Supplementary information

Ethics statement and study registration

The study protocol was compiled according to the Declaration of Helsinki and the Ethics Guidelines for Clinical Research published by the Ministry of Health, Labour, and Welfare of Japan. Approval for this study was obtained from the Ethics Committee of Yokohama City University Hospital on 16 October 2014. All patients provided written informed consent for participation in the study. The protocol and informed consent forms were approved by the institutional ethics committees at each participating institution. This study is registered in the UMIN Clinical Trials Registry (ID: UMIN000016229).

Study design

A total of 84 patients who were diagnosed with CRC by colonoscopy were enrolled in the study. The study was conducted at Yokohama City University Hospital, Yokohama, Japan. Exclusion criteria included the following: patients with insufficient samples including inadequate sample preparation; informed consent was not provided; history of malignant disease (excluding carcinoma *in situ* that had already been resected); history of familial adenomatous polyposis; history of hereditary non-polyposis colorectal cancer; and history of inflammatory bowel disease. Patients who received antibiotics within a month of the study were also excluded. Fresh CRC tissue and saliva samples were obtained from a total of 14 eligible patients.

Patients

Clinicopathological characteristics of the 14 patients are shown in Online Supplementary

Table 1. The mean age was 69.4 years (range, 54–88 years); 10 were male and four were female. All lesions were pathologically diagnosed as tubular adenocarcinoma. The longitudinal length of the lesions measured 15–80 mm (mean ± standard deviation: 35.4±15.9 mm). The lesion sites included six lesions in the ascending colon, four lesions in the sigmoid colon, and four lesions in the rectum. The clinical stage of the lesions was distributed from stage 0 to stage IV.[1]

Sample collection

Tissue specimens of colon lesions were obtained from patients who underwent endoscopy. Saliva was collected before or after endoscopy. All samples were collected in a tube and immediately stored under anaerobic conditions using AnaeroPack Anaero (Mitsubishi Gas Chemical, Tokyo, Japan) and chilled on ice until use.

Isolation of *F. nucleatum* from specimens

Specimens were cultured within 24 h of specimen collection. CRC specimens were homogenized using a 200-µl yellow tip in 100 µl of anaerobic diluent A (see below for further details). Saliva samples were diluted ten-fold and CRC samples were diluted five-fold with anaerobic diluent A. Fifty microlitres of each sample was then spread onto *Fusobacterium* selective agar [2] (see below for further details) and incubated under anaerobic conditions using AnaeroPack Anaero for 48–72 h at 37°C. Then, for purification, approximately 50 colonies were randomly picked and streaked on Eggerth-Gagnon (EG) agar (see below for further details) plates cultured under anaerobic conditions for 48–72 h at 37°C. When more than one colony was identified during the purification process, the colony was added to the list of isolates as another isolate. After

cultivation, all isolates were analysed by polymerase chain reaction (PCR) with *Fusobacterium* spp.- and *F. nucleatum*-specific primers, as follows. First, part of each colony was suspended in 20 μl of BL buffer (20 mM Tris-HCl (pH 8.0), 500 μM EDTA-2Na (Kanto Chemical, Tokyo, Japan), 0.5% (v/v) Tween 20 (Kanto Chemical), 0.25% (v/v) Nonidet P-40 (Pierce Surfact-Amps, Rockford, IL, USA)) [3] including 100 μg/ml Proteinase K (Roche Diagnostics, Mannheim, Germany) and incubated at 60°C for 20 min followed by 95°C for 5 min. Second, each lysate was then analysed by PCR using oligonucleotide primers Fs619F (5′-CGCAGAAGGTGAAAGTCCTGTAT-3′) and Fs719R (5′-TGGTCCTCACTGATTCACACAGA-3′) (approximately 100-bp amplicon), which are specific for the 23S rRNA gene of *Fusobacterium* spp.,[4] and Fs619F and Fn866R (5′-AAGGCACGCCATCACCCAAATG-3′) (approximately 250-bp amplicon), which are specific for *F. nucleatum*. To design the Fn866R primer, 23S rRNA sequences of genus *Fusobacterium* were obtained from the SILVA ribosomal RNA database (https://www.arb-silva.de/) [5] and aligned using MUSCLE software.[6] The primer was designed to include a 2–3 base mismatch within three bases of the 3′ -end in *Fusobacterium* species, with the exception of *F. nucleatum*. PCR was performed in a total volume of 15 μl of PCR mixture containing 0.05 U of Paq5000 DNA polymerase (Agilent Technologies, Santa Clara, CA, USA), 200 μM of dNTPs, 0.2 μM of each primer, 0.5 μl of cell lysate, and 1.5 μl of 10× Paq5000 reaction buffer. The following cycling parameters were used: 94°C for 2 min, 30 cycles of 94°C for 20 s, 58°C for 20 s, and 72°C for 20 s, and an additional cycle of 72°C for 2 min. Amplified products were separated by electrophoresis on a 2% agarose gel. Isolates with double-positive results for the two abovementioned PCRs were identified as *F. nucleatum*.

Components of anaerobe diluent A

Anaerobic diluent A	
KH_2PO_4	4.5 g
Na_2HPO_4	6 g
L-Cysteine·HCl·H_2O	0.5 g
Tween 80	0.5 g
Bacto-agar (Difco)	1 g
0.1% Resazurin	1 ml
Distilled water	1000 ml

N_2/CO_2 (80%/20%) composite gas was injected into glass test tubes using a deoxygenized gas pressure and replace injector, AG-2 (Sanshin Industrial Co. Ltd, Yokohama, Japan). Test tubes were then sealed using butyl rubber stoppers before autoclaving.

Components of EG agar

EG agar	
Lab-Lemco Powder (Oxoid)	2.4 g
Proticasepeptone No.3 (Bacto)	10 g
Yeast extract (Bacto)	5 g
Glucose	4 g
Soluble starch	0.5 g
L-Cystine	0.2 g
Na_2HPO_4	4 g
Bacto-agar (Bacto)	15 g
L-Cysteine·HCl·H_2O	0.5 g
antifoaming agent (Antifoam SI, Wako)	0.25 ml
Defibrinated horse blood	50 ml
Distilled water	up to 950 ml

EG agar plates or tubes were prepared as follows. 1) L-cysteine was added to 50 ml of 1 N HCl and mixed thoroughly. 2) The remaining components were then added to a final volume of 950 ml. 3) pH was adjusted to 7.6–7.8. 4) After autoclaving and cooling to 55°C, 50.0 ml of horse blood was aseptically added. 5) Solution was then mixed thoroughly and poured into sterile Petri dishes or distributed into sterile tubes.

Components of *Fusobacterium* selective agar

***Fusobacterium* selective agar**	
EG agar	100 ml
Selective agent (100× stock solution)	1 ml
The selective agent is filtrated (0.22 μm) and added when the autoclaved EG agar is about 55°C.	
Selective agent (100× stock solution)	
Crystal violet	70 mg
Vancomycin Hydrochloride	50 mg
Neomycin sulfate	300 mg
Nalidixic acid	250 mg
Distilled water	100 ml

16S ribosomal RNA (rRNA) gene sequencing for *F. nucleatum* subspecies identification

The 16S rRNA fragment was amplified by PCR in a total volume of 25 µl of PCR mixture containing 1× Ex Taq Buffer, 0.025 U of Ex Taq DNA Polymerase Hot Start version (Takara Bio Inc., Shiga, Japan), 200 µM of dNTPs, 0.2 µM of primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-ATTAGATACCCTGGTAGTCC-3′), and 10 ng of genomic DNA. The following thermal cyclic conditions were used: 94°C for 2 min, followed by 30 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min; and 72°C for 2 min. The amplicon was sequenced towards each end using primers 787F (5′-ATTAGATACCCTGGTAGTCC-3′) and 920R (5′-GTCAATTCCTTTGAGTTT-3′). DNA sequencing was performed by Eurofins Genomics (Tokyo, Japan). The obtained sequences were merged using ChromasPro software (http://technelysium.com.au/wp/chromaspro/) and compared with the 16S rRNA sequence database using nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Identification to the subspecies level was defined as the top BLAST hit with maximum sequence similarity of at least 99%.

AP-PCR

Isolated *F. nucleatum* colonies grown on EG agar plates were collected and washed twice in 1 ml ice-cold PBS by centrifugation at 15,000 rpm for 1 min at 4°C. The cell pellet was resuspended in 600 µl of DNA extraction buffer (10 mM Tris-HCl, pH 8.0, containing 1 mM EDTA-2Na 0.06% (v/v) sodium dodecyl sulphate and 100 µg/ml Proteinase K (Roche Diagnostics)), and then incubated at 37°C with occasional inversion. After clear lysates were obtained (approximately 0.5–2 h), lysates were purified using an equal

volume of TE-saturated phenol and phenol-chloroform-isoamyl alcohol (25:24:1). The purified solution was added to 0.03× volume of 3M sodium acetate, 3 μl of Ethachinmate (Nippon Gene, Tokyo, Japan) as a DNA carrier, and 0.7× volume of isopropanol, and centrifuged at 15,000 rpm for 15 min at 4°C. The DNA pellet was washed twice with 1 ml 70% ethanol, dried at room temperature, and dissolved in nuclease-free water. The concentration of extracted genomic DNA was determined using the Quant-iT dsDNA HS Assay kit and the Qubit 2.0 fluorimeter (Life Technologies, Carlsbad, CA, USA). AP-PCR was performed in a total volume of 25 μl of PCR mixture containing 1× Ex Taq Buffer, 0.025 U of Ex Taq DNA Polymerase Hot Start version (Takara Bio Inc.), 200 μM of dNTPs, 0.4 μM of primer (D8635 or D11344), and 10 ng of genomic DNA. Primer sequences were as follows: D8635: 5′ -GAGCGGCCAAAGGGAGCAGAC-3′; D11344: 5′-AGTGAATTCGCGGTGAGATGCCA-3′.[7] Amplification was performed according to a reported method with slight modification.[8] Briefly, PCR conditions were as follows: 95°C for 5 min; five cycles at 94°C for 3 min, 37°C for 3 min, and 72°C for 3 min; 30 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 3 min; and 72°C for 10 min. Amplified products were electrophoresed in 1.5% Tris-Borate EDTA agarose gel. The gel was stained with ethidium bromide and digitally photographed (Sayaca-Imager; DRC Co., Ltd, Tokyo, Japan) under UV light. Gene Ladder Wide 1 (0.1–20 kbp) (Nippon gene) was loaded as a DNA size marker.

REFERENCES

1 Sobin LH, Gospodarowicz MK, Wittekind C, et al. TNM classification of malignant tumours. 7th ed. Chichester, West Sussex, UK ; Hoboken, NJ: Wiley-Blackwell 2009:94-

105.

2 Mitsuoka T. A Color Atlas of Anaerobic Bacteria. Tokyo: Sobunsha 1980:pp.320–21.

3 Elbeltagy A, Nishioka K, Suzuki H, et al. Isolation and characterization of endophytic bacteria from wild and traditionally cultivated rice varieties. *Soil Science and Plant Nutrition* 2000;46:617–29.

4 Suzuki N, Yoshida A, Saito T, et al. Quantitative microbiological study of subgingival plaque by real-time PCR shows correlation between levels of Tannerella forsythensis and Fusobacterium spp. *Journal of clinical microbiology* 2004;42:2255–7.

5 Pruesse E, Quast C, Knittel K, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic acids research* 2007;35:7188–96.

6 Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 2004;32:1792–7.

7 Haraldsson G, Holbrook WP, Kononen E. Clonal similarity of salivary and nasopharyngeal Fusobacterium nucleatum in infants with acute otitis media experience. *J Med Microbiol* 2004;53:161–5.

8 George KS, Reynolds MA, Falkler WA, Jr. Arbitrarily primed polymerase chain reaction fingerprinting and clonal analysis of oral Fusobacterium nucleatum isolates. *Oral Microbiol Immunol* 1997;12:219–26.