SUPPLEMENTARY INFORMATION

Bacterial whole genome sequencing

High-quality genomic DNA extracted from bacteria culture was subjected to library preparation and whole genome sequencing using PacBio 3rd generation sequencing platform (Pacific Biosciences, Menlo Park, CA). Raw sequencing reads were filtered using circular consensus sequencing under default parameters (version 4.1). The filtered reads were then subjected to *de novo* genome assembly with Canu under default parameters, including read correction, read trimming, and contig construction. The quality of genome assembly was evaluated by CheckM. Genome annotation was performed by Rapid Annotation using Subsystem Technology (RAST).

Histological assessment

Paraffin-embedded colon tissues were sectioned (4 μm) and stained by haematoxylin and eosin staining. Histological assessment of tumours was evaluated by pathologists who were blinded to the group information. Images were captured by light microscope (Axio Imager 2, Zeiss, Jena, Germany) equipped with Metafer Automatic Slide Scanning and Imaging System (version 3.12.7; MetaSytems, Altlussheim, Germany).

Ki-67 immunohistochemical staining

Colon sections (4 μm) were deparaffinised, blocked, and incubated with anti-Ki-67 primary antibody (1:500; #ab16667, Abcam) at 4°C overnight. Signals were developed by IHC Select Immunoperoxidase Secondary Detection System (Merck) according to manufacturer's instructions. Haematoxylin was used for counterstaining, and rabbit serum diluted to the same concentration of

Su ACY, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2023-330835

primary antibody was used as negative control. Images were captured by light microscope equipped with Metafer Automatic Slide Scanning and Imaging System. The proportion of Ki-67 positive cells in each random field was measured by Image J with plugin IHC Profiler.

TUNEL staining

Colon sections (4 μm) were deparaffinised, rehydrated, and fixed in 4% paraformaldehyde for 15 minutes. 100 ml of proteinase K (20 mg/mL) was added to each slide and incubated at room temperature for 30 minutes. After PBS wash, slides were incubated with 100 mL of equilibration buffer at room temperature for 10 minutes, and with 100 mL of Tdt reaction mix at 37°C for 60 minutes. The reaction was stopped by 2X SSC and slides were blocked by 0.3% hydrogen peroxide for 5 minutes. 100 mL of streptavidin HRP (1:500 in PBS) was then added to slides and incubated at room temperature for 30 minutes. 100 mL of DAB solution was used to develop signals.

Bacteria detection

PCR was conducted on genomic DNA extracted from mouse stools using Premix Taq DNA Polymerase (Hot-Start Version; Takara, Kusatsu, Japan) in Veriti 96-Well Fast Thermal Cycler (Thermo Fisher Scientific, Rockford, IL). Each PCR mix contained 100 ng of sample DNA with a reaction volume of 10 μL. PCR cycle conditions are as follow: 1) 95°C for 5 minutes; 2) 95°C for 30 seconds, 57°C for 15 seconds, and 72°C for 30 seconds for 35-40 cycles; and 3) 72°C for 5 minutes. PCR products were subjected to agarose gel electrophoresis at 90V for 50 minutes, and captured by ChemiDoc XRS+ System (Bio-Rad, Hercules, CA). Primers used for bacteria detection are listed in **Supplementary Table 2**. A subset of PCR products was subjected to Sanger

sequencing performed by Sangon Biotech (Shanghai, China). The sequencing data were analysed using SnapGene Viewer (version 4.1) and BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Cell culture and CRC patient-derived organoid culture

Two human CRC cell lines HCT116 and HT29 as well as a human normal colon epithelial cell line NCM460, were purchased from American Type Culture Collection (Manassas, VA). Cells were cultured in Dulbecco's Modified Eagle's Medium (Gibco, Grand Island, NY) supplemented with 10% foetal bovine serum and 1% penicillin/streptomycin, and maintained at 37ºC in a humidified incubator with 5% carbon dioxide.

CRC organoids were derived from two CRC patients (#74 and #828) from Princess Margaret Living Biobank (Toronto, Canada), embedded into Matrigel (Corning), and cultured in DMEM/F12 (Gibco) with GlutaMAX (Invitrogen, Carlsbad, CA), N2 and B27 supplements (Invitrogen), 10 mmol/L HEPES, 1.25 mmol/L N-acetyl cysteine, glutamine, 1% penicillin/streptomycin, 10 mmol/L SB202190-monohydrochloride, R-spondin-1, Noggin, WNT3A, and 50 ng/mL epithelial growth factor (Invitrogen). Culture medium was changed every 3 days. After treatment of 7-10 days, images were captured and the surface area of organoids in each random field was measured by ImageJ (version 1.53a). The experiment was conducted with 9-12 organoids per group. To investigate candidate proteins, cells or organoids were directly treated with commercial αMAN (10 ng/mL in PBS; #M7257, Sigma-Aldrich). The activity of αMAN was determined using α-mannosidase Assay Kit (#ab272519, Abcam Cambridge, MA).

Cell viability and colony formation assays

Cell viability assay was performed using 3-(4,5-dimethylthiazol- 2-yl)-2,5-diphenyltetrazolium (MTT, 5 mg/mL; Invitrogen). 1,000 cells per well were seeded onto 96-well plates for treatment. Cell viability was determined by measuring absorbance at a wavelength of 570 nm (OD_{570}) to evaluate the amount of MTT formazan products for 4-5 consecutive days. The experiment was conducted with 5-10 replicates per group. For colony formation assay, 1,000 cells per well were seeded onto 6-well plates. Culture medium was changed every 3 days. After treatment of 10 days, colonies were fixed by ice-cold methanol and stained by 0.5% crystal violet. Colonies with more than 50 cells were captured and measured by ImageJ. The experiment was conducted with triplicates per group.

Ki-67 immunofluorescent staining

Cells were seeded onto coverslips in 4-well plates. After treatment of 4-5 days, cells were fixed by 4% formaldehyde in PBS and blocked by 1% bovine serum albumin. Cells were then incubated with anti-Ki-67 primary antibody (1:6400; #9449, Cell Signalling Technology, Danvers, MA) at 4°C overnight, followed by Alexa Fluor 488 secondary antibody (1:500; #A-11001, Thermo Fisher Scientific) at room temperature for 60 minutes in dark. Cells were mounted with 5-minutes nuclear counterstaining by DAPI solution (1 mg/mL; Thermo Fisher Scientific). Images were captured by confocal microscope. The proportion of Ki-67 positive cells in each random field was measured by ImageJ with plugin IHC Profiler. The experiment was conducted with triplicates per group.

Bacteria attachment assay

Cells were seeded onto 24-well plates at ≥ 90% of coverage and cell count was recorded. Bacteria culture at growth phase (OD_{600} = 1) was centrifuged and collected after PBS wash. The obtained bacteria were then added to each well (multiplicity of infection = 10, 50, 100), and incubated with cells at 37ºC for 3 hours. After PBS wash, cells were decomposed by adding 400 μL of BHI broth and 100 μL of bacteria-free double-distilled water. Serial dilutions (1:1, 1:10, 1:100, 1:1000) were prepared for all samples and controls. 100 μL of diluted solutions was spread onto BHI agar plate and incubated at 37ºC overnight. Colony-forming unit was recorded to determine the ratio of bacteria-cell attachment. The experiment was conducted with triplicates per group.

Protein extraction and Western blot

Total proteins were extracted from human CRC cell lines (HCT116, HT29) by CytoBuster Protein Extraction Reagent (Merck) supplemented with protease inhibitor and phosphatase inhibitor. Sample mixtures were incubated at 4ºC for 60 minutes. Supernatants were collected upon centrifugation at 16,000 ×g, 4ºC for 10 minutes. Protein concentration was measured by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) according to manufacturer's instructions. 20-40 mg of the extracted proteins were separated by 10% SDS-PAGE gel electrophoresis and transferred onto PVDF membranes with pore size of 0.22 or 0.45 μm (Merck). Membranes were then blocked by 5% bovine serum albumin for 60 minutes, and incubated with primary anti-PCNA antibody (1:2000; #2586, Cell Signalling Technology) at 4°C overnight, followed by 60-minutes incubation of anti-mouse IgG, HRP-linked secondary antibody (1:10,000; #7476, Cell Signalling Technology) at room temperature. After substantial wash by detergent (Tris-buffered saline with

Su ACY, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2023-330835

0.1% Tween-20), proteins of interest were visualised by Clarity Wstern ECL Substrate (Bio-Rad) using ChemiDoc XRS+ System. β-actin was used as the housekeeping control.

Silver staining

Proteins were resolved by 5% SDS-PAGE and silver staining was performed according to manufacturer's instructions (Thermo Fisher Scientific). In brief, visually intense bands at 100kDa fraction were collected for in-gel digestion and decolourisation. Gel pieces were incubated with 800 µL of ACT for 10 minutes, 600 µL of 5mM DTT at 55°C for 30 minutes, and 800 µL of ACN for 10 minutes. Gels were then alkylated by 600 µL of 15mM IAA for 40 minutes in dark, followed by digestion using 600 µL of trypsin buffer at 37ºC for 30 minutes. The digested peptides were isolated by 800 µL of extraction buffer, vacuum dried, and resuspended in 0.1% formic acid for liquid chromatography-mass spectrometry.

Nano LC-MS and analysis

5 mL of total peptides were subjected to nanoflow ultra-performance LC-MS instrument (EASY-nLC1200) coupled to Q Extractive HF-X Quadrupole-Orbitrap MS System (Thermo Fisher Scientific). Separation was performed using reversed phase column (100 mm ID x 15 cm, Reprosil-Pur 120 C18-AQ, 1.9 mm; Dr. Maisch, Ammerbuch, Germany) under two mobile phases (phase A, H_2O with 0.1% formic acid, 2% acetonitrile; phase B, 80% acetonitrile, 0.1% formic acid) with a 60-minute gradient at a flow rate of 300 nL/min. Data-dependent acquisition was conducted at a resolution of 120,000 (200 m/z) and m/z range of 350-1600 for MS1; and at a resolution of 15,000 with dynamic first mass for MS2. The automatic gain control target for MS1

was 3e6 with maximum injection time of 50 milliseconds; and 1e5 for MS2 with maximum injection time of 110 milliseconds. The top 10 most intense ions were fragmented at normalised collision energy of 27% and isolation window of 1.2 m/z. The dynamic exclusion time window was 45 seconds, and peaks with charge > 6 were excluded. Candidate proteins were screened by the following criteria: (1) matching molecular weight corresponding to the position of protein bands on SDS-PAGE gel after silver staining (100-140kDa or 140-180kDa); and (2) molecular weight > 100kDa, Sequest-HT score > 100, coverage > 30%.