MATERIALS AND METHODS

STUDY DESIGN

This study was conceptualized to investigate the underlying mechanisms of *H. pylori* induced colorectal carcinogenesis.

Tumor mouse models *Apc*$^{+/min}$, initially obtained from Jackson Laboratories, and *Apc*$^{+/1638N}$ mice, provided by Prof. Klaus-Peter Janssen (Klinikum rechts der Isar, München) (1), were bred under specific pathogen-free conditions at our animal facility at the Technical University of Munich. Both female and male mice were used and co-housed with littermate controls.

C57BL/6 mice, *Apc*$^{+/1638N}$ mice and wild type littermates were re-derived germ-free from conventional mice by Prof. Bleich and Dr. Basic (Hannover medical school, Hannover).

Female C57BL/6 mice were purchased from Envigo RMS GmbH at an age of 6 weeks and acclimatized to our animal facility for 1-2 weeks prior infection. Mice were fed with a standard diet and water ad libitum and maintained under a 12-hour light-dark cycle. DNA extracted from mouse tails was used for genotyping. All animal experiments were conducted in compliance with European guidelines for the care and use of laboratory animals and were approved by the Bavarian Government (Regierung von Oberbayern, Az.55.2-1-54-2532-161-2017).

87 fresh colonoscopy biopsies were collected within the framework of the ColoBAC study of the CRC1371 (Dept. of Surgery and II. Medical Dept. Klinikum rechts der Isar, Technical University of Munich, Germany). 67 FFPE colon biopsies were obtained from the Klinikum Bayreuth. Both studies were approved by the respective ethics committees (Klinikum rechts der Isar #322/18, Klinikum Bayreuth #241_20Bc).

H. pylori status of colonoscopy biopsies was determined in serum samples using the recomwell Helicobacter IgG kit (Mikrogen) according to manufacturer's instructions. *H. pylori* status of FFPE biopsies was determined histologically in corresponding gastric biopsies by Prof. Michael Vieth (Table S2).

H. PYLORI INFECTION

H. pylori pre-mouse Sidney Strain 1 (PMSS1) was cultured on Wilkins-Chalgren (WC) Dent (containing vancomycin, trimethoprim, cefsoludin and amphotericin) agar plates in a microaerophilic atmosphere (5% O2, 10% CO2). 6–8-week old mice were orally gavaged twice within 72 hours with 2 × 10^8 *H. pylori* PMSS1 in 200 µl brain-heart-infusion (BHI) medium containing 20% fetal calve serum (FCS). Infection status was determined by plating homogenized stomach, intestinal and colonic tissue on WC Dent plates supplemented with 200 g/ml bacitracin, 10 g/ml nalidixic acid and 3 g/ml polymycin B, and counting colony-forming units (CFU).

H. PYLORI ERADICATION

After 4 weeks of infection, *H. pylori* eradication was performed with an antibiotic cocktail containing clarithromycin (Eberth) (7.15 mg/kg/day), metronidazole (Carl Roth) (14.2 mg/kg/day) and the proton-pump inhibitor omeprazole (Carl Roth) (400 µmol/kg/day) by oral gavage twice daily for 7 consecutive days. Omeprazole was dissolved in 200µl 2.5% Hydroxy-propyl-methyl-cellulose (Sigma-Aldrich) with pH adjusted to 9. Antibiotics, dissolved in 200µl PBS, were administered 45 minutes after omeprazole (2).

STOOL TRANSFER

SPF *Apc*$^{+/1638N}$ mice and wild type littermates were infected at an age of 6-8 weeks and after 12 or 24 weeks of infection, stool pellets were collected from these “donor” mice, dissolved in 0.1ml PBS/g stool and administered via oral gavage to germ-free C57BL/6 or *Apc*$^{+/1638N}$ “recipient” mice.

HISTOLOGY AND IMMUNOHISTOCHEMISTRY

Small intestine and colon were longitudinally opened and flushed with phosphate-buffered saline (PBS). Solid neoplastic lesions were assessed and measured macroscopically by two independent examiners. Dissected tissue was fixed in 4% formaldehyde, embedded in paraffin and 4µm thick sections were used for staining. For immunohistochemical stainings, antigen retrieval was achieved with 10 mM sodium citrate (pH 6) or 1 mM EDTA (pH 8), and primary antibodies were applied overnight at 4°C (Table 1). Horseradish peroxidase (HRP) coupled secondary antibodies (Promega) and diaminobenzidine (DAB) (CellSignaling) were used to detect signal. Periodic acid Schiff (PAS) (Carl Roth) staining was performed to assess the quantity of mucus producing goblet cells. Stomach, intestinal and colonic sections were blindly quantified by two independent researchers by measuring the area of a functional unit (stomach gland, intestinal crypt/villus unit or colonic crypt) and counting positive cells per mm^2, using Aperio ImageScope (Leica BioSystems).

Table 1. Antibodies used for immunohistochemical evaluation.

Target	Clone	Origin/Target	Antigen retrieval	Dilution	Company
CD3	SP7	Rabbit mAB	sodium citrate	1:150	Thermo Fisher
Ki67	D3B5	Rabbit mAB	sodium citrate	1:400	Cell Signaling
pSTAT3	D3A7	Rabbit mAB	EDTA	1:200	Cell Signaling

CHIPCYTOMETRY AND AUTOMATIC IMAGE QUANTIFICATION

Murine intestinal and colonic cross-sections were preserved with O.C.T. compound (Tissue Tek) in cryomolds (Tissue Tek) and kept frozen at -80°C. For ChipCytometry, 7µm thick sections were cut on a Cryostat (Leica), fixed in Fixation Buffer (ZELLKRAFTWERK) for 45 minutes and subsequently transferred to CellSafe Chips (ZELLKRAFTWERK). ChipCytometry on human FFPE biopsies was performed according to the procedure described in Jarosch, Köhlen et al (3). Briefly, tissue sections were rehydrated on coverslips and antigen retrieval was performed using TRIS-EDTA buffer (pH 8.5) and then transferred to CellSafe Chips (ZELLKRAFTWERK). Alternating cycles of staining, immunofluorescence detection and photobleaching were performed for various markers (Table 2). Automated image processing was performed as described in Jarosch, Köhlen et al., which includes segmentation of cells, removing of outliers and spatial spill over correction (3). The resulting cell – marker matrix was analyzed using FlowJo software (V10.8.0), which enabled absolute quantification of cells.

Table 2. Antibodies used for ChipCytometry.

Epitope	Fluorochrome	Clone	Dilution	Company	Catalog #
anti-mouse CD3	PerCP/Cy5.5	17A2	1:200	BioLegend	100218
anti-mouse CD4	eFluor450	RM4-5	1:80	BioLegend	100531
anti-mouse CD45	FITC	30-F11	1:100	BioLegend	103108
anti-mouse Foxp3	PE	FJK-16s	1:80	eBioscience	14-5773-82
Pan-cytokeratin	AlexaFluor 488	C-11	1:100	BioLegend	628602
Hoechst	BUV395	-	1:50.000	ThermoScientific	H3570
α-SMA	eFluor570	1A4	1:750	eBioscience	41-9760-80
anti-human CD3	unconjugated	SP7	1:150	ThermoScientific	RM-9107-S1
anti-human CD4	AlexaFluor 488	polyclonal	1:50	R&D Systems	FAB8165G
anti-human CD45	PerCP/Cy5.5	HI30	1:80	BioLegend	304028
anti-human Foxp3	PE	236A/E7	1:30	eBioscience	563791
2nd anti-rabbit	PE	Polyclonal	1:300	BioLegend	406421

LAMINA PROPRIA AND INTRAEPITHELIAL LYMPHOCYTE ISOLATION AND FLOW CYTOMETRY

Harvested intestinal tissue was cut open longitudinally after removing Peyer's Patches and adjacent tissue. Subsequently, tissue was treated with 30mM EDTA, filtered supernatants were collected as intraepithelial lymphocytes and remaining tissue was digested with 0.5mg/mL collagenase from Clostridium histolyticum Type IV (Sigma Aldrich) and 10µg/mL DNase I (Applichem). Filtered and centrifuged lamina propria cell suspensions were density separated using a Percoll gradient (Thermo Fisher).

Fresh human biopsies were collected in Hank's Balanced Salt Solution w/o Mg^{2+}/Ca^{2+} (HBSS) and digested with 0.1% collagenase from Clostridium histolyticum Type IV (Sigma Aldrich) for 30 min. at 37°C. Digestion was stopped by adding 20 mL HBSS and centrifugation twice. Isolated lymphocytes were frozen in Dimethyl sulfoxide (Applichem) + 20 % FCS at -80° C. Single cell suspensions were blocked with anti-mouse CD16/CD32 or anti human TruStain FcX and live/dead staining performed with Zombie Aqua (BioLegend) in PBS. Surface antibodies (Table 3) were diluted according to titration experiments and cells stained for 30 min at 4° C. For transcription factors, Foxp3 Transcription Factor Staining Buffer Set (eBioscience) was used according to manufacturer's instructions. For stimulation with whole *H. pylori* lysate, cells were stimulated for 12 hours with 20µg/mL PMSS1 lysate at 37° C and protein transport inhibitor Golgi Plug (BD Biosciences) added 1:1000 after 7 hours for a total of 5 hours. Stimulated cells were stained with intracellular cytokine staining kit according manufacturer's instructions (BD Biosciences). Stained single cell suspensions were acquired on a CytoFlex S (Beckman Coulter) and analyzed using FlowJo software (V10.8.0).

Table 3. Antibodies used for flow cytometry.

Epitope	Fluorochrome	Clone	Dilution	Catalog #	Company
anti-mouse CD16/CD32	-	93	1:500	14-0161-86	eBioscience
anti-mouse CD45	AlexaFluor 700	30-F11	1:400	103128	BioLegend
anti-mouse CD3ε	FITC	500A2	1:200	152304	BioLegend
anit-mouse CD4	BV605	RM4-5	1:250	100548	BioLegend
anti-mouse CD4	eFluor450	RM4-5	1:250	48-0042-82	eBioscience
anti-mouse CD8a	APC-H7	53.6-7	1:250	560182	BD Biosciences
anti-mouse FoxP3	eFluor450	FJKL-16s	1:200	45-5773-82	eBioscience
anti-mouse RORyt	PE	B2D	1:100	12-6988-82	eBioscience
anti-mouse IL-17A	APC	eBio17B7	1:150	17-7177-81	eBioscience
anti-human CD45	AlexaFluor 700	2D1	1:50	368514	BioLegend
anti-human CD3	FITC	OKT3	1:300	317306	BioLegend
anti-human CD4	PB450	SK3	1:30	344620	BioLegend
anti-human CD8	APC-H7	SK1 (RUO)	1:150	560179	BD Biosciences
anti-human FoxP3	PerCP/Cy5.5	PCH101		35-4776-42	eBioscience
Human TruStain FcX	-	-	1:500	422302	BioLegend

SINGLE CELL RNA SEQUENCING

Intestinal and colonic tissue was harvested and cells were isolated as described in section "Lymphocyte isolation and Flow cytometry". Single cell suspensions were stained with anti-mouse CD45 PB450 (Clone: 30-F11, BioLegend, #103126), anti-mouse EPCAM APC (Clone: G8.8, BioLegend, #118214) and Propidium Iodide (PI). CD45+ PI- and EPCAM+ PI- cells were sorted.

For cell hashing TotalSeq-B anti-mouse Hashtags 1, 2 and 5 to 8 (M1/42; 30-F11, Biolegend, 155831, 155833, 155839, 155841, 155843, 155845) were used at a dilution of 1:50.

Single cell RNA Sequencing was performed with 10X Genomics, according to manufacturer's instructions (Chromium™ Single Cell 3' GEM v3 kit). Sorted cells were centrifuged and resuspended in mastermix and 37.8 µl of water, before 70 µl of the cell suspension was transferred to the chip. QC was performed with a high sensitivity DNA Kit (Agilent) on a Bioanalyzer 2100, and libraries were quantified with the Qubit dsDNA HS assay kit (life technologies).

Libraries were pooled according to their minimal required read counts (20.000 reads/cell for gene expression libraries). Illumina paired end sequencing was performed with 150 cycles on a NovaSeq 6000.

Annotation was performed using cellranger (V5.0.0, 10X genomics) against the murine reference genome GRCm38 (mm10-2020-A). All subsequent analysis was performed using SCANPY V1.6 (4). Preprocessing was performed following the guidelines of best practice in single-cell RNA-seq analysis (5) and involved less than 15% mitochondrial genes, regressing out cell cycle, mitochondrial genes and total counts. The data was normalized per cell count and logarithmized. Genes used for gene scores are listed in online supplemental table 1 and the scores were computed with the SCANPY build in function "sc.tl.score_genes".

RNA velocities were calculated using velocyto (6) and analyzed with scVelo (V 0.2.3) (7).

16S RRNA SEQUENCING

Bacterial DNA extraction and 16S rRNA sequencing was either performed as described previously (8, 9) by the Core Facility Microbiome of the ZIEL Institute for Food & Health (Technical University of Munich) or as follows: small intestine and colon tissue were homogenized with a Precellys® 24 homogeniser (Avantor). Phenol chloroform DNA isolation and ethanol precipitation were performed following modified protocols of P.J. Turnbaugh et al., 2009 and E. G. Zoetendal et al. 2006 (10, 11). Subsequently, the V3/V4 region of the *16S rRNA* gene was amplified and double indexed using barcoding primers modeled after Kozich *et al*. PCR fragments were purified using magnetic AMPure XP beads (Beckman Coulter, USA) according to manufacturer's instructions (12). The final pooled library was sequenced on an Illumina MiSeq with Reagent Kit v3 (Ilumina) for 600 cycles of paired-end reads.

Raw sequences were analyzed using the Qiime2 platform (v2021.4) (13). In detail, denoising, removing of chimeras and generation of Amplicon Sequence Variants (ASVs) was performed with dada2. Subsequently, a phylogenetic tree was generated and diversity measures were calculated. Chao1 index was used to determine community alpha diversity. Taxonomic

classification was performed with a qiime2 feature classifier trained on the SILVA132 99% OTUs, specifically targeting the V3 region. Linear discriminant analysis effect size (LEfSe) determining differentially abundant features was performed on the online interface at http://huttenhower.sph.harvard.edu/lefse/, developed by Segata et al. (14).

QUANTITATIVE PCR

Stomach, small intestine and colon tissue were homogenized with Precellys® 24 homogeniser (Avantor) and RNA isolation was performed with a Maxwell 48 RSC simply RNA Tissue Kit on a Maxwell RSC Instrument (Promega). cDNA was synthesized with Moloney Murine Leukemia Virus Reverse Transcriptase RNase H- Point Mutant (Promega). Gene expression was assessed with GoTaq qPCR Mastermix (Promega) on a CFX384 system (Bio-Rad). The quantitative PCR consisted of 40 cycles of amplification with 15 sec denaturation at 95 ˚C, 1 min annealing and amplification at 60 ˚C. According to the ΔΔCT method, CT values were normalized to *Gapdh* and to uninfected controls, in order to determine fold changes in gene expression. The sequences of primers used are summarized in Table 3.

Table 3. Primer sequences used for qPCR.

Gene	Forward sequence	Reverse sequence
Gapdh	GCCTTCTCCATGGTGGTGAA	GCACAGTCAAGGCCGAGAAT
Foxp3	AGGAGCCGCAAGCTAAAAGC	TGCCTTCGTGCCCACTGT

STATISTICAL ANALYSIS

Statistical analysis was conducted on biological replicates as stated in the figure legends. Depending on Gaussian distribution, statistical significance between two groups was determined with unpaired student's t-test or Mann-Whitney-U test and for analysis among more than two groups, ordinary one-way analysis of variance with Tukey's multiple-comparisons test or Kruskal-Wallis-test with Dunn's multiple-comparisons test. P values below 0.05 were considered significant. Exact p values are stated when relevant. Statistical analysis was carried out using Prism 8 (GraphPad Software).

1. Janssen KP, Alberici P, Fsihi H, Gaspar C, Breukel C, Franken P, et al. APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression. Gastroenterology. 2006;131(4):1096-109.
2. van Zanten SJ, Kolesnikow T, Leung V, O'Rourke JL, Lee A. Gastric transitional zones, areas where Helicobacter treatment fails: results of a treatment trial using the Sydney strain mouse model. Antimicrob Agents Chemother. 2003;47(7):2249-55.
3. Jarosch S, Köhlen J, Sarker RSJ, Steiger K, Janssen KP, Christians A, et al. Multiplexed imaging and automated signal quantification in formalin-fixed paraffin-embedded tissues by ChipCytometry. Cell Reports Methods. 2021;1(7).
4. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15.
5. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol. 2019;15(6):e8746.
6. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature. 2018;560(7719):494-8.
7. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020;38(12):1408-14.
8. Reitmeier S, Kiessling S, Neuhaus K, Haller D. Comparing Circadian Rhythmicity in the Human Gut Microbiome. STAR Protoc. 2020;1(3):100148.
9. Reitmeier S, Kiessling S, Clavel T, List M, Almeida EL, Ghosh TS, et al. Arrhythmic Gut Microbiome Signatures Predict Risk of Type 2 Diabetes. Cell Host Microbe. 2020;28(2):258-72 e6.
10. Zoetendal EG, Heilig HG, Klaassens ES, Booijink CC, Kleerebezem M, Smidt H, et al. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat Protoc. 2006;1(2):870-3.
11. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. Nature. 2009;457(7228):480-4.
12. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013;79(17):5112-20.
13. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-7.
14. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12(6):R60.