

This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Screening of novel serum biomarkers for gastric cancer in coastal populations using a protein microarray

yi, yongdong; Rubin, Nan; Lu, Jianhua; Liang, Danna; Zhao, Shengsheng; Wang, Xuhui; Zhang, Hongbo; Chen, Bozhen; Chen, Jiangnan; Zheng, Zhiqiang; You, Tao; Chen, Tanzhou; Chen, Xiaodong; Wang, Wenqian; Lin, Limiao; Chen, Yiming; Liu, Shuai; Huang, Yinpeng; Yu, Yaojun; Lu, Mingdong; Li, Pihong; Huang, He; Zhou, Gongting; Lin, Xianhui; Wu, Hao; shen, xian; Sun, Weijian

Published in:
Cancer Science

DOI:
[10.1111/cas.15876](https://doi.org/10.1111/cas.15876)

Published: 01/08/2023

Document Version
Final published version

Document License
CC BY

[Link to publication](#)

Please cite the original version:

yi, Y., Rubin, N., Lu, J., Liang, D., Zhao, S., Wang, X., Zhang, H., Chen, B., Chen, J., Zheng, Z., You, T., Chen, T., Chen, X., Wang, W., Lin, L., Chen, Y., Liu, S., Huang, Y., Yu, Y., ... Sun, W. (2023). Screening of novel serum biomarkers for gastric cancer in coastal populations using a protein microarray. *Cancer Science*, 114(8), 3396-3410. <https://doi.org/10.1111/cas.15876>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Screening of novel serum biomarkers for gastric cancer in coastal populations using a protein microarray

Yongdong Yi¹  | Rubin Nan¹  | Jianhua Lu¹ | Danna Liang¹ | Shengsheng Zhao¹ | Xuhui Wang¹ | Hongbo Zhang² | Bozhen Chen¹ | Jiangnan Chen¹ | Zhiqiang Zheng¹ | Tao You¹ | Tanzhou Chen³ | Xiaodong Chen³ | Wenqian Wang¹  | Limiao Lin¹ | Yiming Chen¹ | Shuai Liu⁴ | Yinpeng Huang¹ | Yaojun Yu¹ | Mingdong Lu¹ | Pihong Li¹ | He Huang¹ | Gongting Zhou¹ | Xianhui Lin¹ | Hao Wu¹ | Xian Shen³  | Weijian Sun¹ 

¹The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China

²Pharmaceutical Sciences Laboratory, Åbo Akademi University, Turku, Finland

³Wenzhou Medical University First Affiliated Hospital, Wenzhou, China

⁴Taizhou Hospital of Zhejiang Province, Zhejiang Linhai, China

Correspondence

Weijian Sun, Department of Gastrointestinal Surgery, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Zhejiang Wenzhou 325000, China.
Email: fame198288@126.com

Xian Shen, Department of General Surgery, Wenzhou Medical University First Affiliated Hospital, Zhejiang Wenzhou 325000, China.
Email: shenxian1120@126.com

Hao Wu, Department of Gastroenterology, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Zhejiang Wenzhou 325000, China.
Email: walkerhao@126.com

Funding information

Key Laboratory of Minimally Invasive Techniques & Rapid Rehabilitation of

Abstract

Gastric cancer (GC) has high rates of morbidity and mortality, and this phenomenon is particularly evident in coastal regions where local dietary habits favor the consumption of pickled foods such as salted fish and vegetables. In addition, the diagnosis rate of GC remains low due to the lack of diagnostic serum biomarkers. Therefore, in this study, we aimed to identify potential serum GC biomarkers for use in clinical practice. To identify candidate biomarkers of GC, 88 serum samples were first screened using a high-throughput protein microarray to measure the levels of 640 proteins. Then, 333 samples were used to validate the potential biomarkers using a custom antibody chip. ELISA, western blot, and immunohistochemistry were then used to verify the expression of the target proteins. Finally, logistic regression was performed to select serum proteins for the diagnostic model. As a result, five specific differentially expressed proteins, TGF β RIII, LAG-3, carboxypeptidase A2, Decorin and ANGPTL3, were found to have the ability to distinguish GC. Logistic regression analysis showed

Abbreviations: ACG, atrophic chronic gastritis; AGC, advanced gastric cancer; ANGPTL3, angiopoietin-like proteins 3; AUC, area under the ROC curve; B2M, Beta2-microglobulin; BC, breast cancer; bIG-H3, Transforming growth Factor-Beta-Induced Protein/TGF β I; BMI, body mass index; CA9, carbonic anhydrase 9; CPA2, carboxypeptidase A2; CCR, colorectal cancer; DCN, Decorin; DEP, differential expression protein; EGC, early gastric cancer; ErbB2, human epidermal growth factor receptor 2/HER2; EDA-A2, Ectodysplasin A2; FC, fold change; GH, growth hormone; GDNF, Glial cell-derived neurotrophic factor; GH, Growth Hormone; HC, healthy controls; HCC, hepatocellular carcinoma; IACUC, Institutional Animal Care and Use Committee; IHC, Immunohistochemistry; iTRAQ, relative and absolute quantitation; LAG-3, lymphocyte activating gene 3; LH, luteinizing hormone; LOX-1, Lectin-like oxidized low-density lipoprotein receptor-1; MPIF-1, myeloid progenitor inhibitory factor-1; NPV, negative predictive value; PCNA, proliferating cell nuclear antigen; PPV, positive predictive value; qRT-PCR, quantitative reverse transcription polymerase chain reaction; RGM-B, Repulsive guidance molecule B; ROC, receiver operating characteristic; SAM, Significance Analysis of Microarrays; TGF β RIII, transforming growth factor- β receptor III; VCAM-1, vascular cellular adhesion molecule-1.

Yongdong Yi, Rubin Nan, Jianhua Lu contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

Digestive System Tumor of Zhejiang Province, Grant/Award Number: 21SZDSYS04; Lin He's New Medicine and Clinical Translation Academician Workstation Research Fund, Grant/Award Number: 18331202 and 19331104; Medical Technology and Education of Zhejiang Province of China, Grant/Award Number: 202045547; National Natural Science Foundation of China, Grant/Award Number: 81972261; Natural Science Foundation of Zhejiang Province, Grant/Award Number: LY18H030009; Wenzhou Medical University, Grant/Award Number: KYW202006; Wenzhou Science and Technology Plan Project, Grant/Award Number: 2020Y0593, Y20210189, ZY2020007 and ZY2021003

that the combination of carboxypeptidase A2 and TGF β RIII had superior potential for diagnosing GC (area under the ROC curve [AUC]=0.801). The results suggested that these five proteins alone and the combination of carboxypeptidase A2 and TGF β RIII may be used as serum markers for the diagnosis of GC.

KEYWORDS

biomarkers, gastric cancer, protein microarray, proteomics, serum

1 | INTRODUCTION

Gastric cancer (GC) is a malignancy that arises from the epithelial cells of the gastric mucosa and is the most frequent digestive tract tumor. It is also a prevalent cancer globally, with high morbidity and mortality rates.^{1,2} In coastal areas, the prevalence of GC remains high, but the rate of clinical diagnosis remains low. Presently, pathological examination based on gastroscopy is the standard for GC diagnosis.³ However, this invasive procedure is time consuming and cumbersome.^{3,4} In contrast, a serological examination can detect the existence of malignant tumors at an early stage, when the number of tumor cells is $\sim 10^8$ or the diameter of the tumor body is ~ 0.5 cm. Furthermore, a serological examination is less invasive, quicker, and more convenient. It can also dynamically detect quantitative changes and is, therefore, easier to use and better accepted by clinicians and patients.⁵ Currently, there are few convenient and accurate serum biomarkers available for clinical diagnosis of GC.^{6,7} Therefore, identifying novel serum biomarkers is crucial to improve the accuracy of GC diagnosis.

In clinical settings, blood samples are often used to initiate the search and discovery of biomarkers. The levels of proteins associated with stomach cancer in the blood can vary by several orders of magnitude.⁸ To overcome this challenge, antibody-based technologies are the primary solution, as they offer higher sensitivity. Planar microarrays, a quantitative array platform technology based on antibodies, require only small samples and are compatible with high-throughput analysis.^{9,10} This technology is particularly useful for detecting cytokines and other serum markers. Therefore, we propose using protein planar microarrays to identify novel serum protein markers in coastal populations for the diagnosis of GC.

In this study, we collected serum samples from three hospitals located on the coast of China and used a novel protein microarray to screen and quantify 640 serum proteins to identify potential serum biomarkers of GC. In total, 28 differentially expressed proteins were identified, and we validated these using a custom protein antibody chip. Five differentially expressed proteins, LAG-3, TGF β RIII, CPA2, DCN and ANGPTL3, showed the ability to distinguish GC from

healthy people. We used ELISAs to detect the expression of these five proteins in the serum of GC mice, and the results were similar to those obtained from human serum. The expression levels of CPA2, DCN, ANGPTL3, and TGF β RIII were elevated in the GC group, whereas LAG-3 expression was decreased. Western blot results in GC cell lines showed that the expression levels of CPNA2, DCN, ANGPTL3, and LAG-3 were elevated, while TGF β RIII expression decreased. We performed IHC detection of these five proteins in an orthotopic transplantation mouse model of an implanted human GC cell line, and the expression levels of the proteins in the tumor tissues were confirmed to be different. Logistic regression analysis of the five proteins indicated that the combination of all five showed the greatest specificity for diagnosing GC, with an AUC of 0.828. The combination of CPA2 and TGF β RIII showed sufficient sensitivity for diagnosing GC, with an AUC of 0.801. Therefore, we believe that these five proteins alone or the combination of CPA2 and TGF β RIII can be used as serum markers for the diagnosis of GC.

2 | MATERIALS AND METHODS

2.1 | Patient information, sample collection and sample preparation

Samples were obtained from three independent subject cohorts for this study: a 640-protein array screening cohort with 88 samples, a validation cohort with, in total, 333 subjects, and a 31-sample IHC cohort. All subjects signed informed consent forms prior to their inclusion in this study. Blood samples were acquired from the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, the First Affiliated Hospital of Wenzhou Medical University, and the Taizhou Hospital of Zhejiang Province. The study was approved by the Institutional Review Board of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, the First Affiliated Hospital of Wenzhou Medical University, and Taizhou Hospital of Zhejiang Province. Patients who had received previous treatment were excluded from this study.

During the sample collection visits from 2020 to 2022, subjects were examined, and data related to age, sex, height, weight, BMI, tumor markers (AFP, CEA, CA125, CA153, and CA199, CA724), CBC, coagulation function, liver and kidney function, and pathology (histological type, depth of invasion, TNM) were recorded.

After registration, 5 mL of whole blood were collected from each subject into a Vacutainer vacuum blood collection tube (Becton, Dickinson and Co., Franklin Lakes, NJ, USA) without anticoagulants, preservatives, or separating agents. The tubes were centrifuged at 1000×g for 10 min, placed in a −20°C freezer for 30 min, and then transferred to −80°C for storage.

2.2 | Protein chips and custom antibody chip screening

A 640 human-protein chip (catalog number QAH-CAA-640, RayBiotech Inc., Peachtree Corners, Georgia, USA) was acquired from RayBiotech. QAH-CAA-640 is a high-throughput antibody chip containing 640 proteins, including inflammation, chemotaxis, growth, angiogenesis, apoptosis, and cycle proteins. For marker screening, high-throughput non-targeted detection was more advantageous. From the protein chip screening, differentially expressed proteins were selected, and biomarkers meeting the following screening criteria were identified as differentially expressed: $\log_2FC > \log_2(1.2)$, adjusted p -value < 0.05 , mean fluorescence signal value per group > 150 , and SAM analysis bias greater than predefined (default 1.2). According to the above criteria, in total, 28 proteins were identified.

For the validation cohort, we designed a custom antibody chip (catalog number QAH-CUST, RayBiotech Inc.) targeting the 28 proteins identified in the primary screening. Each protein was analyzed in quadruplicate on each array. Sera from 333 subjects were analyzed using the custom antibody chips. This group included 39 healthy subjects, 128 cases of early cancer, 80 subjects with advanced GC, 13 cases of gastritis, 22 subjects with liver cancer, 32 patients with colorectal cancer, and 19 cases of breast cancer (Table 1).

Array processing was performed according to the manufacturer's instructions. Briefly, serum samples were diluted 1:100 in Tris-buffered saline containing 0.1% Tween 20 detergent. The protein array was incubated with diluted serum overnight at 4°C. A biotin-labeled antibody was added and incubated for 2 h after extensive washing, and then Alexa Fluor 555-conjugated streptavidin was added to each well for 1 h at room temperature. Wells were washed with Wash Buffer II. The spot intensities of the slides were analyzed using a InnoScan 300 Scanner (Innopsys, Carbonne, France). Q-Analyzer Software (RayBiotech, Inc.) was used to perform data visualization and obtain the protein content of each serum sample.

2.3 | Data processing and figure generation

The data obtained from the protein array screening assay were utilized to create a heatmap that clustered proteins together based on

TABLE 1 Validation cohort for custom antibody chips.

	<i>n</i>	Male	Female	Ages
HC	39	25	14	51.4 ± 14.8
EGC	128	90	38	63.9 ± 9.6
AGC	80	62	18	69.3 ± 10.1
HCC	22	18	4	60 ± 11.3
CCR	32	19	13	61.8 ± 11.7
BC	19	0	19	52.7 ± 12.4
ACG	13	7	6	—

Note: Ages are expressed as the means ± SDs.

Abbreviations: ACG, atrophic chronic gastritis; AGC, advanced gastric cancer; BC, breast cancer; CCR, colorectal cancer; EGC, early gastric cancer; HC, healthy controls; HCC, hepatocellular carcinoma.

their expression patterns. The clustering analysis and heatmap generation data were imported into R for analysis.

After testing and analyzing the significantly differentially expressed proteins, a volcano plot was created with log FC as the abscissa and the negative logarithm of adjusted p -values (adj. p .Val) or p -value as the ordinate. The R.5.5 Comparison plotting function “ggplot2” and the data package *ggfortify* were used for comparison across groups.

2.4 | Cell culture

Three GC cell lines, BGC-823, MKN-45, and AGS, and one human gastric epithelial cell line (GES-1), were used in this study. The cells were cultured in DMEM (HyClone, Thermo Fisher Scientific, Waltham, MA, USA) or RPMI-1640 (HyClone, Thermo Fisher Scientific) containing 10% FBS (Gibco-Invitrogen Corp.) at 37°C in an atmosphere with 5% CO₂.

2.5 | Western blotting

To measure the expression of TGFβ RIII, LAG-3, CPA2, DCN and ANGPTL3, cells were washed with PBS, and total proteins were isolated from cell lines using RIPA buffer (Beyotime). A sample of total proteins (~30 μg) extracted from each cell line was boiled at 100°C for 5 min and then separated using SDS-PAGE on a 12% polyacrylamide gel. The separated protein bands were transferred onto a PVDF membrane (Millipore) that was subsequently blocked with Quickblock Blocking Buffer (Beyotime; Cat#P0252). The membrane was then incubated overnight with primary antibodies at 4°C. After incubation, the membrane was washed three times with TBST buffer and then incubated with an HRP-conjugated secondary antibody (1:5000 dilution; Beyotime) for 2 h at room temperature. The membrane was then washed three more times with TBST buffer, and the protein blots were visualized using ECL-Plus reagent (Millipore). GAPDH was used as a loading control in all western blot studies. The antibodies used for the western blot studies were as follows: anti-ANGPTL3 (1:1000 dilution; Proteintech; Cat#11964-1-AP), anti-DCN

(1:1000 dilution; ABclonal; Cat#A1669), anti-CPA2 (1:1000 dilution; Santa Cruz; Cat#sc-515,450), anti-TGF β RIII (1:1000 dilution; Santa Cruz; Cat#sc-74,511), anti-LAG-3 (1:1000 dilution; Proteintech; Cat#80867-1-RR), and anti-GAPDH (1:1000 dilution; Proteintech; Cat#60004-1-Ig).

2.6 | Gastric cancer models

All animal experiments were conducted in accordance with the guidelines of the IACUC. The Institutional Animal Care and Use Committee of Zhejiang Laboratory Animal Center approved all experiments involving animals. Nude mice were divided randomly into two groups: GC and healthy, with seven animals per group. The midabdominal line was opened to expose the stomach, and the filling site was chosen as the part of the greater curvature of the stomach with fewer blood vessels. Micro-shearing was used to create a cavity of ~2 mm, into which BGC-823 cells were tamped. The opening was closed, and the abdomen was closed layer by layer. The entire wound was disinfected. Healthy mice were treated with 3% pentobarbital sodium. During the 7-day experimental period, mice were fed with normal mouse chow. At 1 week after surgery, tumor growth was monitored using *in vivo* fluorescence imaging. Tumor tissue specimens were collected from mice for use in IHC assays, and mouse sera were collected for use in ELISA validation.

2.7 | IHC analysis of tissue samples

After the GC specimens were harvested, they were fixed in formalin for 48 h at 4°C. The collected specimens were wrapped in paraffin and fixed, and then cut into 5- μ m thick slices. The slide-mounted tissue specimens were incubated in xylene for 5 min, washed twice in 100% ethanol for 10 min and then washed in 95% ethanol for 10 min. Antigen unmasking was performed, and the slides were then blocked with 3% hydrogen peroxide for 30 min at room temperature. Subsequently, the sections were incubated overnight with the primary antibody against proliferating cell nuclear antigen (PCNA) at 4°C. Finally, the tissue sections were stained with DAB chromogen, and the cell nucleus was stained with hematoxylin. LAG-3, TGF β RIII, ANGPTL3, DCN, and CPA2 were detected following a standard H&E protocol. Visualization of stained nuclei was performed using DAB staining and the slides were digitally scanned using an Olympus IX51 microscope (Olympus, Tokyo, Japan).

2.8 | Gastric cancer tissue microarrays

Gastric cancer-specific tissue microarrays (TMAs) containing 31 different paired samples (GC tissues and paracancerous tissues) were used in this study. The tissue samples in this study were all collected from the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University. All patients signed informed

consent forms and the collection of tissue samples for research was approved by the ethics committee.

2.9 | ELISA validation

To verify the serum expression results of TGF β RIII, LAG-3, CPA2, DCN and ANGPTL3, serum samples from GC mice were analyzed using a TGF β RIII, LAG-3, CPA2, DCN and ANGPTL3 detection kit. Briefly, the serum was diluted two-fold and incubated overnight on ELISA plates at 4°C. After washing the plate, the 80-fold diluted antibody was added to the device and mixed with the serum, which was then incubated for 1 h at room temperature. After another washing, HRP-streptavidin was added to the plate wells, and the reaction was stopped by adding sulfuric acid. The optical density was measured using a microplate reader (Biotek, Winooski, VT, USA, ELx800NB), and the concentration value was calculated using SigmaPlot 12.0 software.

2.10 | Statistics

All data in the present study were plotted and analyzed using GraphPad Prism 9, Microsoft Excel or R, and moderated *t*-statistics were used to analyze the data. SAM was used to identify differentially expressed biomarkers or genes. SAM compares a *t*-like statistic, *d*_i, observed across groups with an estimate obtained from randomized permutations of the samples. Biomarkers with deviations larger than the predefined threshold (default 1.2) were identified as differentially expressed. SAM analysis was performed using the R programming language V3.6.3 and implemented using R package *siggenes* 1.60.0. Sensitivity, specificity, and AUC were calculated using *easyROC* software.¹¹ To reflect the diagnostic value of different differential expression protein (DEP) combinations, we performed logistic regression models on the AUC.

3 | RESULTS

3.1 | Study design

In the present study, we performed an integrated analysis of the proteome in the serum of GC patients (Figure 1). Protein chip screening was performed using 640-human-protein chips to analyze 88 serum samples to identify serum proteins expressed differentially between GC patients and healthy controls. In total, 28 target proteins were identified in the microarray. Subsequently, a customized antibody chip was used to validate the 28 candidate proteins, and the results showed five proteins that were expressed differentially in the serum samples. ELISA validation was performed to verify the GC mouse serum expression results of these five proteins. Western blotting studies were conducted to determine the expression of TGF β RIII, LAG-3, CPA2, DCN, and ANGPTL3 in different cell lines. IHC was used to confirm the potential clinical relevance of the relationship between serum protein

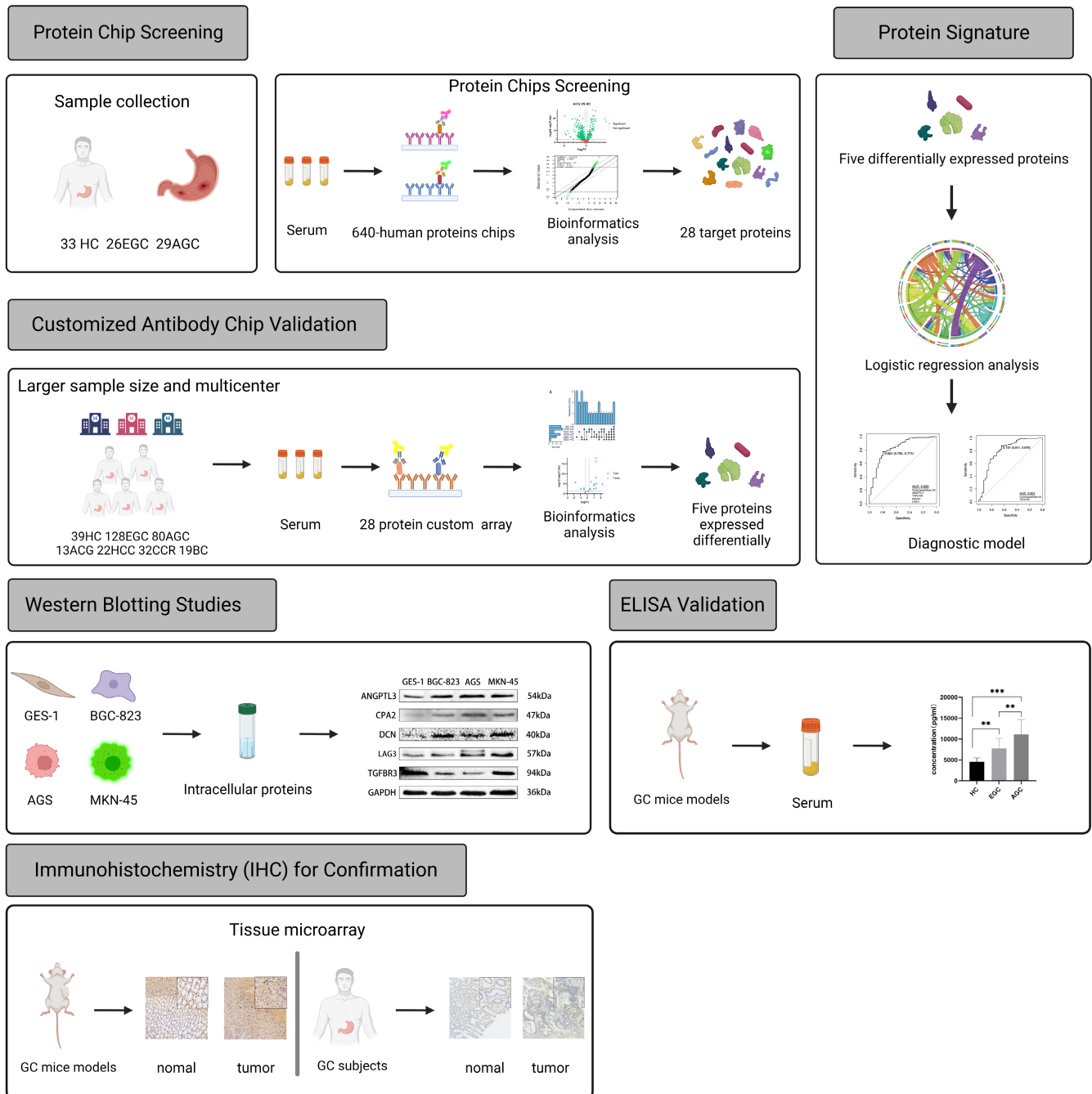


FIGURE 1 Workflow of the study.

levels and the protein levels in tissues. Logistic regression analysis was performed to design satisfactory diagnostic models, and the combination of CPA2 and TGF β RIII showed potential for diagnosing GC.

3.2 | Serum proteomic analysis of GC patients

For the primary screening of candidate biomarkers, serum samples from 88 subjects were analyzed using 640-human proteins chips. This sample group, acquired from the Second Affiliated

Hospital and Yuying Children's Hospital of Wenzhou Medical University, included 33 HCs, 26 cases of early GC and 29 cases of advanced GC. Overall, 64% of the subjects were male and 36% female. Their ages ranged from 20 to 86 years (Table 2). The results of the 88 samples for 640 proteins were used to build the principal component analysis (PCA) plots to show the differences between HC, patients with EGC, and patients with AGC (Figure 2A). As expected, there were differences in protein expression among the three groups. Then expression screening of the proteins using moderated *t*-statistics yielded the protein

TABLE 2 Gastric cancer cohort used for protein array screening.

Variables	HC n=33	EGC n=26	AGC n=29
Sex, n (%)			
Male	20 (22.7%)	16 (18.2%)	22 (25.0%)
Female	13 (14.8%)	10 (11.4%)	7 (8.0%)
Age			
Range	49.33±13.39	63.04±9.64	65.48±11.76
BMI			
Underweight	0	2	1
Normal	21	15	17
Overweight	12	9	11

Note: The normal range of BMI for males is 18.5–24, with lower than 18.5 being underweight and higher than 24 being overweight. The BMI normal range for females is 17.5–23.5, with lower than 17.5 being underweight and higher than 23.5 being overweight. Ages are expressed as means ± SDs.

Abbreviations: AGC, advanced gastric cancer; BMI, body mass index; EGC, early gastric cancer; HC, healthy controls.

concentration multiples and adj.*p*.Val of the GC and HC, which were used to build a volcano plot (Figure 2B,C). Differentially expressed proteins were defined as $|\log_2FC| > 0.263$ and an adj.*p*.Val < 0.05 . The results showed that there were nine differentially expressed proteins between EGC and HC, and 354 differentially expressed proteins between AGC and HC.

Next, we used more stringent criteria to determine which serum proteins should be selected to produce custom chips for further validation. We conducted a more comprehensive statistical analysis of the results to eliminate errors and exclude differentially expressed proteins with poor effect. Using SAM, a permutation-based analysis method, we compared HC with EGC and AGC, as shown in Figure 2D,E. In total, 28 target proteins were identified through statistical comparison and SAM analysis (Table 3), including 25 EGC targets and 15 AGC targets (Tables S1 and S2), with 12 identical target proteins. These 28 target proteins were then subjected to further validation with a larger sample size.

3.3 | Selection of protein candidates using a customized antibody chip

We customized specific antibody arrays to evaluate the reliability of the 28 novel biomarkers obtained from the initial screening. Sera from 333 subjects, including 39 healthy subjects, 128 cases of early GC, 80 subjects with advanced GC, 13 cases of gastritis, 22 subjects with liver cancer, 32 patients with colorectal cancer, and 19 cases of breast cancer, were analyzed using the custom antibody chips. Serum samples from 11 early GC patients and 40 advanced GC patients were collected at the Taizhou Hospital of Zhejiang Province, while serum samples from 26 early GC patients and 18

advanced GC patients were collected at the First Affiliated Hospital of Wenzhou Medical University. The remaining serum samples were obtained from the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University (Table 1). The results of protein expression detection were analyzed using moderated *t*-statistics and were plotted as volcano plots (Figure 3A,B). The protein expression results were plotted as PCA plots and heatmaps (Figure 3C–F). The results showed that increasing the sample size yielded results similar to those of the first screening, but with clearer intergroup differences.

3.4 | Identification of potential biomarkers for GC

In the validation using customized antibody chips, we simultaneously tested serum samples from patients with HCC, BC, CCR, and ACG (Figure S1, Table S3). After significance analysis of the data, we drew an UpSet diagram (Figure 4A) to summarize the experimental results, which clearly showed the comparison between the results of different proteins in each group. Three proteins, RGM-B, LAG-3, and TGFβ RIII, showed heterogeneity in AGC. Three proteins, CPA2, DCN, and ANGPTL3, showed heterogeneity in GC (Figure 4B–F). After correlation analyses of the data from two protein arrays, RGM-B showed an opposite trend and was excluded from the study. The AUC values (Table 4) obtained using ROC analysis for the remaining five proteins also reflected the high specificity and sensitivity of these proteins. Therefore, we believe that these five proteins can be used as potential biomarkers for GC.

3.5 | Validation of the five potential biomarkers using ELISA

To verify the results of the protein array, we established a gastric orthotopic transplantation model in nude mice with GC and collected their serum for analysis using ELISA kits for the five potential biomarkers. The results showed that the expression levels of LAG-3, TGFβ RIII, CPA2, DCN, and ANGPTL3 were similar to those obtained from the protein microarray, and the differences were statistically significant ($p < 0.05$; Figure 5).

3.6 | The target proteins were expressed differently in cell lines

To determine the expression of TGFβ RIII, LAG-3, CPA2, DCN and ANGPTL3 in cell lines, proteins were isolated from GC cell lines (BGC-823, MKN-45 and AGS) and one human gastric epithelial cell line (GES-1). The western blot results showed that the expression levels of CPA2, DCN and ANGPTL3, LAG-3 were elevated in GC cell lines, while TGFβ RIII expression was downregulated. Expression levels of five proteins were similar to the IHC results (Figure 6).

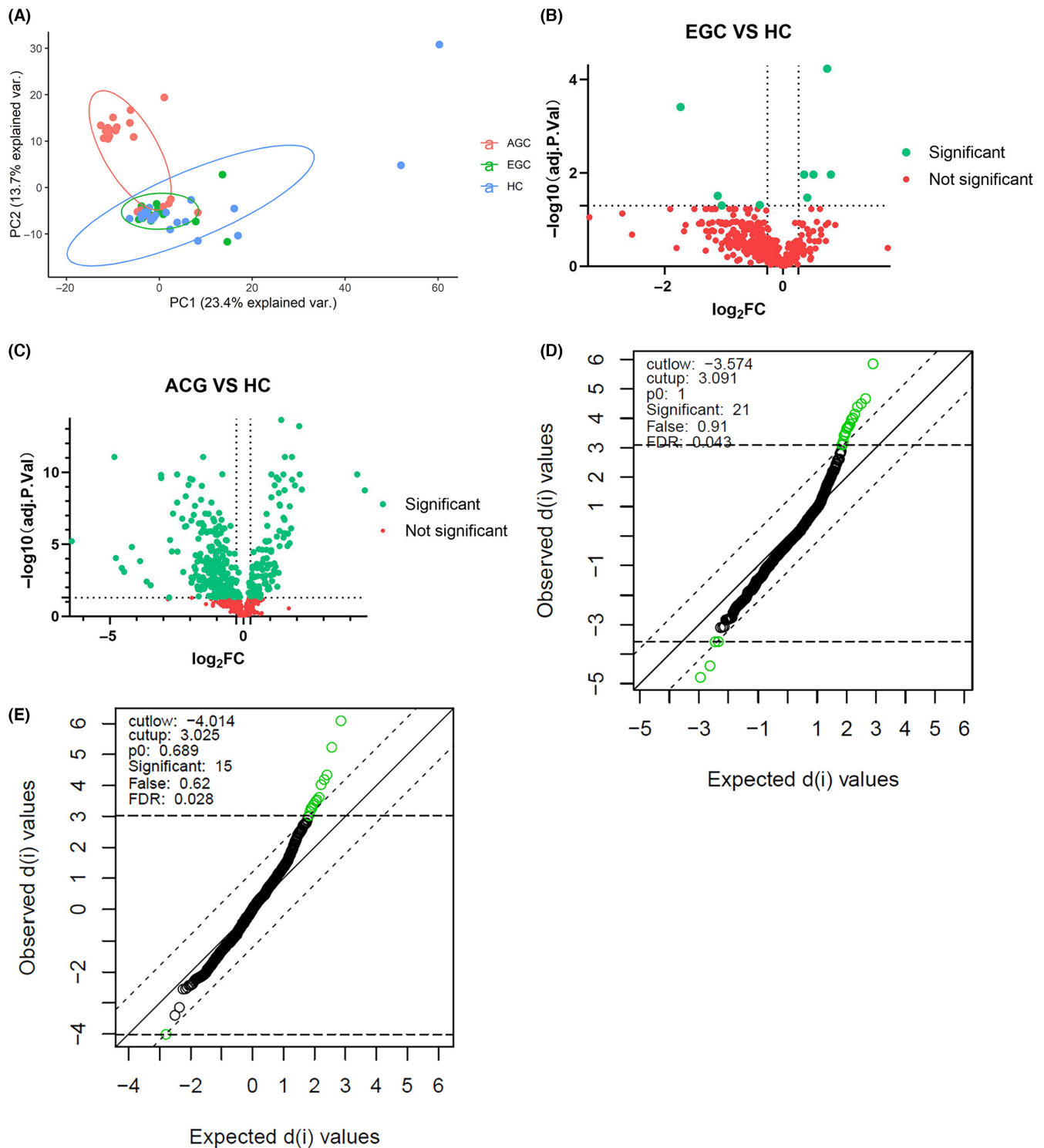


FIGURE 2 A-E, The screening results of the 640-serum protein array of the patients with GC and HC.

3.7 | The expressions of TGF β RIII, LAG-3, CPA2, DCN and ANGPTL3 in tumor tissues

Immunohistochemical analysis was performed on GC tissues to examine the potential clinical relevance of the relationship between serum protein levels and the protein levels in tissues. We established a gastric orthotopic transplantation model in nude mice with GC,

and obtained orthotopic tumor tissues, on which IHC was performed (Figure 7A,B,D) to detect the expression of LAG-3, TGF β RIII, CPA2, DCN and ANGPTL3. The results of CPA2, DCN and ANGPTL3 were similar to their expression in serum, as these proteins showed higher expression levels in GC tissues than in healthy tissues. However, the expression levels of TGF β RIII and LAG-3 in tumor tissues were contrary to the serum results. At the same time, IHC analysis was

TABLE 3 Expression trend results of 28 differentially expressed proteins in the primary screening and validation cohorts.

AGC vs. HC Protein	Log FC	
	Primary screening	Validation cohorts
ErbB2 (human epidermal growth factor receptor 2)	-0.351	-1.066
LAG-3	-2.949	-4.755
LH (luteinizing hormone)	-0.300	0.492
LOX-1 (lectin-like oxidized low-density lipoprotein receptor-1)	-0.251	-4.410
MPIF-1 (myeloid progenitor inhibitory factor-1)	-0.500	0.918
ANGPTL3 (angiopoietin-like protein 3)	1.050	1.367
B2M (beta2-microglobulin)	0.011	1.207
bIG-H3 (transforming growth factor-beta-induced protein/TGFBI)	0.128	1.002
CA9 (carbonic anhydrase 9)	1.372	1.770
Cathepsin B	0.416	2.125
Cystatin B	0.158	0.344
Decorin	0.386	0.302
EDA-A2 (ectodysplasin A2)	0.954	1.260
GDNF (glial cell-derived neurotrophic factor)	0.023	-2.862
P-Cadherin	0.028	0.485
RGM-B (repulsive guidance molecule B)	2.229	-0.648
TGFβ RIII (transforming growth factor-β receptor III)	1.298	1.823
Trappin-2	0.135	1.062
VCAM-1	1.277	2.032

Abbreviations: AGC, advanced gastric cancer; FC, fold change; HC, healthy control; ErbB2, human epidermal growth factor receptor 2/HER2; LAG-3, lymphocyte activating gene 3; LH, luteinizing hormone; LOX-1, lectin-like oxidized low-density lipoprotein receptor-1; MPIF-1, myeloid progenitor inhibitory factor-1; ANGPTL3, angiopoietin-like protein 3; B2M, beta2-microglobulin; bIG-H3, transforming growth factor-beta-induced protein/TGFBI; CA9, carbonic anhydrase 9; EDA-A2, ectodysplasin A2; GDNF, glial cell-derived neurotrophic factor; RGM-B, repulsive guidance molecule B; TGFβ RIII, transforming growth factor-β receptor III; VCAM-1, vascular cellular adhesion molecule-1.

performed on the GC microchips. The results for LAG-3, TGFβ RIII, CPA2, DCN and ANGPTL3 were similar to the results for mouse tumor tissues (Figure 7C,E).

3.8 | The combination of TGFβ RIII, LAG-3, CPA2, DCN and ANGPTL3 exhibits sufficient sensitivity for diagnosing GC

ROC curves were drawn using TGFβ RIII, LAG-3, CPA2, DCN and ANGPTL3 (Figure S3). Among them, CPA2 showed the highest AUC (0.775; Figure 8A). We then combined multiple proteins into

panels to evaluate if any combination of blood proteins exhibited further improvement in diagnostic potential. Using logistic regression analysis, we found that a combination of five proteins had the highest diagnostic specificity (AUC=0.828) in distinguishing GC from HC. The combination of CPA2 and TGFβ RIII showed sufficient sensitivity for diagnosing GC, with an AUC of 0.801 (Figure 8B-F). The results of other combinations are shown in Figure S3.

4 | DISCUSSION

Epidemiological studies had shown that Zhejiang, a seaside capital in China, has a high morbidity rate for GC.^{12,13} Currently, the main method for GC screening is upper gastrointestinal endoscopy, which requires advanced instruments and specialized operators, making it unsuitable for repeated examinations and population-wide screenings. Other screening methods, such as *Helicobacter pylori* detection, serum pepsinogen detection, and gastrin detection, have high false-positive rates and low sensitivity, necessitating further research. Also, there have been recent studies on the detection of GC biomarkers with different technologies. For instance, Shao et al. used circ-RNA microarray and found the AUC of circ-RNA hsa_circ_0000190 for GC diagnosis was 0.799.¹⁴ In addition, Wang et al. identified serum-circulating exosomes using qRT-PCR and the ROC analyses yielded the AUC values of 0.786 for miR-106a-5p and 0.769 for miR-19b-3p.¹⁵ Lee et al. reported an AUC of 0.797 for DEK using isobaric tags for iTRAQ.¹⁶ Serological examination is less harmful, faster, and more convenient. In this study, we used protein arrays to detect serum proteins from a seaside population and found that the diagnosis rate was improved by multibiomarker joint analysis, with an AUC value of 0.801.

At the initial screening stage, we used this platform to screen the serum of GC patients and identified 28 differentially expressed proteins. In the validation set, we custom-built a special glass-based antibody array capable of detecting all 28 proteins simultaneously, based on the primary screening results. Five biomarkers obtained by custom antibody chip, CPA2, ANGPTL3, TGFβ RIII, DCN and LAG-3, exhibited the best AUC values of 0.775, 0.747, 0.717, 0.639, and 0.618, respectively. Among the five targets, we could not find any literature related to GC for TGFβ RIII or CPA2 proteins. Finally, we used logistic regression analysis to combine multiple proteins into panels, and the combination of CPA2 and TGFβ RIII showed sufficient sensitivity for diagnosing GC, with an AUC value of 0.801.

Lymphocyte activating gene 3 (LAG3) is a membrane glycoprotein that belongs to the immunoglobulin (Ig) superfamily and functions as an immune checkpoint receptor. Soluble LAG3 (sLAG3) is most likely to be derived from proteolytic cleavage of surface LAG3, as reported in previous studies.^{17,18} It has been reported that non-cleavable LAG-3 has a potent inhibitory effect on T-cell proliferation and cytokine production, and cell surface cleavage serves as an important negative feedback mechanism to moderate its

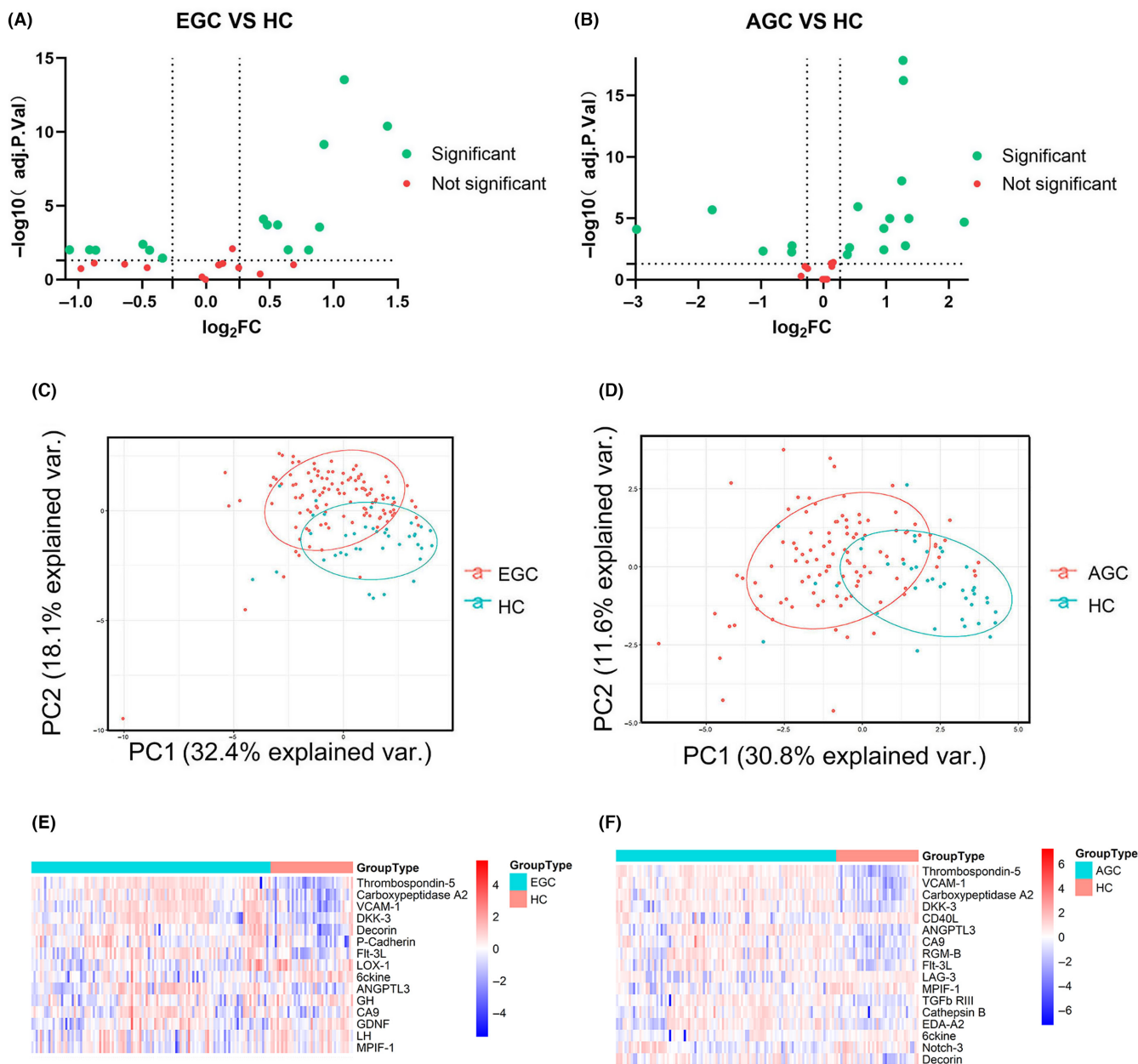


FIGURE 3 A–F, A custom antibody chip was used to validate 28 differentially expressed proteins.

function.¹⁷ We speculate that GC tumors may produce more non-cleavable LAG-3, resulting in lower levels of sLAG3 in the serum, although further research is needed to confirm this hypothesis. Lv et al. reported that, compared with peritumoral tissues, tumor tissues contained more intensive LAG-3⁺ cell infiltration than peritumoral tissues in each subtype of GC patients,¹⁹ which is consistent with our results for tumor tissues.

Transforming growth factor- β receptor III is encoded by the *TGF β RIII* gene on chromosome 1p31–32. Previous studies have shown that the role of TGF β RIII in tumors is complex and varies depending on the type and stage of the cancer. TGF β RIII has two different forms: dissolved in body fluids and expressed on cell membranes. When it is on the cell membrane, it mainly assists in activating the TGF β 1 signaling pathway to inhibit tumor growth and metastasis,

and acts as a tumor suppressor gene. When it is dissolved in body fluids, it mainly stimulates tumor growth and metastasis, and has the effect of promoting oncogenes. For example, TGF β RIII has tumor-suppressive effects in breast cancer, and has tumor-promoting effects in colorectal cancer and glioma.^{20–22} This may be the reason for the different expression of TGF β RIII in serum and tissues of GC.

CPA2 is a secreted pancreatic procarboxypeptidase that cleaves C-terminal amide or ester bonds of peptides with free terminal carboxyl groups. It has been reported that CPA2 is downregulated in pancreatic cancer and plays a fundamental role in the pathogenesis of acute or chronic pancreatitis. In addition, inflammation in acute pancreatitis could upregulate and activate CPA2.²³

DCN is a prototype member of the small leucine-rich proteoglycan (SLRP) family, which is expressed in various tissues and found in

FIGURE 4 A–F, Characteristics of five specific differentially expressed serum proteins and their expression in each group.

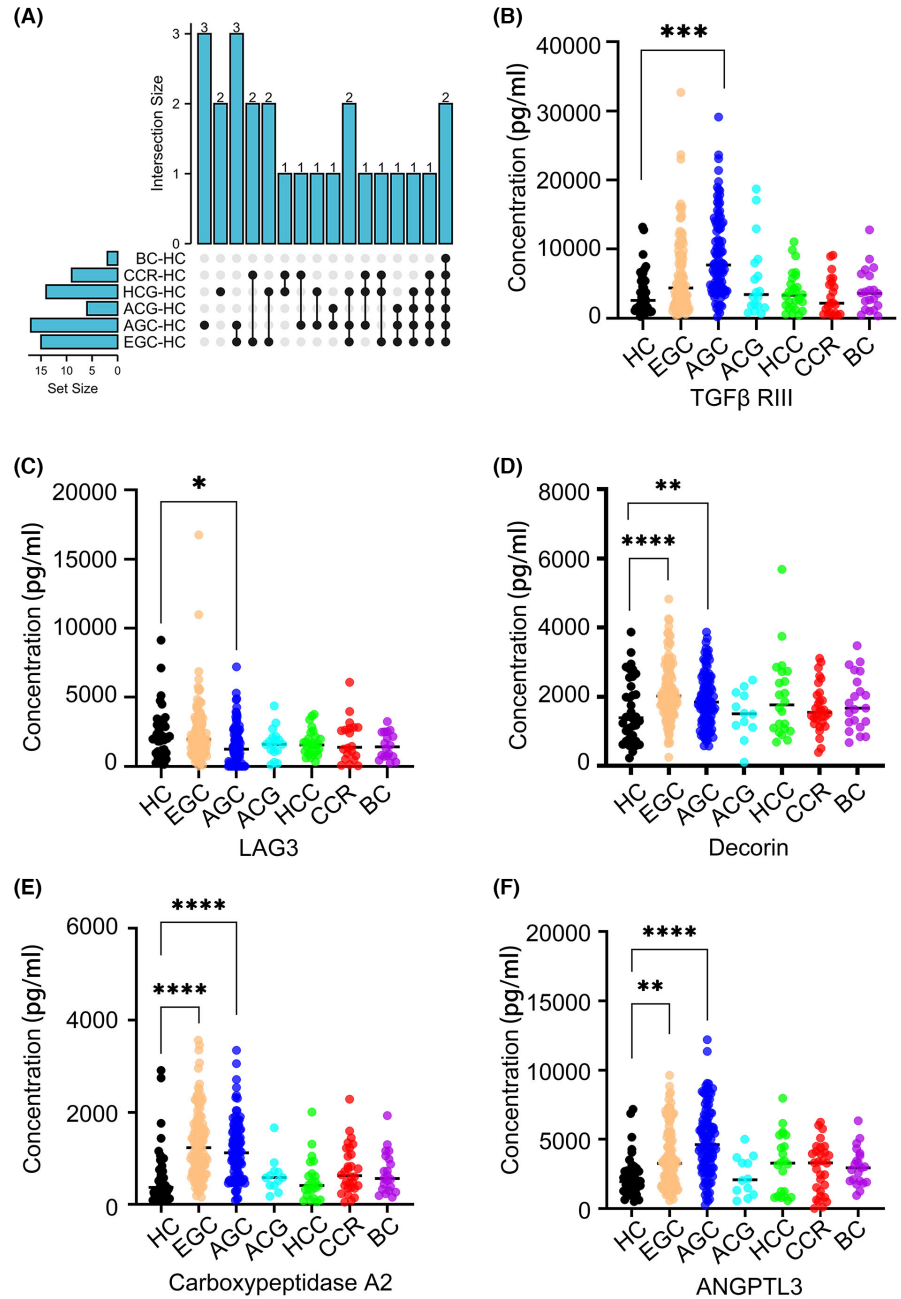


TABLE 4 Statistical features of five differentially expressed proteins.

	LAG-3	TGFβ RIII	Carboxypeptidase A2	Decorin	ANGPTL3
Specificity	0.635	0.827	0.933	0.663	0.683
Sensitivity	0.718	0.641	0.538	0.59	0.872
PPV (positive predictive value)	0.959	0.918	0.913	0.908	0.915
NPV (negative predictive value)	0.256	0.568	0.260	0.277	0.222
Accuracy	0.594	0.867	0.668	0.703	0.574
AUC	0.688	0.797	0.76	0.697	0.809

Abbreviation: AUC, area under the ROC curve; PPV, positive predictive value; NPV, negative predictive value.

the matrix of various cancers. Originally identified as a natural inhibitor of TGFβ, DCN is also an receptor tyrosine kinase (RTK) inhibitor. It triggers cell cycle arrest and apoptosis, as well as antimetastasis

and antiangiogenesis processes.²⁴ Basak et al. reported that DCN was an effective inhibitor of the TGFβ pathway and was negatively correlated with the GC stage.²⁵

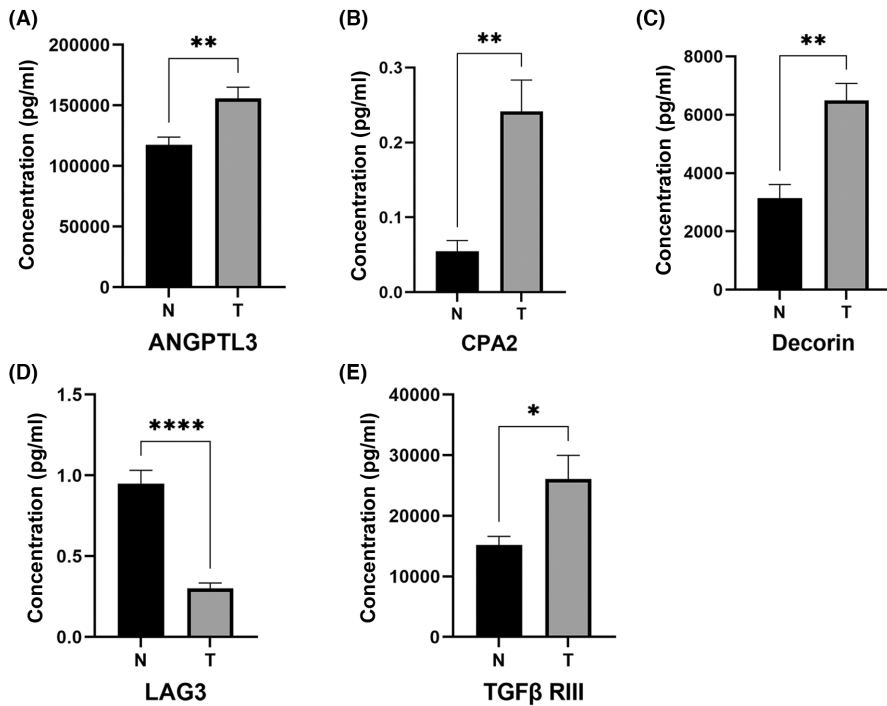


FIGURE 5 A-E, Expression of target proteins in serum of gastric cancer mice by ELISA.

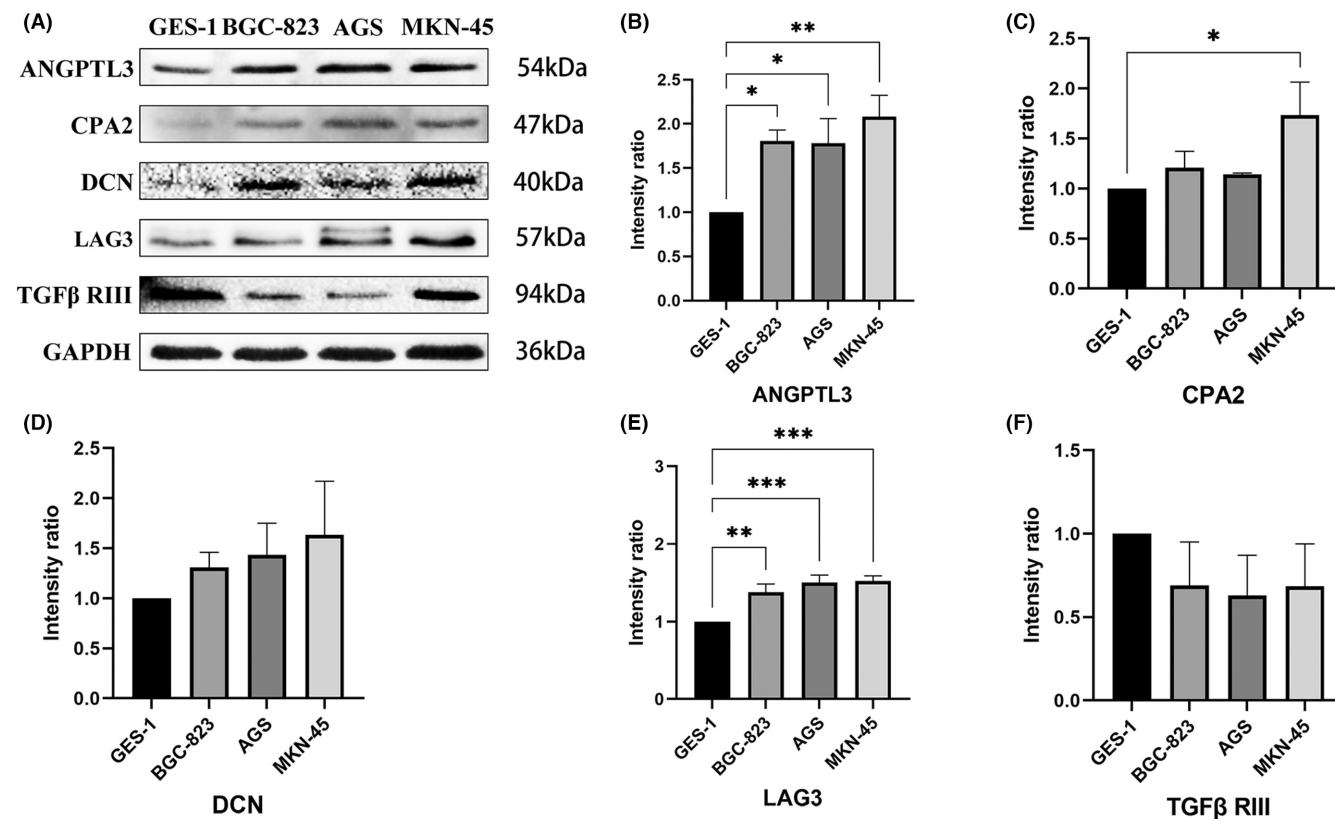


FIGURE 6 A-F, Expression levels of five target proteins in gastric cell lines.

ANGPTL proteins are a series of secreted glycoproteins with structures similar to angiotensin. ANGPTL3 has been reported to be significantly upregulated in cell lines derived from oral squamous cell carcinoma (OSCC) compared with normal tissues.²⁶ The expression and serum levels of ANGPTL3 may be promising non-invasive biomarkers in the diagnosis of chronic hepatitis and HCC.²⁷ High expression of ANGPTL3 has

been found to be associated with shorter overall survival time in GC patients.²⁸

These targets have great prospects as tumor markers of GC. One or a combination of several proteins may have higher sensitivity and specificity than current tumor markers for the diagnosis of GC.

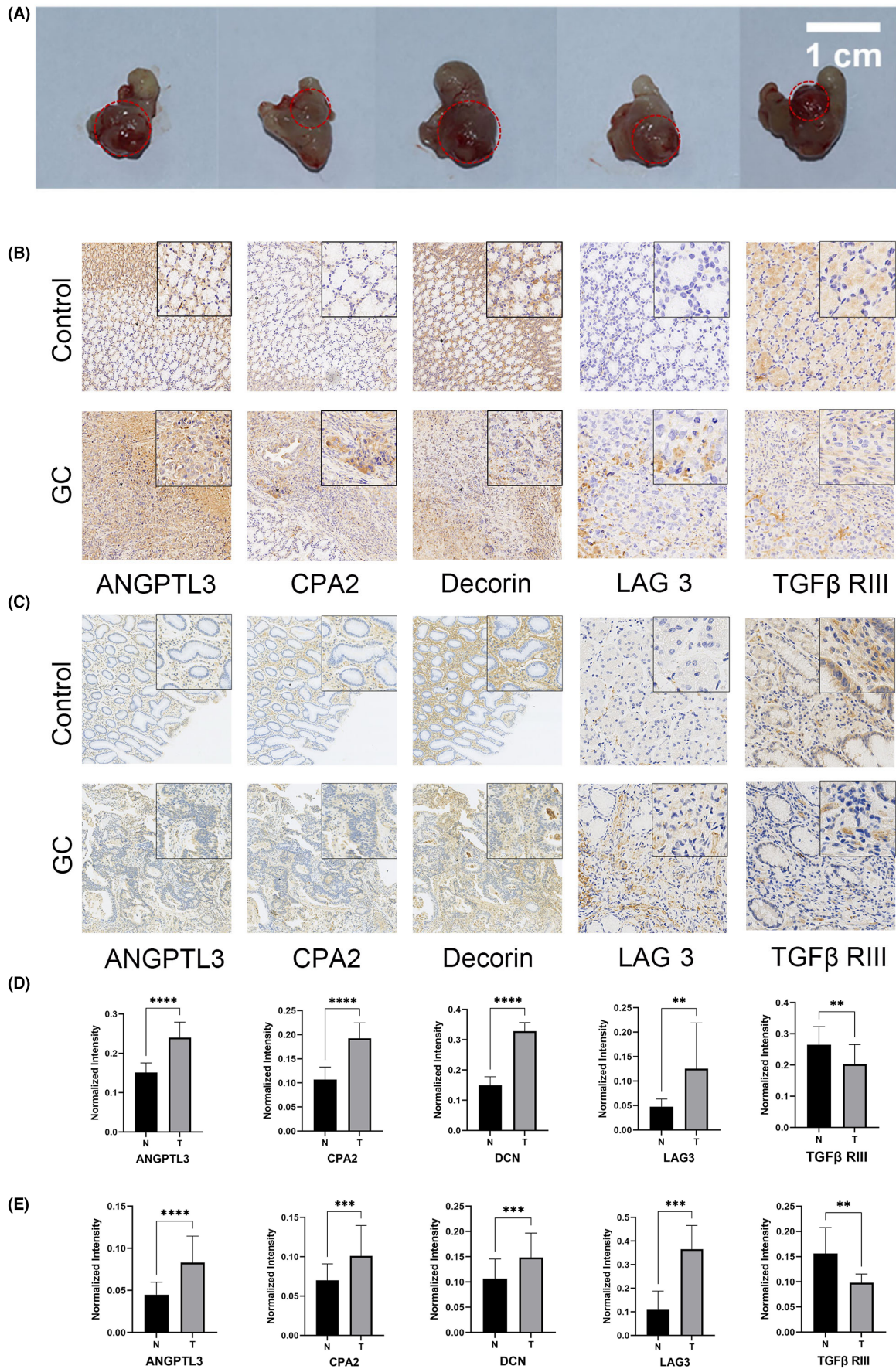


FIGURE 7 A-E, IHC results of five differentially expressed proteins in mouse and human gastric cancer sections.

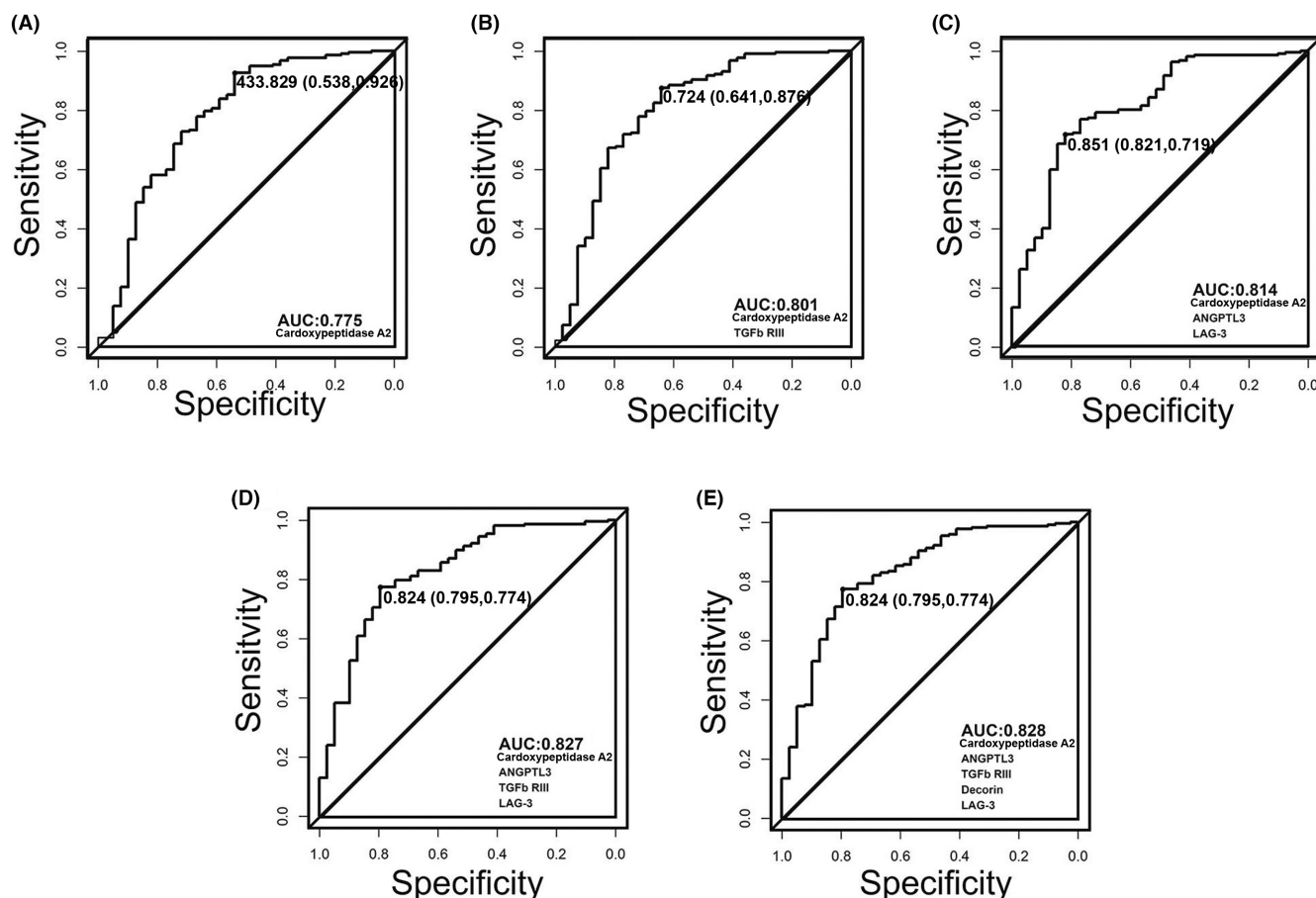


FIGURE 8 A–E, Carboxypeptidase A2; TGFβ RIII; LAG-3; Decorin; and ANGPTL3 were found to be the most effective in discriminating GC from HC.

Interestingly, we also screened two proteins, B2M and biG-H3, which showed differences only in HCC (Figure S2C). They may be serum biomarkers for liver cancer diagnosis. We also found that LH was expressed differently in EGC and HCC compared with HC, and GH was expressed differently in EGC and CCR (Figure S2A,B). There is a possibility that these proteins will be used in the diagnosis of GC in the future.

In the present study, we identified several serum proteins previously reported to be related to GC in our GC patient sample. This result provides independent support for the effectiveness of our screening method.

There are, however, some limitations to this study. The antibody chip is based on the principle of antigen–antibody specific binding. Therefore, using this array, only proteins for which corresponding antibodies are available can be detected. Furthermore, the 640-protein array we used only included secretory proteins or secretory protein receptors, which were limited and pre-selected, potentially excluding some promising biomarker candidates from the study onset.

In this study, we utilized an independent case cohort and different types of control groups, such as other stomach diseases and tumors, to verify the reliability of the biomarkers. However, the number of serum samples from patients with other stomach

diseases and tumors was small, which may have influenced the detection results. In addition, the samples were collected only from three hospitals along the coast of China, which could be influenced by ethnicity and region. Furthermore, a prospective study is valuable for further evaluation of the diagnostic/screening value of protein combinations. The significant increase or decrease in the levels of some serum proteins after surgery can clarify the role of these proteins in the progression of GC. Finally, in the regression model, we did not choose the five protein aggregation models because we believed that the combination of two proteins considered both diagnostic accuracy and the economic burden on clinical patients.

Despite the aforementioned limitations, the results of this study may contribute to a breakthrough in serum biomarkers of GC and provide new ideas for the diagnosis of the disease.

AUTHOR CONTRIBUTIONS

Guarantors of integrity of the entire study: Yongdong Yi, Rubin Nan, Jianhua Lu and Weijian Sun. Study concepts and design: Weijian Sun, Hao Wu and Xian Shen. Literature research: Yongdong Yi, Danna Liang, Hongbo Zhang. Clinical studies: Yongdong Yi, Danna Liang, Hongbo Zhang, Shengsheng Zhao, Bozhen Chen, Jiangnan Chen, Zhiqiang Zheng, Tao You, Tanzhou Chen, Xiaodong

Chen, Wenqian Wang, Yiming Chen, Shuai Liu, Yinpeng Huang, Yaojun Yu, Mingdong Lu, Pihong Li, He Huang, and Gongting Zhou. Experimental studies/data analysis: Yongdong Yi, Limiao Lin, Rubin Nan, Xuhui Wang Hongzheng Li, Jianhua Lu. Statistical analysis: Yongdong Yi, Rubin Nan, Xuhui Wang and Jianhua Lu. Manuscript preparation: Yongdong Yi, Rubin Nan, Xuhui Wang and Jianhua Lu. Manuscript editing: Yongdong Yi, Rubin Nan, Xuhui Wang and Jianhua Lu.

ACKNOWLEDGMENTS

None.

FUNDING INFORMATION

This work was supported by the National Natural Science Foundation of China (Grant Number 81972261), the Major Research Projects of Wenzhou Science and Technology Plan Project (Grant Numbers ZY2020007, Y20210189, ZY2021003 and 2020Y0593), the Key Project Cultivation Project of Wenzhou Medical University (Grant Number KYYW202006), the Key Laboratory of Minimally Invasive Techniques & Rapid Rehabilitation of Digestive System Tumor of Zhejiang Province (Grant Number 21SZDSYS04), Lin He's New Medicine and Clinical Translation Academician Workstation Research Fund (Grant Numbers 18331202, and 19331104), the Natural Science Foundation of Zhejiang Province (Grant Number LY18H030009), and the Project of Medical Technology and Education of Zhejiang Province of China (Grant Number 202045547).

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest.

DISCLOSURE

All authors have read this manuscript, approved the content and agreed to its publication.

ETHICS STATEMENT

Approval of the research protocol by an Institutional Reviewer Board: Yes.

Informed Consent: Informed consent was obtained from subject(s) and/or guardian(s).

Registry and the Registration No. of the study/trial: Blood samples and stomach tissues were acquired from the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, the First Affiliated Hospital of Wenzhou Medical University (2021-K-42-02), and Taizhou Hospital of Zhejiang Province (K20220103).

Animal Studies: The experiments concerning animals were approved by the Institutional Animal Care and Use Committee of Zhejiang Laboratory Animal Center (ZJCLA-IACUC-20050024).

ORCID

Yongdong Yi  <https://orcid.org/0009-0004-1431-4809>

Rubin Nan  <https://orcid.org/0000-0003-2375-6946>

Wenqian Wang  <https://orcid.org/0000-0003-2925-4306>

Xian Shen  <https://orcid.org/0000-0001-7974-830X>

Weijian Sun  <https://orcid.org/0000-0001-5789-1287>

REFERENCES

- Kim Y, Byeon SJ, Hur J, et al. High delta-like ligand 4 expression correlates with a poor clinical outcome in gastric cancer. *J Cancer*. 2019;10(14):3172-3178. doi:10.7150/jca.30257
- Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet*. 2020;396(10251):635-648. doi:10.1016/S0140-6736(20)31288-5
- Pasechnikov V, Chukov S, Fedorov E, Kikuste I, Leja M. Gastric cancer: prevention, screening and early diagnosis. *World J Gastroenterol*. 2014;20(38):13842-13862. doi:10.3748/wjg.v20.i38.13842
- Hamashima C. Benefits and harms of endoscopic screening for gastric cancer. *World J Gastroenterol*. 2016;22(28):6385-6392. doi:10.3748/wjg.v22.i28.6385
- Dong Z, Zhang X, Chen X, Zhang J. Significance of serological gastric biopsy in different gastric mucosal lesions: an observational study. *Clin Lab*. 2019;65(12):2141-2148. doi:10.7754/Clin.Lab.2019.190433
- Matsuoka T, Yashiro M. Biomarkers of gastric cancer: current topics and future perspective. *World J Gastroenterol*. 2018;24(26):2818-2832. doi:10.3748/wjg.v24.i26.2818
- Shimada H, Noie T, Ohashi M, Oba K, Takahashi Y. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the task force of the Japanese gastric cancer association. *Gastric Cancer*. 2014;17(1):26-33. doi:10.1007/s10120-013-0259-5
- Vanarsa K, Soomro S, Zhang T, et al. Quantitative planar array screen of 1000 proteins uncovers novel urinary protein biomarkers of lupus nephritis. *Ann Rheum Dis*. 2020;79(10):1349-1361. doi:10.1136/annrheumdis-2019-216312
- Lin Y, Huang R, Chen L, et al. Identification of interleukin-8 as estrogen receptor-regulated factor involved in breast cancer invasion and angiogenesis by protein arrays. *Int J Cancer*. 2004;109(4):507-515.
- Sanchez-Carbayo M, Socci ND, Lozano JJ, Haab BB, Cordon-Cardo C. Profiling bladder cancer using targeted antibody arrays. *Am J Pathol*. 2006;168(1):93-103.
- Goksuluk DKS, Zararsiz G, Karaagaoglu AE. easyROC: An interactive web-tool for ROC curve analysis using R language environment. *R J*. 2016;8(2):213. doi:10.32614/RJ-2016-042
- Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66(2):115-132. doi:10.3322/caac.21338
- Wei W, Zeng H, Zheng R, et al. Cancer registration in China and its role in cancer prevention and control. *Lancet Oncol*. 2020;21(7):e342-e349. doi:10.1016/S1470-2045(20)30073-5
- Chen S, Li T, Zhao Q, Xiao B, Guo J. Using circular RNA hsa_circ_0000190 as a new biomarker in the diagnosis of gastric cancer. *Clin Chim Acta*. 2017;466:167-171. doi:10.1016/j.cca.2017.01.025
- Wang N, Wang L, Yang Y, Gong L, Xiao B, Liu X. A serum exosomal microRNA panel as a potential biomarker test for gastric cancer. *Biochem Biophys Res Commun*. 2017;493(3):1322-1328. doi:10.1016/j.bbrc.2017.10.003
- Lee K-F, Tsai M-M, Tsai C-Y, et al. DEK is a potential biomarker associated with malignant phenotype in gastric cancer tissues and plasma. *Int J Mol Sci*. 2019;20(22):5689. doi:10.3390/ijms20225689
- Li N, Wang Y, Forbes K, et al. Metalloproteases regulate T-cell proliferation and effector function via LAG-3. *EMBO J*. 2007;26(2):494-504.
- Li N, Workman CJ, Martin SM, Vignali DAA. Biochemical analysis of the regulatory T cell protein lymphocyte activation gene-3 (LAG-3; CD223). *J Immunol (Baltimore, Md: 1950)*. 2004;173(11):6806-6812.
- Zeng H, Zhou Q, Wang Z, et al. Stromal LAG-3+ cells infiltration defines poor prognosis subtype muscle-invasive bladder cancer with

- immuno-evasive contexture. *Journal for ImmunoTherapy of Cancer*. 2020;8(1):e000651. doi:10.1136/jitc-2020-000651
20. Elderbroom JL, Huang JJ, Gatzka CE, et al. Ectodomain shedding of T β RIII is required for T β RIII-mediated suppression of TGF- β signaling and breast cancer migration and invasion. *Mol Biol Cell*. 2014;25(16):2320-2332. doi:10.1091/mbc.E13-09-0524
21. Gatzka CE, Holtzhausen A, Kirkbride KC, et al. Type III TGF- β receptor enhances colon cancer cell migration and anchorage-independent growth. *Neoplasia*. 2011;13(8):758-770.
22. Burghardt I, Schroeder JJ, Weiss T, Gramatzki D, Weller M. A tumor-promoting role for soluble T β RIII in glioblastoma. *Mol Cell Biochem*. 2021;476(8):2963-2973. doi:10.1007/s11010-021-04128-y
23. Song Y, Wang Q, Wang D, et al. Label-free quantitative proteomics unravels carboxypeptidases as the novel biomarker in pancreatic ductal adenocarcinoma. *Transl Oncol*. 2018;11(3):691-699. doi:10.1016/j.tranon.2018.03.005
24. Baghy K, Reszegi A, Tátrai P, Kovalszky I. Decorin in the tumor microenvironment. *Adv Exp Med Biol*. 2020;1272:17-38. doi:10.1007/978-3-030-48457-6_2
25. Basak D, Jamal Z, Ghosh A, et al. Reciprocal interplay between asporin and decorin: implications in gastric cancer prognosis. *PLoS One*. 2021;16(8):e0255915. doi:10.1371/journal.pone.0255915
26. Koyama T, Ogawara K, Kasamatsu A, et al. ANGPTL3 is a novel biomarker as it activates ERK/MAPK pathway in oral cancer. *Cancer Med*. 2015;4(5):759-769. doi:10.1002/cam4.418
27. El-Shal AS, Zidan HE, Rashad NM, Wadea FM. Angiopoietin-like protein 3 and 4 expression 4 and their serum levels in hepatocellular carcinoma. *Cytokine*. 2017;96:75-86. doi:10.1016/j.cyto.2017.03.006
28. Tang C, Chen E, Peng K, et al. Mining the role of angiopoietin-like protein family in gastric cancer and seeking potential therapeutic targets by integrative bioinformatics analysis. *Cancer Med*. 2020;9(13):4850-4863. doi:10.1002/cam4.3100

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Yi Y, Nan R, Lu J, et al. Screening of novel serum biomarkers for gastric cancer in coastal populations using a protein microarray. *Cancer Sci*. 2023;114:3396-3410. doi:10.1111/cas.15876