

ARTICLE



Molecular Diagnostics

Artificial intelligence-driven microRNA signature for early detection of gastric cancer: discovery and clinical functional exploration

Jiachun Lu^{1,2,8}, Yuqi Chen^{3,8}, Xin Liu^{1,2}, Jiayu Wang^{1,2}, Yuxin He^{1,2}, Tongguo Shi²✉, Weichang Chen^{1,2,4}✉ and Wenying Yan^{1,5,6,7}✉

© The Author(s), under exclusive licence to Springer Nature Limited 2025

BACKGROUND: Gastric cancer (GC) is a leading cause of cancer-related deaths worldwide, with late-stage diagnoses frequently leading to poor outcomes. This underscores the need for effective early-stage gastric cancer (ESGC) diagnostics.

METHODS: We introduce ESGCmiRD, an innovative artificial intelligence-driven strategy that identifies a miRNA signature for ESGC detection by integrating robust expression patterns, ESGC relevance, and regulatory capabilities of microRNA (miRNA) based on multiple networks. Expression and biological roles of miRNAs in GC were validated and explored via bioinformatics analysis and in vitro studies. miRNA-target interaction was confirmed by dual-luciferase reporter assay. Molecular docking predicted miRNA-drug binding affinities, assessing the miRNA signature's therapeutic potential.

RESULTS: ESGCmiRD identified a blood miRNA signature (miR-320b, miR-222-3p, miR-181a-5p, miR-103a-3p, miR-107) for ESGC detection, demonstrated high diagnostic accuracy with AUC values of 0.986, 0.977, 0.815, and 0.811 in the test and three validation sets (GSE211692, TCGA-STAD, and our cohort), respectively. The five miRNAs were overexpressed in ESGC plasma and directly target PTEN, promoting GC cell proliferation, migration, and invasion. Molecular docking suggested Paclitaxel had the strongest potential interaction with these miRNAs.

CONCLUSION: This method identifies a robust miRNA signature for ESGC detection and sheds light on gastric carcinogenesis mechanisms, opening doors for potential therapeutic strategies.

British Journal of Cancer (2025) 132:957–972; <https://doi.org/10.1038/s41416-025-02984-9>

INTRODUCTION

Gastric cancer (GC) is a globally prevalent malignant tumour and is the fourth leading cause of cancer-related mortality [1]. Typically, GC is diagnosed at an advanced stage, where the prognosis for advanced gastric cancer (AGC) is poor, with a survival rate of only 20% [2]. However, early therapeutic intervention can significantly improve survival rates, potentially increasing them to 70% [3, 4]. Endoscopy is currently the gold standard for diagnosing GC due to its precision [5]. Despite its reliability, endoscopy has limitations, including its high cost, invasiveness, and associated risks [6]. In the clinical diagnosis of GC, biomarkers such as carcinoembryonic antigen (CEA), carbohydrate antigens (CA) 19-9, CA 125, and pepsinogen are utilised. However, these markers have limited specificity and sensitivity, impeding their effectiveness in the early detection of GC [7]. Therefore, the development of new diagnostic models for early gastric cancer holds significant clinical

implications for improving patient outcomes, reducing medical costs, and enhancing patients' quality of life. They will facilitate early detection and treatment, thereby potentially significantly increasing the cure rate and reducing reliance on invasive diagnostic methods.

Liquid biopsy has emerged as a transformative approach in oncology, enabling non-invasive, real-time monitoring of tumour biomarkers, such as circulating tumour cells, cell-free DNA, circulating tumour DNA, non-coding RNAs (ncRNAs), and exosomes—from bodily fluids like blood, urine, and cerebrospinal fluid [8]. In contrast to conventional tissue biopsy, liquid biopsy offers a less invasive and more practical means of tracking cancer progression, treatment efficacy, and drug resistance [9–11]. With the swift advancement of high-throughput sequencing and AI-driven computational techniques, its clinical potential for early diagnosis, prognostic

¹Department of Gastroenterology, The First Affiliated Hospital of Soochow University, Suzhou, China. ²Jiangsu Institute of Clinical Immunology, The First Affiliated Hospital of Soochow University, Suzhou, China. ³Department of Gastroenterology, The Fourth Affiliated Hospital of Soochow University, Suzhou, China. ⁴Jiangsu Key Laboratory of Clinical Immunology, Soochow University, Suzhou, China. ⁵School of Basic Medical Sciences, Suzhou Medical College of Soochow University, Suzhou, China. ⁶Suzhou Key Lab of Multimodal Data Fusion and Intelligent Healthcare, Suzhou, China. ⁷Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Suzhou, China. ⁸These authors contributed equally: Jiachun Lu, Yuqi Chen. ✉email: shitg@suda.edu.cn; weichangchen@126.com; wyyan@suda.edu.cn

Received: 24 September 2024 Revised: 1 March 2025 Accepted: 12 March 2025

Published online: 15 April 2025

evaluation, and identification of chemotherapy resistance in gastric cancer patients is immense [7, 12].

MicroRNAs (miRNAs), a class of non-coding RNAs, have emerged as promising diagnostic, prognostic, and predictive markers owing to their stability, tissue specificity, and abnormal expression patterns in diseased states. For instance, miR-21, a star miRNA, has been reported to play important roles in the diagnosis and prognosis of GC [13]. A six-exosome-RNA panel inclusive of miRNAs can predict the therapeutic efficacy of neoadjuvant chemotherapy in patients with GC [14]. The miR-200/183 family has also shown potential in predicting GC progression and prognosis [15]. In a study involving subjects from Singapore and Korea, So JBY et al. developed a 12-miRNA biomarker assay for the detection of GC [16]. Abe S. et al. reported that a combination of four serum miRNAs, namely miR-4257, miR-6785-5p, miR-187-5p, and miR-5739, could be a useful diagnostic biomarker for detecting early GC in a Japanese population [17]. Overall, miRNAs have emerged as remarkable biomarkers in the field of liquid biopsy and provide a fresh viewpoint for non-invasive disease monitoring, particularly in the detection of GC.

In this study, as depicted in Fig. 1, we proposed an AI-based strategy called ESGCmiRD for detecting early-stage gastric cancer (ESGC) by integrating robust expression patterns, ESGC relevance, and regulatory capabilities of miRNA based on topological and biological traits across multiple network types. A five-miRNA signature was developed based on ESGCmiRD and subsequently verified through multiple public databases and serum cohorts. Additionally, we delved into the biological functions and potential applications of the molecular markers implicated in gastric carcinogenesis. This holistic methodology offers a potent instrument for early GC diagnosis, thereby enhancing our comprehension of the disease's underlying mechanisms.

METHODS AND MATERIALS

MiRNA expression datasets of ESGC

To discover and validate miRNA biomarkers for ESGC, miRNA expression profiles from 10,071 patient samples were retrieved from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. In our study, the early stages of gastric cancer (GC) were defined as patients classified as Stage I or Stage II according to the TNM staging system. The datasets retrieved included two blood datasets from GEO (GSE164174 and GSE211692) and two tissue datasets (GSE158315 from GEO and TCGA-STAD). The detailed information of these datasets, including the number of patients, sample types, miRNA profiling methods, and other relevant characteristics, are listed in Table S1.

Identification of robust differential expression miRNAs in ESGC

To identify miRNAs that exhibit robust and stable differential expression in ESGC, we initially employed the R package limma [18] to sift through the aforementioned four datasets, searching for miRNAs displaying significant expression disparities. We set a cut-off criterion, requiring a fold change exceeding 1.5 and a *p*-value less than 0.05. Furthermore, to enhance the robustness of the results, we integrated gender and age as covariates in the differential analysis. The samples were stratified into two groups according to a threshold of 65 years of age. Following this initial screening, to integrate insights derived from these diverse datasets, we implemented the robust rank aggregation (RRA) algorithm [19]. This method was applied to the differentially up-regulated miRNAs identified from both blood and tissue datasets, ultimately yielding robust differentially expressed miRNAs specific to ESGC samples.

Weighted miRNA co-expression network construction and analysis

We employed weighted gene co-expression network analysis (WGCNA) methodologies [20] to construct the miRNA co-expression network and pinpoint modules strongly linked to ESGC. The parameters guiding network construction were as follows: a soft threshold of 9, a minimum module size of 30, and a merge cut height of 0.25 (Fig. S1). Modules with a

p-value below 0.05 and a correlation coefficient exceeding 0.6 were designated as critical modules in ESGC. In the final step, we computed the gene significance and module membership for all miRNAs within these key modules. MiRNAs were then filtered using thresholds of $MM > 0.6$ (*p*-value < 0.05) and $GS > 0.45$ (*p*-value < 0.05), aiming to isolate those potentially associated with ESGC.

Evaluation of the miRNA regulatory abilities based on the miRNA regulatory network

Initially, we updated the human miRNA regulatory network [21, 22] by integrating data from PITA [23] and miRTarbase [24]. Following this, we computed our previously established *RNs* scores for the miRNAs within the network, aiming to assess their regulatory potencies in cancer status [25]. Given that the *RNs* scores encompass both expression data and topological features of the miRNA regulatory network, we incorporated miRNA and mRNA expression data from TCGA-STAD. Additionally, we calculated topological network parameters such as degree, closeness, and eigenvector centrality for each miRNA using the R package igraph [26].

Functional enrichment analysis

To gain deeper insights into the potential functions of miRNA-target genes, we conducted functional enrichment analyses based on Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) utilising the R package “clusterProfiler” [27]. Furthermore, we considered GO terms and KEGG pathways significantly enriched if their Benjamin-Hochberg (BH) adjusted *p*-value and *q*-value were both less than 0.05.

Diagnostic model construction based on AI methods

Key miRNAs were utilised to construct a diagnostic model specifically aimed at detecting ESGC among non-GC samples. To build these models, we employed six different AI algorithms: decision tree, k-nearest neighbours (KNN), naive Bayes, random forest, support vector machines (SVM), and extreme gradient boosting (XGBoost). The GSE164174 dataset was randomly partitioned into a 70% training set and a 30% testing set. The performance of the models was rigorously evaluated using the area under the curve (AUC) metric, sensitivity, specificity, and confusion matrix.

Plasma sample collection

Plasma samples were gathered from the First Affiliated Hospital of Soochow University (Suzhou, China). The study protocol received approval from the Institutional Review Board of the First Affiliated Hospital of Soochow University (Approval No. 2021068). Prior written consent was secured from both healthy volunteers and patients. The criteria for inclusion of the GC patients were (1) age ≥ 18 years, no gender restriction, and histopathologically confirmed diagnosis of gastric cancer; (2) no surgical treatment, neoadjuvant therapy such as radiotherapy and chemotherapy, and translational therapy prior to enrolment; (3) pathological staging of stage I-IV; (4) the patients underwent blood sample collection; and (5) complete clinical data. For the GC group, TNM staging followed the pathological staging standards set by the American Joint Committee on Cancer (8th edition). On the morning prior to surgery, whole blood samples (10–20 mL) were drawn from subjects using EDTA-coated tubes, with detailed information labelled on each tube to prevent any mix-ups. Plasma was separated by centrifugation (2500 rpm, 5 min) at room temperature on the day following collection. Individual plasma samples were uniquely numbered and preserved at -80°C . A total of 183 patients' plasma samples were included in this study, and Table S2 summarises the clinical characteristics of all participants.

RNA extraction and quantitative real-time PCR (qRT-PCR)

The extraction and purification of miRNA from plasma samples were carried out using the MolPure[®] Serum/Plasma miRNA Kit (#19332ES50, Yeasen, Shanghai, China). Following this, 7 μL of miRNA from each sample was reverse transcribed into cDNA with the help of the miDETECT A Track miRNA qPCR Kit (#10712, RiboBio, Guangzhou, China). Specific primers for the candidate miRNAs were synthesised by RiboBio for qPCR analysis. Additionally, an equal volume of cel-miR-39-3p was added as an external control before the miRNA extraction step (RiboBio). Cell and tissue RNA extraction was performed by Super FastPure Cell RNA Isolation Kit (#RC102-01, Vazyme, Nanjing, China), 1 μg of RNA from each sample was reverse transcribed into cDNA with the help of the HiScript IV 1st Strand cDNA Synthesis Kit (#R412-01, Vazyme). The relative gene expression levels were determined using the $2^{-\Delta\Delta\text{CT}}$ method [28] and adjusted by log10.

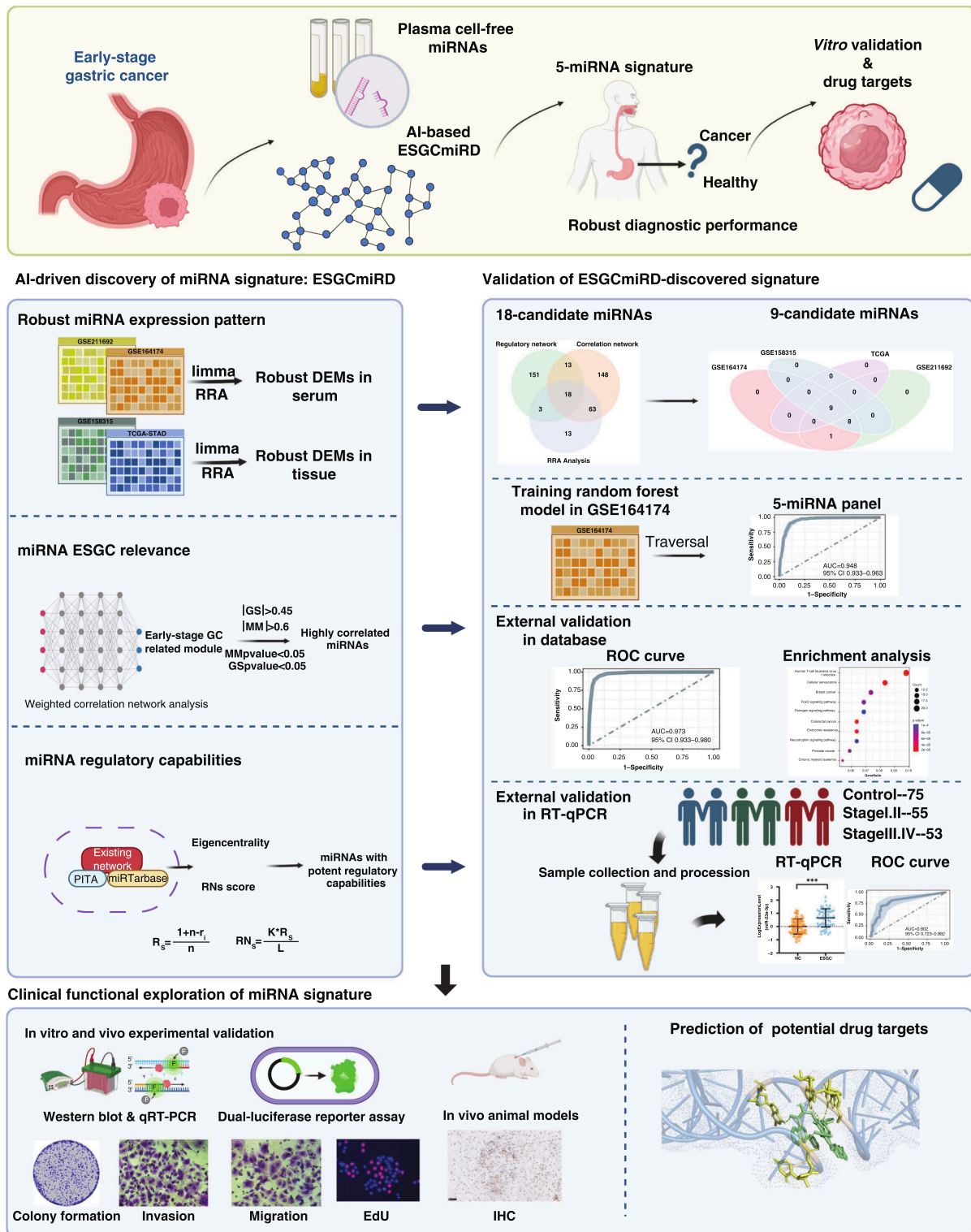


Fig. 1 Overall study design flowchart. ESGCmiRD, an AI-based method integrating multi-scale network and differential expression analysis, identified a five-miRNA diagnostic panel for gastric cancer. This panel was validated in external datasets and a plasma cohort. It promotes cancer progression via the PTEN/AKT axis, providing insights for diagnosis and therapy. RRA Robust Rank Aggregation, MM Module membership, GS Gene significance.

Cell culture and transfection

AGS and MKN-28 cell lines were procured from the American Type Culture Collection (Manassas, VA, USA). The cells were maintained in RPMI-1640 medium (Biological Industries, Beit Haemek, Israel) fortified with 10% fetal bovine serum (Biological Industries) and 1% penicillin-streptomycin-

amphotericin B (NCM Biotech, Suzhou, China). The cultures were kept under conditions of 5% CO₂ and a temperature of 37 °C.

PTEN siRNAs (designated as si-PTEN-1, si-PTEN-2, and si-PTEN-3), a negative control siRNA, along with mimics and inhibitors for hsa-miR-103a-3p, hsa-miR-107, hsa-miR-181a-5p, has-miR-320b, and has-miR-222-3p, as

well as a control mimic/inhibitor (Genepharma, Shanghai, China), were introduced into GC cells using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), following the manufacturer's guidelines. The sequences of the mimics and inhibitors are detailed in the Supplementary file: Table S3.

Transwell migration and invasion assay

Transwell plates (8.0 µm pore size and a PET membrane, Falcon, USA) were employed to conduct transwell migration and invasion assays. In summary, 400 µL of RPMI-1640 medium enriched with 20% fetal bovine serum was introduced to the lower chamber. Following an incubation period of 24–48 h at 37 °C, the membranes were immobilised using 4% paraformaldehyde for 15 min at ambient temperature and subsequently stained with crystal violet (Beyotime, Shanghai, China) for a duration of 30 min. For the transwell invasion studies, the membranes in the upper chamber underwent coating with Matrigel (Corning, Corning, NY, USA) at a 1:30 dilution. Ultimately, the stained cells were examined in five randomly selected fields on each filter membrane through the use of a light microscope (Olympus, Tokyo, Japan).

Colony formation assay

GC cells were seeded into 12-well plates at a density of 1000 cells per well and incubated at 37 °C for a duration of 14 days. Subsequently, the cells were fixed using 4% paraformaldehyde for 15 min at room temperature. The cell colonies were then stained with crystal violet (Beyotime). Upon completion of staining, the plates were inverted and photographed on a luminous white background.

EdU incorporation assay

Cells were plated in 24-well plates and incubated overnight. Following this, the cells were exposed to a medium containing EdU working solution (Beyotime) and incubated at 37 °C for 2 h. After fixation with a 4% paraformaldehyde solution (Beyotime), the cells underwent osmotic treatment with a 0.3% Triton X-PBS solution. They were then stained with click additive solution, and the nuclei were counterstained with Hoechst dye (Beyotime). Finally, positive cells were enumerated using a fluorescence detector.

Protein extraction and western blot analysis

SDS lysis buffer (Beyotime) supplemented with protease inhibitors (Beyotime) was utilised to lyse the collected cells. The protein concentration was determined using a BCA protein assay kit (Beyotime). Proteins were separated by 10–12% SDS-PAGE and then transferred to PVDF membranes. Following blocking with 5% skimmed milk for 1 h at room temperature, the membranes were incubated with primary antibodies overnight at 4 °C. The membranes were then exposed to the secondary antibody for 1 h at room temperature and washed three times with TBST. Subsequently, the membranes were visualised with ECL reagent (Vazyme, Nanjing, China) using ChemiDoc™ MP imaging equipment (Bio-Rad). The antibodies employed for western blot analysis in this study were as follows: mouse anti-human PTEN (Proteintech, Wuhan, China), rabbit anti-human p-AKT (Cell Signaling Technology, Danvers, USA), rabbit anti-human AKT (Cell Signaling Technology) and mouse anti-human GAPDH (Proteintech, Wuhan, China).

Dual-luciferase reporter assay

The PmirGLO vector was constructed by cloning the PTEN 3'-UTR, which contains miRNA recognition sequences, into the PmirGLO dual-luciferase miRNA targeting expression plasmid. Additionally, a mutant plasmid was created by modifying the seed region of the miRNA binding site. Following co-transfection of PTEN 3'-UTR-WT-PmirGLO or 3'-UTR-MT-PmirGLO with miRNA mimics for 48 h, luciferase activity was assayed using the Dual-Luciferase Reporter Assay System (Yeasen, Shanghai, China) and a Biotek luminometer. The relative luciferase activity was determined by normalising the firefly luciferase activity to the renilla luciferase activity.

In vivo mouse experiments

All animal experiments were performed in accordance with the institutional guidelines of the Soochow Animal Care and Use Committee (No.202412A0046). Six-week-old female NSG mice were obtained from the Shanghai Laboratory Animal Centre (Shanghai, China). Mice were randomly divided into two groups: the agomir-NC group and the agomir group ($n = 3$ in each group). Each mouse was injected subcutaneously with 8×10^6 MKN28 cells in 100 µL PBS. Once the subcutaneous tumour grew to

4 × 4 mm, mice in the agomir group were injected with hsa-miR-103a-3p, hsa-miR-107, hsa-miR-181a-5p, hsa-miR-320b, hsa-miR-222-3p of agomir at a dose of 1×5 nmol/pupil (Genepharma, Shanghai, China) and an injection volume of 50 µL per mouse, administered every three days for two consecutive weeks. The control group was injected with the same amount of PBS. The volume of transplanted tumours was measured every three days with a caliper. The formula used to calculate the tumour volume was: $\text{volume (mm}^3\text{)} = 0.5 \times L(\text{mm}) \times S^2(\text{mm}^2)$, where S and L are the minimum and maximum diameters of the tumour, respectively. On day 14 of the experiment, the mice were euthanised and the tumour tissues were excised and weighed.

Immunohistochemistry(IHC) Staining

Subcutaneously transplanted tumour tissue from a mouse model was used for IHC. An embedding machine was used to melt paraffin for dipping fixed and dehydrated tissues. After wax immersion and embedding, tissue sections were prepared using a microtome. The sections were then deparaffinised, rehydrated, and heated to 96 °C in a pan containing a repair solution. After cooling, the sections were inactivated with endogenous enzymes and sealed with 10% goat serum sealant. The mouse anti-human PTEN (Proteintech) was then added and incubated overnight at 4 °C, followed by the addition of a secondary antibody. All sections were stained with 3,3'-diaminobenzidine (DAB) solution and counterstained with haematoxylin. After dehydration and sealing with rubber, the sections were examined under the microscope and photographed.

Molecular docking

To investigate the interactions between miRNAs and drugs, molecular docking is carried in two steps as follows. Firstly, drugs for gastric cancer and their structures were searched from DrugBank database [29]. miRNA sequence data were obtained from miRBase [30], and their structures were predicted via Alphafold3 [31]. Secondly, molecular docking simulations were performed using AutoDock Vina [32]. The docking results were visualised with PyMOL to facilitate analysis and interpretation.

Statistical analysis

The statistical analysis for this study was performed using R software (v 4.2.1) and GraphPad Prism 10. Data visualisation was achieved with the help of the R package "ggplot2" and the Gephi-0.10.1 software. The data are expressed as mean ± standard deviation (SD). To compare differences between groups, we employed the Student's *t*-test or one-way ANOVA. Statistical significance was established at p -value < 0.05.

RESULTS

Robust differentially up-regulated miRNAs in ESGC

To identify miRNA signatures specific to ESGC, we conducted a comprehensive differential expression analysis across multiple datasets to pinpoint robustly up-regulated miRNAs in ESGC. As shown in Fig. S2A, S2B, in two blood datasets, GSE164174 and GSE211692, we detected 501 and 669 differentially up-regulated miRNAs, respectively. Similarly, for the other two tissue datasets, GSE158315 and TCGA-STAD, we found 88 and 137 differentially up-regulated miRNAs, respectively (Fig. S2C–S2D). To further validate the reliability of the identified differentially expressed miRNAs, we conducted covariance analyses by incorporating age and sex as covariates. Following these analyses, the list of differentially expressed genes obtained was found to highly overlap with that from the initial analysis (Fig. S2E). Consequently, the initial results were retained for subsequent analyses. Through intersecting the data from the blood datasets, we identified 471 miRNAs, while from the tissue datasets, we obtained 18 miRNAs. We further refined this selection using RRA analysis, resulting in 93 and 4 robust up-regulated miRNAs from the blood and tissue datasets, respectively (Fig. 2a, b).

Weighted miRNA co-expression network analysis reveals highly ESGC-relevant miRNAs

Using the WGCNA, we constructed a miRNA co-expression network and conducted functional module analysis. Figure 2c

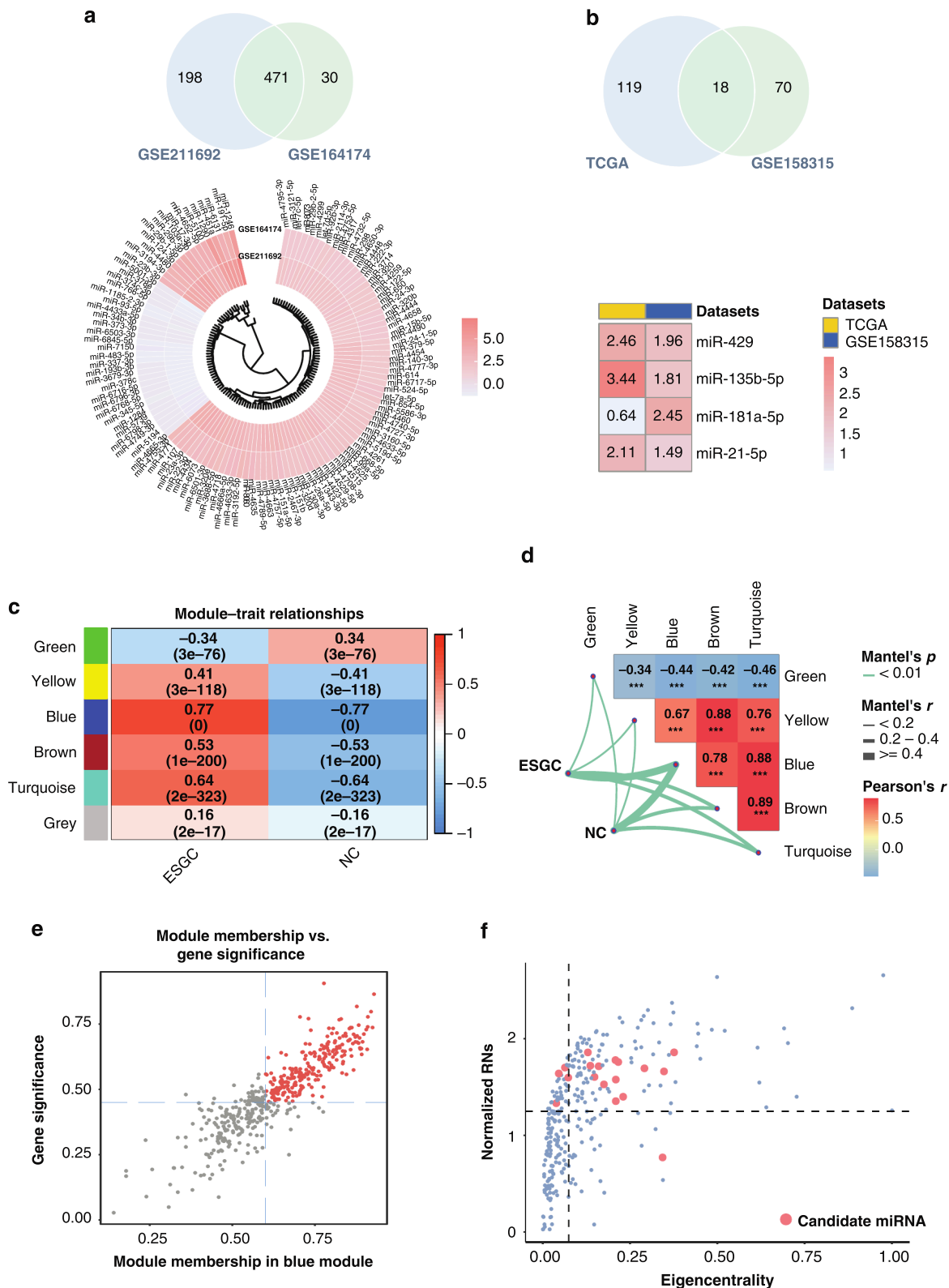


Fig. 2 Identification of miRNA signature candidates for ESGC detection by ESGCmiRD. **a** Venn diagram of significantly up-regulated miRNAs in the serum dataset and heatmap of their RRA scores. **b** Venn diagram of significantly up-regulated miRNAs in the tissue dataset and heatmap of their RRA scores. **c** Heatmap of Pearson correlation between modules and ESGC. Values represent correlation coefficients between modules and disease states. **d** Mantel test between module and ESGC and heatmap of correlation among modules. The line width represents the correlation of the mantel test, and the green line represents the significance of the mantel test (p -value < 0.01). The values in each cell represent the correlation coefficients between the modules. **e** Scatterplot of MM and GS from the blue module with screening thresholds of $MM > 0.6$, $GS > 0.45$. **f** Scatterplot of normalised RNs and eigenvector centrality of miRNAs in GC-specific miRNA regulatory network.

illustrates the identification of six distinct modules within this network. We then examined the correlation between these modules and disease status (ESGC or normal control). Among the six modules, the blue module exhibited the strongest positive correlation with ESGC ($COR = 0.77$, p -value < 0.001) and, conversely, the strongest negative correlation with normal samples ($COR = -0.77$, p -value < 0.001). Furthermore, our analysis revealed that all modules positively correlated with ESGC and also demonstrated a positive correlation with each other (Fig. 2d). Based on these findings, we selected the blue module for further screening. Ultimately, we pinpointed 242 miRNAs that constitute the core of the blue module (with an $MM > 0.6$ and a p -value < 0.05) and that demonstrated significant correlations with ESGC itself ($GS > 0.45$, p -value < 0.05), thereby identifying them as highly relevant miRNAs to ESGC, as illustrated in Fig. 2e.

Identification of miRNAs with potent regulatory capacity in ESGC via AI-based network analysis

Initially, we refined our pre-existing human miRNA regulatory network, significantly expanding its scope to encompass 820 miRNAs and 12,735 mRNAs. Next, we integrated TCGA-STAD miRNA and mRNA expression profiles into this network, thereby constructing a GC-specific miRNA regulatory network, which includes 303 miRNAs and 11,842 mRNAs. Subsequently, we calculated key network metrics such as degree, closeness, betweenness, and eigenvector centrality for the miRNAs within this specialised network (see Table S4 for details). Furthermore, we applied our previously developed AI-driven network scoring system, termed *RNs*, to evaluate each miRNA within the network. Through this comprehensive analysis, we identified a total of 185 miRNA candidates whose *RNs* and eigenvector centrality values surpassed the median thresholds for these parameters (Fig. 2f). These miRNAs are considered to possess significant regulatory capabilities in ESGC.

miRNAs biomarker acquisition and expression validation by qRT-PCR

By considering the robust up-regulation, ESGC relevance, and regulatory capabilities of miRNAs, we screened out 18 miRNA biomarkers (Fig. 3a). To ensure the up-regulation of these miRNA biomarkers in ESGC, we further filtered the 18 miRNAs to nine miRNAs whose expression was increased in all four aforementioned datasets (Fig. 3b). Then to validate the up-regulation and non-invasive potential of these miRNAs, we determined the expression levels of the nine miRNA biomarkers in plasma samples from a newly collected cohort of 55 ESGC patients and 75 healthy controls. The expression levels of nine miRNAs (miR-107, miR-103a-3p, miR-222-3p, miR-320b, miR-181a-5p, miR-191-5p, miR-23a-3p, miR-92b-3p, miR-24-3p) were significantly higher in plasma samples of ESGC patients compared to healthy controls (Fig. 3c), corroborating our previous observations.

Construction and validation of ESGC miRNA-based diagnostic model

Utilising the GSE164174 microarray expression dataset, we constructed AI-based models based on the expression of the above nine miRNAs through six distinct methods as described in the 'Methods' section. As shown in Fig. S3, the random forest model emerged as the most effective, achieving an AUC of 0.944 (95% CI: 0.930–0.959; sensitivity: 0.930; specificity: 0.834) in the test set when compared to the other models. The sensitivity and specificity of all the six models in training set and test set were enumerated in Table S5. It is evident that the random forest model possesses relatively high specificity and sensitivity values. Based on these comprehensive comparisons and considering the characteristics of our research problem, the random forest model emerges as the most rational and effective choice. Furthermore, the importance ranking of miRNAs was also calculated and

presented in Fig. S3C. To verify the model's applicability across different datasets, we used GSE211692, the TCGA database, and our proprietary serum cohort as external validation sets. The model's performance remained consistently high, with an AUC of 0.989 (95% CI: 0.987–0.991; sensitivity: 0.826; specificity: 0.984) in GSE211692, 0.743 (95% CI: 0.667–0.819; sensitivity: 0.586; specificity: 0.756) in TCGA, and 0.813 (95% CI: 0.739–0.888; sensitivity: 0.400; specificity: 0.920) in our serum cohort. The ROC curves are shown in Fig. S4.

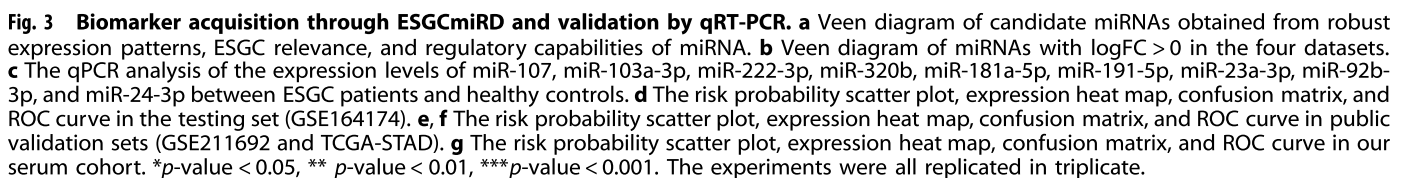
Aiming for a streamlined yet powerful diagnostic tool, we refined our approach to a panel of five miRNAs through an exhaustive search among all possible combinations from the initial nine miRNAs. The optimal panel, consisting of miR-320b, miR-222-3p, miR-181a-5p, miR-103a-3p, and miR-107, proved highly effective, achieving an AUC of 0.986 (95% CI: 0.978–0.994; sensitivity: 0.975; specificity: 0.958) in GSE164174 (Fig. 3d). Similarly, the panel performed well in other datasets: AUC of 0.977 (95% CI: 0.973–0.980; sensitivity: 0.782; specificity: 0.973) in GSE211692 (Fig. 3e), 0.815 (95% CI: 0.750–0.881; sensitivity: 0.726; specificity: 0.711) in TCGA (Fig. 3f), and 0.811 (95% CI: 0.736–0.887; sensitivity: 0.400; specificity: 0.908) in our serum cohort (Fig. 3g). Moreover, in our cohort, our model also demonstrates superior diagnostic performance when compared to the clinically utilised gastric cancer biomarkers, including CEA, CA19-9, and CA72-4. Particularly, the AUC of our model was significantly higher (Fig. S5A).

These findings underscore the diagnostic potential of our refined miRNA panel in accurately distinguishing ESGC patients from healthy individuals. Moreover, we also investigated the signature's role in AGC patients; the results are shown in Fig. S6.

The five-miRNA signature directly targets PTEN and activates the AKT signalling pathway

To explore how the panel of five miRNAs influences GC development, their target genes were selected from our gastric cancer-specific miRNA regulatory network (Fig. S7A), and functional enrichment analyses were separately performed on the targets of each miRNA in the diagnostic model (Fig. S7B–S7F). We found that the targets of miR-103a-3p and miR-107 were closely related to the p53 signalling pathway, Wnt signalling pathway, and Hippo signalling pathway; the targets of miR-181a-5p and miR-320b were closely associated with the MAPK signalling pathway and FoxO signalling pathway; and miR-222-3p's targets was closely linked to apoptosis. These results reveal a possible pathway through which the five miRNAs play a role in gastric cancer.

Among the various targets of our five-miRNA signatures, PTEN emerged as a notable candidate, being predicted as a target for all five miRNAs (Fig. 4a). To verify these predicted regulatory relationships between the miRNAs and PTEN, we conducted a series of in vitro experiments. Figure 4b illustrated the hybridisation models between the 3'-UTR of human PTEN and each of the five miRNAs. Potential PTEN 3'-UTR binding sites for the five miRNAs were predicted by TargetScan as well as PITA (Fig. 4c). As shown in Fig. 4d, transfection with miRNA mimics significantly decreased the luciferase activity in luciferase plasmids containing the wild-type PTEN 3'-UTR in MKN28 cells, as compared to the negative control mimics. Importantly, mutating the miRNA binding sites eliminated all miRNA-mediated regulatory effects. As shown in Fig. 4e, the mimics of these five miRNAs significantly increased the levels of these miRNAs in both MKN28 and AGS cells, while the inhibitors of these five miRNAs had the opposite effects. qRT-PCR and Western blot analyses further confirmed that the overexpression of these miRNAs led to a downregulation, while their knockdown resulted in an up-regulation, of both PTEN mRNA and protein levels in MKN28 and AGS cells (Fig. 4f, g). Collectively, these findings strongly suggest that our five-miRNA panel directly targets PTEN. Meanwhile, after transfection with



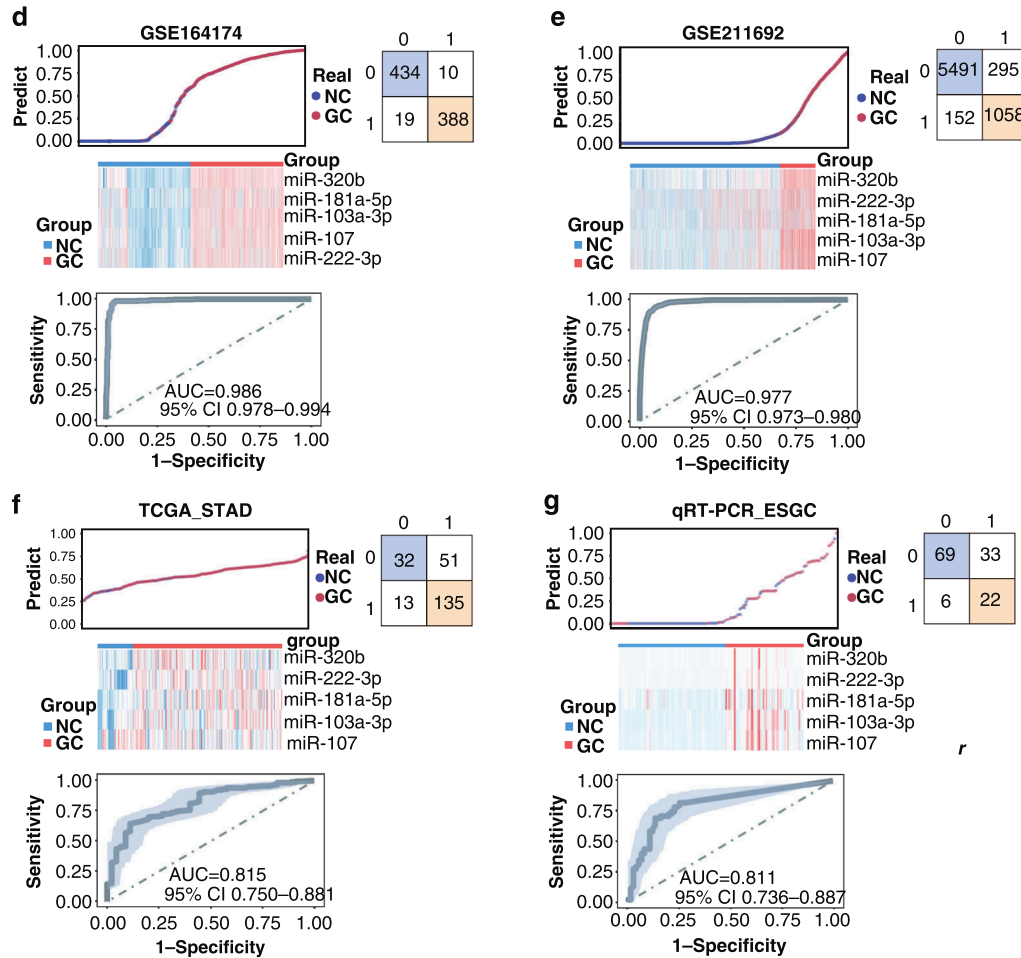


Fig. 3 Continued

miRNA mimics, the level of p-AKT (Ser473) increased, while no significant change in total AKT expression was observed. Conversely, after transfection with miRNA inhibitor, the level of p-AKT decreased, and again, no significant change in total AKT expression was observed. Furthermore, the ratio of p-AKT/AKT was significantly higher in the miRNA mimic group than in the negative control group, whereas it was significantly lower in the miRNA inhibitor group (Fig. 4h). These results suggest that the miRNAs may further activate the AKT pathway by targeting PTEN, thereby affecting the activation level of AKT.

PTEN was involved in the five miRNAs-mediated regulation of cell proliferation, migration, and invasion in GC

To investigate the biological function of PTEN in gastric cancer, we employed small interfering RNA to suppress PTEN expression in MKN28 and AGS cells. As shown in Fig. 5a, three distinct PTEN siRNAs effectively downregulated PTEN protein levels in both MKN28 and AGS cells. Among these PTEN siRNAs, si-PTEN-3 exhibited the strongest inhibitory effect on PTEN expression in both cell lines. Consequently, we selected si-PTEN-3 for subsequent functional analyses. Through clonal formation and EdU assays, we observed that PTEN knockdown significantly enhanced the proliferation rate of MKN28 and AGS cells (Fig. 5b, c). Furthermore, PTEN suppression facilitated both the migration and invasion capabilities of gastric cancer cells (Fig. 5d).

To further validate whether these miRNAs regulated these processes via PTEN, we performed rescue experiments by co-

transfecting with five miRNA inhibitors and si-PTEN-3. The miRNA inhibitors partially reversed the enhanced proliferation, migration, and invasion caused by PTEN knockdown (Fig. 5b, d). These results suggest that miRNAs regulate gastric cancer cell behaviour, at least in part, through the PTEN/AKT axis.

The five miRNA signatures promote tumour growth in vivo

After establishing the subcutaneous xenograft model according to the experimental design (Fig. 5e) and measuring the tumour size to plot the growth curves, we found that the tumour growth was significantly accelerated in the miRNA agomir-injected mouse group, with significant increases in tumour volume and weight (Fig. 5f, g). After collecting the tumours for test, we found that miR-103a-3p, miR-107, miR-181a-5p, miR-320b, and miR-222-3p were significantly increased, while PTEN was significantly decreased (Fig. 5h, i). These findings further validated the biological regulatory functions of the five miRNA signatures.

The five-miRNA signatures exhibiting potential therapeutic implications

Given the extensive body of research on the crucial roles of miRNAs in cancer and their potential clinical applications, miRNAs have emerged as promising targets for small-molecule drug discovery [33–35]. In this study, we also explored the possibility of the identified five miRNAs as drug targets. Consequently, molecular docking was conducted to evaluate the binding affinities of four GC-related drugs, namely Fluorouracil, Irinotecan,

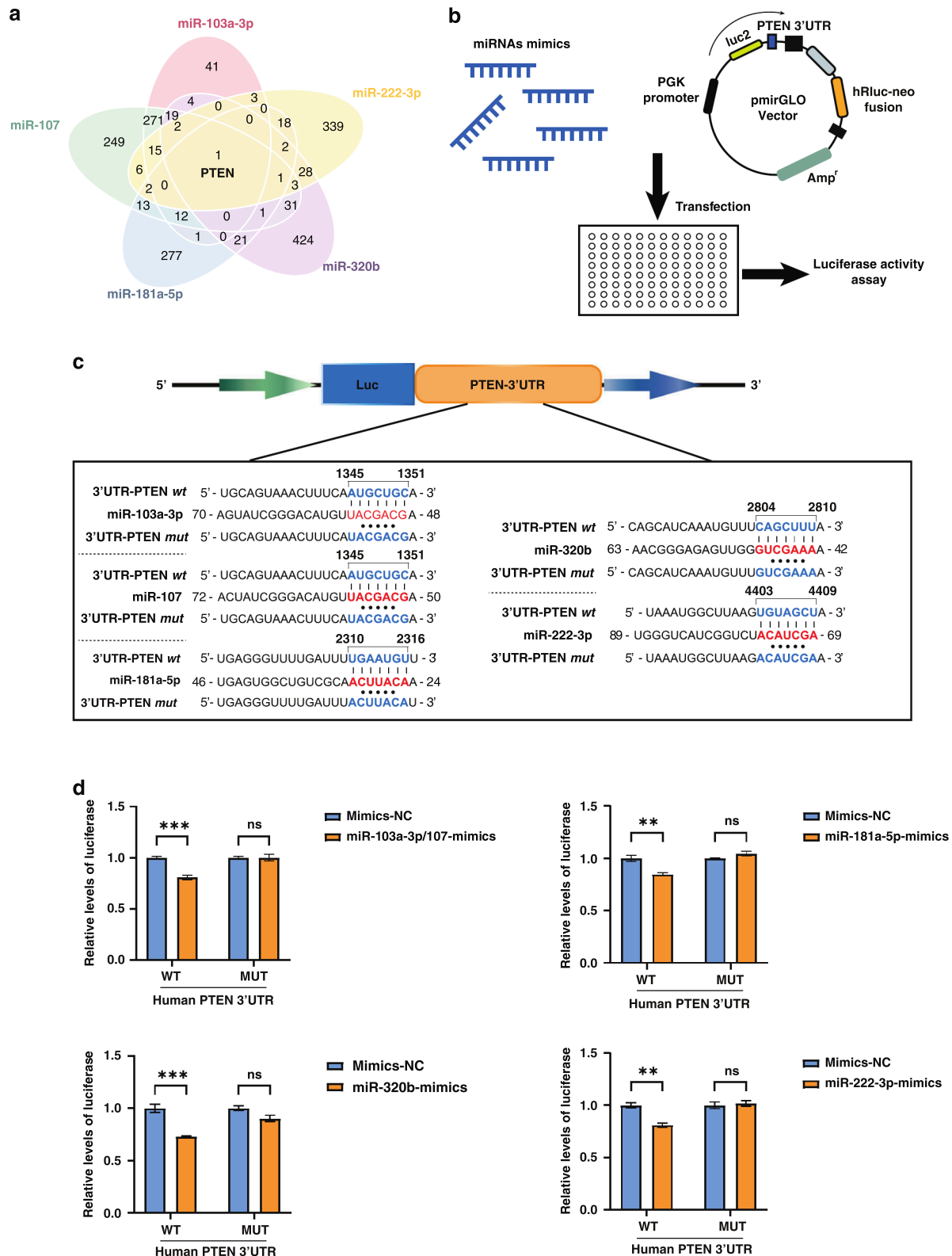


Fig. 4 In vitro experiments verified that five miRNAs co-target PTEN. **a** Venn diagram of the target gene of five miRNAs. **b** The hybridisation models between the 3'-UTR of human PTEN and five miRNAs. **c** The potential PTEN 3'-UTR binding site for 5 miRNAs and its mutated sequence was inserted into the C-terminus of the luciferase gene to generate luciferase reporter plasmids carrying PTEN wild-type (PTEN-3'-UTR-WT) or mutated (PTEN-3'-UTR MUT). **d** Luciferase assay of MKN28 cells transfected with the PTEN-3'-UTR-WT reporter or PTEN-3'-UTR MUT reporter and miRNAs mimics or mimics controls. **e** Validation of five miRNAs mimics/inhibitor transfection efficiency in MKN28 and AGS cells. **f** Validation of PTEN expression by qRT-PCR after transfection with five miRNAs mimics/inhibitors. **g** Validation of PTEN expression by western blot after transfection with five miRNAs mimics/inhibitors. **h** Validation of p-AKT, AKT expression, and p-AKT/AKT by Western Blot after transfection with five miRNAs mimics/inhibitors. * p -value < 0.05, ** p -value < 0.01, *** p -value < 0.001, ns: p -value $\geq$ 0.05. The experiments were all replicated in triplicate.

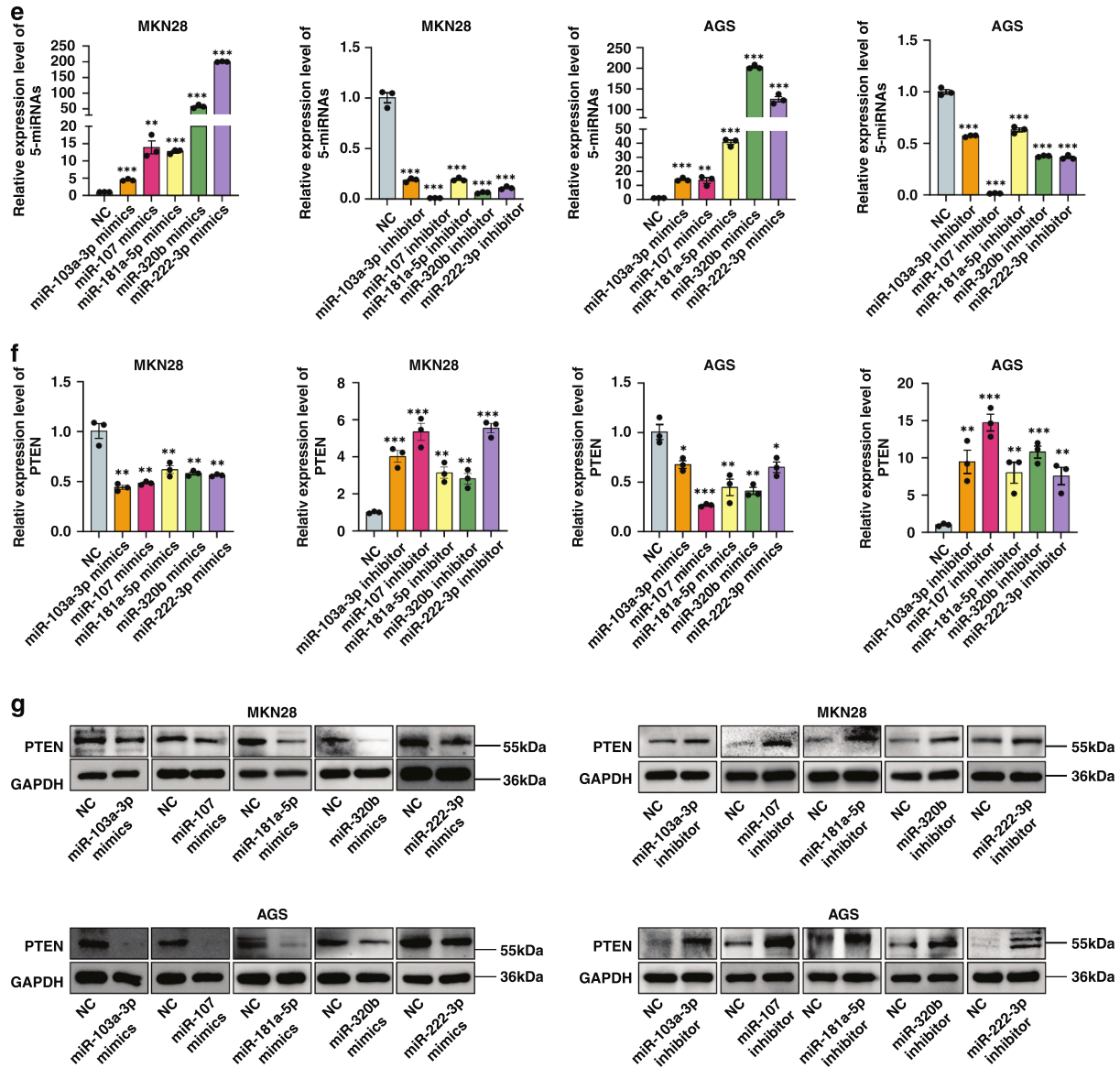


Fig. 4 Continued

Paclitaxel, and Uracil, to the five miRNA precursors. A lower affinity score indicates stronger binding ability. In this study, a binding energy of less than -7 kcal/mol was used as the screening criterion. Paclitaxel exhibited the lowest binding energy with the five miRNA precursors (Fig. 6 and Table S6), demonstrating strong potential interactions between them. The results of the remaining three drugs with miRNAs are listed in Table S7.

DISCUSSION

GC represents a significant global health burden, thereby necessitating the development of effective early diagnostic models and molecular markers to enhance patient prognosis. In this study, we put forward an AI-based strategy, designated as ESGCmiRD, aimed at detecting ESGC by integrating robust expression patterns, ESGC relevance, and regulatory capabilities of miRNA based on topological and biological traits across multiple network types. Employing ESGCmiRD, we identified a five-miRNA signature consisting of miR-320b, miR-222-3p, miR-181a-5p, miR-103a-3p, and miR-107. In different validation sets,

the signature demonstrated exhibited impressive diagnostic performance, achieving an AUC of 0.977 in GSE211692, 0.815 in TCGA-STAD, and 0.811 in our serum cohort, respectively. The performance in the TCGA-STAD dataset was less satisfactory, possibly due to the tissue origin of its samples.

Multiple research groups have explored the potential of using serum miRNAs as biomarkers for gastric cancer detection. Specifically, So JBY et al. developed a panel consisting of 12 miRNAs for detecting GC in Singaporean and Korean subjects [16]. This panel achieved an AUC of 0.92 in their verification cohorts and 0.848 in a prospective study, albeit with lower performance compared to our serum validation dataset GSE211692. Abe S. et al., on the other hand, identified a combination of four serum miRNAs that yielded an impressive AUC of 0.998 in their Japanese population validation set [17]. Nevertheless, both studies were limited to their respective datasets and lacked cross-validation across diverse datasets. Our external validation, conducted on three distinct datasets, underscores the reliability and broad applicability of our five-miRNA signature, highlighting the clinical importance of these identified miRNA biomarkers in detecting

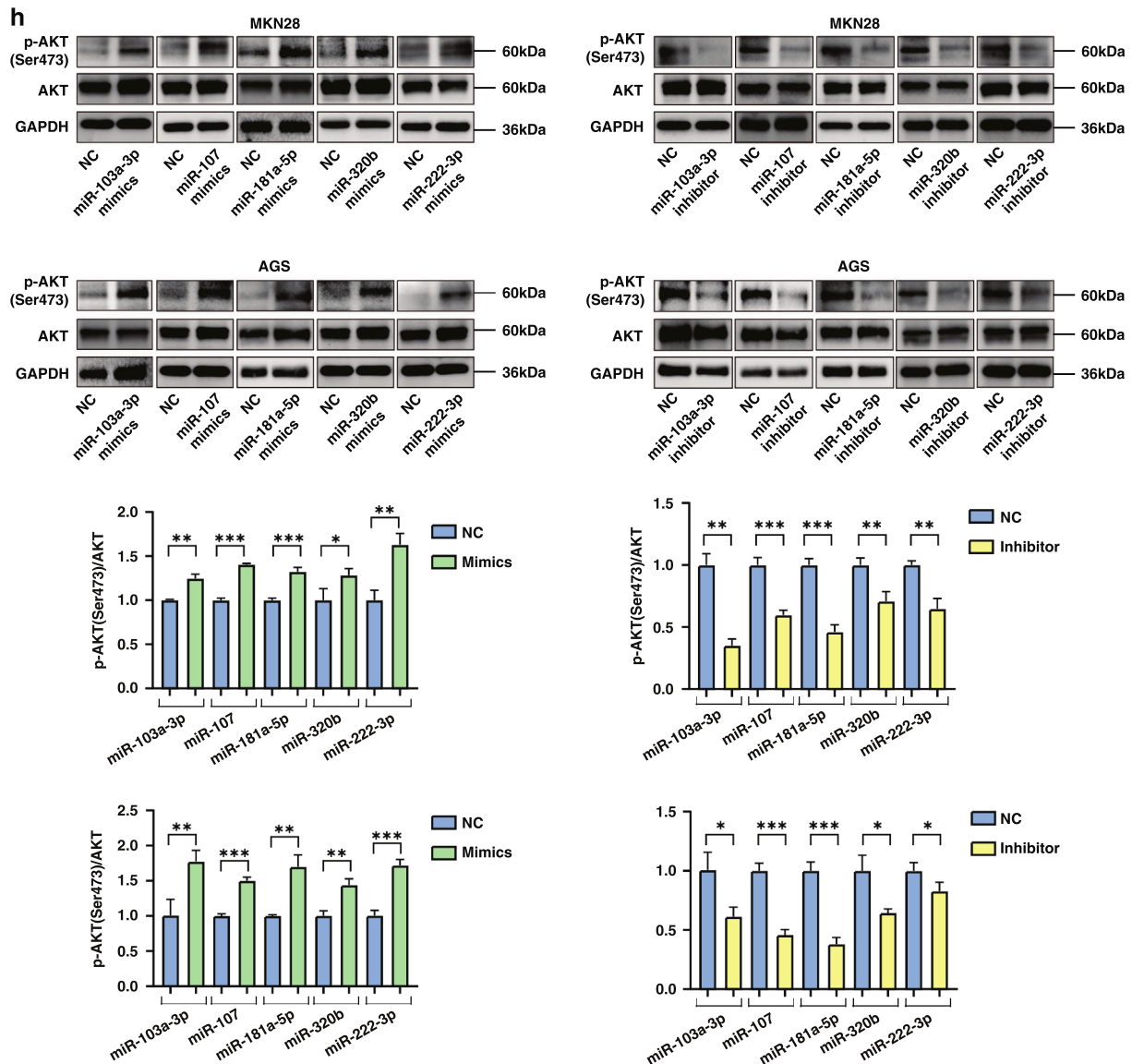


Fig. 4 Continued

ESGC. Furthermore, we also inspected their relationship with the presence or absence of HBV infection by conducting a comparison of the prediction probability based on the five-miRNA panel between the HBV(+) group and the HBV(−) group. As presented in Fig. S5B, there was no significant disparity in the prediction probability between the two groups, suggesting that there is no association of the five-miRNA panel with HBV infection.

We further investigated the roles of a five-miRNA signature in gastric carcinogenesis. This miRNA signature comprises five specific miRNAs: miR-103a-3p, miR-107, miR-181a-5p, miR-222-3p, and miR-320b. Our analysis revealed that the plasma levels of all these miRNAs were elevated in ESGC samples compared to non-cancer control samples in our cohort. To gain deeper insights, we explored the functions of these miRNAs' targets. Functional enrichment analysis indicated that these miRNAs regulate several cancer-related pathways, including the gastric cancer pathway, p53 signalling pathway, Hippo pathway, and Wnt signalling pathway. Although the precise mechanisms by which these miRNAs regulate these pathways in cancer remain to be fully elucidated, previous studies support our findings. Notably, miR-103a-3p has been reported to promote tumour glycolysis in

colorectal cancer by modulating the Hippo pathway [36]. MiR-107 can mediate p53 regulation of hypoxic signalling and tumour angiogenesis [37], and inhibit EGFR downstream signalling in gastric cancer [38].

PTEN is a crucial tumour suppressor gene [39] that plays a significant role in regulating the Hippo and PI3K/Akt pathways, which are known to be involved in gastric tumorigenesis [40, 41]. In our subsequent in vitro experiment, we discovered that all five miRNAs directly target PTEN. Moreover, we observed that the five miRNAs increased the expression of p-AKT, suggesting that these miRNAs may further activate the AKT pathway by targeting PTEN. Knockdown of PTEN was found to significantly enhance the proliferation, migration, and invasion of GC cells. In vivo experiments, we found that they promote growth. Therefore, it is highly plausible that the identified five-miRNA signature promotes gastric tumorigenesis by directly targeting PTEN, underscoring their potential as key biomarkers for ESGC diagnosis and as possible therapeutic targets.

In our study, as a final step, we sought to determine the therapeutic potential of the identified five-miRNA signature by evaluating their binding affinities with GC-related drugs. This

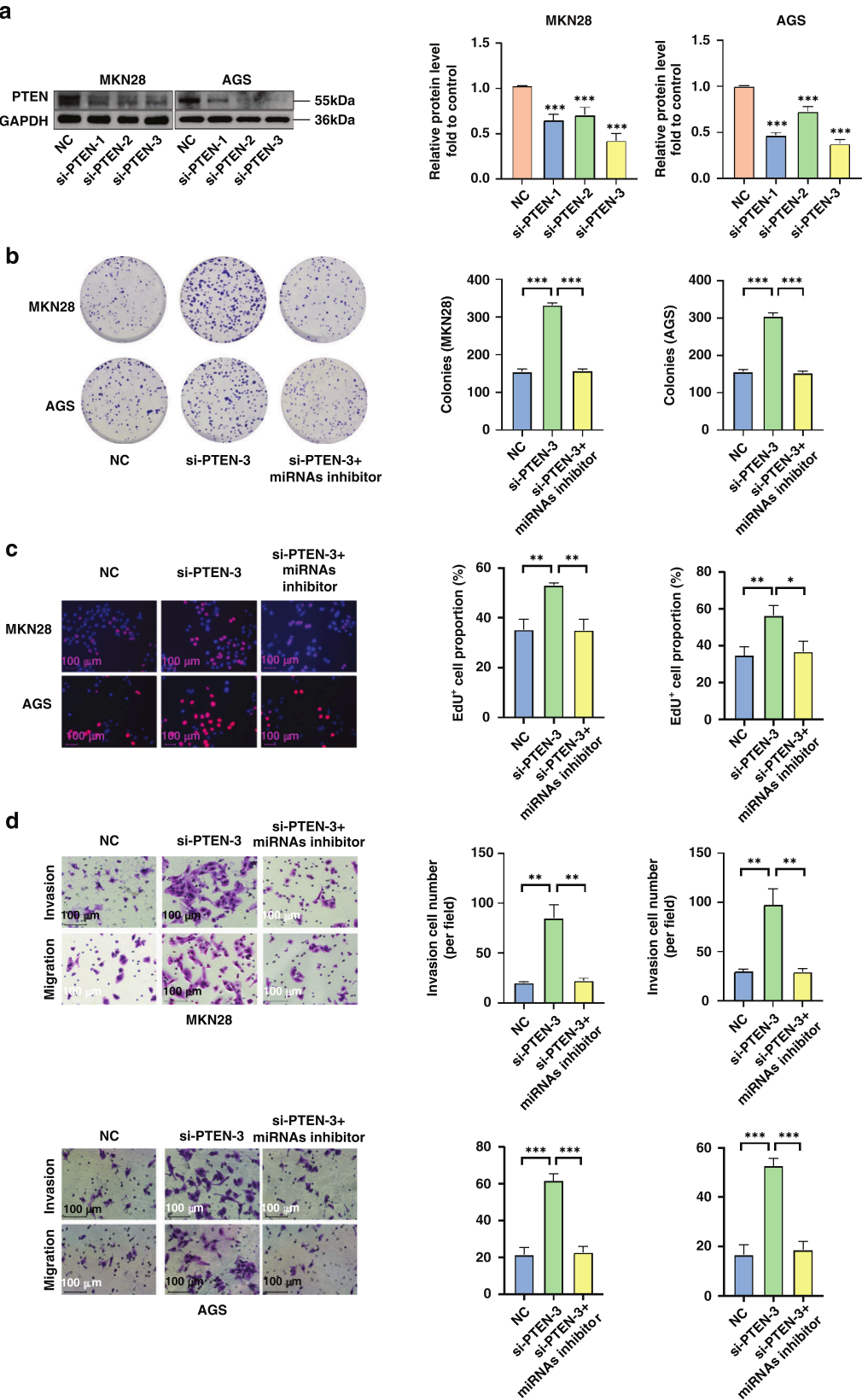


Fig. 5 PTEN was involved in the five miRNAs-mediated regulation of Cell Proliferation, Migration, and Invasion in GC. **a** Validation of miRNA knockdown efficiency in MKN28 and AGS. **b** Colony formation assays were performed to assess colony formation on PTEN knockdown and co-transfection with miRNA inhibitors MKN-28 and AGS cells. **c** EdU analysis was performed on PTEN knockdown and co-transfection with miRNA inhibitors MKN-28 and AGS cells to measure their proliferative capacity. The image scale is 100 μ m. The percentage of EdU-positive cells in transfected GC cells was statistically analysed and shown in the bar graph. **d** Transwell migration and invasion assays were performed to detect the migration and invasion ability of PTEN knockdown and co-transfection with miRNA inhibitors MKN-28 and AGS cells. The image scale is 100 μ m. The experiments were all replicated in triplicate. **e** Flowchart for subcutaneous xenograft experiments in mice. **f** Representative images of NC and Agomir group mice with subcutaneous xenograft tumours. **g** Growth curves and tumour weight of NC and Agomir group mice with subcutaneous xenograft tumours. **h** The qRT-PCR analysis of the expression levels of miR-103a-3p, miR-107, miR-181a-5p, miR-320b, miR-222-3p, and PTEN in xenograft tissues from the Agomir or NC groups. **i** Representative IHC images of PTEN in xenograft tissue from the Agomir and NC groups. * p -value < 0.05, ** p -value < 0.01, *** p -value < 0.001.

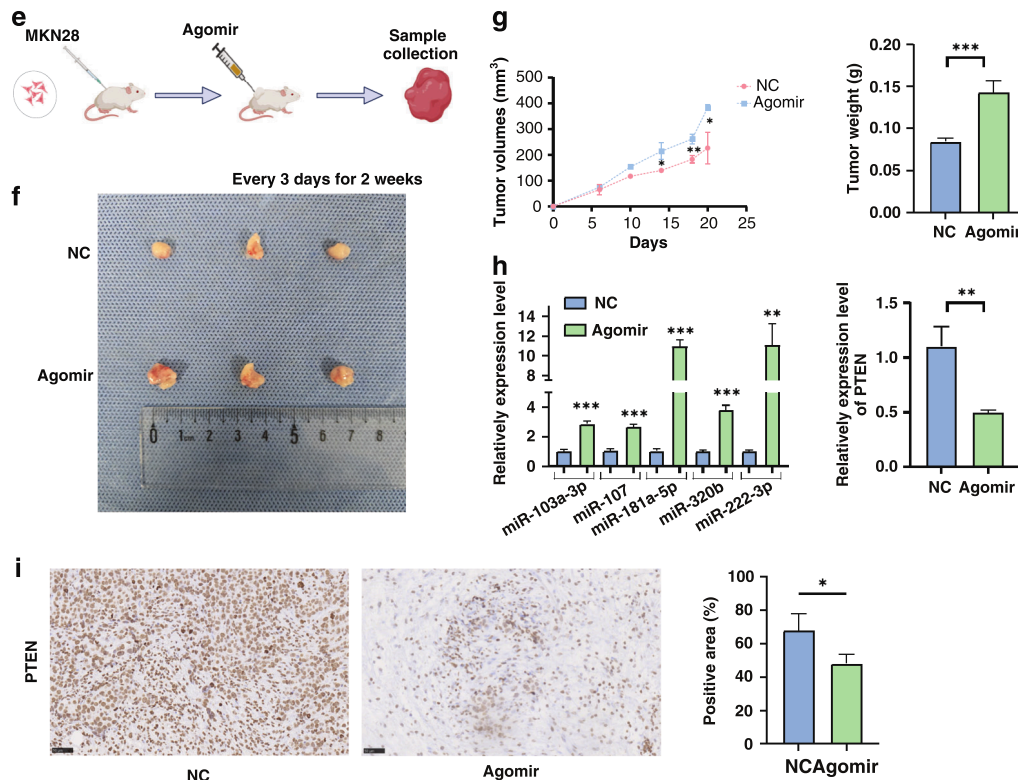


Fig. 5 Continued

exploration was motivated by the emerging recognition of miRNAs as promising targets for small-molecule drug discovery [35]. Among the drugs tested, Paclitaxel—a widely used chemotherapy agent in the treatment of multiple cancers [42, 43]—stood out due to its strong binding affinities with all five miRNAs. This finding suggests that these miRNAs may represent novel targets for Paclitaxel, thereby revealing a new mechanism underlying its therapeutic effect in GC. Furthermore, five-miRNAs in blood potentially offer a new indicator for monitoring and overcoming drug resistance.

This study's findings provide significant insights into the potential clinical applications of miRNAs in GC diagnosis and therapy. However, limitations and obstacles must be acknowledged. Firstly, reliance on past datasets has inherent constraints, such as potential biases that can affect research understanding and applicability. Prospective clinical trials with well-defined cohorts and standardised protocols are needed to validate the clinical utility and reliability of miRNA biomarkers in GC diagnosis. Secondly, the complexity and heterogeneity of GC pose a significant obstacle in biomarker identification and verification. Research should focus on miRNA markers specific to each subtype and their clinical relevance to enhance diagnostic

precision and specificity. Meanwhile, although liquid biopsy is a less invasive and convenient cancer detection method, it has limitations. Variations in sample collection, processing methods, and analysis approaches can lead to bias and affect the reproducibility and reliability of miRNA measurements. Therefore, extensive multicenter cohorts and prospective clinical studies are essential to thoroughly evaluate the clinical application of miRNA-based diagnostic models. A comprehensive evaluation of miRNA assays' performance in real-world clinical settings, including sensitivity, specificity, reproducibility, and cost-effectiveness, is necessary to support their integration into standard clinical practice.

CONCLUSIONS

In conclusion, we have proposed an AI-driven approach to pinpoint a distinct five-miRNA signature for ESGC detection. This miRNA signature plays a pivotal role in gastric cancer by directly targeting the PTEN gene, a key tumour suppressor. This strategy not only advances our understanding of gastric cancer biology but also opens up new possibilities for early diagnosis and potential therapeutic interventions.

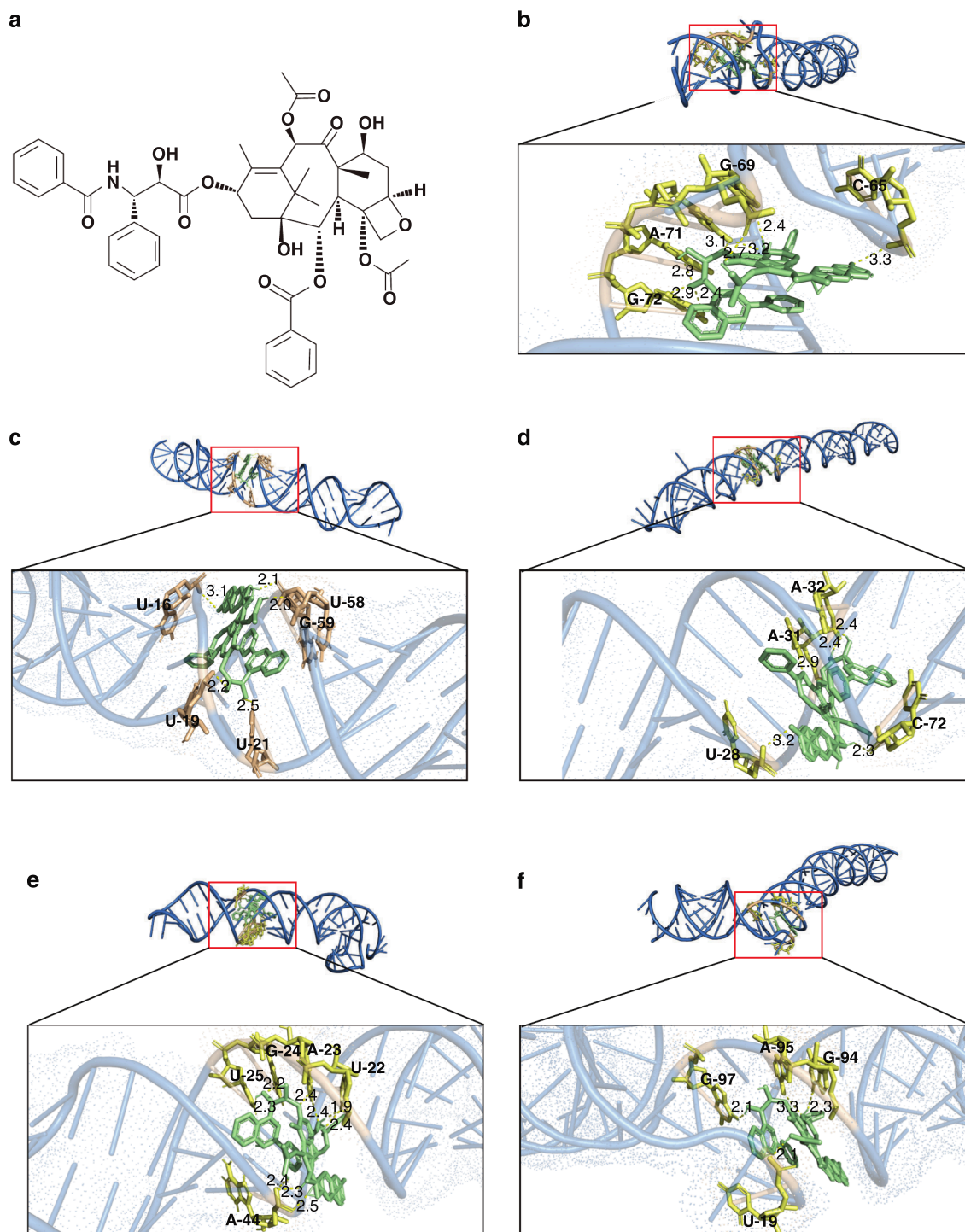


Fig. 6 Visualisation of docking conformations of paclitaxel and five miRNAs. **a** 2D structure of paclitaxel. **b** Binding interactions of paclitaxel with pre-miR-103a-3p. **c** Binding interactions of paclitaxel with pre-miR-107. **d** Binding interactions of paclitaxel with pre-miR-181a-5p. **e** Binding interactions of paclitaxel with pre-miR-320b. **f** Binding interactions of paclitaxel with pre-miR-222-3p. Numbered nucleotides represent their positions in the pre-miRNA sequence.

DATA AVAILABILITY

All data generated or analysed are included in this article and its supplementary information files. The code can be found at <https://github.com/ljc7878/ESGCMiRD.git>.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71:209–49.
2. Li Y, Feng A, Zheng S, Chen C, Lyu J. Recent estimates and predictions of 5-year survival in patients with gastric cancer: a model-based period analysis. *Cancer Control: J Moffitt Cancer Cent.* 2022;29:10732748221099227.
3. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet.* 2020;396:635–48.
4. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *Cancer J Clin.* 2021;71:264–79.
5. Huang Z-B, Zhang H-T, Yu B, Yu D-H. Cell-free DNA as a liquid biopsy for early detection of gastric cancer. *Oncol Lett.* 2021;21:3.

6. Ma S, Zhou M, Xu Y, Gu X, Zou M, Abudushalamu G, et al. Clinical application and detection techniques of liquid biopsy in gastric cancer. *Mol Cancer*. 2023;22:7.
7. Guo X, Peng Y, Song Q, Wei J, Wang X, Ru Y, et al. A liquid biopsy signature for the early detection of gastric cancer in patients. *Gastroenterology*. 2023;165:402–13.e13.
8. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic—implementation issues and future challenges. *Nat Rev Clin Oncol*. 2021;18:297–312.
9. Lone SN, Nisar S, Masoodi T, Singh M, Rizwan A, Hashem S, et al. Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments. *Mol Cancer*. 2022;21:79.
10. Nikanjam M, Kato S, Kurzrock R. Liquid biopsy: current technology and clinical applications. *J Hematol Oncol*. 2022;15:131.
11. Alix-Panabières C, Pantel K. Liquid biopsy: from discovery to clinical application. *Cancer Discov*. 2021;11:858–73.
12. Zhang Z, Wu H, Chong W, Shang L, Jing C, Li L. Liquid biopsy in gastric cancer: predictive and prognostic biomarkers. *Cell Death Dis*. 2022;13:903.
13. Farasati Far B, Vakili K, Fathi M, Yaghoobpoor S, Bhia M, Naimi-Jamal MR. The role of microRNA-21 (miR-21) in pathogenesis, diagnosis, and prognosis of gastro-intestinal cancers: a review. *Life Sci*. 2023;316:121340.
14. Guo T, Tang X-H, Gao X-Y, Zhou Y, Jin B, Deng Z-Q, et al. A liquid biopsy signature of circulating exosome-derived mRNAs, miRNAs and lncRNAs predict therapeutic efficacy to neoadjuvant chemotherapy in patients with advanced gastric cancer. *Mol Cancer*. 2022;21:216.
15. Yan W, Chen Y, Hu G, Shi T, Liu X, Li J, et al. MiR-200/183 family-mediated module biomarker for gastric cancer progression: an AI-assisted bioinformatics method with experimental functional survey. *J Transl Med*. 2023;21:163.
16. So JBY, Kapoor R, Zhu F, Koh C, Zhou L, Zou R, et al. Development and validation of a serum microRNA biomarker panel for detecting gastric cancer in a high-risk population. *Gut*. 2021;70:829–37.
17. Abe S, Matsuzaki J, Sudo K, Oda I, Katai H, Kato K, et al. A novel combination of serum microRNAs for the detection of early gastric cancer. *Gastric Cancer*. 2021;24:835–43.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
19. Kolde R, Laur S, Adler P, Vilo J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics*. 2012;28:573–80.
20. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma*. 2008;9:559.
21. Lin Y, Qi X, Chen J, Shen B. Multivariate competing endogenous RNA network characterization for cancer microRNA biomarker discovery: a novel bioinformatics model with application to prostate cancer metastasis. *Precis Clin Med*. 2022;5:pbac001.
22. Lin Y, Chen F, Shen L, Tang X, Du C, Sun Z, et al. Biomarker microRNAs for prostate cancer metastasis: screened with a network vulnerability analysis model. *J Transl Med*. 2018;16:134.
23. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. *Nat Genet*. 2007;39:1278–84.
24. Huang H-Y, Lin Y-C-D, Cui S, Huang Y, Tang Y, Xu J, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res*. 2022;50:D222–30.
25. Yan W, Liu X, Wang Y, Han S, Wang F, Liu X, et al. Identifying drug targets in pancreatic ductal adenocarcinoma through machine learning, analyzing bio-molecular networks, and structural modeling. *Front Pharmacol*. 2020;11:534.
26. Csardi G, Nepusz T. The Igraph software package for complex network research. *InterJournal. Complex Systems*; 2006. p. 1695.
27. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J Integr Biol*. 2012;16:284–7.
28. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(delta delta C(T)) method. *Methods*. 2001;25:402–8.
29. Knox C, Wilson M, Klinger CM, Franklin M, Oler E, Wilson A, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. *Nucleic Acids Res*. 2024;52:D1265–75.
30. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019;47:D155–62.
31. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630:493–500.
32. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31:455–61.
33. Ling H, Fabbri M, Calin GA. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. *Nat Rev Drug Discov*. 2013;12:847–65.
34. Kim T, Croce CM. MicroRNA: trends in clinical trials of cancer diagnosis and therapy strategies. *Exp Mol Med*. 2023;55:1314–21.
35. Winkle M, El-Daly SM, Fabbri M, Calin GA. Noncoding RNA therapeutics—challenges and potential solutions. *Nat Rev Drug Discov*. 2021;20:629–51.
36. Sun Z, Zhang Q, Yuan W, Li X, Chen C, Guo Y, et al. MiR-103a-3p promotes tumour glycolysis in colorectal cancer via hippo/YAP1/HIF1A axis. *J Exp Clin Cancer Res*. 2020;39:250.
37. Leslie PL, Franklin DA, Liu Y, Zhang Y. p53 regulates the expression of LRP1 and apoptosis through a stress intensity-dependent microRNA feedback loop. *Cell Rep*. 2018;24:1484–95.
38. Wang P, Zhou Y, Wang J, Zhou Y, Zhang X, Liu Y, et al. miR-107 reverses the multidrug resistance of gastric cancer by targeting the CGA/EGFR/GATA2 positive feedback circuit. *J Biol Chem*. 2024;300:107522.
39. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*. 1997;275:1943–7.
40. Xu W, Yang Z, Xie C, Zhu Y, Shu X, Zhang Z, et al. PTEN lipid phosphatase inactivation links the hippo and PI3K/Akt pathways to induce gastric tumorigenesis. *J Exp Clin Cancer Res*. 2018;37:198.
41. Messina B, Lo Sardo F, Scalera S, Memeo L, Colarossi C, Mare M, et al. Hippo pathway dysregulation in gastric cancer: from *Helicobacter pylori* infection to tumor promotion and progression. *Cell Death Dis*. 2023;14:21.
42. Baird RD, Tan DS, Kaye SB. Weekly paclitaxel in the treatment of recurrent ovarian cancer. *Nat Rev Clin Oncol*. 2010;7:575–82.
43. Yoon HH. Ramucirumab plus paclitaxel for gastric cancer in China. *Lancet Gastroenterol Hepatol*. 2021;6:975–6.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my supervisor, Prof. Weichang Chen, for your careful guidance and unwavering support, which greatly benefited me in the process of selecting the topic, designing the research methodology and writing the thesis. Your rigorous attitude and profound knowledge have provided endless insights into my research and pushed me to continuously strive for excellence. Meanwhile, I would like to thank my supervisors, Prof Yan Wenying and Prof Shi Tongguo, who have given me important guidance and support in my research journey, and whose high-level insights and rich research experience have enabled me to avoid many misunderstandings in my research and provided valuable support for my academic growth. During the sample collection process of this project, I am especially grateful to all the colleagues who participated in the sample collection work. Without their hard work and precision, this study could not have been carried out smoothly. Finally, I would like to thank all those who have helped and supported me throughout my academic career.

AUTHOR CONTRIBUTIONS

Jiachun Lu: writing—original draft, formal analysis, investigation, methodology, software, visualisation, data curation. Yuqi Chen: conceptualisation, methodology, investigation, resources. Jiayu Wang: conceptualisation, methodology, investigation. Yuxin He: methodology, investigation. Xin Liu: methodology, investigation. Tongguo Shi: project administration, methodology, supervision, validation, writing—review & editing. Weichang Chen: conceptualisation, funding acquisition, project administration, supervision, writing—review & editing. Wenying Yan: conceptualisation, methodology, project administration, supervision, funding acquisition, writing—review & editing.

FUNDING

This research was supported by the Key Research and Development Programme of Jiangsu Province (BE2020656); Medical and Health Science and Technology Innovation Project of Suzhou (SKY2022010, SKYD2022097); Foundation of Suzhou Medical College of Soochow University (MP13405423, MX13401423); National Natural Science Foundation of China (82270561, 82073156); Jiangsu Provincial Medical Key Discipline (ZDXK202246); and the Priority Academic Programme Development of Jiangsu Higher Education Institutions (SKYD2022097).

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institution and with the 1964 Helsinki Declaration. Written informed consent was obtained from individual or guardian participants. The study was approved by the Bioethics Committee of the First Affiliated Hospital of Soochow University (Approval No. 2021068). All animal

experiments were performed in accordance with the institutional guidelines of the Soochow Animal Care and Use Committee (No.202412A0046).

CONSENT FOR PUBLICATION

The work described has not been previously published, the consent to publish has been obtained from all co-authors, and the publication has been approved by the competent authorities of the institution in which the work was carried out.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41416-025-02984-9>.

Correspondence and requests for materials should be addressed to Tongguo Shi, Weichang Chen or Wenying Yan.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.