Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS Check for updates

Tongue coating microbiota-based machine learning for diagnosing digestive system tumours

Yubo Ma[a]*, Zhengchen Jiang ⬤[b,c]*, Yanan Wang[b], Libin Pan[d], Kang Liu[a], Ruihong Xia[a], Li Yuan ⬤[c,e,f] and Xiangdong Cheng ⬤[b,c,e]

[a]The Second Clinical Medical College, Zhejiang Chinese Medical University, Hangzhou, China; [b]Department of Gastric Surgery, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China; [c]Zhejiang Key Lab of Prevention, Diagnosis and Therapy of Upper Gastrointestinal Cancer, Zhejiang Cancer Hospital, Hangzhou, China; [d]Department of Pharmacy, Zhejiang Cancer Hospital, Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China; [e]Zhejiang Provincial Research Center for Upper Gastrointestinal Tract Cancer, Zhejiang Cancer Hospital, Hangzhou, China; [f]Department of Integrated Chinese and Western Medicine, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China

ABSTRACT
Background: Digestive system tumours (DSTs) often diagnosed late due to nonspecific symptoms. Non-invasive biomarkers are crucial for early detection and improved outcomes.
Patients and Methods: We collected tongue coating samples from 710 patients diagnosed with DST and 489 healthy controls (HC) from April 2023, to December 2023. Microbial composition was analyzed using 16S rRNA sequencing, and five machine learning algorithms were applied to assess the diagnostic potential of tongue coating microbiota.
Results: Alpha diversity analysis showed that the microbial diversity in the tongue coating was significantly increased in DST patients. LEfSe analysis identified DST-enriched genera Alloprevotella and Prevotella, contrasting with HC-dominant taxa Neisseria, Haemophilus, and Porphyromonas (LDA >4). Notably, when comparing each of the four DST subtypes with the HC group, the proportion of Haemophilus in the HC group was significantly higher, and it was identified as an important feature for distinguishing the HC group. Machine learning validation demonstrated superior diagnostic performance of the Extreme Gradient Boosting (XGBoost) model, achieving an AUC of 0.926 (95% CI: 0.893–0.958) in internal validation, outperforming the other four machine learning models.
Conclusion: Tongue coating microbiota shows promise as a non-invasive biomarker for DST diagnosis, supported by robust machine learning models.

Introduction

Cancer was one of the leading causes of death among humans. In 2022, it was estimated that there were approximately 20 million new cancer cases and 9.7 million cancer-related deaths worldwide [1]. Among these, over one-quarter of malignant tumours were digestive system tumours (DSTs), including esophageal cancer (EC), gastric cancer (GC), pancreatic cancer, liver cancer, gallbladder cancer and bile duct cancer, and colorectal cancer (CRC) [2]. Notably, the incidence rates of CRC and GC ranked third and fifth among all cancers, respectively [1]. Most patients with DST usually presented with no symptoms or only non-specific symptoms in the early stages, often resulting in diagnosis at advanced stages or even distant metastasis, leading to poor treatment outcomes [3–6]. In clinical practice, the diagnosis of DSTs (such

as CRC and GC) primarily relied on endoscopy and biopsy [7,8], which were invasive procedures. The diagnosis of liver cancer primarily depended on imaging and histological results, but these methods also caused certain harm to patients [9]. Therefore, it was crucial to identify biomarkers associated with the occurrence and development of DST for early screening and diagnosis.

Currently, an increasing number of studies have confirmed that changes in the oral microbiome are not only direct precursors to oral diseases such as periodontitis and oral cancer [10,11], but oral microbial dysbiosis can also lead to systemic diseases, particularly digestive system diseases [12,13]. The tongue coating is an integral part of the oral environment, providing a unique habitat for diverse microbial communities. As the tongue is anatomically connected to the upper digestive tract, its microbiota may serve as a valuable indicator of digestive

system health [14–16]. Many studies have reported the association between oral microbiota and the occurrence and prognosis of digestive system tumours. Yang et al. [17] found an enrichment of *Neisseria mucosa* and *Prevotella pleuritidis* in the oral microbiota of GC patients, while *Mycoplasma orale* and *Eubacterium yurii* were reduced. Periodontal pathogens *Tannerella forsythia* and *P. gingivalis* were associated with a higher risk of EC [18]. Additionally, *P. gingivalis* and *A. actinomycetemcomitans* in the oral cavity were related to a higher risk of pancreatic cancer [19,20]. *F. nucleatum* was associated with poor prognosis in pancreatic and gastric cancers [21,22]. Flemer et al. [23] demonstrated that the differences in the abundance of oral microbiota could distinguish colorectal cancer from healthy controls. Therefore, we considered that there might be an inseparable connection between DSTs and the oral microbiome [12,24], but no studies had directly evaluated the relationship between the oral microbiome and DST.

We conducted a prospective study at Zhejiang Cancer Hospital to determine whether the oral microbiome was associated with the risk of DST. We directly profiled the tongue coating microbiota by 16S rRNA gene sequencing of tongue coating samples collected at diagnosis from 710 DST patients, including 107 ECs, 122 hepatopancreatobiliary (HPB) cancers, 419 GCs, and 62 CRCs, alongside 489 healthy controls (HC) in the cohort. Furthermore, we used various model algorithms, including Generalized Linear Model (GLM), RandomForest (RF), Decision Tree (DT), Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), to distinguish each DST from HC. Our findings would provide an in-depth understanding of the relationship between tongue coating microbiota and DST and develop new preventive or non-invasive diagnostic methods for patients with DST.

Materials and methods

Tongue coating sample collection

In this prospective study, patients aged between 18 and 80 years, histologically confirmed to have digestive system tumours, were deemed eligible. Any patients who had previously received cancer treatment (including chemotherapy, biotherapy, immunotherapy, radiotherapy, or surgery) or had oral diseases were excluded. Oral disease status was assessed by inquiring whether participants had been diagnosed with periodontal disease or undergone treatment for oral conditions (e.g. oral ulcers, periodontitis, or oral candidiasis) in the past 3 months. Additionally, patients with two or more concurrent malignant tumours were also excluded. Healthy control samples were donated by healthy volunteers with no history of malignant tumours or oral diseases. To minimize potential confounding factors, healthy controls were selected using a propensity score matching method. Propensity scores were calculated based on age, sex, smoking history, and alcohol consumption history, all treated as binary variables. Matching was performed at a 1:1 ratio without replacement, using a caliper width of 0.05 of the standard deviation of the logit of the propensity score, resulting in the selection of 489 healthy controls. All participants were recruited from Zhejiang Cancer Hospital between 1 April 2023, and 31 December 2023, and those who had received antibiotics or probiotic treatment within 4 weeks were also excluded (ClinicalTrial.gov: NCT05794841).

Tongue coating samples from cancer patients were collected on the morning of their surgery. Noncancerous volunteers also had their samples collected in the morning. Before consuming any food or water, participants used disposable swabs to collect tongue coating samples. Prior to sample collection, participants rinsed their mouths three times with sterile water. Professional operators used throat swabs to collect samples from the posterior central area to the anterior central area of the tongue dorsum (each swab scraped three times, totalling two swabs). The swabs were then immediately placed into a storage tube and transferred to a freezer at −80°C.

16S rRNA gene sequencing

Microbial DNA was extracted using an E.Z.N.A. Tissue DNA Kit (D3396–01; Omega, Norcross, Georgia, USA) following the manufacturer's instructions, with additional lysis steps to optimize microbial DNA extraction. Specifically, 80-mesh carborundum, 1 mm grinding beads, and 5 mm steel beads were added to the samples, followed by bead-beating for 3 minutes. Lysozyme was also added during the lysis process to enhance the breakdown of bacterial cell walls. The DNAs were quantified using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA), and molecular size was estimated using agarose gel electrophoresis. Primers targeting the hypervariable V3-V4 region of the 16S rRNA gene were used to amplify the extracted DNA samples. The forward primer was 5'-CCTACGGGNGGCWGCAG-3" and the reverse primer was 5"-GACTACHVGGGTATCTAATCC-3'. AxyPrep PCR Clean-up Kit (AP-PCR-500 G; Corning, NY, USA) was used to separate, extract and purify the PCR products, and the products were quantified using a Quant-iT PicoGreen dsDNA Reagent (P7581, Thermo Scientific, Waltham, MA, USA). After quality determination, libraries passing quality control were sequenced with Novaseq sequencer for 2 × Two terminal sequencing of 250 bp at LC-Bio Co., Ltd.

Quality filtering and denoising

Raw sequence quality was assessed using FastQC (v0.11.9). Adapter sequences and low-quality bases were trimmed using Trimmomatic (v0.39), with the first 20 bases removed due to a drop in quality in the initial cycles. Additional trimming was applied at the leading and trailing ends when base quality scores fell below Q30 and Q25, respectively. We employed QIIME2 for FASTQ annotation and microflora analysis. Low-quality sequences were filtered out, and the DADA2 pipeline was used for error correction, denoising, and chimera removal, generating a read count table of amplicon sequence variants (ASVs). The sequencing depth per sample had a minimum of 7768 reads, a maximum of 79,085 reads, a median of 55,623 reads, and a mean of 54,518.79 reads, with a standard deviation of 11,395.11 reads.

ASVs were taxonomically classified using the Naive Bayes consensus taxonomy classifier trained on the SILVA 16S rRNA database (v138), with species-level annotation assigned at 99% identity. To minimize sparsity-driven biases, ASVs present in <5% of samples or with <1000 total reads were excluded. Potential contaminants were identified and removed using the R package Decontam (version 1.10.0) in both frequency and prevalence modes. After all filtering steps, a total of 807 ASVs remained for downstream analysis.

Bioinformatic and statistic analysis

Quantitative variables conforming to normal distribution were presented as the mean ± SD analyzed by the Student's test and analysis of variance (ANOVA), while Quantitative variables of non-normal distribution were presented as median and interquartile ranges (25th and 75th percentiles) and analyzed by the Mann-Whitney U or Kruskal–Wallis test. Categorical variables were presented as rate or percentage, and chi-square test or Fisher test were used to analysis. Alpha and beta diversity analyses were conducted at the ASV level. ASVs were then classified against the SILVA 16S rRNA database (v138), and both relative abundance and count ASV tables were created. Alpha diversity was evaluated using the Chao-1, Simpson, Shannon, and Pielou's evenness indices. These indices were calculated using the phyloseq R package. Inter-group differences in alpha diversity indices were assessed using the Mann-Whitney U test (for two-group comparisons) with the wilcox.test function in R. The calculation of Aitchison distance first applied centered log-ratio (CLR) transformation to the count data using the microbiome or compositions R package, followed by the computation of Euclidean distance based on the CLR-transformed data. To handle zero values in the raw count data, we added a pseudo-count of 1 prior to the CLR transformation. The permutational multivariate analysis of variance (PERMANOVA) test in the Vegan (v2.6.4) R package was utilized to test the significance of differences in species composition and community structure of grouped samples. Beta diversity was visualized using PCoA plots generated in R (v4.2.1). Venn diagrams were generated using the VennDiagram package in R. An ASV was considered present in a group if its relative mean abundance exceeded 0.1%. The Venn diagrams illustrated the shared and unique ASVs among samples from different groups, highlighting the differences in microbial community composition.

The specific characterization of the tongue coating microbiota to distinguish taxonomic types was performed by the linear discriminant analysis (LDA) effect size (LEfSe) method with the microbiomeMarker package in R. Based on a normalized relative abundance matrix, LEfSe used the Kruskal – Wallis rank-sum test to detect features with significantly different abundance levels between assigned taxa and performs an LDA to estimate the effect size of each feature and then visualizes the results using taxonomic bubble diagrams and cladograms, with additional visualization and manipulation facilitated by the ggtree, phyloseq, and stringr packages.

Based on species abundance, the correlation coefficient (Spearman correlation coefficient) between each genus was calculated, and the correlation coefficient matrix was obtained. Gephi (v0.9.2) was used for the visualization of networks with significant correlations between genera.

PICRUSt2 (https://github.com/picrust/picrust2) was implemented to predict the enriched KEGG pathways between microbial community in participants with HC, GC, EC, HPB and CRC in this study. Used heatmaps to visualize the top 50 KEGG pathways, which were hierarchically clustered based on participant groups (HC, GC, EC, HPB, CRC). A one-way ANOVA was conducted to test for significant differences in pathway abundance among the different groups.

Machine learning for DST classification

The training cohort was randomly split into 80% training data (960 cases) and 20% internal validation data (239 cases) to ensure adequate model training and reliable subsequent validation. After denoising, a total of 807 ASVs, representing 93 microbial genera, were retained. The genus-level data, normalized by relative abundance, were used to construct the machine learning diagnostic models. We employed five commonly used and high-performing machine

learning algorithms: Generalized Linear Model (GLM), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost). Five-fold cross-validation on the training dataset was performed to optimize model performance and prevent overfitting. For each DST and HC sample, the algorithm generated a cancer score ranging from 0 to 1, with higher scores indicating a greater likelihood of digestive system tumors.

Receiver Operating Characteristic (ROC) curves were plotted, and the Area Under the ROC Curve (AUC) was calculated to quantify the models' ability to distinguish DST from HC samples. Higher AUC values indicated better discriminatory power. All data preprocessing, feature selection, model training, and performance evaluation were conducted using R (v4.2.1).

Results

Characteristics of study populat

We prospectively collected tongue coating samples from 1199 participants, including 710 patients with digestive system tumours (DSTs) (419 with GC, 62 with CRC, 107 with EC, and 122 with HPB) and 489 matched healthy controls (HC). Table 1 summaries the baseline characteristics of the participants. Compared with HC, there were no significant differences in age, gender, smoking, or drinking history among patients with DST (all $p > 0.05$). For EC, the proportion of male patients was significantly higher than in the HC group (86.0% vs. 66.1%, $p < 0.001$), and the corresponding smoking and drinking habits were significantly increased. Moreover, the mean age of EC patients was higher than that of the HC group (65.0 years vs. 63.4 years, $p = 0.036$). Additionally, the proportion of CRC patients with a history of drinking was significantly lower than in the HC group (14.5% vs 26.6%, $p = 0.039$). Patients with HPB tended to be younger ($p = 0.02$).

Changes in tongue coating microbiota diversity and characteristics in DST patients

To illustrate microbial community richness and species diversity, we used alpha diversity indices, including the Chao1 index, Pielou index, Shannon index, and Simpson index, all calculated at the ASV level. There was no significant difference in the Chao1 index between the DST and HC groups ($p > 0.05$, Figure 1a). The microbial community evenness in DST patients was significantly higher (increased Pielou index, $p < 0.05$, Figure 1b). Furthermore, the microbial community diversity in DST patients was significantly elevated (increased Shannon and Simpson indices, $p < 0.05$, Figure 1c, d). After applying centered log-ratio (CLR) transformation to the tongue coating microbiome data to eliminate the covariate effects of abundance variation between samples, beta diversity analysis (based on Aitchison distance) revealed a significant difference in the microbial community structure between the DST and HC groups ($p = 0.001$, Figure 1e). These findings suggested that DST patients exhibited increased microbial community evenness and diversity, potentially indicating distinct community structural characteristics.

We evaluated the bacterial distribution based on the relative abundance of different taxa. The ten dominant genera in the HC and DST groups were *Prevotella*, *Neisseria*, *Veillonella*, *Fusobacterium*, *Haemophilus*, *Alloprevotella*, *Porphyromonas*, *Streptococcus*, *Leptotrichia*, and *Actinomyces*, accounting for 80.9% and 83.6% of the total ASVs, respectively (Figure 1f, Supplementary Table S1). We performed Venn analysis on ASVs with an average relative abundance greater than 0.001 in each group. The result showed that the DST and HC groups shared 137 ASVs, while 62 ASVs were unique to the DST group and only 3 ASVs were unique to the HC group (Figure 3g). LEfSe discriminant analysis revealed significant differences in five bacterial genera between the HC and DST groups (Figure 1h, i), all of which were

Table 1. Clinical characteristics of DST patients and healthy controls.

	DST					HC	p value (T vs. HC)				
	All	GC	CRC	EC	HPB[a]		All	GC	CRC	EC	HPB
Gender											
Male (%)	493 (69.4)	268 (64.0)	45 (72.6)	92 (86.0)	88 (72.1)	323 (66.1)	0.217	0.510	0.304	<0.001*	0.201
Female (%)	217 (30.6)	151 (36.0)	17 (27.4)	15 (14.0)	34 (27.9)	166 (33.9)					
Age (SD)	63.2 (10.7)	63.6 (11.1)	64.0 (10.3)	65.0 (7.8)	60.0 (11.1)	63.4 (7.4)	0.787	0.735	0.571	0.036*	<0.001*
<65%)	346 (48.7)	195 (46.5)	26 (41.9)	49 (45.8)	76 (62.3)	247 (50.5)	0.545	0.233	0.203	0.377	0.020*
≥65%)	364 (51.3)	224 (53.5)	36 (58.1)	58 (54.2)	46 (37.7)	242 (49.5)					
Smoking											
No (%)	466 (65.6)	286 (68.3)	47 (75.8)	42 (39.3)	91 (74.6)	338 (69.1)	0.207	0.780	0.280	<0.001*	0.237
Yes (%)	244 (34.4)	133 (31.7)	15 (24.2)	65 (60.7)	31 (25.4)	151 (30.9)					
Drinking											
No (%)	485 (68.3)	289 (69.0)	53 (85.5)	43 (40.2)	100 (82.0)	359 (73.4)	0.057	0.140	0.039*	<0.001*	0.051
Yes (%)	225 (31.7)	130 (31.0)	9 (14.5)	64 (59.8)	22 (18.0)	130 (26.6)					

DST, digestive system tumour; GC, gastric cancer; CRC, colorectal cancer; EC, Esophageal cancer; HPB, hepatopancreatobiliary cancer; HC, healthy control. HPB include 100 cases of hepatocellular carcinoma (HCC), 14 cases of pancreatic ductal adenocarcinoma (PDAC), and 8 cases of cholangiocarcinoma (CCA). *$p < 0.05$.
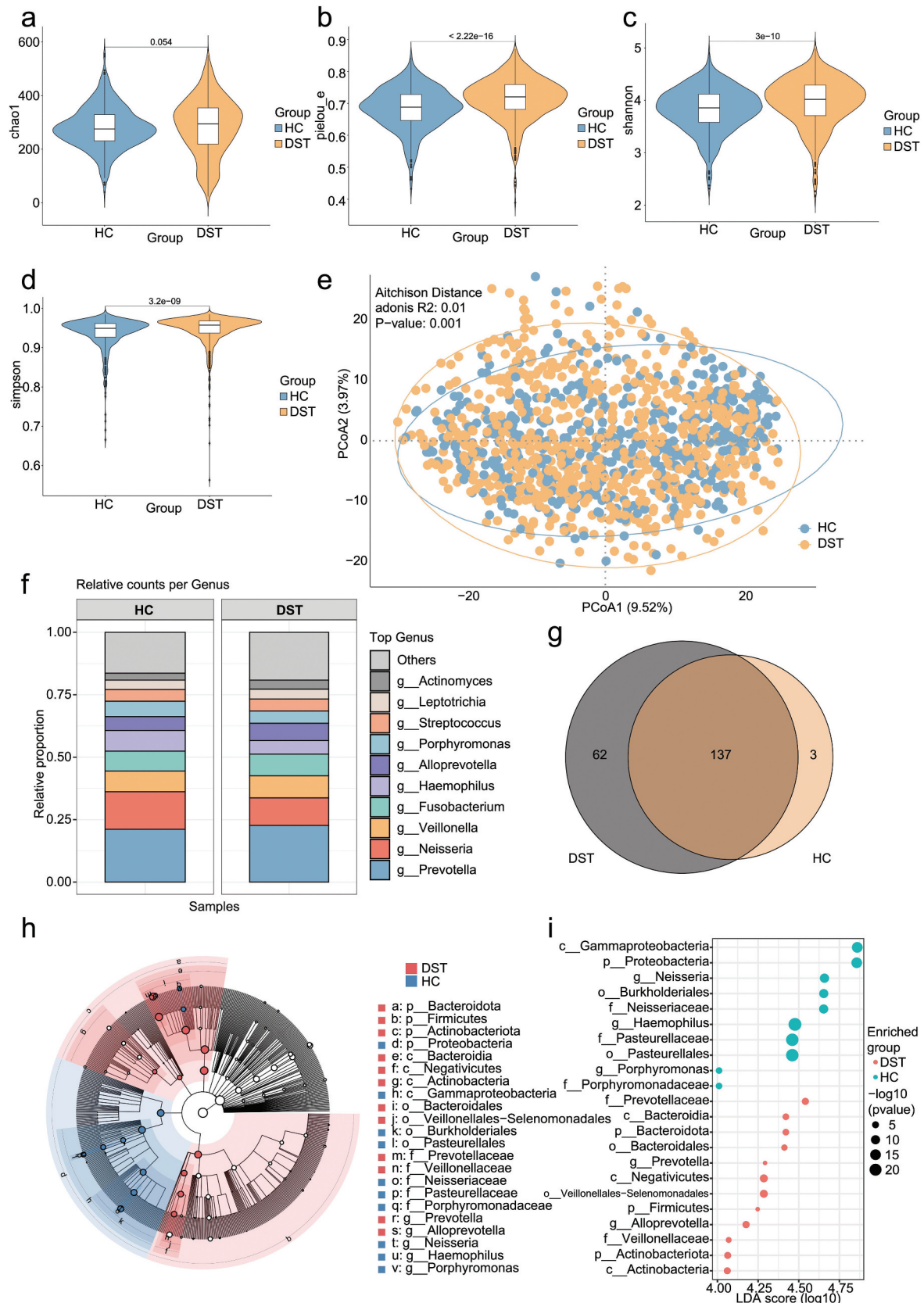
Figure 1. The microbial communities on the tongue coating of patients with DST have changed compared to HC group. Alpha diversity analysis of the HC and DST groups. (a) Chao11 index, (b) Simpson index, (c) Pielou's index, and (d) Shannon index for each group in tongue coating samples (Wilcoxon rank-sum test). (e) Principal coordinate analysis (PCoA) of the tongue coating bacteria could significantly separate the HC and DST groups into different clusters based on the Aitchison distance (PERMANOVA). (f) Composition of the predominant bacterial genera in the tongue coating samples. (g) The Venn diagram shows the sharing of ASVs with an average relative abundance greater than 0.001 in the tongue coating microbiomes between the HC and DST groups. LEfSe analysis for comparing microbial variations in HC and DST at the genus level. (h) LEfSe cladogram representing differentially abundant taxa ($p < 0.05$); (i) LDA scores are calculated based on the LEfSe of differentially abundant taxa among groups, and only taxa with LDA scores of > 4 are presented. DST, digestive system tumour; HC, healthy control.

among the top 10 in relative abundance. Compared with the HC group, the proportions of *Alloprevotella* and *Prevotella* were significantly higher in the tongue coating microbiota of DST patients, while the proportions of *Neisseria*, *Haemophilus*, and *Porphyromonas* were significantly lower (LDA >4, $p < 0.05$). These findings suggested that the differential richness of the tongue coating microbiota, including *Alloprevotella* and *Neisseria*, could distinguish between HC and DST patients.

Changes in tongue coating microbiota diversity and characteristics in patients with different DST subtypes

We further analysed the differences in tongue coating microbiota among DST subtypes (including GC, CRC, EC, and HPB) and the HC group. The Chao1 index of each DST subtype was higher than that of the HC group (Figure 2a, $p < 0.05$). There were no significant differences in microbiota diversity (Shannon index) among the different DST subtypes, but the microbial diversity of all DST subtypes was significantly higher than that of the HC group (Figure 2b, all $p < 0.05$). Additionally, in terms of beta diversity, analyses based on Aitchison distance revealed significant differences in microbial community structure among the four DST subtypes and the HC group (Figure 2c, $p < 0.05$). The relative abundance of dominant genera also showed some differences between the HC group and each DST subtype (Figure 2d, Supplementary Table S2). We performed a Venn analysis of ASVs with an average relative abundance greater than 0.001 in each group. The results showed that 103 ASVs were shared among the five groups, with 8 ASVs unique to the GC group, 13 ASVs unique to the CRC group, 3 ASVs unique to the HPB group, 8 ASVs unique to the EC group, and 3 ASVs unique to the HC group (Figure 2e). These unique ASVs might have reflected the specific impacts of different digestive system tumours on the tongue coating microbiome, suggesting their potential value as biomarkers.

Key bacterial taxa distinguishing each DST from the HC group were identified based on LEfSe discriminant analysis. Compared to HCs, the proportions of *Campylobacter* and *Fusobacterium* were significantly higher in the tongue coating of GC patients (LDA >4, Figure 2f, g). Notably, *Fusobacterium* was significantly increased in both the GC and CRC groups, whereas it was significantly decreased in the EC group (Figure 2f, g, Supplementary figure S1a, b). The proportion of *Alloprevotella* was significantly higher in the tongue coating of CRC, EC, and HPB patients (Supplementary Figure S1). In the LEfSe discriminant analyses comparing each of the four DSTs with the HC group, we found that the proportion of *Haemophilus* was significantly higher in the tongue coating of HC group (Supplementary

Figure S1), suggesting that *Haemophilus* might be a key genus for identifying the HC group.

Microecological network differences in tongue coating microbiota between HC and DST patients

Compared with the healthy control group, the correlation network structure of tongue coating microbiota in patients with digestive system tumours was more complex. The tongue coating microbiota in the HC group showed fewer and weaker associations, reflecting a stable microbial ecosystem in a healthy state. In contrast, the complexity and diversity of the tongue coating microbiota community were significantly increased in patients with DSTs, especially in the EC and CRC groups. However, the tongue microbiota community in GC patients appeared to be simplified, with reduced microbial associations (Figure 3).

In the HC group, microorganisms from the phyla *Bacteroidetes* and *Campilobacterota* might have exhibited strong associations, reflecting their stable symbiotic relationship in a healthy state. However, in the GC group, these correlations weakened or disappeared, potentially due to changes in the gastric environment caused by GC, such as alterations in acidity and immune responses, which could have disrupted the stability of the tongue coating microbiome ecosystem. In the CRC group, certain microbes from the phylum *Fusobacteria* (e.g. some *Fusobacterium* and *Leptotrichia*) formed new correlations with the phylum *Firmicutes*, which were absent in the HC group. This suggested that *Fusobacteria* and *Firmicutes* might have played a potential role in the oral microbiome of CRC patients. In the HPB group, the correlations involving the phylum *Actinobacteria* were significantly enhanced, a change that was not apparent in the GC, CRC, EC, or HC groups, suggesting that *Actinobacteria* could have been one of the specific markers for HPB patients. This indicated that the tongue coating microbiota in patients with different DST subtypes exhibited significant differences in structure and associations.

Identification of microbial function capacity associated with DST

We compared the potential functional capacities of tongue coating microbiota between different DST subtypes and the HC group by mapping and classifying microbial genes using the KEGG database. We analysed the functional differences in the tongue coating microbiota of the five groups using KEGG Orthology (Suplementary figure S2a). KEGG Orthology analysis revealed the enrichment of putative ABC transport system permease protein, putative ABC transport system ATP-binding

Figure 2. The microbial communities on the tongue coating have changed in participants with GC, EC, HPB, and CRC, compared to healthy controls. Alpha diversity analysis of the GC, EC, HPB, CRC, and HC groups, showing (a) Chao1 index, and (b) Shannon index for each group in tongue coating samples (Wilcoxon rank-sum test). (c) Principal coordinate analysis (PCoA) of the tongue coating bacteria could significantly separate the GC, EC, HPB, CRC, and HC groups into different clusters based on the Aitchison distance (PERMANOVA). (d) Composition of the predominant bacterial genera in the tongue coating samples. (e) The Venn diagram shows the sharing of ASVs with an average relative abundance greater than 0.001 in the tongue coating microbiomes between the HC group and each DST subtype. LEfSe analysis for comparing microbial variations in HC and GC at the genus level. (f) LEfSe cladogram representing differentially abundant taxa ($p < 0.05$); (g) LDA scores are calculated based on the LEfSe of differentially abundant taxa among groups, and only taxa with LDA scores of > 4 are presented. GC, gastric cancer; EC, esophageal cancer; CRC, colorectal cancer; HPB, hepatopancreatobiliary cancer; HC, healthy control.

protein, polar amino acid transport system substrate-binding protein, transketolase, and iron complex transport system substrate-binding protein in the gastric cancer group, primarily reflecting

aspects of transmembrane transport and energy dependence. In the HPB group, methionyl aminopeptidase, FKBP-type peptidyl-prolyl cis-trans isomerase, periplasmic protein TonB, ABC-2 type
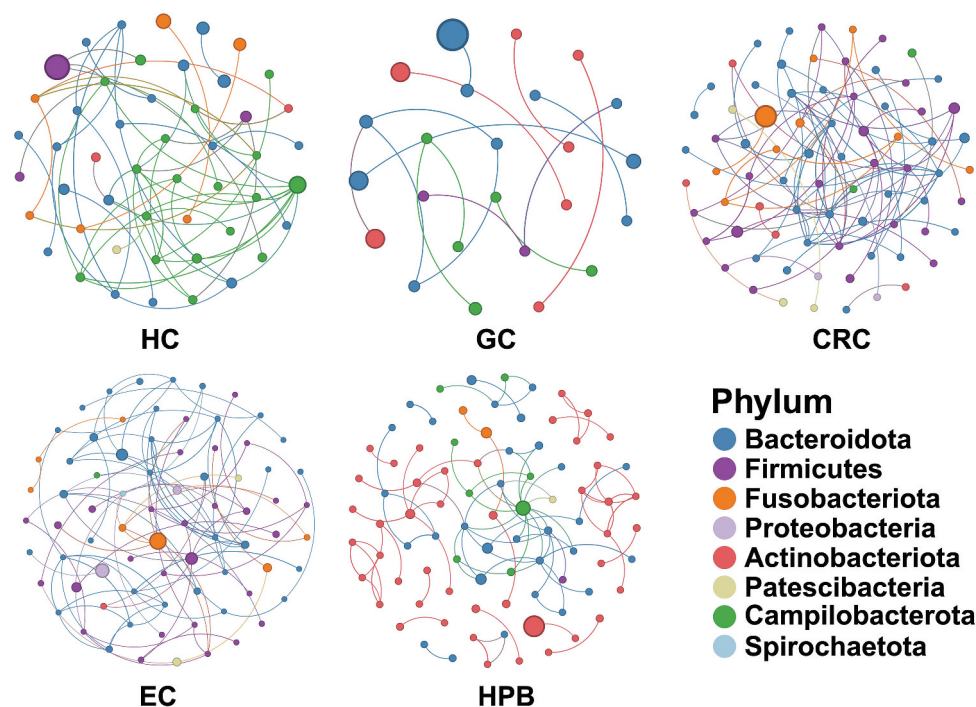
Figure 3. Bacterial interaction network diagram of HC, GC, CRC, EC and HPB. Different nodes represent different genera, and node size represents the average relative abundance of the genus. Genera within the same phylum are represented by the same color. GC, gastric cancer; EC, esophageal cancer; CRC, colorectal cancer; HPB, hepatopancreatobiliary; HC, healthy control.

transport system permease protein, and iron complex outer membrane receptor protein were enriched, mainly reflecting aspects of protein synthesis and modification as well as substance transport and metabolism. In the CRC group, alanine or glycine symporter, AGCS family, ATP-binding cassette, subfamily B, bacterial, and peptide/nickel transport system substrate-binding protein were enriched, mainly reflecting aspects of transmembrane transport and nutrient acquisition. The EC group showed no significantly enriched KO genes. These findings indicated that the tongue coating microbiota of patients with digestive system tumours generally exhibited enhanced functions in transmembrane transport and nutrient synthesis and acquisition.

We analysed the functional differences in the tongue coating microbiota of the five groups using KEGG Metabolism (Supplementary figure S2b). KEGG Metabolism analysis showed that starch and sucrose metabolism, lysine biosynthesis, oxidative phosphorylation, and phenylalanine, tyrosine, and tryptophan biosynthesis were enriched in the tongue coating of the EC group, reflecting characteristics of nutrient competition. In the GC group, selenocompound metabolism and porphyrin metabolism were enriched, reflecting antioxidative and anti-inflammatory effects as well as energy metabolism. In the HPB group, pyrimidine metabolism, one-carbon pool by folate, peptidoglycan biosynthesis,

and drug metabolism-other enzymes were enriched, reflecting aspects of cell proliferation, immunity, and drug resistance. In the CRC group, D-amino acid metabolism and nicotinate and nicotinamide metabolism were enriched, reflecting characteristics of energy metabolism. These findings indicated that the metabolic functions of the tongue-coated microbiota in patients with digestive system tumours had changed.

Machine learning models for DST diagnosis

We developed five machine learning models based on 93 microbial genera from tongue coating samples to assess the diagnostic performance for predicting DST, including GLM, RF, DT, SVM, and XGBoost models. We randomly divided the training cohort (1199 cases) into 80% training data (960 cases) and 20% internal validation data (239 cases). Five-fold cross-validation based on the training dataset was used to optimize model performance and prevent overfitting. In the training dataset, all five machine learning models demonstrated excellent diagnostic performance. Among them, the XGBoost model had an AUC value of 0.978 (95% confidence interval (CI): 0.97–0.986), the SVM model had an AUC value of 0.993 (95% CI: 0.988–0.998), the RF model had an AUC value of 0.921 (95% CI: 0.905–0.936), the GLM model had an AUC value of 0.945 (95% CI: 0.931–0.959), and the DT model had an AUC value of 0.941 (95% CI: 0.926–0.957) (Figure 4a).

In the independent internal validation dataset, the XGBoost model showed even better predictive performance, with an AUC value of 0.926 (95% CI: 0.893–0.958), outperforming the other four machine learning models (Figure 4b). Further analysis of the model scores for different DST subtypes (such as EC, GC, HPB, and CRC) showed that the model scores for the four digestive system tumors remained generally consistent, and the scores for different machine learning models were stable, with average scores mostly ranging from 0.7 to 0.9 (Figure 4c). For the HC group, the average score for most machine learning models was around 0.2, with only the RF model having a relatively higher average score of 0.37 (Figure 4c). These results indicated that the XGBoost model exhibited the best diagnostic performance and predictive accuracy, while the scores for different digestive system tumors across the machine learning models showed good stability.

Discussion

The oral cavity is the second largest microbial ecosystem in the human body [25]. Both traditional wisdom and modern scientific research indicate that changes in the balance of oral bacteria may signal pathological conditions. These changes are associated not only with oral diseases such as halitosis, caries, and periodontitis but also with systemic diseases affecting the respiratory, circulatory, endocrine systems, and even malignancies [26–28]. This study, by analyzing the diversity and characteristic differences of tongue coating microbiota, revealed for the first time its potential role in digestive system tumours. Based on 16S rRNA gene sequencing data, we developed five stable classifiers based on tongue coating microbiota, with the XGBoost model showing outstanding performance in predicting DST. These findings were of significant clinical importance, suggesting that tongue-coating
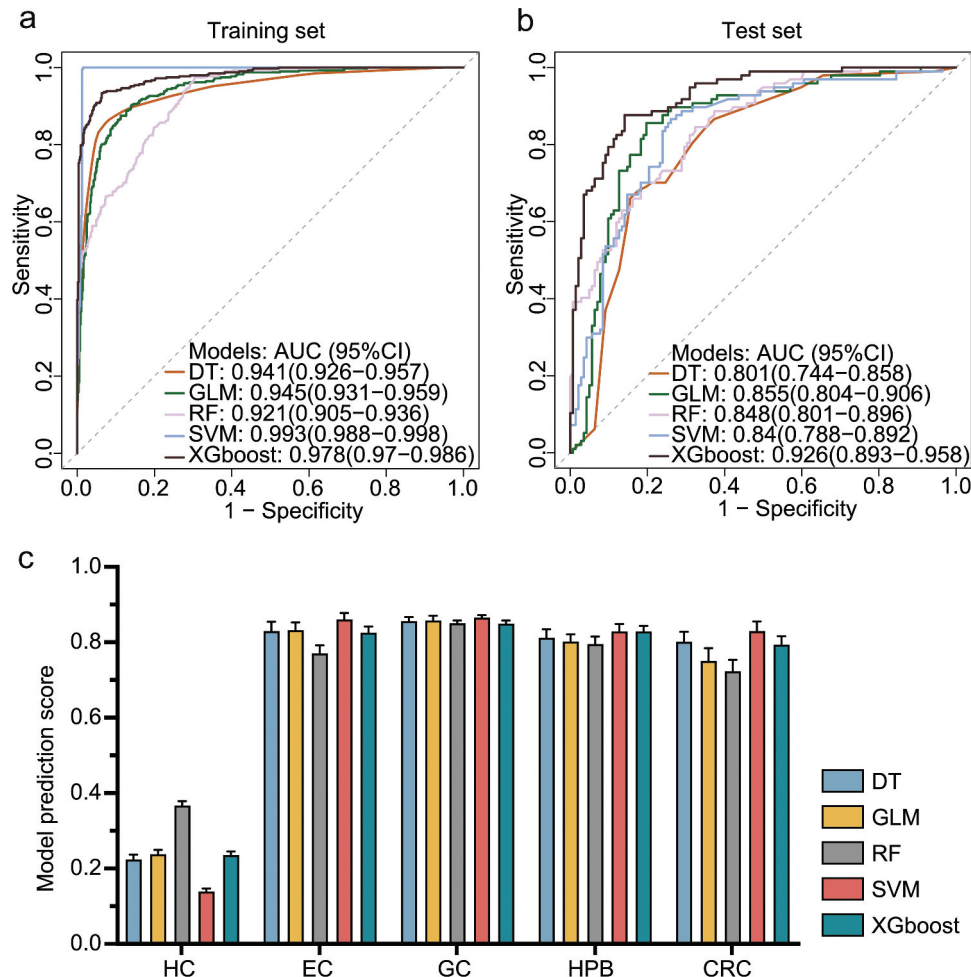

Figure 4. Performance evaluation of five machine learning models in the training dataset and independent internal validation dataset, and their scores for different digestive system tumors. (a) ROC curves and AUC values of five machine learning models in the training dataset; (b) ROC curves and AUC values of five machine learning models in the independent internal validation dataset; (c) scores of different machine learning models for different DST subtypes and HC group, with bar charts showing the average score and standard error for each group. ROC, receiver operating Characteristic; AUC, area under the Curve; CI, confidence interval; GC, gastric cancer; EC, esophageal cancer; CRC, colorectal cancer; HPB, hepatopancreatobiliary cancer; HC, healthy control.

microbiota could serve as a non-invasive biomarker for early detection and risk assessment of DST.

Compared to the HC group, the tongue coating microbiome diversity in DST patients was significantly increased (Shannon and Simpson indices were elevated). These changes might have reflected the remodeling of the tongue coating microbiome ecosystem in DST patients. According to the LEfSe results, the abundance of the genera *Neisseria*, *Haemophilus*, and *Porphyromonas* was lower in the DST group. Several oral microbiome studies have also found that the relative abundance of these three genera is higher in healthy individuals than in cancer patients [29,30]. In a study on pancreatic head carcinoma (PHC), *Haemophilus* and *Porphyromonas* were enriched in the tongue coating microbiota of healthy controls [14]. Wu et al. [31] also confirmed that *Neisseria* and *Porphyromonas* proportions were lower in the tongue coatings of GC patients. *Neisseria* was considered a protective microbe that reduced the risk of cancers such as pancreatic adenocarcinoma (PDAC) [32], oral cancer [33] and GC [30,34]. There was less research on the genus *Porphyromonas*, although *P. gingivalis* had been shown to promote the occurrence of many tumours and was associated with poor prognosis in patients [12]. Additionally, we found that *Alloprevotella* was significantly increased in the tongue coating of overall DST and consistently in CRC, EC, and HPB groups. Fan et al. [19] showed that oral *Alloprevotella* was associated with an increased risk of pancreatic cancer, and similar results were found in cardia cancer [31]. *Prevotella* also increased the risk of DST, which had been confirmed by many studies [15,31].

Changes in tongue-coating microbiota could have led to changes in metabolic pathways and functions. We performed functional prediction based on the KEGG database. We found that the tongue coating microbiota in DST patients generally enhanced genes related to transmembrane transport and nutrient synthesis and acquisition. *Streptococcus*, *Alloprevotella*, and *Prevotella* were usually associated with polysaccharide degradation and fermentation [35–37], which involved related ABC transporters. Phenylalanine, tyrosine, and tryptophan biosynthesis were associated with *Alloprevotella* [38], consistent with our analysis results of the EC group. These amino acid metabolites might directly affect the growth and invasiveness of tumour cells [39]. Our previous study on the tongue coating proteome of GC patients found that microbial-derived proteins linked to ABC transporters were significantly reduced [40]. The 16S rRNA results showed that putative ABC transport system permease protein and putative ABC transport system ATP-binding protein were increased in the tongue coating of the GC group. This could indicate an adaptive response of the microbial community to the altered environment in GC, potentially aiming to maintain essential functions such as nutrient acquisition and cellular homeostasis in the face of stressors imposed by the tongue coating microenvironment. Furthermore, in the GC group, porphyrin metabolism was abnormally enhanced in tongue coating metabolic pathways. Porphyrin could overdrive cancer cells to reprogram metabolism to support abnormal growth [41]. We found significant enhancement of pyrimidine metabolism only in the HPB group. Catalytic degradation of pyrimidine maintained a mesenchymal-like state driven by epithelial-mesenchymal transition in hepatocellular carcinoma [42]. Nicotinate and nicotinamide metabolism were enhanced in the tongue coatings of CRC. Nicotinate and nicotinamide were precursors of NAD^+, and increased NAD^+ levels could promote glycolysis and fuel cancer cells [43].

Digestive system tumours were characterized by limited treatment options, poor prognosis, and frequent recurrence and metastasis at advanced stages [44–46]. There was an urgent need for a reliable tool for early cancer screening. Although endoscopy was considered the gold standard for digestive system tumour screening, it was a relatively expensive and invasive procedure that caused some degree of harm to patients [45,46]. Clinical serum tumour markers such as α-fetoprotein (AFP), carcinoembryonic antigen (CEA), carbohydrate antigen (CA) 125, CA72–4, CA19–9, and CA242 were commonly used for cancer screening [47–49]. However, these markers exhibited low sensitivity and lacked the desired specificity. Tongue diagnosis, as a unique method in traditional Chinese medicine, utilized significant differences in tongue coating microbiota under different disease conditions, potentially providing a method for cancer screening and early diagnosis [29]. In this study, five machine learning models were developed based on 93 microbial genera from tongue coating samples to evaluate their diagnostic performance in predicting DST. In the independent internal validation dataset, the XGBoost model showed an AUC value of 0.926 (95% CI: 0.893–0.958), outperforming the other four machine learning models. As an efficient machine learning algorithm, XGBoost demonstrated excellent predictive performance with lower computational costs and generated a more streamlined model [50]. The model scores for different DST subtypes tended to be stable. Compared to endoscopy, this tongue coating microbiota-based screening method was non-invasive and convenient, overcoming the limitations of requiring specialized equipment and medical personnel for endoscopic examination.

Of course, this study had several limitations that needed to be addressed in future research. First, the

heterogeneity of cancer meant that the study could not observe differences between patients with different tumour sites and/or different disease stages. Second, tongue coating 16S rRNA data could not precisely reach the species level, leaving conclusions at the genus level. Additionally, the case–control study was conducted only in the eastern Chinese population, with a small sample size for hepatobiliary and pancreatic tumours. The clinical applicability of oral tongue microbiota in diagnosing DSTs needed further validation in larger prospective cohorts. Finally, we included all high-abundance genera (with a threshold of fewer than 1100 non-expressing samples out of 1199) in the model. This could have complicated the model; hence, we needed to select representative key genera through screening for future modeling to improve model applicability.

In conclusion, this study systematically analyzed the diversity and characteristic differences of the tongue coating microbiota, revealing its potential diagnostic value for digestive system tumours. The machine learning models we developed, particularly the XGBoost model, demonstrated excellent performance in predicting DSTs. These results provide a solid foundation for the clinical application of tongue coating microbiota as a non-invasive biomarker, and they offer important guidance for future research directions.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Authors' contributions

Conceptualization: Z-CJ, LY, X-DC. Data curation: Y-BM, KL, R-HX. Formal analysis: Z-CJ, Y-BM. Funding acquisition: Z-CJ, LY, X-DC. Investigation: Y-NW, L-BP. Methodology: Z-CJ, Y-BM. Software: Z-CJ, L-BP. Supervision: LY, X-DC. Visualization: Z-CJ, L-BP. Writing – original draft: Z-CJ, Y-BM. Writing – review & editing: LY, X-DC. All authors read and approved the final manuscript.

Availability of data and material

All raw reads were stored in NCBI Sequence Read Archive (SRA) database, and the accession number is PRJNA1236253.

Ethics approval and consent to participate

All tongue coating samples were obtained according to a protocol approved by the ethics committee of participating centers (IRB-2023-100), and informed consent for use was obtained from all participants.

ORCID

Zhengchen Jiang http://orcid.org/0000-0003-4765-8830
Li Yuan http://orcid.org/0000-0002-6245-9437
Xiangdong Cheng http://orcid.org/0000-0002-5099-490X

References

[1] Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–263. doi: 10.3322/caac.21834

[2] Li S, Li Q, Ren Y, et al. HSV: the scout and assault for digestive system tumors. Front Mol Biosci. 2023;10:1142498. doi: 10.3389/fmolb.2023.1142498

[3] Marrelli D, Polom K, de Manzoni G, et al. Multimodal treatment of gastric cancer in the west: where are we going? World J Gastroenterol. 2015;21(26):7954–7969. doi: 10.3748/wjg.v21.i26.7954

[4] Grossberg AJ, Chu LC, Deig CR, et al. Multidisciplinary standards of care and recent progress in pancreatic ductal adenocarcinoma. CA Cancer J Clin. 2020;70(5):375–403. doi: 10.3322/caac.21626

[5] Huang FL, Yu SJ. Esophageal cancer: risk factors, genetic association, and treatment. Asian J Surg. 2018;41(3):210–215. doi: 10.1016/j.asjsur.2016.10.005

[6] Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. JAMA. 2021;325 (7):669–685. doi: 10.1001/jama.2021.0106

[7] Zhang Y, Wang Y, Zhang B, et al. Methods and biomarkers for early detection, prediction, and diagnosis of colorectal cancer. Biomed Pharmacother. 2023;163:114786. doi: 10.1016/j.biopha.2023.114786

[8] Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet. 2020;396(10251):635–648. doi: 10.1016/S0140-6736(20)31288-5

[9] Villanueva A, Longo DL. Hepatocellular carcinoma. N Engl J Med. 2019;380(15):1450–1462. doi: 10.1056/NEJMra1713263

[10] Lu H, He L, Xu J, et al. Well-maintained patients with a history of periodontitis still harbor a more dysbiotic microbiome than health. J Periodontol. 2020;91:1584–1594 doi: 10.1002/JPER.19-0498

[11] Lim Y, Totsika M, Morrison M, et al. Oral microbiome: a new biomarker reservoir for oral and oropharyngeal

cancers. Theranostics. 2017;7(17):4313–4321. doi: 10. 7150/thno.21804

[12] Stasiewicz M, Karpinski TM. The oral microbiota and its role in carcinogenesis. Semin Cancer Biol. 2022;86:633–642. doi: 10.1016/j.semcancer.2021.11.002

[13] Xiao J, Fiscella KA, Gill SR. Oral microbiome: possible harbinger for children's health. Int J Oral Sci. 2020;12 (1):12. doi: 10.1038/s41368-020-0082-x

[14] Lu H, Ren Z, Li A, et al. Tongue coating microbiome data distinguish patients with pancreatic head cancer from healthy controls. J Oral Microbiol. 2019;11:1563409 doi: 10.1080/20002297.2018.1563409

[15] Chen Q, Huang X, Zhang H, et al. Characterization of tongue coating microbiome from patients with colorectal cancer. J Oral Microbiol. 2024;16:2344278 doi: 10.1080/20002297.2024.2344278

[16] Yuan L, Yang L, Zhang S, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. EClinicalMedicine. 2023;57:101834. doi: 10.1016/j.eclinm.2023.101834

[17] Yang Y, Long J, Wang C, et al. Prospective study of oral microbiome and gastric cancer risk among Asian, African American and European American populations. Intl J Cancer. 2022;150(6):916–927. doi: 10.1002/ijc.33847

[18] Peters BA, Wu J, Pei Z, et al. Oral microbiome composition reflects prospective risk for esophageal cancers. Cancer Res. 2017;77:6777–6787 doi: 10.1158/0008-5472.CAN-17-1296

[19] Fan X, Alekseyenko AV, Wu J, et al. Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. Gut. 2018;67(1):120–127. doi: 10.1136/gutjnl-2016-312580

[20] Tan Q, Ma X, Yang B, et al. Periodontitis pathogen porphyromonas gingivalis promotes pancreatic tumorigenesis via neutrophil elastase from tumor-associated neutrophils. Gut Microbes. 2022;14 (1):2073785. doi: 10.1080/19490976.2022.2073785

[21] Mitsuhashi K, Nosho K, Sukawa Y, et al. Association of fusobacterium species in pancreatic cancer tissues with molecular features and prognosis. Oncotarget. 2015;6(9):7209–7220. doi: 10.18632/oncotarget.3109

[22] Morgan AD, Seely KD, Hagenstein LD, et al. Bacterial involvement in progression and metastasis of adenocarcinoma of the stomach. Cancers (Basel). 2022;14 (19):14. doi: 10.3390/cancers14194886

[23] Flemer B, Warren RD, Barrett MP, et al. The oral microbiota in colorectal cancer is distinctive and predictive. Gut. 2018;67(8):1454–1463. doi: 10.1136/gutjnl-2017-314814

[24] Lan Z, Liu WJ, Cui H, et al. The role of oral microbiota in cancer. Front Microbiol. 2023;14:1253025. doi: 10.3389/fmicb.2023.1253025

[25] Backhed F, Fraser CM, Ringel Y, et al. Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. Cell Host Microbe. 2012;12:611–622 doi: 10.1016/j.chom.2012.10.012

[26] Gao L, Xu T, Huang G, et al. Oral microbiomes: more and more importance in oral cavity and whole body. Protein Cell. 2018;9:488–500 doi: 10.1007/s13238-018-0548-1

[27] Baker JL, Welch JLM, Kauffman KM, et al. The oral microbiome: diversity, biogeography and human health. Nat Rev Microbiol. 2024;22(2):89–104. doi: 10.1038/s41579-023-00963-6

[28] Chen Y, Chen X, Yu H, et al. Oral microbiota as promising diagnostic biomarkers for gastrointestinal cancer: a systematic review. Onco Targets Ther. 2019;12:11131–11144 doi: 10.2147/OTT.S230262

[29] Han S, Yang X, Qi Q, et al. Potential screening and early diagnosis method for cancer: tongue diagnosis. Int J Oncol. 2016;48:2257–2264 doi: 10.3892/ijo.2016.3466

[30] Huang K, Gao X, Wu L, et al. Salivary microbiota for gastric cancer prediction: an exploratory study. Front Cell Infect Microbiol. 2021;11:640309. doi: 10.3389/fcimb.2021.640309

[31] Wu J, Xu S, Xiang C, et al. Tongue coating microbiota community and risk effect on gastric cancer. J Cancer. 2018;9(21):4039–4048. doi: 10.7150/jca.25280

[32] Wei AL, Li M, Li GQ, et al. Oral microbiome and pancreatic cancer. World J Gastroenterol. 2020;26 (48):7679–7692. doi: 10.3748/wjg.v26.i48.7679

[33] Zeng B, Tan J, Guo G, et al. The oral cancer microbiome contains tumor space-specific and clinicopathology-specific bacteria. Front Cell Infect Microbiol. 2022;12:942328 doi: 10.3389/fcimb.2022.942328

[34] Dias-Jacome E, Libanio D, Borges-Canha M, et al. Gastric microbiota and carcinogenesis: the role of non-Helicobacter pylori bacteria - a systematic review. Rev Esp Enferm Dig. 2016;108(9):530–540. doi: 10.17235/reed.2016.4261/2016

[35] Chen J, Li Z, Wang X, et al. Isomaltooligosaccharides sustain the growth of prevotella both in vitro and in animal models. Microbiol Spectr. 2022;10:e0262121 doi: 10.1128/spectrum.02621-21

[36] Virginio Junior GF, Reis ME, da Silva AP, et al. Does algae β-glucan affect the fecal bacteriome in dairy calves? PLOS ONE. 2021;16(9):e0258069. doi: 10.1371/journal.pone.0258069

[37] Culurgioni S, Harris G, Singh AK, et al. Structural basis for regulation and specificity of fructooligosaccharide import in Streptococcus pneumoniae. Structure. 2017;25(1):79–93. doi: 10.1016/j.str.2016.11.008

[38] Zeng Z, Lv B, Tang YE, et al. Effects of dietary selenized glucose on intestinal microbiota and tryptophan metabolism in rats: assessing skatole reduction potential. Environ Res. 2024;252:118874. doi: 10.1016/j.envres.2024.118874

[39] Platten M, Nollen EAA, Rohrig UF, et al. Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond. Nat Rev Drug Discov. 2019;18(5):379–401. doi: 10.1038/s41573-019-0016-5

[40] Xu S, Xiang C, Wu J, et al. Tongue coating bacteria as a potential stable biomarker for gastric cancer Independent of lifestyle. Dig Dis Sci. 2021;66 (9):2964–2980. doi: 10.1007/s10620-020-06637-0

[41] Adapa SR, Hunter GA, Amin NE, et al. Porphyrin overdrive rewires cancer cell metabolism. Life Sci Alliance. 2024;7(7):e202302547. doi: 10.26508/lsa.202302547

[42] Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. Mol Metab. 2020;35:100962. doi: 10.1016/j.molmet.2020.02.005

[43] Yaku K, Okabe K, Hikosaka K, et al. NAD metabolism in cancer therapeutics. Front Oncol. 2018;8:622. doi: 10.3389/fonc.2018.00622

[44] Stoffel EM, Brand RE, Goggins M. Pancreatic cancer: changing epidemiology and new approaches to risk assessment, early detection, and prevention. Gastroenterology. 2023;164(5):752–765. doi: 10.1053/j.gastro.2023.02.012

[45] Jain S, Maque J, Galoosian A, et al. Optimal strategies for colorectal cancer screening. Curr Treat Options Oncol. 2022;23:474–493. doi: 10.1007/s11864-022-00962-4

[46] Necula L, Matei L, Dragu D, et al. Recent advances in gastric cancer early diagnosis. World J Gastroenterol. 2019;25(17):2029–2044. doi: 10.3748/wjg.v25.i17.2029

[47] Ke X, Liu W, Shen L, et al. Early screening of colorectal precancerous lesions based on combined measurement of multiple serum tumor markers using artificial neural network analysis. Biosensors (Basel). 2023;13(7):685. doi: 10.3390/bios13070685

[48] Johnson P, Zhou Q, Dao DY, et al. Circulating biomarkers in the diagnosis and management of hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2022;19(10):670–681. doi: 10.1038/s41575-022-00620-y

[49] Ge L, Pan B, Song F, et al. Comparing the diagnostic accuracy of five common tumour biomarkers and CA19–9 for pancreatic cancer: a protocol for a network meta-analysis of diagnostic test accuracy. BMJ Open. 2017;7:e018175. doi: 10.1136/bmjopen-2017-018175

[50] Sheridan RP, Wang WM, Liaw A, et al. Extreme gradient boosting as a method for quantitative structure–activity relationships. J Chem Inf Model. 2016;56 (12):2353–2360. doi: 10.1021/acs.jcim.6b00591