Parvimonas micra , an oral pathobiont associated with colorectal cancer, epigenetically reprograms human primary intestinal epithelial cells

Emma Bergsten, Denis Mestivier, Francoise Donnadieu, Thierry Pedron, Landry Tsoumtsa, Emmanuel Lemichez, Olivier Gorgette, Stevenn Volant, Abiba Doukani, Philippe Sansonetti, et al.

Parvimonas micra, an oral pathobiont associated with colorectal cancer, epigenetically reprograms human primary intestinal epithelial cells

Emma Bergsten,[1,2,4] Denis Mestivier,[2,3] Francoise Donnadieu,[1] Thierry Pedron,[1,10] Landry Tsoumtsa,[4] Emmanuel Lemichez,[4] Olivier Gorgette,[6] Stevenn Volant,[7] Abiba Doukani,[8] Philippe J. Sansonetti,[1,9,*] Iradj Sobhani,[2,5,*] and Giulia Nigro[1,11,*,#]

Authors/affiliations

[1] Unité de Pathogénie Microbienne Moléculaire, INSERM U1202, Institut Pasteur, Paris 75015, France.
[2] Équipe universitaire EC2M3-EA7375, Université Paris- Est (UPEC), Créteil 94000, France.
[3] Plateforme de Bio-informatique, Institut Mondor de Recherche Biomédicale (IMRB / INSERM U955), Université Paris-Est, Créteil 94000 France.
[4] Unité des Toxines Bactériennes, CNRS UMR2001, INSERM U1306, Institut Pasteur, Paris 75015, France.
[5] Service de Gastroentérologie, CHU Henri Mondor APHP, Créteil 94000, France.
[6] Unité de bio-imagerie ultra-structurale (UtechS UBI), Centre de Recherche et de Ressources Technologies (C2RT), Institut Pasteur, Paris 75015, France.
[7] Bioinformatics and Biostatistics Hub, Institut Pasteur, Paris 75015, France.
[8] Sorbonne Université, Inserm, Unité Mixte de Service Production et Analyse de données en Sciences de la Vie et en Santé, UMS PASS, Plateforme P3S, F-75005, Paris, France.
[9] Chaire de Microbiologie et Maladies Infectieuses, Collège de France, 75231 Paris, France.
[10] Present address: Unité Bactériophage, Bactérie, Hôte, Institut Pasteur, Paris 75015, France.
[11] Present address: Microenvironment and Immunity Unit, INSERM U1224, Institut Pasteur, Paris 75015, France.
* Correspondence: philippe.sansonetti@pasteur.fr, iradj.sobhani@aphp.fr, giulia.nigro@pasteur.fr
# Lead contact

Summary

Recently, an intestinal dysbiotic microbiota in the feces of colorectal cancer (CRC) patients with an enrichment of bacteria belonging to the oral microbiota has been described. Here we characterized and investigated one of these oral pathobionts, the Gram-positive anaerobic coccus *Parvimonas micra*. We identified two phylotypes (A and B) exhibiting different phenotypes and adhesion capabilities. We observed a strong association of phylotype A with CRC, with its higher abundance in feces and in tumoral tissue compared to the normal homologous colonic mucosa, which was associated with a distinct methylation status of patients. By developing an *in vitro* hypoxic co-culture system of human primary colonic cells with anaerobic bacteria, we showed that *P. micra* phylotype A induces modifications in DNA

methylation of the promoters of several tumor-suppressor genes, oncogenes, and genes involved in epithelial-mesenchymal transition, providing evidence of its possible role in carcinogenesis.

Keywords: Colorectal cancer, colonic epithelial primary cells, Dysbiosis, *Parvimonas micra,* DNA methylation.

Introduction

Colorectal cancer (CRC) is a multifactorial disease due to various genetic and environmental factors that contribute to tumor formation and disease development. Genetic predispositions for CRC, such as Lynch syndrome (also called human no polyposis colonic cancer-HNPCC) or familial adenomatous polyposis (FAP), only accounts for a minority of CRC, representing less than 5% of all cases (Jiao et al., 2014). The overwhelming majority of CRCs are sporadic cancers, i.e., the cause is unknown. Several lifestyle environmental factors increase the risk of CRC including physical inactivity (Samad et al., 2005), overweight (Kyrgiou et al., 2017) and high consumption of red meat, processed meat and unsaturated fatty acids (Chan et al., 2011). The identification of lifestyle factors supports the hypothesis that the increase in CRC incidence is strictly related to environmental changes. Among environmental factors, the role of microorganisms in cancer has been increasingly recognized and a global imbalance, called dysbiosis, of gut microbiota compared to healthy individuals was observed in several cancers such as biliary, hepatic, breast cancers, or CRC (Schwabe and Jobin, 2013). Thus, the gut microbiota has emerged as an important carcinogenic environmental factor, in particular for CRC, due to the gut microbiota proximity and constant crosstalk with the intestinal epithelium (Soderholm and Pedicord, 2019). Different bacterial species, for example *Bacteroides fragilis*, *Fusobacterium nucleatum (F. nucleatum)*, pathogenic strains of *Escherichia coli,* or *Streptococcus gallolyticus*, have been associated with CRC development and mechanistic studies have tried understand the roles of these bacteria in carcinogenesis (Goodwin et al., 2011; Pasquereau-Kotula et al., 2018; Pleguezuelos-Manzano et al., 2020; Rubinstein et al., 2013). Recently, our team and others have identified, by metagenomic studies from humans' feces, not only the intestinal bacterium *F. nucleatum* associated with CRC but also bacteria belonging to the oral microbiota, such as *Gemella morbillorum*, *Solobacterium moorei*, *Porphyromonas gingivalis, and Parvimonas micra* (*P. micra*) (Koliarakis et al., 2019; Schmidt et al., 2019; Sobhani et al.,

2019). The overabundance of these bacteria in the colonic niche may be an etiological factor in CRC (Saffarian et al., 2019), but the underlying mechanisms remain to be elucidated. Here, we chose to focus on the candidate bacterium, *P. micra*.

P. micra, the only described species within its genus, is a Gram-positive anaerobic coccus and commensal of the oral cavity (Rams et al., 1992). *P. micra* could be considered a pathobiont (Kamada *et al.*, 2013) because it is often isolated from oral polymicrobial infections associated with periodontitis (Socransky et al., 1999), from superficial polymicrobial infections, including wounds, ulcers, and skin abscesses (Bowler and Davies, 1999b, 1999a; Edmiston et al., 1990; Mousa, 1997; Sanderson et al., 1979), as well as from deep tissue infections in the brain, lung and reproductive organs (Murdoch et al., 1988).

P. micra has been poorly studied due to difficulties in its cultivation and laboratory identification by traditional methods and the polymicrobial nature of the infective zones from which it is usually isolated. The aim of this work was to better characterize *P. micra*, define its carriage in CRC patients, and analyze *in vitro* the nature of the crosstalk between this oral pathobiont and colonic epithelial cells. We identified two phylotypes of *P. micra* (A and B) and found that phylotype A is strongly associated with human CRC, and particularly in patients with a positive cumulative methylation index, and *in vitro* induces DNA methylation modifications in colonic primary epithelial cells.

Results

P. micra exists in two phylotypes with phenotypic and genetic diversity

To explore *P. micra* diversity, twenty-seven clinical isolates from various types of infections (blood, systemic, and oral cavity) were collected (Table S1), while, unfortunately, the high bacterial diversity in feces precluded *P. micra* isolation from CRC patient fecal samples. *P. micra* colonies on blood agar plates appeared heterogeneous. A contact haemolytic activity was observed in twenty isolates (20/27) (Figure 1A, Figure S1A) while *P. micra* colonies from solid agar plates were found to be either “compact” (7/27) or “non-compact” (20/27) (Figure 1A, Figure S1A). Sedimentation assays in liquid of two representative isolates: ATCC 33270 (“non-compact”) and HHM BlNa17 (“compact”), (named hereafter *PmA* and *PmB*, respectively), showed differences in the aggregation rate (Figure S1B).

Phenotypical differences observed among the twenty-seven isolates could reflect genetic diversity. Indeed, sequencing of the full 16S rDNA of these isolates revealed two distinct phylogenetic groups: A and B (Figure 1A). *Finegoldia magna (F. magna)*, the most closely related bacterial species to *P. micra*, was used to set the root of the phylogenetic tree. Interestingly, all *P. micra* isolates from phylotype A were haemolytic whereas isolates belonging to phylotype B were not. Likewise, most phylotype B isolates were "compact" whereas phylotype A isolates were not. To deepen our analysis, the ten *P. micra* genomes available on the NCBI database were clustered based on the presence/absence of genes. Grouping of isolates confirmed the existence of phylotypes A and B (Figure S1C).

Recent clinical studies using 16S rDNA sequencing from tissue had shown the presence of *Parvimonas* sequences in tumoral and normal homologous colonic mucosa of CRC patients, suggesting that this microorganism may be able to adhere to colonocytes or the colonic cellular matrix (Flemer et al., 2017; Gao et al., 2017). We therefore measured the ability a representative isolate from each phylotype, *PmA* (phylotype A) and *PmB* (phylotype B), to adhere to different extracellular matrix (ECM) proteins and three human colonic cell lines. *PmA* adhered to the ECM proteins collagen I, collagen IV, fibronectin and laminin, as well as to Matrigel® but not to fibrinogen, whereas *PmB* did not adhere to any of these ECM proteins (Figure 1B). Similarly, *PmA* adhered 6-, 2-, and 3-fold more to Caco-2, HT-29, and HCT116, respectively as compared to *PmB* ($p<0.05$) (Figure 1C), indicating heterogeneity of adhesion capacity in *P. micra* phylotypes.

To explore whether differences in the aggregation phenotype could be due to specific bacterial surface structures, ultrathin sections of the two phylotypes were analyzed by positive contrast transmission electron microscopy. Spiky surface structures were observed only in *PmB* (non haemolytic, "compact") and not in *PmA* (haemolytic, "non-compact") (Figure 1D).

These data show that two phenotypically and genetically different phylotypes can be described for *P. micra*.

***P. micra* phylotype A is associated with CRC**

Differences in the *in vitro* adhesion capacity of the two phylotypes might reflect a different colonization ability *in vivo*. To assess this hypothesis, we first improved the resolution of our whole genome sequencing (WGS) metagenomic analysis of the *P. micra* sequences carried in CRC patients' feces

(Zeller et al., 2014) focusing on the taxonomic assignation of raw data in order to precisely investigate microbial composition.

We observed an enrichment (27% versus 1.1%) of *P. micra* in the feces of CRC patients compared to control healthy individuals. Three taxa of *P. micra* were detected: *'Parvimonas micra*' from phylotype A, and '*Parvimonas* sp. oral taxon 393 - F0440' and '*Parvimonas* sp. oral taxon 110 - F0139' from phylotype B. All 27% of *Parvimonas* positive patients carried the phylotype A in their feces, except in one patient where reads of both phylotypes were detected.

We then analyzed *Parvimonas* carriage in feces of CRC patients and control individuals, as well as from paired tumor and normal homologous tissues (adjacent to the tumor), using 16S rDNA sequencing of the V3-V4 region. As expected, the occurrence of *Parvimonas* in feces was enriched in CRC patients (81%) as compared to control individuals (60%) (Figure 2A), while *Parvimonas* carriage increased at early stages of CRC (I & II) according to TNM (tumor-node-metastasis) staging of CRC (Figure 2A), an observation consistent with previous WGS results. In 78% of CRC patients, *Parvimonas* sequences were detected in both tumoral and normal homologous samples with a significant enrichment of the bacterium in the tumoral tissues (means of 2.2% versus 0.53% respectively; $p<0.01$) (Figure 2B). Prevalence of *Parvimonas* in tissues showed no difference according to the TNM staging, with 74% in stages I-II versus 84% in stages III-IV and a respective abundance of 1.34% and 1.41% (Figure 2B).

The V3-V4 regions of the 16S rDNA was then analyzed to discriminate between *Parvimonas* phylotypes A and B. Forty-four percent of control individuals carried the phylotype A in their feces compared to 61% of the CRC patients ($p<0.01$), whereas no difference in the prevalence of the phylotype B was observed (22% and 27%, respectively) (Figure 2C). An increase in the abundance of the phylotype A, but not of the phylotype B, was also observed in tumor tissues compared to homologous normal tissues ($p<0.05$) (Figure 2D).

Previously, we demonstrated a strong correlation between *Parvimonas micra* carriage in feces and a positive cumulative methylation index (CMI) in the blood based on the methylation levels of promoters in the WIF1 gene involved in carcinogenesis and PENK, and NPY, two neuromediators genes (Sobhani et al., 2019). Thus, we wondered whether the phylotypes herein described, were correlated with the CMI. Analysis of 16S rDNA sequencing data from feces showed that phylotype A was correlated with a positive CMI ($p<0.01$) but this was not the case for phylotype B (Figure 2E). *Parvimonas* carriage in tissues was

also correlated with a positive CMI ($p<0.05$), although the number of samples was not sufficient to discriminate between the two phylotypes (Figure 2F).

These results indicate that *Parvimonas micra* phylotype A sub-species is associated with CRC.

P. micra **impact on human colonic primary epithelial cells**

Colorectal tumor is defined as a group of anarchically proliferating cells with DNA mutations and aberrant DNA methylation. To determine whether *P. micra* has an impact on host cells, we first developed a compatible co-culture model between this oxygen-sensitive anaerobic microorganism and primary human colonic epithelial cells. Colonic samples were obtained from normal site (far from tumor tissue) of two colonic tumor resections. Colonic crypts were isolated and cultured as organoids (Figure 3A), allowing amplification of initial material through stem cell proliferation. To develop a system where bacteria could be in contact with the apical pole of primary cells, reflecting a physiological situation, organoid fragments were seeded onto a Transwell® insert system and grown to confluency up to a polarized monolayer (Figure 3B), as previously described (Noel et al., 2017). Cell growth and formation of a tight monolayer were monitored over time by phase-contrast microscopy and transepithelial electrical resistance (TEER) measurements (Figures 3C and D). Under these culture conditions, monolayers were fully confluent three days after organoids fragments were seeded on the inserts (Figure 3D).

P. micra has been described as an anaerobic microorganism (Kremer et al., 1997) and little is known about its oxygen tolerance. To determine compatible oxygenation conditions with eukaryotic cell growth requirements, *P. micra* was cultivated at various oxygen concentrations (0, 2, or 21% O_2) and viability was assessed by Colony Forming Units (CFU) counting at different time points. *P. micra's* viability dropped by 70% in less than ten minutes in aerobic conditions (21% O_2), whereas in hypoxic conditions (2% O_2) it was able to grow albeit at half the rate compared to full anaerobic conditions (Figure S2A). Thus, primary cells were grown for four days in aerobic conditions to allow monolayer formation, followed by three days in hypoxia (2% O_2). The primary cell monolayer's integrity was maintained in these conditions (Figures 3C and D) and showed all characteristics of a differentiated epithelium as assessed by the presence of tight junctions, differentiated epithelial cells, colonocytes, goblet cells, and enteroendocrine cells (Figure 3E).

To investigate the possible involvement of *P. micra* in early oncogenic processes, we co-cultivated primary cell monolayers with the two reference isolates, *PmA* and *PmB,* as well as *F. magna*, which is the closest phylogenetic taxon to *P. micra*, as a control. One day prior to co-culture, cells were placed in hypoxic conditions (2% O_2) to allow pre-adaptation. On the fourth day of the culture, cells were co-cultured with bacteria using a multiplicity of infection (M.O.I.) of 250 and incubated at 2% hypoxic conditions for 48 hours prior to analysis (Figure 3F). Bacterial growth was monitored over time by CFU counting of co-culture supernatants. After 48 hours of co-culture, *PmA* showed a 5-fold increase but did not impact cell viability, while *PmB* and *F. magna* showed no significant growth but had a more toxic effect, inducing a 30% decrease in cell viability (Figure 3G). Cell proliferation was measured by Ki-67 quantification upon immunofluorescent staining and no significant differences were observed between the various groups (Figure S2C, F). Because goblet cells are known to be impacted by intestinal microbes (Deplancke and Gaskins, 2001), we quantified their proportion and no significant differences were observed (Figure S2D, G). Moreover, because virulent bacteria can stimulate NF-κB and induce DNA breaks (Nougayrède et al., 2006) the proportion of cells with DNA double-strand breaks was measured and was found unchanged (Figure S2E, H). Yet, as compared to the NS condition, a significant increase in nuclear NF-κB was observed with both *PmA* ($p<0.01$) and *PmB* ($p<0.05$), but not *F. magna*, (Figure 3H). Thus, in the time frame of the experiments, *P. micra* did not induce DNA breaks despite activation of the NF-κB pathway and the possible induction of an inflammatory program in exposed cells. These results suggested that *P. micra* might contribute to host epithelial cell mutagenesis through a NF-κB pathway-mediated epigenetic effect.

P. micra **phylotype A induces DNA methylation changes**

Global DNA methylation was measured using a 5-methyl-cytosine dosage assay. *In vitro* co-cultures of both *P. micra* phylotypes showed a significant increase in global DNA methylation of primary cells as compared to NS condition ($p<0.05$) or to *F. magna* (Figure 4A). To identify affected genes, a genome-wide DNA methylation analysis was performed on the human colonic primary cells exposed or not to *P. micra* phylotypes or *F. magna*, using the Infinium MethylationEPIC BeadChip from Illumina. The methylation status of >850,000 CpG sites was assessed at all colonocyte genome targeting sites already known to be methylation-sensitive (Data set ready on Mendeley Data with the following DOI:

10.17632/nwj3bgbg5m.1). These were then classified by context (in or outside of CpG islands) and regions: body genes (3' or 5' UTR, first exon, exon bond, and body), promoters (200 or 1500 bases upstream of the transcriptional start site [TSS200 and TSS1500]) or intergenic. Using the LIMMA paired test we identified hyper- and hypo-methylated CpG sites from colonic primary cells challenged with bacteria compared to non-stimulated cells. CpG sites were then sorted into four categories: CpGs in promoters and within CpG islands; CpGs in promoters and outside CpG islands; CpGs in body genes (promoters region excluded) and within CpG islands; as well as CpGs in body genes outside of CpG islands. Because aberrant DNA methylation in promoters of tumor suppressor genes (TSGs) is a hallmark of CRC, combined p-values of differentially methylated CpG sites within the same gene were determined based on these four categories. We found that DNA methylation modifications were induced by all three bacteria while *PmB* promoted the highest number of hypo- or hyper-methylated genes (Figure 4B and C). Focusing on CpGs located in promoters and within CpG islands, we identified 15, 90, and 39 differentially methylated genes between non-stimulated cells as compared to cells co-cultivated with *PmA*, *PmB*, and *F. magna,* respectively. Venn diagram revealed six genes in common between *PmB* and *F. magna*, only one in-between the two phylotypes of *P. micra,* and two genes for *PmA* and *F. magna* (Figure 4D). Strikingly, 60% of the modified genes were either oncogenes, tumor suppressor genes, or genes involved in epithelial-mesenchymal transition (EMT), for PmA versus only 10% for *PmB* and 18% for *F. magna* (Figure 4E). In addition, cells co-cultivated with *PmA* presented hyper-methylation of several tumor suppressor gene promoters such as SCIN, HACE1, TSPAN13, FBXO32, IGFBP7, SIX1 or CXXC5. Except for the KIAA0494 gene that codes for an uncharacterized protein, all of the genes with a modification of the methylation status of their promoters induced by *PmA* are involved in carcinogenesis, particularly in EMT processes (Table S3) or cytoskeleton remodeling. These results suggest that *P. micra* phylotype A induces modifications in the expression of a specific gene set through the selective epigenetic modulation of gene promoters among which we identified well-characterized carcinogenesis regulators.

Discussion

Recently, we and others observed an epidemiological association between colorectal cancer and several oral bacteria such as *F. nucleatum*, *Porphyromonas gingivalis*, *Solobacterium moorei*, *Peptostreptococcus stomatis*, *Gemella morbillorum,* and *P. micra*, through 16S rDNA sequencing (Baxter

et al., 2014; Flemer et al., 2017) or metagenome analysis (Yu et al., 2017; Schmidt et al., 2019; Sobhani et al., 2019), of fecal samples. Some of these oral bacteria were also associated with colonic tissues of CRC patients (Gao et al., 2015; Nakatsu et al., 2015; Flemer et al., 2017), where they were more abundant in tumors than in adjacent healthy tissues (Nakatsu et al., 2015; Gao et al., 2017). In the present study, we focused on *P. micra,* at the subspecies level, and quantitatively measured its carriage in feces and in association with colonic tissues in a large cohort from Henri Mondor hospital (Directed by Prof. I. Sobhani) that includes normal colonoscopy, adenomatous and CRC patients. We observed, as others have (Xu et al., 2020), that *P. micra* is enriched in CRC from early stages (I and II of TNM classification), but not in feces of patients with adenomas, benign tumors considered as precursor of CRC (Fearon and Vogelstein, 1990). Since *P. micra* sequences were associated with adenoma tissues, we hypothesized that *P. micra* could be involved in the early steps of colon carcinogenesis. If *P. micra* is involved in promoting and/or accelerating carcinogenesis, *P. micra*, a common habitant of the oral cavity, would also have to be adapted for growth in the carcinogenesis-modified gut microenvironment. Indeed, oral microbiota such as *P. micra, F. nucleatum, P. stomatis,* and *G. morbilorum* were observed in biofilms-like structures on colonic mucosa of CRC patients and healthy subjects, (Dejea and Sears, 2016; Dejea et al., 2014; Drewes et al., 2017) that could be a way for these pathobionts to colonize the colonic mucosa (Horiuchi et al., 2020; Koliarakis et al., 2019; Li et al., 2017).

A recent *in vivo* study reported that $Apc^{Min/+}$ mice orally gavaged with *P. micra* exhibited a significantly higher tumor burden and were associated with altered immune responses and enhanced inflammation (Zhao et al., 2020). However, the mechanisms and the bacterial factors involved remain unknown. To move from association to causality, the first step is to select the most relevant genotype/phylotype associated with CRC and to develop *in vitro* models to assess the interaction of *P. micra* with host cells. Here we succeeded in developing a physiologically-relevant, low-oxygen *in vitro* co-culturing model of bacteria with primary human colonic intestinal cells to assess the effect of *P. micra* on colonic cells. Using this novel methodology, we established an epidemiological association of *P. micra* phylotype A with CRC. Notably, this cutting-edge cellular model offers a novel means to interrogate the effects of similar other pathobionts on colonocytes. Using this novel methodology, no impact of *P. micra* was observed on cell proliferation or differentiation, nor detected significant DNA damaging capacity. In contrast, *P. micra* co-culture induced the activation of the central transcription factor NF-κB, a key modulator of inflammation

and immune responses known to be constitutively activated in most cancers. Congruent with this observation, *P. micra*, like other oral pathobionts, have been shown to be associated with the CMS1 subtype of CRC, which makes up 14% of all CRC cases (Purcell et al., 2017), and with an over-activation of genes involved in immune responses. The CMS1 tumors, also called “immune subtype”, are characterized by a strong immune cell infiltration of CD8+ cytotoxic T cells, CD4+ T helper cells, and natural killer cells (Becht et al., 2016; Guinney et al., 2015). *P. micra* and inflammatory responses therefore seem to be key aspects in colon carcinogenesis.

Apart from specific mutations that characterize cancerous cells, epigenetic modifications of DNA and methylation of tumor suppressor genes (TSGs) promotors are essential in colonic carcinogenesis. We recently showed that CRC-associated dysbiotic feces transplanted to mice caused epigenetic changes similar to those observed in human tumors and the occurrence of murine colonic crypt aberrations (Sobhani et al., 2019). Furthermore, the CMS1 CRC subtype, in which *P. micra* and other oral bacteria enrichment has been observed (Purcell et al., 2017), is associated with the phenotype of hypermethylation of CpG islands in tumor suppressor gene promoters (high CIMP) (Guinney et al., 2015). Consistent with this hypothesis, Xia *et al.* observed an associations between enrichments of *F. nucleatum* or *Parvimonas spp.* and promoter methylation of several TSGs in tumoral tissue, (Xia et al., 2020), although they did not focus on *Parvimonas* subspecies. It is therefore tempting to speculate that *P. micra* might have a similar driver role in CRC by inducing hypermethylation in the promoter regions of keys TSGs.

In the present study, we report for the first time that *P. micra* increased global DNA methylation of target host cells using an *in vitro* co-culture model of human primary colonic cells. Notably, by comparing different *P. micra* phylotypes, we established a *P. micra* phylotype A-dependent signature of methylated promoters whose gene function converges towards the regulation of the cytoskeleton (Table S3) and include TSGs or genes involved in epithelial-mesenchymal transition (EMT) (Table S3). For example, Scinderin gene (SCIN) coding for a Ca^{2+}dependent actin-severing and capping protein, is involved in the regulation of actin cytoskeleton and known to be overexpressed in CRC (Lin et al., 2019); Tetraspanin 13 (TSPAN13) is a tumor suppressor gene coding for a transmembrane signal transduction protein that regulates cell development, motility, and invasion (Lou et al., 2017); DIAPH3 gene, is a major regulator of actin cytoskeleton involved in cell mobility and adhesion (Rana et al., 2018); Semaphorin 3F

(SEMA3F), is a tumor suppressor gene coding a secreted protein involved in cytoskeletal collapse and loss of migration (Shimizu et al., 2008); and SASH1 is a TSG coding for a scaffold protein involved in the TLR4 signaling, and known to interact with the actin cytoskeleton to maintain stable cell-cell adhesion (Martini et al., 2011). Notably, the hypermethylation of the TSG *HACE1* gene promoter was also observed upon co-culture with *P. micra* phylotype A. HACE1 is an E3 ubiquitin ligase controlling RAC1 stability, a small GTPase of the Ras superfamily involved in cell motility (Torrino et al., 2011) and the TSG HACE1's promoter was previously shown to be hypermethylated in CRC (Zhang et al., 2007) with a broad down-regulated expression in about 50% of all primary human tumors (Hibi et al., 2008). Loss of *HACE1* expression predisposes animals to DSS-induced colitis and AOM-induced colon carcinogenesis (Tortola et al., 2016). Hence, phylotype A of *P. micra* may cause epigenetic modifications that prepare or enhance cell transformation through cytoskeleton rearrangement.

In conclusion, here we described the *P. micra* phylotype A as the most prevalent CRC-associated Parvimonas, characterized by haemolytic capacity and adherent properties. Bacterial haemolysins have been shown to cause epigenetic marks in histones (Hamon et al., 2007) but studies on DNA methylation changes are lacking.

Clearly, further investigations on cross-talks between pathobionts, including *P. micra*, and the human colonic epithelium are warranted. The next step will be to identify bacterial effectors that trigger adhesion and directly or indirectly mediate epigenetic alterations in host cells.

Acknowledgments

The authors thank patients for having contributed to the present research program. We thank the bacteriology Departments of Cochin, Pitié Salpétrière, and Mondor Hospitals from APHP for providing *P. micra* clinical isolates (particularly Drs Asmaa Tazi, Alexandra Aubry and Biba Nebbad) and the Collection of Institut Pasteur for providing reference strains. We thank Claude Parsot and Shaynoor Dramsi for their dedication and inspiration in the development of this work, and Pamela Schnupf for careful reading of the manuscript. We acknowledge the URC of Henri Mondor and of St Antoine hospitals art APHP, Center for Translational Science (CRT)-Cytometry and Biomarkers Unit of Technology and Service (CB UTechS) at Institut Pasteur for their supports. This study was performed with the financial support from the Université Paris Est Creteil (UPEC), PHRC (Vatnimad) of the French government, SNFGE (French society of

Gastroenterology Commad Support 2019), and from ITMO Cancer AVIESAN (Alliance Nationale pour les Sciences de la Vie et de la Santé, National Alliance for Life Sciences & Health) within the framework of the Cancer Plan (HTE201601) with the coordination of Mathias Chamaillard, who was supported by INCA, and executed within the frame of the Oncomix research program between Institut Pasteur and AP-HP.

Authors contributions

Clinical and translational concept IS; basic conceptualization of *in vitro* study PS, GN, and EB; methodology GN, EB, CP, SD, EL; investigation FD, TP, CP, GN and EB; formal analysis AD, DM, SV, GN and EB; writing – original draft GN and EB; writing – review & editing GN, PS and IS; funding acquisition and resources EL, PS, and IS; and supervision PS, IS and GN.

Declaration of interests

The authors declare no competing interests regarding these data and materials.

Figures Legends

Figure 1: *P. micra* exists in two phylotypes with phenotypic and genetic diversity.

A) Phylogenetic tree based on full length 16S rDNA sequences of 27 *P. micra* clinical isolates (named *PMX)* and 22 reference sequences from the NCBI database. Root: *Finegoldia magna* ATCC 29328. The infection type from which the 27 clinical isolates are derived are indicated by a color code; haemolysis and colony compaction abilities are indicated as present (Y) or absent (N). B) Adhesion capacity of *P. micra* ATCC 33270 (*PmA)* and HHM BlNa17 (*PmB)* to extracellular matrix proteins. Optical density at 595 nm represents measurement of bacterial adhesion. C) Adhesion capacity of *PmA* and *PmB* to different human colonic cell lines (TC7, HT-29, and HCT116) after 1 hour of co-culture. Bacterial adhesion was quantified by the analysis of fluorescent images of cells co-cultured with the two phylotypes and is reported as arbitrary fluorescent units (x100). Data are represented as mean ± SEM from three independent experiments. Mann-Whitney test **p<0.01; *p<0.05. D). Positive contrast transmission electron microscopy on ultrafine sections of *PmA* (top) and *PmB* (bottom). Spiky surface structures (arrow) were observed only in *PmB* and not in *PmA*. Scale: 200 nm. See also Figure S1 and Table S1.

Figure 2: *P. micra* phylotype A is associated with CRC.

Parvimonas abundance, determined by 16S rDNA sequencing of the V3-V4 region, in controls, adenoma (Ad.), or sporadic CRC patients at early (CRC I & II) and late (CRC III & IV) stages of carcinogenesis in A) feces (Test Mann-Whitney ***p<0.001; **p<0.01) or B) mucosa: on the left, normal homologous mucosa versus tumoral tissue of CRC patients (Wilcoxon's Signed Wilcoxon Rank Test for paired samples **p<0.01); on the right, the same results only in tumoral tissue and according to CRC stages development. C) *P. micra* A and B phylotypes abundance in control and CRC patients' feces, or D) associated with normal homologous mucosa and tumoral tissue of CRC patients. *P. micra* phylotypes association with CMI (cumulative methylation index) methylation score of the WIF1, PENK, and NPY genes promoters in E) controls and sporadic CRC patients' feces, or F) associated with tumoral tissue of CRC patients. Results are expressed as a percentage of count (number of sequences assigned to *Parvimonas* per number of total bacterial sequences X100) mean +/- SEM. Statistical analysis of C, D, E, and F using the Mann-Whitney test **p<0.01, *p<0.05. See Table S2.

Figure 3: *P. micra* impact on human colonic primary epithelial cells. A) Colonic organoids obtained from normal tissue of CRC patients and cultured in Matrigel® with growth factors-enriched medium. B) Schematic representation of Transwell® permeable insert used to cultivate monolayers of primary colonic epithelial cells derived from organoids. C) Representative phase contrast images of organoid monolayers at seeding (d-5), after 3 days of culture in aerobic conditions (d-2) and after 7 days of culture, including 4 days in aerobic conditions and 3 days in hypoxic conditions (d2). D) Transepithelial electrical resistance (TEER) at different days of culture. TEER experiments were performed on cells from two donors, for at least 4 experiments per donor. Results are expressed as a mean of Ohm per cm2 +/- SEM. Mann-Whitney test **p<0.01. E) The differentiation of the monolayers after 4 days in aerobic conditions and 3 days in hypoxic conditions (2% O_2) was assessed by confocal microscopy. Nuclei (DAPI) are shown in grey, actin (phalloidin) in red, microvilli (anti-villin) in yellow, tight junctions (anti-occludin) in dark blue, Colonocytes (anti-KRT20) in cyan, Goblet cells (anti-muc2) in green, and enteroendocrine cells (anti-ChrA) in magenta. Main panels, XY projection; right panels, YZ projection; bottom panels, XZ projection. F) Co-culture experimental design with *P. micra* ATCC 33270 (*PmA)*, HHM BINa17 (*PmB)*, and *F. magna*. G) Cell viability after 48 hours of co-culture under hypoxic conditions at 37°C in 5% CO_2. Living cells were quantified by flow cytometry and presented as the ratio of living cells in stimulated versus non-stimulated samples (NS). H) Quantification of NF-κB positive nuclei upon immunofluorescent labeling with an anti-p65 antibody and nuclear staining. Results are expressed as ratios of p65-positive nuclei over the total number of nuclei and normalized by the NS condition. In G and H, results are from cells derived from two donors with at least 3 independent experiments per donor and the data are represented as mean +/- SEM. Mann-Whitney test ***p<0.001, **p<0.01, *p<0.05. See also Figure S2.

Figure 4. *P. micra* phylotype A induces DNA methylation changes. A) Global DNA methylation dosage by quantification of 5-methyl-cytosine (5m-C) residues compared to total cytosines, after 48 hours of co-culture of bacteria with human colonics monolayers. Results are expressed as the percentages of 5-mC normalized to the non-stimulated samples (NS). Experiments were performed on cells from two donors, with at least 3 experiments per donor and data are represented as mean +/- SEM. Mann-Whitney test *p<0.05. B and C) Gene number and distribution of hypo- and hyper-methylated genes between *P.*

micra ATCC 33270 (*PmA*), *P. micra* HHM BlNA17 (*PmB*) phylotypes, or *F. magna* ATCC 29328 and the NS condition. Differentially methylated genes were obtained by calculating a combined p-value based on differentially methylated CpG sites located at specific genomic regions (TSS, in gene promoters; or non-TSS, not at promoter sites) and regions within CpG islands or not. D) Venn diagram of differentially methylated genes when only promoters (TSS) and CpG islands were considered. E) Heat map of fold-change of the top 15 most differentially methylated genes (hypo- and hyper-methylated) in the comparison of each bacterium- *versus* non-stimulated condition. CpG sites only in promoters (TSS) and in CpG islands were considered. Whether these genes belonged to oncogenes, tumor suppressor genes (TSGs) or epithelial-mesenchymal transition (EMT) databases, is indicated by a cross. See also Table S3.

Figure S1: Phenotypic and genetic characterization of *P. micra*. A) Colony appearance of ATCC 33270 (*PmA*) (top) and HHM BlNa17 (*PmB*) (bottom) on blood agar plates: i) after 48 hours of culture under anaerobic conditions at 37°C; ii) when the colonies were removed to observe the haemolytic zone below the colonies; and iii) when colonies were analyzed via the Gram stain. B) Sedimentation assay performed on *PmA* and *PmB*. On the right, representative pictures of the suspension (top *PmA*, bottom *PmB*) obtained after 2 hours at 4°C without shaking. C) Clustering of *P. micra* phylotype A and B whole genomes according to the presence (red) or absence (blue) of genes.

Figure S2: A) *P. micra* oxygen sensitivity. Representative growth curves of *P. micra* ATCC 33270 (*PmA*) under anaerobic (0% O_2), aerobic (21% O_2) and hypoxic (2% O_2) conditions. At different incubation times, the cultures were plated on horse blood plates and returned to anaerobic conditions for 48 hours. CFUs were counted to estimate bacterial viability. The results are expressed as percentage of CFU obtained towards the inoculum grown at 0% O_2. B) *P. micra* viability by CFU counts after 48 hours of co-culture with human colonic primary cells in hypoxic conditions (mean +/- SEM). C, D, and E. Quantification of proliferative cells, cells harboring double-stranded DNA breaks, and goblet cells by immunofluorescence using respectively an anti-Ki67, anti-γH2ax, and anti-Muc2 labelling after 48 hours of co-culture with *P. micra* ATCC 33270 (*PmA*), *P. micra* HHM BlNa17 (*PmB*), *F. magna* ATCC 29328 or the *E. coli Pks*+ strain IHE3034 (used as a positive control for double-stranded breaks). NS, non-stimulated. Results are

422 expressed as percentages of positive cells relative to the total number of nuclei and normalized using the
423 non-treated sample (NS). Experiments were performed on cells from two donors, with at least 3
424 experiments per donor. The data are represented as mean +/- SEM. Mann-Whitney test ***p<0.001. F,
425 G, and H are representative images of Ki67, γH2ax, and Muc2 immunostaining, respectively.
426
427 **Table S1:** Description of *P. micra* clinical isolates obtained from several hospitals in Paris and from
428 different infectious locations.
429
430 **Table S2**: Clinical data.
431
432 **Table S3: *PmA* induces DNA methylation changes in promoters of genes involved in**
433 **carcinogenesis.** CRC: colorectal cancer; TSG: tumor suppressor gene; EMT: epithelial-mesenchymal
434 transition; IGFs: insulin-like growth factors; SCF: SKP1-CUL1-F-box protein; TRAPP: transport protein
435 particle; ESCC: esophageal squamous cell carcinoma.

Materials and Methods

Patient recruitment at the Créteil Henri Mondor Hospital

Participants were selected from different cohorts recruited with informed consent between 2004 and 2018 by the endoscopy department at Henri Mondor hospital (Créteil) where patients had been referred for colonoscopy; detailed description in (Sobhani et al., 2011). Participants with previous colon or rectal surgery, colorectal cancer, inflammatory bowel diseases, or with a genetic form of CRC were excluded. Individuals exposed to antibiotics or probiotics within four weeks before collection or suffering from acute diarrhea were also excluded from the study. Tumor Node Metastasis (TNM) stages of colonic neoplasia were determined by radiological examinations and analyses of the surgical specimen by the anatomopathological department of Henri Mondor Hospital. Clinical parameters of the patients, such as body mass index (BMI), age, sex, and disease history were referenced. The cumulative methylation index (CMI) score was previously determined (Sobhani et al., 2011). This score was calculated from the methylation status of three genes (*wif1*, *penk* and *npy*) involved in colorectal carcinogenesis and was considered negative for CMI<2 or positive for CMI≥2. Clinical data are summarized in Table S2.

Fecal and tissue samples

Fresh feces were collected between 2 weeks and 3 days prior to colonoscopy and ten grams were frozen at -20°C for 4 hours and then stored at -80°C until use. Paired samples of colorectal tumor tissue and homologous normal mucosa (more than 15 cm from the margin of tumor resection) were collected within 30 minutes after surgical resection and immediately frozen in liquid nitrogen and stored at -80°C until further use.

DNA extraction and quantification

DNA extraction from fecal samples was performed using the G'NOME DNA isolation kit® (MP Biomedicals) according to the supplier's instructions, with modifications as described in (Furet et al., 2009). DNA extraction from tissue samples was performed from eight 50 µm cryosections of nitrogen frozen tissue using the QIAamp PowerFecal DNA Kit® (Qiagen) following the supplier's instructions with an additional lysis step (0.1 mm diameter silica beads were added to the lysis solution provided and shaking was performed at maximum speed for 10 minutes in a vibratory shaker). Fecal and tissue DNA concentrations were determined by Qubit® fluorometer (ThermoFisher) and stored at -20°C until use.

Whole genome sequencing (WGS) of fecal samples

Metagenome sequencing was performed by the high-throughput platforms BIOMICS (Institut Pasteur, Paris, France) and EMBL (Heidelberg, Germany). Sequencing was performed in pairs, using the HiSeq 2000/platform 2500 equipment, over a length of 100 bp of DNA and at a sequencing depth of 5 Gbp (Zeller et al., 2014). The raw data have already been reported in several papers (Schmidt et al., 2019; Sobhani et al., 2011; Zeller et al., 2014). For this study, one hundred and sixty-six fecal samples from the CCR1 and DETECT cohorts were considered. The Diamond/MEGAN6 bioinformatics pipeline (Bağcı et al., 2019) was used for metagenomic assignment. Sequences were filtered for an average quality (Phred score) greater than 20 over a window of two consecutive bases and a length greater than 100 bp using Trimmomatic software (version 0.35). Good quality sequences were translated in the six possible reading frames and aligned to the NR reference library (RefSEQ non-redundant protein database) in which proteins/peptides that are more than 99% similar are combined into a single organism-associated group with a specific identifier (O'Leary et al., 2016). A sequence can be assigned to several taxa. The MEGAN6 LCA (Lowest Common Ancestor) algorithm (Huson et al., 2016) was used to resolve multi-mapping reads: when a sequence was assigned to two (or more) taxa of different phylogeny, it was assigned to the top taxonomic level in common.

16S rDNA sequencing of tissue samples

For this study, 71 pairs of homologous tissue samples (matched tumor and healthy tissue) were analyzed. After amplification of the 16S rRNA gene V3-V4 region, pairwise sequencing was performed to a length of 250 bp on the Illumina MiSeq platform. The resulting raw database was cleaned of sequences corresponding to human or phage sequences. Adapters and primers, as well as 5' and 3' ends with a Phred quality score over a 2 bp sliding window of less than 30 were removed using Trimmomatic (version 0.35) software. Paired sequences were merged using the FLASH2 tool (version 2.2.00). Taxonomic assignments were performed using MALT/MEGAN6 bioinformatic pipeline (Huson et al., 2016) with default parameters and the SILVA rRNA database version 123 (Quast et al., 2012). OTUs were clustered at a threshold of 97% identity.

Colonic cell cultures

Cancer cell lines: human colon cell lines HT-29, HCT116 and Caco-2 (subclone TC7), were grown in DMEM ("Dulbecco Modified Eagle Medium") 1 g/L glucose (Gibco), supplemented with 20% (vol/vol)

493 decomplemented fetal bovine serum (FBS) (Gibco), 1% non-essential amino acids (Gibco), 1%
494 GlutaMAX (Gibco), at 37°C in 10% CO_2.

495 **Human colonic organoids**

496 Human colonic surgical resection specimens were obtained from Henri Mondor Hospital from two patients
497 who had undergone colectomy surgery from rectal adenocarcinoma and had given their informed consent
498 (agreement N°2012-37). Tissues samples from the normal site (far from tumor tissue) were sterilely
499 washed using PBS supplemented with gentamicin (50 µg/mL), Normoxin (1 mg/mL) and amphotericin B
500 (2 µg/mL), to obtain bacteria-free normal colonic mucosa and kept in 0.1% PBS-BSA at 4°C for the
501 duration of transport (approximately 4 hours). Crypts were isolated according to the protocol of Sato et
502 al, (Sato et al., 2011), with some modifications. The epithelium was stripped of underlying muscularis and
503 serosa, washed several times in cold PBS until the supernatant was clear, and cut into 5 mm^2 fragments.
504 The fragments were incubated for 20 minutes on ice in cold chelation buffer consisting of 5.6 mmol/L
505 Na_2HPO_4, 8.0 mmol/L KH_2PO_4, 96 mmol/L NaCl, 1.6 mmol/L KCl, 44 mmol/L sucrose, 54.8 mmol/L D-
506 sorbitol, and 0.5 mmol/L DL-dithiothreitol in distilled water plus 2 mmol/L EDTA. After removal of the
507 supernatant and addition of cold chelation buffer without EDTA, the tissue fragments were vigorously
508 resuspended by several passes through a 10 mL serological pipette. After sedimentation of the tissue
509 fragments, the supernatant, containing the crypts, was recovered. The resuspension/sedimentation
510 procedure was repeated twice. The tissue fragments were again incubated in chelation buffer with EDTA
511 for 15 minutes and the resuspension/sedimentation procedure was repeated three times. Fractions
512 containing crypts were pooled and centrifuged at 300g for 5 minutes. The crypt pellet was resuspended
513 in Matrigel® growth factor-reduced medium (Corning) diluted to 75% in culture medium (see composition
514 below). Four 25 µl drops of the Matrigel-crypt mixture were placed per well in 12-well plates with
515 approximately 100 crypts per drop. The plates were incubated for 15 minutes at 37°C to allow
516 polymerization of the Matrigel, then 800 µL of culture medium was added and the plates were incubated
517 at 37°C and 5% CO_2. The culture medium consisted of AdvancedDMEM/F12 (Gibco), 10 mM HEPES
518 (Gibco), 1X GlutaMAX (Gibco), 100 υ/mL penicillin, 100 µg/mL streptomycin (Gibco), 1X N2
519 (ThermoFisher), 1X B27 (ThermoFisher), 1 mM N-acetyl-L-cysteine (Sigma), 100 ng/mL Noggin (R&D
520 systems), 50 ng/mL recombinant human EGF (R&D systems), 150 ng/mL Wnt-3A (R&D systems), 1
521 µg/mL recombinant human R-spondin-1 (R&D systems), 500 nM A83-01 (R&D systems), 10 mM

nicotinamide (Sigma), 10 µM SB202190 (Sigma), 10 nM Gastrin I (Sigma), 3 µM CHIR99021 (Biogems), 10 µM Y27632 (Sigma) and 10% decomplemented fetal bovine serum (vol/vol) (Gibco). The culture medium was changed every two days, without Y27632. Each week, organoids were split with a 1:3 ratio. The medium was replaced by cold fresh medium and the Matrigel drops and the organoids were dissociated by pipetting several times through a 200 µl tip and collect in a tube. After a spin for 5 minutes at 300g at 4°C, the pellet was washed with fresh cold medium, resuspended with a p200 pipette several times, and spun again. The pellet of organoid fragments was resuspended and diluted in Matrigel® as described above. After expansion, organoids were frozen and stored at -80°C in freezing medium composed of AdvancedDMEM/F12, 10% fetal bovine serum, 10% DMSO (PAN Biotech), 10 µM Y-27632 and 3 µM CHIR99021.

Generation of human colonic epithelial monolayer

Monolayers of organoid-derived primary cells were generated as previously described (Noel et al., 2017). Briefly, the twenty-four well plates with 0.4 µm pores (Corning) Transwell® insert system were used as culture support for primary cells in monolayers. Inserts were previously coated with 50 µl of human collagen IV at 30 µg/ml (Millipore) overnight at 4°C, then washed with DMEM. Approximately three hundred organoids at five days of growth were dissociated into fragments (using Cell Recovery (Corning)), resuspended in 200 µl of culture medium and put on each insert. The culture medium consisted of DMEM/F12 (Gibco), 10 mM HEPES (Gibco), 1X GlutaMAX (Gibco), 100 u/mL penicillin, 100 µg/mL streptomycin (Gibco), 1X N2 (ThermoFisher) and B27 (ThermoFisher) supplements, 1 mM N-acetyl-L-cysteine (Sigma), 100 ng/mL noggin (R&D systems), 50 ng/mL recombinant human EGF (R&D systems), 150 ng/mL Wnt-3A (R&D systems), 1 µg/mL recombinant human R-spondin-1 (R&D systems), 500 nM A83-01 (R&D systems), 10 mM nicotinamide (Sigma), 10 µM SB202190 (Sigma), 10 nM Gastrin I (Sigma), 3 µM CHIR99021 (Biogems), 10 µM Y27632 (Sigma) and 10% decomplemented fetal bovine serum (vol/vol) (Gibco). Monolayers were incubated at 37°C in 5% CO_2. After three days of culture, the medium was changed without addition of Y-27632. Cell growth and monolayer closure were monitored daily by observation with an inverted brightfield microscope (IX81, Olympus) and by recording the transepithelial electrical resistance (Millipore).

Bacterial strains, culture and characterization

Escherichia coli pks+ strain IHE3034 was used as a positive control for double-strand DNA breaks. *Shigella flexneri* 5a BS176, lacking the virulence plasmid, was used as a negative control for bacterial adhesion and autoaggregation negative control. The reference strains ATCC 33270 of *P. micra (PmA)* and ATCC 29328 of *Finegoldia magna* were obtained from the Pasteur Institute Collection (CIP). In this study, twenty-seven clinical isolates of *P. micra*, including HHM BINa17 (*PmB*), were collected from the bacteriology departments of three Parisian hospitals: Henri Mondor, Cochin, and Pitié Salpetrière (Table S1). The identification of the different isolates was confirmed by mass spectrometry and 16S rDNA sequencing analysis (Eurofins Genomics) and sequences were deposited on GenBank (Table S1). 16S rDNA whole sequence from positions 112 to 1302 were aligned with the ClustalW program (EMBL) and the phylogenetic tree was calculated using the PhyML program. Pan-genomic comparison: available complete *P. micra* genomes were downloaded from the NCBI database. Analysis of pangenomes and comparison of common or strain-specific gene carriage was performed using the Roary program (Page et al., 2015) with a 95% identity threshold and the use of default parameters.

Culture:

Bacteria were grown under anaerobic conditions in TGY-V enriched broth (Trypticase peptone 30 g/L, D-glucose 10 g/L, Yeast extract 20 g/L, L-cysteine-HCl 0.5 g/L, haemin 5 mg/L, vitamin B12 5 µg/L, menadione sodium bisulfite 500 µg/L, Thiamine 1 mg/L, nicotinic acid 1 mg/L, riboflavin 500 µg/L, p-aminobenzoic acid 100 µg/L, biotin 25 mg/L, calcium pantothenate 1 mg/L, pyridoxamine 2HCl 500 µg/L, folic acid 500 µg/L) or on Columbia agar solid support enriched with 5% horse blood (COH) (Biorad). Anoxic conditions were generated in Gaspack (BD) or anaerobic cabinet (Don Whitley DG250 Workstation). Hemolysis ability and colonies compactions were assessed by observing CFUs on COH plates after 48 hours at 37°C in anaerobic conditions: i) contact haemolytic activity was defined by the presence of an haemolytic zone located underneath the colony and slightly beyond (Figure S1A) and ii) colonies compaction was defined as “non-compact” or “compact”: the “non-compact” colonies spread out on agar upon contact with a loop, whereas the compact colonies remained associated in clumps suggesting a strong interbacterial aggregative property. For the sedimentation assay, bacteria were grown on blood agar plates under anaerobic conditions at 37°C for 48 hours, resuspended in PBS, and incubated at 4°C statically. The optical density (600 nm) in the upper half of the test tubes was measured periodically (Figure S1B).

Transmission electron microscopy:

Bacteria from a 4 day-old cultures were fixed by adding glutaraldehyde (2.5% final concentration) to the culture for 1 hour at room temperature (RT). The bacteria were washed twice with HEPES 0.1 M, pH 7.5 and postfixed with 1% osmium tetroxide in HEPES 0.1 M pH 7.5 for 1h30 at RT. After three washes in distilled water, the bacteria were dehydrated in a graded ethanol series (25, 50, 75, 95 and 100% ethanol), and gradually infiltrated in Epon resin. Ultrathin sections (60 nm) were obtained on a FC6/UC6 ultramicrotome (Leica). Sections were transferred to 200 Mesh Square Copper grids (CF-200-CuO, Delta Microscopy) formvar and carbon coated, stained with 4% uranyl acetate, and counterstained with Reynold's lead citrate for 20 minutes. Images were recorded with TECNAI SPIRIT 120 Kv, with a bottom-mounted EAGLE 4KX4K Camera.

Bacteria adhesion

Adhesion to extracellular matrix proteins: MaxiSorp 96-well flat-bottom plates (ThermoFisher) were coated overnight at 4°C under agitation with 1 µg per well, in triplicate, with the following extracellular matrix proteins: collagen I (Gibco), collagen IV (Millipore), fibronectin (Gibco), fibrinogen (Sigma), laminin (Sigma), and Matrigel® (consisting of a mixture of proteins, mainly laminin, collagen IV, entactin, and heparan sulfate proteoglycan) (Corning). Bacteria were grown on COH plates for 48 hours at 37°C under anaerobic conditions, collected and suspended in PBS at an OD of three units. The bacterial suspension was placed in the wells and the plate was centrifuged at 300g for 5 minutes. After incubation for 1 hour at RT, the wells were washed three times with PBS to remove non-adherent bacteria and then stained with 0.1% crystal violet (Sigma). The plates were incubated for 30 minutes at RT and washed three more times with PBS. The dye was dissolved with 20:80 acetone-ethanol and the optical density at 595 nm was read by a spectrometer (Infinite M200Pro-TECAN).

P. micra adhesion to colonic cell lines: human colonic cell lines (TC7, HT-29, and HCT116) were grown on glass coverslips in 24-well plates to confluence. In parallel, *P. micra* was grown on COH plates for 48 hours at 37°C under anaerobic conditions. Bacteria were washed in PBS, re-suspended in DMEM with 1 g/L glucose (Gibco), and put on the cells using a M.O.I. of 1. The plates were centrifuged for 5 minutes at 300g and incubated for 1 hour at 37°C. After incubation, the cells were washed three times in PBS, fixed in 4% paraformaldehyde (PFA) (Electron Microscopy Sciences) for 30 minutes at RT, washed again three times in PBS, and then stored in 0.1% PBS-BSA at 4°C until labeling. Non-specific sites were

blocked with 1% PBS-BSA for 1 hour at RT. The cells were then incubated with rabbit anti-*Parvimonas* serum (on order by Covalab) at 1:1000 in 1% PBS-BSA for 30 minutes at room temperature. After washing, the cells were permeabilized with PBS- 0.2% triton for 30 minutes. Anti-rabbit secondary antibody coupled to Alexa Fluor 488 (ThermoFisher) at 1:400 and phalloidin coupled to Alexa Fluor 568 (Invitrogen) at 1:200 were used in 1% PBS-BSA for 45 minutes at RT. After washing, nucleic acids were labeled with 1 µg/mL DAPI (ThermoFisher) for 5 minutes. The coverslip with the cells were washed again and mounted on a slide with Prolong Gold (ThermoFisher). Images were taken using a slide scanner with 40X objective (AxioScan, Zeiss) and processed using Fiji (Schindelin et al., 2012). Bacterial adhesion was quantified as area of the anti-*Parvimonas* fluorescence signal over an entire field normalized by the total number of nuclei and reported as arbitrary fluorescent units x100.

Co-culture of colonic cells and *P. micra*

Bacteria were grown in TGY-V medium under anaerobic conditions at 37°C for 48 hours, followed by a 1:10 subculture for 24 hours to reach exponential phase. Bacterial density was adjusted to an M.O.I. of 250 in cell culture medium and then put on the cells. Twenty-four hours prior to bacterial exposure, the colonic cells were placed in 2% O_2 in H35 hypoxic cabinet (Don Whitley) to allow acclimatization. Co-cultures with primary cells were performed for 48 hours at 37°C under hypoxic conditions.

Flow cytometry

Cell monolayers were dissociated with TrypLE™ Express Enzyme (Gibco) at 37°C for ten to twenty minutes. The cells were recovered in DMEM/F12 at RT and then centrifuged for five minutes at 300g. After washing the pellet in PBS, cells were stained with the Live/Dead fixable cell stain kit (ThermoFisher) for twenty minutes on ice, following the supplier's instructions. Cells were analyzed using a FACS Attune (ThermoFisher) and data were analyzed using FlowJo software (v10.1).

Ki-67, γH2ax, Muc2 and NF-κB immunofluorescence

Following 48 hours co-culture with *P. micra*, cells were washed with PBS, fixed (4% PFA, 30 minutes), washed again, and stored (0.1% PBS-BSA) at 4°C until labeling. Cells were permeabilized with 0.5% Triton-PBS for 30 minutes at RT, washed in 0.1% PBS-Triton, and blocked with 1% bovine serum albumin (BSA)-0.1% Triton for 30 minutes. The cells were then incubated either i) with a 1:200 dilution of an antibody against the proliferation marker Ki-67 coupled to Alexa Fluor 666 (clone SolA15, eBioscience) and a 1:500 dilution of an antibody against the phosphorylated form of H2A histone family member X

(γH2Ax) (clone JBW301, Merck) and 1:200 phalloidin coupled to Alexa Fluor 568; ii) a 1:50 dilution of an antibody against nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) p65 (ab7970 Abcam) and 1:200 phalloidin coupled to Alexa Fluor 568; or iii) with a 1:100 dilution of an anti-mucin2 (Muc2) antibody (clone Ccp58, Abcam) and 1:200 phalloidin coupled to Alexa Fluor 568, in 0.1% Triton-BSA overnight at 4°C. After washing in 0.1% PBS-Triton, the cells were incubated with 1:400 secondary antibodies coupled to Alexa Fluor 488 (ThermoFisher) in 0.1% PBS-Triton for 45 minutes at RT. Cells were washed again in 0.1% PBS-Triton and the nucleic acids were labeled with 1 µg/mL DAPI for 5 minutes before cells were mounted on a slide with ProlongGold (ThermoFisher). Fluorescence images were taken using a confocal fluorescence microscope (Zeiss) equipped with an Opterra system (Bruker), and the number of cells was analyzed as the number of positive cells/total cells using Fiji (Schindelin et al., 2012) or Imaris 9.8 software (URL: http://www.bitplane.com/imaris/imaris) software.

DNA methylation studies

Human cells co-cultured with *P. micra* were subjected to DNA extraction using the DNeasy Blood & Tissue Kit® (Qiagen) according to the manufacturer's instructions and DNA was quantified using a Qubit® fluorometer.

Global DNA methylation level: MethylFlash methylated DNA Quantification Fluorometer Kit (Epigentek) was used to detect the global methylation level in DNA isolated from primary colonic cells co-cultured with *P. micra* ATCC 33270 (*PmA*), HHM BlNa17 (*PmB*) and *F. magna*. Assays were performed in duplicate. As instructed by the manufacturer, 100 ng of the isolated genomic DNA was bound to the assay well. The capture antibody, detection antibody, and enhancer solution were then added consecutively to the wells before the fluoro-developing solution was added and the relative fluorescence units (RFU) were measured (TECAN). The percentage of 5-methylcytosine (5-mC) relative to the total amount of cytosines in the sample was calculated to represent the global methylation dosage and was reported as a ratio to the non-stimulated samples.

DNA methylation profiling: DNA methylation profiling was performed on eight independent co-cultures of *PmA*, *PmB*, and *F. magna* with human primary colonic cells (four experiments per patient tissue). The Infinium MethylationEPIC BeadChip Kit (850K) (Illumina, San Diego, CA, USA) was used for a genome-wide methylation profiling to determine the DNA methylation status of >850,000 CpG sites (Moran et al., 2016). A total of 500 ng of genomic DNA from cells in co-cultures were bisulfite-treated using the

ZymoResearch EZ DNA Methylation kit (Zymo Research Corp, Irvine, CA, USA). The Infinium HD Methylation Assay (bisulfite modification, amplification, fragmentation, precipitation, hybridization, wash, extension, staining, and imaging) was performed according to the manufacturer's explicit specifications. The quality was supported by multiple quality control (QC) measures, including tests for proper bisulfite conversion, staining, and specificity of the internal controls, as determined by Illumina GenomeStudio software. Average-beta values (proxy for methylated DNA level between 0, unmethylated, and 1, fully methylated) were normalized to internal controls and corrected by background subtraction. Non-autosomal CpGs (n=14,522) and CpG probes with suboptimal detection ($p<0.05$ in at 80% of samples) (n= 29,316), as well as single nucleotide polymorphisms (SNPs) associated CpGs (n=59), were removed from our analyses. Beta values were transformed into M-values as described by (Du et al., 2010), to be able to use the Limma statistical model (Wettenhall and Smyth, 2004) package R Limma, version 4.0.2. The four conditions (Non-stimulated NS; ATCC; BlNa and FM) and individuals were kept as variables. Pairwise comparisons were obtained from a contrast matrix. No adjustment method was used. For each gene (n=25,560), a t-test on log-fold-change was performed on four categories of probes depending on their CpG context and gene localization: probes in CpG island context and 1/ localized in genes promoters (TSS1500+TSS200) or 2/ in body gene; and probes out of CpG island context 3/ in promoters or 4/ in body gene. Genes with >2probes, p-value <0.05, and with a log-FC >0,1 (absolute value) were considered significant. Using three different databases, the genes were classified as oncogenes, tumor suppressor genes, or involved in Epithelial-Mesenchymal Transition (ONGene (Liu et al., 2017), TSGene 2.0 (Zhao et al., 2016) or dbEMT 2.0 (Zhao et al., 2019).

References

Bağcı, C., Beier, S., Górska, A., and Huson, D.H. (2019). Introduction to the analysis of environmental sequences: Metagenomics with MEGAN. In Methods in Molecular Biology, (Humana Press Inc.), pp. 591–604.

Becht, E., De Reyniès, A., Giraldo, N.A., Pilati, C., Buttard, B., Lacroix, L., Selves, J., Sautès-Fridman, C., Laurent-Puig, P., and Fridman, W.H. (2016). Immune and stromal classification of Colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapy. Clin. Cancer Res. *22*, 4057–4066.

Bowler, P.G., and Davies, B.J. (1999a). The Microbiology of Acute and Chronic Wounds. Wounds *11*, 72–78.

Bowler, P.G., and Davies, B.J. (1999b). The microbiology of infected and noninfected leg ulcers. Int. J. Dermatol. *38*, 573–578.

Chan, D.S.M., Lau, R., Aune, D., Vieira, R., Greenwood, D.C., Kampman, E., and Norat, T. (2011). Red and processed meat and colorectal cancer incidence: Meta-analysis of prospective studies. PLoS One *6*, e20456.

Dejea, C.M., and Sears, C.L. (2016). Do biofilms confer a pro-carcinogenic state? Gut Microbes *7*, 54–57.

Dejea, C.M., Wick, E.C., Hechenbleikner, E.M., White, J.R., Mark Welch, J.L., Rossettid, B.J., Peterson, S.N., Snesrud, E.C., Borisy, G.G., Lazarev, M., et al. (2014). Microbiota organization is a distinct feature of proximal colorectal cancers. Proc. Natl. Acad. Sci. U. S. A. *111*, 18321–18326.

Deplancke, B., and Gaskins, H.R. (2001). Microbial modulation of innate defense: Goblet cells and the intestinal mucus layer. Am. J. Clin. Nutr. *73*, 1131S-1141S.

Drewes, J.L., White, J.R., Dejea, C.M., Fathi, P., Iyadorai, T., Vadivelu, J., Roslani, A.C., Wick, E.C., Mongodin, E.F., Loke, M.F., et al. (2017). High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. Npj Biofilms Microbiomes *3*.

Du, P., Zhang, X., Huang, C.C., Jafari, N., Kibbe, W.A., Hou, L., and Lin, S.M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics *11*, 1–9.

Edmiston, C.E., Walker, A.P., Krepel, C.J., and Gohr, C. (1990). The nonpuerperal breast Infection: Aerobic and anaerobic microbial recovery from acute and chronic disease. J. Infect. Dis. *162*, 695–699.

Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759–767.

Flemer, B., Lynch, D.B., Brown, J.M.R., Jeffery, I.B., Ryan, F.J., Claesson, M.J., O'Riordain, M., Shanahan, F., and O'Toole, P.W. (2017). Tumour-associated and non-tumour-associated microbiota in colorectal cancer. Gut *66*, 633–643.

Furet, J.P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., Doré, J., and Corthier, G. (2009). Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. FEMS Microbiol. Ecol. *68*, 351–362.

Gao, R., Kong, C., Huang, L., Li, H., Qu, X., Liu, Z., Lan, P., Wang, J., and Qin, H. (2017). Mucosa-associated microbiota signature in colorectal cancer. Eur. J. Clin. Microbiol. Infect. Dis. *36*, 2073–2083.

Goodwin, A.C., Destefano Shields, C.E., Wu, S., Huso, D.L., Wu, X.Q., Murray-Stewart, T.R., Hacker-

727 Prietz, A., Rabizadeh, S., Woster, P.M., Sears, C.L., et al. (2011). Polyamine catabolism contributes to
728 enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. Proc. Natl. Acad. Sci. U. S. A. *108*,
729 15354–15359.
730 Guinney, J., Dienstmann, R., Wang, X., De Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman,
731 P., Nyamundanda, G., Angelino, P., et al. (2015). The consensus molecular subtypes of colorectal
732 cancer. Nat. Med. *21*, 1350–1356.
733 Hamon, M.A., Batsché, E., Régnault, B., To, N.T., Seveau, S., Muchardt, C., and Cossart, P. (2007).
734 Histone modifications induced by a family of bacterial toxins. Proc. Natl. Acad. Sci. U. S. A. *104*, 13467–
735 13472.
736 Hibi, K., Mizukami, H., Shirahata, A., Goto, T., Sakata, M., Saito, M., Ishibashi, K., Kigawa, G., Nemoto,
737 H., and Sanada, Y. (2009). Aberrant methylation of the UNC5C gene is frequently detected in advanced
738 colorectal cancer. Anticancer Res. *29*, 271–273.
739 Horiuchi, A., Kokubu, E., Warita, T., and Ishihara, K. (2020). Synergistic biofilm formation by Parvimonas
740 micra and Fusobacterium nucleatum. Anaerobe *62*.
741 Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., and Tappu, R.
742 (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome
743 Sequencing Data. PLoS Comput. Biol. *12*, e1004957.
744 Jiao, S., Peters, U., Berndt, S., Brenner, H., Butterbach, K., Caan, B.J., Carlson, C.S., Chan, A.T., Chang-
745 Claude, J., Chanock, S., et al. (2014). Estimating the heritability of colorectal cancer. Hum. Mol. Genet.
746 *23*, 3898–3905.
747 Kamada, N., Chen, G.Y., Inohara, N., and Núñez, G. (2013). Control of pathogens and pathobionts by
748 the gut microbiota. Nat. Immunol. *14*, 685–690.
749 Koliarakis, I., Messaritakis, I., Nikolouzakis, T.K., Hamilos, G., Souglakos, J., and Tsiaoussis, J. (2019).
750 Oral bacteria and intestinal dysbiosis in colorectal cancer. Int. J. Mol. Sci. *20*, 4146.
751 Kremer, B.H.A., Magee, J.T., Van Dalen, P.J., and Van Steenbergen, T.J.M. (1997). Characterization of
752 smooth and rough morphotypes of Peptostreptococcus micros. Int. J. Syst. Bacteriol. *47*, 363–368.
753 Kyrgiou, M., Kalliala, I., Markozannes, G., Gunter, M.J., Paraskevaidis, E., Gabra, H., Martin-Hirsch, P.,
754 and Tsilidis, K.K. (2017). Adiposity and cancer at major anatomical sites: Umbrella review of the literature.
755 BMJ *356*, 477.
756 Li, S., Konstantinov, S.R., Smits, R., and Peppelenbosch, M.P. (2017). Bacterial Biofilms in Colorectal
757 Cancer Initiation and Progression. Trends Mol. Med. *23*, 18–30.
758 Lin, Q., Li, J., Zhu, D., Niu, Z., Pan, X., Xu, P., Ji, M., Wei, Y., and Xu, J. (2019). Aberrant scinderin
759 expression correlates with liver metastasis and poor prognosis in colorectal cancer. Front. Pharmacol.
760 *10*, 1183.
761 Liu, Y., Sun, J., and Zhao, M. (2017). ONGene: A literature-based database for human oncogenes. J.
762 Genet. Genomics *44*, 119–121.
763 Lou, J., Huang, J., Dai, X.Y., Xie, Y., Dong, M.J., Chen, B., Zhao, J., Zhou, H., Zhou, B., and Yu, H.
764 (2017). Knockdown of tetraspanin 13 inhibits proliferation of colorectal cancer cells. Int. J. Clin. Exp. Med.
765 *10*, 6387–6395.
766 Martini, M., Gnann, A., Scheikl, D., Holzmann, B., and Janssen, K.P. (2011). The candidate tumor

suppressor SASH1 interacts with the actin cytoskeleton and stimulates cell-matrix adhesion. Int. J. Biochem. Cell Biol. *43*, 1630–1640.

Mazmanian, S.K., Cui, H.L., Tzianabos, A.O., and Kasper, D.L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell *122*, 107–118.

Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics *8*, 389–399.

Mousa, H.A.L. (1997). Aerobic, anaerobic and fungal burn wound infections. J. Hosp. Infect. *37*, 317–323.

Murdoch, D.A., Mitchelmore, I.J., and Tabaqchali, S. (1988). Peptostreptococcus Micros in Polymicrobial Abscesses. Lancet *331*, 594.

Noel, G., Baetz, N.W., Staab, J.F., Donowitz, M., Kovbasnjuk, O., Pasetti, M.F., and Zachos, N.C. (2017). A primary human macrophage-enteroid co-culture model to investigate mucosal gut physiology and host-pathogen interactions. Sci. Rep. *7*.

Nougayrède, J.P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., Buchrieser, C., Hacker, J., Dobrindt, U., and Oswald, E. (2006). Escherichia coli induces DNA double-strand breaks in eukaryotic cells. Science *313*, 848–851.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. *44*, D733–D745.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., and Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics *31*, 3691–3693.

Pasquereau-Kotula, E., Martins, M., Aymeric, L., and Dramsi, S. (2018). Significance of Streptococcus gallolyticus subsp. gallolyticus association with colorectal cancer. Front. Microbiol. *9*, 614.

Pleguezuelos-Manzano, C., Puschhof, J., Rosendahl Huber, A., van Hoeck, A., Wood, H.M., Nomburg, J., Gurjao, C., Manders, F., Dalmasso, G., Stege, P.B., et al. (2020). Mutational signature in colorectal cancer caused by genotoxic pks + E. coli. Nature *580*, 269–273.

Purcell, R. V., Visnovska, M., Biggs, P.J., Schmeier, S., and Frizelle, F.A. (2017). Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer. Sci. Rep. *7*, 11590.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596.

Rams, T.E., Feik, D., Listgarten, M.A., and Slots, J. (1992). Peptostreptococcus micros in human periodontitis. Oral Microbiol. Immunol. *7*, 1–6.

Rana, M.K., Aloisio, F.M., Choi, C., and Barber, D.L. (2018). Formin-dependent TGF-β signaling for epithelial to mesenchymal transition. Mol. Biol. Cell *29*, 1465–1475.

Rubinstein, M.R., Wang, X., Liu, W., Hao, Y., Cai, G., and Han, Y.W. (2013). Fusobacterium nucleatum Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/β-Catenin Signaling via its FadA Adhesin. Cell Host Microbe *14*, 195–206.

Saffarian, A., Mulet, C., Regnault, B., Amiot, A., Tran-Van-Nhieu, J., Ravel, J., Sobhani, I., Sansonetti,

P.J., and Pédron, T. (2019). Crypt- and mucosa-associated core microbiotas in humans and their alteration in colon cancer patients. MBio *10*, e01315-19.

Samad, A.K.A., Taylor, R.S., Marshall, T., and Chapman, M.A.S. (2005). A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Color. Dis. *7*, 204–213.

Sanderson, P.J., Wren, M.W.D., and Baldwin, A.W.F. (1979). Anaerobic organisms in postoperative wounds. J. Clin. Pathol. *32*, 143–147.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmidt, T.S.B., Hayward, M.R., Coelho, L.P., Li, S.S., Costea, P.I., Voigt, A.Y., Wirbel, J., Maistrenko, O.M., Alves, R.J.C., Bergsten, E., et al. (2019). Extensive transmission of microbes along the gastrointestinal tract. Elife *8*, e42693.

Schwabe, R.F., and Jobin, C. (2013). The microbiome and cancer. Nat. Rev. Cancer *13*, 800–812.

Shimizu, A., Mammoto, A., Italiano, J.E., Pravda, E., Dudley, A.C., Ingber, D.E., and Klagsbrun, M. (2008). ABL2/ARG tyrosine kinase mediates SEMA3F-induced RhoA inactivation and cytoskeleton collapse in human glioma cells. J. Biol. Chem. *283*, 27230–27238.

Sobhani, I., Tap, J., Roudot-Thoraval, F., Roperch, J.P., Letulle, S., Langella, P., Gérard, C., van Nhieu, J.T., and Furet, J.P. (2011). Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One *6*, e16393.

Sobhani, I., Bergsten, E., Couffin, S., Amiot, A., Nebbad, B., Barau, C., de'Angelis, N., Rabot, S., Canoui-Poitrine, F., Mestivier, D., et al. (2019). Colorectal cancer-associated microbiota contributes to oncogenic epigenetic signatures. Proc. Natl. Acad. Sci. U. S. A. *116*, 24285–24295.

Socransky, S.S., Haffajee, A.D., Ximenez-Fyvie, L.A., Feres, M., and Mager, D. (1999). Ecological considerations in the treatment of Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis periodontal infections. Periodontol. 2000 *20*, 341–362.

Soderholm, A.T., and Pedicord, V.A. (2019). Intestinal epithelial cells: at the interface of the microbiota and mucosal immunity. Immunology *158*, 267–280.

Torrino, S., Visvikis, O., Doye, A., Boyer, L., Stefani, C., Munro, P., Bertoglio, J., Gacon, G., Mettouchi, A., and Lemichez, E. (2011). The E3 ubiquitin-ligase HACE1 catalyzes the ubiquitylation of active Rac1. Dev. Cell *21*, 959–965.

Tortola, L., Nitsch, R., Bertrand, M.J.M., Kogler, M., Redouane, Y., Kozieradzki, I., Uribesalgo, I., Fennell, L.M., Daugaard, M., Klug, H., et al. (2016). The Tumor Suppressor Hace1 Is a Critical Regulator of TNFR1-Mediated Cell Fate. Cell Rep. *15*, 1481–1492.

Wettenhall, J.M., and Smyth, G.K. (2004). limmaGUI: A graphical user interface for linear modeling of microarray data. Bioinformatics *20*, 3705–3706.

Xia, X., Wu, W.K.K., Wong, S.H., Liu, D., Kwong, T.N.Y., Nakatsu, G., Yan, P.S., Chuang, Y.M., Chan, M.W.Y., Coker, O.O., et al. (2020). Bacteria pathogens drive host colonic epithelial cell promoter hypermethylation of tumor suppressor genes in colorectal cancer. Microbiome *8*, 108.

Xu, J., Yang, M., Wang, D., Zhang, S., Yan, S., Zhu, Y., and Chen, W. (2020). Alteration of the abundance of Parvimonas micra in the gut along the adenoma-carcinoma sequence. Oncol. Lett. *20*, 106.

Zeller, G., Tap, J., Voigt, A.Y., Sunagawa, S., Kultima, J.R., Costea, P.I., Amiot, A., Böhm, J., Brunetti,

F., Habermann, N., et al. (2014). Potential of fecal microbiota for early-stage detection of colorectal cancer. Mol. Syst. Biol. *10*, 766.

Zhang, L., Anglesio, M.S., O'Sullivan, M., Zhang, F., Yang, G., Sarao, R., Nghiem, M.P., Cronin, S., Hara, H., Melnyk, N., et al. (2007). The E3 ligase HACE1 is a critical chromosome 6q21 tumor suppressor involved in multiple cancers. Nat. Med. *13*, 1060–1069.

Zhao, L., Zhou, Y., Zhao, R., Coker, O.O.O., Zhang, X., Chu, E.S., Wei, H., Wu, W.K., Wong, S.H., Sung, J.J., et al. (2020). Parvimonas Micra Promotes Intestinal Tumorigenesis in Conventional Apcmin/+ Mice and in Germ-Free Mice. Research Square doi:10.21203/RS.3.RS-25974/V1.

Zhao, M., Kim, P., Mitra, R., Zhao, J., and Zhao, Z. (2016). TSGene 2.0: An updated literature-based knowledgebase for Tumor Suppressor Genes. Nucleic Acids Res. *44*, D1023–D1031.

Zhao, M., Liu, Y., Zheng, C., and Qu, H. (2019). dbEMT 2.0: An updated database for epithelial-mesenchymal transition genes with experimentally verified information and precalculated regulation information for cancer metastasis. J. Genet. Genomics *46*, 595–597.

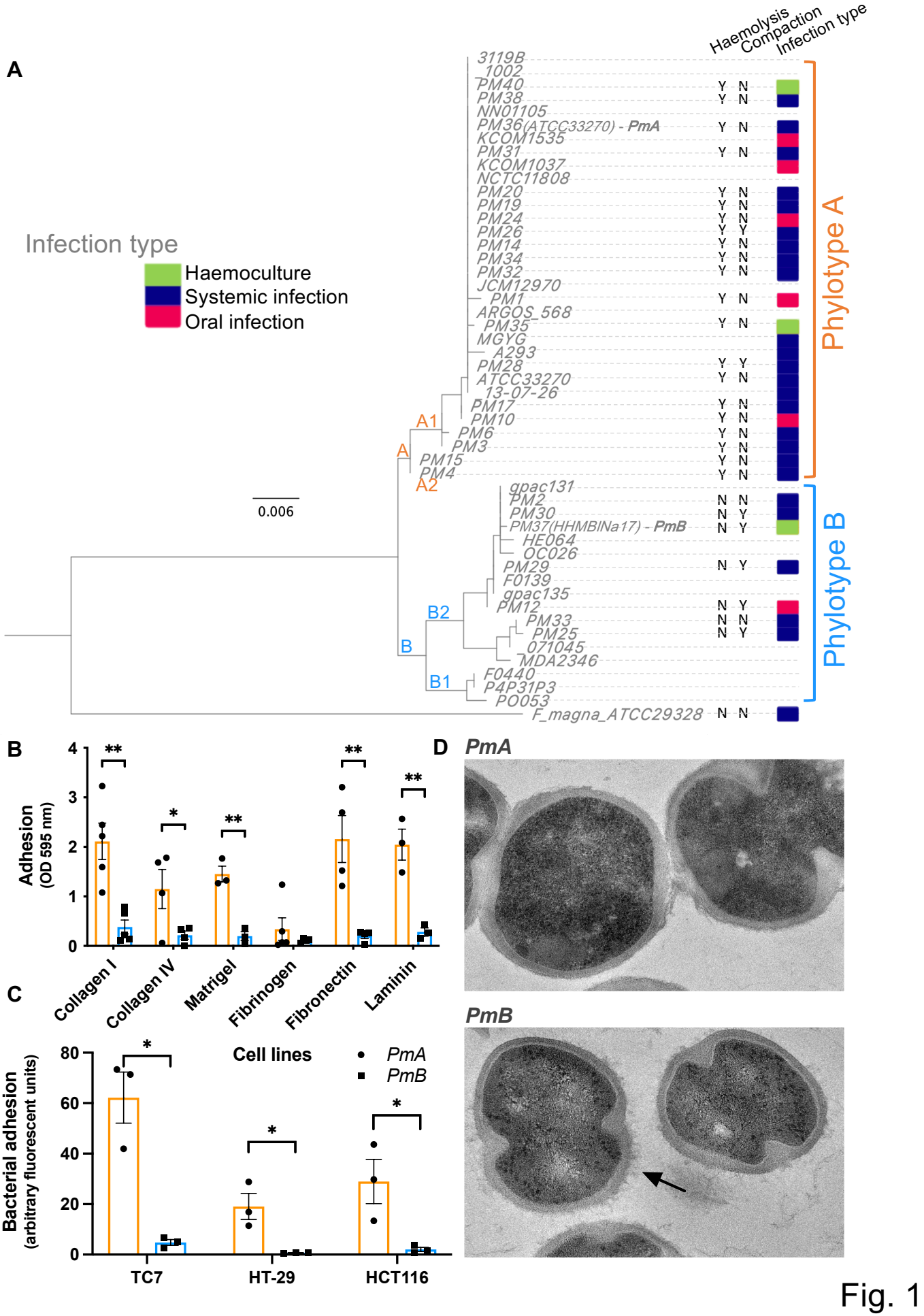

Fig. 1

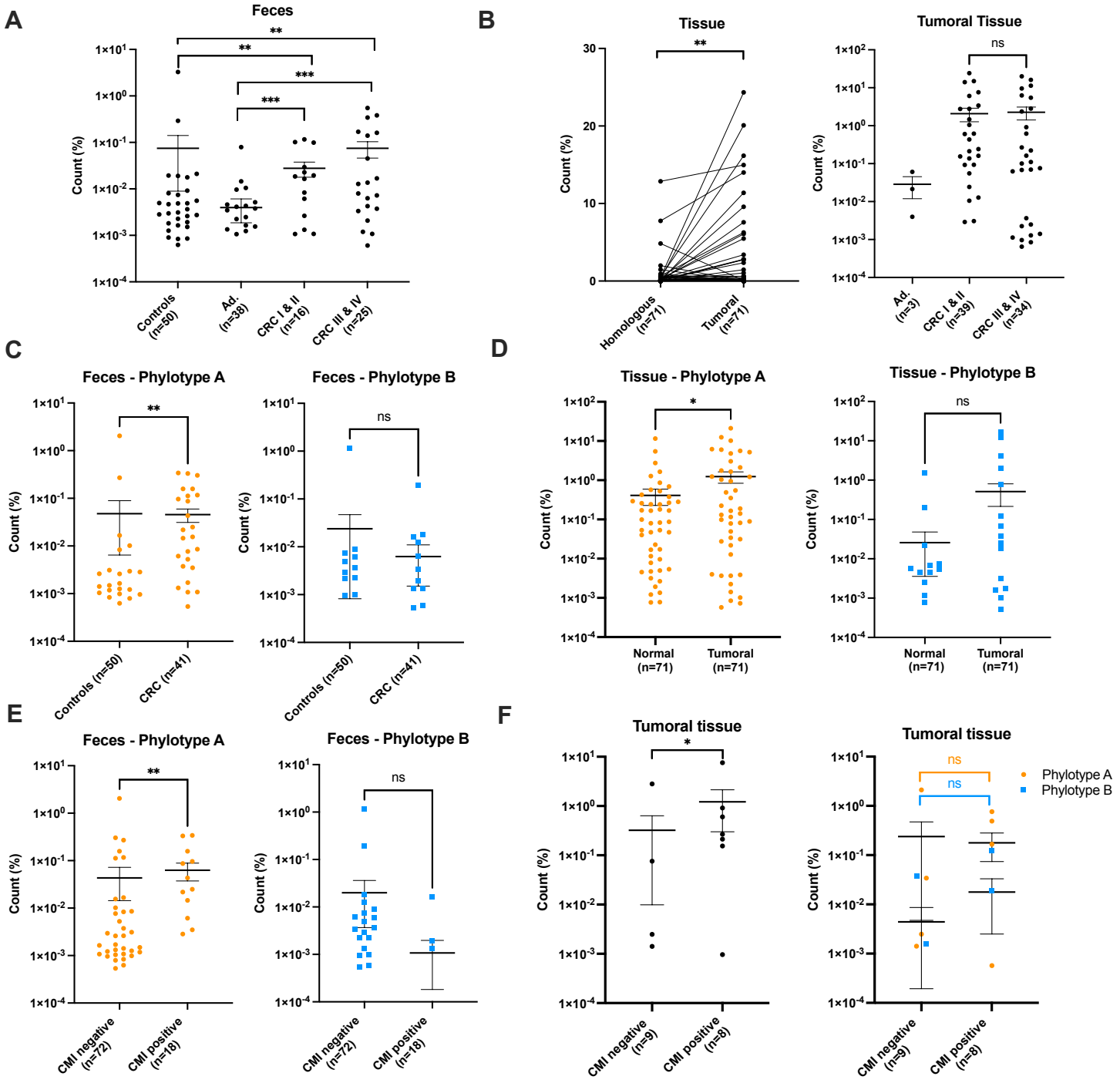

Fig. 2

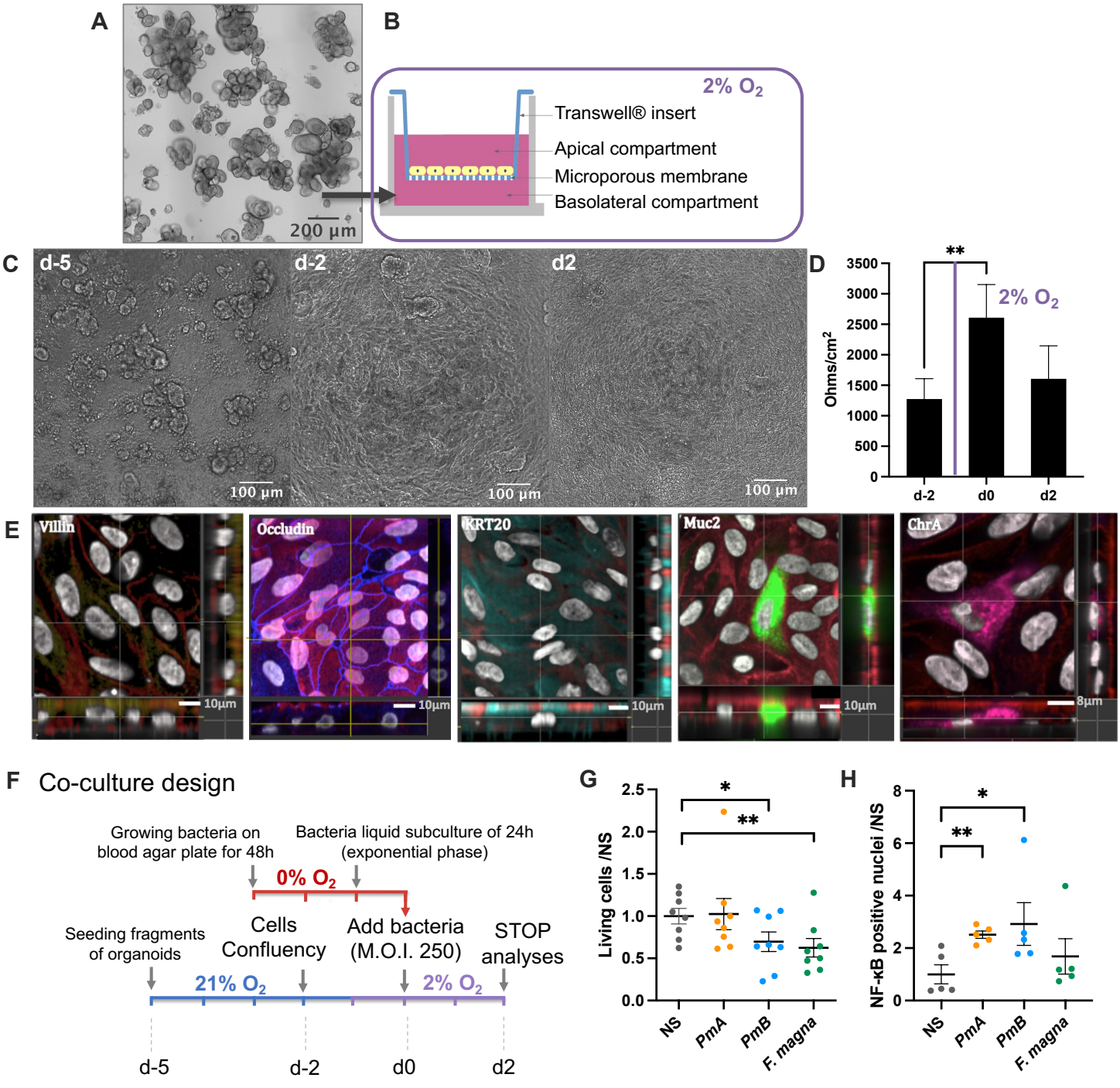
A
B
200 µm
2% O2
Transwell® insert
Apical compartment
Microporous membrane
Basolateral compartment
C
d-5
d-2
d2
100 µm
100 µm
100 µm
D
**
2% O2
Ohms/cm2
3500
3000
2500
2000
1500
1000
500
0
d-2
d0
d2
E
Villin
Occludin
KRT20
Muc2
ChrA
10µm
10µm
10µm
10µm
8µm
F Co-culture design
Growing bacteria on blood agar plate for 48h
Bacteria liquid subculture of 24h (exponential phase)
0% O2
Seeding fragments of organoids
Cells Confluency
Add bacteria (M.O.I. 250)
STOP analyses
21% O2
2% O2
d-5
d-2
d0
d2
G
Living cells /NS
2.5
2.0
1.5
1.0
0.5
0.0
*
**
NS
PmA
PmB
F. magna
H
NF-κB positive nuclei /NS
8
6
4
2
0
*
**
NS
PmA
PmB
F. magna

Fig. 3

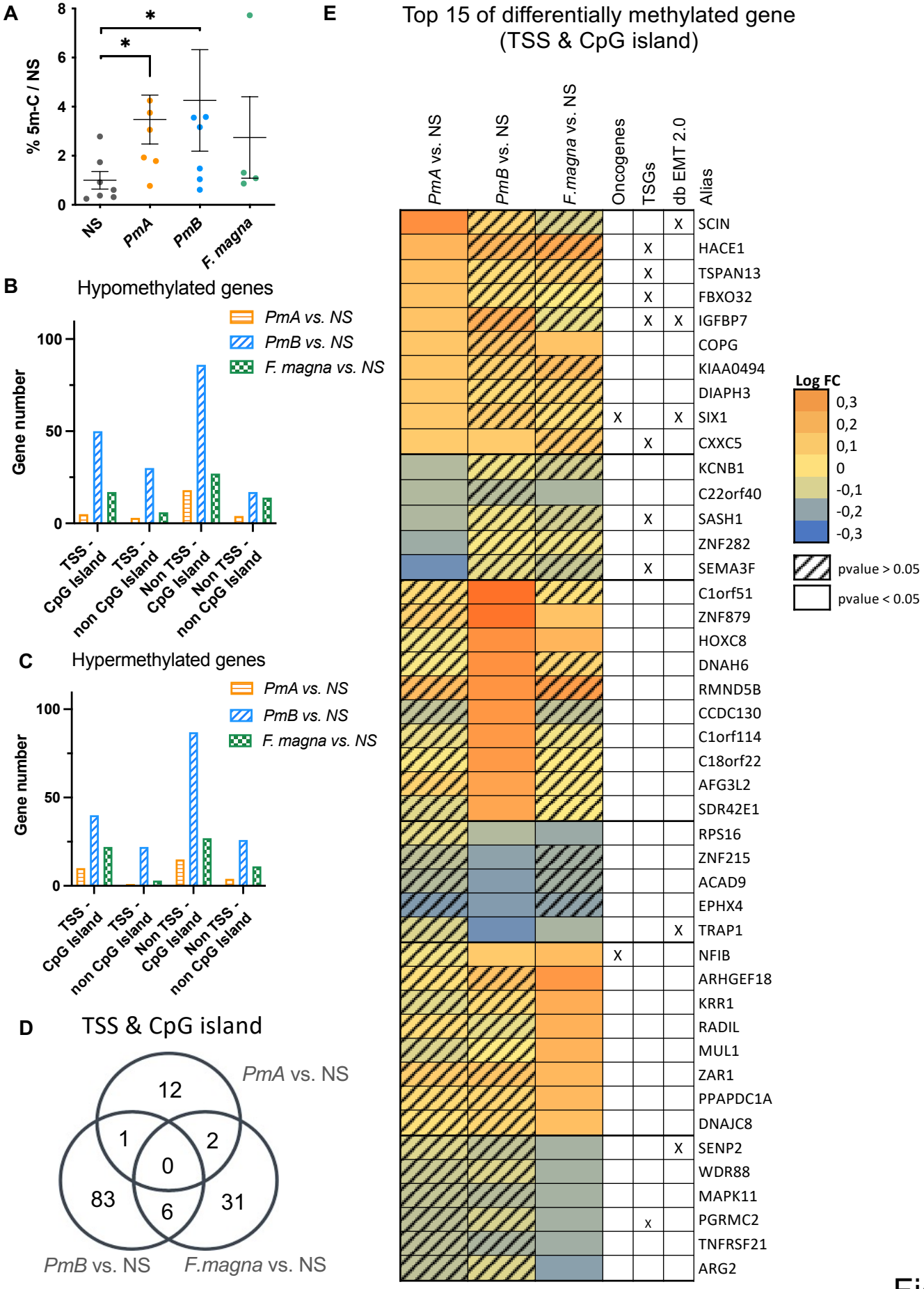
A
% 5m-C / NS
NS
PmA
PmB
F. magna
B
Hypomethylated genes
PmA vs. NS
PmB vs. NS
F. magna vs. NS
Gene number
TSS - CpG Island
TSS - non CpG Island
Non TSS - CpG Island
Non TSS - non CpG Island
C
Hypermethylated genes
D
TSS & CpG island
PmA vs. NS
PmB vs. NS
F.magna vs. NS
12
1
2
0
83
6
31
E
Top 15 of differentially methylated gene
(TSS & CpG island)
PmA vs. NS
PmB vs. NS
F.magna vs. NS
Oncogenes
TSGs
db EMT 2.0
Alias
SCIN
HACE1
TSPAN13
FBXO32
IGFBP7
COPG
KIAA0494
DIAPH3
SIX1
CXXC5
KCNB1
C22orf40
SASH1
ZNF282
SEMA3F
C1orf51
ZNF879
HOXC8
DNAH6
RMND5B
CCDC130
C1orf114
C18orf22
AFG3L2
SDR42E1
RPS16
ZNF215
ACAD9
EPHX4
TRAP1
NFIB
ARHGEF18
KRR1
RADIL
MUL1
ZAR1
PPAPDC1A
DNAJC8
SENP2
WDR88
MAPK11
PGRMC2
TNFRSF21
ARG2
Log FC
0,3
0,2
0,1
0
-0,1
-0,2
-0,3
pvalue > 0.05
pvalue < 0.05

Fig. 4

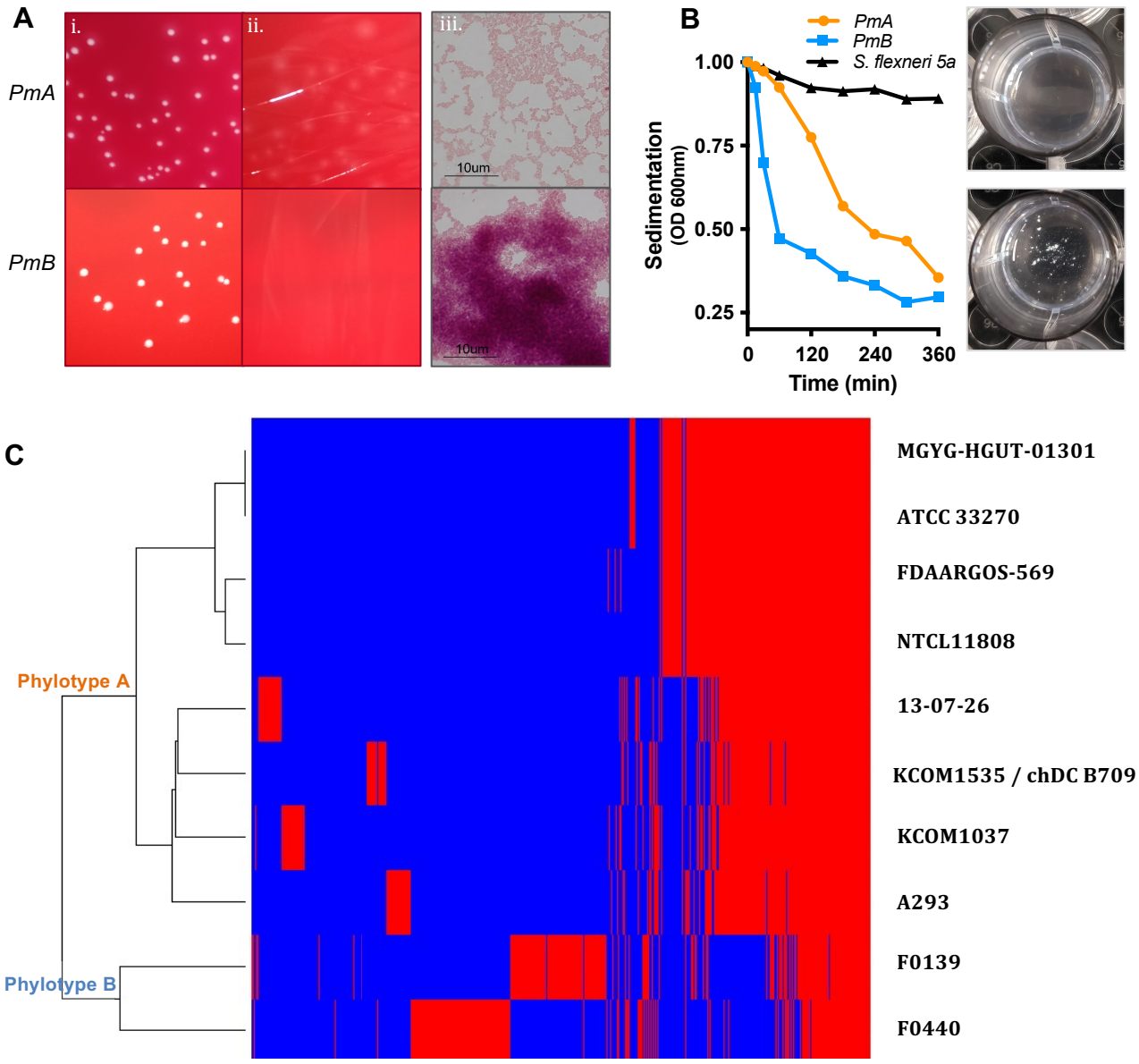
A
i.
ii.
iii.
PmA
PmB
10um
10um
B
PmA
PmB
S. flexneri 5a
Sedimentation (OD 600nm)
1.00
0.75
0.50
0.25
0
120
240
360
Time (min)
C
Phylotype A
Phylotype B
MGYG-HGUT-01301
ATCC 33270
FDAARGOS-569
NTCL11808
13-07-26
KCOM1535 / chDC B709
KCOM1037
A293
F0139
F0440

Fig. S1

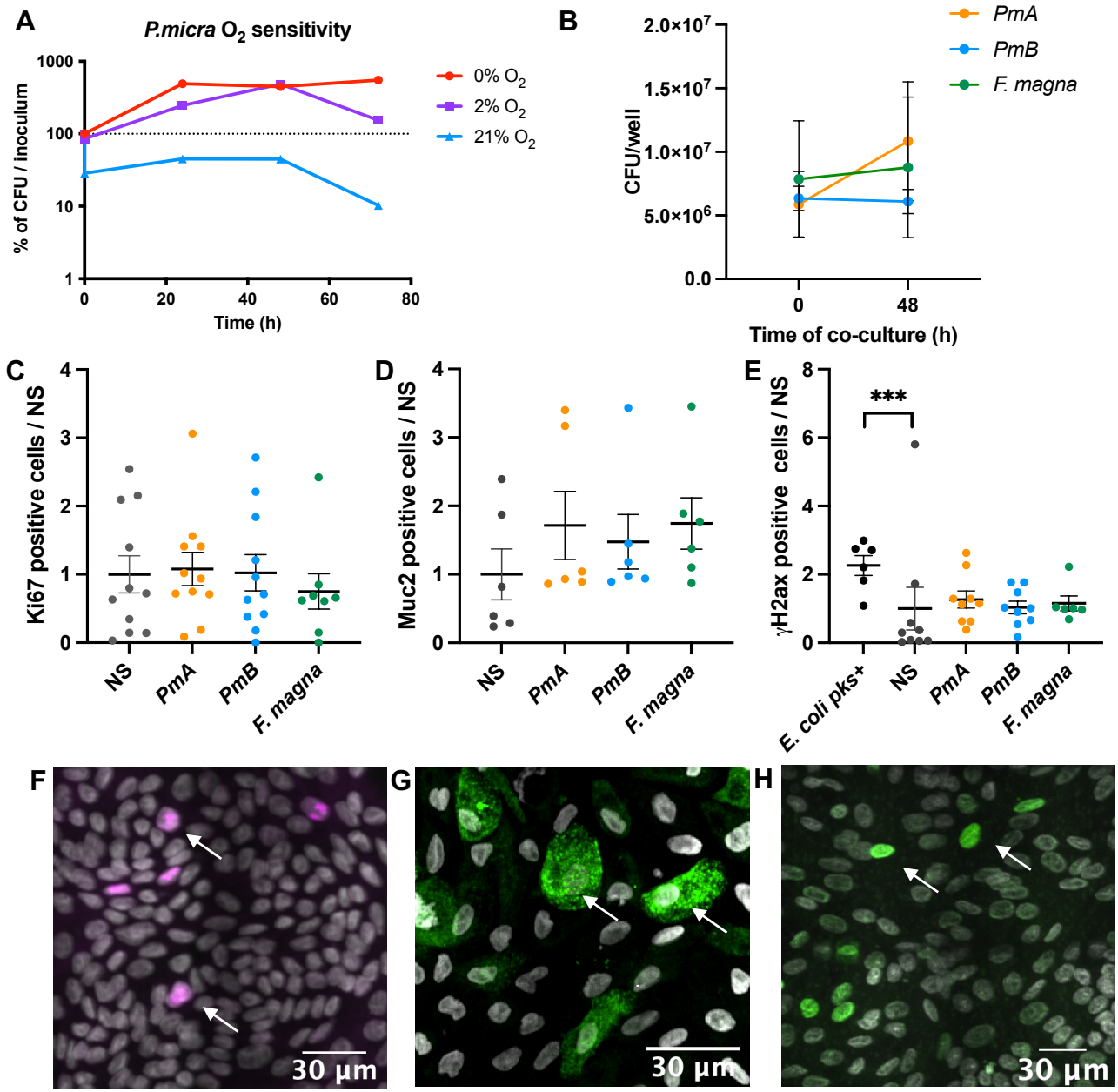
A
P.micra O2 sensitivity
% of CFU / inoculum
Time (h)
0% O2
2% O2
21% O2
B
CFU/well
Time of co-culture (h)
PmA
PmB
F. magna
C
Ki67 positive cells / NS
NS
PmA
PmB
F. magna
D
Muc2 positive cells / NS
E
γH2ax positive cells / NS
***
E. coli pks+
F
G
H
30 µm

Fig. S2

Clinical isolate number	Origin	Source of isolation	Hemolysis	Colonies compaction	Phylogenetic Group	GenBank accession numbers
PM 1	Pitié Salpêtrière	Dental cellulite	+	-	A1	OM287472
PM 10	Pitié Salpêtrière	Dental cellulite	+	-	A1	OM287474
PM 14	Pitié Salpêtrière	Cerebral abscess	+	-	A1	OM287469
PM 15	Pitié Salpêtrière	Cerebral abscess	+	-	A2	OM287478
PM 17	Pitié Salpêtrière	Cerebral abscess	+	-	A1	OM287473
PM 19	Pitié Salpêtrière	Cervical abscess	+	-	A1	OM287466
PM 20	Pitié Salpêtrière	Cervical adenopathy	+	-	A1	OM287465
PM 24	Pitié Salpêtrière	Lower canaliculus	+	-	A1	OM287467
PM 26	Pitié Salpêtrière	Cerebral abscess	+	+	A1	OM287468
PM 28	Cochin	Femur	+	+	A1	OM287464
PM 3	Pitié Salpêtrière	Cerebral abscess	+	-	A1	OM287476
PM 31	Henri Mondor	Empyema brain abscesses	+	-	A1	OM287463
PM 32	Cochin	Hip articular	+	-	A1	OM287471
PM 34	Cochin	Hip	+	-	A1	OM287470
PM 35	Henri Mondor	Hemoculture	+	-	A1	OM287462
PM 36 / ATCC 33270 / *PmA*	CIP Pasteur	Purulent pleurisy	+	-	A1	OM287461
PM 38	CIP Pasteur	Abdominal wound	+	-	A1	OM287460
PM 4	Pitié Salpêtrière	Cerebral abscess	+	-	A2	OM287477
PM 40	Henri Mondor	Haemoculture	+	-	A1	OM287459
PM 6	Pitié Salpêtrière	Cerebral abscess	+	-	A1	OM28747
PM 25	Pitiée Salpêtrière	Cerebral abscess	-	+	B2a	OM287480
PM 33	Cochin	Knee joint	-	-	B2a	OM287479
PM 12	Pitié Salpêtrière	Maxillary sinus	-	+	B2b	OM287481
PM 2	Pitié Salpêtrière	Cervical collection	-	-	B2b	OM287484
PM 29	Cochin	Urinary tract	-	+	B2b	OM287482
PM 30	Cochin	Femur	-	+	B2b	OM287483
PM 37 / HHM BINA17 / *PmB*	Henri Mondor	Haemoculture	-	+	B2b	OM287485

Table S1: Description of *P. micra* clinical isolates obtained from several hospitals in Paris and from different infectious locations.

Samples	**Feces** **n= 166**	**Tissue (tumor and homologous)** **n=71**
Status	Control n= 88; CRC n= 78	CRC n=71
Tumor localization	Right or transversal colon n= 27 Sigmoid or left Colon n= 48, ind=3	Right or transversal colon n= 35 Sigmoid or left Colon n= 41
TNM stages	I or II n= 27; III or IV n= 48, ind=3	I or II n= 40; III or IV n=31
Gender (Female/male)	F: n= 70; M: n= 76	F: n= 23; M: n=48
Ages (mean +/- SEM)	Control: 60.48 +/- 1.13 CRC: 66.23 +/-1.60	CRC: 65.46 +/- 1.22
BMI (kg/m2) mean +/- SEM	Control: 25.18 +/- 0.4 CCR: 25.67 +/- 0.81	CRC: 25.2 +/- 0.5
Cumulative methylation index (Negative/Positive)	N: n= 119; P: n= 43	N: n= 8; P: n= 7

Table S2: Clinical data. Ind, indetermined.

Names, Alias	Described role	Link with CRC or cancers	References
SCIN (Scinderin)	Ca(2+)-dependent actin-severing and capping protein. Regulation of actin cytoskeleton.	Overexpressed in CRC. Inhibits cell proliferation and tumorigenesis.	*Lin 2019*
HACE1 (HECT domain E3 ubiquitin protein ligase 1)	E3 ubiquitin ligase involved in specific tagging of target proteins, leading to their subcellular localization or proteasomal degradation.	TSG. Down-regulated by DNA methylation in CRC. Loss or knockout of HACE1 enhanced tumor growth, invasion, and metastasis; in contrast, the overexpression of HACE1 can inhibit the development of tumors.	*Li 2019* *Hibi 2008*
TSPAN13 (Tetraspanin 13)	Transmembrane signal transduction protein that plays a role in the regulation of cell development, activation, growth, motility and invasion.	TSG. Downregulation inhibits proliferation of CRC cells.	*Lou 2017*
FBXO32 (F-Box Protein 32) Astogin-1	Component of a SCF E3 ubiquitin-protein ligase complex which mediates the ubiquitination and subsequent proteasomal degradation of target proteins.	TSG. Under-expressed in CRC. Induce cell differentiation. Upstream regulator of EMT.	*Yuan 2018* *Sahy 2017*
IGFBP7 (Insulin Like Growth Factor Binding Protein 7)	Protein coding gene that regulate IGFs. Stimulates prostacyclin production and cell adhesion. Promotes cancer cell growth and migration.	TSG. Silencing induce metastasis.	*Ruan 2007* *Suzuki 2010*
KIAA0494 (EF-Hand Calcium Binding Domain 14)	Uncharacterized protein. Predicted membrane protein.	-	
DIAPH3 (Diaphanous Related Formin 3)	Protein coding gene involve in actin remodeling. Regulate cell movement and adhesion.	Deficiency enhances cell motility, invasion and metastasis in many cancers.	*Hager 2012* *Rana 2018*
SIX1 (sine oculis homeobox 1)	Transcription factor involve in regulation of cell proliferation, apoptosis, embryonic development and tumorigenesis.	Oncogene. Overexpressed in CRC, overexpression of Six1 dramatically promotes CRC tumor growth and metastasis in vivo.	*Wu 2014* *Xu 2017*
KCNB1 (Potassium Voltage-Gated Channel Subfamily B Member 1)	Voltage-gated potassium channel Kv2.1. Contributes to the pronounced pro-apoptotic potassium current surge during neuronal apoptotic cell death in response to oxidative injury.	KCNB1 polymorphisms correlate to CRC treatment and patient's outcome. Kv2.1 occasionally forms complexes with other voltage-gated potassium alpha-subunits (i.e. Kv9.3), witch its silencing potently inhibiting proliferation in human colon cells.	*Li 2015,* *Fahra 2020*
SASH1 (SAM and SH3 domain-containing protein 1)	Scaffold protein involved in the TLR4 signaling. Stimulate cytokine production and endothelial cell migration in response to binding pathogens	TSG. Downregulated in CRC. Downregulation expression was correlated with the formation of metachronous distant metastasis. Loss of SASH1 induces EMT. SASH1 inhibits metastasis formation In Vivo.	*Rimkus 2006,* *Franke 2019*
ZNF282 (Zinc Finger Protein 282)	Transcription factor binding the U5RE (U5 repressive element) of HLTV-1 (human T cell leukemia virus type 1) with a repressive effect. Co-activator of the estrogen receptor alpha.	Knockdown reduced migration, invasion and tumorigenesis of ESCC *in vitro* and reduced the tumorigenicity of ESCC xenograft in nude mouse.	*Yeo 2014*
SEMA3F (Semaphorin 3F)	Secreted signaling protein that are involved in axon guidance during neuronal development. Act in an autocrine fashion to induce apoptosis, inhibit cell proliferation and survival. Regulator of actin cytoskeleton.	TSG. Down-regulated by DNA methylation in CRC tissues and CRC cell lines. Associated with progressive phenotypes of CRC. Overexpression reduced proliferation, adhesion, and migratory capability of colon cancer cells. SEMA3F-overexpressing cells exhibited diminished tumorigenesis when transplanted in nude mice and reduced liver metastases.	*Gao 2015,* *Wu 2011,* *Bielenberg 2004*

Table S3: *PmA* induces DNA methylation changes in promoters of genes involved in carcinogenesis. CRC: colorectal cancer; TSG: Tumor suppressor gene; EMT: epithelial-mesenchymal transition; IGFs: insulin-like growth factors; SCF: SKP1-CUL1-F-box protein; TRAPP: Transport Protein Particle; ESCC: Esophageal squamous cell carcinoma.