# *Fusobacterium sphaericum sp. nov.*, isolated from a human colon tumor, is prevalent in various human body sites and induces IL-8 secretion from colorectal cancer cells

Martha A. Zepeda-Rivera*[1], Yannick Eisele*[2,3], Alexander Baryiames[2], Hanrui Wu[2], Kaitlyn D. LaCourse[2], Dakota S. Jones[1], Hans Hauner[3,4], Floyd E. Dewhirst[5,6], Samuel S. Minot[7], Christopher D. Johnston[1#], Susan Bullman[2#]

[1]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA
[2]Human Biology Division, Fred Hutchinson Cancer Center, Seattle, WA, USA
[3]Technical University of Munich, School of Medicine, Munich, Germany
[4]Else Kroener-Fresenius-Center for Nutritional Medicine, School of Life Sciences, Technical University of Munich, Munich, Germany
[5] Forsyth Institute, Cambridge, MA, USA
[6] Department of Oral Medicine, Infection and Immunity, Harvard School of Dental Medicine, Boston, MA, USA
[7]Data Core, Shared Resources, Fred Hutchinson Cancer Center, Seattle, WA, USA

*Contributed equally to this work.

**Keywords**: Fusobacterium, novel species, pangenome, comparative genomics, intestinal microbiota, colorectal cancer

[#]**Address Correspondence to**:

Susan Bullman, PhD, sbullman@fredhutch.org

Christopher D. Johnston, PhD, johnston@fredhutch.org

***Abstract***

Cancerous tissue is a largely unexplored microbial niche that provides a unique environment for the colonization and growth of specific bacterial communities, and with it, the opportunity to identify novel bacterial species. Here, we report distinct features of a novel *Fusobacterium* species, *F. sphaericum sp. nov.* (*Fs*), isolated from primary colon adenocarcinoma tissue. We acquire the complete, closed genome of this organism and phylogenetically confirm its classification into the *Fusobacterium* genus. Phenotypic and genomic analysis of *Fs* reveal that this novel organism is of coccoid shape, rare for *Fusobacterium* members, and has species-distinct gene content. *Fs* displays a metabolic profile and antibiotic resistance repertoire consistent with other *Fusobacterium* species. *In vitro*, *Fs* has adherent and immunomodulatory capabilities, as it intimately associates with human colon cancer epithelial cells and promotes IL-8 secretion. Analysis of the prevalence and abundance of *Fs* in ~1,750 human metagenomic samples shows that it is a moderately prevalent member of the human oral cavity and stool. Intriguingly, analysis of ~1,270 specimens from patients with colorectal cancer demonstrate that *Fs* is significantly enriched in colonic and tumor tissue as compared to mucosa or feces. Our study sheds light on a novel bacterial species that is prevalent within the human intestinal microbiota and whose role in human health and disease requires further investigation.

## *Introduction*

Molecular, imaging, and culture-based approaches have demonstrated that human tumors are inhabited by a broad range of bacterial species, with microbial community composition varying by cancer type[1,2]. Tissue-associated microbiomes are not readily amenable to metagenomic analysis due to the overabundance of host nucleic acids. Therefore, characterization of tumor-associated microbial communities is often restricted to amplicon-based sequencing approaches with limited phylogenetic resolution. To overcome this challenge, classical microbiology culture-based approaches from patient tissue specimens have reemerged as valuable tools to isolate members of tissue-associated microbiomes. Subsequent whole genome sequencing and comparative genomics allows for the detailed phylogenetic identification of these members.

In patients with colorectal cancer (CRC), unbiased genomic analyses have revealed an enrichment of *Fusobacterium* in CRC tumors relative to noncancerous colorectal tissue[3–7]. Previous work from our group and others demonstrated that within patient CRC tumors, *Fusobacterium* colonizes regions with immune and epithelial functions supportive of cancer progression[8], that it persists with the primary tumor during metastasis[9], and that microbiome modulation targeting it decreases cancer cell proliferation and tumor growth[9,10]. Furthermore, high *Fusobacterium* loads are associated with high-grade dysplasia[3], advanced disease stage[11], and poorer patient prognosis[12]. Through a culturing-based approach, we isolated and characterized a novel *Fusobacterium* species from the resected tumor of a treatment naïve female patient with microsatellite stable (MSS), stage III, right-sided colon cancer. Phylogenetic appraisal of this organism places this bacterium within the *Fusobacterium* genus but demonstrates that it is distinct from known species. Phenotypic analysis of this novel organism revealed a coccobacilloid cellular morphology, leading us to

propose this species as *Fusobacterium sphaericum sp. nov.* (*Fs*). In this present study, we find that while the genome structure and gene content of *Fs* is predominantly unique to this species, its metabolic profile and antibiotic resistance repertoire are consistent with other *Fusobacterium* members. Furthermore, our analyses indicate that most predicted virulence factors are related to metabolism, cell-cell adherence, and immunomodulation. *In vitro* co-culture with human colon cancer epithelial cells confirmed *Fs* adherence and indicated that its presence significantly increases IL-8 secretion. Assessment of the prevalence and abundance of *Fs* in 1,973 human metagenomic samples shows that is a moderately prevalent member of the human oral cavity and stool. Intriguingly, in 1,273 specimens from patients with CRC, *Fs* is found to be significantly enriched in colonic and tumor tissue as compared to mucosa or feces.

## ***Results***

### **Phylogenetic and morphologic characterization of a novel *Fusobacterium* species**

*Fusobacterium* isolation was performed on a human colorectal cancer (CRC) tumor from a treatment naïve female patient with microsatellite stable (MSS), stage III, right-sided colon cancer. Using long-read single-molecule real-time (SMRT)[13] sequencing we generated a complete and closed genome for isolate SB021, with two putative chromosomes and two additional extrachromosomal elements (Fig. 1a, Supplementary Table 1). BLASTn analysis of the 16S rRNA gene of SB021 confirmed its classification into the *Fusobacterium* genus yet indicated that all related hits fell below the established 16S rRNA gene species similarity threshold of 98.65%[14], suggesting that SB021 represents a novel *Fusobacterium* species (Supplementary Table 2). Comparative phylogenetics of SB021 and other *Fusobacterium spp.* strains (Table 1), both by the 16S rRNA gene and a reference-free whole-genome approach[15], indicate that SB021 forms a

distinct branch within the same clade as *Fusobacterium perfoetens* (Fig. 1b-c). To quantify the relatedness between SB021 and *F. perfoetens*, we calculated the average nucleotide identity (ANI) and found it to be 75.27%, well below the established species threshold of 95%[14] (Fig. 1d, Supplementary Fig. 1, Supplementary Table 3), supporting the classification of SB021 as a novel *Fusobacterium* species.

The cellular morphology of *Fusobacterium* cells is known to vary[16,17]. As SB021 is phylogenetically closely related to *F. perfoetens*, we assessed whether like *F. perfoetens*, SB021 cells have a short, ovoid coccobacillus morphology. Macroscopically, SB021 cells form distinct, vaguely round, raised opaque colonies with central indentations following 48 hours of anaerobic culture. The colonies are viscous and malodorous and measure 3-5 millimeters in diameter (Fig. 1e). Microscopically, SB021 cells are short, gram-negative coccobacilli, with an average length of 1.60μm and an average width of 1.25μm (Fig. 1f). Transmission electron microscopy further revealed a spherical shape and a granulated cytoplasm (Fig. 1g). These results suggest that members within this *Fusobacterium* clade share similar cellular phenotypes. Given the coccobacilloid morphology of SB021 cells, we propose this novel species as *Fusobacterium sphaericum sp. nov.* (*Fs*).

**_Fs_ genomic attributes predict metabolic and virulence capabilities**

As significant genomic heterogeneity has been reported within the *Fusobacterium* genus, with genome size, architecture, and content differing between species of varying pathogenic capabilities[18], we sought to characterize the genomic content of *Fs*. To compare the genetic content of *Fs* to other *Fusobacterium spp*., we implemented the analysis and visualization platform for

‘omics data (Anvi’o) workflow for microbial pangenomics[19]. This genus-level pangenomic analysis allows us to identify all genes present in the *Fusobacterium* genus (“pangenome”) and discern between gene content that is shared amongst most species (“core genome”), is shared amongst subsets of species (“accessory genomes”) or is unique to individual members (“singletons”). Across eighteen representative *Fusobacterium* genomes (Table 1), only 0.05% of gene clusters compose the core (12/25,256), indicative of the extreme genetic heterogeneity across this genus (Fig. 2a). *Fs* harbored 9.24% of identified gene clusters (2,333/25,256), but of these, 93.6% (2,183/2,333) were unique to *Fs*, while only 4.4% (133/2,333) were shared with its closest phylogenetic neighbor, *F. perfoetens* (Fig. 2b).

As the genetic content of *Fs* is predominantly unique to this species, we assessed the 2,333 identified gene clusters. KEGG ortholog analysis[20] revealed that mapped gene clusters (52.6%) were predominantly involved in metabolic functions and pathways (Supplementary Table 4). *In vitro* metabolic profiling via the API 20A Gallery System indicated that *Fs* can ferment the monosaccharides glucose and mannose, the disaccharide lactose, the trisaccharide raffinose, and the coumarin glucoside esculin (Supplementary Fig. 2a). This is in agreement with previous reports indicating that all *Fusobacterium spp.* can ferment glucose[21], and that *Fusobacterium spp.* within the same clade as *Fs* are able to metabolize mannose[22] (*F. mortiferum*, *F. ulcerans*, and *F. varium*), lactose (*F. mortiferum*), raffinose (*F. mortiferum*) and esculin (*F. mortiferum*)[22,23]. Primary metabolic pathways assessment via gutSMASH[24] indicated that *Fs* uniquely harbors a putative operon predicted to function in the metabolism of gallic acid, a phenolic compound extensively used in food additives with described anticarcinogenic, antimutagenic, and antioxidant properties[25] (Supplementary Fig. 3). Enzymatic profiling via the API ZYM panel additionally showed positive

activity for alkaline phosphatase, C4-esterase, C8-esterase-lipase (weak), acid phosphatase (weak), napthol-AS-BI-phosphohydrolase (weak) and α-Galactosidase (strong) (Supplementary Fig. 2b). This strong α-Galactosidase activity may be a distinct feature of *Fs*, as activity for this enzyme has only been observed weakly in *F. mortiferum* and *F. necrogenes*[22,23].

Notably, *Fs* harbors fifteen multidrug resistance efflux pumps (Supplementary Table 4) and further analysis with the comprehensive antibiotic resistance database (CARD)[26] identified 183 putative antibiotic resistance genes (Supplementary Table 5), suggesting that *Fs* is resistant to a suite of antibiotics. Our group has previously demonstrated that antibiotic-mediated microbiome modulation targeting *Fusobacterium* decreases cancer cell proliferation and tumor growth[9]. As such, we sought to characterize the antibiotic resistance profile of *Fs*. *In vitro* antibiotic susceptibility testing demonstrated that *Fs* is susceptible to various classes of antibiotics, including metronidazole (MIC 0.502 μg/ml; nitroimizadole class), colistin (MIC 0.875 μg/ml; polymyxin class), and penicillin (MIC 0.079 μg/ml; beta-lactam class). Consistent with other *Fusobacterium spp*., *Fs* is resistant to erythromycin (macrolide class) and vancomycin (glycopeptide class) antibiotics. Curiously, *Fs* susceptibility to aminoglycosides varied, with susceptibility to gentamycin (MIC 62.2 μg/ml) and resistance to streptomycin and kanamycin (Table 2).

Given that *Fs* was isolated directly from a patient CRC tumor, we next queried the presence of virulence factors important for host-bacterial interactions. We found that *Fs* lacks known *Fusobacterium* virulence factors[27–35] (Fig. 2c) but of the 2,333 identified gene clusters, 28.3% of them are homologous to known virulence factors in other bacterial pathogens[36] (Supplementary Table 4). Of this subset, the majority function to modulate the immune system (28.7%), to

scavenge and metabolize nutritional sources (16%), and to adhere to host cells (12.1%) (Fig. 2d). This suggests that *Fs* has the genetic capability to intimately interact with eukaryotic cells and modulate their immune responses.

**_Fs_ adheres to colon cancer epithelial cells and increases IL-8 secretion**

To evaluate the interaction between *Fs* and colon epithelial cells, we co-cultured *Fs* with cell lines derived from human colorectal carcinomas, HCT 116 and HT-29. As *F. nucleatum* has previously been reported to adhere to and invade cancer epithelial cells[9], we used *F. nucleatum* SB010 as a positive control. Confocal laser scanning microscopy of fluorescent immunocytochemically stained cells demonstrated that *Fs* strongly adheres to epithelial cell surfaces in distinct chain-like clusters with both HCT 116 and HT-29 (Fig. 3a). However, we did not observe intracellular *Fs* suggesting that, unlike *F. nucleatum*, *Fs* adheres to but does not invade colonic epithelial cells.

To investigate whether this physical interaction between *Fs* and colonic epithelial cells modulates eukaryotic immune responses, we assessed chemokine and cytokine induction in HCT 116 cells post co-culturing. In the presence of *Fs* cells, we observed an increase in various pro-inflammatory chemokines, most notably IL-8, with both live (2.59-fold increase) and dead (2.29-fold increase) bacterial cells (Fig. 3b, Supplementary Fig.4), indicating that the immunomodulatory capabilities of *Fs* are independent of viability. IL-8 expression is elevated in colonic tissue of patients with CRC[37], enhances colon cancer cell growth and migration to the liver and lungs[38,39], and has previously been reported to be stimulated by the presence of *F. nucleatum* in an adhesion-dependent manner[40]. To assess whether physical interaction between *Fs* and colonic epithelial cells is required for IL-8 secretion, we co-incubated *Fs* supernatant with HCT 116 cells and detected a

1.27-fold increase in IL-8 (Fig. 3b). These results suggest that unlike *F. nucleatum*, a direct bacterial-eukaryotic cell interaction with *Fs* is not required to stimulate secretion of IL-8 and other pro-inflammatory chemokines and cytokines (Fig. 3b, Supplementary Fig. 4). Furthermore, quantitative comparison of IL-8 production in the presence of *Fs* or *F. nucleatum* further showed that co-incubation with *Fs* stimulated higher IL-8 levels at both 4-hour (*Fs*: ca. 13,403 pg/ml; *F. nucleatum*: ca.1,732 pg/ml) and 24-hour (*Fs*: ca. 12,004 pg/ml; *F. nucleatum* ca. 5,551 pg/ml) time points (Fig. 3c).

**_Fs_ is prevalent across human body sites and enriched in the colorectal tissue of CRC patients**

Given that *Fs* is a novel *Fusobacterium* species with adherent and immunomodulatory capabilities, we sought to determine its prevalence in human specimens from various body sites. Alignment of the *Fs* genome to metagenomic raw reads from 1,751 human specimens from the Integrative Human Microbiome Project (iHMP/HMP2) indicated that *Fs* is primarily detected in stool (161/419=38.42%), oral (304/1087 = 0.28%), nasal (2/49=0.04%), and vaginal (7/196= 0.04%) specimens (Fig. 4a). Although *Fs* was detected in 27.1% of patient specimens analyzed (*n*=474/1,751), the proportion of reads confidently aligning to its genome was typically less than $10^{-7}$, suggesting that *Fs* is a prevalent, but low abundant member of the human microbiota, particularly in stool and within the oral cavity.

To assess the prevalence and abundance of *Fs* in patients with CRC, the *Fs* 16S rRNA gene was aligned to gene amplicon data from 1,273 patient specimens (Supplementary Table 6). Our results indicate that while *Fs* is detected across all sample types (colorectal tissue 8/191=4.19%, mucosal tissue 1/228=0.44%, stool 3/854=0.35%), it is significantly enriched in colorectal tissue as

compared to mucosal tissue (p=0.0083) or stool (p<0.0001) (Fig. 4b). High precision analysis with DADA2[41] to improve downstream measures of diversity and dissimilarity revealed that the relative abundance (RA) of *Fs* in twelve *Fs* positive samples varies, with up to 24.4% RA in a single tumor tissue specimen (Fig. 4c-d). Collectively, these findings shed light on a novel bacterial species that is prevalent within the human intestinal microbiota, enriched in human CRC tissues, and whose role in human health and disease requires further investigation.

## ***Discussion***

Owing to the abundance of host-associated nucleic acids in complex specimens, tissue-associated microbiome analysis is currently restricted to amplicon-based sequencing. As such, applying traditional culturing-based approaches to tumor specimens provides a valuable opportunity to expand our understanding and classification of tumor associated microbiota[42]. Here, we report the isolation and characterization of a novel *Fusobacterium* species, *F. sphaericum sp. nov*. (*Fs*), from a primary colon adenocarcinoma of a female patient (Fig.1). Phylogenetic and pangenomic analyses reveal that *Fs* is most closely related to *F. perfoetens* (ATCC 29250), but support that *Fs* is a distinct species with predominantly unique gene content (Fig. 1b-d, Fig.2b).

Motivated by the putative cell-cell adherent and immunomodulatory capabilities of *Fs* genetic content (Fig. 2d), we investigated the interactions between *Fs* and colonic epithelial cells *in vitro*. Transmission electron microscopy and confocal microscopy of fluorescent immunocytochemically stained cells from *Fs* co-culture with human colon epithelial cell lines demonstrated that *Fs* adheres to but does not appear to invade eukaryotic cells, in contrast to the invasive capabilities of other *Fusobacterium* species, such as *F. nucleatum* (Fig. 3a). *Fs* cells

furthermore induced eukaryotic IL-8 secretion at higher levels than *F. nucleatum* (Fig. 3b-c). Recent studies have demonstrated increased secretion of IL-8 colorectal cancer (CRC) tissue and liver metastases[37,38] with a significant upregulation in the presence of *F. nucleatum*[40]. However, induction of IL-8 secretion by *F. nucleatum* is dependent on Fap2-mediated adhesion to host cells[40]. As *Fs* lacks Fap2 (Fig. 2c) and was not observed to invade colonic epithelial cells (Fig. 3a), further studies to identify the *Fs* factors that induce IL-8 secretion are warranted. As *Fs* induction of IL-8 secretion was comparable independent of bacterial cell viability, and to a lower extent with cell-free supernatants (Fig. 3b), it is possible such factors are secreted or present in outer membrane vesicles. Demonstrating that *Fs* induces IL-8 secretion adds a novel species to a growing body of evidence for a distinct role of microbes contributing to a pro-inflammatory environment within human tissue. In addition to chemokine induction, *Fs* may potentially metabolize anti-carcinogenic compounds, such as gallic acid[25] (Supplementary Fig. 3), to facilitate a pro-tumorigenic environment.

We demonstrate that in human metagenomic samples (n=1,751) *Fs* is a moderately prevalent but low abundant member of the human microbiota, particularly in the oral cavity and stool (Fig. 4a). In humans, *Fusobacterium* species composition differs more between the oral cavity and the gastrointestinal tract of the same individual than it does between individuals, with higher compositional variation in oral communities[43]. Future genomic studies including *Fs*, a species present in both niches, will be required to reveal its intricate relationships with other microbial members to decipher its impact on the human host. Our analyses also detected *Fs* in the human nasal cavity and vagina. While the detection of *Fs* in nasal samples may be a result of similarities in microbial communities in the nose and anatomically adjacent oral cavity, the presence of *Fs* in

the vaginal niche remains to be studied. It should also be noted that in these analyses, the overall prevalence of detection (27.1%) is likely an underestimate, due to the limit of detection for low abundant members in metagenomic sequencing, and that measurements of relative abundance cannot be converted in a straightforward manner to a proportion of genome copies since the genome sizes are unknown. Using 16S-based alignments, we additionally demonstrate that in patients with colorectal cancer, *Fs* is significantly enriched in colonic tissue (tumor and adjacent normal) as compared to mucosal tissue or stool (n=1,273 specimens) (Fig. 4b-d). Whether the presence of *Fs* in colorectal tissue may contribute to a model of dysbiosis and chronic bacterial infection in the context of CRC requires further investigation in animal models and longitudinal studies in large human cohorts.

In conclusion, in this study we have emphasized and demonstrated the importance of reductionist approaches, such as traditional microbial culturing, to isolate and identify members of tissue-associated microbiomes as they are not yet amenable to metagenomics. Next generation sequencing techniques, such as long-read SMRT-Sequencing[13], allows for the assembly of high-quality, closed, and complete genomes from which phylogenetic characterization, putative functions and virulent capabilities can be predicted through comparative genomic analyses. Deciphering the role of identified novel microbial species within tumor-associated bacterial communities on specific disease outcomes remains crucial for informing future preventative and therapeutic interventions.

***Figures***

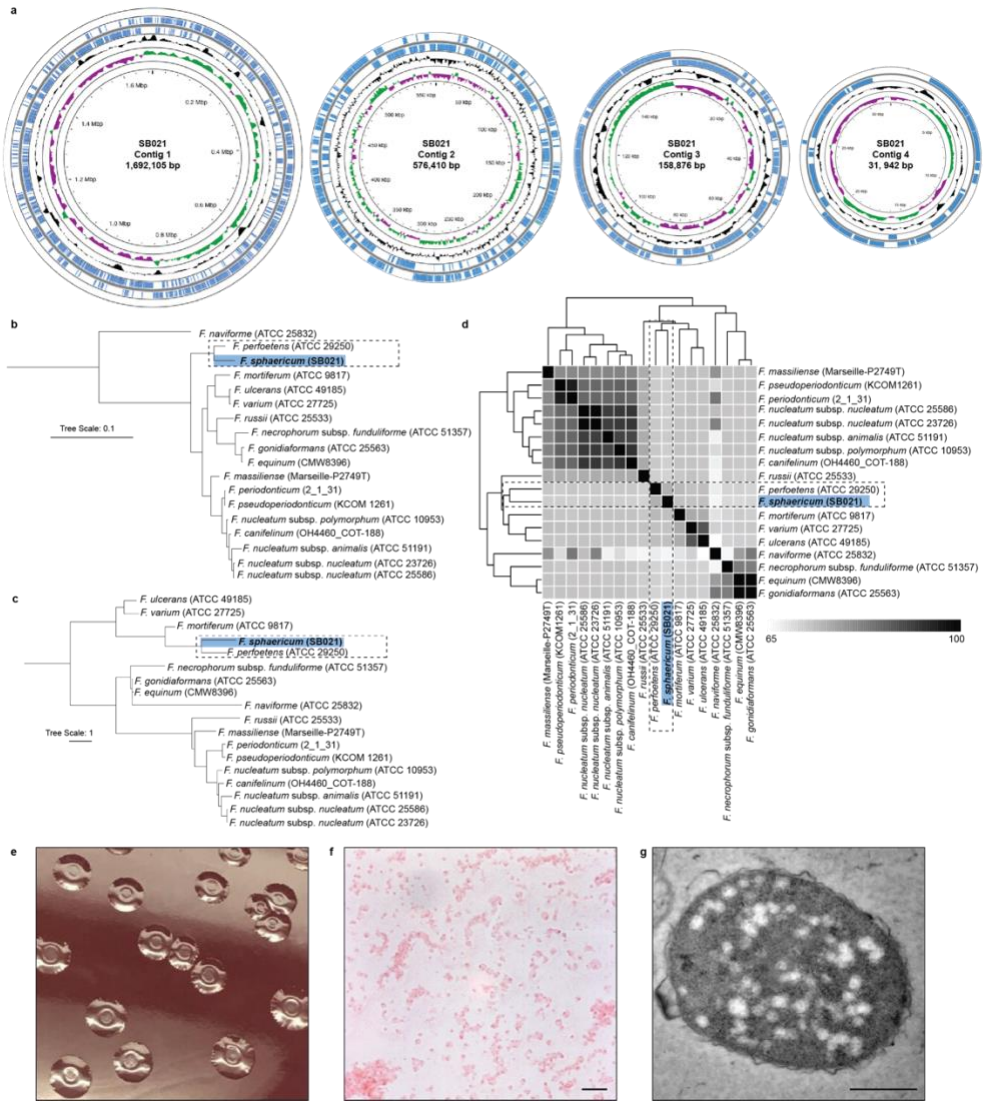



**Fig. 1: Taxonomic and Morphologic Analysis of *Fusobacterium sphaericum sp. nov.* (SB021)**

**a)** Proksee circular genome architecture of *F. sphaericum sp. nov.* SB021 contigs. From outwards to inwards, the first and second circles show annotated coding sequences on positive strand and negative strand, respectively. Third circle indicates average GC content, with areas higher than average GC content as outward projections and areas lower than average as inward projections. Fourth circle shows GC skew, with above zero values in green and below zero values in purple. **b)** A 16S rRNA gene-based dendrogram illustrating the phylogenetic relationship between *F.*

*sphaericum sp. nov.* (blue) and other *Fusobacterium spp*. **c)** kSNP[15] maximum-likelihood whole-genome phylogenetic tree with *F. sphaericum sp. nov.* highlighted in blue. **d)** Clustered average nucleotide identity (ANI) matrix with *F. sphaericum sp. nov.* highlighted in blue. ANI values are reported in Supplementary Table 1. White to black scale indicates and ANI from 65% to 100%, respectively. **e)** Digital photography of *F. sphaericum sp. nov.* SB021 colonies grown on fastidious anaerobe agar enriched with 10% defibrinated horse blood. **f)** Bright-field microscopy of Gram-stained *F. sphaericum sp. nov.* SB021. Images acquired at 100X magnification with oil immersion lense. Image scale bar is 10 µm. **g)** Transmission electron micrographs of an *F. sphaericum sp. nov.* SB010 cell. Image scale bar is 500 µm.

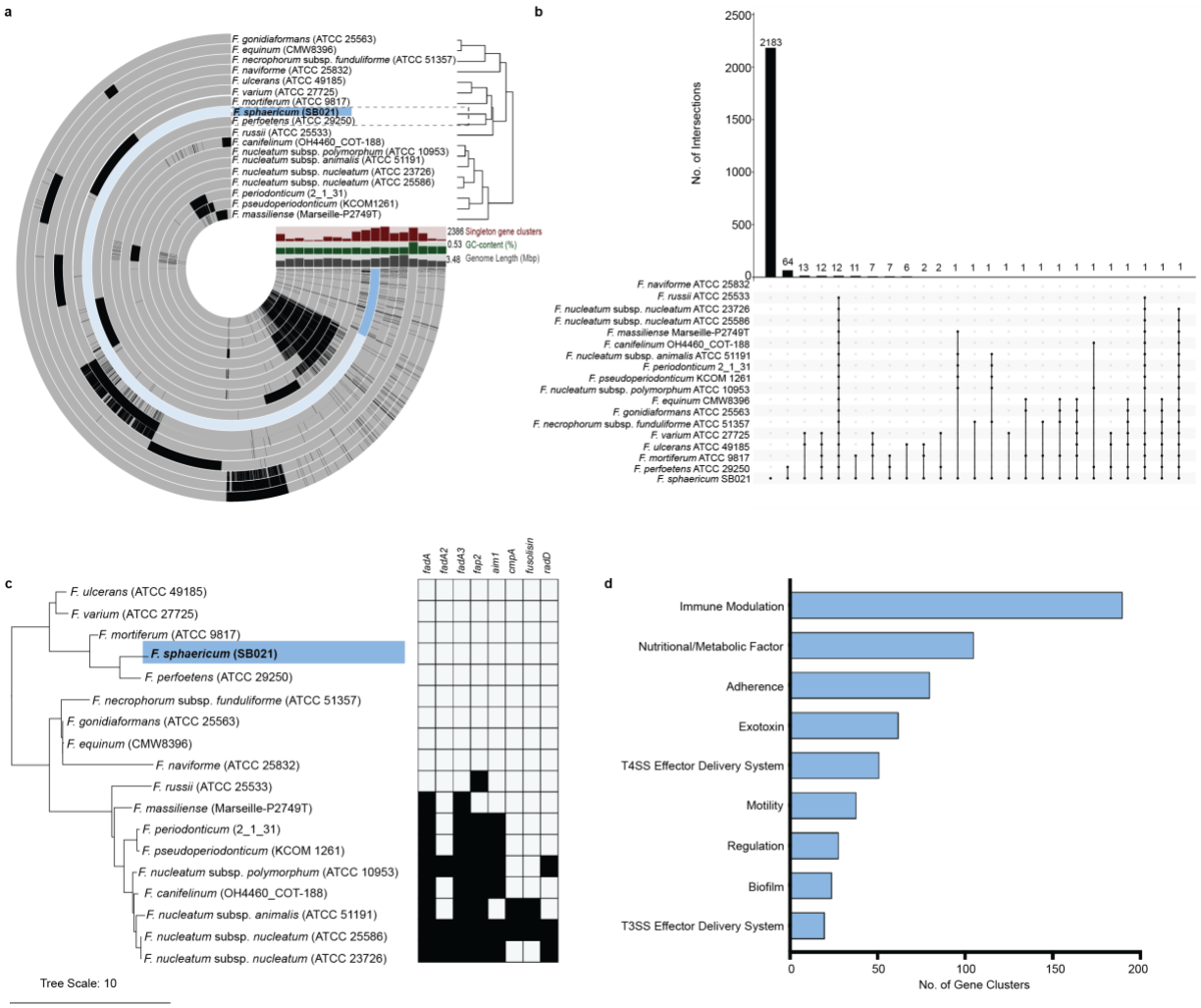
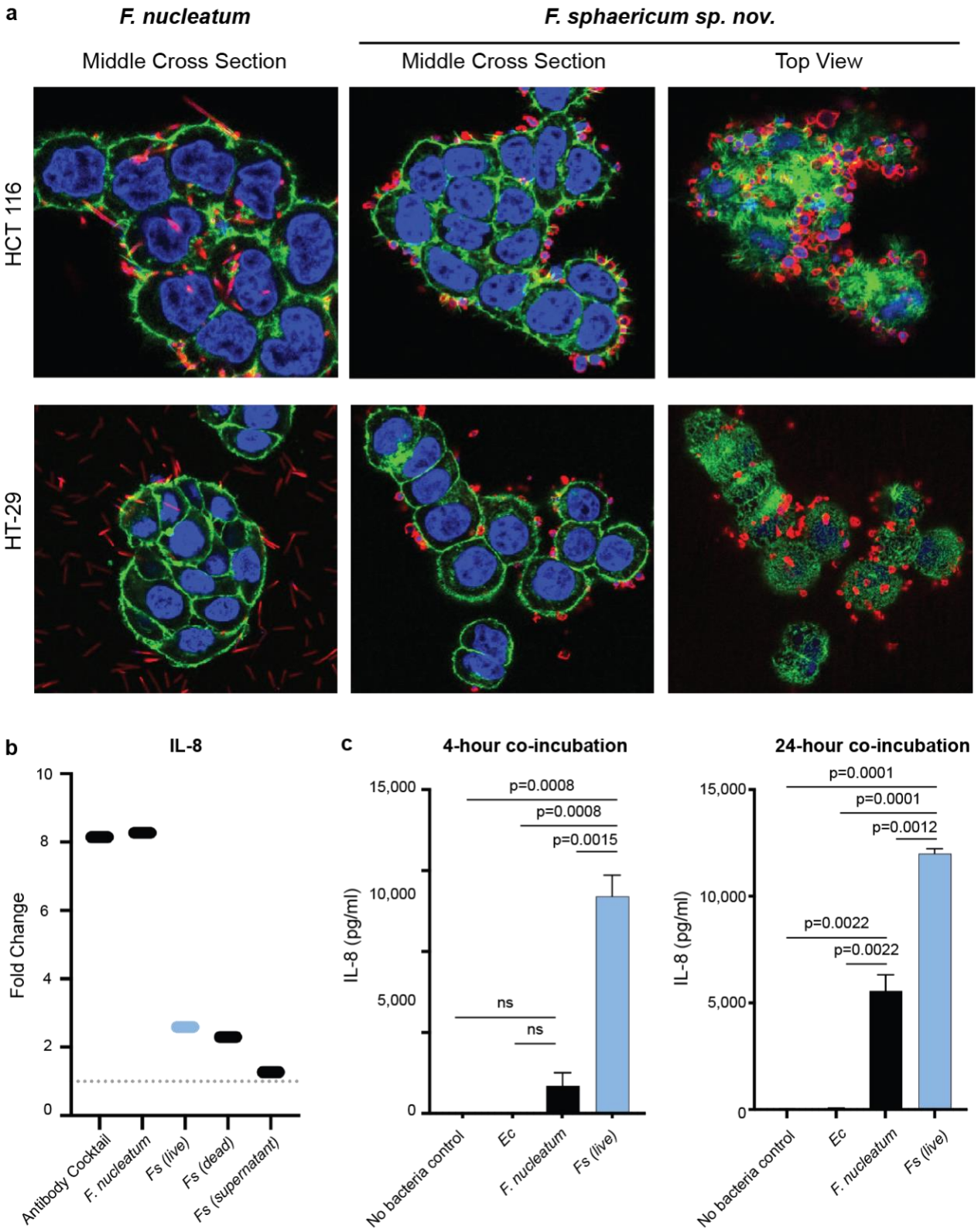



**Fig. 2: *Fusobacterium sphaericum sp. nov.* (SB021) Gene Content Analysis**

**a)** Anvi'o pangenomic analysis[19] of *F. sphaericum sp. nov.* SB021 (blue) as compared to seventeen *Fusobacterium spp.* genomes (Table 1). Each circle represents an individual genome, with identified gene clusters in black. Genomes are ordered by ANI clustering (Fig. 1d). **b)** Upset plot depicting the presence of gene clusters identified in *F. sphaericum sp. nov.* across other *Fusobacterium spp.* genomes. Bar height (top) indicates number of gene clusters shared between combinations of genomes (bottom). **c)** Presence (black) versus absence (white) plot of canonical virulence factors across *Fusobacterium spp.* genomes, ordered by a kSNP[15] maximum-likelihood whole-genome phylogenetic tree with *F. sphaericum sp. nov.* highlighted in blue. **d)** *F. sphaericum*

*sp. nov.* anvi'o-identified gene clusters were analyzed using the virulence factor database (VFDB)[36]. Graph shows the number of gene clusters in each of the top functional categories of putative virulence factors (Supplementary Table 4).


a
F. nucleatum
F. sphaericum sp. nov.
Middle Cross Section
Middle Cross Section
Top View
HCT 116
HT-29
b
IL-8
Fold Change
10
8
6
4
2
0
Antibody Cocktail
F. nucleatum
Fs (live)
Fs (dead)
Fs (supernatant)
c
4-hour co-incubation
IL-8 (pg/ml)
15,000
10,000
5,000
0
p=0.0008
p=0.0008
p=0.0015
ns
ns
No bacteria control
Ec
F. nucleatum
Fs (live)
24-hour co-incubation
IL-8 (pg/ml)
15,000
10,000
5,000
0
p=0.0001
p=0.0001
p=0.0012
p=0.0022
p=0.0022
No bacteria control
Ec
F. nucleatum
Fs (live)

313

**Fig. 3: *Fusobacterium sphaericum sp. nov.* adheres to colonic epithelial cells and stimulates IL-8 secretion**

**a)** Confocal laser scanning microscopy of colon cancer epithelial cell lines, HCT 116 and HT-29, co-incubated with *F. nucleatum* SB010 or *F. sphaericum sp. nov.* (*Fs*) SB021. Bacterial cell staining: FM4-64X (red) and DAPI (blue). Eukaryotic cell staining: DAPI (blue), actin (green). **b)** Comparison of IL-8 secretion measured via the Multi-Analyte ELISArray from HCT 116 cells co-incubated with different *Fusobacterium* species or antibody cocktail control, normalized to a no-bacteria control. Bacterial species used included *F. nucleatum* SB010 and *Fs* SB021 live cells or oxygen-exposed dead cells. An additional co-incubation with *Fs* supernatant alone was also performed. Additional cytokine and chemokine measurements are shown in Supplementary Fig. 4. **c)** Quantitative levels of IL-8 secretion measured via the Single-Analyte ELISArray from HCT 116 cells incubated alone or co-incubated with *Esherichia coli* TOP10, *F. nucleatum* SB010, or *Fs* SB021 at 4-hour and 24-hour time points. Statistical analysis performed via a one-way ANOVA. For **a-c**, all co-incubations were at a multiplicity of infection (MOI) of 100.

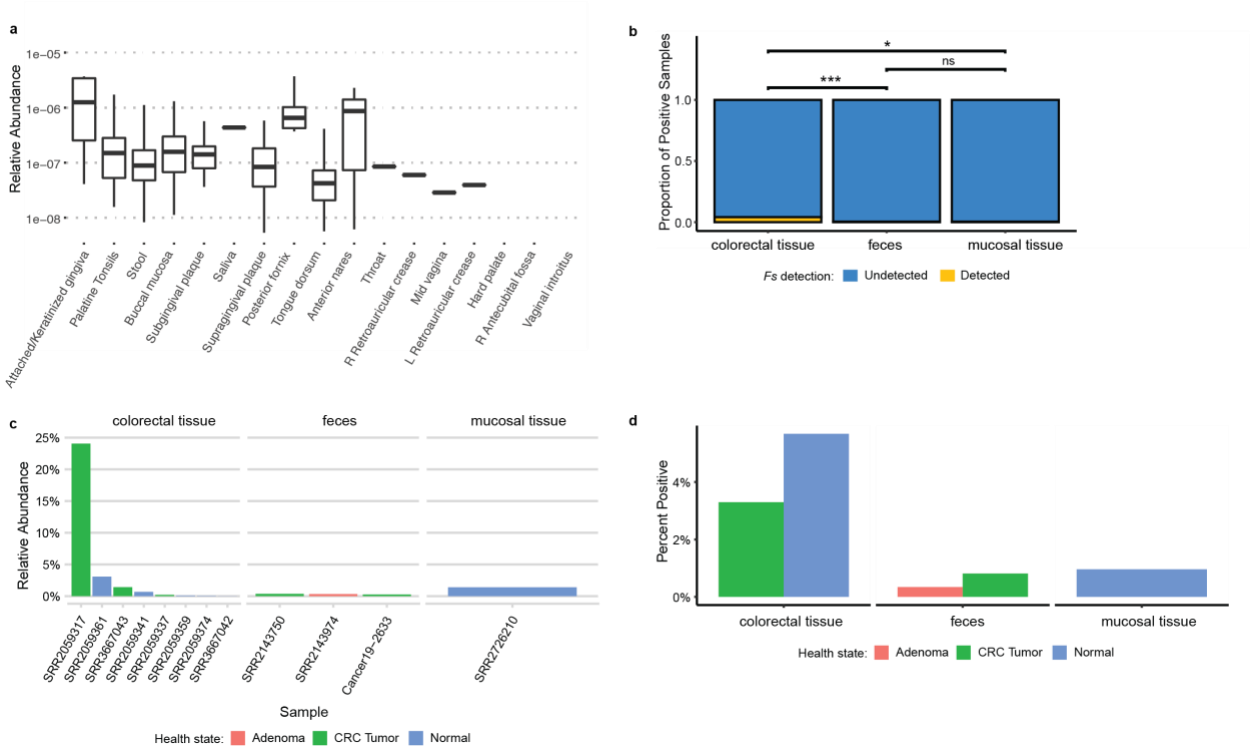



**Fig. 4: *Fusobacterium sphaericum sp. nov.* is prevalent across human body sites**

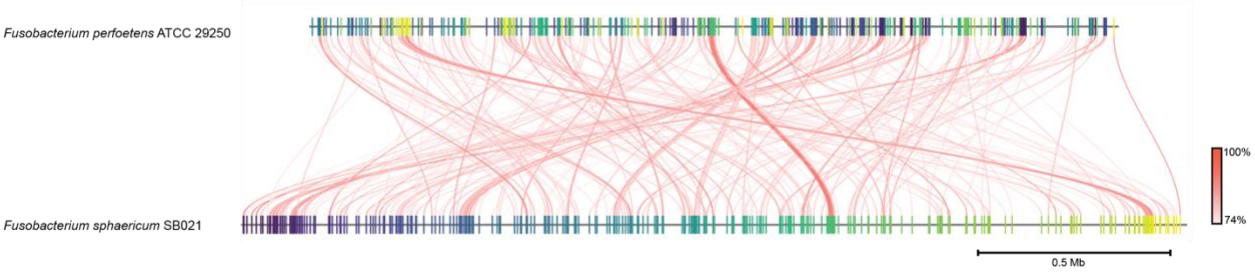
**a)** *F. sphaericum sp. nov.* (*Fs*) SB021 genome was aligned to metagenomic raw reads for 1,751 specimens across different human body sites from the Integrative Human Microbiome Project (iHMP/HMP2). **b-d)** The *Fs* 16S rRNA gene was aligned to gene amplicon data from 1,273 CRC patient specimens. Plots shows **b)** Proportion of *Fs* positive and *Fs* negative samples, **c)** a panel of single *Fs* positive samples (n=12, RA>0) and **d)** the percentage of *Fs* positive samples (n=12, RA>0), by sample type (colorectal tissue, feces, mucosal tissue) and patient health status (adenoma: red, CRC tumor: green, normal: blue).

## Table Legends

**Table 1: Genomes used in this study**
Table indicates each *Fusobacterium* genome used for comparative analyses in this study with their corresponding genome accession numbers, genome length, GC content, and number of contigs in the assembly noted.

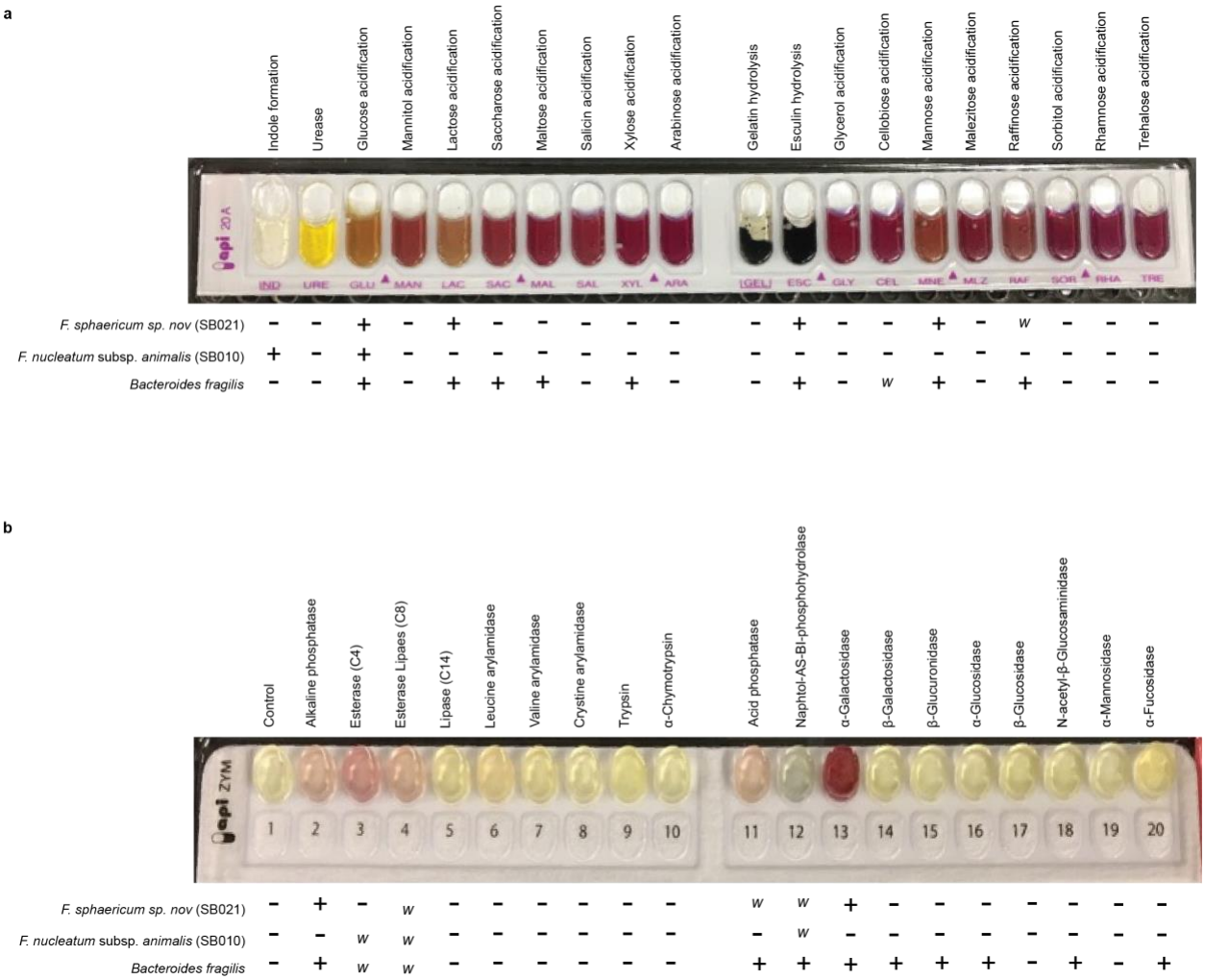
**Table 2: *F. sphaericum sp. nov.* antibiotic resistance profile**
Table shows results for *F. sphaericum sp. nov.* susceptibility or resistance to nine tested antibiotic compounds.

***Supplementary Figures***



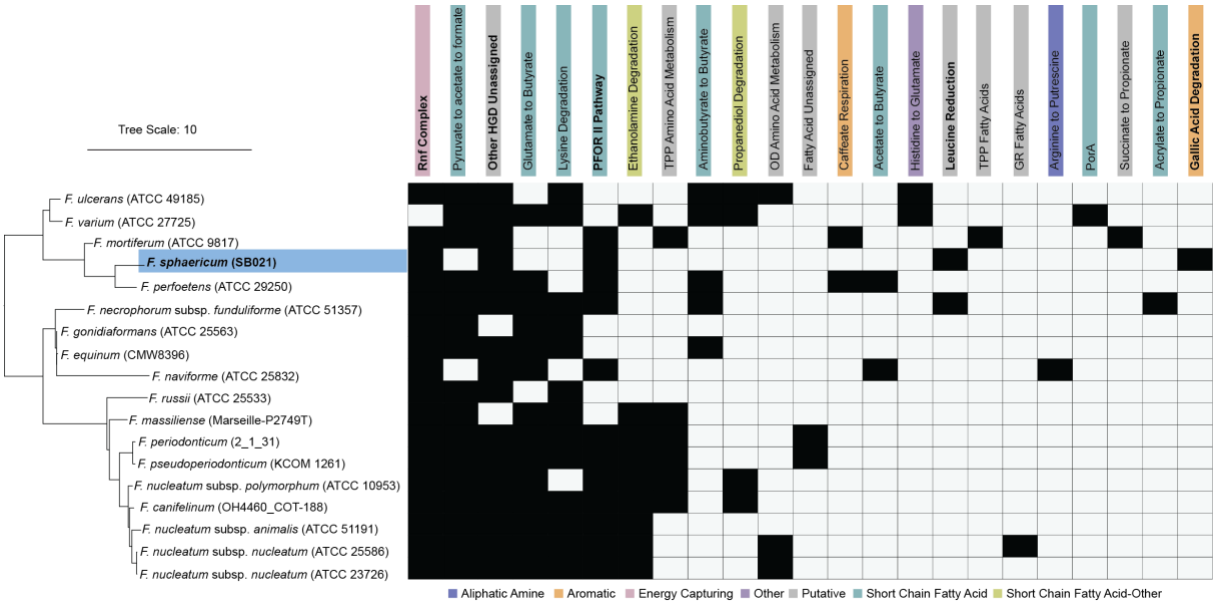

**Supplementary Fig. 1: Genomic architecture of *Fusobacterium sphaericum sp. nov.* SB021 as compared to *Fusobacterium perfoetens* ATCC 29250.**

Visualization of fastANI alignment between *Fusobacterium sphaericum sp. nov.* SB021 and its nearest phylogenetic neighbor, *Fusobacterium perfoetens* ATCC 29250, where each red line segment denotes a reciprocal mapping of the 275 orthologous matches between the two genomes, indicating evolutionarily conserved regions.



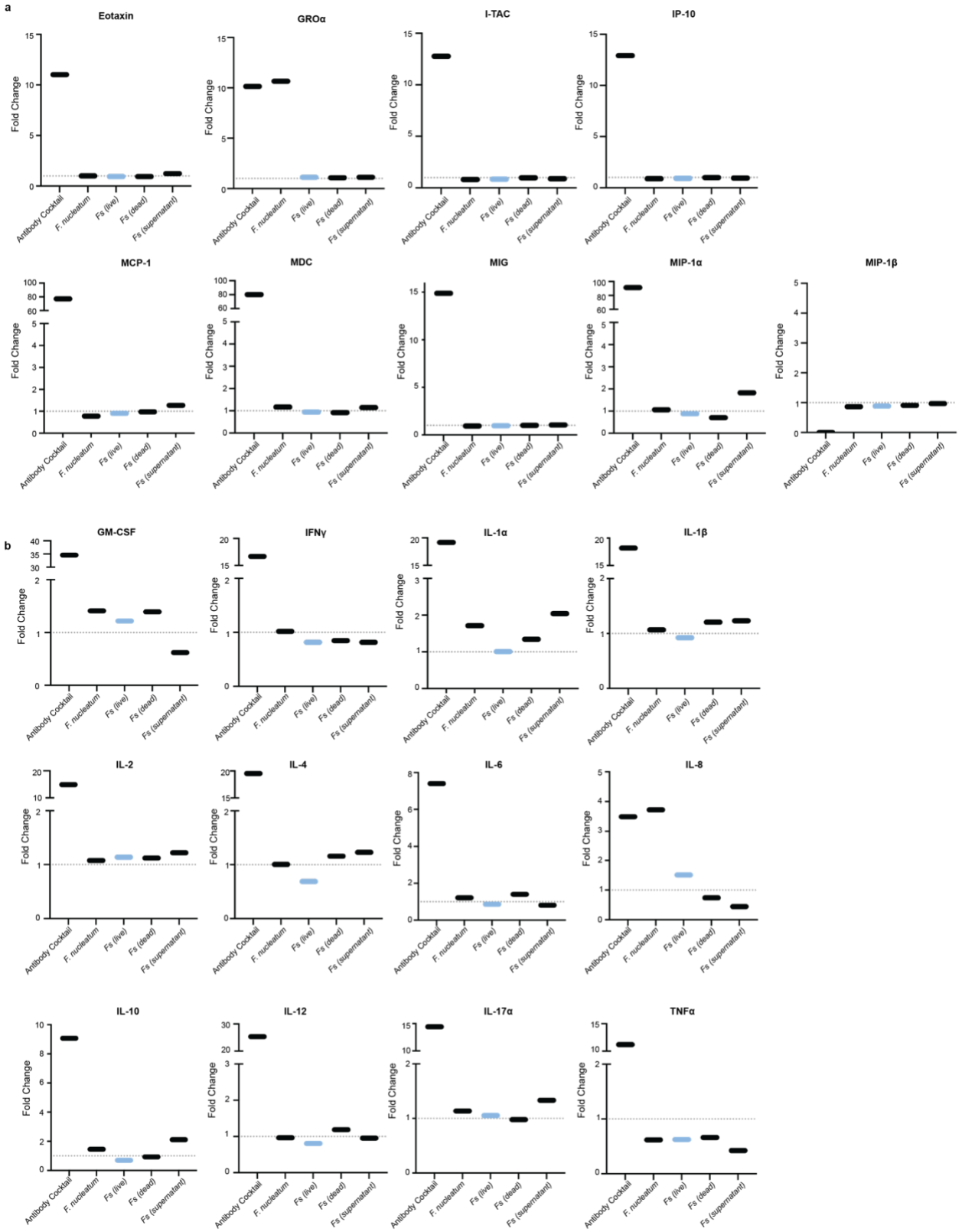

**Supplementary Fig. 2: Metabolic characteristics for *Fusobacterium sphaericum sp. nov.***

*F. sphaericum sp. nov.* **a)** metabolic and **b)** chemotaxonomic characteristics assayed by API 20A and API ZYM detection kits, respectively. Images are representative results for *F. sphaericum* SB021. Results for *F. sphaericum sp. nov.* SB021, *Fusobacterium nucleatum* subspecies *animalis* SB010, and *Bacteroides fragilis* are summarized below each API strip, where (+) indicates a positive result, (-) a negative result, and "*w*" a weakly positive result.




**Supplementary Fig. 3: Primary metabolic pathways in *Fusobacterium sphaericum* SB021**

Presence (black) versus absence (white) plot of primary metabolic pathways identified by gutSMASH[24] across *Fusobacterium spp.* genomes, ordered by a kSNP[15] maximum-likelihood whole-genome phylogenetic tree with *F. sphaericum sp. nov.* SB021 highlighted in blue.




**Supplementary Fig. 4: Cytokine and chemokine induction in HCT 116 cells co-incubated**

**with *F. sphaericum sp. nov.***

Measurements of **a**) common chemokine and **b**) inflammatory cytokine levels measured via the Multi-Analyte ELISArray from HCT 116 cells co-incubated with different *Fusobacterium* species at a multiplicity of infection (MOI) of 100 or antibody cocktail control, normalized to a no-bacteria control. Bacterial species used included *F. nucleatum* SB010 and *F. sphaericum sp. nov.* (*Fs*) SB021 live cells or oxygen-exposed dead cells. An additional co-incubation with *Fs* supernatant alone was also performed.

## Supplementary Table Legends

**Supplementary Table 1: RAST annotation of *F. sphaericum sp. nov.* genome**

The *F. sphaericum sp. nov.* (SB021) genome was annotated by the Rapid Annotation using Subsystem Technology (RAST) pipeline. Table indicates resulting genes for each contig, with their length and predicted product.

**Supplementary Table 2: BLASTn analysis of *F. sphaericum sp. nov.* 16S rRNA gene**

Table shows BLASTn top 100 results for the *F. sphaericum sp. nov.* 16S rRNA gene by decreasing percent identity.

**Supplementary Table 3: Average nucleotide identity (ANI) scores between *Fusobacterium* genomes**

Table shows the pairwise average nucleotide identity (ANI) scores between each *Fusobacterium* genome (Table 1).

**Supplementary Table 4: Analysis of *F. sphaericum sp. nov.* anvi'o-identified gene clusters**

The *Fusobacterium* pangenome was characterized using the analysis and visualization platform for microbial 'omics (Anvi'o)[19]. For each gene cluster (GC) identified in *F. sphaericum sp. nov*., the table shows its representive amino acid sequence, its mapped KO identifier by KEGG ortholog mapping via KofamKOALA[20], and any identified hits by the virulence factor database[36] (VFDB).

**Supplementary Table 5: CARD analysis of *F. sphaericum sp. nov.* genome**

Table shows results for the analysis of the *F. sphaericum sp. nov.* genome against the comprehensive antibiotic resistance database[26] (CARD).

## *Methods*

### Bacterial isolation and culturing

Written informed consent was obtained from patient as approved by the Fred Hutchinson Cancer Center Institutional Review Board. Primary tumor tissue was snap frozen and obtained from a biorepository. A small cross section of the tumor was cut from the main tissue mass, lightly scored, and then spread onto a small section on fastidious anaerobe agar (FAA) plates (Oxoid, ThermoFisher Scientific, Waltham, MA, USA) supplemented with 10% defibrinated horse blood (DHB; Lampire Biological Laboratories, Fisher Scientific, Pipersville, PA, USA). Using a sterile loop, the section was spread on the plates for the isolation of single colonies and subsequently, plates were incubated at 37°C under anaerobic conditions. Every 48 hours the plates were checked for new growth and the identity of colonies was determined by colony PCR. For all subsequent experiments, *F. sphaericum sp. nov* was grown for 48 hours on FAA+10% DHB plates supplemented with josamycin (3 μg/ml), vancomycin (4 μg/ml), and norfloxacin (1μg/ml) (Sigma Aldrich, USA) and incubated at 37°C under anaerobic conditions (AnaeroGen Gas Generating Systems, Oxoid, ThermoFisher Scientific, USA).

### 16S rRNA bacterial colony PCR

Single colonies were collected with a 1μl loop and suspended in 20μl of nuclease free water (Promega, Madison WI, USA)**.** The same 1μl loop was then used to streak a FAA + 10% DHB plate to subculture and isolate pure colonies. The suspensions were placed in a thermocycler at 98°C for 20 minutes and then cooled to 4°C. 2μl of the lysate was subsequently used in a 20μL PCR reaction (Promega, Madison WI, USA) with forward primer 342F (5'- CTA CGG GGG GCA GCA G -3') and reverse primer 1492R (5'- TAC GGY TAC CTT GTT ACG ACT T -3') at a final concentration of 1μM. PCR products were shipped for Sanger Sequencing and sequencing results

were analyzed using NCBI's BLASTn for species level identification. Identified strains were stocked in 60% TSB with 40% glycerol (v/v) at -80°C.

**High-molecular weight genomic DNA (gDNA) extraction**

High molecular weight genomic DNA was extracted using the MasterPure™ Gram Positive DNA Purification Kit (Epicentre, Lucigen, USA). Cells from two plates were re-suspend in 1.5mL 1X PBS and harvested by centrifugation. Pellets were subjected to manufacturer instructions modified by doubling all reagent volumes and removing vortexing steps to prevent DNA shearing. HMW gDNA was quantified using a Qubit fluorometer (ThermoFisher Scientific, USA).

**PacBio Single-Molecule Real-Time Sequencing and Genome Assembly**

Single molecule real-time sequencing (SMRT-Seq)[13] was carried out on a PacBio Sequel instrument (Pacific Biosciences, USA) or a PacBio Sequel II instrument (Pacific Biosciencies, USA) at the University of Minnesota Genomics Center. Sequencing reads were processed using Microbial Assembly pipeline within Pacific Biosciences' SMRTAnalysis pipeline version 9.0.0.92188. Additional assembly was performed using Flye assembler version 2.8 as needed (https://github.com/fenderglass/Flye). Functional predictions for genes was conducted with the Rapid Annotations using Subsystems Technology (RAST)[44,45] tool (Supplementary Table 1).

**Phylogenetic classification**

The full-length 16S rRNA gene sequence for *F. sphaericum sp. nov.* SB021 was extracted from its whole genome sequence and analyzed using NCBI BLASTn (Supplementary Table 2). For available *Fusobacterium* type strains, the 16S rRNA gene sequence was downloaded and all resulting sequences were aligned via MEGA X[46] using the MUSCLE clustering algorithm from which a maximum-likelihood dendrogram was generated. kSNP3[15] with a kmer size of 17, resulting in a fraction of core kmers (FCK) of 0.113, was used to generate a maximum-likelihood

whole-genome, reference-free phylogeny of these strains. Final tree were generated using the interactive tree of life (iTOL) tool, version 5[47]. Average nucleotide identity was calculated using the EZ BioCloud ANI calculator[48].

**Pangenomic Analysis**

Pangenome analysis was performed using the Analysis and Visualization platform for microbial 'omics (Anvi'o) workflow[19] with thresholds set to a minbit of 0.9 and an MCL of 2. All F. sphaericum sp. nov. gene clusters were analyzed via KofamKOALA[20] for KEGG ortholog mapping for their putative functions and analyzed against the Virulence Factor Database (VFDB)[36] to identify putative virulence-associated genes.

**Bacterial Gram Staining**

Gram staining was performed on *F. sphaericum sp. nov.* uniform bacterial smears after growth at 37°C for 48 hours followed by toxic oxygen exposure using a Gram Stain Kit (Remel, Lenexia, KS, USA). The stains were imaged with TissueFAXS microscope system (TissueGnostics, Vienna, Austria) with 100x oil immersion at Fred Hutchinson Cancer Center, Seattle, WA, USA.

**Metabolic Profiling**

The metabolic profile of *F. sphaericum sp. nov.* was determined with the API 20A Gallery System (bioMérieux, Marcy-l'Étoile, France) by utilization of colonies after 24h anaerobic culture. The colonies were resuspended in phosphate-buffered saline (PBS, Corning, Corning, NY, USA) to obtain an $OD_{600}$ (optical density at 600nm) of 0.4-0.5 (=3 McFarland) for inoculation of pre-coated cupules on the incubation strip. Results were read after 24 hours and additional procedures were carried out as indicated by the manufacturer. The enzymatic profile was investigated similarly with the API Zym Gallery System (bioMérieux, Marcy-l'Étoile, France), except for an $OD_{600}$ of 0.7

(=5-6 McFarland) and an incubation period of 4-4.5 hours. Both API 20A and API Zym were performed with 3 replicates.

**Antibiotic Sensitivity Testing**

In vitro testing of antibiotic susceptibility of F. sphaericum sp. nov. was performed on harvested bacteria after 24 hours incubation at 37°C under anaerobic conditions. Cells were resuspended in PBS to an $OD_{600}$ of 0.5 and used to seed bacterial lawns on FAA+10% DHB plates. Antibiotic reading strips (MIC test strips; Liofilchem Diagnostics, Waltham, MA, USA) were placed on the lawns with sterile tweezers. All experiments were performed under anaerobic conditions and in triplicates for each antibiotic agent including negative controls (no bacteria). Results were read after 24 hours of incubation.

**Mammalian Cell Culture**

Mammalian cell culture was performed on the cell lines HCT-116 (CCL-247; ATCC, Manassas, VA, USA) and HT-29 (HTB-38; ATCC, Manassas, VA, USA). Mammalian cells were grown in T75 culture flasks using McCoy's 5A culture medium (Iwakata and Grace Modification) (Corning, Corning NY, USA) with 10% fetal bovine serum (FBS; Sigma-Aldrich, St. Louis, MO, USA), and antibiotics (1x Pen/Strep: penicillin (5000 units/ml), streptomycin (5000 µg/ml), Gibco, Carlsbad, CA, USA) and incubated constantly in a 37°C/5% $CO_2$ air humidified incubator. Cells were checked microscopically for monitoring purposes (growth rate, confluency) and split upon reaching confluency (>=80%). Trypsin-EDTA (Gibco, Carlsbad, CA, USA) was added to the flask to dissociate cells from flask surface, followed by neutralization of the dissociation agent with growth medium. Subsequently, the cell suspension was centrifuged for 3 minutes at 302 RCF at room temperature. The cell pellet was resuspended in sterile medium and seeded into new flasks. Passage numbers were recorded and did not exceed $n$=10. Prior to co-culture experiments, cells

were split as described above and resuspended in antibiotic-free sterile growth medium, added to 6-well plates to a final concentration of $1.5x10^6$ cells per well and incubated for 24 hours before experimental testing.

**Scanning Transmission Electron Microscopy (S/TEM)**

*F. sphaericum sp. nov.* was cultured as described above. Cell size of *F. sphaericum sp. nov.* was determined via scanning transmission electron microscopy (S/TEM). We co-incubated HCT 116 human colorectal carcinoma cell line (CCL-247; ATCC, Manassas, VA, USA) at low passage (<10) with *F. sphaericum sp. nov.* ($OD_{600}$ = 1.0) for 4 hours followed by fixation with ½ strength Karnovsky's fixative (2.5% glutaraldehyde and 2% paraformaldehyde in 0.1M cacodylate buffer) and further sample preparation on the same day. Specimens were viewed on a JEOL JEM 1400 transmission electron microscope (JEOL, Tokyo, Japan) with an accelerating voltage of 120kV. Images were acquired with a Gatan Rio 4k x 4k CMOS Camera (Gatan, AMETEK, Berwyn, PA, USA) at Fred Hutchinson Cancer Center, Seattle, WA, USA.

**Confocal Microscopy of *Fusobacterium* co-cultures with human cell lines**

*Fusobacterium* strains, *F. nucleatum* subspecies *animalis* SB010 and *F. sphaericum sp. nov.* SB021, were grown for 48 hours at 37°C in anaerobic conditions and resuspended in 1 mL of PBS at a final concentration of $5x10^8$ cells/mL. Bacterial membranes were stained in this suspension with 50 µl of 100 mg/mL FM 4-64FX (Molecular Probes) and washed twice in PBS by centrifugation for 7,000 x g at RT. Cells from human adenocarcinoma tumor cell lines, HCT 116 and HT-29, were grown and passaged as described above and resuspended in 1 mL of McCoys 5A + 10% FBS. $6x10^4$ HCT 116 or HT29 cells were seeded with an MOI of 100 bacterial cells in 6-well plates. Bacteria were centrifuged onto HCT 116 and HT-29 at 300 x g for 5 minutes at room temperature (RT), and plates were immediately placed in a 37°C $CO_2$ incubator and cultured

together for 4 hours. Cell lines were then washed 3x with PBS (all washes are 5 min in duration at RT) and then fixed in 4% paraformaldehyde in PBS for 30 min at RT. Following fixation, wells were washed 3x in PBS and then permeabilized with 0.2% (*v/v*) Triton X-100 in PBS for 4 mins at RT. Cells were washed 3x in PBS and then stained for 20 min at RT with 2 drops/mL of NucBlue Fixed Cell Strain ReadyProbes (Invitrogen, Carlsbad, CA, USA) and ActinGreen 488 Ready Probes (Invitrogen) to stain DNA and actin, respectively. Cells were washed a final 3x with PBS and then the coverslips were mounted onto glass slides in ProLong Gold antifade mounting medium (Invitrogen). Samples were viewed with a Leica SP8 confocal laser-scanning microscope (Leica, Wetzlar, Germany) for image acquisition. Representative confocal micrographs of 1024 x 1024 pixels (pixel size: 103.3 nm) were acquired and assembled using Fiji (with Bio-Formats Plugin)[49].

**Analysis of Chemokine and Cytokine Induction in HCT 116**

To determine the immunologic impact of *Fusobacterium* presence on HCT 116 human colon epithelial cells, we utilized the Multi-Analyte ELISArray kits protocol version 1.4 from SABiosciences (Qiagen, Valencia, CA, USA) for common human chemokines and inflammatory human cytokines according to the manufacturer's instructions. Tests were performed with the following samples after a 4-hour co-incubation with HCT 116 (low passage, <10): *F. sphaericum sp. nov.* at MOI of 1:100 (live and oxygen-exposed), *F. sphaericum sp. nov.* cell-free supernatant, *F. nucleatum* subsp. *animalis* SB010 at MOI 1:100 (live), a negative control (HCT-116 only) and a positive control (cocktail of antibodies). Supernatants were centrifuged at 1,000 x g for 10 minutes to remove any particulate material, filter sterilized and 50μl of each experimental sample were added to the array with the specific capture antibodies (IL-1α, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-12, IL-17α, IFN-γ, TNF-α and GM CSF; IL-8, MCP-1, RANTES, MIP-1a, IP-10, I-TAC, MIG,

Eotaxin, TARC, MDC, GROα), followed by an incubation at room temperature (RT) for 2 hours. Subsequently, the wells were buffer washed and 100µl of diluted biotinylated detection antibodies were added for a 1h-incubation at RT in the dark. After another washing procedure, the wells were filled with 100µl of Avidin-horseradish peroxidase (HRP) and incubated at RT for 30 minutes, again in the dark. We then added development and stop solutions to detect absorbance changes immediately after termination of the experiment with a microplate reader (Synergy H4 Hybrid Reader, BioTek) at 570nm and 450nm, visualized with Gen5 Synergy H4 software (Version 3.08, BioTek). Raw data of the absorbance reading were normalized to the cell counts of HCT 116 cells incubated without bacteria. This experiment was executed in duplicates.

For quantification of immunologic effects, we utilized the Single-Analyte ELISArray kit protocol version 1.4. from SABiosciences (Qiagen, Valencia, CA, USA) for the chemokine IL-8 according to the manufacturer's instructions. Co-cultures were set up as described above with *F. sphaericum sp. nov.* at MOI of 1:100 (live), *F. sphaericum sp. nov.* cell-free supernatant, *F. nucleatum* subsp. *animalis* SB010 at MOI 1:100 (live), *Escherichia coli* TOP10, or a positive control (HCT-116 only) and incubated for 4-hours or 24-hours. Supernatants were centrifuged at 1000 x g for 10 minutes and filter sterilized subsequently. A serial dilution of the antigen standard was prepared and 50µl of each sample was added to the array, followed by an incubation for 2 hours at RT. Similar to the Multi-Analyte kit, the array was then incubated with detection antibodies and subsequently Avidin-HRP to determine the attachment of detection antibodies to the chemokine. After the color development, changes of absorbance were read at 570nm and 450nm and raw data were equally normalized. This experiment was executed in duplicates. Statistical comparison was

calculated by applying a one-way ANOVA using GraphPad Prism v7.0 Software (GraphPad Software).

**Prevalence Analysis within CRC patient Specimens (16S rRNA Amplicon Sequencing Data)**

To assess the potential associations of *F. sphaericum sp. nov*. with colorectal tissue, mucosa or feces in human specimens from healthy individuals and patients with colorectal cancer (CRC), we identified studies and corresponding BioProjects from publicly available data on NCBI BioProject using the keyword phrases ("fecal/stool", "16S rRNA amplicon metagenome"). Studies were also identified from data cited in recent research publications. Our inclusion criteria for studies were available paired-end, raw 16S rRNA Amplicon Sequencing data and complete corresponding metadata of the individuals, detailing health/disease state, material type, and body site for each individual run. Studies containing single-ended data, or data generated from 454 and shotgun sequencing, or focusing on conditions other than CRC (e.g. inflammatory bowel diseases (IBD), Chron's disease (CD), ulcerative colitis (UC), obesity) were excluded. The ascension list and metadata file for each BioProject was downloaded from NCBI's run selector and runs were downloaded and split using SRA Toolkit v2.9.2 (http://ncbi.github.io/sra-tools/), other studies' data were acquired through the original published dataset. The complete list of the *n*=7 BioProjects comprising *n*=1273 samples utilized for this analysis is provided in Supplementary Table 6.

DADA2 (Divisive Amplicon Denoising Algorithm 2) is an R package to infer exact amplicon sequence variants (ASVs) from high-throughput amplicon sequencing data[41]. In this study, DADA2 was used to process paired-end, raw 16S rRNA Amplicon data from selected studies. For each study data, low-quality bases from both ends were trimmed and reads with more than 2 expected errors were discarded. Subsequently, reads were deduplicated and ASVs were inferred.

Naive Bayesian classifier method implemented in DADA2 was applied to assign taxonomy to genus level, trained on the Silva database (version 132)[50]. Species level assignments were processed based on exact matching between ASVs and the Silva database with an additional *F. sphaericum sp. nov.* 16S rRNA sequence. ASVs assigned to eukaryotes or mitochondria were excluded and ASVs assigned only to the bacteria kingdom level were also removed from our downstream analysis. Relative abundance for each ASV was then calculated based on the number of reads in the corresponding ASV versus total number of reads.

To investigate *F. sphaericum sp. nov*. potential associations with tissue, mucosa and feces, we created a single dataframe comprising every ASV from the list of BioProjects. Accompanying metadata was added to this dataframe to generate a list containing taxonomic assignments, relative abundance, and other metadata for each ASV. Subsequently, the sheet was used to generate a binary outcome (true/false) list for the presence or absence of *F. sphaericum sp. nov*., while presence was defined as a relative abundance > 0. Statistical comparisons between sample types were calculated by applying a two-proportion Z-test.

***References***


1. Nejman, D. *et al.* The human tumor microbiome is composed of tumor type–specific intracellular bacteria. *Science* **368**, 973–980 (2020).
2. Poore, G. D. *et al.* Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020).

3. Flanagan, L. *et al.* Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur. J. Clin. Microbiol. Infect. Dis.* **33**, 1381–1390 (2014).
4. Kostic, A. D. *et al.* Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res.* **22**, 292–298 (2012).
5. Castellarin, M. *et al.* Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299–306 (2012).
6. Viljoen, K. S., Dakshinamurthy, A., Goldberg, P. & Blackburn, J. M. Quantitative Profiling of Colorectal Cancer-Associated Bacteria Reveals Associations between Fusobacterium spp., Enterotoxigenic Bacteroides fragilis (ETBF) and Clinicopathological Features of Colorectal Cancer. *PLOS ONE* **10**, e0119462 (2015).
7. Li, Y.-Y. *et al.* Association of *Fusobacterium nucleatum* infection with colorectal cancer in Chinese patients. *World J. Gastroenterol.* **22**, 3227 (2016).
8. Galeano Niño, J. L. *et al.* Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature* **611**, 810–817 (2022).
9. Bullman, S. *et al.* Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. *Science* **358**, 1443–1448 (2017).
10. LaCourse, K. D. *et al.* The cancer chemotherapeutic 5-fluorouracil is a potent Fusobacterium nucleatum inhibitor and its activity is modified by intratumoral microbiota. *Cell Rep.* **41**, 111625 (2022).
11. Mima, K. *et al.* Fusobacterium nucleatum in Colorectal Carcinoma Tissue According to Tumor Location: *Clin. Transl. Gastroenterol.* **7**, e200 (2016).

12. Mima, K. *et al. Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. *Gut* **65**, 1973–1980 (2016).
13. Eid, J. *et al.* Real-Time DNA Sequencing from Single Polymerase Molecules. **323**, 7 (2009).
14. Kim, M., Oh, H.-S., Park, S.-C. & Chun, J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **64**, 346–351 (2014).
15. Gardner, S. N., Slezak, T. & Hall, B. G. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome: Table 1. *Bioinformatics* **31**, 2877–2878 (2015).
16. Brennan, C. A. & Garrett, W. S. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. *Nat. Rev. Microbiol.* **17**, 156–166 (2019).
17. Van Assche, P. F. & Wilssens, A. T. Fusobacterium perfoetens (Tissier) Moore and Holdeman 1973: Description and Proposed Neotype Strain. *Int. J. Syst. Bacteriol.* **27**, 1–5 (1977).
18. Manson McGuire, A. *et al.* Evolution of Invasion in a Diverse Set of Fusobacterium Species. *mBio* **5**, e01864-14 (2014).
19. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).
20. Aramaki, T. *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
21. Robrish, S. A., Oliver, C. & Thompson, J. Amino acid-dependent transport of sugars by Fusobacterium nucleatum ATCC 10953. *J. Bacteriol.* **169**, 3891–3897 (1987).

22. De Witte, C. *et al.* Detection, isolation and characterization of Fusobacterium gastrosuis sp. nov. colonizing the stomach of pigs. *Syst. Appl. Microbiol.* **40**, 42–50 (2017).

23. Park, S.-N. *et al.* Fusobacterium pseudoperiodonticum sp. nov., Isolated from the Human Oral Cavity. *Curr. Microbiol.* **76**, 659–665 (2019).

24. Pascal Andreu, V. *et al.* gutSMASH predicts specialized primary metabolic pathways from the human gut microbiota. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01675-1.

25. Jiang, Y. *et al.* Gallic Acid: A Potential Anti-Cancer Agent. *Chin. J. Integr. Med.* **28**, 661–671 (2022).

26. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **51**, D690–D699 (2023).

27. Bachrach, G., Rosen, G., Bellalou, M., Naor, R. & Sela, M. N. Identification of a Fusobacterium nucleatum 65 kDa serine protease. *Oral Microbiol. Immunol.* **19**, 155–159 (2004).

28. Ogawa, A. T., Brasil de Souza, T. de A. C., de Uzeda, M., Jankevicius, J. V. & Jankevicius, S. I. Characterization of Proteolytic Activities of Fusobacterium nucleatum. *J. Endod.* **32**, 521–523 (2006).

29. Doron, L. *et al.* Identification and Characterization of Fusolisin, the Fusobacterium nucleatum Autotransporter Serine Protease. *PLoS ONE* **9**, e111329 (2014).

30. Coppenhagen-Glazer, S. *et al.* Fap2 of Fusobacterium nucleatum Is a Galactose-Inhibitable Adhesin Involved in Coaggregation, Cell Adhesion, and Preterm Birth. *Infect. Immun.* **83**, 1104–1113 (2015).

31. Kaplan, C. W. *et al. Fusobacterium nucleatum* Apoptosis-inducing Outer Membrane Protein. *J. Dent. Res.* **84**, 700–704 (2005).

32. Kaplan, C. W., Lux, R., Haake, S. K. & Shi, W. The *Fusobacterium nucleatum* outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm. *Mol. Microbiol.* **71**, 35–47 (2009).

33. Lima, B. P., Shi, W. & Lux, R. Identification and characterization of a novel *Fusobacterium nucleatum* adhesin involved in physical interaction and biofilm formation with *Streptococcus gordonii*. *MicrobiologyOpen* **6**, e00444 (2017).

34. Han, Y. W. *et al.* Identification and Characterization of a Novel Adhesin Unique to Oral Fusobacteria. *J. Bacteriol.* **187**, 5330–5340 (2005).

35. Umaña, A. *et al.* Utilizing Whole *Fusobacterium* Genomes To Identify, Correct, and Characterize Potential Virulence Protein Families. *J. Bacteriol.* **201**, e00273-19, /jb/201/23/JB.00273-19.atom (2019).

36. Liu, B., Zheng, D., Zhou, S., Chen, L. & Yang, J. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res.* **50**, D912–D917 (2022).

37. Chen, L.-C. *et al.* Alteration of Gene Expression in Normal-Appearing Colon Mucosa of ***APC min*** Mice and Human Cancer Patients. *Cancer Res.* **64**, 3694–3700 (2004).

38. Rubie, C. Correlation of IL-8 with induction, progression and metastatic potential of colorectal cancer. *World J. Gastroenterol.* **13**, 4996 (2007).

39. Lee, Y. S. *et al.* Interleukin-8 and its receptor CXCR2 in the tumour microenvironment promote colon cancer growth, progression and metastasis. *Br. J. Cancer* **106**, 1833–1841 (2012).

40. Fusobacterium nucleatum host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. *Sci. Signal.* (2020).
41. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
42. Lagier, J.-C. *et al.* The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clin. Microbiol. Rev.* **28**, 237–264 (2015).
43. Richardson, M. *et al.* Analysis of 16S rRNA genes reveals reduced Fusobacterial community diversity when translocating from saliva to GI sites. *Gut Microbes* **12**, 1814120 (2020).
44. Overbeek, R. *et al.* The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **42**, D206–D214 (2014).
45. Aziz, R. K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75 (2008).
46. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
47. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
48. Yoon, S.-H., Ha, S., Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**, 1281–1286 (2017).
49. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
50. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

***Data Availability***

The whole genome sequence of *F. sphaericum sp. nov.* strain SB021 has been deposited in GenBank under the accession number SAMN15202580.

***Acknowledgements***

We would like to thank the Cellular Imaging Shared Resource of the Fred Hutchinson Cancer Center/University of Washington Cancer Consortium (P30 CA015704). Scientific Computing Infrastructure at Fred Hutchinson Cancer Center was funded by ORIP grant S10OD028685.

***Funding***

Research reported in this publication was supported by the National Institute of Dental and Craniofacial Research of the National Institutes of Health under award number R01 DE027850 (to C.D.J.), the National Cancer Institute under award number R00 CA229984-03 (to S.B.), start-up funds provided by the FHCC (to S.B and C.D.J), a study scholarship from the Konrad Adenauer Foundation (Y.E.), and the Washington Research Foundation Postdoctoral Fellowship (to M.Z.R).

***Contributions***

Y.E., H.H., C.D.J. and S.B. designed the study. M.Z.R, Y.E., C.D.J., and S.B. wrote the paper. A.B. and K.D.L. processed patient tissue specimen. A.B. and K.D.L. carried out microbial isolation and culture. M.Z.R., D.S.J. and C.D.J. performed DNA extraction, genome sequencing, and assembly. M.Z.R., Y.E., A.B., H.W., and S.S.M. performed computational and statistical analysis. Y.E., A.B., and K.D.L. carried out *in vivo* and *in vitro* assays. S.B. and C.D.J. obtained

funding and supervised computational and wet lab experiments. All authors read and provided edits to the paper and contributed to the final version.

***Competing Interests***

S.B. has consulted for GlaxoSmithKline and BiomX. C.D.J. has consulted for Series Therapeutics and Azitra. S.B. is an inventor on US patent application no. PCT/US2018/042966, submitted by the Broad Institute and Dana-Farber Cancer Institute, which covers the targeting of *Fusobacterium* for the treatment of colorectal cancer. K.D.L. is currently employed by NanoString Technologies. The remaining authors declare no competing interests.