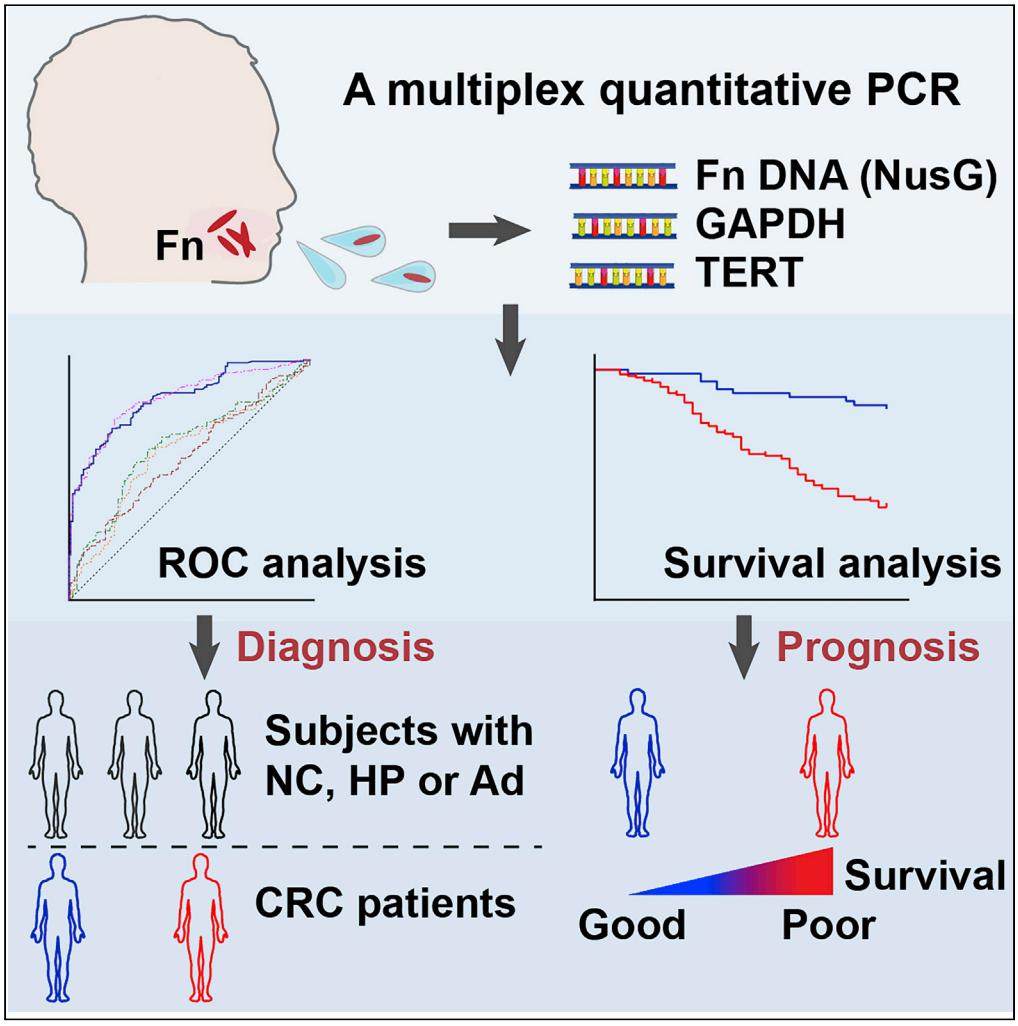
iScience

CellPress
OPEN ACCESS

Article

Salivary *Fusobacterium nucleatum* serves as a potential biomarker for colorectal cancer



Xin Zhang, Yaping Zhang, Xinru Gui, ..., Ziqi Shang, Yiwei Xin, Yi Zhang

yizhang@sdu.edu.cn

Highlights

Fusobacterium nucleatum DNA level is increased in saliva of colorectal cancer patients

Salivary *F. nucleatum* DNA is a biomarker for colorectal cancer diagnosis

Salivary *F. nucleatum* DNA is an independent prognostic factor

KEGG identified relationships to ECM-receptor interaction and focal adhesion pathways

Zhang et al., iScience *25*, 104203
May 20, 2022
https://doi.org/10.1016/j.isci.2022.104203

iScience

CellPress
OPEN ACCESS

Article

Salivary *Fusobacterium nucleatum* serves as a potential biomarker for colorectal cancer

Xin Zhang,[1,2] Yaping Zhang,[1] Xinru Gui,[1] Yanli Zhang,[3] Zhenhong Zhang,[2,4] Wendan Chen,[1] Xiaowei Zhang,[2,4] Yanxiang Wang,[2,4] Mengjiao Zhang,[1] Ziqi Shang,[1] Yiwei Xin,[1] and Yi Zhang[1,2,5,*]

SUMMARY

***Fusobacterium nucleatum* (Fn) is primarily colonized in the oral cavity. Recently, Fn has been closely associated with the tumorigenesis of colorectal cancer (CRC). Here, we showed that the relative level of Fn DNA was increased in the saliva of the CRC group compared with the normal colonoscopy, hyperplastic polyp, and adenoma groups. Receiver operating characteristic curve analysis illustrated that Fn DNA was superior to carcinoembryonic antigen and carbohydrate antigen 19-9 in CRC diagnosis. Moreover, levels of Fn DNA were associated with the overall survival and disease-free survival of CRC patients, which was an independent factor for prognostic prediction. Transcriptome sequencing identified 1,287 differentially expressed mRNAs in tumor tissues between CRC patients with high-Fn and low-Fn infection. Kyoto encyclopedia of genes and genomes analysis showed that ECM-receptor interaction and focal adhesion were the top two significant pathways. Overall, salivary Fn DNA may be a noninvasive diagnostic and prognostic biomarker for CRC patients.**

INTRODUCTION

Colorectal cancer (CRC) ranks fourth in terms of incidence, and it is also the second leading cause of cancer-related deaths worldwide (Bray et al., 2018). CRC is usually diagnosed at advanced stages because of the late appearance of symptoms, and thus it lacks an effective treatment option, leading to a poor prognosis (Kanth and Inadomi, 2021). Diagnostic strategies currently available for CRC patients rely more on colonoscopy, an invasive, uncomfortable, and potentially harmful procedure, and some heterogeneous tumors can be neglected by such procedure (Ladabaum et al., 2020; Longstreth et al., 2020). The traditional serum tumor biomarkers, such as carcinoembryonic antigen (CEA), carbohydrate antigen (CA) 19-9, and so on, have limited sensitivity and specificity, leading to false negative or overdiagnosis (Rao et al., 2021). Meanwhile, some patients detected at an early stage often suffer consequent overtreatment or do not receive timely treatment, resulting in shortened overall survival (OS) because of no reliable factors for its prognostic prediction (Quasar Collaborative Group, 2007; Sinicrope et al., 2021). Therefore, it is urgently necessary to identify noninvasive markers to improve the early diagnosis and prognostic prediction of CRC.

As a gram-negative, nonspore-forming anaerobic bacterial strain, *F. nucleatum* (Fn) has species-specific reservoirs in the human oral cavity and gastrointestinal tract (Cho and Blaser, 2012; Signat et al., 2011). Fn is first identified as an adhesive and symbiotic bacterial strain, which is well-known as a periodontal pathogen (Lamont et al., 2018). Subsequently, whole-genome sequencing and transcriptome sequencing analyses provided the earliest evidence that Fn is enriched in CRC tissues (Castellarin et al., 2012). In recent years, increasing evidence indicates that Fn plays an important role in the carcinogenesis of CRC. For example, Fn has been shown to increase the proliferation and invasive capabilities of CRC cells by activating TLR4 signaling to NF-κB (Yang et al., 2017). Moreover, Fn even can promote the development of colonic neoplasia under the treatment of some chemotherapeutic drugs by inducing autophagy (Yu et al., 2017). Some oncogenic microRNAs are abnormally increased in CRC with infection of Fn, indicating they play a crucial role in Fn induced CRC (Feng et al., 2019; Guo et al., 2020; Yang et al., 2017; Yu et al., 2017). Several studies demonstrate that Fn functions as an inhibitor of antitumor T cell-mediated adaptive immunity in the CRC microenvironment (Borowsky et al., 2021; Mima et al., 2015). It has also been reported that Fn may encode virulence factors, such as FadA, Fap2, and MORN2 proteins, to induce CRC development (Ranjbar et al., 2021). Consistent with these lines of experimental evidence, clinical studies demonstrate that the abundance of Fn in CRC tissues is significantly increased compared with the adjacent normal

[1]Department of Clinical Laboratory, Qilu Hospital of Shandong University, Jinan 250012, China

[2]Molecular Diagnostics Center, Qilu Hospital of Shandong University, Jinan 250012, China

[3]Department of Clinical Laboratory, Shandong Provincial Third Hospital, Jinan 250031, China

[4]Jinan AXZE Medical Test Laboratory, Jinan 250102, China

[5]Lead contact

*Correspondence: yizhang@sdu.edu.cn

https://doi.org/10.1016/j.isci.2022.104203

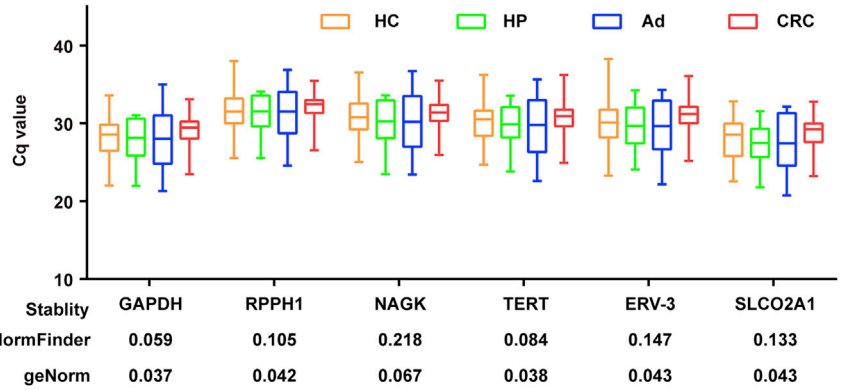


Figure 1. The expression and evaluation of six reference genes in saliva samples among HC, HP, Ad, and CRC patients.

tissues, and a high prevalence of Fn is associated with advanced stage, metastasis, recurrence, and short OS of CRC patients (Eisele et al., 2021; Mima et al., 2016; Serna et al., 2020; Yamamoto et al., 2021).Therefore, Fn, as a potential oncobacterium, may be helpful for the early detection and prognostic prediction of CRC.

Given that Fn colonizing CRC tissues may originate from the oral cavity (Abed et al., 2020; Komiya et al., 2019) and saliva is easy to acquire with noninvasive and painless approaches, we developed a multiplex quantitative real-time PCR (qPCR) method to detect the levels of Fn DNA (*NusG*) and human reference genes in saliva. Moreover, the relative levels of salivary Fn DNA were evaluated in the training subjects with normal colonoscopy (HC), hyperplastic polyp (HP), adenoma (Ad), and CRC. Finally, an independent test set was used to validate its clinical significance, and the potential functional mechanism *in vivo* was also explored. To the best of our knowledge, we, for the first time, reported that salivary Fn DNA could be used as a biomarker for the diagnosis and prognosis of CRC.

RESULTS

Evaluation of the candidate reference genes in saliva

A total of 80 cases were randomly selected from the training cohort (n = 20 for each group) to evaluate the expressions of candidate reference genes in saliva. Figure 1 shows that there were no significant differences in the expressions of *GAPDH*, *RPPH1*, *NAGK*, *TERT*, *ERV-3*, and *SLC O 2A1* at the DNA level among HC, HP, Ad, and CRC groups. Moreover, we employed NormFinder and geNorm algorithms to evaluate the variable stability of candidate reference genes. Both the programs identified *GAPDH* as the most stably expressed reference gene with the lowest stability scores, and the combination of *GAPDH* and *TERT* was selected as the most stable pair for the evaluation of salivary DNA (Figure 1).

Levels of salivary Fn DNA in the training cohort

We developed a multiplex qPCR method for the simultaneous detection of Fn DNA and reference genes (*GAPDH* and *TERT*) in saliva, and the standard curves showed good linearity between Cq values and the log of sample concentrations (all $R^2 > 0.99$; Figure 2A). The amplification efficiencies of Fn DNA, *GAPDH*, and *TERT* were suitable. The relative level of Fn DNA was calculated as the ratio of Fn DNA (*NusG* gene) level to the geometric mean of *GAPDH* and *TERT* levels. Figure 2B shows that there was a dramatic difference in the level of salivary Fn DNA among HC, HP, Ad, and CRC patients in the training group. Moreover, the level of Fn DNA was independent of age and sex (Table S1). Further post hoc multiple comparisons showed that the level of salivary Fn DNA was extremely elevated in the CRC group compared with the HC, HP, and Ad groups, and it was also increased in the Ad group compared with the HC and HP groups (Figure 2B). In CRC patients, the level of salivary Fn DNA was increased with the increase of the TNM stage (Figure 2C).

Relationship between salivary Fn DNA and clinicopathologic characteristics in CRC patients

Table 1 shows that the level of salivary Fn DNA was significantly associated with tumor location, regional lymph node metastasis, distant metastasis, and CA19-9 levels. No significant differences were observed

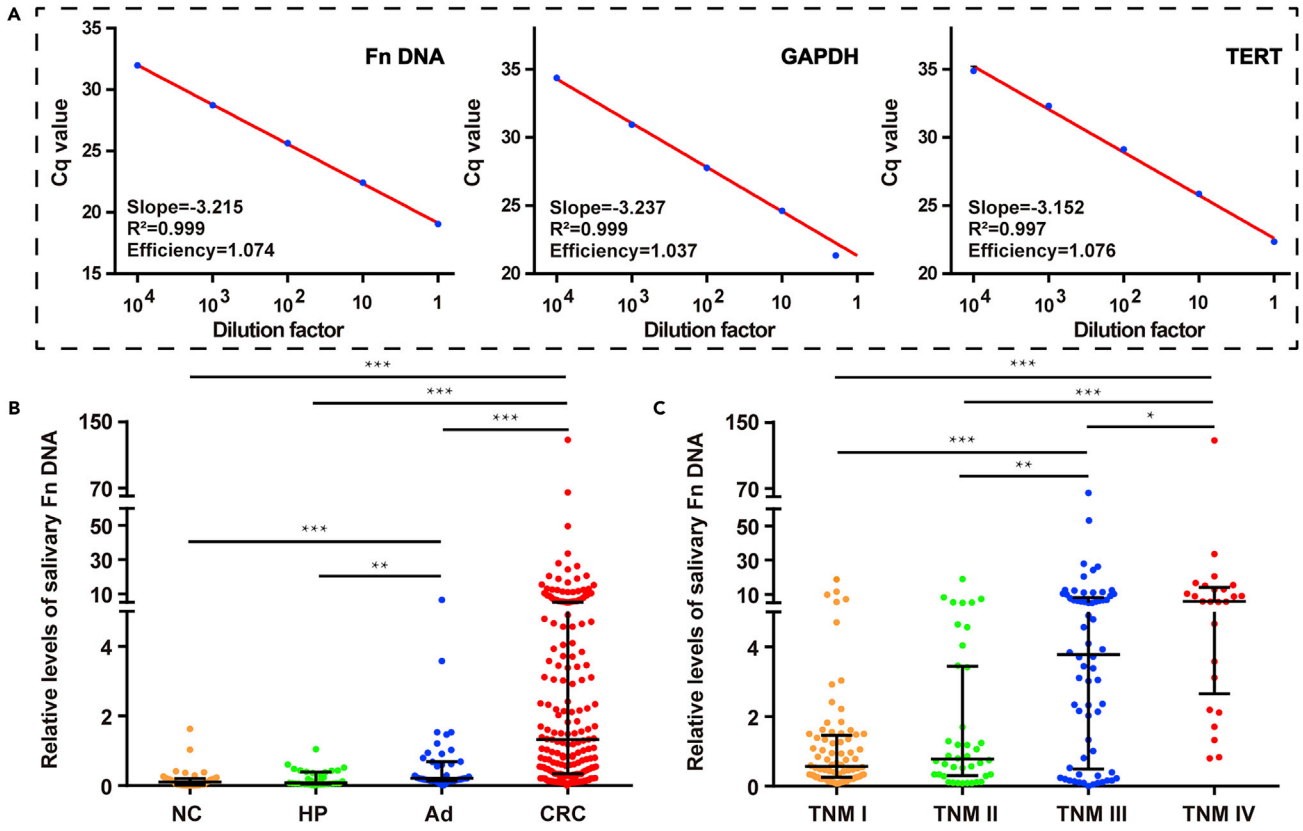


Figure 2. Relative levels of salivary Fn DNA in subjects in the training cohort

(A) Standard curve for Fn DNA, GAPDH, and TERT amplifications. Standard curves were generated 10-fold serial dilutions of DNA from DNA samples containing Fn DNA, GAPDH, and TERT. Efficiency of qPCR amplification is calculated from the slope of the standard curve.

(B) Comparison analysis of salivary Fn DNA levels among NC (n = 41), HP (n = 33), Ad (n = 43), and CRC (n = 207).

(C) Comparison analysis of salivary Fn DNA levels among CRC patients with different TNM stages. The relative level of Fn DNA was calculated as the ratio of NusG gene level to the geometric mean of GAPDH and TERT levels. *p < 0.05, **p < 0.01, ***p < 0.001 (Mann-Whitney U test). Data represent the median (IQR).

when CRC cases were stratified by age, gender, tumor size, tumor differentiation, local invasion, and CEA levels. Moreover, CRC cases with regional lymph node metastasis, distant metastasis, and high CA19-9 levels showed increased levels of salivary Fn DNA.

Diagnostic performance of salivary Fn DNA for CRC

ROC analysis illustrated that salivary Fn DNA could be a potential biomarker for distinguishing CRC (n = 207) from others (n = 117) with an area under the curve (AUC) of 0.841 (95% CI (CI) 0.797 to 0.879) (Figure 3A). When Youden's index reached a maximum value, the corresponding cutoff value of salivary Fn DNA was 0.437 for the diagnosis of CRC, achieving a sensitivity of 71.5% and a specificity of 82.1%.

To better evaluate the diagnostic performance of salivary Fn DNA, we also detected the serum levels of CEA and CA19-9, the most commonly used noninvasive biomarkers for the diagnosis of CRC in clinical practice. As expected, the serum levels of CEA and CA19-9 in CRC patients were significantly higher in the CRC group compared with the HC, HP, and Ad groups (Table S1). However, the AUC values of CEA, CA19-9, and their combination for the diagnosis of CRC were dramatically lower compared with Fn (Figure 3A). Besides, the AUC of the combination of Fn DNA, CEA, and CA19-9 was not significantly different compared with salivary Fn DNA only (Figure 3A). When the combination of salivary Fn DNA, CEA, and CA19-9 was used, the AUC was not significantly elevated compared with Fn alone. Table S2 shows the diagnostic performance characteristics of Fn DNA, CEA, and CA19-9 for CRC, such as sensitivity, specificity, positive predictive value, and negative predictive value. More importantly, ROC analysis was performed

Table 1. Associations between salivary Fn DNA levels and clinicopathologic characteristics

Parameters	Case	Levels of salivary Fn DNA[a]	*p* value[b]
Age			0.234
<63	102	1.343 (0.309–4.659)	
≥63 (median)	105	1.266 (0.533–6.438)	
Gender			0.567
Male	111	1.324 (0.294–5.134)	
Female	96	1.327 (0.409–5.318)	
Tumor location			0.017
Colon	95	1.828 (0.476–7.135)	
Rectum	112	0.945 (0.327–3.538)	
Tumor size			0.869
<4 cm	68	1.422 (0.341–3.995)	
≥4 cm	139	1.297 (0.337–5.584)	
Differentiation			0.354
Well	45	3.021 (0.388–6.874)	
Moderate	118	1.074 (0.342–3.959)	
Poor	44	1.386 (0.217–7.743)	
Local invasion			0.161
T1-T2	78	0.819 (0.236–5.179)	
T3-T4	129	1.608 (0.475–5.197)	
Regional lymph nodes metastasis			<0.001
No	116	0.651 (0.270–1.504)	
Yes	91	4.662 (1.329–10.180)	
Distant metastasis			<0.001
No	182	1.022 (0.299–4.055)	
Yes	25	5.856 (2.654–14.110)	
CEA levels			0.321
<5 ng/mL	161	1.266 (0.337–5.669)	
≥5 ng/mL	46	1.360 (0.414–2.267)	
C19-9 levels			0.025
<39 U/mL	185	1.181 (0.316–4.846)	
≥39 U/mL	22	2.693 (1.153–6.587)	

[a]The data are shown as median (IQR).
[b]Compared with *Mann-Whitney U* test.

for CRC in the CEA-negative group or CA19-9-negative group, and the AUC for salivary Fn DNA was 0.823 (95% CI, 0.773–0.866) and 0.831 (95% CI, 0.784–0.871), respectively (Figures 3B and 3C). A similar situation occurred in the group negative for both CEA and CA19-9, showing an AUC of 0.831 (95% CI, 0.784–0.871; Figure 3D).

Because the level of Fn DNA was increased in Ad patients, we assessed its diagnostic performance for Ad. Figure 3E shows that Fn DNA only yielded an AUC of 0.672 (95% CI, 0.579–0.756) when distinguishing Ad (n = 43) from HP and NC (n = 74). Besides, there were no significant differences in the diagnosis of Ad among Fn DNA, CEA, CA19-9, and their combination.

To examine the capacity of salivary Fn DNA to predict metastasis, CRC patients were classified into the metastatic group (n = 91) and nonmetastatic group (n = 116). ROC analysis indicated the usefulness of salivary Fn DNA as a predictive marker for metastasis in CRC, with an AUC of 0.763 (95% CI, 0.699 to 0.819), which was visibly larger compared with CEA and CA19-9 individually or their combination (Figure 3F).

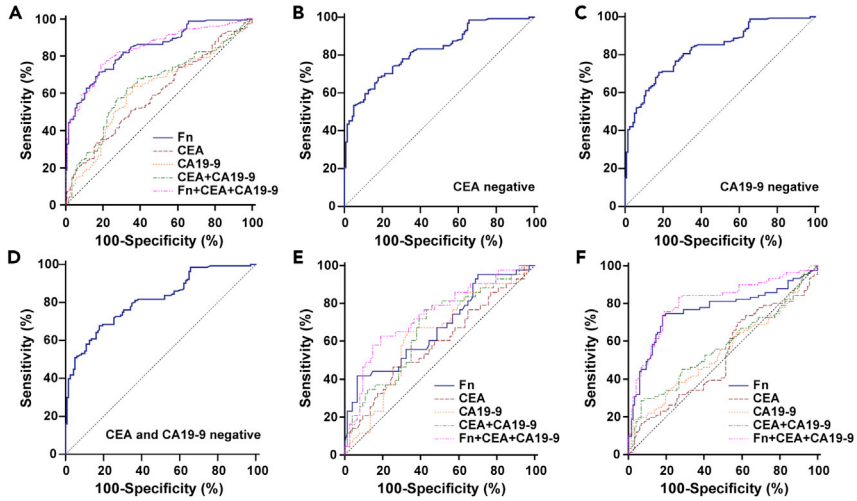


Figure 3. Diagnostic performance of Fn DNA, CEA, and CA19-9 in the training set

(A) ROC curve analysis for the detection of CRC.

(B) ROC curve analysis for the detection of CRC with CEA negative using Fn DNA.

(C) ROC curve analysis for the detection of CRC with CA19-9 negative using Fn DNA.

(D) ROC curve analysis for the detection of CRC with both CEA and CA19-9 negative using Fn DNA.

(E) ROC curve analysis for discriminating Ad from NC and HP.

(F) ROC curve analysis for discriminating metastatic CRC patients from nonmetastatic ones.

Correlation between salivary Fn DNA and CRC patients' survival

Based on the aforementioned cutoff value (0.437) for the diagnosis of CRC, CRC patients were divided into low-Fn and high-Fn DNA groups. Kaplan-Meier curves showed that patients with high levels of Fn DNA had poorer OS and disease-free survival (DFS) than those with low levels of Fn DNA (Figures 4A and 4D). However, we observed that there were no significant differences in OS and DFS between patients who were negative or positive for CEA and CA19-9 (Figures 4B, 4C, 4E, and 4F).

We further investigated the prognostic performance of Fn using the Cox proportional-hazards model. The univariate Cox analyses revealed that the OS was significantly associated with lymph node metastasis, distant metastasis, and Fn, whereas DFS was related to tumor location, lymph node metastasis, distant metastasis, and Fn (Table 2). Multivariate Cox-regression analysis demonstrated that only Fn and lymph node metastasis were the independent prognostic factors in CRC patients (Table 2).

Evaluation of salivary Fn DNA in the test cohort

Similar to the training cohort, the level of salivary Fn DNA in the CRC group were markedly higher compared with the NC, HP, and Ad groups in the test cohort ($p < 0.001$), whereas no significant differences were observed among NC, HP, and Ad groups ($p > 0.05$) (Figure 5A). The ROC curve also showed that salivary Fn DNA could strongly discriminate against CRC from NC, HP, and Ad, with an AUC of 0.860 (95%IC 0.774–0.922) (Figure 5B). Using a cutoff value of 0.437 obtained from the training set, the sensitivity, specificity, positive predictive value, and negative predictive value for CRC in the test cohort were 86.7%, 67.2%, 54.2%, and 91.8%, respectively.

We tested whether the levels of Fn DNA in the saliva of CRC patients could reflect its expression in corresponding tumor tissues. Figure 5C shows that the levels of salivary Fn DNA were significantly correlated with their levels in tumor tissues ($r > 0.5$, $p < 0.001$). The positive rates of Fn DNA in saliva and tumor tissues were 93.3% (28/30) and 76.7% (23/30), respectively. However, the positive rate in matched blood specimens was only 3.3% (1/30).

Fn-associated transcriptomic profile in CRC tissues

To explore the molecular mechanism by which Fn might contribute to the development of CRC, RNA-Seq was performed in 18 CRC tumor tissues collected from the test cohort. Among them, 10 cases demonstrated high levels of Fn DNA, and eight cases showed low levels of Fn DNA in both saliva and tumor

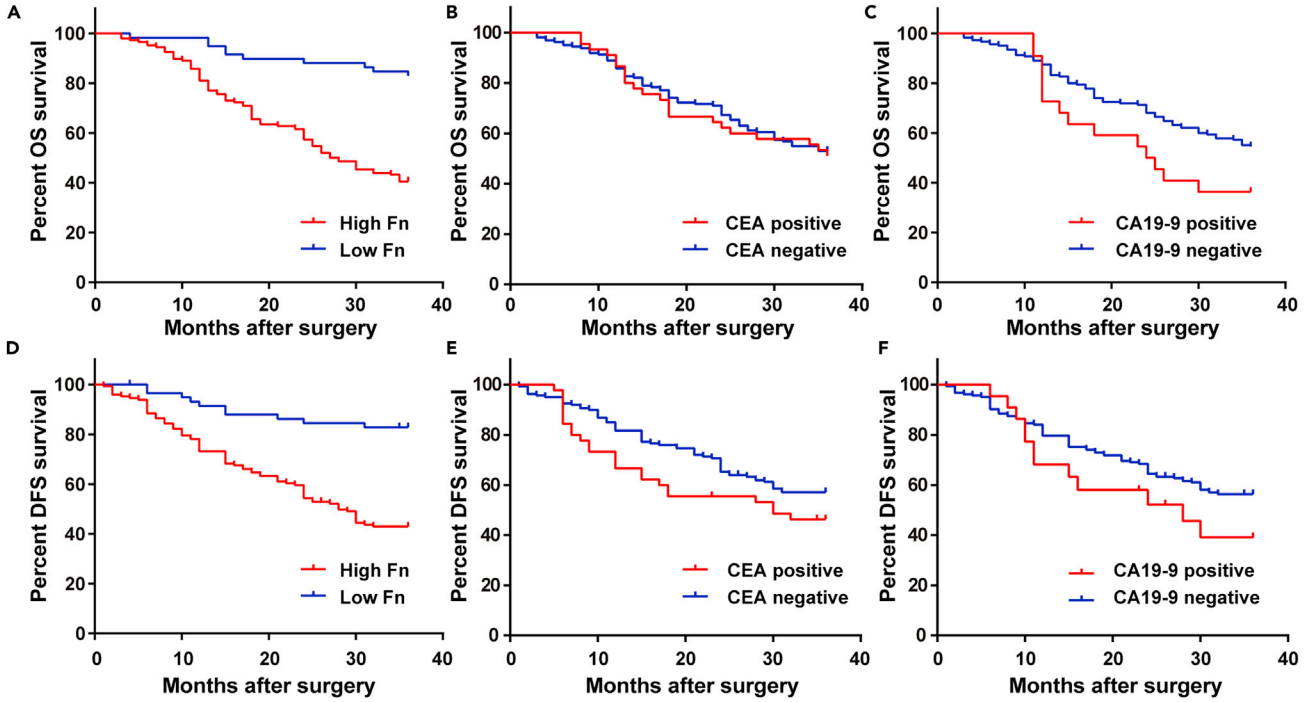


Figure 4. Survival analysis for salivary Fn DNA and traditional serum markers, CEA and CA19-9, in CRC patients

(A–F). Kaplan-Meier curves for OS in CRC patients stratified according to levels of salivary Fn DNA (A), CEA (B), or CA19-9(C); Kaplan-Meier curves for OS in CRC patients stratified according to levels of salivary Fn DNA (D), CEA (E), or CA19-9(F); CRC patients were classified as high-Fn and low-Fn DNA subgroups according to the cutoff value (0.109) or as CEA or CA19-9 negative and positive subgroups based on the standard cutoff values (5 ng/mL or 39 U/mL, respectively).

tissues. Principal component analysis (PCA) was carried out on transcriptomic profiling using ggord to investigate the distribution patterns in patients with high and low levels of Fn DNA. Figure 5D shows that the high-Fn and low-Fn DNA groups were separated into two parts, and the transcriptomic expression profile of CRC patients with high Fn infection was distinguished from those with low Fn infection.

A total of 1,287 differentially expressed mRNAs were identified between the high-Fn and low-Fn DNA groups. Among them, 619 were upregulated and 668 were downregulated in the high-Fn group (Figure 5E). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analyses identified 59 significant pathways based on the differentially expressed genes (Table S3), of which ECM-receptor interaction and focal adhesion were the top two significant biological functions. Moreover, Figure 5F presents the top 10 representative enriched KEGG pathways.

DISCUSSION

In the present study, we obtained several findings. First, we found that the levels of salivary Fn DNA were significantly higher in CRC patients using a multiplex qPCR. Second, compared with traditional serum tumor markers, salivary Fn DNA was a more appropriate biomarker for the diagnosis and prognostic prediction of CRC. Third, salivary Fn DNA could reflect its infection status in CRC tissue and might contribute to the development of CRC through ECM-receptor interaction and focal adhesion pathways.

It has become increasingly clear that the abundance of intratumoral Fn is associated with the pathogenesis of CRC (Huang et al., 2020; Wang et al., 2021). Fn DNA is predominantly found in CRC tissues while not in adjacent normal tissues, especially in advanced tumors (Yang et al., 2017). Several studies have also shown that Fn DNA is more frequently detected in the feces of CRC patients compared with healthy controls (Liang et al., 2021; Zhang et al., 2019). Consistent with those findings, our data from both the training and test cohorts showed that the level of Fn DNA was significantly increased in the saliva of CRC patients. Notably, elevated levels of salivary Fn DNA were also observed in patients with colorectal adenoma. This finding might be attributed to the fact

Table 2. COX analysis of prognostic factors predicting disease free survival and overall survival in CRC patients

	Univariate analysis		Multivariate analysis	
Parameters	HR (95%IC)	*p* value	HR (95%IC)	*p* value
Disease free survival				
Gender (Male VS Female)	0.958 (0.633–1.449)	0.838		
Age (<63 years VS ≥63 years)	0.799 (0.527–1.210)	0.289		
Location (rectal VS colon)	1.391 (0.920–2.103)	0.118		
Differentiation (Well VS Moderate VS Poor)	0.839 (0.609–1.156)	0.284		
Tumor size (<4 cm VS ≥4 cm)	1.425 (0.893–2.273)	0.138		
T stage (T1+T2 VS T3+T4)	1.509 (0.965–2.359)	0.071		
Lymph nodes metastasis (Negative VS Positive)	2.929 (1.907–4.499)	<0.001	2.444 (1.537–3.886)	<0.001
Distant metastasis (Negative VS Positive)	2.780 (1.630–4.740)	<0.001	1.237 (0.695–2.203)	0.469
CEA (Negative VS Positive)	1.501 (0.946–2.381)	0.085		
CA19-9(Negative VS Positive)	1.557 (0.846–2.864)	0.155		
Fn (Negative VS Positive)	4.301 (2.225–8.312)	<0.001	3.648 (1.867–7.129)	<0.001
Overall survival				
Gender (Male VS Female)	0.895 (0.602–1.330)	0.582		
Age (<63 years VS ≥63 years)	0.958 (0.644–1.423)	0.830		
Location (rectal VS colon)	1.539 (1.035–2.288)	0.033	1.438 (0.965–2.144)	0.075
Differentiation (Well VS Moderate VS Poor)	0.904 (0.662–1.236)	0.529		
Tumor size (<4 cm VS ≥4 cm)	1.092 (0.711–1.679)	0.687		
T stage (T1+T2 VS T3+T4)	1.378 (0.903–2.102)	0.137		
Lymph nodes metastasis (Negative VS Positive)	3.295 (2.168–5.007)	<0.001	2.805 (1.785–4.406)	<0.001
Distant metastasis (Negative VS Positive)	2.942 (1.794–4.826)	<0.001	1.155 (0.675–1.977)	0.599
CEA (Negative VS Positive)	1.102 (0.691–1.758)	0.684		
CA19-9 (Negative VS Positive)	1.682 (0.954–2.966)	0.072		
Fn (Negative VS Positive)	4.788 (2.486–9.221)	<0.001	4.001 (2.055–7.792)	<0.001

HR, Hazard ratio; CI, Confidence interval.

that Fn infection is an early event in CRC development, and its influence can be observed throughout all stages of colorectal neoplasia development in normal-adenoma-carcinoma sequence (Flanagan et al., 2014; Nakatsu et al., 2015). Therefore, it would be helpful for screening early-stage lesions of CRC with noninvasive detection of this bacterial strain. Therefore, we hypothesized that salivary Fn might serve as a potential noninvasive biomarker for the early detection of CRC.

At present, CEA and CA19-9 are the most commonly used noninvasive tumor markers for CRC in clinical practice. However, both of them show limited diagnostic capability. We found that the sensitivities of CEA and CA19-9 were only 22.2 and 10.1%, respectively, and their AUCs were less than 0.7 in distinguishing CRC from others. When at a similar specificity, salivary Fn DNA presented an improved sensitivity, which was associated with a markedly larger AUC. In addition, Fn DNA had an equivalent diagnostic capacity in CEA-negative or CA19-9-negative individuals. These findings indicated the great significance of salivary Fn DNA for the diagnosis of CRC even when conventional markers are negative, leading to an accurate diagnosis of CRC.

In the present study, CRC patients with a high abundance of salivary Fn were associated with regional lymph node metastasis and distant metastasis, which represent a more biologically aggressive cancer subtype. ROC analysis also showed that Fn DNA had some capacities in distinguishing CRC metastasis, which might help optimize the therapeutic decision-making process of clinicians. Especially for T1 carcinomas, colonoscopy can be used instead of more traumatic surgery if there is no risk of metastasis (Kudo et al., 2021; Saitoh et al., 2016).

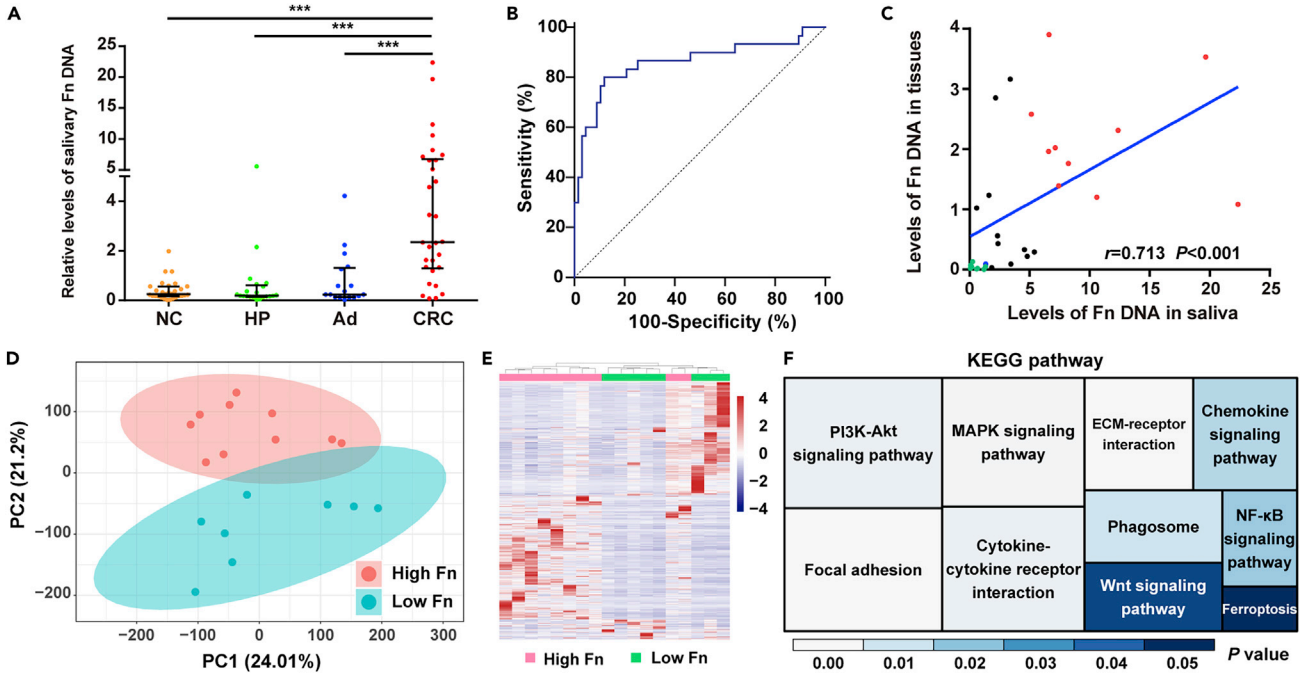


Figure 5. Evaluation of salivary Fn DNA in test subjects and its functional analysis

(A) Comparison analysis of salivary Fn DNA levels among NC (n = 29), HP (n = 21), Ad (n = 17), and CRC (n = 30). ***p < 0.001 (Mann-Whitney U test). Data represent the median (IQR).

(B) ROC curve analysis for the detection of CRC.

(C) The correlation analysis of Fn DNA between saliva and tumor tissues in CRC patients. Spearman test.

(D) PCA for CRC patients with high-Fn DNA levels in both saliva and tumor tissues and those with low-Fn DNA levels in both saliva and tumor tissues.

(E) Heatmap analysis of differentially expressed mRNAs between CRC patients with high-Fn and low-Fn DNA levels in both saliva and tumor tissues.

(F) Treemap of 10 enriched KEGG pathways for the differentially expressed genes shown in (E).

Existing data support that there is a positive relationship between the level of tissue Fn and worse prognosis in gastroenterological cancers, particularly CRC (Mima et al., 2016; Salvucci et al., 2021; Yamaoka et al., 2018), esophageal cancer (Yamamura et al., 2016), and gastric cancer (Boehm et al., 2020). A recent meta-analysis also predicts that high levels of Fn in tumor tissues suggest a poor prognosis (Huangfu et al., 2021). In this research, CRC patients were divided into low-Fn and high-Fn DNA groups according to the cutoff value for CRC diagnosis. Consistent with the aforementioned findings, our data showed that the high levels of Fn DNA in saliva were associated with poor OS and DFS of CRC patients, which was an independent factor for prognostic prediction. Besides, the traditional tumor biomarkers, CEA and CA19-9, had no significant application value in prognostic prediction. Therefore, we provided a potential noninvasive biomarker for the prognostic prediction of CRC.

A previous study has indicated that CRC-associated Fn comes from the oral microbial community, and oral fusobacteria reach the colon tumor via a hematogenous route (Abed et al., 2020). To test this hypothesis, we examined the amount of Fn in saliva, tumor tissue, and blood samples of CRC patients. Our data showed that the level of Fn in the saliva was positively associated with its abundance in CRC tissue, whereas Fn DNA was only detected in one of 30 serum samples. Therefore, we thought that the level of Fn in the oral cavity reflected its tissue infection status in CRC patients. Moreover, Fn entering the circulatory system might be only a short transient, because no growth and reproduction occurred there. In addition, we detected Fn DNA in 93.3 and 76.7% of the analyzed saliva and tumor tissue samples, respectively, suggesting that salivary Fn had higher sensitivity compared with tissue and blood detections.

Fn has been reported to be involved in the proliferation and migration of CRC via orchestrating a molecular network of some signaling pathways (Hu et al., 2021; Kong et al., 2021; Yang et al., 2017). However, these mechanistic insights most come from cell lines, which may not truly reflect its function in the human body. In the present study, we employed 18 CRC tissues with high or low Fn infection. Through KEGG analysis, 59 significant pathways were identified, some of which have already been demonstrated in the development

of CRC. For example, ECM-receptor interaction and focal adhesion are the top two significant KEGG pathways. These two signaling pathways are the possible molecular mechanisms underlying the metastasis of CRC (Machackova et al., 2020). Our clinical data also showed that there was a close relationship between the level of salivary Fn and regional lymph node metastasis or distant metastasis. The other representative enriched KEGG pathways, such as the Wnt signaling pathway, MAPK signaling pathway, cytokine-cytokine receptor interaction, and so on, have also been reported in some studies. Fn confers the CRC progression through activation of the Wnt/β-catenin signaling pathway (Li et al., 2021). AKT/MAPK and NF-κB signaling pathways have been found to be activated after Fn infection (Kang et al., 2019). Similarly, cytokine-cytokine receptor interaction has also been reported in Fn-positive esophageal cancers. Moreover, Fn-induced chemokines, such as CCL20, may contribute to the development and metastasis of tumors (Xu et al., 2021a; Yamamura et al., 2016). Phagosome was found in our enriched KEGG pathways. This might be attributed to the fact that Fn can induce autophagy to promote the metastasis and chemoresistance of CRC cells (Chen et al., 2020; Yu et al., 2017). Interestingly, Fn is involved in the regulation of ferroptosis, which is a new type of programmed cell death (Xu et al., 2021b). However, further studies are required to elucidate the mechanism by which Fn affects tumor behavior by modulating ferroptosis.

In conclusion, salivary Fn DNA could be used as a potential biomarker for the detection of CRC, and we also revealed the potential mechanism of Fn in the development of CRC. Compared with currently available detection methods, salivary Fn DNA showed certain advantages, such as its higher sensitivity and safety.

Limitations of the study

Although the findings are promising, there are some limitations in our current work. First, CRC patients usually present with a decreased autoimmune function (Di Caro et al., 2014), and Fn may have been altered. However, the clinical information has not included the indexes that reflect the immune status. Second, although the levels of salivary Fn DNA were increased in Ad patients, it was not suitable for the diagnosis of Ad. Moreover, whether Ad patients with high levels of Fn have a high risk for CRC still remains largely unknown. This limitation is unavoidable, as the Ad patients usually take timely treatment. Third, it is unclear to what extent Fn can be affected in patients with poor survival who have a radical operation and would therefore benefit most from Fn DNA detection. Fourth, in our limited test cohort, the patients were not followed up for enough time. Besides, the prognostic significance of Fn will be further validated in our further work.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Subjects
- METHOD DETAILS
  - Sample preparation
  - DNA and RNA extraction
  - qPCR
  - CEA and CA19-9 assays
  - RNA sequence (RNA-Seq)
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104203.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (81972005, 82172339), Natural Science Foundation of Shandong Province (ZR2020MH238, ZR2021MH110), and Shandong Medical and Health Technology Development Project (2018WS327).

AUTHOR CONTRIBUTIONS

Y.P.Z., X.R.G., Y.L.Z., Z.H.Z., W.D.C., and X.W.Z. conducted the experiments. X.Z. and Y.X.W. contributed to the statistical analysis. M.J.Z., Z.Q.S., and Y.W.X. collected the clinical information of samples. X.Z. contributed to the study design and wrote the paper. Y.Z. contributed to study supervision and revising the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 8, 2021
Revised: February 14, 2022
Accepted: April 1, 2022
Published: May 20, 2022

REFERENCES

Abed, J., Maalouf, N., Manson, A.L., Earl, A.M., Parhi, L., Emgard, J.E.M., Klutstein, M., Tayeb, S., Almogy, G., Atlan, K.A., et al. (2020). Colon cancer-associated Fusobacterium nucleatum may originate from the oral cavity and reach colon tumors via the circulatory system. Front. Cell Infect. Microbiol. *10*, 400. https://doi.org/10.3389/fcimb.2020.00400.

Boehm, E.T., Thon, C., Kupcinskas, J., Steponaitiene, R., Skieceviciene, J., Canbay, A., Malfertheiner, P., and Link, A. (2020). Fusobacterium nucleatum is associated with worse prognosis in Lauren's diffuse type gastric cancer patients. Sci. Rep. *10*, 16240. https://doi.org/10.1038/s41598-020-73448-8.

Borowsky, J., Haruki, K., Lau, M.C., Dias Costa, A., Vayrynen, J.P., Ugai, T., Arima, K., da Silva, A., Felt, K.D., Zhao, M., et al. (2021). Association of Fusobacterium nucleatum with specific T-cell subsets in the colorectal carcinoma microenvironment. Clin. Cancer Res. *27*, 2816–2826. https://doi.org/10.1158/1078-0432.CCR-20-4009.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *68*, 394–424. https://doi.org/10.3322/caac.21492.

Castellarin, M., Warren, R.L., Freeman, J.D., Dreolini, L., Krzywinski, M., Strauss, J., Barnes, R., Watson, P., Allen-Vercoe, E., Moore, R.A., and Holt, R.A. (2012). Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res. *22*, 299–306. https://doi.org/10.1101/gr.126516.111.

Chen, Y., Chen, Y., Zhang, J., Cao, P., Su, W., Deng, Y., Zhan, N., Fu, X., Huang, Y., and Dong, W. (2020). Fusobacterium nucleatum promotes metastasis in colorectal cancer by activating autophagy signaling via the upregulation of CARD3 expression. Theranostics *10*, 323–339. https://doi.org/10.7150/thno.38870.

Cho, I., and Blaser, M.J. (2012). The human microbiome: at the interface of health and disease. Nat. Rev. Genet. *13*, 260–270. https://doi.org/10.1038/nrg3182.

Di Caro, G., Bergomas, F., Grizzi, F., Doni, A., Bianchi, P., Malesci, A., Laghi, L., Allavena, P., Mantovani, A., and Marchesi, F. (2014). Occurrence of tertiary lymphoid tissue is associated with T-cell infiltration and predicts better prognosis in early-stage colorectal cancers. Clin. Cancer Res. *20*, 2147–2158. https://doi.org/10.1158/1078-0432.CCR-13-2590.

Eisele, Y., Mallea, P.M., Gigic, B., Stephens, W.Z., Warby, C.A., Buhrke, K., Lin, T., Boehm, J., Schrotz-King, P., Hardikar, S., et al. (2021). Fusobacterium nucleatum and clinicopathologic features of colorectal cancer: results from the ColoCare study. Clin. Colorectal Cancer *20*, e165–e172. https://doi.org/10.1016/j.clcc.2021.02.007.

Feng, Y.Y., Zeng, D.Z., Tong, Y.N., Lu, X.X., Dun, G.D., Tang, B., Zhang, Z.J., Ye, X.L., Li, Q., Xie, J.P., and Mao, X.H. (2019). Alteration of microRNA-4474/4717 expression and CREB-binding protein in human colorectal cancer tissues infected with Fusobacterium nucleatum. PLoS One *14*, e0215088. https://doi.org/10.1371/journal.pone.0215088.

Flanagan, L., Schmid, J., Ebert, M., Soucek, P., Kunicka, T., Liska, V., Bruha, J., Neary, P., Dezeeuw, N., Tommasino, M., et al. (2014). Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. Eur. J. Clin. Microbiol. Infect. Dis. *33*, 1381–1390. https://doi.org/10.1007/s10096-014-2081-3.

Quasar Collaborative Group (2007). Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. Lancet *370*, 2020–2029. https://doi.org/10.1016/S0140-6736(07)61866-2.

Guo, S., Chen, J., Chen, F., Zeng, Q., Liu, W.L., and Zhang, G. (2020). Exosomes derived from Fusobacterium nucleatum-infected colorectal cancer cells facilitate tumour metastasis by selectively carrying miR-1246/92b-3p/27a-3p and CXCL16. Gut. https://doi.org/10.1136/gutjnl-2020-321187.

Hu, L., Liu, Y., Kong, X., Wu, R., Peng, Q., Zhang, Y., Zhou, L., and Duan, L. (2021). Fusobacterium nucleatum facilitates M2 macrophage polarization and colorectal carcinoma progression by activating TLR4/NF-kappaB/S100A9 cascade. Front. Immunol. *12*, 658681. https://doi.org/10.3389/fimmu.2021.658681.

Huang, X., Hong, X., Wang, J., Sun, T., Yu, T., Yu, Y., Fang, J., and Xiong, H. (2020). Metformin elicits antitumour effect by modulation of the gut microbiota and rescues Fusobacterium nucleatum-induced colorectal tumourigenesis. EBioMedicine *61*, 103037. https://doi.org/10.1016/j.ebiom.2020.103037.

Huangfu, S.C., Zhang, W.B., Zhang, H.R., Li, Y., Zhang, Y.R., Nie, J.L., Chu, X.D., Chen, C.S., Jiang, H.P., and Pan, J.H. (2021). Clinicopathological and prognostic significance of Fusobacterium nucleatum infection in colorectal cancer: a meta-analysis. J. Cancer *12*, 1583–1591. https://doi.org/10.7150/jca.50111.

Kang, W., Jia, Z., Tang, D., Zhang, Z., Gao, H., He, K., and Feng, Q. (2019). Fusobacterium nucleatum facilitates apoptosis, ROS generation, and inflammatory cytokine production by activating AKT/MAPK and NF-kappaB signaling pathways in human gingival fibroblasts. Oxid Med. Cell Longev. *2019*, 1681972. https://doi.org/10.1155/2019/1681972.

Kanth, P., and Inadomi, J.M. (2021). Screening and prevention of colorectal cancer. BMJ *374*, n1855. https://doi.org/10.1136/bmj.n1855.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360. https://doi.org/10.1038/nmeth.3317.

Komiya, Y., Shimomura, Y., Higurashi, T., Sugi, Y., Arimoto, J., Umezawa, S., Uchiyama, S., Matsumoto, M., and Nakajima, A. (2019). Patients with colorectal cancer have identical strains of Fusobacterium nucleatum in their colorectal cancer and oral cavity. Gut *68*, 1335–1337. https://doi.org/10.1136/gutjnl-2018-316661.

Kong, C., Yan, X., Zhu, Y., Zhu, H., Luo, Y., Liu, P., Ferrandon, S., Kalady, M.F., Gao, R., He, J., et al. (2021). Fusobacterium nucleatum promotes the development of colorectal cancer by activating a cytochrome P450/Epoxyoctadecenoic acid Axis via TLR4/Keap1/NRF2 signaling. Cancer Res. *81*, 4485–4498. https://doi.org/10.1158/0008-5472.CAN-21-0453.

Kudo, S.E., Ichimasa, K., Villard, B., Mori, Y., Misawa, M., Saito, S., Hotta, K., Saito, Y., Matsuda, T., Yamada, K., et al. (2021). Artificial intelligence system to determine risk of T1 colorectal cancer metastasis to lymph node.

Gastroenterology *160*, 1075–1084.e2. https://doi.org/10.1053/j.gastro.2020.09.027.

Ladabaum, U., Dominitz, J.A., Kahi, C., and Schoen, R.E. (2020). Strategies for colorectal cancer screening. Gastroenterology *158*, 418–432. https://doi.org/10.1053/j.gastro.2019.06.043.

Lamont, R.J., Koo, H., and Hajishengallis, G. (2018). The oral microbiota: dynamic communities and host interactions. Nat. Rev. Microbiol. *16*, 745–759. https://doi.org/10.1038/s41579-018-0089-x.

Li, X., Huang, J., Yu, T., Fang, X., Lou, L., Xin, S., Ji, L., Jiang, F., and Lou, Y. (2021). Fusobacterium nucleatum promotes the progression of colorectal cancer through cdk5-activated Wnt/beta-catenin signaling. Front. Microbiol. *11*, 545251. https://doi.org/10.3389/fmicb.2020.545251.

Liang, J.Q., Wong, S.H., Szeto, C.H., Chu, E.S., Lau, H.C., Chen, Y., Fang, J., Yu, J., and Sung, J.J. (2021). Fecal microbial DNA markers serve for screening colorectal neoplasm in asymptomatic subjects. J. Gastroenterol. Hepatol. *36*, 1035–1043. https://doi.org/10.1111/jgh.15171.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

Longstreth, G.F., Anderson, D.S., Zisook, D.S., Shi, J.M., and Lin, J.C. (2020). Low rate of cancer detection by colonoscopy in asymptomatic, average-risk subjects with negative results from fecal immunochemical tests. Clin. Gastroenterol. Hepatol. *18*, 2929–2936.e1. https://doi.org/10.1016/j.cgh.2020.01.029.

Machackova, T., Vychytilova-Faltejskova, P., Souckova, K., Trachtova, K., Brchnelova, D., Svoboda, M., Kiss, I., Prochazka, V., Kala, Z., and Slaby, O. (2020). MiR-215-5p reduces liver metastasis in an experimental model of colorectal cancer through regulation of ECM-receptor interactions and focal adhesion. Cancers (Basel) *12*. https://doi.org/10.3390/cancers12123518.

Mima, K., Nishihara, R., Qian, Z.R., Cao, Y., Sukawa, Y., Nowak, J.A., Yang, J., Dou, R., Masugi, Y., Song, M., et al. (2016). Fusobacterium nucleatum in colorectal carcinoma tissue and patient prognosis. Gut *65*, 1973–1980. https://doi.org/10.1136/gutjnl-2015-310101.

Mima, K., Sukawa, Y., Nishihara, R., Qian, Z.R., Yamauchi, M., Inamura, K., Kim, S.A., Masuda, A., Nowak, J.A., Nosho, K., et al. (2015). Fusobacterium nucleatum and T cells in colorectal carcinoma. JAMA Oncol. *1*, 653–661. https://doi.org/10.1001/jamaoncol.2015.1377.

Nakatsu, G., Li, X., Zhou, H., Sheng, J., Wong, S.H., Wu, W.K., Ng, S.C., Tsoi, H., Dong, Y., Zhang, N., et al. (2015). Gut mucosal microbiome across stages of colorectal carcinogenesis. Nat. Commun. *6*, 8727. https://doi.org/10.1038/ncomms9727.

Ranjbar, M., Salehi, R., Haghjooy Javanmard, S., Rafiee, L., Faraji, H., Jafarpor, S., Ferns, G.A., Ghayour-Mobarhan, M., Manian, M., and Nedaeinia, R. (2021). The dysbiosis signature of Fusobacterium nucleatum in colorectal cancer-cause or consequences? A systematic review. Cancer Cell Int. *21*, 194. https://doi.org/10.1186/s12935-021-01886-z.

Rao, H., Wu, H., Huang, Q., Yu, Z., and Zhong, Z. (2021). Clinical value of serum CEA, CA24-2 and CA19-9 in patients with colorectal cancer. Clin. Lab. *67*. https://doi.org/10.7754/Clin.Lab.2020.200828.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

Saitoh, Y., Inaba, Y., Sasaki, T., Sugiyama, R., Sukegawa, R., and Fujiya, M. (2016). Management of colorectal T1 carcinoma treated by endoscopic resection. Dig. Endosc. *28*, 324–329. https://doi.org/10.1111/den.12503.

Salvucci, M., Crawford, N., Stott, K., Bullman, S., Longley, D.B., and Prehn, J.H.M. (2021). Patients with mesenchymal tumours and high Fusobacteriales prevalence have worse prognosis in colorectal cancer (CRC). Gut. https://doi.org/10.1136/gutjnl-2021-325193.

Serna, G., Ruiz-Pace, F., Hernando, J., Alonso, L., Fasani, R., Landolfi, S., Comas, R., Jimenez, J., Elez, E., Bullman, S., et al. (2020). Fusobacterium nucleatum persistence and risk of recurrence after preoperative treatment in locally advanced rectal cancer. Ann. Oncol. *31*, 1366–1375. https://doi.org/10.1016/j.annonc.2020.06.003.

Signat, B., Roques, C., Poulet, P., and Duffaut, D. (2011). Fusobacterium nucleatum in periodontal health and disease. Curr. Issues Mol. Biol. *13*, 25–36.

Sinicrope, F.A., Chakrabarti, S., Laurent-Puig, P., Huebner, L., Smyrk, T.C., Tabernero, J., Mini, E., Goldberg, R.M., Zaanan, A., Folprecht, G., et al. (2021). Prognostic variables in low and high risk stage III colon cancers treated in two adjuvant chemotherapy trials. Eur. J. Cancer *144*, 101–112. https://doi.org/10.1016/j.ejca.2020.11.016.

Wang, S., Liu, Y., Li, J., Zhao, L., Yan, W., Lin, B., Guo, X., and Wei, Y. (2021). Fusobacterium nucleatum acts as a pro-carcinogenic bacterium in colorectal cancer: from association to causality. Front. Cell Dev. Biol. *9*, 710165. https://doi.org/10.3389/fcell.2021.710165.

Xu, C., Fan, L., Lin, Y., Shen, W., Qi, Y., Zhang, Y., Chen, Z., Wang, L., Long, Y., Hou, T., et al. (2021a). Fusobacterium nucleatum promotes colorectal cancer metastasis through miR-1322/CCL20 axis and M2 polarization. Gut Microbes *13*, 1980347. https://doi.org/10.1080/19490976.2021.1980347.

Xu, S., He, Y., Lin, L., Chen, P., Chen, M., and Zhang, S. (2021b). The emerging role of ferroptosis in intestinal disease. Cell Death Dis. *12*, 289. https://doi.org/10.1038/s41419-021-03559-1.

Yamamoto, S., Kinugasa, H., Hirai, M., Terasawa, H., Yasutomi, E., Oka, S., Ohmori, M., Yamasaki, Y., Inokuchi, T., Harada, K., et al. (2021). Heterogeneous distribution of Fusobacterium nucleatum in the progression of colorectal cancer. J. Gastroenterol. Hepatol. *36*, 1869–1876. https://doi.org/10.1111/jgh.15361.

Yamamura, K., Baba, Y., Nakagawa, S., Mima, K., Miyake, K., Nakamura, K., Sawayama, H., Kinoshita, K., Ishimoto, T., Iwatsuki, M., et al. (2016). Human microbiome Fusobacterium nucleatum in esophageal cancer tissue is associated with prognosis. Clin. Cancer Res. *22*, 5574–5581. https://doi.org/10.1158/1078-0432.CCR-16-1786.

Yamaoka, Y., Suehiro, Y., Hashimoto, S., Hoshida, T., Fujimoto, M., Watanabe, M., Imanaga, D., Sakai, K., Matsumoto, T., Nishioka, M., et al. (2018). Fusobacterium nucleatum as a prognostic marker of colorectal cancer in a Japanese population. J. Gastroenterol. *53*, 517–524. https://doi.org/10.1007/s00535-017-1382-6.

Yang, Y., Weng, W., Peng, J., Hong, L., Yang, L., Toiyama, Y., Gao, R., Liu, M., Yin, M., Pan, C., et al. (2017). Fusobacterium nucleatum increases proliferation of colorectal cancer cells and tumor development in mice by activating toll-like receptor 4 signaling to nuclear factor-kappaB, and up-regulating expression of MicroRNA-21. Gastroenterology *152*, 851–866.e24. https://doi.org/10.1053/j.gastro.2016.11.018.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.

Yu, T., Guo, F., Yu, Y., Sun, T., Ma, D., Han, J., Qian, Y., Kryczek, I., Sun, D., Nagarsheth, N., et al. (2017). Fusobacterium nucleatum promotes chemoresistance to colorectal cancer by modulating autophagy. Cell *170*, 548–563.e16. https://doi.org/10.1016/j.cell.2017.07.008.

Zhang, X., Zhu, X., Cao, Y., Fang, J.Y., Hong, J., and Chen, H. (2019). Fecal Fusobacterium nucleatum for the diagnosis of colorectal tumor: a systematic review and meta-analysis. Cancer Med. *8*, 480–491. https://doi.org/10.1002/cam4.1850.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Patient samples (saliva and serum) in the training cohort	Qilu Hospital of Shandong University	N/A
Patient samples (saliva, serum and tumor tissue) in the test cohort	Shandong Provincial Third Hospital	N/A
Critical commercial assays		
QIAamp DNA Mini Kit	Qiagen	51306
TRIzol	Invitrogen	15596018
2×AceQ qPCR Probe Master Mix	Vazyme	Q112-03
Elecsys CEA kit	Roche Diagnostics	07027028190
Elecsys CA19-9 kit	Roche Diagnostics	07027079190
Ribo-off rRNA Depletion Kit	Vazyme	N406-02
VAHTS Universal V8 RNA-seq Library Prep Kit for Illumina	Vazyme	NR605-02
VAHTS RNA Clean Beads	Vazyme	N412-02
VAHTS DNA Clean Beads	Vazyme	N411-02
Deposited data		
RNA sequence	NCBI	SRA: PRJNA777960
Oligonucleotides		
qPCR primers (Table S4)	This paper	N/A
qPCR probes (Table S4)	This paper	N/A
Software and algorithms		
Principal component analysis	ggord	https://github.com/fawda123/ggord
Trimgalore v0.6.6	Trimgalore	https://github.com/FelixKrueger/TrimGalore
FastQC v0.11.9	FastQC	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
hisat2 v2.2.1	Kim et al. (2015)	http://daehwankimlab.github.io/hisat2/
featureCounts v2.0.3	Liao et al. (2014)	http://subread.sourceforge.net/
edgeR v3.32.1	Robinson et al. (2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
clusterProfiler v3.18.1	Yu et al. (2012)	https://guangchuangyu.github.io/software/clusterProfiler/
R (version 3.6.3)	R Core Team (2020)	https://cran.r-project.org
GraphPad Prism, 9.1	GraphPad Software	https://www.graphpad.com/
MedCalc, 9.3	MedCalc statistical software	https://www.medcalc.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled bythe lead contact, Yi Zhang (yizhang@sdu.edu.cn)

Materials availability

This study did not generate new unique reagents.

Data and code availability

- RNA-seq data have been deposited at SRA (PRJNA777960). Any additional information required to re-analyze the data reported in this paper is available from the lead contact upon request.
- All data reported in this paper will be shared by the lead contact upon request
- This paper does not report original code.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Subjects

The experimental protocols were approved by the local ethical committees (No. KYLL-2019-2-013), and written informed consent was obtained from all participants. In the training cohort, pre-operative saliva samples were collected from 324 cases with HC (n = 41), HP (n = 33), Ad (n = 43), and CRC (n = 207) undergoing colonoscopy examination in Qilu Hospital of Shandong University between March 2017 and June 2018. CRC patients had complete medical records and were followed up at regular intervals until death or June 2021. The patients with incomplete medical records and those who were lost to follow-up, or withdrew the consent were excluded from this study. In the test cohort, an independent cohort consisting of subjects with HC (n = 30), HP (n = 12), Ad (n = 18), and CRC (n = 30) was recruited from Shandong Provincial Third Hospital. In this cohort, paired serum and tumor tissue samples were collected from CRC patients besides saliva. All pathological tissues were confirmed by histopathological analyses. CRC cases were diagnosed and staged according to the 7th AJCC/UICC tumor-node-metastasis (TNM) system. Table S1 included the sample size, age and gender.

METHOD DETAILS

Sample preparation

Saliva was collected using Salivette® with a cotton swab from each subject. Briefly, subjects were asked to chow the swab for about 60 s after rinsing the mouth with water. The swab with absorbed saliva was returned to the tube and centrifuged at 1,000 g for 2 min to get a clear saliva sample. Blood was collected by vena puncture, and serum was separated by centrifugation at 1,600 g for 10 min. Saliva and serum samples were stored at −80°C until further analyses. Tumor tissues were washed with Hank's solution three times and then immediately frozen in liquid nitrogen.

DNA and RNA extraction

DNA was extracted using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). RNA was isolated from tissues using the TRIzol reagent (Invitrogen, Eugene, OR, USA). The DNA/RNA concentration was determined by a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA integrity was evaluated using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA).

qPCR

Probe-based qPCR was performed using 2×AceQ qPCR Probe Master Mix (Vazyme, Nanjing, China) in a reaction volume of 30 μl, with 15 μl AceQ qPCR Probe Master Mix, 0.6 μl forward/reverse primer for Fn DNA (10 μm), 0.3 μl TaqMan probe for Fn DNA (10 μm), 0.6 μl forward/reverse primer for reference genes (5 μm), 0.3 μl TaqMan probe for reference genes (5 μm), and 10.5 μl DNA. The primers and TaqMan probes were designed based on the *nusG* gene of Fn and human reference genes (*GAPDH*, *RPPH1*, *NAGK*, *TERT*, *ERV-3*, and *SLC O 2A1*), and their sequences were shown in Table S4. PCR amplification was on a Bio-Rad CFX96 real-time PCR cycler (Bio-Rad, Hercules, CA, USA), with an initial denaturation step at 95°C for 5 min, and 40 cycles at a melting temperature of 95°C for 10 s and an annealing temperature of 60°C for 30 s. All assays were conducted in triplicate. No-template reactions were performed as negative controls. The PCR products were confirmed by sanger sequence (Figure S1).

GAPDH and *TERT* were selected as reference genes for normalizing qPCR as described in result part. The relative level of Fn DNA was recorded as the ratio of Q (Fn DNA) to the geometric mean of Q (*GAPDH*) and Q (*TERT*). Q was calculated according to the following formula: $(\text{Effiency}+1)^{-\Delta Cq}$, where $\triangle$ Cq = [Cq (test)-Cq (calibrator)]. The effiencies of Fn DNA, GAPDH and TERT amplification were shown in Figure 2.

CEA and CA19-9 assays

The levels of CEA and CA19-9 were determined using electrochemiluminescence with the Roche Cobas e601 fully automatic analyzer (Roche Diagnostics, Mannheim, Germany) equipped with dedicated reagents. In brief, 6ul serum was incubated with biotinylated CEA/CA19-9 monoclonal antibody and Ru labeled CEA/CA19-9 monoclonal antibody to form antigen-antibody sandwich complex. Then the complex was incubated with magnetic bead particles coated with streptavidin. The reaction solution was sucked into the measuring cell, the magnetic beads were adsorbed on the electrode surface through electromagnetic action, and the substances not combined with the magnetic beads were removed. A certain voltage was applied to the electrode to make the complex chemiluminescence, and the luminous intensity was measured by a photoelectric multiplier. According to the manufacturer's recommendations, the cutoff values of CEA and CA19-9 were defined as 5 ng/mL and 39 U/mL, respectively.

RNA sequence (RNA-Seq)

RNA-seq libraries were prepared using Ribo-off rRNA Depletion Kit (Vazyme, Nanjing, China) and VAHTS Universal V8 RNA-seq Library Prep Kit for Illumina (Vazyme, Nanjing, China). In brief, total RNA (500 ng) extracted from CRC tissues was hybridized with ribosomal RNA (rRNA) Probe, followed by digestion with RNase H at 37°C for 30 min, and digestion with DNase I at 37°C for 30 min. Then the ribosomal-depleted RNA was purified with VAHTS RNA Clean Beads (Vazyme, Nanjing, China), and fragmented at 94°C for 5 min, followed by double strand cDNA synthesis, adapter ligation, two-rounds of purification for 250-350 bp insert size selection, and PCR (using 14 cycles). Finally, libraries were quantified using Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), assessed using LabChip Gx Touch 24 (PerkinElmer, Waltham, MA, USA), and sequenced on a NovaSeq 6000 sequencer (Illumina, San Diego, CA, USA) using 150 bp paired-end reads.

The raw data were trimmed with Trimgalore v0.6.6 and verified using FastQC v0.11.9. The clean reads aligned to the GRCH38 genome with hisat2 v2.2.1 (Kim et al., 2015). The reads were counted with featureCounts v2.0.3 (Liao et al., 2014) using genome annotation from Gencode v38 (https://www.gencodegenes.org/). The expression levels of all gene were normalized by calculating FPKM (fragment per kilobase of transcript per million mapped reads). Differential expression genes (DE-Gs) analysis between Fn low and high groups was performed with edgeR v3.32.1 (Robinson et al., 2010). And, the genes with $p < 0.05$ and absolute fold change ≥ 2 were considered as DE-Gs. DE-Gs were then subjected to enrichment analysis of KEGG pathways with clusterProfiler v3.18.1 (Yu et al., 2012).

QUANTIFICATION AND STATISTICAL ANALYSIS

Kruskal-Wallis test was performed for global comparison of Fn DNA, CEA, or CA19-9 levels among multiple groups, and further post-hoc multiple comparisons were carried out using the *Mann-Whitney U* test. The logistic regression model was adopted to combine markers and generate predicted probability values. The correlation between Fn levels in tissue and saliva was analyzed by the *Spearman* test. Survival curves were constructed using the *Kaplan-Meier* method and compared by log rank test. The Cox model was used to identify the independent prognostic factors. The AUC with 95%CI was computed on the receiver operating characteristic (ROC) curve. The cutoff value was also calculated based on the Youden index (sensitivity + specificity-1). Analyses were performed with GraphPad Prism, 9.1 and MedCalc, 9.3, all tests were two-sided, and $p < 0.05$ was considered as statistically significant.