MECHANISM OF INTESTINAL COLONIZATION BY SYMBIOTIC BACTERIA

Thesis by

Sung-Eun Melanie Lee

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2013

(Defended September 17, 2012)

ACKNOWLEDGEMENTS

I was very green and naïve when I first came to Caltech as an MD/PhD candidate looking for a place to call home and pursue a PhD degree. I was a young and inexperienced student eager to learn but not quite sure how to go about doing science and I have come a long way since then. This thesis would not be possible without the help, guidance, support, and inspiration from far too many people. First and foremost, I would like to thank my advisor, Sarkis Mazmanian. Over the years, he taught me everything I know about being an independent scientist: His mentoring began with teaching me the most basic bench work techniques (I learned how to streak bacteria and pour plates from him), then for the next few years, he tirelessly challenged me to pursue my project with rigorous thinking, bold and innovative approaches, and creativity. I really appreciate his dedication to teaching. One-on-one meetings with his postdocs and students always have been his top priority and for me, it was always a source of inspiration and direction. Sarkis never missed a chance to mentor me in various aspects: I learned a great deal of grant writing, manuscript writing, and politics in science by assisting him with his R21 and R01 writing, as well as various manuscripts from our lab or reviewing manuscripts from other labs. I can probably recount at least ten anecdotes from Sarkis' graduate and postdoc years which he meant as a lesson to impart and he reminds me of how my father would try to teach me lessons. I have also learned quite a bit about management and mentoring other people as he entrusted me with three SURF students and two volunteer students over the years. More than anything, I am grateful to his financial and moral support, encouragement, and affirmation throughout my thesis project, which has

definitely been a rocky road and would not have been possible without his belief in me and the project.

There are many people to thank from the Mazmanian lab. Sara McBride, who wears as many hats as one person possibly can, really has been an important lab-mom figure whose passion and dedication for the lab made a real difference in our work environment. I met Greg Donaldson closer to the end of my PhD, a young but precocious and experienced Biology student who shares the same passion for the symbiotic bacteria and how they have co-evolved with the human host. Every discussion we had left me with something to think about, and towards the end I found myself learning from him about my project. I can't think of a better person to carry out the project after my departure. Yue Shen, Janet Chow, and Arya Khosravi were my fellow founding members of the Mazmanian lab and I couldn't have asked for better friends and colleagues to go through this journey together. Gil Sharon and Hiutung Chu, new postdocs in the lab, really helped me through my last stage of PhD with their eager support, advice, and encouragement as they went through the same process in the recent past. Thank you for rescuing me from the unexpected pregnant mouse culling incident and your truly helpful feedback on my writing.

I want to thank all of my collaborators. I worked with Harry Choi for applying HCR technology to imaging bacteria. He was a meticulous and hard-working person from whom I learned a great deal. Igor Antoshechkin, JJ Kang, Adam Rosenthal, and Jingli Zhang have all helped me with performing RNAseq experiments and analysis. Having no experience with programming and bioinformatics analysis, I owe a great deal to them for performing my RNAseq experiment. Zbigniew Mikulski, whom I never met

in person has been a fantastic collaborator, and he came through when I needed someone with an imaging expertise the most. Even though not all projects came to fruition, I am deeply indebted to everyone I have encountered throughout my project for their generosity with time and knowledge. Working with these people have been some of my favorite memories from the PhD experience.

Lastly, I want to thank family members and my husband Sotiris Masmanidis. My parents and my parents-in-law have always given me constant love, encouragement and prayer. They are my rock. My proudest accomplishment during PhD, however, is not any paper I published or getting a screen to work – it is meeting my husband Sotiris. And he deserves a huge acknowledgement for this thesis even though he was never directly involved. I am not sure how he did it but ever since I met Sotiris, my research project blossomed, I became much more efficient and focused and I believe Sarkis can even vouch for it. His undivided love and support gave me confidence and his exemplary work ethic and discipline were definitely positive influences on me. Thank you for the gift of beautiful marriage, stability, support, and opening a new chapter in my life as a wife and a mother.

ABSTRACT

All animals live in symbiosis. Shaped by eons of co-evolution, host-bacterial associations have developed into prosperous relationships creating mechanisms for mutual benefits to both microbe and host. No better example exists in biology than the astounding numbers of bacteria harbored by the lower gastrointestinal tract of mammals. This community of symbionts establishes a life-long habitat within the distal gut and profoundly impacts host health. Although many recent investigations have led to determination of the microbiota composition, molecular mechanisms mediating establishment and maintenance of the microbial community within the gut is poorly described.

We use gnotobiotic mice to elucidate mechanisms of colonization by *Bacteroides*, one of the most numerically prominent genera in the intestine. We generate mutant strains of *Bacteroides fragilis* that lack the ability to express multiple capsular polysaccharides and demonstrate defect in colonization in competition with wild-type strain, suggesting a role for surface sugar architecture during host-symbiont mutualism. Through a functional *in vivo* genetic screen of colonization, we identify a novel operon from the genome of *B. fragilis* that is highly conserved among many sequenced intestinal *Bacteroides* and that mediate a species-specific colonization profile. We have named this genetic locus the commensal colonization factor (*ccf*). *B. fragilis* deleted in the *ccf* genes exhibit colonization defects in both germ-free and complex microbiota harboring mice. The *ccf* genes of *B. fragilis* are up-regulated during gut colonization, preferentially at the mucosal surface, supporting an *in vivo* function. Indeed, deletion of *ccf* genes leads to reduced mucosal association and a defect in bacterial occupation of colonic crypts of

mice. The ability of *B. fragilis* to repopulate the gut after antibiotic perturbation or gastroenteritis requires expression of *ccf*, suggesting the niche within colonic crypts represents a colonization reservoir for the gut microbiota following environmental stress. These findings suggest a novel and evolutionarily conserved mechanism for persistent gut colonization by the *Bacteroides* species.

TABLE OF CONTENTS

Chapter 1:

Introduction

1. The human gut microbiota

Immediately upon a sterile birth, mammals are coordinately colonized with complex communities of bacteria on all environmentally exposed surfaces (Ley et al., 2006). After weaning, a compositionally stable and numerically vast microbial ecosystem is formed in the gastrointestinal tract (Palmer et al., 2007). During adulthood, the human gut harbors up to 100 trillion (10^{14}) microbes. This is equivalent to 10^{11}-10^{12} cells/g colonic content with a biomass of >1 kg, which is the highest density recorded for any microbial habitat (O'Hara et al., 2006; Whitman et al., 1998). Our knowledge of the composition of the human gut microbiota was limited to culture-based studies until recent initiatives utilizing high-throughput sequencing of bacterial 16S rRNA genes made available a detailed inventories of the normal human gut microbiome comparing individuals and populations across different age groups and geographic regions (The Human Microbiome Project Consortium, 2012; Yatsunenko et al., 2012; Eckburg et al., 2005; Zoetendal et al., 2002; Xu et al., 2007). Such comprehensive sequence-based microbiome profiling studies not only increased our understanding of the population composition, dynamics and ecology of the gut microbiota but gave us evolutionary insight into the host-microbe mutualism. Although there are at least 55 divisions of Bacteria and 13 different divisions of Archaea on Earth, the human distal gut microbial community is dominated by members of just two bacterial divisions, *Bacteroidetes* and *Firmicutes*, and one member of Archaea, *Methanobrevibacter smithii* (Eckburg et al., 2005). In contrast to the low levels of deep diversity, the microbial diversity in mammalian gut exhibits high levels of variation at the level of species and strains. Such extreme fan-like phylogenetic architecture is unique to the mammalian gut microbiota

and may reflect a long history of co-evolution between microbes and their hosts whereby a few early successful colonizers have maintained an exclusive membership within the gut and have undergone diversification into species and strains (Backhed et al., 2005; Dethlefsen et al., 2007; Ley et al., 2006).

The gut ecosystem is dynamic. Within a given intestinal habitat, some microbial members may be bona fide "residents" (autochthonous components) that have established commensal status in the gut and persist throughout the lifetime of the host. However, detection of a particular bacterial species from the fecal bacterial community does not necessarily mean that the bacteria are permanent inhabitants of the gut ecosystem. Some microbes found in the gut can be described more accurately as "transient" components (allochthonous members) of the gut microbial communities which originate from ingested food, water and various components of our environment. The difference between the mucosa-associated bacterial communities in the colon and the luminal and fecal bacterial communities can be attributed in part to the different nature of the microbial inhabitance found in different niches within the gut at a given time (Zoetendal et al., 2002). In order to fully understand the human biology with respect to the human microbiota, it is crucial to determine which compartment of the microbiota is genuinely autochthonous to the human gut and discover the molecular traits that enable the resident bacterial species to establish and maintain colonization in a highly variable and competitive environment.

2. Study of host-bacterial mutualism using gnotobiology

Gut microbiota sequencing efforts led us to obtain insights into the ecological and evolutionary pressures shaping the community structure. However, snap-shot profiling approach alone cannot address questions regarding specific requirements for a given microbial species to be recognized by the host as a commensal organism and be tolerated. One of the most successful experimental approaches to dissect the intricate crosstalk between host and commensal bacteria can be found in the use of gnotobiotic animal models. Traditionally performed with rodents, the presence and composition of the microbiota is experimentally manipulated using germ-free host animals. Comparative studies of germ-free and conventionally colonized animals have demonstrated the profound impact of the gut microbiota on the host biology ranging from the mucosal and systemic immunity and morphological integrity of the intestinal epithelial cells to the metabolic, absorptive and endocrine functions (Round and Mazmanian, 2009; Smith et al., 2007). Experiments comparing the gene expression profiles of the intestinal epithelial layer between colonized and germ-free animals have shown direct host response to intestinal bacteria at the molecular level (Backhed et al., 2004; Hooper et al., 2003; Hooper et al., 2001; Salzman et al., 2007; Stappenbeck et al., 2002). Germ-free animals mono-associated with a single microbial species or purified microbial compounds provided a powerful tool for dissecting which microbial signals are required for mutualistic interaction with the host (Bry et al., 1996; Hooper et al., 1999; Mazmanian et al., 2005; Freitas et al., 2005).

One of the pioneering works by McFall-Ngai and colleagues discovered a “mutualism factor” from an elegant model of host-bacterial symbiosis between the

marine bacterium *Vibrio fischeri* and its squid host *Euprymna scolopes* (Nyholm and McFall-Ngai, 2003; Nyholm and McFall-Ngai, 2004). In this partnership, the squid provides a nutrient-rich environment for its luminescent symbiont in a specialized light organ. Upon bacterial colonization, the squid undergoes a dramatic morphogenesis which is associated with symbiosis. Interestingly, much of the host morphogenetic program can be induced by conserved components of bacterial cell walls, lipopolysaccharide (LPS) and peptidoglycan (PGN) (Koropatnick et al., 2004). This is an example of host-microbial mutualism driven by microbial factors which are none other than ligands of innate immune system called pattern recognition receptors (PRRs), also known as microbial-associated molecular patterns (MAMPs). Although MAMPs are commonly linked with pathogenesis in animals, it appears that microbes can use similar molecular signals to mediate beneficial interaction with the host in a context-dependent manner (Cheesman and Guillemin, 2007). In summary, gnotobiotic experimental strategies have been instrumental for identifying some of the key molecular mechanisms underlying mutualism between host and microbes.

3. The *Bacteroides* as important members of the microbiota

Bacteroidetes comprises one of the most numerically abundant Gram-negative phyla in the mammalian gastrointestinal tract GI (Palmer et al., 2007; Ley et al., 2008). One of the first evidences of gut microbiota playing an active role in the host biology was demonstrated by mono-association of germ-free mice with *Bacteroides thetaiotaomicron* which then induced carbohydrate decorations of the intestinal epithelium to mediate the normal architectural development of host tissue (Bry et al., 1996). DNA microarray

profiling of the global host transcriptional responses in germ-free mice mono-associated with *B. thetaiotaomicron* revealed a number of genes involved with regulation of postnatal maturation (adenosine deaminase), nutrient uptake and metabolism (SGLT-1, colipase, L-FABP), processing of angiogenesis (angiogenin-3) and fortification of the host epithelial barrier (DAF, CRP-ductin, Sprr2a) (Hooper et al., 2001). These studies demonstrated a remarkable example of a single species of commensal organism restoring many of the structural, metabolic and developmental defects of a previously germ-free host. *B. thetaiotaomicron* has also been shown to induce the production of anti-microbial immune defenses during colonization, to protect their host from invading pathogens (Cash et al., 2006). *Bacteroides fragilis* is a gut commensal that aids in host health by directing the cellular and physical maturation of the developing immune system of the host (Mazmanian et al., 2005) and by protecting animals from inflammatory disease such as colitis and multiple sclerosis (Mazmanian et al., 2008; Lee et al., 2011; Ochoa-Reparaz et al., 2010). *B. fragilis* produces an immunomodulatory molecule that can reprogram the host immune system to produce anti-inflammatory molecules required for protection from disease in animals (Round and Mazmanian, 2010). In summary, many *Bacteroides* species have been well documented for their beneficial interactions with the host and emerged as model organisms for the study of host-bacterial mutualism.

4. Mucosal surface colonization by commensal bacteria

Host-microbe interactions often begin with colonization of the mucosal surfaces. Understanding the host-microbe interaction at the mucosal surface is fundamental to uncovering colonization mechanism of commensal bacteria in the gut. However, studying

the interactions between host and bacteria through the thickness of the mucus layer *in vivo* is a technically formidable task. Our current understanding of how certain bacteria interact with mucus layer is mainly limited to *in vitro* bacterial mucin-binding studies which may not recapitulate inside the host intestinal tract.

The mucus gel layer of the large intestine is a dense matrix of polysaccharides (and proteins) derived mainly from the epithelium's goblet cell lineage. Its thickness and mucin composition vary along the length of the gut (Matsuo et al., 1997). Traditionally, the mucus gel layer is considered a buffer between the highly immunogenic luminal content (commensals and pathogens alike) and the host epithelial layer to protect both the host and the gut bacteria (Deplancke and Gaskins, 2001). On the contrary, the mucus layer can represent a habitat and source of nutrients for the bacterial communities that colonize mucosal surfaces (Sonnenburg et al., 2004). The ability of mucus to support the growth of bacteria is evident from numerous *in vitro* studies in which bacteria have been shown to grow readily in mucus preparations (Jonsson et al., 2001; McCormick et al., 1988). Using streptomycin-treated mouse model, several mucus-derived sugars were described as major carbon sources required for *E. coli* colonization of the gut (Chang et al., 2004). Initially, a whole-genome transcriptional profiling of *E. coli* strain MG1655 during growth *ex vivo* on cecal mucus was conducted. Several nutritional genes corresponding to catabolic pathways for nutrients found in mucus were induced. Each pathway was systematically knocked out and the mutants were tested for fitness in mouse intestine colonization. Competitive colonization between wild-type MG1655 and isogenic mutants that lack the ability to catabolize various nutrients confirmed that carbohydrate catabolism plays a dominant role in the initiation and maintenance of *E. coli*

colonization of the mouse intestine (Chang et al., 2004). Nutrient availability within the colon mucus layer creates an attractive ecological niche for bacteria and thus provides at least one likely mechanism of gut colonization by commensal bacteria.

Analysis of carbohydrate structures along the length of the gastrointestinal tract of two humans showed that although their mucin-associated glycans were diverse, their region-specific glycosylation patterns were well conserved (Robbe et al., 2003). These glycoproteins and mucoproteins on the mucosal surface of the host gut can serve as receptor sites for attachment and adherence by commensal bacteria (Baranov and Hammarstrom, 2004; Granato et al., 2004). High diversity among the glycans with conserved spatial patterns found on the gut mucosal surface strongly suggests a mechanism of host-driven (maybe as a result of bacteria modulating the host) regulation of gut microbial community composition by directing members of the microbiota to distinct host niches. Host-derived glycan substrates that serve as nutrient source or docking sites for specific microbes may describe a mechanism for establishment and maintenance of a spatially diversified gut microbiota.

5. The role of bacterial surface structures during colonization

The microbial envelope represents a functional organelle decorated with molecules that provide critical interactions between microorganisms and their environment. In addition to proteins, bacterial surfaces are covered with abundant amounts of lipopolysaccharides, exopolysaccharides and capsular polysaccharides. Capsule production has been shown to be required for bacterial virulence by numerous pathogens, and polysaccharides are the key components of many vaccines developed to

prevent bacterial infections (Johri et al., 2006; Lee et al., 2001). Conversely, the biologic significance of capsular polysaccharide production during beneficial host-bacterial commensalism has only recently been investigated (Comstock and Kasper, 2006). Unlike pathogens that usually possess few capsular polysaccharides (0–2 per strain), analysis of the genomes of all sequenced intestinal *Bacteroides* species reveal a unique feature: the presence of numerous genomic loci that encode for capsular polysaccharides (8–13 per strain) (Coyne and Comstock, 2008). It is believed that the multiple capsular polysaccharides (and perhaps other surface structures) of *Bacteroides* create systems for altering the physical properties of bacterial surfaces (Xu et al., 2007). Several reports have predicted that the expression of multiple capsular polysaccharides by *B. fragilis* provide functions that are critical for host-bacterial symbiosis (Krinos et al., 2001; Kuwahara et al., 2004; Cerdeno-Tarraga et al., 2005). It appears that *Bacteroides* have invested heavily in the development of dynamic molecular mechanisms to establish residence within the mammalian ecosystem.

There are many examples of adherence mechanisms employed by bacterial pathogens during infection; including molecules such as pili, fimbriae, surface proteins, teichoic acids, lipopolysaccharides and other carbohydrates (Pizarro-Cerda and Cossart, 2006). Unlike pathogens which generally cause acute infections, symbiotic bacteria have evolved to associate with their hosts for prolonged periods of time. However, mechanisms involved in sustained colonization by beneficial bacteria remain largely undescribed. The following chapters in this thesis describe studies seeking to identify and characterize the ‘colonization factors’ of the important human symbiont, *B. fragilis*. Though very speculative, knowing how beneficial bacteria maintain long-term

association with their mammalian hosts may provide means to ‘engineer’ delivery of health-promoting bacteria as therapeutics for various human diseases.

REFERENCES

Backhed, F., Ding, H., Wang, T., Hooper, L.V., Koh, G.Y., Nagy, A., Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A *101,*15718–15723.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. Science *307, 1915*–1920.

Baranov, V., and Hammarstrom, S. (2004). Carcinoembryonic antigen (CEA) and CEA-related cell adhesion molecule 1 (CEACAM1), apically expressed on human colonic M cells, are potential receptors for microbial adhesion. Histochem Cell Biol *121*, 83–89.

Bry, L., Falk, P.G., Midtvedt, T., and Gordon, J.I. (1996). A model of host-microbial interactions in an open mammalian ecosystem. Science *273*, 1380–1383.

Cash, H.L., Whitham, C.V., Behrendt, C L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science *313*, 1126–1130.

Cerdeno-Tarraga, A.M., Patrick, S., Crossman, L.C., Blakely, G., Abratt, V., Lennard, N., Poxton, I., Duerden, B., Harris, B., Quail, M.A., Barron, A., Clark, L., Corton, C., Doggett, J., Holden, M.T., Larke, N., Line, A., Lord, A., Norbertczak, H., Ormond, D., Price, C., Rabbinowitsch, E., Woodward, J., Barrell, B., and Parkhill, J. (2005). Extensive DNA inversions in the B. fragilis genome control variable gene expression. Science *307*, 1463–1465.

Chang, D.E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., Anderson, A.B., Grissom, J.E., Laux, D.C., Cohen, P.S., and Conway, T. (2004). Carbon nutrition of Escherichia coli in the mouse intestine. Proc Natl Acad Sci U S A *101*, 7427–7432.

Cheesman, S.E., and Guillemin, K. (2007). We know you are in there: conversing with the indigenous gut microbiota. Res Microbiol *158*, 2–9.

Comstock, L.E., and Kasper, D.L. (2006). Bacterial glycans: key mediators of diverse host immune responses. Cell *126*, 847–850.

Coyne, M.J., and Comstock, L.E. (2008). Niche-specific features of the intestinal bacteroidales. J Bacteriol *190*, 736–742.

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.

Dethlefsen, L., McFall-Ngai, M., and Relman, D.A. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature *449*, 811–818.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635–1638.

Freitas, M., Axelsson, L.G., Cayuela, C., Midtvedt, T., and Trugnan, G. (2005). Indigenous microbes and their soluble factors differentially modulate intestinal glycosylation steps in vivo. Use of a "lectin assay" to survey in vivo glycosylation changes. Histochem Cell Biol *124*, 423–33.

Granato, D., Bergonzelli, G.E., Pridmore, R.D., Marvin, L., Rouvet, M., and Corthesy-Theulaz, I.E. (2004). Cell surface-associated elongation factor Tu mediates the attachment of Lactobacillus johnsonii NCC533 (La1) to human intestinal cells and mucins. Infect Immun *72*, 2160–2169.

Hooper, L.V., Stappenbeck, T.S., Hong, C.V., and Gordon, J.I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. Nat Immunol *4*, 269–273.

Hooper, L.V., Wong, M.H., Thelin, A., Hansson, L., Falk, P.G., and Gordon J.I. (2001). Molecular analysis of commensal host-microbial relationships in the intestine. Science *291*, 881–884.

Hooper, L.V., Xu, J., Falk, P.G., Midtvedt, T., and Gordon, J.I. (1999). A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc Natl Acad Sci U S A *96*, 9833–9838.

Johri, A.K., Paoletti, L.C., Glaser, P., Dua, M., Sharma, P.K., Grandi, G., and Rappuoli, R. (2006). Group B Streptococcus: global incidence and vaccine development. Nat Rev Microbiol *4*, 932–942.

Jonsson, H., Strom, E., and Roos, S. (2001). Addition of mucin to the growth medium triggers mucus-binding activity in different strains of Lactobacillus reuteri in vitro. FEMS Microbiol Lett *204*, 19–22.

Koropatnick, T.A., Engle, J.T., Apicella, M.A., Stabb, E.V., Goldman, W.E., and McFall-Ngai, M.J. (2004). Microbial factor-mediated development in a host-bacterial mutualism. Science *306*, 1186–1188.

Krinos, C.M., Coyne, M.J., Weinacht, K.G., Tzianabos, A.O., Kasper D.L., and Comstock, L. E. (2001). Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature *414*, 555–558.

Kuwahara, T., Yamashita, A., Hirakawa, H., Nakayama, H., Toh, H., Okada, N., Kuhara, S., Hattori, M., Hayashi, T., and Ohnishi, Y. (2004). Genomic analysis of Bacteroides

fragilis reveals extensive DNA inversions regulating cell surface adaptation. Proc Natl Acad Sci U S A *101*, 14919–14924.

Lee, C.J., Lee, L.H., Lu, C.S., and Wu, A. (2001). Bacterial polysaccharides as vaccines--immunity and chemical characterization. Adv Exp Med Biol *491*, 453–471.

Ley, R.E., Peterson, D.A., and Gordon J.I. (2006). Ecological and evolutional forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., and Gordon, J.I. (2008). Evolution of mammals and their gut microbes. Science *320*, 1647–1651.

Lee, Y.K., Menezes, J.S., Umesaki, K., and Mazmanian, S.K. (2011). Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. Proc Natl Acad Sci U S A *108 Suppl 1*, 4615–4622.

Matsuo, K., Ota, H., Akamatsu, T., Suglyama, A., and Katsuyama, T. (1997). Histochemistry of the surface mucous gel layer of the human colon. Gut *40*, 782–789.

Mazmanian, S.K., Liu, C.H., 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.

Mazmanian, S.K., Round, J.L., and Kasper, D.L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. Nature *453*, 620–625.

McCormick, B.A., Stocker, B.A., Laux, D.C., and Cohen, P.S. (1988). Roles of motility, chemotaxis, and penetration through and growth in intestinal mucus in the ability of an avirulent strain of Salmonella typhimurium to colonize the large intestine of streptomycin-treated mice. Infect Immun, 1988. *56*, 2209–2217.

Nyholm, S.V., and McFall-Ngai, M.J. (2003). Dominance of Vibrio fischeri in secreted mucus outside the light organ of Euprymna scolopes: the first site of symbiont specificity. Appl Environ Microbiol *69*, 3932–3937.

Nyholm, S.V., and McFall-Ngai, M.J. (2004). The winnowing: establishing the squid-vibrio symbiosis. Nat Rev Microbiol *2*, 632–642.

Ochoa-Reparaz, J., Mielcarz, D.W., Ditrio, L.E., Burroughs, A.R., Begum-Haque, S., Dasgupta, S., Kasper, D.L., and Kasper, L.H. (2010). Central nervous system demyelinating disease protection by the human commensal Bacteroides fragilis depends on polysaccharide A expression. J Immunol *185*, 4101–4108.

O'Hara, A.M., and Shanahan, F. (2006). The gut as a forgotten organ. EMBO Rep *7*, 688–693.

Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the human infant intestinal microbiota. PLoS Biol *5*, e177.

Pizarro-Cerda, J., and Cossart, P. (2006). Bacterial adhesion and entry into host cells. Cell *124*, 715–727.

Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.P., and Michalski, J.C. (2003). Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the intestinal tract. J Biol Chem *278*, 46337–46348.

Round, J.L., and Mazmanian, S.K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol *9*, 313–323.

Round, J.L., and Mazmanian, S.K. (2010). Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A *107*, 12204–12209.

Salzman, N.H., Underwood, M.A., and Bevins, C.L. (2007). Paneth cells, defensins, and the commensal microbiota: a hypothesis on intimate interplay at the intestinal mucosa. Semin Immunol *19*, 70–83.

Smith, K., McCoy, K.D., and Macpherson, A.J. (2007). Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. Semin Immunol *19*, 59–69.

Sonnenburg, J.L., Angenent, L.T., and Gordon, J.I. (2004). Getting a grip on things: how do communities of bacterial symbionts become established in our intestine? Nat Immunol *5*, 569–573.

Stappenbeck, T.S., Hooper, L.V., and Gordon, J.I. (2002). Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. Proc Natl Acad Sci U S A *99*, 15451–15455.

The Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998). Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A *95*, 6578–6583.

Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho, P.M., Minx, P., Latreille, P., Cordum, H., Van Brunt, A., Kim, K., Fulton, R.S., Fulton, L.A., Clifton, S.W., Wilson, R.K., Knight, R.D., and Gordon, J.I. (2007). Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol *5*, e156.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R., and Gordon, J.I. (2012). Human gut microbiome viewed across age and geography. Nature *486*, 222–227.

Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D., and de Vos, W.M. (2002). Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl Environ Microbiol *68*, 3401–3407.

Chapter 2:

Regulation of surface architecture by symbiotic bacteria mediates host colonization

Cui Hua Liu*, S. Melanie Lee*, Jordan M. VanLare, Dennis L. Kasper, and Sarkis K. Mazmanian

*These authors contributed equally to the work.

This chapter has been published in the *Proceedings of the National Academy of Sciences of the U S A* (2008) 105: 3951–3956.

ABSTRACT

Microbes occupy countless ecological niches in nature. Sometimes these environments may be on or within another organism, as is the case in both microbial infections and symbiosis of mammals. Unlike pathogens that establish opportunistic infections, hundreds of human commensal bacterial species establish a life-long cohabitation with their hosts. Although many virulence factors of infectious bacteria have been described, the molecular mechanisms used during beneficial host–symbiont colonization remain almost entirely unknown. The novel identification of multiple surface polysaccharides in the important human symbiont *Bacteroides fragilis* raised the critical question of how these molecules contribute to commensalism. To understand the function of the bacterial capsule during symbiotic colonization of mammals, we generated *B. fragilis* strains deleted in the global regulator of polysaccharide expression and isolated mutants with defects in capsule expression. Surprisingly, attempts to completely eliminate capsule production are not tolerated by the microorganism, which displays growth deficits and subsequent reversion to express capsular polysaccharides. We identify an alternative pathway by which *B. fragilis* is able to reestablish capsule production and modulate expression of surface structures. Most importantly, mutants expressing single, defined surface polysaccharides are defective for intestinal colonization compared with bacteria expressing a complete polysaccharide repertoire. Restoring the expression of multiple capsular polysaccharides rescues the inability of mutants to compete for commensalism. These findings suggest a model whereby display of multiple capsular polysaccharides provides essential functions for bacterial colonization during host–symbiont mutualism.

INTRODUCTION

We live in a microbial world. Immediately upon birth, humans coordinately assemble a complex bacterial microbiota on almost all environmentally exposed surfaces (Ley et al., 2006). Although it has been appreciated for decades that humans harbor multitudes of commensal bacteria, recent studies have begun to reveal the extraordinary diversity and complexity of the ecosystem we provide to microorganisms. Advances in genomic technologies have demonstrated that we harbor dozens of bacterial species in our stomachs, hundreds on our skin and oral cavity, and thousands within our lower gastrointestinal tract (Bik et al., 2006; Gao et al., 2007; Gill et al., 2006). The magnitude of these interactions and the evolutionary forces that drive them must exert profound influences on the biology of both microbe and man. The gastrointestinal tract provides an excellent example of the complex interactions between the microbiota and the host (Hooper and Gordon, 2001). Bacteria dominate this biological niche, both numerically and in terms of diversity. Of the multitudes of bacterial species that colonize the mammalian gastrointestinal tract ($> 10^{13}$ organisms from > 1,000 different species), those of the genus *Bacteroides* are among the most numerically prominent in humans (Eckburg et al., 2005). For decades, bacteria have been known to perform the essential function of metabolizing complex carbohydrates subsequently used by their mammalian hosts; *Bacteroides* species have been shown to be essential for this function (Hooper et al., 2002; Xu et al., 2003). Analysis of the genome sequences of the human *Bacteroides* (*Bacteroides thetaiotaomicron*, *Bacteroides vulgatus*, *Bacteroides distasonis*, and *Bacteroides fragilis*) reveals that this genus has evolved numerous glycosidases for carbohydrate degradation (Xu et al., 2007). *B. thetaiotaomicron* induces carbohydrate

decorations of the intestinal epithelium to mediate the normal architectural development of host tissue (Bry et al., 1996). Furthermore, *B. fragilis* was first described to produce multiple surface capsular polysaccharides (Krinos et al., 2001). It has been recently revealed that all studied *Bacteroides* contain numerous genomic loci for capsular polysaccharide production, a unique and distinguishing feature of this genus of bacteria. Studies show that various *Bacteroides* species share with their human host a mammalian-evolved biochemical pathway for the addition of sugar modifications to surface proteins and polysaccharides (Coyne et al., 2005). Moreover, we have recently demonstrated that *B. fragilis* elaborates an important immunomodulatory polysaccharide that instructs the normal development of the host immune system (Mazmanian et al., 2005). Thus, the *Bacteroides* have dedicated a significant proportion of their biology to the production and functions of capsular polysaccharides during coevolution with mammals.

Decades of research have assigned various functions to surface polysaccharides of pathogens, including biofilm production, tissue adherence, and antiphagocytic activity during immune evasion (Mazmanian and Kasper, 2006). Capsule production has been shown to be required for bacterial virulence in numerous animal models of disease, and polysaccharides are the key components of many vaccines developed to prevent pathogenic bacterial infections (Johri et al., 2006; Lee et al., 2001). Conversely, the biologic significance of capsular polysaccharide production during beneficial host–bacterial commensalism has only recently been suggested (Comstock and Kasper, 2006). It is believed that the multiple capsular polysaccharides (and perhaps other surface structures) of the *Bacteroides* create systems for altering the physical properties of bacterial surfaces (Xu et al., 2007). Several reports have predicted that the expression of

multiple capsular polysaccharides by *B. fragilis* provides functions that are critical for host–bacterial symbiosis (Krinos et al., 2001; Kuwahara et al., 2004; Cerdeno-Tarraga et al., 2005). However, this notion currently remains without experimental corroboration. In the report contained herein, we examine the role of capsular polysaccharide production during the relationship between *B. fragilis* and its mammalian host. By generating bacterial mutants in the regulation of capsule expression, we find that production of at least one capsular polysaccharide is required for the viability of the microorganism. Most importantly, inhibiting the ability of *B. fragilis* to modulate its surface architecture renders it unable to compete for colonization of the gastrointestinal tract of animals. It appears that *Bacteroides* have invested heavily in the development of dynamic molecular mechanisms to establish and maintain residence within the mammalian ecosystem. These results provide the foundation for understanding our evolutionary cohabitation with the microbial world around and within us.

RESULTS

Capsular polysaccharide expression is critical for *B. fragilis.*

The eight known capsular polysaccharide (CPS) biosynthesis loci of *B. fragilis* are scattered throughout the genome as distinct polycistronic operons of 11–22 genes (PSA–PSH) (Kuwahara et al., 2004; Coyne et al., 2003). Mpi (multiple promoter invertase), a member of the serine site-specific DNA recombinase family, acts as a global transcriptional regulator of polysaccharide expression through inverting promoters between the ‘‘on’’ and ‘‘off’’ orientations (Figure 2.1A) (Coyne et al., 2003). We introduced a genomic mutation in the *mpi* gene by homologous recombination in a PSA deletion background, because PSA is the most abundantly expressed polysaccharide [produced by 79% of all bacterial cells in culture (Krinos et al., 2001)]. We subsequently screened mutants by PCR using promoter-specific primers to isolate clones with all Mpi-regulated promoters in the ‘‘locked-off’’ orientation. Similar to previous findings (Coyne et al., 2003), cultures of wild-type *B. fragilis* are heterogeneous with regard to polysaccharide expression and are in phase variation displaying both ‘‘on’’ and ‘‘off’’ promoters (Figure 2.1B and Supplementary Table 2.1). Upon mutagenesis, we isolated one mutant [*B. fragilis* CPM1 (capsular polysaccharide mutant)] with seven of the eight characterized promoters in the ‘‘off’’ orientation at the time of *mpi* deletion (PSC is not controlled by Mpi). When *mpi* was introduced into CPM1 by transcomplementation, each of the Mpi-controlled polysaccharide promoters regained the ability to invert (Figure 2.1B). We phenotypically verified the presence or absence of CPS synthesis with immunoblot analysis using antisera specific for each molecule. A complete loss of polysaccharide expression from all Mpi-controlled loci was observed in strain CPM1

compared with wild-type *B. fragilis* (Figure 2.1C).

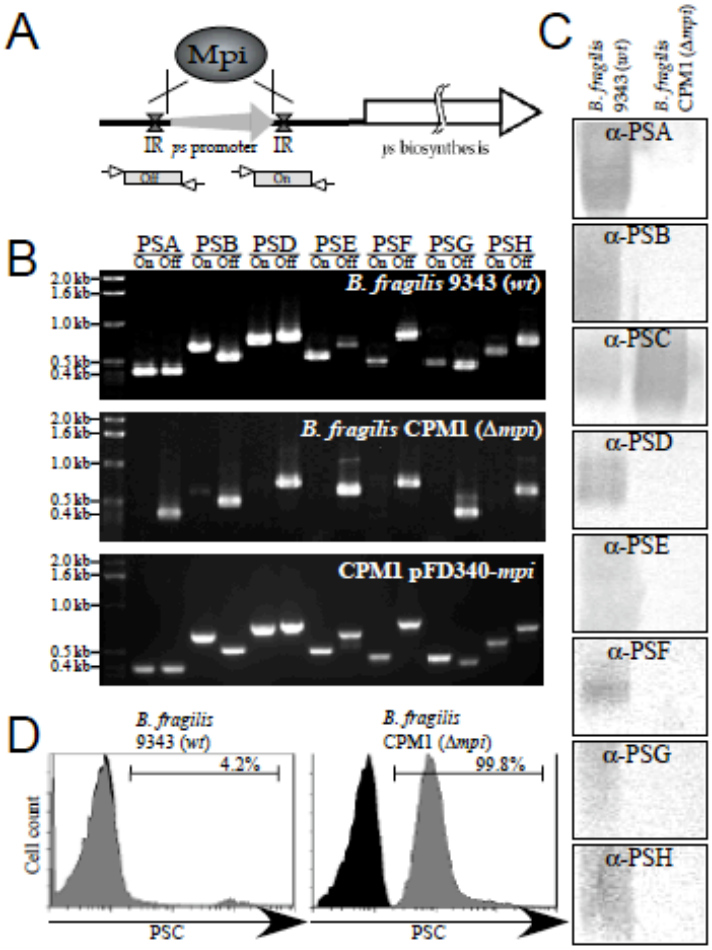


Figure 2.1 | Deletion of the global regulator of polysaccharide promoter inversion in *B. fragilis*. (A) Schematic representing promoters flanked by inverted repeats (IR) of polysaccharide biosynthesis loci. Mpi induces recombination at the IR sites to invert promoters. Boxes below represent PCR products distinguishing "on" and "off" orientations. (B) PCR products generated from chromosomal DNA of wild-type *B. fragilis* 9343, an isolated Δ*mpi* mutant (CPM1), and CPM1 strain complemented with the *mpi* gene (CPM1 pFD340-*mpi*). Unlike wild-type, all Mpi-controlled capsular polysaccharide (CPS) promoters are locked "off" in CPM1. Genetic rescue of *mpi* restores phase variation in CPM1. (C) Immunoblot analysis demonstrates that expression of PSC is not controlled by Mpi. Note the absence of all Mpi-regulated CPS and the overexpression of PSC. (D) Flow cytometry analysis of surface CPS demonstrates that all CPM1 cells express PSC (99.8%) compared to 4.2% in wild-type cultured cells. Black histograms represent the control antibody; gray histograms (α-PSC) show PSC on the surface of *B. fragilis* CPM1.

Expression of the remaining known capsular polysaccharide (PSC) in *B. fragilis* is predicted to not be controlled by promoter inversion, although it shares other critical transcriptional regulatory elements with the seven Mpi-regulated CPS (Krinos et al.,

2001). Indeed, immunoblot analysis showed an increase in PSC production for CPM1 over that for wild- type cultures (Figure 2.1C). Consistent with this observation, PSC was expressed in 100% of cells for the Mpi mutant strain CPM1 but in only 4.2% of wild-type cultures when measured by flow cytometry with PSC-specific antiserum (Figure 2.1D). Thus, expression of PSC appears to represent a default mechanism for capsule expression in the absence of Mpi-regulated polysaccharides.

Capsule mutants spontaneously revert to express capsular polysaccharides.

To characterize the role of polysaccharide production in *B. fragilis*, we introduced a PSC deletion vector into CPM1 to generate mutants defective in the production of PSC. The strain bearing this deletion (named CPM2) unexpectedly showed a dramatic defect in culture growth compared with that of wild-type *B. fragilis* or the mutant strain CPM1, as assessed by optical density and plating for colony-forming units (Figure 2.2A and data not shown). Furthermore, the mutant strain aggregated upon *in vitro* growth in culture (Supplementary Materials and Methods and Supplementary Figure 2.1). As previously shown for *B. fragilis* and other organisms, the absence of a capsule layer may expose surface adhesive molecules leading to aggregation (Patrick et al., 1996; Schembri et al., 2005). We next investigated the possible mechanism(s) that may explain the growth attenuation resulting from deletion of polysaccharide biosynthesis in *B. fragilis*. When CPM2 was grown in laboratory culture media and sequentially passaged for 5 days, we observed a progressive restoration of growth (Figure 2.2A). Furthermore, passaged cultures displayed a recovery of CPS synthesis, as demonstrated by the presence of high-molecular-weight species after immunoblot analysis of whole-cell lysates using antisera

raised against whole bacteria (Figure 2.2B). Thus, the absence of polysaccharide production and the attenuation of growth are either genetically or phenotypically linked, demonstrating the importance of polysaccharide molecules to *B. fragilis*. The 5-day-passaged strain that exhibits wild-type growth characteristics may have acquired a genetic mutation to overcome the selective pressure of a growth defect; accordingly, we termed this revertant strain CPM3.

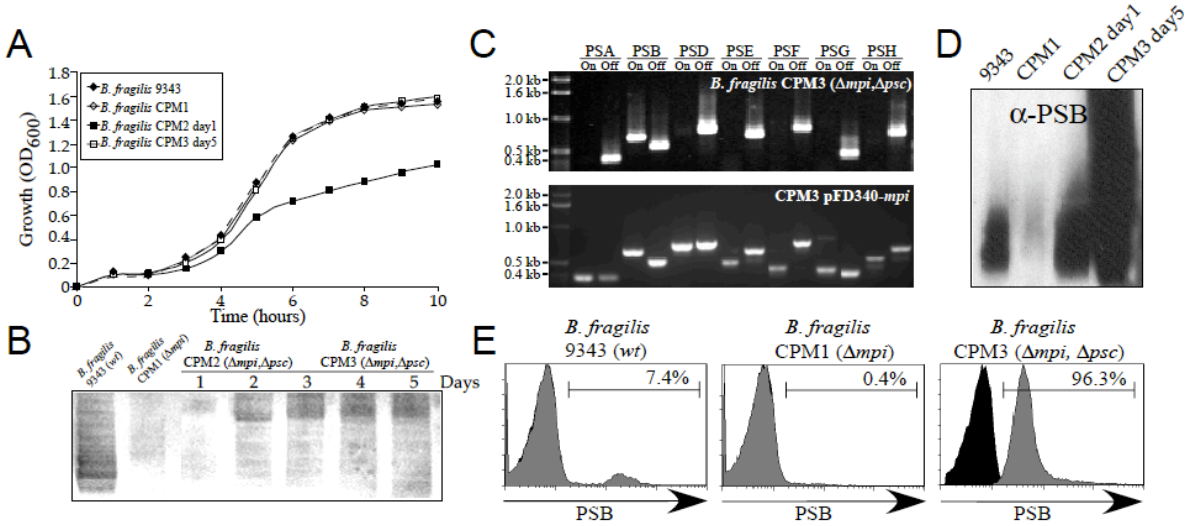


Figure 2.2 | Phenotypic reversion of *B. fragilis* mutants defective in CPS production. (A) Growth curve of bacterial cultures after specific deletion of PSC from the *mpi* mutant CPM1 shows severe growth attenuation of CPM2 (day1). A spontaneous reversion after serial passage of the bacterial cultures for 5 days restores *in vitro* growth (CPM3 day5). (B) Immunoblot analysis shows the overall reduction in CPS production in *mpi* mutant strains (CPM1 and CPM2) and the subsequent stepwise increase in polysaccharide production in CPM3 coincident with growth restoration. (C) Promoter analysis of CPM3 after 5 days of passage demonstrates that the PSB promoter undergoes reversion to the ‘‘on’’ orientation. CPM3 complemented with the *mpi* gene (CPM3 pFD340-*mpi*) displays restoration of phase variation at every promoter. (D) Immunoblot analysis with specific antisera shows the loss of PSB expression in CPM1 and a spontaneous recovery of PSB production under the selective pressure imposed by *psc* deletion (CPM2 day1). A phenotypic enrichment of PSB expression is found in CPM3. (E) Flow cytometry analysis of surface-displayed PSB reveals expression in nearly all bacterial cells after reversion in CPM3 but in only 7.4% of wild-type cultured cells. Black histograms represent control antibody staining; gray histograms (α-PSB) show PSB on the surface of *B. fragilis* CPM3.

Promoter Inversion Occurs in the Absence of Mpi.

What mediates polysaccharide expression in an *mpi* mutant strain? Initially, we

tested the hypothesis that alternate promoters directed in the ‘‘off’’ orientation may be driving transcription of capsular polysaccharides in CPM3. Promoter fusions to transcriptional reporters determined that ‘‘off’’-oriented promoters were incapable of mediating polysaccharide expression (Supplementary Figure 2.2). To examine the possible existence of a pathway secondary to Mpi that is capable of catalyzing promoter inversion, we screened the promoters of polysaccharide operons in CPM3 for their ‘‘on’’ or ‘‘off’’ orientation. We were surprised to observe that the promoter upstream of the PSB biosynthesis locus displayed a reversion to the ‘‘on’’ orientation in CPM3 (the *mpi* mutant background), but only after PSC deletion (Figure 2.2C: Upper). Only the PSB promoter reverted in 8 of 10 trials (PSD and PSH once each); thus, there appears to be a requirement for production of at least one capsular polysaccharide, with a strong (but not absolute) bias for PSB production in the absence of Mpi. Immunoblot analysis of whole-cell lysates using antiserum specific to PSB demonstrated the phenotypic production of PSB (Figure 2.2D). CPM3 exhibited recovery of all Mpi-regulated promoter inversions when *mpi* was provided *in trans* (Figure 2.2C: Lower). To quantify the frequency of reversion, we used flow cytometry to enumerate the proportion of bacteria expressing PSB on their surfaces. Unlike wild-type cultures, in which only 7.4% of cells in a mixed population express PSB, nearly all CPM3 revertants assembled PSB as part of their bacterial envelopes (Figure 2.2E). A previously developed assay using cleavage of PCR products from invertible promoters demonstrated that the PSB promoter in strain CPM3 is found exclusively in the ‘‘on’’ orientation (Supplementary Figure 2.3) (Krinos et al., 2001). Quantitative-PCR digestion and flow cytometry showed that every bacterial cell expresses PSB, thus strongly suggesting that production of at least one capsular

polysaccharide is essential for *B. fragilis* viability. This reversion resulting in the expression of PSB occurred only after deletion of a non-Mpi-controlled polysaccharide (PSC). A further link between the expression of PSB and PSC was shown through the significant increase in PSC-expressing bacteria upon deletion of only PSB (*B. fragilis* 9343ΔPSB), as measured by flow cytometry (Supplementary Figure 2.4). Taken together, these data demonstrate that *B. fragilis* has developed a profound and unusual requirement for the production of at least one capsular polysaccharide.

***B. fragilis* employs multiple pathways for capsular polysaccharide expression.**

The finding that the PSB promoter reverted to the ‘‘on’’ orientation in CPM3 suggests that this strain acquires a ‘‘gain-of-function’’ phenotype to invert promoters under the physiologic stress of multiple polysaccharide deletion. To test whether CPM3 has the ability to catalyze promoter recombination in the absence of *mpi*, we introduced a reporter plasmid with the PSB promoter in the ‘‘off’’ orientation into several *B. fragilis* strains. We measured inversion of the PSB promoter within the plasmid from ‘‘off’’ to ‘‘on’’ orientation by PCR. The increased activity of promoter inversion in CPM3 was evidenced by recombination of the PSB promoter reporter, similar to the function found in wild-type bacteria (Figure 2.3A). CPM1, also deleted in *mpi* but not displaying a selective pressure to revert to polysaccharide production, is incapable of flipping the ‘‘off’’- positioned PSB promoter. Furthermore, analysis of reporter plasmids for the remaining polysaccharide promoters illustrates that they are incapable of flipping to the ‘‘on’’ orientation in CPM3 (Supplementary Figure 2.5), corroborating the specificity of CPM3’s activity for the PSB promoter.

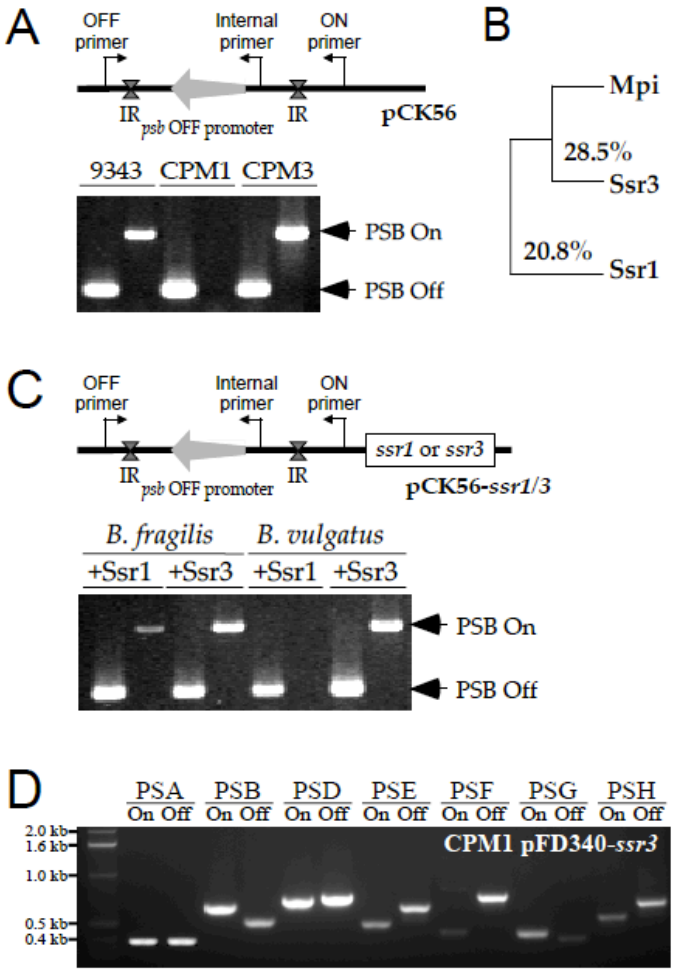


Figure 2.3 | An Mpi-independent pathway for promoter inversion. (A) Upon introduction of reporter plasmid pCK56 (containing the PSB promoter in the ‘‘off’’ orientation), wild-type *B. fragilis* and CPM3 are competent to mediate conversion of the ‘‘off’’ promoter to the ‘‘on’’ position. CPM1 (Δ*mpi*) lacks the ability to invert the locked-off *psb* promoter. (B) Phylogenetic analysis demonstrating the relatedness of the three serine site-specific recombinases (Ssr) in *B. fragilis* NCTC9343 generated by CLUSTALW analysis. The percent identity values indicate direct sequence comparisons of Ssr3 and Ssr1 to Mpi, respectively. (C) PCR products from a reporter plasmid containing the *psb* promoter in the ‘‘off’’ orientation and each of the Mpi homologues, Ssr1 and Ssr3. The *psb* promoter inverts in *B. vulgatus* only in the presence of Ssr3. Ssr3 overexpression in wild-type *B. fragilis* shows an increase in inversion activity. (D) PCR promoter orientation assay of CPM1 complemented with *ssr3* (CPM1 pFD340-*ssr3*). All Mpi-regulated promoters in the ‘‘off’’ orientation undergo phase variation when Ssr3 is constitutively expressed. Compare complementation to the expression of Mpi (Figure 2.1B).

In addition to *mpi*, the genome sequence of the prototype *B. fragilis* strain NCTC9343 has been shown to contain two other serine site-specific recombinase genes:

ssr1 and *ssr3* (also known as *finB*) (Figure 2.3B) (Cerdeno-Tarraga et al., 2005; Coyne et al., 2003). We hypothesized that one of these homologues may serve the nonredundant function of catalyzing recombination at invertible promoters in the absence of Mpi. To determine whether either gene product could mediate promoter inversion, we used reporter plasmids with the PSB promoter in the ‘‘off’’ orientation that also expressed either Ssr1 or Ssr3. These reporter constructs were subsequently introduced into *B. vulgatus*, a related species previously shown to be incapable of inverting the polysaccharide promoters of *B. fragilis* (Coyne et al., 2003). As demonstrated in Figure 2.3C, Ssr1 was unable to mediate PSB promoter inversion in *B. vulgatus*, as assessed by PCR. In contrast, the expression of Ssr3 was sufficient to catalyze this reaction. The ability of Ssr3 to facilitate promoter inversion is consistent with the previous observation that this molecule binds to invertible promoters upstream of polysaccharide biosynthesis operons in *B. fragilis* (Patrick et al., 2003).

We next investigated whether Ssr3 was capable of mediating inversion of the other capsular polysaccharide promoters in *mpi* deletion mutants. *ssr3* was cloned into the expression vector pFD340 and was ectopically overexpressed in CPM1 (with all Mpi-regulated invertible promoters in the ‘‘locked-off’’ orientation). As shown by PCR analysis using promoter-specific primers, Ssr3 is capable of catalyzing recombination at all Mpi-regulated invertible promoters in the absence of Mpi (Figure 2.3D). The recovery of phase variation is also recapitulated in the mutant strain CPM3 (Supplementary Figure 2.6). In addition to capsular polysaccharides, it has been shown that Mpi regulates promoter inversion at six additional genomic loci, the products of which all encode for unknown proteins. Mpi is capable of catalyzing promoter inversion at all of these sites;

Ssr3 is able to flip five of the six promoters, again demonstrating its ability to substitute for the lack of Mpi (Supplementary Figure 2.7). None of these ORFs encode for predicted surface proteins, and, unlike capsular polysaccharides, none are affected by the deletion of PSC (data not shown). Taken together, in addition to Mpi, we have revealed that a second enzyme is capable of mediating promoter inversion in *B. fragilis* and appears to provide a ‘‘fail-safe’’ function to ensure production of capsular polysaccharides.

Mutants expressing a single capsular polysaccharide are defective for intestinal colonization of animals.

It has been proposed that control of multiple polysaccharide loci through promoter inversion allows for the generation of extensive surface diversity during bacterial colonization of the mammalian gastrointestinal tract (Krinos et al., 2001; Kuwahara et al., 2004; Cerdeno-Tarraga et al., 2005). We reasoned that if the purpose of this system is to generate a multiplicity of surface structures (256 possible combinations of the 8 known polysaccharides), the homogeneous polysaccharide-expressing strains of *B. fragilis* we have created may display defects during colonization. Initially, germ-free mice (animals born and raised in the absence of microbial contamination) were readily colonized to similar levels by wild-type and mutant strains upon mono-association with bacteria (Figure 2.4A). Thus, all strains are competent for growth in animals. However, germ-free animals provide a model for the direct comparison of two bacterial strains for initial colonization (mimicking the events after a sterile birth) and thus accurately reflect each strain’s capacity for competitive colonization without the confounding effects of a complete microbiota. We co-associated germ-free animals with wild-type and mutant *B.*

fragilis strains and compared their ability to establish intestinal colonization. Competition experiments between wild-type and CPM1 strains demonstrated that wild-type *B. fragilis* quickly outcompeted the single-polysaccharide (PSC)-producing CPM1 strain in terms of intestinal colonization (Figure 2.4B). After 7 days, 95% of the bacteria recovered from the competition experiment in germ-free animals were wild-type organisms. Coculture of wild-type bacteria with CPM1 during serial passage in laboratory media consistently resulted in no growth defect at any time point (Figure 2.4C); this finding illustrates that the colonization phenotype observed is due to factors found in the intestinal environment of animals. The same phenotype for co-colonization is also observed for CPM3, which expresses only PSB (Supplementary Figure 2.8), as well as for a previously characterized PSA-only-expressing strain Δ*mpi* mut44 (data not shown) (Coyne et al., 2003). Thus, expression of PSA, PSB, or PSC alone is insufficient to allow bacterial competition against wild-type *B. fragilis*.

We next sought to determine whether the colonization defect for CPM1, a single-polysaccharide-producing mutant of *B. fragilis*, could be reversed by restoring expression of the variable surface polysaccharides with ectopic expression of either Mpi or Ssr3. In competitive co-colonization experiments, CPM1 strains complemented with Mpi or Ssr3 rapidly outcompeted the PSC-only-producing mutant CPM1 (Figure 2.4 D and E). This result provides additional evidence that single-polysaccharide-expressing *B. fragilis* mutants exhibit severe defects in host colonization when challenged by the presence of a *B. fragilis* strain that can vary its surface polysaccharide expression. In summary, examinations within this article demonstrate that *B. fragilis* has a strong propensity to elaborate at least one surface polysaccharide for in vitro growth, and expression of a

single polysaccharide appears insufficient to allow bacterial colonization of the mammalian host comparable with that by wild-type bacteria.

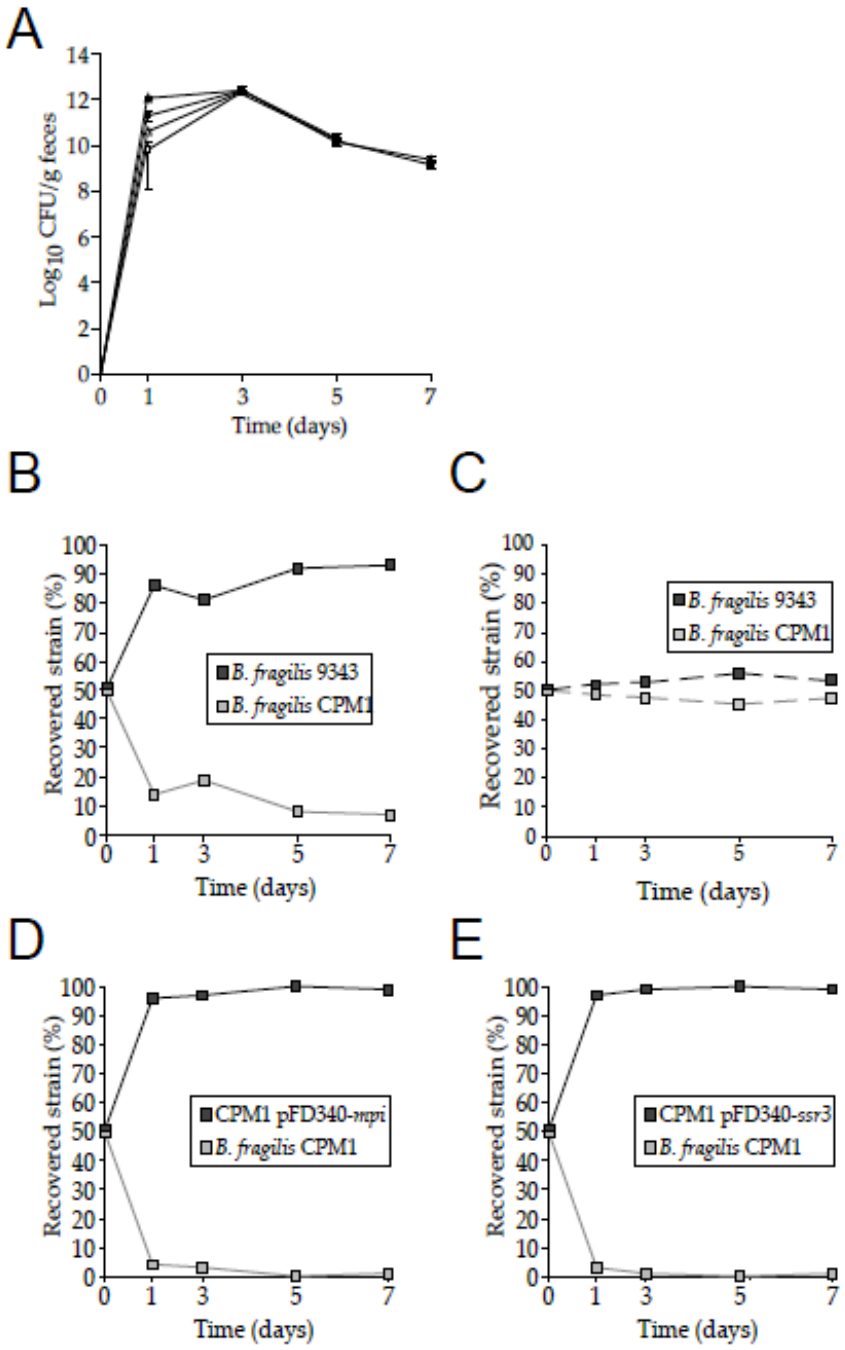


Figure 2.4 | Comparison of germ-free animal colonization by *B. fragilis* mutants with limited diversity of surface polysaccharides. (A) Counts of colony-forming units (CFU) recovered from feces of germ-free animals after oral inoculation of bacteria show the ability of both wild-type and mutant strains to colonize laboratory animals during mono-association. Filled

squares, wild-type *B. fragilis* 9343; open triangles, *B. fragilis* CPM1; open squares, *B. fragilis* CPM3. (B) Percentages of each strain recovered from feces of animals after dual colonization of wild-type *B. fragilis* and *B. fragilis* CPM1 strain, premixed at a 1:1 ratio before gavage. Upon plating on selective media, wild-type bacteria quickly outcompete CPM1. (C) Percentages of each strain recovered from laboratory media after daily serial passages during coculture demonstrate that CPM1 displays no growth defect in culture competition with wild-type *B. fragilis*. Overnight cultures were subcultured daily into fresh media at a 1:100 dilution. (D) Percentages of each strain recovered from feces of animals after dual colonization of CPM1 and CPM1 complemented with *mpi*. (E) Percentages of each strain recovered from feces of animals after dual colonization of CPM1 and CPM1 complemented with *ssr3 in trans*.

DISCUSSION

The intestinal microbiota of humans contains 10 times more cells than the human body and 100 times the number of genes than the human genome (O'Hara and Shanahan, 2006). Although we are beginning to understand the identities of microorganisms that inhabit the gastrointestinal tract, the mechanisms they use to establish colonization are almost entirely unknown. The *Bacteroides* represent one of the most numerically prominent constituents of the human microbiota, and all members of this genus encode multiple capsule loci and contain Mpi homologues (Xu et al., 2007). Are capsular polysaccharides key molecular components that mediate various interactions between symbiotic *Bacteroides* and their mammalian hosts? The importance of these molecules is suggested by our finding that *B. fragilis* requires expression of at least one capsular polysaccharide for *in vitro* growth. When we selected for a mutant that was genetically unable to express all seven Mpi-regulated polysaccharides, we observed that every viable cell phenotypically expressed the default polysaccharide PSC. When we deleted PSC in this strain, we were able to recover only a poorly growing strain that overcame its growth defect and reverted to express another polysaccharide. To our knowledge, this is the first demonstration that the lack of capsule expression adversely affects bacterial growth. We demonstrate that *B. fragilis* contains an alternative recombinase (Ssr3) that is capable of inverting polysaccharide promoters. We were, however, unable to generate mutants of Ssr3 (because it is present on a multicopy plasmid) to directly assess whether it is responsible for the CPM3 phenotype. Also, genomic sequencing of the PSB promoter or the *ssr3* gene and upstream regulatory elements yielded no mutations between any of the CPM mutants and wild-type bacteria (data not shown). Importantly, Ssr3 can compensate

for the absence of Mpi in both catalyzing promoter inversion and complementing defects in animal colonization. An inhibition of growth upon capsule deletion, a redundancy of mechanisms to ensure capsule biosynthesis, and defects in colonization upon limiting the capsular polysaccharide repertoire attest to the evolutionary importance for regulation of surface architecture by *B. fragilis* during colonization.

The bacterial cell envelope can be viewed as the structural interface between microorganisms and their countless environments, mediating essential functions required for microbial attachment and colonization (Mazmanian et al., 2001). The many examples of adherence mechanisms used by bacterial pathogens during infection include molecules such as pili, fimbriae, and surface proteins (Pizarro-Cerda and Cossart, 2006). However, mechanisms involved in the establishment of colonization by multitudinous and important symbiotic organisms remain largely uncharacterized. Pioneering work by Comstock and colleagues (Krinos et al., 2001) first demonstrated that *B. fragilis* contains at least eight distinct capsular polysaccharides; the recent findings from genome sequences that three other *Bacteroides* species contain multiple capsule structures indicate that this is a unique and distinguishing feature of these important human commensals (Xu et al., 2007). We demonstrate that *B. fragilis*, a numerically prominent symbiotic organism of the human microbiota, is attenuated for intestinal colonization when it can no longer express a diverse repertoire of surface structures. Specifically, our data show that *B. fragilis* strains able to synthesize only a single polysaccharide cannot compete with wild-type bacteria for survival in the gastrointestinal tract of germ-free animals. Restoration of surface diversity in these mutants through complementation with either Mpi or its alternative homologue, Ssr3, rescues defects in both capsular

polysaccharide production and intestinal colonization. All sequenced *Bacteroides* contain orthologs of Mpi, presumably to create extensive surface diversity in these human commensals. Thus, our results with the well studied model organism *B. fragilis* may extend to other numerically significant commensal bacteria that inhabit the gastrointestinal tract of humans. Because the gut presents a dynamic and changing environment to colonizing organisms (i.e., nutrient changes, immune responses, bacteriophage attacks), perhaps sustained association of commensal bacteria with mammals requires intricate and dynamic processes unlike expression of a single toxin or adhesin used by many pathogens to establish acute infections. The regulation of surface architecture during intestinal colonization appears to be critically involved for establishing the commensal association of *B. fragilis* with its mammalian host. Recent efforts to define the ‘‘normal’’ sequence of colonization by bacterial species after birth have resulted in a deeper understanding of the evolutionary partnership between humans and their microbiota (Palmer et al., 2007). Our findings provide the framework to compel future studies to reveal the mechanisms by which the bacterial envelope contributes to the vital process of host–bacterial mutualism.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media.

Bacterial strains and plasmids are described in Supplementary Table 2.2.

Generation of strains CPM1 and CPM2.

To create the *B. fragilis*Δ*mpi* mutant CPM1, the plasmid pLEC80 (Supplementary Table 2.2) was mobilized fromDH5α cells by helper plasmid RK231 and conjugally transferred into a *B. fragilis* 9343ΔPSA mutant strain to generate a deletion of *mpi* and *tsr19*, a tyrosine site-specific recombinase that is not involved in Mpi-mediated capsular polysaccharide promoter inversion (Roche-Hakansson et al., 2007). Colonies selected for erythromycin resistance (Em^r), indicating integration of the suicide vector into the host chromosome, were passaged for 5 days and then plated on nonselective medium (BHIS). The resulting colonies were replica-plated to BHIS containing Em, and Em^s (sensitive) colonies were screened by PCR to distinguish wild-type revertants from strains with the desired mutation. To create mutant CPM2, a portion of the PSC locus was deleted from CPM1 by using plasmid pMJC2Δ.1 (Supplementary Table 2.2). Complementation studies with *mpi* and *ssr3* were performed by cloning each gene into the *B. fragilis–Escherichia coli* shuttle plasmid pFD340 (Supplementary Table 2.2).

Flow cytometry.

For surface staining, cell preparations were washed twice in ice-cold FACS buffer (PBS with 2% FBS) and resuspended in 100 μl of the same buffer. A total of 1×10^7 cells were incubated with rabbit polyclonal antisera to PSB and PSC at a 1:100 dilution for 30

min at 4°C. Cells were then washed, incubated with goat anti-rabbit fluorochrome-conjugated secondary antibodies, and analyzed by flow cytometry with an FC500 cytometer.

Germ-free animal colonization experiments.

Male Swiss–Webster germ-free mice were purchased from Taconic Farms. Animals were screened for bacterial, viral, and fungal contamination by Gram staining, degenerate PCR, RapID ANA II System analysis, and plating of fecal samples under aerobic and anaerobic conditions. Animals were housed in gnotobiotic chambers or microisolator cages, and all food and bedding were sterilized by autoclave. Erythromycin (10 μg/ml) and gentamicin (100 μg/ml) were added to sterile drinking water. All *Bacteroides* strains (naturally resistant to gentamicin) contained the plasmid pFD340 conferring erythromycin resistance. Eight- to 13-week-old mice were orally inoculated with ~ 1×10^8 colony-forming units of bacteria harvested from a log-phase culture and resuspended in PBS with 1.5% NaHCO3. At each time point after bacterial introduction to animals, fresh fecal samples were collected, serially diluted, and plated for colony-forming units. To discern between strains, we introduced a plasmid pFD340-*cat* (Supplementary Table 2.2), which confers chloramphenicol resistance. One hundred individual colonies were patched onto BHIS agar with erythromycin and chloramphenicol to determine the ratio between strains during colonization. All animals were cared for under established protocols using Institutional Animal Care and Use Committee guidelines of Harvard Medical School and the California Institute of Technology.

REFERENCES

Bik, E.M., Eckburg, P.B., Gill, S.R., Nelson, K.E., Purdom, E.A., Francois, F., Perez-Perez, G., Blaser, M.J., and Relman, D.A. (2006). Molecular analysis of the bacterial microbiota in the human stomach. Proc Natl Acad Sci U S A *103*, 732–737.

Bry, L., Falk, P.G., Midtvedt, T., and Gordon, J.I. (1996). A model of host-microbial interactions in an open mammalian ecosystem. Science *273*, 1380–1383.

Cerdeno-Tarraga, A.M., Patrick, S., Crossman, L.C., Blakely, G., Abratt, V., Lennard, N., Poxton, I., Duerden, B., Harris, B., Quail, M.A., Barron, A., Clark, L., Corton, C., Doggett, J., Holden, M.T., Larke, N., Line, A., Lord, A., Norbertczak, H., Ormond, D., Price, C., Rabbinowitsch, E., Woodward, J., Barrell, B., and Parkhill, J. (2005). Extensive DNA inversions in the B. fragilis genome control variable gene expression. Science *307*, 1463–1465.

Comstock, L.E., and Kasper, D.L. (2006). Bacterial glycans: key mediators of diverse host immune responses. Cell *126*, 847–850.

Coyne, M.J., and Comstock, L.E. (2008). Niche-specific features of the intestinal bacteroidales. J Bacteriol *190*, 736–742.

Coyne, M.J., Reinap, B., Lee, M.M., and Comstock, L.E. (2005). Human symbionts use a host-like pathway for surface fucosylation. Science *307*, 1778–1781.

Coyne, M.J., Weinacht, K.G., Krinos, C.M., and Comstock, L.E. (2003). Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proc Natl Acad Sci U S A *100*, 10446–10451.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635–1638.

Gao, Z., Tseng, C.H., Pei, Z., and Blaser, M.J. (2007). Molecular analysis of human forearm superficial skin bacterial biota. Proc Natl Acad Sci U S A *104*, 2927–2932.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M., and Nelson, K.E. (2006). Metagenomic analysis of the human distal gut microbiome. Science *312*, 1355–1359.

Hooper, L.V., and Gordon, J.I. (2001). Commensal host-bacterial relationships in the gut. Science *292*, 1115–1118.

Hooper, L.V., Midtvedt, T., and Gordon, J.I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu Rev Nutr *22*, 283–307.

Johri, A.K., Paoletti, L.C., Glaser, P., Dua, M., Sharma, P.K., Grandi, G., and Rappuoli, R. (2006). Group B Streptococcus: global incidence and vaccine development. Nat Rev Microbiol *4*, 932–942.

Krinos, C.M., Coyne, M.J., Weinacht, K.G., Tzianabos, A.O., Kasper D.L., and Comstock, L. E. (2001). Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature *414*, 555–558.

Kuwahara, T., Yamashita, A., Hirakawa, H., Nakayama, H., Toh, H., Okada, N., Kuhara, S., Hattori, M., Hayashi, T., and Ohnishi, Y. (2004). Genomic analysis of Bacteroides fragilis reveals extensive DNA inversions regulating cell surface adaptation. Proc Natl Acad Sci U S A *101*, 14919–14924.

Lee, C.J., Lee, L.H., Lu, C.S., and Wu, A. (2001). Bacterial polysaccharides as vaccines--immunity and chemical characterization. Adv Exp Med Biol *491*, 453–471.

Ley, R.E., Peterson, D.A., and Gordon J.I. (2006). Ecological and evolutional forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

Mazmanian, S.K., and Kasper, D.L. (2006). The love-hate relationship between bacterial polysaccharides and the host immune system. Nat Rev Immunol *6*, 849–858.

Mazmanian, S.K., Liu, C.H., 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.

Mazmanian, S.K., Ton-That, H., and Schneewind, O. (2001). Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus. Mol Microbiol *40*, 1049–1057.

O'Hara, A.M., and Shanahan, F. (2006). The gut as a forgotten organ. EMBO Rep *7*, 688–693.

Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the human infant intestinal microbiota. PLoS Biol *5*, e177.

Patrick, S., McKenna, J.P., O'Hagan, S., and Dermott, E. (1996). A comparison of the haemagglutinating and enzymic activities of Bacteroides fragilis whole cells and outer membrane vesicles. Microb Pathog *20*, 191–202.

Patrick, S., Parkhill, J., McCoy, L.J., Lennard, N., Larkin, M.J., Collins, M., Sczaniecka, M., and Blakely, G. (2003). Multiple inverted DNA repeats of Bacteroides fragilis that control polysaccharide antigenic variation are similar to the hin region inverted repeats of Salmonella typhimurium. Microbiology *149*, 915–924.

Pizarro-Cerda, J., and Cossart, P. (2006). Bacterial adhesion and entry into host cells. Cell *124*, 715–727.

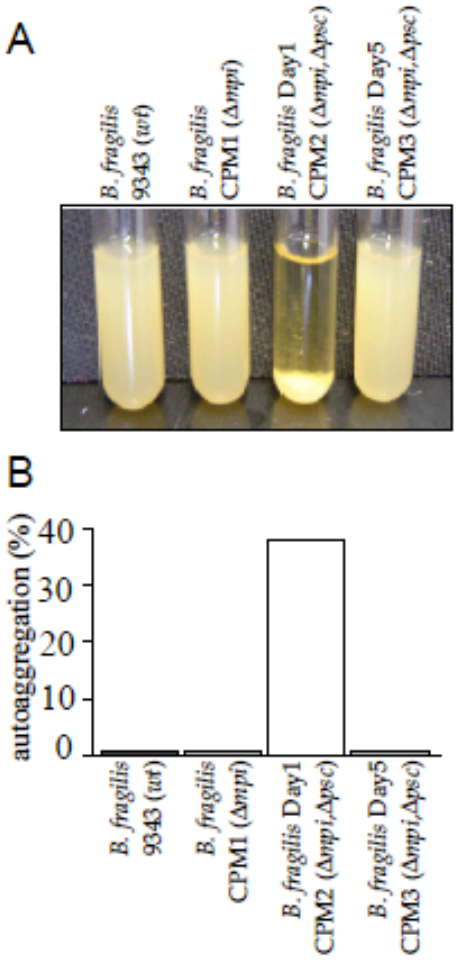
Roche-Hakansson, H., Chatzidaki-Livanis, M., Coyne, M.J., and Comstock, L.E. (2007). Bacteroides fragilis synthesizes a DNA invertase affecting both a local and a distant region. J Bacteriol *189*, 2119–2124.

Schembri, M.A., Blom, J., Krogfelt, K.A., and Klemm, P. (2005). Capsule and fimbria interaction in Klebsiella pneumoniae. Infect Immun *73*, 4626–4633.

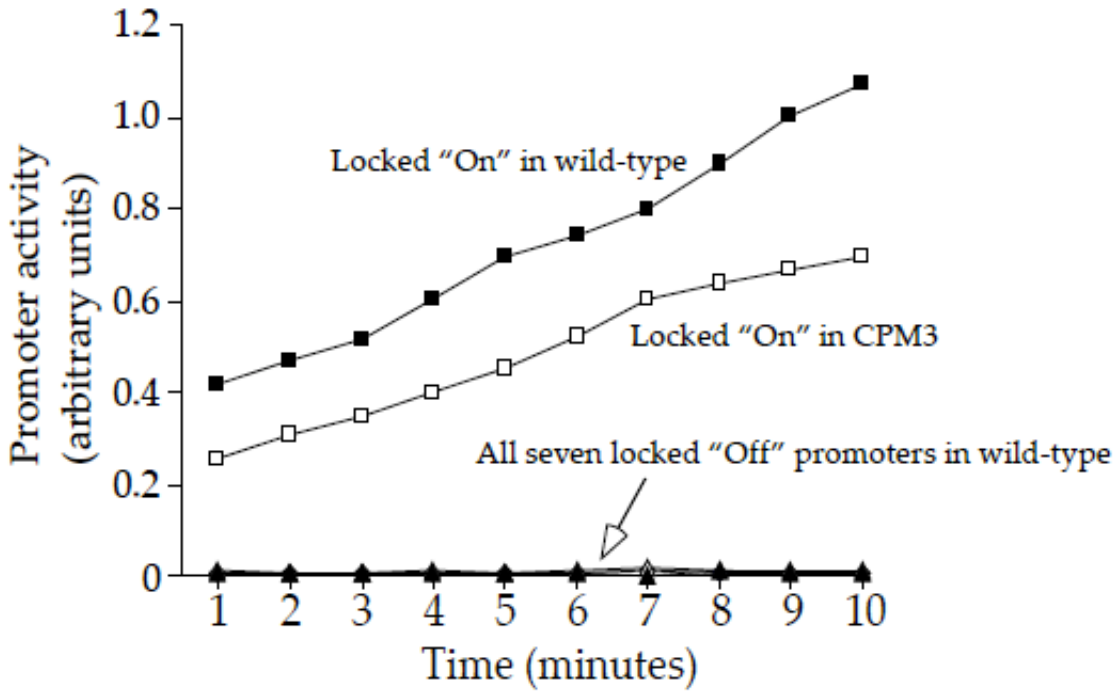
Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., and Gordon, J.I. (2003). A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science *299*, 2074–2076.

Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho, P.M., Minx, P., Latreille, P., Cordum, H., Van Brunt, A., Kim, K., Fulton, R.S., Fulton, L.A., Clifton, S.W., Wilson, R.K., Knight, R.D., and Gordon, J.I. (2007). Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol *5*, e156.

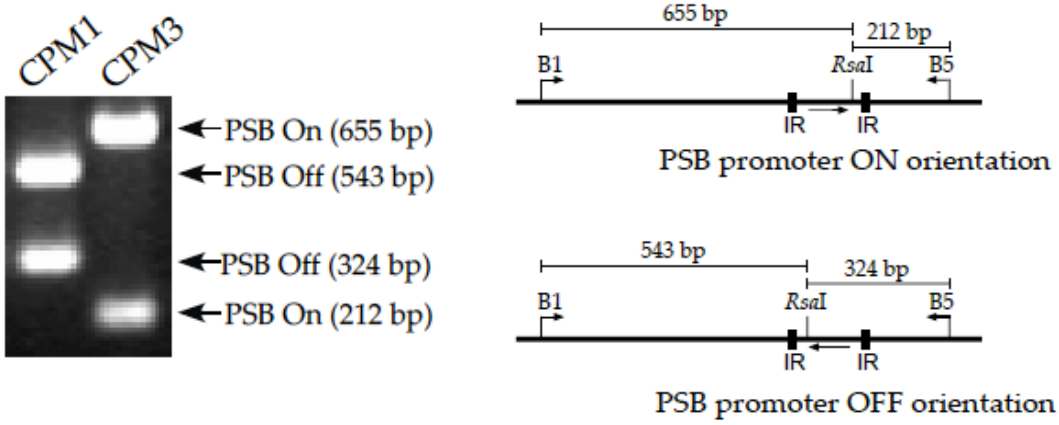
Supplementary Figure 2.1 | (A) Wild-type and CPM1 (Δ*mpi*) strains display no aggregation in overnight liquid culture whereas CPM2 (Δ*mpi*, Δ*psc*) aggregates to each other and sinks to the bottom of the culture tube. Five-day passage in culture allows this phenotype to revert (day-5 CPM3). (B) The autoaggregation phenotype of day-1 CPM2 liquid culture and the subsequent reversion of the phenotype in day-5 CPM3 liquid culture are quantitatively determined by optical density measurements.



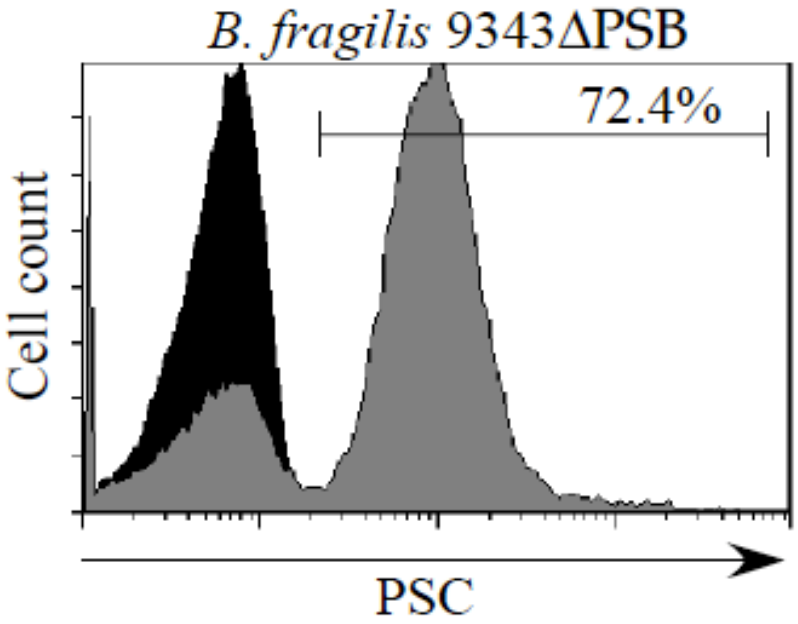
Supplementary Figure 2.2 | Transcriptional reporter assay shows that invertible promoters do not possess activity when locked in the "off" orientation. Transcriptional reporter fusions of all seven Mpi-regulated polysaccharide promoters in the "locked-off" orientation (i.e., mutated in the inverted repeat sites) with the *xylE* gene of *Escherichia coli* cannot support transcription in wild-type *B. fragilis*. Furthermore, introduction of the same "locked-off" reporter constructs into strain CPM3 once again results in lack of XylE expression (data not shown). These results demonstrate the absence of an additional promoter with the opposite orientation. "Locked-on" promoter fusions in wild-type and CPM3 organisms serve as assay controls.



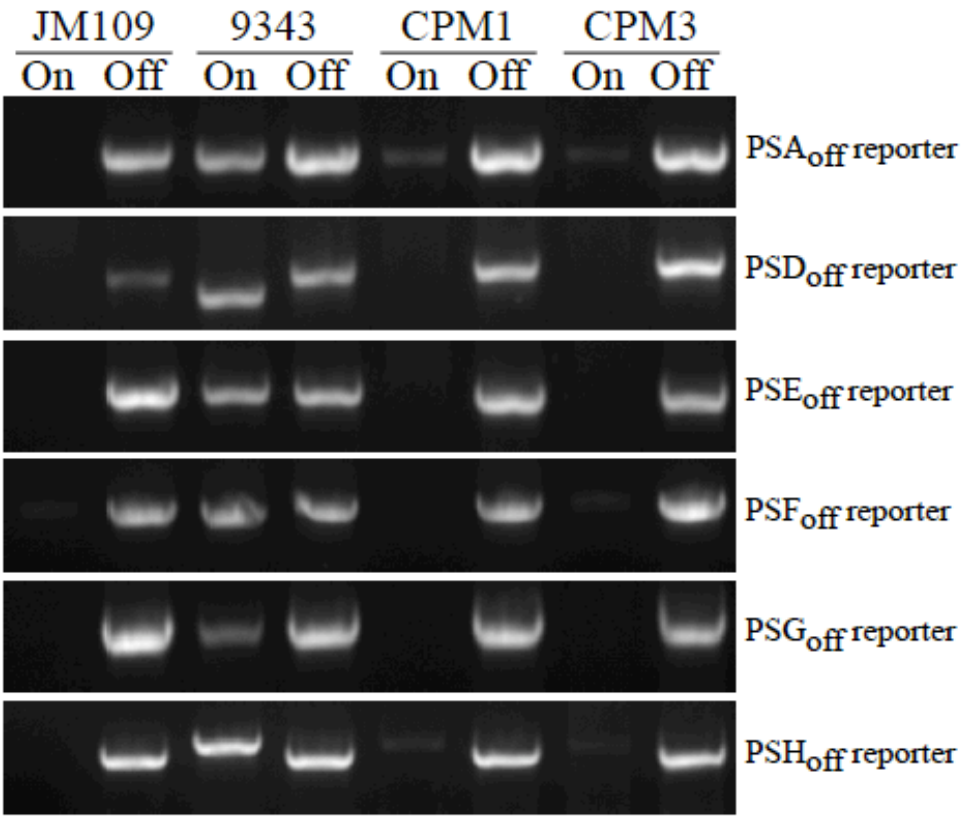
Supplementary Figure 2.3 | EtBr-stained agarose gel electrophoresis analysis of the PSB invertible promoter region from CPM1 and CPM3 chromosomal DNA is amplified with PCR using primers B1 and B5 (Supplementary Table 2.1). The preferred orientation of the promoter can be quantitatively determined when digested with restriction endonuclease (RsaI), which cleaves the fragment asymmetrically between the inverted repeat elements. PCR-digestion assay for the PSB invertible promoter region demonstrates that, whereas the *psb* promoter is completely in the "off" orientation in strain CPM1 (Δ*mpi*), as also shown by immunoblot and flow cytometry, serially passaged CPM3 (Δ*mpi*, Δ*psc*) mutant strain displays only "on" promoters.



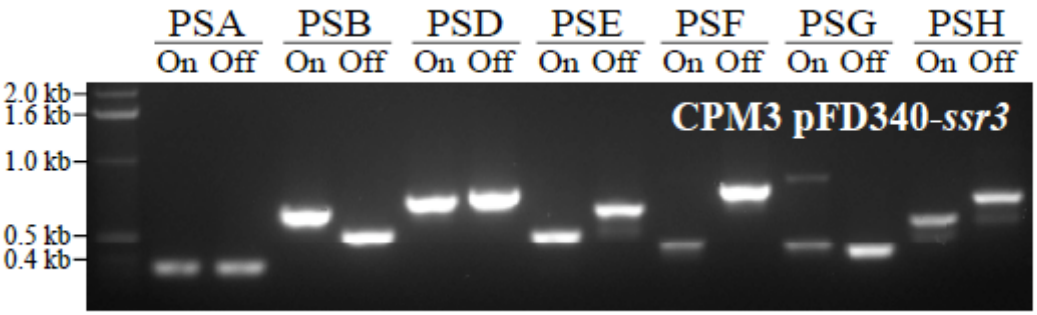
Supplementary Figure 2.4 | Flow cytometry analysis of surface polysaccharides shows PSC expression in > 70% of bacterial cells after the specific mutation of only the *psb* locus (*B. fragilis* 9343ΔPSB) but in only 4.2% of wild-type cultured cells (Figure 2.1D). Thus, the expression of PSB and PSC is phenotypically linked. Black histogram represents control staining; gray histogram shows PSC on the surface of *B. fragilis* 9343ΔPSB.



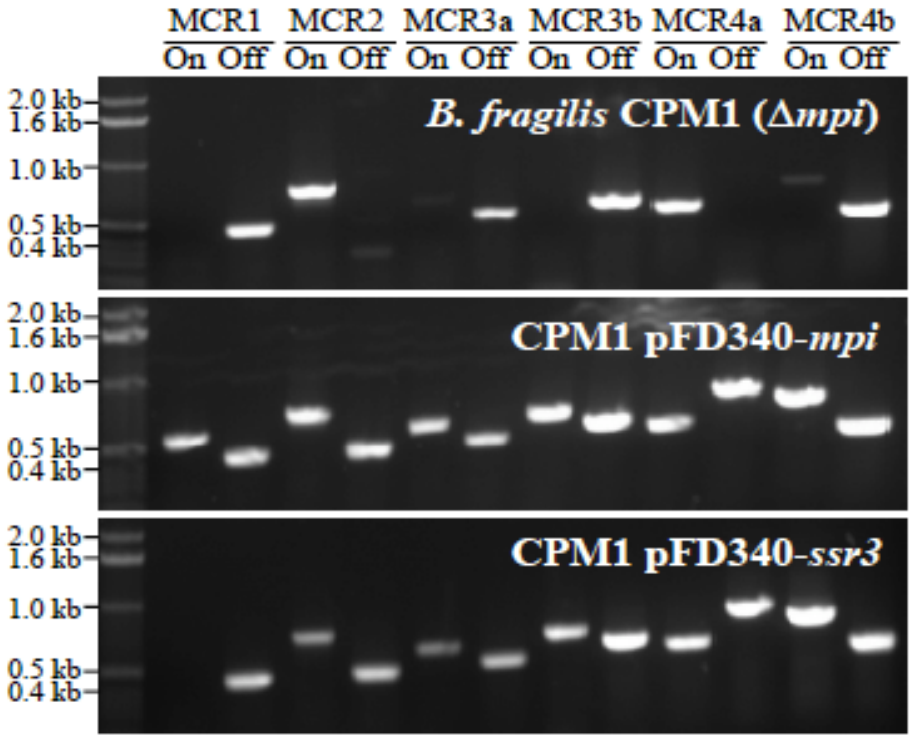
Supplementary Figure 2.5 | PCR analysis of the reporter plasmids containing one of the six invertible capsular polysaccharide promoters (*psa*, *psd*, *pse*, *psf*, *psg*, and *psh*) in the "off" orientation. Upon introduction into bacteria, all six of the "off"-oriented invertible CPS promoters flip to the "on" orientation in wild-type *B. fragilis* but not in the Δ*mpi* mutant strains, CPM1 and CPM3. The recombination of the invertible promoter in CPM3 is specific to *psb*. The orientation of the promoter within the reporter constructs was analyzed from JM109 as a negative control.



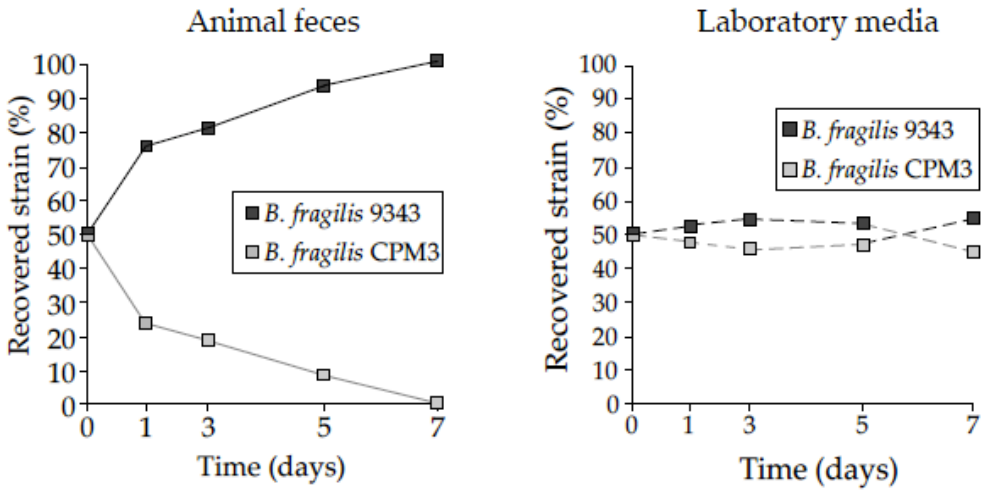
Supplementary Figure 2.6 | EtBr-stained agarose gel electrophoresis of PCR products generated from chromosomal DNA of *B. fragilis* mutant CPM3 (Δ*mpi*, Δ*psc*) complemented with *ssr3* gene (CPM3 pFD340-*ssr3*) with specific primers for both the "on" and the "off" orientations for each polysaccharide promoter. All Mpi-regulated promoters in the "off" orientation previously demonstrated by the PCR analysis of the chromosomal DNA of CPM3 (Figure 2.2C) undergo phase variation at each promoter when Ssr3 is constitutively expressed at a high level.



Supplementary Figure 2.7 | PCR orientation assay of *B. fragilis* mutant CPM1 (Δ*mpi*), CPM1 complemented with *mpi* gene (CPM1 pFD340-*mpi*), and CPM1 complemented with *ssr3* gene (CPM1 pFD340-*ssr3*) with specific primers for both the "on" and the "off" orientations at each of the six Mpi-controlled regions (Coyne et al., 2003). In the *mpi* mutant strain, all MCRs (Mpi-controlled regions) are locked to either "on" or "off" orientation. These promoters undergo phase variation at each promoter when Mpi is expressed ectopically (CPM1 pFD340-*mpi*). When Ssr3 is expressed under the pFD340 promoter, all MCRs with an exception of the MCR1 recover their ability to invert (CPM1 pFD340-*ssr3*).



Supplementary Figure 2.8 | CPM3 (expressing only PSB) is outcompeted by wild-type bacteria in animal colonization but displays no *in vitro* growth defect in co-culture with wild-type strain. For animal cocolonization, bacterial strains were premixed at a 1:1 ratio of mutant to wild-type and gavaged into germ-free animals. At specified time points, fecal samples were diluted in BHIS media and plated. The ratio between the two strains from the recovered bacteria was analyzed by strain-specific PCR. For *in vitro* co-culture experiment, bacterial strains were premixed before culture at a 1:1 ratio and inoculated into BHIS laboratory media. Overnight cultures were subcultured daily into fresh media (at a 1:100 dilution) for serial passage and plated. The percentage of each strain recovered from laboratory media was analyzed by strain-specific PCR. All bacteria were used only after serial passage in culture for 5 days to allow for phenotypic reversion of CPM2.



Supplementary Table 2.1 | Sequences of primers used for PCR in this study (All primers listed 5'→3')

B. fragilis chromosomal DNA PCR promoter orientation assay primers

Primer name	Sequence	Purpose
A3	ACCTTTTTTCGACCTTTTCTAAAAATC	Inversion PSA promoter
A4	ACAAAGGTAAGGCACATTTTTATAAC	Inversion PSA promoter
C473	AGAAAACTCCTGGTCCTTCTTTG	Inversion PSA promoter
B2	ACAGACTCCTTACCTTTGTTCAATCAAACG	Inversion PSB promoter
B4	CCGGAATGCTCTGGCATATTTTTTCAGCTC	Inversion PSB promoter
B6	GCGCTCAATACACCGGCAATACGAATAAC	Inversion PSB promoter
D5	GCAGTTATGAAAATACCCTCTATCTTTGCG	Inversion PSD promoter
D6	GTCACCTTTGATACGGACAAACTCACCCTC	Inversion PSD promoter
D7	GTTTTTTCCATCTCAGTTTCATGGCTTCAG	Inversion PSD promoter
E1	GCCTTTTCCGTTGCTTACTG	Inversion PSE promoter
E3	CGTTGAGGATAACAGCAGCA	Inversion PSE promoter
E4	AGGTATAAACTAAATTTTGATGTGCAA	Inversion PSE promoter
F1	CGTTTCATGTAAGGCGGATT	Inversion PSF promoter
F2	CAGAAGAGAACAGAAAACAAAATCA	Inversion PSF promoter
F3	CCAGTTCAAAGCGGAAGAAG	Inversion PSF promoter
G1	TTTGCTTGTTGTCCGTTTTG	Inversion PSG promoter
G2	TCGAAACATAAAAGCAGACAGA	Inversion PSG promoter
G3	ACCGCATAGCGTCAGTCTCT	Inversion PSG promoter
H1	CTTTGCCAGTTCCCGTATGT	Inversion PSH promoter
H2	TGATGAAATTCAGAACCGGATA	Inversion PSH promoter
H3	CGCTCGTTCTTGACGATGTA	Inversion PSH promoter
MCR1-1	TGTATCCGGAGGAATGTGTTATAC	Inversion MCR1 promoter
MCR1-2	CGAGTGATTTATTACGAAAAAGGA	Inversion MCR1 promoter
MCR1-3	CAACCTATAATTTTCCGGAGTAGG	Inversion MCR1 promoter
MCR2-1	GGTATTAAATGAGTAGCCCACCAC	Inversion MCR2 promoter
MCR2-2	CTGTGCGAGAACTAAAGATTGTGT	Inversion MCR2 promoter
MCR2-3	AGCAGTATATTTTGGAAGCGTAGG	Inversion MCR2 promoter
MCR3a-1	GTCGTACAGGCTATCCGAGACTAT	Inversion MCR3a promoter
MCR3a-2	TGTACAAAATGCAAAGGGAATAGA	Inversion MCR3a promoter
MCR3a-3	TTAGGGTGGAAGGAGTATTTTCTG	Inversion MCR3a promoter
MCR3b-1	GAAGGTAAAACCGATGTCAATAGC	Inversion MCR3b promoter
MCR3b-2	GGGAGCCGATAAACTTCTGATAAT	Inversion MCR3b promoter
MCR3b-3	TACGTAATTGCTGCCTACTTTACG	Inversion MCR3b promoter
MCR4a-1	CACTCATGTAGTTCACCATTTTGG	Inversion MCR4a promoter
MCR4a-2	CTGAAGGCCAAAGGAAGATATAAA	Inversion MCR4a promoter
MCR4a-3	CCTGCCGTACTATTGAACGAGTAT	Inversion MCR4a promoter
MCR4b-1	CTACTAGTGATGGGGGTACAGGAG	Inversion MCR4b promoter
MCR4b-2	AACATCATTTTTACTCTCCGCACT	Inversion MCR4b promoter
MCR4b-3	AGCCGACTGGAAGACATAGTTATT	Inversion MCR4b promoter

Primers used for PSB promoter PCR-digestion assay

B1	CCGGAATGCTCTGGCATATTTTTTCAGCTC
B5	GCGGCAAACCAATACTTCTGTTGTTCAGA

pCK56 PCR promoter orientation assay primers

C7	GGGGACATTGTCTCTCTTTC
VarB-F	CGCAAAGATTACTTTTCCATTTACTTCG
L103	GTTCGTTTGATTGAACAAAGGTAAGG

Primers used to clone *ssr1* and *ssr3* (bold: 5' addition)

ssr1 F	**CCGGATCC**CAAACTTAGATACCATCAAAGAGAACC
ssr1 R	**TGGGATCC**AAAAATGCGGAAATAGCTCAGT
ssr3 F	**AGGGATCC**TTTTGCGAATAGTTGAATTTATGA
ssr3 R	**CAGGATCC**AGCCCGACCTGAAAGGTAAC

Primers used to amplify the promoter upstream sequences for XylE assay (bold: 5' addition)

PSA F	**ATCGGGATCC**TTCGTTTATAAACACAAAGG
PSA R	**ATCGGGATCC**GGATATTAAAAGTCGAGAG
PSB F	**ATCGGGATCC**TTCGTTTGATTGAACAAAGG
PSB R	**ATCGGGATCC**TATACTTGGGGAAATTTTC
PSD F	**ATCGGGATCC**ATCGTCTAATTGAACCAAAGATAG
PSD R	**ATCGGGATCC**ATTCTATCAGCGCTTTTA
PSE F	**ATCGGGATCC**TTCGTTTAATAGCACAAAGGTAT
PSE R	**ATCGGGATCC**TTGGTTGACGAGAGGT
PSF F	**ATCGGGATCC**TTCGTTTAATTGAACGCAAAGATA
PSF R	**ATCGGGATCC**GTTCTTTGGGGACACCGG
PSG F	**ATCGGGATCC**TTCGTCTATTTGAACGCAAAGATAA
PSG R	**ATCGGGATCC**TATTTCAGAAAAGGACATTC
PSH F	**ATCGGGATCC**TTCGTTCAATGGAACAAAGGTAA
PSH R	**ATCGGGATCC**TCGGTTTTTTTTGCGTCATTTC

Primers used in germ-free animal colonization competition and in vitro competition experiment

CPM1-specific PCR primers	
ssr2-D1	AGTACTGATAACTCCGGTGACTCC
ssr2-D6	ATGACATAGATAATGGGGAAGAGG
CPM2-specific PCR primers	
C121	TATCCTGATGTTCTGCTTTTCCG
C127	GGAGGATGTTTGAATTGGTGG

Supplementary Table 2.2 | Plasmids and strains used in this study

Strain or plasmid	Description	Reference or source
E. coli DH5a	F$^-$ f80d*lac*ZDM15 D(*lacZYA-argF*) *U169 deoR recA1 endA1 hsdR17*($r_K^- m_K^+$) *phoA supE44 l*$^-$ *thi-1 gyrA96 relA1*	1
Bacteroides fragilis NCTC 9343	Type strain	2
B. fragilis 9343DPSA	9343 mutant with chromosomal deletion of PSA consecutive biosynthesis genes *wzx-wcfS*	3
B. fragilis 9343DPSA D*mpi*.off (CPM1)	9343DPSA mutant with chromosomal deletion of *mpi-tsr19* using pLEC80 deletion vector; invertible promoters for PSA, PSB, and PSD-H all in "off" orientation	This study
B. fragilis 9343DPSA DPSCD*mpi*.off (CPM2)	9343DPSAD*mpi*.off mutant with chromosomal deletion of PSC consecutive biosynthesis genes *wcfD-L, wzy,* and *orf5*	This study
B. vulgatus ATCC 8482	Type strain	4
PCR 2.1	Invitrogen TA-cloning vector; Km^r Amp^r	Invitrogen
pNJR6	*Bacteroides* suicide vector; *mob*$^+$ Tra$^-$ Km^r (*E. coli*) Em^r (*Bacteroides*)	5
R751	Mobilizable mating plasmid to move constructs from *E. coli* to *B. fragilis*; Tra$^+$ Tp^r	6
RK231	Mobilizable mating plasmid to move constructs from *E. coli* to *B. fragilis* and *B. vulgatus*, RK2 derivative; Tra$^+$ Tet^r Km^r	7
pCK56	Modified pFD340 plasmid containing PSB invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pCK56-*ssr1*	Modified pCK56 plasmid containing the *ssr1* gene PCR amplified from *B. fragilis* NCTC 9343 cloned into a *BamHI* site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pCK56-*ssr3*	Modified pCK56 plasmid containing the *ssr3* gene PCR amplified from *B. fragilis* NCTC 9343 cloned into a *BamHI* site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pLEC80	*Bacteroides* D*mpi-tsr19* deletion vector; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pMJC2D.1	*Bacteroides* DPSC deletion vector; Km^r (*E. coli*) Em^r (*Bacteroides*)	8
pFD340	*E. coli/B. fragilis* shuttle vector, IS*4351* promoter; Amp^r (*E. coli*) Em^r (*Bacteroides*)	9
pFD340-*mpi*	Modified pFD340 plasmid containing the *mpi* gene PCR amplified from *B. fragilis* NCTC9343 cloned into a *BamHI* site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*ssr3*	Modified pFD340 plasmid containing the *ssr3* gene PCR amplified from *B. fragilis* pBF9343 cloned into a *BamHI* site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*cat*	Modified pFD340 plasmid containing the *cat* gene PCR amplified from E. Coli K12/pACYC184 (accession #: X06403) cloned into a *SmaI* site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*psa*.off	Modified pFD340 plasmid containing PSA invertible promoter region in the "off" position cloned into the	This study

	PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	
pFD340-*psd*.off	Modified pFD340 plasmid containing PSD invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*pse*.off	Modified pFD340 plasmid containing PSE invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*psf*.off	Modified pFD340 plasmid containing PSF invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*psg*.off	Modified pFD340 plasmid containing PSG invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study
pFD340-*psh*.off	Modified pFD340 plasmid containing PSH invertible promoter region in the "off" position cloned into the PstI site; Amp^r (*E. coli*) Em^r (*Bacteroides*)	This study

1. Bethesda Research Laboratories. (1986) *Focus* 8: 9.
2. Johnson JL. (1978) *Int J Syst Bacteriol* 28: 245–268.
3. Coyne MJ, Tzianabos AO, Mallory BC, Carey VJ, Kasper DL, and Comstock LE. (2001) *Infect Immun* 69: 4342–4350.
4. Cato EP, and Johnson JL. (1976) *Int J Syst Bacteriol* 26: 230–237.
5. Stevens AM, Shoemaker NB, and Salyers AA. (1990) *J Bacteriol* 172: 4271–4279.
6. Shoemaker NB, Getty CE, Salyers AA, and Gardner JF. (1985) *J Bacteriol* 162: 626–632.
7. Guiney D, Hasegawa P, and Davis C. (1984) *PNAS* 83: 7203–7206.
8. Comstock LE, Coyne MJ, Tzianabos AO, Pantosti A, Onderdonk AB, and Kasper DL. (1999) *Infect Immun* 67:3525–3532.
9. Smith CJ, Rogers MB, and McKee ML. (1992) *Plasmid* 27: 141–154.

Supplementary Materials and Methods

Bacterial Growth. *Bacteroides* strains were grown anaerobically on brain-heart infusion plates supplemented with hemin (5 μg/ml) and vitamin K1 (0.5 μg/ml) (BHIS). Gentamicin (200 μg/ml), erythromycin (5 μg/ml), and chloramphenicol (20 μg/ml) were added where appropriate.

Autoaggregation Assay. All strains of *B. fragilis* were grown at 37°C in BHIS laboratory media. The cells were adjusted in BHIS to an OD of 0.5 at 600 nm and were split into two tubes of equal volume. After incubation at 37°C for 2 h, 2 ml of the upper suspension was transferred from one tube, and the OD was measured at 600 nm; in addition, 2 ml of the total bacterial suspension was taken from the other tube after mixing, and the OD was measured. Aggregation was expressed as {1 - (OD upper suspension/OD total bacterial suspension)} × 100%.

PCRs. Primers used for all PCRs are listed in Supplementary Table 2.1.

Functional Promoter Analysis and XylE Assay. *xylE* reporter plasmid pLEC23 was used to assay for functional promoter activity (Krinos et al., 2001). Transcriptional fusions of all polysaccharide promoters locked in the "off" orientation (including a 300-bp upstream sequence) were amplified and cloned into the BamHI site of pLEC23 (primers listed in Supplementary Table 2.1). Promoter activity was examined by spectroscopy measuring the XylE-dependent colorimetric hydrolysis of the substrate catechol 2,3-dioxygenase.

PCR-Digestion Assay for PSB Promoter Orientation. To quantitatively determine the preferred orientation of the PSB promoter, PCR digestion assay was performed as described previously (Krinos et al., 2001). Briefly, PCR was performed with the primers B1 and B5, using chromosomal DNA from *B. fragilis* CPM1 and CPM3, which amplifies the PSB inverted region and some flanking DNA. The PCR product was digested with a restriction endonuclease (RsaI) that cleaves asymmetrically between the inverted repeat elements. The relative quantities of the resulting fragments directly correlate with the promoter orientations. These assays were performed a minimum of three times and yielded similar results.

Chapter 3:

Commensal colonization factors promote specificity and stability of the gut microbiome

S. Melanie Lee, Zbigniew Mikulski, Gregory P. Donaldson, Silva Boyajian, Klaus Ley, and Sarkis K. Mazmanian

This chapter has been submitted to *Nature*.

ABSTRACT

Mammals harbor a complex gut microbiome, comprised of bacteria that provide nutritional, metabolic and immunologic benefits during colonization (Ley et al., 2006). Despite recent advances in sequence-based microbiome profiling (The Human Microbiome Project Consortium, 2012), virtually nothing is known about the molecular processes employed by symbiotic bacteria to permanently colonize the gastrointestinal (GI) tract. The gut can harbor a tremendous diversity of microbes (Yatsunenko et al., 2012); however, we surprisingly find that mono-association of germ-free mice with single *Bacteroides* species saturates colonization. To identify bacterial mechanisms for species-specific saturable colonization, we devised an *in vivo* genetic screen and discovered a novel operon that is highly conserved among the intestinal *Bacteroides*, one of the most prominent gut bacterial genera of humans. We named this genetic locus the *c*ommensal *c*olonization *f*actors (*ccf*), which consist of transcriptional regulatory genes that control expression of a putative outer membrane protein complex. Deletion of the *ccf* genes in the model symbiont, *Bacteroides fragilis*, results in colonization defects in both germ-free and conventional mice. The *ccf* genes of *B. fragilis* are up-regulated during gut colonization, preferentially at the mucosal surface, supporting an *in vivo* function. Indeed, deletion of *ccf* genes leads to a defect in bacterial occupation of colonic crypts of mice and reduced mucosal association of mice. Remarkably, *B. fragilis* colonization of the gut after antibiotic depletion or gastroenteritis requires the CCF system, suggesting the niche within colonic crypts represents a reservoir for the gut microbiota following environmental stress. These findings reveal that intestinal *Bacteroides* have evolved

species-specific molecular interactions with the host that mediate stable and resilient gut colonization.

INTRODUCTION

Immediately upon birth, mammals enter an organized program of colonization to form complex microbial communities on environmentally exposed surfaces (Ley et al., 2006). Recent international microbiome sequencing initiatives are revealing detailed inventories of the partnerships we share with diverse bacteria, across various body sites and human populations (The Human Microbiome Project Consortium, 2012; Qin et al., 2010). We have learned that the relationships forged between humans and microbes over millennia create resilient microbiomes specific to individual hosts (Palmer et al., 2007; Dethlefsen et al., 2008). Accordingly, disturbances in the community composition of the gut microbiota (known as dysbiosis) have been linked to increased human disorders such as obesity, inflammatory bowel disease (IBD), malignancies and autism (Sekirov et al., 2010). The 'microbiota hypothesis' proposes that modern lifestyle changes (e.g., Western diets, widespread antibiotic use, Caesarean sections) have reduced our exposure to health-promoting symbionts while enriching for disease-causing pathobionts (Round and Mazmanian, 2009). Although microbiome profiling has revealed critical links between gut bacteria and health, discovery of the molecular forces controlling intestinal colonization may be critical for addressing disturbing trends in human disease based on dysbiosis.

RESULTS

Bacteroidetes is one of the most numerically abundant Gram-negative phyla in the mammalian gastrointestinal tract (Eckburg et al., 2005). *Bacteroides thetaiotaomicron* induces glycosylation of the intestinal epithelium to mediate architectural development of the gut (Bry et al., 1996), produces glycohydrolases that break down dietary carbohydrates for host nutrient utilization and promotes anti-microbial defenses to invading pathogens (Cash et al., 2006). *Bacteroides fragilis* is a gut mutualist that aids in host health by directing the cellular maturation of the developing immune system of the host (Mazmanian et al., 2005) and by protecting animals from inflammation in experimental models of IBD and multiple sclerosis (Mazmanian et al., 2008; Round and Mazmanian, 2010; Ochoa-Reparaz et al., 2010). Colonization by specific commensal bacteria may confer health in humans, and microbiomes with increased *Bacteroidetes* appear to be protected from some diseases (Takaishi et al., 2008; Turnbaugh et al., 2009; Penders et al., 2007). However, microbial processes by which these (or any) prominent human symbionts mediate life-long colonization of the gut remain unknown.

To explore the mechanisms and dynamics of host-microbial symbiosis, we simulated microbiome assembly by sequentially introducing various *Bacteroides* species to germ-free mice, and monitoring colonization over time. Animals were readily colonized with *B. fragilis* followed by *B. thetaiotaomicron* (Figure 3.1A) or *Bacteroides vulgatus* (Figure 3.1B), and altering the sequence of microbial exposure did not affect results (Figure 3.1C). Remarkably however, animals colonized with *B. fragilis* and subsequently exposed to the same species (marked by an antibiotic resistance gene) were resistant to super-colonization, with the challenging strain never reaching maximum levels and being

cleared over several days (Figure 3.1D). This property of 'colonization resistance' following mono-association was conserved in three other *Bacteroides* species examined (Figures 3.1E, Supplementary Figure 3.1), but not *Escherichia coli* (Supplementary Figure 3.2A). As conventional mice typically harbor 10^{11}-10^{12} colony-forming units (CFU) per gram of cecal content (Ley et al., 2006) (100-fold greater than *Bacteroides* in mono-association), there appears to be no limitation for space or nutrients under these conditions. We thus hypothesized that each individual *Bacteroides* colonizes the gut via saturating a finite quantity of unknown host products (which could represent receptors, nutrients, *etc*.). If this were the case, we speculated that displacing the existing strain would open 'niches' for the challenge strain; indeed, treatment of *B. fragilis* mono-associated mice with erythromycin promoted sustained colonization by the erythromycin-resistant challenge strain (Figure 3.1F). These data suggest that *Bacteroides* colonize the gut in a species-specific and saturable manner, suggesting a highly co-evolved mechanism for symbiosis between mammals and their microbiota.

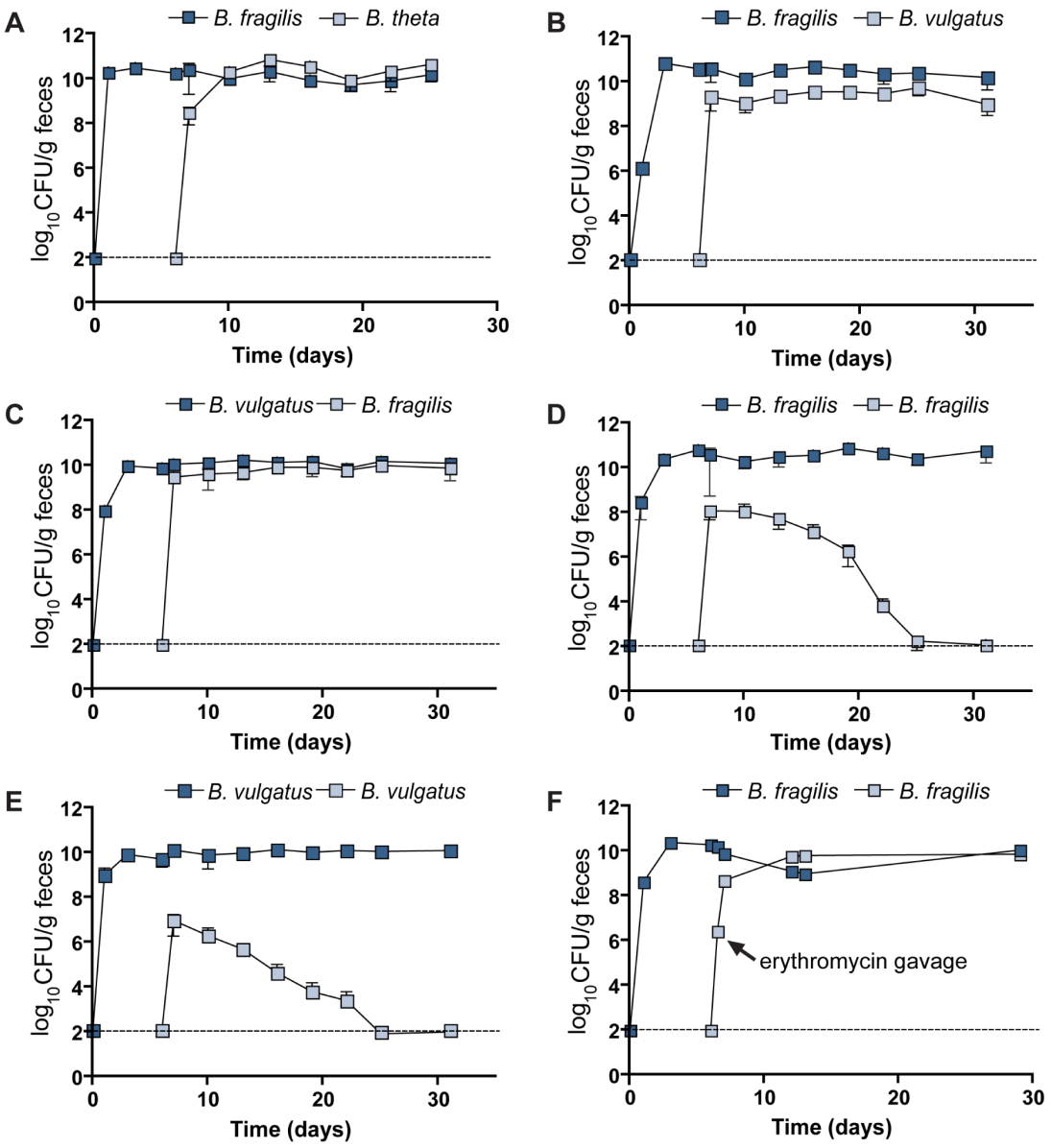


Figure 3.1 | Many *Bacteroides* species occupy finite and species-specific niche in the host intestine. Bacterial colonization level measured by CFU/g of feces over time. Germ-free Swiss Webster mice are mono-associated with *B. fragilis* for 6 days and subsequently challenged orally with ~ 10^8 CFU of (A) *B. thetaiotaomicron*; (B) *B. vulgatus*; (D) *B. fragilis*. Germ-free Swiss Webster mice are mono-associated with *B. vulgatus* for 6 days and subsequently challenged orally with ~ 10^8 CFU of (C) *B. fragilis*; (E) *B. vulgatus*. Dashed line indicates the limit of detection at 100 CFU/g feces. Results are representative of at least 2 independent trials (n=3 animals/group). (F) Germ-free mouse mono-associated with erythromycin sensitive *B. fragilis* is orally challenged with ~ 10^8 CFU of erythromycin-resistant *B. fragilis.* Erythromycin treatment was administered orally by gavage (0.5 mg) and in drinking water (10 µg/ml) 12 hours after challenge. Results are representative of 3 independent trials (n=1 animal/group).

We sought to determine the molecules required for saturable-niche colonization. We have previously shown that a *B. fragilis* mutant that could not phase vary its multiple capsular polysaccharides was defective for establishment of colonization (Liu et al., 2008). However, mice mono-associated with the capsular polysaccharide mutant appeared to be saturated for the *B. fragilis* niche similar to wild-type bacteria (Supplementary Figure 3.2B). Based on our discovery of self-exclusion by *Bacteroides,* we devised a functional *in vivo* screen to identify genetic factor(s) sufficient to mediate species-specific colonization. Mice were mono-associated with *B. vulgatus*, then challenged with *B. vulgatus* clones that contained a genomic library from *B. fragilis* (the scheme is depicted in Figure 3.2A). We reasoned that only those clones containing genes that conferred colonization by *B. fragilis* would persist, with the remainder being cleared from the gut. We screened 2,100 clones (in pools of 96 per mouse) each containing 9-10 kilobases of DNA, providing a 3.8-fold coverage of the *B. fragilis* genome and 98% probability that a given DNA sequence is present in the library (see Equation 3.1). Remarkably, only two clones sustained colonization in animals and the rest were cleared. Sequencing the two isolates mapped to the identical genomic locus of *B. fragilis* with the minimal genetic element common to both DNA fragments containing five novel hypothetical open reading frames: BF3583, BF3582, BF3581, BF3580 and BF3579 (Figure 3.2B).

$$N = \frac{\ln(1-P)}{\ln(1-\frac{i}{G})} = \frac{\ln(1-P)}{\ln(1-\frac{9.5\times10^{3}}{5.2\times10^{6}})} = 2100$$

$$P = 0.979$$

Equation 3.1 | Optimal number of clones in the genomic library (Sambrook et al., 1989). P denotes the probability of isolating a particular DNA sequence. The total number of clones screened (N) is 2100; average insert size (i) is 9.5 kb and the *B. fragilis* genome size (G) is 5.2×10^{6} bp.

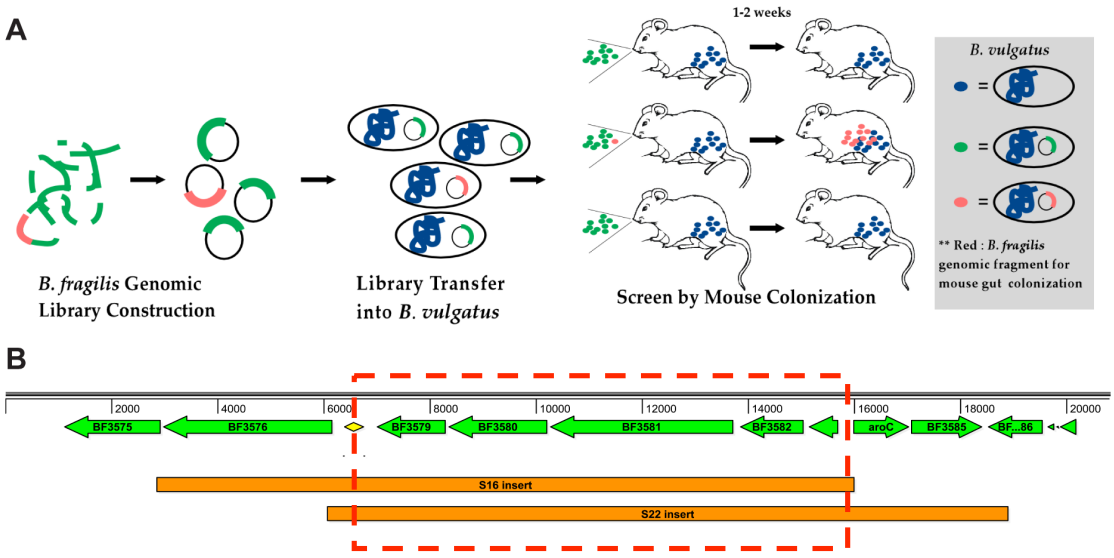


Figure 3.2 | (A) Schematic of functional *in vivo* screen of the *B. fragilis* genome for mouse gut colonization. 9-10 kb fragments of *B. fragilis* genomic DNA generated by partial digestion with Sau3AI are ligated into *E. coli-Bacteroides* shuttle plasmid pFD340-*cat*BII (Cm^{r}). Each individual clone is conjugally transferred into *B. vulgatus* generating a library of *B. vulgatus* carrying a unique *B. fragilis* genomic DNA fragment. The library is screened for *B. fragilis*-specific niche colonization phenotype in animals mono-associated with *B. vulgatus*. (B) The minimal genetic element common to the two clones (S16 and S22) that display colonization phenotype from the functional screen contains 5 previously unknown genes: BF3583, BF3582, BF3581, BF3580 and BF3579.

Based on the protein sequence, BF3583 and BF3582 show similarity to a sigma (σ) factor/anti-σ factor gene pair that may control transcription of three downstream genes, which appear to be in an operon according to several operon databases such as MicrobesOnline (Price et al., 2005) and DOOR (Database of prOkaryotic OpeRons)

(Dam et al., 2007; Mao et al., 2009) (Supplementary Figure 3.3). BF3581 sequence shows strong homology to the SusC family of outer membrane proteins. BF3580 is a SusD homolog, an outer membrane lipoprotein often paired with SusC. Together, the Sus-like system is believed to bind and import a range of sugar molecules (Koropatkin et al., 2012; Martens et al., 2009; Shipman et al., 2000; Schauer et al., 2008). BF3579 encodes for a putative chitobiase, suggesting a possible polysaccharide substrate for this system (Kawada et al., 2008). Comparative genomic analysis using the genes identified from the screen to query JGI's Integrated Microbial Genomes database (http://img.jgi.doe.gov/cgi-bin/w/main.cgi) revealed conservation of similar clusters of genes among sequenced *Bacteroidetes* species (Figure 3.3). Homologous loci were identified within the genomes of *B. fragilis*, *B. thetaiotaomicron, B. ovatus* and *B. vulgatus*, all having displayed the colonization resistance phenotype (Figure 3.4). Therefore, we have identified a novel pathway in *B. fragilis* sufficient for species-specific colonization; moreover, this process may be shared by many intestinal *Bacteroides*. The genes are named *ccfA-E,* for *c*ommensal *c*olonization *f*actors.

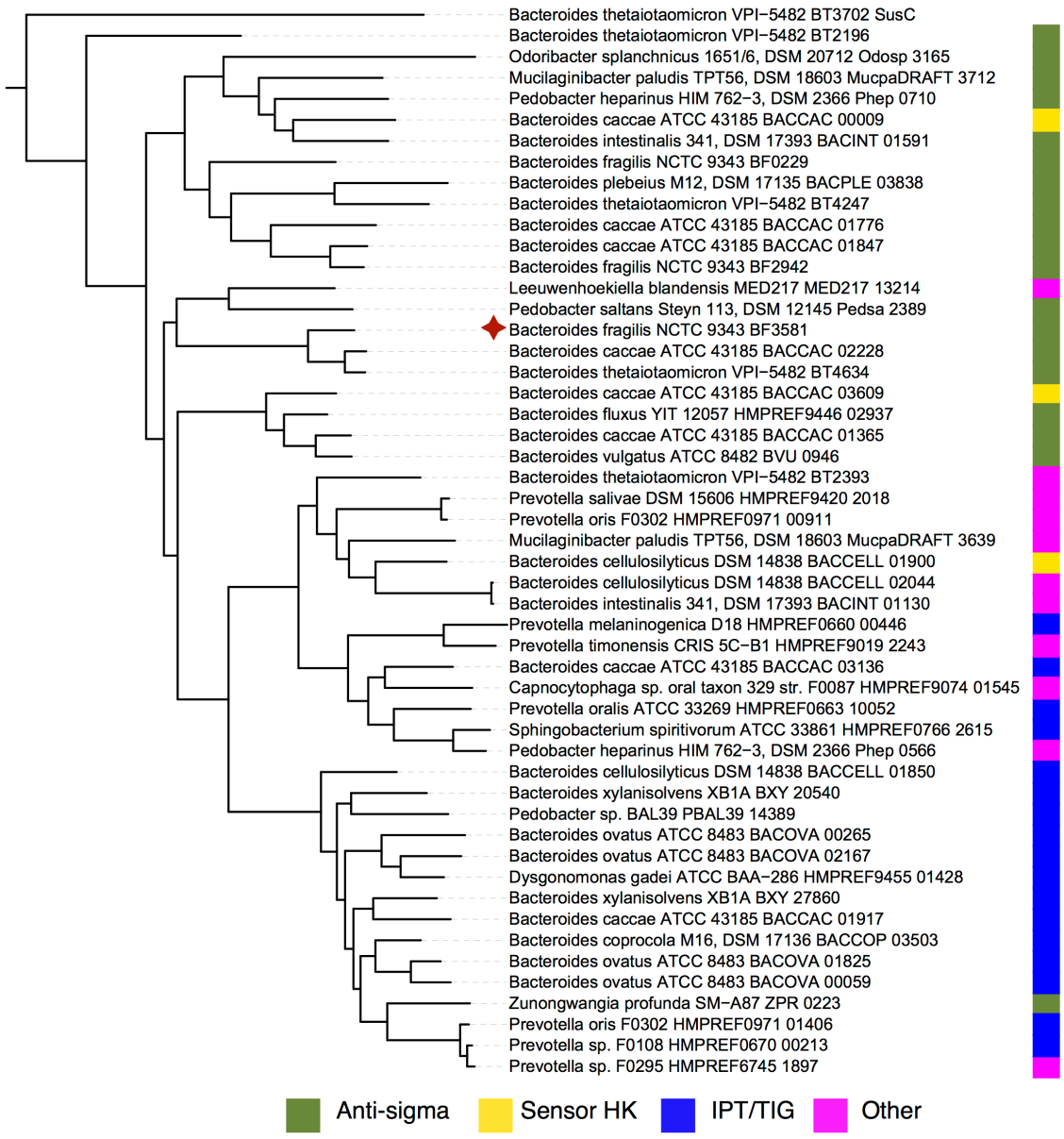


Figure 3.3 | Phylogenetic tree based on BF3581 and BF3580 homology in the phylum *Bacteroidetes*. Labels indicate the locus ID of the BF3581 homologues. Amino acid sequences of BF3581and BF3580 are used to query sequenced genomes using JGI's Integrated Microbial Genomes site (Markowitz et al., 2012). Genomic context adjacent to BF3581 and BF3580 homologues with expected values less than 10^{-10} are examined for operon organization resemblance to BF3579-BF3583. For inclusion in the tree, candidate operons are required to have a downstream BF3579 homologue with at least one of either the F5/F8 Type C domain (pfam00754) or the domain of unknown function, DUF1735 (pfam08522), which are present in BF3579. Also, candidate operons are required to have a BF3582 homologue. These include anti-sigma factors, IPT/TIG domain-containing transmembrane proteins (pfam01833), sensor histidine kinases, and transmembrane 6-bladed NHL-repeat beta-propellers (pfam01436). The legend refers to the type of transmembrane regulatory domain found upstream of the BF3581 homologue. The *sus* operon (BT3702) from *B. thetaiotaomicron* is included as an outgroup.

Operons from redundant strains of single species and the 22 draft genomes of unnamed *Bacteroides* sp. Isolates are not included. Amino acid sequences of BF3581 and BF3580 homologues are concatenated and locally aligned using MUSCLE (Edgar, 2004). The tree is constructed using PhyML (Guindon et al., 2010) and displayed using iTOL (Letunic and Bork, 2007).

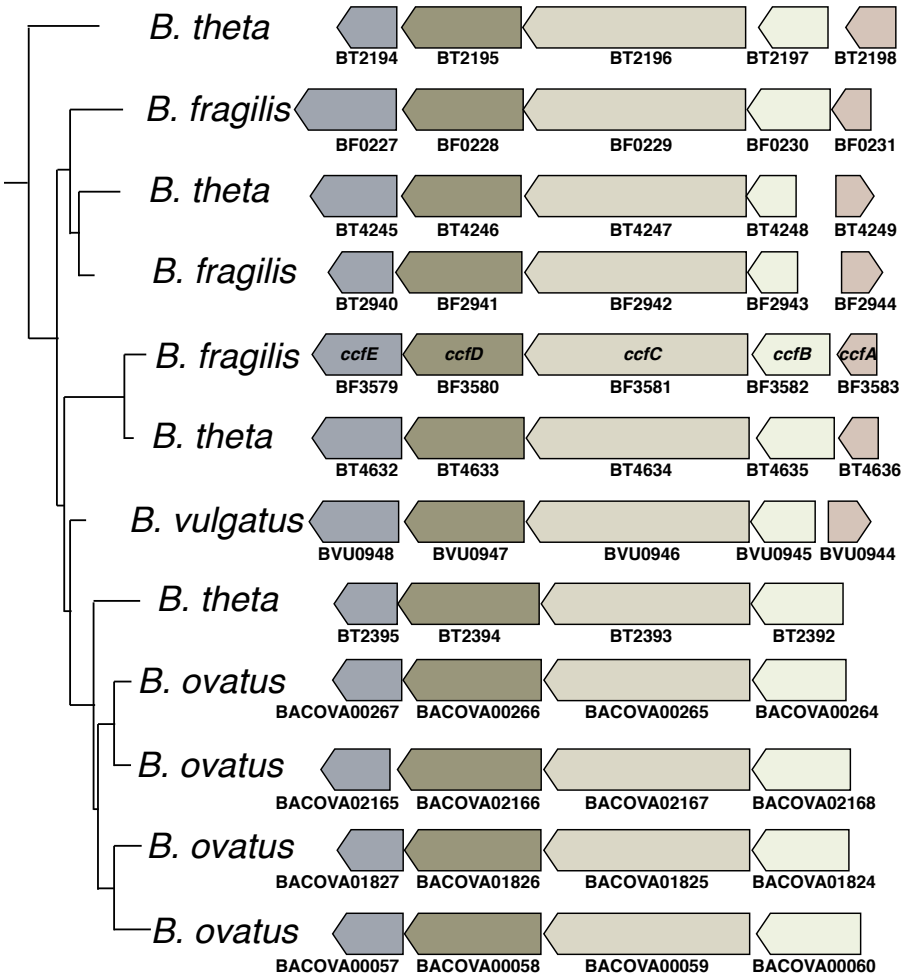


Figure 3.4 | Comparative genomic analysis for homologues of the operon identified from the *in vivo* screen. Homologous operons in *B. fragilis*, *B. thetaiotaomicron*, *B. ovatus*, and *B. vulgatus* were selected from a phylogenetic tree constructed by querying the genomes in the phylum *Bacteroidetes* (Figure 3.3).

To test if the putative structural genes (*ccfC-E*) are required for gut colonization, we generated in-frame deletion mutants of *B. fragilis*; Δ*ccfC*, Δ*ccfD* and Δ*ccfE*. All strains exhibited normal morphology on solid agar medium and unimpaired growth in laboratory culture (data not shown). Germ-free mice were associated with individual

mutant strains, and subsequently challenged with WT *B. fragilis*. As previously shown, animals mono-colonized with WT *B. fragilis* completely cleared the challenge strain after 30 days (Figure 3.5; 1st bars). However, animals mono-associated with Δ*ccfC* or Δ*ccfD* were permissive to colonization by WT bacteria (Figure 3.5; 2nd and 3rd bars), unlike the *ccfE* mutant (Figure 3.5; 4th bars). A deletion mutant in all three genes (*B. fragilis*ΔCCF) also allowed WT *B. fragilis* to colonize (Figure 3.5; 5th bars). Similarly, *ccfC-E* mutant in *B. vulgatus* (ΔBVU0946-BVU0948) also permitted WT *B. vulgatus* colonization (Supplementary Figure 3.4; 1st and 2nd bars), demonstrating this system may be universal to this genus. When *B. fragilis*ΔCCF mono-associated animals were challenged with the same mutant strain, the challenging bacteria were cleared (Figure 3.5; 6th bars). As there are multiple similar gene clusters in *B. fragilis*, this finding suggests that the CCF system we identified is necessary for colonization resistance, but redundant mechanisms may be mediated by the other *ccf* gene homologues in its absence (Figure 3.4). *B. vulgatus*ΔCCF strain also prevented further colonization by the same mutant strain (Supplementary Figure 3.4; 3rd bar), again suggesting functional redundancy during colonization. This interpretation is consistent with the fact that challenge of *B. fragilis*Δ*ccfC* or Δ*ccfD* with WT bacteria led to colonization that was two logs lower than mono-association (Figure 3.5; 2nd and 3rd bars). Further, mono-colonization with a *B. vulgatus* strain expressing the *B. fragilis ccf* genes (*ccfA-E*) was unable to saturate *B. fragilis*-specific niche in the gut and permitted WT *B. fragilis* colonization (Supplementary Figure 3.4; 5th bar). These findings suggest that a single *ccf* locus is responsible for partial saturation of a broader species-specific niche (which could be host receptors, nutrients or spatial access to a

niche, etc.). Based on its requirement for saturable colonization, we focused further studies on the primary *ccf* operon identified from the functional screen.

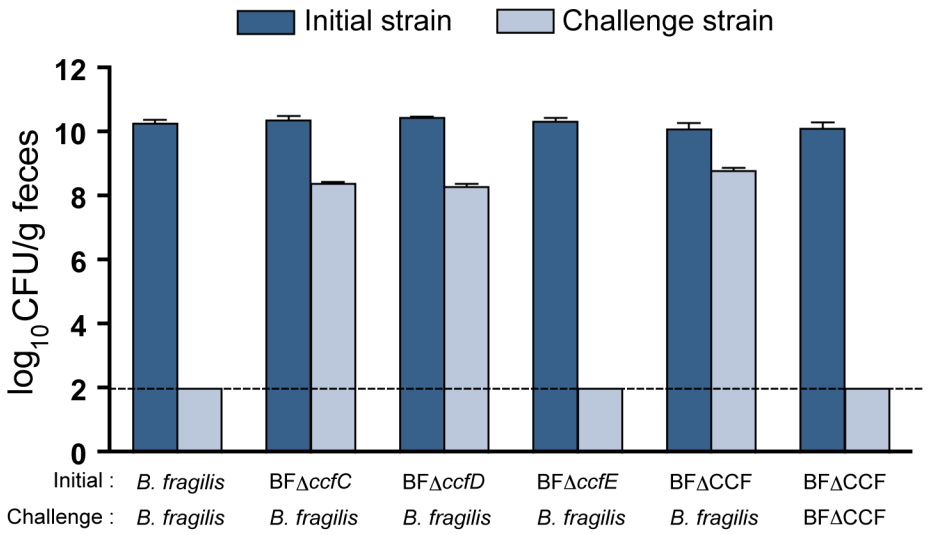


Figure 3.5 | The *ccf* genes mediate saturable-niche colonization by *B. fragilis*. Germ-free Swiss Webster mice are mono-associated with WT or mutant strains deleted in the putative structural *ccf* gene(s) for 6–7 days and subsequently challenged orally with ~ 10^8 CFU of WT *B. fragilis* or mutant strain deleted in *ccfC-E* genes (BFΔCCF). 30 days after challenge, initial and challenge strains are quantified. Dashed line indicates the limit of detection at 100 CFU/g feces (n=3 animals/group).

Virulence factors of pathogens are expressed upon animal infection. We speculated that the *ccf* genes may be induced in bacteria during animal colonization. Based on the previous observation that a population of *B. fragilis* associates with mucosal tissues (Round et al., 2011), we show that *ccfB-E* are preferentially expressed by bacteria in contact with the colon, with lower levels in luminal cecal content and feces (Figure 3.6A). There was virtually no expression in laboratory culture, suggesting a role for the host factor(s) to activate the CCF system. *ccfA* and *ccfB* have strong homology to a bacterial σ factor/anti-σ factor pair, which regulate transcription by controlling RNA polymerase recruitment to specific promoters. In-frame deletion of *ccfA* led to highly

reduced expression of all genes in the operon during animal colonization (Figure 3.6B). Accordingly, mono-colonization of germ-free mice with Δ*ccfA* mutant was permissive for challenge with WT bacteria, demonstrating a functional defect in the saturable-niche occupancy (Figure 3.6C). Interestingly, an in-frame deletion mutation in *ccfB* also resulted in attenuated gene expression throughout the operon (Supplementary Figure 3.5A), and a defect in saturable colonization of the gut (Supplementary Figure 3.5B). Thus, *in vivo* expression of the *ccf* operon near gut tissue may be critical for colonization fitness. If true, then challenging WT *B. fragilis* mono-associated mice with WT *B. fragilis* harvested from the gut of donor mice may enhance colonization. In contrast to laboratory grown bacteria (see Figure 3.1D), sustained colonization was conferred to bacteria recovered directly from animals (Figure 3.6D). These data suggest that sensing of host factors (possibly near the gut epithelium) activates transcription of the *ccf* genes and may promote persistent host association.

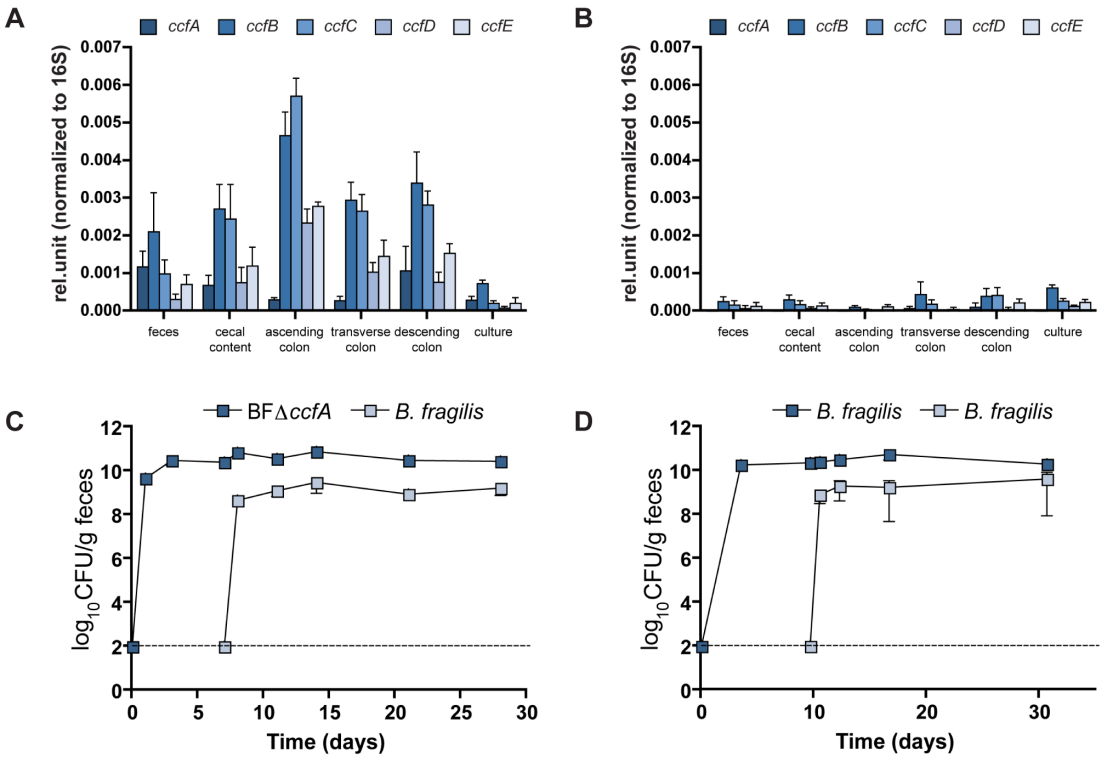


Figure 3.6 | *ccf* genes are preferentially expressed at the mucosal surface of the gut. (A) qRT-PCR of the *ccf* gene expression levels normalized against 16S rRNA. Total RNA is extracted from feces, cecal content and colon tissues of the animals mono-associated with WT *B. fragilis* for 2 weeks (n=3 animals/group) and from laboratory culture. (B) qRT-PCR of the *ccf* gene expression levels from BFΔ*ccfA* mono-associated animal feces, cecal content and colon tissues (n=3 animals/group) and culture. (C) Fecal bacterial colonization levels measured from germ-free mice mono-associated with BFΔ*ccfA* and subsequently challenged with ~ 10^8 CFU of WT *B. fragilis*. (D) Fecal bacterial colonization level of germ-free mice mono-associated with *B. fragilis* and subsequently challenged orally with ~ 10^8 CFU of *B. fragilis* harvested from cecal content of another animal (n=3 animals/group, 1 donor).

Increased expression of the *ccf* genes near colon tissue prompted us to investigate whether *B. fragilis*ΔCCF displays a defect in mucosal association. Since each strain can mono-colonize germ-free mice equally, we used a competitive colonization approach to examine if the CCF system promotes bacterial localization to colonic tissue. Mice were mono-associated with either the WT or the *ccf* deletion strain, and both groups were subsequently challenged with WT *B. fragilis*. 24 hours after challenge, we observed the same relative numbers for challenge strains in feces of both groups (Figure 3.7A, B & C;

1st and 2nd bars). However, colon tissue-associated bacterial levels of the challenge strain were higher in animals pre-associated with the *ccf* mutant (Figure 3.7A, B & C; 3rd and 4th bars). An increased tissue association in animals pre-colonized with the *ccf* mutant compared to WT bacteria suggests a colonization defect by the mutant strain specifically at the mucosal surface, but not in the lumen. Since luminal contents are constantly being expelled from the gut, we speculate that the CCF system mediates long-term colonization by localizing *B. fragilis* to the colonic epithelium.

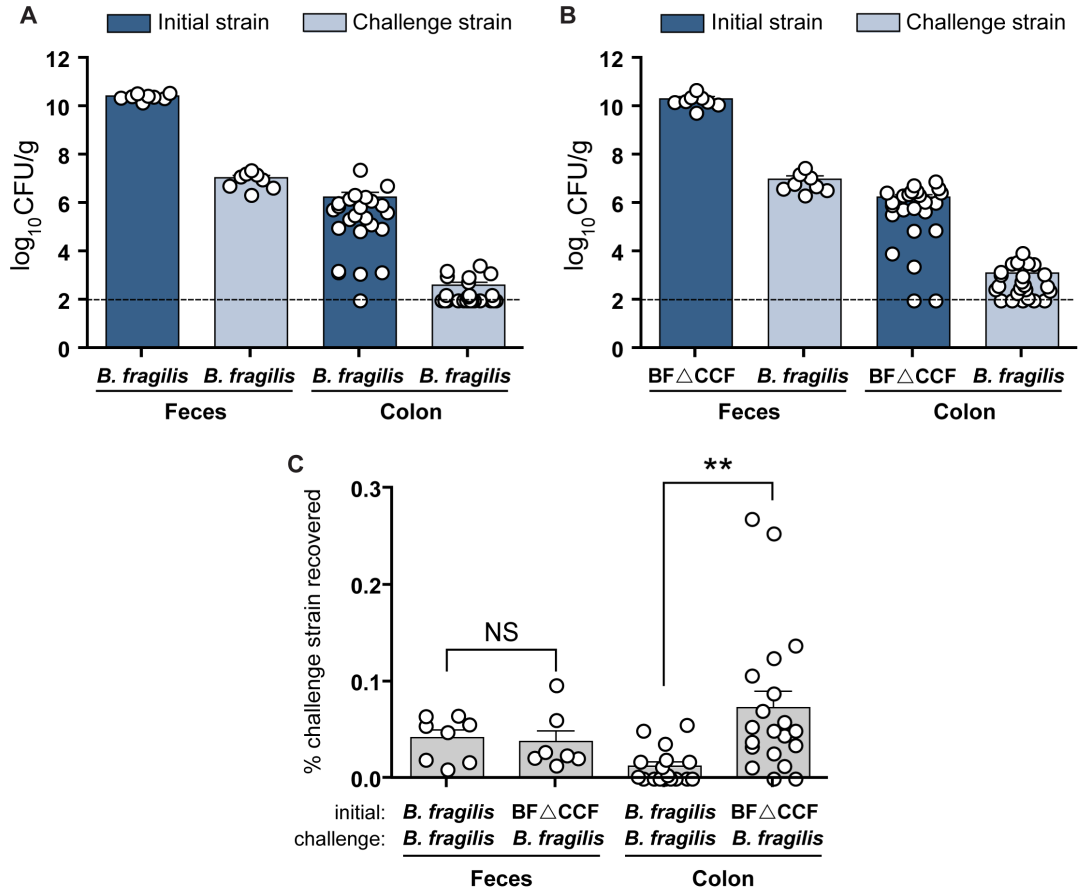


Figure 3.7 | Germ-free Swiss Webster mice are mono-associated with either (A) WT *B. fragilis* or (B) *B. fragilis* ΔCCF for 7 days and orally challenged with 10^8 CFU of laboratory culture WT *B. fragilis*. 1 day after challenge, feces and colon tissues (1 cm representative piece from ascending, transverse and descending colon) are harvested, mucus layer scraped off from the tissue and homogenized in PBS and plated on selective media for CFU determination (A, B). (C) Percentage of challenge strain from total bacteria is calculated by dividing the challenge strain counts from (A) and (B) with the sum of initial and challenge strain counts for each sample (n=8 animals/group). ** indicates p value less than 0.01.

Symbiotic gut bacteria were historically believed to reside exclusively in the GI lumen and not make intimate contact with host tissue (Hooper, 2009). However, we have shown that *B. fragilis* occupies the colonic crypts of mono-associated mice (Round et al., 2011) and recent studies have revealed numerous species colonize intestinal crypts of SPF mice (Pedron et al., 2012). Increased production of the *ccf* genes near mucosal tissue prompted us to explore the intriguing hypothesis that the CCF system mediates crypt occupancy by *Bacteroides*. We mono-colonized mice with *B. fragilis* and visualized bacterial localization in colon tissue using whole-mount confocal microscopy. A close look at a single crypt by analyzing z-stack images revealed a distinct cluster of WT *B. fragilis* deep within the channel of the crypt (Figure 3.8A). When WT *B. fragilis* mono-associated animal colon was compared to that of *B. fragilis*ΔCCF mono-associated mice, we saw drastically fewer crypt occupancies by the mutant bacteria (Figure 3.8B), suggesting that CCF system mediates specific crypt colonization by the bacteria. Two-photon imaging of colon explants clearly demonstrated presence of WT *B. fragilis* on the surface of the epithelium and inside the crypt (Figure 3.9A). While both strains of *B. fragilis* associate with the surface of the epithelium, only WT bacteria are able to penetrate deep into the colonic crypts of mice (Figure 3.9A). Measuring the distance from the apical surface of the epithelium to the fluorescent signal revealed significantly greater penetration from the tissue surface by WT bacteria (Figure 3.9B). Collectively, these data suggest that the CCF system allows *B. fragilis* to reside in a specific niche within crypts during symbiotic colonization.

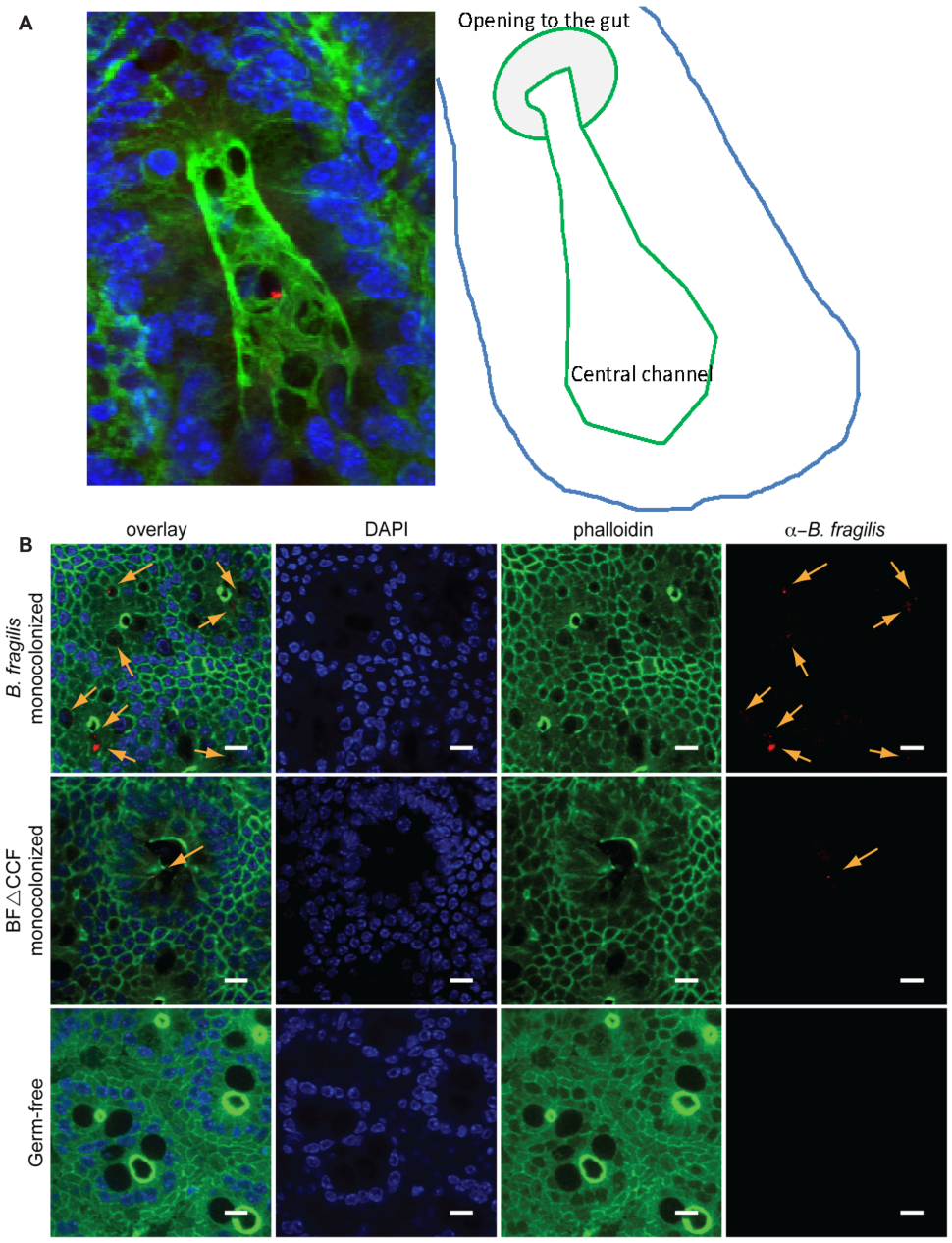


Figure 3.8 | *B. fragilis* colonization of the colon crypt is mediated by CCF. (A) 4% PFA fixed whole-mount colon tissues from WT *B. fragilis* mono-associated mice. The colon crypt is visualized by nuclear counterstaining with DAPI (blue) and F-actin counterstaining with phalloidin (green). Bacteria are visualized by IgY polyclonal antibody raised against *B. fragilis* (red). A z stack is reconstructed and bacteria harboring crypt is visualized from the side-view with a cartoon schematic on the right. (B) Confocal micrographs of 4% PFA fixed whole-mount colon tissues from germ-free, WT *B. fragilis* mono-associated or *B. fragilis*ΔCCF mono-associated mice stained for bacteria (red, arrows) with DAPI (blue) and phalloidin (green) counterstaining. Images are representative of seven different sites analyzed from at least two different colons. Scale bar: 10 μm.

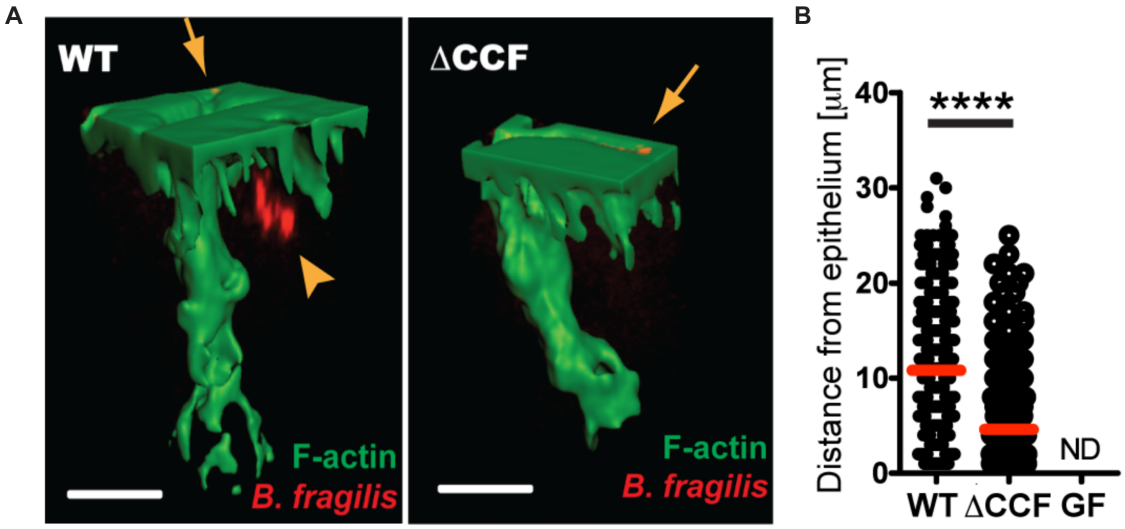


Figure 3.9 | CCF mediates bacterial occupancy of a deep crypt space. (A) 3D reconstructions of *B. fragilis* mono-associated mouse colon crypts using two-photon microscopy. Fluorescence signals are detected on the apical surface of the epithelium (arrows) for both strains and in the crypt space (arrowhead) by WT bacteria. Scale bar: 10 μm. (B) Quantification of *B. fragilis* localization in the crypt space measured by the distance of the immune-reactivity signal from the epithelial surface. GF: germ-free mice. ND: not detected. **** indicates p value less than 0.0001.

We next investigated the effects of the CCF system in the context of a complex microbiota, as germ-free mice represent an unnatural microbial status. *Bacteroides* species do not readily colonize most strains of specific pathogen-free (SPF) mice, namely BALB/c, Swiss Webster and C57BL/6, despite oral administration of high inoculums (Supplementary Figure 3.6A, C and data not shown). Furthermore, challenge of mono-colonized mice with a complex mouse microbiota led to clearance of the WT. *B. fragilis* (Supplementary Figure 3.6B, D). As we previously observed empirically that *B. fragilis* can colonize SPF Rag-/- mice, which lack mature B and T lymphocytes, and SPF non-obese diabetic (NOD) mice, immunocompetent animal model for type 1 diabetes, we introduced either WT or *ccf* mutant *B. fragilis* at equal inoculums into separate groups of animals. Only WT *B. fragilis* stably colonized Rag-/- mice, whereas *B. fragilis*ΔCCF established a significantly lower colonization in the gut (Figure 3.10A). Co-inoculation of

equal numbers of WT and *ccf* mutant bacteria into Rag-/- mice also resulted in rapid clearance of the mutant strain from the gut (Figure 3.10B), demonstrating a cell intrinsic defect that could not be complemented in trans by WT bacteria. NOD animals were also preferentially colonized by WT *B. fragilis* compared to *ccf* mutant strain, when given to separate groups of animals (Figure 3.10C) or together in equal amount (Figure 3.10D). These data show that deletion of the *ccf* genes severely compromises *B. fragilis* colonization of mice with a complex microbiota, independent of the adaptive immune system. A profiling study comparing Rag-/-, NOD and C57BL/6 SPF microbiota, though beyond the scope of this study, may reveal specific microbes absent in Rag-/- or NOD mice that directly compete with *B. fragilis* for saturable colonization.

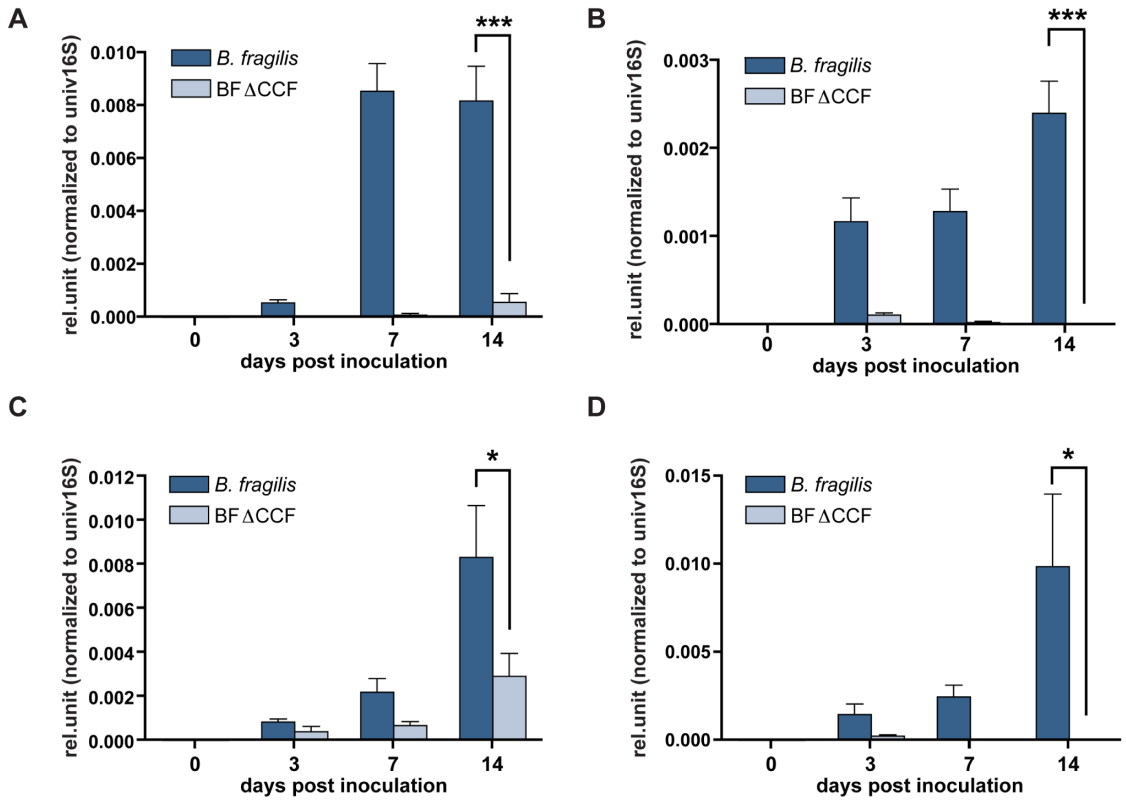


Figure 3.10 | *B. fragilis* colonization of the animal harboring a complex microbiota is attenuated in the absence of CCF. (A) SPF Rag-/- mice are given either WT *B. fragilis* or *B. fragilis*ΔCCF by single oral gavage and the colonization level assessed by qPCR (n=4 animals/group). *** indicates p value less than 0.001. (B) SPF Rag-/- mice are given equal dose

of WT *B. fragilis* and *B. fragilis*ΔCCF by single oral gavage and the colonization level assessed by qPCR (n=4 animals/group). *** indicates p value less than 0.001. (C) SPF NOD mice are given either WT *B. fragilis* or *B. fragilis*ΔCCF by single oral gavage and the colonization level assessed by qPCR (n=4 animals/group). * indicates p values less than 0.05. (D) SPF NOD mice are given equal dose of WT *B. fragilis* and *B. fragilis*ΔCCF by single oral gavage and the colonization level assessed by qPCR (n=4 animals/group). * indicates p values less than 0.05.

During symbiosis with mammals, the microbiota may be confronted by rapid environmental changes with potentially adverse consequences to bacteria, such as antibiotic intake, dietary changes, gastrointestinal infection/inflammation, *etc*. We speculate that *Bacteroides* species evolved the *ccf* genes to promote long-term host colonization. Gastroenteritis is commonly experienced by most humans multiple times throughout life, and is known to perturb the microbiota (Sekirov and Finlay, 2009). To test if resiliency of *B. fragilis* colonization is CCF-dependent, we used *Citrobacter rodentium* infection of SPF mice to mimic human gastrointestinal infection (Schauer and Falkow, 1993; Mundy et al., 2005). Using an antibiotic treatment protocol that does not sterilize the gut but promote colonization of SPF mice by *Bacteroides* (Bloom et al., 2011), we were able to simultaneously colonize mice with equivalent levels of WT and *ccf* mutant bacteria. Mice were subsequently infected orally with *C. rodentium*, and colonization of *B. fragilis* was monitored. Consequently, WT bacteria declined in number following the 2-week course of infection, but returned to maximal levels 3-4 weeks post infection (Figure 3.11A). Importantly, the *B. fragilis*ΔCCF strain was completely cleared from the mouse gut following gastroenteritis but not when animals were given PBS as control (Figure 3.11A, B). These results reveal that the CCF system establishes colonization resiliency by gut *Bacteroides* following disruption of the microbiome, representing the seminal example for a bacterial process that mediates stable host-microbial association.

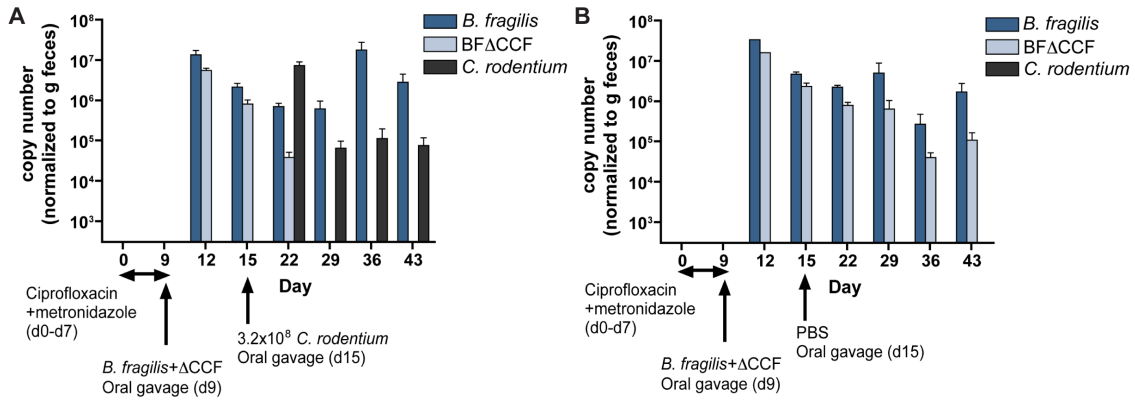


Figure 3.11 | The CCF system mediates *B. fragilis* persistence after perturbation of microbiota. (A) Real-time qPCR analysis of bacterial colonization levels in feces from SPF SW mice that were antibiotic treated to allow co-association with WT *B. fragilis* and *B. fragilis*ΔCCF. After 6 days, mice were infected with *Citrobacter rodentium* (n=4 animals/group). Results are representative of 2 independent experiments. (B) Real-time qPCR analysis of bacterial colonization levels in feces from SPF SW mice that were antibiotic treated and co-associated with WT *B. fragilis* and *B. fragilis*ΔCCF. After 6 days, mice were given PBS (n=4 animals/group). Results are representative of 2 independent experiments. All error bars indicate standard error of the mean (SEM).

DISCUSSION

Evolution has bound microbes and man in an inextricable relationship, and recent evidence reveals specificity and stability of the human microbiome (The Human Microbiome Project Consortium, 2012; Yatsunenko et al., 2012). However, the molecular mechanism(s) by which stable microbiomes assemble are unknown. Our findings reveal a novel, evolutionarily conserved outer membrane complex required for persistent colonization of the mammalian gut by the numerically prominent *Bacteroides*. Homology to the Sus family of proteins suggests a role for CCF in uptake and utilization of glycans. *Sus* genes represent the archetype for sugar acquisitions mechanisms by intestinal bacteria, and are widely duplicated in gut *Bacteroides* (88 gene clusters in *B. thetaiotaomicron*, encompassing 18% of the genome). Although Sus systems in *Bacteroides* mediate foraging of host mucus as an energy source (Koropatkin et al., 2012), their contribution to microbial colonization have not been investigated. Our discovery of CCF-dependent saturable colonization within, but not between, species suggests a model whereby *Bacteroides* evolved specific Sus-like systems to utilize unique and limiting sugars for association with the host. Future work will aim to identify the substrates for various CCF systems. Based on the findings that *ccf* genes are preferentially expressed in proximity to mucosal tissues and *B. fragilis* associates with colonic crypts, we find it likely that host factors (and not dietary substrates) may induce expression of the CCF system. Perhaps unique glycan structures found within colonic crypts serve as a carbon source for growth and/or adherence to host tissue. Finally, we propose that the *ccf* genes evolved to promote stable and resilient colonization, as crypt-associated bacterial reservoirs may represent ‘founder’ cells that repopulate the gut

following disruption of the microbiome by antibiotics or enteric infections. Understanding the evolutionary forces that mediate commensal colonization may advance therapies to correct dysbiosis, providing potentially novel therapies for numerous human diseases.

MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions.

Bacterial strains and plasmids are described in Supplementary Table 3.1. *Bacteroides* strains were grown anaerobically at 37°C for two days in brain-heart infusion broth supplemented with 5 μg/ml hemin and 0.5 μg/ml vitamin K (BHIS), with gentamicin (200 μg/ml), erythromycin (5 μg/ml), chloramphenicol (10 μg/ml) and tetracycline (2 μg/ml) added where appropriate. *Escherichia coli* JM109 containing recombinant plasmids were grown in LB with ampicillin (100 μg/ml) or kanamycin (30 μg/ml). *Citrobacter rodentium* DBS100 strain was grown in LB at 37°C for 24 hours.

Mice.

Germ-free Swiss Webster mice were purchased from Taconic Farms (Germantown, NY) and bred in flexible film isolators. For gnotobiotic colonization experiments, germ-free mice were transferred to freshly autoclaved microisolator cages with sterile food and bedding, supplemented with 10 μg/ml of erythromycin and 100 μg/ml of gentamicin in drinking water. Male SPF (Specific Pathogen-Free) C57BL/6 mice and Swiss Webster mice were purchased from Taconic Farms. Male SPF NOD/ShiLtJ mice and Rag-/- C57BL/6 mice were purchased from the Jackson Laboratory. All procedures were performed in accordance with the approved protocols using IACUC guidelines of the California Institute of Technology.

Construction of chromosomal library and screen.

Genomic DNA was isolated from overnight culture of *B. fragilis* using a commercial kit (Wizard® Genomic DNA Purification Kit, Promega). 20 μg of genomic DNA was incubated with 4U of Sau3AI for 5, 10, 15, or 20 minutes at 37°C in 50 μl volume and the partially digested genomic DNA was separated by electrophoresis on 0.7% agarose gel. 9–10 kb fragment DNA was excised and recovered from the agarose gel (Zymoclean™ Gel DNA Recovery Kit, Zymo Research). Insert DNA was ligated to *Bgl*II site of plasmid vector (pFD340-*cat*BII, Supplementary Table 4.1), transformed into *E. coli* and amplified on LB-ampicillin plate. Individual clones from the plasmid library were mobilized from *E. coli* to *B. vulgatus* by conjugal helper plasmid RK231 generating a library of *B. vulgatus* hosting *B. fragilis* chromosomal DNA fragments consisting of approximately ~ 2100 clones. To screen the library *in vivo*, pools of 96 clones (10^6 CFU of each clone) were gavaged into 22 germ-free Swiss Webster mice (10^8 CFU per animal) pre-colonized with *B. vulgatus* pFD340 for 1–2 weeks. Two weeks after gavage, fresh fecal samples were plated on BHIS agar plate containing chloramphenicol to select for clones with colonization phenotype.

Generation of *ccfA, ccfB, ccfC, ccfD, ccfE* and *ccfC-E* (ΔCCF) deletion mutants.

~ 2 kb DNA segments flanking the region to be deleted were PCR amplified using primers listed in Supplementary Table 3.2. Reverse primer of the left flanking DNA and forward primer of the right flanking DNA were designed to be partially complementary at their 5' ends by 18–21 bp. Fusion PCR was performed using the left and right flanking DNA (~ 300 ng each after gel purification) as DNA template and forward primer of the left flanking DNA and reverse primer of the right flanking DNA (Wurch et al., 1998).

The fused PCR product was cloned into *Bam*HI or *Sal*I site of the *Bacteroides* conjugal suicide vector pNJR6 and mobilized into *B. fragilis*. Colonies selected for erythromycin resistance (Em^r), indicating integration of the suicide vector into the host chromosome were passaged for five days and then plated on nonselective medium (BHIS). The resulting colonies were replica plated to BHIS containing Em, and Em^s (erythromycin sensitive) colonies were screened by PCR to distinguish wild-type revertants from strains with the desired mutation. The same strategy was employed to generate Δ*ccfC-E* deletion mutant from *B. vulgatus*.

Quantitative RT-PCR.

Total RNA was extracted from mid-log-phase bacterial culture using ZR Fungal/Bacterial RNA MiniPrep™ (Zymo Research), feces and cecal content from mice using ZR Soil/Fecal RNA MicroPrep™ (Zymo Research), and mouse colon tissues after removing luminal content by gently scraping the mucosal surface and PBS rinse using Trizol (Invitrogen). cDNA was made using an iSCRIPT cDNA synthesis kit per manufacturer's instructions (Bio-Rad). All q-PCR reactions were performed in ABI PRISM 7900HT Fast Real-Time PCR System (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems). Gene-specific primers are described in Supplementary Table 3.2.

Gnotobiotic animal colonization experiments.

8–12 week old germ-free Swiss Webster mice were either mono-associated or sequentially bi-associated with *Bacteroides* strains by oral gavage (~ 10^8 CFU of each

bacterial strain harvested from a log-phase culture and resuspended in PBS with 1.5% $NaHCO_3$). All *Bacteroides* strains used to colonize germ-free animals were resistant to gentamicin inherently, and to erythromycin by plasmid. Unless otherwise indicated, the initial strains carried pFD340-*cat* (chloramphenicol resistant; Cm^r) and the challenge strains, pFD340-*tetQ* (tetracycline resistant; Tet^r). At each time point, fresh fecal samples were collected, homogenized and serially diluted in PBS and plated on selective media for CFU per g of feces.

Colon whole-mount immunofluorescence experiments.

Colons were fixed in buffered 4% paraformaldehyde, washed with PBS and subjected to indirect immunofluorescence. Tissues were made permeable by incubation with 0.5% (wt/vol) saponin, 2% (vol/vol) FBS, and 0.09% (wt/vol) azide in PBS for at least 18 hours. The same buffer was used for subsequent incubations with antibodies. Colon fragments (2 mm x 2 mm) were incubated with a primary polyclonal chicken IgY anti-*B. fragilis* antibodies for 12–16 hours at room temperature followed by 1–2 hour incubation at 37°C. Following PBS washes, samples were reacted with goat anti-chicken IgY secondary antibodies (Alexa Fluor 488 or Alexa Fluor 633, 2 μg/ml, Molecular Probes), fluorescently labeled phalloidin (fluorescein or AF568, 2 U/ml, Molecular Probes) and DAPI (2 μg/ml, Molecular Probes). Tissues were mounted in Prolong Gold (Invitrogen) and allowed to cure for at least 48 hours prior to imaging. In some experiments, anti-*B. fragilis* antibodies were preabsorbed on tissue fragments (4 mm x 4 mm) derived from germ-free mice for 18 hours at room temperature.

Fluorescence microscopy.

An SP5 resonant laser-scanning confocal and two-photon microscope (both scanning heads mounted on the same DM 6000 upright microscope, Leica Microsystems) with a 40× oil objective (numerical aperture 1.4) or 63× oil objective (numerical aperture 1.4) were used for fluorescence microscopy. Images used for 3D reconstructions were acquired using dual confocal – two photon mode. For confocal imaging, 488-nm and 543-nm excitation wavelengths were used for Alexa Fluor 488-labeled bacteria and Alexa Fluor 568-labeled phalliodin, and signals were detected with internal photomultiplier tubes. 2-photon imaging was performed with 4 nondescanned detectors (Leica Microsystems) and a Chameleon Ultra Ti: Sapphire laser (Coherent) tuned at 700–800 nm for acquisition. Emitted fluorescence was split with 3 dichroic mirrors (496 nm, 560 nm and 593 nm) and passed through filter (Semrock) at 585/40 nm. Images (512x512) acquired with a 0.5 μm Z step were smoothed by median filtering at kernel size 3 × 3 pixels. 3D reconstructions of crypts and bacteria were performed using Imaris software (version 7.5.1 ×64; Bitplane AG). Crypt structures were visualized by DAPI and phalloidin signals. Images used for quantification were acquired with FluoView FV10i confocal microscope (Olympus) using 60x (NA 1.35) oil objective. Frames of 512x512 pixels were acquired with 1 μm Z steps in the crypt length axis. Images were processed using ImageJ software (NIH). Background was subtracted (rolling ball method), images were smoothed by median filtering (3x3 pixels), images were segmented by threshold and position of the signal in the Z stack was recorded. Data did not follow normal distribution and were analyzed by non-parametrical two-sided Mann Whitney test.

SPF animal colonization experiments.

7–8 week old male SPF mice (C57BL/6, Swiss Webster, NOD, and Rag-/-) were given a single inoculum of 10^8 CFU of either WT *B. fragilis*, *B. fragilis*ΔCCF, or 1:1 mixture of the two strains by oral gavage. At each time point, bacterial genomic DNA from fecal samples were isolated using a commercial kit (ZR Fecal DNA MiniPrep™, Zymo Research) following the manufacturer's instructions and the relative densities of bacteria were determined by qPCR using strain-specific primers (Supplementary Table 3.2).

Antibiotic treatment.

8–10 week old germ-free Swiss Webster mice were co-associated with WT *B. fragilis* (Cm^r) and *B. fragilis*ΔCCF (Tet^r) mixed at 1:1 ratio for 1 week. Animals were treated with ciprofloxacin for 5 days dissolved in drinking water (1 mg/ml; Hikma Pharmaceuticals). Bacterial CFU of each strain was determined by plating on selective media.

***Citrobacter rodentium* infection.**

8 week old female SPF Swiss Webster mice were treated with metronidazole (100 mg/kg) by oral gavage every 24 hours and ciprofloxacin dissolved in drinking water (0.625 mg/ml) for seven days; mice were transferred to a fresh sterile cage every two days. Two days after the cessation of antibiotic treatment, mice were orally gavaged a single inoculum of 1:1 mixture of WT *B. fragilis* and *B. fragilis*ΔCCF (~ $5x10^8$ CFU total per animal). 6–7 days after *B. fragilis* gavage, mice were either infected orally with ~

$5x10^8$ CFU of overnight culture *C. rodentium* or PBS-gavaged as control. The relative densities of bacteria were determined by fecal bacterial DNA extraction and qPCR.

REFERENCES

Antonopoulos, D.A., Huse, S.M., Morrison, H.G., Schmidt, T.M., Sogin, M.L., and Young, V.B. (2009). Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. Infect Immun *77*, 2367–2375.

Bloom, S.M., Bijanki, V.N., Nava, G.M., Sun, L., Malvin, N.P., Donermeyer, D.L., Dunne, W.M., Jr., Allen, P.M., and Stappenbeck, T.S. (2011). Commensal Bacteroides species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. Cell Host Microbe *9*, 390–403.

Bry, L., Falk, P. G., Midtvedt, T., and Gordon, J. I. (1996). A model of host-microbial interactions in an open mammalian ecosystem. Science *273*, 1380–1383.

Cash, H.L., Whitham, C.V., Behrendt, C L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science *313*, 1126–130.

Dam, P., Olman, V., Harris, K., Su, Z., and Xu, Y. (2007). Operon prediction using both genome-specific and general genomic information. Nucleic Acids Res *35,* 288–298.

Dethlefsen, L., Huse, S., Sogin, M.L., and Relman, D.A. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol *6*, e280.

Dethlefsen, L., and Relman, D.A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proc Natl Acad Sci U S A *108 Suppl 1*, 4554–4561.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635–1638.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792–1797.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol *59*, 307–321.

Hooper, L.V. (2009). Do symbiotic bacteria subvert host immunity? Nat Rev Microbiol *7*, 367–374.

Kawada, M., Chen, C.C., Arihiro, A., Nagatani, K., Watanabe, T., and Mizoguchi, E. (2008). Chitinase 3-like-1 enhances bacterial adhesion to colonic epithelial cells through the interaction with bacterial chitin-binding protein. Lab Invest *88*, 883–895.

Koropatkin, N.M., Cameron, E.A. & Martens, E.C. (2012). How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol *10,* 323–335.

Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics *23*, 127–128.

Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

Liu, C.H., Lee, S.M., Vanlare, J.M., Kasper, D.L., and Mazmanian, S.K. (2008). Regulation of surface architecture by symbiotic bacteria mediates host colonization. Proc Natl Acad Sci U S A *105*, 3951–3956.

Mao, F., Dam, P., Chou, J., Olman, V., and Xu, Y. (2009). DOOR: a database for prokaryotic operons. Nucleic Acids Res *37*, D459–D463.

Markowitz, V.M., Chem, I.M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., Huang, J., Williams, P., Huntemann, M., Anderson, I., Mavromatis, K., Ivanova, N.N., and Kyrpides, N.C. (2012). IMG: the Integrated Microbial Genomes database and comparative anlysis system. Nucleic Acids Red *40*, D115–D122.

Martens, E.C., Koropatkin, N.M., Smith, T.J. & Gordon, J.I. (2009). Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. J Biol Chem *284*, 24673–24677.

Mazmanian, S.K., Liu, C.H., 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.

Mazmanian, S.K., Round, J.L., and Kasper, D.L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. Nature *453*, 620–625.

Mundy, R., MacDonald, T.T., Dougan, G., Frankel, G., and Wiles, S. (2005). Citrobacter rodentium of mice and man. Cell Microbiol *7*, 1697–1706.

Ochoa-Reparaz, J., Mielcarz, D.W., Ditrio, L.E., Burroughs, A.R., Begum-Haque, S., Dasgupta, S., Kasper, D.L., and Kasper, L.H. (2010). Central nervous system demyelinating disease protection by the human commensal Bacteroides fragilis depends on polysaccharide A expression. J Immunol *185*, 4101–4108.

Palmer, C., Bik, E.M., Digiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the Human Infant Intestinal Microbiota. PLoS Biol *5*, e177.

Pedron, T., Mulet, C., Dauga, C., Frangeul, L., Chervaux, C., Grompone, G., and Sansonetti, P.J. (2012). A crypt-specific core microbiota resides in the mouse colon. MBio *3*.

Penders, J., Stobberingh, E.E., van den Brandt, P.A., and Thijs, C. (2007). The role of the intestinal microbiota in the development of atopic disorders. Allergy *62*, 1223–1236.

Price, M.N., Huang, K.H., Alm, E.J., and Arkin, A.P. (2005). A novel method for accurate operon predictions in all sequenced prokaryotes. Nucleic Acids Res *33*, 880–892.

Qin, J. *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59–65.

Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., and Mazmanian, S.K. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. Science *332*, 974–977.

Round, J.L., and Mazmanian, S.K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol *9*, 313–323.

Round, J.L., and Mazmanian, S.K. (2010). Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A *107*, 12204–12209.

Sambrook, J., Maniatis, T., and Fritsch, E.F. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schauer, D.B., and Falkow, S. (1993). Attaching and effacing locus of a Citrobacter freundii biotype that causes transmissible murine colonic hyperplasia. Infect Immun *61*, 2486–2492.

Schauer, K., Rodionov, D.A., de Reuse, H. (2008). New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem Sci *33*, 330–338.

Sekirov, I., and Finlay, B.B. (2009). The role of the intestinal microbiota in enteric infection. J Physiol *587*, 4159–4167.

Sekirov, I., Russell, S.L., Antunes, L.C., and Finlay, B.B. (2010). Gut microbiota in health and disease. Physiol Rev *90*, 859–904.

Shipman, J.A., Berleman, J.E., Salyers, A.A. (2000). Characterization of four outer membrane proteins involved in binding starch to the cell surface of Bacteroides thetaiotaomicron. J Bacteriol *182*, 5365–5372.

Takaishi, H., Matsuki, T., Nakazawa, A., Takada, T., Kado, S., Asahara, T., Kamada, N., Sakuraba, A., Yajima, T., Higuchi, H., Inoue, N., Ogata, H., Iwao, Y., Nomoto, K., Tanaka, R., and Hibi, T. (2008). Imbalance in intestinal microflora constitution could be

involved in the pathogenesis of inflammatory bowel disease. Int J Med Microbiol *298*, 463–472.

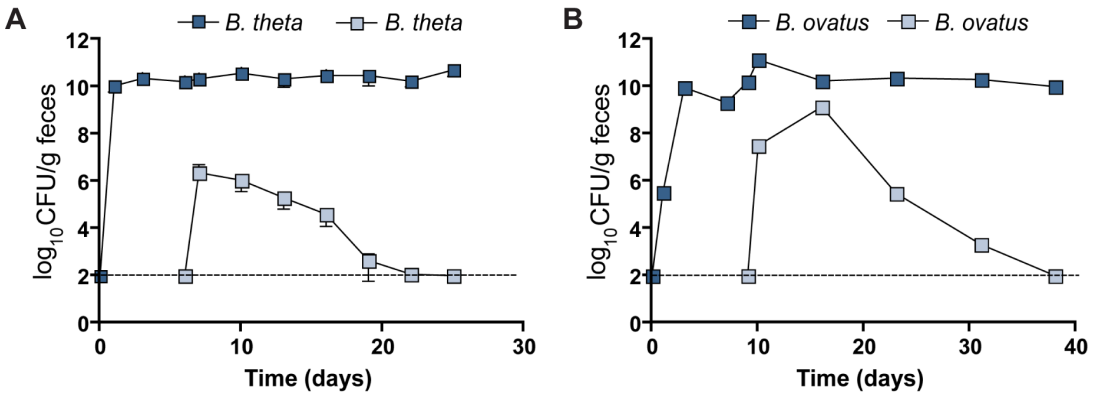
The Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., Egholm, M., Henrissat, B., Heath, A.C., Knight, R., and Gordon, J.I. (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480–484.

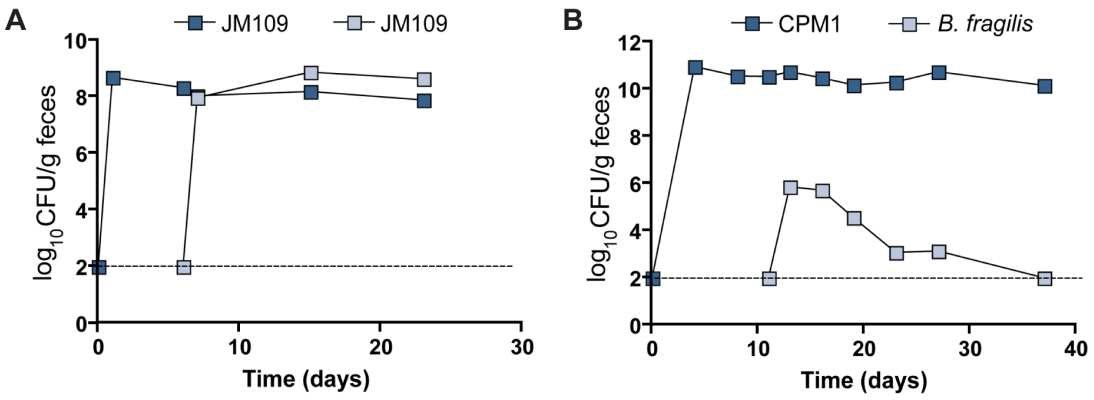
Wurch, T., Lestienne, F., and Pauwels, P.J. (1998). A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes. Biotechnol Tech *12*, 653–657.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R., and Gordon, J.I. (2012). Human gut microbiome viewed across age and geography. Nature *486*, 222–227.

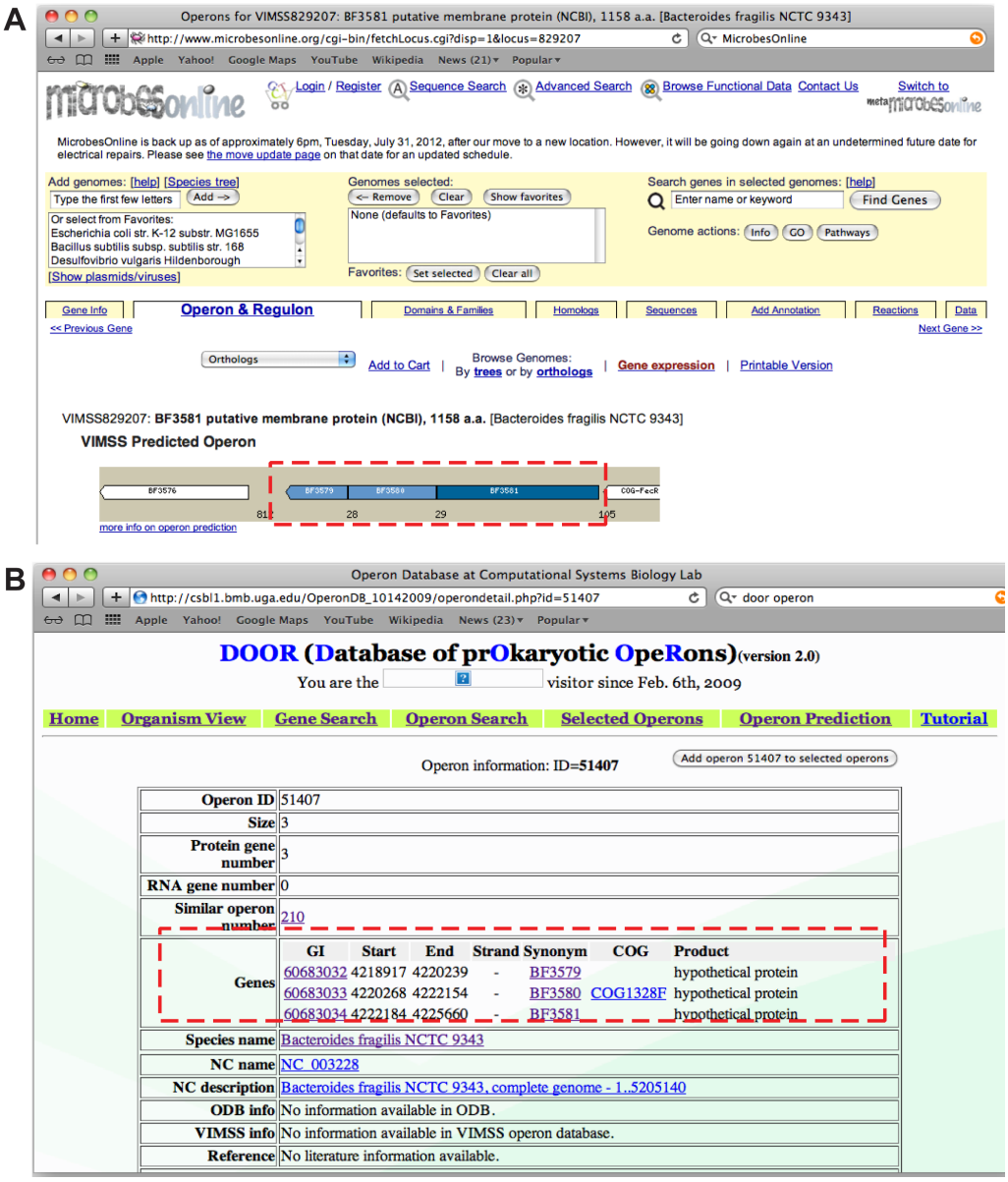
Supplementary Figure 3.1 | (A) Germ-free Swiss Webster mice are mono-associated with *B. thetaiotaomicron* containing pFD340-*cat* (chloramphenicol resistant) for 6 days and subsequently challenged orally with ~ 10^8 CFU of *B. thetaiotaomicron* containing pFD340-*tetQ* (tetracycline resistant). CFU is determined by homogenizing fresh feces in PBS and plating serial dilutions on BHIS agar plate with either chloramphenicol or tetracycline. Dashed line indicates the limit of detection at 100 CFU/g feces (n=3 animals/group). (B) Germ-free Swiss Webster mice are mono-associated with *B. ovatus* containing pFD340-*cat* (chloramphenicol resistant) for 9 days and subsequently challenged orally with ~ 10^8 CFU of *B. ovatus* containing pFD340-*tetQ* (tetracycline resistant). CFU is determined by homogenizing fresh feces in PBS and plating serial dilutions on BHIS agar plate with either chloramphenicol or tetracycline. Dashed line indicates the limit of detection at 100 CFU/g feces (n=3 animals/group).



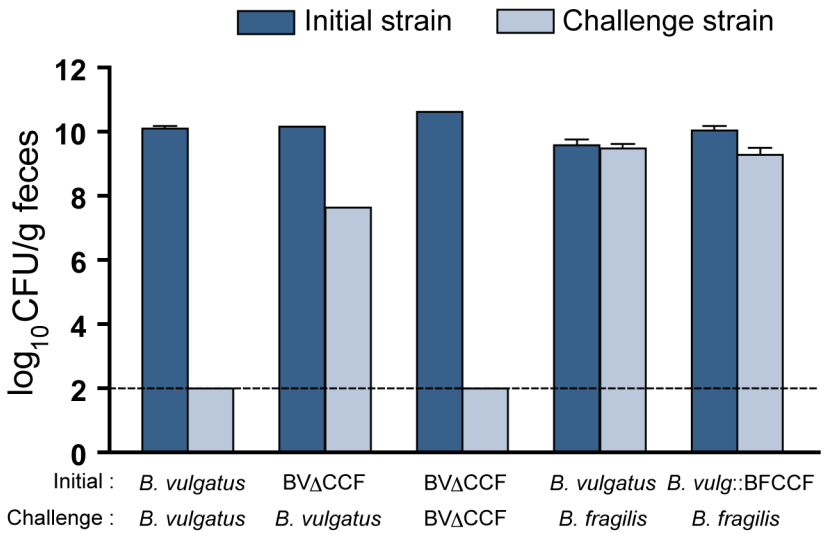
Supplementary Figure 3.2 | (A) Germ-free Swiss Webster mice are mono-associated with *E. coli* JM109 containing pNJR6 (kanamycin resistant) for 6 days and subsequently challenged orally with ~ 10^8 CFU of *E. coli* JM109 containing pFD340 (ampicillin resistant). CFU is determined by serial dilution plating of fecal homogenate on LB agar plate with either ampicillin or kanamycin. (B) Germ-free Swiss Webster mice are mono-associated with a capsular polysaccharide mutant *B. fragilis* strain (CPM1) containing pFD340 (Em-resistant, Cm-sensitive) for 11 days and subsequently challenged orally with ~ 10^8 CFU of wild-type *B. fragilis* containing pFD340-*cat* (Em^r, Cm^r). CFU is determined by serial dilution plating of fecal homogenate on BHIS plate with either Em or Cm. Dashed line indicates the limit of detection at 100 CFU/g feces.



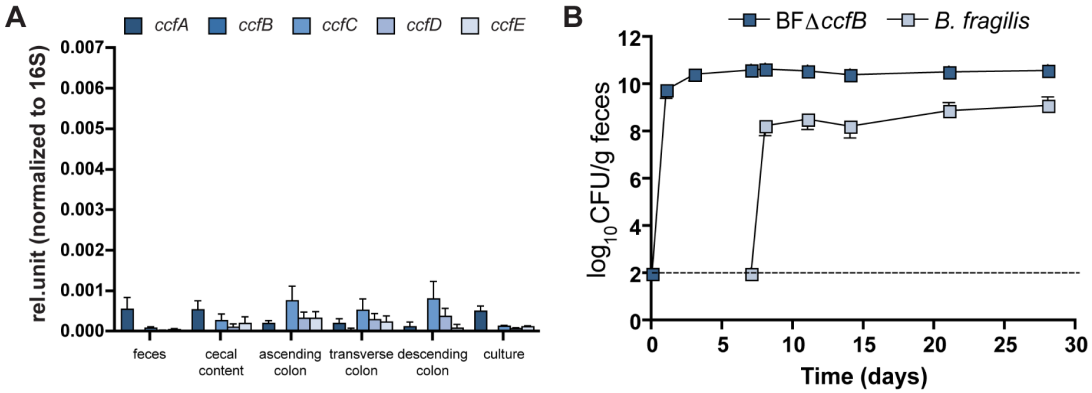
Supplementary Figure 3.3 | (A) MicrobesOnline Operon Predictions. BF3581 was used as a query gene, shown in navy blue in the operon viewer. Shown in light blue, BF3580 and BF3579 are predicted to be in the same operon. The arrowhead indicates the transcription direction. The numbers below the genes indicate the intergenic distance between two adjacent genes. (B) An operon (ID: 51407) was found in DOOR (Database of prOkaryotic OpeRons) by gene query using BF3581. BF3580 and BF3579 also belong to the same operon.



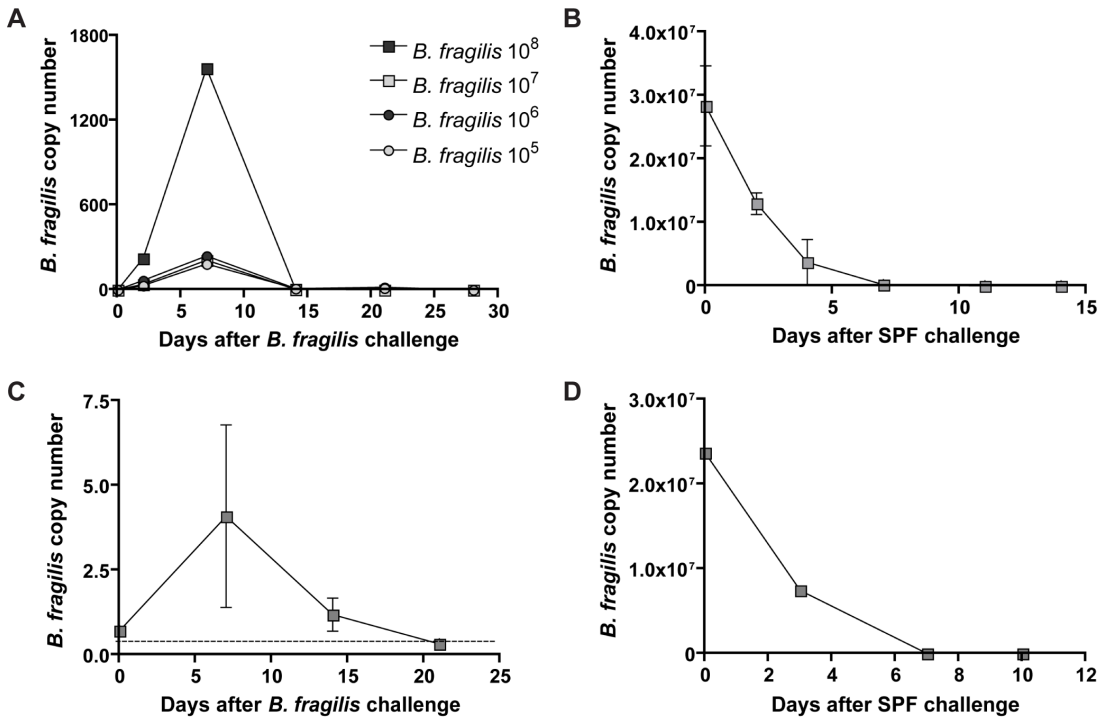
Supplementary Figure 3.4 | Germ-free Swiss Webster mice were mono-associated with *B. vulgatus* (wild-type, mutant strains deleted in the putative structural *ccf* genes (BVΔCCF) or complemented with the *ccfA-E* genes from *B. fragilis* (*B. vulg*::BFCCF)) for 6–7 days and subsequently challenged orally with ~ 10^8 CFU of various challenge strains (*B. vulgatus*, BVΔCCF or *B. fragilis*). CFU of the initial strains and the challenge strains were enumerated from feces by microbiological plating with antibiotic selection 30 days after challenge. Dashed line indicates the limit of detection at 100 CFU/g feces.



Supplementary Figure 3.5 | (A) qRT-PCR of the *ccf* gene expression levels from BFΔ*ccfB* mono-associated animal feces, cecal content and colon tissues (n=3 animals/group) and culture. (B) Fecal bacterial colonization levels measured from germ-free mice mono-associated with BFΔ*ccfB* and subsequently challenged with 10^8 CFU of WT *B. fragilis*.



Supplementary Figure 3.6 | *B. fragilis* colonization level as relative copy number was determined by fecal DNA qPCR for the following animal experiments: (A) SPF Swiss Webster mice were challenged with 10^8, 10^7, 10^6, or 10^5 CFU of *B. fragilis* by oral gavage. (B) Germ-free Swiss Webster mice were mono-associated with *B. fragilis* for 2 weeks and orally challenged with cecal content diluted in PBS from SPF Swiss Webster mice. (C) SPF C57BL/6 mice were challenged with 10^8 CFU of *B. fragilis* by oral gavage. (D) C57BL/6 mice mono-associated with *B. fragilis* from birth were orally challenged with cecal content diluted in PBS from SPF C57BL/6 mice.



Supplementary Table 3.1 | Strains and plasmids used in this study.

Strain or plasmid	Description	Reference or source
E. cloni® 10G	F- *mcrA Δ(mrr-hsdRMS-mcrBC) endA1 recA1 Φ80dlacZΔM15 ΔlacX74 araD139 Δ(ara,leu)7697 galU galK rpsL nupG λ- tonA*; Standard cloning strain	Lucigen
E. coli JM109	F`[*traD36 proA+B+ laclq* Δ*(lacZ)M15*] Δ*(lac-proAB) glnV44 (supE44) e14- (McrA-) thi gyrA96 (NalR) endA1 hsdR17(rk- mk+) relA1 recA1*; standard cloning strain	Zymo Research
Bacteroides fragilis NCTC 9343	Type strain	[1]
B. fragilis 9343ΔPSA Δ*mpi*.off (CPM1)	*B. fragilis* 9343ΔPSA mutant with chromosomal deletion of 534- of 591-bp in the *mpi* gene; invertible promoters for PSA, PSB, and PSD-H all in "off" orientation	[2]
B. fragilis 9343Δ*ccfA*	*B. fragilis* 9343 mutant with chromosomal deletion in *ccfA* gene (BF3583); 312 bp of the 579-bp *ccfA* gene removed	This study
B. fragilis 9343Δ*ccfB*	*B. fragilis* 9343 mutant with chromosomal deletion in *ccfB* gene (BF3582); 945 bp of the 1221-bp *ccfB* gene removed	This study
B. fragilis 9343Δ*ccfC*	*B. fragilis* 9343 mutant with chromosomal deletion in *ccfC* gene (BF3581); 2148 bp of the 3477-bp *ccfC* gene removed	This study
B. fragilis 9343Δ*ccfD*	*B. fragilis* 9343 mutant with chromosomal deletion in *ccfD* gene (BF3580); 1086 bp of the 1887-bp *ccfD* gene removed	This study
B. fragilis 9343Δ*ccfE*	*B. fragilis* 9343 mutant with chromosomal deletion in *ccfE* gene (BF3579); 1006 bp of the 1323-bp *ccfE* gene removed	This study
B. fragilis 9343Δ*ccfC-E* (*B. fragilis*ΔCCF)	*B. fragilis* 9343 mutant with chromosomal deletion of CCF operon in biosynthesis genes *ccfC-E* (BF3581-79); 6059 bp removed	This study
B. vulgatus ATCC 8482	Type strain	[3]
*B. vulgatus*Δ*ccfC-E* (*B. vulgatus*ΔCCF)	*B. vulgatus* mutant with chromosomal deletion of CCF operon in biosynthesis genes *ccfC-E* (BVU946-8); 6304 bp removed	This study
B. vulgatus::BFCCF	*B. vulgatus* complemented with *B fragilis* CCF operon. Clone S16 isolated from the *in vivo* chromosomal library screen.	This study
B. thetaiotaomicron ATCC 29148	Type strain; VPI-5482	[4]
B. ovatus ATCC 8483	Type strain	[3]
Citrobacter rodentium DBS100	Type strain; ATCC 51459	[5]
pNJR6	*Bacteroides* suicide vector; mob^+ Tra^- Km^r (*E. coli*) Em^r (*Bacteroides*)	[6]
R751	Mobilizable mating plasmid to move constructs from *E. coli* to *Bacteroides*; Tra^+ Tp^r	[7]
RK231	Mobilizable mating plasmid to move constructs from *E. coli* to *Bacteroides*, RK2 derivative; Tra^+ Tet^r Km^r	[8]
pFD340	*E. coli-Bacteroides* shuttle vector, IS4351 promoter; Amp^r (*E. coli*) Em^r (*Bacteroides*)	[9]

pFD340-*cat*	Modified pFD340 plasmid containing *cat* gene PCR amplified from *E. Coli* K12/pACYC184 (accession #: X06403) cloned into a *Sma*I site; Amp^r (*E. coli*) Cm^rEm^r (*Bacteroides*)	[2]
pFD340-*cat*BII	*E.coli-Bacteroides* shuttle vector containing IS4351-*cat* cassette PCR amplified from pFD340-*cat*, *Bgl*II restriction site encoded at the 5'-end by PCR primer, and cloned into *Bam*HI/*Pst*I digested and blunted pFD340 backbone; Amp^r (*E. coli*) Cm^rEm^r (*Bacteroides*)	This study
pFD340-*tetQ*	Modified pFD340 plasmid containing *tetQ* gene PCR amplified from *Parabacteroides merdae* ATCC 43184 cloned into *Bam*HI/*Kpn*I site; Amp^r (*E. coli*) Tet^rEm^r (*Bacteroides*)	This study

1. Johnson, J.L., *Taxonomy of Bacteroides .1. Deoxyribonucleic-Acid Homologies among Bacteroides-Fragilis and Other Saccharolytic Bacteroides-Species.* International Journal of Systematic Bacteriology, 1978. **28**(2): 245–256.
2. Liu, C.H., et al., *Regulation of surface architecture by symbiotic bacteria mediates host colonization.* Proc Natl Acad Sci U S A, 2008. **105**(10): 3951–3956.
3. Cato, E.P. and J.L. Johnson, *Reinstatement of Species Rank for Bacteroides-Fragilis, Bacteroides-Ovatus, Bacteroides-Distasonis, Bacteroides-Thetaiotaomicron, and Bacteroides-Vulgatus—Designation of Neotype Strains for Bacteroides-Fragilis (Veillon and Zuber) Castellani and Chalmers and Bacteroides-Thetaiotaomicron (Distaso) Castellani and Chalmers.* International Journal of Systematic Bacteriology, 1976. **26**(2): 230–237.
4. Xu, J., et al., *A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.* Science, 2003. **299**(5615): 2074–2076.
5. Schauer, D.B. and S. Falkow, *Attaching and effacing locus of a Citrobacter freundii biotype that causes transmissible murine colonic hyperplasia.* Infect Immun, 1993. **61**(6): 2486–2492.
6. Stevens, A.M., N.B. Shoemaker, and A.A. Salyers, *The region of a Bacteroides conjugal chromosomal tetracycline resistance element which is responsible for production of plasmidlike forms from unlinked chromosomal DNA might also be involved in transfer of the element.* J Bacteriol, 1990. **172**(8): 4271–4279.
7. Shoemaker, N.B., et al., *Evidence that the clindamycin-erythromycin resistance gene of Bacteroides plasmid pBF4 is on a transposable element.* J Bacteriol, 1985. **162**(2): 626–632.
8. Guiney, D.G., P. Hasegawa, and C.E. Davis, *Plasmid transfer from Escherichia coli to Bacteroides fragilis: differential expression of antibiotic resistance phenotypes.* Proc Natl Acad Sci U S A, 1984. **81**(22): 7203–7206.
9. Smith, C.J., M.B. Rogers, and M.L. McKee, *Heterologous gene expression in Bacteroides fragilis.* Plasmid, 1992. **27**(2): 141–154.

Supplementary Table 3.2 | Sequences of primers used for PCR in this study.

Primers used for cloning recombinant genes and generating deletion constructs (bold: 5' addition)

Primer name	Sequence	Purpose
IS4351-F2	**AAAGATCT**GAAAGAGAGACAATGTCCCC	clone IS4351-*cat*
cat2-X	**AACTCGAG**CGAATTTCTGCCATTCATCCG	clone IS4351-*cat*
tetQ-F2	**AAGGATCC**GTAATCGTTATGCGGCAGTAATAATA TACA	clone *tetQ*
tetQ-R	**AAGGTACC**GAGCTCGTCTATTTTTTTATTGCCAAG	clone *tetQ*
delBF3583L_F	**CTGTCGAC**CGAGGGAAGCATCACTTCAT	delete *ccfA* – left flank
delBF3583L_R	**TTCCCGGTAT**TCTCCCAGACAGCGAGAGAT	delete *ccfA* – left flank
delBF3583R_F	**GTCTGGGAGA**ATACCGGGAATATAGCCATGC	delete *ccfA* – right flank
delBF3583R_R	**ATGTCGAC**TTGTGATACGTCCGTCGGTA	delete *ccfA* – right flank
delBF3582L_F	**AAGGATCC**CGCACTGTTTTTGGGATCTT	delete *ccfB* – left flank
delBF3582L_R	**ATGGCCGA**CGAGACGATTTACGGGATGT	delete *ccfB* – left flank
delBF3582R_F	**AATCGTCTCG**TCGGGCCATACGTAATTCTC	delete *ccfB* – right flank
delBF3582R_R	**TTGGATCC**TCCGATCCCTTTTGTTTGAG	delete *ccfB* – right flank
delBF3581L_F	**GTGGATCC**TAGTTAAACTGACCGAACGATTGA	delete *ccfC* – left flank
delBF3581L_R	**TGCCATTAC**TTTTACCCGGAATAAAATTCTTGA	delete *ccfC* – left flank
delBF3581R_F	**CGGGTAAAA**GTAATGGCACCTATGGTAGGATTC	delete *ccfC* – right flank
delBF3581R_R	**TTGGATCC**TGTGAATGTTTATAGGCAGAAGGA	delete *ccfC* – right flank
delBF3580L_F	**GTGGATCC**GGCTGATTTTATCAGAGTTCCTGT	delete *ccfD* – left flank
delBF3580L_R	**TCGTCAGGC**TATTATTTTCCGTTTGGCAGATTT	delete *ccfD* – left flank
delBF3580R_F	**AAAATAATA**GCCTGACGAATGTATTTGTAACAG	delete *ccfD* – right flank
delBF3580R_R	**TTGGATCC**ACTGTAGGGGTAGATCTCGCTATG	delete *ccfD* – right flank
delBF3579L_F	**AAGGATCC**TTGGCATATCCGGAATTCAT	delete *ccfE* – left flank
delBF3579L_R	**TAGGCGAAAAG**CGATCGGTCAGTTTGGTTTT	delete *ccfE* – left flank
delBF3579R_F	**TGACCGATCG**CTTTTCGCCTACATTATAAGATTGC	delete *ccfE* – right flank
delBF3579R_R	**ATGGATCC**GCGTCGACCAGTCCAATTAT	delete *ccfE* – right flank
delBF3579L_F	**AAGGATCC**TTGGCATATCCGGAATTCAT	delete *ccfC-E* – left flank
delBF3579L_Rb	**TGCCATTAC**CGATCGGTCAGTTTGGTTTT	delete *ccfC-E* – left flank
delBF3581R_Fb	**TGACCGATCG**GTAATGGCACCTATGGTAGGATTC	delete *ccfC-E* – right flank
delBF3581R_R	**TTGGATCC**TGTGAATGTTTATAGGCAGAAGGA	delete *ccfC-E* – right flank
delBVU946L_F	**CTGTCGAC**GGATTTCGTCTTGCACAGGT	delete *ccfC-E* – left flank (*B. vulgatus*)
delBVU946L_R	**TGGGGAATCC**ACGTTGCTGCCCTCAAATAC	delete *ccfC-E* – left flank (*B. vulgatus*)
delBVU948R_F	**GCAGCAACGT**GGATTCCCCACTGCTACAAA	delete *ccfC-E* – right flank (*B. vulgatus*)
delBVU948R_R	**ATGTCGAC**TCGACTCCGTAGATCCCATC	delete *ccfC-E* – right flank (*B. vulgatus*)

Primers used for quantitative PCR

Primer name	Sequence	Target	Reference or source
BF3583 QF	GGAATTTGCATGACACTTAT	*B. fragilis ccfA*	This study
BF3583 QR	CTGAGAGGTTTCATCTTCTG	*B. fragilis ccfA*	This study
BF3582 QF	AGTGTCCCCACTTCATCGTC	*B. fragilis ccfB*	This study
BF3582 QR	TGAAACTTTTGCCGGAGAAT	*B. fragilis ccfB*	This study
BF3581 QF	GATGAACTGATAGCCCATTA	*B. fragilis ccfC*	This study
BF3581 QR	TAGCGATGACTAAAGGTGTT	*B. fragilis ccfC*	This study

BF3580 QF	CGGTTATATGCTTTTCAAAC	*B. fragilis ccfD*	This study
BF3580 QR	CAAATAGAAATCTGCCAAAC	*B. fragilis ccfD*	This study
BF3579 QF	TGCTATTTGCACGGGTAACA	*B. fragilis ccfE*	This study
BF3579 QR	CCGAAACTCCGATTCTTCAT	*B. fragilis ccfE*	This study
Bfragilis 16S F	TGATTCCGCATGGTTTCATT	*B. fragilis* 16S rRNA	[10]
Bfragilis 16S R	CGACCCATAGAGCCTTCATC	*B. fragilis* 16S rRNA	[10]
BF3581 QF3	CACCGATACCCTGCGTAAAT	WT *B. fragilis* specific	This study
BF3581 QR3	GGCGGACTGGTAACGATAAA	WT *B. fragilis* specific	This study
delCCF QF	CGGTGCTAACGTTGTCGTAA	*B. fragilis*ΔCCF specific	This study
delCCF QR	ATTTTAGTGCGGCATCCTGA	*B. fragilis*ΔCCF specific	This study
CfcH QF	GGTAAATCCACCACCCTGAA	*C. rodentium* specific	This study
CfcH QR	GTATTCCACGGGGTCTTCAA	*C. rodentium* specific	This study
UniF334	ACTCCTACGGGAGGCAGCAGT	Universal 16S	[11]
UniR514	ATTACCGCGGCTGCTGGC	Universal 16S	[11]

10. Odamaki, T., et al., *Distribution of different species of the Bacteroides fragilis group in individuals with Japanese cedar pollinosis.* Appl Environ Microbiol, 2008. **74**(21): 6814–6817.
11. Amann, R.I., et al., *Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations.* Appl Environ Microbiol, 1990. **56**(6): 1919–1925.

Chapter 4:

Conclusions and future directions

***B. fragilis* capsular polysaccharides mediate gut colonization**

In Chapter 2, we showed that expression of at least one of eight capsular polysaccharides is critical for *in vitro* growth of *B. fragilis* and an attempt to abrogate expression of all capsular polysaccharides leads to spontaneous phenotypic reversion of the bacteria to express one of the polysaccharides, PSB, through promoter inversion. Clearly, *B. fragilis* invested a great deal of evolutionary energy into expressing multiple polysaccharides as part of its surface architecture and the ability to vary its capsule sugar molecules gives the organism advantage in establishing colonization in the host gut. We speculated that this dynamic and diverse surface sugar repertoire is important for long-term and stable host colonization by *B. fragilis*.

There are many possible explanations for why WT strain outcompetes CPM1 strain (PSC-only expressing strain) during colonization. The ability to vary its capsular polysaccharide molecules may be important for camouflaging its surface and evading immune response when the bacteria is in intimate contact with the host cells. Different polysaccharides may be required for surviving and transiting through different microenvironments within the gut. Introducing heterogeneity within the *B. fragilis* population may increase their fitness *in vivo*. Or, there may be a dominant polysaccharide out of the eight known molecules that is responsible for direct interaction with the host mediating stable colonization and it is missing from the CPM1 strain. To address these questions, we can generate a variety of mutant strains expressing a single capsular polysaccharide (PSA-only, PSB-only, etc.) and test their colonization fitness in competition. If there were indeed a dominant polysaccharide molecule directly responsible for colonization, this approach would reveal the identity of the molecule and

open doors to further dissection of colonization mechanism of a model symbiont. Similar mutants can be generated using other *Bacteroides* species that express phase-variable multiple capsular polysaccharides (Coyne and Comstock, 2008) to show conservation of the mechanism among other gut symbionts of this genus.

Species-specific and saturable gut colonization by *Bacteroides*

In Chapter 3, we discovered that many of the intestinal *Bacteroides* species colonize the host gut by saturation of a limited niche which is a previously unknown phenomenon and unique to the *Bacteroides* genus (although we have only tested *E. coli* JM109 as a control). Remarkably, *B. fragilis* mono-associated mice resisted additional colonization by the isogenic *B. fragilis* strain and the same was observed in *B. thetaiotaomicron, B. ovatus* and *B. vulgatus*. Different *Bacteroides* species, however, were able to both colonize the same animal. From these observations, we concluded that *Bacteroides* species occupy limited and saturable niche in the gut that is specific to the species and different species do not compete for the same niche. To our knowledge, it is the first time such an observation has been made with regards to commensal microbe and host interaction.

This finding opens up many additional questions and future studies. *Bacteroides* is one of many major bacterial genera that make up the mammalian gut microbiota (Eckburg et al., 2005) and it would be important to assess if the other groups such as *Clostridium*, *Lactobacillus, Bifidobacterium*, and *E. coli* also establish long-term colonization by occupying saturable, species-specific niche using gnotobiotic animal

model. Consequently, we will learn whether this mechanism is universal to a broader range of commensal organisms or if it is unique to the *Bacteroides* genus.

The colonization studies strongly implicate presence of distinct species-specific niches in the intestine for *Bacteroides*. Visualization of the *Bacteroides* biogeography within the host intestine can be accomplished by immunofluorescence or fluorescence *in situ* hybridization techniques by developing probes that specifically recognize different species. This approach will elucidate whether *Bacteroides* species segregate spatially within the intestine or occupy distinct niches through an entirely different mechanism (i.e., nutrient or host receptor binding), which will subsequently provide insight into which host-derived factors may serve as nutrient substrates or docking sites for specific microbes and regulate gut microbial community composition.

Identification of a Sus-like system mediating *Bacteroides* colonization

In Chapter 3, we identified a novel molecular mechanism mediated by a set of genes in an operon which we named the *c*ommensal *c*olonization *f*actors (*ccf*). By sequence homology, the CCF system is conserved among many *Bacteroides* species, many of which display saturable colonization phenotype, and at the core of the operon is the SucC/D family of proteins, one of the largest cohort of outer membrane protein complexes conserved among the *Bacteroidetes* that bind and import a wide range of polysaccharides (Kuwahara et al., 2004; Cerdeno-Tarraga et al., 2005; Schauer et al., 2008; Martens et al., 2009).

The starch utilization system (Sus) has been extensively studied from *B. thetaiotaomicron* (Koropatkin et al., 2012). Whole-genome transcriptional profiling of *B.*

thetaiotaomicron disclosed that the bacteria expresses different sets of carbohydrate utilization genes (37 SusC and 16 SusD paralogs were up-regulated) when introduced into the germ-free mouse gut compared to when grown in broth consisting of minimal medium with glucose (MM-G) (Sonnenburg et al., 2005). The glycan foraging behavior of the gut symbiont was further explored by comparing the bacterial gene expressions in germ-free mice maintained either on the standard polysaccharide-rich chow diet or on the simple sugar diet devoid of fermentable polysaccharide. Whereas the polysaccharide-rich diet induced up-regulation of carbohydrate utilization genes such as xylanases, arabinosidases, and pectate lyases, glucose and sucrose only diet led to increased expression of a different subset of genes involved in retrieving carbohydrates from mucus glycans such as hexosaminidases, α-fucosidases and a sialidase. These genes may also serve to mediate bacterial attachment to mucus glycan to avoid bacterial washout from the gut (Xu et al., 2003; Xu and Gordon, 2003). The ability to salvage energy from nutrients that are otherwise non-digestible by the host provides an evolutionary driving force for the bacteria to maintain residency in the host intestine.

The CCF system reveals a novel, evolutionarily conserved molecular pathway by which prominent human *Bacteroides* species maintain persistent colonization of the mammalian gut. Homology to the Sus family of proteins suggests that the putative outer membrane complex we have identified is involved in uptake and utilization of glycans. Our discovery of saturable colonization within, but not between, species suggests that individual *Bacteroides* members utilize unique and limiting sugars for association with the host. We have identified a conserved clade of the *sus* genes, and our data suggest a model that (at least a subset of) Sus systems evolved to sense and utilize limiting glycan

nutrients by individual *Bacteroides* species to promote intestinal colonization and identification of the chemical substrate(s) for the CCF system will serve as a crucial future direction.

REFERENCES

Cerdeno-Tarraga, A.M., Patrick, S., Crossman, L.C., Blakely, G., Abratt, V., Lennard, N., Poxton, I., Duerden, B., Harris, B., Quail, M.A., Barron, A., Clark, L., Corton, C., Doggett, J., Holden, M.T., Larke, N., Line, A., Lord, A., Norbertczak, H., Ormond, D., Price, C., Rabbinowitsch, E., Woodward, J., Barrell, B., and Parkhill, J. (2005). Extensive DNA inversions in the B. fragilis genome control variable gene expression. Science *307*, 1463–1465.

Coyne, M.J., and Comstock, L.E. (2008). Niche-specific features of the intestinal Bacteroidales. J Bacteriol *190*, 736–742.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635–1638.

Koropatkin, N.M., Cameron, E.A. & Martens, E.C. (2012). How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol *10,* 323–335.

Kuwahara, T., Yamashita, A., Hirakawa, H., Nakayama, H., Toh, H., Okada, N., Kuhara, S., Hattori, M., Hayashi, T., and Ohnishi, Y. (2004). Genomic analysis of Bacteroides fragilis reveals extensive DNA inversions regulating cell surface adaptation. Proc Natl Acad Sci U S A *101*, 14919–14924.

Martens, E.C., Koropatkin, N.M., Smith, T.J. & Gordon, J.I. (2009). Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. J Biol Chem *284*, 24673–24677.

Schauer, K., Rodionov, D.A., de Reuse, H. (2008). New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem Sci *33*, 330–338.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science *307*, 1955–1959.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., and Gordon, J.I. (2003). A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science *299*, 2074–2076.

Xu, J., and Gordon, J.I. (2003). Inaugural Article: Honor thy symbionts. Proc Natl Acad Sci U S A *100*, 10452–10459.

Appendix I:

The human commensal *Bacteroides fragilis* binds intestinal mucin

Julie Y. Huang, S. Melanie Lee, and Sarkis K. Mazmanian

Anaerobe 17 (2011) 137–141

Contents lists available at ScienceDirect

Anaerobe

journal homepage: www.elsevier.com/locate/anaerobe

Molecular Biology, Genetics and Biotechnology

The human commensal *Bacteroides fragilis* binds intestinal mucin

Julie Y. Huang [1], S. Melanie Lee, Sarkis K. Mazmanian*

Division of Biology, California Institute of Technology, 1200 E. California Bl., Mail Code 156-29, Pasadena, CA 91125, USA

ARTICLE INFO

Article history:
Received 28 October 2010
Received in revised form
14 May 2011
Accepted 16 May 2011
Available online 2 June 2011

Keywords:
Bacteroides fragilis
Gut commensal
Mucin
Adhesion
Microbiota

ABSTRACT

The mammalian gastrointestinal tract harbors a vast microbial ecosystem, known as the microbiota, which benefits host biology. *Bacteroides fragilis* is an important anaerobic gut commensal of humans that prevents and cures intestinal inflammation. We wished to elucidate aspects of gut colonization employed by *B. fragilis*. Fluorescence in situ hybridization was performed on colonic tissue sections from *B. fragilis* and *Escherichia coli* dual-colonized gnotobiotic mice. Epifluorescence imaging reveals that both *E. coli* and *B. fragilis* are found in the lumen of the colon, but only *B. fragilis* is found in the mucosal layer. This observation suggests that physical association with intestinal mucus could be a possible mechanism of gut colonization by *B. fragilis*. We investigated this potential interaction using an *in vitro* mucus binding assay and show here that *B. fragilis* binds to murine colonic mucus. We further demonstrate that *B. fragilis* specifically and quantitatively binds to highly purified mucins (the major constituent in intestinal mucus) using flow cytometry analysis of fluorescently labeled purified murine and porcine mucins. These results suggest that interactions between *B. fragilis* and intestinal mucin may play a critical role during host-bacterial symbiosis.

Following a sterile birth, the gastrointestinal (GI) tracts of humans and all mammals coordinately assemble a diverse multitude of microorganisms, collectively known as the microbiota. It has been acknowledged for decades that many of these microorganisms live symbiotically with their hosts, performing beneficial functions such as metabolizing complex carbohydrates and providing essential nutrients [1]. Recent studies have shown that the microbiota critically augments the development and function of the immune system (reviewed in [2] and [3]). Although the microbial diversity in the mammalian gut is vast (with an estimated 500–1000 species of microorganisms present in the human), organisms belonging to the genus *Bacteroides* represent one of the most abundant microbial taxa in both mice and humans [4]. *Bacteroides fragilis* is a ubiquitous Gram-negative anaerobic bacterium that inhabits the lower GI tract of most mammals [5]. Recent findings have revealed that this organism possesses the ability to direct the cellular and physical maturation of the host immune system and to protect its host from experimental colitis [6–8]. Therefore, the contributions of the microbiota to human health appear to be profound.

We wanted to understand how *B. fragilis* colonizes the mammalian gut. *B. fragilis* expresses at least eight distinct surface capsular polysaccharides (CPS), and previous studies have shown that CPS expression by the bacterium is required for successful intestinal colonization [9,10]. How these molecules mediate the initial interactions with the host, and whether they are involved in long-term persistence in the gut are currently unknown. Several mechanisms of gut colonization by symbiotic bacteria have been studied. Some organisms form biofilms, composed of a polymeric matrix secreted by the bacteria, which adhere to the epithelial layer. Others interact with components of the mucosal layer (reviewed in [11]). Mucus is a viscous stratum which separates epithelial cells from the lumen of the gut and acts as a crucial barrier against infection by pathogens. Various membrane-bound or secreted glycoproteins called mucins associate with one another to form the gel-like mucus. Interactions between gut bacteria and mucus have been hypothesized to be important for the assembly and stability of the microbiota [12]. Accordingly, we sought to determine whether or not *B. fragilis* binds intestinal mucus and purified mucin.

Initially, we visualized the spatial localization in the colon of 2 different commensal bacteria to determine potential differences in association with the mucus layer *in vivo*. Wild-type *B. fragilis* NCTC9343 was grown anaerobically in brain-heart infusion (BHI) supplemented with hemin (5 μg/ml) and vitamin K (0.5 μg/ml), and *Escherichia coli* BL21 was grown aerobically in BHI at 37 °C. Bacteria were grown to OD_{600} of 0.7–0.8, and 1×10^8 colony forming units (CFUs) were orally inoculated into germ-free Swiss Webster mice by gavage. Following 1 week of colonization, mice were sacrificed and colon tissue was fixed in Carnoy's solution and embedded in

* Corresponding author. Tel.: +1 626 395 2356.
E-mail address: sarkis@caltech.edu (S.K. Mazmanian).
[1] Present address: Dept. of Microbiology and Immunology, Fairchild Science Building D300, 299 Campus Drive, Stanford University, Stanford, CA 94305, USA.

1075-9964/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.anaerobe.2011.05.017

paraffin wax for sectioning. Fluorescence in situ hybridization was performed on tissue sections mounted on glass slides using a 6-carboxyfluorescein (6-FAM)-labeled oligonucleotide probe for *E. coli* (EnterbactB [AAGCCACGCCTCAAGGGCACAA]) and a Cy3-labeled oligonucleotide probe for *B. fragilis* (Bfra602 [GAGCCG-CAAACTTTCACAA]) (Integrated DNA Technologies, Inc.). Briefly, slides were deparaffinized, dried, and hybridized with both probes at 5 ng/μl concentration each for 2 h at 46 °C in hybridization buffer (0.9 M NaCl, 15% formamide, 20 mM Tris–HCl (pH 7.4), and 0.01% sodium dodecyl sulfate (SDS)). Slides were washed for 15 min at 48 °C in wash buffer (20 mM Tris–HCl (pH 7.4), 318 mM NaCl, and 0.01% SDS). For visualization of the colon epithelial cell nuclei, the slides were counterstained with 4′,6′-diamidino-2-phenylindole (DAPI). The autofluorescence background allowed visualization of the tissue structures. The slides were examined with an Axioplan microscope (Zeiss, Oberkochen, Germany) using a 100× oil immersion objective. Epifluorescence images of a cross section through the colon of gnotobiotic mice that were dual-colonized with both *E. coli* and *B. fragilis* reveal that both bacteria are found in the lumen of the gut in high abundance (Fig. 1). Surprisingly however, only *B. fragilis* is found in the mucus layer that lies between the lumen and the gut epithelium tissue (Fig. 1). The spatial segregation of *B. fragilis* and *E. coli* across the colon mucus barrier suggests that *B. fragilis* may interact with mucus *in vivo* and this may be important for sustained colonization of commensal *B. fragilis*. Furthermore, these results reveal that not all bacteria are equally able to penetrate the mucus layer, suggesting dedicated mucus associating functions for *B. fragilis*.

To test the hypothesis that *B. fragilis* colonization of the distal gut is mediated by mucus binding, a standard mucus binding assay was used to determine if live bacteria are able to bind a crude, intestinal mucus preparation. Crude mucus was isolated from the colon and cecum of conventionally-colonized Swiss Webster mice as described in Cohen et al. [13]. Briefly, colonic and cecal mucus was scraped into HEPES-Hanks' Buffer (pH 7.4 with Calcium Chloride and Magnesium Chloride). Next, non-soluble material was removed by centrifuging once at 12,000 × g for 10 min at 4 °C, and then once at 26,500 × g for 15 min at 4 °C. The final concentration of the crude mucus solution was determined by the Bradford assay. The mucus was diluted with HEPES-Hanks' Buffer to 1 mg/ml. 0.2 ml of mucus was added into the wells of a 24-well tissue culture plate and incubated overnight at 4 °C. Controls included wells containing 0.2 ml of a 1 mg/ml solution of Bovine Serum Albumin (BSA, which served as a specificity control) or 0.2 ml of HEPES-Hanks' Buffer (which served as a negative control). The wells were washed with HEPES-Hanks' Buffer to remove non-immobilized proteins. The plate was UV-sterilized for 10 min and was then ready for use in the mucus binding assay. 1×10^8 CFUs of bacteria were added to both the immobilized mucus and the BSA control, and incubated at 37 °C for 1 h. Wells were washed with HEPES-Hanks' Buffer, treated with 0.05% trypsin for 10 min at room temperature to liberate bacteria. One milliliter of cold BHI was added to quench the trypsin activity. Samples were serially diluted and plated for CFUs. Fig. 2A shows that *B. fragilis* binds to crudely purified mucus *in vitro*, as determined by recovered CFUs. The BSA- and buffer-containing wells illustrate low background binding. A mutant strain of *B. fragilis* (CPM1), which only expresses one of the eight CPS [9], is able to bind mucus as effectively as wild-type *B. fragilis*, suggesting that CPS expression does not mediate mucus binding. Therefore, *B. fragilis* specifically binds intestinal mucus via a mechanism that appears not to involve expression of multiple surface polysaccharides.

Next, a mucus binding competition assay was performed to determine if the interaction between *B. fragilis* and mucus is saturable. We reasoned that as *B. fragilis* is pre-coated with higher concentrations of excess mucus, fewer putative receptors would be available to bind immobilized mucus in the well. Briefly, 1×10^8 CFUs of *B. fragilis* were incubated with excess mucus at 37 °C for 2 h under aerobic conditions with shaking. Bacteria were washed and added to

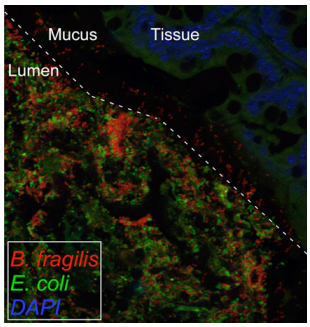


Fig. 1. Colon tissue section from a *B. fragilis* and *E. coli* dual-colonized Swiss Webster mouse. Epifluorescence image of bacteria visualized by FISH, and the epithelial cells counterstained with DAPI (blue) to visualize DNA. Both *E. coli* (green) and *B. fragilis* (red) are found in the lumen but only *B. fragilis* is found in the mucus layer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

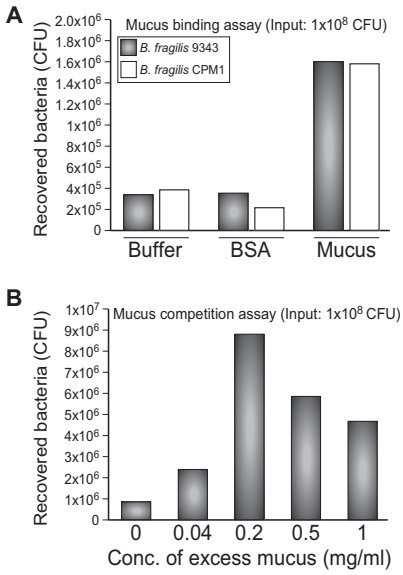


Fig. 2. *B. fragilis* binds intestinal mucus. (A) Number of *B. fragilis* (in CFUs) recovered after 1 h incubation in wells with an immobilized mucus layer, an immobilized BSA layer, or buffer only. Of the 1×10^8 CFUs incubated, 1.6×10^6 (1.6%) bound to immobilized mucus. The CPM1 mutant binds mucus similarly to wild-type bacteria. These data are representative of four independent trials. (B) Number of bacteria recovered from mucus binding assay after a 2 h pre-incubation with different concentrations of excess mucus. These data are representative of three independent trials.

wells of a 24-well tissue culture plate containing either immobilized mucus, BSA, or nothing (prepared as above). After 1 h, samples were treated with trypsin, serially diluted, and plated for CFUs. Unexpectedly, pre-incubation with excess mucus appeared to increase *B. fragilis* binding to mucus with a bi-phasing profile (Fig. 2B). Binding to immobilized mucus reached a peak when *B. fragilis* was pre-incubated with 0.2 mg/ml of excess mucus. Pre-incubation of bacteria with excess mucus at concentrations higher than 0.2 mg/ml resulted in a decrease in mucus binding, yet binding was still higher than without pre-incubation with mucus. Pre-incubation of bacteria with 0.4 mg/ml and 1 mg/ml of BSA did not affect binding, once again showing that the *B. fragilis*-mucus interaction is specific (data not shown). These results suggest that bacteria pre-incubated with mucus (and not BSA) have an increased ability to bind immobilized mucus until putative receptors are saturated at the highest mucus concentrations. Further experiments are required to determine if dedicated molecules on the bacterial surface mediate mucus binding.

In addition to mucin, intestinal mucus is known to contain host molecules, such as anti-microbial peptides, immunoglobulin A (IgA) antibodies, and lysozyme [13]. We wished to determine if mucus binding by *B. fragilis* was specific to mucin. As murine colonic mucin is not commercially available, we purified mucins from Swiss Webster mice based on the protocol by Shekels et al. [14] with a few modifications. Fig. 3 illustrates a schematic of this modified protocol and the analysis of mucin purity. We then tested the purified mucin and BSA for specific binding by *B. fragilis*. Purified mucin and the BSA control were labeled with Thermo Scientific DyLight Amine-Reactive Fluor 488, and unbound fluorophores were removed from the sample via dialysis against PBS. *B. fragilis* was pre-incubated with either unlabeled BSA or PBS and was subsequently incubated with labeled mucin or labeled BSA for 30 min at room temperature. The bacteria were washed after each incubation to remove non-adherent material. Percentage of mucin-binding bacteria in each sample was determined by flow cytometry. When *B. fragilis* was incubated with fluorescently labeled BSA, no binding was detected (Fig. 4A). However, when *B. fragilis* was incubated with labeled mucin, a significant number of *B. fragilis* was detected by flow cytometry. Pre-incubation with BSA did not diminish the percentage of *B. fragilis* adherent to mucin (Fig. 4A). Taken together, *B. fragilis* binds specifically to purified murine colonic mucin and not to BSA.

B. fragilis colonizes the intestines of most mammalian species studied to date [5]. In order to determine if mucin interactions extend beyond the murine host, we examined the ability of *B. fragilis* to bind porcine mucin. Starting with partially purified porcine gastric mucin purchased from Sigma Aldrich, we purified mucin to homogeneity using the same protocol as described above. Fig. 4B shows that a significant amount of mucin-binding bacteria were detected by flow cytometry, indicating that *B. fragilis* bound the fluorescently labeled purified porcine mucin. Both approaches we used in this study to demonstrate mucus binding resulted in only a small portion of

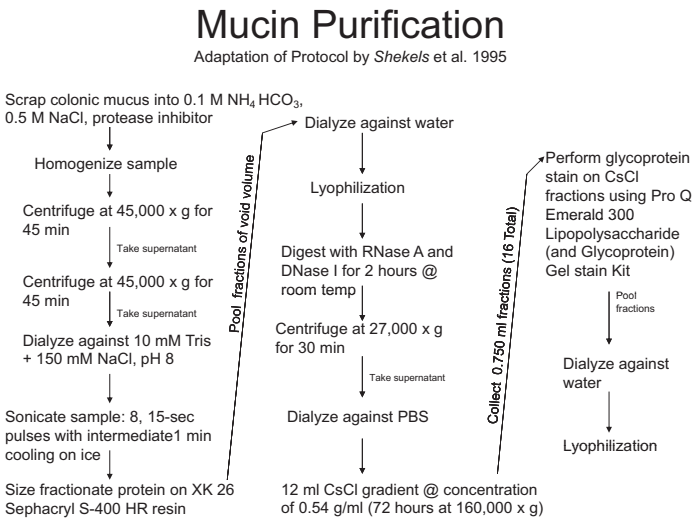


Fig. 3. Schematic of the mucin purification protocol. Briefly, crude mucus was scraped from the colon and cecum of conventionally-colonized 8-week-old male Swiss Webster mice into 0.1 M NH_4HCO_3, 0.5 M NaCl, and a cocktail of protease inhibitors on ice. The sample was then homogenized and centrifuged at 45,000 × g for 45 min at 4 °C. Non-soluble material was removed before centrifuging again at 45,000 × g for another 45 min at 4 °C. The supernatant was taken and dialyzed against 10 mM Tris pH 8.0 + 150 mM NaCl for about 24 h. Next, the sample was sonicated at eight 15-s pulses with intermediate 1 min cooling on ice on a Brason Sonicator at speed 3 to break up large aggregates and then centrifuged once more at 45,000 × g for 45 min. Next HPLC was employed whereby the supernatant was size fractionated on an XK 26/70 column containing Sephacryl S-400 resin (equilibrated in 10 mM Tris, pH 8.0). The void volume (which contained the large mucin glycoproteins) was collected and dialyzed against water for about 36 h and then lyophilized. The lyophilized glycoproteins were resuspended in a solution containing RNase A and DNase I and digested for 2 h at room temperature. After the digestion, the sample was centrifuged at 27,000 × g for 30 min at 4 °C and the supernatant was dialyzed against phosphate buffered saline (PBS) for 36 h. Cesium chloride was added to the dialyzed supernatant to a final concentration of 0.54 g/ml, and then centrifuged at 160,000 × g for 72 h. One milliliter fractions were collected and analyzed with the Pro Q Emerald Glycoprotein Staining Kit to determine which fractions contained the purified mucins. The mucin-containing fractions were pooled, dialyzed against water for 24 h, lyophilized, and then stored at −20 °C. Positive fractions from gel filtration chromatography were identified by absorbance readings at 280 nm. CsCl fractions and final product were assayed to contain mucin by glycoprotein staining (data not shown).

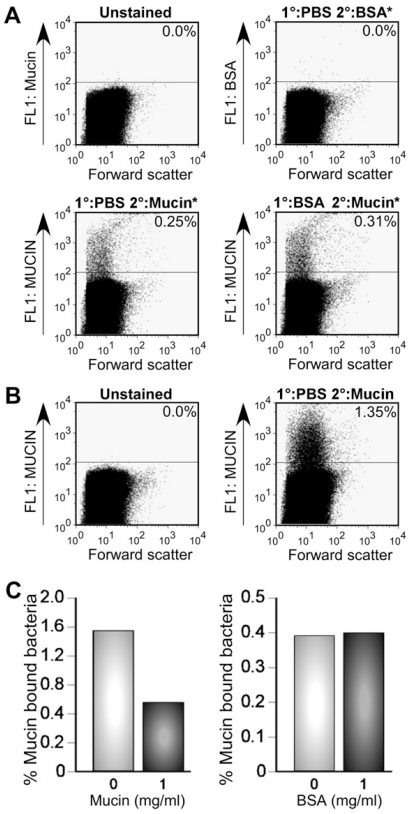


Fig. 4. *B. fragilis* binds soluble murine and porcine mucin. (A) Flow cytometry plots indicating percentage of *B. fragilis* bound to fluorescently labeled murine colonic mucin. Cells were either pre-incubated with BSA or not (1°), and secondary incubations were with fluorescently labeled BSA or mucin (denoted by asterisk). Percentages represent bacteria bound to fluorescently-labeled mucin relative to total number of bacteria analyzed per sample. These data are representative of two independent trials. (B) Percentage of *B. fragilis* bound to fluorescently labeled porcine mucin with no pre-incubation. These data are representative of two independent trials. (C) Percentage of *B. fragilis* bound to fluorescently labeled porcine gastric mucin following pre-incubation with unlabeled mucin (left) or unlabeled BSA (right). Porcine mucin was purchased from a commercial source and purified as described in Fig. 3 from the RNase/DNase digestion step. These data are representative of two independent trials.

bacterial binding (~1.6% for the immobilized plate assay and ~1.5% for the soluble mucin binding assay). This is consistent with the known ability of *B. fragilis* to be highly phase variable whereby only a portion of the bacterial population express a given surface molecule [15]. Fig. 4C shows that pre-incubation with 1.0 mg/ml of unlabeled mucin was able to compete with the fluorescently-labeled mucin, resulting in a lower percentage of bacteria binding to the fluorescently labeled mucin. Pre-incubation with BSA shows no inhibition (Fig. 4C), serving as a specificity control. Our results show that *B. fragilis* specifically binds porcine mucin in addition to murine mucin.

B. fragilis has emerged as a model symbiont for the study of host–microbial interactions with the immune system [3]. The mechanism by which *B. fragilis* maintains long-term colonization of the mammalian intestine remains unknown. Associations with mucus may involve bacterial binding, and/or nutrient utilization of mucin for bacterial growth. If binding to mucus is involved during the colonization process *in vivo*, we predict that *B. fragilis* would express defined and dedicated receptor(s) with specific affinity for mucin. Along these lines, the *B. fragilis* genome and other sequenced *Bacteroides* species express numerous homologs of the SusC/SusD proteins, which are known to bind starch and other carbohydrates that decorate the mucin glycoproteins [16]. Furthermore, SusC/SusD proteins of *B. fragilis* were recently shown to be phase variable [17]. This property is similar to the phase variability of capsular polysaccharides, whereby only a small fraction of bacteria express any one of the eight CPS of *B. fragilis* [9]. If mucus binding is also phase variable, this would explain why only a small percentage of bacteria invade the mucus layer (as shown in Fig. 1), and why only a small fraction of bacteria bind mucus and mucin *in vitro* (as shown in Figs. 2 and 4). A non-mutually exclusive function for mucus binding may be the use of host derived sugars as a carbon source. Several studies have shown that *B. fragilis* can degrade mucin and utilize it as a nutrient source for growth [18,19]. In fact, *B. fragilis* can utilize porcine mucin as a sole source for carbon and nitrogen [20], and structural analysis of the SusD homolog of *Bacteroides thetaiotaomicron* (also found in *B. fragilis*) suggests it binds sugars liberated from mucin glycoproteins [21]. Therefore, mucus binding may serve as a physical mechanism for sustained colonization, as a means to degrade and import nutrients into the bacterial cell for growth, or both. We have shown here that *B. fragilis* specifically binds intestinal mucin (although *B. fragilis* may also bind to other components in the mucus) and associates with the mucus layer *in vivo*. These findings, along with previous work, suggest that specific interactions between *B. fragilis* and mucus are relevant for *in vivo* colonization of animals. The identity of dedicated mucin-binding receptor(s), and a molecular mechanism during long-term association of the mammalian gut, await discovery.

Acknowledgments

We are grateful to Dr. William Clemons, Jr. (Caltech) and Justin Chartron (Caltech) for help with mucin purification. S.K.M. is a Searle Scholar. Work in the laboratory of the authors is supported by funding from the National Institutes of Health (DK078938, DK083633), Damon Runyon Cancer Research Foundation and the Crohn's and Colitis Foundation of America to S.K.M.

References

[1] Hooper LV, Gordon JI. Commensal host-bacterial relationships in the gut. Science 2001;292:1115–8.
[2] Macpherson AJ, Harris NL. Interactions between commensal intestinal bacteria and the immune system. Nat Rev Immunol 2004;4:478–85.
[3] Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol 2009;9:313–23.
[4] Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 2006;124:837–48.
[5] Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. Science 2008;320:1647–51.
[6] Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell 2005;122:107–18.
[7] Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 2008;453:620–5.
[8] Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A 2010;107:12204–9.
[9] Liu CH, Lee SM, Vanlare JM, Kasper DL, Mazmanian SK. Regulation of surface architecture by symbiotic bacteria mediates host colonization. Proc Natl Acad Sci U S A 2008;105:3951–6.
[10] Coyne MJ, Chatzidaki-Livanis M, Paoletti LC, Comstock LE. Role of glycan synthesis in colonization of the mammalian gut by the bacterial symbiont *Bacteroides fragilis*. Proc Natl Acad Sci U S A 2008;105:13099–104.
[11] Macfarlane S, Dillon JF. Microbial biofilms in the human gastrointestinal tract. J Appl Microbiol 2007;102:1187–96.

[12] Sonnenburg JL, Angenent LT, Gordon JI. Getting a grip on things: how do communities of bacterial symbionts become established in our intestine? Nat Immunol 2004;5:569–73.
[13] Cohen PS, Laux DC. Bacterial adhesion to and penetration of intestinal mucus in vitro. Methods Enzymol 1995;253:309–14.
[14] Shekels LL, Lyftogt C, Kieliszewski M, Filie JD, Kozak CA, Ho SB. Mouse gastric mucin: cloning and chromosomal localization. Biochem J 1995;311:775–85.
[15] Kuwahara T, Yamashita A, Hirakawa H, Nakayama H, Toh H, Okada N, et al. Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation. Proc Natl Acad Sci U S A 2004;101:14919–24.
[16] Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, Martens EC, et al. Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol 2007;5:e156.
[17] Nakayama-Imaohji H, Hirakawa H, Ichimura M, Wakimoto S, Kuhara S, Hayashi T, et al. Identification of the site-specific DNA invertase responsible for the phase variation of SusC/SusD family outer membrane proteins in *Bacteroides fragilis*. J Bacteriol 2009;191:6003–11.
[18] Roberton AM, Stanley RA. In vitro utilization of mucin by *Bacteroides fragilis*. Appl Environ Microbiol 1982;43:325–30.
[19] Tsai HH, Sunderland D, Gibson GR, Hart CA, Rhodes JM. A novel mucin sulphatase from human faeces: its identification, purification and characterization. Clin Sci (Lond) 1992;82:447–54.
[20] Macfarlane GT, Gibson GR. Formation of glycoprotein degrading enzymes by *Bacteroides fragilis*. FEMS Microbiol Lett 1991;61:289–93.
[21] Koropatkin N, Martens EC, Gordon JI, Smith TJ. Structure of a SusD homologue, BT1043, involved in mucin O-glycan utilization in a prominent human gut symbiont. Biochemistry 2009;48:1532–42.

Appendix II:

The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota

June L. Round, S. Melanie Lee, Jennifer Li, Gloria Tran, Bana Jabri, Talal A. Chatila, and Sarkis K. Mazmanian

from NIH (DK30292, DK70977, DK078669, and UL1 RR024992), the Crohn's and Colitis Foundation of America, and NIH Institutional Training Grant T32-A1007172 (to B.D.M.). The National Center for Biotechnology Information Sequence Read Archive accession number for the bacterial 16S rRNA and shotgun data sets related to our studies of mammalian and Calorie Restriction Society member human fecal microbiomes is SRA030940. Sequence data have also been deposited in MG-Rast with project accession numbers qiime:625, qiime:626, qiime:627, and qiime:628.

Supporting Online Material
www.sciencemag.org/cgi/content/full/332/6032/970/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S11
References (*19*–*34*)

5 October 2010; accepted 13 April 2011
10.1126/science.1198719

The Toll-Like Receptor 2 Pathway Establishes Colonization by a Commensal of the Human Microbiota

June L. Round,[1]* S. Melanie Lee,[1] Jennifer Li,[1] Gloria Tran,[1] Bana Jabri,[2] Talal A. Chatila,[3] Sarkis K. Mazmanian[1]*

Mucosal surfaces constantly encounter microbes. Toll-like receptors (TLRs) mediate recognition of microbial patterns to eliminate pathogens. By contrast, we demonstrate that the prominent gut commensal *Bacteroides fragilis* activates the TLR pathway to establish host-microbial symbiosis. TLR2 on $CD4^+$ T cells is required for *B. fragilis* colonization of a unique mucosal niche in mice during homeostasis. A symbiosis factor (PSA, polysaccharide A) of *B. fragilis* signals through TLR2 directly on $Foxp3^+$ regulatory T cells to promote immunologic tolerance. *B. fragilis* lacking PSA is unable to restrain T helper 17 cell responses and is defective in niche-specific mucosal colonization. Therefore, commensal bacteria exploit the TLR pathway to actively suppress immunity. We propose that the immune system can discriminate between pathogens and the microbiota through recognition of symbiotic bacterial molecules in a process that engenders commensal colonization.

Throughout our lives, we continuously encounter microorganisms that range from those essential for health to those causing death (*1*). Consequently, our immune system is charged with the critical task of distinguishing between beneficial and pathogenic microbes. Toll-like receptors (TLRs) are evolutionarily conserved molecules that promote immune responses, and TLR signaling by innate immune cells is indispensable for proper activation of the immune system during infections. T cells also express functional TLRs (*2*–*4*), and TLR signaling has furthermore been shown to restrain immune responses (*5*). As symbionts and pathogens produce similar molecular patterns that are sensed by TLRs, the mechanisms by which our immune system differentiates between the microbiota and enteric infections remain unknown.

Whereas the intestinal microbiota contains hundreds of bacterial species and is integral to human health (*6*), the mucosal immune system employs an arsenal of responses to control enteric pathogens. Germ-free mice lack proinflammatory T helper 17 (T_H17) cells in the gut (*7*, *8*) (Fig. 1A), and only select symbiotic bacteria can induce T_H17 cells (*9*, *10*). Most microbes express common TLR ligands (e.g., peptidoglycan, unmethylated CpG, and lipoproteins); therefore, how do symbionts avoid triggering intestinal immunity in their mammalian hosts? We examined the hypothesis that the human gut commensal *Bacteroides fragilis* evolved molecular mechanisms to suppress T_H17 responses during homeostatic colonization. As predicted previously (*7*, *10*, *11*), we found that *B. fragilis* mono-associated animals did not induce T_H17 cell development in the colon compared to germ-free controls (Fig. 1A). The beneficial contributions of *B. fragilis* require a single immunomodulatory molecule named polysaccharide A (PSA), which prevents and cures inflammatory disease (*12*–*14*). Colonization with *B. fragilis* in the absence of PSA (*B. fragilis*ΔPSA), however, resulted in significant T_H17 cell responses in the gut (Fig. 1, A and B). Colonic lamina propria (LP) lymphocytes from *B. fragilis*ΔPSA mono-associated animals displayed increased secretion of interleukin-17A (IL-17A) in vitro (fig. S1) and elevated transcription of *Il17a* and the T_H17-specific lineage differentiation factor RORγt (*Rorc*) (Fig. 1, C and D) (*15*). Differences in T_H1 responses were not observed during *B. fragilis*ΔPSA colonization (fig. S2). Cells from *B. fragilis* mono-associated animals produced low amounts of IL-17A during in vitro T_H17 skewing assays, whereas $CD4^+$ T cells from *B. fragilis*ΔPSA animals display elevated IL-17A production (fig. S3). Administration of purified PSA to *B. fragilis*ΔPSA mono-associated animals suppresses T_H17 immunity (Fig. 1E). Thus, *B. fragilis* actively restrains T_H17 cell responses during colonization.

Recent studies have shown that certain gut bacteria can promote regulatory T cell (T_{reg}) induction (*11*, *14*). T_{regs} expressing the transcription factor Foxp3 (forkhead box P3) suppress proinflammatory T_H17 cell reactions. To test whether T_{regs} prevent immune responses during *B. fragilis* colonization, we reconstituted germ-free *Rag1*$^{-/-}$ mice with bone marrow from Foxp3-DTR (diphtheria toxin receptor) donors, which allowed for specific ablation of T_{regs} by administration of diphtheria toxin (DT) (*16*). Mice were mono-associated with *B. fragilis* to induce T_{reg} development (Fig. 1F). DT treatment of mice resulted in depletion of $Foxp3^+$ T cells (Fig. 1F), with a concomitant increase in T_H17 responses (Fig. 1G and fig. S4), which suggests that $Foxp3^+$ T_{regs} are required for suppression of T_H17 cells during *B. fragilis* colonization.

PSA is an immunomodulatory bacterial molecule that shapes host immune responses (*17*). Induction of IL-10 and interferon-γ (IFN-γ) from $CD4^+$ T cells by PSA requires TLR2 signaling (*14*, *18*). We sought to determine the mechanism whereby *B. fragilis* suppresses T_H17 cell responses by testing whether PSA functions through TLR2 signaling by dendritic cells (DCs) and/or $CD4^+$ T cells. PSA elicited a significant increase in IL-10 and IFN-γ production from mixed cultures of wild-type DCs and wild-type $CD4^+$ T cells in vitro (Fig. 2A and fig. S5). When *Tlr2*$^{-/-}$ T cells were cocultured with wild-type DCs, however, PSA-induced IL-10 production was reduced, whereas IFN-γ expression was not affected (Fig. 2A and fig. S5), which indicated that PSA required TLR2 expression on T cells to promote IL-10 production. IL-10 responses to PSA were specific to T cells (fig. S6). Consistent with previous findings (*18*), proinflammatory IFN-γ production was dependent on TLR2 signaling by DCs (fig. S5); however, IL-10 production was unaffected in cultures containing wild-type $CD4^+$ T cells and *Tlr2*$^{-/-}$ DCs (Fig. 2A). Therefore, TLR2 expression by T lymphocytes is necessary for IL-10 production by PSA.

$CD4^+$ T cells produce IL-10 in response to PSA in the absence of antigen-presenting cells (APCs) (Fig. 2B). Moreover, PSA induces IL-10 expression from purified T cells in a dose-dependent manner, whereas other TLR2 ligands do not (fig. S7). TLR2 can function as either a homodimer or a heterodimer with TLR1 or TLR6 (*19*). PSA could induce high amounts of IL-10 from wild-type, *Tlr1*$^{-/-}$, and *Tlr6*$^{-/-}$ $CD4^+$ T cells; however, IL-10 production was lost only from *Tlr2*$^{-/-}$ $CD4^+$ T cells and T cells deleted in the TLR adapter molecule MyD88 (Fig. 2B). To determine T_{reg} suppression as a function of cell-

[1]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA. [2]Department of Pathology, Department of Pediatrics, Department of Medicine, University of Chicago, Chicago, IL 60637, USA. [3]Division of Immunology, Allergy and Rheumatology, Department of Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA.

*To whom correspondence should be addressed. E-mail: jround@caltech.edu (J.L.R.); sarkis@caltech.edu (S.K.M.)

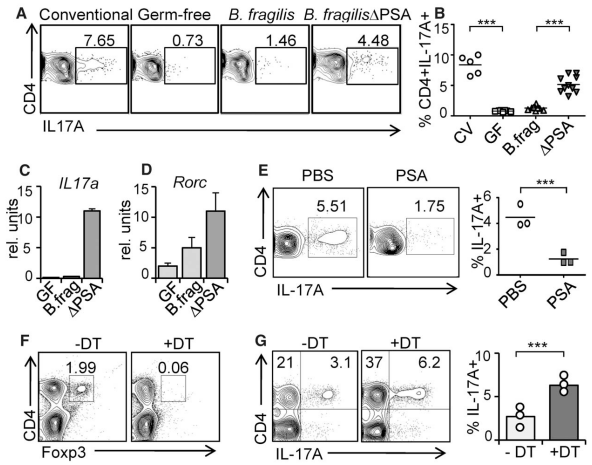


Fig. 1. PSA actively suppresses T_H17 cell development during *B. fragilis* colonization through Foxp3$^+$ T_{regs}. (**A**) Colonic lamina propria lymphocytes (LPLs) were harvested and stained with antibodies against CD4 and IL-17A and analyzed by flow cytometry. Numbers indicate the percentage of CD4$^+$IL-17A$^+$ (T_H17) cells. Conventional mice are specific pathogen free. (**B**) Compiled data from three independent experiments as in (A). CV, conventional; GF, germ-free; B.frag, *B. fragilis*; ΔPSA, *B. fragilis*ΔPSA. ***$P < 0.001$, two-way analysis of variance. (**C** and **D**) CD4$^+$ T cells were isolated from the mesenteric lymph nodes of the indicated animals. RNA was collected and used as a template to determine the relative levels of IL-17A (C) and RORγt (D) transcript. Error bars represent SDs from triplicate samples. The data are representative of three independent experiments. (**E**) *B. fragilis*ΔPSA mono-associated mice were treated with either phosphate-buffered saline (PBS) or PSA, and the LPLs were isolated and the percentage of CD4$^+$ IL-17A–producing cells was determined by flow cytometry. Each symbol represents an individual animal (n = 3 to 4 mice per group). ***$P < 0.001$. (**F** and **G**) Germ-free *Rag1*$^{-/-}$ animals were reconstituted with bone marrow from Foxp3-DTR mice and then mono-associated with *B. fragilis*. Animals were treated with either PBS (–DT) or diphtheria toxin (+DT) as described (*16*). Colonic LPLs were harvested after T_{reg} ablation and restimulated with phorbol 12-myristate 13-acetate (PMA)–ionomycin and brefeldin A. Cells were stained with antibodies to CD4, Foxp3, and IL-17A and analyzed by flow cytometry. (Right) Symbols represent T cell proportions from individual mice within a single experiment (n = 3 to 4 mice per group) and are representative of two independent trials. ***$P < 0.001$.

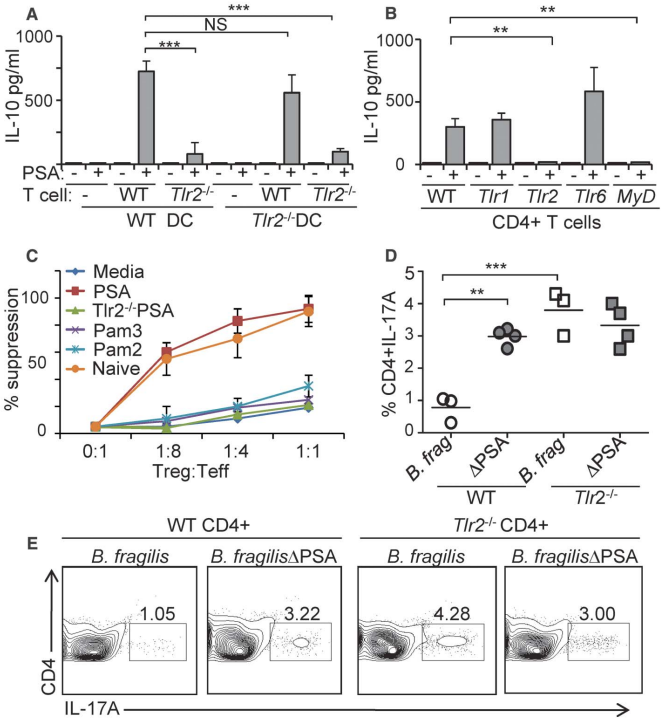


Fig. 2. PSA signals through TLR2 on CD4$^+$ T cells to suppress T_H17 cell responses. (**A**) Bone marrow–derived dendritic cells from wild-type (WT) or *Tlr2*$^{-/-}$ mice were incubated with splenic CD4$^+$ T cells. Amounts of secreted IL-10 were determined by enzyme-linked immunosorbent assay (ELISA). Error bars represent SDs from two independent assays performed in quadruplicate. ***$P < 0.001$; NS, not significant. (**B**) CD4$^+$ T cells were isolated from WT mice, and the indicated knockout mice and cells were stimulated as in (A). MyD is *Myd88*$^{-/-}$. IL-10 was assayed by ELISA. **$P < 0.01$. Error bars represent SDs for quadruplicate samples and are representative of two independent trials. (**C**) CD4$^+$Foxp3$^+$ T_{regs} were purified from Foxp3^{EGFP} mice (*26*) and *Tlr2*$^{-/-}$ X Foxp3^{EGFP} mice and stimulated with plate-bound antibodies against CD3 and recombinant TGF-β, with PSA or with the indicated TLR ligands. Equal numbers of live cultured T_{regs} were then incubated with CFSE (carboxyfluorescein diacetate succinimidyl ester)–pulsed responder cells (CD4$^+$Foxp3$^-$). Percent suppression is determined by the ratio of proliferating responder cells in each condition relative to proliferation in the absence of added T_{regs}. Error bars are SDs from a single experiment performed in duplicate and are representative of two independent trials. (**D** and **E**) Germ-free *Rag1*$^{-/-}$ animals were reconstituted with CD4$^+$ T cells from WT or *Tlr2*$^{-/-}$ mice and then mono-associated with either WT *B. fragilis* or *B. fragilis*ΔPSA. Colonic LPLs were isolated and analyzed for T_H17 cell proportions by flow cytometry. Plots are gated on CD4$^+$ cells. (D) Each symbol represents an individual animal (n = 3 to 4 mice per group), and data are representative of two independent trials. **$P < 0.01$; ***$P < 0.001$.

intrinsic TLR2 signaling, we measured proliferation of responder T cells after coculture with T_{regs} stimulated in vitro with PSA, or classical TLR2 ligands Pam3CysK or Pam2CysK. The proportion of Foxp3$^+$ T cells was equivalent under all conditions (fig. S8). PSA-treated T_{regs} displayed increased suppressive capacity compared to media or other TLR2 ligand–treated T_{regs} (Fig. 2C). IL-2 was not produced by T_{regs} during treatment with any of the TLR2 ligands, and IL-2 neutralization had no effect on in vitro suppression (fig. S9). Notably, the suppressive capacity of PSA-treated T_{regs} was lost when Foxp3$^+$ T cells were deficient in TLR2 (Fig. 2C). PSA likely directed the development of inducible T_{regs} by promoting the expression of IL-10, transforming growth factor–β2 (TGF-β2), and Foxp3 from purified Foxp3$^+$ T cells in a TLR2-dependent fashion (fig. S10). Collectively, these studies show that unlike other TLR2 ligands, PSA enhances T_{reg} function and gene expression in the absence of APCs through TLR2 signaling directly on CD4$^+$Foxp3$^+$ Treg cells.

We next determined whether PSA signals through TLR2 on Foxp3$^+$ T_{regs} during *B. fragilis* colonization of animals to suppress T_H17 cell responses. Germ-free *Rag1*$^{-/-}$ mice were reconstituted with purified CD4$^+$ T cells from wild-type or *Tlr2*$^{-/-}$ animals. Animals were subsequently mono-associated with either *B. fragilis* or *B. fragilis*ΔPSA. Unlike wild-type animals, IL-10–producing Foxp3$^+$ T_{regs} were not induced in the gut of *B. fragilis* mono-colonized animals that were reconstituted with *Tlr2*$^{-/-}$ CD4$^+$ T cells (fig. S11). Although we observed minimal T_H17 cell development in wild-type mice mono-associated with *B. fragilis*, T_H17 cell responses were significantly increased during colonization of animals that were reconstituted with *Tlr2*$^{-/-}$ CD4$^+$ T cells (Fig. 2, D and E). Furthermore, TLR2-deficient CD4$^+$ T cells from *B. fragilis* mono-associated mice produced more IL-17A compared to mice reconstituted with wild-type CD4$^+$ T cells (fig. S12). Collectively, these data demonstrate that *B. fragilis* requires TLR2 to induce Foxp3$^+$ T_{regs} during intestinal colonization and actively suppresses T_H17 responses through engagement of TLR2 specifically on T cells.

The intestinal microbiota occupies both mucosal and luminal niches during normal colonization; however, the biogeographic distributions of specific microbial species are poorly characterized. We reasoned that the functions of PSA were driven by an evolutionary impetus to prevent deleterious mucosal immune reactions to *B. fragilis*, enabling bacteria to associate with host tissue. Intriguingly, we discovered a population of bacteria that intimately associates with the intestinal epithelium (Fig. 3A). Whole-mount preparations of medial colons were probed for bacteria after labeling with *B. fragilis*–specific antisera (fig. S13). Three-dimensional reconstruction of confocal microscopic images revealed microcolonies of *B. fragilis* residing within colonic crypts (Fig. 3A). The amounts of *B. fragilis* associated with host tissue represented a fraction of total bacteria (Fig. 3, B and C), but likely are an important population that are in close proximity to the host immune system. We speculated that the amounts of tissue-associated bacteria would be sensitive to host immune responses such as T_H17 cell induction. Notably, animals colonized with *B. fragilis*ΔPSA displayed profoundly reduced numbers of tissue-associated bacteria when compared to animals colonized with wild-type *B. fragilis* (Fig. 3D). Treatment of *B. fragilis*ΔPSA–colonized animals with purified PSA corrected this defect and increased colonization of *B. fragilis*ΔPSA to wild-type bacterial levels (Fig. 3D). Only tissue-associated bacteria were affected, because no differences were observed in the amounts of bacteria in the gut lumen by either strain (fig. S14) (*17*). Collectively, these data identify a previously unappreciated mucosal niche for *B. fragilis* and reveal that PSA is required for maintaining host-bacterial symbiosis at the epithelial surface of the gut.

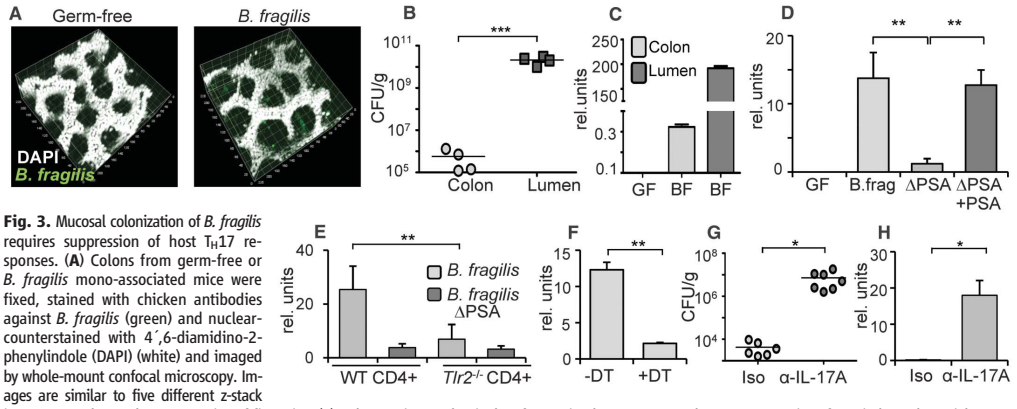


Fig. 3. Mucosal colonization of *B. fragilis* requires suppression of host T_H17 responses. (**A**) Colons from germ-free or *B. fragilis* mono-associated mice were fixed, stained with chicken antibodies against *B. fragilis* (green) and nuclear-counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (white) and imaged by whole-mount confocal microscopy. Images are similar to five different z-stack images per colon and representative of five mice. (**B**) Colon sections or luminal contents from *B. fragilis* mono-associated mice were homogenized and serially diluted to obtain live bacterial counts. CFUs (colony-forming units) per gram of tissue were determined after microbiologic plating. Each symbol represents an individual animal (n = 3 to 4 animals per group) and is representative of three independent trials. ***P < 0.001. (**C**) Quantitative reverse transcription–polymerase chain reaction (qRT-PCR) analysis was performed with *Bacteriodes*-specific primers on RNA extracted from colon tissue or luminal contents. GF, germ-free; BF, *B. fragilis*. Error bars represent SDs from individual mice in the same experiment and are representative of two independent trials. (**D**) qRT-PCR analysis for *B. fragilis* was performed on RNA extracted from colon homogenates from indicated animals. The bar furthest to the right shows colonization of *B. fragilis*ΔPSA in animals orally treated with purified PSA. GF, germ-free; B.frag, *B. fragilis*; ΔPSA, *B. fragilis*ΔPSA. Data are shown for four animals per group and are representative of two independent trials. **P < 0.01. (**E**) Germ-free *Rag1*$^{-/-}$ animals were reconstituted with *Tlr2*$^{-/-}$ or WT CD4$^+$ T cells and colonized with either WT *B. fragilis* or *B. fragilis*ΔPSA. Colons were prepared and analyzed as in (D). **P < 0.01. (**F**) Germ-free *Rag1*$^{-/-}$ animals were reconstituted with Foxp3-DTR bone marrow and colonized with *B. fragilis*. Two months after reconstitution animals were treated with either PBS (–DT) or with diphtheria toxin (+DT), and colons were prepared as described in (D). **P < 0.01. (**G** and **H**) Neutralization of IL-17A increases *B. fragilis* colonization. Germ-free animals were colonized with *B. fragilis*ΔPSA and treated with either an antibody that neutralizes IL-17A (α-IL-17A) or an isotype control (Iso). Colon homogenates were analyzed by live bacterial plating (G) or qRT-PCR (H) as described in (B) and (C). Each symbol in (G) represents an individual animal. Error bars in (H) show the SDs from individual animals and are compiled data from two independent trials with three or four animals per group. *P < 0.05.

Our findings suggest that PSA induces T_{regs} through TLR2 signaling to suppress T_H17 cell responses and promote mucosal colonization by *B. fragilis*. To test this model, we measured colonization levels of *B. fragilis* in *Rag1*$^{-/-}$ mice reconstituted with TLR2-deficient CD4$^+$ T cells. Tissue association by wild-type *B. fragilis* in the colon was reduced to the levels of *B. fragilis*ΔPSA in these mice (Fig. 3E and fig. S15). Moreover, Foxp3$^+$ T_{reg} ablation in *B. fragilis* mono-associated animals resulted in significantly reduced amounts of tissue-associated *B. fragilis* (Fig. 3F), without affecting bacterial numbers in the lumen of the gut (fig. S16). Finally, to functionally determine the role of IL-17 responses in mucosal association, we treated *B. fragilis*ΔPSA mono-associated animals with a neutralizing antibody to IL-17A. Whereas the amounts of *B. fragilis*ΔPSA in isotype control–treated animals remained low, neutralization of IL-17A resulted in a 1000-fold increase in tissue-associated bacteria (Fig. 3, G and H). These data indicate that IL-17 suppression by PSA is required by *B. fragilis* during association with its host. Therefore, unlike pathogens that trigger inflammatory responses through TLRs to clear infections, symbiotic colonization by *B. fragilis* is actually enhanced via the TLR pathway. We conclude that PSA evolved to engender host-bacterial mutualism by inducing mucosal tolerance through TLR2 activation of T_{reg} cells.

The gastrointestinal tract represents a primary portal for entry by numerous pathogens. Toll-like receptors recognize MAMPs (microbial-associated molecular patterns) expressed by bacteria and coordinate a cascade of innate and adaptive immune responses that control infections (*20*). Although TLRs have classically been studied on innate immune cells, recent reports have demonstrated their expression by T cells in both mice and humans (*4*, *21*–*23*). As bacteria contain universally conserved MAMPs, how do commensal microbes, unlike pathogens, avoid triggering TLR activation? It is historically believed that the microbiota is excluded from the mucosal surface (*24*). However, certain symbiotic bacteria tightly adhere to the intestinal mucosa (*9*–*11*), and thus immunologic ignorance may not explain why inflammation is averted by the microbiota. Our study provides new insight into the mechanisms by which the immune system distinguishes between pathogens and symbionts. The functional activity of PSA on T_{regs} contrasts with the role of TLR2 ligands of pathogens, which elicit inflammation, and thus reveals an unexpected function for TLR signaling during homeostatic intestinal colonization by the microbiota. Although engagement of TLR2 by previously identified ligands is known to stimulate microbial clearance of pathogens, TLR signaling by PSA paradoxically allows *B. fragilis* persistence on mucosal surfaces. These results identify PSA as the incipient member of a new class of TLR ligands termed "symbiont-associated molecular patterns (SAMPs)" that function to orchestrate immune responses to establish host-commensal symbiosis. On the basis of the importance of the microbiota to mammalian health (*25*), evolution appears to have created molecular interactions that engender host-bacterial mutualism. In conclusion, our findings suggest that animals are not "hard-wired" to intrinsically distinguish pathogens from symbionts, and that microbial-derived mechanisms have evolved to actively promote immunologic tolerance to symbiotic bacteria. This concept suggests a reconsideration of how we define self versus nonself.

References and Notes

1. R. E. Ley, D. A. Peterson, J. I. Gordon, *Cell* **124**, 837 (2006).
2. S. Manicassamy, B. Pulendran, *Semin. Immunol.* **21**, 185 (2009).
3. M. Fukata *et al.*, *J. Immunol.* **180**, 1886 (2008).
4. J. M. Reynolds *et al.*, *Immunity* **32**, 692 (2010).
5. I. Caramalho *et al.*, *J. Exp. Med.* **197**, 403 (2003).
6. P. B. Eckburg *et al.*, *Science* **308**, 1635 (2005).
7. I. I. Ivanov *et al.*, *Cell Host Microbe* **4**, 337 (2008).
8. K. Atarashi *et al.*, *Nature* **455**, 808 (2008).
9. V. Gaboriau-Routhiau *et al.*, *Immunity* **31**, 677 (2009).
10. I. I. Ivanov *et al.*, *Cell* **139**, 485 (2009).
11. K. Atarashi *et al.*, *Science* **331**, 337 (2011).
12. S. K. Mazmanian, J. L. Round, D. L. Kasper, *Nature* **453**, 620 (2008).
13. J. Ochoa-Repáraz *et al.*, *Mucosal Immunol.* **3**, 487 (2010).
14. J. L. Round, S. K. Mazmanian, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12204 (2010).
15. Materials and methods are available as supporting material on *Science* Online.
16. J. M. Kim, J. P. Rasmussen, A. Y. Rudensky, *Nat. Immunol.* **8**, 191 (2006).
17. S. K. Mazmanian, C. H. Liu, A. O. Tzianabos, D. L. Kasper, *Cell* **122**, 107 (2005).
18. Q. Wang *et al.*, *J. Exp. Med.* **203**, 2853 (2006).
19. C. A. Janeway Jr., R. Medzhitov, *Annu. Rev. Immunol.* **20**, 197 (2002).
20. R. Medzhitov, *Nature* **449**, 819 (2007).
21. H. Liu, M. Komai-Koma, D. Xu, F. Y. Liew, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7048 (2006).
22. R. P. Sutmuller *et al.*, *J. Clin. Invest.* **116**, 485 (2006).
23. S. Babu, C. P. Blauvelt, V. Kumaraswami, T. B. Nutman, *J. Immunol.* **176**, 3885 (2006).
24. L. V. Hooper, *Nat. Rev. Microbiol.* **7**, 367 (2009).
25. K. Smith, K. D. McCoy, A. J. Macpherson, *Semin. Immunol.* **19**, 59 (2007).
26. W. Lin *et al.*, *Nat. Immunol.* **8**, 359 (2007).

Acknowledgments: We thank S. W. McBride and Y. Shen (California Institute for Technology) for help with bacterial colonization and germ-free studies. We are grateful to A. Rudensky [Memorial Sloan-Kettering Cancer Center and Howard Hughes Medical Institute (HHMI)] for the gift of Foxp3-DTR mice and L. Hooper (University of Texas Southwestern and HHMI) for germ-free *Rag1*$^{-/-}$ mice. We thank members of the Mazmanian laboratory for their critical review of the manuscript. B.J. acknowledges support from the Crohn's and Colitis Foundation of America (CCFA) (award 2831) and T.A.C acknowledges support from the NIH (grant AI 080002). J.L.R. is a Merck Fellow of the Jane Coffin Childs Memorial Fund. S.K.M. is a Searle Scholar. This work was supported by funding from the NIH (grants DK 078938, DK 083633, AI 088626), the Damon Runyon Cancer Research Foundation, and the CCFA (award 2405) to S.K.M. J.L.R. and S.K.M. have a patent application (PCT/US2008/082928) on the use of PSA as a therapy for inflammatory bowel disease. The authors have no competing financial interests related to this publication.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1206095/DC1
Materials and Methods
Figs. S1 to S16

15 February 2011; accepted 6 April 2011
Published online 21 April 2011;
10.1126/science.1206095

A Packing Mechanism for Nucleosome Organization Reconstituted Across a Eukaryotic Genome

Zhenhai Zhang,[1]* Christian J. Wippo,[2]* Megha Wal,[1] Elissa Ward,[1] Philipp Korber,[2]† B. Franklin Pugh[1]†

Near the 5′ end of most eukaryotic genes, nucleosomes form highly regular arrays that begin at canonical distances from the transcriptional start site. Determinants of this and other aspects of genomic nucleosome organization have been ascribed to statistical positioning, intrinsically DNA-encoded positioning, or some aspect of transcription initiation. Here, we provide evidence for a different explanation. Biochemical reconstitution of proper nucleosome positioning, spacing, and occupancy levels was achieved across the 5′ ends of most yeast genes by adenosine triphosphate–dependent trans-acting factors. These transcription-independent activities override DNA-intrinsic positioning and maintain uniform spacing at the 5′ ends of genes even at low nucleosome densities. Thus, an active, nonstatistical nucleosome packing mechanism creates chromatin organizing centers at the 5′ ends of genes where important regulatory elements reside.

Statistical positioning depends on the presence of a genomic barrier within a linear array of nucleosomes (*1*). Nucleosomes within the array will passively align at regular intervals from the barrier, independent of sequence or other external factors, rather than arrange randomly. Nucleosome organization in vivo displays patterns that are consistent with statistical