Identification of DEGs and Transcription Factors Involved in *H. pylori*-associated inflammation and their relevance with Gastric Cancer

Honghao Yin [1, 2, 3], **Aining Chu** [1, 2, 3], **Songyi Liu** [1, 2, 3], **Yuan Yuan** [Corresp., 1, 2, 3], **Yuehua Gong** [Corresp. 1, 2, 3]

[1] Tumor Etiology and Screening Department of Cancer Institute and General Surgery, the First Hospital of China Medical University, Shenyang, LiaoNing, China

[2] Key Laboratory of Cancer Etiology and Prevention in Liaoning Education Department, the First Hospital of China Medical University, Shenyang, LiaoNing, China

[3] Key Laboratory of GI Cancer Etiology and Prevention in Liaoning Province, the First Hospital of China Medical University, Shenyang, LiaoNing, China

Corresponding Authors: Yuan Yuan, Yuehua Gong
Email address: yuanyuan@cmu.edu.cn, yhgong@cmu.edu.cn

Background: Previous studies have indicated that chronic inflammation linked to *H. pylori* infection is the leading causes for gastric cancer (GC). However, the exact mechanism is not entirely clear until now.

Purpose: To identify the key molecules and TFs involved in *H. pylori* infection and to provide new insights into *H. pylori*-associated carcinogenesis and lay the groundwork for the prevention of GC.

Results: GO and KEGG analysis revealed that The DEGs of Hp^+-NAG were mainly associated with Immune response, Chemokine activity, Extracellular region and Rheumatoid arthritis pathway. The DEGs of Hp^+-AG-IM were related to Apical plasma membrane, Intestinal cholesterol absorption, Transporter activity and Fat digestion and absorption pathway. In Hp^+-NAG network, TNF, CXCL8, MMP9, CXCL9, CXCL1, CCL20, CTLA4, CXCL2, C3, SAA1 and FOXP3, JUN had statistical significance between normal and cancer in TCGA database. In Hp^+-AG-IM network APOA4, GCG, CYP3A4, XPNPEP2 and FOXP3, JUN differed statistically in the comparison of normal and cancer in TCGA database. FOXP3 were negatively associated with overall survival, and the association of JUN were positively.

Conclusion: The current study identified key DEGs and their transcriptional regulatory networks involved in *H. pylori*-associated NAG, AG-IM and GC and found that patients with higher expressed FOXP3 or lower expressed JUN had shorter overall survival time. Our study provided new directions for inflammation-associated oncogenic transformation involved in *H. pylori* infection.

Identification of DEGs and Transcription Factors Involved in *H. pylori*-associated inflammation and Their Relevance with Gastric Cancer

Honghao Yin, Aining Chu, Songyi Liu, Yuan Yuan, Yuehua Gong**

[1]Tumor Etiology and Screening Department of Cancer Institute and General Surgery, the First Hospital of China Medical University, Shenyang 110001, China

[2]Key Laboratory of Cancer Etiology and Prevention in Liaoning Education Department, the First Hospital of China Medical University, Shenyang 110001, China

[3]Key Laboratory of GI Cancer Etiology and Prevention in Liaoning Province, the First Hospital of China Medical University, Shenyang 110001, China

*Co-Correspondence should be addressed to

Dr. Gong Yuehua, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Hospital of China Medical University, No.155 NanjingBei Street, Heping District, Shenyang, Liaoning Province, P.R. China 110001， Telephone: +86-024-83282153; fax: +86-024-83282383. Email：yhgong@cmu.edu.cn

Dr. Yuan Yuan, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Hospital of China Medical University, No.155 NanjingBei Street, Heping District, Shenyang, Liaoning Province, P.R. China 110001， Telephone:+86-024-83282153;fax: +86-024-83282292.Email：yuanyuan@cmu.edu.cn

Abstract:

Background: Previous studies have indicated that chronic inflammation linked to *H. pylori* infection is the leading causes for gastric cancer (GC). However, the exact mechanism is not entirely clear until now.

Purpose: To identify the key molecules and TFs involved in *H. pylori* infection and to provide new insights into *H. pylori*-associated carcinogenesis and lay the groundwork for the prevention of GC.

Results: GO and KEGG analysis revealed that The DEGs of Hp^+-NAG were mainly associated with Immune response, Chemokine activity, Extracellular region and Rheumatoid arthritis pathway. The DEGs of Hp^+-AG-IM were related to Apical plasma membrane, Intestinal cholesterol absorption, Transporter activity and Fat digestion and absorption pathway. In Hp^+-NAG network, TNF, CXCL8, MMP9, CXCL9, CXCL1, CCL20, CTLA4, CXCL2, C3, SAA1 and FOXP3, JUN had statistical significance between normal and cancer in TCGA database. In Hp^+-AG-IM network APOA4, GCG, CYP3A4, XPNPEP2 and FOXP3, JUN differed statistically in the comparison of normal and cancer in TCGA database. FOXP3 were negatively associated with overall survival, and the association of JUN were positively.

Conclusion: The current study identified key DEGs and their transcriptional regulatory networks involved in *H. pylori*-associated NAG, AG-IM and GC and found that patients with higher expressed FOXP3 or lower expressed JUN had shorter overall survival time. Our study provided new directions for inflammation-associated oncogenic transformation involved in *H. pylori* infection.

Keywords: *H. pylori*, inflammation, gastric cancer, DEGs , transcription factor, regulatory network

Introduction

Gastric cancer (GC) is one of the most common malignancies, and ranks the second in the world in terms of cancer mortality (Chmiela et al. 2017; Dadashzadeh et al. 2017; Van Cutsem et al. 2016). Helicobacter pylori (*H. pylori*) infection can induce inflammatory, affect the growth, differentiation, renewal, integrity of mucosa, and lead to gastric injury. Several previous studies have indicated that chronic inflammation linked to *H. pylori* infection is one of the leading causes of GC (Sipponen et al. 2015). Thus, investigating the inflammation mechanisms of *H.pylori* infection is of great importance for the occurrence and progression of GC.

According to the *Correa*'s model (Correa 1992), *H. pylori* infection was firmly related to intestinal-type GC through the process of non-atrophic gastritis (NAG), atrophic gastritis (AG), intestinal metaplasia (IM), atypical hyperplasia. In the NAG stage, infection with *H. pylori* is characterized by infiltration of lymphocytes, polymorphonuclear leukocytes, and macrophages in the gastric mucosa. Over time, gastric mucosa would suffer a loss of glandular cells and be replaced by intestinal and fibrous tissues eventually, which is manifested as AG or AG-IM. In these processes, *H.pylori* can induce the expression of pro-inflammatory factors, chemokines, inflammatory regulatory factors and contribute to gastric disorder(Ernst et al. 2000). Current research indicate that chronic NAG and AG-IM are associated with the development of GC(Matysiak-Budnik et al. 2006). Also, the existing intervention trials have shown that *H. pylori* eradication in the NAG and AG-IM stage is helpful to the prevention of GC (Kuipers et al. 2006). However, until now it is not entirely clear about the key genes involved in the *H. pylori*-related inflammation.

Gene expression is determined at both transcriptional and post-transcriptional levels. Transcription factors (TFs) regulate gene expression by site-specific binding to chromosomal

DNA, thereby preventing or promoting the transcription by RNA polymerase. Studies have shown that TFs vary during different inflammatory stages of *H. pylori* infection. For example, activator protein-1 (AP-1) and cAMP-response element-binding protein (CREB) modulate early inflammatory responses, while nuclear factor-κB (NF-κB) and interferon-sensitive response element (ISRE) contact with inflammatory processes of AG (Sokolova et al. 2017). Thus, searching for key TFs involved in the inflammatory response of *H. pylori* is of great importance for the development of GC .

As the availability of multi-level expression data for diseases and normal tissues increases, new opportunities for extraction and integration of large data sets, such as gene expression omnibus (GEO) and The Cancer Genome Atlas (TCGA), may help to provide a more comprehensive understanding of the pathogenesis of *H. pylori* infection. Here, we used online bioinformatics resources to identify the key molecules involved in *H. pylori*-related gastric inflammation and the TFs regulatory networks. Our study intended to provide new insights into *H. pylori*-associated carcinogenesis and lay the foundation for GC prevention.

Materials and methods

Microarray Data

Two sets of microarray data from the public database GEO were used in this study. For the data set with the GEO accession number GSE27411, three cases of no *H. pylori* infection (Hp^--No), three cases of *H. pylori* infection without corpus-predominant AG (Hp^+-NAG) and three cases of *H. pylori* infection with corpus-predominant AG (Hp^+-AG-IM) were included. For the data set with accession number GSE60662, four replicates of the control were included as Hp^--No, four replicates of mild gastritis and four replicates of severe gastritis as Hp^+-NAG, and four replicates of IM as Hp^+-AG-IM.

Data Processing

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) was undertook to compare multiple sets of samples and to identify differentially expressed genes (DEGs) in the GEO series (Barrett et al. 2013). FDR <0.05 and |logFC | > 1 were considered statistically significant.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Enrichment Analyses

GO analysis is a major bioinformatics tool for annotating genes and gene products. It contains terms under three categories: cellular component, molecular function, and biological process. To claim the different underlying biological processes of DEGs involved in H. pylori-related inflammation, GO biological process enrichment analysis was performed using Gene Ontology Consortium (http://www.geneontology.org) and KEGG pathway enrichment analysis was used to find the potential pathways of H. pylori-related inflammation by David database (https://david.ncifcrf.gov/) (Dennis et al. 2003). The cut-off criteria of significant GO terms and KEGG pathways was FDR < 0.05.

Protein-protein interaction (PPI) networks of key DEGs and TFs

The Retrieval of Interacting Genes (STRING) database tool (string-db.org) was used to figure out the interactive relationships of DEGs, and only interactions with a combined score>0.4 were considered as significant and retained. The key DEGs were identified by degree$\geqslant$15, which were calculated using the online tool Centiscape 2.2. PROMO database that can use species-specific searches to detect known transcription regulatory elements (Messeguer et al. 2002). We obtained the DNA sequence from 2000bp upstream to 100bp downstream of the transcription start site of the DEGs from University of California Santa Cruz (UCSC) genome

113 browser database. After entering above sequences into the PROMO database with zero fault
114 tolerance, we obtained all the TFs that could regulate the key DEGs. PPI networks of TFs-key
115 DEGs were visualized and analyzed by Cytoscape 3.4.0 (Scardoni et al. 2009).

116 ***TCGA database Analysis of key DEGs and TFs***

117 The Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/) provides
118 genomic information on 33 types of cancer. In the database, there are 18 GC specimens with *H.*
119 *pylori* positive and 32 normal specimens without *H. pylori* infection (see Table 1 for details).
120 Further, we downloaded RNA expression data and compared the differences of the key DEGs
121 and TFs between *H. pylori* positive GC and normal groups using the Mann-Whitney U test. P
122 <0.05 was considered statistically significant.

123 ***Survival analysis***

124 Kmplot (www.kmplot.com) provided customizable functions such as patient survival
125 analysis (Á et al. 2018). To determine the possible relationship of the key DEGs and TFs with
126 GC prognosis, we performed survival analysis of 882 GC patients in Kmplot. And P <0.05 was
127 considered statistically significant. Figure 1 depicted the flow diagram of all above
128 bioinformatics analysis.

129 ## Result

130 ***Screening of DEGs involved in H. pylori-associated inflammation***

131 Comparing Hp^--No with Hp^+-NAG in GSE27411, there were 191 downregulated and 323
132 upregulated genes. In terms of Hp^--No and Hp^+-NAG in GSE60662, there were 743

downregulated and 1682 upregulated genes. After the intersection, there were 97 high-expressed genes and 14 low-expressed genes screened out.

Comparing Hp^+-NAG with Hp^+-AG-IM in GSE27411, there were 235 downregulated and 508 upregulated genes. In terms of Hp^+-NAG and Hp^+-AG-IM in GSE60662，there were 1376 downregulated and 1364 upregulated genes. After the intersection, there were 342 genes of high expression and 43 genes of low expression screened out .

The cellular functions and pathway analysis of DEGs involved in H. pylori-associated inflammation

As can be seen from Figure 2，GO terms of Hp^+-NAG participated in cell component of extracellular region, space MHC class II protein complex, integral component of lumenal side of endoplasmic reticulum membrane, and transport vesicle membrane. About biological processes, these genes enriched in immune response, inflammatory response, antigen processing and presentation of peptide or polysaccharide antigen via MHC class II and cell chemotaxis. In addition, molecular function suggested enrichment mainly at chemokine activity, MHC class II receptor activity, peptide antigen binding, CXCR chemokine receptor binding, CCR6 chemokine receptor binding. According to KEGG pathway analysis, the most significant pathways were rheumatoid arthritis, staphylococcus aureus infection, asthma, graft-versus-host disease, allograft rejection and so on.

As shown in Figure 3, GO terms of Hp^+-AG-IM participated in cell component of apical plasma membrane, extracellular exosome, brush border, brush border membrane, integral component of membrane. For Biological processes, these genes enriched in intestinal cholesterol absorption, cholesterol homeostasis, retinoid metabolic process, cholesterol efflux, xenobiotic metabolic process. In addition, molecular function suggested enrichment mainly at transporter activity, phospholipid binding, cholesterol transporter activity, protein homodimerization activity, ATPase activity, coupled to transmembrane movement of substances. According to KEGG

158 pathway analysis, the most significant pathways were fat digestion and absorption, metabolic
159 pathways, drug metabolism, protein digestion and absorption, metabolism of xenobiotics by
160 cytochrome P450 and so on.

161 ***Construction of DEGs-TFs PPI networks***

162 As can be seen from Table 2, the key genes of Hp^+-NAG were TNF, CXCL8, MMP9,
163 CXCL9, CXCL1, CCL20, LCN2, CTLA4, FPR1, CXCL2, C3, SAA1, and all of which were
164 high expression. TFs regulated these key DEGs were FOXP3, TP53, ESR1, JUN and FOSB.
165 Figure 4 showed the PPI network of DEGs-TFs involved in Hp^+-NAG.

166 As shown in Table 3，the key genes of Hp^+-AG-IM were APOB, SLC2A2, FABP1, APOA4,
167 NR1H4, APOC3, DGAT1, APOA1, GCG, CYP3A4, DPP4, GLUL, SI, XPNPEP2, MGAM,
168 SLC15A1. Among them, GLUL were low-expressed and others with high expression. And TFs
169 regulated these key DEGs were TBP, NR3C1, FOXP3, ESR1, JUN. Figure 5 showed the PPI
170 network of DEGs-TFs involved in Hp^+-AG-IM.

171 ***The Relevance of key DEGs and TFs with GC in TCGA database***

172 Next, we analyzed above genes between 18 GC with *H. pylori* and 32 normal without *H.*
173 *pylori* in TCGA database. The results indicated that, the expressed differences of TNF, CXCL8,
174 MMP9, CXCL9, CXCL1, CCL20, CTLA4, CXCL2, C3, SAA1 and FOXP3, JUN in Hp^+-NAG
175 network, had statistical significance between normal and cancer（P<0.05）. APOA4, GCG,
176 CYP3A4, XPNPEP2 and FOXP3, JUN of Hp^+-AG-IM network differed between normal and
177 cancer（P<0.05）.

178 ***Survival analysis of key DEGs and TFs in Kmplot***

To further analyze the prognostic characteristics of key DEGs and TFs, survival analysis was performed by Kmplot software. As shown in Figure 6, FOXP3 was negatively associated with overall survival, and the association of JUN were positive.

Discussion：

NAG and AG-IM caused by *H. pylori* infection are closely related to gastric carcinogenesis. However, the key genes and transcriptional regulatory networks in this process are not apparent. In this paper, we used GEO and TCGA database to analyze the key DEGs and TFs involved in *H. pylori*-related inflammation and GC. The present study would provide new insights into the early prevention of gastric diseases caused by *H. pylori*.

Firstly, by comparing Hp^--No with Hp^+-NAG samples, we obtained 111 DEGs, which were mainly related to immune response，inflammatory response，extracellular region and space, MHC class II protein complex, chemokine activity and so on. Through KEGG enrichment, they primarily concentrated on rheumatoid arthritis, staphylococcus aureus infection, allograft rejection and so on. In TCGA database, the expression of TNF, CXCL8, MMP9, CXCL9, CXCL1, CCL20, CTLA4, CXCL2, C3, SAA1 and FOXP3, JUN were differed between cancer and normal, suggesting that these genes may be related to both NAG inflammation and GC. Except JUN, these genes were all high expressed in GC group. CXCL and CXCR are members of endogenous ligands or receptor families of chemokines, and current studies have believed that they are strictly correlated with many kinds of cancers (Pevida et al. 2014; Wyler et al. 2014). *H. pylori* could upregulate TNFα to induce CCL20 expression in gastric epithelial cells, which were positively associated with the degree of inflammation (Wu et al. 2007). Cytotoxic T lymphocyte‐associated antigen‐4 (CTLA‐4), is an essential negative regulator expressed on regulatory T cells (Tregs) and activated T cells (Hayakawa et al. 2016). During *H. pylori*

202 infection, CTLA-4 engagement would reduce immune response and promote the development of
203 stomach inflammation (Watanabe et al. 2004). Some studies have asserted that *H. pylori* induces
204 macrophages to release TNF and CXCL8 (Tavares et al. 2018), thereby suppressing immunity
205 and promoting tumorigenesis and development (Lin et al. 2019). However, other CXCL family
206 members identified in this study, such as CXCL8, CXCL9, CXCL1, and CXCL2, are currently
207 less described in *H. pylori* infection. CXCL9 was shown to upregulate PD-L1 during gastric
208 carcinogenesis by activating STAT and PI3K-Akt pathways (Zhang et al. 2018). CXCL1
209 improved MMP-2/9 expression through the integrin β1/FAK/AKT signaling pathway and
210 promoted lymph node metastasis of GC (Wang et al. 2017). CXCL2 increased bladder cancer
211 progression by recruiting myeloid-derived suppressor cells. It has been reported that the
212 inflammation of *H. pylori* may improve MMP-9 expression (Slomiany et al. 2016). *Sung et al*
213 demostrated that SAA was induced from lung cancer cells by the interaction with monocyte
214 macrophages, in return, inducing MMP-9 from monocyte macrophages, thereby promoting the
215 occurrence and development of lung adenocarcinoma (Sung et al. 2011) . *Yuan et al* showed
216 local C3 deposition in the tumor microenvironment was a relevant immune signature for
217 predicting prognosis of GC. It may aberrantly activate JAK2/STAT3 pathway, then allowing
218 tumor progression. Foxp3 is considered to be a hallmark of the forkhead transcription factor
219 family (Guo et al. 2016). However, it is unclear how Foxp3 participates in the process of *H.*
220 *pylori*-associated inflammation. Our study found that CCL20, CXCL1, CXCL9, and MMP9 may
221 be regulated by FOXP3. JUN is a TF member of the AP-1 family, which are crucial regulators of
222 improving cell proliferation and differentiation (Shaulian 2010). However, our research found
223 the decreased JUN expression in GC, which might be induced by the dedifferentiation process
224 during tumorigenesis.

225 Comparing Hp^+-NAG with Hp^+-AG-IM, 385 DEGs were screened out. These genes were
226 mostly related to apical plasma membrane，extracellular exosome，intestinal cholesterol

absorption，and so on. Through KEGG enrichment analysis, they principally concentrated in fat digestion and absorption, metabolic pathways, drug metabolism, and so on. It is worth noting that the expression of APOA4, GCG, CYP3A4, XPNPEP2, and FOXP3, JUN were different between cancer and normal samples in TCGA database. APOA4 was reported to be closely related to urinary bladder cancer (Soukup et al. 2019). CYP3A4 is currently indicated for the treatment of ovarian and breast cancer (Fiszer-Maliszewska et al. 2018; Liu et al. 2019). *Bian et al* found that GCG affected the development and progression of colon cancer(Bian et al. 2019). *Li et al* demonstrated that XPNPEP2 was associated with lymph node metastasis in prostate cancer patients (Li et al. 2019). However, the relationships between these genes and *H. pylori*-related AG-IM and GC are still unclear. But interestingly, we found that these genes were all involved in metabolically changes. GCG was related to glucose metabolism; other genes were closely associated with lipid metabolism. At present, the relationship between metabolic regulation and cancer have made significant progress (Xiao et al. 2017). In our study, they showed a trend of increasing in NAG and then decreasing in AG and GC, which may be closely associated with the occurrence of GC.

Hu, et al et al screened genes involved in the Hp^{+}-GC group than in the *H. pylori*$^{-}$-GC group, furthermore verified the results in TCGA database(Hu et al. 2018). They did not analyze differential expressed genes during the dynamic progression from NAG, AG-IM and GC. They found TP53 was upregulated, and CCDC151, CHRNB2, GMPR2, HDGFRP2 and VSTM2L were downregulated in the *H.pylori*-positive GC group. By our screening, we also found upregulated TP53 and downregulated CHRNB2, VSTM2L in Hp^{+}-GC (P <0.05). But they were not the DEGs in Hp^{+}-NAG or Hp^{+}-AG-IM group. It suggests that these genes may be involved in Hp-associated GC, with more significant changes in cancer tissues, and may not play the most critical role in the process from inflammation to carcinogenesis.

Further, we explored the correlation of DEGs/TFs with GC prognosis in Kmplot database. It showed that patients with higher expressed FOXP3 or lower expressed JUN had shorter overall survival time. *Wyler et al* have claimed that the median overall survival rate of GC patients with high FOXP3 expression is significant lower than that of patients with low expression (Wyler et al. 2014). Furthermore, *Ma GF et al* found that FOXP3 expression in tumor cells indicated a good prognosis, while high expression in the stroma indicated a poor prognosis (Ma et al. 2014). It implied us that the prognosis of patients may be adjusted by examining the position of FOXP3 expression. Alternately, some studies have shown that JUN expression is associated with poor prognosis (Zhang et al. 2018). JUN generally regulates cell differentiation and has a decreased expression with decreasing differentiation. In our study, JUN expression fluctuated from AG-IM to GC. However, GC patients with lower JUN expression had a shorter survival time. The above results showed that FOXP3, JUN involved in Hp-related NAG, AG-IM, GC, and also closely related to the prognosis of GC. That is to say, as TFs, they may be involved in the transformation process of *H. pylori* infection-related inflammation to cancer.

In conclusion, the current study revealed key DEGs and their transcriptional regulatory networks involved in *H. pylori*-associated NAG, AG and GC. TNF, CXCL8, MMP9, CXCL9, CXCL1, CCL20, CTLA4, CXCL2, C3, SAA1 and FOXP3, JUN were key DEGs and TFs of NAG, related with *H. pylori*-infected GC. APOA4, GCG, CYP3A4, XPNPEP2 and FOXP3, JUN constituted a regulatory network of key DEGs and TFs, and were involved in AG-IM and GC. More importantly, FOXP3 and JUN were closely connected with the survival of patients with GC. Our study provided new directions for inflammation-associated oncogenic transformation of *H. pylori* infection.

References:

Á N, A L, O M, and B G. 2018. Author Correction: Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific reports* 8:11515.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, and Soboleva A. 2013. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 41:D991-995. 10.1093/nar/gks1193

Bian Q, Chen J, Qiu W, Peng C, Song M, Sun X, Liu Y, Ding F, Chen J, and Zhang L. 2019. Four targeted genes for predicting the prognosis of colorectal cancer: A bioinformatics analysis case. *Oncology Letters*. 10.3892/ol.2019.10866

Chmiela M, Karwowska Z, Gonciarz W, Allushi B, and Staczek P. 2017. Host pathogen interactions in Helicobacter pylori related gastric cancer. *World J Gastroenterol* 23:1521-1540. 10.3748/wjg.v23.i9.1521

Correa P. 1992. Human gastric carcinogenesis: a multistep and multifactorial process--First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res* 52:6735-6740.

Dadashzadeh K, Peppelenbosch MP, and Adamu AI. 2017. Helicobacter pylori Pathogenicity Factors Related to Gastric Cancer. *Can J Gastroenterol Hepatol* 2017:7942489. 10.1155/2017/7942489

Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, and Lempicki RA. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:P3.

Ernst PB, and Gold BD. 2000. The disease spectrum of Helicobacter pylori: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. *Annu Rev Microbiol* 54:615-640. 10.1146/annurev.micro.54.1.615

Fiszer-Maliszewska L, Laczmanski L, Dolinska A, Jagas M, Kolodziejska E, Jankowska M, and Kusnierczyk P. 2018. Polymorphisms of ABCB1, CYP3A4 and CYP3A5 Genes in Ovarian Cancer and Treatment Response in Poles. *Anticancer Res* 38:1455-1459. 10.21873/anticanres.12370

Guo G, He Z, and Shi Z. 2016. Correlation between FOXP3 expression and gastric cancer. *Oncol Lett* 12:1554-1558. 10.3892/ol.2016.4752

Hayakawa S, Okada S, Tsumura M, Sakata S, Ueno Y, Imai K, Morio T, Ohara O, Chayama K, and Kobayashi M. 2016. A Patient with CTLA-4 Haploinsufficiency Presenting Gastric Cancer. *J Clin Immunol* 36:28-32. 10.1007/s10875-015-0221-x

Hu Y, He C, Liu JP, Li NS, Peng C, Yang-Ou YB, Yang XY, Lu NH, and Zhu Y. 2018. Analysis of key genes and signaling pathways involved in Helicobacter pylori-associated gastric cancer based on The Cancer Genome Atlas database and RNA sequencing data. *Helicobacter* 23:e12530. 10.1111/hel.12530

Kuipers EJ, and Sipponen P. 2006. Helicobacter pylori eradication for the prevention of gastric cancer. *Helicobacter* 11 Suppl 1:52-57. 10.1111/j.1478-405X.2006.00425.x

Li F, Dai Y, Xu H, Huang K, Zhou Y, Luo D, Ma D, Xi L, Lv M, and Ma X. 2019. XPNPEP2 is associated with lymph node metastasis in prostate cancer patients. *Sci Rep* 9:10078. 10.1038/s41598-019-45245-5

Lin C, He H, Liu H, Li R, Chen Y, Qi Y, Jiang Q, Chen L, Zhang P, Zhang H, Li H, Zhang W, Sun Y, and Xu J. 2019. Tumour-associated macrophages-derived CXCL8 determines immune evasion through autonomous PD-L1 expression in gastric cancer. *Gut* 68:1764-1773. 10.1136/gutjnl-2018-316324

Liu X, Huang X, Zhang S, Niu F, Ouyang Y, Shou Z, and Liu J. 2019. Correlations between CYP3A4 polymorphism and susceptibility to breast cancer in Chinese Han population. *Int J Clin Oncol* 24:179-188. 10.1007/s10147-018-1346-8

Ma GF, Miao Q, Liu YM, Gao H, Lian JJ, Wang YN, Zeng XQ, Luo TC, Ma LL, Shen ZB, Sun YH, and Chen SY. 2014. High FoxP3 expression in tumour cells predicts better survival in gastric cancer and its role in tumour microenvironment. *Br J Cancer* 110:1552-1560. 10.1038/bjc.2014.47

Matysiak-Budnik T, and Megraud F. 2006. Helicobacter pylori infection and gastric cancer. *Eur J Cancer* 42:708-716. 10.1016/j.ejca.2006.01.020

Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, and Alba MM. 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 18:333-334. 10.1093/bioinformatics/18.2.333

Pevida M, Lastra A, Meana A, Hidalgo A, Baamonde A, and Menendez L. 2014. The chemokine CCL5 induces CCR1-mediated hyperalgesia in mice inoculated with NCTC 2472 tumoral cells. *Neuroscience* 259:113-125. 10.1016/j.neuroscience.2013.11.055

Scardoni G, Petterlini M, and Laudanna C. 2009. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* 25:2857-2859. 10.1093/bioinformatics/btp517

Shaulian E. 2010. AP-1 — The Jun proteins: Oncogenes or tumor suppressors in disguise? *Cellular Signalling* 22:894-899. 10.1016/j.cellsig.2009.12.008

Sipponen P, and Maaroos HI. 2015. Chronic gastritis. *Scand J Gastroenterol* 50:657-667. 10.3109/00365521.2015.1019918

Slomiany BL, and Slomiany A. 2016. Helicobacter pylori-elicited induction in gastric mucosal matrix metalloproteinase-9 (MMP-9) release involves ERK-dependent cPLA2 activation and its recruitment to the membrane-localized Rac1/p38 complex. *Inflammopharmacology* 24:87-95. 10.1007/s10787-016-0261-8

Sokolova O, and Naumann M. 2017. NF-kappaB Signaling in Gastric Cancer. *Toxins (Basel)* 9. 10.3390/toxins9040119

Soukup V, Capoun O, Pesl M, Vavrova L, Sobotka R, Levova K, Hanus T, Zima T, and Kalousova M. 2019. The significance of calprotectin, CD147, APOA4 and DJ-1 in non-

invasive detection of urinary bladder carcinoma. *Neoplasma* 66:1019-1023. 10.4149/neo_2019_190124N74

Sung HJ, Ahn JM, Yoon YH, Rhim TY, Park CS, Park JY, Lee SY, Kim JW, and Cho JY. 2011. Identification and validation of SAA as a potential lung cancer biomarker and its involvement in metastatic pathogenesis of lung cancer. *J Proteome Res* 10:1383-1395. 10.1021/pr101154j

Tavares R, and Pathak SK. 2018. Induction of TNF, CXCL8 and IL-1beta in macrophages by Helicobacter pylori secreted protein HP1173 occurs via MAP-kinases, NF-kappaB and AP-1 signaling pathways. *Microb Pathog* 125:295-305. 10.1016/j.micpath.2018.09.037

Van Cutsem E, Sagaert X, Topal B, Haustermans K, and Prenen H. 2016. Gastric cancer. *Lancet* 388:2654-2664. 10.1016/s0140-6736(16)30354-3

Wang Z, Wang Z, Li G, Wu H, Sun K, Chen J, Feng Y, Chen C, Cai S, Xu J, and He Y. 2017. CXCL1 from tumor-associated lymphatic endothelial cells drives gastric cancer cell into lymphatic system via activating integrin beta1/FAK/AKT signaling. *Cancer Lett* 385:28-38. 10.1016/j.canlet.2016.10.043

Watanabe K, Murakami K, Sato R, Okimoto T, Maeda K, Nasu M, Nishizono A, and Fujioka T. 2004. CTLA-4 blockade inhibits induction of Helicobacter pylori-associated gastritis in mice. *Clin Exp Immunol* 135:29-34. 10.1111/j.1365-2249.2004.02338.x

Wu YY, Tsai HF, Lin WC, Hsu PI, Shun CT, Wu MS, and Hsu PN. 2007. Upregulation of CCL20 and recruitment of CCR6+ gastric infiltrating lymphocytes in Helicobacter pylori gastritis. *Infect Immun* 75:4357-4363. 10.1128/IAI.01660-06

Wyler L, Napoli CU, Ingold B, Sulser T, Heikenwalder M, Schraml P, and Moch H. 2014. Brain metastasis in renal cancer patients: metastatic pattern, tumour-associated macrophages and chemokine/chemoreceptor expression. *Br J Cancer* 110:686-694. 10.1038/bjc.2013.755

Xiao S, and Zhou L. 2017. Gastric cancer: Metabolic and metabolomics perspectives (Review). *Int J Oncol* 51:5-17. 10.3892/ijo.2017.4000

Zhang C, Li Z, Xu L, Che X, Wen T, Fan Y, Li C, Wang S, Cheng Y, Wang X, Qu X, and Liu Y. 2018. CXCL9/10/11, a regulator of PD-L1 expression in gastric cancer. *BMC Cancer* 18:462. 10.1186/s12885-018-4384-8

Figure 1

Depicted the flow diagram of all above bioinformatics analysis.

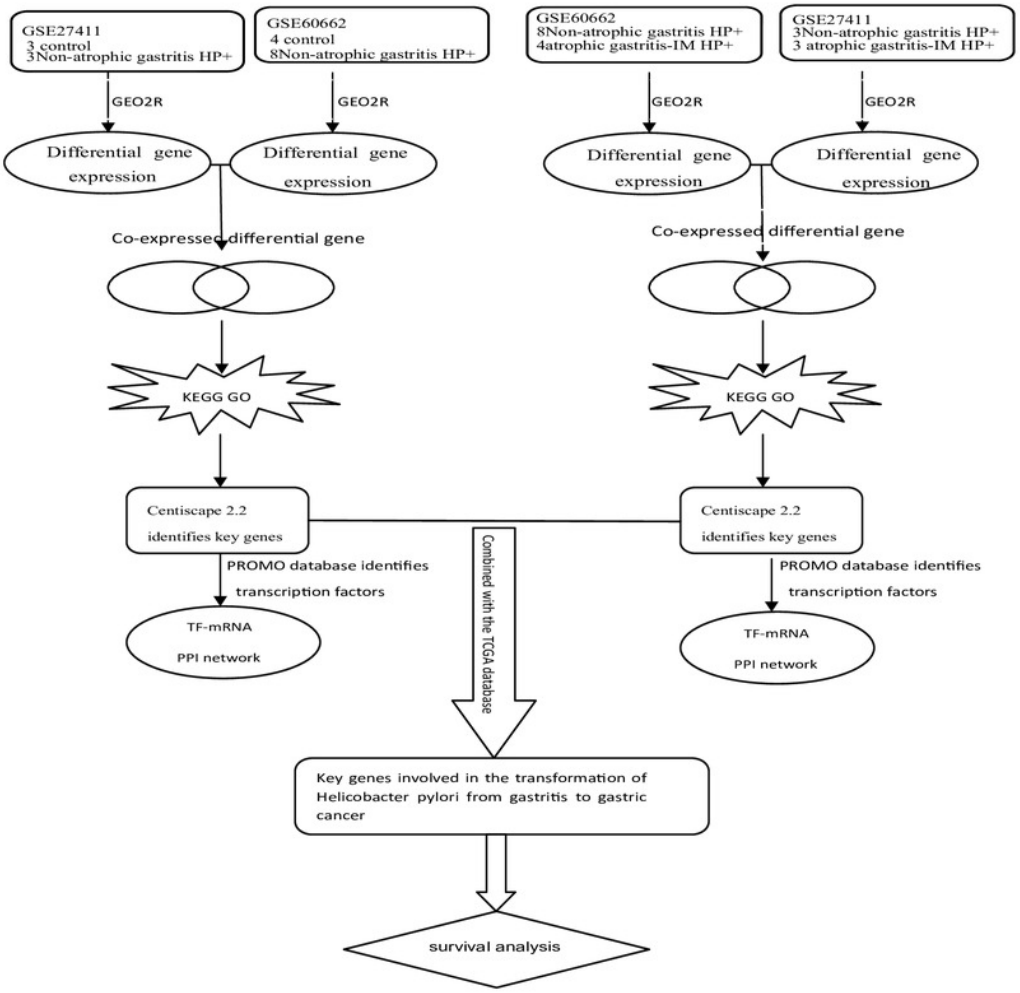

Figure 2

The functions and pathway analysis of DEGs in *H.pylori*-associated NAG.

Results of GO and KEGG enrichment analysis of the 111 genes between Hp^--No and Hp^+-NAG.Ordinate is the enriched functions and pathway, and abscissa is the ratio of the DEGs. The area of the displayed graphic is proportional to the number of genes assigned to the term and the color corresponds to the *P* value.

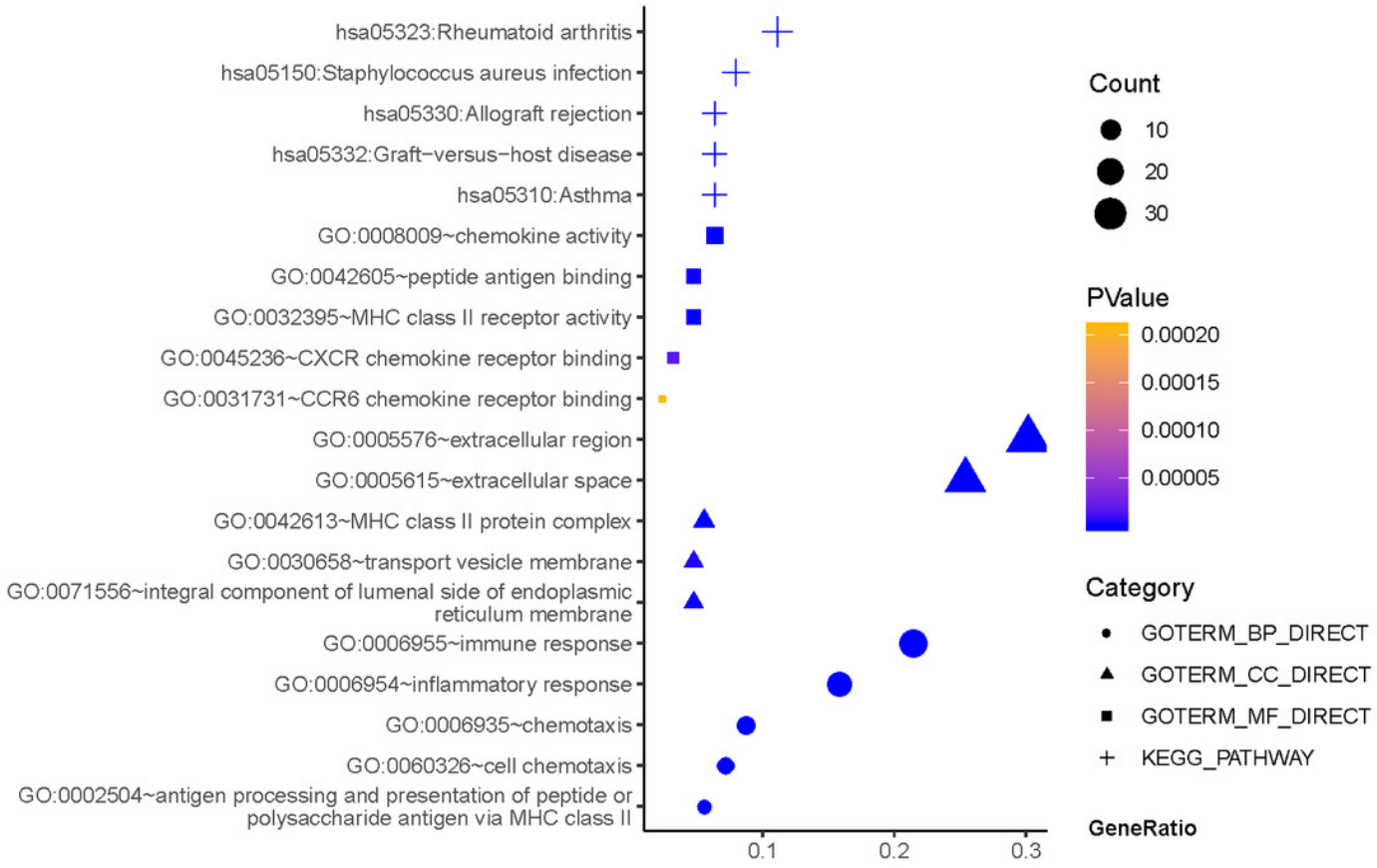

Figure 3

The functions and pathway analysis of DEGs in *H.pylori*-associated AG-IM.

Results of GO and KEGG enrichment analysis of the 385 genes between Hp^+-NAG and Hp^+-AG-IM. Ordinate is the enriched functions and pathway, and abscissa is the ratio of the DEGs. The area of the displayed graphic is proportional to the number of genes assigned to the term and the color corresponds to the *P* value.

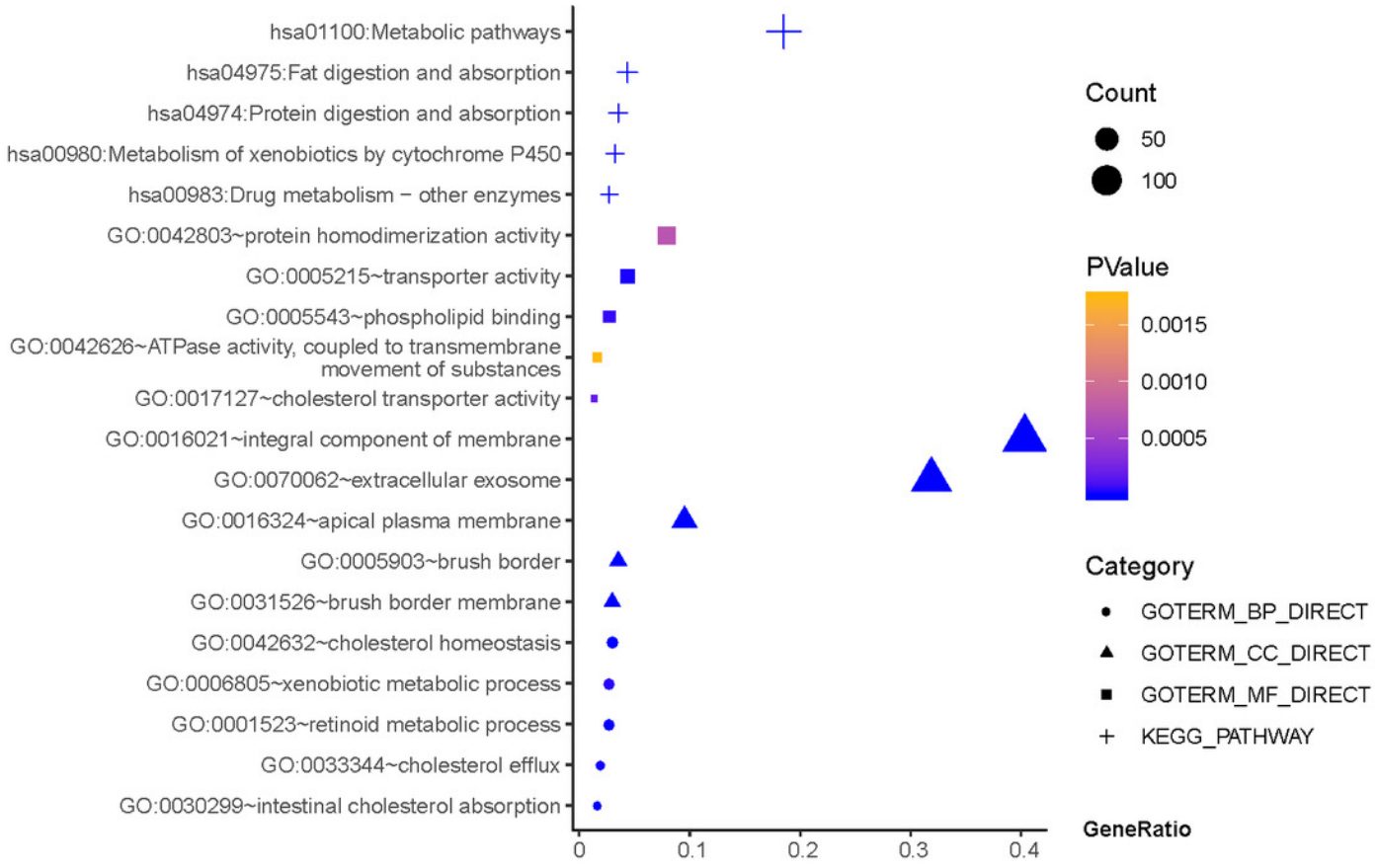

Figure 4

DEGs-TFs regulatory network involved in Hp^{+}-NAG.

The network consisting of 83 nodes and 370 edges was extracted from the whole PPI network. Key nodes in the network are highlighted in different colors and shape: blue dots corresponds to the up-regulated gene, yellow dots corresponds to the down-regulated gene and red square indicate TFs, size increasing with degree. Red edges indicate transcriptional regulatory relationships.

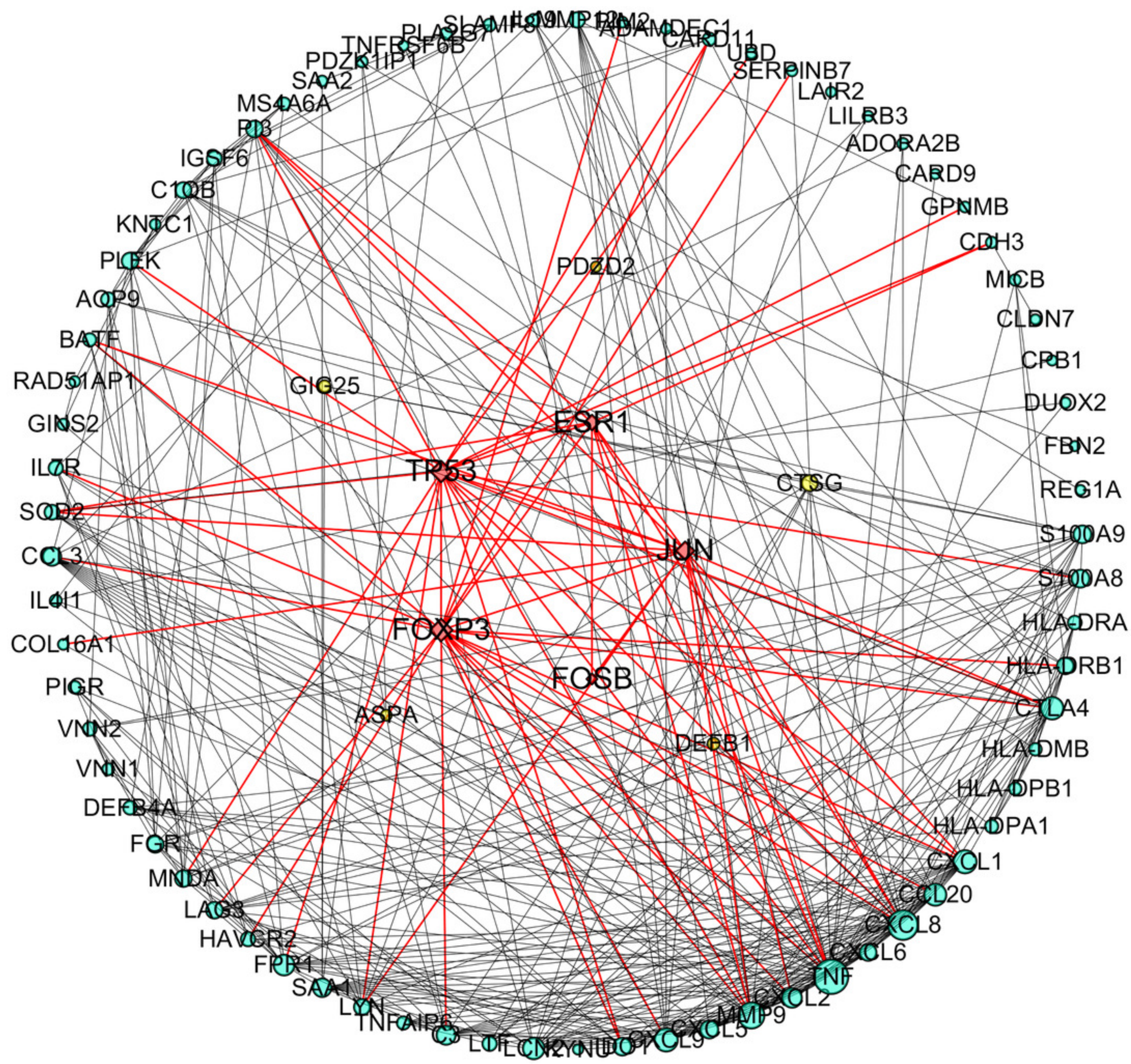
ESR1
TP53
JUN
FOXP3
FOSB
CTSG
GIG25
PDZD2
ASPA
DEFB1
CARD11
UBD
SERPINB7
LAIR2
LILRB3
ADORA2B
CARD9
GPNMB
CDH3
MICB
CLDN7
CPB1
DUOX2
FBN2
REG1A
S100A9
S100A8
HLA-DRA
HLA-DRB1
CTLA4
HLA-DMB
HLA-DPB1
HLA-DPA1
CXCL1
CCL20
CXCL8
CXCL6
CXCL2
MMP9
CXCL5
CXCL9
LCN2
TNFAIP6
LYN
SAA1
FPR1
HAVCR2
LAG3
MNDA
FGR
DEFB4A
VNN1
VNN2
PIGR
COL16A1
IL4I1
CCL3
SOD2
IL7R
GINS2
RAD51AP1
BATF
AQP9
PLEK
KNTC1
C1QB
IGSF6
PI3
MS4A6A
SAA2
PDZK1IP1
TNFRSF6B

Figure 5

DEGs-TFs regulatory network involved in Hp^{+}-AG-IM.

The giant network consisting of 166 nodes and 741 edges was extracted from the whole PPI network. Key nodes in the giant network are highlighted in different colors and shape: blue dots corresponds to the up-regulated gene, yellow dots corresponds to the down-regulated gene and red square indicate TFs, size increasing with degree. Red edges indicate transcriptional regulatory relationships.

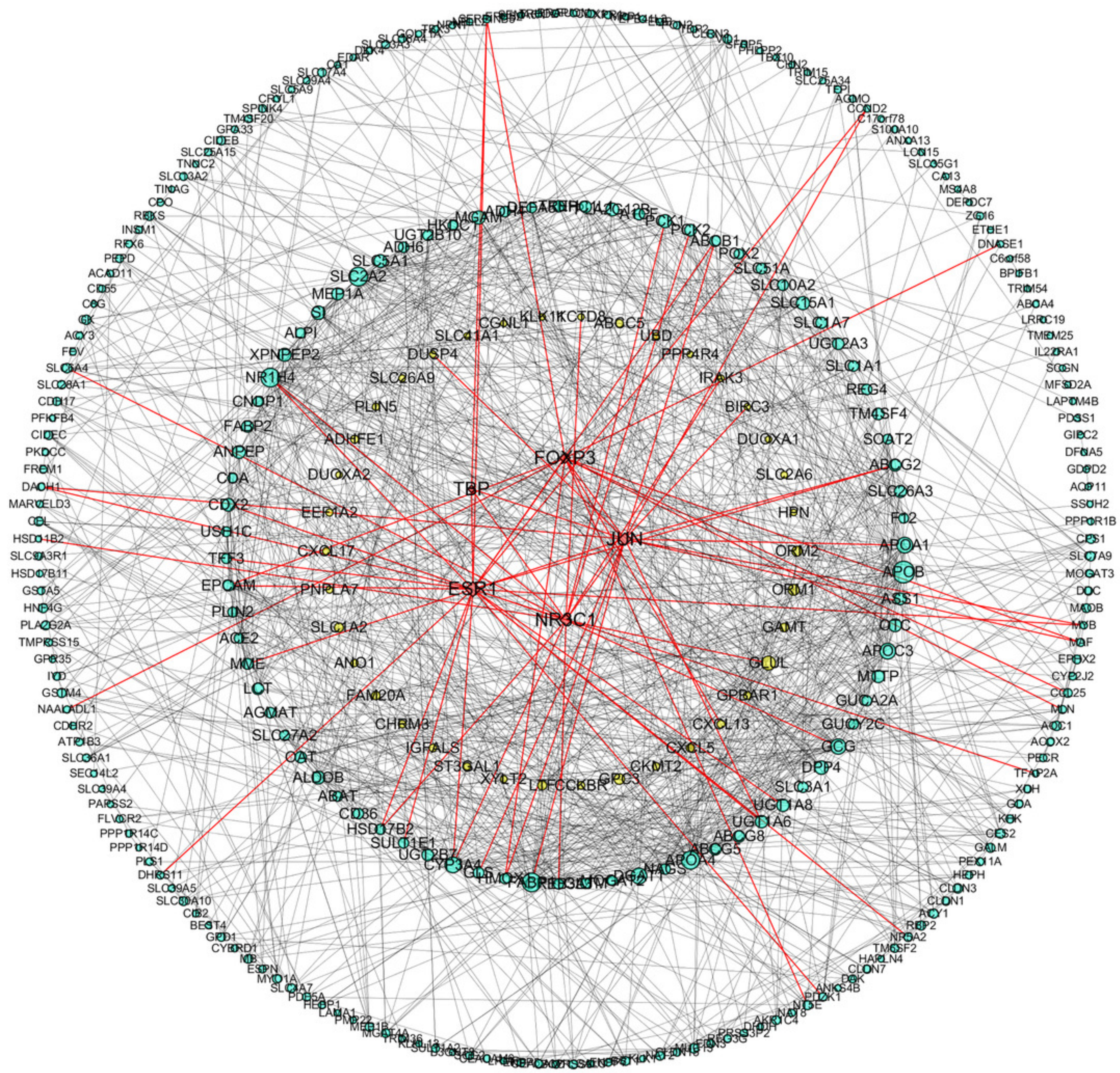
FOXP3
TBP
JUN
ESR1
NR3C1
XPNPEP2
NR1H4
CNDP1
FABP2
ANPEP
CDA
CDX2
TFF3
EPCAM
PLIN2
ACE2
MME
LCT
AGMAT
SLC27A2
OAT
ALDOB
ABAT
CD36
HSD17B2
SULT1E1
CYP3A4
SLC26A9
PLIN5
DUSP4
DUOXA2
EEF1A2
CXCL17
PNPLA7
SLC1A2
ANO1
FAM20A
CHRM3
IGFALS
ST3GAL1
CKMT2
GPC3
CXCL13
CXCL5
GPBAR1
GLUL
GAMT
ORM1
ORM2
HPN
SLC2A6
DUOXA1
BIRC3
IRAK3
PPP1R4
UBD
ABCC5
SLC1A7
SLC1A1
REG4
TM4SF4
SOAT2
ABCG2
SLC26A3
APOA1
APOB
ASS1
OTC
APOC3
MTTP
GUCA2A
GUCY2C
GCG
DPP4
SLC3A1
UGT1A8
UGT1A6
ABCG8
ABCG5
SLC5A1
SLC2A2
MEP1A
ADH6
UGT2B10
SLC51A
SLC10A2
SLC15A1
UGT2A3
SI
ALPI

Figure 6

The Kaplan-Meier survival curve of 882 gastric cancer (GC) patients based on FOXP3(A), JUN(B) in Kmplot software.

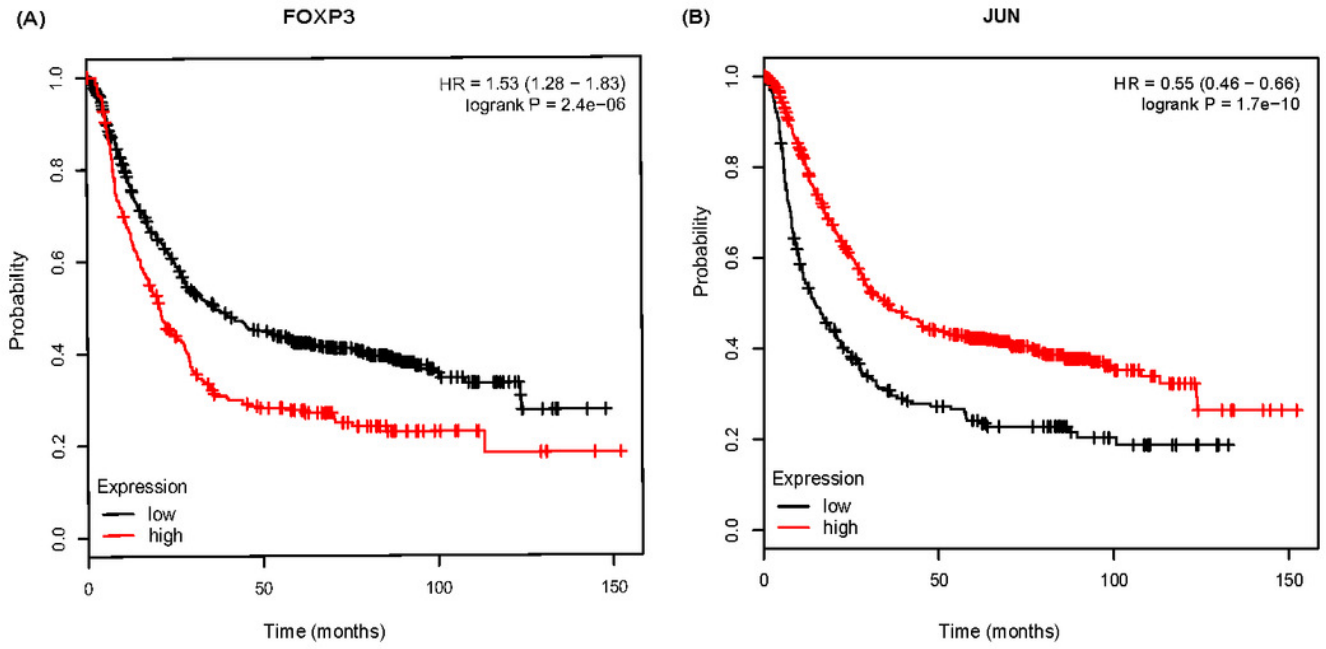

Table 1(on next page)

Information about cases of GC and normal in TCGA

1 **Table 1. Information about cases of GC and normal in TCGA**

	Total	*H. pylori* (+) GC	Normal tissue
Average age(year)	66.20	63.61	68.78
Gender			
Male	35(70%)	13(26%)	22(44%)
female	15(30%)	5(10%)	10(20%)
Race category			
Asian	9(18%)	2(4%)	7(14%)
White	28(56%)	11(22%)	17(34%)
Others	13(26%)	5(10%)	8(16%)
Cancer type			
Intestinal type	10(55.6%)	10(55.6%)	0
Diffuse type	7(38.9%)	7(38.9)	0
Not otherwise specified	1(5.5%)	1(5.5%)	0
Disease stage			
Stage I	1(5.5%)	1(5.5%)	0
Stage II	4(22.3%)	4(22.3%)	0
Stage III	11(61.1%)	11(61.1%)	0
Stage IV	2(11.1%)	2(11.1%)	0

2

Table 2(on next page)

key DEGs involved in *Hp*$^{+}$-NAG

1 **Table 2. key DEGs involved in *Hp*$^{+}$-NAG**

Gene Name	Degree	Betweenness	Closeness
TNF	**38**	**1332.769**	**0.00885**
CXCL8	**30**	**427.4495**	**0.008**
MMP9	**24**	**839.5228**	**0.007519**
CXCL9	**22**	**639.6061**	**0.007692**
CXCL1	**20**	**115.4851**	**0.007353**
CCL20	**20**	**117.0942**	**0.007299**
LCN2	**20**	**694.3924**	**0.007194**
CTLA4	**19**	**285.9193**	**0.007194**
FPR1	**18**	**180.3151**	**0.00641**
CXCL2	**15**	**27.46918**	**0.006897**
C3	**15**	**109.0515**	**0.006579**

2

Table 3(on next page)

key DEGs involved in Hp^{+}-AG-IM

1 **Table 3. key DEGs involved in Hp^{+}-AG-IM**

2

Name	Degree	Betweenness	Closeness
APOB	**42**	**6129.43**	**0.001372**
SLC2A2	**34**	**5955.64**	**0.001355**
FABP1	**33**	**6241.264**	**0.001381**
APOA4	**32**	**5013.443**	**0.001297**
NR1H4	**32**	**2808.839**	**0.001316**
APOC3	**26**	**1692.219**	**0.001279**
DGAT1	**26**	**2460.765**	**0.001225**
APOA1	**25**	**982.1235**	**0.001224**
GCG	**24**	**7636.908**	**0.001323**
CYP3A4	**24**	**4010.06**	**0.001267**
DPP4	**20**	**2526.871**	**0.001318**
GLUL	**20**	**5105.64**	**0.001255**
SI	**18**	**3977.184**	**0.001304**
XPNPEP2	**16**	**1895.259**	**0.001215**
MGAM	**15**	**4515.383**	**0.001241**
SLC15A1	**15**	**2751.226**	**0.001235**