SCIENCE CHINA PRESS

SCIENCE CHINA Life Sciences
life.scichina.com link.springer.com
https://doi.org/10.1007/s11427-024-2642-8

RESEARCH PAPER

Novel DNA methylation markers for early detection of gastric cardia adenocarcinoma and esophageal squamous cell carcinoma

Zhiyuan Fan[1], Jiajie Hao[2], Feifan He[1], Hao Jiang[3], Jinwu Wang[4], Minjuan Li[5], Xinqing Li[1], Ru Chen[1] & Wenqiang Wei[1,6*]

[1]Office of National Central Cancer Registry, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;
[2]State Key Laboratory of Molecular Oncology, Center for Cancer Precision Medicine, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;
[3]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China;
[4]Department of Pathology, Linzhou Cancer Hospital, Linzhou 456550, China;
[5]Department of Orthopedic Surgery, Beijing Jishuitan Hospital, Capital Medical University, Beijing 100035, China;
[6]Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing 211166, China

*Corresponding author (email: weiwq@cicams.ac.cn)

Received 29 April 2024; Accepted 29 May 2024; Published online 30 August 2024

Gastric cardia adenocarcinoma (GCA) and esophageal squamous cell carcinoma (ESCC) present significant health challenges in China, often diagnosed at advanced stages with poor prognoses. However, effective biomarkers for early detection remain elusive. This study aimed to integrate methylome and transcriptome data to identify DNA methylation markers for the early detection of GCA and ESCC. In the discovery stage, we conducted Infinium MethylationEPIC array analysis on 36 paired GCA and non-tumor adjacent tissues (NAT), identifying differentially methylated CpG sites (DMCs) between GCA/ESCC and NAT through combined analyses of in-house and publicly available data. In the validation stage, targeted pyrosequencing and quantitative real-time RT-PCR were performed on paired tumor and NAT samples from 50 GCA and 50 ESCC patients. In the application stage, an independent set of 438 samples, including GCA, ESCC, high- and low-grade dysplasia (HGD/LGD), and normal controls, was tested for selected DMCs using pyrosequencing. Our analysis validated three GCA-specific, two ESCC-specific, and one tumor-shared DMCs, exhibiting significant hypermethylation and decreased expression of target genes in tumor samples compared with NAT. Leveraging these DMCs, we developed a GCA-specific 4-marker panel (cg27284428, cg11798358, cg07880787, and cg00585116) with an area under the receiver operating characteristic curve (AUC) of 0.917, effectively distinguishing between cardia HGD/GCA patients and cardia LGD/normal controls. Similarly, an ESCC-specific 3-marker panel (cg14633892, cg04415798, and cg00585116) achieved an AUC of 0.865 in distinguishing esophageal HGD/ESCC cases. Furthermore, integrating cg00585116, age, and alcohol consumption yielded a tumor-shared logistic model with good discrimination for two cancer/HGD (AUC, 0.767; 95% confidence interval, 0.720–0.813). The mean AUC of the model after 5-fold cross-validation was 0.764. In summary, our study identifies novel DNA methylation markers capable of accurately distinguishing GCA/ESCC and HGD from LGD and normal controls. These findings offer promising prospects for targeted DNA methylation assays in future minimally invasive cancer screening methods.

gastric cardia adenocarcinoma | esophageal squamous cell carcinoma | high-grade dysplasia | DNA methylation markers | early detection

INTRODUCTION

Gastric cancer (GC) and esophageal cancer (EC) are prevalent upper gastrointestinal malignancies, ranking as the fifth and seventh leading causes of cancer mortality worldwide, respectively (Bray et al., 2024). Both cancers can be categorized into two major subtypes with distinct epidemiological features. GC has two major topographic subsites: gastric cardia adenocarcinoma (GCA), defined as cancer involving the esophagogastric junction with epicenter within 2 cm into the proximal stomach, and gastric non-cardia adenocarcinoma (GNCA) originating from more distal regions of the stomach (Rice et al., 2017; Thrift and El-Serag, 2020). Over recent decades, there has been a global trend of increasing GCA incidence (Arnold et al., 2020; Gao et al., 2023; Parfitt et al., 2006). Furthermore, esophageal squamous cell carcinoma (ESCC) accounts for nearly 84% of the global EC burden, outnumbering esophageal adenocarcinoma (Arnold et al., 2020). Notably, the incidence rates of GCA and ESCC were simultaneously very high at over 20 per 100,000 in some locations, especially in the Asian populations (such as Linzhou, Cixian and Shexian in China), indicating that there might be some shared etiology between these two cancers (He et al., 2024; Li et al., 2023).

Due to the absence of alarm symptoms, most GCA and ESCC patients are typically diagnosed at advanced stages, resulting in poor prognoses (Maharjan and Kauppila, 2022; Zeng et al., 2018). Thus, early detection of these cancers, when curative treatments are still feasible, is crucial for improving patient

outcomes and extending survival (Wang et al., 2023). While endoscopic screening has proven effective in reducing morbidity and mortality associated with GCA and ESCC, its widespread use is limited by high costs and invasiveness (Chen et al., 2021). Early detection of these malignancies remains challenging. Hence, there is an urgent need to identify effective biomarkers for early detection of these two anatomically adjacent but histologically distinct tumors.

Aberrant DNA methylation (DNAme), a well-established epigenetic hallmark of cancer, plays an essential role in human tumor development, including GCA and ESCC (Cao et al., 2020; Guo et al., 2016). Compared with other genetic alterations, aberrant DNAme has shown superior efficacy in early cancer detection due to its tissue-specific nature, early emergence during carcinogenesis, and relative stability over time (Xi et al., 2022). Both genetics and environmental factors can influence DNAme, potentially leading to distinct methylation patterns in GCA and GNCA (Li et al., 2015). However, most previous studies have treated GC as a single entity without distinguishing between subsites, thereby lacking comprehensive characterization of the genome-wide methylation profiles of GCA (Li et al., 2020; Usui et al., 2023). Although prior genome-wide methylation studies have shown promise for early detection of ESCC, they have been limited by the lack of integrative analyses of DNA methylome and transcriptome in the same patients or by ignoring the diagnostic value of DNAme for precancerous lesions (Talukdar et al., 2021; Xi et al., 2022). Moreover, there is limited evidence regarding the potential shared mechanisms of DNAme in the pathogenesis of GCA and ESCC.

To address these limitations, we first conducted a comprehensive analysis by integrating methylome and transcriptome profiles from in-house and publicly available data, identifying tumor-specific and tumor-shared differentially methylated CpG sites (DMCs) between GCA/ESCC tumors and non-tumor adjacent tissues (NAT). We then validated these candidate markers in an independent validation set, utilizing matched methylation and expression data from the same patients. Lastly, we assessed the performance of our methylation markers for early detection of GCA, ESCC and precancerous lesions in an independent application set.

RESULTS

Characteristics of the study participants

The demographic characteristics of the participants are summarized in Table 1. In the discovery set, the mean age of GCA patients was 65.4 years (standard deviation (SD), 8.4 years), with a predominance of male patients (77.8%, 28/36). In addition, 47.2% and 44.4% of GCA patients reported tobacco and alcohol consumption, respectively. In the validation set, the mean age was 65.3 years (SD, 7.5 years) for GCA patients and 67.2 years (SD, 5.8 years) for ESCC patients. The majority of patients were male (82% GCA and 54% ESCC), smokers (68% GCA and 60% ESCC), and drinkers (66% GCA and 54% ESCC). Moreover, 22% of GCA cases and 58% of ESCC cases were at early stages (stage I and II). The application set comprised 438 subjects with varying grades of dysplasia. Patients with high-grade dysplasia (HGD)

Table 1. The characteristics of participants

Parameters	Discovery set	Validation set		Application set						
	GCA	GCA	ESCC	CHGD/GCA	NC/CLGD	*P* value	EHGD/ESCC	NE/ELGD	*P* value	*P* value[a)]
	(*n*=36)	(*n*=50)	(*n*=50)	(*n*=129)	(*n*=80)		(*n*=176)	(*n*=53)		
Age, mean±SD, *y*	65.4±8.4	65.3±7.5	67.2±5.8	64.8±7.9	57.6±5.9	<0.001	63.6±6.9	62.9±5.3	0.675	<0.001
Gender, *n* (%)						0.03			<0.001	<0.001
Male	28 (77.8)	41 (82.0)	27 (54.0)	93 (72.1)	46 (57.5)		111 (63.1)	19 (35.8)		
Female	8 (22.2)	9 (18.0)	23 (46.0)	36 (27.9)	34 (42.5)		65 (36.9)	34 (64.2)		
Tobacco smoking, *n* (%)						0.303			0.001	0.005
Yes	17 (47.2)	34 (68.0)	30 (60.0)	61 (47.3)	32 (40.0)		82 (46.6)	11 (20.8)		
No	19 (52.8)	16 (32.0)	20 (40.0)	68 (52.7)	48 (60.0)		94 (53.4)	42 (79.2)		
Alcohol intake, *n* (%)						0.212			<0.001	<0.001
Yes	16 (44.4)	33 (66.0)	27 (54.0)	41 (31.8)	19 (23.8)		76 (43.2)	6 (11.3)		
No	20 (55.6)	17 (34.0)	23 (46.0)	88 (68.2)	61 (76.3)		100 (56.8)	47 (88.7)		
Stage, *n* (%)										
Early stage (stage I–II)	6 (16.7)	11 (22.0)	29 (58.0)	25 (26.0)			54 (54.0)			
Late stage (stage III–IV)	30 (83.3)	39 (78.0)	21 (42.0)	71 (74.0)			46 (46.0)			
Normal control, *n* (%)					56 (70.0)			13 (24.5)		
LGD, *n* (%)					24 (30.0)			40 (75.5)		
HGD, *n* (%)				33 (25.6)			76 (43.2)			
cancer, *n* (%)				96 (74.4)			100 (56.8)			

a) Cancer/HGD compared with normal/LGD. Cancer includes GCA and ESCC, HGD includes CHGD and EHGD, LGD includes CLGD and ELGD, and Normal control includes NC and NE. CHGD, cardia high-grade dysplasia; CLGD, cardia low-grade dysplasia; NC, normal cardia; EHGD, esophageal high-grade dysplasia; ELGD, esophageal low-grade dysplasia; NE, normal esophagus.

https://doi.org/10.1007/s11427-024-2642-8

and cancer were older, had a higher proportion of males, and reported more frequent tobacco and alcohol use compared with those with low-grade dysplasia (LGD) and normal tissues (P<0.05). An overview of our study design and participant flow is presented in Figure S1 in Supporting Information.

Discovery of DNAme markers to distinguish GCA and ESCC from normal tissues

To identify DNAme markers for distinguishing GCA from NAT, we conducted genome-wide DNAme analyses on 36 paired GCA and NAT samples using HM850K methylation array. A total of 40,332 CpGs exhibited differential methylation ($|\Delta\beta|$>0.15 and adjusted P<0.05) between GCA and NAT samples, with 36.0% (14,501/40,332) of these CpGs hypermethylated (Figure 1A). Among them, 4,920 DMCs located in promoter regions of 2,063 unique genes were selected. We further compared the expression patterns of these genes between 62 paired GCA and NAT samples obtained from GSE29272. We observed that hypermethylation of 216 CpGs inversely correlated with the expression levels of their 67 target genes, which had at least 1.5-fold higher expression in NAT samples compared with GCA tissues (Figure 1B). Utilizing

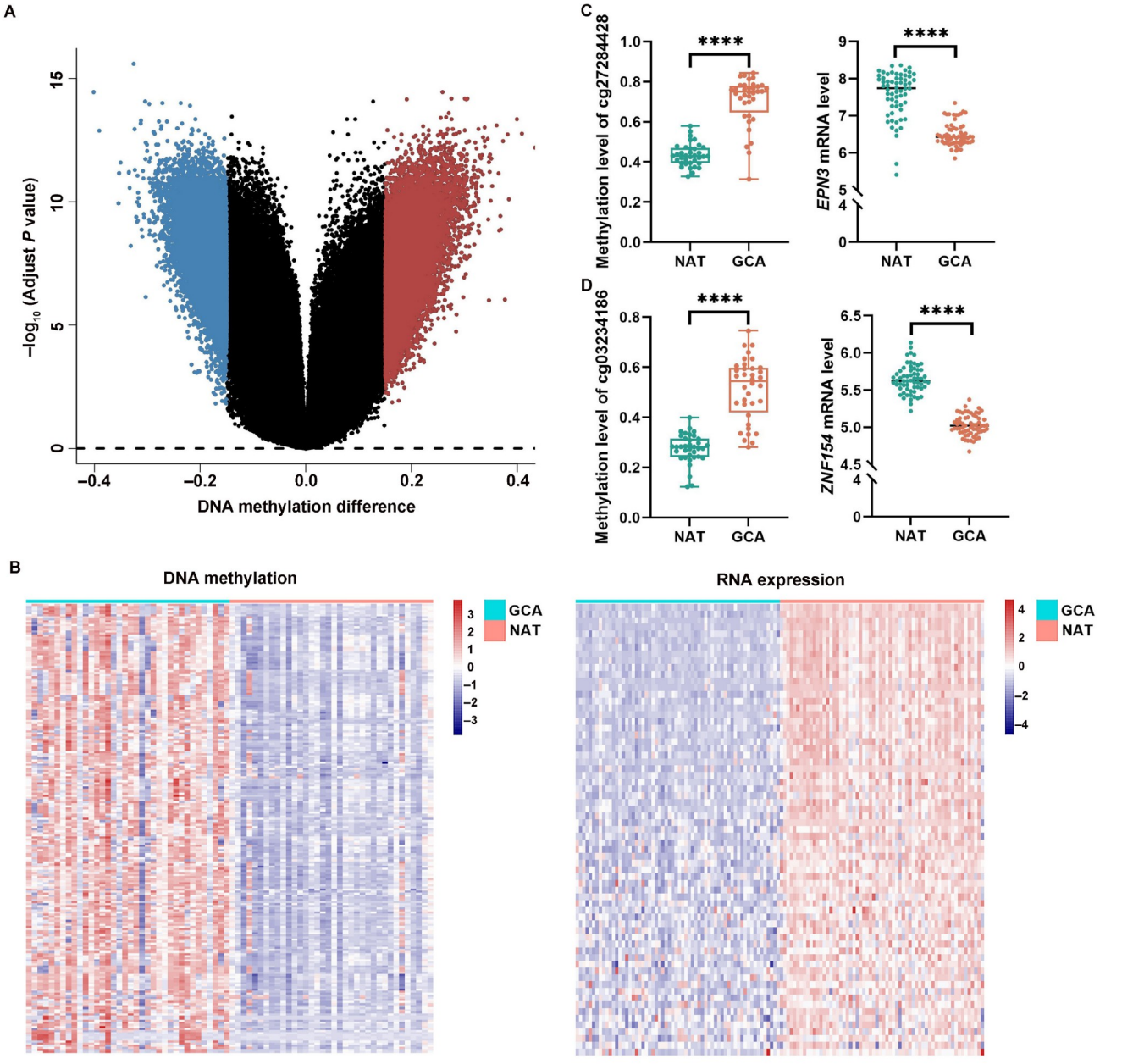

Figure 1. Identification of DNA methylation biomarkers to distinguish GCA from normal tissue. A, Volcano plot of DMCs between GCA and NAT. Blue, the significantly hypomethylated CpGs; red, the significantly hypermethylated CpGs. B, Heatmap of identified 216 DMCs (left) and 67 target genes (right). The row represents individual CpGs/genes, and the column represents individual samples. Blue, GCA; red, NAT samples. The color in the heatmap represents the methylation levels of CpGs and expression levels of target genes. Red, hypermethylated CpGs/higher expressed genes; Blue, hypomethylated CpGs/lower expressed genes. C and D, Differential DNA methylation level of cg27284428 and expression level of target gene *EPN3* (C), cg03234186 and target gene *ZNF154* (D). A methylation level of 0 represents no methylation, and 1 represents full methylation. ****, P<0.0001

an overlap of Least Absolute Shrinkage and Selector Operation (LASSO) and random forest (RF) algorithm, we identified two DMCs (cg27284428 in *EPN3* and cg03234186 in *ZNF154*) for further analysis (Figure 1C and D).

We next identified DMCs exhibiting higher methylation and lower expression of their respective genes in ESCC than in NAT samples. Analysis of two cohorts from the Gene Expression Omnibus (GEO) DNAme datasets (GSE164083 and GSE121932) revealed 16,860 hypermethylated CpGs located in the promoter regions of 5,629 unique genes ($|\Delta\beta|>0.15$ and adjusted $P<0.05$). Further integrative analysis of DNAme and gene expression datasets (GSE38129 and GSE20347) identified 1,188 DMCs located in 369 unique genes with an inverse methylation-expression correlation (Figure 2A). By combining the DMCs selected through the LASSO and RF methods, we identified six ESCC-specific DMCs with good discrimination power, including cg04415798 in *PAX9*, cg15590153 in *PGM1*, cg07542675 in *ST5*, cg22012981 in *ACOX2*, cg27015161 in *NCOA1*, and cg14633892 in *ARHGEF3* (Figure 2B and C; Figure S2 in Supporting Information).

To explore potential etiological relationships between GCA and ESCC, we examined tumor-shared markers. A comparison between 216 DMCs identified in our GCA discovery set and 1,188 DMCs identified from GEO datasets of ESCC revealed an overlap of four DMCs, including cg07880787 in *SORBS2*, cg00585116 in *PSCA*, cg27284428 and cg11798358 in *EPN3* (Figure S3 in Supporting Information).

Validation of candidate markers using matched multi-omics data

To validate the results from the discovery stage, we conducted pyrosequencing and quantitative real-time RT-PCR (qRT-PCR) on 50 paired GCA and NAT samples (Figure 3A–D; Figure S4A in Supporting Information). All selected DMCs (cg27284428, cg11798358, cg07880787, cg00585116, cg03234186) exhibited significantly higher methylation levels in both early-stage and late-stage GCA compared with NAT samples, with methylation levels increasing significantly with disease severity (P for trend <0.01). The qRT-PCR results revealed significantly lower expression levels of *EPN3* ($P<0.0001$), *SORBS2* ($P=0.0012$), and *PSCA* ($P<0.0001$) in GCA compared with NAT samples. A

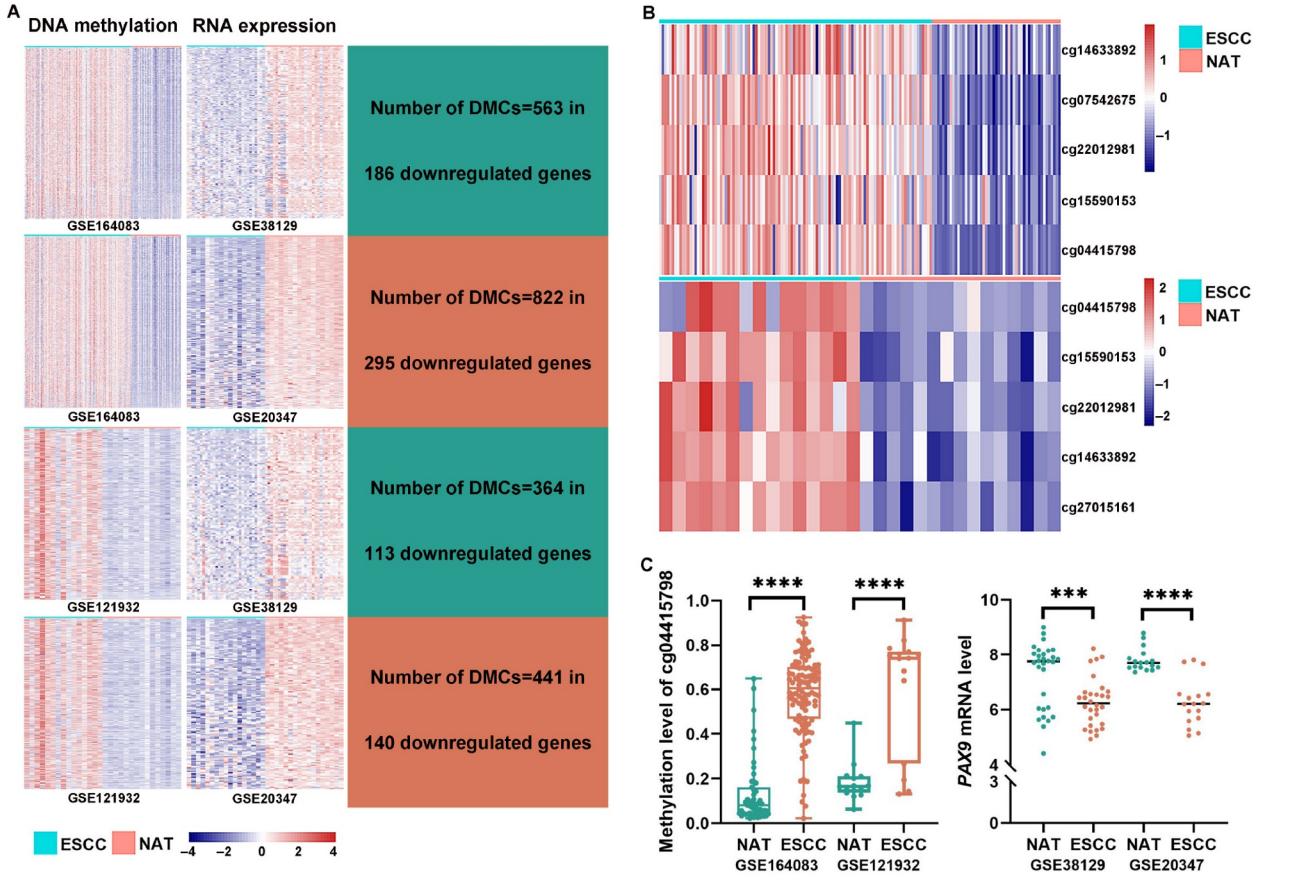

Figure 2. Identification of DNA methylation biomarkers to distinguish ESCC from normal tissue. A, Heatmap of identified 1,188 DMCs (left) and 369 target genes (right). There are four clusters: 563 hypermethylated CpGs identified in GSE164083 located in 186 downregulated genes identified in GSE38129; 822 hypermethylated CpGs identified in GSE164083 located in 295 downregulated genes identified in GSE20347; 364 hypermethylated CpGs identified in GSE121932 located in 113 downregulated genes identified in GSE38129; 441 hypermethylated CpGs identified in GSE121932 located in 140 downregulated genes identified in GSE20347. The row represents individual CpGs/genes, and the column represents individual samples. Blue, ESCC; red, NAT samples. The color in the heatmap represents the methylation levels of CpGs and expression levels of target genes. Red, hypermethylated CpGs/higher expressed genes; Blue, hypomethylated CpGs/lower expressed genes. B, Heatmap of six DMCs, with five DMCs covered by HM850K array (top) and five DMCs by HM450K array (bottom). The row represents individual CpGs, and the column represents individual samples. Blue, ESCC; red, NAT samples. The color in the heatmap represents the methylation levels of the CpGs. Red, hypermethylated CpGs; Blue, hypomethylated CpGs. C, Differential DNA methylation level of cg04415798 and expression level of the target gene *PAX9*. A methylation level of 0 represents no methylation, and 1 represents full methylation. ***, $P<0.001$; ****, $P<0.0001$

https://doi.org/10.1007/s11427-024-2642-8

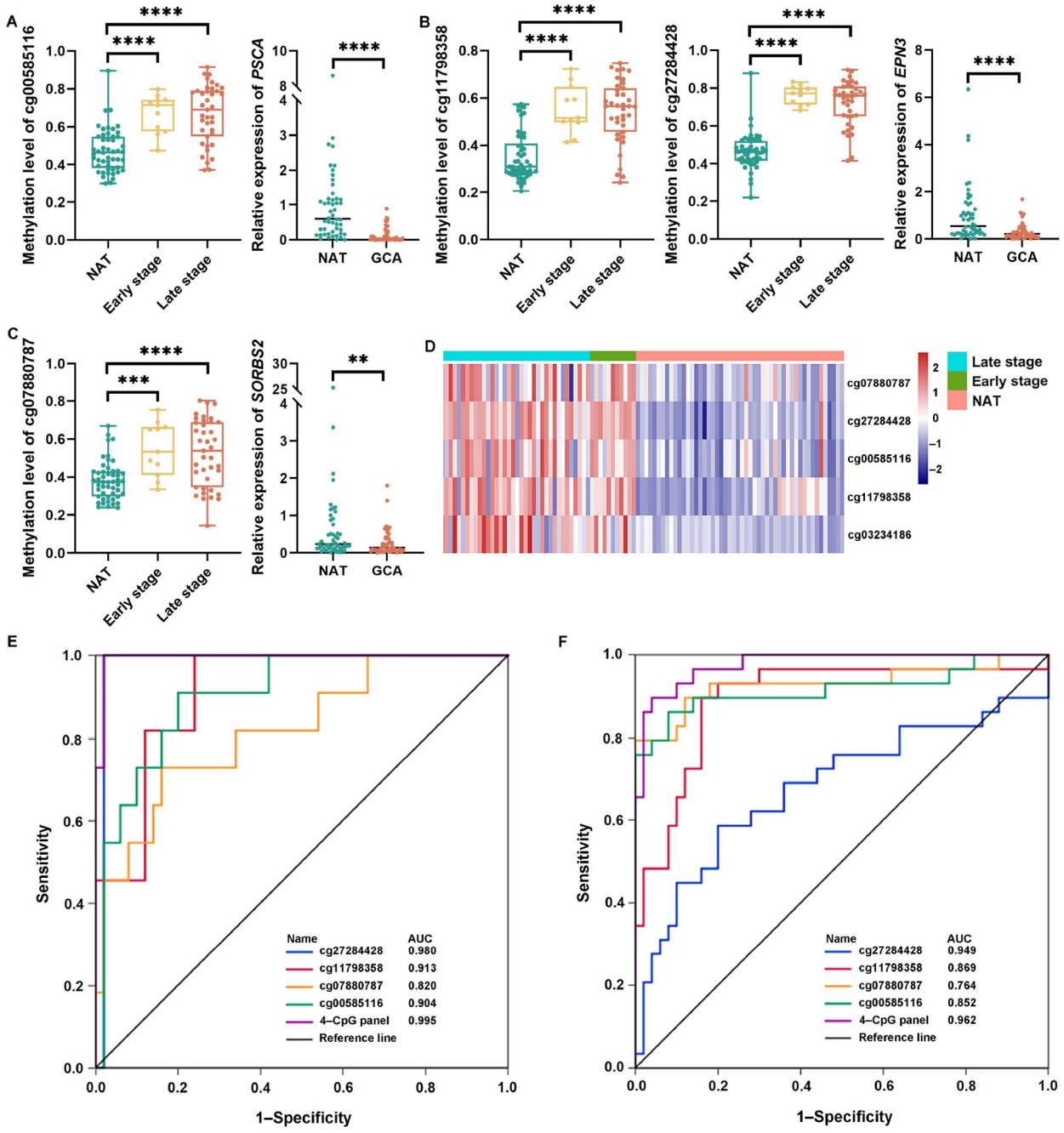

Figure 3. Validation of candidate markers in GCA patients. A–C, DNA methylation values of cg00585116 and relative mRNA expression of target gene *PSCA* (A), cg11798358, cg27284428 and target gene *EPN3* (B), cg07880787 and target gene *SORBS2* (C) in NAT and GCA with different clinical stages. A methylation level of 0 represents no methylation, and 1 represents full methylation. mRNA expression of genes normalized to housekeeping gene β-actin. D, Heatmap of DMCs in NAT and GCA with different clinical stages. The row represents individual DMCs, and the column represents individual samples. Blue, late-stage GCA; green, early-stage GCA; red, NAT samples. The color in the heatmap represents the methylation levels of the CpGs. Red, hypermethylated CpGs; blue, hypomethylated CpGs. E and F, ROC curves showing diagnostic performance of DMCs and marker panel for early-stage GCA (E) and all-stage GCA (F). **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$

similar trend was observed for *ZNF154*, although statistical significance was not reached ($P=0.2823$). Furthermore, we assessed the classification accuracy of cg27284428, cg11798358, cg07880787, and cg00585116, whose methylation and expression level changes were consistent with the discovery stage. For early-stage GCA detection, the area under receiver operating characteristic (ROC) curve (AUC) values for the four DMCs were all higher than 0.8, with cg27284428 exhibiting the highest AUC of 0.980 (Figure 3E). When applied to all-stage GCA diagnosis, the four DMCs also exhibited good discrimination power, with AUC values ranging from 0.764 to 0.949 (Figure 3F). Combining these four DMCs achieved an AUC of 0.995 (95% confidence interval (CI), 0.982–1.000) and 0.962 (95% CI, 0.920–1.000) for early-stage and all-stage GCA, respectively.

Next, we conducted pyrosequencing analysis and qRT-PCR on

50 paired ESCC and NAT samples (Figure 4A–D, Figure S4B–G in Supporting Information). Except for cg07880787, all selected DMCs exhibited significantly higher methylation levels in both early-stage and late-stage tumors compared with NAT samples (P for trend <0.01), generally in agreement with the discovery stage. The qRT-PCR results indicated significant downregulation of *PSCA* (P<0.0001), *PAX9* (P<0.0001), and *SORBS2* (P<0.0001) in ESCC tumor samples, with *ARHGEF3* showing marginal statistical significance (P=0.0594). No significantly lower expression was observed in ESCC tissues for the other five genes. Therefore, cg04415798 (*PAX9*), cg00585116 (*PSCA*), and cg14633892 (*ARHGEF3*) were validated with higher methylation and lower expression levels in ESCC tumor tissues, consistent with the discovery stage findings. ROC analysis revealed that all three validated DMCs could effectively distinguish ESCC from NAT (P<0.001). The AUC values for early-stage

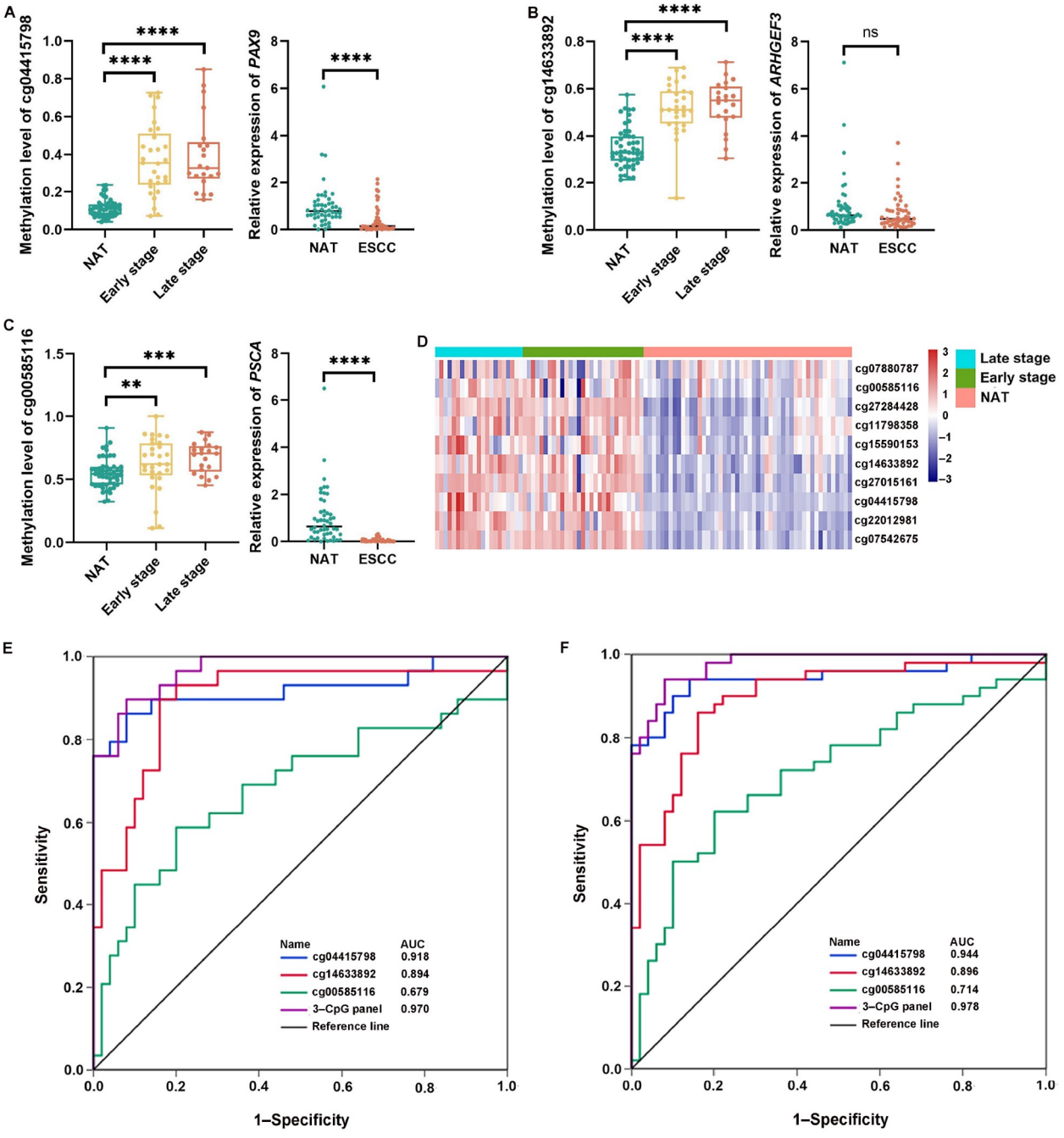

Figure 4. Validation of candidate markers in ESCC patients. A–C, DNA methylation values of cg04415798 and relative mRNA expression of target gene *PAX9* (A), cg14633892 and target gene *ARHGEF3* (B), cg00585116 and target gene *PSCA* (C) in NAT and ESCC with different clinical stages. A methylation level of 0 represents no methylation, and 1 represents full methylation. mRNA expression of genes normalized to housekeeping gene β-actin. D, Heatmap of DMCs in NAT and ESCC with different clinical stages. The row represents individual DMCs, and the column represents individual samples. Blue, late-stage ESCC; green, early-stage ESCC; red, NAT samples. The color in the heatmap represents the methylation levels of the CpGs. Red, hypermethylated CpGs; blue, hypomethylated CpGs. E and F, ROC curves showing diagnostic performance of DMCs and marker panel for early-stage ESCC (E) and all-stage ESCC (F). **, P<0.01; ***, P<0.001; ****, P<0.0001

https://doi.org/10.1007/s11427-024-2642-8

ESCC diagnosis ranged from 0.679 to 0.918, with cg04415798 achieving the highest AUC (Figure 4E). For all-stage ESCC detection, AUC values ranged from 0.714 to 0.944, with cg04415798 again demonstrating the highest AUC (Figure 4F). The 3-CpG panel showed superior diagnostic value with an AUC of 0.970 (95% CI, 0.939–1.000) and 0.978 (95% CI, 0.958–0.999) for detecting early-stage and all-stage ESCC, respectively.

Taken together, we validated three GCA-specific (cg27284428, cg11798358, and cg07880787), two ESCC-specific (cg04415798 and cg14633892), and one tumor-shared (cg00585116) DMCs, which were significantly hypermethylated with lower expression of their respective genes in tumor compared with NAT samples.

Diagnostic performance of DNAme markers for cancerous and precancerous lesions

To evaluate the performance of DNAme markers for detecting cancerous and precancerous lesions, we examined the methylation levels of our selected DMCs in an independent set of formalin-fixed, paraffin-embedded (FFPE) tissues across histological stages of cardia and esophagus disease. The methylation levels of candidate DMCs differed significantly between normal and lesion tissues in both cardia and esophagus disease ($P<0.05$; Figure 5A and B). Similarly, the tumor-shared marker cg00585116 exhibited significantly higher methylation levels in the combined cancer, HGD, and LGD groups compared with the normal control group ($P<0.05$; Figure 5C). Moreover, for consecutive intraepithelial neoplasia and cancer, methylation levels of all candidate DMCs increased significantly with histological severity.

We then assessed the classification accuracy of candidate markers when used individually and in combination with a marker panel. All four individual markers could significantly distinguish cardia HGD (CHGD) and GCA cases from cardia LGD (CLGD) and normal controls ($P<0.001$), with AUC values ranging from 0.729 to 0.887 (Figure S5A in Supporting Information). The 4-CpG panel outperformed any single marker, achieving an AUC of 0.917 (95% CI, 0.879–0.955; Figure 5D), with a sensitivity of 82.2% and specificity of 87.5% at the J-threshold ($\geqslant$0.68). At this threshold, the panel classified 90.6% (87/96) of GCA as methylation positive, along with 57.6% (19/33) of CHGD, 37.5% (9/24) of CLGD, and 1.8% (1/56) of normal control samples (Table S1 in Supporting Information). Similarly, the AUC of the three individual markers for detecting esophageal HGD (EHGD) and ESCC ranged from 0.724 to 0.826, with cg14633892 achieving the highest AUC (Figure S5B in Supporting Information). Using the J-threshold ($\geqslant$0.82), the 3-CpG panel provided a sensitivity of 74.4% and a specificity of 90.6%, with an AUC of 0.865 (95% CI, 0.812–0.917; Figure 5E). At this threshold, this panel classified 91.0% (91/100) of ESCC, 52.6% (40/76) of EHGD, 12.5% (5/40) of esophageal LGD (ELGD), and 0% (0/13) of normal samples as methylation positive (Table S2 in Supporting Information).

To develop a tumor-shared diagnostic model, we performed a combined analysis of the two disease entities. Univariable and multivariable logistic regression analysis identified age, alcohol drinking and tumor-shared marker cg00585116 as independent predictors for GCA, ESCC and HGD (Table S3 in Supporting Information). The variance inflation factor for each predictor was less than five, indicating no multicollinearity. A nomogram incorporating these three independent predictors was generated, yielding an AUC of 0.767 (95% CI, 0.720-0.813) for predicting GCA, ESCC and HGD (Figure 6A and B). Internal cross-validation showed a mean AUC was 0.764, slightly smaller than the full-data estimate. The calibration plot illustrated good agreement of the predictive nomogram with actual observations (Figure S6A in Supporting Information). Furthermore, the decision curve analysis (DCA) indicated that when the threshold probability was within a range from 0.02 to 0.99, the nomogram offered a net benefit over the "treat all" or "treat none" strategies (Figure S6B in Supporting Information).

DISCUSSION

In this study, we first identified functionally relevant DMCs between GCA/ESCC tumor and NAT through integrated analyses of genome-wide DNA methylome and transcriptome data. We further validated one tumor-shared and five tumor-specific markers in an independent series of GCA/ESCC and NAT samples. Using these DNAme markers, we constructed tumor-specific marker panels and a tumor-shared diagnostic model capable of discriminating cancer and HGD from normal control and LGD. Collectively, our findings underscore the potential value of DNAme markers in the early detection of GCA and ESCC.

DNAme markers were chosen in this study because they occur at predictable sites and can be quantitatively assayed. Previous studies have reported abnormal DNAme at gene promoters in GCA, such as *RASSF5A*, *GADD45A*, and *GADD45G*, implicating its role in carcinogenesis (Guo et al., 2013; Han et al., 2015). A recent pilot study analyzed genome-wide DNAme patterns in eight GCA patients using the HM850K array (Lin et al., 2023). Although the identified DMCs from that study were not selected in our analysis due to variations in filtering criteria and sample sources, they exhibited similar methylation patterns to our results, reinforcing the credibility of our results. In our study, we analyzed a larger sample size of 36 GCA patients using the HM850K array and identified five functionally relevant DMCs (cg27284428, cg11798358, cg07880787, cg00585116, cg03234186), which have not been previously reported. We validated that all selected CpG markers were differentially methylated in both early-stage and late-stage GCA compared with NAT samples, with methylation levels increasing significantly with disease severity. Additionally, 75% of their host genes exhibited opposite directionality of expression and methylation level changes, supporting the inverse relationship between DNAme and gene expression. *EPN3*, one of the host genes, encodes a family member of accessory proteins involved in clathrin-mediated endocytosis (Spradling et al., 2001). *EPN3* knockdown has been shown to cause resistance to DNA damage-induced apoptosis both in vitro and in vivo, and mRNA levels of *EPN3* were downregulated in GC tissues compared with normal tissues, consistent with our results (Mori et al., 2017). Another host gene, *PSCA*, encodes a glycosylphosphatidylinositol-anchored cell membrane glycoprotein, which exhibits cell-proliferation inhibition activity in vitro and is frequently silenced in GC tissues, further supporting our current findings (Study Group of Millennium Genome Project for Cancer et al., 2008; Xu et al., 2020). Although the role of *SORBS2* in GCA development remains unclear, *SORBS2* expression was significantly decreased in hepatocellular carcinoma and correlated with metastasis, advanced clinical stage, and poor prognosis (Han et al., 2019).

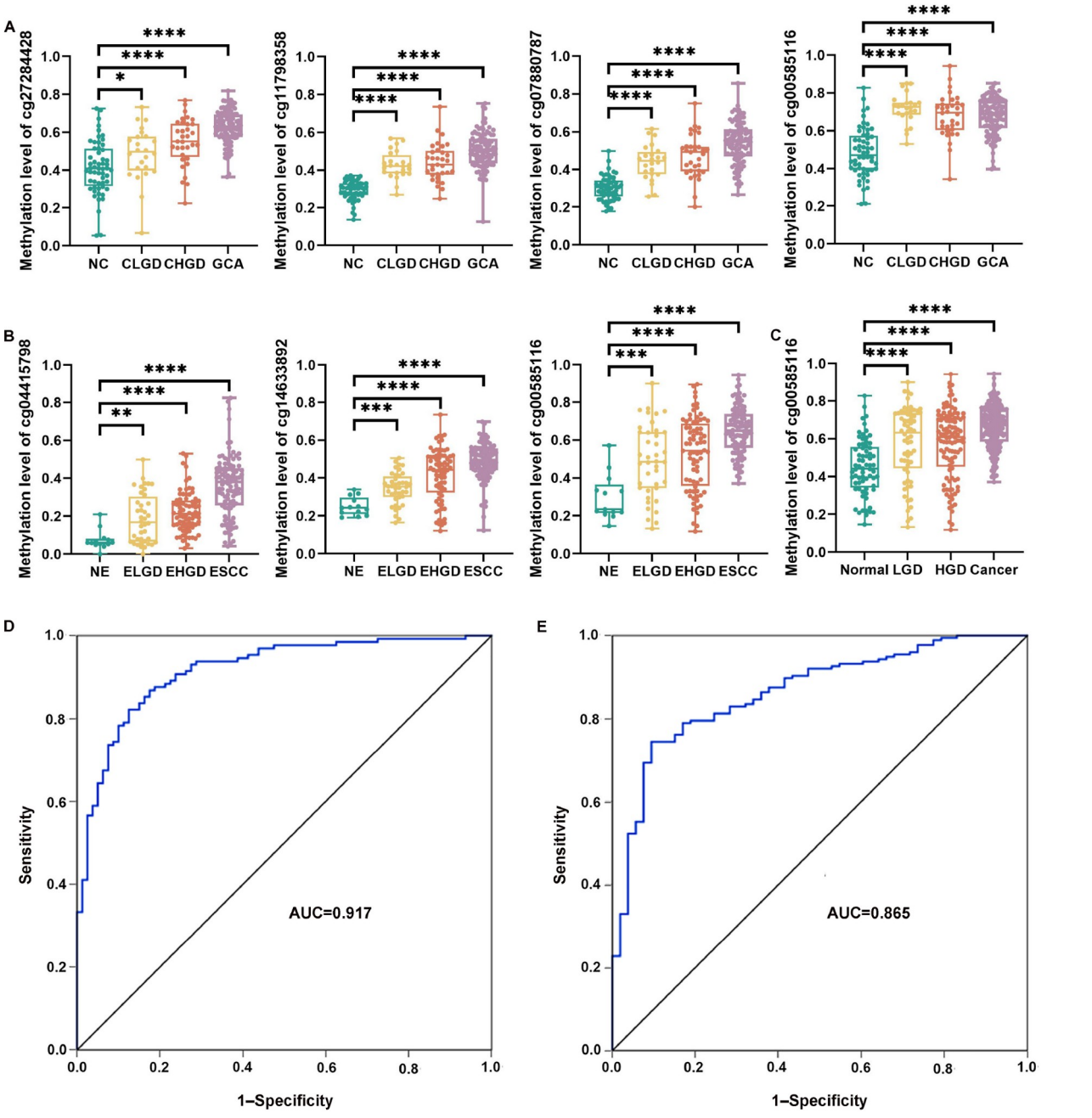

Figure 5. Early detection performance of DMCs. A, Methylation levels of cg27284428, cg11798358, cg07880787, and cg00585116 among NC, CLGD, CHGD and GCA. B, Methylation levels of cg04415798, cg14633892, and cg00585116 among NE, ELGD, EHGD and ESCC. C, Methylation levels of cg00585116 among normal control, LGD, HGD and cancer. A methylation level of 0 represents no methylation, and 1 represents full methylation. D, ROC curve for GCA-specific panel based on cg27284428, cg11798358, cg07880787 and cg00585116 for diagnosing CHGD and GCA. E, ROC curve for ESCC-specific panel based on cg04415798, cg14633892, and cg00585116 for diagnosing EHGD and ESCC. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$

Notably, our 4-CpG panel achieved AUCs of 0.995 and 0.962 for discriminating early-stage and all-stage GCA from NAT samples in an independent set, respectively, indicating that DMCs could serve as valuable early detection biomarkers for GCA.

Genome-wide methylation-based biomarkers for ESCC diagnosis have been investigated in several studies. A large methylome-wide study yielded a set of CpG markers with good discrimination for ESCC across multi-country, high-risk populations (Talukdar et al., 2021). However, Chinese patients were not included for analysis, despite China having the highest incidence of ESCC worldwide. In this study, we conducted a combined analysis of its released data with another methylation dataset from a Chinese population. Moreover, we validated three selected DMCs (cg04415798, cg14633892, and cg00585116) and their host genes in an independent set of Chinese ESCC patients. All three selected DMCs showed significant differences between ESCC and NAT samples in the two studies that provided the DNAme data used in the discovery stage, supporting the robustness of our

https://doi.org/10.1007/s11427-024-2642-8

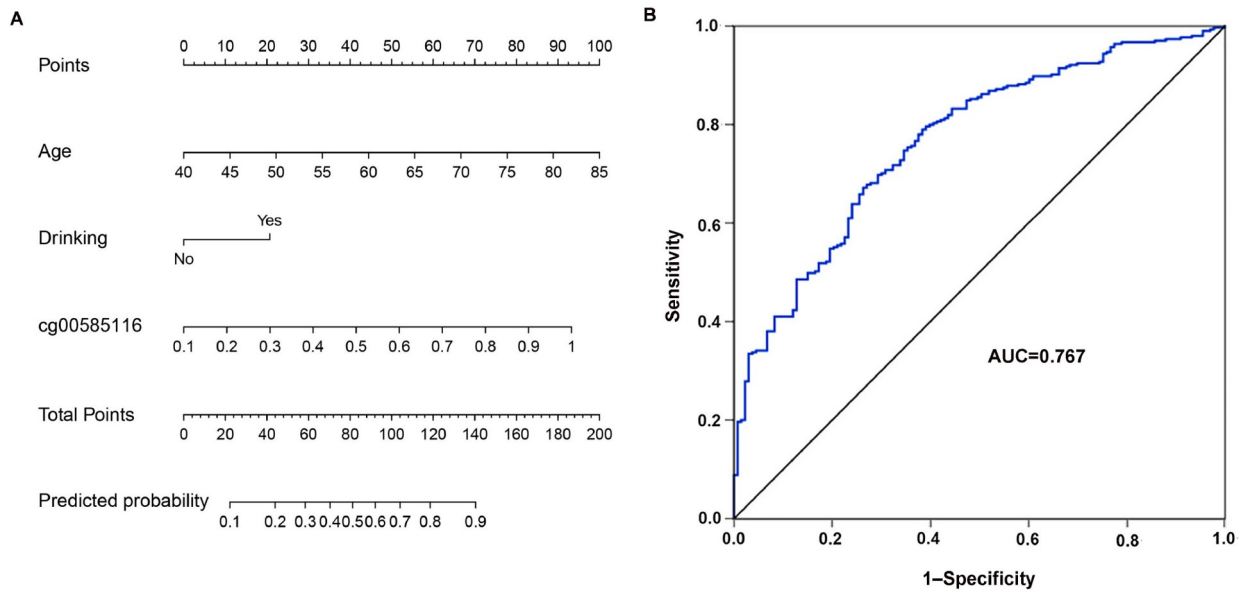

Figure 6. Tumor-shared logistic regression model. A, Nomogram of tumor-shared logistic regression model to predict probability of cancer (GCA/ESCC) and HGD. Find each predictor's point on the uppermost point scale and add them up. The total point projected to the bottom scale indicates the probability of cancer and HGD. B, ROC curves for tumor-shared logistic regression model

reported results (Chen et al., 2020; Talukdar et al., 2021). The methylation levels of the three DMCs demonstrated a gradational trend with advancing clinical stage, indicating their prognostic value in the clinical setting. Among the host genes, decreased expression of *PSCA* is a frequent event reported in as many as 86.2% of ESCC cases, and functional studies showed that *PSCA* could inhibit cell cycle progression and promote cell differentiation (Zhang et al., 2016). Our findings of hypermethylation and downregulation of *PSCA* in both ESCC and GCA suggest that *PSCA* may play a potential role in shared etiologic mechanisms, although the detailed roles and mechanisms require further exploration. The transcription factor *PAX9* has been reported to be involved in squamous cell differentiation and carcinogenesis of the esophageal epithelium, with promoter hypermethylation associated with *PAX9* silencing in ESCC (Bhol et al., 2021; Talukdar et al., 2021). In addition, *PAX9* deficiency was demonstrated to promote cell proliferation, delay cell differentiation, and alter gene expression in vivo (Xiong et al., 2018). This is in accordance with our observation, reinforcing the notion that promoter hypermethylation is crucial in the decreased *PAX9* expression contributing to ESCC pathogenesis. The role of *ARHGEF3* in ESCC has not been reported and warrants further investigations. Based on systematically selected DMCs, we developed a 3-CpG marker panel that effectively discriminates ESCC from NAT samples. A recently reported panel of 12 DMCs for ESCC diagnosis showed comparable discrimination with an AUC of 0.966, while our 3-CpG panel, with fewer markers, may be more suitable and applicable for clinical use (Xi et al., 2022).

In contrast to most previous studies that identified DNAme markers for GCA or ESCC diagnosis, our study focused on both cancerous and precancerous lesions (Lin et al., 2023; Qin et al., 2019). We found that methylation levels of all selected DNAme markers increased significantly with increasing severity of histological diagnosis, reinforcing the idea that DNA hypermethylation is a frequent and early event in both GCA and ESCC carcinogenesis (Fan et al., 2022; Liao et al., 2024). Importantly, we observed higher methylation levels in a subset of LGD and HGD, similar to the high methylation levels in cancer patients. One interesting possibility could be that precancerous lesions with high methylation levels might represent a subgroup at higher risk of progressing to cancer. Large-scale and longitudinal prospective studies are needed to test this hypothesis. Our tumor-specific DNAme marker panels demonstrated good discriminatory ability for cancer and HGD, indicating their potential application in the early detection of GCA and ESCC. These findings provide evidence for clinical decision making based on objective and reproducible biomarkers.

The identification of a tumor-shared DNAme marker and the development of a model for simultaneously diagnosing GCA, ESCC and HGD are notable for several reasons. Firstly, the high disease burden and co-occurrence of GCA and ESCC in many regions underscore the necessity for simultaneous screening. However, early detection relies on endoscopy, which is not feasible for population-level implementation, highlighting the importance of identifying shared biomarkers for risk stratification before endoscopy. Secondly, GCA and ESCC share common epidemiological features in high-risk geographic regions, such as alcohol consumption and smoking, which may be involved in cancer development by modulating DNAme (Chamberlain et al., 2022; Zhou et al., 2022). Thirdly, advancements in minimally invasive non-endoscopic sampling devices allow for cell collection from the esophagus and cardia (Gao et al., 2023). Recent progress has indicated that combining non-endoscopic sampling devices with DNAme markers enables an efficient screening method for Barrett's esophagus (Chettouh et al., 2018; Wang et al., 2019). Our current findings suggest potential use of non-endoscopic sampling devices with DNAme markers for simultaneous screening of GCA and ESCC. The development of minimally invasive and cost-effective screening procedures remains a key focus of ongoing research.

Several limitations should be considered in this study. First, integrating DNAme and gene expression data from different

sources in the discovery stage may introduce heterogeneity in the pooled patient samples. Although we validated the candidate markers in an independent set using matched methylation and expression data from the same patients, some relevant biomarkers might have been missed. Second, GCA and ESCC patients in the application stage were individuals recruited from outpatient clinics, with most diagnosed due to presenting symptoms. This might cause an overestimation of the performance of our model in the real world, where asymptomatic and screened individuals are more likely to be early-stage patients. Third, this study was conducted at a single academic center with a relatively modest sample size, limiting the generalizability of our results. Finally, while our findings indicated that DNAme was associated with a significantly increased risk of GCA and ESCC, the case-control nature of this study impairs the ability to determine the direction of causation. Future prospective multicenter studies are needed to validate the hypothesis that aberrant DNAme predisposes individuals to an increased cancer risk.

Our study identified DMCs with significant potential for early detection of ESCC and GCA, paving the way for the development of a diagnostic kit for routine clinical practice. The kit could potentially improve current screening practices by enabling non-invasive or minimally invasive methods, facilitating early detection and improving patient outcomes. Future research should focus on validating these biomarkers in larger and more diverse populations to ensure their effectiveness and reliability across different demographics. Additionally, investigating the biological mechanisms underlying the observed differential methylation patterns could provide deeper insights into the pathogenesis of ESCC and GCA, potentially identifying new therapeutic targets.

In conclusion, we identified and validated novel DNAme markers that hold promise for efficient screening of GCA and ESCC. We hypothesize that combining these markers with minimally invasive samples from sponge cytology testing or liquid biopsy could help target high-risk populations and optimize healthcare resources management.

MATERIALS AND METHODS

Participants and sample preparation

Our study design comprised three sequential cross-sectional stages (Figure S1 in Supporting Information). This study included a discovery set of 36 GCA patients in the discovery stage, followed by an independent validation set of 50 GCA and 50 ESCC patients in the validation stage. Patients who underwent surgical therapy and had not received chemotherapy or radiotherapy prior to surgical removal of tumor tissue were recruited from two cancer hospitals in Linzhou between 2021 and 2022. Fresh-frozen tumor and NAT samples (≥4 cm from the tumor margin) were collected from each patient immediately after surgery, with pathological diagnoses independently reviewed by two pathologists.

In the application stage, an independent set of FFPE slides was analyzed, comprising 69 normal tissues, 64 LGD, 109 HGD, and 196 cancer (96 GCA and 100 ESCC) samples. Among these, 242 subjects with normal epithelia, LGD and HGD were recruited from Linzhou between 2017 and 2020, all of whom underwent endoscopic examination as we previously described (Chen et al., 2019). The 196 cancer patients were recruited from outpatient clinics between 2021 and 2022. Samples were collected from endoscopic biopsy or surgically excised tissues and processed into FFPE tissues. Hematoxylin and eosin-stained slides were prepared for each sample and examined by two expert pathologists to ensure the histological diagnosis.

This study was approved by the Institutional Review Board of the Cancer Hospital of the Chinese Academy of Medical Sciences (Approval No. 16-171/1250). Informed consent was obtained from all participants before specimen collection.

Genome-wide DNAme analysis

Genomic DNA was extracted from fresh-frozen tissues of GCA patients using the E.Z.N.A. Tissue DNA Kit (OMEGA Bio-Tek, USA), following the manufacturer-provided protocol. The isolated DNA was then subjected to quality control by agarose gel electrophoresis and quantified by NanoDrop. Bisulfite conversion was conducted on 500 ng genomic DNA with the EZ DNA Methylation-Gold kit (Zymo Research, USA), following the manufacturer's instructions. All samples passed quality control tests. Genome-wide DNAme profiling was conducted using the Infinium Methylation EPIC BeadChip (HM850K array, Illumina, USA), which measures methylation status at approximately 850,000 CpG sites in the human genome of GCA patients. For ESCC, raw DNAme data from GSE164083 and GSE121932 were obtained from the GEO database.

The raw Illumina data were processed using the Bioconductor R package minfi. Probes with missing values or detection P values≥0.05 in more than 5% of samples were excluded from further analysis. The methylation data were then normalized using the subset-quantile within array normalization (SWAN) method. Methylation levels at each CpG site were represented as β values, calculated as (methylated allele intensity (M)/unmethylated allele intensity (U)+methylated allele intensity (M)+100). DMCs were selected using an algorithm in IMA Bioconductor. A CpG site was considered differentially methylated if the absolute difference in mean β values ($|\Delta\beta|$, tumor *vs.* NAT) was >0.15, with adjusted P<0.05. Only DMCs located in the promoter regions (transcription start site (TSS) 200, TSS1500, and 5′ UTR) were selected for further analysis.

Gene expression analysis

Differentially expressed genes between paired tumor and NAT samples were identified using raw gene expression microarray data from three datasets (GSE29272, GSE38129, and GSE20347) downloaded from the GEO database. Normalization and background correction of the raw data were performed using the Robust Multichip Average algorithm. Missing values were imputed using the *k*-nearest neighbor method implemented in the R package impute. Differential expression analysis was conducted using empirical Bayes methods in the limma statistics package. Genes that were downregulated with statistical significance were further analyzed, with selection criteria including a $\log_2$ fold change <−0.585 (fold change: 1.5) and a false discovery rate adjusted P<0.05.

qRT-PCR

Total RNA was isolated from fresh-frozen tissue samples using the MJzol Animal RNA Extraction Kit (MagBeads; Majorbio,

https://doi.org/10.1007/s11427-024-2642-8

Shanghai, China), and RNA concentrations were determined using the NanoDrop spectrophotometer. Reverse transcription was performed using the HiFiScript cDNA Synthesis Kit (CWBiotech, Beijing, China) according to the manufacturer's protocol. qRT-PCR was conducted in triplicate in 96-well plates using TB Green Premix Ex Taq (TaKaRa, Japan). The housekeeping gene β-actin served as the internal control, and relative gene expression was determined using the $2^{-(\Delta\Delta C_t)}$ method. The primer sequences of the target genes are provided in Table S4 in Supporting Information.

Bisulfite pyrosequencing

Genomic DNA from fresh-frozen tissues and FFPE tissues was isolated using the MagaBio plus General Genomic DNA Purification Kit (BIOER, Hangzhou, China) and the HiPure FFPE DNA Kit (Magen, Guangzhou, China), respectively, following the manufacturer's instructions. The isolated genomic DNA was treated with sodium bisulfite using the EZ DNA Methylation-Gold kit (Zymo Research, USA). The bisulfite-treated DNA served as a template for amplifying the target regions of interest. The PCR products (10 μL) were sequenced by pyrosequencing on the PyroMark Q48 (QIAGEN, Germany) following the manufacturer's protocol. The primer sequences for pyrosequencing are listed in Table S5 in Supporting Information.

Statistical analysis

Comparisons of methylation and gene expression levels between paired tumor and NAT were performed using paired *t*-test or Wilcoxon signed-rank test. Differences in methylation levels among different histological categories were evaluated by unpaired *t* test or Mann-Whitney *U* test. The Cuzick test for trend was employed to evaluate the relationship between DNAme levels and disease progression. Categorical variables were compared using the Chi-square or Fisher's exact test as appropriate. LASSO algorithm and RF analysis were performed using the glmnet and randomForest packages, respectively. The LASSO regression identified DMCs with non-zero coefficients as potential biomarkers, and the RF method selected the top 10 DMCs by ranking their importance based on the mean decrease in Gini values.

For analysis of individual disease entities, logistic regression was used to evaluate the diagnostic performance of individual biomarkers and their combined discriminatory ability as a marker panel. For combined analysis of both disease entities, predictor variables with $P<0.10$ in univariable analysis were selected into multivariable logistic regression model using backwards stepwise selection. A nomogram was constructed based on the factors included in the model, and the AUC (equal to concordance statistic) and 95% CI were calculated to assess the performance. A calibration curve was created by bootstraps with 1,000 resamples to evaluate the agreement between model predictions and observations. A DCA was performed to assess the clinical usefulness of the nomogram. A 5-fold cross-validation was used to internally validate the stability of the model. These analyses focused on distinguishing between "normal/LGD" and "HGD/cancer" groups, as patients with HGD or cancer are recommended to receive endoscopic or surgical therapy (Wang and Wei, 2020).

All statistical analyses were conducted using R software version 4.2.0, GraphPad Prism 9.5, and Statistical Package for Social Science (SPSS) version 26.0. All reported *P* values were two-sided, with a significance threshold set at 0.05.

Data availability

The methylation data generated in this study are deposited in the OMIX, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (http://bigd.big.ac.cn/omix, accession number OMIX004975). Other data supporting the findings of the present study are available from the corresponding author upon reasonable request.

Compliance and ethics

The author(s) declare that they have no conflict of interest.

Acknowledgement

This work was supported by CAMS Innovation Fund for Medical Sciences (2021-I2M-1–010), National Key Research and Development Program of China (2016YFC0901404), and the National Natural Science Foundation of China (81974493). We thank the Gene Expression Omnibus database for providing public platform and researchers for their valuable data sets. WQ. W. and ZY.F. conceived and designed the study. ZY.F., FF.H., MJ.L., XQ.L., JW.W., R.C., and H.J. collected data and performed quality control. ZY.F. and JJ.H. performed the experiments and analyzed the data. ZY.F. completed the initial drafting of the manuscript. JJ.H. and WQ.W. critically revised the manuscript. All authors read and approved the final manuscript.

Supporting information

The supporting information is available online at https://doi.org/10.1007/s11427-024-2642-8. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

Arnold, M., Ferlay, J., van Berge Henegouwen, M.I., and Soerjomataram, I. (2020). Global burden of oesophageal and gastric cancer by histology and subsite in 2018. Gut 69, 1564–1571.

Bhol, C.S., Patil, S., Sahu, B.B., Patra, S.K., and Bhutia, S.K. (2021). The clinical significance and correlative signaling pathways of paired box gene 9 in development and carcinogenesis. Biochim Biophys Acta Rev Cancer 1876, 188561.

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., and Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74, 229–263.

Cao, W., Lee, H., Wu, W., Zaman, A., McCorkle, S., Yan, M., Chen, J., Xing, Q., Sinnott-Armstrong, N., Xu, H., et al. (2020). Multi-faceted epigenetic dysregulation of gene expression promotes esophageal squamous cell carcinoma. Nat Commun 11, 3675.

Chamberlain, J.D., Nusslé, S., Chapatte, L., Kinnaer, C., Petrovic, D., Pradervand, S., Bochud, M., Harris, S.E., Corley, J., Cox, S.R., et al. (2022). Blood DNA methylation signatures of lifestyle exposures: tobacco and alcohol consumption. Clin Epigenet 14, 155.

Chen, R., Liu, Y., Song, G., Li, B., Zhao, D., Hua, Z., Wang, X., Li, J., Hao, C., Zhang, L., et al. (2021). Effectiveness of one-time endoscopic screening programme in prevention of upper gastrointestinal cancer in China: a multicentre population-based cohort study. Gut, 70: 251–260.

Chen, R., Ma, S., Guan, C., Song, G., Ma, Q., Xie, S., Wang, M., Shao, D., Li, X., and Wei, W. (2019). The national cohort of esophageal cancer-prospective cohort study of Esophageal Cancer and Precancerous Lesions based on High-Risk Population in China (NCEC-HRP): study protocol. BMJ Open 9, e027360.

Chen, Y., Liao, L.D., Wu, Z.Y., Yang, Q., Guo, J.C., He, J.Z., Wang, S.H., Xu, X.E., Wu, J.Y., Pan, F., et al. (2020). Identification of key genes by integrating DNA methylation and next-generation transcriptome sequencing for esophageal squamous cell carcinoma. Aging 12, 1332–1365.

Chettouh, H., Mowforth, O., Galeano-Dalmau, N., Bezawada, N., Ross-Innes, C., MacRae, S., Debiram-Beecham, I., O'Donovan, M., and Fitzgerald, R.C. (2018). Methylation panel is a diagnostic biomarker for Barrett's oesophagus in endoscopic biopsies and non-endoscopic cytology specimens. Gut 67, 1942–1949.

Fan, Z., Qin, Y., Zhou, J., Chen, R., Gu, J., Li, M., Zhou, J., Li, X., Lin, D., Wang, J., et al. (2022). Feasibility of using *P16* methylation as a cytologic marker for esophageal squamous cell carcinoma screening: a pilot study. Cancer Med 11, 4033–4042.

Gao, Y., Xin, L., Lin, H., Yao, B., Zhang, T., Zhou, A.J., Huang, S., Wang, J.H., Feng, Y. D., Yao, S.H., et al. (2023). Machine learning-based automated sponge cytology for

screening of oesophageal squamous cell carcinoma and adenocarcinoma of the oesophagogastric junction: a nationwide, multicohort, prospective study. Lancet Gastroenterol Hepatol 8, 432–445.
Guo, W., Dong, Z., Guo, Y., Chen, Z., Kuang, G., and Yang, Z. (2013). Methylation-mediated repression of GADD45A and GADD45G expression in gastric cardia adenocarcinoma. Int J Cancer 133, 2043–2053.
Guo, Y., Zhu, T., Guo, W., Dong, Z., Zhou, Z., Cui, Y., and Zhao, R. (2016). Aberrant CpG island shore region methylation of CAV1 is associated with tumor progression and poor prognosis in gastric cardia adenocarcinoma. Arch Med Res 47, 460–470.
Han, L., Dong, Z., Wang, C., Guo, Y., Shen, S., Kuang, G., and Guo, W. (2015). Decreased expression and aberrant methylation of RASSF5A correlates with malignant progression of gastric cardia adenocarcinoma. Mol Carcinog 54, 1722–1733.
Han, L., Huang, C., and Zhang, S. (2019). The RNA-binding protein SORBS2 suppresses hepatocellular carcinoma tumourigenesis and metastasis by stabilizing RORA mRNA. Liver Int 39, 2190–2203.
He, S., Xia, C., Li, H., Cao, M., Yang, F., Yan, X., Zhang, S., Teng, Y., Li, Q., and Chen, W. (2024). Cancer profiles in China and comparisons with the USA: a comprehensive analysis in the incidence, mortality, survival, staging, and attribution to risk factors. Sci China Life Sci 67, 122–131.
Li, M., Park, J.Y., Sheikh, M., Kayamba, V., Rumgay, H., Jenab, M., Narh, C.T., Abedi-Ardekani, B., Morgan, E., de Martel, C., et al. (2023). Population-based investigation of common and deviating patterns of gastric cancer and oesophageal cancer incidence across populations and time. Gut 72, 846–854.
Li, T., Chen, X., Gu, M., Deng, A., and Qian, C. (2020). Identification of the subtypes of gastric cancer based on DNA methylation and the prediction of prognosis. Clin Epigenet 12, 161.
Li, Y., Liang, J., and Hou, P. (2015). Hypermethylation in gastric cancer. Clin Chim Acta 448, 124–132.
Liao, X., Lin, R., Zhang, Z., Tian, D., Liu, Z., Chen, S., Xu, G., and Su, M. (2024). Genome-wide DNA methylation and transcriptomic patterns of precancerous gastric cardia lesions. J Natl Cancer Institute 116, 681–693.
Lin, R., Qian, Y., Zhang, J., Xia, D., Guo, D., Hong, L., Qing, B., Xu, M., Huang, Y., Lin, W., et al. (2023). Genome-wide DNA methylation profiling of gastric cardia cancer. J Gastro Hepatol 38, 290–300.
Maharjan, U., and Kauppila, J.H. (2022). Survival trends in gastric cancer patients between 1987 and 2016: a population-based cohort study in Finland. Gastric Cancer 25, 989–1001.
Mori, J., Tanikawa, C., Ohnishi, N., Funauchi, Y., Toyoshima, O., Ueda, K., and Matsuda, K. (2017). EPSIN 3, a novel p53 target, regulates the apoptotic pathway and gastric carcinogenesis. Neoplasia 19, 185–195.
Parfitt, J.R., Miladinovic, Z., and Driman, D.K. (2006). Increasing incidence of adenocarcinoma of the gastroesophageal junction and distal stomach in Canada—an epidemiological study from 1964 to 2002. Can J Gastroenterol 20, 271–276.
Qin, Y., Wu, C.W., Taylor, W.R., Sawas, T., Burger, K.N., Mahoney, D.W., Sun, Z., Yab, T.C., Lidgard, G.P., Allawi, H.T., et al. (2019). Discovery, validation, and application of novel methylated DNA markers for detection of esophageal cancer in plasma. Clin Cancer Res 25, 7396–7404.
Rice, T.W., Ishwaran, H., Ferguson, M.K., Blackstone, E.H., and Goldstraw, P. (2017). Cancer of the esophagus and esophagogastric junction: an eighth edition staging primer. J Thorac Oncol 12, 36–42.
Spradling, K.D., McDaniel, A.E., Lohi, J., and Pilcher, B.K. (2001). Epsin 3 is a novel extracellular matrix-induced transcript specific to wounded epithelia. J Biol Chem 276, 29257–29267.
Study Group of Millennium Genome Project for Cancer, Sakamoto, H., Yoshimura, K., Saeki, N., Katai, H., Shimoda, T., Matsuno, Y., Saito, D., Sugimura, H., Tanioka, F., et al. Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. Nat Genet, 2008, 40: 730–740.
Talukdar, F.R., Soares Lima, S.C., Khoueiry, R., Laskar, R.S., Cuenin, C., Sorroche, B. P., Boisson, A.C., Abedi-Ardekani, B., Carreira, C., Menya, D., et al. (2021). Genome-wide DNA methylation profiling of esophageal squamous cell carcinoma from global high-incidence regions identifies crucial genes and potential cancer markers. Cancer Res 81, 2612–2624.
Thrift, A.P., and El-Serag, H.B. (2020). Burden of gastric cancer. Clin Gastroenterol Hepatol 18, 534–542.
Usui, G., Matsusaka, K., Huang, K.K., Zhu, F., Shinozaki, T., Fukuyo, M., Rahmutulla, B., Yogi, N., Okada, T., Minami, M., et al. (2023). Integrated environmental, lifestyle, and epigenetic risk prediction of primary gastric neoplasia using the longitudinally monitored cohorts. EBioMedicine 98, 104844.
Wang, G.Q., and Wei, W.Q. (2020). Technology Scheme for Upper Gastrointestinal Cancer Early Detection and Treatment (in Chinese). Beijing: People's Medical Publishing House.
Wang, Y., Yan, Q., Fan, C., Mo, Y., Wang, Y., Li, X., Liao, Q., Guo, C., Li, G., Zeng, Z., et al. (2023). Overview and countermeasures of cancer burden in China. Sci China Life Sci 66, 2515–2526.
Wang, Z., Kambhampati, S., Cheng, Y., Ma, K., Simsek, C., Tieu, A.H., Abraham, J.M., Liu, X., Prasath, V., Duncan, M., et al. (2019). Methylation biomarker panel performance in esophacap cytology samples for diagnosing barrett's esophagus: a prospective validation study. Clin Cancer Res 25, 2127–2135.
Xi, Y., Lin, Y., Guo, W., Wang, X., Zhao, H., Miao, C., Liu, W., Liu, Y., Liu, T., Luo, Y., et al. (2022). Multi-omic characterization of genome-wide abnormal DNA methylation reveals diagnostic and prognostic markers for esophageal squamous-cell carcinoma. Sig Transduct Target Ther 7, 53.
Xiong, Z., Ren, S., Chen, H., Liu, Y., Huang, C., Zhang, Y.L., Odera, J.O., Chen, T., Kist, R., Peters, H., et al. (2018). PAX9 regulates squamous cell differentiation and carcinogenesis in the oro-oesophageal epithelium. J Pathol 244, 164–175.
Xu, L., Qiu, H., Yuan, S., Chen, Y., Zhou, Z., and Chen, Y. (2020). Downregulation of PSCA promotes gastric cancer proliferation and is related to poor prognosis. J Cancer 11, 2708–2715.
Zeng, H., Chen, W., Zheng, R., Zhang, S., Ji, J.S., Zou, X., Xia, C., Sun, K., Yang, Z., Li, H., et al. (2018). Changing cancer survival in China during 2003–15: a pooled analysis of 17 population-based cancer registries. Lancet Glob Health 6, e555–e567.
Zhang, L.Y., Wu, J.L., Qiu, H.B., Dong, S.S., Zhu, Y.H., Lee, V.H.F., Qin, Y.R., Li, Y., Chen, J., Liu, H.B., et al. (2016). PSCA acts as a tumor suppressor by facilitating the nuclear translocation of RB1CC1 in esophageal squamous cell carcinoma. Carcinogenesis 37, 320–332.
Zhou, X., Wang, L., Xiao, J., Sun, J., Yu, L., Zhang, H., Meng, X., Yuan, S., Timofeeva, M., Law, P.J., et al. (2022). Alcohol consumption, DNA methylation and colorectal cancer risk: results from pooled cohort studies and Mendelian randomization analysis. Int J Cancer 151, 83–94.