Journal of Biosciences and Medicines, 2026, 14(4), 81-89
https://www.scirp.org/journal/jbm
ISSN Online: 2327-509X
ISSN Print: 2327-5081

Core Gene Screening and Correlation Analysis in Gastric Cancer

Jiahui Wang[1,2,3], Huaidong Xu[4], Runze Wu[1,2,3], Dingpeng Ban[1,2,3], Erxi Xu[1,2,3], Huiying Li[1,2,3], Gongzu Wang[1,2,3], Shengchao Yang[1,2,3], Haidong Zhou[1,2,3], Hongbo Li[5*]

[1]Department of Sport Medicine, Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China
[2]Biomedical Materials Engineering Research Center for Bone and Joint Degenerative Diseases, Baise, China
[3]Graduate School of Youjiang Medical University for Nationalities, Baise, China
[4]Department of Critical Care Medicine, Affiliated Hospital of Chongqing University, Chongqing, China
[5]Tianlin County People's Hospital, Baise, China
Email: wjh010828@163.com, 1695573477@qq.com, *bobosiji009@163.com

How to cite this paper: Wang, J.H., Xu, H.D., Wu, R.Z., Ban, D.P., Xu, E.X., Li, H.Y., Wang, G.Z., Yang, S.C., Zhou, H.D. and Li, H.B. (2026) Core Gene Screening and Correlation Analysis in Gastric Cancer. *Journal of Biosciences and Medicines*, **14**, 81-89.
https://doi.org/10.4236/jbm.2026.144007

Received: January 29, 2026
Accepted: April 10, 2026
Published: April 13, 2026

Abstract

Objective: To screen and analyze differentially expressed genes (DEGs) in gastric carcinoma (GC) using a bioinformatics approach. Methods: Data were retrieved from the GEO microarray public database in NCBI, and the microarray dataset GSE49051 was selected for analysis. The R language limma package was used to screen differentially expressed mRNAs (DEmRNAs), and the data were subjected to normalization processing. A Venn diagram was employed to identify the common DEGs between the two groups. The R language clusterProfiler package was used to perform Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses on these common DEGs. The STRING database was utilized for protein-protein interaction (PPI) analysis, and the results were imported into Cytoscape software to generate the PPI interaction network, identify core modules, and determine hub genes. Known GC-related genes were downloaded from the OMIM database and intersected with the core module genes identified by MCODE, yielding the APOH gene. Furthermore, the genes downloaded from the OMIM database were intersected with the screened DEGs, resulting in five genes: APOH, AHSG, APOA2, AMBP, and HP. The hub genes were then imported into the BioGPS database, which revealed no tissue-specific genes. Conclusion: The identified DEGs and associated signaling pathways contribute to understanding the molecular mechanisms underlying GC pathogenesis and may provide a basis for the early diagnosis of GC.

Keywords

Gastric Cancer, Bioinformatics Analysis, Differentially Expressed Gene,

Signaling Pathways

1. Introduction

Gastric carcinoma (GC) is a common malignant tumor originating from gastric epithelial cells, characterized by a complex pathogenesis and rapid disease progression, posing a serious threat to human life and health. It features high incidence, high mortality, high disability rates, and a heavy medical burden, ranking fifth in global malignant tumor incidence and third in mortality. As a country with a high incidence of gastric cancer, China accounts for approximately 42% of new cases worldwide, and about 80% of patients are diagnosed at an advanced local stage, making treatment difficult and prognosis poor. Epidemiological data show that the incidence in men is about twice that in women, with the peak incidence concentrated in the 60 - 80 age group. In recent years, the incidence among younger populations has shown an upward trend, with the proportion of patients under 30 gradually increasing.

Gastric cancer is a disease caused by the interaction of multiple factors, and its pathogenesis involves genetics, infection, environment, lifestyle, and underlying diseases, among other aspects. Regarding genetic factors, 5% - 10% of patients exhibit familial clustering, with individuals with a positive family history having a 2 - 3 times higher risk of developing the disease than the general population; 1% - 3% of patients carry specific pathogenic genes, such as CDH1 gene mutations that can lead to hereditary diffuse gastric cancer, with carriers having a lifetime risk of up to 70% - 80%. Among infectious factors, *Helicobacter pylori* (Hp) infection is a key trigger, as its secreted urease, vacuolating toxin, and other substances damage the gastric mucosal barrier, leading to chronic inflammation. Long-term stimulation can result in intestinal metaplasia, dysplasia, and ultimately malignant transformation. Additionally, exposure to aflatoxins, bacterial overgrowth due to reduced gastric acid, and the generation of nitrites are closely associated with the occurrence of gastric cancer.

Advanced age is a significant risk factor for the disease, as DNA damage accumulates and repair functions decline with age, significantly increasing the risk of gene mutations. Geographical distribution varies significantly, with areas such as Linqu in Shandong, Zhuanghe in Liaoning, and Changle in Fujian in China having significantly higher incidence rates due to long-term consumption of high-salt preserved foods. In contrast, incidence rates are relatively lower in Europe and America, a difference related to dietary structure, *Helicobacter pylori* infection rates, and environmental factors, among others. In terms of underlying diseases, chronic atrophic gastritis, gastric ulcers, gastric polyps, remnant stomach, hypertrophic gastritis, and pernicious anemia can cause repeated damage and repair of the gastric mucosa, promoting abnormal cell proliferation and malignant transformation.

Lifestyle factors are also noteworthy. Irregular eating habits disrupt the rhythm of gastric acid secretion, while long-term high-salt and red meat diets increase the burden on the gastric mucosa and promote nitrite formation. Obesity leads to hormonal imbalances, and smoking and alcohol abuse directly damage the gastric mucosa. Nitrites and aflatoxins in preserved foods are recognized carcinogens. Additionally, long-term staying up late and excessive psychological stress can interfere with neuroendocrine regulation, reduce immune function, and impair the defense mechanisms of the gastric mucosa. Early-stage gastric cancer often presents with subtle symptoms, such as upper abdominal discomfort and loss of appetite, which are nonspecific and easily missed. Advanced stages may exhibit typical symptoms such as abdominal pain, hematemesis, and melena, but often at a point when the optimal treatment window has been missed. Currently, gastroscopy combined with pathological biopsy is the gold standard for diagnosis, but due to low early screening coverage, most patients are diagnosed at an intermediate or advanced stage. Treatment methods include surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy. Improving early diagnosis rates and optimizing comprehensive treatment strategies remain key challenges in the field of gastric cancer prevention and treatment.

In summary, this study conducts bioinformatics analysis on gastric cancer and control population gene chip data from the public gene chip database (Gene Expression Omnibus, GEO) to explore potential molecular biological functions involved, providing a theoretical reference for elucidating its molecular mechanisms.

2. Materials and Methods

2.1. Data Acquisition

Raw data from the patient dataset chip GSE49051 were obtained by searching the public Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information using "gastric carcinoma" as the keyword. GSE49051 includes specimens from 3 gastric carcinoma patients and 3 normal individuals. The specimens taken for gastric cancer were gastric cancer tumor tissues, and those taken from normal individuals were normal gastric tissues. Total RNA was extracted from the specimens. The microarray platform used was GPL10332 Agilent-026652 Whole Human Genome Microarray 4x44K v2 (Feature Number version).

Differential expression analysis was performed on the gastric carcinoma mRNA chip data using the R language limma package to identify differentially expressed mRNAs (DEmRNAs). The screening criteria were set as a statistical significance of $P < 0.05$ and an absolute value of the log fold change $|\log_2 FC| > 4$ means that the screened mRNAs must exhibit an actual expression fold change greater than or equal to 16-fold (2^4) or less than or equal to 1/16-fold (2^{-4}) between gastric cancer tissues and normal tissues, aiming to focus on genes that undergo drastic changes and may play core functional roles.

2.2. GO Function and KEGG Pathway Enrichment Analysis

GO function enrichment analysis was conducted to investigate the biological processes involving the gastric carcinoma DEmRNAs, using the R clusterProfiler package. Simultaneously, KEGG pathway enrichment analysis was performed.

2.3. PPI and Core Module Analysis

The differential genes were analyzed using the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) 11.0 online tool. The results from STRING were imported into Cytoscape 3.9.1 software. Plugins such as CytoNCA, Cytohubba, and MCODE were utilized to construct protein-protein interaction networks, identify co-expression core genes, and analyze modules.

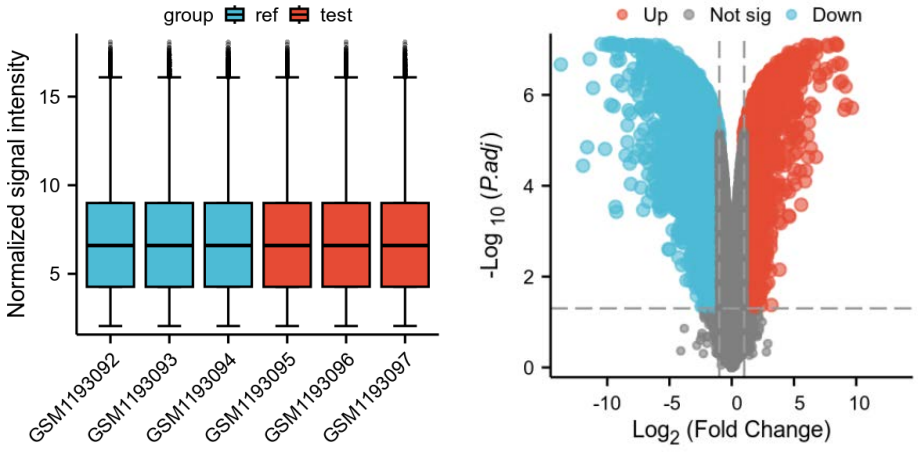
2.4. Hub Gene Localization

Tissue localization of the Hub genes was performed using the online database BioGPS (http://biogps.org/#goto=welcome). A gene was considered significantly tissue-specific if its highest expression level in one tissue was more than twice the second-highest expression level, indicating its potential as a common expression marker gene for GC.

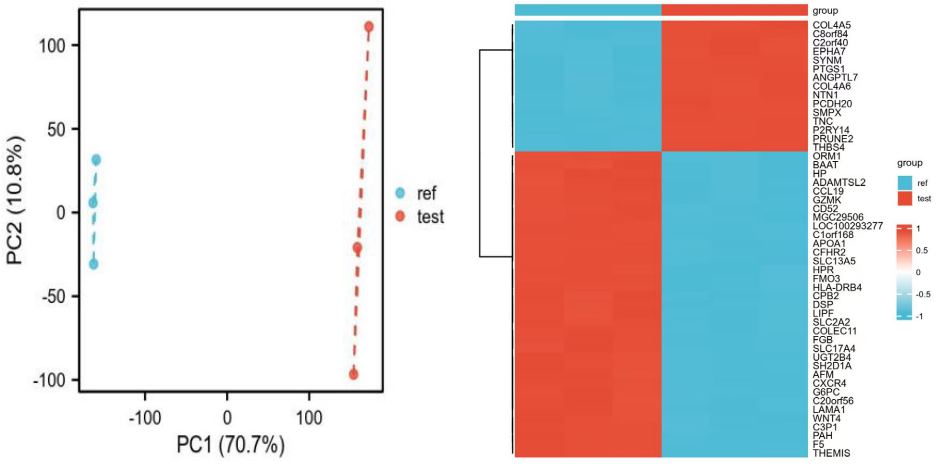
3. Results

3.1. Identification of Differentially Expressed mRNAs

Analysis of the GSE49051 dataset using the R package limma revealed that the expression levels of most genes were largely consistent, indicating the suitability of the data for further analysis. Figure 1(a) presents a boxplot of the samples. From this analysis, a total of 603 differentially expressed genes (DEGs) were identified between gastric carcinoma patients and normal individuals. Figure 1(b) shows the volcano plot depicting the distribution of these DEGs. Principal Component Analysis (PCA) of the sample data is displayed in Figure 1(c), and the heatmap of the DEG distribution is presented in Figure 1(d).



(a) (b)



(c) (d)

Note: (a) Box plot of the data chip, (b) Volcano plot of the data chip, (c) PCA analysis plot of the data chip, (d) Heatmap of differentially expressed gene distribution.

Figure 1. Box plot and volcano plot of gene expression.

3.2. GO and KEGG Enrichment Analysis of Intersection Genes

The R clusterProfiler package was employed to perform GO and KEGG pathway enrichment analysis on the common target genes of gastric cancer. Analyses were conducted on Biological Processes (BP), Cellular Components (CC), Molecular Functions (MF), and Pathways in KEGG. The results are as follows:

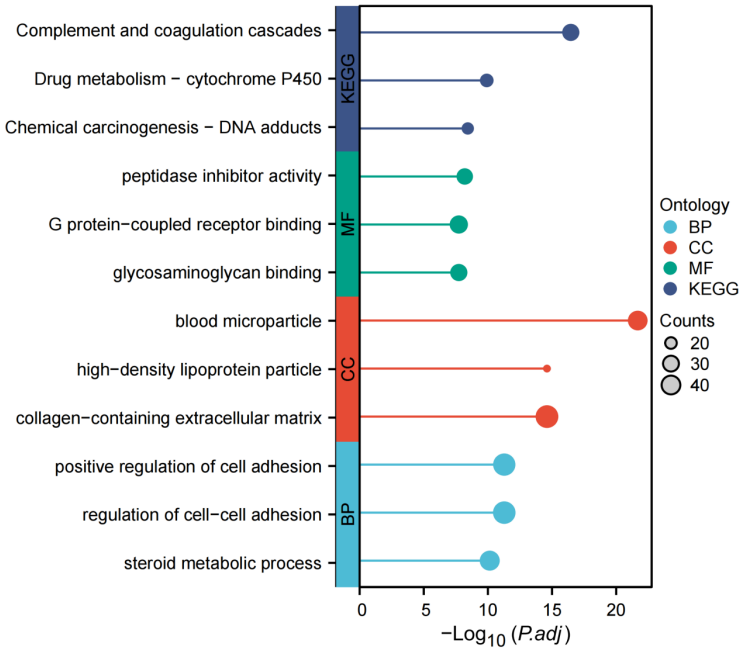


Figure 2. GO functional enrichment analysis and KEGG pathway analysis.

1) BP was primarily enriched in positive regulation of cell adhesion, regulation of cell-cell adhesion, and steroid metabolic process, as shown in Figure 2;

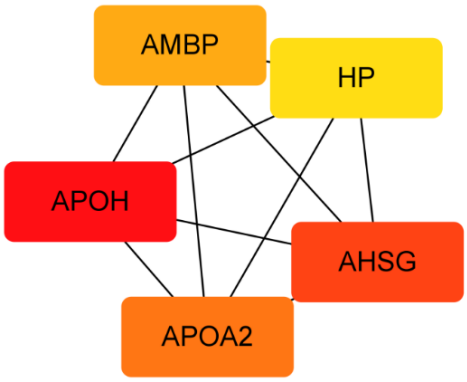
2) CC was primarily enriched in blood microparticle, high-density lipoprotein particle, and collagen-containing extracellular matrix, as shown in Figure 2;

3) MF was primarily enriched in peptidase inhibitor activity, G protein-coupled receptor binding, and glycosaminoglycan binding, as shown in Figure 2;

4) KEGG pathways were primarily enriched in Complement and coagulation cascades, Drug metabolism, cytochrome P450, and Chemical carcinogenesis, DNA adducts, as shown in Figure 2.

3.3. Construction of Common Expression PPI Network and Prediction of Core Genes

The 603 common DEmRNAs were analyzed using the STRING online tool with a minimum interaction score set to 0.4. The STRING results were imported into Cytoscape 3.9.1 software to construct the PPI interaction diagram. The CytoNCA plugin, employing the BC algorithm, was used to rank genes with higher betweenness centrality (BC) scores in clockwise order based on their scores. The MCODE algorithm within the Cytoscape plugin identified common expression gene modules, resulting in a total of 7 core modules with scores greater than 4. Among these, using 10 algorithms such as MCC, DMNC, and MNC in the CytoHubba plugin, five hub genes were identified (Figure 3), ranked by score as follows: APOH, AHSG, APOA2, AMBP, and HP. The final screening strategy for the five core Hub genes (APOH, AHSG, APOA2, AMBP, HP) does not simply take the union of the results from ten algorithms (which may lead to an excessive number of candidate genes) or their intersection (which might be too stringent and omit important genes). Instead, it first summarizes the top 50 genes independently output by each algorithm, and then calculates the comprehensive score of each gene across the algorithms. Ultimately, the weighted scores from all algorithms are summed, and the top five genes with the highest total scores are selected as Hub genes. This method effectively integrates the advantages of different topological indicators, enhances the stability and biological consistency of Hub gene identification, and avoids the bias that may arise from using a single algorithm.



Note: Red indicates a high enrichment score, with lighter colors representing lower scores. The connecting lines represent interactions between genes.

Figure 3. Hub gene core module.

4. Discussion

By screening the OMIM (Online Mendelian Inheritance in Man) database of human genes and genetic disorders, we identified genes reported in the literature to be associated with Gastric Carcinoma (GC). Ultimately, we obtained the intersection between the differential genes and core modules of GC and the reported genes from the OMIM database using Venn diagrams, resulting in five core genes: APOH, AHSG, APOA2, AMBP, and HP.

APOH (Apolipoprotein H), also known as β2-glycoprotein I (β2-GPI), plays a role in lipid metabolism, coagulation, and immune responses. The currently available literature does not directly detail the specific role and mechanism of APOH in gastric cancer. However, dysregulated lipid metabolism is a key characteristic of tumorigenesis and progression [1] [2]. For instance, Apolipoprotein C1 (APOC1) has been shown to regulate gastric cancer progression by inducing epithelial-mesenchymal transition (EMT) via the JAK/STAT pathway [3], and the Epidermal Growth Factor (EGF)/APOC1/CPT1A axis also plays a role in gastric cancer metabolism and therapeutic targeting [4]. Although APOH belongs to the apolipoprotein family alongside APOA2 and APOE, its precise function in gastric cancer requires further investigation.

AHSG (Alpha-2-HS-glycoprotein), also known as Fetuin-A, is a multifunctional molecule involved in various biological processes, including mineralization, tumor growth, and inflammation [5]. In gastric cancer, it has been explored as a potential diagnostic biomarker [6]. While AHSG has been identified as a gene promoting tumor proliferation, migration, and invasion in lung adenocarcinoma and is an independent predictor of poor prognosis [7], and in bladder cancer, it can promote bladder cancer cell proliferation by regulating the TGF-β signaling pathway [8], its specific mechanism and role in gastric cancer require further in-depth study. Recent serum proteomics research in gastric cancer and precancerous lesions has revealed protein signatures related to systemic inflammatory responses and metaplastic differentiation, which may include AHSG [9].

APOA2 (Apolipoprotein A-II) is one of the major apolipoproteins of high-density lipoprotein (HDL), composed of 77 amino acid residues [10]. In gastric cancer, aberrant expression of APOA2 is associated with disease progression and poor prognosis. Studies have found that APOA2 expression is upregulated in gastric cancer tissues [10]. For example, a bioinformatics analysis combined with in vitro experiments revealed that APOA2 may influence gastric cancer development by promoting tumor cell proliferation and migration [11]. The expression level of APOA2 is significantly correlated with the prognosis of gastric cancer patients, suggesting its potential as a biomarker for gastric cancer. Furthermore, APOA2 isoforms have also been studied as serum biomarkers for early detection of pancreatic cancer [11], indicating their potential as biomarkers in other gastrointestinal tumors.

The AMBP gene encodes the Alpha-1-microglobulin/bikunin precursor protein, which is enzymatically cleaved to produce two separate proteins: alpha-1-

microglobulin and bikunin. These proteins possess various biological functions, including immunomodulation and protease inhibition. In gastric cancer-related gene analyses, the specific role and mechanism of the AMBP gene are not directly detailed in the provided literature. However, studies suggest that both products have significant biological functions, including regulating immune responses and protease activity. In gastric cancer, aberrant expression of AMBP is associated with tumor invasiveness. Research indicates that AMBP expression is downregulated in gastric cancer tissues, and its low expression correlates with deeper tumor invasion depth and later TNM stages, suggesting a poorer prognosis [12]. Nonetheless, the potential role of AMBP in gastric cancer still requires further experimental and clinical studies for confirmation.

HP (Haptoglobin) is an acute-phase response protein whose primary function is to bind free hemoglobin, preventing iron loss and kidney damage. In gastric cancer, changes in HP expression and its polymorphisms are associated with tumor risk and progression. Serum HP levels are often significantly elevated in gastric cancer patients and correlate with tumor burden and malignancy. Mechanistically, HP may shape a tumor-favorable microenvironment by promoting the polarization of tumor-associated macrophages, regulating angiogenesis, and participating in oxidative stress responses. Furthermore, HP gene polymorphisms (particularly the Hp1-1, Hp2-1, and Hp2-2 types) are also considered to be associated with individual susceptibility to gastric cancer and prognostic differences. Therefore, HP is not only a marker of inflammation but also an active molecule involved in gastric cancer progression.

5. Summary and Outlook

The aforementioned genes are involved in the occurrence and development of GC to varying degrees, with several being particularly associated with inflammation and immune dysfunction. This provides new insights for future in-depth research into the role and connection of the immune system in the pathogenesis of these conditions.

In summary, through data mining, investigating target genes related to GC serves a dual purpose: offering new early warning indicators for GC and providing a reference for future research aimed at elucidating the mechanisms of GC.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Wang, C., Yang, Z., Xu, E., Shen, X., Wang, X., Li, Z., *et al.* (2021) Apolipoprotein C-II Induces EMT to Promote Gastric Cancer Peritoneal Metastasis via PI3K/Akt/mTOR Pathway. *Clinical and Translational Medicine*, **11**, e522. https://doi.org/10.1002/ctm2.522

[2] Li, F., Han, M., Gao, X., Du, X. and Jiang, C. (2024) APOA1 mRNA and Serum

APOA1 Protein as Diagnostic and Prognostic Biomarkers in Gastric Cancer. *Translational Cancer Research*, **13**, 2141-2154. https://doi.org/10.21037/tcr-23-1966

[3] Zhang, J., Yang, Z., Li, Z.J., Li, R.Q., Li, D., Feng, M., Guan, W.X. and Liu, X.N. (2023) Apolipoprotein C1 Regulates Gastric Cancer Progression by Inducing Epithelial-Mesenchymal Transition via the JAK/STAT Pathway. https://doi.org/10.21203/rs.3.rs-2659063/v1

[4] Du, F., Li, P., Chen, G., *et al*. (2024) EGF/APOC1/CPT1A Axis: A Novel Pathway in Gastric Cancer Metabolism and Therapeutic Targeting. https://doi.org/10.21203/rs.3.rs-4832902/v1

[5] Odiase, P., Ma, J., Ranganathan, S., Ogunkua, O., Turner, W.B., Marshall, D., *et al*. (2024) The Role of Fetuin-A in Tumor Cell Growth, Prognosis, and Dissemination. *International Journal of Molecular Sciences*, **25**, Article 12918. https://doi.org/10.3390/ijms252312918

[6] Shi, F., Wu, H., Qu, K., Sun, Q., Li, F., Shi, C., *et al*. (2018) Identification of Serum Proteins AHSG, FGA and APOA-I as Diagnostic Biomarkers for Gastric Cancer. *Clinical Proteomics*, **15**, Article No. 18. https://doi.org/10.1186/s12014-018-9194-0

[7] Xing, X., Cao, F., Gao, L. and Song, M. (2023) AHSG, a Gene Promoting Tumour Proliferation, Migration and Invasion, Is an Independent Prognostic Factor for Poor Overall Survival in Lung Adenocarcinoma. *Molecular Biology Reports*, **50**, 7659-7666. https://doi.org/10.1007/s11033-023-08623-x

[8] Dong, Y., Ding, D., Gu, J., Chen, M. and Li, S. (2022) α-2 Heremans Schmid Glycoprotein (AHSG) Promotes the Proliferation of Bladder Cancer Cells by Regulating the TGF-β Signalling Pathway. *Bioengineered*, **13**, 14282-14298. https://doi.org/10.1080/21655979.2022.2081465

[9] Gong, Y., Lou, Y., Han, X., Chen, K., Zhao, Y., Zhang, H., *et al*. (2024) Serum Proteomic Profiling of Precancerous Gastric Lesions and Early Gastric Cancer Reveals Signatures Associated with Systemic Inflammatory Response and Metaplastic Differentiation. *Frontiers in Molecular Biosciences*, **11**, Article 1252058. https://doi.org/10.3389/fmolb.2024.1252058

[10] Bao, M., Xu, J. and Liu, J. (2022) Overexpression of APOA2 Is Associated with Progression and Poor Prognosis in Gastric Cancer. https://doi.org/10.21203/rs.3.rs-1361258/v1

[11] Xu, P., Zhang, Q., Zhai, J., Chen, P., Deng, X., Miao, L., *et al*. (2025) APOA1 Promotes Tumor Proliferation and Migration and May Be a Potential Pan-Cancer Biomarker and Immunotherapy Target. *Translational Oncology*, **55**, Article ID: 102344. https://doi.org/10.1016/j.tranon.2025.102344

[12] Huang, H., Han, Y., Gao, J., Feng, J., Zhu, L., Qu, L., *et al*. (2013) High Level of Serum AMBP Is Associated with Poor Response to Paclitaxel-Capecitabine Chemotherapy in Advanced Gastric Cancer Patients. *Medical Oncology*, **30**, Article No. 748. https://doi.org/10.1007/s12032-013-0748-8