C. Wang, J.J. Wang, W. Meng, J. Yang, Q.H. Wang, W.W. Mao, J. Ma, Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma, Mol. Oncol. 13 (2019) 2227–2245.
[40] T.M. Malta, A. Sokolov, A.J. Gentles, T. Burzykowski, L. Poisson, J.N. Weinstein, B. Kamińska, J. Huelsken, L. Omberg, O. Gevaert, A. Colaprico, P. Czerwińska, S. Mazurek, L. Mishra, H. Heyn, A. Krasnitz, A.K. Godwin, A.J. Lazar, J.M. Stuart, K.A. Hoadley, P.W. Laird, H. Noushmehr, M. Wiznerowicz, Machine learning identifies stemness features associated with oncogenic dedifferentiation, Cell 173 (2018) 338–354.e315.
[41] Z. Liu, Q. Zhao, Z.-X. Zuo, S.-Q. Yuan, K. Yu, Q. Zhang, X. Zhang, H. Sheng, H.-Q. Ju, H. Cheng, F. Wang, R.-H. Xu, Z.-X. Liu, Systematic analysis of the aberrances and functional implications of ferroptosis in cancer, iScience 23 (2020) 101302.
[42] L. Yi, G. Wu, L. Guo, X. Zou, P. Huang, Comprehensive analysis of the PD-L1 and immune infiltrates of m(6)A RNA methylation regulators in head and neck squamous cell carcinoma, molecular therapy, Nucleic acids 21 (2020) 299–314.
[43] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin Epigenetics 11 (2019) 123.
[44] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.
[45] J. Fu, K. Li, W. Zhang, C. Wan, J. Zhang, P. Jiang, X.S. Liu, Large-scale public data reuse to model immunotherapy response and resistance, Genome Med. 12 (2020) 21.
[46] P. Geeleher, N.J. Cox, R.S. Huang, Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines, Genome Biol. 15 (2014) R47.
[47] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (2011) R41.
[48] A. Schlattl, S. Anders, S.M. Waszak, W. Huber, J.O. Korbel, Relating CNVs to transcriptome data at fine resolution: assessment of the effect of variant size, type, and overlap with functional regions, Genome Res. 21 (2011) 2004–2013.
[49] L. Liu, Y. Zhao, G. Fan, T. Shuai, B. Li, Y. Li, Helicobacter pylori infection enhances heparanase leading to cell proliferation via mitogen-activated protein kinase signalling in human gastric cancer cells, Mol. Med. Rep. 18 (2018) 5733–5741.

[50] R. Cristescu, J. Lee, M. Nebozhyn, K.M. Kim, J.C. Ting, S.S. Wong, J. Liu, Y.G. Yue, J. Wang, K. Yu, X.S. Ye, I.G. Do, S. Liu, L. Gong, J. Fu, J.G. Jin, M.G. Choi, T. S. Sohn, J.H. Lee, J.M. Bae, S.T. Kim, S.H. Park, I. Sohn, S.H. Jung, P. Tan, R. Chen, J. Hardwick, W.K. Kang, M. Ayers, D. Hongyue, C. Reinhard, A. Loboda, S. Kim, A. Aggarwal, Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes, Nat. Med. 21 (2015) 449–456.
[51] L. Li, X. Wang, Identification of gastric cancer subtypes based on pathway clustering, npj Precis. Oncol. 5 (2021) 46.
[52] S. Urakawa, M. Yamasaki, T. Makino, Y. Kurokawa, K. Yamamoto, K. Goto, M. Haruna, M. Hirata, A. Morimoto-Okazawa, A. Kawashima, K. Iwahori, T. Mizushima, E. Sato, M. Mori, Y. Doki, H. Wada, The impact of ICOS(+) regulatory T cells and Helicobacter pylori infection on the prognosis of patients with gastric and colorectal cancer: potential prognostic benefit of pre-operative eradication therapy, Cancer Immunol. Immunother. 70 (2021) 443–452.
[53] M. Wiese-Szadkowska, A. Helmin-Basa, A. Eljaszewicz, L. Gackowska, M. Januszewska, I. Motyl, M. Andryszczyk, J. Wieczynska, J. Michalkiewicz, Selected commensal bacteria change profiles of Helicobacter pylori-induced T cells via dendritic cell modulation, Helicobacter 24 (2019) e12614.
[54] J. Wu, X. Zhao, Z. Lin, Z. Shao, A system level analysis of gastric cancer across tumor stages with RNA-seq data, Mol. Biosyst. 11 (2015) 1925–1932.
[55] J. Woo, J.W. Lim, H. Kim, Astaxanthin inhibits integrin α5 expression by suppressing activation of JAK1/STAT3 in Helicobacter pylori-stimulated gastric epithelial cells, Mol. Med. Rep. 27 (2023) 127.
[56] S. Gui, Q. Wang, L. Bao, X. He, Z. Wang, L. Liu, L. Wu, Y. Zhao, J. Zhou, Y. Xie, Effects of Helicobacter pylori on the expression of the FTO gene and its biological role in gastric cancer, Oncol. Lett. 25 (2023) 143.
[57] J. Baj, I. Korona-Głowniak, A. Forma, A. Maani, E. Sitarz, M. Rahnama-Hezavah, E. Radzikowska, P. Portincasa, Mechanisms of the epithelial-mesenchymal transition and tumor microenvironment in Helicobacter pylori-induced gastric cancer, Cells 9 (2020) 1055.
[58] A. Dongre, R.A. Weinberg, New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer, Nat. Rev. Mol. Cell Biol. 20 (2019) 69–84.
[59] W.Q. Li, J.Y. Zhang, J.L. Ma, Z.X. Li, L. Zhang, Y. Zhang, Y. Guo, T. Zhou, J.Y. Li, L. Shen, W.D. Liu, Z.X. Han, W.J. Blot, M.H. Gail, K.F. Pan, W.C. You, Effects of Helicobacter pylori treatment and vitamin and garlic supplementation on gastric cancer incidence and mortality: follow-up of a randomized intervention trial, Bmj 366 (2019) l5016.
[60] T.H. Chiang, W.J. Chang, S.L. Chen, A.M. Yen, J.C. Fann, S.Y. Chiu, Y.R. Chen, S.L. Chuang, C.F. Shieh, C.Y. Liu, H.M. Chiu, H. Chiang, C.T. Shun, M.W. Lin, M. S. Wu, J.T. Lin, C.C. Chan, D.Y. Graham, H.H. Chen, Y.C. Lee, Mass eradication of Helicobacter pylori to reduce gastric cancer incidence and mortality: a long-term cohort study on Matsu Islands, Gut 70 (2021) 243–250.
[61] R.A. Hussein, M.T.S. Al-Ouqaili, Y.H. Majeed, Detection of Helicobacter Pylori infection by invasive and non-invasive techniques in patients with gastrointestinal diseases from Iraq: a validation study, PLoS One 16 (2021) e0256393.
[62] Y. Shi, H. Zheng, M. Wang, S. Ding, Influence of Helicobacter pylori infection on PD-1/PD-L1 blockade therapy needs more attention, Helicobacter 27 (2022) e12878.
[63] V.F.M. Segers, L. Dugaucquier, E. Feyen, H. Shakeri, G.W. De Keulenaer, The role of ErbB4 in cancer, Cell. Oncol. 43 (2020) 335–352.
[64] L.M. Lucas, V. Dwivedi, J.I. Senfeld, R.L. Cullum, C.P. Mill, J.T. Piazza, I.N. Bryant, L.J. Cook, S.T. Miller, J.H.t. Lott, C.M. Kelley, E.L. Knerr, J.A. Markham, D. P. Kaufmann, M.A. Jacobi, J. Shen, D.J. Riese 2nd, The yin and Yang of ERBB4: tumor suppressor and oncoprotein, Pharmacol. Rev. 74 (2022) 18–47.
[65] J. Xu, L. Gong, Z. Qian, G. Song, J. Liu, ERBB4 promotes the proliferation of gastric cancer cells via the PI3K/Akt signaling pathway, Oncol. Rep. 39 (2018) 2892–2898.
[66] H. Wang, L. Shen, Y. Li, J. Lv, Integrated characterisation of cancer genes identifies key molecular biomarkers in stomach adenocarcinoma, J. Clin. Pathol. 73 (2020) 579–586.
[67] H. Tao, C. Wang, Y. Zhu, C. Lu, X. Zhou, Role of Delta/Notch-like epidermal growth factor-related receptor in gastric cancer patients and cells and its clinical significance, Anti Cancer Drugs 33 (2022) 1175–1181.
[68] M. Terashima, M. Kobayashi, M. Motomiya, N. Inoue, T. Yoshida, H. Okano, N. Iwasaki, A. Minami, I. Matsuoka, Analysis of the expression and function of BRINP family genes during neuronal differentiation in mouse embryonic stem cell-derived neural stem cells, J. Neurosci. Res. 88 (2010) 1387–1393.
[69] Kishore R., Arnaboldi V., Van Slyke C.E., Chan J., Nash R.S., Urbano J.M., Dolan M.E., Engel S.R., Shimoyama M., Sternberg P.W., Genome Resources T.A.O., Automated Generation of Gene Summaries at the Alliance of Genome Resources, Database (Oxford) 2020 (2020) baaa037.
[70] T. Shen, W. Wang, W. Zhou, I. Coleman, Q. Cai, B. Dong, M.M. Ittmann, C.J. Creighton, Y. Bian, Y. Meng, D.R. Rowley, P.S. Nelson, D.D. Moore, F. Yang, MAPK4 promotes prostate cancer by concerted activation of androgen receptor and AKT, J. Clin. Invest. 131 (2021) e135465.
[71] W. Wang, D. Han, Q. Cai, T. Shen, B. Dong, M.T. Lewis, R. Wang, Y. Meng, W. Zhou, P. Yi, C.J. Creighton, D.D. Moore, F. Yang, MAPK4 promotes triple negative breast cancer growth and reduces tumor sensitivity to PI3K blockade, Nat. Commun. 13 (2022) 245.
[72] X. Zeng, S. Jiang, S. Ruan, Z. Guo, J. Guo, M. Liu, C. Ye, J. Dong, MAPK4 silencing together with a PARP1 inhibitor as a combination therapy in triple-negative breast cancer cells, Mol. Med. Rep. 24 (2021) 548.
[73] S. Tian, L. Lou, M. Tian, G. Lu, J. Tian, X. Chen, MAPK4 deletion enhances radiation effects and triggers synergistic lethality with simultaneous PARP1 inhibition in cervical cancer, Journal of experimental & clinical cancer research : CR 39 (2020) 143.
[74] S.Y. Du, X.X. Huang, N.M. Li, C.Y. Lv, C.H. Lv, M.L. Wei, Z. Gao, Y.P. Zhang, MiR-127-3p inhibits proliferation of ovarian cancer in rats through down-regulating MAPK4, Eur. Rev. Med. Pharmacol. Sci. 24 (2020) 10383–10390.
[75] C.L. Zhou, H.L. Su, H.W. Dai, Thrombopoietin is associated with a prognosis of gastric adenocarcinoma, Rev. Assoc. Med. Bras. 66 (1992) 590–595, 2020.
[76] L. Rossini, Y. Hashimoto, H. Suzuki, M. Kurita, M. Gianfriddo, C. Scali, R. Roncarati, D. Franceschini, G. Pollio, L. Trabalzini, G.C. Terstappen, M. Matsuoka, A. Caricasole, VSTM2L is a novel secreted antagonist of the neuroprotective peptide Humanin, FASEB J 25 (2011) 1983–2000.
[77] Y. Li, J. Zhang, Y. Cai, H. Liu, W. Yang, Y. Xu, M. Huang, VSTM2L contributes to anoikis resistance and acts as a novel biomarker for metastasis and clinical outcome in ovarian cancer, Biochem. Biophys. Res. Commun. 658 (2023) 107–115.
[78] Z. Wang, H. Chen, L. Peng, Y. He, J. Wei, X. Zhang, DNER and GNL2 are differentially m6A methylated in periodontitis in comparison with periodontal health revealed by m6A microarray of human gingival tissue and transcriptomic analysis, J. Periodontal. Res. (2023) 529–543.
[79] W. Guan, S. Li, Z. Zhang, H. Xiao, J. He, J. Li, X. He, J. Luo, Y. Liu, L. Lei, J. Ma, L. Chen, C. Chen, Promotor methylation status of MAPK4 is a novel epigenetic biomarker for prognosis of recurrence in patients with thymic epithelial tumors, Thorac Cancer 13 (2022) 2844–2853.
[80] Z. Zeng, D. Xie, J. Gong, Genome-wide identification of CpG island methylator phenotype related gene signature as a novel prognostic biomarker of gastric cancer, PeerJ 8 (2020) e9624.
[81] J.C. Atherton, The pathogenesis of Helicobacter pylori-induced gastro-duodenal diseases, Annu. Rev. Pathol. 1 (2006) 63–96.
[82] A. Ralser, A. Dietl, S. Jarosch, V. Engelsberger, A. Wanisch, K.P. Janssen, M. Middelhoff, M. Vieth, M. Quante, D. Haller, D.H. Busch, L. Deng, R. Mejías-Luque, M. Gerhard, Helicobacter pylori promotes colorectal carcinogenesis by deregulating intestinal immunity and inducing a mucus-degrading microbiota signature, Gut 72 (2023) 1258–1270.
[83] M. Noori, F. Fayyaz, N. Rezaei, Impact of Helicobacter pylori infection on the efficacy of immune checkpoint inhibitors for cancer treatment: a meta-analysis, Immunotherapy 15 (2023) 657–667.
[84] R.A. Hussein, M.T. Al-Ouqaili, Y.H. Majeed, Association between alcohol consumption, cigarette smoking, and Helicobacter pylori infection in Iraqi patients submitted to gastrointestinal endoscopy, Journal of Emergency Medicine, Trauma & Acute Care 2022 (2022) 12.
[85] H. Liu, J. Weng, A pan-cancer bioinformatic analysis of RAD51 regarding the values for diagnosis, prognosis, and therapeutic prediction, Front. Oncol. 12 (2022) 858756.
[86] H. Liu, J. Weng, A comprehensive bioinformatic analysis of cyclin-dependent kinase 2 (CDK2) in glioma, Gene (2022) 146325.
[87] H. Liu, T. Tang, Pan-cancer genetic analysis of cuproptosis and copper metabolism-related gene set, Front. Oncol. 12 (2022) 952290.
[88] H. Liu, Y. Li, Potential roles of Cornichon Family AMPA Receptor Auxiliary Protein 4 (CNIH4) in head and neck squamous cell carcinoma. Cancer Biomarkers 35, 2022, pp. 439–450.
[89] H. Liu, J.P. Dilger, J. Lin, A pan-cancer-bioinformatic-based literature review of TRPM7 in cancers, Pharmacol. Ther. (2022) 108302.
[90] H. Liu, Pan-cancer profiles of the cuproptosis gene set, Am. J. Cancer Res. 12 (2022) 4074–4081.

[91] Y. Li, H. Liu, Clinical powers of aminoacyl tRNA synthetase complex interacting multifunctional protein 1 (AIMP1) for head-neck squamous cell carcinoma. Cancer Biomarkers 35, 2022, pp. 359–374.
[92] H. Liu, Li Y., Potential roles of Cornichon Family AMPA Receptor Auxiliary Protein 4 (CNIH4) in head and neck squamous cell carcinoma, Cancer Biomarkers 35 (2022) 439–450.
[93] H. Liu, T. Tang, Pan-cancer genetic analysis of disulfidptosis-related gene set, Cancer Genetics 278-279 (2023) 91–103.
[94] H. Liu, Expression and potential immune involvement of cuproptosis in kidney renal clear cell carcinoma, Cancer Genetics 274–275 (2023) 21–25.
[95] L. Hengrui, An example of toxic medicine used in Traditional Chinese Medicine for cancer treatment, J. Tradit. Chin. Med. 43 (2023) 209–210.
[96] H. Liu, T. Tang, A bioinformatic study of IGFBPs in glioma regarding their diagnostic, prognostic, and therapeutic prediction value, Am J Transl Res 15 (2023) 2140–2155.