Bioinformatics analysis for the identification of differentially expressed genes and related signaling pathways in \textit{H. pylori}-CagA transfected gastric cancer cells

Dingyu Chen*, Chao Li, Yan Zhao, Jianjiang Zhou, Qinrong Wang and Yuan Xie*

Key Laboratory of Endemic and Ethnic Diseases, Ministry of Education, Guizhou Medical University, Guiyang, China

*These authors contributed equally to this work.

ABSTRACT

\textbf{Aim.} \textit{Helicobacter pylori} cytotoxin-associated protein A (CagA) is an important virulence factor known to induce gastric cancer development. However, the cause and the underlying molecular events of CagA induction remain unclear. Here, we applied integrated bioinformatics to identify the key genes involved in the process of CagA-induced gastric epithelial cell inflammation and can eration to comprehend the potential molecular mechanisms involved.

\textbf{Materials and Methods.} AGS cells were transected with pcDNA3.1 and pcDNA3.1::CagA for 24 h. The transfected cells were subjected to transcriptome sequencing to obtain the expressed genes. Differentially expressed genes (DEG) with adjusted $P$ value < 0.05, $|\log FC| > 2$ were screened, and the R package was applied for gene ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The differential gene protein–protein interaction (PPI) network was constructed using the STRING Cytoscape application, which conducted visual analysis to create the key function networks and identify the key genes. Next, the Kaplan–Meier plotter survival analysis tool was employed to analyze the survival of the key genes derived from the PPI network. Further analysis of the key gene expressions in gastric cancer and normal tissues were performed based on The Cancer Genome Atlas (TCGA) database and RT-qPCR verification.

\textbf{Results.} After transfection of AGS cells, the cell morphology changes in a hummingbird shape and causes the level of CagA phosphorylation to increase. Transcriptomics identified 6882 DEG, of which 4052 were upregulated and 2830 were downregulated, among which q-value < 0.05, FC > 2, and FC under the condition of $\leq 2$. Accordingly, 1062 DEG were screened, of which 594 were upregulated and 468 were downregulated. The DEG participated in a total of 151 biological processes, 56 cell components, and 40 molecular functions. The KEGG pathway analysis revealed that the DEG were involved in 21 pathways. The PPI network analysis revealed three highly interconnected clusters. In addition, 30 DEG with the highest degree were analyzed in the TCGA database. As a result, 12 DEG were found to be highly expressed in gastric cancer, while seven DEG were related to the poor prognosis of gastric cancer. RT-qPCR verification results...
showed that Helicobacter pylori CagA caused up-regulation of BPTF, caspase3, CDH1, CTNNB1, and POLR2A expression.

**Conclusion.** The current comprehensive analysis provides new insights for exploring the effect of CagA in human gastric cancer, which could help us understand the molecular mechanism underlying the occurrence and development of gastric cancer caused by Helicobacter pylori.

**Subjects** Biochemistry, Bioinformatics, Cell Biology, Microbiology, Molecular Biology

**Keywords** Helicobacter pylori, CagA, Gastric cancer, Transcriptomics, Bioinformatics analysis

**INTRODUCTION**

Gastric cancer is the fifth-most common malignant tumor and the third-most common cause of death worldwide (Park et al., 2018). The development of gastric cancer involves multiple aspects, including the host factors, environmental factors, and Helicobacter pylori infection. Among these, *H. Pylori* infection is known to cause chronic inflammation of the gastric mucosa, which in turn causes atrophic gastritis, gastric cancer, and various other gastrointestinal diseases. Reportedly, *H. Pylori* is a very common infective agent of the stomach across the world, and this infection has been closely related to the development of gastric cancer and its malignant precursors (Valenzuela et al., 2015). Presently, the mechanism of *H. Pylori*-induced damage to the gastric mucosa is not well understood. The assumed possible mechanisms include the damages caused by *H. pylori* colonization and toxin production, the host’s immune response, and the abnormal gastric acid secretion. Accumulating body of work supports that specific virulence factors in *H. Pylori* have a strong correlation with gastric cancer (Chmiela et al., 2017), including CagA and vacuolar cytotoxin A (VacA). CagA is a 128-145-kilodalton (kDa) protein that is composed of a structured N-terminal region and an intrinsically disordered/unstructured C-terminal tail. Variations in the molecular weight of CagA are due to the structural polymorphisms in its C-terminal region, which exist in distinct strains of *H. pylori*. Once injected into the host gastric epithelial cells, CagA is localized to the inner leaflet of the plasma membrane. CagA is encoded by the *H. pylori* cag pathogenic island and injected into gastric epithelial cells via T4SS, where it undergoes tyrosine phosphorylation at the Glu-Pro-Ile-Tyr-Ala (EPIYA) motif in its C-terminal region and then acts as a carcinogenic scaffold protein which physically interacts with different host-signaling proteins. From the sequence flanking the EPIYA motif, four distinct EPIYA segments have been identified in the CagA protein: EPIYA-A, EPIYA-B, EPIYA-C, and EPIYA-D. The EPIYA-C segment is present in variable numbers of copies among distinct Western CagA variants, typically represented in tandem between one to three times. The EPIYA-repeat region of CagA found in East Asian countries also possesses EPIYA-A and EPIYA-B segments but, instead of the tandem EPIYA-C segment, contains a distinct EPIYA-containing segment termed EPIYA-D (47 amino acids), and the CagA protein is referred to as East Asian CagA or ABD-type CagA. Due to the variation of the sequence flanking the tyrosine (Y) residue, the distinct EPIYA segments are tyrosine-phosphorylated selectively by different kinases. EPIYA-A and EPIYA-C or EPIYA-B...
and EPIYA-D are preferably phosphorylated in combination in Western CagA and East Asian CagA, respectively. Therefore, there may be a stepwise event in which EPIYA-C or EPIYA-D is phosphorylated by SFKs at the start of an infection followed by phosphorylation of EPIYA-A or EPIYA-B by c-Abl at a subsequent time. Deregulation of SHP2, the pro-oncogenic PTPase involved in the regulation of cell growth, motility, and morphology. East Asian CagA exhibits a stronger ability to bind/deregulate SHP2 and a greater capability to induce SHP2-dependent morphological changes in gastric epithelial cells than Western CagA. Collectively, the findings reveal that the East Asian CagA-specific EPIpYA-D motif is qualitatively very different from the Western CagA-specific EPIpYA-C motif in terms of the biological activity required for deregulation of the SHP2 oncoprotein, which may causatively account for the higher incidence of gastric cancers in East Asian countries than in Western countries (Takahashi-Kanemitsu, Knight & Hatakeyama, 2020). CagA affects the proliferation and apoptosis of cells through various regulation and signaling pathways, ultimately promoting gastric mucosal carcinogenesis (Takahashi-Kanemitsu, Knight & Hatakeyama, 2020).

Past studies have demonstrated that the non-physiological scaffolding of CagA in cells promote the malignant transformation of normal cells by conferring onto them cancer markers with multiple phenotypes. In chronic inflammation, CagA’s in vivo carcinogenic activity is further enhanced. Because H. pylori infection triggers a pro-inflammatory response in the host cell, the resultant feed-forward stimulation loop enhances the carcinogenic effects of CagA and cause inflammation in the gastric mucosa, where CagA is injected. Considering the need for clarification on these aspects, we attempted to explore the molecular mechanisms of CagA-induced gastric epithelial cells to seek effective molecular targets in order to provide a basis for early clinical diagnosis, prevention, and treatment of gastric cancer (Cover, 2016). Then, we applied integrated bioinformatics to identify the key genes involved in the process of CagA-induced gastric epithelial cell inflammation and canceration to comprehend the potential molecular mechanisms involved.

**MATERIALS AND METHODS**

**pcDNA3.1::CagA plasmid vector transfection of AGS cells**

The CagA plasmid pcDNA3.1(+)cagA and the empty vector pcDNA3.1(+)EGFP were purchased from Nobel Biotech (Shanghai, China). AGS cells were obtained from the ATCC. AGS cells were incubated in RPMI-1640 medium (Gibco, Grand Island, NY, USA) supplemented with 10% heat-inactivated fetal bovine serum (Gibco), 100 U/ml of penicillin, and 100 g/ml of streptomycin at 37 °C in a humidified incubator (NSE, Brunswick, NJ, USA) containing 5% CO2. AGS cells were seeded in 6-well plates respectively at a density of 5 × 10^6 cells/well, grown to whose confluence reached at 60–70%, then the cells were transfected with 3 µg plasmid and 5 µl Lipofectamine 2000 (Invitrogen, USA) in 125 µl Opti-MEM™ medium (Gibco, USA) followed by the addition of 1,875 µl Opti-MEM™ medium according to the manufacturer. After 24 h the transfection efficiency was evaluated by observation under a fluorescence microscope, and the relevant cell samples were collected. The CagA expression was verified by western blotting.
Differential gene collection and screening
The vectors pcDNA3.1::CagA and pcDNA3.1 were transfected into AGS cells respectively. After 24 h, cell samples were collected and sent to NOVOgene (Beijing, China) transcriptome for sequencing to obtain the differentially expressed genes between the two. By adjusting $P < 0.05$, $|\logFC| > 2$, the genes with significant differences are listed.

Analysis of gene ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of differentially expressed genes (DEG)
The GO enrichment analysis is a commonly used method for large-scale functional analysis. The gene functions can be classified as biological processes (BP), molecular functions (MF), and cellular components (CC). KEGG is a widely used database that stores the information on a large number of related genomes, biological pathways, diseases, chemicals, and drugs. We applied the R package with data package, visualization, and integrated discovery to perform GO enrichment analysis and KEGG pathway analysis on the DEG in this study, with $P < 0.05$ considered as statistically significant, and passed the “ggplot2” of R package to visually generate histograms and lists (Tang et al., 2020).

PPI network construction and key cluster identification
The PPI networks of DEG were constructed using the STRING database (Gu et al., 2019) (http://string-db.org), which is a software application that is commonly used to identify interactions, assess potential PPI relationships, and identify previously determined differences. Briefly, the DEG were mapped into the STRING database. The PPI networks were then visualized by the Cytoscape software (Gu et al., 2019) (https://cytoscape.org/). The software predicts the network, with each node as a gene. The network visualization helps identify the interactions and pathway relationships among the proteins encoded by DEG in gastric cancer. The corresponding protein in the central node could be a core protein or a key candidate gene with important physiological regulatory functions. According to the Cytoscape visualization network of molecular interactions, the Molecular Complex Detection (MCODE) plug-in is used to identify densely interconnected clusters, based on the following selection criteria: degree $\geq 2$, node score $\geq 0.2$, K-core $\geq 2$, max depth $= 100$ (Li, 2019).

Selection of key genes and their expression analysis in Hp infection status
The top 30 central genes with the most connections in the PPI network are defined as key genes. The differential expression of Hp infection and uninfected tissues was analyzed with reference to the TCGA database ($P < 0.05$ is considered to indicate statistical significance).

Survival analysis of the key genes
The Kaplan–Meier plotter database (Ma, Zhou & Zheng, 2020) (http://www.kmplot.com) is an online tool that can be used to evaluate 54,675 genes under the conditions of 10,461 cancer samples. We used this database to perform a survival analysis ($P < 0.05$ was considered to indicate statistical significance). Functional enrichment analysis of the genes that were highly expressed in gastric cancer was performed.
**Identification of RT-qPCR**

The cagA gene knockout mutant strain Hp/cagA cm was constructed by Sangon Biotech (Shanghai, China). AGS cells were seeded in 6-well plates respectively at a density of $5 \times 10^6$ cells/well, grown to whose confluence reached at 60–70%, then the cells were infected with Hp/ΔcagA::Cm and Wild type Hp/cagA + with a multiplicity of infection (MOI) of 30, respectively. 24 h later, these cells were harvested to investigate BPTF, CASP3, CDH1, CTNNB1 and POLR2A mRNA levels by RT-qPCR.

**Western blot**

The total protein was extracted according to the instructions of the lysate kit. After quantification by the BCA protein quantification kit, SDS-PAGE electrophoresis for 2 h, membrane transfer for 2 h, 1XTBST (0.05% Tween20) solution containing skimmed milk powder was blocked at room temperature for 2 h, and CagA primary antibody was added (1: 1,000), p-CagA (1: 1,000), GAPDH (1: 5,000) Incubate overnight at 4 °C, wash the membrane with 1×TBST (0.05% Tween20) solution 3 times, 10 min/time, add two Incubate at room temperature for 2 h with anti (1:10,000), wash the membrane with 1×TBST as above, add a chemiluminescence reagent for color development, and expose and image with a chemiluminescence imager.

**RESULTS**

**Transfection of pcDNA3.1::CagA plasmid into AGS cells and verification by western blotting**

After 24 h of transfection of pcDNA3.1::CagA, The efficiency of fluorescent transfection is estimated to be >70% (Fig. 1A). The morphology of the cells was observed under the microscope. It was found that compared with the control group (Control) and the empty vector group (pcDNA3.1), the morphology of the cells in the cagA transfection group (cagA) changed significantly. The shape of the cell changed from obtuse to long fusiform, spindle-shaped, irregular, and the polarity of the cell disappeared, showing a ‘hummingbird change’ (Fig. 1B). Western blot verification showed that CagA and p-CagA protein was successfully expressed in pcDNA3.1::cagA transfected group (Fig. 1C).

**Screening DEG**

The application of transcriptomics identified 6,882 common genes, of which 4052 were upregulated and 2830 were downregulated (Fig. 2). At adjusted $P < 0.05$, |logFC| >2, 1062, DEG with statistical significance were further screened to identify 594 upregulated and 468 downregulated genes (Table 1).

**GO term enrichment analysis of DEG**

Using the R package with data package, visualization, and integrated discovery, GO enrichment analysis was performed on 1062 DEG with different meanings. Our results revealed that 151 DEG participated in BP, 56 in CC, and 40 in MF. With respect to BP, the DEG were significantly enriched in the mRNA catabolic process, covalent chromatin modification, and histone modification. With respect to CC, they were mainly enriched in focal adhesion, cell-substrate adherens junction, and cell-substrate junction. With respect...
Figure 1  Detection of pcDNA3.1::cagA plasmid transfection efficiency and CagA expression in AGS cells. pcDNA3.1::cagA plasmid was transfected into AGS cells for 24 h, the transfection efficiency observed under a microscope and measure the protein levels of CagA and phosphorylated CagA by Western blotting. (A) Fluorescence showed that the transfection efficiency reached > 70%; (B) cagA plasmid transfected group shows hummingbird-like changes in cell morphology. (C) The protein levels of CagA and p-CagA were showed in pcDNA3.1::cagA transfected group. The control group represents the untreated group; pcDNA3.1 group represents the AGS cells transfected with empty vector pcDNA3.1; pcDNA3.1::cagA group represents the AGS cells transfected with pcDNA3.1::cagA.

KEGG pathway analysis of DEG
Using the R package with data package, visualization, and integrated discovery, KEGG enrichment analysis was performed on 1062 DEG with different meanings. Our results revealed that a total of 21 pathways were enriched, mainly ribosome, ubiquitin-mediated proteolysis, and cancer pathways (Fig. 4, Table 3).

Construction of PPI network and identification of key genes
STRING and Cytoscape analyses identified a total of 845 DEG participating in the PPI network, with 5,571 edges (Fig. 5A), 471 upregulation, and 374 downregulation. Through the MCODE plug-in, the first three densely interconnected clusters of the PPI network were analyzed. Cluster 1 consisted of 67 nodes and 1,098 edges. The enrichment results indicated that the genes included in Cluster 1 of the PPI were mainly enriched in the terms extracellular exosome” and “poly(A) RNA binding”. Cluster 2 was composed of 20
Figure 2  Differential expression of data between two sample sets. The red points represent upregulated genes screened on the basis of |fold change| > 2.0 and a corrected P value of < 0.05. The green points represent downregulation of the gene expression screened on the basis of |fold change| > 2.0 and a corrected P-value of < 0.05. The blue points represent genes with no significant difference. FC indicates the fold change.

Full-size DOI: 10.7717/peerj.11203/fig-2

Figure 3  GO enrichment analysis of DEGs. GO analysis categorized DEG into three functional groups: molecular function, biological processes, and cell composition.

Full-size DOI: 10.7717/peerj.11203/fig-3

nodes and 13 edges. The enrichment results indicated that the genes included in Cluster 2 were mainly enriched in the terms “nuclear-transcribed mRNA catabolic process” and “acetylation”. Cluster 3 was composed of 15 nodes and 92 edges. The enrichment results indicated that the genes included in Cluster 3 were mainly enriched in the terms
### Table 1  Filter out statistically significant DEGs.

| DEGs          | Gene names |
|---------------|------------|
| Upregulated:  | ATM AHNAK2 MYCBP2 UBE4A KIAA1109 MIEF1 LSM12 FRYL SVYN1 GSN RERE STK36 ACTN1 SMG1 KMT2A PIK3CB HAS3 ACOT7 CHTF8 KIF21B ANO1 ADCY3 PPP5K1 HUWE1 ZNF184 ADGRG1 ABR MUC1 FRYL PTTF1 KMT2D BICRA DHCR24 TNRC18 BCORL1 YLPM1 SYNE1 DSTD1 YLPM1 COL1A1 XIAP NCO2R2 BRD4 TRIM32 CTPD1 HSPA8 EIF3F PRRL2 KDM6B LMNA NUP214 LAMA5 AKHGEF1 ATXNL2 HSPG2 NCOA6 AZA22A PIGB BPTF AKT1S1 PCBP2 MCM3AP NR2F2 ATN1 MYO10 AFF1 FNDC3A ILF3 VPS13D FGD4 MYSM1 NOL8 LASL1 ANP32E FOXM1 DGKZ CDC73 CALU SUPT20H GLT8D1 SLC20A1 NEDD4 OXAX1 CTNNAL1 PCBP2 GART ALG11 RERE1 DD5X PTCH1 TYW3 ZNF703 TJP2 HMBOX1 G6orf106 NPIPB12 CLASP1 ERGIC3 PEAK1 LRP1 ARNTL CTNNB1 DDX11 SCMH1 ZNF615 DENN3 RAPPH1 KMT2D TRIM16 SMC4 SUPT5H SLC39A4 YWHAZ ZFP36L1 LRC8D MED13L PDR2A SRCAP SMC1A TRMT1 PEX5 PARV3 BIRC2 LCOR TNK2 BCL9L CTNNAL1 RARG AA1K1 TENT2 CACHD1 Z3CHAV1 UBE2V1 UB4R CEPP5 SPEN DXXDC1 DNM2 AKAP13 DLGAP4 PTEN23 MLXIP STAT1 CAP1 CAMSAAP2 ADP SRCAP KASNL1 TMEM150A 5HGF10L TRC1515 SH3D15 RPRM MLXIP NPIPB5 P2RX4 TEF PPP6C ZNF615 ENTP3D PROX1 VKORC1L1 EP400 SH3YLI DFP2 SEPT2 MUC1 RABEP2 ZNFSLX S9LA7 PGRMC2 CPEB3 TCTN3 RAD23B KDM6B CELS82 WIZ NRDC CEACAM1 YAP1 RALBP1 SEC1A1 UAP2L2 STAT5 CLTC FOXP1 FBXW1 CNOT6 KMT2C ARHGEF11 SMARCA4 TBC1D7 ZSWIM8 ABT81 POM121 VPS37B CLASP2 SRGAP1 SLC35A3 RNPS1 SLC30A6 SEPT9 CCNE2 USP20 TGRFBRAP1 RPL21 STX16 BSA3 RBM26 CREBBP HADC1 BTBD3 URGCP FOXP1 CSD1E1 ZNF28 PIGF RPSF7 IFT140 SRCAP GSN SECISBP2 LTBP3 TJP2 GPRR5 PRSS23 PLAG2A DEND4B4 CREBBP ASXL2 CABIN1 ESY2T ATPT24 MYCBP2 PRAG1 KASNL1 SORB2B ARID1A MEFD2 RPL34 RALGPS1 CEPP5 BFAET2 TACC2 ZNF10 IGFI1 HRDAC8 ACN11 ZFDN6 ZFHX3 HNRNP13 MIZ1I RHOC ZNF638 SYNO PT1K1 HNRNP3A SBBP3 ATG2A PPRC4 PLEC IGSF9 IP6K1 NFATC3 PIGN SAMD12 VOP11 AC901057.6 ASPSR1 HNRNP3A HECTD4 ATXNL2 DOCLK PLXNB2 CEPP4 NPIPB3 EGFR2 AP3B1 MUC10 ZNF720 HARS EPB41L2 SEPT7 RNFL45 MPND PLCXD1 MEX MX2 ZNF567 CACNG8 KMT2D KAT6B FASN RHRDF2 RNFLF28 SFTD2 TANC2 TANCPC2 MED12 ARID1A PIK3CB ZFPM1 GRN2D ZFHX2 FGYY LEPR NPIPB4 C6orf132 TRIOBP ANKRD11 OXSR1 FAM210A SMIM8 FAM98B STK4 AFEDN AHDC1 AL354822.1 AFAF1 PU5 TBC1D11 HNRNP2A2B1 SEPT9 NDK1 YIFP1 TSPAN14 ADAR ESR1 NPIIPS3 ATXN10 BRD3OS KMT2B DNYC2H1 TRIM7 USP3 DMTN BCLL1B COBL ZC3H12C FAM47E XB1 PCBP2 R3HC11L MRTFA ARHGEF9 ER11 RN4F CASP3 ELM4 PTPN21 RJN23 ATRX SYNE2 SCAFI BROX DSP NPIIPS3 SMARCA1I USP48 ASCC1 ZNF107 DCAF13 PDCD4 PKEC1 USP49 ANO1 ARHGA21 ME1 TAF1C PXD1 SMD4B MYBBP1A NEK11 TAB2 MAVS NEURL4 KDM6A ELMSAN1 PILRB ZNF332 TRIM66 VPS35L TXNDC1 Z3CHX3 PHLB2 MYO10 LRAT1 MATN2 PDR2B CEPP51 CEPP52 BRD3 NPIPB5 ZNF528 HSRP1 EP400 PARG CNOT9 HOXC6 ZNF561 HNRNPM TNKS AK1 CMIP ZFEN326 CTBP2 R3HCC1L MRTFA PKP4 TRRAP HELZ2 TRIM4 LITAF SLCC2A18 MUC1 UBE2D2 PPRG RUVB1L2 SLCC2A3 ATRX TMECAM14 UCKL1 TTC2A1 KIAA1211 NPIPB3 PCMI4 SMARCA4 PLXNB2 KDM4B HNRNPUL1 POM121C MEST BCL11A SMAD5 ACVR1B RANGAP1 SEC16A CEACAM1 ABR7B ADCY3 GRK6 MOCR2 RBM17 NAV2 ANKPK1 SON FBXL8 CDRT4 RRM2 PLCB1 PDCDHGB5 CRTAP CREBBP URM1 CANT1 DCAKD HIPK2 PLEC ZNF184 SMARCA4 AHNAK MEFD2 ALG11 C1QTNF3 HNRNP6 SEZL62 BCAM LIMA1 USP6NL ELF1 AFDN ARID4B PABPC4 PTPRF CCNK WDHDI UBE2K SH3D19 BCLAF3 USP16 PRDM2 HADH ADAM20 GEMIN2 GRAMFL1A POU5F2 ZNF565 PLEKHA7 TMEM80 HTATSF1 PHKB MIDN CREBL3 B2P110 KCNC3 PDCD12A TSMF BCL9L COLE2A71 BDR1 KDM6A TCTN1 KMT2A SETD1A RAB12 AP3D1 MICAL1 RHO1 RBM33 CPT2 MEIS2 ITGB1 GFBP1 COP2 KRB2 ST3GAL3 MPLKIP TMCC3 ZBTB43 NCO2R1 RAVER1 ZBED3 EPHX1 FAS KDM4C USP48 KIAA1217 NEAT5 PCMTD2 NPIPB3 SOCS4 PPIE UVSSA HMG2 MYCBP2 TAP2 RUNX1 TRIO ABCB8 LTBP3 MMS19 TAF5L USP21 KIAA1549 DIDO1 GRMDA1 PLEKHB2 FTO CLK2 MAML1 ITS1 NF4F TRIM3 SUCLA2 RPL18 HNRNPD PML CHD1 DTX2 RUNX1 PALLD TBP23C

(continued on next page)
### Table 1 (continued)

| DEGs | Gene names |
|------|-------------|
| Downregulated: | DHCR24 GRSF1 MRNIP RALBP1 HACD3 RN6F LSM7 MAT3 PDAIA6 TAF5L RPL15 HBPI IFNGR1 NEDD1 SCAND1 TCEA1 SLA2 PSMC2 CNTD1 SNX22 FAM161B ISG15 CAPN12 NACA BEST1 ACSBG1 RPS27 FAM166A CEP85 SAP18 MORF41 KMO ARFGAP1 MMACHC NOL8 EML5 TYK2 ATP5ME CPM GSTM3 EXOC3 MXD3 SLC9A6 PLEKHB2 TCEAL4 PTGDR2 NPC2 ERLIN2 RSPH9 RNF38 ENTPD1 MEMO1 ESRP1 ANKRD6 UBA52 RPL21 IAMN1 NDRG1 RPL8 ZFAND6 PSMC3 |

**Notes.**

DEGs, differentially expressed genes.

---

“The transcription and chromatin regulator”, and one node was a DEG (Figs. 5B–5D; Tables 4–6). The first 30 genes in the connectivity evaluation in the PPI network were **Hub genes** (degree ≥ 53) (Table 7).

### Key gene expression analysis in Hp infection status

The DEG identified in the PPI network (≥ 53) was analyzed in the TCGA database to assess the correlation with Helicobacter pylori infection. A total of 14 DEGs were highly expressed in positive Helicobacter pylori infection (P < 0.05) and were up-regulated by CagA, namely, ATM, BPTF, CDH1, CTNNB1, HSPA8, HDAC1, POLR2A, ISG15, RPL8, RNP1, RPL30, RPS27, RUVBL1 and CASP3. RT-qPCR verification results showed that Helicobacter pylori CagA caused up-regulation of BPTF, CASP3, CDH1, CTNNB1 and POLR2A expression (P < 0.05) (Figs. 6A–6S).

Chen et al. (2021), *PeerJ*, DOI 10.7717/peerj.11203
Table 2  GO analysis of DEGs.

| Term                        | Description                                               | Count | P-value       |
|-----------------------------|------------------------------------------------------------|-------|---------------|
| GO:0019080                  | viral gene expression                                      | 41    | 2.00E–13      |
| GO:0019083                  | viral transcription                                        | 39    | 2.30E–13      |
| GO:0000956                  | nuclear-transcribed mRNA catabolic process                | 42    | 2.37E–13      |
| GO:000184                   | nuclear-transcribed mRNA catabolic process, nonsense-     | 30    | 1.00E–11      |
|                             | mediated decay                                             |       |               |
| GO:0006402                  | mRNA catabolic process                                     | 54    | 1.00E–11      |
| GO:0006613                  | cotranslational protein targeting to membrane              | 28    | 2.82E–11      |
| GO:0006614                  | SRP-dependent cotranslational protein targeting to         | 27    | 6.77E–11      |
|                             | membrane                                                  |       |               |
| GO:0006401                  | RNA catabolic process                                      | 54    | 2.36E–10      |
| GO:0045047                  | protein targeting to ER                                    | 27    | 1.09E–09      |
| GO:0070972                  | protein localization to endoplasmic reticulum             | 30    | 1.37E–09      |
| GO:0072599                  | establishment of protein localization to endoplasmic      | 27    | 2.08E–09      |
|                             | reticulum                                                 |       |               |
| GO:0016569                  | covalent chromatin modification                            | 57    | 5.86E–09      |
| GO:0016570                  | histone modification                                       | 55    | 9.37E–09      |
| GO:0006612                  | protein targeting to membrane                              | 33    | 4.53E–08      |
| GO:0006605                  | protein targeting                                         | 51    | 1.11E–07      |
| GO:0006338                  | chromatin remodeling                                      | 28    | 3.94E–06      |
| GO:0090150                  | establishment of protein localization to membrane          | 40    | 4.60E–06      |
| GO:0019058                  | viral life cycle                                           | 39    | 9.44E–06      |
| GO:1904837                  | beta-catenin-TCF complex assembly                          | 11    | 9.72E–06      |
| GO:0006413                  | translational initiation                                  | 28    | 1.16E–05      |
| GO:0030522                  | intracellular receptor signaling pathway                   | 33    | 0.000112261   |
| GO:0043401                  | steroid hormone mediated signaling pathway                 | 25    | 0.000193979   |
| GO:0071383                  | cellular response to steroid hormone stimulus             | 30    | 0.000231361   |
| GO:0018205                  | peptidyl-lysine modification                              | 40    | 0.000377348   |
| GO:0031647                  | regulation of protein stability                            | 32    | 0.000377348   |
| GO:0030099                  | myeloid cell differentiation                              | 40    | 0.00113214    |
| GO:0034332                  | adherens junction organization                             | 20    | 0.00117253    |
| GO:0050792                  | regulation of viral process                               | 25    | 0.001423449   |
| GO:0009755                  | hormone-mediated signaling pathway                         | 27    | 0.001423449   |
| GO:0043484                  | regulation of RNA splicing                                | 19    | 0.001726434   |
| GO:0043900                  | regulation of multi-organism process                      | 38    | 0.002773222   |
| GO:0034330                  | cell junction organization                                 | 30    | 0.003325059   |
| GO:0043903                  | regulation of symbiosis, encompassing mutualism through  | 25    | 0.003664061   |
|                             | parasitism                                                 |       |               |
| GO:0016573                  | histone acetylation                                        | 20    | 0.003664061   |
| GO:0030518                  | intracellular steroid hormone receptor signaling pathway  | 18    | 0.003887898   |
| GO:0019079                  | viral genome replication                                  | 17    | 0.004734418   |
| GO:0048525                  | negative regulation of viral process                      | 15    | 0.004893681   |
| GO:0018393                  | internal peptidyl-lysine acetylation                      | 20    | 0.005107113   |

(continued on next page)
| Term                      | Description                                                                 | Count | P-value       |
|---------------------------|------------------------------------------------------------------------------|-------|---------------|
| GO:0043967                | histone H4 acetylation                                                       | 12    | 0.005107113   |
| GO:0002181                | cytoplasmic translation                                                      | 15    | 0.005107113   |
| GO:0033143                | regulation of intracellular steroid hormone receptor signaling pathway       | 13    | 0.005220493   |
| GO:0016331                | morphogenesis of embryonic epithelium                                        | 19    | 0.00555893    |
| GO:0016578                | histone deubiquitination                                                     | 7     | 0.005647965   |
| GO:0051052                | regulation of DNA metabolic process                                          | 38    | 0.006801205   |
| GO:0006475                | internal protein amino acid acetylation                                      | 20    | 0.006801205   |
| GO:1905331                | negative regulation of morphogenesis of an epithelium                       | 6     | 0.007269559   |
| GO:0042692                | muscle cell differentiation                                                  | 35    | 0.007269559   |
| GO:0048545                | response to steroid hormone                                                 | 35    | 0.007269559   |
| GO:0030521                | androgen receptor signaling pathway                                          | 11    | 0.007885335   |
| GO:0018394                | peptidyl-lysine acetylation                                                 | 20    | 0.007885335   |
| GO:0072175                | epithelial tube formation                                                    | 17    | 0.009185473   |
| GO:0034329                | cell junction assembly                                                       | 25    | 0.00934069    |
| GO:1903901                | negative regulation of viral life cycle                                      | 13    | 0.010119749   |
| GO:0030111                | regulation of Wnt signaling pathway                                         | 33    | 0.010318481   |
| GO:0035148                | tube formation                                                               | 18    | 0.011328993   |
| GO:0006473                | protein acetylation                                                         | 22    | 0.012697644   |
| GO:0044782                | cilium organization                                                          | 34    | 0.01287833    |
| GO:0006354                | DNA-templated transcription, elongation                                      | 15    | 0.013214194   |
| GO:0030177                | positive regulation of Wnt signaling pathway                                 | 20    | 0.014794079   |
| GO:006984                 | ER-nucleus signaling pathway                                                 | 9     | 0.014859166   |
| GO:0045637                | regulation of myeloid cell differentiation                                   | 25    | 0.014859166   |
| GO:0008380                | RNA splicing                                                                | 39    | 0.014859166   |
| GO:0034968                | histone lysine methylation                                                  | 15    | 0.014859166   |
| GO:0001843                | neural tube closure                                                          | 13    | 0.014920111   |
| GO:0034333                | adherens junction assembly                                                   | 13    | 0.014920111   |
| GO:0051147                | regulation of muscle cell differentiation                                    | 20    | 0.01537082    |
| GO:0060606                | tube closure                                                                 | 13    | 0.016202371   |
| GO:0051348                | negative regulation of transferase activity                                  | 27    | 0.01736378    |
| GO:0060765                | regulation of androgen receptor signaling pathway                            | 7     | 0.017685114   |
| GO:0001841                | neural tube formation                                                        | 14    | 0.018715149   |
| GO:0032204                | regulation of telomere maintenance                                          | 12    | 0.018715149   |
| GO:0043901                | negative regulation of multi-organism process                               | 19    | 0.018715149   |
| GO:0051054                | positive regulation of DNA metabolic process                                 | 23    | 0.018715149   |
| GO:0060271                | cilium assembly                                                              | 32    | 0.019274508   |
| GO:0060070                | canonical Wnt signaling pathway                                              | 30    | 0.019418493   |
| GO:0021915                | neural tube development                                                      | 18    | 0.02189728    |
| GO:0010171                | body morphogenesis                                                          | 9     | 0.022689191   |
| GO:0032784                | regulation of DNA-templated transcription, elongation                        | 9     | 0.022689191   |

(continued on next page)
| Term          | Description                                           | Count | P-value       |
|--------------|-------------------------------------------------------|-------|---------------|
| GO:1903311   | regulation of mRNA metabolic process                  | 29    | 0.02307114    |
| GO:0044319   | wound healing, spreading of cells                     | 7     | 0.02307114    |
| GO:00950505  | epiboly involved in wound healing                      | 7     | 0.02307114    |
| GO:2000781   | positive regulation of double-strand break repair      | 7     | 0.02307114    |
| GO:0060766   | negative regulation of androgen receptor signaling pathway | 5     | 0.02307114    |
| GO:0000723   | telomere maintenance                                  | 18    | 0.02307114    |
| GO:0000209   | protein polyubiquitination                             | 28    | 0.023399292   |
| GO:1903900   | regulation of viral life cycle                         | 17    | 0.023399292   |
| GO:0007044   | cell-substrate junction assembly                        | 13    | 0.023399292   |
| GO:0014020   | primary neural tube formation                          | 13    | 0.023399292   |
| GO:0090504   | epiboly                                               | 7     | 0.026209669   |
| GO:0001838   | embryonic epithelial tube formation                    | 15    | 0.026209669   |
| GO:006479    | protein methylation                                    | 19    | 0.026209669   |
| GO:0008213   | protein alkylation                                     | 19    | 0.026209669   |
| GO:0036124   | histone H3-K9 trimethylation                           | 5     | 0.028504092   |
| GO:0016571   | histone methylation                                    | 16    | 0.028504092   |
| GO:0034976   | response to endoplasmic reticulum stress               | 26    | 0.028504092   |
| GO:0036498   | IRE1-mediated unfolded protein response                | 10    | 0.028504092   |
| GO:0046782   | regulation of viral transcription                      | 10    | 0.028504092   |
| GO:0001837   | epithelial to mesenchymal transition                   | 16    | 0.02997652    |
| GO:0072665   | protein localization to vacuole                        | 10    | 0.031694802   |
| GO:0071824   | protein-DNA complex subunit organization               | 26    | 0.031924183   |
| GO:2000779   | regulation of double-strand break repair               | 11    | 0.032301863   |
| GO:0016049   | cell growth                                           | 38    | 0.033335915   |
| GO:0018023   | peptidyl-lysine trimethylation                         | 8     | 0.03675333    |
| GO:0006330   | regulation of response to interferon-gamma             | 6     | 0.03675333    |
| GO:0060334   | regulation of interferon-gamma-mediated signaling pathway | 6     | 0.03675333    |
| GO:0018022   | peptidyl-lysine methylation                            | 15    | 0.039268      |
| GO:0002011   | morphogenesis of an epithelial sheet                   | 9     | 0.034329329   |
| GO:1903391   | regulation of adherens junction organization           | 10    | 0.037116501   |
| GO:0051098   | regulation of binding                                  | 31    | 0.037419782   |
| GO:0034504   | protein localization to nucleus                        | 24    | 0.037419782   |
| GO:0033144   | negative regulation of intracellular steroid hormone receptor signaling pathway | 7     | 0.037419782   |
| GO:0043543   | protein acylation                                      | 23    | 0.03974544    |
| GO:0055007   | cardiac muscle cell differentiation                    | 15    | 0.040008254   |
| GO:1903706   | regulation of hemopoiesis                              | 37    | 0.040008254   |
| GO:0006913   | nucleocytoplasmic transport                            | 9     | 0.040008254   |
| GO:006352    | DNA-templated transcription, initiation                | 23    | 0.040379329   |
| GO:0002067   | glandular epithelial cell differentiation              | 8     | 0.040379329   |
| GO:0007041   | lysosomal transport                                    | 13    | 0.040379329   |

(continued on next page)
| Term                        | Description                                                                 | Count | P-value          |
|-----------------------------|-----------------------------------------------------------------------------|-------|-----------------|
| GO:0032200                  | telomere organization                                                       | 18    | 0.040379329     |
| GO:0051895                  | negative regulation of focal adhesion assembly                              | 5     | 0.040379329     |
| GO:0097242                  | amyloid-beta clearance                                                      | 7     | 0.040379329     |
| GO:0031503                  | protein-containing complex localization                                      | 25    | 0.040379329     |
| GO:0007045                  | cell-substrate adherens junction assembly                                   | 11    | 0.040379329     |
| GO:0048041                  | focal adhesion assembly                                                     | 11    | 0.040379329     |
| GO:0060560                  | developmental growth involved in morphogenesis                              | 22    | 0.040379329     |
| GO:0051169                  | nuclear transport                                                           | 29    | 0.040822401     |
| GO:0010172                  | embryonic body morphogenesis                                                | 4     | 0.040822401     |
| GO:0048096                  | chromatin-mediated maintenance of transcription                             | 4     | 0.040822401     |
| GO:0070933                  | histone H4 deacetylation                                                    | 4     | 0.040822401     |
| GO:1900112                  | regulation of histone H3-K9 trimethylation                                  | 4     | 0.040822401     |
| GO:0043921                  | modulation by host of viral transcription                                   | 6     | 0.04145549      |
| GO:0052472                  | modulation by host of symbiont transcription                                | 6     | 0.04145549      |
| GO:0001959                  | regulation of cytokine-mediated signaling pathway                           | 18    | 0.041669901     |
| GO:0071156                  | regulation of cell cycle arrest                                             | 13    | 0.042842707     |
| GO:2000058                  | regulation of ubiquitin-dependent protein catabolic process                 | 16    | 0.042842707     |
| GO:0042176                  | regulation of protein catabolic process                                      | 31    | 0.042914363     |
| GO:0006623                  | protein targeting to vacuole                                                | 7     | 0.042921458     |
| GO:0051261                  | protein depolymerization 13 0.044702762                                    |       |                 |
| GO:0051099                  | positive regulation of binding                                               | 18    | 0.044702762     |
| GO:0051972                  | regulation of telomerase activity                                           | 8     | 0.044702762     |
| GO:0072666                  | establishment of protein localization to vacuole                            | 8     | 0.044702762     |
| GO:2000059                  | negative regulation of ubiquitin-dependent protein catabolic process        | 8     | 0.044702762     |
| GO:1902115                  | regulation of organelle assembly                                            | 19    | 0.044826787     |
| GO:0050684                  | regulation of mRNA processing                                               | 15    | 0.046220077     |
| GO:0052312                  | modulation of transcription in other organism involved in symbiotic interaction | 6     | 0.046220077     |
| GO:0051893                  | regulation of focal adhesion assembly                                        | 9     | 0.047313553     |
| GO:0090109                  | regulation of cell-substrate junction assembly                              | 9     | 0.047313553     |
| GO:0015931                  | nucleobase-containing compound transport                                    | 22    | 0.047313553     |
| GO:0051236                  | establishment of RNA localization                                           | 19    | 0.048534128     |
| GO:0008347                  | glial cell migration                                                         | 8     | 0.048610574     |
| GO:0097193                  | intrinsic apoptotic signaling pathway                                       | 25    | 0.048695169     |
| GO:0022625                  | cytosolic large ribosomal subunit                                          | 22    | 7.58E–12        |
| GO:0005925                  | focal adhesion                                                             | 54    | 1.22E–10        |
| GO:0005924                  | cell-substrate adherens junction                                           | 54    | 1.22E–10        |
| GO:0030055                  | cell-substrate junction                                                     | 54    | 1.35E–10        |
| GO:0022626                  | cytosolic ribosome                                                         | 26    | 2.99E–10        |
| GO:0015934                  | large ribosomal subunit                                                     | 24    | 5.69E–08        |
| GO:0044445                  | cytosolic part                                                             | 34    | 3.89E–07        |

(continued on next page)
| Term | Description                                      | Count | P-value   |
|------|--------------------------------------------------|-------|-----------|
| GO:0044391 | ribosomal subunit                             | 28    | 1.77E–06  |
| GO:0000123 | histone acetyltransferase complex              | 17    | 9.25E–06  |
| GO:0031248 | protein acetyltransferase complex              | 18    | 9.25E–06  |
| GO:1902493 | acetyltransferase complex                      | 18    | 9.25E–06  |
| GO:0035097 | histone methyltransferase complex              | 16    | 5.31E–05  |
| GO:0016363 | nuclear matrix                                 | 18    | 6.56E–05  |
| GO:0034399 | nuclear periphery                              | 19    | 0.000262  |
| GO:0070603 | SWI/SNF superfamily-type complex               | 14    | 0.000262  |
| GO:0042788 | polysomal ribosome                             | 9     | 0.00027895|
| GO:0098984 | neuron to neuron synapse                       | 34    | 0.000596952|
| GO:0014069 | postsynaptic density                           | 32    | 0.000714655|
| GO:0032279 | asymmetric synapse                             | 32    | 0.000864164|
| GO:0030496 | midbody                                         | 21    | 0.000875088|
| GO:0005840 | ribosome                                        | 28    | 0.000907132|
| GO:0000790 | nuclear chromatin                               | 35    | 0.000907132|
| GO:0099572 | postsynaptic specialization                    | 33    | 0.000926357|
| GO:1904949 | ATPase complex                                  | 15    | 0.000926357|
| GO:0034708 | methyltransferase complex                       | 16    | 0.000926357|
| GO:0005667 | transcription factor complex                    | 34    | 0.000926357|
| GO:0016607 | nuclear speck                                   | 36    | 0.000940948|
| GO:0005938 | cell cortex                                     | 30    | 0.001070189|
| GO:0101002 | ficolin-1-rich granule                          | 21    | 0.001630215|
| GO:0044798 | nuclear transcription factor complex            | 22    | 0.001853344|
| GO:0000812 | Swr1 complex                                    | 5     | 0.002106877|
| GO:0000118 | histone deacetylase complex                     | 10    | 0.004130179|
| GO:0042470 | melanosome                                      | 14    | 0.004493787|
| GO:0048770 | pigment granule                                 | 14    | 0.004493787|
| GO:0005844 | polysome                                        | 11    | 0.005579034|
| GO:0090575 | RNA polymerase II transcription factor complex   | 18    | 0.005891446|
| GO:1904813 | ficolin-1-rich granule lumen                    | 15    | 0.006649382|
| GO:1902562 | H4 histone acetyltransferase complex            | 8     | 0.00862081|
| GO:0031252 | cell leading edge                               | 33    | 0.00862081|
| GO:0005643 | nuclear pore                                    | 12    | 0.00862081|
| GO:0005635 | nuclear envelope                                | 36    | 0.012865041|
| GO:0099092 | postsynaptic density, intracellular component   | 5     | 0.013474404|
| GO:0030027 | lamellipodium                                   | 19    | 0.014296793|
| GO:0070461 | SAGA-type complex                               | 6     | 0.017456175|
| GO:0044455 | mitochondrial membrane part                     | 21    | 0.018871465|
| GO:0099091 | postsynaptic specialization, intracellular      | 5     | 0.026165033|
| GO:0044666 | MLL3/4 complex                                  | 4     | 0.026793714|
| GO:0005913 | cell-cell adherens junction                     | 13    | 0.026869727|

(continued on next page)
| Term                       | Description                               | Count | P-value   |
|----------------------------|-------------------------------------------|-------|-----------|
| GO:0000792                 | heterochromatin                           | 10    | 0.028802567 |
| GO:0099738                 | cell cortex region                        | 7     | 0.028802567 |
| GO:0031965                 | nuclear membrane                          | 24    | 0.041189729 |
| GO:0000124                 | SAGA complex                              | 4     | 0.042262875 |
| GO:0071565                 | nBAF complex                              | 4     | 0.042262875 |
| GO:0044309                 | neuron spine                              | 16    | 0.045301703 |
| GO:0017053                 | transcriptional repressor complex         | 10    | 0.045301703 |
| GO:0097346                 | INO80-type complex                        | 5     | 0.04846594  |
| GO:0016605                 | PML body                                  | 11    | 0.048596868 |
| GO:0003713                 | transcription coactivator activity        | 44    | 1.54E–07   |
| GO:0035257                 | nuclear hormone receptor binding          | 26    | 5.03E–06   |
| GO:0045296                 | cadherin binding                          | 41    | 5.22E–06   |
| GO:0051427                 | hormone receptor binding                  | 27    | 4.05E–05   |
| GO:0003735                 | structural constituent of ribosome        | 28    | 5.75E–05   |
| GO:0050839                 | cell adhesion molecule binding            | 49    | 0.000164258 |
| GO:0061659                 | ubiquitin-like protein ligase activity    | 28    | 0.000564585 |
| GO:0061630                 | ubiquitin protein ligase activity         | 27    | 0.000675924 |
| GO:0042393                 | histone binding                           | 25    | 0.000675924 |
| GO:0047485                 | protein N-terminus binding                | 17    | 0.001209051 |
| GO:0019787                 | ubiquitin-like protein transferase activity | 39  | 0.001859565 |
| GO:0035258                 | steroid hormone receptor binding          | 15    | 0.001859565 |
| GO:0004842                 | ubiquitin-protein transferase activity    | 36    | 0.004671669 |
| GO:0003730                 | mRNA 3′-UTR binding                       | 14    | 0.005656635 |
| GO:0030374                 | nuclear receptor transcription coactivator activity | 11  | 0.016358235 |
| GO:0031267                 | small GTPase binding                      | 38    | 0.016779445 |
| GO:0017016                 | Ras GTPase binding                        | 37    | 0.016779445 |
| GO:0044389                 | ubiquitin-like protein ligase binding      | 29    | 0.016779445 |
| GO:0001085                 | RNA polymerase II transcription factor binding | 18  | 0.017774698 |
| GO:0003714                 | transcription corepressor activity        | 24    | 0.017774698 |
| GO:0033613                 | activating transcription factor binding    | 12    | 0.024777396 |
| GO:0031625                 | ubiquitin protein ligase binding          | 27    | 0.024777396 |
| GO:0042800                 | histone methyltransferase activity (H3-K4 specific) | 5   | 0.024777396 |
| GO:0003779                 | actin binding                             | 36    | 0.026046781 |
| GO:0005088                 | Ras guanyl-nucleotide exchange factor activity | 16  | 0.026427777 |
| GO:0004402                 | histone acetyltransferase activity        | 10    | 0.026427777 |
| GO:0046965                 | retinoid X receptor binding               | 5     | 0.027715133 |
| GO:0055106                 | ubiquitin-protein transferase regulator activity | 5   | 0.027715133 |
| GO:0050681                 | androgen receptor binding                 | 8     | 0.028197009 |
| GO:0061733                 | peptide-lysine-N-acetyltransferase activity | 10  | 0.029296991 |
| GO:0042974                 | retinoic acid receptor binding            | 6     | 0.031275892 |
| GO:0005089                 | Rh guanyl-nucleotide exchange factor activity | 11  | 0.032203632 |
| GO:0016887                 | ATPase activity                           | 32    | 0.037839109 |
| GO:0070577                 | lysine-acetylated histone binding         | 5     | 0.038501268 |

(continued on next page)
Survival analysis of key genes

The Kaplan–Meier plotter bioinformatics analysis platform was used to investigate the prognostic value of genes in 14 potential centers, including data from 875 gastric cancer patients for overall survival analysis. Our results show that under high expression ($P < 0.05$), a total of 7 genes are associated with poor prognosis of gastric cancer ($P < 0.05$), namely ATM, BPTF, CDH1, POLR2A, RNP1, BPL30 and RPS27 (Figs. 7A–7G).

**DISCUSSION**

The development of gastric cancer is an extremely complicated biological process, involving the abnormal expression of various tumor-related genes, activation of various tumor-related pathways, and inactivation of tumor suppressor genes. The causative gene is silent and inactive. In fact, evidence prove that the tumor is induced by genetic and epigenetic changes (Belinsky, 2004; Herman & Baylin, 2003; Jones & Baylin, 2002). Helicobacter pylori is closely related to gastric cancer, and Helicobacter pylori CagA is involved in multiple cellular processes related to carcinogenesis (Hatakeyama, 2017). In combination with public biological databases (such as GO and KEGG), the development of a high-throughput detection technology would facilitate systematic exploration of a list of DEG throughout the genome (Ma, Zhou & Zheng, 2020) and comb through the related BP. The application of informatics provides a good means to comprehend the mechanisms of occurrence and development of gastric cancer at the molecular level.

In this study, we compared 1062 genes with significant differences between the pcDNA3.1::CagA and pcDNA3.1 group via bioinformatics. Of these genes, 594 were upregulated and 468 were downregulated. Functional enrichment revealed that these genes participated in multiple signaling pathways, including the Notch signaling pathway and Wnt signaling pathway. The notch signaling pathway is a signal transduction system that repeatedly regulates cell proliferation and apoptosis. We found that the Notch signaling pathway was closely related to cell differentiation, proliferation, apoptosis, adhesion, and the transformation of epidermal cells into the mesenchyme; this pathway is essential for the normal development of most tissues (Leong & Karsan, 2006; Luo, Renault & Rando, 2005; Maillard, Fang & Pear, 2005; Zanotti & Canalis, 2016). Past studies have demonstrated that this pathway plays an important role in regulating the cell cycle as well (Bhattacharya et al., 2017; Herranz & Milán, 2008; Seidel & Kimble, 2015). In a large number of hematopoietic and solid tumors, the Notch pathway undergoes genetically alteration. The activation or inhibition of the pathway depends on the background and
Table 3  KEGG pathway analysis of DEGs.

| Pathway                                         | P-value   | Genes                                                                 |
|-------------------------------------------------|-----------|----------------------------------------------------------------------|
| Ribosome                                        | 2.39E-10  | RPL18, RPL36A, RPL13, RPL15, RPL35, RPL36, RPL37, RPL38, RPS2, RPL30, RPS27, MRPL13, RPL31, RPL34, RPL8, RPL5, RPL11, MRPL33, RPS23, RPL35A, RPL27, RPL28, RPS7, RPL23, RPL13A, RPL21, RPL37A, UBA52 |
| Adherens junction                               | 1.39E-04  | PARD3, PTPRF, CREBBP, CSNK2B, CTNND1, ACTN1, CDH1, CTNNB1, CTNNB1, CDC42, IGF1R, CSNK2A1, AFDN |
| Ubiquitin mediated proteolysis                  | 9.53E-04  | SYVN1, XIAP, UBE4A, PML, SKP1, BRIC2, STUB1, FANCL, TRIM37, FBXW7, UBE2D2, HUWE1, UBE2K, UBA3, TRIM32, NEDD4L, FBXW11 |
| Bacterial invasion of epithelial cells          | 0.001284745 | CDC42, PT2, SEPT2, PIK3CB, ARPC5L, CDH1, CLTC, CTNNNA1, ITGB1, CTNNB1, DNMT, SEPT9 |
| Viral carcinogenesis                            | 0.002245962 | HIST1H4L, YWHAZ, HIST1H2BC, PIK3CB, CREBBP, UBR4, ACTN1, CDK4, PKM, CCNE2, CDC42, HDAC4, CASP3, HDAC1, GSN, GTF2A2, CREB3L2, RBP, HDAC8, CHD4, SYK |
| Pathways in cancer                              | 0.005009934 | ADCY3, FGFR2, WNT5B, XIAP, PPAR, PML, CDH1, ITGB1, TGFB1, CTNNB1, CCNE2, IGF1R, CDC42, PT2, CASP3, RALB, FAS, RUNX1, PLCC1, CTBP2, RALBP1, PIK3CB, CREBBP, CDH4, CTNNNA1, STAT1, BRIC2, ARHGFE, HDAC1, LAM5, PTCH1, GSP1 |
| Huntington’s disease                            | 0.005587064 | DNAH11, UQCR2C, COX7A2, CREBBP, PPAR, CLTC, NDUF1A, NDUF2B, NDUF2A, SOD2, NDUF5, NDUF5, NRF1, CASP3, HDAC1, CREB3L2, PLCC1, UQCRB |
| Non-alcoholic fatty liver disease               | 0.006537343 | UQCR2C, COX7A2, PIK3CB, LEPR, NDUF2B, NDUF1A, NDUF2B, TGFB1, NDUF2B, CDC42, NDUF5, CASP3, NDUF5, XBP1, MLXIP, FAS, UQCRB |
| Herpes simplex infection                       | 0.007640717 | MAVS, CREBBP, PML, CSNK2B, HCFC1, SKP1, ARNTL1, STAT1, TAB2, POLR2A, TYK2, CASP3, TAF5L, CSNK2A1, TAP2, FAS, IFN8R, IFN8R |
| Fatty acid metabolism                           | 0.008015791 | CPT2, ACADM, HACD3, HACD4, FASN, HADH, HADHA, ACSBG1 |
| Hepatitis B                                     | 0.010764364 | MAVS, YWHAZ, PIK3CB, CREBBP, MAP2K4, HSPG2, CDK4, STAT1, TGFB1, STAT6, CCNE2, CASP3, CREB3L2, FAS, NFATC3 |
| Measles                                         | 0.012473037 | MAVS, PIK3CB, CSNK2B, CDK4, STAT1, TAB2, TYK2, CCNE2, CSNK2A1, FAS, IFN8R2, IFN8R1, IFN8R1, HSP8A, ADAR |
| Transcriptional misregulation in cancer         | 0.015781479 | ASPSCR1, FUS, HIST1H3J, KDM6A, KMT2A, PPAR, PML, AFI1, DDX5, HMG2, ATM, MEN1, IGF1R, PT2, HDAC1, RUNX1 |
| Toxoplasmosis                                   | 0.017710762 | TYK2, CASP3, XIAP, LAMA5, STAT1, BRIC2, TAB2, IFN8R2, ITGB1, TGFB1, HSP8A, IFN8R1 |

(continued on next page)
the activation status of other potential oncogenic pathways. There are several different patterns of abnormal regulatory pathways and their targets in cancer (Ranganathan, Weaver & Capobianco, 2011; Vasquez-Del Carpio et al., 2011; Weaver et al., 2014). These patterns include the activation and inactivation mutations, receptor/ligand overexpression, epigenetic regulation, and the effects of post-translational modifications (Wang et al., 2007). Wnt is a secreted glycoprotein that can regulate diverse biological functions (MacDonald, Tamai & He, 2009). Wnt signaling is one of the main regulators of embryonic development, tissue renewal, and regeneration in multicellular organisms (Sidrat et al., 2021).
This signaling pathway controls several aspects of the development process, including cell proliferation, apoptosis, cell migration, and cell polarity during the development and maintenance of adult stem cells. Cell proliferation and apoptosis are often associated with tumor formation and development (Bordonaro, 2020; Foulquier et al., 2018; Yang et al., 2016). Inappropriate activation of the Wnt pathway is also a major factor influencing the human carcinogenesis (Martin-Orozco et al., 2019) involving 13 enriched genes.

The PPI network analysis provided the interaction network with 845 genes, and the first 3 clusters with a high correlation were analyzed through the MCODE plug-in. Cluster 1 genes mainly participated in the extracellular exosome pathway, cluster 2 genes mainly participated in nuclear-transcribed mRNA catabolic processes, and cluster 3 genes were mainly involved in transcription. Some of the past studies have demonstrated that extracellular exosomes are involved in the development of tumors. The results of GO enrichment in these clusters indicate their partial relationship to tumors, suggesting that the signal molecules regulated by the Oriental strain CagA may participate in the possible molecular mechanism of tumor development.

The 30 key genes with the highest screening in the PPI network were analyzed through data, and 14 genes were highly expressed in Helicobacter pylori-positive gastric cancer patients (according to the TCGA database analysis, including ATM, BPTF, CDH1, CTNNB1, HSPA8, HDAC1, POLR2A), ISG15, RPL8, RNP1, RPL30, RPS27, RUVBL1 and CASP3). Finally, use the Kaplan–Meier plotter tool to predict the relationship between...
### Table 4  Differential genes in Cluster 1.

| Gene name | MCODE_Score | Expression |
|-----------|-------------|------------|
| PSMC2     | 22          | down       |
| RPL13     | 25.68403361 | down       |
| KCNC3     | 23          | up         |
| RPL21     | 25.68403361 | down       |
| RAD23B    | 27          | up         |
| RPL35     | 25.68403361 | down       |
| SRP14     | 26.93349754 | down       |
| RPS23     | 25.68403361 | down       |
| PLEC      | 26.45564516 | up         |
| RPL37A    | 25.68403361 | down       |
| RPL38     | 25.68403361 | down       |
| ISG15     | 26.93349754 | down       |
| RPL15     | 25.68403361 | down       |
| ERI1      | 23          | up         |
| UBA52     | 25.68403361 | down       |
| RPL11     | 25.68403361 | down       |
| RPL28     | 25.68403361 | down       |
| UBE2D2    | 23          | up         |
| RPL34     | 25.68403361 | up         |
| NEDD4L    | 23          | up         |
| RPL36     | 25.68403361 | down       |
| RNF213    | 23          | up         |
| RNF111    | 23          | up         |
| RPS2      | 25.68403361 | down       |
| RPL36A    | 25.68403361 | down       |
| SMG1      | 27          | up         |
| TRIM37    | 23          | down       |
| RPL35A    | 25.68403361 | down       |
| RPS27     | 25.68403361 | down       |
| RPL27     | 25.68403361 | down       |
| LAS1L     | 25          | up         |
| RNPS1     | 27          | up         |
| RPL8      | 25.68403361 | down       |
| RPL23     | 25.68403361 | down       |
| RPS7      | 25.68403361 | up         |
| UBA3      | 23          | down       |
| RPL30     | 25.68403361 | down       |
| RPL13A    | 25.68403361 | down       |
| RPL18     | 25.68403361 | up         |
| RPL5      | 25.68403361 | down       |
| RPL31     | 25.68403361 | down       |

(continued on next page)
### Table 4 (continued)

| Gene name | MCODE_Score | Expression |
|-----------|-------------|------------|
| RNF6      | 23          | down       |
| HUWE1     | 23          | up         |
| GART      | 21.92028986 | up         |
| EIF3L     | 27          | up         |
| UBE2K     | 23          | up         |
| EXOSC8    | 25          | down       |
| BTF3L4    | 25          | down       |
| SECISBP2  | 26          | up         |
| TRIM32    | 23          | up         |
| NACA      | 27          | down       |
| RNF4      | 23          | up         |
| UBE2V1    | 23          | up         |
| TCEB1     | 23          | down       |
| TRIM4     | 23          | up         |
| LTN1      | 23          | up         |
| GAN       | 23          | down       |
| UBR4      | 23          | up         |
| UBE4A     | 23          | up         |
| STUB1     | 23          | down       |
| FBXL8     | 23          | up         |
| RPL37     | 25.94117647 | down       |
| FBXW11    | 23          | up         |
| MRPL13    | 26          | down       |
| SSR1      | 27          | down       |
| SKP1      | 23          | down       |
| FBXW7     | 23          | up         |

### Table 5 Differential genes in Cluster 2.

| Gene name | MCODE_Score | Expression |
|-----------|-------------|------------|
| SRCAP     | 11.1        | up         |
| ARID1A    | 11.1        | up         |
| ATM       | 10.3956044  | down       |
| TRRAP     | 11.25146199 | up         |
| KMT2B     | 12.35       | up         |
| EP400     | 11.28947368 | up         |
| SMARCA4   | 10.11111111 | up         |
| HDAC1     | 10.46769231 | up         |
| KMT2D     | 11.28947368 | up         |
| BPTF      | 10.12681159 | up         |
| KDM6A     | 11.01578947 | up         |
| CHD1      | 10.31578947 | up         |
| KDM6B     | 9.991666667 | up         |
| BRD4      | 10.69264069 | up         |
| KMT2C     | 10.69264069 | up         |
Table 6  Differential genes in Cluster 3.

| Gene name   | MCODE_Score | Expression |
|-------------|-------------|------------|
| KPNB1       | 18.90952381 | down       |
| HNRNPUL1    | 17          | up         |
| DDX5        | 17          | up         |
| SRRT        | 17          | down       |
| RBM17       | 17          | up         |
| PFDN5       | 16.90058484 | down       |
| RUVBL1      | 15.89542484 | up         |
| HNRNPA3     | 17          | up         |
| PCBP2       | 17          | up         |
| FUS         | 17          | up         |
| HNRNPD      | 17          | up         |
| HNRNPM      | 17          | up         |
| HNRNPH2     | 17          | up         |
| HNRNPA2B1   | 17          | up         |
| PCF11       | 17          | down       |
| HNRNPU      | 17          | up         |
| POLR2A      | 17          | up         |
| PKM         | 17          | down       |
| LSM7        | 17          | down       |
| SRRM2       | 17          | up         |

them and the poor prognosis of the patient. We have noticed that the high survival rate
of these 7 genes is very low, which is related to the poor prognosis of gastric cancer,
including genes ATM, BPTF, CDH1, POLR2A, RNP1, BPL30 and RPS27. The enrichment
analysis of these 7 genes showed that they are related to the binding of P53, the binding
of transcription factors and transcriptional regulation. After verification by RT-qPCR, the
results showed that CagA of Helicobacter pylori only caused the up-regulation of 5 genes,
including BPTF, CASP3, CDH1, CTNNB1 and POLR2A. Compared with survival analysis,
BPTF, CDH1 and POLR2A have high gene expression and low survival rate. Past studies
have reported that CDH1 gene mutations are associated with diffuse gastric cancer. This
gene encodes E-cadherin, a transmembrane cadherin, and cell adhesion molecules that
depend on this gene are involved in the formation of cell junctions and the maintenance
of epithelial integrity (Cho et al., 2017; Figueiredo et al., 2019; Li, 2019; Van der Post et al.,
2015). CDH1 is involved in mediating cell adhesion, migration, epithelial cell proliferation
and cell cycle (Han et al., 2019; Pal et al., 2020). CDH1 germline mutations are associated
with the encoded tumor suppressor protein E-cadherin, which is the genetic cause of
hereditary diffuse gastric cancer (Van der Post et al., 2015). Among the other seven genes,
BPTF is the core subunit of the nucleosome remodeling factor (NURF) complex and plays
an important role in chromatin remodeling. This gene can directly activate oncogenic
signals or coordinate activation with other key protein factors, thereby affecting tumor
progression (Zhao et al., 2019). Human POLR2A encodes the highly conserved RPBI
protein, which is the largest of the 12 subunits of the essential RNA polymerase II (pol
II) enzyme. This protein complex is responsible for the transcription of pol II encoded by
all proteins. Further studies have shown that the sustained release of pol II bound to the
promoter, the truncated RPBI encoding and the shortened C-terminal domain will affect
Table 7  Thirty hub genes in the PPI network constructed by STRING (degree ≥ 53).

| Gene symbol | Gene description                                      | Degree | Express |
|-------------|--------------------------------------------------------|--------|---------|
| UBA52       | ubiquitin A-52 residue                                 | 131    | down    |
| HDAC1       | histone deacetylase 1                                 | 94     | up      |
| CTNNB1      | catenin beta 1                                        | 89     | up      |
| POLR2A      | RNA polymerase II subunit A                           | 85     | up      |
| HSPA8       | heat shock protein family A (Hsp70) member 8           | 79     | up      |
| CREBBP      | CREB binding protein                                  | 71     | up      |
| CDH1        | cadherin 1                                            | 69     | up      |
| CDC42       | cell division cycle 42                                | 69     | down    |
| SMARCA4     | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4 | 67     | up      |
| ATM         | ATM serine/threonine kinase                           | 67     | down    |
| RPS2        | ribosomal protein S2 [Homo sapiens]                   | 66     | down    |
| RUVBL1      | RuvB like AAA ATPase 1                                | 66     | up      |
| RPL11       | ribosomal protein L11                                 | 63     | down    |
| RPL5        | ribosomal protein L5                                  | 62     | down    |
| RPL8        | ribosomal protein L8                                  | 62     | down    |
| RPL27       | ribosomal protein L27                                 | 62     | down    |
| ISG15       | ISG15 ubiquitin like modifier                         | 61     | down    |
| RPL31       | ribosomal protein L31                                 | 59     | down    |
| RPL15       | ribosomal protein L15                                 | 59     | down    |
| RPL23       | ribosomal protein L23                                 | 58     | down    |
| RPS27       | ribosomal protein S27                                 | 58     | down    |
| RNPS1       | RNA binding protein with serine rich doma in 1         | 58     | up      |
| RPL13A      | ribosomal protein L13a                                | 56     | down    |
| RPL30       | ribosomal protein L30                                 | 55     | down    |
| RPL35A      | ribosomal protein L30                                 | 55     | down    |
| BPTF        | bromodomain PHD finger transcription                   | 55     | up      |
| RPS7        | ribosomal protein S7                                  | 54     | up      |
| RPL34       | ribosomal protein L34                                 | 54     | up      |
| RPL13       | ribosomal protein L13                                 | 54     | down    |
| CASP3       | caspase 3                                             | 53     | up      |

transcriptional regulation and cell cycle (Haijes et al., 2019). Based on the above analysis, BPTF, CDH1, POLR2A may be important target genes and signal molecules regulated by CagA of Helicobacter pylori and have a poor clinical prognosis. Notch and Wnt may be important signaling pathways regulated by Helicobacter pylori CagA, and play an important role in CagA regulating tumor signal molecules. Through bioinformatics analysis of the target genes and signaling pathways regulated by Helicobacter pylori CagA, and exploring the mechanism of CagA, we found that the target genes are related to the occurrence of multiple tumors in the signaling pathway. In past studies, Helicobacter pylori has a greater relationship with gastric cancer. This provides a theoretical basis for future exploration of the possible molecular mechanism of Helicobacter pylori CagA causing gastric cancer. At present, the interaction between these molecules lacks support, and experimental evidence
Figure 6  Analyze the expression and expression verification of key genes in Hp infection status according to the TCGA database. (A–N) Red color indicate expression in Hp infection status, blue color indicate expression in uninfected status. (O–S) The mRNA levels of BPTF, CASPASE3, CDH1, CTNNB1 and POLR2A by RT-qPCR. *P < 0.05. The Hp/cagA+ infected group compared with the Hp/ cagA−::Cm infected group*Compared between Hp/cagA+ and Hp ΔcagA group, P < 0.05.

Full-size DOI: 10.7717/peerj.11203/fig-6
Figure 7 Kaplan–Meier analyses indicated the overall survival of central genes expressed in patients with gastric cancer. (A–G) $P < 0.05$ was considered to be statistically significant. HR, hazard ratio.

Full-size DOI: 10.7717/peerj.11203/fig-7

is needed to clarify the underlying mechanism. The rise and development of the field of bioinformatics has accelerated the development of biology. Bioinformatics tools provide opportunities to deal with big data that cannot be managed manually (Wroblewski & Peek Jr, 2016)

CONCLUSION

DEG of the *H. pylori* CagA plasmid group and the empty vector (negative control) group were obtained via high-throughput sequencing, followed by bioinformatics analysis using the R software, Cytoscape, and related databases. For this purpose, first, 1062 DEG with statistical significance were identified, of which 594 were upregulated and 468 were downregulated. GO enrichment and KEGG pathway analysis revealed that DEG was mainly enriched in the Wnt pathway, Notch pathway, Adhesive connection, and other pathways in cancer. To provide a theoretical basis for studying the biological processes of gastric cancer,
we successfully constructed DEG PPI network, screened out 30 key genes with a relatively high degree, and further studied the network to understand the interaction among DEG. Comprehensive analysis of TCGA database, RT-qPCR and Kaplan–Meier plotter showed that Helicobacter pylori CagA can cause the up-regulation of genes BPTF, CDH1, POLR2A, and their high expression is attributable to poor clinical results. Through data analysis, these genes may be induced and regulated by Helicobacter pylori CagA. These findings enable us to understand the downstream target gene molecules and signal pathways regulated by Helicobacter pylori CagA, and provide a theoretical basis for studying the mechanism of Helicobacter pylori CagA. The target genes and signal pathways obtained in this study are related to the occurrence and development of tumors. These findings enable us to further explore and understand the basic molecular mechanism of Helicobacter pylori CagA regulating the tumorigenesis and development of target genes and signaling pathways.

ACKNOWLEDGEMENTS
The authors are grateful to the editor, the associate editor, and the reviewer. We thank the Novo Gene for technical assistance and BMCSCI for editing this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding
The work was supported by the National Natural Science Foundation of China (31660031, 31760328, 31960028), the Project of Science and Technology of Guiyang (ZhuKeHe[2017]30-4), the Key Project of Science and Technology of Guizhou Province (QianKeHe JC [2020]1Z010), and the Central Government Guides Local Science and Technology Development Projects of Guizhou (grant no. [2019] 4008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
National Natural Science Foundation of China: 31660031, 31760328, 31960028.
Project of Science and Technology of Guiyang: ZhuKeHe[2017]30-4.
Key Project of Science and Technology of Guizhou Province: 2020]1Z010.
The Central Government Guides Local Science and Technology Development Projects of Guizhou: [2019] 4008.

Competing Interests
The authors declare there are no competing interests.

Author Contributions
• Dingyu Chen performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
• Chao Li performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
• Yan Zhao performed the experiments, prepared figures and/or tables, and approved the final draft.
• Jianjiang Zhou and Yuan Xie conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
• Qinrong Wang analyzed the data, prepared figures and/or tables, and approved the final draft.

Data Availability
The following information was supplied regarding data availability:
Raw data are available in the Supplementary Files.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.11203#supplemental-information.

REFERENCES

Belinsky SA. 2004. Gene-promoter hypermethylation as a biomarker in lung cancer. Nature Reviews Cancer 4(9):707–717 DOI 10.1038/nrc1432.

Bhattacharya A, Li K, Quiquand M, Rimesso G, Baker NE. 2017. The Notch pathway regulates the second mitotic wave cell cycle independently of bHLH proteins. Developmental Biology 431(2):309–320 DOI 10.1016/j.ydbio.2017.08.035.

Bordonaro M. 2020. Hypothesis: retinoblastoma protein inactivation mediates effects of histone deacetylase inhibitor-induced Wnt hyperactivation in colorectal cancer cells. Journal of Cancer 11(3):668–677 DOI 10.7150/jca.37864.

Chmiela M, Karwowska Z, Gonciarz W, Allushi B, Stączek P. 2017. Host pathogen interactions in Helicobacter pylori related gastric cancer. World Journal of Gastroenterology 23(9):1521 DOI 10.3748/wjg.v23.i9.1521.

Cho SY, Park JW, Liu Y, Park YS, Kim JH, Yang H, Um H, Ko WR, Lee BI, Kwon SY, Ryu SW, Kwon CH, Park DY, Lee JH, Lee SI, Song KS, Hur H, Han SU, Chang H, Kim SJ, Kim BS, Yook JH, Yoo MW, Kim BS, Lee IS, Kook MC, Thiessen N, He A, Stewart C, Dunford A, Kim J, Shih J, Saksena G, Cerniack AD, Schumacher S, Weiner AT, Kastritis E, Getz G, Yang EG, Ryu MH, Bass AJ, Kim HK. 2017. Sporadic early-onset diffuse gastric cancers have high frequency of somatic CDH1 alterations, but low frequency of somatic RHOA mutations compared with late-onset cancers. Gastroenterology 153(2):536–549 DOI 10.1053/j.gastro.2017.05.012.

Cover TL. 2016. Helicobacter pylori diversity and gastric cancer risk. mBio 7(1):e01869 DOI 10.1128/mBio.01869-15.

Figueiredo J, Melo S, Carneiro P, Moreira AM, Fernandes MS, Ribeiro AS, Guilford P, Paredes J, Seruca R. 2019. Clinical spectrum and pleiotropic nature of CDH1 germline mutations. Journal of Medical Genetics 56(4):199–208 DOI 10.1136/jmedgenet-2018-105807.
Foulquier S, Daskalopoulos EP, Lluri G, Hermans KCM, Deb A, Blankesteijn WM. 2018. WNT signaling in cardiac and vascular disease. Pharmacological Reviews 70(1):68–141 DOI 10.1124/pr.117.013896.

Gu Y, Feng Q, Liu H, Zhou Q, Hu A, Yamaguchi T, Xia S, Kobayashi H. 2019. Bioinformatic evidences and analysis of putative biomarkers in pancreatic ductal adenocarcinoma. Helixyon 5(8):e02378 DOI 10.1016/j.helixyon.2019.e02378.

Haijes HA, Koster MJE, Rehmann H, Li D, Hakonarson H, Cappuccio G, Hancarova M, Lehalle D, Reardon W, Schaeper GB, Lehman A, Van de Laar IMBH, Tesselaar CD, Turner C, Goldenberg A, Patrier S, Thevenon J, Pinelli M, Brunetti-Pierri N, Prchalová D, Havlovicová M, Vlckova M, Sedláček Z, Lopez E, Ragoussis V, Pagnamenta AT, Kini U, Vos HR, Van Es RM, Van Schaik RFMA, Van Essen TAJ, Kibaek M, Taylor JC, Sullivan J, Shashi V, Petrovskii S, Fagerberg C, Martin DM, Van Gassen KLI, Pfundt R, Falk MJ, Mc Cormick EM, Timmers HTM, Van Hasselt PM. 2019. De novo heterozygous POLR2A variants cause a neurodevelopmental syndrome with profound infantile-onset hypotonia. American Journal of Human Genetics 105(2):283–301 DOI 10.1016/j.ajhg.2019.06.016.

Han T, Jiang S, Zheng H, Yin Q, Xie M, Little MR, Yin X, Chen M, Song SJ, Beg AA, Pandolfi PP, Wan L. 2019. Interplay between c-Src and the APC/C co-activator Cdh1 regulates mammary tumorigenesis. Nature Communications 10(1):3716–3716 DOI 10.1038/s41467-019-11618-7.

Hatakeyama M. 2017. Structure and function of Helicobacter pylori CagA, the first-identified bacterial protein involved in human cancer. Proceedings of the Japan Academy. Series B, Physical and Biological Sciences 93(4):196–219 DOI 10.2183/pjab.93.013.

Herman JG, Baylin SB. 2003. Gene silencing in cancer in association with promoter hypermethylation. New England Journal of Medicine 349(21):2042–2054 DOI 10.1056/nejmra023075.

Herranz H, Milán M. 2008. Signalling molecules, growth regulators and cell cycle control in Drosophila. Cell Cycle 7(21):3335–3337 DOI 10.4161/cc.7.21.6996.

Jones PA, Baylin SB. 2002. The fundamental role of epigenetic events in cancer. Nature Reviews Genetics 3(6):415–428 DOI 10.1038/nrg816.

Leong KG, Karsan A. 2006. Recent insights into the role of Notch signaling in tumorigenesis. Blood 107(6):2223–2233 DOI 10.1182/blood-2005-08-3329.

Li X. 2019. Study of CDH1 gene germ-line mutation and promoter methylation status in hereditary diffuse gastric cancer. Basic & Clinical Medicine 56(6):370–379.

Luo D, Renault VM, Rando TA. 2005. The regulation of Notch signaling in muscle stem cell activation and postnatal myogenesis. Seminars in Cell and Developmental Biology 16(4-5):612–622 DOI 10.1016/j.semcdb.2005.07.002.

Ma X, Zhou L, Zheng S. 2020. Transcriptome analysis revealed key prognostic genes and microRNAs in hepatocellular carcinoma. PeerJ 8:e8930 DOI 10.7717/peerj.8930.

MacDonald BT, Tamai K, He X. 2009. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Developmental cell 17(1):9–26 DOI 10.1016/j.devcel.2009.06.016.
Maillard I, Fang T, Pear WS. 2005. Regulation of lymphoid development, differentiation, and function by the notch pathway. Annual Review of Immunology 23(1):945–974 DOI 10.1146/annurev.immunol.23.021704.115747.

Martin-Orozco E, Sanchez-Fernandez A, Ortiz-Parra I, Ayala-SanNicolas M. 2019. WNT signaling in tumors: the way to evade drugs and immunity. Frontiers in Immunology 10:2854–2854 DOI 10.3389/fimmu.2019.02854.

Pal D, Torres AE, Stromberg BR, Messina AL, Dickson AS, De K, Willard B, Venere M, Summers MK. 2020. Chk1-mediated phosphorylation of Cdh1 promotes the SCF(βTRCP)-dependent degradation of Cdh1 during S-phase and efficient cell-cycle progression. Cell Death & Disease 11(4):298–298 DOI 10.1038/s41419-020-2493-1.

Park JY, Forman D, Waskito LA, Yamaoka Y, Crabtree JE. 2018. Epidemiology of helicobacter pylori and CagA-positive infections and global variations in gastric cancer. Toxins 10(4):163 DOI 10.3390/toxins10040163.

Ranganathan P, Weaver KL, Capobianco AJ. 2011. Notch signalling in solid tumours: a little bit of everything but not all the time. Nature Reviews Cancer 11(5):338–351 DOI 10.1038/nrc3035.

Seidel HS, Kimble J. 2015. Cell-cycle quiescence maintains Caenorhabditis elegans germline stem cells independent of GLP-1/Notch. eLife 4:e10832 DOI 10.7554/elife.10832.

Sidrat T, Khan AA, Idrees M, Joo MD, Xu L, Lee KL, Kong IK. 2020. Role of Wnt signalling during in-vitro bovine blastocyst development and maturation in synergism with PPAR δ signaling. Cells 9(4):923 DOI 10.3390/cells9040923.

Takahashi-Kanemitsu A, Knight CT, Hatakeyama M. 2020. Molecular anatomy and pathogenic actions of Helicobacter pylori CagA that underpin gastric carcinogenesis. Cellular and Molecular Immunology 17(1):50–63 DOI 10.1038/s41423-019-0339-5.

Tang S, Jing H, Huang Z, Huang T, Lin S, Liao M, Zhou J. 2020. Identification of key candidate genes in neuropathic pain by integrated bioinformatic analysis. Journal of Cellular Biochemistry 121(2):1635–1648 DOI 10.1002/jcb.29398.

Tepekoy F, Akkoyunlu G, Demir R. 2015. The role of Wnt signaling members in the uterus and embryo during pre-implantation and implantation. Journal of Assisted Reproduction and Genetics 32(3):337–346 DOI 10.1007/s10815-014-0409-7.

Tischkowitz M, Joe ST, Van Dijck B, Van Grieken NC, Roviello F, Seruca R, Van Hillegersberg R, Van Sandick JW, Vehof R, Van Krieken JH, Fitzgerald RC. 2015. Hereditary diffuse gastric cancer: updated clinical guidelines with an emphasis
on germline CDH1 mutation carriers. *Journal of Medical Genetics* 52(6):361–374 DOI 10.1136/jmedgenet-2015-103094.

Vasquez-Del Carpio R, Kaplan FM, Weaver KL, Van Wye JD, Alves-Guerra MC, Robbins DJ, Capobianco AJ. 2011. Assembly of a Notch transcriptional activation complex requires multimerization. *Molecular and Cellular Biology* 31(7):1396–1408 DOI 10.1128/MCB.00360-10.

Wang J, Qin H, Liang J, Zhu Y, Liang L, Zheng M, Han H. 2007. The transcriptional repression activity of KyoT2 on the Notch/RBP-J pathway is regulated by PIAS1-catalyzed sumoylation. *Journal of Molecular Biology* 370(1):27–38 DOI 10.1016/j.jmb.2007.04.010.

Weaver KL, Alves-Guerra MC, Jin K, Wang Z, Han X, Ranganathan P, Zhu X, DaSilva T, Liu W, Ratti F, Demarest RM, Tzimas C, Rice M, Vasquez-Del Carpio R, Dhamane N, Robbins DJ, Capobianco AJ. 2014. NACK is an integral component of the Notch transcriptional activation complex and is critical for development and tumorigenesis. *Cancer Research* 74(17):4741–4751 DOI 10.1158/0008-5472.CAN-14-1547.

Wroblewski LE, Peek Jr RM. 2016. Helicobacter pylori, and Cancer, and the Gastric Microbiota. *Advances in Experimental Medicine and Biology* 908:393–408 DOI 10.1007/978-3-319-41388-4_19.

Yang K, Wang X, Zhang H, Wang Z, Nan G, Li Y, Zhang F, Mohammed MK, Haydon RC, Luu HH, Bi Y, He TC. 2016. The evolving roles of canonical WNT signaling in stem cells and tumorigenesis: implications in targeted cancer therapies. *Laboratory Investigation; A Journal of Technical Methods and Pathology* 96(2):116–136 DOI 10.1038/labinvest.2015.144.

Zanotti S, Canalis E. 2016. Notch signaling and the skeleton. *Endocrine Reviews* 37(3):223–253 DOI 10.1210/er.2016-1002.

Zhao X, Zheng F, Li Y, Hao J, Tang Z, Tian C, Yang Q, Zhu T, Diao C, Zhang C, Chen M, Hu S, Guo P, Zhang L, Liao Y, Yu W, Chen M, Zou L, Guo W, Deng W. 2019. BPTF promotes hepatocellular carcinoma growth by modulating hTERT signaling and cancer stem cell traits. *Redox Biology* 20:427–441 DOI 10.1016/j.redox.2018.10.018.