Screening of independent prognostic long non-coding RNA for gastric cancer in TCGA

Qiang Zhang
Zhongshan Hospital Xiamen University
Qiongyun Chen
Zhongshan Hospital Xiamen University
Yinyin Lv
Zhongshan Hospital Xiamen University
Xuan Dong
Zhongshan Hospital Xiamen University
Xiaoqing Huang
Zhongshan Hospital Xiamen University
Changsheng Yan
Zhongshan Hospital Xiamen University
Hongzhi Xu (xuhongzhi@xmu.edu.cn)
Zhongshan Hospital Xiamen University https://orcid.org/0000-0002-2725-5359

Primary research

Keywords: TCGA, Gastric cancer, LncRNA, ENSG00000224363, Prognosis

Posted Date: July 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-37844/v1

Abstract

Background The global incidence of gastric cancer (GC) ranks the fourth among cancers and its 5-year survival is less than 25%. LncRNAs are vital regulators involved in pathological processes of cancer. It is urgent to screen the prognostic lncRNA in GC. Method Expression file and clinical data of GC were downloaded from TCGA. Differentially expressed lncRNAs were calculated by edger R package, followed by the prognosis analysis. COX analysis was conducted to compute the independent factor of GC. Potential signaling pathways that the screened lncRNAs enriched in were evaluated by gene set enrichment analysis (GSEA). At last, Pearson analysis was conducted to predict the possible mechanism of lncRNA in GC process. Result ENSG00000224363 was an unfavorable prognostic factor to OS (overall survival) and DFS (disease-free survival) of GC as COX regression analyzed. GSEA analysis indicated that ENSG00000224363 may regulate cell cycle, apoptosis and autophagy of GC cells. Conclusion LncRNA ENSG00000224363 is overexpressed in GC, serving as an independent unfavorable prognostic factor.

Background

Globally, the incidence of gastric cancer (GC) ranks the fourth among cancers. The number of Asian GC patients accounts for 60% of the total. In particular, the number of Chinese GC patients is on the top place of Asian cases, and its 5-year survival is less than 25.0%. Therefore, it is of great significance to find novel diagnostic and prognostic biomarkers of GC, aiming to effectively improve the survival rate (1, 2, 3).

Long non-coding RNA (lncRNA) is a non-coding RNA molecule with over 200 nucleotides in length (4). LncRNA is widely expressed in various organisms, but it cannot be translated to proteins. However, it is capable of transcriptionally, post-transcriptionally or translationally regulating gene expressions. Recent studies have shown the multiple functions of lncRNAs (5) in epigenetic modification (6, 7), transcriptional regulation (8, 9), RNA editing (10), protein translation (11) and other important cellular processes. In addition, lncRNAs are also crucial in tumor cell activities like cell maintenance and differentiation (12), tumor proliferation (13), tumor metastasis (14), drug resistance (15), etc.

SP1-induced LINC00673 is reported to be an oncogene in GC by interacting with LSD1 and EZH2 (16). HOTAIR promotes GC metastasis through downregulating PCBP1 (17). Besides, TUG1, PVT1, HOXA11-AS and GAPLINC have been reported to participate in GC development (18, 19, 20). The ENCODE project and GENCODE annotation have revealed the prevalence of thousands of lncRNAs. However, only few of them have been assigned with biological functions (21, 22). Therefore, screening prognostic lncRNAs in GC is of significance.

The Cancer Genome Atlas (TCGA) project is originally composed of the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI), which is free to the public. Clinical data and gene expression profile of 39 kinds of tumors involving 10,000 patients are collected in TCGA database. Using GC profile downloaded from TCGA, we aim to screen functional lncRNAs that can be used as prognostic factor of GC.

Methods

Data collection

Using the R package (Bioconductor/TCGAbiolinks), RNA-Seq raw count, as well as clinical data of GC patients (age, gender, tumor grade, TNM staging, OS and DFS) were downloaded from TCGA database (https://tcga-data.nci.nih.gov/tcga/). A total of 407 expression files containing 32 normal tissues and 375 tumor tissues were collected. Among them, OS was recorded from 373 patients and DFS was recorded from 303 patients.

Downloaded datasets containing complete clinical data of GC patients were subjected to survival analysis or correlation analysis. Based on the median level of ENSG00000224363 in GC samples, they were classified to high-expression group (> median of ENSG00000224363) and low-expression group (≤ median of ENSG00000224363).

Gene set enrichment analysis (GSEA)

Datasets (c2.cp.kegg.v5.2.symbols.gmt and c5.bp.v5.2.symbols.gmt) were obtained from molecular signatures database (MsigDB) on the GSEA website. Enrichment analysis was conducted by default-weighted enrichment using the GSEA 2.2.3 software. The number of random combination was set at 1,000 times.

Statistical methods

Statistical Product and Service Solutions 22.0 (SPSS 22.0) was used for statistical analysis. Differentially expressed genes in downloaded TCGA datasets were calculated using the Edger package. $\chi 2$ test and Fisher's exact probability test were used for analyzing the correlation between ENSG00000224363 and pathology of GC patients. Kaplan-Meier (K-M) and Log-rank tests were adopted for survival analysis. Univariate and multivariate Cox proportional hazard models were introduced to analyze risk factors of GC survival. Cor package was used to calculate Pearson correlation of all genes with ENSG00000224363. $p<0.05$ indicated a statistically significant difference. Significantly enriched gene sets were judged as per gene sets with $p<0.05$ and the false discovery rate (FDR) <0.25 in GSEA.

Results

2480 lncRNAs were up-regulated and 707 were down-regulated in GC

A total of 50455 genes were detected to be expressed in GC tissues from the downloaded TCGA database, in which 14,464 were lncRNAs (fold change of cut-off value ≥2 and p <0.05). Particularly, 2480 up-regulated lncRNAs and 707 lncRNAs were obtained (Fig 1A).

48 prognostic lncRNAs in GC

Survival R package was utilized for screening prognostic lncRNAs of DFS and OS of GC. Firstly, GC patients were divided to two groups based on the median expression of alternative lncRNAs. Prognosis curve was drawn using the K-M method. LncRNAs with log-rank *p* value <0.05 were output. Finally, there were 234 lncRNAs screened out to be the prognostic factors for OS, of which 48 lncRNAs were prognostic factors for DFS (data were not all showed, Fig 1B, 1C, 1D and 1E).

ENSG00000224363 was an independent unfavorable prognostic factor for DFS and OS of GC

The univariate Cox analyses of DFS and OS were conducted in the 48 screened lncRNAs, respectively. Eleven lncRNAs were associated with OS and 16 were associated with DFS (Table 1). Only 4 lncRNAs were both associated with DFS and OS of GC, which were subjected to the multivariate Cox analyses of DFS, in which 2 were associated with DFS of GC (Table 2-5). At last, these 2 lncRNAs were enrolled in multivariate Cox analyses of OS and only lncRNA ENSG00000224363 was obtained, which was an independent prognostic risk factor for DFS and OS of GC (Table 6- 7). Chi-square test showed that ENSG00000224363 was associated with lymph node metastasis of GC (Table 8). However, ENSG00000224363 was not correlated to age, tumor grade, TNM staging and other indicators of GC patients.

ENSG00000224363 could predict DFS and OS of GC

To analyze the prognostic potential of ENSG00000224363 in GC, we divided GC patients in the TCGA database into 14 groups according to tumor stage, grade, tumor remnant, depth of tumor local infiltration, lymph node metastasis, gender and age. Correlation between ENSG00000224363 expression with DFS and OS in each group was calculated, respectively. Among them, ENSG00000224363 was a risk factor for DFS in the female group (Fig 2A), male group (Fig 2B), low-grade group (Fig 2C), high-grade group (Fig 2D), early-stage group (Fig 2E), advanced-stage group (Fig 2F), no distant metastasis group (Fig 2G), no lymph node metastasis group (Fig 2H), lymph node metastasis group (Fig 2I), no tumor residual group (Fig 2J), local tumor deep infiltration group (Fig 2K) and older group (Fig 2L). In addition, ENSG00000224363 expression was a risk factor for OS in the female group (Fig 3A), male group (Fig 3B), older group (Fig 3C), no distant metastasis group (Fig 3D), high-grade group (Fig 3E), lymph node metastasis group (Fig 3F), no tumor residual group (Fig 3G), local tumor deep infiltration group (Fig 3H) and advanced-stage group (Fig 3I).

ENSG00000224363 mainly regulated cell cycle, apoptosis and autophagy of GC

Subsequently, potential biological signaling that ENSG00000224363 enriched in was analyzed by KEGG and GO analyses using GSEA software. KEGG results showed that ENSG00000224363 mainly regulated cell apoptosis (Fig 4A), cell cycle (Fig 4B), DNA replication (Fig 4C), and Wnt (Fig 4G), P53 (Fig 4F), mTOR (Fig 4E) and ErbB pathways (Fig 4D). GO analysis results showed that ENSG00000224363 mainly regulated cell apoptosis pathway (Fig 4H-J), cell cycle (Fig 4K) and autophagy (Fig 4L) in GC.

Correlation between ENSG00000224363 and genes involved in GC progression

Correlation between the whole genome and ENSG00000224363 was analyzed using Cor R package. It is found that cyclin dependent kinase 3 (CDK3) (Fig 5A), CDK15 (Fig 5B), cyclin dependent kinase-like 3 (CDKL3) (Fig 5C) and CDKL4 (Fig 5D) were positively correlated with ENSG00000224363 expression. However, cyclin dependent kinase inhibitor 1A (CDKN1A) (Fig 5E) and CDKN3 (Fig 5F) were negatively correlated with ENSG00000224363 expression. In addition, ENSG00000224363 was positively correlated with cell apoptosis inhibitor molecules [Caspase12 (Fig 5G) and Caspase14 (Fig 5H)], invasiveness and metastasis molecules [matrix metalloproteinase-21 (MMP-21) (Fig 5L) and MMP26 (Fig 5M)], and key factors of ErbB (Fig 5I), mitogen-activated protein kinase (MAPK) (Fig 5J and 5K) and Wnt (Fig 5N-P). pathways.

Discussion

LncRNA is generally transcribed in eukaryotic cells with barely or no protein-encoding ability (23, 24). It regulates gene expressions in the form of RNA through pre-transcriptional, transcriptional, and post-transcriptional level (4, 25). Recent studies have shown that lncRNA is involved in many important processes, such as X chromosome silencing, genomic imprinting, chromatin modification, transcriptional activation, transcriptional interference and intracranial transport. LncRNA is also closely related to tumors and non-neoplastic diseases (26, 27, 28, 29). The present study showed that differentially expressed lncRNA was related to the occurrence, development, invasiveness, metastasis and prognosis of GC, which may be used as a diagnostic marker and therapeutic target (30, 31, 32).

Dysregulated lncRNAs were firstly analyzed in the present study between normal gastric tissues and GC tissues in TCGA. There were 3187 lncRNAs to be analyzed. To identify the relationship between dysregulated lncRNAs and prognosis of GC, K-M analysis of DFS and OS was introduced. There were 48 lncRNAs identified to be associated with both OS and DFS of GC. Among them, 47 were unfavorable factors and the remaining were protective ones.

Cox hazard rate model was widely used to assess the clinical outcome of patients since 1972. It has the advantage to analyze the prognostic values of multiple factors (33).

After the prognostic lncRNAs were obtained, they were enrolled in the univariate Cox analysis of OS and DFS. Analysis data showed that lncRNA ENSG00000224363 was an independent prognostic risk factor for DFS and OS of GC. Its high level predicted an earlier recurrence and worse outcome of GC patients. Later, chi-square test revealed that ENSG00000224363 was positively associated with lymph node metastasis.

To assess the prognostic potential of ENSG00000224363 in GC, patients were divided into 14 groups. The results revealed that ENSG00000224363 predicted DFS in 12 groups and OS in 9 groups. GSEA analysis indicated that the main function of ENSG00000224363 was enriched in cell cycle, apoptosis and autophagy of GC (34, 35, 36).

Correlation analysis showed that ENSG00000224363 was positively correlated with key tumor-driving genes and negatively correlated with tumor-suppressor genes. For example, the CDK family, which are key regulators in promoting cell cycle and modulating transcription (37, 38), were positively correlated with ENSG00000224363 expression. Besides, CDKN1A (p21), an inhibitor of CDK family, can arrest cell cycle arrest and eventually inhibit cell growth (39, 40). In addition, the caspase family members that are capable of inhibiting cell apoptosis, were positively correlated with ENSG00000224363 (41, 42). Invasiveness is a vital trigger for tumor progression (43, 44). Expression level of MMP family was in accordance with ENSG00000224363. Meanwhile, expression levels of key regulators in the ErbB4, MAPK family and Wnt family were also coordinated with ENSG00000224363 (45, 46, 47).

This study for the first time demonstrated that lncRNA ENSG00000224363 was up-regulated in GC and it was an independent prognostic factor for DFS and OS of GC. It also revealed the possible mechanisms of ENSG00000224363 in regulating GC process. However, *in vitro* experiments are lacked, and our findings should be validated at the cytological level in the future.

Conclusion

LncRNA ENSG00000224363 is up-regulated in GC, serving as an independent unfavorable prognostic factor.

Abbreviations

TCGA: The Cancer Genome Atlas

GSEA: Gene set enrichment analysis

DFS: disease free survival

OS: overall survival

LncRNA: long noncoding RNA

NHGRI: the National Human Genome Research Institute

NCI: the National Cancer Institute

KEGG: Kyoto Encyclopedia of Genes and Genomes

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 81370591 and 81770558)

Author information

Affiliations

Department of Gastroenterology, Zhongshan Hospital, Xiamen University, Xiamen, Fujian, China

Institute for Microbial Ecology, School of Medicine, Xiamen University, Xiamen, Fujian, China

Qiang Zhang, Qiongyun Che ,Yinyin Lv, Xuan Dong, Changsheng Yan, Hongzhi Xu

Department of Chinese Traditional Medicine, Zhongshan Hospital, Xiamen University, Xiamen, Fujian, China

Xiaoqing Huang

Contributions

QZ and QC performed the experiments. YL and XD analyzed the data. XH and CY wrote the paper. HX supervised the whole experimental work and revised the manuscript. All authors reviewed and approved the manuscript

Corresponding author

Correspondence to Hongzhi Xu

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians. (2011) 61(2):69–90. doi: 10.3322/caac.20107.
2. Villanueva MT. Combination therapy: update on gastric cancer in East Asia. Nature reviews Clinical oncology. 2011;8(12):690. doi:10.1038/nrclinonc.2011.171.
3. Lin Y, Ueda J, Kikuchi S, Totsuka Y, Wei WQ, Qiao YL, Inoue M. Comparative epidemiology of gastric cancer between Japan and China. World journal of gastroenterology. 2011;17(39):4421–8. doi:10.3748/wjg.v17.i39.4421.
4. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136(4):629–41. doi:10.1016/j.cell.2009.02.006.
5. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annual review of biochemistry. (2012) 81:145–166. doi: 10.1146/annurev-biochem-051410-092902.
6. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008;322(5902):750–6. doi:10.1126/science.1163045.
7. Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science. 2011;331(6013):76–9. doi:10.1126/science.1197349.
8. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature. 2007;445(7128):666–70. doi:10.1038/nature05519.
9. Mariner PD, Walters RD, Espinoza CA, Drullinger LF, Wagner SD, Kugel JF, et al. Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Molecular cell. 2008;29(4):499–509. doi:10.1016/j.molcel.2007.12.013.
10. Peters NT, Rohrbach JA, Zalewski BA, Byrkett CM, Vaughn JC. RNA editing and regulation of Drosophila 4f-rnp expression by sas-10 antisense readthrough mRNA transcripts. Rna. 2003;9(6):698–710. doi:10.1261/rna.2120703.
11. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature. 2012;491(7424):454–7. doi:10.1038/nature11508.
12. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature. 2011;477(7364):295–300. doi:10.1038/nature10398.
13. Coccia EM, Cicala C, Charlesworth A, Ciccarelli C, Rossi GB, Philipson L, et al. Regulation and expression of a growth arrest-specific gene (gas5) during growth, differentiation, and development. Molecular cellular biology. 1992;12(8):3514–21. doi:10.1128/mcb.12.8.3514.
14. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010;464(7291):1071–6. doi:10.1038/nature08975.

15. Tsang WP, Kwok TT. Riboregulator H19 induction of MDR1-associated drug resistance in human hepatocellular carcinoma cells. Oncogene. 2007;26(33):4877–81. doi:10.1038/sj.onc.1210266.
16. Huang M, Hou J, Wang Y, Xie M, Wei C, Nie F, et al. Long Noncoding RNA LINC00673 Is Activated by SP1 and Exerts Oncogenic Properties by Interacting with LSD1 and EZH2 in Gastric Cancer. Molecular therapy: the journal of the American Society of Gene Therapy. 2017;25(4):1014–26. doi:10.1016/j.ymthe.2017.01.017.
17. Zhang ZZ, Shen ZY, Shen YY, Zhao EH, Wang M, Wang CJ, et al. HOTAIR Long Noncoding RNA Promotes Gastric Cancer Metastasis through Suppression of Poly r(C)-Binding Protein (PCBP) 1. Molecular cancer therapeutics. 2015;14(5):1162–70. doi:10.1158/1535-7163.MCT-14-0695.
18. Hu Y, Wang J, Qian J, Kong X, Tang J, Wang Y, et al. Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. Cancer research. 2014;74(23):6890–902. doi:10.1158/0008-5472.CAN-14-0686.
19. Zhang E, He X, Yin D, Han L, Qiu M, Xu T, et al. Increased expression of long noncoding RNA TUG1 predicts a poor prognosis of gastric cancer and regulates cell proliferation by epigenetically silencing of p57. Cell death disease. 2016;7:e2109. doi:10.1038/cddis.2015.356.
20. Kong R, Zhang EB, Yin DD, You LH, Xu TP, Chen WM, et al. Long noncoding RNA PVT1 indicates a poor prognosis of gastric cancer and promotes cell proliferation through epigenetically regulating p15 and p16. Molecular cancer. 2015;14:82. doi:10.1186/s12943-015-0355-8.
21. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74. doi:10.1038/nature11247.
22. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome research. 2012;22(9):1760–74. doi:10.1101/gr.135350.111.
23. Ye J, Renault VM, Jamet K, Gilson E. Transcriptional outcome of telomere signalling. Nature reviews Genetics. 2014;15(7):491–503. doi:10.1038/nrg3743.
24. Shay JW, Wright WE. Telomerase therapeutics for cancer: challenges and new directions. Nature reviews Drug discovery. 2006;5(7):577–84. doi:10.1038/nrd2081.
25. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. Cancer cell. 2016;29(4):452–63. doi:10.1016/j.ccell.2016.03.010.
26. Chen YG, Satpathy AT, Chang HY. Gene regulation in the immune system by long noncoding RNAs. Nature immunology. 2017;18(9):962–72. doi:10.1038/ni.3771.
27. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiological reviews. 2016;96(4):1297–325. doi:10.1152/physrev.00041.2015.
28. Huarte M. The emerging role of lncRNAs in cancer. Nature medicine. 2015;21(11):1253–61. doi:10.1038/nm.3981.
29. Satpathy AT, Chang HY. Long noncoding RNA in hematopoiesis and immunity. Immunity. 2015;42(5):792–804. doi:10.1016/j.immuni.2015.05.004.

30. Yang X, Gao L, Guo X, Shi X, Wu H, Song F, et al. A network based method for analysis of lncRNA-disease associations and prediction of lncRNAs implicated in diseases. PloS one. 2014;9(1):e87797. doi:10.1371/journal.pone.0087797.
31. Mei D, Song H, Wang K, Lou Y, Sun W, Liu Z, et al. Up-regulation of SUMO1 pseudogene 3 (SUMO1P3) in gastric cancer and its clinical association. Medical oncology. 2013;30(4):709. doi:10.1007/s12032-013-0709-2.
32. Yang F, Xue X, Bi J, Zheng L, Zhi K, Gu Y, et al. Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. Journal of cancer research clinical oncology. 2013;139(3):437–45. doi:10.1007/s00432-012-1324-x.
33. Cox DR. Regression models and Life tables. Journal of the Royal Statistical Society, Series B. (1972) 20:187–220. doi: 10.1007/978-1-4612-4380-9_37.
34. Brown JS, O'Carrigan B, Jackson SP, Yap TA. Targeting DNA Repair in Cancer: Beyond PARP Inhibitors. Cancer discovery. 2017;7(1):20–37. doi:10.1158/2159-8290.CD-16-0860.
35. Freed-Pastor W, Prives C. Targeting mutant p53 through the mevalonate pathway. Nature cell biology. 2016;18(11):1122–4. doi:10.1038/ncb3435.
36. Galluzzi L, Bravo-San Pedro JM, Demaria S, Formenti SC, Kroemer G. Activating autophagy to potentiate immunogenic chemotherapy and radiation therapy. Nature reviews Clinical oncology. 2017;14(4):247–58. doi:10.1038/nrclinonc.2016.183.
37. Malumbres M. Cyclin-dependent kinases. Genome biology. 2014;15(6):122. doi:10.1186/gb4184.
38. Pitts TM, Davis SL, Eckhardt SG, Bradshaw-Pierce EL. Targeting nuclear kinases in cancer: development of cell cycle kinase inhibitors. Pharmacology therapeutics. 2014;142(2):258–69. doi:10.1016/j.pharmthera.2013.12.010.
39. Liu G, Lozano G. p21 stability: linking chaperones to a cell cycle checkpoint. Cancer cell. 2005;7(2):113–4. doi:10.1016/j.ccr.2005.01.019.
40. Weiss RH. p21Waf1/Cip1 as a therapeutic target in breast and other cancers. Cancer cell. 2003;4(6):425–9. doi:10.1016/S1535-6108(03)00308-8.
41. Esteban-Fernandez de Avila B, Ramirez-Herrera DE, Campuzano S, Angsantikul P, Zhang L, Wang J. Nanomotor-Enabled pH-Responsive Intracellular Delivery of Caspase-3: Toward Rapid Cell Apoptosis. ACS nano. 2017;11(6):5367–74. doi:10.1021/acsnano.7b01926.
42. Xu L, Li X, Chu ES, Zhao G, Go MY, Tao Q, et al. Epigenetic inactivation of BCL6B, a novel functional tumour suppressor for gastric cancer, is associated with poor survival. Gut. 2012;61(7):977–85. doi:10.1136/gutjnl-2011-300411.
43. Yin Y, Grabowska AM, Clarke PA, Whelband E, Robinson K, Argent RH, Tobias A, Kumari R, Atherton JC, Watson SA. Helicobacter pylori potentiates epithelial:mesenchymal transition in gastric cancer: links to soluble HB-EGF, gastrin and matrix metalloproteinase-7. Gut. 2010;59(8):1037–45. doi:10.1136/gut.2009.199794.
44. Crawford HC, Krishna US, Israel DA, Matrisian LM, Washington MK, Peek RM Jr. Helicobacter pylori strain-selective induction of matrix metalloproteinase-7 in vitro and within gastric mucosa.

Gastroenterology. 2003;125(4):1125–36. doi:10.1016/s0016-5085(03)01206-x.

45. Chaturvedi R, Asim M, Piazuelo MB, Yan F, Barry DP, Sierra JC, et al. Activation of EGFR and ERBB2 by Helicobacter pylori results in survival of gastric epithelial cells with DNA damage. Gastroenterology. 2014;146(7):1739–51 e1714. doi:10.1053/j.gastro.2014.02.005.
46. Mitsuno Y, Yoshida H, Maeda S, Ogura K, Hirata Y, Kawabe T, et al. Helicobacter pylori induced transactivation of SRE and AP-1 through the ERK signalling pathway in gastric cancer cells. Gut. 2001;49(1):18–22. doi:10.1136/gut.49.1.18.
47. Chang HR, Nam S, Kook MC, Kim KT, Liu X, Yao H, et al. HNF4alpha is a therapeutic target that links AMPK to WNT signalling in early-stage gastric cancer. Gut. 2016;65(1):19–32. doi:10.1136/gutjnl-2014-307918.

Tables

Table 1
The Univariate COX analysis of LncRNAs in DFS and OS.

Gene	Univariate analysis of DFS		Univariate analysis of OS	
	HR	P value	HR	P value
RP11-289H16.1	1.492	0.001	1.346	0.001
RP11-59D5__B.2	1.153	0.020	1.161	0.004
ABCA9-AS1	1.170	0.016	1.112	0.044
POT1-AS1	1.209	0.048	1.112	0.044

Table 2
Univariate and multivariate COX regression analyses ABCA9-AS1 for DFS of patients in study cohort.

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.525	0.993	0.972–1.015	0.380	0.990	0.967–1.013
Gender(Female vs. Male)	0.001	2.441	1.418–4.203	0.011	2.048	1.179–3.558
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.091	0.828	0.666–1.031	0.251	0.875	0.696–1.099
Tumor Grade (G1 vs. G2 vs. G3)	0.319	1.265	0.797–2.008	0.695	1.100	0.683–1.771
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	4.063	2.019–8.175	0.003	2.858	1.421–5.750
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Female)	0.029	2.224	1.087–4.547	0.689	1.091	0.711–1.676
(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Male)	0.062	1.366	0.984–1.897			
Distant metastasis(No vs. Yes)	0.521	1.477	0.449–4.858	0.677	1.333	0.345–5.145
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)	< 0.000	1.480	1.200-1.825	0.004	1.378	1.107–1.716
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)	0.503	1.105	0.826–1.478	0.554	0.981	0.609–1.305
ABCA9-AS1(High vs. Low)	0.016	1.170	1.030–1.328	0.045	1.145	1.003–1.308

Table 3
Univariate and multivariate COX regression analyses POT1-AS1 for DFS of patients in study cohort.

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.525	0.993	0.972–1.015	0.764	0.996	0.973–1.020
Gender(Female vs. Male)	0.001	2.441	1.418–4.203	0.010	2.072	1.191–3.606
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.091	0.828	0.666–1.031	0.180	0.855	0.681–1.075
Tumor Grade (G1 vs. G2 vs. G3)	0.319	1.265	0.797–2.008	0.474	1.195	0.734–1.943
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	4.063	2.019–8.175	0.003	2.817	1.407–5.639
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Female)	0.029	2.224	1.087–4.547	0.431	1.211	0.752–1.951
(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Male)	0.062	1.366	0.984–1.897			
Distant metastasis(No vs. Yes)	0.521	1.477	0.449–4.858	0.867	1.120	0.295–4.248
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)	< 0.000	1.480	1.200-1.825	0.002	1.400	1.132–1.732
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)	0.503	1.105	0.826–1.478	0.594	0.916	0.665–1.264
POT1-AS1(High vs. Low)	0.048	1.209	1.002–1.459	0.069	1.191	0.986–1.439

Table 4
Univariate and multivariate COX regression analyses RP11-59D5__B.2 for DFS of patients in study cohort.

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.525	0.993	0.972–1.015	0.519	0.992	0.970–1.016
Gender(Female vs. Male)	0.001	2.441	1.418–4.203	0.006	2.140	1.239–3.696
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.091	0.828	0.666–1.031	0.138	0.840	0.667–1.058
Tumor Grade (G1 vs. G2 vs. G3)	0.319	1.265	0.797–2.008	0.341	1.270	0.776–2.079
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	4.063	2.019–8.175	0.006	2.688	1.335–5.414
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Female)	0.029	2.224	1.087–4.547	0.456	1.199	0.744–1.932
(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Male)	0.062	1.366	0.984–1.897			
Distant metastasis(No vs. Yes)	0.521	1.477	0.449–4.858	0.800	1.192	0.306–4.648
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)	< 0.000	1.480	1.200-1.825	0.004	1.377	1.107–1.713
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)	0.503	1.105	0.826–1.478	0.417	0.876	0.637–1.206
RP11-59D5__B.2(High vs. Low)	0.02	1.153	1.023-1.300	0.085	1.109	0.986–1.248

Table 5
Univariate and multivariate COX regression analyses ENSG00000224363 for DFS of patients in study cohort

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.53	0.99	0.972–1.015	0.49	0.99	0.969–1.015
Gender(Female vs. Male)	0.00	2.44	1.418–4.203	0.01	2.10	1.213–3.627
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.09	0.83	0.666–1.031	0.66	0.95	0.744–1.207
Tumor Grade (G1 vs. G2 vs. G3)	0.32	1.27	0.797–2.008	0.60	1.14	0.706–1.833
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	4.06	2.019–8.175	0.01	2.46	1.220–4.970
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Female)	0.03	2.22	1.087–4.547	0.53	1.14	0.754–1.733
(Stage I vs. Stage II vs. Stage III vs. Stage IV)(Male)	0.06	1.37	0.984–1.897			
Distant metastasis(No vs. Yes)	0.52	1.48	0.449–4.858	0.62	1.40	0.378–5.167
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)	< 0.000	1.48	1.200-1.825	0.00	1.38	1.112–1.703
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)	0.50	1.11	0.826–1.478	0.31	0.83	0.569–1.195
ENSG00000224363(High vs. Low)	0.00	1.49	1.184–1.879	0.01	1.41	1.105–1.802

Table 6
Univariate and multivariate COX regression analyses ABCA9-AS1 for OS of patients in study cohort

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.044	1.02	1.001–1.040	0.017	1.025	1.004–1.046
Gender(age ≤ 60)	0.637	1.207	0.553–2.637	0.496	1.071	0.879–1.304
(age > 60)	0.042	1.687	1.018–2.793			
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.638	0.957	0.795–1.151	0.045	1.538	1.009–2.344
Tumor Grade (G1 vs. G2 vs. G3)	0.053	1.453	0.996–2.121	0.232	1.259	0.863–1.836
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	3.293	1.969–5.507	0.007	2.200	1.238–3.911
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)	< 0.000	1.780	1.377–2.301	0.226	1.258	0.868–1.824
Distant metastasis(No vs. Yes)(age ≤ 60)(Female)	0.017	11.161	1.527–81.571	0.073	2.017	0.937–4.343
(age ≤ 60)(Male)	0.082	3.880	0.842–17.876			
(age > 60)(Female)	0.004	7.200	1.906–27.192			
(age > 60) (Male)	0.558	1.534	0.366–6.435			
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)(age ≤ 60)(Female)	0.124	1.622	0.876–3.003	0.001	1.358	1.135–1.626
(N0 vs. N1 vs. N2 vs. N3)(age ≤ 60) (Male)	0.045	1.621	1.010–2.603			
(N0 vs. N1 vs. N2 vs. N3)(age > 60) (Female)	0.004	1.809	1.204–2.718			
(N0 vs. N1 vs. N2 vs. N3)(age > 60) (Male)	0.036	1.282	1.016–1.616			
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)(age ≤ 60)	0.849	1.049	0.644–1.708	0.650	1.080	0.776–1.503

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
(T1 vs. T2 vs. T3 vs. T4)(age > 60) (Female)	0.185	1.495	0.825–2.710			
(T1 vs. T2 vs. T3 vs. T4)(age > 60) (Male)	0.012	1.566	1.105–2.218			
ABCA9-AS1(High vs. Low)	0.044	1.112	1.003–1.234	0.480	1.041	0.932–1.163

Table 7
Univariate and multivariate COX regression analyses ENSG00000224363 for OS of patients in study cohort.

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
Age(≤ 60 vs. >60)	0.044	1.02	1.001–1.040	0.029	1.023	1.002–1.043
Gender(age ≤ 60)	0.637	1.207	0.553–2.637	0.126	1.392	0.911–2.126
(age > 60)	0.042	1.687	1.018–2.793			
Location(Antrum vs. Cardia vs. Fundus vs. Gastroesophageal Junction)	0.638	0.957	0.795–1.151	0.392	1.090	0.895–1.329
Tumor Grade (G1 vs. G2 vs. G3)	0.053	1.453	0.996–2.121	0.144	0.324	0.909–1.929
Residual tumor(R0 vs. R1 vs.R2)	< 0.000	3.293	1.969–5.507	0.009	2.089	1.198–3.643
Stage(Stage I vs. Stage II vs. Stage III vs. Stage IV)	< 0.000	1.780	1.377–2.301	0.001	1.587	1.211–2.079
Distant metastasis(No vs. Yes)(age ≤ 60)(Female)	0.017	11.161	1.527–81.571	0.210	1.690	0.744–3.839
(age ≤ 60)(Male)	0.082	3.880	0.842–17.876			
(age > 60)(Female)	0.004	7.200	1.906–27.192			
(age > 60) (Male)	0.558	1.534	0.366–6.435			
Lymph node metastasis(N0 vs. N1 vs. N2 vs. N3)(age ≤ 60)(Female)	0.124	1.622	0.876–3.003	0.236	1.157	0.909–1.474
(N0 vs. N1 vs. N2 vs. N3)(age ≤ 60) (Male)	0.045	1.621	1.010–2.603			
(N0 vs. N1 vs. N2 vs. N3)(age > 60) (Female)	0.004	1.809	1.204–2.718			
(N0 vs. N1 vs. N2 vs. N3)(age > 60) (Male)	0.036	1.282	1.016–1.616			
Depth of invasion(T1 vs. T2 vs. T3 vs. T4)(age ≤ 60)	0.849	1.049	0.644–1.708	0.682	1.073	0.767–1.501

Variables	Univariate analysis			Multivariate analysis		
	P value	HR	95% CI	P value	HR	95% CI
(T1 vs. T2 vs. T3 vs. T4)(age > 60) (Female)	0.185	1.495	0.825–2.710			
(T1 vs. T2 vs. T3 vs. T4)(age > 60) (Male)	0.012	1.566	1.105–2.218			
ENSG00000224363(High vs. Low)	0.001	1.346	1.121–1.617	0.004	1.317	1.090–1.592

Table 8
The relation between ENSG00000224363 and clinical data.

Variables	Number of cases	expression level		*P* value
		Low	High	
Age				
≤ 60	121	62	59	0.767
> 60	250	124	126	
Location				
Antrum	138	69	69	0.970
Cardia	48	24	24	
Fundus	130	64	66	
Gastroesophageal Junction	41	22	19	
Gender				
Femal	134	73	61	0.210
Male	241	115	126	
Grade				
G1 + G2	147	79	68	0.241
G3	219	104	115	
Residual tumor				
No	298	152	146	0.250
Yes	30	12	18	
Stage				
I/II	164	89	75	0.135
III/IV	188	87	101	
Distant metastasis				
No	330	162	168	0.151
Yes	25	16	9	
Lymph node metastasis				
N0 + N1	208	114	94	0.037
N2 + N3	149	65	84	

Variables	Number of cases	expression level		*P* value
		Low	High	
Depth of invasion				
T1 + T2	99	57	42	0.083
T3 + T4	268	127	141	

Figures

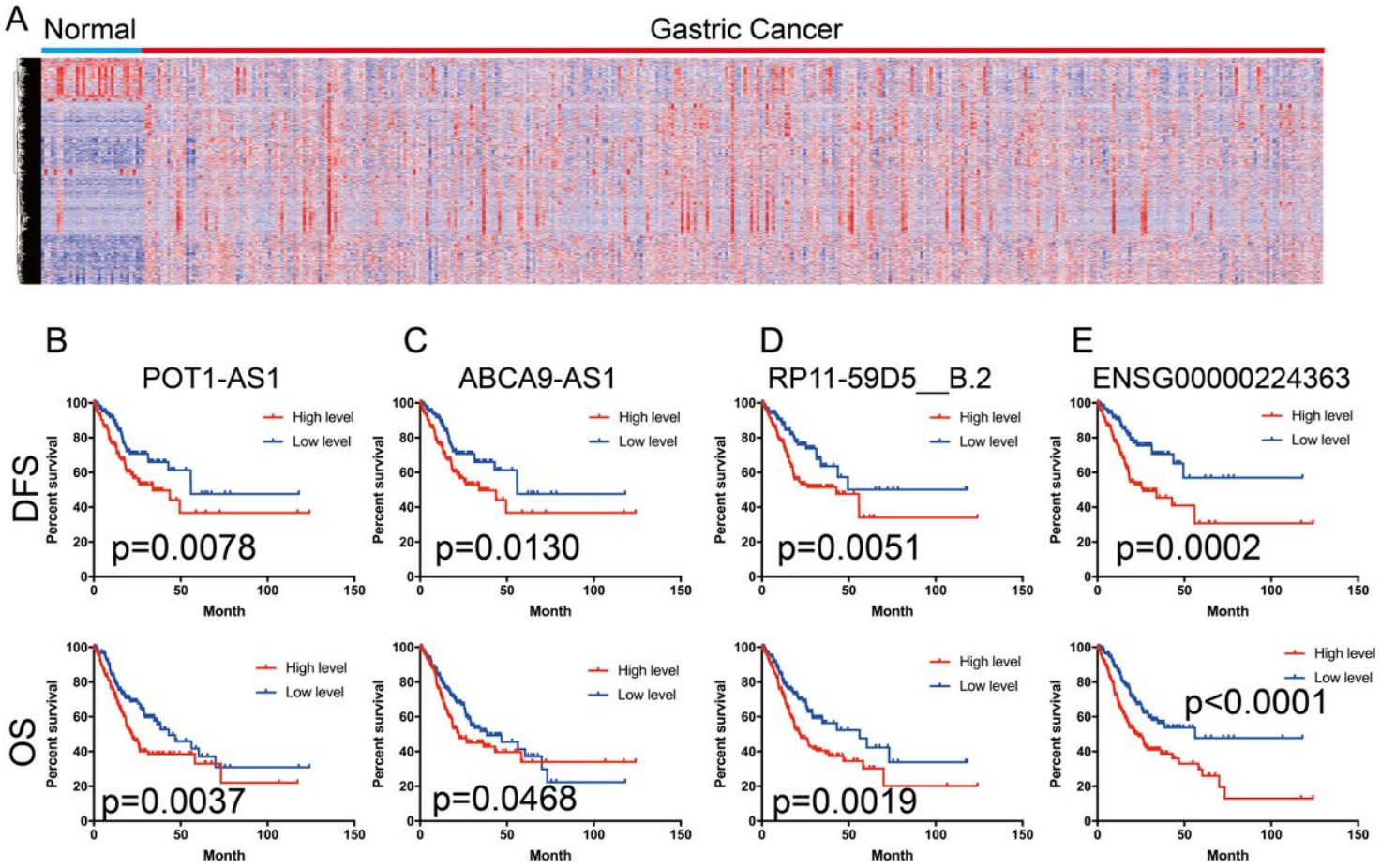

Figure 1

The dis-regulated expressed lncRNAs in gastric cancer. A. The heatmap of dis-regulated lncRNAs in gastric cancer. B-E. The prognostic curve (up, DFS; down, OS) of POT1-AS1, ABCA9-AS1, RP11-59D5_B.2, ENSG00000224363.

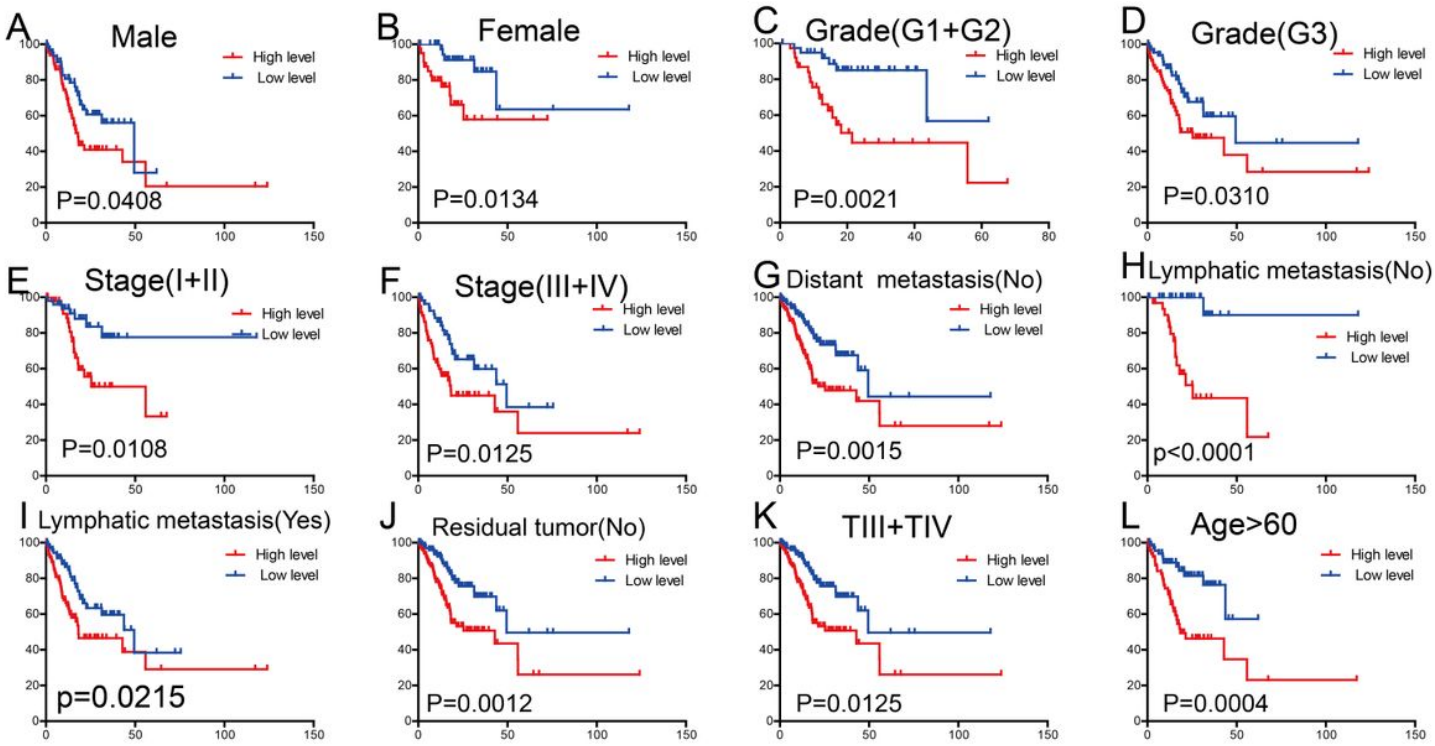

Figure 2

ENSG00000224363 could predict DFS of gastric cancer. A. The DFS prognostic curve of male group. B. The DFS prognostic curve of female group. C. The DFS prognostic curve of low grade (G1+G2) group. D. The DFS prognostic curve of high grade (G3) group. E. The DFS prognostic curve of early stage (stage I+ stage II) group. F. The DFS prognostic curve of late stage (stage III+ stage IV) group. G. The DFS prognostic curve of no distant metastasis group. H. The DFS prognostic curve of no lymphatic metastasis group. I. The DFS prognostic curve of lymphatic metastasis group. J. The DFS prognostic curve of no residual tumor group. K. The DFS prognostic curve of deep invasion group. L. The DFS prognostic curve of old age group.

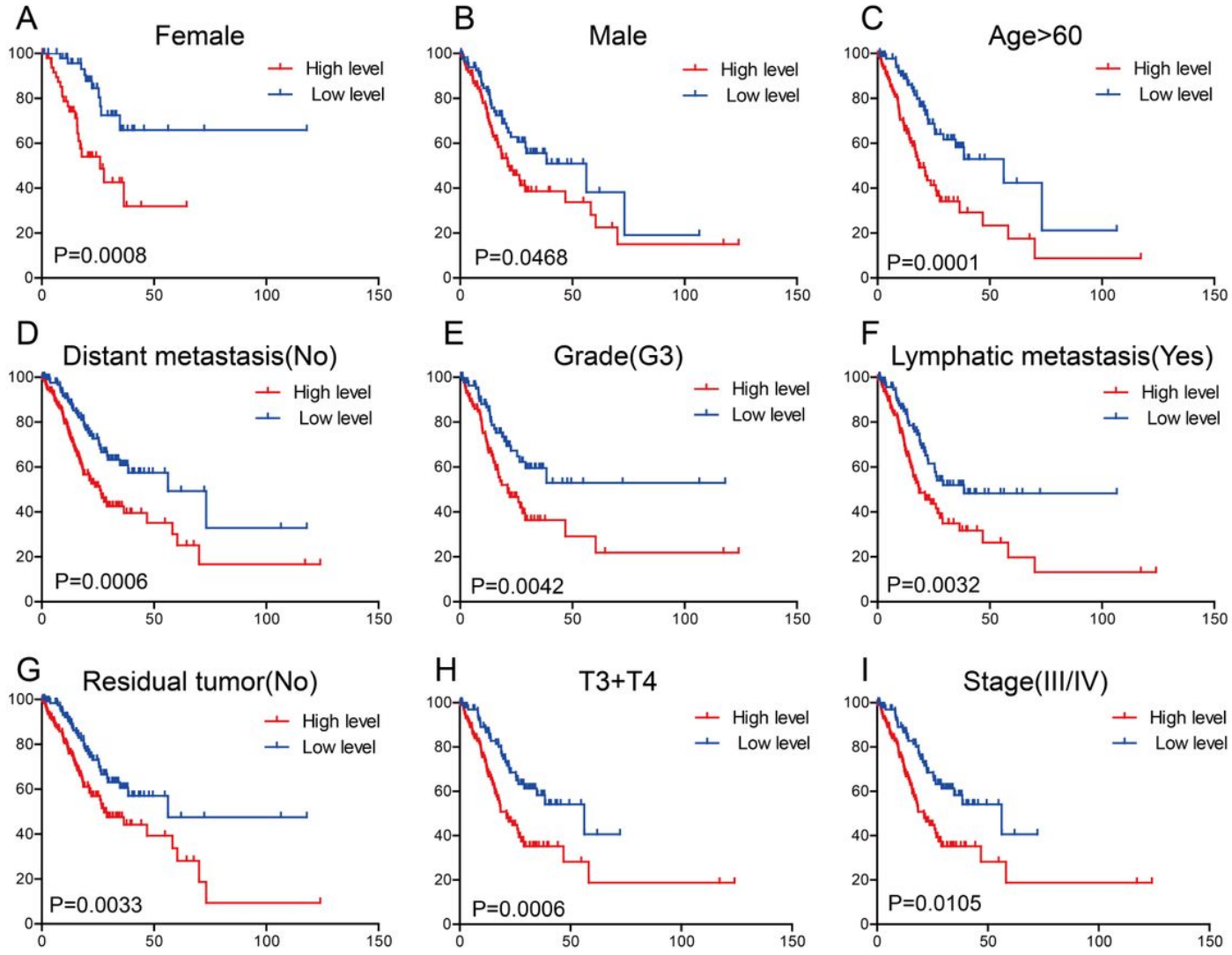

Figure 3

ENSG00000224363 could predict OS of gastric cancer. A. The OS prognostic curve of female group. B. The OS prognostic curve of male group. C. The OS prognostic curve of old age (age>60 years) group. D. The OS prognostic curve of no distant metastasis group. E. The OS prognostic curve of high grade (G3) group. F. The OS prognostic curve of lymphatic metastasis group. G. The OS prognostic curve of no residual tumor group. H. The OS prognostic curve of deep invasion group. I. The OS prognostic curve of late stage (stage III+ stage IV) group.

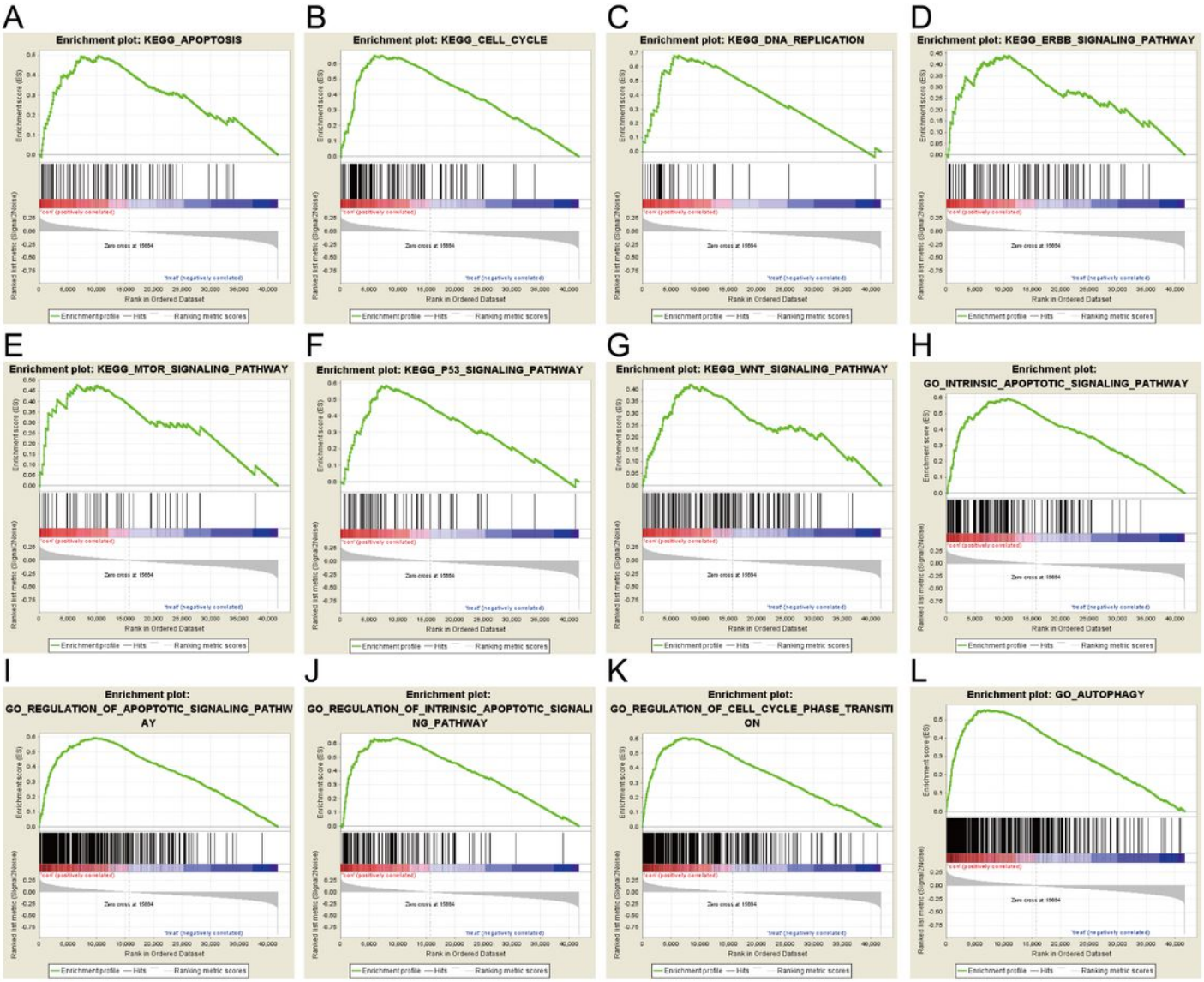

Figure 4

ENSG00000224363 regulated cell cycle, apoptosis and autophagy. (A, H, I, J) KEGG and GO analysis showed that ENSG00000224363 could regulate cell apoptosis. (B, C, K) KEGG and GO analysis showed that ENSG00000224363 could regulate cell cycle. D. KEGG analysis showed that ENSG00000224363 could regulate ErbB signal pathway. E. KEGG analysis showed that ENSG00000224363 could regulate mTOR signal pathway. F. KEGG analysis showed that ENSG00000224363 could regulate P53 signal pathway. G. KEGG analysis showed that ENSG00000224363 could regulate WNT signal pathway. L. GO analysis showed that ENSG00000224363 could regulate autophagy.

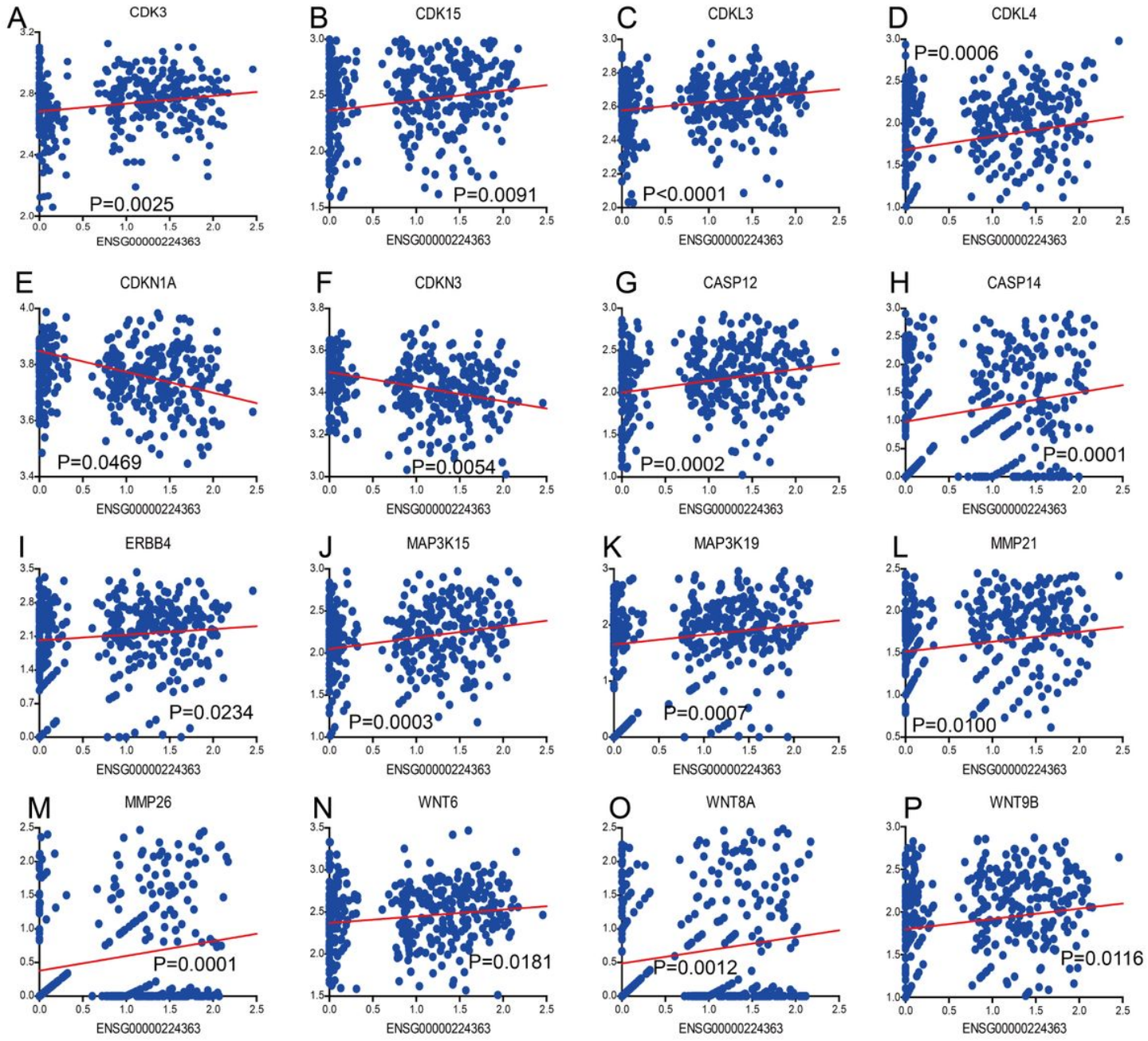

Figure 5

The correlation between key factors and ENSG00000224363. (A, B, C and D) ENSG00000224363 was positive correlated with CDK family (CDK3, CDK5, CDKL3, CDKL4). (E and F) ENSG00000224363 was negative correlated with CKI family (CDKN1A, CDKN3). (G and H) ENSG00000224363 was positive correlated with caspase family (CASP12, CASP14). I. ENSG00000224363 was positive correlated with ERBB4. (J and K) ENSG00000224363 was positive correlated with MAPK family (MAP3K15, MAP3K19). (L and M) ENSG00000224363 was positive correlated with MMP family (MMP21, MMP26). (N, O and P) ENSG00000224363 was positive correlated with WNT family (WNT6, WNT8A, WNT9B)