Discovering Functions of Host-Associated Microbiomes using Metabolic Footprinting

By

CONNOR REID TIFFANY
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Andreas J. Bäumler, Chair

Charles L. Bevins

Douglas C. Nelson

Committee in Charge

2023

Acknowledgement and Dedication

I would like to thank my PhD mentor, Dr. Andreas Baumler for providing so much support and guidance throughout my PhD while allowing me to independently choose and develop projects that interested me the most. I would also like to extend my thanks to all my colleagues and collaborators who have helped me grow as a scientist, including Baumler lab alumni and colleagues at the INRAE of Jouy en Josas and the INRIA at Bordeaux during my time working in France.

Next, I would like to thank my friends and family who have always been supportive of me and celebrated my accomplishments (because I'm terrible at doing that myself). Most of all, my big sister Caitlin. I wouldn't be the scientist, or person I am today without her.

Table of Contents

Abstract

Advances in DNA sequencing technologies have galvanized research efforts into understanding the effects of host-associated microbiomes on host health. While a staggering amount of data has been generated to describe taxonomic differences between “healthy” and “dysbiotic” microbiomes, the functions performed by microbiomes during homeostasis remain incompletely understood. This in turn makes understanding dysbiosis in the context of different etiologies difficult. Thus, it is paramount to develop new frameworks to expand our knowledge on the beneficial functions of host-associated microbiomes. To accomplish that goal, we created metabolic footprinting, a methodological framework which utilizes gnotobiotics, comparative-untargeted metabolomics, and growth assays to determine what microbes consume *in vivo*. Using metabolic footprinting, we determined that commensal *Clostridia* perform an important digestive function for the mammalian host by consuming sugar alcohols in the large bowel. Sugar alcohol intolerance is commonly reported by patients with irritable bowel syndrome or inflammatory bowel disease, suggesting the loss of *Clostridia* may be an important factor in these etiologies. Additionally, we used metabolic footprinting to discover that the human fungal pathogen *Candida albicans* consumes small sugars in the gut. Interestingly, *C. albicans* required oxygen to utilize these sugars *in vitro*, and we found that the gut microbiota prevented the expansion of *C. albicans* in the gut by restricting access to oxygen *in vivo*. The finding that oxygen is a critical resource for *C. albicans* in the large bowel could improve prophylaxis aimed at preventing invasive

candidiasis in susceptible patients. Collectively, the results in this thesis demonstrate that metabolic footprinting is a useful tool for elucidating the functions performed by host-associated microbiomes.

Chapter 1

Introduction: Dysbiosis: from fiction to function

Connor R. Tiffany and Andreas J. Bäumler

Department of Medical Microbiology and Immunology, School of Medicine, University of California Davis, Davis, CA, USA.

Published in

Tiffany CR, Bäumler AJ. Dysbiosis: from fiction to function. Am J Physiol Gastrointest Liver Physiol. 2019 Nov 1;317(5):G602-G608. doi: 10.1152/ajpgi.00230.2019. Epub 2019 Sep 11. PMID: 31509433; PMCID: PMC6879887.

Abstract

Advances in data collection technologies reveal that an imbalance (dysbiosis) in the composition of host-associated microbial communities (microbiota) is linked to many human illnesses. This association makes dysbiosis a central concept for understanding how the human microbiota contributes to health and disease. However, it remains problematic to define the term dysbiosis by cataloguing microbial species names. Here we discuss how incorporating the germ-organ concept, ecological assumptions and immunological principles into a theoretical framework for microbiota research provides a functional definition for dysbiosis. The generation of such a framework suggests that the next logical step in microbiota research will be to illuminate the mechanistic underpinnings of dysbiosis, which often involves a weakening of immune mechanisms that balance our microbial communities.

Text

In the wake of advances in data collection and analysis technologies, the 21st century has seen the birth of a new discipline devoted to studying microbial communities inhabiting our body (the human microbiota). Pioneering work cataloguing microbial species reveals that an imbalance in the human microbiota is linked to various non-communicable diseases (11), thereby raising the hope that investment into microbiota research might yield new therapeutic approaches for many human illnesses. However, there is a growing consensus that information gained from cataloguing microbial species names is reaching an asymptote, where additional work will increase the size of available databases without bringing about revolutionary advances in understanding what constitutes a healthy microbiome (54). Considering this problem in microbiome research, we should reflect on how we arrived at this juncture and assess where the field may be headed.

Cataloguing the microbiome to define homeostasis and dysbiosis

One reason for the strong emphasis microbiota research has placed on cataloguing microbial genes and species names is the concept that host-associated microbial communities form an entity that is separate from the rest of the body (54). This viewpoint is reflected in an early definition of the term microbiome credited to Joshua

Lederberg, who suggested that the expression should be used to describe the collective genome of our resident microbes (28). Building on this concept, the microbiome is commonly defined as the microbiota, its genes and gene products, which puts the focus entirely on the microbes (**Fig. 1.1A**). This inward-looking view of host-associated microbial communities nurtures an expectation that cataloguing microorganisms, microbial genes and microbiota-derived metabolites in sufficient depth will provide a complete understanding of the microbiome eventually. To accomplish this goal, the Human Microbiome Project, launched in 2007 by the National Institute of Health, set out to identify elements that are common to a healthy microbiome (30, 31). Initial efforts to catalogue genes present in the large intestine failed to define a core microbiome (74), which prompted calls for additional studies with increased sample sizes (44). This call was heeded through the generation of 42 terabytes of multi-omic data by the Human Microbiome Project, which yielded a wealth of community resources (31). However, the remarkable depth of microbiota measurements achieved by the Human Microbiome Project in 2019 did not reveal what constitutes a healthy microbiome (54), which illustrates that this approach has limitations. As cataloguing microbial species and their genes is not revealing what represents a balanced microbial community, some go so far as to suggest that the balance concept is a holdover from prescientific thought, as it cannot be measured and is therefore not useful for microbiome research (52).

Our inability to define balance in turn makes it problematic to specify what constitutes an imbalance in the microbiota, commonly referred to as dysbiosis, a major

organizing concept in microbiome research (52). Dysbiosis can be described as a compositional and functional alteration in the microbiota in individuals with disease compared to healthy subjects (35). Dysbiosis can feature a loss of beneficial microorganisms, an expansion of potentially harmful microbes and/or a loss of overall microbial diversity (53). But since researchers still do not agree on what constitutes a healthy microbiome, it is not clear how to define an impaired one (54). Despite its elusiveness, the term dysbiosis is gaining popularity among scientists, as indicated by a rapidly growing number of articles using this expression (**Fig. 2**), including some published by the authors of this review. Recently, after publishing an article discussing the dysbiosis concept (38), one of us received an email from an eminent microbiome researcher urging us not to use this "*unscientific*" expression because "*the gut microbiota is dynamic and changing and the term dysbiosis cannot be applied to it, it has no units and cannot even be defined*". We believe this criticism points to the larger underlying problem that data-driven research has been ineffective in translating terabytes of information into a theoretical framework. As a result, the terminology and organizing concepts generated through technology-based empiricism remain elusive (7, 52). Increasing the amount of available multi-omic microbiome profiling data from terabytes to petabytes holds little promise for clarifying how to define dysbiosis, which raises the question where microbiome science should go from here.

A case for changing the course of microbiome science by widening its scope

As cataloguing the microbiome in greater depth is not likely to provide decisive breakthroughs, the elephant in the room becomes that microbiome research has to move in a new direction to meet expectations to generate novel interventions for treating diseases that are linked to dysbiosis. At the conclusion of the Human Microbiome Project, Gilbert and Lynch wrote a perspective on the future direction of microbiota research, which highlights the necessity to incorporate community ecology as a framework for microbiome studies (24). There is a growing appreciation for the need to understand the ecological and evolutionary relationships microbes have with each other and with their hosts (54). Notably, studies investigating both the human host and the microbiota are rare, but neglecting the host omits how we interact with our resident microbial communities (41). Additionally, the host should be considered a foundation species because of its substantial role in structuring the ecosystem through nutrient flow, the immune system, and by comprising many of its habitats. Thus, omitting the host from microbiome studies is inadmissible within this framework. Incorporating ecological principles into microbiota research provides an important impetus for expanding research beyond the microbes to include an influence from the host, because in ecological terms the microbiome is defined as the microbiota and its environment (69) (**Fig. 1.1B**). This is a much broader definition than the one suggested by Joshua Lederberg (28), as it includes the environment inhabited by the microbiota, which is shaped by the human host (10, 17). In other words, incorporating ecology as a theoretical framework suggests that the microbiome is a host-microbe chimera composed of the microbiota and host factors that influence the microbial environment (8). Thus, whereas cataloguing microbial species and their genes can shine

light on an important piece of the microbiome puzzle (**Fig. 1.1A**), this approach alone cannot provide a complete picture of the microbiome (**Fig. 1.1B**), which explains why further increasing the depth of microbiota measurements is not likely to create future breakthroughs. To obtain a complete picture of the microbiome it will be necessary to change our approach by incorporating how the host shapes the environment inhabited by the microbiota into a conceptual framework.

The behavior of the host, including choice of habitat and diet, influences microbiota composition by altering host habitat patches in the intestine (47). In addition, host cells and host tissues employ numerous control mechanisms that serve as habitat filters for sculpting the microbiota (**Fig. 1.1B**). Some host control mechanisms limit the size of microbial communities, which includes physical defenses, such as the mucus escalator in the lung (55), or non-specific chemical defenses, such as gastric acid (61). Paneth cells of the small bowel can influence the biogeography of microbial communities by secreting antimicrobial peptides, such as human defensin 5, which limits growth of segmented filamentous bacteria (phylum *Firmicutes*) that occupy a spatial niche in direct contact with the epithelial surface (60). Similarly, MUC2 mucin secretion by goblet cells shapes the biogeography in the large intestine by forming a densely-packed inner mucus layer that is largely devoid of bacteria (32). Finally, some host control mechanisms function in shaping the species composition of the microbiota, which is exemplified by the breakdown of host-derived glycogen by human alpha-amylase in vaginal mucosal fluid to nourish members of the genus *Lactobacillus* (phylum *Firmicutes*) (66), a taxon that dominates the

microbiota in the female reproductive tract (13, 56). Whereas this list of examples is by no means complete, it will suffice to illustrate that the host has control over the environment inhabited by the microbiota (10). The question arising from this concept is for what purpose does the host control the size, biogeography and species composition of the microbiota.

Acknowledging host control over the ecosystem defines dysbiosis functionally

Ecological theory suggests that host control over the microbial environment shapes the microbiota to be beneficial (17). In some cases, host control mechanisms that limit the size of microbial communities or keep microorganisms at a distance from the epithelium might serve simply to prevent harm to mucosal surfaces. However, microbial communities also provide direct benefit to the host and these functions need to be maintained to preserve homeostasis. For instance, obligate anaerobic bacteria belonging to the classes *Clostridia* (phylum *Firmicutes*) and *Bacteroidia* (phylum *Bacteroidetes*), which dominate the microbiota of the large intestine (30), perform an important digestive function for the host. Complex carbohydrates, which escape digestion by host enzymes in the upper gastrointestinal tract, are broken down in the colon by *Clostridia* and *Bacteroidia* into fermentation products, such as short-chain fatty acids, which are absorbed by the host for nutrition (14, 16, 72). Furthermore, microbial metabolites, such as short-chain fatty acids, also provide benefit by contributing to immune education (2, 4,

21, 64). Due to the benefit it provides to the host, the colonic microbiota could be viewed as an organ-like collection of microbes (11, 45, 51). To ensure functionality of its microbial organ, the host limits the amount of oxygen diffusing into the intestinal lumen by maintaining the colonic epithelium in a state of physiological hypoxia (<1% O_2) (75). Hypoxia is actively maintained through epithelial PPARγ (peroxisome proliferator-activated receptor gamma)-signaling, which increases cellular oxygen consumption by activating mitochondrial oxidative phosphorylation (9). Since oxygen can diffuse across biological membranes, epithelial hypoxia limits the amount of oxygen entering the intestinal lumen, which helps to maintain anaerobiosis, thereby driving a dominance of obligate anaerobic *Clostridia* and *Bacteroidia* (9, 58) (**Fig. 3**). This series of events enables the host to maintain functionality of its microbial organ by shaping the composition of the gut microbiota towards a dominance of species capable of performing digestive functions that confer benefit to the host, a concept termed the germ-organ theory (7). Notably, the germ-organ theory provides a functional definition for what constitutes a healthy microbiome in the colon by suggesting that the host maintains homeostasis by limiting the availability of a critical resource, oxygen, to preserve the digestive function of its microbial organ (7).

Changes in energy metabolism that reduce oxygen consumption in the colonic epithelium are associated with a loss of epithelial hypoxia, which in turn disrupts anaerobiosis in the colon (38). A disruption of anaerobiosis is associated with an expansion of facultative anaerobic bacteria belonging to the phylum *Proteobacteria* (57),

which can use oxygen to outgrow obligate anaerobic *Clostridia* and *Bacteroidia* in the colon (9, 25, 29, 42, 58). An expansion of facultative anaerobic *Proteobacteria* is a microbial signature of dysbiosis in the colon (62) (**Fig. 3**). As *Proteobacteria* lack the capability of breaking down complex carbohydrates (16), their expansion provides no digestive benefit to the host. The emerging picture suggests that epithelial hypoxia is a host component of the microbiome that maintains homeostasis in the colon, whereas changes in epithelial energy metabolism that impair this host control mechanism trigger dysbiosis (39). Thus, the colonic epithelium functions as a control switch that governs a shift between homeostatic and dysbiotic microbial communities (38). These considerations identify dysbiosis in the colon as a microbial signature of a weakened epithelial control over the microbial ecosystem (39), a functional definition that becomes measurable, for example, by quantifying epithelial oxygen consumption (9). In conclusion, by expanding the definition of the microbiome to include host control mechanisms that shape the microbial ecosystem (**Fig. 1.1B**), it becomes possible to define homeostasis and dysbiosis functionally. However, epithelial hypoxia is a host control mechanism that is specific to the colon and digestive functions of the microbiota are limited to the gastrointestinal tract, which raises the question whether functional definitions for homeostasis and dysbiosis can be adapted to other body surfaces.

A quest for causation as the next phase in microbiome research

A benefit conferred by the microbiota that is relevant to all body surfaces is its ability to repel pathogens, a property known as colonization resistance (36, 37). Colonization resistance is part of the larger phenomenon of microbiota resistance, an ability to remain unchanged in the face of a disturbance, which is mediated through mechanisms that are entrenched in community ecology (65). To coexist, each member of the microbiota must occupy a different nutrient-niche defined by critical resources that the occupant can utilize better than any other microorganism present in the microbial community (19). Each occupant of a nutrient-niche gains an advantage over new arrivals that compete for the same resources through niche preemption or niche modification, a principle known as historical contingency (20, 67). Colonization resistance against pathogens is mediated through priority effects resulting from niche preemption or niche modification (40, 70, 71). Whereas niche preemption signifies a competition for resources, niche modification encompasses the production of microbiota-derived metabolites that antagonize pathogens (1, 22). Notably, our microbial organ and host control mechanisms that maintain its functionality could be viewed as being part of our immune system, because colonization resistance against pathogens represents a non-specific immune function (8). This part of the immune system has been dubbed microbiota-nourishing immunity, because some host control mechanisms literally nourish the microbiota, a feature that distinguishes this part of the immune system from sterilizing immunity, which functions in detection and removal of microorganisms from host tissue, but never provides nourishment to microbes (8).

Microbiota-nourishing immunity encompasses the microbiota and host control mechanisms that shape the microbial ecosystem, thereby covering all relevant aspects of the microbiome (8) (**Fig. 1.1B**). This theoretical framework therefore enables microbiome science to adopt well-established immunological concepts, such as immunocompetence and immunodeficiency. For example, using this theoretical framework, a healthy microbiome could be defined simply as a state of immunocompetence in microbiota-nourishing immunity. Conversely, a weakening of microbiota-nourishing immunity could be viewed as an immunodeficiency associated with dysbiosis (8). Consistent with the immunodeficiency concept, a disruption of the microbial component of microbiota-nourishing immunity through antibiotic therapy is associated with opportunistic infections caused by *Clostridioides difficile* (phylum *Firmicutes*) (34). Similarly, a weakening of a host component of microbiota-nourishing immunity in cystic fibrosis, which is caused by mutations in the *CFTR* (cystic fibrosis transmembrane-conductance regulator) gene that impair the mucus escalator, is linked to opportunistic infections with *Pseudomonas aeruginosa* (phylum *Proteobacteria*) (26). Notably, accepting the immunodeficiency concept suggests that dysbiosis is a symptom of a weakening in microbiota-nourishing immunity (8). This concept steers microbiome research in a new direction, because it suggests that a description of dysbiosis should always be followed by studies aimed at identifying the nature of the underlying immunodeficiency. Thus, after the initial cataloguing of microbial genes and species names is complete, the vision of microbiota-nourishing immunity provides microbiome

science with the next goal to seize as it enters its second phase of identifying potential treatment targets for diseases associated with dysbiosis.

Concluding remarks

Increasing the depth of microbiome measurements has failed to provide genuinely needed definitions for homeostasis and dysbiosis (52), which reflects the lack of an overarching conceptual framework in microbiome science (7). Here we discuss the idea that incorporating concepts from community ecology, immunology and the germ-organ theory into microbiome research provides a functional definition for dysbiosis and helps to guide microbiome science towards investigating an underlying weakening in microbiota-nourishing immunity. When cataloguing microbial species names one should keep in mind that not every microbial change is necessarily functionally consequential for the host and that the presence and expression of functionally relevant metabolic pathways might be conserved across species barriers. Our theoretical framework suggests that the resident microbiota is dysbiotic if it exhibits impaired functionality, as indicated by a weakening in colonization resistance or a decline in its ability to perform digestive tasks, immune education or other beneficial duties.

A remaining ambiguity of the terminological framework is that it does not distinguish between dysbiosis being the cause or the result of disease (52). Dysbiosis can be the cause of disease if microbiota-nourishing immunity becomes weakened by a direct

disruption of its microbial component, which is a common side-effect of antibiotic treatment. For instance, fecal microbiota transplants are effective in treating opportunistic *C. difficile* infections (5), because this therapy targets the underlying cause, which is an antibiotic-mediated disruption of the microbiota. However, in case of illnesses in which a host component of microbiota-nourishing immunity is weakened, dysbiosis is secondary to an immunodeficiency of the host (8), which supports the idea that dysbiosis should refer to the result of disease, rather than its cause (52). Between these two extremes lie conditions in which dysbiosis is secondary to an underlying host immunodeficiency, but the resulting change in the microbial community structure exacerbates disease. Thus, it will remain important to determine experimentally in each case whether dysbiosis associated with a human illness makes a causal contribution to the disease process, for example by adopting Koch's postulates for microbiome research (49, 63, 73). Determining the nature of the immunodeficiency that underlies dysbiosis and resolving whether dysbiosis makes a causal contribution to disease will be relevant for identifying new therapeutic targets to treat illnesses associated with changes in the microbiota composition.

Acknowledgements

Work in A.J.B.'s lab is supported by USDA/NIFA award 2015-67015-22930 and by Public Health Service Grants AI044170, AI096528, AI112445 and AI112949.

Figures

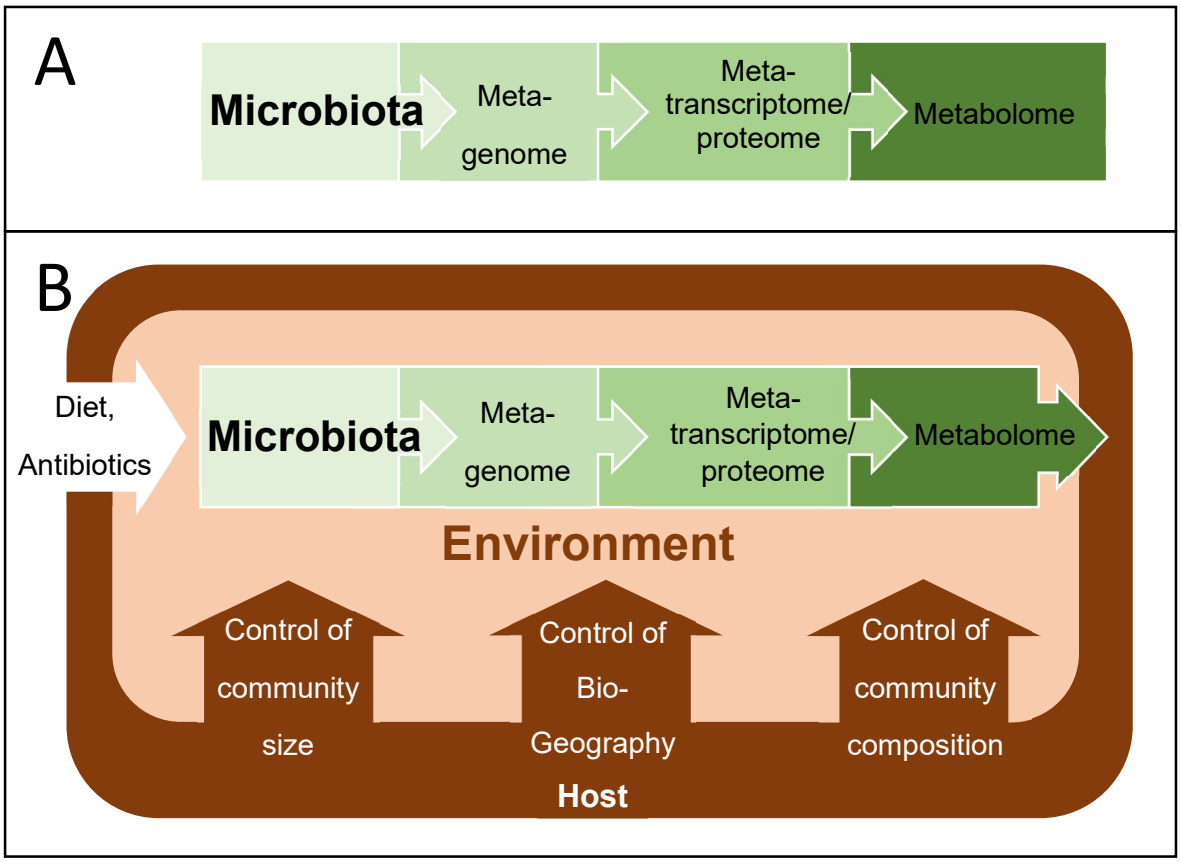

Figure 1.1

Figure 1.1: Graphical representation of different microbiome concepts. (**A**) The organ metaphor suggests that the microbiota contains our second genome, which spawned the concept that the microbiome is composed of the microbiota, its genes and gene products. (**B**) Ecological frameworks define the microbiome as the microbiota and its environment. The host shapes the microbial environment using habitat filters that control the size, biogeography and species composition of the microbiota. Thus, ecologically the microbiome can be viewed as a host-microbe chimera.

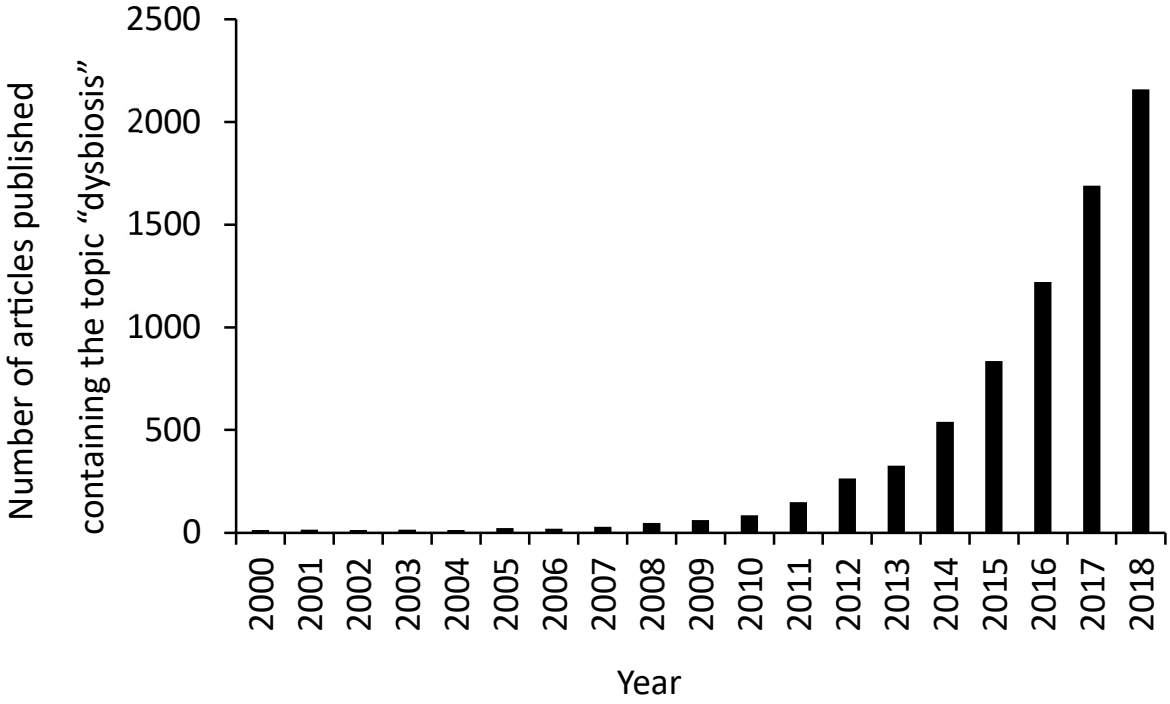
Number of articles published containing the topic "dysbiosis"
2500
2000
1500
1000
500
0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Year

Figure 1.2

Figure 1.2: **Increasing popularity of the term dysbiosis in the scientific literature.** The graph shows the annual number of articles deposited in the *Web of Science* database that contain the search topic ‘dysbiosis’.

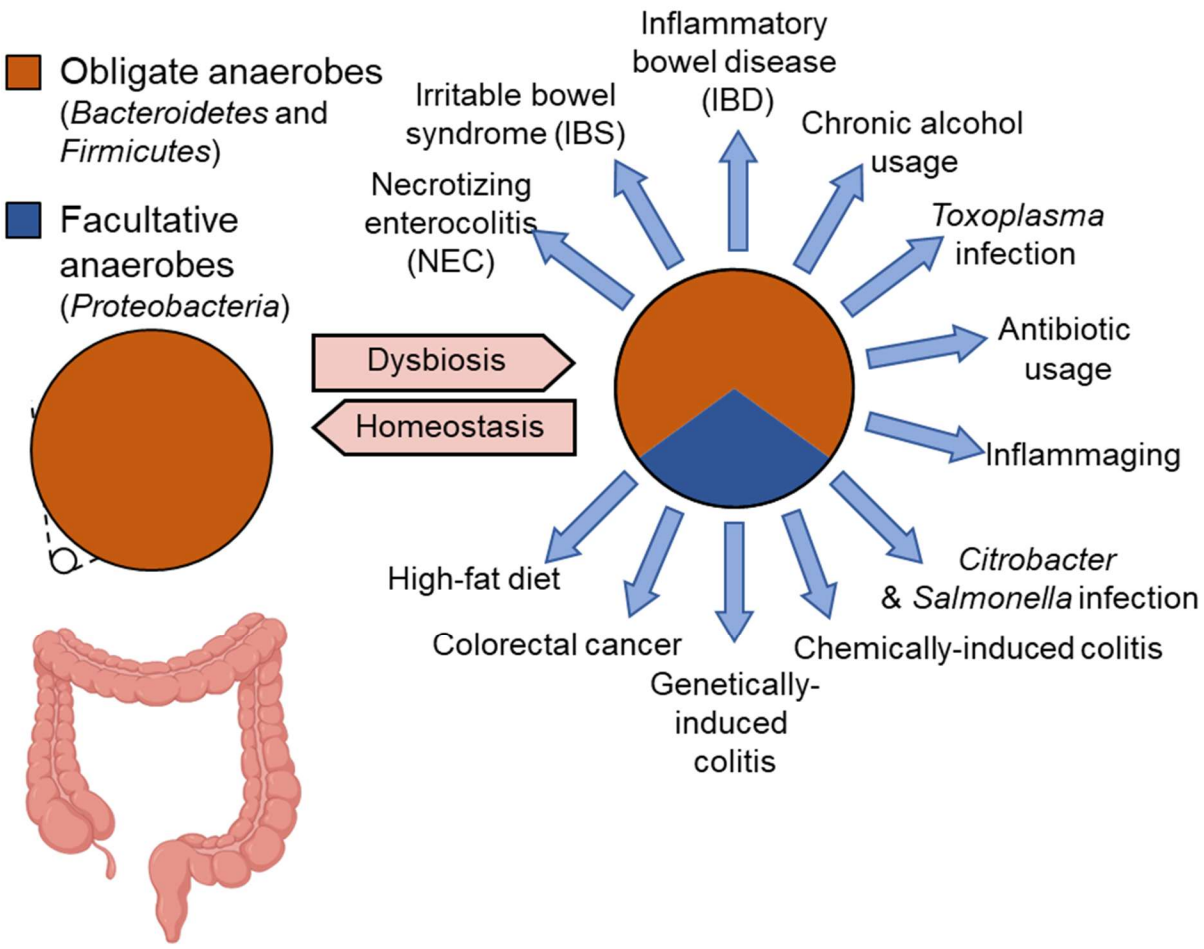
Obligate anaerobes
(*Bacteroidetes* and
Firmicutes)
Facultative
anaerobes
(*Proteobacteria*)
Dysbiosis
Homeostasis
Inflammatory
bowel disease
(IBD)
Irritable bowel
syndrome (IBS)
Chronic alcohol
usage
Necrotizing
enterocolitis
(NEC)
Toxoplasma
infection
Antibiotic
usage
Inflammaging
Citrobacter
& *Salmonella* infection
High-fat diet
Colorectal cancer
Genetically-
induced
colitis
Chemically-induced colitis

Figure 1.3

Figure 1.3: An expansion of *Proteobacteria* is a microbial signature of dysbiosis in the fecal microbiota. The fecal microbiota of healthy individuals is dominated by obligate anaerobic bacteria belonging to the phyla *Firmicutes* and *Bacteroidetes* (30). A dysbiotic expansion in the fecal microbiota of facultative anaerobic bacteria of the phylum *Proteobacteria* is observed in patients with necrotizing enterocolitis (50), irritable bowel syndrome (12, 33), inflammatory bowel disease (18), colorectal cancer (3) or in individuals consuming a high-fat diet (46), with chronic alcohol usage (15), or undergoing inflammaging (48). A dysbiotic expansion of *Proteobacteria* in the large intestine is also observed in mouse models of chemically-induced colitis (43), genetically-induced colitis (23), antibiotic treatment (6, 59) and infection with *Salmonella enterica* (68), *Citrobacter rodentium* (43) or *Toxoplasma gondii* (27).

References

1. Antunes LC, McDonald JA, Schroeter K, Carlucci C, Ferreira RB, Wang M, Yurist-Doutsch S, Hira G, Jacobson K, Davies J, Allen-Vercoe E, and Finlay BB. Antivirulence activity of the human gut metabolome. *MBio* 5: e01183-01114, 2014.

2. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, deRoos P, Liu H, Cross JR, Pfeffer K, Coffer PJ, and Rudensky AY. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* 504: 451-455, 2013.

3. Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM, Fan TJ, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor AA, and Jobin C. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338: 120-123, 2012.

4. Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, Cheng G, Yamasaki S, Saito T, Ohba Y, Taniguchi T, Takeda K, Hori S, Ivanov, II, Umesaki Y, Itoh K, and Honda K. Induction of colonic regulatory T cells by indigenous Clostridium species. *Science* 331: 337-341, 2011.

5. Bakken JS, Borody T, Brandt LJ, Brill JV, Demarco DC, Franzos MA, Kelly C, Khoruts A, Louie T, Martinelli LP, Moore TA, Russell G, Surawicz C, and Fecal Microbiota Transplantation W. Treating Clostridium difficile infection with fecal microbiota transplantation. *Clin Gastroenterol Hepatol* 9: 1044-1049, 2011.

6. Bohnhoff M, Drake BL, and Miller CP. Effect of streptomycin on susceptibility of intestinal tract to experimental Salmonella infection. *Proc Soc Exp Biol Med* 86: 132-137, 1954.

7. Byndloss MX, and Baumler AJ. The germ-organ theory of non-communicable diseases. *Nat Rev Microbiol* 16: 103-110, 2018.

8. Byndloss MX, Litvak Y, and Baumler AJ. Microbiota-nourishing Immunity and Its Relevance for Ulcerative Colitis. *Inflamm Bowel Dis* 25: 811-815, 2019.

9. Byndloss MX, Olsan EE, Rivera-Chávez F, Tiffany CR, Cevallos SA, Lokken KL, Torres TP, Byndloss AJ, Faber F, Gao Y, Litvak Y, Lopez CA, Xu G, Napoli E, Giulivi C, Tsolis RM, Revzin A, Lebrilla CB, and Bäumler AJ. Microbiota-activated PPAR-g signaling inhibits dysbiotic Enterobacteriaceae expansion. *Science* 357: 570-575, 2017.

10. Byndloss MX, Pernitzsch SR, and Baumler AJ. Healthy hosts rule within: ecological forces shaping the gut microbiota. *Mucosal Immunol* 11: 1299-1305, 2018.

11. Cani PD. Gut microbiota - at the intersection of everything? *Nat Rev Gastroenterol Hepatol* 14: 321-322, 2017.

12. Carroll IM, Ringel-Kulka T, Siddle JP, and Ringel Y. Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society* 24: 521-530, e248, 2012.

13. Chen C, Song X, Wei W, Zhong H, Dai J, Lan Z, Li F, Yu X, Feng Q, Wang Z, Xie H, Chen X, Zeng C, Wen B, Zeng L, Du H, Tang H, Xu C, Xia Y, Xia H, Yang H, Wang J, Wang J, Madsen L, Brix S, Kristiansen K, Xu X, Li J, Wu R, and Jia H. The microbiota

continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat Commun* 8: 875, 2017.

14. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud DJ, and Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* 54: 2325-2340, 2013.

15. Dubinkina VB, Tyakht AV, Odintsova VY, Yarygin KS, Kovarsky BA, Pavlenko AV, Ischenko DS, Popenko AS, Alexeev DG, Taraskina AY, Nasyrova RF, Krupitsky EM, Shalikiani NV, Bakulin IG, Shcherbakov PL, Skorodumova LO, Larin AK, Kostryukova ES, Abdulkhakov RA, Abdulkhakov SR, Malanin SY, Ismagilova RK, Grigoryeva TV, Ilina EN, and Govorun VM. Links of gut microbiota composition with alcohol dependence syndrome and alcoholic liver disease. *Microbiome* 5: 141, 2017.

16. El Kaoutari A, Armougom F, Gordon JI, Raoult D, and Henrissat B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* 11: 497-504, 2013.

17. Foster KR, Schluter J, Coyte KZ, and Rakoff-Nahoum S. The evolution of the host microbiome as an ecosystem on a leash. *Nature* 548: 43-51, 2017.

18. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, and Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences of the United States of America* 104: 13780-13785, 2007.

19. Freter R, Brickner H, Fekete J, Vickerman MM, and Carey KE. Survival and implantation of Escherichia coli in the intestinal tract. *Infect Immun* 39: 686-703, 1983.

20. Fukami T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst* 46: 1-23, 2015.

21. Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, Nakanishi Y, Uetake C, Kato K, Kato T, Takahashi M, Fukuda NN, Murakami S, Miyauchi E, Hino S, Atarashi K, Onawa S, Fujimura Y, Lockett T, Clarke JM, Topping DL, Tomita M, Hori S, Ohara O, Morita T, Koseki H, Kikuchi J, Honda K, Hase K, and Ohno H. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504: 446-450, 2013.

22. Garcia C, Tebbji F, Daigneault M, Liu NN, Kohler JR, Allen-Vercoe E, and Sellam A. The Human Gut Microbial Metabolome Modulates Fungal Growth via the TOR Signaling Pathway. *mSphere* 2: 2017.

23. Garrett WS, Gallini CA, Yatsunenko T, Michaud M, DuBois A, Delaney ML, Punit S, Karlsson M, Bry L, Glickman JN, Gordon JI, Onderdonk AB, and Glimcher LH. Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell host & microbe* 8: 292-300, 2010.

24. Gilbert JA, and Lynch SV. Community ecology as a framework for human microbiome research. *Nat Med* 25: 884-889, 2019.

25. Gillis CC, Hughes ER, Spiga L, Winter MG, Zhu W, Furtado de Carvalho T, Chanin RB, Behrendt CL, Hooper LV, Santos RL, and Winter SE. Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. *Cell Host Microbe* 23: 54-64 e56, 2018.

26. Govan JR, and Deretic V. Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. *Microbiol Rev* 60: 539-574, 1996.

27. Haag LM, Fischer A, Otto B, Plickert R, Kuhl AA, Gobel UB, Bereswill S, and Heimesaat MM. Intestinal microbiota shifts towards elevated commensal Escherichia coli loads abrogate colonization resistance against Campylobacter jejuni in mice. *PLoS One* 7: e35988, 2012.

28. Hooper LV, and Gordon JI. Commensal host-bacterial relationships in the gut. *Science* 292: 1115-1118, 2001.

29. Hughes ER, Winter MG, Duerkop BA, Spiga L, Furtado de Carvalho T, Zhu W, Gillis CC, Buttner L, Smoot MP, Behrendt CL, Cherry S, Santos RL, Hooper LV, and Winter SE. Microbial Respiration and Formate Oxidation as Metabolic Signatures of Inflammation-Associated Dysbiosis. *Cell Host Microbe* 21: 208-219, 2017.

30. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207-214, 2012.

31. Integrative HMPRNC. The Integrative Human Microbiome Project. *Nature* 569: 641-648, 2019.

32. Johansson ME, Phillipson M, Petersson J, Velcich A, Holm L, and Hansson GC. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A* 105: 15064-15069, 2008.

33. Krogius-Kurikka L, Lyra A, Malinen E, Aarnikunnas J, Tuimala J, Paulin L, Makivuokko H, Kajander K, and Palva A. Microbial community analysis reveals high

level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea-predominant irritable bowel syndrome sufferers. *BMC gastroenterology* 9: 95, 2009.

34. Kuijper EJ, Coignard B, Tull P, difficile ESGfC, States EUM, European Centre for Disease P, and Control. Emergence of Clostridium difficile-associated disease in North America and Europe. *Clin Microbiol Infect* 12 Suppl 6: 2-18, 2006.

35. Levy M, Kolodziejczyk AA, Thaiss CA, and Elinav E. Dysbiosis and the immune system. *Nat Rev Immunol* 17: 219-232, 2017.

36. Libertucci J, and Young VB. The role of the microbiota in infectious diseases. *Nat Microbiol* 4: 35-45, 2019.

37. Litvak Y, and Baumler AJ. The founder hypothesis: A basis for microbiota resistance, diversity in taxa carriage, and colonization resistance against pathogens. *PLoS Pathog* 15: e1007563, 2019.

38. Litvak Y, Byndloss MX, and Baumler AJ. Colonocyte metabolism shapes the gut microbiota. *Science* 362: 2018.

39. Litvak Y, Byndloss MX, Tsolis RM, and Baumler AJ. Dysbiotic Proteobacteria expansion: a microbial signature of epithelial dysfunction. *Curr Opin Microbiol* 39: 1-6, 2017.

40. Litvak Y, Mon KKZ, Nguyen H, Chanthavixay G, Liou M, Velazquez EM, Kutter L, Alcantara MA, Byndloss MX, Tiffany CR, Walker GT, Faber F, Zhu Y, Bronner DN, Byndloss AJ, Tsolis RM, Zhou H, and Baumler AJ. Commensal Enterobacteriaceae Protect against Salmonella Colonization through Oxygen Competition. *Cell Host Microbe* 25: 128-139 e125, 2019.

41. Llorens-Rico V, and Raes J. Tracking humans and microbes. *Nature* 569: 632-633, 2019.

42. Lopez CA, Miller BM, Rivera-Chávez F, Velazquez EM, Byndloss MX, Chávez-Arroyo A, Lokken KL, Tsolis RM, Winter SE, and Bäumler AJ. Virulence factors enhance Citrobacter rodentium expansion through aerobic respiration. *Science* 353: 1249-1253, 2016.

43. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, and Finlay BB. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe* 2: 204, 2007.

44. Marchesi JR. Human distal gut microbiome. *Environ Microbiol* 13: 3088-3102, 2011.

45. Marchesi JR, Adams DH, Fava F, Hermes GD, Hirschfield GM, Hold G, Quraishi MN, Kinross J, Smidt H, Tuohy KM, Thomas LV, Zoetendal EG, and Hart A. The gut microbiota and host health: a new clinical frontier. *Gut* 65: 330-339, 2016.

46. Martinez-Medina M, Denizot J, Dreux N, Robin F, Billard E, Bonnet R, Darfeuille-Michaud A, and Barnich N. Western diet induces dysbiosis with increased E coli in CEABAC10 mice, alters host barrier function favouring AIEC colonisation. *Gut* 63: 116-124, 2014.

47. Miller ET, Svanback R, and Bohannan BJM. Microbiomes as Metacommunities: Understanding Host-Associated Microbes through Metacommunity Ecology. *Trends Ecol Evol* 33: 926-935, 2018.

48. Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, Cresci A, Silvi S, Orpianesi C, Verdenelli MC, Clavel T, Koebnick C, Zunft HJ, Dore J, and Blaut M. Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Appl Environ Microbiol* 72: 1027-1033, 2006.

49. Neville BA, Forster SC, and Lawley TD. Commensal Koch's postulates: establishing causation in human microbiota research. *Curr Opin Microbiol* 42: 47-52, 2018.

50. Normann E, Fahlen A, Engstrand L, and Lilja HE. Intestinal microbial profiles in extremely preterm infants with and without necrotizing enterocolitis. *Acta paediatrica* 102: 129-136, 2013.

51. O'Hara AM, and Shanahan F. The gut flora as a forgotten organ. *EMBO Rep* 7: 688-693, 2006.

52. Olesen SW, and Alm EJ. Dysbiosis is not an answer. *Nat Microbiol* 1: 16228, 2016.

53. Petersen C, and Round JL. Defining dysbiosis and its influence on host immunity and disease. *Cell Microbiol* 16: 1024-1033, 2014.

54. Proctor L. Priorities for the next 10 years of human microbiome research. *Nature* 569: 623-625, 2019.

55. Quie PG. Lung defense against infection. *J Pediatr* 108: 813-816, 1986.

56. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, and Forney

LJ. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108 Suppl 1: 4680-4687, 2011.

57. Rigottier-Gois L. Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis. *ISME J* 7: 1256-1261, 2013.

58. Rivera-Chavez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, Xu G, Velazquez EM, Lebrilla CB, Winter SE, and Baumler AJ. Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella. *Cell Host Microbe* 19: 443-454, 2016.

59. Saito K. [Studies on the habitation of pathogenic Escherichia coli in the intestinal tract of mice. I. Comparative experiments on the habitation of each type of resistant pathogenic Escherichia coli under an administration of streptomycin]. *Paediatria Japonica* 65: 385-393, 1961.

60. Salzman NH, Hung K, Haribhai D, Chu H, Karlsson-Sjoberg J, Amir E, Teggatz P, Barman M, Hayward M, Eastwood D, Stoel M, Zhou Y, Sodergren E, Weinstock GM, Bevins CL, Williams CB, and Bos NA. Enteric defensins are essential regulators of intestinal microbial ecology. *Nat Immunol* 11: 76-83, 2010.

61. Sarker SA, Ahmed T, and Brussow H. Hunger and microbiology: is a low gastric acid-induced bacterial overgrowth in the small intestine a contributor to malnutrition in developing countries? *Microb Biotechnol* 10: 1025-1030, 2017.

62. Shin NR, Whon TW, and Bae JW. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol* 33: 496-503, 2015.

63. Singh VP, Proctor SD, and Willing BP. Koch's postulates, microbial dysbiosis and inflammatory bowel disease. *Clin Microbiol Infect* 22: 594-599, 2016.

64. Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly YM, Glickman JN, and Garrett WS. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* 341: 569-573, 2013.

65. Sommer F, Anderson JM, Bharti R, Raes J, and Rosenstiel P. The resilience of the intestinal microbiota influences health and disease. *Nat Rev Microbiol* 15: 630-638, 2017.

66. Spear GT, French AL, Gilbert D, Zariffard MR, Mirmonsef P, Sullivan TH, Spear WW, Landay A, Micci S, Lee BH, and Hamaker BR. Human alpha-amylase present in lower-genital-tract mucosal fluid processes glycogen to support vaginal colonization by Lactobacillus. *J Infect Dis* 210: 1019-1028, 2014.

67. Sprockett D, Fukami T, and Relman DA. Role of priority effects in the early-life assembly of the gut microbiota. *Nat Rev Gastroenterol Hepatol* 15: 197-205, 2018.

68. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, Chaffron S, Macpherson AJ, Buer J, Parkhill J, Dougan G, von Mering C, and Hardt WD. Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol* 5: 2177-2189, 2007.

69. Tipton L, Darcy JL, and Hynson NA. A Developing Symbiosis: Enabling Cross-Talk Between Ecologists and Microbiome Scientists. *Front Microbiol* 10: 292, 2019.

70. Vannette RL, and Fukami T. Historical contingency in species interactions: towards niche-based predictions. *Ecol Lett* 17: 115-124, 2014.

71. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston MR, Lopez CA, Litvak Y, Liou MJ, Faber F, Bronner DN, Tiffany CR, Byndloss MX, Byndloss AJ, and Baumler AJ. Endogenous Enterobacteriaceae underlie variation in susceptibility to Salmonella infection. *Nat Microbiol* 4: 1057-1064, 2019.

72. Velazquez OC, Lederer HM, and Rombeau JL. Butyrate and the colonocyte. Production, absorption, metabolism, and therapeutic implications. *Adv Exp Med Biol* 427: 123-134, 1997.

73. Vonaesch P, Anderson M, and Sansonetti PJ. Pathogens, microbiome and the host: emergence of the ecological Koch's postulates. *FEMS Microbiol Rev* 42: 273-292, 2018.

74. Yang X, Xie L, Li Y, and Wei C. More than 9,000,000 unique genes in human gut bacterial community: estimating gene numbers inside a human body. *PLoS One* 4: e6074, 2009.

75. Zheng L, Kelly CJ, and Colgan SP. Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A Review in the Theme: Cellular Responses to Hypoxia. *Am J Physiol Cell Physiol* 309: C350-360, 2015.

Chapter 2

omu, a Metabolomics Count Data Analysis Tool for Intuitive Figures and Convenient Metadata Collection

Connor R Tiffany, Andreas J. Bäumler*

Department of Medical Microbiology and Immunology, School of Medicine, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA.

Published in

Tiffany CR, Bäumler AJ. omu, a Metabolomics Count Data Analysis Tool for Intuitive Figures and Convenient Metadata Collection. Microbiol Resour Announc. 2019 Apr 11;8(15):e00129-19. doi: 10.1128/MRA.00129-19. PMID: 30975806; PMCID: PMC6460029.

Abstract

Metabolomics is a powerful tool for measuring the functional output of the microbiota. Currently, there are few established workflows for analysis downstream of metabolite identification. Here, we introduce omu, an R package designed for assigning compound hierarchies, and linking compounds to corresponding enzyme and gene annotations for organisms of interest.

Text

The omu R package is designed to analyze processed metabolomics count data. The central idea behind omu is assigning hierarchical metadata from the Kyoto Encyclopedia of Genes and Genomes (KEGG)(1) to each metabolite in order to help the user create intuitive figures to visualize their data. To do this, omu provides a suite of graphing and statistical functions, centered around the use of assign_hierarchy, which provides the metadata to each metabolite based on a KEGG identifier in the users data. omu comes with an example dataset of the fecal metabolome of NOS2-defienct C57BL/6J mice three days after mock-treatment (gavage with sterile water) or oral gavage with a single dose of streptomycin (20mg/animal). This dataset was used to generate figures in this paper.

Initially, users can use assign_hierarchy to provide metadata for each metabolite that has a KEGG compound number provided (table1). KEGG data are annotated by compound Class, Subclass_1, Subclass_2, Subclass_3, and Subclass_4. This annotation provides several options for the user to analyze and visualize their data. For example, users can subset to a compound class they are particularly interested in, such as carbohydrates. The user can use the omu_summary function to perform a Student's *t* test on each metabolite between two experimental groups, provide a fold-change value for each compound, and use this data in conjunction with the hierarchical compound annotation to create intuitive figures. The count_fold_changes function can be used to provide a count table of every metabolite within a KEGG compound category that

significantly increased or decreased between groups to make bar plots (fig2.1a) that show how many compounds increased or decreased between experiment groups by a metabolite class or subclass. Alternatively, since omu_summary also provides fold-change data, the user can incorporate effect size into a figure by creating a volcano plot using the plot_volcano function, which allows the user to highlight points in the plot by their metadata (fig2.1b).

The omu R package can also help the user generate hypotheses from their metabolomics data. The KEGG_gather function can get all known enzyme orthology data and gene data associated with each metabolite from the KEGG database as long as the computer is connected to the internet. This produces high dimensional data, but assign_hierarchy also provides metadata for enzymes, and for genes in the form of organism data, allowing the user to reduce the data to items of interest. For example, if a user is interested in genes *Pseudomonas spp.* has to metabolize organic acids, KEGG_gather function in conjunction with the assign_hierarchy function can be used to generate a table containing this information (table2.1).

In summary, omu is a novel metabolomics analysis tool that helps users describe their data by incorporating metabolite metadata into intuitive figures and creating tables with genes and enzymes associated with metabolites of interest. omu is available for download on the CRAN repository https://cran.r-project.org/web/packages/omu/index.html , and was built using R(2), devtools(3), dplyr(4), ggfortify(5), ggplot2(6), KEGGREST(7), knitr(8), magrittr(9), plyr(10), reshape2(11), rmarkdown(12), roxygen2(13), stringr(14), and tidyr(15). Detailed

instructions for using omu can be found at https://cran.r-project.org/web/packages/omu/vignettes/Omu_vignette.html.

Acknowledgements

Work in AJB's lab is supported by USDA/NIFA award 2015-67015-22930 and by Public Health Service Grants AI044170, AI096528, AI112445 and AI112949.

Figures

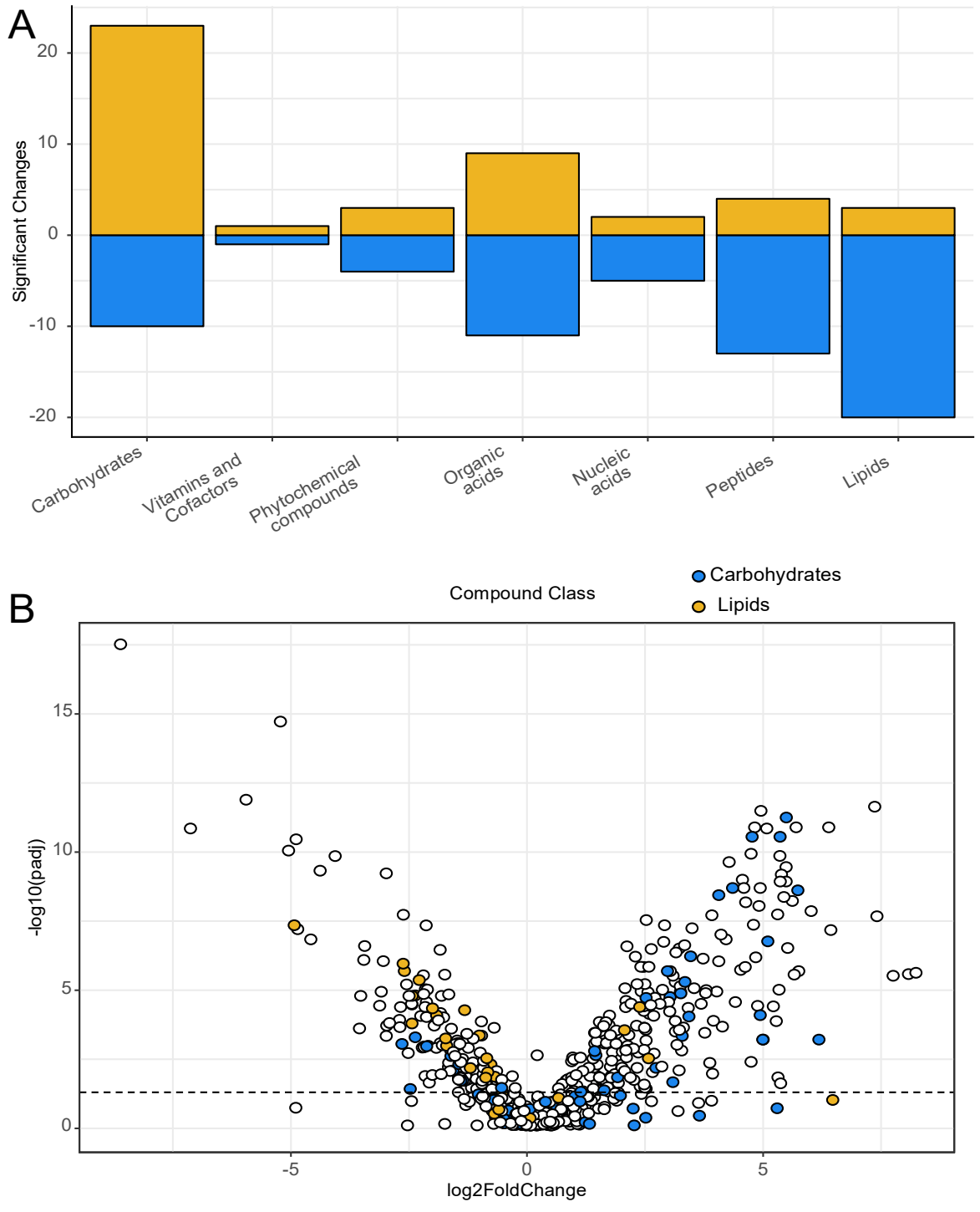

Figure 2.1

Figure 2.1: Data visualization in omu. (a) Bar plot showing the number of significantly different metabolites between treatment groups by class from the example dataset that comes with the R package. (b) Volcano plot comparing two treatment groups from the example dataset. Highlighted points in the plot correspond to metabolite classes.

Genes	Metabolite	KEGG	Sample	Class	Subclass_1	Subclass_2	Species.Strain.Serotype
PA4091(hpaA)	4_hydroxyphenylacetic_acid	C00642	32307	Organic acids	none	none	Pseudomonas aeruginosa PAO1
N297_4221(hpaB)	4_hydroxyphenylacetic_acid	C00642	32307	Organic acids	none	none	Pseudomonas aeruginosa PAO1-VE13
N296_4221(hpaB)	4_hydroxyphenylacetic_acid	C00642	32307	Organic acids	none	none	Pseudomonas aeruginosa PAO1-VE2
PA14_11000(hpaA)	4_hydroxyphenylacetic_acid	C00642	32307	Organic acids	none	none	Pseudomonas aeruginosa UCBPP-PA14
PSPA7_1007(hpaB)	4_hydroxyphenylacetic_acid	C00642	32307	Organic acids	none	none	Pseudomonas aeruginosa PA7
PP4_31900(gbd)	4_hydroxybutyric_acid	C00989	315	Organic acids	Carboxylic acids	Hydroxycarboxylic acids	Pseudomonas putida NBRC 14164
APT59_05510	4_hydroxybutyric_acid	C00989	315	Organic acids	Carboxylic acids	Hydroxycarboxylic acids	Pseudomonas oryzihabitans
PverR02_11545	4_hydroxybutyric_acid	C00989	315	Organic acids	Carboxylic acids	Hydroxycarboxylic acids	Pseudomonas veronii
BJP27_20245	4_hydroxybutyric_acid	C00989	315	Organic acids	Carboxylic acids	Hydroxycarboxylic acids	Pseudomonas psychrotoleranns

Table 2.1

Table 2.1: Data table created from the example dataset, showcasing the metadata collection that omu can perform.

References

1. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. 1999. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res.
2. Team RDC, R Development Core Team R. 2016. R: A Language and Environment for Statistical Computing. R Found Stat Comput.
3. Wickham H, Hester J, Chang W. 2018. devtools: Tools to Make Developing R Packages Easier.
4. Wickham H, Francois R, Henry L, Müller K. 2018. dplyr: A Grammar of Data Manipulation. R Packag version 075.
5. Tang Y, Horikoshi M, Li W. 2016. ggfortify: Unified Interface to Visualize Statistical Results of Popular R Packages. R J 82.
6. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
7. Tenenbaum D. 2018. KEGGREST: Client-side REST access to KEGG.
8. Xie Y. 2018. knitr: A General-Purpose Package for Dynamic Report Generation in R.
9. Bache SM, Wickham H. 2014. magrittr: A Forward-Pipe Operator for R.
10. Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. J Stat Softw 40:1–29.

11. Wickham H. 2012. reshape2: Flexibly reshape data: a reboot of the reshape package. R Packag version.

12. Allaire JJ, Xie Y, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W. 2018. rmarkdown: Dynamic Documents for R.

13. Wickham H. 2018. stringr: Simple, Consistent Wrappers for Common String Operations.

14. Wickham H, Danenberg P, Eugster M. 2014. R: Roxygen2. CRAN.

15. Wickham H, Henry L. 2018. tidyr: Easily Tidy Data with “spread()” and “gather()” Functions.

Chapter 3

The metabolic footprint of *Clostridia* and *Erysipelotrichia* reveals their role in depleting sugar alcohols in the cecum

Connor R. Tiffany[1], Jee-Yon Lee[1], Andrew W.L. Rogers[1], Erin E. Olsan[1,2], Pavel Morales[3], Franziska Faber[1,4,*] and Andreas J. Bäumler[1,*]

[1] Department of Medical Microbiology and Immunology, School of Medicine, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA.

[2] Current address: Department of Biological Sciences, California State University Sacramento, 6000 J Street, Sacramento, CA 95819, USA

[3] Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA.

[4] Current address: Institute for Molecular Infection Biology (IMIB), Faculty of Medicine, University of Würzburg, Josef-Schneider-Street 2/D15, 97080 Würzburg, Germany

Published in

Tiffany, C.R., Lee, JY., Rogers, A.W.L. *et al.* The metabolic footprint of *Clostridia* and *Erysipelotrichia* reveals their role in depleting sugar alcohols in the cecum. *Microbiome* **9**, 174 (2021). https://doi.org/10.1186/s40168-021-01123-9

ABSTRACT

Background: The catabolic activity of the microbiota contributes to health by aiding in nutrition, immune education and niche protection against pathogens. However, the nutrients consumed by common taxa within the gut microbiota remain incompletely understood.

Methods: Here we combined microbiota profiling with an un-targeted metabolomics approach to determine whether depletion of small metabolites in the cecum of mice correlated with the presence of specific bacterial taxa. Causality was investigated by engrafting germ-free or antibiotic-treated mice with complex or defined microbial communities.

Results: We noted that a depletion of *Clostridia* and *Erysipelotrichia* from the gut microbiota triggered by antibiotic treatment was associated with an increase in the cecal concentration of sugar acids and sugar alcohols (polyols). Notably, when we inoculated germ-free mice with a defined microbial community of 14 *Clostridia* and 3 *Erysipelotrichia* isolates, we observed the inverse, with a marked decrease in the concentrations of sugar acids and polyols in cecal contents. The carbohydrate footprint produced by the defined microbial community was similar to that observed in gnotobiotic mice receiving a cecal microbiota transplant from conventional mice. Supplementation with sorbitol, a polyol used as artificial sweetener, increased cecal sorbitol concentrations in antibiotic-treated mice, which was abrogated after inoculation with a *Clostridia* isolate able to grow on sorbitol *in vitro.*

Conclusions: We conclude that consumption of sugar alcohols by *Clostridia* and *Erysipelotrichia* species depletes these metabolites from the intestinal lumen during homeostasis.

Keywords: Microbiota; alcoholic sugars; polyols; FODMAPs; *Clostridia*; *Erysipelotrichia.*

INTRODUCTION

The gut microbiota plays an important role in the homeostasis and health of its host organism. A large amount of research has focused on how specific constituents of the microbiota interact with the host to influence immune regulation and metabolism [1], but the metabolic capabilities of the gut microbiota remains incompletely understood. Previous research suggests that while phylogeny can be a good predictor of complex metabolic traits requiring large amounts of genes, such as photosynthesis, it is not necessarily a good predictor of simple metabolic traits like carbohydrate utilization [2]. The relationship between phylogeny and metabolic traits in bacteria is complicated further by the widespread occurrence of lateral gene transfer [3], gene loss, and convergent evolution [4]. Nonetheless, some metabolic pathways are conserved among phylogenetic groupings within the gut microbiome [5].

For example, members of the *Enterobacteriaceae*, a family within the phylum *Proteobacteria*, lack the capabilities to break down complex carbohydrates and instead rely on respiratory catabolism of fermentation products using terminal electron acceptors such as nitrate and oxygen [6, 7]. Many *Bacteroides spp*., obligate anaerobes within the phylum *Bacteroidetes*, are primary fermenters of fiber and mucin-derived polysaccharides [8], encoding a large repertoire of genes associated with glycolytic gene clusters [9], while producing much of the short-chain fatty acid propionate found in the murine gut [10]. *Prevotella spp*. within the phylum *Bacteroidetes* can also ferment many plant polysaccharides [11] and the abundance of this genus positively correlates with increased

fiber intake in humans [12], but unlike *Bacteroides, Prevotella spp*. do not produce propionate [10]. Bacteria within *Clostridia*, a class of spore-forming, obligate anaerobes in the phylum *Firmicutes*, have been linked to the production of the short-chain fatty acid butyrate [13, 14]. This metabolite fulfills an important role in the gut, acting as a signaling molecule and carbon source for host colonic epithelial cells (colonocytes) [15, 16]. By driving colonocyte metabolism towards high oxygen consumption through mitochondrial beta oxidation, butyrate helps maintain the epithelium in a state of physiologic hypoxia [17], thereby preserving anaerobiosis in the lumen of the large bowel [18]. *Clostridia* have also been shown to comprise the bulk of diversity within the human gut microbiome [19], further highlighting the importance of this taxon within the ecosystem of the large intestine. While there is increasing information on the metabolic outputs of *Clostridia* and their function in mediating gut homeostasis, less is known about what nutrients they consume *in vivo*.

Clostridia were recently divided into two classes, *Clostridia* and *Erysipelotrichia*, which form two sister groups within the phylum *Firmicutes* [20]. In light of this change in nomenclature, previous work on butyrate production by *Clostridia* also included *Erysipelotrichia* isolates [21]. Here we combined germ-free mice engrafted with a defined consortium of *Clostridia* and *Erysipelotrichia* isolates with an un-targeted metabolomics approach to determine how these taxa alter the metabolite landscape in the cecum.

RESULTS

Streptomycin treatment shifts the cecal metabolome by increasing sugar alcohols and sugar acids

To investigate how antibiotic-treatment changes the murine cecal metabolome, mice (C57BL/6J) were mock-treated or received a single dose of streptomycin intragastrically and cecal contents were collected three days later. The soluble fraction of the cecal contents was then analyzed by un-targeted gas chromatography time-of-flight mass spectrometry (metabolite profiling). Principle component analysis of the samples showed a distinct clustering of mice by treatment group, indicating that streptomycin treatment changes the composition of the murine cecal metabolome (**Fig. 3.1a**). Consistent with a previous report [22], differential abundance of metabolites between streptomycin-treated mice and mock-treated mice revealed that the concentration of many compounds annotated as carbohydrates significantly increased with streptomycin treatment, while the concentration of many compounds annotated as lipids, peptides, or organic acids decreased (**Fig. 3.1b and 3.1c**). Interestingly, most of the carbohydrates that increased with streptomycin treatment were identified as sugar acids (galactonic acid, glyceric acid, lactobionic acid, ribonic acid, saccharic acid, threonic acid, and xylonic acid) or sugar alcohols (1,5-anhydroglucitol, erythritol, galacitonol, lyxitol, myo-inositol, pentitol, sorbitol, threitol, and xylitol) (**Fig. 3.1d and Supplementary Table 3.1**). This increase in

the amount of sugar alcohols and sugar acids raised the possibility that streptomycin treatment might deplete microbes that normally consume these nutrients in the cecum.

Streptomycin treatment depletes *Clostridia* and *Erysipelotrichia* from the gut microbiota

To investigate whether changes in the cecal metabolite profile were linked to specific changes in the community composition, we extracted total genomic DNA from contents of the proximal colon, assembled libraries, and then performed 16S rRNA gene amplicon sequencing (microbiota profiling). Weighted unifrac analysis of the samples showed distinct clustering by treatment group (**Fig. 3.2a**), and a significant difference in weighted unifrac distance was observed between streptomycin-treated and mock-treated mice (**Fig. 3.2b**), indicating that the community structure was different between groups. *Clostridia* and *Erysipelotrichia* were the only taxonomic groups to both contain significantly decreased amplicon sequence variants (ASVs) in streptomycin-treated mice relative to mock-treated mice according to negative binomial regression analysis, explaining the difference between the gut microbiota of streptomycin-treated mice and mock-treated mice according to linear discriminant analysis (**Fig. 3.2c**). Conversely, *Gammaproteobacteria* and *Bacilli* were significantly enriched in streptomycin-treated mice. Consistent with previous reports [18, 23, 24], we observed a decrease in the relative abundance of *Clostridia* in streptomycin-treated mice (**Fig. 3.2d**). Within the

Erysipelotrichia, three taxa dropped below the limit of detection in the streptomycin treated group, whereas the most abundant taxon significantly increased (**Fig. 3.2e and Fig. 3.S.1**), resulting in an elevated overall relative abundance of *Erysipelotrichia* in streptomycin-treated mice (**Fig. 3.2d**). Differential abundance analysis revealed that the class *Clostridia* had the most significantly different amplicon sequence variants (ASVs), with most of these decreasing after streptomycin treatment (**Fig. 3.2e**). Whereas *Clostridia* were not the most abundant taxon, this class comprised approximately two thirds of the ASV diversity within the microbiota of mock-treated mice (**Fig. 3.2f**). Thus, streptomycin treatment altered the community structure by depleting *Clostridia* and *Erysipelotrichia* (**Fig. 3.2c-3.2e**), with the former being the taxonomically most diverse class within the gut microbiota (**Fig. 3.2f**).

Clostridia and *Erysipelotrichia* deplete sugar alcohols and sugar acids

Having established a correlation between *Clostridia* and *Erysipelotrichia* depletion and an increase in sugar alcohols and sugar acids, we next wanted to determine whether *Clostridia* and *Erysipelotrichia* caused this shift in the cecal metabolome in the absence of potential confounding factors, such antibiotic treatment. To that end, we inoculated germ-free Swiss Webster mice with a consortium originally described as 17 commensal *Clostridia* isolated from a healthy volunteer [21]. According to current nomenclature, this consortium consists of fourteen *Clostridia* isolates and three *Erysipelotrichia* isolates [25]

(**Figure 3.3a and Supplementary Table 3.2**). Five days after colonization, samples were collected and analyzed using microbiota and metabolite profiling. Microbiota profiling revealed that ten *Clostridia* isolates and three *Erysipelotrichia* isolates engrafted at levels detectable by 16S ribosomal RNA gene amplicon sequencing (**Figure 3.3b**). Principle component analysis showed a distinct clustering pattern between mock-treated germ-free mice and those given the consortium of commensal *Clostridia* and *Erysipelotrichia* isolates (**Fig. 3.4a**). Remarkably, association of germ-free mice with the commensal consortium (**Fig. 3.4b, 3.4c**) produced an equivalent but inverse phenotype to that observed after depleting *Clostridia* and *Erysipelotrichia* by treatment with streptomycin (**Figure 3.1b, 3.1c**). All significantly different metabolites identified as carbohydrates decreased when germ-free mice were associated with a consortium of commensal *Clostridia* and *Erysipelotrichia* isolates (**Fig. 3.4b, 3.4c)**, which was accompanied by a concomitant increase in a plurality of significantly different lipids, peptides and organic acids. Within carbohydrates, sugar alcohols (erythritol, galactinol, lyxitol, myo-inositol, pentitol, sorbitol, ribitol, and xylitol) and sugar acids (digalacturonic acid, galactonic acid, galacturonic acid, glyceric acid, threonic acid, and xylonic acid) contained the largest number of significantly decreased compounds (**Fig. 4d and Supplementary Table 3.3**).

Comparing metabolite profiling data from streptomycin-treated mice with those from gnotobiotic mice engrafted with commensal *Clostridia* and *Erysipelotrichia* isolates is complicated by the fact that conventional and gnotobiotic mice differed in their genotype (C57BL/6J versus Swiss Webster) and the composition of their chow. To compare metabolite profiles in the absence of these confounding variables, germ-free Swiss

Webster mice received a cecal microbiota transplant from conventionally housed C57BL/6J mice. Five days after engraftment, cecal samples were collected and analyzed using metabolite profiling. Principle component analysis revealed clustering of samples from mice receiving a cecal microbiota transplant from conventional (C57BL/6) mice with those from mice engrafted with the consortium of commensal *Clostridia* and *Erysipelotrichia* isolates (**Fig. 4a**). Notably, only four carbohydrates (including the alcoholic sugars hexitol, pentitol, and threitol) were depleted in germ-free mice receiving a cecal microbiota transplant from conventional (C57BL/6) mice compared to germ-free mice engrafted consortium of commensal *Clostridia* and *Erysipelotrichia* isolates (**Fig. 4e, 4f and Supplementary Table 3.3**). These data suggested that the carbohydrate footprint of a defined microbial community of human *Clostridia* and *Erysipelotrichia* isolates was similar to that produced by a complex mouse microbiota.

Polyol utilization by *Clostridia* and *Erysipelotrichia* species *in vitro* and *in vivo*

This result demonstrated that a consortium of commensal *Clostridia* and *Erysipelotrichia* isolates depletes sugar alcohols and sugar acids in the cecum, but did not provide direct evidence that these bacteria could utilize these carbon sources for growth. Polyols, such as sorbitol, mannitol or xylitol, are of particular interest, because they are used as artificial sweeteners in a broad range of processed food products. Thus, we performed an *in vitro* growth assay using no carbon defined media (NCDM)

supplemented with D-gluconic, glucose or one of the following sugar alcohols: xylitol, D-sorbitol, D-mannitol, myo-inositol, and meso-erythritol. Consumption of various combinations of sugar alcohols was observed for six *Clostridia* and one *Erysipelotrichia* isolate, with all seven isolates being able to grow on sorbitol as a carbon source (**Fig. 3.5a**).

Sorbitol is commonly used as a supplement to sweeten food. We thus wanted to determine whether consumption of polyols by *Clostridia* could be observed in streptomycin treated mice receiving dietary sorbitol supplementation. To this end, mice were maintained on a defined diet containing low concentrations of FODMAPs (Teklad Custom Diet TD.110675, Envigo). Mice were then mock-treated or treated with streptomycin and received 5% sorbitol in their drinking water. Absolute sorbitol concentrations determined by a colorimetric assay were close to the limit of detection in cecal contents of mock-treated mice. Sorbitol supplementation did not increase the concentration of sorbitol in cecal contents of mock-treated mice, suggesting that an intact microbiota can deplete this dietary supplement (**Fig. 3.5b**). In contrast, disruption of the microbiota by streptomycin treatment resulted in a marked increase in cecal sorbitol levels in mice receiving sorbitol supplementation. To investigate whether consumption of sorbitol by microbes was responsible for depleting this dietary supplement, the experiment was repeated with streptomycin-treated mice that were mock-inoculated or engrafted with *Anaerostipes caccae* (strain 9 from the defined microbial community, originally classified as *Clostridium indolis*) (**Figure 3.3a and Supplementary Table 3.2**), which was able to grow *in vitro* using sorbitol as a carbon source (**Fig. 3.5a**). Notably,

engraftment of streptomycin-treated mice with *A. caccae* resulted in a marked depletion of sorbitol (**Fig. 3.5c**), thus further supporting the idea that consumption of polyols by *Clostridia* results in a depletion of these metabolites from the metabolome in the murine cecum.

DISCUSSION

Members of the class *Clostridia* perform critical functions within the mammalian gut, including maintenance of hypoxia within the colon [18], providing colonization resistance to enteric pathogens [23, 26], and contributing to the immune education of the mammalian host organism [27]. It is known that *Clostridia* and their sister group, the *Erysipelotrichia*, are producers of short-chain fatty acids in the gut, including butyrate [13, 14]. We observed a significant decrease in most organic acids in the gut metabolome following streptomycin treatment, and an increase when *Clostridia* and *Erysipelotrichia* were given to germ-free mice. It is thus possible that these organic acids represent *Clostridia* and *Erysipelotrichia*-derived fermentation products, but the un-targeted metabolomics approach performed in this study was unable to detect short-chain fatty acids, because of poor retention and separation of these metabolites in the columns used. We thus focused our analysis on metabolites depleted by *Clostridia* and *Erysipelotrichia*, which was indicative of nutrients consumed by members of these taxa. Notably, our data suggested that when commensal *Clostridia* and *Erysipelotrichia* were depleted in the murine large bowel, most of the detected sugar alcohols and sugar acids significantly increased; whereas; when germ-free mice were associated with a defined community of 14 human *Clostridia* and 3 human *Erysipelotrichia* isolates, the inverse held true. Furthermore, *Clostridia* and *Erysipelotrichia* isolates (i.e., strains 1, 3, 7, 9, and 29) were able to grow on sugar alcohols as sole carbon sources *in vitro* and had a combined relative abundance of approximately 12% within the defined community of 14 human

Clostridia and 3 human *Erysipelotrichia* isolates, thus explaining their ability to deplete polyols in gnotobiotic mice.

Sugar alcohols, such as sorbitol, xylitol or mannitol, are of interest because of their use as low calorie sweeteners in foods [28]. They belong to a class of sugars known as polyols, which fall under the umbrella of fermentable dietary oligosaccharides, disaccharides, monosaccharides, and polyols (FODMAPs) that are poorly absorbed in the small intestine [29]. As humans lack active transporters for sugar alcohols in the small intestine, transport relies on passive diffusion, which is why these nutrients are poorly absorbed in the small intestine [29]. As a result, the bulk of dietary sugar alcohols reaches the colon, where they are fermented by the microbiota into short-chain fatty acids [30]. Sugar alcohols are naturally occurring in many fruits and vegetables, making them a prevalent constituent of the human diet [31], but excessive consumption is associated with transient bloating or diarrhea [32]. Excessive consumption of sugar alcohols also is associated with enhanced gas production and diarrhea in patients with irritable bowel syndrome [33, 34] and a diet low in FODMAPs alleviates symptoms in these patients [35]. Our results show that a *Clostridia* species, *A. caccae*, that was able to grow on sorbitol *in vitro*, was able to deplete this polyol when it was provided to antibiotic-treated mice as a dietary supplement.

Interestingly, a previous study suggests that sugar alcohols could be a growth factor for opportunistic enteric pathogens. *Clostridioides difficile* (*C. difficile*) is an opportunistic pathogen within the class *Clostridia,* which causes pseudomembranous colitis in individuals undergoing antibiotic therapy [36]. The metabolome of mice treated

with the broad-spectrum antibiotic cefoperazone contains a higher abundance of sugar alcohols relative to the metabolome of mock treated animals, similar to what we observed after streptomycin treatment, and this treatment renders mice susceptible to *C. difficile* infection [37]. Additionally, the sugar alcohols sorbitol and mannitol are sufficient for *C. difficile* growth in NCDM *in vitro* [37] and increase toxin production [38]. Furthermore, depletion of alcoholic sugars by the gut microbiota contributes to colonization resistance against *C. difficile* [39]. Similarly, colonization resistance against opportunistic *Escherichia coli* infection is lost when mice receive supplementation with galacitol, a sugar alcohol [40]. Thus, depletion of sugar alcohols by endogenous *Clostridia* and *Erysipelotrichia* species might contribute to colonization resistance against opportunistic enteric pathogens.

Conclusions

Our results identify members of the classes *Clostridia* and *Erysipelotrichia* as the main taxa in the gut microbiota that consume dietary FODMAPs during homeostasis. This finding points to *Clostridia* and *Erysipelotrichia* as a possible cause of transient bloating observed after excessive consumption of FODMAPs.

MATERIALS AND METHODS

Bacterial strains and culture techniques:

A community of 14 human *Clostridia* isolates and 3 human *Erysipelotrichia* isolates was kindly provided by K Honda [21, 41], and bacteria were cultured as described previously [25].

Animal Experiments:

Female C57BL/6 mice, aged 8 weeks, were obtained from The Jackson Laboratory (C57BL/6J mice) and maintained on gamma irradiated Teklad global 18% protein rodent diet 2918 (Envigo). C57BL/6 mice were treated with 20 mg/animal of streptomycin via oral gavage or mock treated and sacrificed three days post treatment, and cecal contents were collected, freeze-dried, flash frozen in liquid nitrogen, and stored at -80°C until further processing for metabolomics. Proximal colon contents were collected, snap frozen, and stored at -80°C until further processing for DNA isolation and sequencing.

Sorbitol supplementation: C57BL/6 mice were maintained on gamma irradiated Teklad Custom Diet TD.110675 (Envigo) and treated with 20 mg/animal of streptomycin via oral gavage or mock treated. One day later mice received drinking water

supplemented with 5% sorbitol or no supplementation and were inoculated with 200µl of a *Clostridia* culture by oral gavage. Cecal contents were collected three days after streptomycin treatment.

Microbiota transfer: Male and female germ-free Swiss-Webster mice were bred in house using gamma irradiated chow (Purina Diet 5066; Charles River) and used for the experiment at 8 weeks of age. For transfer of human *Clostridia* and *Erysipleotrichia* strains into germ-free mice, each strain was individually inoculated into Egerth Gagnon medium and incubated for 48 hours. A mixture of equal volumes for each strain was prepared, and each animal was given 200µl of inoculum via oral gavage [25]. For microbiota transfer from conventional C57BL/6J mice, the entire cecal content of a mouse was collected into 5 mL pre-reduced sterile PBS inside an anaerobic chamber, homogenized by vortexing and set down for 10 min. to let particles settle. Immediately after, 200µl of the homogenate supernatant was used to inoculate germ-free Swiss-Webster mice by oral gavage. For the duration of the experiment, mice were kept in sterile cages inside a biological safety cabinet. The mice were sacrificed 5 days after inoculation, and cecal contents were collected, freeze-dried, flash frozen in liquid nitrogen, then stored at -80°C until further processing for metabolomics; proximal colon content was collected, flash frozen in liquid nitrogen, then stored at -80°C until further processing for DNA isolation and sequencing.

In vitro growth assay of _Clostridia_ and _Erysipleotrichia_ isolates.

To assess the ability of the Clostridia and *Erysipleotrichia* consortium to grow on various individual carbohydrates, we used no carbon defined media (NCDM), which is a modified form of no carbon minimal media (NCMM) [37], but NCDM utilizes below vitamins and amino acids at the following concentrations: Thiamine, pantothenate, nicotinic acid, riboflavin, pyridoxine, p-aminobenzoic acid, folic acid, biotin, and vitamin B12 were added in the form of ATCC Vitamin Supplement (ATCC, Manassas, Virginia) to a final concentration of 5% (v/v). Calcium, magnesium, manganese, iron, and cobalt salts were added in the form of ATCC Trace Mineral Supplement (ATCC, Manassas, Virginia) to a final concentration of 2% (v/v). Indicated carbohydrates were used at a final concentration of 0.5% (w/v). We further modified NCDM to accommodate slower growing strains by using Bacto Casamino acids (ThermoFisher) at a concentration of 4.575 g/L, supplemented with cysteine (400mg/L), methionine (27mg/L), alanine (72 mg/L), and tryptophan (30 mg/L). Vitamin K2 was added at a concentration of 200 µL/L. Individual strains were grown anaerobically in a volume of 2mL of above medium containing glucose for 24 hours. The entire culture volume was then spun down at 12,500 rpm for 3 minutes, the supernatant was discarded, and the bacterial pellet was resuspended in 750 µL of reduced PBS. 25 µL of the resuspended pellet of each strain was then used to inoculate 2mL volumes of NCDM containing D-gluconic acid, glucose, *myo*-inositol, D-mannitol, D-sorbitol, xylitol, or no added carbohydrates in quadruplicate. After 72 hours of incubation at 37°C, growth was assessed by removing the cultures from the anaerobic chamber and reading the optical density at 600 nm. Un-inoculated NCDM was used as a blank reference.

16S rRNA gene amplicon sequencing sample preparation, library preparation, and sequencing:

Nucleic acid extraction was done using the DNeasy Blood & Tissue Kits from Qiagen. DNA concentration was recorded using a Qubit a 2.0 fluorometer. Samples were normalized to 20 ng/μL before library preparation. The 16sV4 Region was amplified with the 515f-806R primer pair. Each 10 μL amplicon-PCR reaction consisted of 7 μL of TailorMix 2X SYBR Green qPCR Master Mix (SeqMatic, Fremont, CA, USA), 1 μL of 10 μM of each primer and 1 μL of DNA template. Each sample was denatured at 95°C for 10 minutes before undergoing 25 cycles of 95°C for 15 seconds, 60°C for 1 minute and 72°C for 1.5 minutes and a final extension at 72°C for 10 minutes. The amplicons were pooled, purified with the DNA/RNA Purification Beads (SeqMatic, Fremont, CA, USA) and resuspended in 30 μL of 10mM Tris, pH 8.5. The final purified library was quantified with the 2200 TapeStation (Agilent, Santa Clara, CA, USA). The Purified Library pool was quantified using Qubit and 2200 Tapestation and then diluted to 20 pM. The pool was then denatured at 94°C on the MJ Research PTC-200 Thermal Cycler and hybridized to Illumina's HT1 buffer. Final quantification was done using qPCR on the Roche Lightcycler 96 system. Using results from the final qPCR, the pool was loaded into a 300 cycle Illumina MiSeq Cartridge at 5 pM and sequenced to generate 150 Paired end reads. Sequencing was performed via Illumina MiSeq.

For 16S rRNA amplicon library preparation and sequencing of colon contents from germ-free mice inoculated with the 14 human *Clostridia* isolates and 3 human *Erysipelotrichia* isolates, Primers 319F (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG(spacer)*GT*ACTCCTACGGGAGGCAGCAGT) and 806R (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG(spacer)*CC*GGACTACNVGGGTWTCTAAT were used to amplify the V3-V4 domain of the 16S rRNA using a two-step PCR procedure. In step one of the amplification procedure, both forward and reverse primers contained an Illumina tag sequence (bold), a variable length spacer (no spacer, C, TC, or ATC for 319F; no spacer, G, TG, ATG for 806R) to increase diversity and improve the quality of the sequencing run, a linker sequence (italicized), and the 16S target sequence (underlined). Each 25 µl PCR reaction contained 1 Unit Kapa2G Robust Hot Start Polymerase (Kapa Biosystems), 1.5 mM $MgCl_2$, 0.2 mM final concentration dNTP mix, 0.2 µM final concentration of each primer and 1µl of DNA for each sample. PCR conditions were: an initial incubation at 95°C for 3 min, followed by 25 cycles of 95°C for 45 seconds, 50°C for 30 seconds, 72°C for 30 seconds and a final extension of 72°C for 3 minutes. In step two, each sample was barcoded with a unique forward and reverse barcode combination using forward primers (AATGATACGGCGACCACCGAGATCTACACNNNNNNNNTCGTCGGCAGCGTC) with an Illumina P5 adapter sequence (bold), a unique 8 nt barcode (N), a partial matching sequence of the forward adapter used in step one (underlined), and reverse primers (CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTCTCGTGGGCTCGG)) with an

Illumina P7 adapter sequence (bold), unique 8 nt barcode (N), and a partial matching sequence of the reverse adapter used in step one (underlined). The PCR reaction in step two contained 1 Unit Kapa2G Robust Hot Start Polymerase (Kapa Biosystems), 1.5 mM $MgCl_2$, 0.2 mM final concentration dNTP mix, 0.2 µM final concentration of each uniquely barcoded primer and 1ul of the product from the PCR reaction in step one diluted at a 10:1 ratio in water. PCR conditions were: an initial incubation at 95°C for 3 min, followed by 8 cycles of 95°C for 30 seconds, 58°C for 20 seconds, 72°C for 20 seconds and a final extension of 72°C for 3 minutes.

The final product was quantified on the Qubit instrument using the Qubit Broad Range DNA kit (Invitrogen) and individual amplicons were pooled in equal concentrations. The pooled library was cleaned utilizing Ampure XP beads (Beckman Coulter) then the band of interest was further subjected to isolation via gel electrophoresis on a 1.5% Blue Pippin HT gel (Sage Science). The library was quantified via qPCR followed by 300-bp paired-end sequencing using an Illumina MiSeq instrument in the Genome Center DNA Technologies Core, University of California, Davis.

16S rRNA gene amplicon sequencing analysis:

Sequencing reads were demultiplexed using QIIME 1.8 [42], and non-biological nucleotides were trimmed using Trimmomatic [43]. 16S rRNA sequencing reads were subsequently processed and assembled into amplicon sequence variants (ASV) using

dada2 [44] in R. First, reads with more than 2 expected errors were removed. Dereplication and sample inference were then performed on forward and reverse reads, prior to merging. A sequence table was constructed from merged reads, and chimeric reads were subsequently removed. Taxonomy was assigned to reads to the genus level using the dada2 formatted rdp training dataset 14 which can be found here: https://zenodo.org/record/158955#.XJqlnxNKjUI. The R package phyloseq [45] was then used in downstream analysis of the data, including the generation of a phyloseq object, relative abundance bar plots, and the principle coordinate analysis plot. For the weighted unifrac analysis, the R package msa [46] was used to generate a multiple sequence alignment from the assembled reads with the following parameters: method="ClustalW", type="dna", order="input". The R package phangorn [47] was used to generate a maximum likelihood tree from the sequence alignment using a general time reversible (GTR) model with the following parameters: model="GTR", optInv=TRUE, optGamma=TRUE, rearrangement = "stochastic", control = pml.control(trace = 0). The R package vegan was used to perform permanova analysis on weighted unifrac distances with default parameters and the R package ggplot2 was used to graph boxplots of the weighted unifrac distances. The diversity pie chart was generated by computing the ratio of ASVs within each taxonomic class level present to the total number using the R package plyr and then graphed with ggplot2. The relative abundance pie chart was generated by using phyloseq and then graphed using ggplot2. For linear discriminant analysis, data were reformatted in R, written to a tab separated text file, and then uploaded to the LEfSe galaxy server [48]

where the default statistical parameters were used in the analysis to generate LDA scores and the LDA cladogram. Differential abundance analysis of taxa was performed using the R package deseq2 [25516281] with the parameters: test="Wald", fitType="parametric", cooksCutoff = FALSE. The R package omu [49] was then used to generate fold change frequency tables and differential abundance bar plots from the deseq2 modelled data.

For 16S rRNA sequencing analysis of colon contents from ex-germ-free mice inoculated with the 14 human *Clostridia* isolates and 3 human *Erysipelotrichia* isolates, reads were assigned strain level taxonomy by creating a local nucleotide blast database of the corresponding genomes using command line blast, and then querying each of the 26 ASVs against the database. ASVs that matched the same strain (due to dada2 identifying multiple 16S gene copies from the same taxon as unique ASVs) were aggregated prior to the construction of the abundance boxplot using a combination of phyloseq and ggplot2.

Phylogeny of human *Clostridia* and *Erysipelotrichia* isolates

The genome for strain 4 was omitted from this analysis because the sequence provided in NCBI was flagged as chimeric. Analysis of the chimeric sequence of strain 4 suggested that it was derived from two strains belonging to the same species, *Hungatella hathewayi.* To determine the taxonomy of the remaining 13 *Clostridia* strains and 3

Erysipelotrichia strains, we employed a phylogenetic approach. 63 *Clostridia* and *Erysipelotrichia* genomes were curated, along with 3 *Alicyclobacilus* genomes that were included as an outgroup (**Supplementary Table 6**). Protein alignments of the 66 curated genomes, along with the 16 strains from the defined community of *Clostridia* and *Erysipelotrichia* isolates, were generated using clustal omega [50] for the following single copy genes: 30S ribosomal protein S3, 30S ribosomal protein S7, 30S ribosomal protein S9, 30S ribosomal protein S10, 30S ribosomal protein S12, 30S ribosomal protein S13, 30S ribosomal protein S17, 50S ribosomal protein L3, 50S ribosomal protein L14, 50S ribosomal protein L15, and 50S ribosomal protein L16. Alignments were then concatenated, providing a total alignment length of 1822 amino acids. We then used RaxML [51], with the PROTGAMMAWWAG rate of heterogeneity model to generate 20 maximum likelihood trees in order to determine the best scoring tree. Non-parametric boostrapping was subsequently performed to determine the confidence of branch placements. 150 bootstraps were sufficient for the model to reach convergence. Bootstraps were then used to draw bipartitions to the best fit tree determined by the maximum likelihood model. The tree was imported into R, and visualized using the package ggtree [52].

GC-TOF mass spectrometry:

Data were acquired using the following chromatographic parameters, with more details to be found in [53]. Column: Restek corporation Rtx-5Sil MS (30 m length x 0.25 mm internal diameter with 0.25 µm film made of 95% dimethyl/5%diphenylpolysiloxane) Mobile phase: Helium Column temperature: 50-330°C Flowrate: 1 mL min-1 Injection volume: 0.5 µL Injection: 25 splitless time into a multi-baffled glass liner Injection temperature: 50°C ramped to 250°C by 12°C s -1 Oven temperature program: 50°C for 1 min, then ramped at 20°C min-1 to 330°C, held constant for 5 min. The analytical GC column is protected by a 10 m long empty guard column which is cut by 20 cm intervals whenever the reference mixture QC samples indicate problems caused by column contaminations. We have validated that at this sequence of column cuts, no detrimental effects are detected with respect to peak shapes, absolute or relative metabolite retention times or reproducibility of quantifications. We use automatic liner exchanges after each set of 10 injections which we could show to reduce sample carryover for highly lipophilic compounds such as free fatty acids. Mass spectrometry parameters are used as follows: a Leco Pegasus IV mass spectrometer is used with unit mass resolution at 17 spectra s -1 from 80-500 Da at -70 eV ionization energy and 1800 V detector voltage with a 230°C transfer line and a 250°C ion source.

Metabolomics data processing:

Raw data files are preprocessed directly after data acquisition and stored as ChromaTOF-specific *.peg files, as generic *.txt result files and additionally as generic ANDI MS *.cdf files. ChromaTOF vs. 2.32 is used for data preprocessing without smoothing, 3 s peak width, baseline subtraction just above the noise level, and automatic mass spectral deconvolution and peak detection at signal/noise levels of 5:1 throughout the chromatogram. Apex masses are reported for use in the BinBase algorithm. Result *.txt files are exported to a data server with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database. The BinBase algorithm (rtx5) used the settings: validity of chromatogram (<10 peaks with intensity >10^7 counts s -1), unbiased retention index marker detection (MS similarity>800, validity of intensity range for high m/z marker ions), retention index calculation by 5th order polynomial regression. Spectra are cut to 5% base peak abundance and matched to database entries from most to least abundant spectra using the following matching filters: retention index window ±2,000 units (equivalent to about ±2 s retention time), validation of unique ions and apex masses (unique ion must be included in apexing masses and present at >3% of base peak abundance), mass spectrum similarity must fit criteria dependent on peak purity and signal/noise ratios and a final isomer filter. Failed spectra are automatically entered as new database entries if s/n >25, purity <1.0 and presence in the biological study design class was >80%. All thresholds reflect settings for ChromaTOF v. 2.32. Quantification is reported as peak height using the unique ion as default, unless a different quantification ion is manually set in the BinBase administration software BinView. A quantification report table is produced for all

database entries that are positively detected in more than 10% of the samples of a study design class (as defined in the miniX database) for unidentified metabolites. A subsequent post-processing module is employed to automatically replace missing values from the *.cdf files. Replaced values are labeled as ‘low confidence’ by color coding, and for each metabolite, the number of high-confidence peak detections is recorded as well as the ratio of the average height of replaced values to high-confidence peak detections. These ratios and numbers are used for manual curation of automatic report data sets to data sets released for submission. These data were then normalized to the mTIC value (sum of the peak heights of the known metabolites).

Metabolomics data analysis:

Metabolomics data were loaded into R, and the following analysis was performed using the R package omu [49]. Compound hierarchy was assigned from the KEGG [54] hierarchy databases in omu’s system data. The data were normalized using the natural log, a student’s t test was performed to calculate p values, a benjamini-hochburg procedure was used to perform false discovery rate correction for type 2 errors, and log2 fold change of group means was calculated to visualize effect size. Fold change frequency tables were generated to make bar plots. Volcano plots were generated using FDR corrected p values and log2 fold change between group means. Principle component analysis was performed on natural log transformed data.

D-sorbitol measurement

Mouse cecal contents were homogenized in 1 mL of sterile PBS, centrifuged at 300g for 10 minutes at 4C and then the supernatants were collected. The supernatants were filtered through a 10-kDa spin column (Biovision, Milpitas, CA). D-sorbitol concentrations were determined by D-sorbitol assay kit (Biovision, Milpitas, CA) according to the manufacturer's instruction.

DECLARATIONS

Ethics approval and consent to participate

All experiments in this study were approved by the Institutional Animal Care and Use Committee at the University of California at Davis.

Consent for publication

Not applicable

Availability of data and material

Illumina sequences obtained in the present study were deposited in the Sequence Read Archives (SRA) NCBI database under BioProject ID: PRJNA560082. Untargeted metabolomics data are provided in Supplementary Tables 1, 3 and 4.

Competing interests

The authors declare that they have no competing interests

Funding

Work in A.J.B.'s laboratory was supported by USDA/NIFA award 2015-67015-22930 (A.J.B.), by Crohn's and Colitis Foundation of America Senior Investigator Award # 650976 and by Public Health Service Grants AI044170, AI096528, AI112445, AI112949, AI146432 and AI153069. A.W.L.R. was supported by Public Health Service Grant AI060555. E.E.O. was supported by Public Health Service Grants TR001860 and TR001861.

Authors' contributions

CRT analyzed metabolomics data and microbiota profiling data and wrote the manuscript. AWLR performed and analyzed in vitro growth experiments. EEO helped perform mouse experiments. FF performed mouse experiments and analyzed the data. AJB was a major contributor in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to acknowledge the mass spectrometry-based Metabolomics Service Core at the West Coast Metabolomics Center for expert sample analysis. We would like to acknowledge Briana Young for expert technical assistance.

Figures

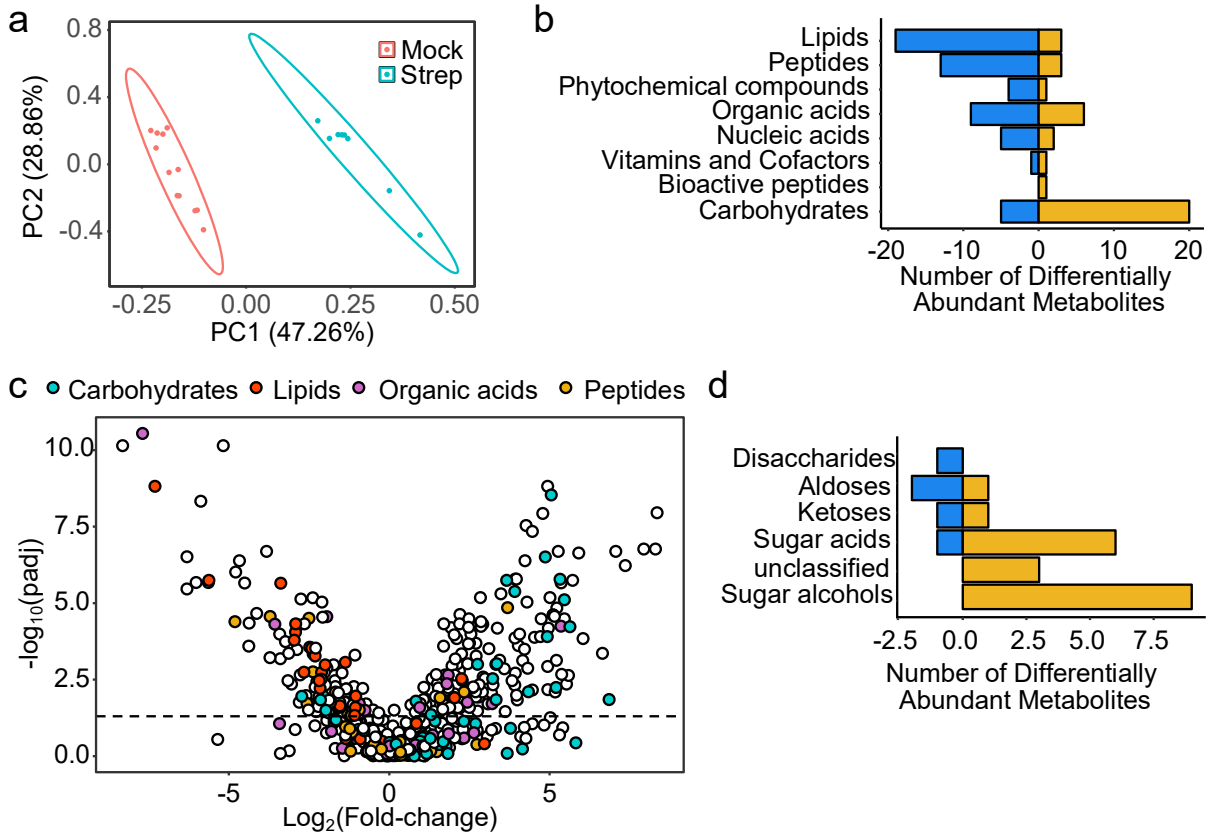

Figure 3.1

Figure 3.1: Streptomycin treatment alters the mouse cecal metabolome, characterized by an increase in carbohydrates. Mice were mock-treated (n = 12) or received a single dose of streptomycin (n = 8) and the cecal metabolome was analyzed three days later (mice of each group were housed in three different cages each). (a) PCA plot showing variation of the mouse cecal metabolome by treatment group. (b) Bar plot displaying the number of significantly increased and decreased metabolites by metabolite class in streptomycin-treated mice relative to mock-treated mice (FDR corrected *P* value = < 0.05). (c) Volcano plot of metabolites colored by class. Metabolites with a positive fold-change value increased with streptomycin treatment while metabolites with a negative-fold change value decreased with streptomycin treatment. The dashed line is set at an FDR corrected *P* value of 0.05. (d) Bar plot displaying the number of differentially significantly increased and decreased carbohydrates by metabolite subclass in streptomycin treated mice relative to mock-treated mice (FDR corrected *P* value =< 0.05).

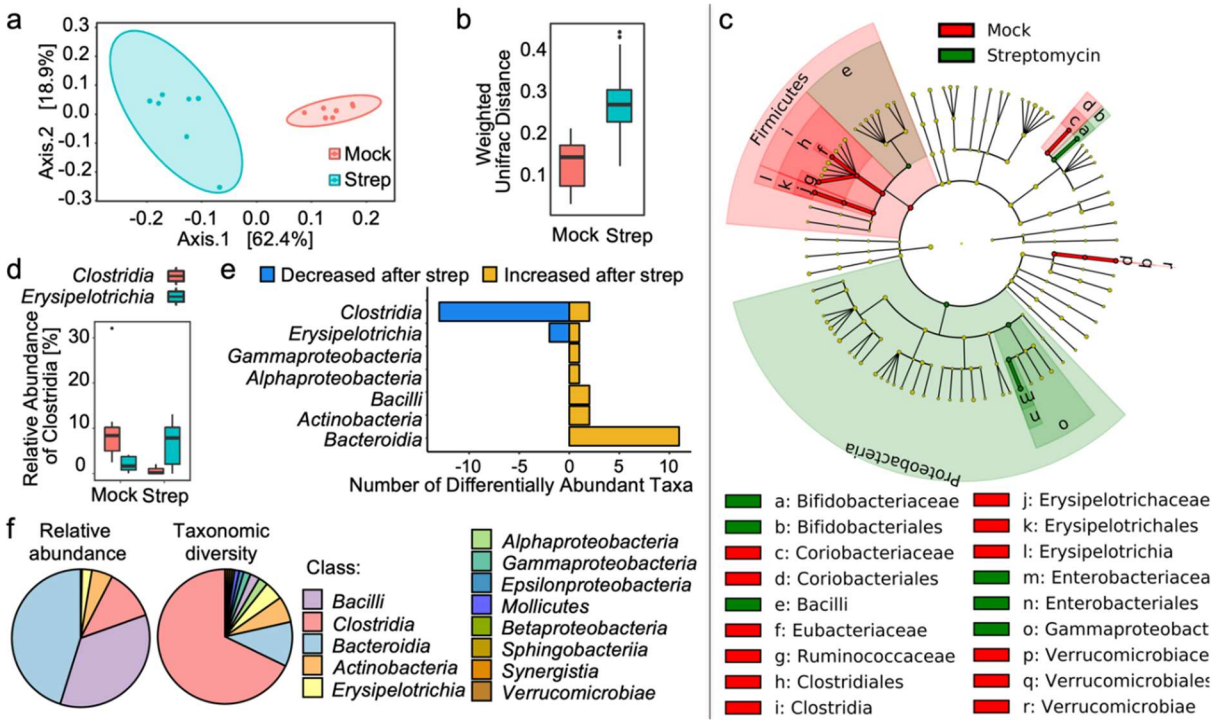
a
Axis.2 [18.9%]
Axis.1 [62.4%]
Mock
Strep
b
Weighted Unifrac Distance
Mock Strep
c
Mock
Streptomycin
Firmicutes
Proteobacteria
a: Bifidobacteriaceae
b: Bifidobacteriales
c: Coriobacteriaceae
d: Coriobacteriales
e: Bacilli
f: Eubacteriaceae
g: Ruminococcaceae
h: Clostridiales
i: Clostridia
j: Erysipelotrichaceae
k: Erysipelotrichales
l: Erysipelotrichia
m: Enterobacteriacea
n: Enterobacteriales
o: Gammaproteobact
p: Verrucomicrobiace
q: Verrucomicrobiales
r: Verrucomicrobiae
d
Clostridia
Erysipelotrichia
Relative Abundance of Clostridia [%]
Mock Strep
e
Decreased after strep
Increased after strep
Clostridia
Erysipelotrichia
Gammaproteobacteria
Alphaproteobacteria
Bacilli
Actinobacteria
Bacteroidia
Number of Differentially Abundant Taxa
f
Relative abundance
Taxonomic diversity
Class:
Bacilli
Clostridia
Bacteroidia
Actinobacteria
Erysipelotrichia
Alphaproteobacteria
Gammaproteobacteria
Epsilonproteobacteria
Mollicutes
Betaproteobacteria
Sphingobacteriia
Synergistia
Verrucomicrobiae

Figure 3.2

Figure 3.2: Streptomycin treatment of mice induces a shift in community structure of the large bowel microbiome. Mice were mock-treated (n = 8) or received a single dose of streptomycin (n = 8) and bacterial communities in the colon contents were analyzed three days later. (a) Principle coordinate analysis of 16S rRNA gene amplicon sequencing of colon contents from mice. (b) Box plot of weighted unifrac distance between mock and streptomycin treated mice (*P* value = 0.002). (c) LEFSE cladogram displaying taxa enriched in mock-treated mice (red) or in streptomycin-treated mice (green). Taxa enriched within higher level taxa are indicated by a darker shading. (d) Box plots showing the relative abundance of *Clostridia* and *Erysipelotrichia* in mock-treated and streptomycin-treated mice. (e) Significantly increased and decreased amplicon sequence variants by class in streptomycin-treated mice relative to mock-treated mice (FDR corrected *P* value =<0.05). (f) Relative abundance (left pie chart) and taxonomic diversity of ASVs (right pie chart) in mock-treated mice at the class level.

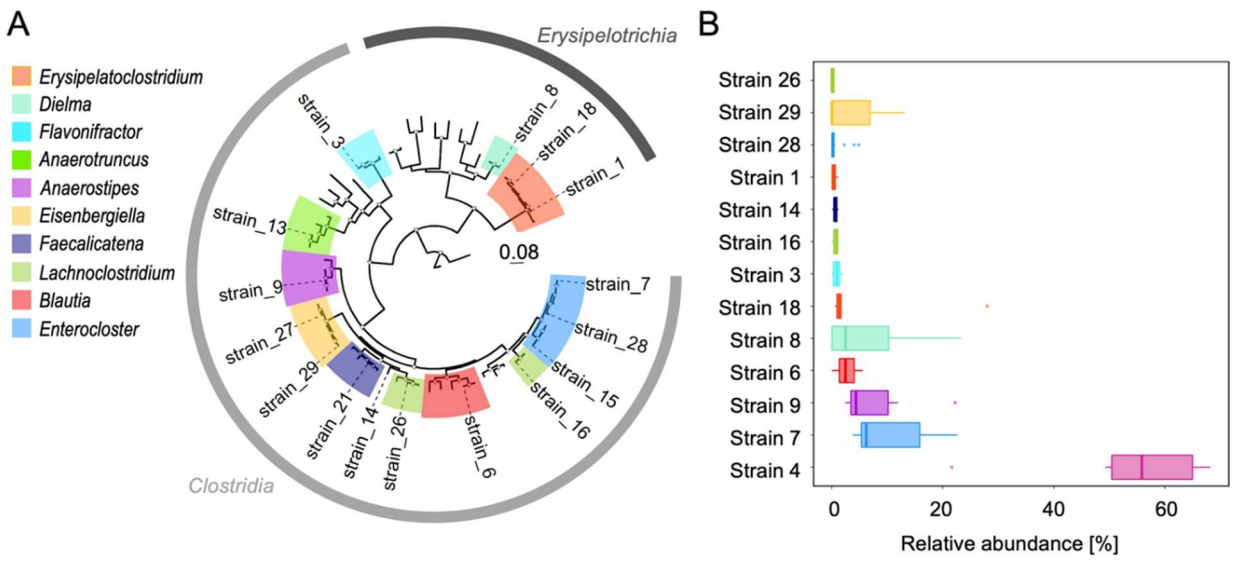
A
Erysipelatoclostridium
Dielma
Flavonifractor
Anaerotruncus
Anaerostipes
Eisenbergiella
Faecalicatena
Lachnoclostridium
Blautia
Enterocloster
Erysipelotrichia
Clostridia
strain_3
strain_8
strain_18
strain_1
strain_13
strain_9
strain_27
strain_29
strain_21
strain_14
strain_26
strain_6
strain_7
strain_28
strain_15
strain_16
0.08
B
Strain 26
Strain 29
Strain 28
Strain 1
Strain 14
Strain 16
Strain 3
Strain 18
Strain 8
Strain 6
Strain 9
Strain 7
Strain 4
0
20
40
60
Relative abundance [%]

Figure 3.3

Figure 3.3: Engraftment of germ-free mice with a defined community of human *Clostridia* and *Erysipelotrichia* isolates. (A) Placement of the members of a defined microbial community (Strains 1, 3, 6, 7, 8, 9, 13, 14, 15, 16, 18, 21, 26, 27, 28, and 29) on a maximum likelihood phylogeny built from the alignments of 11 total single copy 30S and 50S ribosomal proteins, which were concatenated together for a total alignment length of 1822 amino acids. The tree encompasses 66 taxa, three of which being *Alicyclobacillus* species belonging to the class *Bacilli* that were included as an outgroup. Strain 4 from the defined microbial community was excluded from analysis because NCBI flagged its deposited genomic sequence as chimeric, containing sequences from two strains belonging to the same species, *Hungatella hathewayi*. Sequences that did not belong to strains within the defined microbial community were obtained from [55]. 150 non-parametric bootstraps were performed to determine confidence of branch placements (white triangle at a node indicate above 70% confidence). (B) Germ-free mice were engrafted with a community of 14 *Clostridia* and 3 *Erysipelotrichia* isolates and microbiota profiling performed on colon contents obtained five days later.

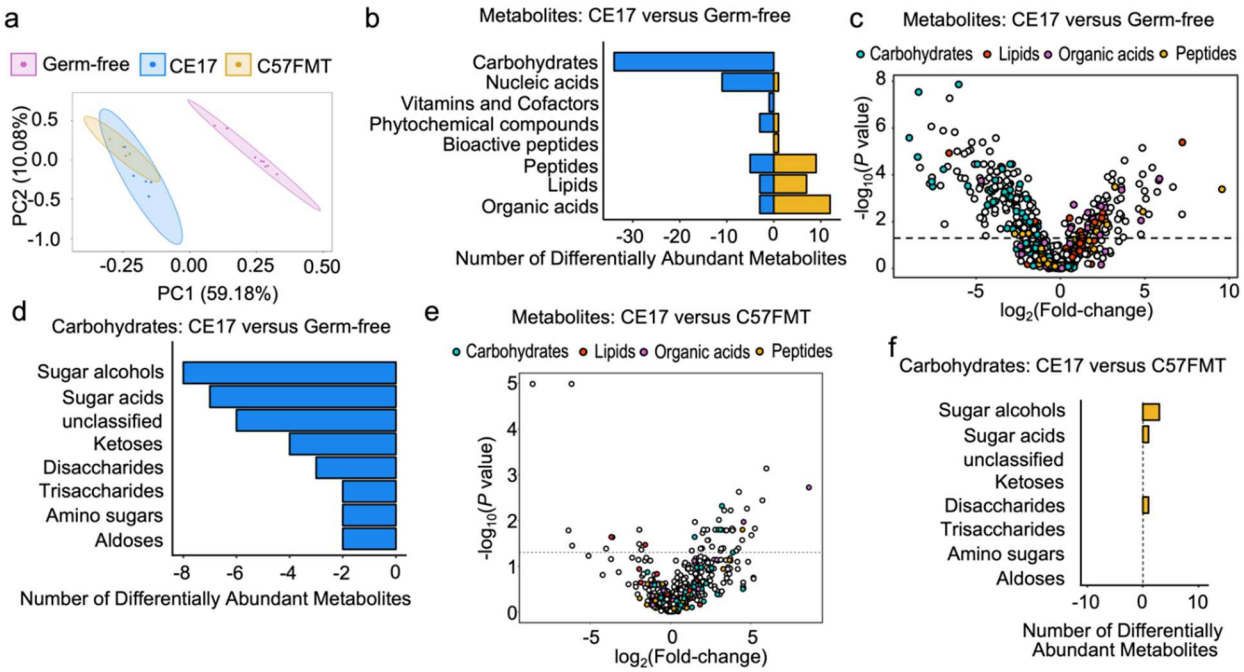
a
Germ-free
CE17
C57FMT
PC2 (10.08%)
PC1 (59.18%)
b
Metabolites: CE17 versus Germ-free
Carbohydrates
Nucleic acids
Vitamins and Cofactors
Phytochemical compounds
Bioactive peptides
Peptides
Lipids
Organic acids
Number of Differentially Abundant Metabolites
c
Metabolites: CE17 versus Germ-free
Carbohydrates
Lipids
Organic acids
Peptides
-log10(P value)
log2(Fold-change)
d
Carbohydrates: CE17 versus Germ-free
Sugar alcohols
Sugar acids
unclassified
Ketoses
Disaccharides
Trisaccharides
Amino sugars
Aldoses
Number of Differentially Abundant Metabolites
e
Metabolites: CE17 versus C57FMT
Carbohydrates
Lipids
Organic acids
Peptides
-log10(P value)
log2(Fold-change)
f
Carbohydrates: CE17 versus C57FMT
Sugar alcohols
Sugar acids
unclassified
Ketoses
Disaccharides
Trisaccharides
Amino sugars
Aldoses
Number of Differentially Abundant Metabolites

Figure 3.4

Figure 3.4: Colonization of germ-free mice with Clostridia induces a shift in the cecal metabolome. Comparative analysis of the cecal metabolome of germ-free mice (n = 6), gnotobiotic mice associated with a defined microbial community containing 14 human *Clostridia* and 3 human *Erysipelotrichia* isolates (CE17) (n = 6), and gnotobiotic mice receiving a cecal microbiota transplant from conventional C57BL/6J mice (C57FMT) (n = 4). (a) Principle component analysis of the cecal metabolome in the indicated groups of mice. (b) Bar plot of significantly increased and decreased metabolites by metabolite class in gnotobiotic mice associated with a defined microbial community (CE17) relative to germ-free mice (FDR corrected *P* value = < 0.05). (c) Volcano plot of metabolites colored by class. Metabolites with a positive fold-change value increased in mice engrafted with a defined microbial community compared to germ-free mice, while metabolites with a negative fold-change value decreased. (d) Bar plot of significantly increased and decreased carbohydrates by metabolite subclass in gnotobiotic mice associated with a defined microbial community (CE17) relative to germ-free mice (FDR corrected *P* value = < 0.05). The dashed line is set at an FDR corrected *P* value of 0.05. (e) Volcano plot of metabolites colored by class. Metabolites with a positive fold-change value increased in mice engrafted with a defined microbial community compared to gnotobiotic mice receiving a cecal microbiota transplant from conventional C57BL/6J mice (C57FMT), while metabolites with a negative fold-change value decreased. (f) Bar plot of significantly increased and decreased metabolites by metabolite class in gnotobiotic mice associated with a defined microbial community (CE17) relative to

gnotobiotic mice receiving a cecal microbiota transplant from conventional C57BL/6J mice (C57FMT) (FDR corrected *P* value = < 0.05).

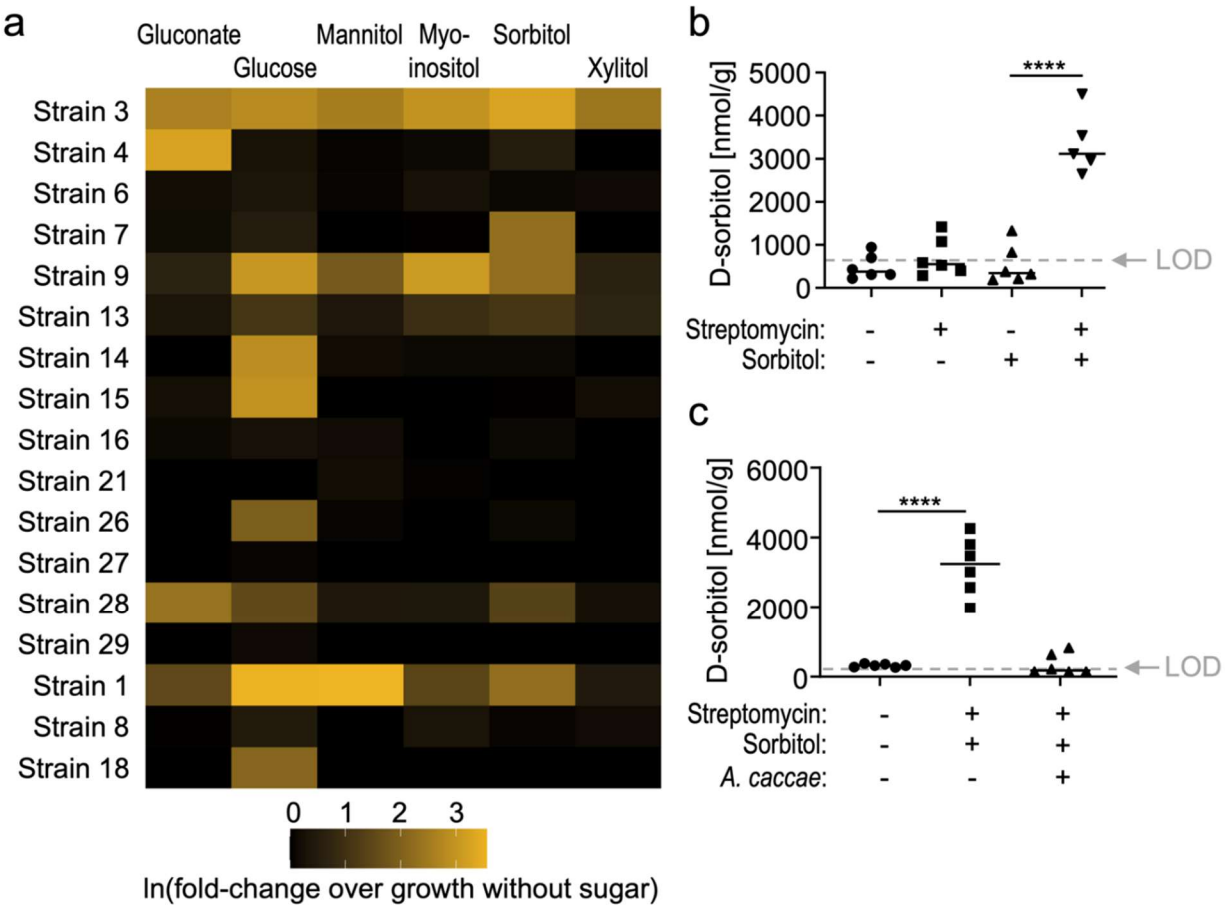
a
Gluconate
Glucose
Mannitol
Myo-inositol
Sorbitol
Xylitol
Strain 3
Strain 4
Strain 6
Strain 7
Strain 9
Strain 13
Strain 14
Strain 15
Strain 16
Strain 21
Strain 26
Strain 27
Strain 28
Strain 29
Strain 1
Strain 8
Strain 18
0 1 2 3
ln(fold-change over growth without sugar)
b
D-sorbitol [nmol/g]
5000
4000
3000
2000
1000
0
****
LOD
Streptomycin: - + - +
Sorbitol: - - + +
c
D-sorbitol [nmol/g]
6000
4000
2000
0
****
LOD
Streptomycin: - + +
Sorbitol: - + +
A. caccae: - - +

Figure 3.5

Figure 3.5: *In vivo* depletion of sorbitol by Sorbitol-eaters. (a) Heat map showing growth of individual *Clostridia* and *Erysipelotrichia* isolates on no carbon defined media (NCDM) supplemented with the indicated sugars. (b and c) Conventional C57BL/6J mice were mock treated or treated with a single dose of streptomycin and received drinking water supplemented with or without 5% sorbitol. (b) The concentration of sorbitol in ceca contents was measured by ELISA 3 days after streptomycin treatment. (c) One day after streptomycin treatment, mice were mock inoculated or inoculated with *A. caccae* (designated “strain 9” in panel a). The concentration of sorbitol in cecal contents was measured by a colorimetric assay 3 days after streptomycin treatment. LOD, limit of detection; ****, $P < 0;0001$.

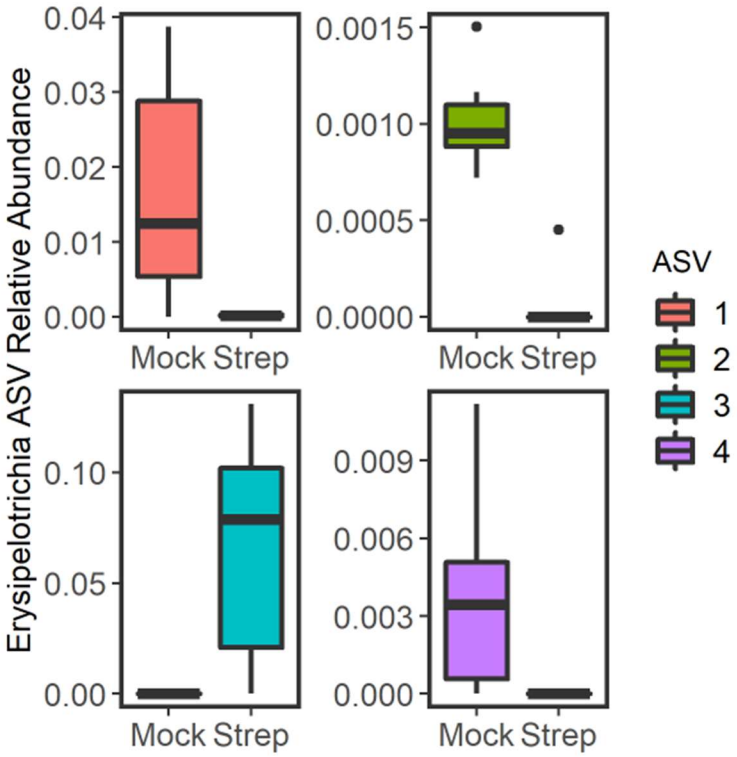
Erysipelotrichia ASV Relative Abundance
0.04
0.03
0.02
0.01
0.00
Mock Strep
0.0015
0.0010
0.0005
0.0000
Mock Strep
0.10
0.05
0.00
Mock Strep
0.009
0.006
0.003
0.000
Mock Strep
ASV
1
2
3
4

Figure 3.S.1

Figure 3.S.1: Amplicon sequence variants (ASVs) belonging to *Erysipelotrichia* which significantly changed abundance after antibiotic treatment (related to Fig. 3.2). Mice were mock-treated (*n* = 8) or received a single dose of streptomycin (*n* = 8) and bacterial communities in the colon contents were analyzed 3 days later. Box plots showing the relative abundance of *Erysipelotrichia* ASVs in mock-treated and streptomycin-treated mice.

Metabolite	Strep.mean	Mock.mean	Fold_Change	padj
1,5-anhydroglucitol	11699.63	1546.583	7.564822	4.89E-05
erythritol	21997.12	664	33.12819	2.90E-09
galactinol	434435.8	3750.167	115.8444	0.014197
galactonic acid	19531	963.6667	20.26738	0.007944
glyceric acid	11815.12	35179.83	0.335849	0.025473
lactobionic acid	9494.5	311.6667	30.46363	0.000125
lyxitol	49792	4968.667	10.0212	0.000997
myo-inositol	136575.4	3715.667	36.75662	0.005713
pentitol	7158.125	474.5	15.08562	4.16E-06
ribonic acid	39540.13	987.4167	40.04401	1.66E-06
saccharic acid	17418.5	350.5	49.69615	5.87E-05
sorbitol	106775.1	2430	43.94037	7.76E-06
threitol	5630.75	445.4167	12.64153	1.79E-06
threonic acid	5692.75	557.0833	10.21885	0.000962
xylitol	42092.63	1444.417	29.1416	3.10E-07
xylonic acid	5014.5	536.4167	9.348143	0.002935

SupplementaryTable 3.1

Supplementary Table 3.1: Table of significantly depleted carbohydrates (related to Fig. 3.1).

Strain	Old Nomenclature	Updated Nomenclature
1	Clostridium saccharogumia	Erysipelatoclostridium saccharogumia
3	Lachnospiraceae_7_1_58FAA	Flavonifractor plautii
4	Clostridium hathewayi	Hungatella hathewayi
6	Blautia producta	Blautia producta
7	Clostridium bolteae	Enterocloster bolteae
8	Clostridiaceae JC13	Dielma fastidiosa
9	Clostridium indolis	Anaerostipes caccae
13	Anaerotruncus colihominis	Anaerotruncus colihominis
14	Ruminococcus sp. ID8	Unclassified Lachnospiraceae
15	Clostridium asparagiforme	Enterocloster asparagiformis
16	Clostridium sp. 7_3_54FAA	Lachnoclostridium symbiosum
18	Clostridium ramosum	Erysipelatoclostridium ramosum
21	Eubacterium contortum	Faecalicatena fissicatena
26	Clostridium scindens	Lachnoclostridium scindens
27	Lachnospiraceae_3_1_57FAA_CT1	Eisenbergiella massiliensis
28	Clostridiales 1_7_47FAA	Enterocloster spp.
29	Lachnospiraceae_3_1_57FAA_CT1	Eisenbergiella tayi

Supplementary Table 3.2

Supplementary Table 3.2: Key including old CE17 nomenclature and updated CE17 nomenclature. (reference to Fig 3.3 and Fig 3.5).

Metabolite	C17 Mean	Germ Free Mean	C17/GF Fold Change	Adjusted Pvalue
digalacturonic acid	1189.8	17027.88	0.069874	0.001338
erythritol	846.6	6693.125	0.126488	0.000448
galactinol	4126	1384386	0.00298	2.88E-08
galactonic acid	1400.4	62401.25	0.022442	0.000299
galacturonic acid	555.6	6127.5	0.090673	5.62E-05
glyceric acid	4101	19642.75	0.208779	0.004349
lactobionic acid	275.6	17898	0.015398	1.38E-08
lyxitol	2008	244648.1	0.008208	5.83E-05
myo-inositol	2691.8	44464	0.060539	0.00038
pentitol	1379.2	3471.375	0.397307	0.017015
ribitol	1425	9910.875	0.143781	0.021147
sorbitol	2569	503660.2	0.005101	0.000196
threonic acid	2036.2	7112.75	0.286275	0.010194
xylitol	2241.2	32355	0.069269	0.00017
xylonic acid	929.8	3694.375	0.25168	0.001071

Supplementary Table 3.3

Supplementary Table 3.3: Significantly depleted carbohydrates in germ-free mice which received the CE17 compared to germ-free mice (reference to Fig. 3.2).

Accession	Genus	Species
GCA_000144625.1	Lacrimispora	saccharolytica
GCF_000526995.1	Lacrimispora	indolis
GCF_000687555.1	Lacrimispora	aerotolerans
GCF_003609635.1	Lacrimispora	algidixylanolytica
GCF_014131675.1	Anaerostipes	caccae
GCA_000332875.2	Anaerostipes	hadrus
GCF_005280655.1	Anaerostipes	rhamnosivorans
GCF_016586355.1	Anaerostipes	butyraticus
GCF_014131715.1	Blautia	producta
GCF_000157975.1	Blautia	hydrogenotrophica
GCF_002222595.2	Blautia	Hansenii
GCF_013302295.1	Blautia	Schinkii
GCF_002234575.2	Enterocloster	Bolteae
GCF_013304225.1	Enterocloster	clostridioformis
GCF_000158075.1	Enterocloster	asparagiformis
GCF_900102595.1	Enterocloster	lavalensis
GCF_000466485.1	Lachnoclostridium	symbiosum
GCF_004295125.1	Lachnoclostridium	scindens
GCF_000156515.1	Lachnoclostridium	hylemonae
GCF_001881565.1	Eisenbergiella	tayi
GCF_003435485.1	Eisenbergiella	massiliensis
GCF_003478085.1	Eisenbergiella	OF01-20
GCF_000154565.1	Anaerotruncus	colihominis
GCF_900199635.1	Anaerotruncus	massiliensis
GCF_902363605.1	Anaerotruncus	rubiinfantis
GCF_003149105.1	Faecalicatena	contorta
GCF_001028025.1	Faecalicatena	fissicatena
GCF_003149245.1	Faecalicatena	orotica
GCF_004345805.1	Flavonifractor	plautii
GCF_002161215.1	Flavonifractor	sp. An10
GCF_000179635.2	Ruminococcus	Albus
GCF_002834225.1	Ruminococcus	bromii
GCF_000468015.1	Ruminococcus	callidus
GCF_001312825.1	Ruminococcus	champanellensis
GCF_000518765.1	Ruminococcus	flavefaciens
GCF_000425525.1	Ruminococcus	gauvreauii

GCF_000154485.1	Erysipelatoclostridium	ramosum
GCF_000686665.1	Erysipelatoclostridium	saccharogumia
GCF_900102365.1	Erysipelatoclostridium	cocleatum
GCF_000313565.1	Dielma	fastidiosa
GCF_000154805.1	Erysipelatoclostridium	spiroforme
GCF_012317185.1	Erysipelatoclostridium	innocuum
GCF_000157995.1	Holdemania	filiformis
GCF_902374945.1	Holdemania	massiliensis
GCF_000156655.1	Holdemania	biformis
GCF_000384195.1	Allobaculum	stercoricanis
GCF_000425005.1	Solobacterium	moorei
GCF_000177375.1	Bulleidia	extructa
GCF_001545095.1	Erysipelothrix	larvae
GCF_000160815.2	Erysipelothrix	rhusiopathiae
GCF_001544355.1	Alicyclobacillus	acidiphilus
GCF_014647315.1	Alicyclobacillus	cellulosilyticus
GCF_004366795.1	Alicyclobacillus	sacchari
GCF_000508865.1	C17	Strain 1
GCF_000508885.1	C3	Strain 3
GCF_000508925.1	C6	Strain 6
GCF_000508945.1	C7	Strain 7
GCF_000508965.1	C8	Strain 8
GCF_000508985.1	C9	Strain 9
GCF_000509005.1	C13	Strain 13
GCF_000509025.1	C14	Strain 14
GCF_000509045.1	C15	Strain 15
GCF_000509065.1	C16	Strain 16
GCF_000509085.1	C18	Strain 18
GCF_000509105.1	C21	Strain 21
GCF_000509125.1	C26	Strain 26
GCF_000509145.1	C27	Strain 27
GCF_000509165.1	C28	Strain 28
GCF_000509185.1	C29	Strain 29

Supplementary Table 3.4

Supplementary Table 3.4: Reference list of genomes used to perform phylogenetic analysis of the CE17 (reference to Fig 3.3).

References

1. Byndloss MX, Baumler AJ. The germ-organ theory of non-communicable diseases. Nat Rev Microbiol. 2018;16(2):103-10; doi: 10.1038/nrmicro.2017.158.
2. Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. ISME J. 2013;7(4):830-8; doi: 10.1038/ismej.2012.160.
3. Doolittle WF. Phylogenetic classification and the universal tree. Science. 1999;284(5423):2124-9.
4. Snel B, Bork P, Huynen MA. Genomes in flux: the evolution of archaeal and proteobacterial gene content. Genome Res. 2002;12(1):17-25; doi: 10.1101/gr.176501.
5. Andreu VP, Augustijn HE, Chen L, Zhernakova A, Fu J, Fischbach MA, et al. A systematic analysis of metabolic pathways in the human gut microbiota. bioRxiv. 2021:2021.02.25.432841; doi: 10.1101/2021.02.25.432841.
6. Rivera-Chavez F, Baumler AJ. The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. Annu Rev Microbiol. 2015;69:31-48; doi: 10.1146/annurev-micro-091014-104108.
7. Rivera-Chavez F, Lopez CA, Baumler AJ. Oxygen as a driver of gut dysbiosis. Free Radic Biol Med. 2017;105:93-101; doi: 10.1016/j.freeradbiomed.2016.09.022.
8. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and

Enhances Pathogen Susceptibility. Cell. 2016;167(5):1339-53 e21; doi: 10.1016/j.cell.2016.10.043.

9. Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, Klopp C, et al. Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. Genome Res. 2010;20(11):1605-12; doi: 10.1101/gr.108332.110.

10. Jacobson A, Lam L, Rajendram M, Tamburini F, Honeycutt J, Pham T, et al. A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. Cell Host Microbe. 2018;24(2):296-307 e7; doi: 10.1016/j.chom.2018.07.002.

11. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex carbohydrates in the gut. Gut Microbes. 2012;3(4):289-306; doi: 10.4161/gmic.19897.

12. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science. 2011;334(6052):105-8; doi: 10.1126/science.1208344.

13. Pryde SE, Duncan SH, Hold GL, Stewart CS, Flint HJ. The microbiology of butyrate formation in the human colon. FEMS Microbiol Lett. 2002;217(2):133-9; doi: 10.1111/j.1574-6968.2002.tb11467.x.

14. Vital M, Howe AC, Tiedje JM. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. MBio. 2014;5(2):e00889; doi: 10.1128/mBio.00889-14.

15. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK, et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell Metab. 2011;13(5):517-26; doi: 10.1016/j.cmet.2011.02.018.

16. Velazquez OC, Lederer HM, Rombeau JL. Butyrate and the colonocyte. Production, absorption, metabolism, and therapeutic implications. Adv Exp Med Biol. 1997;427:123-34.

17. Zheng L, Kelly CJ, Colgan SP. Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A Review in the Theme: Cellular Responses to Hypoxia. Am J Physiol Cell Physiol. 2015;309(6):C350-60; doi: 10.1152/ajpcell.00191.2015.

18. Byndloss MX, Olsan EE, Rivera-Chávez F, Tiffany CR, Cevallos SA, Lokken KL, et al. Microbiota-activated PPAR-g signaling inhibits dysbiotic Enterobacteriaceae expansion. Science. 2017;357:570-5.

19. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, et al. Diversity of the human intestinal microbial flora. Science. 2005;308(5728):1635-8; doi: 10.1126/science.1110591.

20. Ludwig W. SKH, Whitman W. B. Revised road map to the phylum Firmicutes In: De Vos P. GGM, Jones D., Krieg N. R., Ludwig W., Rainey F. A., Schleifer K.-H., Whitman W. B. , editor. Bergey's Manual of Systematic Bacteriology. New York: Springer; 2009. p. 1-13.

21. Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, et al. Induction of colonic regulatory T cells by indigenous Clostridium species. Science. 2011;331(6015):337-41; doi: 10.1126/science.1198469.

22. Antunes LC, Han J, Ferreira RB, Lolic P, Borchers CH, Finlay BB. Effect of antibiotic treatment on the intestinal metabolome. Antimicrob Agents Chemother. 2011;55(4):1494-503; doi: 10.1128/AAC.01664-10.

23. Rivera-Chavez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, et al. Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella. Cell Host Microbe. 2016;19(4):443-54; doi: 10.1016/j.chom.2016.03.004.

24. Gillis CC, Hughes ER, Spiga L, Winter MG, Zhu W, Furtado de Carvalho T, et al. Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. Cell Host Microbe. 2018;23(1):54-64 e6; doi: 10.1016/j.chom.2017.11.006.

25. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, et al. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature. 2013;500(7461):232-6; doi: 10.1038/nature12331 nature12331 [pii].

26. Kim YG, Sakamoto K, Seo SU, Pickard JM, Gillilland MG, 3rd, Pudlo NA, et al. Neonatal acquisition of Clostridia species protects against colonization by bacterial pathogens. Science. 2017;356(6335):315-9; doi: 10.1126/science.aag2029.

27. Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature. 2013;504(7480):446-50; doi: 10.1038/nature12721.

28. Belloir C, Neiers F, Briand L. Sweeteners and sweetness enhancers. Curr Opin Clin Nutr Metab Care. 2017;20(4):279-85; doi: 10.1097/MCO.0000000000000377.

29. Gibson PR, Shepherd SJ. Personal view: food for thought--western lifestyle and susceptibility to Crohn's disease. The FODMAP hypothesis. Aliment Pharmacol Ther. 2005;21(12):1399-409; doi: 10.1111/j.1365-2036.2005.02506.x.

30. Beaugerie L, Flourie B, Marteau P, Pellier P, Franchisseur C, Rambaud JC. Digestion and absorption in the human intestine of three sugar alcohols. Gastroenterology. 1990;99(3):717-23.

31. Wang YM, van Eys J. Nutritional significance of fructose and sugar alcohols. Annu Rev Nutr. 1981;1:437-75; doi: 10.1146/annurev.nu.01.070181.002253.

32. Gibson PR, Shepherd SJ. Evidence-based dietary management of functional gastrointestinal symptoms: The FODMAP approach. J Gastroenterol Hepatol. 2010;25(2):252-8; doi: 10.1111/j.1440-1746.2009.06149.x.

33. Grabitske HA, Slavin JL. Low-digestible carbohydrates in practice. J Am Diet Assoc. 2008;108(10):1677-81; doi: 10.1016/j.jada.2008.07.010.

34. Ong DK, Mitchell SB, Barrett JS, Shepherd SJ, Irving PM, Biesiekierski JR, et al. Manipulation of dietary short chain carbohydrates alters the pattern of gas

production and genesis of symptoms in irritable bowel syndrome. J Gastroenterol Hepatol. 2010;25(8):1366-73; doi: 10.1111/j.1440-1746.2010.06370.x.

35. Halmos EP, Power VA, Shepherd SJ, Gibson PR, Muir JG. A diet low in FODMAPs reduces symptoms of irritable bowel syndrome. Gastroenterology. 2014;146(1):67-75 e5; doi: 10.1053/j.gastro.2013.09.046.

36. Bartlett JG, Moon N, Chang TW, Taylor N, Onderdonk AB. Role of Clostridium difficile in antibiotic-associated pseudomembranous colitis. Gastroenterology. 1978;75(5):778-82.

37. Theriot CM, Koenigsknecht MJ, Carlson PE, Jr., Hatton GE, Nelson AM, Li B, et al. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. Nat Commun. 2014;5:3114; doi: 10.1038/ncomms4114.

38. Hofmann JD, Biedendieck R, Michel AM, Schomburg D, Jahn D, Neumann-Schaal M. Influence of L-lactate and low glucose concentrations on the metabolism and the toxin formation of Clostridioides difficile. PLoS One. 2021;16(1):e0244988; doi: 10.1371/journal.pone.0244988.

39. Ghimire S, Roy C, Wongkuna S, Antony L, Maji A, Keena MC, et al. Identification of Clostridioides difficile-Inhibiting Gut Commensals Using Culturomics, Phenotyping, and Combinatorial Community Assembly. mSystems. 2020;5(1); doi: 10.1128/mSystems.00620-19.

40. Oliveira RA, Ng KM, Correia MB, Cabral V, Shi H, Sonnenburg JL, et al. Klebsiella michiganensis transmission enhances resistance to

Enterobacteriaceae gut invasion by nutrition competition. Nat Microbiol. 2020;5(4):630-41; doi: 10.1038/s41564-019-0658-4.

41. Narushima S, Sugiura Y, Oshima K, Atarashi K, Hattori M, Suematsu M, et al. Characterization of the 17 strains of regulatory T cell-inducing human-derived Clostridia. Gut Microbes. 2014;5(3):333-9; doi: 10.4161/gmic.28572.

42. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6; doi: 10.1038/nmeth.f.303.

43. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20; doi: 10.1093/bioinformatics/btu170.

44. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581-3; doi: 10.1038/nmeth.3869.

45. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013;8(4):e61217; doi: 10.1371/journal.pone.0061217.

46. Bodenhofer U, Bonatesta E, Horejs-Kainrath C, Hochreiter S. msa: an R package for multiple sequence alignment. Bioinformatics. 2015;31(24):3997-9; doi: 10.1093/bioinformatics/btv494.

47. Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics. 2011;27(4):592-3; doi: 10.1093/bioinformatics/btq706.

48. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12(6):R60; doi: 10.1186/gb-2011-12-6-r60.

49. Tiffany CR, Baumler AJ. omu, a Metabolomics Count Data Analysis Tool for Intuitive Figures and Convenient Metadata Collection. Microbiol Resour Announc. 2019;8(15); doi: 10.1128/MRA.00129-19.

50. Sievers F, Higgins DG. Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol. 2014;1079:105-16; doi: 10.1007/978-1-62703-646-7_6.

51. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312-3; doi: 10.1093/bioinformatics/btu033.

52. Yu G. Using ggtree to Visualize Data on Tree-Like Structures. Curr Protoc Bioinformatics. 2020;69(1):e96; doi: 10.1002/cpbi.96.

53. Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee DY, Lu Y, et al. Quality control for plant metabolomics: reporting MSI-compliant studies. Plant J. 2008;53(4):691-704; doi: 10.1111/j.1365-313X.2007.03387.x.

54. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30; doi: 10.1093/nar/28.1.27.

55. Lang JM, Darling AE, Eisen JA. Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. PLoS One. 2013;8(4):e62510; doi: 10.1371/journal.pone.0062510.

Chapter 4

Epithelial hypoxia maintains colonization resistance against *Candida albicans*

Hannah P. Savage[1, *], Derek J. Bays[3, *], Connor Tiffany[1, *], Mariela A. F. Gonzalez[1], Eli. J.Bejarano[1], Henry Nguyen[1], Hugo L. P. Masson[1], Thaynara P. Carvalho[1,4], Renato L.Santos[1,4], Krystle L. Reagan[5], George R. Thompson[3], and Andreas J.Bäumler[1]

[1] Department of Medical Microbiology and Immunology, School of Medicine, University of California Davis, Davis, CA 95616, USA

[2] Current address: Department of Pathology Microbiology and Immunology, School of Veterinary Medicine, University of California Davis, Davis, CA 95616, USA

[3] Department of Internal Medicine, Division of Infectious Diseases, School of Medicine, University of California Davis, Sacramento, CA 95817, USA.

[4] Departamento de Clinica e Cirurgia Veterina´ria, Escola de Veterina´ria da Universidade Federal de Minas Gerais, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, Belo Horizonte, MG, Brazil

[5] Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California-Davis, Davis, CA 95615, USA.

* Contributed equally

In revision at Cell Host & Microbe

ABSTRACT

Antibiotic treatment sets the stage for intestinal domination by *Candida albicans*, but the resources driving this bloom remain poorly defined. We show that oxygen is a critical resource required for post-antibiotic *C. albicans* expansion. The metabolic footprint produced by *C. albicans* in gnotobiotic mice suggested that the yeast consumes simple sugars *in vivo*, but *in vitro* growth on the respective metabolites required the presence of oxygen. The ability of probiotic *Escherichia coli* to prevent a post-antibiotic bloom of *C. albicans* required oxygen respiration. *Clostridia* species limit oxygen availability in the large intestine by producing butyrate, which activates PPAR-γ signaling to maintain epithelial hypoxia. Streptomycin treatment depleted *Clostridia*-derived butyrate to increase epithelial oxygenation, but the PPAR-γ agonist 5-aminosalicylic acid (5-ASA) functionally replaced *Clostridia* species to restore epithelial hypoxia and colonization resistance against *C. albicans.* We conclude that 5-ASA treatment restores colonization resistance against *C. albicans* by limiting access to oxygen.

INTRODUCTION

The human colon is host to a large microbial community, the gut microbiota, which is dominated by bacteria belonging to the classes *Clostridia* and *Bacteroidia* (Human Microbiome Project, 2012). Fungi are minority species within the gut microbiota, which includes the opportunistic pathogen *Candida albicans* in approximately 60% of individuals (Pappas et al., 2018). A non-specific defense mechanism, termed colonization resistance, limits growth of *C. albicans* in the large intestine during homeostasis, resulting in a low abundance of the opportunist in the feces. However, a disruption of the gut microbiota, particularly through administration of broad-spectrum antibiotics with anaerobic activity, ablates colonization resistance and leads to intestinal dominance by *C. albicans* (AlonsoMonge et al., 2021; Pappas et al., 2018). A bloom of *C. albicans* in the gastrointestinal tract during antibiotic therapy is the most common etiology of invasive disease (candidemia) in patients with hematologic malignancies (Zhai et al., 2020). Candidemia is a leading cause of nosocomial infections in the United States (Magill et al., 2018; Wilson et al., 2002) and carries a high mortality rate (up to 49%) (Gudlaugsson et al., 2003; Morgan et al., 2005; Viscoli et al., 1999; Wilson et al., 2002). In cases of invasive candidiasis originating from the gut (Alonso-Monge et al., 2021; Zhai et al., 2020), preventing intestinal domination of intestinal *C. albicans* during receipt of antibiotics could be a novel approach to reducing invasive disease in patients with hematologic malignancy. The development of this prophylactic strategy requires an improved understanding of the factors conferring colonization resistance against *C. albicans*.

The importance of an intact microbiota in preventing an expansion of *C. albicans* in the gastrointestinal tract can be modeled in mice (Clark, 1971; Kennedy and Volz, 1983; Mason et al., 2012; Samonis et al., 1994). Antibiotic treatment depletes microbiotaderived short-chain fatty acids and increases levels of carbohydrates and sugar alcohols, which correlates with increased gastrointestinal colonization of *C. albicans* in mice (Guinan et al., 2019; Gutierrez et al., 2020). Carbohydrates and sugar alcohols promote growth of *C. albicans in vitro* (Gutierrez et al., 2020) but are depleted from intestinal contents by *Clostridia* species in antibiotic naïve mice (Tiffany et al., 2021). *Clostridia* species are also the main producers of the short-chain fatty acid butyrate (Louis and Flint,

2009; Vital et al., 2014), which inhibits growth of *C. albicans in vitro* (Nguyen et al., 2011). An antibiotic-induced depletion of *Clostridia* species correlates with decreased colonization resistance against *C. albicans* in rodent models and has been shown to increase the risk for fungal infections in preterm infants (Huang et al., 2022). This correlative evidence supports the premis that an antibiotic-mediated depletion of *Clostridia* species paves the way for intestinal domination by *C. albicans.* However, probiotic *Clostridia* species that are susceptible to antibiotics cannot restore colonization resistance against *C. albicans* during antibiotic therapy, which is a necessary treatment for many patients with hematologic malignancies (Moreno-Sanchez and Gomez-Gomez, 2022). Since patients often develop invasive candidiasis while they are receiving antibiotic prophylaxis, alternative approaches for restoring colonization resistance are needed to prevent intestinal domination by *C. albicans* in these individuals.

Here we used a mouse model to study which resources drive intestinal domination of *C. albicans* in antibiotic-treated mice. We then investigated whether limiting access to these resources *in vivo* would curb growth of *C. albicans* and restore colonization resistance following antibiotic treatment.

RESULTS

Metabolic footprinting reveals that *C. albicans* catabolizes carbohydrates and sugar alcohols in the gut

To identify resources *C. albicans* uses to support its growth in the gastrointestinal tract of mice, we developed a generalizable framework, termed metabolic footprinting (**Fig. 4.1A**), which consists of the following steps. First, germ-free mice are inoculated under sterile conditions with the microbe of interest. An un-targeted metabolomics screen is conducted on cecal contents to identify metabolites that are decreased in the cecal metabolome of mice that received the microbe relative to mice that remained germ-free, as these metabolites are potential carbon sources utilized by the microbe of interest. *In vitro* growth assays on minimal media are then performed to confirm consumption of a decreased metabolite when it is added as a carbon source.

To perform metabolomic footprinting for *C. albicans* (ATCC28367), germ-free (Swiss Webster) mice were mock-inoculated or inoculated with the opportunistic pathogen and cecal contents collected three days later. The soluble fraction of the cecal contents was then analyzed by un-targeted gas chromatography time-of-flight mass spectrometry (metabolite profiling). For the downstream analysis of un-targeted metabolomics data, we developed a reactive web application called omuShiny (**Fig. 4.1B**).

This reactive web application is an extension of the existing R package omu (Tiffany and Baumler, 2019) with a graphical user interface, which leverages Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) metadata to improve data analysis, visualization, and interpretation. Principal component analysis of the samples showed a distinct clustering of mice by treatment group, indicating that colonization with *C. albicans* changes the composition of the murine cecal metabolome (**Fig. 4.1C**). The metabolic footprint of *C. albicans* revealed that many of the known metabolites that exhibited a reduced concentration during colonization with the opportunistic pathogen were simple carbohydrates, including fructo-oligosaccharides (e.g., 1-kestose), disaccharides (e.g., β-gentiobiose), sugar acids (eg., mucic acid and gluconolactone), and alcoholic sugars (e.g., sorbitol, ribitol and galactinol) (**Fig 4.1D**).

To validate results from metabolic footprinting, we determined whether sorbitol, 1kestose, or β-gentiobiose, which were each reduced in the ceca of mice colonized with *C. albicans* (**Fig. 4.1E**), served as carbon sources to support *in vitro* growth of the opportunistic pathogen. Sorbitol, 1-kestose, or β-gentiobiose promoted aerobic growth of *C. albicans* in yeast nitrogen broth supplemented with cysteine (**Fig. 4.1F**), which supported the idea that a reduced abundance of these metabolites in gnotobiotic mice engrafted with *C. albicans* (**Fig. 4.1E**) was due to consumption by the yeast. The finding that *C. albicans* depleted carbohydrates and alcoholic sugars in the murine cecum (**Fig 4.1D**) was of interest because previous studies showed that the abundance of these carbon sources increases during antibiotic treatment (Gutierrez et al., 2020; Tiffany et al., 2021). Aerobic *in vitro* growth of *C. albicans* on these resources thus suggests that

antibiotic-induced metabolome alterations favor growth of the opportunistic pathogen (Gutierrez et al., 2020). However, aerobic *in vitro* batch culture is not considered a good mimic of growth conditions encountered in the healthy cecum, where oxygen levels are low. Notably, when we repeated the experiment under anaerobic growth conditions, neither sorbitol, 1kestose, nor β-gentiobiose promoted growth of *C. albicans* (**Fig. 4.1F**). This result was remarkable, because it suggested that to deplete sorbitol, 1-kestose, or β-gentiobiose from cecal contents (**Fig. 4.1D and E**), *C. albicans* required an additional resource: oxygen. Importantly, oxygen levels in the large intestine become elevated during antibiotic treatment (Byndloss et al., 2017b). We thus hypothesized that an antibiotic-induced increase in luminal oxygen levels enables *C. albicans* to take advantage of antibioticinduced metabolome alterations in the large intestine to promote an intestinal expansion.

Escherichia coli **limits growth of** ***C. albicans*** **in the antibiotic-treated gut by competing for oxygen**

To test our hypothesis, we first established a mouse model in which antibioticmediated gut microbiota disruption increases susceptibility to *Candida albicans* colonization. Oral streptomycin was chosen because it is not systemically absorbed, which limits effects on the host, but induces changes in the microbiota composition that include a depletion of *Clostridia* species (Byndloss et al., 2017b; Sekirov et al., 2008),

and an increase in the levels of carbohydrates and sugar alcohols in the cecal metabolome (Tiffany et al., 2021). Specific pathogen-free C57BL/6J mice pre-treated with a single dose of streptomycin became colonized after challenge with 10^4 colony-forming units (CFU) of *C. albicans*, whereas antibiotic-naïve mice had to be challenge with a 10,000-fold higher dose to achieve colonization with the opportunistic pathogen (**Fig. 4.2A**). Similarly, in mice pre-colonized with *C. albicans*, treatment with a single dose of streptomycin triggered a fecal bloom of the opportunistic pathogen (**Fig. 4.2B**). These data suggest that treatment of mice with streptomycin disrupts colonization resistance against *C. albicans*.

Our hypothesis predicted that inoculation with a probiotic that consumes oxygen would restore colonization resistance in streptomycin-treated mice. Mice were mocktreated or pre-treated with a single dose of streptomycin. The next day, mice were mockinoculated or inoculated with 10^9 CFU of the probiotic *Escherichia coli* Nissle 1917 (family *Enterobacteriaceae*) (Nissle, 1925), a facultative anaerobic bacterium that respires available oxygen in the murine large intestine (Byndloss et al., 2017b). Mice were challenged the next day with 10^5 CFU *C. albicans* and colonization levels in the cecum were determined two days later. Streptomycin pre-treatment increased recovery of *C. albicans* by 3-4 orders of magnitude compared to antibiotic-naïve mice. Inoculation of streptomycin-pretreated mice with *E. coli* Nissle 1917 markedly lowered the numbers of *C. albicans* that were recovered from feces and cecal contents (**Fig. 4.2C and 4.2D**), suggesting the probiotic improved colonization resistance against the opportunistic pathogen. To causatively link colonization resistance to oxygen respiration by the

probiotic, streptomycin pre-treated mice were inoculated with an aerobic respiration deficient *E. coli* Nissle 1917 derivative (*cydA* mutant), which lacks cytochrome *bd* oxidase, an enzyme required for respiring oxygen under microaerophilic conditions (Cevallos et al., 2021a). Notably, *C. albicans* was recovered in markedly higher numbers from streptomycin pre-treated mice inoculated with an aerobic respiration-deficient *E. coli* Nissle 1917 strain (*cydA* mutant) compared to mice inoculated with the aerobic respiration-proficient wild-type *E. coli* Nissle 1917 (wild-type) (**Fig. 4.2C and 4.2D**), although both wild-type and *cydA* mutant were recovered in similar numbers from fecal contents at the time of *C. albicans* challenge (**Fig. 4.2D**). These data supported the idea that probiotic *E. coli* prevents a bloom of *C. albicans* after streptomycin treatment by competing for oxygen.

The drug 5-aminosalicylic acid restores epithelial hypoxia and improves colonization resistance after streptomycin

With evidence pointing to oxygen as critical resource required for the bloom of *C. albicans* in the antibiotic-treated gut, we next wanted to test whether limiting oxygen availability would restore colonization resistance. Short-chain fatty acids reduce *C. albicans* growth (Guinan et al., 2019; Nguyen et al., 2011) and hyphal formation *in vitro* (Noverr and Huffnagle, 2004), which is thought to help control the opportunist during homeostasis. Disruption of the microbiota by streptomycin treatment depletes short-

chain fatty acids (Meynell, 1963), thereby relieving growth inhibition. However, the *Clostridia* derived short-chain fatty acid butyrate (Louis and Flint, 2009; Vital et al., 2014) also stimulates epithelial PPAR-γ (peroxisome proliferator-activated receptor gamma) signaling to maintain the epithelial surface in a state of physiological hypoxia (Byndloss et al., 2017b; Furuta et al., 2001; Kelly et al., 2015). We thus reasoned that an increase in the luminal oxygen bioavailability might be a second mechanism by which depletion of short-chain fatty acids stimulates growth of *C. albicans* in the large intestine. The presence of short-chain fatty acids is sufficient to limit *in vitro* growth of *C. albicans* (Guinan et al., 2019; Nguyen et al., 2011). Similarly, our results suggest that oxygen limitation was sufficient to limit *C. albicans* growth (**Fig. 4.1F**). We have shown previously that the PPAR-γ agonist 5-aminosalicylic acid (5-ASA) restores an anaerobic environment in the colon of mice with chemically induced colitis by activating PPAR-γ signaling specifically in the intestinal epithelium (Cevallos et al., 2021a). We thus hypothesized that reducing oxygen bioavailability in the antibiotic treated gut by treatment with 5-ASA might be sufficient to restore colonization resistance against *C. albicans.*

To test our hypothesis, we first determined whether 5-ASA treatment would improve colonization resistance against *C. albicans* when mice were colonized with the opportunistic pathogen prior to streptomycin exposure. C57BL/6J mice were precolonized with *C. albicans* and subsequently treated with a single dose of streptomycin to disrupt colonization resistance. Notably, streptomycin treatment triggered a fecal *C. albicans* bloom, but this bloom was blunted by 5-ASA

supplementation (**Fig. 4.3A**). Next, we determined whether 5-ASA would improve colonization resistance when streptomycin pre-treated mice were challenged with *C. albicans.* C57BL/6J mice received normal chow or chow supplemented with 5-ASA and were mock-treated or pre-treated with a single dose of streptomycin to decrease colonization resistance. Mice were challenged the next day with *C. albicans*. Remarkably, 5-ASA supplementation restored colonization resistance in streptomycin pre-treated mice, as indicated by reduced recovery of *C. albicans* from cecal contents at one (**Fig. 4.3B**) or seven (**Fig. 4.3C**) days after challenge. Collectively, these data suggest that 5-ASA improves colonization resistance against *C. albicans* after streptomycin treatment.

5-ASA has anti-inflammatory properties (Rousseaux et al., 2005), and inflammation can alter the gut microbial composition by inducing a bloom of facultatively anaerobic bacteria (Winter et al., 2013). To investigate whether reduced colonization by *C. albicans* was attributable to reduced inflammation in mice treated with 5-ASA, we investigated whether streptomycin pre-treatment or *C. albicans* challenge triggered inflammatory responses in the gut by measuring transcript levels of genes encoding lipocalin-2, TNF- α, or IL-17A one day after challenge with *C. albicans* (**Fig. 4.3D**). Although small increases in transcript levels of *Lcn2* and *Il17a* were observed in streptomycin pretreated mice after challenge with *C. albicans*, these increases were not reversed by 5ASA supplementation and thus did not correlate with the reduced *C. albicans* colonization levels observed in these animals (**Fig. 4.3B**). These data did not

support the idea that 5ASA improved colonization resistance against *C. albicans* by reducing intestinal inflammation.

To test whether 5-ASA changes the intestinal environment by restoring epithelial hypoxia, we visualized epithelial oxygenation using pimonidazole, a 2-nitroimidazole that is reductively activated to specifically bind to hypoxic (< 1% oxygen) cells (Kizaka-Kondoh and Konse-Nagasawa, 2009). Detection of pimonidazole binding using a monoclonal antibody revealed that in antibiotic-naïve mice, the epithelial surface in the cecum was hypoxic. Streptomycin pre-treatment reduced epithelial hypoxia, but 5-ASA supplementation restored epithelial hypoxia in streptomycin pre-treated C57BL/6J mice (**Fig. 4.4A-C**). 5-ASA supplementation did not restore butyrate concentrations in the feces of streptomycin pre-treated mice, ruling out butyrate as a trigger for restoring hypoxia (**Fig. 4.4D**).

Collectively, these data suggested that 5-ASA changes the intestinal environment by restoring epithelial hypoxia after antibiotic treatment, thereby limiting the growth of *C. albicans* in the lumen of the large intestine.

5-aminosalicylic acid restores colonization resistance by functionally replacing *Clostridia* species

A single dose of streptomycin disrupts the microbiota, but acquisition of microbes from the environment leads to a gradual restoration of colonization resistance over time as the microbiota recovers (Gillis et al., 2018; Rivera-Chavez et al., 2016). In contrast, when germ-free mice are engrafted with microbiota from antibiotic-treated mice, the sterile housing conditions prevent acquisition of environmental bacteria, which hinders microbiota recovery. We thus reasoned that a gnotobiotic model would enable us to study whether 5-ASA limits growth of *C. albicans* in mice when the post-antibiotic treated microbiota does not recover over time. C57BL/6J donor mice were mock-treated or received a single dose of streptomycin and cecal contents were collected 48 hours later. These cecal contents were then used to inoculate germ-free Swiss Webster recipient mice by oral gavage (cecal microbiota transfer). Two weeks later, gnotobiotic recipient mice were challenged with *C. albicans* and cecal contents were collected one week later (**Fig. 4.5A**). Gnotobiotic recipient mice engrafted with microbiota from mock-treated donor mice carried a significantly lower *C. albicans* burden in their ceca than gnotobiotic recipient mice engrafted with microbiota from streptomycin-treated donor mice (**Fig 4.5B**), indicating that loss of colonization resistance was not attributable to direct effects of streptomycin on *C. albicans* or on host cells. To mimic microbiota recovery, gnotobiotic mice were engrafted with microbiota from streptomycin-treated donor mice and were inoculated one week later with microbiota from a mock-treated donor mouse. The following week, mice were challenged with *C. albicans*, and the opportunistic pathogen was enumerated in cecal contents collected one week after challenge (**Fig. 4.5A**), which revealed that colonization resistance had been restored

(**Fig 4.5B**). Blinded scoring of histological sections from the cecum revealed that *C. albicans* challenge did not trigger overt inflammation in gnotobiotic mice engrafted with microbiota from streptomycintreated or mock-treated donor mice (**Fig 4.5C and D**).

To determine whether loss of epithelial hypoxia induced by streptomycin treatment in donor mice (**Fig 4.4A-C**) was recapitulated when cecal microbiota was transferred into gnotobiotic recipient mice, we visualized epithelial oxygenation with pimonidazole. Pimonidazole staining revealed that epithelial hypoxia was reduced in gnotobiotic recipient mice engrafted with microbiota from streptomycin-treated donor mice compared to gnotobiotic recipient mice engrafted with microbiota from mock-treated donor mice (**Fig. 4.5E-G**). Consistent with the idea that loss of epithelial hypoxia during streptomycin treatment is due to a depletion of *Clostridia* species (Byndloss et al., 2017b; RiveraChavez et al., 2016), epithelial hypoxia was restored when gnotobiotic recipient mice engrafted with microbiota from streptomycin-treated donor mice were inoculated with a community of 17 human *Clostridia* isolates (Atarashi et al., 2013)(**Fig. 4.5E-G**). Inoculation with a consortium of 17 human *Clostridia* isolates also restored colonization resistance against *C. albicans* (**Fig. 4.5B**), which was consistent with previous work implicating *Clostridia* species in colonization resistance against the yeast (Huang et al., 2022). Notably, when gnotobiotic recipient mice engrafted with microbiota from streptomycin- treated donor mice received chow containing 5-ASA (**Fig. 4.5A**), both epithelial hypoxia (**Fig. 4.5E-G**) and colonization resistance against *C. albicans* (**Fig. 4.5B**) were restored. These data demonstrate that 5-ASA can functionally replace *Clostridia* species to promote epithelial hypoxia and limit intestinal growth of *C. albicans.*

DISCUSSION

Fecal carriage of *C. albicans* is detected in approximately 60% of individuals, but this opportunistic pathogen is typically a minority species within the gut microbiota (Pappas et al., 2018). Factors that limit growth of *C. albicans* in the large intestine include inhibitory concentrations of short-chain fatty acids and a limited availability of critical resources that promote its growth, such as simple sugars (Guinan et al., 2019; Gutierrez et al., 2020). A disruption of the gut microbiota by antibiotic treatment results in a depletion of microbiota-derived short-chain fatty acids and increases the availability of simple sugars in the lumen of the large intestine (Guinan et al., 2019; Gutierrez et al., 2020; Meynell, 1963; Tiffany et al., 2021). It is thought that these changes in the intestinal environment drive intestinal domination by *C. albicans*, which is a common cause of candidemia in patients with hematologic malignancies who receive antibiotic prophylaxis (Zhai et al., 2020). By determining the metabolic footprint of *C. albicans* in gnotobiotic mice, our results provide support for the idea that the opportunistic pathogen supports its growth in the cecum by catabolizing simple sugars, such as sorbitiol, β-gentiobiose and

1-kestose. Importantly, oxygen was required for the consumption of these carbon sources by *C. albicans*. The identification of oxygen as a critical resource that limits intestinal growth of *C. albicans* has important ramifications for the development of strategies to prevent intestinal domination by the opportunistic pathogen.

During homeostasis, the host limits diffusion of oxygen into the intestinal lumen by maintaining the colonic epithelium in a state of physiological hypoxia (Lee et al., 2022; Litvak et al., 2018). An antibiotic-mediated depletion of short-chain fatty acids increases epithelial oxygenation (Kelly et al., 2015), which disrupts anaerobiosis in the lumen of the large intestine (Byndloss et al., 2017b). Here we show that an antibiotic-mediated disruption of the microbiota can be mitigated by inoculation with *Clostridia* species, which restores epithelial hypoxia and reconstitutes colonization resistance against *C. albicans. Clostridia* species belonging to the families *Lachnospiraceae* and *Ruminococcaceae* are the main producers of the short-chain fatty acid butyrate (Louis and Flint, 2009; Vital et al., 2014), which activates epithelial PPAR-γ signaling to maintain epithelial hypoxia (Byndloss et al., 2017b). However, using a probiotic therapy to restore colonization resistance raises health concerns in immunocompromised patients, such as those with hematologic malignancies, who are at risk for developing invasive candidiasis (Redman et al., 2014). Furthermore, patients with hematologic malignancies can develop candidiasis while receiving antibiotic prophylaxis, which would likely clear any antibiotic susceptible probiotic bacteria.

Inhibitory concentrations of short-chain fatty acids, such as butyrate, can limit growth of *C. albicans* (Nguyen et al., 2011). Our results show that oxygen limitation can prevent growth of *C. albicans* on simple sugars, such as sorbitol. These data suggest that a bloom of *C. albicans* in the antibiotic-treated gut could be prevented by either restoring inhibitory concentrations of short-chain fatty acids or by limiting access to oxygen. Here we demonstrate for the first time that a post-antibiotic *C. albicans*

expansion can be restricted by restoring epithelial hypoxia through treatment with the PPAR-γ agonist 5ASA. 5-ASA functionally replaced *Clostridia* species to restore epithelial hypoxia and reinstate colonization resistance against *C. albicans* by limiting oxygen availability. To describe this novel approach, we propose the term “faux-biotics” for chemical microbiota substitutes (e.g., 5-ASA) that replace probiotics (e.g., *Clostridia* species) functionally. Whereas antibiotic prophylaxis interferes with engraftment of probiotics or fecal microbiota transplants, treatment with a faux-biotic offers the benefit that it is not inhibited or eliminated by antibiotics. Furthermore, by restoring epithelial hypoxia in the antibiotic treated gut using 5-ASA, growth of *C. albicans* is restricted by the same mechanism the host uses to maintain a low abundance of the yeast during homeostasis. Since low-level fecal carriage of *C. albicans* is a constellation that has likely existed for a long time, it is unlikely that the opportunistic pathogen can develop resistance against growth restriction imposed by epithelial hypoxia. Thus, reinstating colonization resistance by restoring host functions, such as epithelial hypoxia, could be a therapeutic strategy to end the antimicrobial arms race with opportunistic fungal pathogens.

STAR METHODS

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-pimonidazole monoclonal antibody MAb1	Hypoxyprobe	Hypoxyprobe™-1 Kit, HP1-1000
Cyanine3-labeled goat anti-mouse IgG	Jackson ImmunoResearch	Cat# 115-165-003
Bacterial and virus strains		
Escherichia coli Nissle 1917	DSMZ	DSM 6601
Escherichia coli Nissle 1917 ΔcydA	(Cevallos et al., 2019)	N/A
Erysipelatoclostridium saccharogumia	(Atarashi et al., 2013)	C1
Flavonifractor plautii	(Atarashi et al., 2013)	C3
Hungatella hathewayi	(Atarashi et al., 2013)	C4
Blautia producta	(Atarashi et al., 2013)	C6

Enterocloster bolteae	(Atarashi et al., 2013)	C7
Dielma fastidiosa	(Atarashi et al., 2013)	C8
Anaerostipes caccae	(Atarashi et al., 2013)	C9
Anaerotruncus colihominis	(Atarashi et al., 2013)	C13
Unclassified *Lachnospiraceae*	(Atarashi et al., 2013)	C14
Enterocloster asparagiformis	(Atarashi et al., 2013)	C15
Lachnoclostridium symbiosum	(Atarashi et al., 2013)	C16
Erysipelatoclostridium ramosum	(Atarashi et al., 2013)	C18
Faecalicatena fissicatena	(Atarashi et al., 2013)	C21
Lachnoclostridium scindens	(Atarashi et al., 2013)	C26
Eisenbergiella massiliensis	(Atarashi et al., 2013)	C27

Enterocloster spp.	(Atarashi et al., 2013)	C28
Eisenbergiella tayi	(Atarashi et al., 2013)	C29
Chemicals, peptides, and recombinant proteins		
Streptomycin	Sigma	Cat#59137
Chloramphenicol	ThermoFisher	Cat#BP904
Sorbitol	Sigma	Cat#53889
1-Kestose	Sigma	Cat#72555
β-Gentiobiose	Sigma	Cat#G3000
N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide	Sigma	Cat#00942
Critical commercial assays		
DNAeasy Powersoil Kit	QIAGEN	Cat#:12888-100
TRI Reagent	Molecular Research Center	Cat#TR118
PureLink DNAse	Invitrogen	Cat#18068015
MultiScribe Reverse Transcriptase	ThermoFisher	Cat#4311235
GeneAmp™ 10X PCR Gold Buffer & $MgCl_2$	ThermoFisher	Cat#4306898
Applied Biosystems™ Random Hexamers (50 µM)	ThermoFisher	Cat#N8080127

Applied Biosystems™ RNase Inhibitor	ThermoFisher	Cat#N8080119
Applied Biosystems™ PowerUp™ SYBR™ Green Master Mix for qPCR	ThermoFisher	Cat#A25742
Applied Biosystems™ GeneAmp™ dNTP Blend (100 mM)	ThermoFisher	Cat# N8080261
Experimental models: Organisms/strains		
Candida albicans	ATCC	ATCC 28367
Mus musculus C57BL/6J	The Jackson Laboratory	Cat#000664
Mus musculus Germ-Free Swiss Webster	Bred in house; originally acquired from Taconic	N/A
Mus musculus Germ-Free C57BL/6J	Bred in house; originally acquired from UNC National Gnotobiotic Rodent Resource Center https://www.med.unc.edu/ngrrc/products-services/	N/A

Oligonucleotides		
See supplementary table 2.	See Table S2	See Table S2
Software and algorithms		
omu	Connor R. Tiffany	https://cran.r-project.org/web/packages/omu/index.html
omuShiny	Connor R Tiffany	https://clostridia-enjoyer.shinyapps.io/omuShiny/
KEGG	M Kanehisa , S Goto	https://www.genome.jp/kegg/
Prism	Graph Pad	https://www.graphpad.com/features
R	R project	https://www.r-project.org/
RStudio	Posit	https://posit.co/products/open-source/rstudio/

Tidyverse	Posit	https://www.tidyverse.org/
Shiny	Posit	https://shiny.posit.co/
gridExtra	Baptiste Auguie, Anton Antonov	https://cran.r-project.org/web/packages/gridExtra/index.html
officer	David Gohel et al.	https://cran.r-project.org/web/packages/officer/index.html
openxlsx	Philipp Schauberger et al.	https://cran.r-project.org/web/packages/openxlsx/index.html
shinyFeedback	Andy Merlino, Patrick Howard	https://cran.rstudio.com/web/packages/shinyFeedback/index.html

DT	Yihui Xie et al.	https://cran.r-project.org/web/packages/DT/index.html
shinyWidgets	Victor Perrier et al.	https://cran.r-project.org/web/packages/shinyWidgets/index.html
magick	Jeroen Ooms	https://cran.r-project.org/web/packages/magick/index.html
spsComps	Le Zhang	https://cran.r-project.org/web/packages/spsComps/index.html
thematic	Posit.	https://cran.r-project.org/web/packages/thematic/index.html

ggpubr	Alboukadel Kassambara	https://cran.r-project.org/web/packages/ggpubr/index.html
car	Fox J, Weisberg S	https://cran.r-project.org/web/packages/car/index.html
e1071	David Meyer et al.	https://cran.r-project.org/web/packages/e1071/index.html
colourpicker	Dean Attali, David Griswold	https://cran.r-project.org/web/packages/colourpicker/index.html
ggrepel	Kamil Slowikowski et al.	https://cran.r-project.org/web/packages/ggrepel/index.html

Microsoft Excel	Microsoft	https://www.microsoft.com/en-us/microsoft-365
Other		
Mouse 10% Fat Diet	Teklad Diet	#TD110675
0.25% 5-ASA added to base teklad diet	Teklad Diet	N/A

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. Andreas Bäumler ajbaumler@ucdavis.edu

Materials availability

Reagents generated in this study are available, upon reasonable requests, with a completed material transfer agreement.

Data and Code Availability

Normalized metabolomics spectra data are available as supplementary material with this manuscript (**Supplementary File 4.1**). Raw Spectra data can be found at the NIH metabolomics (database www.metabolomicsworkbench.org) using the values in the

file id row of **Supplementary File 4.1**. Source code for omuShiny can be found at https://github.com/connor-reid-tiffany/omuShiny, while the web app can be found at https://clostridia-enjoyer.shinyapps.io/omuShiny/.

Experimental model and subject details

Mice

For experiments using specific pathogen free (SPF) mice, 7-8 week-old female C57BL/6J mice were obtained from The Jackson Laboratory. Mice were maintained under SPF conditions throughout the duration of the experiment. For gnotobiotic experiments, age-, sex-, and strain-matched male and female C57BL/6J and Swiss-Webster mice, ages 6-10 weeks old, were bred in-house and were housed in Tecniplast Isocages throughout the duration of the experiments. Mice were fed autoclaved water and irradiated Teklad 2018 (2918) mouse chow. For the colonization resistance experiment with *Escherichia coli* Nissle 1917, mice were fed a defined, irradiated Teklad chow (#TD110675).

To colonize specific pathogen free mice with the probiotic *Escherichia coli* Nissle 1917 strain or the isogenic mutant *Escherichia coli* 1917 strain deficient in oxygen respiration (*cydA* mutant), mice were given 20 mg of streptomycin one day prior to *E. coli* inoculation. The *E. coli* inoculums were produced as follows. Both *E. coli* strains were grown anaerobically in Luria-Bertani broth (LB) at 37°C for 24 hours. Cell concentration was then determined by OD600 measurements. The *E. coli* cultures were

then centrifuged to produce a pellet which was resuspended in sterile PBS to achieve a desired concentration of 10^9 CFU/100µL. Mice were then given 100µL of inoculum by oral gavage. To collect feces, mice were placed in 70% ethanol-cleaned beakers cleaned until they produced a fecal pellet, which was then collected in 1 mL PBS, weighed, vortexed, and plated at 10-fold dilutions on either MacConkey agar, or LB agar with either carbenicillin or kanamycin for the competitive index experiment. Cecal contents collected at necropsy were similarly placed in PBS and plated. Plates were incubated at 37°C for 18 hours prior to counting colony forming units.

To colonize the gnotobiotic mice with a cecal microbiota transfer (CMT) from either control or antibiotic-treated SPF mice, SPF mice were treated with 20 mg streptomycin vs. mock treatment and two days later cecal contents were collected, diluted at 1g/10 ml PBS, and stored in 25% glycerol at -80 C until use. Gnotobiotic mice were administered 200 µL of stored mock- or streptomycin-treated CMT by oral gavage. The CMT was given 1 week to colonize prior to further treatment. For the mice receiving first streptomycintreated CMT followed by mock-treated CMT, mice were gavaged after 1 week with 200 µL of stored mock-treated CMT. For the C17 treatment group, mice were colonized with a mix of 17 human Clostridia isolates (Atarashi et al., 2013) known to produce SCFAs. Strains were grown individually for 4 days in Eggerth-Gagnon broth, then equal volumes of broth from each strain were combined and 200 µL were administered by oral gavage to each mouse. For the 5-ASA treatment group, gnotobiotic mice were then placed on a diet containing 0.25% 5-ASA added to a base diet of Teklad

2018 mouse chow (Envigo). One week after the diet change or administration of C17, mice were infected with *C. albicans*.

At necropsy, mice were euthanized by overexposure to carbon dioxide followed by cervical dislocation. All procedures were approved by the UC Davis Animal Care and Use Committee.

Strains and culture conditions

Candida albicans ATCC 28367 was streaked from a glycerol stock onto agar with 1% yeast extract, 2% peptone, 2% dextrose (YPD) + chloramphenicol at 100 ug/mL four days prior to infection, and re-streaked at two days prior to infection. To prepare the inoculum, colonies from the agar were resuspended in PBS, concentration was determined by OD600, and samples were diluted to the desired final concentration. Inocula were vortexed immediately prior to infection. Gnotobiotic mice were administered 10^5 colony forming units (CFU) *C. albicans* in 100 µL phosphate buffered saline (PBS) by oral gavage. Specific pathogen free mice were given 10^5 CFU (Fig. 3A, 5C, 6A-C) or 10^6 CFU (Fig. 3B-C) *C. albicans* in 100 µL PBS by oral gavage, unless otherwise indicated. Inoculum CFUs were confirmed by plating on YPD agar + chloramphenicol to determine concentration.

One-2 days prior to or 1-5 days after infection with *C. albicans*, as indicated for each experiment, mice were randomized to receive 20 mg streptomycin in 100 µL sterile

water by oral gavage or no treatment. For mice colonized with *C. albicans* prior to streptomycin treatment, colonization was confirmed, via measuring CFU/g feces, to be comparable between groups at the time of antibiotic administration.

C. albicans burden after infection was determined in feces and cecal contents by enumerating CFU on YPD agar + chloramphenicol. Prior to infection, mice were not colonized with any fungal organisms that grow on YPD agar. To collect feces, mice were placed in 70% ethanol-cleaned beakers cleaned until they produced a fecal pellet, which was then collected in 1 mL PBS, weighed, vortexed, and plated at 10-fold dilutions on YPD agar + chloramphenicol. Cecal contents collected at necropsy were similarly placed in PBS and plated. Plates were counted approximately 24 hours later. A presumptive identification of *C. albicans* was made for white, glistening, circular, smooth colonies growing on YPD agar containing chloramphenicol. Any colony growing on YPD agar + chloramphenicol with an atypical appearance was confirmed to be *Candida* spp. by Gram

stain.

For *in vitro* growth on different carbon sources, *Candida albicans* ATCC 28367 was cultured on yeast peptone dextrose agar plates for 48 hours at 30°C. Culture from the agar plate was then resuspended in sterile PBS. OD600 measurements were taken to dilute the PBS solution to a final concentration of 10^8 cells/mL of PBS. Minimal media for the candida growth assays was created using yeast nitrogen base (YNB) medium without amino acids, supplemented with a 50μM concentration of cysteine. For the

different growth conditions, we added either no sugar, 0.5% sorbitol, 5% 1-kestose, or 5% beta-gentiobiose to the YNB media. 198 µL of media was added to 96 well maxisorb plates for each growth condition in triplicate, and 2 µL of the 10^8 cells/mL candida inoculum was added to each well for a final concentration of 10^6 cells/well. Plates were then inoculated in a victor nivo plate reader under either aerobic, or anaerobic conditions in a hypoxia chamber for 48 hours at 30°C, with OD600 measurements taken every 2 hours after 30 seconds of shaking.

Method details

Development of omuShiny

A shiny app extension of the existing R package omu designed for untargeted metabolomics data analysis, termed omuShiny, was developed using the following packages: Shiny, devtools, and roxygen2. These additional packages are dependencies for omuShiny: rstatix, shinyFeedback, shinyWidgets, omu, tidyverse, openxlsx, DT, officer, thematic , colourpicker, e1071, car, magick, spsComps, ggpubr, ggrepel, and gridExtra.

GC-TOF mass spectrometry

Cecal contents from gnotobiotic mice which were either kept germ-free or infected with 10^6 CFU of *Candida albicans* ATCC 28367 were collected 3 days following infection, flash frozen in liquid nitrogen, then stored at -80°C until further processing. Data were acquired using the following chromatographic parameters. Column: Restek corporation Rtx-5Sil MS (30 m length × 0.25 mm internal diameter with 0.25-μm film made of 95% dimethyl/5%diphenylpolysiloxane). Mobile phase: Helium. Column temperature: 50–330 °C. Flow rate: 1 mL min−1. Injection volume: 0.5 μL. Injection: 25 splitless time into a multi-baffled glass liner. Injection temperature: 50 °C ramped to 250 °C by 12 °C s−1. Oven temperature program: 50 °C for 1 min, then ramped at 20 °C min−1 to 330 °C, held constant for 5 min. The analytical GC column is protected by a 10-m long empty guard column which is cut by 20-cm intervals whenever the reference mixture QC samples indicate problems caused by column contaminations. We have validated that at this sequence of column cuts, no detrimental effects are detected with respect to peak shapes, absolute or relative metabolite retention times or reproducibility of quantifications. We use automatic liner exchanges after each set of 10 injections which we could show to reduce sample carryover for highly lipophilic compounds such as free fatty acids. Mass spectrometry parameters are used as follows: a Leco Pegasus IV mass spectrometer is used with unit mass resolution at 17 spectra s−1 from 80–500 Da at − 70 eV ionization energy and 1800 V detector voltage with a 230 °C transfer line and a 250 °C ion source.

Metabolomics data processing

Raw data files are preprocessed directly after data acquisition and stored as ChromaTOF-specific *.peg files, as generic *.txt result files and additionally as generic ANDI MS *.cdf files. ChromaTOF vs. 2.32 is used for data preprocessing without smoothing, 3 s peak width, baseline subtraction just above the noise level, and automatic mass spectral deconvolution and peak detection at signal/noise levels of 5:1 throughout the chromatogram. Apex masses are reported for use in the BinBase algorithm. Result *.txt files are exported to a data server with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database. The BinBase algorithm (rtx5) used the settings: validity of chromatogram (< 10 peaks with intensity > 107 counts s−1), unbiased retention index marker detection (MS similarity > 800, validity of intensity range for high m/z marker ions), and retention index calculation by 5th order polynomial regression. Spectra are cut to 5% base peak abundance and matched to database entries from most to least abundant spectra using the following matching filters: retention index window ± 2000 units (equivalent to about ± 2 s retention time), validation of unique ions and apex masses (unique ion must be included in apexing masses and present at > 3% of base peak abundance), mass

spectrum similarity must fit criteria dependent on peak purity and signal/noise ratios and a final isomer filter. Failed spectra are automatically entered as new database entries if s/n > 25, purity < 1.0, and presence in the biological study design class was > 80%. All thresholds reflect settings for ChromaTOF v. 2.32. Quantification is reported as peak height using the unique ion as default, unless a different quantification ion is manually set in the BinBase administration software BinView. A quantification report table is produced for all database entries that are positively detected in more than 10% of the samples of a study design class (as defined in the miniX database) for unidentified metabolites. A subsequent post-processing module is employed to automatically replace missing values from the *.cdf files. Replaced values are labeled as 'low confidence' by color coding, and for each metabolite, the number of high-confidence peak detections is recorded as well as the ratio of the average height of replaced values to high-confidence peak detections. These ratios and numbers are used for manual curation of automatic report data sets to data sets released for submission. These data were then normalized to the mTIC value (sum of the peak heights of the known metabolites).

Metabolomics data analysis

Following processing, metabolomics data were analyzed using our recently developed shiny application, omuShiny. Principal component analysis was performed on natural logarithm transformed metabolite abundances and ellipses denote a 95%

confidence interval on a multivariate t distribution. Univariate statistics were performed using a welch-corrected two-sided t-test on natural logarithm transformed metabolite abundances. P values were adjusted using the Benjamini-Hochberg method to correct for the false discovery rate. Volcano plots and dot plots were derived from data generated by the welch t test in omuShiny. Error bars on the dot plot were calculated using the standard error of the mean.

SCFA analysis

Two fecal pellets per mouse were collected in 200 µL PBS. Samples were vortexed to disrupt particulate matter and then centrifuged at 6,000 rpm for 10 min to pellet any remaining debris. For each sample, 100 µL of supernatant was combined with 10 µL of a solution containing deuterated acetate, propionate, and butyrate so that each deuterated metabolite was at a final concentration of 100 µM. Samples were dried without heat in a vacuum dryer and then stored at -80°C until use.

Dried extracts were then solubilized by sonication in 0.1 mL anhydrous pyridine and then incubated for 20 min at 80°C. An equal amount of N-tert-butyldimethylsilyl-Nmethyltrifluoroacetamide with 1% tert-butyldimethylchlorosilate (Sigma-Aldrich) was added, and the samples were incubated for 1 h at 80°C. Samples were centrifuged at 20,000 g for 1 min to remove leftover particles. One hundred microliters of the supernatant were transferred to an autosampler vial and analyzed by gas

chromatography-mass spectrometry (Agilent 8890 Gas Chromatograph and Agilent 7000D Mass spectrometer). 1 µL of the sample was injected with a 1:50 split ratio at an injection temperature of 250°C on an HP 5ms Ultra Inert (2x15-m-length, 0.25-mm diameter, 0.25 µm film thickness) fused silica capillary column. Helium was used as the carrier gas with a constant flow of 1.2 mL/min. The gas chromatograph (GC) oven temperature started at 50°C for 5 min, rising to 110°C at 10°C/min and holding for 2 min, then raised to 310°C at 40°C/min with a final hold for 4 min. The interface was heated to 300°C. The ion source was used in electron ionization (EI) mode (70 V, 150 µA, 200°C). The dwell time for selected ion monitoring (SIM) events was 50ms. Both acetate, propionate and butyrate were quantified using SIM, with the monitored m/z, and experimentally determined retention times detailed in the following table:

Analyte	Retention time (minutes)	Quantifier ion (m/z)	Qualifier 1 (m/z)	Qualifier 2 (m/z)
Acetate	7.630	117	118	75
Propionate	9.704	131	132	75
Butyrate	11.220	145	146	75
Acetate-d3	7.590	120	162	121
Propionate-d5	9.640	136	178	137
Butyrate-d7	7.590	152	153	194

Efficient recovery of target metabolites was determined using deuterated compounds as internal standards. Quantification was based on external standards comprised of a

series of dilutions of pure compounds, derivatized as described above at the same time as the samples.

Histopathology

At necropsy, cecal tips were collected in individual histology cassettes and immediately placed in 10% phosphate-buffered formalin. After approximately 24 hours, cassettes were moved to 70% ethanol until processing. Samples were embedded in paraffin, mounted on slides, and stained with hematoxylin and eosin (H&E) by the UC Davis VMTH Pathology Lab. Slides were scored in a blinded manner by a board-certified Veterinary Pathologist using standard scoring criteria (**Supplemental Table 4.1**) (Spiga et al., 2017).

Colonocyte isolation and Real Time-PCR

Ceca and colons were opened and placed in PBS on ice after feces were gently removed. Tissues were swirled in D-PBS until clean of fecal material, placed in DPBS with 0.03 M EDTA and 1.5 mM DTT on ice for 20 mins, then transferred to D-PBS with 0.03 M EDTA in 37°C for 10 min. Samples were shaken vigorously until colonocytes were released, remaining tissues were removed, and DPBS tubes containing

colonocytes were spun at 800g for 5 min. at 4°C. After supernatant was removed, the colonocyte pellet was transferred to a cryovial and stored at -80°C.

For mRNA isolation, colonocytes were thawed, TRI Reagent (Molecular Research Center) was added and incubated for 5 min., 200 µL chloroform was added, cells were spun at max speed, 4°C, for 15 min., and the aqueous phase was collected. An equal volume of 95% ethanol was added and samples were loaded onto Econo-Spin Columns (Epoch Life Science). Samples were washed with 3M sodium acetate, treated with PureLink DNase (Invitrogen), washed with 3x 70% ethanol+1% HEPES, and eluted with RNase-free water. RNA was quantified using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). Complementary DNA was generated from 1 µg of RNA using MultiScribe reverse transcriptase (Applied Biosystems) with 10X RT PCR buffer (Applied Biosystems), 25mM MgCl2 (Applied Biosystems), dNTPs (Applied Biosystems), random hexamers (Applied Biosystems), and RNase inhibitor (Applied Biosystems). Samples were run on a PTC-200 Peltier Thermal Cycler (MJ Research) for 25°C (10 min.), 48°C (30 min.), 95°C (5 min.), then 4°C. Real-time PCR was performed with SYBR green (Applied Biosystems) and then primers listed in **Supplemental Table 4.2** on a ViiA 7 realtime PCR system (Applied Biosystems) with the following cycling parameters: 50°C (2 min), 95°C (10 min), 40 cycles of 95°C (15 s) and 60°C (1 min). Results were analyzed using QuantiStudio Real-Time PCR software v1.3 (Applied Biosystems), and $\Delta\Delta C_T$ was calculated with beta-actin as the control gene.

Hypoxia staining and imaging

30-90 minutes prior to necropsy, mice were injected intraperitoneally with 100 mg/kg of pimonidazole HCl (Hypoxyprobe) in PBS. Staining with the Hypoxyprobe kit was done as previously described (Byndloss et al., 2017a; Cevallos et al., 2021b). Briefly, paraffin-embedded tissues mounted on slides and prepared for staining with xylene (2x10 min.) and ethanol (3 min. each in 95%, 80%, and 70%). Samples were treated with Proteinase K 20 mg/mL in TE buffer for 15 min. at 37°C, nonspecific binding was blocked with serum at room temperature for 1 hour, and slides were stained overnight at 4°C with the mouse IgG1 anti-PMDZ monoclonal antibody 4.3.11.3 (Hypoxyprobe). Slides were then stained with Cyanine3-labeled goat anti-mouse IgG (Jackson ImmunoResearch) for 90 min. at room temperature. Between each staining step, slides were washed 3x 5 min. in PBS. Slides were briefly dried and mounted with Shandon Immu-Mount (Thermo Scientific).

Image numbers were randomized, blinded, and three representative images were collected on a Carl Zeiss AxioVision microcope with AxioVision 4.8.1 software (Zeiss) at 20X for scoring and 63X for image samples. Using ImageJ (NIH), the Texas Red channel (Cyanine3) was isolated, 3 representative slices of equal size containing the epitheliallumen border were saved, and the Plot Profile for each was acquired. After unblinding, PMDZ peaks (epithelial-lumen borders) were aligned for each image and

profiles for the 9 slices associated with each mouse were averaged to obtain an average PMDZ Plot Profile and average PMDZ Peak for each mouse.

Statistical analysis

Statistics were performed using GraphPad Prism 9.4.0 (GraphPad). An unpaired Student t test was performed on log-transformed CFU results, fold change in mRNA expression, and on short chain fatty acid levels. To compare peak pimonidazole staining intensities, a Mann-Whitney U test was performed. Throughout all figures, * = $p<0.05$, ** = $p<0.005$, *** = $p<0.0005$, and **** = $p<0.00005$.

ACKNOWLEDGEMENTS

We would like to thank K. Honda for kindly providing 17 human fecal *Clostridia* isolates. We would like to thank Ardeypharm GmbH for kindly providing *E. coli* strain Nissle 1917. Metabolomics sample prep, GC-TOF mass spectrometry, and data processing were performed by the West Coast Metabolomics Center at the University of California, Davis.

This work was supported by Microbiome Tri-institutional Partnership in Microbiome Research (TrIP) Seed funding award number 000582. Work in A.J.B.'s laboratory was supported by award 650976 from the Crohn's and Colitis Foundation of America and by Public Health Service Grants AI044170, AI096528, AI112949, AI146432 and AI153069. This project was also supported by F32 AI161850 (H.P.S.). The project described was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant number UL1 TR000002 and linked award TL1 TR000133 (H.P.S.). The project described was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant number UL1 TR001860 (D.J.B.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author contributions

H.P.S., D.J.B. and C.R.T study design, supervised personnel, performed mouse experiments and analyzed the data. C.R.T. analyzed the metabolomics data.

M.A.F.G., E.J.B., and H.N. helped perform mouse experiments.

H.L.P.M. performed measurements of butyrate concentrations.

T.P.C. and R.L.S. scored histological sections.

A.T., S.M.N., G.R.T., K.L.R., and A.J.B. contributed to obtaining funding for the research. H.P.S., D.J.B., C.R.T., and A.J.B. were major contributors in conceptualization, supervision and writing the manuscript.

Declarations of Interest

Dr. George Thompson is consulting for Astellas, Cidara, F2G, Immy, Mayne, Melinta, Mundipharma, Scynexis, and Pfizer on projects that are not related to this publication. Dr. George Thompson receives research support from Astellas, Cidara, F2G, Mayne, Melinta, Merck, Mundipharma, Scynexis, and Pfizer for projects that are not related to this publication.

Figures

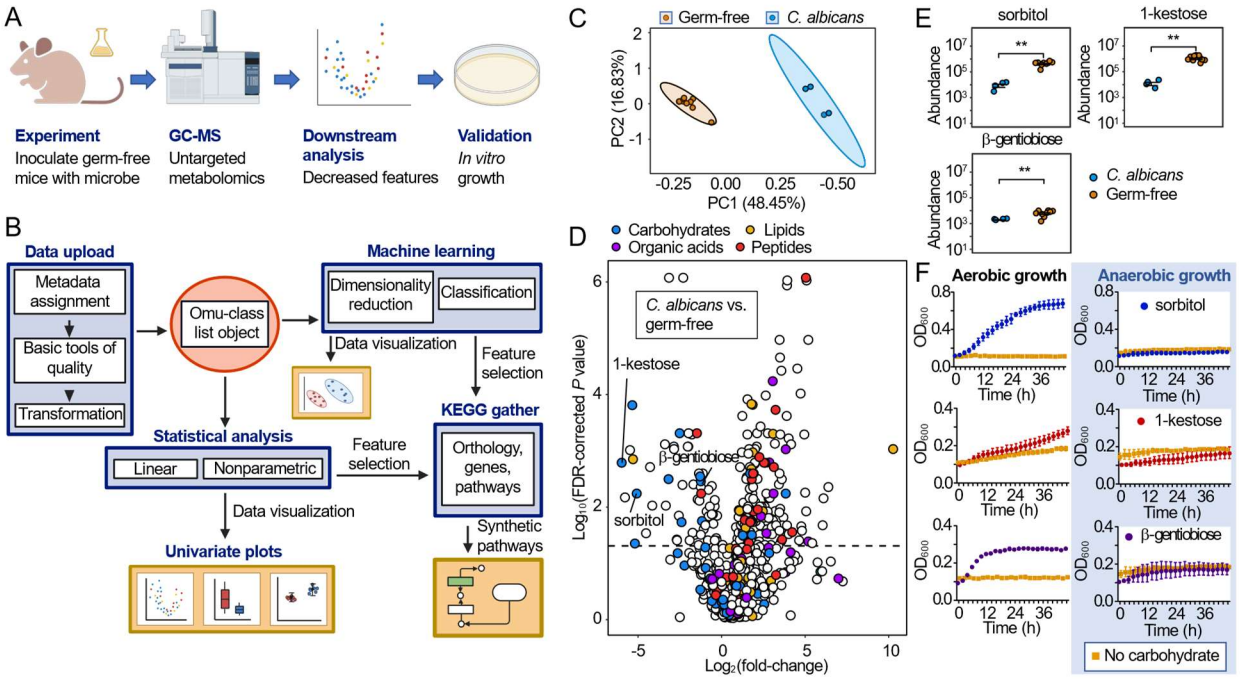

Figure 4.1

Figure 4.1: Metabolic footprinting identifies resources that support *C. albincans* growth in the gastrointestinal tract of mice. (A) Schematic of steps involved in metabolic footprinting. (B) Workflow of metabolomics data analysis with the reactive web application omuShiny. (C-E) Comparative analysis of the cecal metabolome of germ-free mice (n = 6), and gnotobiotic mice three days after engraftment with *C. albicans* (n = 4). (C) Principal component analysis of the cecal metabolome in the indicated groups of mice. Ovals indicate the 95% confidence interval. (D) Volcano plot of metabolites colored by class. The Y-axis shows the decadic logarithm of the false discovery rate (FDR)-corrected *P* value. The dashed line is set at an FDR corrected *P* value of 0.05. Metabolites with a negative fold-change value decreased in mice engrafted with *C. albicans* compared to germ-free mice, while metabolites with a positive fold-change value increased. (E) The graphs show the abundance of the indicated sugars in mice engrafted with *C. albicans* compared to germ-free mice. (C and E) Each dot represents data from one animal. (F) *In vitro* growth of *C. albicans* in yeast nitrogen broth supplemented with cysteine and the indicated sugars under aerobic (left panels) or anaerobic (right panels) growth conditions. GC-MS, gas chromatography time-of-flight mass spectrometry; KEGG, Kyoto Encyclopedia of Genes and Genomes; **, $P < 0.01$.

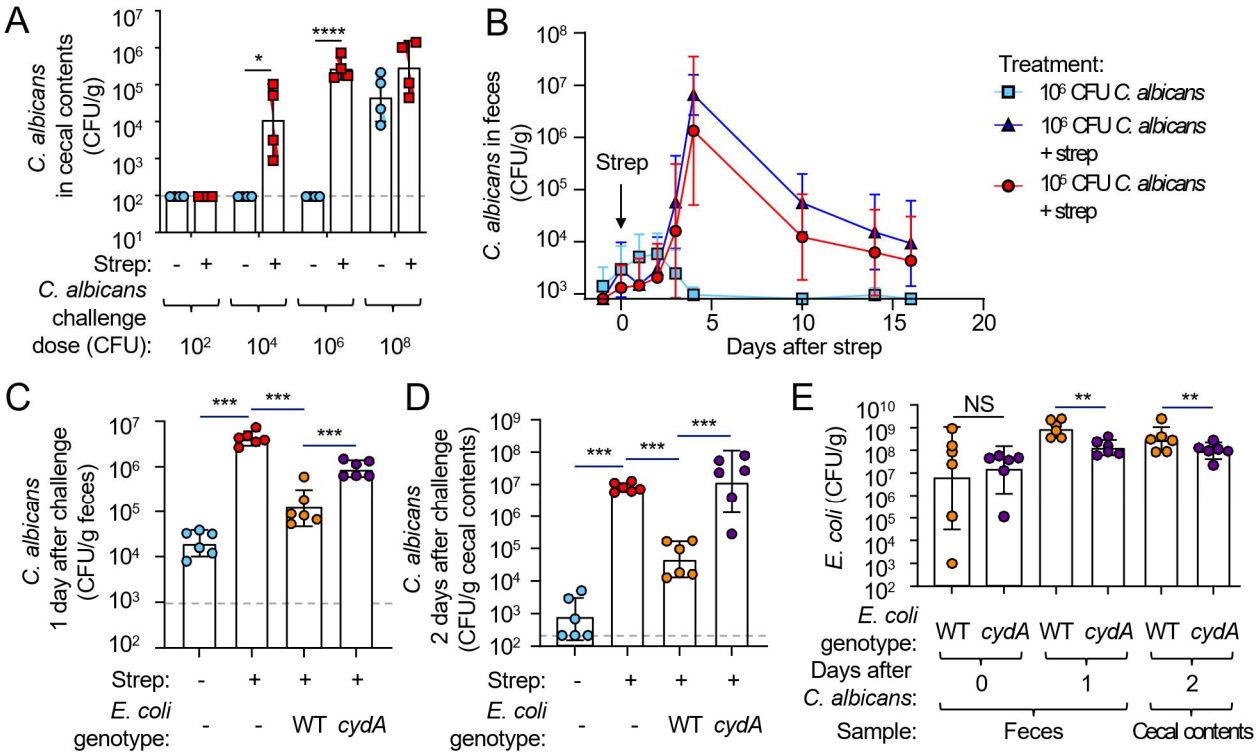
A
C. albicans in cecal contents (CFU/g)
Strep:
C. albicans challenge dose (CFU):
B
C. albicans in feces (CFU/g)
Strep
Days after strep
Treatment:
10⁶ CFU C. albicans
10⁶ CFU C. albicans + strep
10⁵ CFU C. albicans + strep
C
C. albicans 1 day after challenge (CFU/g feces)
E. coli genotype:
D
C. albicans 2 days after challenge (CFU/g cecal contents)
E
E. coli (CFU/g)
Days after C. albicans:
Sample:
Feces
Cecal contents
NS

Figure 4.2

Figure 4.2: *E. coli* requires aerobic respiration to restore colonization resistance against *C. albincans* after streptomycin treatment. (A) Groups (n = 4) of special pathogen-free C57BL/6J mice were either given a single dose of streptomycin (20 mg/animal) by oral gavage (Strep: +) or were mock treated (Strep: -) and were challenged 48 hours later with the indicated doses of *C. albicans*. Shown are geometric means (bars) +/- geometric standard deviation (error bars) of *C. albicans* colony-forming units (CFU) recovered per gram (g) of cecal contents. (B) Special pathogen-free C57BL/6J mice were infected with the indicated dose of *C. albicans* and were either mock-treated (n = 4) or given a single dose of streptomycin (20 mg/animal) by oral gavage five days after *C. albicans* infection (n = 8). Shown are geometric means (symbols) +/- geometric standard deviation (error bars) of *C. albicans* CFU recovered per gram of feces at the indicated time points. (C-E) Special pathogen-free C57BL/6J mice were either mock treated (-) or given a single dose of streptomycin (20 mg/animal) by oral gavage (+). The next day, mice were either mock inoculated, inoculated with 10^9 CFU of *E. coli* Nissle 1917 (WT) or 10^9 CFU of an *E. coli* Nissle 1917 *cydA* mutant (*cydA*). One day later, mice were challenged with 10^5 CFU of *C. albicans* and organs were collected two days later. (C and D) Shown are geometric means (bars) +/- geometric standard deviation (error bars) of *C. albicans* CFU recovered one day after challenge from feces (C) or two days after challenge from cecal contents (D). (E) Shown are geometric means (bars) +/- geometric standard deviation (error bars) of *E. coli* CFU recovered from feces or cecal contents at the indicated time points relative to *C. albicans* challenge. (A and C-E) Each symbol represents data from one animal. NS, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.005$.

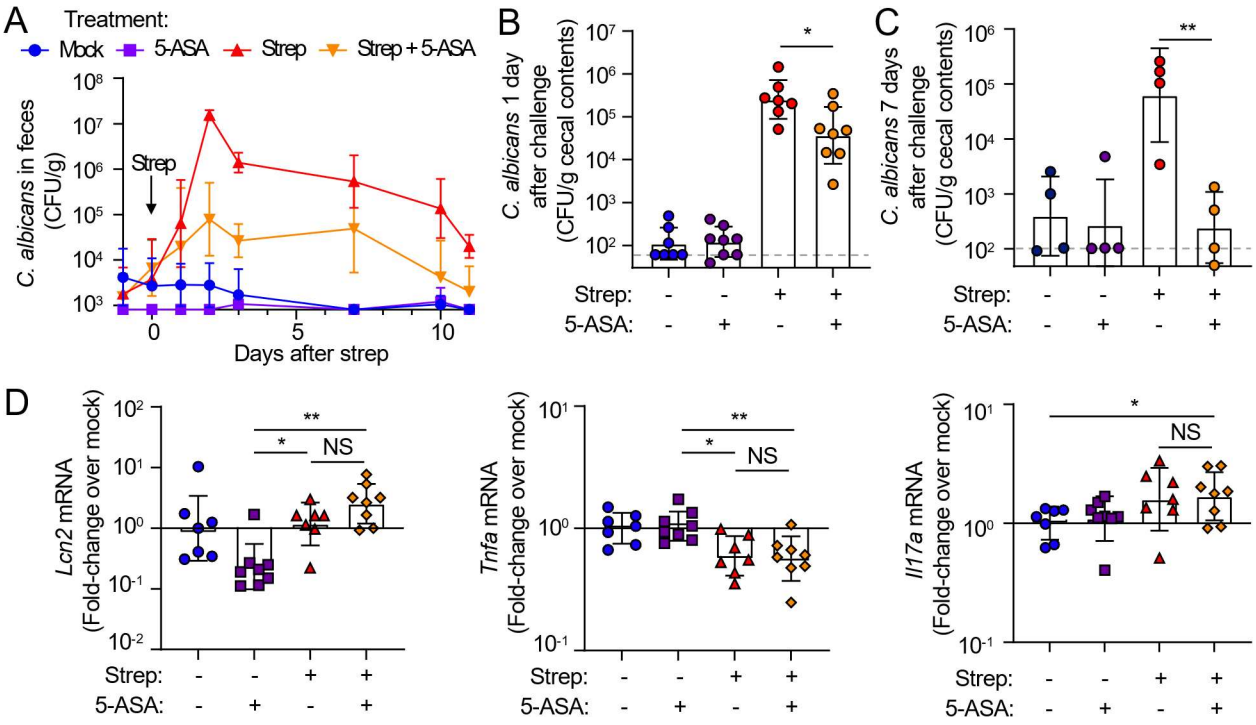
A
Treatment:
Mock
5-ASA
Strep
Strep + 5-ASA
C. albicans in feces (CFU/g)
Strep
Days after strep
B
C. albicans 1 day after challenge (CFU/g cecal contents)
C
C. albicans 7 days after challenge (CFU/g cecal contents)
Strep:
5-ASA:
D
Lcn2 mRNA (Fold-change over mock)
Tnfa mRNA (Fold-change over mock)
Il17a mRNA (Fold-change over mock)
NS

Figure 4.3

Figure 4.3: 5-ASA restores colonization resistance against *C. albicans* after streptomycin treatment. (A) Mice maintained on control chow or chow supplemented with 5-aminosalicylic acid (5-ASA) were challenged with 10^6 colony-forming units (CFU) of *C. albicans*. Three days later, mice were mock-treated or received a single dose of streptomycin (Strep) by oral gavage. Shown are geometric mean +/- geometric standard deviation of *C. albicans* CFU recovered from feces in each group (n = 6 for strep + 5-ASA, n = 3 for remaining groups) at the indicated time points. (B-D) Mice were started on chow supplemented with 5-aminosalicylic acid (5-ASA: +) or control chow (5-ASA: -) and two days later mice received a single dose of streptomycin by oral gavage (Strep: +) or were mock-treated (Strep: -). Each symbol represents data from one animal (i.e., n for each group is provided by the number of symbols). (B and D) Mice were challenged with 10^5 *C. albicans* CFU one day after streptomycin treatment and *C. albicans* CFU in cecal contents were determined the next day. (C) Mice were challenged with 10^6 *C. albicans* CFU two days after streptomycin treatment and *C. albicans* CFU in cecal contents were determined seven days later. (B and C) Shown are geometric means (bars) +/- geometric standard deviation (error bars) of *C. albicans* CFU recovered from cecal contents. (D) RNA was extracted from preparations of colonic epithelia cells collected one day after *C. albicans* challenge. Transcript levels of *Lcn2*, *Tnfa*, and *Il17a* were determined by quantitative real-time PCR. Shown are geometric means (bars) +/- geometric standard deviation (error bars) of fold-changes in transcript levels of *Lcn2*, *Tnfa*, or *Il17a* compared to transcript levels in mock-treated mice on control chow. NS, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$.

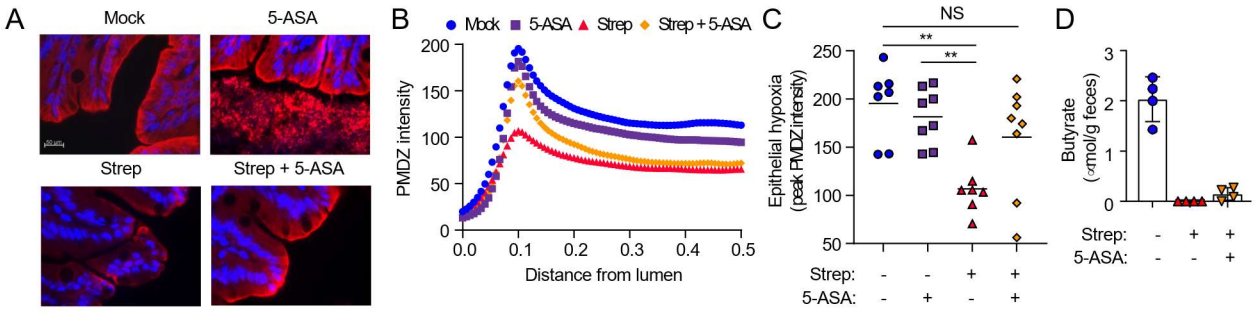
A
Mock
5-ASA
Strep
Strep + 5-ASA
B
Mock
5-ASA
Strep
Strep + 5-ASA
PMDZ intensity
Distance from lumen
C
NS
**
**
Epithelial hypoxia (peak PMDZ intensity)
Strep:
5-ASA:
D
Butyrate (μmol/g feces)
Strep:
5-ASA:

Figure 4.4

Figure 4.4: 5-ASA restores epithelial hypoxia in the large intestine after streptomycin treatment. Mice were started on chow supplemented with 5-aminosalicylic acid (5-ASA: +) or control chow (5-ASA: -) and two days later mice received a single dose of streptomycin by oral gavage (Strep: +) or were mock-treated (Strep: -). Mice were challenged with 10^5 *C. albicans* CFU one day after streptomycin treatment. One day after challenge, mice were injected with pimonidazole HCl (PMDZ) and the cecum was collected one hour later. (A-C) Binding of pimonidazole was detected using hypoxyprobe-1 primary antibody and a Cy-3 conjugated goat anti-mouse secondary antibody (red fluorescence) in histological sections from the cecum that were counterstained with DAPI nuclear stain (blue fluorescence). (A) Representative images are shown. (B) Pimonidazole staining was quantified by measuring mean PMDZ intensities from the lumen (distance of 0.0 arbitrary units) to the border of the colonocytes (distance of 0.1), and into the tissue. (C) The graph shows the peak PMDZ intensity for each mouse (symbols) and the mean peak intensity for each group (lines) (n = 7-8). (D) Butyrate concentrations in cecal contents were determined by gas chromatography-mass spectrometry (n = 4). NS, $P > 0.05$; **, $P < 0.01$.

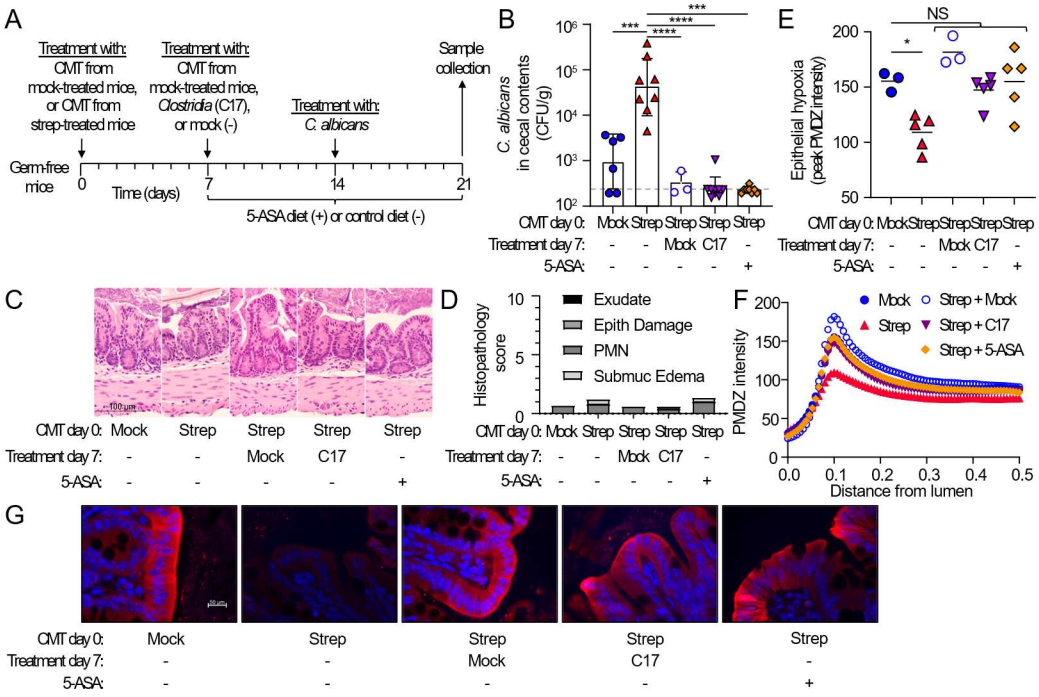
A
Treatment with:
CMT from
mock-treated mice,
or CMT from
strep-treated mice
Treatment with:
CMT from
mock-treated mice,
Clostridia (C17),
or mock (-)
Treatment with:
C. albicans
Sample
collection
Germ-free
mice
0
Time (days)
7
14
21
5-ASA diet (+) or control diet (-)
B
C. albicans
in cecal contents
(CFU/g)
CMT day 0: Mock Strep Strep Strep Strep
Treatment day 7: - - Mock C17 -
5-ASA: - - - - +
E
Epithelial hypoxia
(peak PMDZ intensity)
NS
CMT day 0: Mock Strep Strep Strep Strep
Treatment day 7: - - Mock C17 -
5-ASA: - - - - +
C
100 µm
CMT day 0: Mock Strep Strep Strep Strep
Treatment day 7: - - Mock C17 -
5-ASA: - - - - +
D
Histopathology
score
Exudate
Epith Damage
PMN
Submuc Edema
CMT day 0: Mock Strep Strep Strep Strep
Treatment day 7: - - Mock C17 -
5-ASA: - - - - +
F
PMDZ intensity
Mock
Strep
Strep + Mock
Strep + C17
Strep + 5-ASA
Distance from lumen
G
CMT day 0: Mock Strep Strep Strep Strep
Treatment day 7: - - Mock C17 -
5-ASA: - - - - +

Figure 4.5

Figure 4.5: 5-ASA can functionally replace *Clostridia* species to restore colonization resistance against *C. albicans* after streptomycin treatment. Germ-free Swiss Webster mice received a cecal microbiota transplant from streptomycin-treated C57BL/6J mice (Strep) or from mock-treated C57BL/6J mice (Mock). Seven days later, some mice were inoculated with a community of human *Clostridia* isolates (C17), received a second cecal microbiota transplant from mock-treated C57BL/6J mice (Mock), or were switched to chow supplemented with 5-aminosalicylic acid (5-ASA: +). Seven days later, all mice were challenged with 10^5 *C. albicans* CFU and organs were collected one week later. (A) Schematic of the experimental timeline. (B) Shown are *C. albicans* CFU recovered 7 days after challenge from cecal contents of each animal (symbols; n = 3 for Strep + Mock, n = 6 for Strep + Mock and Strep + C17, n = 8-9 for remaining groups) and geometric means (bars) +/- geometric standard deviation (error bars). (C and D) Blinded sections from the cecum of each animal were scored by a veterinary pathologist. (C) Representative images of colonic sections for each group are shown. (D) Average histopathology score for each group. (E-G) Mice were injected with pimonidazole HCl (PMDZ) and the cecum was collected one hour later. Binding of pimonidazole was detected using hypoxyprobe-1 primary antibody and a Cy-3 conjugated goat anti-mouse secondary antibody (red fluorescence) in histological sections from the cecum that were counterstained with DAPI nuclear stain (blue fluorescence). (E) The graph shows the peak PMDZ intensity for each mouse (symbols) and the mean peak intensity for each group (lines). (F) Pimonidazole staining was quantified by measuring mean PMDZ intensities from the lumen (distance of 0.0 arbitrary units) to the border of the colonocytes

(distance of 0.1), and into the tissue. (G) Representative images are shown. NS, $P > 0.05$; ***, $P < 0.005$; ****, $P < 0.001$.

SUPPLEMENTAL MATERIALS

Supplemental Table 4.1: Histopathology scoring. Scoring criteria for H&E-stained cecal tips.

Score	Exudate	Epithelial damage	Neutrophils	Submucosal edema
0	Absent	Absent	0-5 per hpf	Absent
1	Slight accumulation	Desquamation	6-20 per hpf	Slight (<10%)
2	Mild accumulation	Mild erosion	21-50 per hpf	Mild (10-20%)
3	Moderate accumulation	Marked erosion	51-100 per hpf	Moderate (20-40%)
4	Severe accumulation	Ulceration	>100 per hpf	Severe (>40%)

Supplemental Table 4.2: Primers.

Target genes	Name	Sequence	Reference
Actb	Beta actin	F: TGTCCACCTTCCAGCAGATGT R: AGCTCAGTAACAGTCCGCCTAGA	Gu et al., 2013
Lcn2	Lipocalin 2	F: ACATTTGTTCCAAGCTCCAGGGC R: CATGGCGAACTGGTTGTAGTCCG	Godinez et al., 2008
Tnfa	TNF-α	F: AGCCAGGAGGGAGAACAGAAAC R: CCAGTGAGTGAAAGGGACAGAACC	Wilson et al., 2008
Il17a	IL-17A	F: AACCCCCACGTTTCTCAGCAAAC R: GGACCCCTTTACACCTTCTTTTCATTG	Godinez et al., 2008

REFERENCES

Alonso-Monge, R., Gresnigt, M.S., Roman, E., Hube, B., and Pla, J. (2021). Candida albicans colonization of the gastrointestinal tract: A double-edged sword. PLoS Pathog *17*, e1009710.

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K.*, et al.* (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature *500*, 232-236.

Byndloss, M.X., Olsan, E.E., Rivera-Chavez, F., Tiffany, C.R., Cevallos, S.A., Lokken, K.L., Torres, T.P., Byndloss, A.J., Faber, F., Gao, Y.*, et al.* (2017a). Microbiota-activated PPAR-gamma signaling inhibits dysbiotic Enterobacteriaceae expansion. Science (New York, NY) *357*, 570-575.

Byndloss, M.X., Olsan, E.E., Rivera-Chávez, F., Tiffany, C.R., Cevallos, S.A., Lokken, K.L., Torres, T.P., Byndloss, A.J., Faber, F., Gao, Y.*, et al.* (2017b). Microbiota-activated PPAR-g signaling inhibits dysbiotic Enterobacteriaceae expansion. Science *357*, 570-575.

Cevallos, S.A., Lee, J.Y., Tiffany, C.R., Byndloss, A.J., Johnston, L., Byndloss, M.X., and Baumler, A.J. (2019). Increased Epithelial Oxygenation Links Colitis to an Expansion of Tumorigenic Bacteria. MBio *10*.

Cevallos, S.A., Lee, J.Y., Velazquez, E.M., Foegeding, N.J., Shelton, C.D., Tiffany, C.R., Parry, B.H., Stull-Lane, A.R., Olsan, E.E., Savage, H.P.*, et al.* (2021a). 5-Aminosalicylic

Acid Ameliorates Colitis and Checks Dysbiotic Escherichia coli Expansion by Activating PPAR-gamma Signaling in the Intestinal Epithelium. mBio *12*.

Cevallos, S.A., Lee, J.Y., Velazquez, E.M., Foegeding, N.J., Shelton, C.D., Tiffany, C.R., Parry, B.H., Stull-Lane, A.R., Olsan, E.E., Savage, H.P.*, et al.* (2021b). 5-Aminosalicylic Acid Ameliorates Colitis and Checks Dysbiotic Escherichia coli Expansion by Activating PPAR-γ Signaling in the Intestinal Epithelium. mBio *12*.

Clark, J.D. (1971). Influence of antibiotics or certain intestinal bacteria on orally administered Candida albicans in germ-free and conventional mice. Infect Immun *4*, 731-737.

Furuta, G.T., Turner, J.R., Taylor, C.T., Hershberg, R.M., Comerford, K., Narravula, S., Podolsky, D.K., and Colgan, S.P. (2001). Hypoxia-inducible factor 1-dependent induction of intestinal trefoil factor protects barrier function during hypoxia. J Exp Med *193*, 1027-1034.

Gillis, C.C., Hughes, E.R., Spiga, L., Winter, M.G., Zhu, W., Furtado de Carvalho, T., Chanin, R.B., Behrendt, C.L., Hooper, L.V., Santos, R.L.*, et al.* (2018). Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. Cell Host Microbe *23*, 54-64 e56.

Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., Herwaldt, L., Pfaller, M., and Diekema, D. (2003). Attributable mortality of nosocomial candidemia, revisited. Clin Infect Dis *37*, 1172-1177.

Guinan, J., Wang, S., Hazbun, T.R., Yadav, H., and Thangamani, S. (2019). Antibiotic-induced decreases in the levels of microbial-derived short-chain fatty acids correlate with increased gastrointestinal colonization of Candida albicans. Sci Rep *9*, 8872.

Gutierrez, D., Weinstock, A., Antharam, V.C., Gu, H., Jasbi, P., Shi, X., Dirks, B., Krajmalnik-Brown, R., Maldonado, J., Guinan, J.*, et al.* (2020). Antibiotic-induced gut metabolome and microbiome alterations increase the susceptibility to Candida albicans colonization in the gastrointestinal tract. FEMS Microbiol Ecol *96*.

Huang, D., Li, H., Lin, Y., Lin, J., Li, C., Kuang, Y., Zhou, W., Huang, B., and Wang, P. (2022). Antibiotic-induced depletion of Clostridium species increases the risk of secondary fungal infections in preterm infants. Front Cell Infect Microbiol *12*, 981823.

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

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res *28*, 27-30.

Kelly, C.J., Zheng, L., Campbell, E.L., Saeedi, B., Scholz, C.C., Bayless, A.J., Wilson, K.E., Glover, L.E., Kominsky, D.J., Magnuson, A.*, et al.* (2015). Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. Cell Host Microbe *17*, 662-671.

Kennedy, M.J., and Volz, P.A. (1983). Dissemination of yeasts after gastrointestinal inoculation in antibiotic-treated mice. Sabouraudia *21*, 27-33.

Kizaka-Kondoh, S., and Konse-Nagasawa, H. (2009). Significance of nitroimidazole compounds and hypoxia-inducible factor-1 for imaging tumor hypoxia. Cancer Sci *100*, 1366-1373.

Lee, J.Y., Tsolis, R.M., and Baumler, A.J. (2022). The microbiome and gut homeostasis. Science *377*, eabp9960.

Litvak, Y., Byndloss, M.X., and Baumler, A.J. (2018). Colonocyte metabolism shapes the gut microbiota. Science *362*, eaat9076.

Louis, P., and Flint, H.J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol Lett *294*, 1-8.

Magill, S.S., O'Leary, E., Janelle, S.J., Thompson, D.L., Dumyati, G., Nadle, J., Wilson, L.E., Kainer, M.A., Lynfield, R., Greissman, S.*, et al.* (2018). Changes in Prevalence of Health Care-Associated Infections in U.S. Hospitals. N Engl J Med *379*, 1732-1744.

Mason, K.L., Erb Downward, J.R., Mason, K.D., Falkowski, N.R., Eaton, K.A., Kao, J.Y., Young, V.B., and Huffnagle, G.B. (2012). Candida albicans and bacterial microbiota interactions in the cecum during recolonization following broad-spectrum antibiotic therapy. Infect Immun *80*, 3371-3380.

Meynell, G.G. (1963). Antibacterial mechanisms of the mouse gut. II. The role of Eh and volatile fatty acids in the normal gut. British journal of experimental pathology *44*, 209-219.

Moreno-Sanchez, F., and Gomez-Gomez, B. (2022). Antibiotic Management of Patients with Hematologic Malignancies: From Prophylaxis to Unusual Infections. Curr Oncol Rep *24*, 835-842.

Morgan, J., Meltzer, M.I., Plikaytis, B.D., Sofair, A.N., Huie-White, S., Wilcox, S., Harrison, L.H., Seaberg, E.C., Hajjeh, R.A., and Teutsch, S.M. (2005). Excess mortality, hospital stay, and cost due to candidemia: A case-control study using data from population-based candidemia surveillance. Infect Cont Hosp Ep *26*, 540-547.

Narushima, S., Sugiura, Y., Oshima, K., Atarashi, K., Hattori, M., Suematsu, M., and Honda, K. (2014). Characterization of the 17 strains of regulatory T cell-inducing human-derived Clostridia. Gut microbes *5*, 333-339.

Nguyen, L.N., Lopes, L.C.L., Cordero, R.J.B., and Nosanchuk, J.D. (2011). Sodium butyrate inhibits pathogenic yeast growth and enhances the functions of macrophages. J Antimicrob Chemoth *66*, 2573-2580.

Nissle, A. (1925). Weiteres über grundlagen und praxis der mutaflorbehandlung. DMW-Deutsche Medizinische Wochenschrift *51*, 1809-1813.

Noverr, M.C., and Huffnagle, G.B. (2004). Regulation of Candida albicans morphogenesis by fatty acid metabolites. Infect Immun *72*, 6206-6210.

Pappas, P.G., Lionakis, M.S., Arendrup, M.C., Ostrosky-Zeichner, L., and Kullberg, B.J. (2018). Invasive candidiasis. Nat Rev Dis Primers *4*, 18026.

Redman, M.G., Ward, E.J., and Phillips, R.S. (2014). The efficacy and safety of probiotics in people with cancer: a systematic review. Ann Oncol *25*, 1919-1929.

Rivera-Chavez, F., Zhang, L.F., Faber, F., Lopez, C.A., Byndloss, M.X., Olsan, E.E., Xu, G., Velazquez, E.M., Lebrilla, C.B., Winter, S.E.*, et al.* (2016). Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella. Cell Host Microbe *19*, 443-454.

Rousseaux, C., Lefebvre, B., Dubuquoy, L., Lefebvre, P., Romano, O., Auwerx, J., Metzger, D., Wahli, W., Desvergne, B., Naccari, G.C.*, et al.* (2005). Intestinal antiinflammatory effect of 5-aminosalicylic acid is dependent on peroxisome proliferator-activated receptor-gamma. J Exp Med *201*, 1205-1215.

Samonis, G., Anastassiadou, H., Dassiou, M., Tselentis, Y., and Bodey, G.P. (1994). Effects of broad-spectrum antibiotics on colonization of gastrointestinal tracts of mice by Candida albicans. Antimicrob Agents Chemother *38*, 602-603.

Sekirov, I., Tam, N.M., Jogova, M., Robertson, M.L., Li, Y., Lupp, C., and Finlay, B.B. (2008). Antibiotic-induced perturbations of the intestinal microbiota alter host susceptibility to enteric infection. Infect Immun *76*, 4726-4736.

Spiga, L., Winter, M.G., Furtado de Carvalho, T., Zhu, W., Hughes, E.R., Gillis, C.C., Behrendt, C.L., Kim, J., Chessa, D., Andrews-Polymenis, H.L.*, et al.* (2017). An Oxidative Central Metabolism Enables Salmonella to Utilize Microbiota-Derived Succinate. Cell host & microbe *22*, 291-301.e296.

Tiffany, C.R., and Baumler, A.J. (2019). omu, a Metabolomics Count Data Analysis Tool for Intuitive Figures and Convenient Metadata Collection. Microbiol Resour Announc *8*.

Tiffany, C.R., Lee, J.Y., Rogers, A.W.L., Olsan, E.E., Morales, P., Faber, F., and Baumler, A.J. (2021). The metabolic footprint of Clostridia and Erysipelotrichia reveals their role in depleting sugar alcohols in the cecum. Microbiome *9*, 174.

Viscoli, C., Girmenia, C., Marinus, A., Collette, L., Martino, P., Vandercam, B., Doyen, C., Lebeau, B., Spence, D., Krcmery, V.*, et al.* (1999). Candidemia in cancer patients: a prospective, multicenter surveillance study by the Invasive Fungal Infection Group

(IFIG) of the European Organization for Research and Treatment of Cancer (EORTC). Clin Infect Dis *28*, 1071-1079.

Vital, M., Howe, A.C., and Tiedje, J.M. (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. MBio *5*, e00889.

Wilson, L.S., Reyes, C.M., Stolpman, M., Speckman, J., Allen, K., and Beney, J. (2002). The direct cost and incidence of systemic fungal infections. Value Health *5*, 26-34.

Winter, S.E., Lopez, C.A., and Baumler, A.J. (2013). The dynamics of gut-associated microbial communities during inflammation. EMBO Rep *14*, 319-327.

Zhai, B., Ola, M., Rolling, T., Tosini, N.L., Joshowitz, S., Littmann, E.R., Amoretti, L.A., Fontana, E., Wright, R.J., Miranda, E.*, et al.* (2020). High-resolution mycobiota analysis reveals dynamic intestinal translocation preceding invasive candidiasis. Nat Med *26*, 59-64.

Chapter 5

Concluding Remarks

This thesis presents a novel framework called metabolic footprinting, which was created to discover unknown functions in host-associated microbiomes. Using this framework, we determined that *Clostridia* and *Erysipelotrichia* provide an important digestive function for the host by consuming sugar alcohols in the large bowel, and that the gut microbiota restricts growth of the human fungal pathogen *Candida albicans* by limiting access to oxygen. These studies demonstrate the ability of metabolic footprinting to expand our knowledge of host-associated microbiomes.

Development of the Metabolic Footprinting Framework.

Initially in our review “Dysbiosis: from fiction to function”, we proposed that petabytes of additional sequencing data aimed at performing taxonomic descriptions will not help determine the functions of host-associated microbiomes during homeostasis. A poor understanding of the microbiome during homeostasis in turn makes it difficult to understand dysbiosis in the context of different etiologies. We proposed that pursuing mechanistic research using a holistic approach that incorporates theory and methodology from microbiology, immunology, physiology, and ecology was necessary to develop a better understanding of host-associated microbiomes. To that end, we created a framework called metabolic footprinting. The method begins by inoculating a germ-free organism with the microbe of interest. In the studies presented in this thesis, we utilized germ-free mice. An untargeted metabolomics screen is then performed on

samples to determine what small, soluble metabolites are present. During analysis, the metabolites that are significantly decreased in the group colonized by the microbe of interest relative to germ-free controls are considered candidate nutrient niches. *In-vitro* growth assays on defined minimal media are then performed on candidate metabolites to determine if the microbe can utilize them as sole carbon sources for growth.

For the analysis of untargeted-metabolomics data, we developed an R package named omu, which is the Japanese word for omelet. The name is an homage to the Kyoto Encyclopedia of Genes and Genomes (KEGG), a database with a logo that is an egg filled with nucleotides. This name was chosen because omu leverages metabolite metadata within KEGGs compound BRITE hierarchy database to provide descriptive metadata for features in a metabolomics dataset. These metadata follow a hierarchy of classes and subclasses (e.g. Carbohydrates, Monosaccharides, Sugar alcohols). We also developed an extension of omu named omuShiny, which is a web application with a graphical user interface and does not require any programming expertise.

Metabolic footprinting is an effective framework for discovering functions of host-associated microbiomes. While we demonstrated metabolic footprinting's efficacy in mouse models focused on the gut microbiome, we believe this framework can be generalized to any host-associated microbiome where gnotobiotic models are feasible.

Metabolic footprinting reveals Clostridia and Erysipelotrichia are the main consumers of sugar alcohols in the large bowel of mammals.

Numerous studies have been performed to determine how commensal microbes degrade dietary fiber in the mammalian gut, but the degradation of small, unabsorbed sugars remains poorly understood. In a pilot experiment, we found that antibiotic treatment elevated sugar alcohols in the cecal metabolome of mice and depleted *Clostridia* and *Erysipelotrichia*, two classes of obligate anaerobes that are considered key members of a healthy gut microbiota due to their production of the short-chain fatty acid butyrate. While the depletion of sugar alcohols by the gut microbiota had been previously shown, the specific bacteria performing this function remained unknown. Following our pilot study, we performed metabolic footprinting using a consortium of 17 human *Clostridia* and *Erysipelotrichia* isolates and discovered that *Clostridia* and *Erysipelotrichia* were consuming sugar alcohols in the gut. This result piqued our interest because many patients with either irritable bowel syndrome or inflammatory bowel disease experience exacerbated symptoms after consuming sugar alcohols, with sorbitol being the most common trigger. In a proof-of-concept experiment, we then demonstrated that administration of a sugar-alcohol-consuming Clostridial isolate following antibiotic treatment restored sorbitol degradation in the host. Collectively, our results suggest that *Clostridia* and *Erysipelotrichia* are the primary consumers of alcoholic sugars in the large bowel of mammals, and that antibiotic treatment can ablate sugar alcohol degradation by depleting these commensal bacteria. Future research should investigate whether low abundance of *Clostridia* and *Erysipelotrichia* correlates with sugar alcohol intolerance in humans, and if commensal bacteria within these classes can be used as second-generation probiotics to treat sugar alcohol intolerance.

Our success in identifying a novel, important function of *Clostridia* and *Erysipelotrichia* demonstrates the ability of metabolic footprinting to uncover digestive tasks performed by gut microbiota.

Metabolic footprinting reveals oxygen is a critical resource for expansion of the human fungal pathogen Candida albicans in the gut.

It has been known for decades that gut bacteria provide colonization resistance against the opportunistic human fungal pathogen *Candida albicans*. During homeostasis, *C. albicans* exists as a commensal yeast and minority member of the gut microbiota. However, in patients on antibiotic prohylaxis who are immunocompromised, *C. albicans* expands in the gut to the point of intestinal dominance (>30% relative abundance). *C. albicans* then disseminates outside of the large bowel and germinates into its hyphal form. This disease is called invasive candidiasis, and it is one of the leading nosocomial infections globally with a mortality rate of approximately 49%. While it is known that commensal bacteria are crucial for preventing the intestinal domination of *C. albicans*, the mechanisms responsible for colonization resistance remain poorly understood. Moreover, while much research has focused on what resources *C.albicans* can use for growth *in vitro*, few studies have looked at what resources *C. albicans* consumes *in vivo*.

Using metabolic footprinting, we determined that *C. albicans* was consuming small sugars in the large bowel but was unable to grow on these sugars under anaerobic conditions, suggesting that oxygen is a critical resource for growth on these carbohydrates. Previous research has demonstrated that antibiotic-induced depletion of the gut microbiota ablates physiological hypoxia of the colonic epithelium, leading to an increase in oxygen availability. We thus hypothesized that the gut microbiota may provide colonization resistance against C. albicans by limiting oxygen in the environment.

In accordance with our hypothesis, we found that the probiotic bacterium *Escherichia coli* Nissle 1917 conferred colonization resistance against *C. albicans*, but an isogenic mutant that is unable to respire using oxygen as a terminal electron acceptor did not, suggesting that colonization resistance against *C. albicans* provided by facultative anaerobic bacteria is dependent on competition over oxygen. Additionally, we discovered that the anti-colitis drug 5-aminosalicylic acid (5-ASA) restored both epithelial hypoxia and colonization resistance against *C. albicans* after antibiotic treatment. This result was exciting because patients who are most at risk of developing invasive candidiasis are on antibiotic prophylaxis, which makes the use of probiotics or fecal microbiota transplants to prevent intestinal domination by *C. albicans* unrealistic. Given the findings of our study, we believe future research into the use of 5-ASA as a prophylactic drug to prevent *C. albicans* expansion in vulnerable patients is warranted.

This study uncovered critical resources for *C. albicans* expansion in the gut, and determined that the microbiota prevents expansion of *C. albicans* by limiting oxygen

availability. These results suggest that metabolic footprinting is a useful tool to elucidate mechanisms underpinning microbiota-mediated colonization resistance.