Identification of Key Biomarkers Associated with Survival and Prognosis in Colon Adenocarcinoma

Yanjie Zhou (zhouyj1@jyrmyy.com)
The affiliated Jiangyin Hospital of Southeast University Medical School https://orcid.org/0000-0001-7227-0673

Lu Jiang
Wuxi Center for Disease Control and Prevention

Jiang Lin
The affiliated Jiangyin Hospital of Southeast University Medical School

Wendong Tang
The affiliated Jiangyin Hospital of Southeast University Medical School

Wenqian Jiang
The affiliated Jiangyin Hospital of Southeast University Medical School

Ke Wang
The affiliated Jingyin Hospital of Southeast University Medical School

Research Article

Keywords: Colorectal cancer, Biomarker, Prognostic, SPP1, GRP, GNGT1

Posted Date: April 26th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-444592/v1

Abstract

Background: Colorectal cancer (CRC) has a high rate of relapse and recurrence that result in poor prognosis and unsatisfactory outcomes. Colon adenocarcinoma (COAD) is the most prevalent type of CRC. It is crucial to identify novel molecular biomarkers for the early diagnosis, prognosis evaluation and disease monitoring of COAD.

Methods: Three gene expression profile data were downloaded from the Gene Expression Omnibus(GEO), and the differential expression genes(DEGs) were identified by GEO2R. Gene Ontology (GO) and KEGG pathway enrichment analysis were conducted by WebGestalt online tool. String database and Cytoscape software were used for protein–protein interaction (PPI) network construction and module analysis. The top 20 Hub Genes were screened from the PPI network using MCC algorithm on CytoHubba app of Cytoscape software, and were verified by ONCOMINE database then. The core genes affecting CRC prognosis were screened by GEPIA2 survival analysis web tool. Finally, the expression level and clinical indicators including core genes was analyzed by TCGA-COAD dataset.

Results: In total, 413 differentially expressed genes (DEGs) were identified, and the GO and KEGG enrichment analyses of DEGs were processed. After, the protein–protein interaction (PPI) network was constructed and 20 hub genes were identified. Furthermore, three core genes were selected via survival analysis . Finally, the diagnostic and prognostic value of these core genes was verified by clinical analysis of TCGA-COAD dataset.

Conclusion: SPP1, GRP and GNGT1 were all over-expressed in COAD, and may be regarded as novel diagnostic and prognostic biomarkers for COAD.

Background

Colorectal cancer (CRC) is one of the most common cancer types worldwide, which accounting for about 11% of all diagnosed cancer cases [1, 2]. Additionally, CRC is the second most lethal cancer worldwide [3]. Previous studies have indicated that the pathogenesis of CRC is a multi-step and multi-path evolution process in which chromosomal instability, dysregulation of oncogenes and tumor suppressor, epigenetic alterations, metabolic alterations, abnormal immune response occured [4, 5].Although numerous treatments, including surgery, chemotherapy radiotherapy, and targeted therapy, have been improved, nearly 55% of CRC patients eventually relapse and suffer from the recurrence [6], and the prognosis of patients is poor [7]. Colon adenocarcinoma (COAD) is the most prevalent type of CRC, the incidence and mortality of COAD was 10.2 and 9.2% [3, 8]. The exact mechanisms of CRC development are still poorly understood, therefore, it is crucial to better understand the exact mechanism of tumorigenesis and identify novel molecular biomarkers for early diagnosis, prognosis evaluation, disease monitoring. [9, 10]

The recently adopted high-throughput gene microarray technology has been extensively applied in several fields of medicine and biology, which allows us to share and explore various molecular mechanisms [11].

Furthermore, we can study the key genes and select potential molecular targets and diagnostic markers [12].

In this paper, 413 differentially expressed genes (DEGs) in CRC were identified by GEO2R. Gene Ontology (GO) and KEGG pathway enrichment analysis were conducted by WebGestalt online tool to determine the functional annotation and pathway of DEGs. At the same time, the protein–protein interaction (PPI) network was constructed and 20 hub genes were identified. Three core genes were finally selected by survival analysis, and the association between hub genes and its diagnostic as well as prognostic value in COAD was investigated(Fig. 1).

Methods

Microarray data

Gene Expression Omnibus(GEO) (http://www.ncbi.nlm.nih.gov/geo) functions as a public functional genomics database of high throughout gene expression data, chips and microarrays [13]. Three gene expression profiles in colorectal cancer and normal epithelium were downloaded from GEO, that is GSE21815, GSE35279 and GSE71187. Microarray data of these three datasets were all on account of GPL6480 Platforms, Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version). The GSE21815 dataset contained 132 colorectal cancer samples and 9 normal epithelium samples. The GSE35279 dataset contained 74 colorectal cancer samples and 5 normal epithelium samples. The GSE71187 dataset contained 47 Colorectal cancer tissue specimen and 12 normal colorectal tissue specimen.

Identification of DEGs

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r) is regarded as an interactive online tool designed to compare two or more datasets in a GEO series for the purpose of DEGs identification across experimental conditions. The DEGs between CRC tissues and normal epithelium tissues were identified using GEO2R with the P value < 0.05 which were considered of statistically significance. For the next step, we utilized the “dplyr” R package to identify the DEGs between CRC samples and normal epithelium samples. Adjusted P < 0.05 and |log fold change (FC)| > 2 were chosen as the cutoff threshold. The volcano plot and venn diagram containing these DEGs were drawn with the ”ggplot2” package in R 4.0.4.

Functional enrichment analysis

WEB-based GEne SeT AnaLysis Toolkit(WebGestalt) (http://www.webgestalt.org/) is an online biological information software tool that integrates functional categories derived from centrally and publicly curated databases as well as computational analyses [14]. We analyzed GO and KEGG pathway enrichment analyses for the DEGs by WebGestalt. GO terms enrichment analysis were categorized into biological process (BP), cellular component (CC), and molecular function (MF). Top 10 terms were selected according to FDR.

PPI network construction and module analysis

Search Tool for the Retrieval of Interacting Genes(STING)(http://string-db.org) online database was used to predict the PPI network which may further explain the mechanisms of the occurrence and progression of diseases [15]. By using STRING database, PPI network of DEGs was analyzed and an interaction with a combined score > 0.7 was recognized as statistical significance. The plug-in MCODE (Molecular Complex Detection) app of Cytoscape (an public bioinformatics software, version 3.7.1) is constructed for clustering a network based on topology to determine intensively connected regions [16]. The PPI network was plotted with the application of Cytoscape and the most significant module in the PPI network was narrowed down using MCODE with the following criteria: degree cutoff = 2, node score cutoff = 0.2, k-core = 2, max depth = 100.

Hub genes selection and analysis

The plug-in Cytohubba of Cytoscape is an APP provided with 11 topological analysis methods for ranking nodes in a PPI network by their network features [17]. In the present study, the top 20 hub genes were ranked according to the maximal clique centrality (MCC). Oncomine (http://www.oncomine.com) online database was applied for explore the mRNA expression levels of the hub genes in various kinds of cancers, including CRC [18].

Gene Expression Profilling Interactive Analysis (GEPIA2)(http://gepia2.cancer-pku.cn/#index) online tool was used for mRNA expression level and survival analysis of the hub genes [19]. The survival map of 20 hub genges was plotted by “Survival Analysis” function in GEPIA2. Accordingly, three core genes affecting CRC prognosis were screened.

The verification of the expression level of core genes and their diagnostic and prognostic value was conducted via analysising the TCGA-COAD dataset. The box plot and ROC curve were plotted by R 4.0.4.

Result

Identification of DEGs in colorectal cancer

Via GEO2R online tools, DEGs in three datasets (1357 DEGs in GSE35279, 1493 DEGs in GSE21815, and 1884 DEGs in GSE71187, respectively) were extracted after gene expression profile data processing and standardization with the cutoff standard of P value < 0.05 and |logFC|>2. The volcano plot and venn diagram were drawn with the “ggplot2” package in R 4.0.4. The overlapping DEGs among these three datasets contained 413 genes as shown in the Venn diagram (Fig. 2).

Enrichment analysis for DEGs

To elucidate the biological functions of the overlapping DEGs, we performed functional annotation and pathway enrichment analysis via WebGestalt online tool. Results indicated that the overlapping DEGs in biological process of GO enrichment were markedly associated with tissue development, tube morphogenesis, tissue morphogenesis, tube development, skeletal system development, anatomical

structure formation involved in morphogenesis, animal organ morphogenesis, circulatory system development, cell proliferation and epithelium development (Fig. 3A). As for molecular function of GO enrichment, DEGs were remarkably related to receptor regulator activity, receptor ligand activity, RNA polymerase II regulatory region sequence-specific DNA binding, RNA polymerase II regulatory region DNA binding, extracellular matrix structural constituent, transcription regulatory region sequence-specific DNA binding, sequence-specific double-stranded DNA binding, extracellular matrix structural constituent conferring tensile strength, hormone activity, transcription regulatory region DNA binding(Fig. 3B). In addition to cellular component, collagen-containing, extracellular matrix, basement membrane, extracellular matrix component, complex of collagen trimers, basolateral plasma membrane, plasma membrane region, endoplasmic reticulum lumen supramolecular polymer, supramolecular complex(Fig. 3C). Besides, signaling pathway analysis of KEGG demonstrated that those DEGs played pivotal roles in Wnt signaling pathway, AGE-RAGE signaling pathway in diabetic complications, Central carbon metabolism in cancer, Biosynthesis of amino acids, Breast cancer, Malaria, Gastric cancer, ECM-receptor interaction, TGF-beta signaling pathway, MicroRNAs in cancer (Fig. 3D).

PPI network construction and significant module identification

To evaluate the interrelationships among these DEGs, we first drew the network of DEGs in STRING, we set the maximum number of interacting bodies to 0 and used a confidence score of 0.7 as the cut-off criterion, the PPI network included 388 nodes and 334 edges(Fig. 4A). Additionally, the Molecular Complex Detection (MCODE) app was also employed to select modules of the PPI network in Cytoscape according to node score cut-off = 0.2, degree cut-off = 2, max.depth = 100, and k − core = 2. The top three functional clusters of modules were selected (module 1, MCODE score = 12.000; module 2, MCODE score = 6.333; module 3, MCODE score = 6.167) (Fig. 4B-4D).

KEGG pathway analysis of each module was performed by Enrichr (https://maayanlab.cloud/Enrichr/) online tool. The KEGG pathway analysis of module 1 indicated that these genes were involved in the Neuroactive ligand-receptor interaction, Chemokine signaling pathway, and Cytokine-cytokine receptor interaction; genes in module 2 enriched in Progesterone-mediated oocyte maturation and cell cycle; and genes in module 3 related to the Protein digestion and absorption, ECM-receptor interaction, and carbon metabolism.

Hub genes selection and validation

Based on querying STRING protein information from the public database, we constructed a PPI network of the top 20 hub genes according to the maximal clique centrality (MCC). The top 20 hub genes with maximal clique centrality were as follows: SST, PPBP, SPP1, GCG, INSL5, CXCL8, EDN3, CXCL11, PYY, P2RY1, HTR1D, GPR4, SAA1, IL6, PLCB1, APLN, GRP, ADRA2C, GNGT1, GAL(Fig. 5). Then, mRNA expressions of hub genes in 20 types of cancers were measured and compared to normal tissues by ONCOMINE database, and 15 hub genes were found to be over-expression in colorectal cancer (Fig. 6).

We used GEPIA2 online tool to identify the survival data of 20 hub genes and found that the expression levels of three hub genes, namely SPP1, GRP and GNGT1, were remarkably related to the survival of colon adenocarcinoma(COAD) patients(Fig. 7,8). Furthermore, we validated the expression level of these three hub genes in COAD and normal tissues via TCGA-COAD dataset, and explored their diagnostic value as well as clinical significance. The results showed that expression levels of the three genes were significantly increased in the COAD samples and the over-expression of these three genes was correlated with the TMN stage of COAD patients (Fig. 9–10).

Discussion

CRC is a great threat to people's health, which has been the 3rd most common cancer throughout the world, according to the global burden of disease study in 2018. COAD is the most prevalent type of CRC. In the last decade, increasingly researches have focused on the pathogenesis mechanisms of CRC and the therapeutic strategies have been developed. Nevertheless, CRC mortality and morbidity rates remain high due to postsurgical recurrence and metastasis of primary tumours. Therefore, it is of great importance to identify novel molecular targets for the early diagnosis, treatment, and prognosis of CRC.

Bioinformatics analysis is an effective method to discover the pathogenesis of cancer by studying the occurrence and development of cancer at the molecular level. In the present study, DEGs between CRC and non-cancerous tissues were obtained from three mRNA microarray datasets via GEO2R. The functions of the DEGs were identified by GO and KEGG enrichment analysis. The results demonstrated that the DEGs were enriched in the function items and pathways which were associated with progression and prognosis of CRC, such as the cell proliferation, extracellular matrix structural constituent, Wnt signaling pathway, Central carbon metabolism in cancer, ECM-receptor interaction and TGF-beta signaling pathway. The extracellular matrix (ECM) regulates tissue development and homeostasis, and its dysregulation have been appreciated as key drivers for both development and cancer progression [20, 21]. Aberrant Wnt/β-catenin signaling has often been reported in different cancers [22], particularly CRC [23], and this signaling cascade is central to carcinogenesis which has potential value as a therapeutic target in the treatment of CRC [24]. Transforming growth factor-beta (TGF-β) signaling is one of the important cellular pathways that play key roles for tissue maintenance [25]. In particular, it is important in the context of inflammation and tumorigenesis by modulating cell growth, differentiation, apoptosis, and homeostasis [25, 26].

Furthermore, we constructed a PPI network via STRING online tools and detected 20 hub genes by CytoHubba. The expression and diagnostic value of these hub genes were verified in the ONCOMINE database, and the survival analysis by TCGA database reveals the high expression of SPP1, GRP and GNGT1 were significantly associated with poor cancer survival rates.

The secreted phosphoprotein 1 (SPP1, also known as osteopontin), is an extracellular matrix chemokine-like phosphoglycoprotein that facilitates cell-matrix interaction which is overexpressed in various malignant neoplasms and plays a role in tumourigenesis and metastasis [27]. Wei et al. [28] found that

SPP1 overexpression led to enhanced anchorage-independent growth, cell migration and invasion in KRAS gene mutant cells and overexpression also induced PI3K signalling, expression of Snail and Matrix metallopeptidase 9 (MMP9), and suppressed the expression of E-cadherin in KRAS mutant cells. Zeng et al. [29] found the expression of SPP1 can activate activated Integrin β1/FAK/AKT pathway promotes ovarian cancer progression. Zhang et al. [30] found SPP1 could upregulate PD-L1 which mediates macrophage polarization and facilitates immune escape in lung adenocarcinoma. In our research, we inferred that SPP1 was enriched in ECM-receptor interaction, which was significantly associated with EMT pathway [31].

Gastrin-releasing peptide(GRP) is a small regulatory peptide with homology to bombesin [32]. GRP and its receptor play an important role in a multitude of physiological functions including sensory transmission, regulation of central autonomic pathways, thermoregulation, secretion of pituitary hormones, gastric and pancreatic secretion, and food intake and satiety [33–35].Increasing evidences demonstrated that GRP acts as a mitogen, morphogen and pro-angiogenic factor in certain cancers [36, 37]. Patel et al. [38] found that nonamidated peptides derived from the C terminus of pro-GRP are expressed in significant quantities in CRC cell lines and tumors and stimulate the proliferation of CRC cells and of normal colonic mucosa. Such peptides are attractive targets for novel CRC therapies. Ni et al. [39] found GRP was highly expressed in breast cancer patients with lymph node metastasis. Besides, among the patients with lymph node metastasis, the ones with higher expression of gastrin-releasing peptide had shorter survival time. Overexpression of gastrin-releasing peptide significantly enhanced cell invasiveness. Conversely, a knockdown of gastrin-releasing peptide through the short hairpin RNA approach remarkably reduced MCF-7 cell invasion. Additionally, our study reveals that the high expression of GRP significantly associated with poor cancer survival rates, and may be a novel diagnosis biomarker for colon cancer.

The transducing gamma subunit gene (GNGT1) encodes a member (gamma1) of the family of heterotrimeric G-protein gamma subunits that is specific to rod photoreceptors. [40] Alsaleem et al. [41] found reduced/loss E-cadherin expression was associated with differential expression genes including GNGT1 which regulats PIK3-AKT signaling pathways. Mucaki et al. [42] found GNGT1 is a gene that has an efficient response to chemotherapy by carboplatin. In the current study, GNGT1 was significantly upregulated and high mRNA expression of GNGT1 was associated with poor overall survival in CRC patients.

Although hub genes were verified with different data sources, the main limitation of our research is that our study is only at the level of bioinformatics analysis. It was probably because of SPP1 and GRP were secretory proteins, we couldn’t find positive results from the human protein atlas (THPA) database. Further experimental analysis, including using a cell model and CRC tissues from the clinic, and different experimental methods such as western blot and ELISA to validate the prediction is urgently needed.

Conclusion

In conclusion, with the integrated bioinformatics analysis for gene expression profiles from GEO database, we dug out three core molecules associated with the prognosis of COAD, including SPP1, GRP and GNGT1. These core genes were all over-expressed in COAD, and may be regarded as novel diagnostic and prognostic biomarkers for COAD.

Abbreviations

CRC
Colorectal cancer
COAD
Colon adenocarcinoma
READ
Rectum adenocarcinoma

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets analyzed for this study were obtained from Gene Expression Omnibus(GEO) (http://www.ncbi.nlm.nih.gov/geo), Oncomine (http://www.oncomine.com), The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) and Search Tool for the Retrieval of Interacting Genes (STRING) (http://string-db.org) .

Competing interests

The authors declare that they have no competing interests

Funding

The Wuxi science and Technology Development Fund (grant numbers NZ2019042)

The Jiangyin People's Hospital Specialized Reserve Fund

Authors' contributions

Yanjie Zhou and Lu Jiang designed this study. Wendong Tang and Yanjie Zhou collected data, analyzed the data in this study, and interpreted the findings. Wenqian Jiang drafted the manuscript. Ke Wang and Jiang Lin carried out data management and revised themanuscript. All authors reviewed the final version of the manuscript.

Acknowledgements

Thanks to the fund (Wuxi science and Technology Development Fund and Jiangyin People's Hospital Specialized Reserve Fund) for supporting this research.

References

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019;144:1941–53. doi:10.1002/ijc.31937.
2. Weitz J, Koch M, Debus J, Höhler T, Galle PR, Büchler MW. Colorectal cancer. Lancet. 2005;365:153–65. doi:10.1016/S0140-6736(05)17706-X.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424. doi:10.3322/caac.21492.
4. Aran V, Victorino AP, Thuler LC, Ferreira CG. Colorectal Cancer: Epidemiology, Disease Mechanisms and Interventions to Reduce Onset and Mortality. Clin Colorectal Cancer. 2016;15:195–203. doi:10.1016/j.clcc.2016.02.008.
5. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000;343:78–85. doi:10.1056/NEJM200007133430201.
6. Asano H, Kojima K, Ogino N, Fukano H, Ohara Y, Shinozuka N. Postoperative recurrence and risk factors of colorectal cancer perforation. Int J Colorectal Dis. 2017;32:419–24. doi:10.1007/s00384-016-2694-3.
7. Al Bandar MH, Kim NK. Current status and future perspectives on treatment of liver metastasis in colorectal cancer (Review). Oncol Rep. 2017;37:2553–64. doi:10.3892/or.2017.5531.
8. Masoomi H, Ziogas A, Lin BS, Barleben A, Mills S, Stamos MJ, Zell JA. Population-based evaluation of adenosquamous carcinoma of the colon and rectum. Dis Colon Rectum. 2012;55:509–14. doi:10.1097/DCR.0b013e3182420953.
9. Wu Y, Xu Y. Integrated bioinformatics analysis of expression and gene regulation network of COL12A1 in colorectal cancer. Cancer Med. 2020;9:4743–55. doi:10.1002/cam4.2899.
10. Dai G-P, Wang L-P, Wen Y-Q, Ren X-Q, Zuo S-G. Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. Oncol Lett. 2020;19:388–98. doi:10.3892/ol.2019.11068.

11. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021. doi:10.3322/caac.21660.
12. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol. 2006;24:971–83. doi:10.1038/nbt1235.
13. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41:D991-5. doi:10.1093/nar/gks1193.
14. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res. 2013;41:W77–83. doi:10.1093/nar/gkt439.
15. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45:D362–8. doi:10.1093/nar/gkw937.
16. Bader GD, Hogue CWV. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2. doi:10.1186/1471-2105-4-2.
17. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4:S11. doi:10.1186/1752-0509-8-S4-S11.
18. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia. 2004;6:1–6. doi:10.1016/s1476-5586(04)80047-2.
19. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019;47:W556–60. doi:10.1093/nar/gkz430.
20. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. EMBO Rep. 2014;15:1243–53. doi:10.15252/embr.201439246.
21. Walker C, Mojares E. Del Río Hernández A. Role of Extracellular Matrix in Development and Cancer Progression. Int J Mol Sci. 2018. doi:10.3390/ijms19103028.
22. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. Oncogene. 2017;36:1461–73. doi:10.1038/onc.2016.304.
23. Nie X, Liu H, Liu L, Wang Y-D, Chen W-D. Emerging Roles of Wnt Ligands in Human Colorectal Cancer. Front Oncol. 2020;10:1341. doi:10.3389/fonc.2020.01341.
24. Cheng X, Xu X, Chen D, Zhao F, Wang W. Therapeutic potential of targeting the Wnt/β-catenin signaling pathway in colorectal cancer. Biomed Pharmacother. 2019;110:473–81. doi:10.1016/j.biopha.2018.11.082.
25. Derynck R, Akhurst RJ, Balmain A. TGF-beta signaling in tumor suppression and cancer progression. Nat Genet. 2001;29:117–29. doi:10.1038/ng1001-117.

26. Itatani Y, Kawada K, Sakai Y. Transforming Growth Factor-β Signaling Pathway in Colorectal Cancer and Its Tumor Microenvironment. Int J Mol Sci. 2019. doi:10.3390/ijms20235822.
27. Shevde LA, Samant RS. Role of osteopontin in the pathophysiology of cancer. Matrix Biol. 2014;37:131–41. doi:10.1016/j.matbio.2014.03.001.
28. Wei R, Wong JPC, Lyu P, Xi X, Tong O, Zhang S-D, et al. In vitro and clinical data analysis of Osteopontin as a prognostic indicator in colorectal cancer. J Cell Mol Med. 2018;22:4097–105. doi:10.1111/jcmm.13686.
29. Zeng B, Zhou M, Wu H, Xiong Z. SPP1 promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway. Onco Targets Ther. 2018;11:1333–43. doi:10.2147/OTT.S154215.
30. Zhang Y, Du W, Chen Z, Xiang C. Upregulation of PD-L1 by SPP1 mediates macrophage polarization and facilitates immune escape in lung adenocarcinoma. Exp Cell Res. 2017;359:449–57. doi:10.1016/j.yexcr.2017.08.028.
31. Guo R, Lv Y, Ouyang Y, Liu S, Li D. The Role of miR-497/EIF3A Axis in TGFβ1-Induced Epithelial-Mesenchymal Transition and Extracellular Matrix in Rat Alveolar Epithelial Cells and Pulmonary Fibroblasts. J Cell Biochem. 2017;118:3401–8. doi:10.1002/jcb.25997.
32. Minamino N, Kangawa K, Matsuo H. Neuromedin B: a novel bombesin-like peptide identified in porcine spinal cord. Biochem Biophys Res Commun. 1983;114:541–8. doi:10.1016/0006-291x(83)90814-8.
33. Panula P. Histochemistry and function of bombesin-like peptides. Med Biol. 1986;64:177–92.
34. Gonzalez N, Moody TW, Igarashi H, Ito T, Jensen RT. Bombesin-related peptides and their receptors: recent advances in their role in physiology and disease states. Curr Opin Endocrinol Diabetes Obes. 2008;15:58–64. doi:10.1097/MED.0b013e3282f3709b.
35. Ischia J, Patel O, Shulkes A, Baldwin GS. Gastrin-releasing peptide: different forms, different functions. Biofactors. 2009;35:69–75. doi:10.1002/biof.10.
36. Jensen RT, Battey JF, Spindel ER, Benya RV. International Union of Pharmacology. LXVIII. Mammalian bombesin receptors: nomenclature, distribution, pharmacology, signaling, and functions in normal and disease states. Pharmacol Rev. 2008;60:1–42. doi:10.1124/pr.107.07108.
37. Patel O, Shulkes A, Baldwin GS. Gastrin-releasing peptide and cancer. Biochim Biophys Acta. 2006;1766:23–41. doi:10.1016/j.bbcan.2006.01.003.
38. Patel O, Clyde D, Chang M, Nordlund MS, Steel R, Kemp BE, et al. Pro-GRP-derived peptides are expressed in colorectal cancer cells and tumors and are biologically active in vivo. Endocrinology. 2012;153:1082–92. doi:10.1210/en.2011-1875.
39. Ni C, Zhao X, Sun T, Liu Y, Gu Q, Sun B. Role of gastrin-releasing peptides in breast cancer metastasis. Hum Pathol. 2012;43:2342–7. doi:10.1016/j.humpath.2012.04.007.
40. Scherer SW, Feinstein DS, Oliveira L, Tsui LC, Pittler SJ. Gene structure and chromosome localization to 7q21.3 of the human rod photoreceptor transducin gamma-subunit gene (GNGT1). Genomics. 1996;35:241–3. doi:10.1006/geno.1996.0346.

41. Alsaleem M, Toss MS, Joseph C, Aleskandarany M, Kurozumi S, Alshankyty I, et al. The molecular mechanisms underlying reduced E-cadherin expression in invasive ductal carcinoma of the breast: high throughput analysis of large cohorts. Mod Pathol. 2019;32:967–76. doi:10.1038/s41379-019-0209-9.
42. Mucaki EJ, Zhao JZL, Lizotte DJ, Rogan PK. Predicting responses to platin chemotherapy agents with biochemically-inspired machine learning. Signal Transduct Target Ther. 2019;4:1. doi:10.1038/s41392-018-0034-5.

Figures

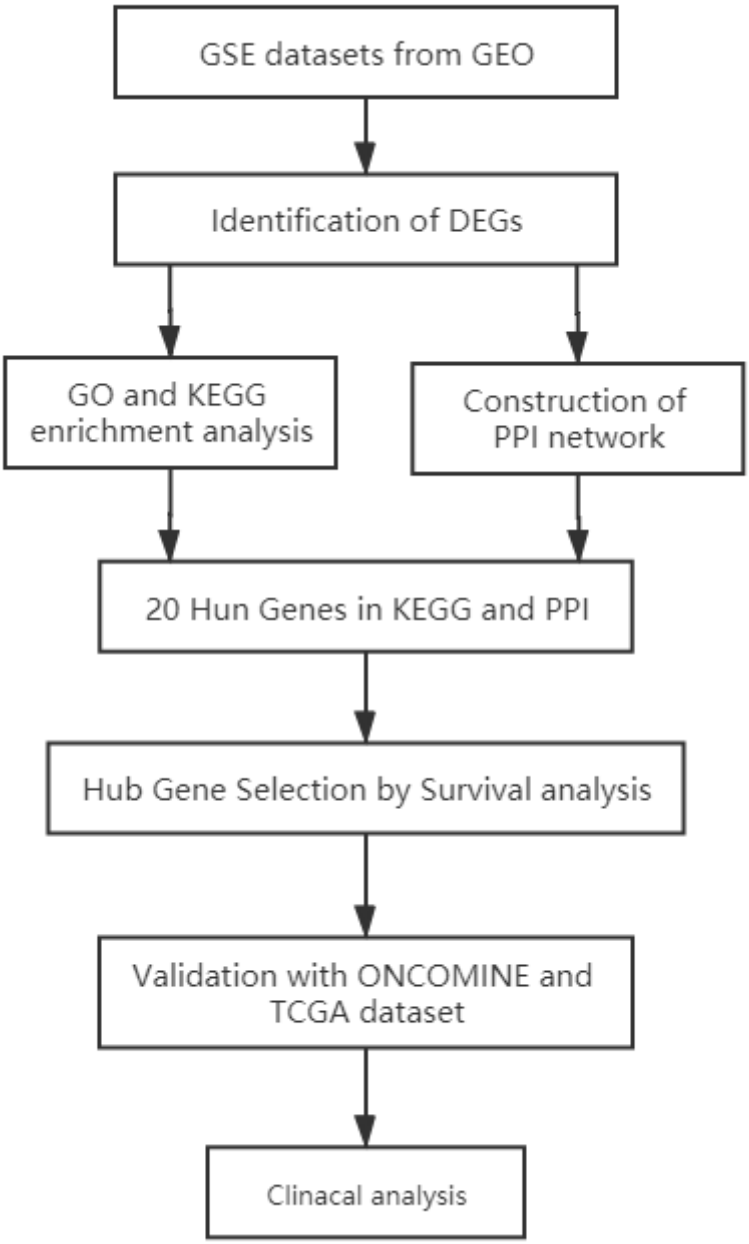

Figure 1

The analysis process of present study.

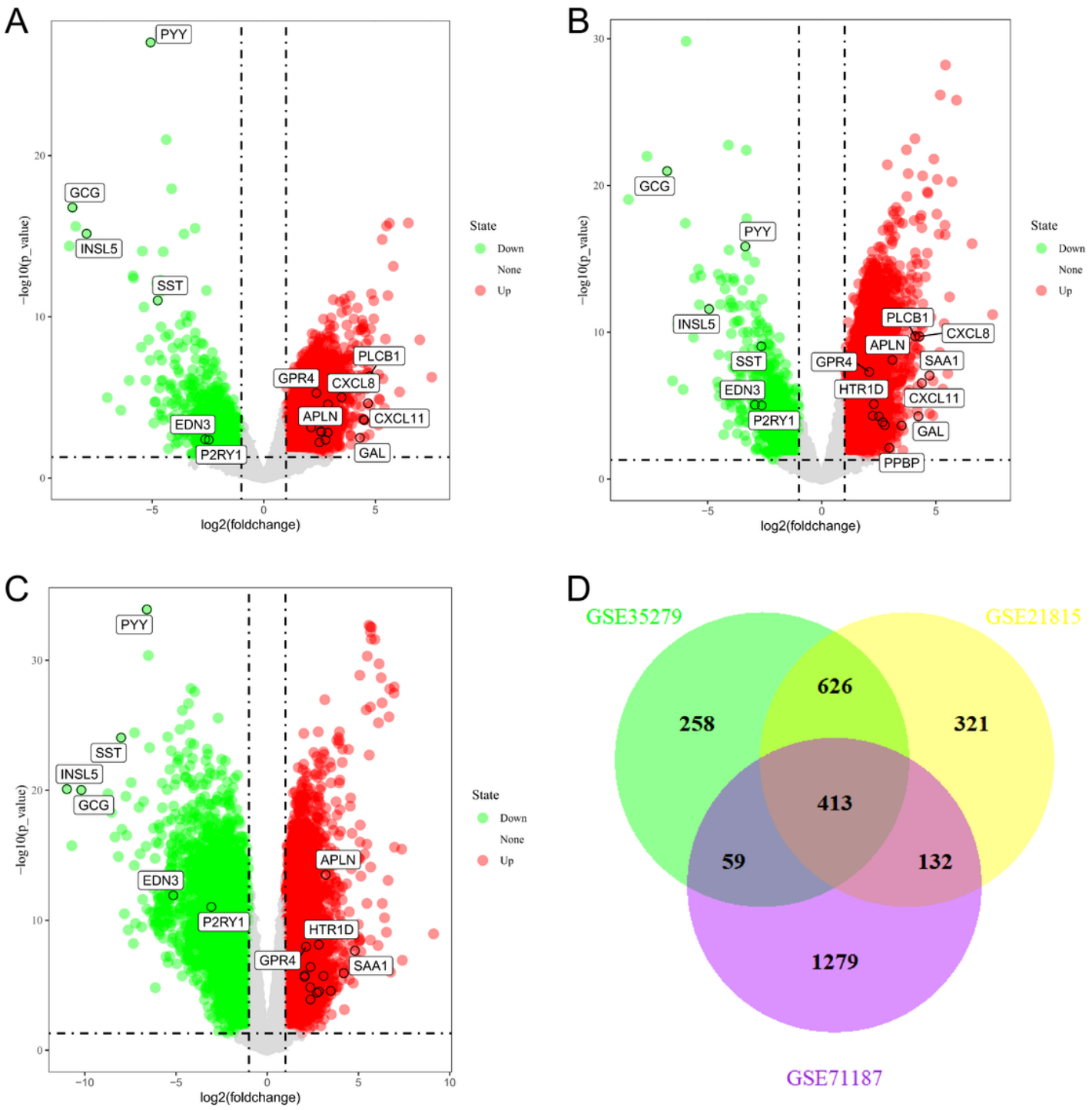

Figure 2

DEGs' volcanic map and VENN diagram. A-C, differential genes were screened from three data sets (GSE35279, GSE21815, and GSE71187), and the distribution of differential genes in each data sets was observed by volcanic map. D, VENN diagram were drawn according to all differential genes(DEGs).

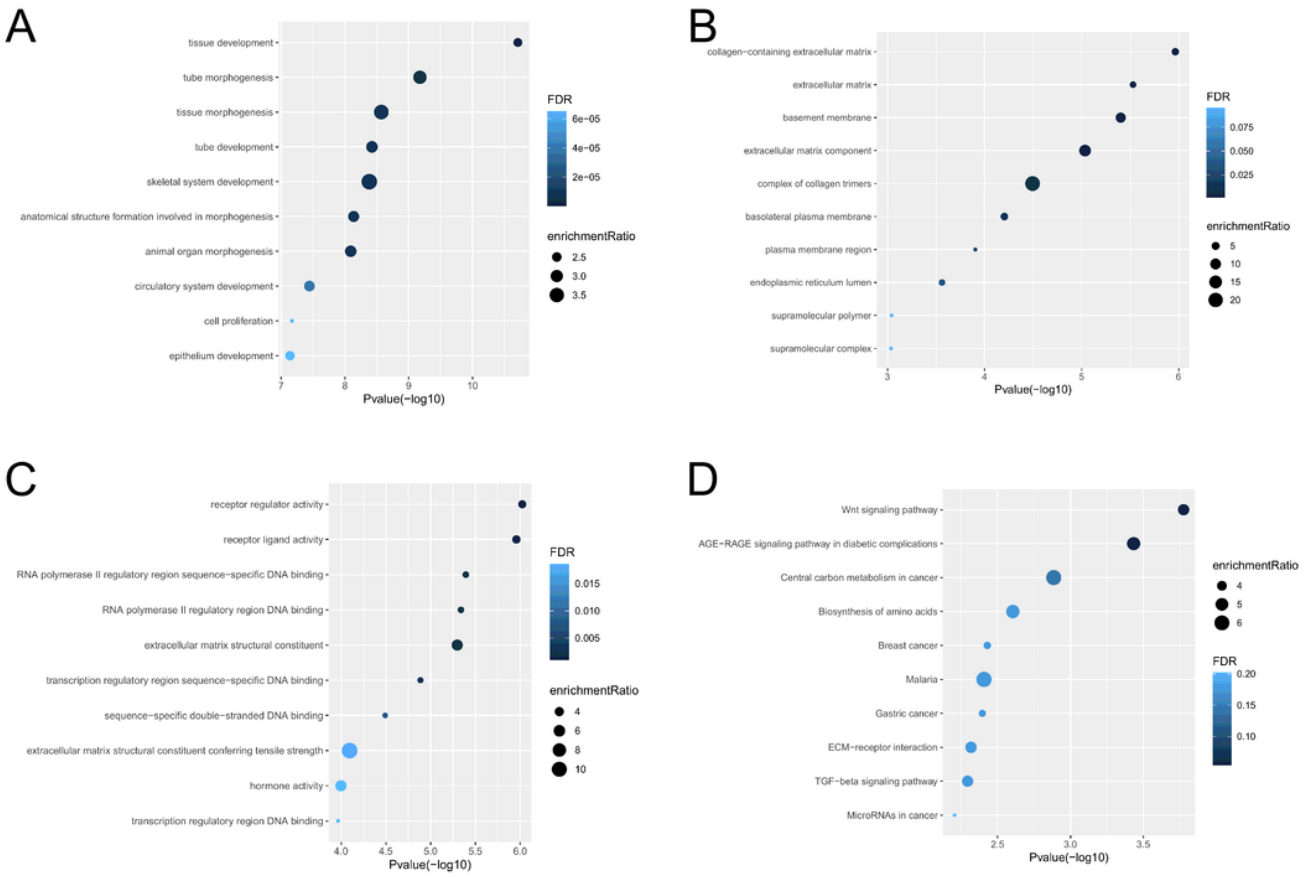

Figure 3

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of DEGs. A, Biological process analyses of DEGs; B, Molecular function analyses of DEGs; C, Cellular components analyses of DEGs;D, KEGG pathway analyses of DEGs. The GO terms and KEGG pathways were ranked by FDR. Top 10 terms were selected according to FDR. Gene ratio: the ratio of the number of enriched genes in each term to the total number of DEGs.

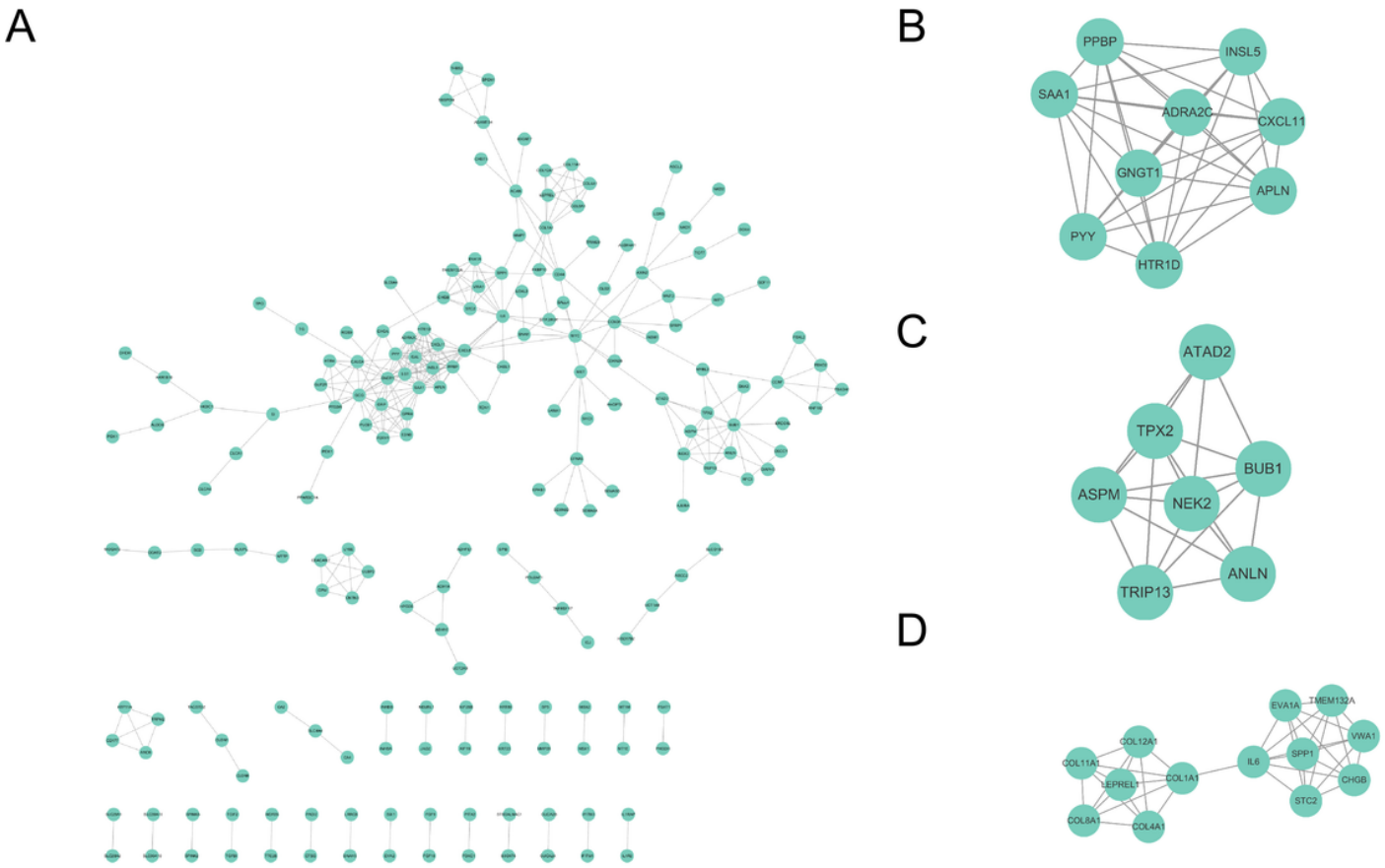

Figure 4

Protein–protein interaction (PPI) network and significant modules of DEGs. A, PPI network contained 388 nodes and 334 edges. B-D, Top three functional clusters of modules selected from PPI network.

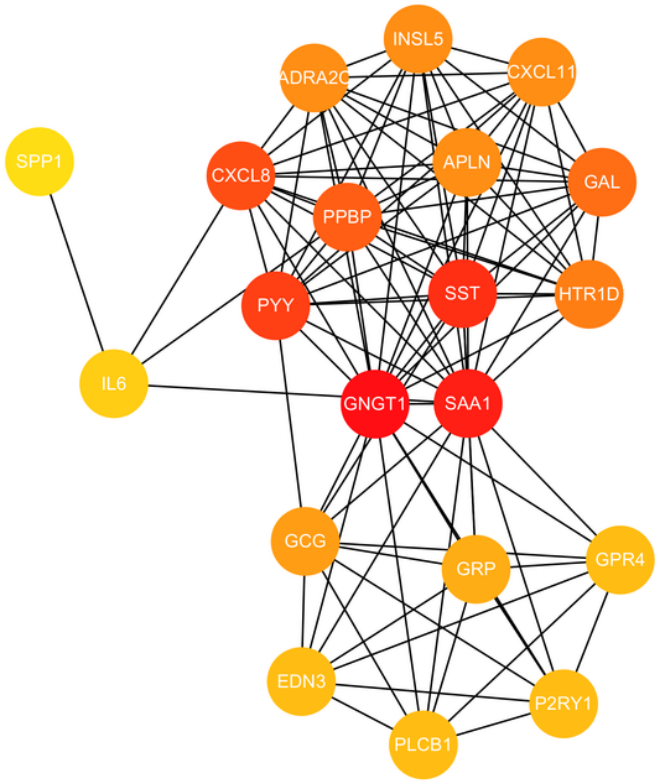

Figure 5

PPI network of the top 20 hub genes according to the maximal clique centrality (MCC)

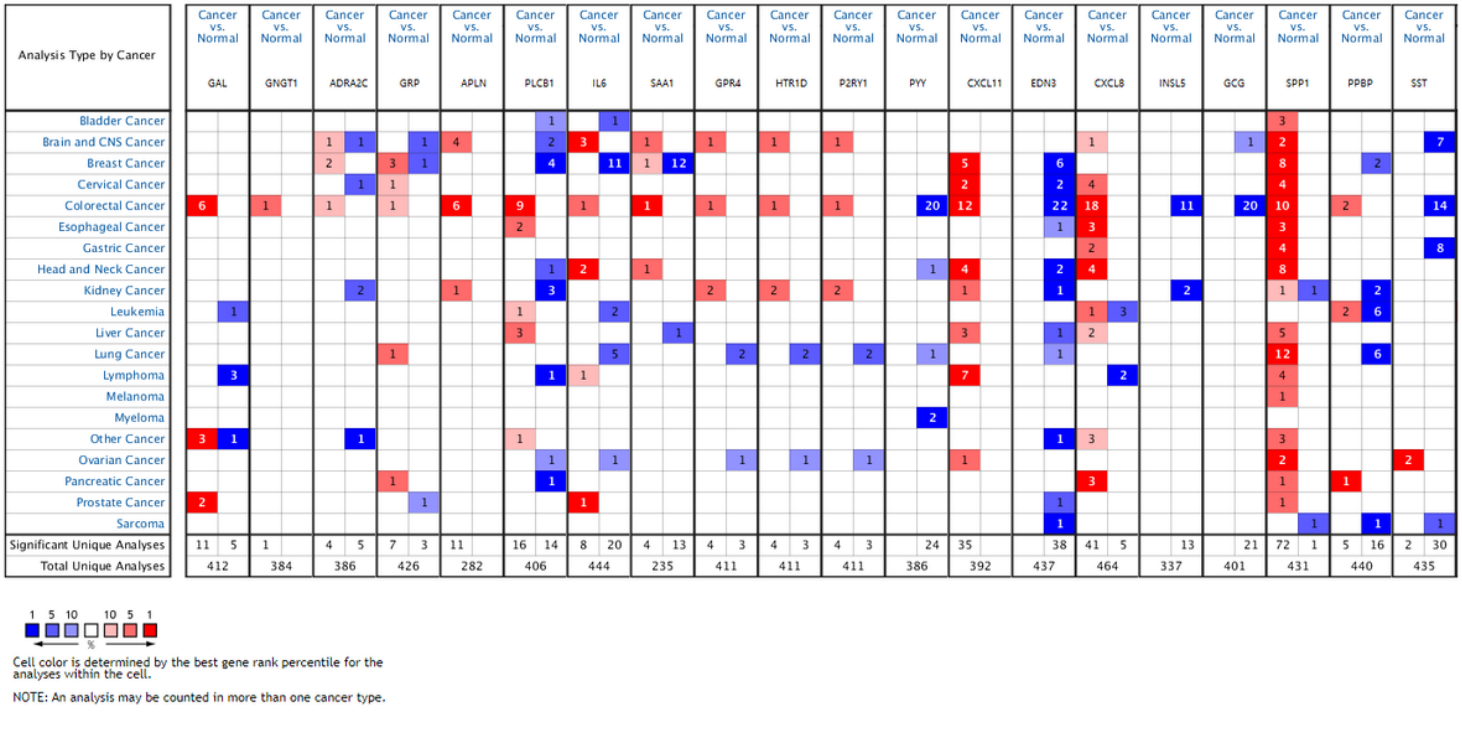

Figure 6

Transcriptional expression of Hub genes in 20 different types of cancer diseases. Difference of transcriptional expression was compared by students' t test. Cut-off of p value and fold change were as follows: p value, 1E-4; fold change, 2; gene rank, 10%; data type, mRNA.

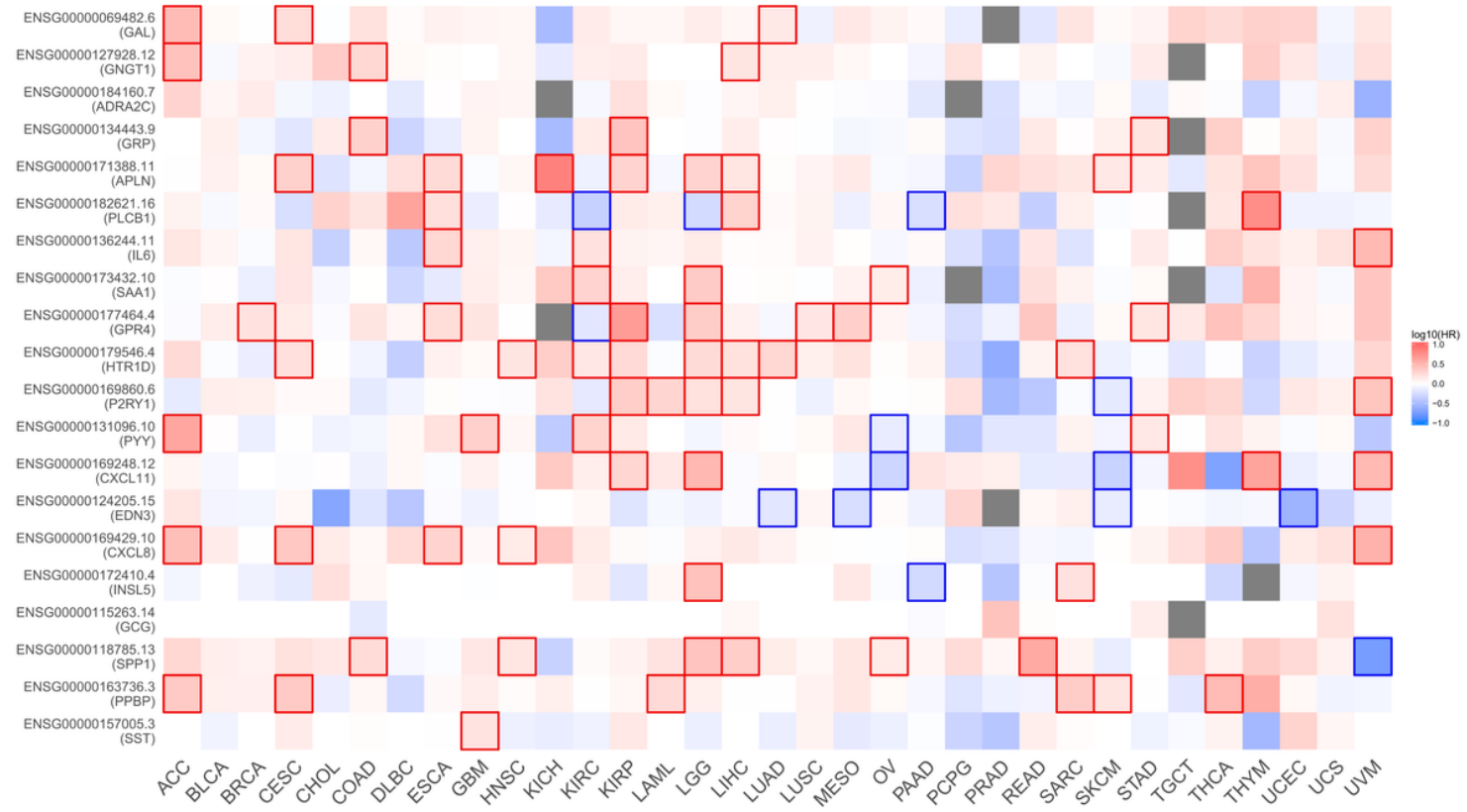

Figure 7

Overall survival analysis of hub genes in pan cancer was performed by utilizing the GEPIA2 Survival Map online tool. Significance Level, 0.05; Group Cutoff, 25% high.

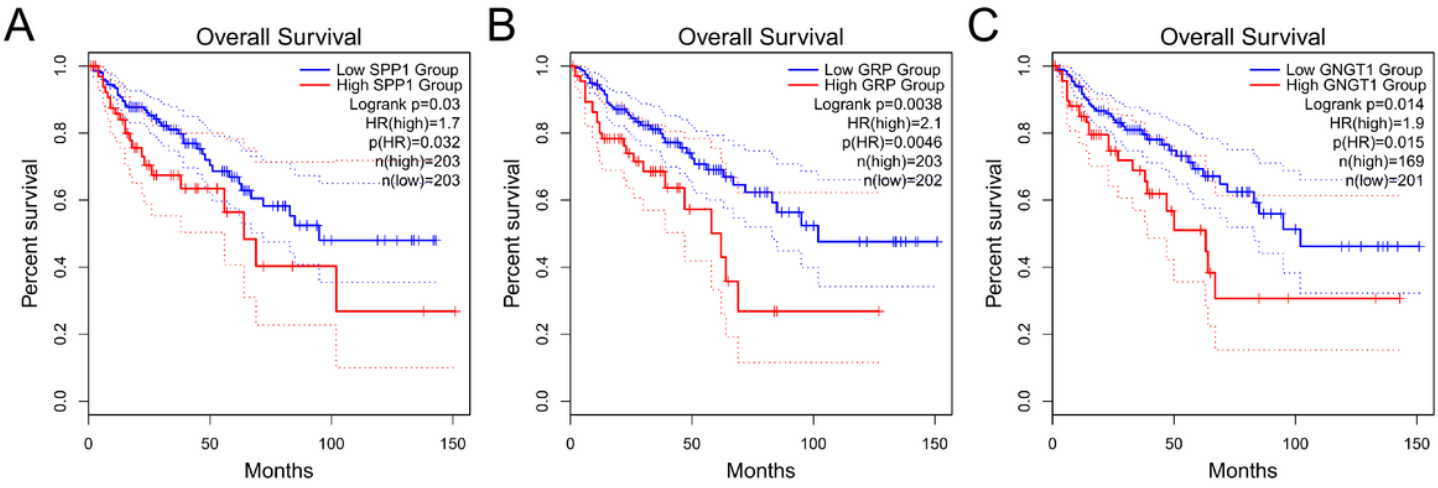

Figure 8

The detection of the survival rate of the chosen hub genes in colon adenocarcinoma (COAD). A-C, Kaplan-Meier curves of SPP1, GRP, and GNGT1. Log-rank test was carried on the relevant results. Significance Level, 0.05; Group Cutoff, 25% high.

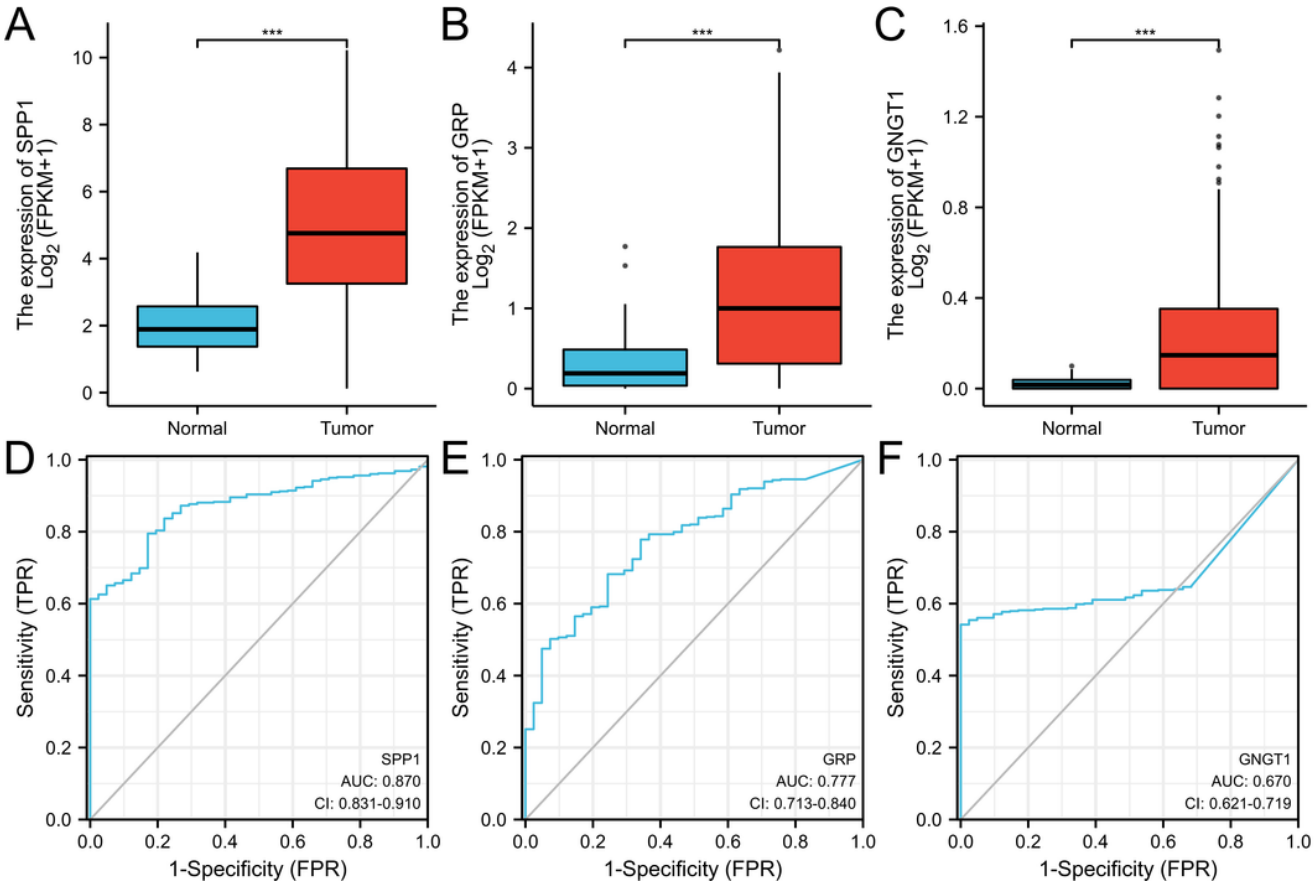

Figure 9

The detection of the expression level the diagnostic value the chosen hub genes. A-C, Box plot of SPP1, GRP, and GNGT1 expression level in COAD,the result show that these core genes were over-expressed in COAD ($p<0.001$). D-F, ROC curves of SPP1, GRP, and GNGT1 in COAD. AUC and CI were carried on the relevant results

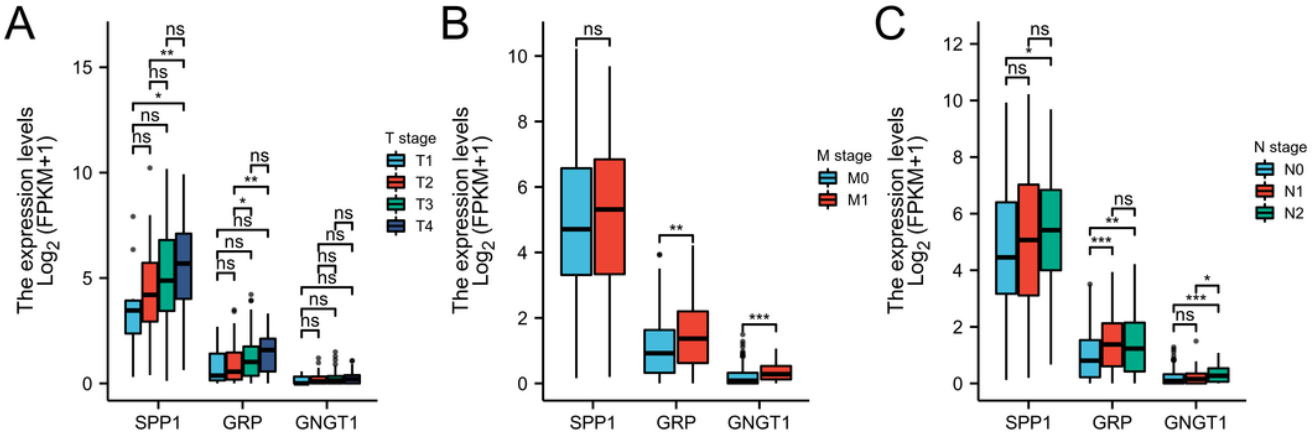

Figure 10

The detection of the relationship between gene expression and TMN stage of COAD.A, Hub gene expression level in different T stage. B, Hub gene expression level in different M stage. C, Hub gene expression level in different N stage.