Title

A bacterial sialidase mediates early-life colonization by a pioneering gut commensal.

Journal

Cell Host & Microbe, 32(2)

Authors

Buzun, Ekaterina

Hsu, Chia-Yun

Sejane, Kristija

et al.

Publication Date

2024-02-14

DOI

10.1016/j.chom.2023.12.014

Peer reviewed

HHS Public Access

Author manuscript

Cell Host Microbe. Author manuscript; available in PMC 2025 February 14.

Published in final edited form as:

Cell Host Microbe. 2024 February 14; 32(2): 181–190.e9. doi:10.1016/j.chom.2023.12.014.

A bacterial sialidase mediates early life colonization by a pioneering gut commensal

Ekaterina Buzun[1], **Chia-Yun Hsu**[1], **Kristija Sejane**[2], **Renee E. Oles**[1], **Adriana Vasquez Ayala**[1], **Luke R. Loomis**[1], **Jiaqi Zhao**[1], **Leigh-Ana Rossitto**[3], **Dominic M. McGrosso**[3], **David J. Gonzalez**[3,4], **Lars Bode**[2,5,6], **Hiutung Chu**[1,6,7,8,9]

[1]Department of Pathology, University of California San Diego, La Jolla, CA, 92093, United States

[2]Department of Pediatrics, University of California San Diego, La Jolla, CA, 92093, United States

[3]Department of Pharmacology, University of California San Diego, La Jolla, CA, 92093, United States

[4]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, 92093, United States

[5]Larsson-Rosenquist Foundation Mother-Milk-Infant Center of Research Excellence (MOMI CORE), University of California San Diego, La Jolla, CA, 92093, United States

[6]Human Milk Institute (HMI), University of California San Diego, La Jolla, CA, 92093, United States

[7]Chiba University-UC San Diego Center for Mucosal Immunology, Allergy and Vaccines (cMAV), University of California, San Diego, La Jolla, CA, 92093, United States

[8]Humans and the Microbiome Program, Canadian Institute for Advanced Research, Toronto, ON M5G 1M1, Canada

[9]Lead Contact

SUMMARY

The early microbial colonization of the gastrointestinal tract can have long-term impacts on development and health. Keystone species, including *Bacteroides spp.*, are prominent in early life and play crucial roles in maintaining the structure of the intestinal ecosystem. However, the process by which a resilient community is curated during early life remains inadequately

Correspondence: hiuchu@ucsd.edu.

Autor Contributions

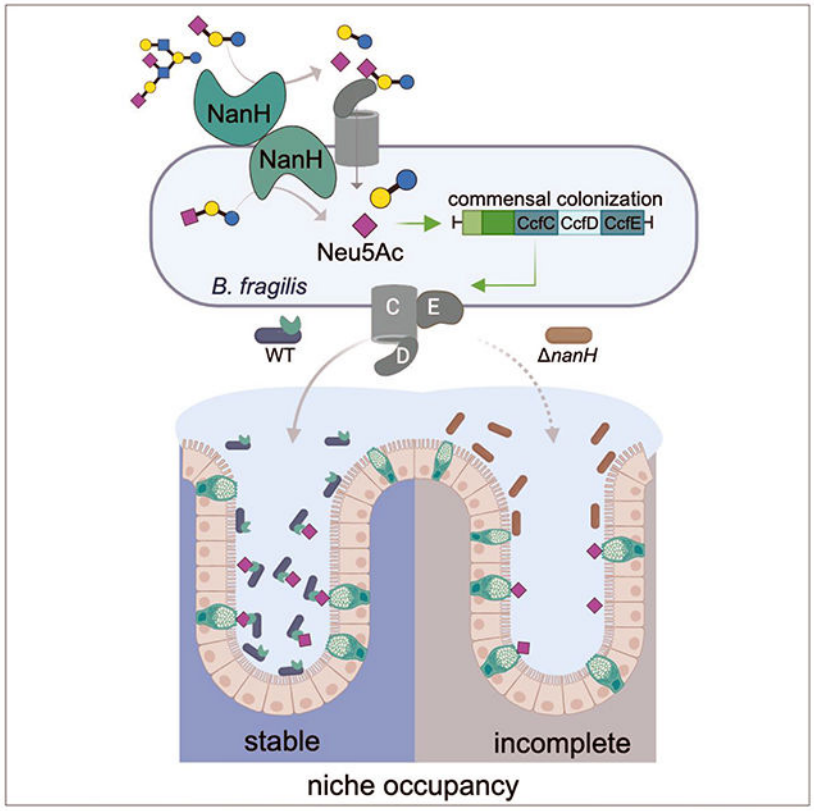
Conceptualization, E.B., H.C.; Methodology, E.B., H.C.; Software, R.E.O., L.A.R., D.M.M., D.J.G.; Formal analysis, E.B., R.E.O, L.A.R., H.C.; Investigations, E.B., C.Y.H., K.S., A.V.A., L.R.L., J.Z.; Resources, D.J.G., L.B., H.C.; Data curation, E.B., R.E.O, L.A.R.; Writing – original draft, E.B., H.C.; Writing – review & editing, E.B., R.E.O., C.Y.H., A.V.A., L.R.L., D.J.G., L.B., H.C.; Visualization, E.B., H.C.; Supervision, D.J.G., L.B., H.C.; Project Administration, H.C.; Funding acquisition, L.B., H.C.

Declaration of interests

The authors declare no competing interests.

understood. We show that a single sialidase, NanH, in *Bacteroides fragilis* mediates stable occupancy of the intestinal mucosa in early life and regulates a commensal colonization program. This program is triggered by sialylated glycans, including those found in human milk oligosaccharides and intestinal mucus. NanH is required for vertical transmission from dams to pups and promotes *B. fragilis* dominance during early life. Furthermore, NanH facilitates commensal resilience and recovery after antibiotic treatment in a defined microbial community. Collectively, our study reveals a co-evolutionary mechanism between the host and microbiota mediated through host-derived glycans to promote stable colonization.

Graphical Abstract



In brief

Buzun et al. demonstrate that human milk oligosaccharides activate a colonization program in a pioneer gut commensal, *Bacteroides fragilis*. Orchestrated by the *B. fragilis* sialidase, NanH, this program mediates the competitive fitness and resilience of B. fragilis during vertical transmission from dams to pups in the murine gut.

INTRODUCTION

The infant gut microbiota profoundly influences physiological development and overall human health.[1] The composition and dynamics of the establishing infant gut microbiota are strongly shaped by the delivery mode[2-4] and consumption of human milk oligosaccharides

(HMOs).[5,6] Emerging studies have highlighted *Bacteroides* species as pioneer species of the infant microbiome that are detected in the first days of life.[7-9] *Bacteroides* possess an extraordinary repertoire of polysaccharide utilization loci (PULs) that facilitate the breakdown of complex host-derived glycans,[10,11] including HMOs.[12] Importantly, HMOs share structural similarities with intestinal mucus O-glycans,[13] and *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* are known to possess PULs mediating foraging of intestinal mucus to support colonization and persistence in the gut.[14-17] While several studies have shown that *Bacteroides* are capable of degrading HMOs,[12,18] the mechanisms directly linking HMO metabolism to commensal colonization of the infant gut remain undetermined.

Here, we define the HMO utilization profile in *B. fragilis* and employ comparative proteomics to unveil the enzymatic system governing HMO metabolism in *B. fragilis*. We identify the induction of a sialidase (NanH) and the commensal colonization factor (CCF) during *B. fragilis* growth on HMOs. Our findings demonstrate that *B. fragilis* sialidase, NanH, is crucial for stable niche occupancy of the murine gut and colonization during the nursing period in pups. Notably, the role of NanH in colonization extends beyond suckling pups, as mutants lacking NanH sialidase exhibit impaired competitive fitness during co-colonization in adult mice. We demonstrate that in a complex community, *B. fragilis* can persist and recover following antibiotic treatment in a NanH-dependent manner. Our results provide the evidence that host-derived glycans, such as HMOs, act as the primary signal for pioneer commensals to initiate gut colonization during early life, facilitating their persistence in the intestinal mucosa.

RESULTS

Defining the HMO metabolizing system in *Bacteroides fragilis*.

We first assessed the ability of *Bacteroides* species to utilize HMOs as a sole carbon source. Using pooled HMOs (pHMOs) purified from donor human milk, we monitored growth of *B. fragilis* (NCTC 9343), *B. thetaiotaomicron* (VPI-8482), *Phocaeicola vulgatus* (formerly *B. vulgatus*; ATCC 8482), *B. ovatus* (ATCC 8483), *B. salyersiae* (DSM 18765), *B. uniformis* (ATCC 8492), and *B. acidifaciens* (JCM 10556). From these *Bacteroides* species, *B. fragilis, B. thetaiotaomicron*, and *P. vulgatus* displayed robust growth, whereas *B. ovatus, B. salyersiae, B. uniformis*, and *B. acidifaciens* exhibited slow growth (Figure 1A). Thin layer chromatography (TLC) analysis revealed that the composition of products remaining from stationary phase cultures varied among *Bacteroides* species (Figure 1B). *B. fragilis* was able to rapidly deplete available HMOs compared to *B. thetaiotaomicron* and *P. vulgatus* (Figure 1B and Figure S1A-F). To further elucidate the HMO utilization profile in *B. fragilis, B. thetaiotaomicron*, and *P. vulgatus* we used high performance liquid chromatography (HPLC) to separate and identify 19 most abundant HMO structures (Figure S1D). This revealed that *B. fragilis* displayed a selective preference for α2-6-sialylated HMOs (Figure 1C), while *B. thetaiotaomicron*, and *P. vulgatus* exhibited distinct utilization profiles (Figure S1E and S1F). Further, *in vitro* co-cultures on pHMOs between *B. fragilis* and *B. thetaiotaomicron* or *P. vulgatus*, along with other common HMO-utilizers, *Bifidobacterium longum* and *Bifidobacterium breve*, demonstrated that *B. fragilis* possesses a highly efficient strategy for HMO metabolism, outcompeting all bacterial strains examined (Figure S1G-S1J).

We then sought to uncover the metabolic system that enables HMO metabolism in *B. fragilis*. Importantly, *B. fragilis* exhibited a biphasic growth on pHMOs (Figure 1A), during which bacteria shift their metabolism to adapt to environmental changes.[19-21] TLC and HPLC analyses of the culture supernatants from the early-, mid-, and late-log growth phases showed that the substrate diversity changed over time (Figures S1K-N). To define the *B. fragilis* protein repertoire orchestrating HMO metabolism, we performed whole cell proteomics on HMO-grown *B. fragilis* at the three growth phases. In total, 2,808 proteins were identified (Figure 1D, Figure S2A, and Table S1). The number of significantly abundant proteins (>2-log fold change compared to glucose control, $P<0.05$,) varied across the three time points, demonstrating a metabolic shift in *B. fragilis* during consumption of HMOs (Figure 1D and Figure S2A). Notably, a group of proteins was significantly abundant in *B. fragilis* over all three growth phases (Figure 1D and 1E). This included components of PUL 30[22] (hereafter HMO PUL 1) which encodes a sialidase BF1806 (NanH, GH33); two β-N-acetylglucosaminidase, BF1807 and BF1811 (GH20); one β-galactosidase, BF1815 (GH2); one putative β-mannosidase, BF1810 (GH2); and two pairs of SusC and SusD proteins (Figure 1D-F). Additionally, the Nan operon (PUL 20[22]), which mediates conversion of Neu5Ac (sialic acid) cleaved by the sialidase NanH into glucose-6-phosphate for downstream glycolysis,[23,24] was also significantly abundant (Figure 1D-F). We confirmed the upregulated expression of the HMO PUL 1 genes, *bf1802* (SusD), *bf1806* (*nanH*, GH33), and *bf1807* (GH20), in *B. fragilis* during growth on pHMOs (Figure S2B). These findings suggest that the HMO PUL 1 is the central enzymatic system mediating HMO breakdown in *B. fragilis*.

Further analysis demonstrated that proteins comprising the N-glycan utilization (Don) operon[25] (PUL 43[22]) resided in the top 0.5% of the total proteome during early- and mid-log growth phases (Figure 1D-F). As a result of the metabolic shift, the Don operon was displaced with proteins involved in intestinal colonization, including the commensal colonization factor (CCF) locus[15] (Figure 1D-F) and the fucRIOAKP operon[26] (Figure S2C). Both the Don and CCF operons enclose putative proteins, BF3415 (DonF), BF3416 (DonG), and BF3579 (CcfE), predicted to function as a carbohydrate binding modules (CBM) 32 in PUL-DB,[22] indicating a potential role in glycan acquisition. Interestingly, out of 100 most upregulated proteins in the late-log phase, 40% accounted for lone SusC-SusD pairs and proteins containing domains of unknown function (DUFs) predicted to act as glycan binding modules (Figure S2D). We therefore propose *B. fragilis* uses a range of binding proteins to capture extracellular HMOs and perform their metabolism intracellularly, while using the sialidase NanH (BF1806) and other enzymes in HMO PUL 1 as a central enzymatic system for depolymerization of HMOs.

NanH sialidase promotes intestinal colonization in a commensal bacterium.

Our work here (Figure 1) demonstrates that *B. fragilis* upregulates a sialidase, NanH (BF1806), in response to pHMOs, in contrast to *B. thetaiotaomicron* (Figure S2E and Table S2) and *P. vulgatus* (Figure S2E and Table S3). Phylogenetic analysis indicated that homologues of BF1806 can be found in 11 *Bacteroides* species and other commensals (Figure S3A). However, these enzymes are enclosed in distinct operons, sharing low similarity with NanH (BF1806) (Figure S3B and S3C). To determine the functional role of

B. fragilis NanH, we generated an isogenic *B. fragilis* mutant lacking BF1806 (Δ*nanH*). No differences in growth were observed in nutrient-rich media (BHI-S) between *B. fragilis* wild type (WT) and the Δ*nanH* mutant (Figure 2A). However, the ability of the Δ*nanH* strain to utilize pHMOs was impaired (Figure 2B). Notably, 20% of pHMOs contain α2-3/6-Neu5Ac decorations and can be separated into sialylated (acidic) and non-sialylated (neutral) HMOs (Figure 2C). We fractionated pooled human milk samples and demonstrated that *B. fragilis* Δ*nanH* displayed a growth defect in pooled sialylated HMOs (Figure 2D), while its ability to metabolize pooled non-sialylated/neutral HMOs was not attenuated (Figure 2E). Consistently, HPLC analyses of the supernatant after growth revealed that Δ*nanH* was unable to utilize 6'SL and 3'SL and exhibited partial degradation of other HMOs in the heterogeneous pHMOs sample (Figure S3D-G). To investigate linkage specificity of the NanH sialidase, we grew *B. fragilis* WT and Δ*nanH* in singular HMOs, 3'SL and 6'SL, harboring α2-3- and α2- 6-Neu5AC decorations, respectively. This revealed that *B. fragilis* Δ*nanH* was unable to grow on 6'SL (Figure 2F) and exhibited significantly delayed growth kinetics on 3'SL (Figure 2G). Further, we reveal that the sialidase activity of whole and lysed *B. fragilis* Δ*nanH* cells was lost against 6'SL and strongly attenuated against 3'SL (Figure S4A and S4B). Complementation of NanH in the *B. fragilis* Δ*nanH* strain restored growth on both 6'SL and 3'SL (Figure S4C and S4D). Combined, our findings demonstrate that *B. fragilis* NanH is a central sialidase, which enables metabolism of sialylated HMOs.

Gastrointestinal mucin glycans are abundantly sialylated and share high structural similarities with sialylated HMOs.[13] Previous studies reported that *B. fragilis* NanH is highly expressed during growth on porcine gastric mucin (PGM) II,[16] and an orthologue of NanH in *Tannerella forsythia* is able to desialylate salivary mucins along with a range of other glycoproteins.[27,28] Thus, we investigated whether *B. fragilis* NanH sialidase may also play a role in niche occupancy within the intestinal mucosa. Through mono-colonization of germ-free (GF) C57BL/6J mice, we observed a steady decrease in bacterial numbers of *B. fragilis* Δ*nanH*, failing to establish full colonization when compared to *B. fragilis* WT (Figure 2H). Moreover, the abundance of *B. fragilis* Δ*nanH* was significantly reduced in the colonic mucus in comparison to *B. fragilis* WT mono-colonized mice (Figure 2I). Further, the ability of the *B. fragilis* Δ*nanH* to utilize mucus harvested from germ-free mice was impaired *in vitro* (Figure S4E), suggesting that NanH mediates mucus degradation and mucosal occupancy in the gut. To assess the competitive fitness of *B. fragilis* Δ*nanH*, we gavaged germ-free mice with *B. fragilis* WT and Δ*nanH* at 1:1 ratio. *B. fragilis* WT displayed a significant competitive advantage over *B. fragilis* Δ*nanH* at 15 days post-co-colonization (Figure 2J and 2K), indicating that NanH contributes to commensal fitness *in vivo*.

Our proteomics data indicated that pHMOs induce expression of the CCF locus (Figure 1D-F), previously shown to facilitate occupancy of the intestinal crypts.[15] We next analyzed relative expression of the key *ccf* genes in the fecal content and colonic mucosa. This revealed a significant reduction in relative expression of *ccfC, ccfD*, and *ccfE* of *B. fragilis* Δ*nanH* colonized mice compared to *B. fragilis* WT (Figure S4F and S4G), suggesting a functional link between the NanH sialidase and the expression of genes comprising the CCF locus. Although the absence of the CCF did not affect growth on pHMOs, the double mutant Δ*nanH*ΔCCF exhibited a reduced carrying capacity on pHMOs compared to *B. fragilis*

Δ*nanH* (Figure S4H). Combined, these data suggest that the NanH sialidase orchestrates metabolism of sialylated host glycans and governs the commensal colonization program in *B. fragilis*.

B. fragilis NanH sialidase determines stable niche occupancy and colonization during early life.

Previous work demonstrated *B. fragilis* can stably occupy its niche within the intestinal mucosa and concurrently permit colonization by other gut *Bacteroides* (e.g., *B. thetaiotaomicron*), but not isogenic *B. fragilis* strains.[15] Access to the nutritional mucosal niche requires the removal of the terminal sialic acid moieties capping the colonic mucus glycans.[13,29] We explored whether *B. fragilis* requires the sialidase NanH to achieve stable niche occupancy of the intestinal mucosa. Germ-free mice were either first mono-associated with *B. fragilis* WT, followed by a challenge with *B. fragilis* Δ*nanH* on day 8 (Figure 3A), or first mono-associated with *B. fragilis* Δ*nanH* followed by *B. fragilis* WT challenge (Figure 3B). When *B. fragilis* WT served as the initial strain to colonize gnotobiotic mice, the WT strain did not allow for the sequential co-colonization by *B. fragilis* Δ*nanH* and gradually cleared the mutant strain (Figure 3A). In contrast, *B. fragilis* Δ*nanH* strain was permissive to the challenge by *B. fragilis* WT (Figure 3B). Enumeration of the *B. fragilis* WT and Δ*nanH* strains 25 days post-challenge revealed that *B. fragilis* Δ*nanH* was less abundant in the feces (Figure 3C) and was undetectable in the mucus of mice that were first colonized with *B. fragilis* WT (Figure 3D). In contrast, both *B. fragilis* WT and Δ*nanH* were present in the feces and mucus of mice that were initially mono-colonized with *B. fragilis* Δ*nanH* (Figure 3C and 3D). This indicates that NanH sialidase plays a role in directing *B. fragilis* towards its niche in the intestinal mucosa.

In contrast to HMOs that are comprised of heterogeneous structures, mouse milk is primarily composed of lactose and sialylated glycans, 3'SL and 6'SL.[30,31] Building on the observation that *B. fragilis* Δ*nanH* exhibited defective growth on sialylated HMOs (Figure 2D, 2F, and 2G), we explored whether NanH facilitates *B. fragilis* colonization during the suckling period. Germ-free breeder pairs were orally gavaged with *B. fragilis* WT and Δ*nanH* at equal ratio and newborn pups were sacrificed at 7, 10, and 14 days of age (Figure 3E). *B. fragilis* WT efficiently colonized colons of the newborn pups, dominating over *B. fragilis* Δ*nanH* (Figure 3E). Further, TLC and HPLC analyses of the stomach contents confirmed that the pups received a diet primarily composed of 3'SL, 6'SL, and lactose during the first 14 days of life (Figure 3F and 3G). These data link the ability to utilize sialylated HMOs to the commensal competitive fitness in the intestinal mucosa during early life.

NanH mediates resilience and competitive fitness of *B. fragilis*.

A number of commensal bacteria possess sialidases (Figure S3A-C), granting them access to host-derived glycans.[24,32] However, the specific role of sialidases during competitive co-colonization has not been extensively studied *in vivo*. Notably, *B. thetaiotaomicron* expresses a homologue of NanH (BT4055)[12,33] (Figure S3A-C) but lacks the nanLET operon enabling metabolism of sialic acid.[23,34] We assessed the ability of *B. fragilis* WT and Δ*nanH* to share the intestinal niche with *B. thetaiotaomicron* during early life. Germ-free

breeder pairs were co-colonized with *B. fragilis* WT and *B. thetaiotaomicron* (Bf WT:Bt) or *B. fragilis* Δ*nanH* and *B. thetaiotaomicron* (Bf Δ*nanH*:Bt) for 2-3 weeks, and the pups were sacrificed on days 7, 9, and 14. We detected *B. fragilis* WT and *B. thetaiotaomicron* in the colons of the newborns 7 days after birth, demonstrating that both strains are vertically transferred (Figure 4A and 4B). At 9 days old, *B. fragilis* WT was present at a higher proportion than *B. thetaiotaomicon* (Figure 4A), whereas *B. fragilis* Δ*nanH* was unable to compete with *B. thetaiotaomicron*, displaying a fitness defect (Figure 4B). Moreover, *B. fragilis* WT continued to dominate over *B. thetaiotaomicron* in 14 days old pups (Figure 4A), while *B. fragilis* Δ*nanH* was still unable to bloom at comparable levels (Figure 4B). The defect in competitive fitness was retained by *B. fragilis* Δ*nanH* during co-colonization with *B. thetaiotaomicron* of adult gnotobiotic mice, as compared to *B. fragilis* WT (Figure 4C, 4D and Figure S4I). These findings further emphasize the importance of *B. fragilis* NanH sialidase during competitive co-colonization of both newborn and adult mice.

We next explored whether NanH mediates commensal resilience, enabling post-antibiotic recovery in the context of a defined microbial community. We orally gavaged adult mice harboring the Simplified Human Intestinal Microbiota (SIHUMIx)[35] with *B. fragilis* WT and Δ*nanH* at 1:1 ratio. Mice were colonized for 10 days to allow for the stabilization of the *B. fragilis* strains into the complex community. We then introduced a broad-spectrum antibiotic, ciprofloxacin, in drinking water for 5 days, after which the mice were maintained on regular drinking water. We observed *B. fragilis* Δ*nanH* was present at a lower proportion than *B. fragilis* WT prior to the ciprofloxacin treatment (Figure 4E and 4F). While *B. fragilis* WT was able to recover from the ciprofloxacin treatment, *B. fragilis* Δ*nanH* was unable to re-establish colonization and gradually reduced in the lumen (Figure 4E and 4F) and intestinal mucosa (Figure 4G). These results demonstrate that the presence of NanH sialidase is critical for *B. fragilis* to re-establish its intestinal niche and compete within the microbial community after antibiotic perturbation.

B. fragilis is a keystone species that is transferred from mother to infant during birth.[7,9] We isolated *B. fragilis* strains from two human infants, and both strains exhibited growth comparable to *B. fragilis* NCTC 9343 on pHMOs (Figure 4H) and BHI-S (Figure S4J). Further, analysis of over 300 publicly available *B. fragilis* genomes, including strains from pediatric donors, revealed that *nanH* is a core gene, with a high sequence identity (98%) (Table S4). These findings provide compelling evidence that the presence and functionality of NanH sialidase are conserved among *B. fragilis* strains, underscoring its importance in the colonization dynamics of this pioneer *Bacteroides* species. Altogether, we demonstrate that sialylated HMOs received pre-weaning direct *B. fragilis* towards its intestinal niche, ensuring a stable and resilient colonization, revealing the intricate interplay between HMOs and microbial colonization during early life.

DISCUSSION

Here, we report that HMOs from human milk play a key role in establishing intestinal colonization of commensal bacteria. Several studies have indicated a positive association between breastfeeding and the successful establishment of *Bacteroides* species in the infant gut.[6,36] We unveil the molecular mechanism of HMO metabolism in *B. fragilis* and show

that defects in the sialidase, NanH, result in the impaired growth on HMOs and reduced intestinal colonization in early life, indicating that these processes are intricately linked.

Based on our findings, we speculate whether *Bacteroides* species are maintained in breastfed infants due to the induction of bacterial systems associated with stable intestinal colonization, including *B. fragilis* NanH and CCF. Recent work has reported a correlation between reduced levels of a sialylated milk oligosaccharide (DSLNT) and the development of inflammatory disorders such as necrotizing enterocolitis breastfed infants.[37,38] Further, the analysis of HMOs from Malawi mothers revealed that lower levels of sialylated HMOs was associated with stunted development.[39] Transplantation of the fecal microbiota of malnourished infants into germ-free mice maintained on bovine sialylated oligosaccharides improved growth and development.[39] Indeed, in this study *B. fragilis* exhibited the most prominent transcriptional response to sialylated oligosaccharides by upregulating genes composing the Nan operon and a SusC/D pair (BF1804/BF1805)[39] from HMO PUL 1, adjacent to NanH (BF1806) described in our study (Figure 1F). Another prominent *Bacteroides, B. thetaiotaomicron*, possesses a GH33 sialidase, enclosed in a PUL homologous to HMO PUL 1, shown to mediate removal of glycosidic decorations shared between complex N-glycans and HMOs.[33] This suggests that, in addition to HMOs and mucus glycans, enzymes in *B. fragilis* HMO PUL 1 might provide access to a variety of host glycoproteins. Further, a metagenomic study revealed an enrichment of bacterial sialidases in human infants younger than 6 months,[40] suggesting that bacterial sialidases may play a key role in the assembly of the infant gut microbiota.

We demonstrate that the genes induced by HMOs confer *B. fragilis* a competitive advantage during early life and the involvement of NanH-mediated colonization in the recovery and resilience of *B. fragilis* following antibiotic treatment. Our findings suggest that the mammalian host curates its indigenous microbiota by providing specific substrates, such as HMOs, within a developmental window to ensure stable colonization of beneficial microbes. Altogether, we propose a model where sialylated HMOs act as the signal for the pioneer gut commensal to occupy the mucosal niche at early life, after which structurally similar mucin glycans maintain a bacterial reservoir, supporting commensal persistence and resilience after weaning.

LIMITATIONS

Here, we demonstrate a key role of *B. fragilis* NanH sialidase during early life colonization in gnotobiotic mouse models. Importantly, the diversity of mouse milk oligosaccharides is significantly reduced compared to the heterogeneous pool of HMOs present in human milk. The abundance of NanH in response to pHMOs was identified *in vitro*, using HMOs pooled from multiple donors. However, the overall transcriptional response in *B. fragilis* would be shaped by multiple host factors and the induction of NanH during early life colonization of human infants remains to be fully explored. Further, our experiments assessing vertical transmission and commensal resilience were performed in a controlled environment using mouse models and a limited number of bacterial species, while human infants are exposed to a wide range of microorganisms, extensively competing for the intestinal niche in the newborn infant. We demonstrate that the growth kinetics of the *B. fragilis* strains isolated

from human infants on pHMOs is similar to the laboratory type strain NCTC 9343, however due to the absence of the metadata associated with the donor samples, we are unable to conclude whether this phenotype could be dependent on the age, gender, race, or the socioeconomic background of the donors. Combined, we demonstrate that NanH sialidase is important for early life colonization and commensal resilience in the context of a defined community, however the role of NanH sialidase in assembly of the human infant gut microbiota warrants further investigation.

STAR ★ METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Hiutung Chu hiuchu@ucsd.edu.

Materials and availability—All unique resources generated in this study may be available from the lead contact under a material transfer agreement.

Data Code Availability

- Raw whole cell proteomic data are publicly available online at https://massive.ucsd.edu under study ID MSV000090386 as of the date of publication. https://github.com/rolesucsd/hmo_proteomics
- Analyzed whole cell proteomics data for *B. fragilis, B. thetaiotaomicron*, and *P. vulgatus* are available in Table S1, Table S2, and Table S3, respectively.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL DETAILS

Murine models.—4-week-old sex-matched germ-free C57BL/6J mice (JAX: 000664) were housed in sterile, sealed positive pressure cages with double HEPA filtration (Allentown) with autoclaved chow (Lab Diet 5010), sterile drinking water, and sterile aspen wood chip bedding (Newco). Environmental conditions were maintained at 21 °C, 15 % humidity, and 12/12-hour light/dark cycle. Health checks were performed daily. All procedures were performed in accordance with the guidelines and approved protocols from the IACUC of UC San Diego. For additional study details, use “methods details” section.

Bacterial strains and growth conditions—Bacterial strains used in this study are listed in the Key Resource Table. Type strains (ATCC) were used for growth experiments unless otherwise stated. *Bacteroides* strains and *Bififdobacterium spp*. were grown in BHI-S (Brain Heart Infusion, BD) medium supplemented with 0.5 % hemin and 0.5 μg/ml vitamin K (Sigma Aldrich) for 16 h at 37 °C under anaerobic conditions (10 % H_2, 10 % CO_2, 80 % N_2; Coy Lab Products). For growths utilizing a sole carbon source, ZMB1 (Zhang-Millis-Block1)[70] defined medium was supplemented with 0.5 % hemin and 0.005 % yeast extract (Sigma Aldrich). Cultures were then normalized to the same OD_{600} inoculated at 2.5 %

(v/v) into ZMB1, containing 15 mg/ml pooled mixed, pooled acidic, pooled neutral HMOs or 5 mg/ml single HMOs (6'SL and 3'SL) as a sole carbon source. Where appropriate gentamicin (200 μg/ml), erythromycin (10 μg/ml), ampicillin (100 μg/ml), chloramphenicol (10 μg/ml), or tetracycline (6 μg/ml) were used for bacterial selection.

Human samples—Human samples were obtained from participants who provided a written consent. Procedures were reviewed and approved by UC San Diego's Institutional Review Board. Fecal samples were collected from human volunteers between 7 and 42 days old. To enrich for *Bacteroides spp*. fecal materials were plated on BHI-S agar (0.05% hemin and 0.5 μg/ml vitamin K (Sigma Aldrich), containing 100 μg/ml of gentamicin, enrichment for *Bifidobacterium spp*. was performed on Gifu Anaerobic Medium (GAM)-S agar (0.05% hemin and 0.5 μg/ml vitamin K; Sigma Aldrich). Plates were incubated in the in the anaerobic conditions (10 % H_2, 10 % CO_2, 80 % N_2; Coy Lab Products) at 37 °C for 2 days. Individual colonies were then picked and the species identity was confirmed by 16S sequencing (EtonBio) from a PCR product amplified with primers provided in Table S5. An overnight culture was grown from a single colony and stored in 25 % glycerol in −80 °C.

METHOD DETAILS

Generation of bacterial growth curves.—Bacterial strains and media were prepared and maintained as described in the "experimental model details" section. Growth assays were performed in triplicates in a 96-well plate and monitored by plate reader (Biotek Synergy HT) under anaerobic conditions at 37 °C, reading OD_{600} at regular intervals. All growth curves are representative of at least three independent experiments and were analyzed with GraphPad Prism 9.0 or higher.

To assess growth on mucus, mucosal scrapings were collected from the colons of germ-free adult mice and resuspended in 1 ml of sterile PBS. The mucus was homogenized with sterile stainless-steel beads (ø3.2mm) using a bullet blender (Next Advance) for 3 mins at speed 8, then diluted with ZMB1 to 18 mg/ml to be used for growth experiments in culture tubes. Ten-fold serial dilutions were plated on BHI-agar (0.05% hemin and 0.5 μg/ml vitamin K (Sigma Aldrich)) and incubated under anaerobic conditions at 37 °C for at least 36 hours. Bacterial growth was assessed by colony forming units per ml of culture (CFU/ml).

Generation of deletion mutants, complementation, and pFD340-conjugates—Deletion of *bf*1806 (*nanH*) was generated by amplifying 1 kb fragments upstream and downstream of the region (Table S5) using NEBNext High Fidelity Master Mix (NEB), then cloned into BamHI-SalI digested pKNOCK-bla-erm[45] using NEBuilder (NEB). The plasmid was conjugally transferred into *B. fragilis* NCTC9343 or *B. fragilis* ΔCCF[15] using *E. coli* S-17 *λ-pir*. Conjugates were selected based on erythromycin resistance and the second recombination event was encouraged by daily passage until erythromycin resistance was lost. Scarless in-frame deletions were confirmed by PCR using primers listed in Table S5 and sequencing (Primordium). Complementation of *nanH* was performed in trans using pFD340.[46] PCR amplified *nanH* was ligated into BamHI-SacI digested pFD340 and expressed under control of the IS4351 promoter.

To create antibiotic resistant strains used for colonization experiments, *E. coli* S-17 *λ-pir* were first transformed with pFD340-*tetQ* (tetracycline resistant, TetR) and pFD340-*cat* (chloramphenicol resistant, CmR) plasmids.[15] These were then conjugally transferred to generate *B. fragilis* WT+pFD340-*tetQ, B. fragilis* Δ*nanH*+pFD340-*cat*, *B. thetaiotaomicron*+pFD340-*tetQ*, *B. thetaiotaomicron*+pFD340-*cat*. The conjugates were selected based on erythromycin (10 mg/ml) resistance and secondary resistance to tetracycline (6 mg/ml) or chloramphenicol (10 mg/ml). During *in vitro* culturing, erythromycin was added to rich media (BHI-S) to ensure plasmid maintenance.

HMO isolation from pooled donor human milk—Pooled HMOs (pHMOs) were isolated from human milk pooled from at least 5 different donors as previously described.[71] After centrifugation, the lipid layer was removed, and proteins were precipitated from the aqueous phase by addition of ice-cold ethanol and subsequent centrifugation. Ethanol was removed from the HMO-containing supernatant by roto-evaporation. Lactose and salts were removed by gel filtration chromatography over a BioRad P2 column (100 cm x 316 mm, Bio-Rad) using a semi-automated fast protein liquid chromatography (FPLC) system. pHMO composition was measured, as described below. Only pHMOs with less than 2% lactose were used for bacterial growth experiments.

HMO analysis—Analysis of HMOs in defined media was performed by high-performance liquid chromatography (HPLC) with fluorescence detection, as previously described for human milk[72] with slight modifications: 10 μL of media was spiked with 12 ng/ml of maltose as an internal standard, lyophilized, and directly labeled with the fluorophore 2-aminobenzamide (2AB). 2AB-labeled HMOs were separated on a TSKgel Amide-80 column (2.0 mm ID x 15 cm, 3m, Tosoh Bioscience) and detected at 360 nm excitation and 425 nm emission. HMO peaks were annotated based on standard retention times and quantified in reference to the internal standard. HMO utilization at indicated timepoints was calculated in reference to HMO concentrations in the media at the beginning of each experiment (t=0).

Thin layer chromatography—Aliquots of 3-12 μl of cell-free supernatant or enzymatic reactions were spotted onto silica coated plates (TLC Silica gel 60 F254, Merck), dried, and resolved in butanol:acetic acid:water (2:1:1). When the mobile phase reached to top of the TLC plate, the plate was dried and resolved again. To visualize sugars the TLC plate was first dried, coated with diphenylamine (DPA) stain,[73] and then heated with a heat gun until the bands were visible. 1 mM of each glucose, galactose, fucose, lactose, sialic acid (Neu5AC) controls were run alongside samples and used for reference.

Experiments with whole bacterial cells—*B. fragilis* WT and Δ*nanH* were grown to mid-exponential phase in BHI-S, washed two times with ZMB1 media and inoculated into a defined ZMB1 medium containing 1.5 % pHMOs. 1 ml of culture was collected from the mid-exponential growth, washed two times with sterile PBS, and resuspended in 1 ml of PBS. 0.5 ml of cells was then pelleted for 1 min at 13,000 x g, resuspended in 0.5 ml of BugBuster (Millipore), and incubated for 5 mins at room temperature to lyse the cells. Cell debris were removed by centrifugation. Whole cells and the cell lysates were then incubated

with 2.5 mg/ml of 6'SL or 3'SL at 37 °C. The aliquots were collected at 0, 2, 4, and 16 hours, the reactions were stopped by boiling at 98 °C for 5 mins and centrifuged for 1 min at 13,000 x g. The supernatants were then analyzed by TLC.

Experiments with gnotobiotic mice—For mono-colonization experiments, 4-week-old mice were orally gavaged with *B. fragilis* wild type (WT) and Δ*nanH* (8 x 10^7 CFUs). For co-colonization experiments, 4-week-old mice were orally gavaged with bacterial suspension containing equal proportions of *B. fragilis* WT and the isogenic Bf Δ*nanH* mutant, *B. fragilis* WT and *B. thetaiotaomicron* WT, or *B. fragilis* Δ*nanH* and *B. thetaiotaomicron* WT (~8 x 10^7 CFUs of each strain). For the sequential colonization experiments, 4–5-week-old germ-free C57BL/6J mice were mono-colonized with the initial strain (8 x 10^7 CFUs) for 7 days, and the challenge strain was introduced by oral gavage on day 8. All strains used for colonization experiments retained pFD340 plasmids, conferring resistance to erythromycin and either tetracycline or chloramphenicol.[15] Gentamicin (100 μg/ml) and erythromycin (10 μg/ml) were added to the drinking water to ensure plasmid maintenance. At each time point, a fresh fecal pellet was collected, resuspended in sterile PBS, 10-fold serially diluted, and plated on BHI-S agar containing either tetracycline (6 μg/ml) or chloramphenicol (10 μg/ml) to allow for differentiation between *Bacteroides species*. Abundance of each strain was enumerated by CFU per milligram of feces.

Experiments in mice colonized with a defined microbial community.—Gnotobiotic C57BL/6J mice were colonized with the Simplified Human Intestinal Microbiota (SIHUMIx)[35] composed of seven bacterial species: *Anaerostipes caccae* (DSMZ 14662), *Bifidobacterium longum* (NCC2705), *Blautia producta* (DSMZ 2950), *Clostridium butyricum* (DSMZ 10702), *Clostridium ramosum* (DSMZ 1402), *Escherichia coli* K-12 (MG1665), *Lactobacillus plantarum* (DSMZ 20174) and bred in gnotobiotic isolators. Age- and sex- matched 5-6 weeks old mice were then inoculated with a single gavage of bacteria mixture of WT and Δ*nanH B. fragilis* (8 x 10^7 CFUs of each strain) and the strains were allowed 10 days for the integration into the community. On day 10, mice were then treated with 0.625 mg/ml of ciprofloxacin (Sigma Aldrich) in drinking water for 5 days, after which they were switched back and maintained on regular autoclaved water for 13 days. At each time point, bacterial CFUs were determined by plating on selective media. In addition, total microbial genomic DNA was extracted using Quick DNA Fecal/soil Microbe miniprep kit (Zymo Research) as per manufacturer's protocol. The relative abundance of *B. fragilis* WT and Δ*nanH* was assessed by absolute qPCR using strain-specific and universal 16S primers provided in Table S5.

Vertical transfer experiments and colonization of suckling pups—Breeding pairs of 8-week-old male and female germ-free *Rag1*−/− (JAX:008449) mice were orally gavaged with a bacterial culture comprising of *B. fragilis* WT and Δ*nanH, B. fragilis* WT and *B. thetaiotaomicron* WT, or *B. fragilis* Δ*nanH* and *B. thetaiotaomicron* WT (8 x 10^7 CFUs of each strain). The breeders were housed in sterile, sealed positive-pressure cages with double HEPA filtration (Allentown), with autoclaved chow (Lab Diet 5010) and drinking water containing gentamicin (100 mg/ml) and erythromycin (10 mg/ml). Upon delivery, the pups were sacrificed at 7, 9-10, and 14 days of age. Colons were excised, and the contents

were resuspended in 300 μl of sterile PBS in pre-weighed tubes. The contents were then homogenized with sterile stainless-steel beads (ø3.2mm) in a bullet blender (Next Advance) for 3 mins at speed 8. The homogenized colonic material was then 10- fold serially diluted and plated on either tetracycline or chloramphenicol containing BHI-S plates to allow for differentiation between the competing species. CFUs per mg were calculated from colony counts for each pup and presented as a ratio relative to the breeders.

Stomach processing for analysis by TLC and HPLC—The stomachs were collected from pups of 0-1, 2, 3, and 4 weeks of age. Images were taken prior to placing the stomachs in 1 ml of sterile PBS. The tissue was homogenized using stainless-steel beads (ø3.2mm) with a bullet blender (Next Advance) for 5 mins at speed 10. Homogenized tissue was then centrifuged at 13,000 x g for 3 mins, the supernatant containing soluble glycans was separated and used for the downstream analysis by TLC or HPLC.

Mucosal scraping for bacterial quantification—Colons were excised and washed with sterile ice-cold PBS to remove luminal contents. The colons were then cut open longitudinally, washed until fecal material was no longer visible, and the mucus was scraped with light pressure using the blunt side of the sterile tweezers. Collected mucus was placed in pre-weighed tubes, containing 500 μl of sterile PBS, and homogenized using sterile stainless-steel beads (ø3.2mm) with a bullet blender (Next Advance) for 3 mins at speed 8. Homogenized mucus was then 10-fold serially diluted and plated on BHI-S with appropriate antibiotic for enumeration by colony counts.

RNA extraction, cDNA synthesis and relative gene expression—Overnight culture of *B. fragilis* was inoculated into a defined medium containing 15 mg/ml pHMOs and 10 mg/ml glucose (Sigma Aldrich). Bacterial pellets were collected from mid-exponential phase (OD_{600}=0.6-1.0), resuspended in RNA later (Thermo Scientific), and stored at −80 °C. For the determination of the transcript levels during mono-colonization, colonic contents or mucosal scrapings were placed in RNA later immediately after collection and stored at −80 °C. RNA was extracted from bacterial pellets, fecal material, or the mucosal lining using a NucleoSpin RNA isolation kit (Macherey-Nagel), 0.5-1 μg was reverse transcribed into cDNA (Superscript IV VILO Master Mix, Thermo Scientific), as per manufacturer's instructions, and diluted 1:5. Transcription levels were assessed by qPCR (SYBR green qPCR Master Mix, Life Technologies) in Quantstudio 5 (Life Technologies), using standard conditions: initial denaturation at 95 °C for 10 min, 40 cycles of 95 °C for 15 sec, 60 °C for 1 min, 72 °C for 1 min, and an interval increasing temperature from 60 °C to 95 °C to generate a melting curve. To determine relative expression *in vivo*, the data were normalized to *gyraseB* using primers listed in Table S5. For *in vitro* experiments, relative expression was normalized to glucose-grown cells.

Quantification by absolute qPCR—For *in vitro* competition experiments, the presence of each bacterial species was determined by qPCR (SYBR green qPCR Master Mix, Life Technologies) from total genomic DNA extracted from culture pellets using the NucleoSpin Microbial DNA kit (Macherey Nagel). For the determination of bacterial abundance in *in vivo* experiments, total microbial DNA was extracted from fecal or mucosal material using

the Quick DNA Fecal/Soil Microbe Miniprep Kit (Zymo research). Each gDNA sample was diluted to 20 ng/μl and the relative abundance of each species was determined from standard curves prepared from gDNA extracted from pure overnight cultures (range: 20 ng, 10 ng, 1 ng, 0.1, ng, 0.01 ng, 0.001 ng) or a fecal pellet collected from SIHUMIx colonized mice without *B. fragilis* for bacterial 16S (range: 80 ng, 45 ng, 20 ng, 10 ng, 1 ng, 0.1 ng, 0.001 ng). Data was analyzed in Quantstudio 5 (Life Technologies), using thermocycling conditions: initial denaturation at 95 °C for 10 mins, 40 cycles of 95 °C for 15 sec, 60 °C for 1 min, 72 °C for 1 min, and an interval increasing temperature from 60 °C to 95 °C to generate a melting curve. Species-specific primers are listed in Table S5.

Quantitative Multiplex and Label-Free Proteomics—Peptide preparation: Overnight cultures of *B. fragilis* NCTC 9343, *B. thetaiotaomicron* VPI5482, and *P. vulgatus* ATCC8482 were inoculated into ZMB1 media containing 15 mg/ml pHMOs or 10 mg/ml glucose as a sole carbon source. All cultures were performed in triplicates. 4 ml of culture was collected at early-exponential (OD_{600} = 0.5), mid-exponential (OD_{600} = 1.0), and late-exponential (OD_{600} = 1.5) growth phases, pellet was washed three times in sterile PBS and stored at −20 °C prior to analysis. Pellets were resuspended in lysis buffer (6M urea, 7% SDS, 50 mM TEAB, titrated to pH 8.1 with phosphoric acid) with protease and phosphatase inhibitors added (Roche, CO-RO and PHOSS-RO) and then sonicated. Proteins were reduced, alkylated, then trapped using ProtiFi S-Trap columns (ProtiFi, C02-mini), digested with sequencing-grade trypsin (Promega, V5113), and eluted according to manufacturer protocols. Eluents were desalted on SepPak C18 columns (WAT054960, Waters). Peptides were quantified using a Pierce Colorimetric Peptide Quantification Assay Kit (23275, Thermo Scientific). 50 μg of each sample was separated for multiplex proteomic analysis, with several samples aliquoted twice to serve as technical duplicates and fill multiplex channels.

Labeling and fractionation—The labeling scheme for multiplex experiment is included in the Supplemental Material. Samples were labeled using Tandem Mass Tag (TMT) 16-plex reagents (A44520; lot number XA341491, Thermo Scientific) following manufacturer protocol, then combined into a single multiplex. The plex was desalted on SepPak C18 columns and dried under a vacuum. The plex was then fractionated using reverse phase high pH liquid chromatography on a 10-40 % acetonitrile (ACN) gradient to increase sequencing depth, as previously described.[74] The resulting 96 fractions were concatenated into 24 fractions by combining alternating wells within each column, and 12 alternating fractions were used for mass spectrometry analysis.[75]

LC-MS/MS—One μg of each fraction was loaded and analyzed on an Orbitrap Fusion Tribrid mass spectrometer with an in-line Easy-nLC 1000 System and an in-house pulled and packed column, as previously described.[75] Peptides were eluted after loading using a gradient ranging from 6 % to 25 % ACN with 0.125 % formic acid over 165 minutes at a flow rate of 300 nL/min. Data were acquired in data-dependent mode with polarity set to positive. MS1 spectra were acquired in the Orbitrap with a scan range of 500-1200 m/z and a mass resolution of 120,000. Ions selected for MS2 analysis were isolated in the quadrupole and detected in the ion trap. MS2 ions were fragmented with high-energy collision–induced

dissociation, and MS3 fragment ions were analyzed in the Orbitrap. All data acquired were centroided.

Data processing and normalization—Raw files were processed using Proteome Discoverer 2.5. Using the SEQUEST algorithm, MS2 data from *B. fragilis* NCTC 9343, *B. thetaiotaomicron* VPI 5482, and *P. vulgatus* ATCC 8482 were queried against Uniprot proteome UP000006731, UP000001414, UP000002861, respectively, downloaded May 2022. Resulting peptide spectral matches were filtered at a 0.01 FDR by the Percolator module against a decoy database. Peptide spectral matches were quantified using MS3 ion intensities, exported, then summed to the protein level. Entire plex was then batch corrected in a multistep process, as previously described.[76]

Bioinformatics analysis—Glycoside hydrolases and PUL boundaries were identified using Cazy database and PUL-DB.[22] Modular organization of proteins was assessed in InterPro.[67] Sequence alignments and percent identities were determined in Clustal Omega.[68] The nature of the N-terminal signal peptides was identified in SignalP 5.0.[69] The phylogenetic tree for NanH was created by running BlastP against the non-redundant protein database and selecting the top hit for the species of interest. The phylogenetic tree was created by Neighbor-Joining and plotted in R.

QUANTIFICATION AND STATISTICAL ANALYSIS

Mouse experiments were performed in accordance with the guidelines from IACUC UC San Diego, using 5 animals per cage. The number of animals used in one experiment (n) is specified in the figure legends, the data are shown as mean ± SD or SEM and are representative of at least two independent experiments. For the vertical transfer experiments the number of pups was pooled from at least three litters from independent breeders, $n \geq 4$ per age group, the exact number of pups is specified in the figure legends. Statistical analysis was performed in GraphPad Prism version 9.0 or higher, non-parametric Mann-Whitney test was used for multiple comparison analyses and unpaired Student's t-test was used for comparison of two groups, $P < 0.05$ was considered as significant. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns – not significant.

The normalized differential protein abundance testing for the whole cell proteomic datasets were performed in R. A Welch's t-test was conducted for each protein between the given time point and the glucose set to find proteins with change in differential abundance in relation to the control. Benjamini-Hochberg correction for multiple testing was applied to the resulting P-values. The data were normalized by $\log_2$ transformation and technical duplicates were averaged. Proteins were considered significant if the corrected P-value ≤ 0.05 and the $\log_2$-fold change ≥ 2. This data was used to create volcano plots in Prism Version 9 (GraphPad). The heatmap showing significant protein hits relative to glucose was analyzed with ANOVA test between all timepoints ($\log_{10}$-fold change ≥ 2, Benjamini-Hochberg corrected P-value ≤ 0.05) and plotted in R.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of the Chu lab for technical support and helpful discussions. We also thank G. Donaldson and A. Khosravi for valuable feedback and discussions. We thank the Collaborative Center for Multiplexed Proteomics and the La Jolla Institute for Immunology Microscopy and Histology Core Facility. NIH S10 OD021831 funded the Zeiss LSM 880. This work was supported by grants from the National Institute of Health (NIH) R00 DK110534, R01 AI167860, and P30 DK120515, and a Seed grant made available through the UC San Diego Larsson-Rosenquist Foundation Mother-Milk-Infant Center of Research Excellence to H.C. L.B. is UC San Diego Chair of Collaborative Human Milk Research endowed by the Family Larsson-Rosenquist Foundation (FLRF), Switzerland. Additional support was provided to H.C. by the Chiba University-UC San Diego Center for Mucosal Immunology, Allergy and Vaccines (cMAV), CIFAR Humans and the Microbiome Program, The Hartwell Foundation, and AMED (JP233fa627003). Support to D.M.M. was provided by R01 DK131005. Support to L.R. was provided by the UCSD Graduate Training Program in Cellular and Molecular Pharmacology through an institutional training grant from the National Institute of General Medical Sciences, T32 GM007752.

REFERENCES

1. Gensollen T, Iyer SS, Kasper DL, and Blumberg RS (2016). How colonization by microbiota in early life shapes the immune system. Science 352, 539–544. 10.1126/science.aad9378. [PubMed: 27126036]
2. Lou YC, Olm MR, Diamond S, Crits-Christoph A, Firek BA, Baker R, Morowitz MJ, and Banfield JF (2021). Infant gut strain persistence is associated with maternal origin, phylogeny, and traits including surface adhesion and iron acquisition. Cell Reports Medicine 2, 100393. 10.1016/j.xcrm.2021.100393. [PubMed: 34622230]
3. Mitchell CM, Mazzoni C, Hogstrom L, Bryant A, Bergerat A, Cher A, Pochan S, Herman P, Carrigan M, Sharp K, et al. (2020). Delivery Mode Affects Stability of Early Infant Gut Microbiota. Cell Reports Medicine 1, 100156. 10.1016/j.xcrm.2020.100156. [PubMed: 33377127]
4. Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, Ross MC, Lloyd RE, Doddapaneni H, Metcalf GA, et al. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature 562, 583–588. 10.1038/s41586-018-0617-x. [PubMed: 30356187]
5. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, et al. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host & Microbe 17, 690–703. 10.1016/j.chom.2015.04.004. [PubMed: 25974306]
6. Moossavi S, Atakora F, Miliku K, Sepehri S, Robertson B, Duan QL, Becker AB, Mandhane PJ, Turvey SE, Moraes TJ, et al. (2019). Integrated Analysis of Human Milk Microbiota With Oligosaccharides and Fatty Acids in the CHILD Cohort. Front Nutr 6, 58. 10.3389/fnut.2019.00058. [PubMed: 31157227]
7. Yassour M, Jason E, Hogstrom LJ, Arthur TD, Tripathi S, Siljander H, Selvenius J, Oikarinen S, Hyöty H, Virtanen SM, et al. (2018). Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. Cell Host & Microbe 24, 146–154.e4. 10.1016/j.chom.2018.06.007. [PubMed: 30001517]
8. Ferretti P, Pasolli E, Tett A, Asnicar F, Gorfer V, Fedi S, Armanini F, Truong DT, Manara S, Zolfo M, et al. (2018). Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. Cell Host Microbe 24, 133–145.e5. 10.1016/j.chom.2018.06.005. [PubMed: 30001516]
9. Valles-Colomer M, Blanco-Míguez A, Manghi P, Asnicar F, Dubois L, Golzato D, Armanini F, Cumbo F, Huang KD, Manara S, et al. (2023). The person-to-person transmission landscape of the gut and oral microbiomes. Nature 614, 125–135. 10.1038/s41586-022-05620-1. [PubMed: 36653448]
10. La Rosa SL, Ostrowski MP, Vera-Ponce de León A, McKee LS, Larsbrink J, Eijsink VG, Lowe EC, Martens EC, and Pope PB (2022). Glycan processing in gut microbiomes. Current Opinion in Microbiology 67, 102143. 10.1016/j.mib.2022.102143. [PubMed: 35338908]
11. Lapébie P, Lombard V, Drula E, Terrapon N, and Henrissat B (2019). Bacteroidetes use thousands of enzyme combinations to break down glycans. Nat Commun 10, 2043. 10.1038/S41467-019-10068-5. [PubMed: 31053724]

12. Marcobal A, Barboza M, Sonnenburg ED, Pudlo N, Martens EC, Desai P, Lebrilla CB, Weimer BC, Mills DA, German JB, et al. (2011). Bacteroides in the Infant Gut Consume Milk Oligosaccharides via Mucus-Utilization Pathways. Cell Host Microbe 10, 507–514. 10.1016/j.chom.2011.10.007. [PubMed: 22036470]

13. Luis AS, and Hansson GC (2023). Intestinal mucus and their glycans: A habitat for thriving microbiota. Cell Host & Microbe 31, 1087–1100. 10.1016/j.chom.2023.05.026. [PubMed: 37442097]

14. Martens EC, Chiang HC, and Gordon JI (2008). Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. Cell Host Microbe 4, 447–457. 10.1016/j.chom.2008.09.007. [PubMed: 18996345]

15. Lee SM, Donaldson GP, Mikulski Z, Boyajian S, Ley K, and Mazmanian SK (2013). Bacterial colonization factors control specificity and stability of the gut microbiota. Nature 501, 426–429. 10.1038/nature12447. [PubMed: 23955152]

16. Pudlo NA, Urs K, Kumar SS, German JB, Mills DA, and Martens EC (2015). Symbiotic Human Gut Bacteria with Variable Metabolic Priorities for Host Mucosal Glycans. mBio 6, e01282–15. 10.1128/mBio.01282-15. [PubMed: 26556271]

17. Donaldson GP, Ladinsky MS, Yu KB, Sanders JG, Yoo BB, Chou W-C, Conner ME, Earl AM, Knight R, Bjorkman PJ, et al. (2018). Gut microbiota utilize immunoglobulin A for mucosal colonization. Science 360, 795–800. 10.1126/science.aaq0926. [PubMed: 29724905]

18. Kijner S, Cher A, and Yassour M (2022). The Infant Gut Commensal Bacteroides dorei Presents a Generalized Transcriptional Response to Various Human Milk Oligosaccharides. Frontiers in Cellular and Infection Microbiology 12.

19. Loomis WF, and Magasanik B (1967). Glucose-Lactose Diauxie in Escherichia coli. Journal of Bacteriology 93, 1397–1401. 10.1128/jb.93.4.1397-1401.1967. [PubMed: 5340309]

20. Chang D-E, Smalley DJ, and Conway T (2002). Gene expression profiling of Escherichia coli growth transitions: an expanded stringent response model. Molecular Microbiology 45, 289–306. 10.1046/j.1365-2958.2002.03001.x. [PubMed: 12123445]

21. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, and Alon U (2006). A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat Methods 3, 623–628. 10.1038/nmeth895. [PubMed: 16862137]

22. Terrapon N, Lombard V, Drula É, Lapébie P, Al-Masaudi S, Gilbert HJ, and Henrissat B (2018). PULDB: the expanded database of Polysaccharide Utilization Loci. Nucleic Acids Research 46, D677–D683. 10.1093/nar/gkx1022. [PubMed: 29088389]

23. Brigham C, Caughlan R, Gallegos R, Dallas MB, Godoy VG, and Malamy MH (2009). Sialic Acid (N-Acetyl Neuraminic Acid) Utilization by Bacteroides fragilis Requires a Novel N-Acetyl Mannosamine Epimerase. J Bacteriol 191, 3629–3638. 10.1128/JB.00811-08. [PubMed: 19304853]

24. Almagro-Moreno S, and Boyd EF (2010). Bacterial catabolism of nonulosonic (sialic) acid and fitness in the gut. Gut Microbes 1, 45–50. 10.4161/gmic.1.1.10386. [PubMed: 21327116]

25. Cao Y, Rocha ER, and Smith CJ (2014). Efficient utilization of complex N-linked glycans is a selective advantage for Bacteroides fragilis in extraintestinal infections. Proceedings of the National Academy of Sciences 111, 12901–12906. 10.1073/pnas.1407344111.

26. Hooper LV, Xu J, Falk PG, Midtvedt T, and Gordon JI (1999). A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc Natl Acad Sci U S A 96, 9833–9838. [PubMed: 10449780]

27. Satur MJ, Urbanowicz PA, Spencer DIR, Rafferty J, and Stafford GP (2022). Structural and functional characterisation of a stable, broad-specificity multimeric sialidase from the oral pathogen Tannerella forsythia. Biochem J 479, 1785–1806. 10.1042/BCJ20220244. [PubMed: 35916484]

28. Thompson H, Homer KA, Rao S, Booth V, and Hosie AHF (2009). An Orthologue of Bacteroides fragilis NanH Is the Principal Sialidase in Tannerella forsythia. J Bacteriol 191, 3623–3628. 10.1128/JB.01618-08. [PubMed: 19304852]

29. Holmén Larsson JM, Thomsson KA, Rodríguez-Piñeiro AM, Karlsson H, and Hansson GC (2013). Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and

Muc2 mucin O-glycan patterns reveal a regiospecific distribution. Am J Physiol Gastrointest Liver Physiol 305, G357–363. 10.1152/ajpgi.00048.2013. [PubMed: 23832516]

30. Fuhrer A, Sprenger N, Kurakevich E, Borsig L, Chassard C, and Hennet T (2010). Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. Journal of Experimental Medicine 207, 2843–2854. 10.1084/jem.20101098. [PubMed: 21098096]
31. Li J, Jiang M, Zhou J, Ding J, Guo Z, Li M, Ding F, Chai W, Yan J, and Liang X (2021). Characterization of rat and mouse acidic milk oligosaccharides based on hydrophilic interaction chromatography coupled with electrospray tandem mass spectrometry. Carbohydrate Polymers 259, 117734. 10.1016/j.carbpol.2021.117734. [PubMed: 33673995]
32. Tailford LE, Owen CD, Walshaw J, Crost EH, Hardy-Goddard J, Le Gall G, de Vos WM, Taylor GL, and Juge N (2015). Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. Nat Commun 6, 7624. 10.1038/ncomms8624. [PubMed: 26154892]
33. Briliūtė J, Urbanowicz PA, Luis AS, Baslé A, Paterson N, Rebello O, Hendel J, Ndeh DA, Lowe EC, Martens EC, et al. (2019). Complex N-glycan breakdown by gut Bacteroides involves an extensive enzymatic apparatus encoded by multiple co-regulated genetic loci. Nat Microbiol 4, 1571–1581. 10.1038/s41564-019-0466-x. [PubMed: 31160824]
34. Almagro-Moreno S, and Boyd EF (2009). Insights into the evolution of sialic acid catabolism among bacteria. BMC Evolutionary Biology 9, 118. 10.1186/1471-2148-9-118. [PubMed: 19470179]
35. Krause JL, Schaepe SS, Fritz-Wallace K, Engelmann B, Rolle-Kampczyk U, Kleinsteuber S, Schattenberg F, Liu Z, Mueller S, Jehmlich N, et al. Following the community development of SIHUMIx – a new intestinal in vitro model for bioreactor use. Gut Microbes 11, 1116–1129. 10.1080/19490976.2019.1702431.
36. Triantis V, Bode L, and van Neerven RJJ (2018). Immunological Effects of Human Milk Oligosaccharides. Front Pediatr 6, 190. 10.3389/fped.2018.00190. [PubMed: 30013961]
37. Autran CA, Kellman BP, Kim JH, Asztalos E, Blood AB, Spence ECH, Patel AL, Hou J, Lewis NE, and Bode L (2018). Human milk oligosaccharide composition predicts risk of necrotising enterocolitis in preterm infants. Gut 67, 1064–1070. 10.1136/gutjnl-2016-312819. [PubMed: 28381523]
38. Masi AC, Embleton ND, Lamb CA, Young G, Granger CL, Najera J, Smith DP, Hoffman KL, Petrosino JF, Bode L, et al. (2021). Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis. Gut 70, 2273–2282. 10.1136/gutjnl-2020-322771. [PubMed: 33328245]
39. Charbonneau MR, O'Donnell D, Blanton LV, Totten SM, Davis JCC, Barratt MJ, Cheng J, Guruge J, Talcott M, Bain JR, et al. (2016). Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition. Cell 164, 859–871. 10.1016/j.cell.2016.01.024. [PubMed: 26898329]
40. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, et al. (2012). Human gut microbiome viewed across age and geography. Nature 486, 222–227. 10.1038/nature11053. [PubMed: 22699611]
41. JOHNSON JL (1978). Taxonomy of the Bacteroides. International Journal of Systematic and Evolutionary Microbiology 28, 245–256. 10.1099/00207713-28-2-245.
42. Cato EP, and Johnson JL (1976). Reinstatement of Species Rank for Bacteroides fragilis, B. ovatus, B. distasonis, B. thetaiotaomicron, and B. vulgatus: Designation of Neotype Strains for Bacteroides fragilis (Veillon and Zuber) Castellani and Chalmers and Bacteroides thetaiotaomicron (Distaso) Castellani and Chalmers. International Journal of Systematic Bacteriology 26, 230–237. 10.1099/00207713-26-2-230.
43. Miyamoto Y, and Itoh K (2000). Bacteroides acidifaciens sp. nov., isolated from the caecum of mice. Int J Syst Evol Microbiol 50 Pt 1, 145–148. 10.1099/00207713-50-1-145. [PubMed: 10826798]
44. Cebolla A, Sousa C, and de Lorenzo V (2001). Rational design of a bacterial transcriptional cascade for amplifying gene expression capacity. Nucleic Acids Res 29, 759–766. 10.1093/nar/29.3.759. [PubMed: 11160899]

45. Alexeyev MF (1999). The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. Biotechniques 26, 824–826, 828. 10.2144/99265bm05. [PubMed: 10337469]
46. Smith CJ, Rogers MB, and McKee ML (1992). Heterologous gene expression in Bacteroides fragilis. Plasmid 27, 141–154. 10.1016/0147-619x(92)90014-2. [PubMed: 1615064]
47. Kolde R. (2012). Pheatmap: pretty heatmaps. R package version 1, 726.
48. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. (2019). Welcome to the Tidyverse. Journal of Open Source Software 4, 1686. 10.21105/joss.01686.
49. Neuwirth E. ColorBrewer palettes. 2022.
50. Dowle M, Srinivasan A, Gorecki J, Chirico M, Stetsenko P, Short T, Lianoglou S, Antonyan E, Bonsch M, Parsonage H, et al. (2023). data.table: Extension of “data.frame”. Version 1.14.8.
51. Fischer B, Neumann S, Gatto L, Kou Q, and Rainer J (2023). mzR: parser for netCDF, mzXML, mzData and mzML and mzIdentML files (mass spectrometry data). Version 2.36.0 (Bioconductor version: Release (3.18)). 10.18129/B9.bioc.mzR 10.18129/B9.bioc.mzR.
52. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47. 10.1093/nar/gkv007. [PubMed: 25605792]
53. Xie Y. (2014). knitr: A Comprehensive Tool for Reproducible Research in R. In Implementing Reproducible Research (Chapman and Hall/CRC).
54. Raffelsberger W. (2023). wrMisc: Analyze Experimental High-Throughput (Omics) Data. Version 1.14.0.
55. Raffelsberger W. (2023). wrGraph: Graphics in the Context of Analyzing High-Throughput Data. Version 1.3.6.
56. Tang Y, Horikoshi M, and Li W (2016). ggfortify: Unified Interface to Visualize Statistical Results of Popular R Packages. The R Journal 8, 474. 10.32614/RJ-2016-060.
57. Kassambara A. (2023). ggpubr: “ggplot2” Based Publication Ready Plots. Version 0.6.0.
58. Zhu Y, Orre LM, Zhou Tran Y, Mermelekas G, Johansson HJ, Malyutina A, Anders S, and Lehtiö J (2020). DEqMS: A Method for Accurate Variance Estimation in Differential Protein Expression Analysis. Mol Cell Proteomics 19, 1047–1057. 10.1074/mcp.TIR119.001646. [PubMed: 32205417]
59. Korotkevich G, Sukhov V, Budin N, and Sergushichev A (2023). fgsea: Fast Gene Set Enrichment Analysis. Version 1.28.0 (Bioconductor version: Release (3.18)). 10.18129/B9.bioc.fgsea 10.18129/B9.bioc.fgsea.
60. Storey JD (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. The Annals of Statistics 31, 2013–2035. 10.1214/aos/1074290335.
61. Torchiano M. (2016). Effsize - a package for efficient effect size computation. (Zenodo). 10.5281/ZENODO.1480624 10.5281/ZENODO.1480624.
62. Fox J, Weisberg S, Price B, Adler D, Bates D, Baud-Bovy G, Bolker B, Ellison S, Firth D, Friendly M, et al. (2023). car: Companion to Applied Regression. Version 3.1–2.
63. Schauberger P, Walker A, Braglia L, Sturm J, Garbuszus JM, and Barbone JM (2023). openxlsx: Read, Write and Edit xlsx Files. Version 4.2.5.2
64. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrlund TS, Bushuiev R, Murray KJ, Hoffmann N, Lu M, et al. (2023). Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nat Biotechnol 41, 447–449. 10.1038/s41587-023-01690-2. [PubMed: 36859716]
65. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, et al. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol 34, 828–837. 10.1038/nbt.3597. [PubMed: 27504778]
66. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, and Henrissat B (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42, D490–495. 10.1093/nar/gkt1178. [PubMed: 24270786]

67. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, et al. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Research 47, D351–D360. 10.1093/nar/gky1100. [PubMed: 30398656]

68. Sievers F, and Higgins DG (2018). Clustal Omega for making accurate alignments of many protein sequences. Protein Science 27, 135–145. 10.1002/pro.3290. [PubMed: 28884485]

69. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, and Nielsen H (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37, 420–423. 10.1038/s41587-019-0036-z. [PubMed: 30778233]

70. Zhang G, Mills DA, and Block DE (2009). Development of chemically defined media supporting high-cell-density growth of lactococci, enterococci, and streptococci. Appl Environ Microbiol 75, 1080–1087. 10.1128/AEM.01416-08. [PubMed: 19074601]

71. Jantscher-Krenn E, Zherebtsov M, Nissan C, Goth K, Guner YS, Naidu N, Choudhury B, Grishin AV, Ford HR, and Bode L (2012). The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats. Gut 61, 1417–1425. 10.1136/gutjnl-2011-301404. [PubMed: 22138535]

72. Berger PK, Hampson HE, Schmidt KA, Alderete TL, Furst A, Yonemitsu C, Demerath E, Goran MI, Fields DA, and Bode L (2023). Stability of Human-Milk Oligosaccharide Concentrations Over 1 Week of Lactation and Over 6 Hours Following a Standard Meal. J Nutr 152, 2727–2733. 10.1093/jn/nxac214. [PubMed: 36111739]

73. Zhang Z, Xie J, Zhang F, and Linhardt RJ (2007). Thin-layer chromatography for the analysis of glycosaminoglycan oligosaccharides. Anal Biochem 371, 118–120. 10.1016/j.ab.2007.07.003. [PubMed: 17679101]

74. Campeau A, Mills RH, Stevens T, Rossitto L-A, Meehan M, Dorrestein P, Daly R, Nguyen TT, Gonzalez DJ, Jeste DV, et al. (2022). Multi-omics of human plasma reveals molecular features of dysregulated inflammation and accelerated aging in schizophrenia. Mol Psychiatry 27, 1217–1225. 10.1038/s41380-021-01339-z. [PubMed: 34741130]

75. Wang Y, Yang F, Gritsenko MA, Wang Y, Clauss T, Liu T, Shen Y, Monroe ME, Lopez-Ferrer D, Reno T, et al. (2011). Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. Proteomics 11, 2019–2026. 10.1002/pmic.201000722. [PubMed: 21500348]

76. Lapek JD, Lewinski MK, Wozniak JM, Guatelli J, and Gonzalez DJ (2017). Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Protein Vpr. Mol Cell Proteomics 16, 1447–1461. 10.1074/mcp.M116.066019. [PubMed: 28606917]

Highlights

- Human milk oligosaccharides initiate commensal colonization during early life
- The *B. fragilis* sialidase, NanH, is induced by HMOs
- *B. fragilis* NanH activity mediates mucosal colonization in pups and adult mice
- NanH facilitates commensal resilience and recovery post-antibiotic treatment

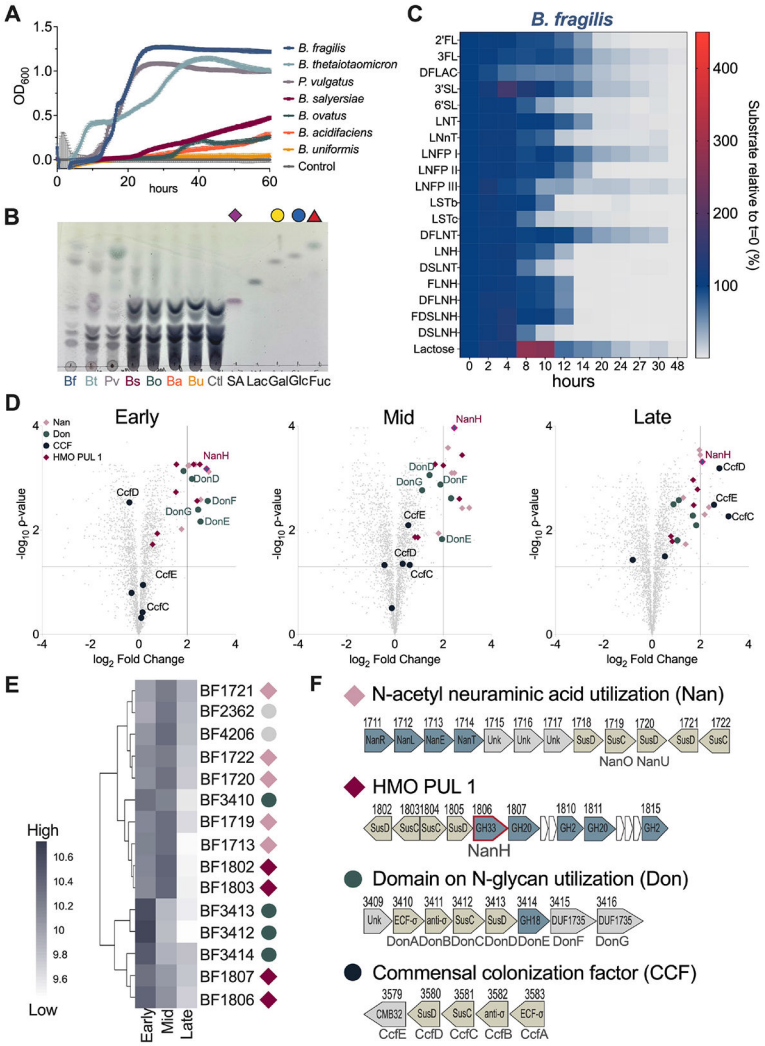


Figure 1. Defining HMO utilization system in *B. fragilis*

(A) Growth of seven *Bacteroides* species: *B. fragilis* NCTC 9343 (Bf), *B. thetaiotaomicron* VPI-8482 (Bt), *P. vulgatus* ATCC 8482 (Pv), *B. ovatus* ATCC 8483 (Bo), *B. salyersiae* DSM 18765 (Bs), *B. uniformis* ATCC 8492 (Bu), and *B. acidifaciens* JCM 10556 (Ba) in defined medium with 1.5 % pooled HMOs (pHMOs). Solid line represents mean ± SD.

(B) TLC of cell-free supernatant collected from stationary phase cultures. Controls: sialic acid/Neu5Ac (SA); lactose (Lac); galactose (Gal), glucose (Glc); fucose (Fuc).

(C) HPLC analysis of the supernatant collected during *B. fragilis* growth on pHMOs over time.

(D) Proteomic analysis of pHMO-grown whole *B. fragilis* cells, pHMOs vs glucose fold change at early-, mid-, and late-log phases. Significance: $\log_2$-fold change ≥2, P-value < 0.05.

(E) Relative abundance of proteins with a max $\log_{10}$-fold change ≥ 2 and a Benjamini Hochenburg corrected P-value < 0.05 across growth phases. Hierarchical clustering was performed using Euclidean distance. Scale bar: protein abundance from high (dark purple) to low (white).

(F) Organization of significantly abundant polysaccharide utilization loci (PULs) in *B. fragilis* during growth on pHMO, gene labeling corresponds to locus tags BFxxxx: N-acetyl neuraminic acid utilization (Nan) operon (pink diamond); HMO PUL1 (maroon diamond); Domain of N-glycan (Don) utilization operon (teal circle); Commensal colonization factor (CCF) operon (black circle).

See also Figure S1, Figure S2, and Tables S1-S3.

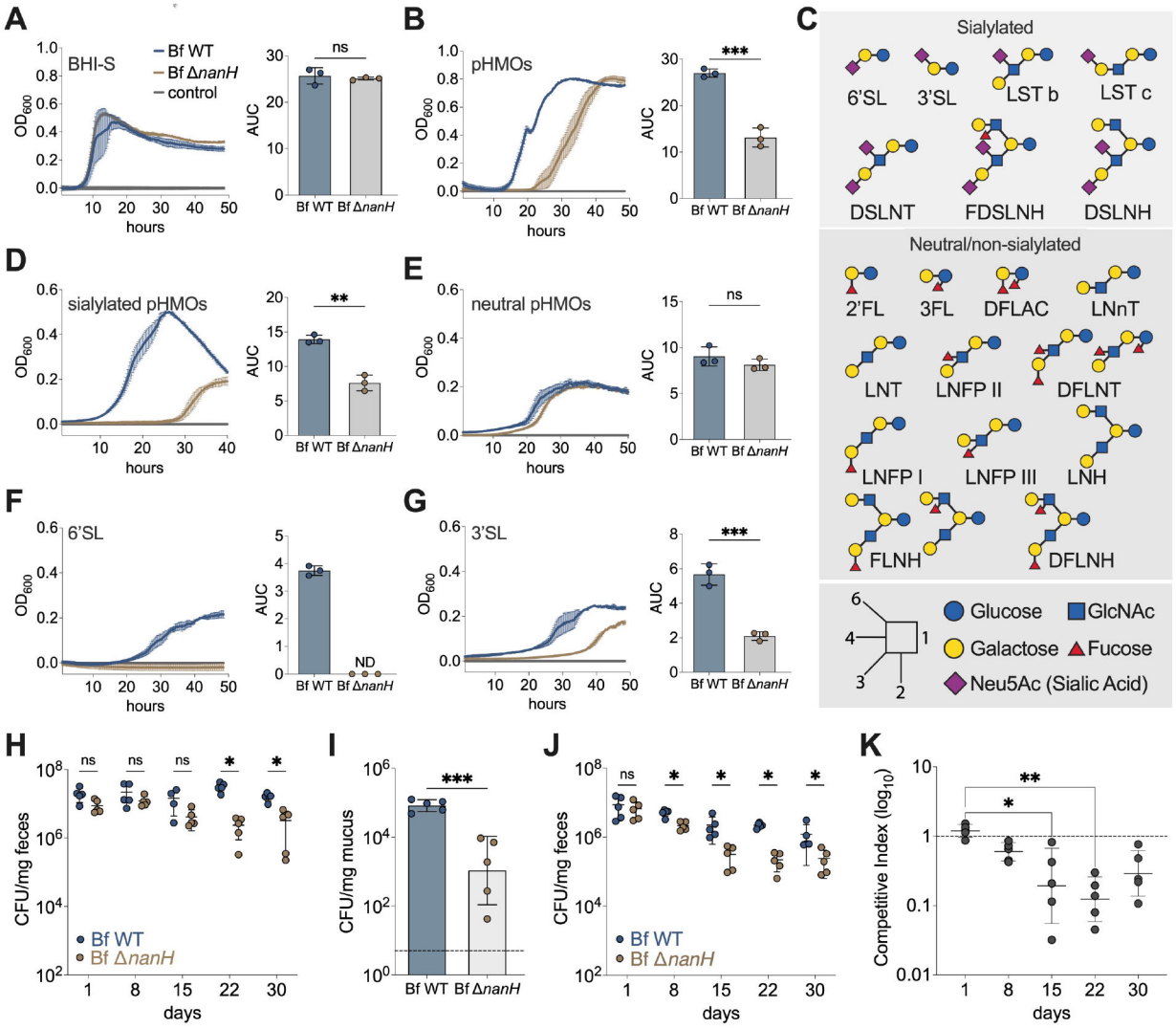


Figure 2. NanH is required for growth on HMOs and promotes intestinal colonization.

(A – G) Growth curves and area under the curve (AUC) of Bf WT (blue) and Bf Δ*nanH* (brown) on (A) BHI-S; (B) Pooled HMOs (pHMOs); (C) Diagrammatic structures of sialylated (top) and non-sialylated (bottom) HMOs. Sugar and linkage keys are shown; (D) Pooled sialylated HMOs; (E) Pooled neutral HMOs; (F) 6'-sialylactose (6'SL); (G) 3'-sialylactose (3'SL). Data is shown as mean ± SD of 3 biological replicates. Solid line represents mean ± SD.

(H – I) Germ-free mice were orally gavaged with either *B. fragilis* WT or Δ*nanH* (n=10). Abundance of each strain was quantified in CFU/mg of (G) feces and (I) colonic mucus. Limit of detection (LOD) = 5 CFU/mg.

(J– K) Germ-free mice (n=5) were orally gavaged with 1:1 mixture of *B. fragilis* WT and Δ*nanH* strains. (J) Quantification of Bf WT and Bf Δ*nanH* in feces of co-colonized mice. (K) The competitive index (log_{10}) as indicated by input/output ratios. Differentiation between the Bf WT and Bf Δ*nanH* was performed as described in Methods. Data are shown as geometric mean ± SD, representative of two independent experiments. *P < 0.05; **P <

0.01; ***P < 0.001; ns – not significant; (Kruskal-Wallis and Mann-Whitney tests); ND – not detected.

See also Figure S3, Figure S4.

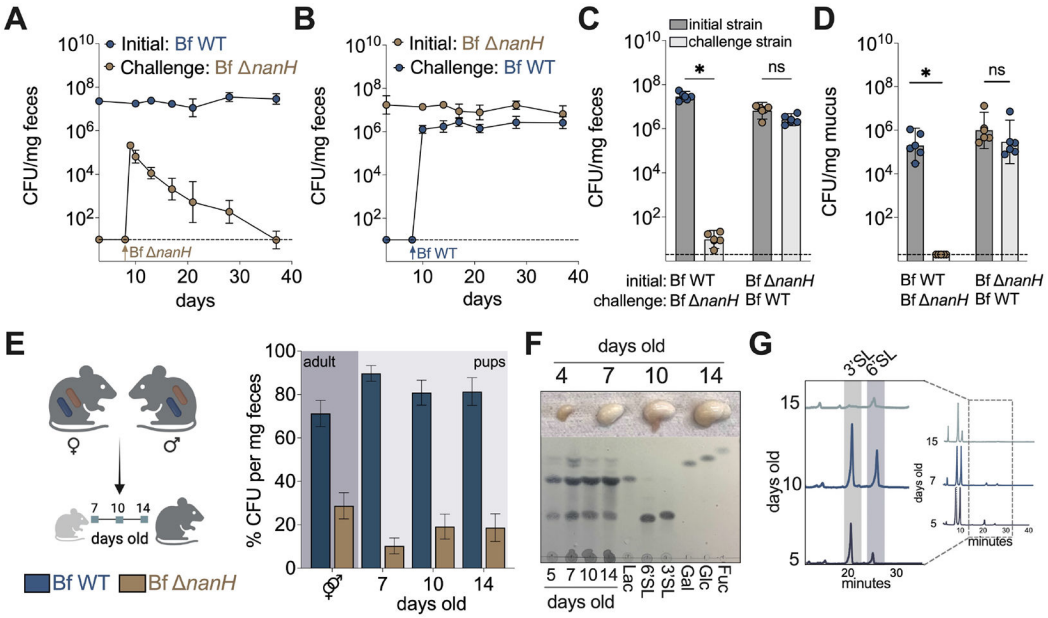


Figure 3. NanH mediates stable niche occupancy and early life colonization.

(A) Germ-free mice were mono-associated with Bf WT (initial) and challenged with Bf Δ*nanH* on day 8; n=5.

(B) Germ-free mice were mono-associated with Bf Δ*nanH* (initial) and challenged with Bf WT on day 8; n=5.LOD = 10 CFU/mg of feces

(C-D) CFU/mg of the initial and the challenge strains in (C) fecal pellets and (D) colonic mucus 27 days post-challenge; n=5. Data are shown as geometric mean ± SD, representative of two independent experiments. LOD = 5 CFU/mg of feces or mucus; *P < 0.05; ns – not significant; Mann-Whitney tests.

(E) Vertical transmission schematic (left) and the abundance of Bf WT and Bf Δ*nanH* (right) in the feces of adult breeders (dark gray) and newborn suckling pups (light gray). n=16 pups.

(F) Images of stomachs collected from suckling pups (top). TLC analysis of the homogenized stomachs (bottom). Controls: 6'SL, 3'SL, lactose (Lac), glucose (Glc), galactose (Gal), and fucose (Fuc).

(G) HPLC analysis of the stomach contents. Peaks in the shaded area indicate the peaks of 3'SL and 6'SL.

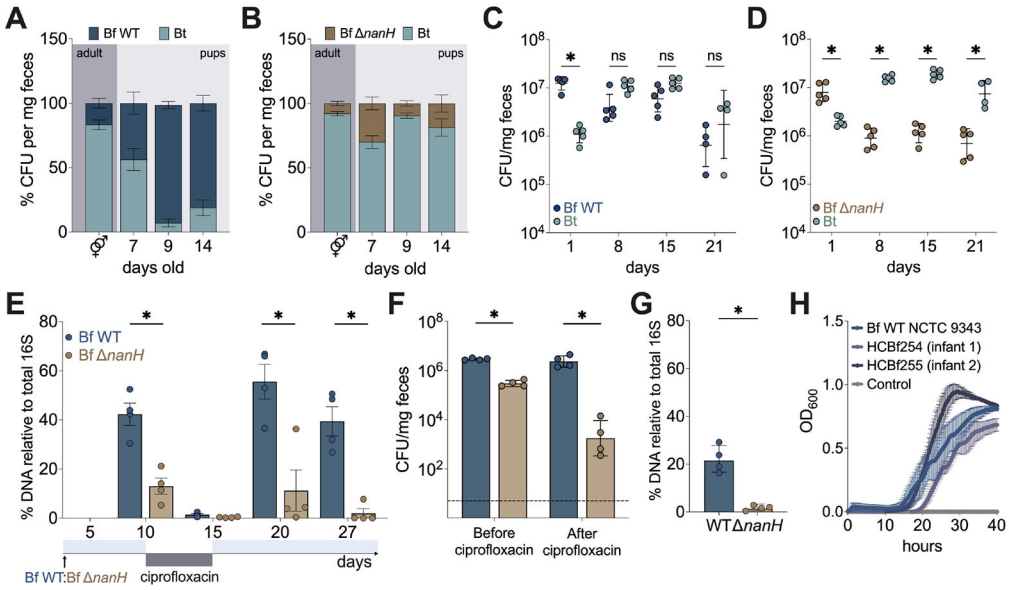


Figure 4. *B. fragilis* NanH defines resilience and competitive fitness.

(A-B) Abundance of (A) Bf WT and *B. thetaiotaomicron* (n=12 pups) and (B) Bf Δ*nanH* and *B. thetaiotaomicron* (n=20 pups) in the feces of adult breeders (♀♂) and newborn suckling pups at designated ages (days old).

(C – D) Germ free mice were inoculated with (C) Bf WT and *B. thetaiotaomicron* WT (Bt) or (D) Bf Δ*nanH* and *B. thetaiotaomicron* WT (Bt) at 1:1 ratio; n=5 per group.

(E – G) Mice harboring the Simplified Human Intestinal Microbiota (SIHUMIx) were challenged with WT and Bf Δ*nanH* (1:1) on day 0. (G) The abundance of Bf WT and Bf Δ*nanH* was determined by qPCR, relative to total microbial 16S before (day 10), after (day 15), and during the recovery period (days 20 and 27) from the ciprofloxacin treatment; n=4.

(F) Fecal load (CFU/mg) of Bf WT and Bf ΔnanH before and after the ciprofloxacin treatment. Differentiation between strains is described in Methods.

(G) Abundance of Bf WT and Bf Δ*nanH* in colonic mucus, determined by qPCR; n= 4. Data show mean ± SD, LOD = 5 CFU/mg. *P < 0.05; ns – not significant; Mann-Whitney test. Data represents at least two independent experiments.

(H) Growth of *B. fragilis* strains isolated from human infants on 15 mg/ml pHMOs. Solid line represents the mean ± SD.

See also Figure S4.

KEY RESOURCE TABLE

REAGENT or SOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
*Bacteroides fragilis*NCTC 9343	Johnson et al., 1978[41]	NCTC 9343
Bacteroides thetaiotaomicron VPI 5482	ATCC	Cat# 29148
Phocaeicola vulgatus ATCC 8482	Cato et al., 1976[42]	ATCC 8482
Bacteroides ovatus ATCC 8483	Cato et al., 1976[42]	ATCC 8483
Bacteroides salyersiae DSM 18765	ATCC	Cat# BAA-997
Bacteroides uniformis ATCC 8492	Cato et al., 1976[42]	ATCC 8492
Bacteroides acidifaciens JCM 10556	Miyamoto et al., 2000[43]	JCM 10556
Bifidobacterium longum NCC2705	ATCC	Cat# 15707
Bifidobacterium breve RA40 E10	Lab isolate	RA40 E10
B. fragilis ΔCCF	Lee et al., 2013[15]	*B. fragilis* lacking *bf*3579-*bf*3583
B. fragilis Δ*nanH*	This study	*B. fragilis* lacking *bf*1806
B. fragilis ΔCCFΔ*nanH*	This study	*B. fragilis* lacking *bf*3579-*bf*3583 and *bf*1806
B. fragilis Δ*nanH*::NanH	This study	*B. fragilis* lacking *bf*1806 and containing *bf*1806 on pFD340
B. fragilis WT + pFD340-tetQ	This study	*B. fragilis* containing pFD340-tetQ
B. fragilis Δ*nanH* + pFD340-cat	This study	*B. fragilis* containing pFD340-cat
B. fragilis Δ*nanH* + pFD340-cat	This study	*B. thetaiotaomicron* WT containing pFD340-tetQ
B. thetaiotaomicron WT + pFD340-cat	This study	*B. thetaiotaomicron* WT containing pFD340-cat
E. coli CC118 *λ-pir*	Cebolla et al., 2001[44]	*Δ(ara-leu) araD ΔlacX74 galE galK phoA20 thi-1 rpsE rpoB argE (Am) recA1 λpir*
E. coli S-17 *λ-pir*	Cebolla et al., 2001[44]	*recA pro hsdR RP4- 2 (Tc::Mu;Km::Tn7) (λ pir)*
Biological samples		
Fecal samples collected from human infants	This study	N/A
Human milk oligosaccharides (HMOs) purified from pooled human donor milk	This study	N/A
Human pooled acidic HMOs	This study	N/A
Human pooled neutral HMOs	This study	N/A
3'sialyllactose	GeneChem	Cat# BO07Th14-02
6'sialyllactose	GeneChem	Cat# G14004
Chemicals, peptides, and recombinant proteins		
Brain Heart Infusion (BHI) broth	Thermo Scientific	Cat# DF0037-07-0
Bacto Agar dehydrated	Thermo Scientific	Cat# DF0140-01-0
Bacto Yeast Extract	Thermo Scientific	Cat# 288620
Luria-Bertani Agar	Thermo Scientific	Cat# DF0445-17-4
Luria-Bertani Broth	Thermo Scientific	Cat# DF0446-17-3
Gifo Anaerobic Media	Thermo Scientific	Cat# NC0690484
Ampicillin	Sigma Aldrich	Cat# A9518
Ciprofloxacin	Thermo Scientific	Cat# J61317-14
Vitamin K1	Sigma Aldrich	Cat# V3501-1G

REAGENT or SOURCE	SOURCE	IDENTIFIER
Hemin	Sigma Aldrich	Cat# H9039-1G
Gentamycin sulfate vet grade	VetOne	Cat# 13985-633-04
Erythromycin	Sigma Aldrich	Cat# E5389
Chloramphenicol	Sigma Aldrich	Cat# C0378
Tetracycline hydrochloride	Sigma Aldrich	Cat# T7660
L-Histidine	Sigma Aldrich	Cat# H8000
L-Isoleucine	Sigma Aldrich	Cat# 12752
L-Leucine	Sigma Aldrich	Cat# L8000
L-Methionine	Sigma Aldrich	Cat# M9625
L-Valine	Sigma Aldrich	Cat# V0500
L-Arginine	Sigma Aldrich	Cat# A5006
Myo-Inositol	Thermo Scientific	Cat# AC122261000
Potassium phosphate dibasic	Thermo Scientific	Cat# P288
Potassium phosphate monobasic	Thermo Scientific	Cat# P285
L-Glutamic Acid	Sigma Aldrich	Cat# G1251
L-Phenylalanine	Sigma Aldrich	Cat# AAA1323814
L-Proline	Thermo Scientific	Cat# AAA1019914
L-Asparagine	Sigma Aldrich	Cat# A4159
L-Aspartic Acid	Thermo Scientific	Cat# AAA1352022
L-Glutamine	Thermo Scientific	Cat# BP379
L-Serine	VWR	Cat# AAA1117914
L-Threonine	Thermo Scientific	Cat# AC138930250
L-Cysteine	Sigma Aldrich	Cat# C7477
L-Alanine	VWR	Cat# AAA1580414
Glycine	Thermo Scientific	Cat# BP381
L-Lysine	Thermo Scientific	Cat# AAA1624918
L-Tryptophan	Sigma Aldrich	Cat# T0254
Calcium pantothenate	Thermo Scientific	Cat# AC243300050
Niacin	Thermo Scientific	Cat# 18-604-780
Pyridoxal HCl	Thermo Scientific	Cat# AAA1785506
$MgSO_4$	Thermo Scientific	Cat# M63-500
$FeSO_4.7H_20$	Thermo Scientific	Cat# I146500
$ZnSO_4.7H_2O$	Thermo Scientific	Cat# Z68-500
Folic Acid	Sigma Aldrich	Cat# F8758-5G
p-Aminobenzoic acid	Thermo Scientific	Cat# AC146212500
Potassium acetate	Thermo Scientific	Cat# BP364-500
Lipoic acid	Thermo Scientific	Cat# L0058
Tween-80	Thermo Scientific	Cat# BP338
Adenine	Thermo Scientific	Cat# AAA1490614
Guanine	Thermo Scientific	Cat# AC120250250
Uracil	Thermo Scientific	Cat# AAA1557018
Xanthine	Thermo Scientific	Cat# AAA1107714

REAGENT or SOURCE	SOURCE	IDENTIFIER
MOPS	Thermo Scientific	Cat# BP308
Tricine	Thermo Scientific	Cat# AC172642500
$MnSO_4.4H_2O$	Thermo Scientific	Cat# AAB2208122
$CaCl_2.2H_2O$	Sigma Aldrich	Cat# C3306
$CoSO_4.6H_2O$	Sigma Aldrich	Cat# C6768
$CuSO_4.5H_2O$	Thermo Scientific	Cat# BP346
Boric acid	Thermo Scientific	Cat# A74-500
K_2SO_4	Thermo Scientific	Cat# P304-500
KI	Thermo Scientific	Cat# P410-100
EDTA	Sigma Aldrich	Cat# E6758
Nitrilotriacetic acid	Thermo Scientific	Cat# AAA1193622
L-Glutathione	Sigma Aldrich	Cat# G4251
$(NH_4)_2SO_4$	Thermo Scientific	Cat# BP212R
NaCl	Thermo Scientific	Cat# S271-1
Tyrosine	Thermo Scientific	Cat# AAA1114118
Biotin	Thermo Scientific	Cat# AC230090010
Thiamine	Sigma Aldrich	Cat# T4625
Riboflavin	Sigma Aldrich	Cat# R7649
NEBNext High-fidelity 2X PCR Master Mix	New England Biolabs	Cat# M0541
NEBuilder HiFiDNA assembly Master mix	New England Biolabs	Cat# E2621
BamHI	New England Biolabs	Cat# R3136
SalI	New England Biolabs	Cat# R0138
SacI	New England Biolabs	Cat# R3156
SYBR Green Master Mix	Life Technologies	Cat# A25778
SuperScript IV VILO Master Mix	Life Technologies	Cat# 11756500
TLC Silica gel 60 F_{254}	Sigma Aldrich	Cat# 1055540001
Butanol	Sigma Aldrich	Cat# 537993
Acetic acid	Sigma Aldrich	Cat# 695092
Diphenylamine	Sigma Aldrich	Cat# 112763
Aniline	Sigma Aldrich	Cat# 242284
Phosphoric acid	VWR	Cat# 2796
Hydrochloric acid	Sigma Aldrich	Cat# 258148
Deoxy-D-Glucose	Sigma Aldrich	Cat# D8375
D-Galactose	Sigma Aldrich	Cat# G0625
L-Fucose	Sigma Aldrich	Cat# F2252
D-lactose	Sigma Aldrich	Cat# L254
Sialic acid (Neu5AC)	Biosynth	Cat# MA00746
Dulbecco's Phosphate Buffered Saline with calcium and magnesium (1X)	Thermo Scientific	Cat# MT21030CV
Bugbuster	Millipore	Cat# 705843
RNAlater stabilization solution	Thermo Scientific	Cat# am7021
Anthranilamide	Thermo Scientific	Cat# 104905000

REAGENT or SOURCE	SOURCE	IDENTIFIER
Sodium cyanoborohydride	Sigma Aldrich	Cat# 156159
Acetic acid	Sigma Aldrich	Cat# A6283
Tandem Mass Tag (TMT) 16plex reagents	Thermo Scientific	Cat# A44520; lot#XA341491
C_4 5 μm Stationary phase	Sepax	Cat# 109045-0000
C_{18} 3 μm Stationary phase	Sepax	Cat# 101183-0000
C_{18} 1.8 μm Stationary phase	Sepax	Cat# 101181-0000
Critical commercial assays		
QIAquick PCR purification kit	Qiagen	Cat# 28106
QIAquick Gel extraction kit	Qiagen	Cat# 28706
Nucleospin Plasmid kit	Macherey Nagel	Cat# 740588
Quick-DNA Fecal/Soil Microbe Miniprep kit	Genesee Scientific	Cat# 11-322
NucleoSpin Microbial DNA kit	Macherey Nagel	Cat# 740235
Nucleospin RNA kit	Macherey Nagel	Cat# 740955
ProtiFi S-Trap columns	ProtiFi	Cat# C02-mini
SepPak C18 columns	Waters	Cat# WAT054960
Pierce Quantitative Colorimetric Peptide Assay	Thermo Scientific	Cat# 23275
Deposited data		
Whole cell proteomic	This study	ID MSV000090386; https://massive.ucsd.edu
Experimental models: Organisms/strains		
Mouse: C57BL/6J	Jackson Laboratory	RRID:IMSR_JAX:000664
Mouse: B6.Cg-*Rag2*$^{tm1.1Cgn}$/J	Jackson Laboratory	RRID:IMSR_JAX:008449
Oligonucleotides		
See Table S5 for oligonucleotides		
Recombinant DNA		
PKNOCK-bla-erm	Alexeyev et al., 1999[45]	*Bacteroides* suicide vector, *mob*+, *tra*−, AmpR (*E. coli*), ErmR (*Bacteroides*)
pFD340	Smith et al., 1992[46]	*Bacteroides* shuttle vector, contains IS4351 promoter, AmpR (*E. coli*), ErmR (*Bacteroides*)
pFD340-cat	Lee et al., 2013[15]	Modified pFD340 plasmid, AmpR (*E. coli*), CmR ErmR (*Bacteroides*)
pFD340-tetQ	Lee et al., 2013[15]	Modified pFD340 plasmid, AmpR (*E. coli*), TetR ErmR (*Bacteroides*)
Software and algorithms		
pheatmap	Kolde et al., 2012[47]	https://github.com/raivokolde/pheatmap
tidyverse	Wickham et al., 2019[48]	https://tidyverse.tidyverse.org/
RColorBrewer	Neuwirth et al., 2022[49]	https://CRAN.R-project.org/package=RColorBrewer
data.table	Dowle et al., 2019[50]	https://github.com/Rdatatable/data.table
mzR	Smith et al., 2013[51]	https://github.com/sneumann/mzR/
limma	Ritchie et al., 2015[52]	https://kasperdanielhansen.github.io/-/genbioconductor/html/limma.html
knitr	Xie et al., 2014[53]	https://CRAN.R-project.org/package=knitr
wrMisc	Raffelsberger, 2023[54]	https://CRAN.R-project.org/package=wrMisc
wrProteo	Raffelsberger, 2023[54]	https://CRAN.R-project.org/package=wrProteo

REAGENT or SOURCE	SOURCE	IDENTIFIER
wrGraph	Raffelsberger, 2023[55]	https://CRAN.R-project.org/package=wrGraph
ggfortify	Tang et al., 2016[56]	https://github.com/sinhrks/ggfortify
ggpubr	Kassambara et al., 2023[57]	https://rpkgs.datanovia.com/ggpubr/
DEqMS	Zhu et al., 2020[58]	https://github.com/yafeng/DEqMS
fgsea	Korotkevich et al., 2023[59]	https://github.com/ctlab/fgsea/
qvalue	Storey, 2003[60]	https://github.com/StoreyLab/qvalue
effsize	Torchiano, 2020[61]	https://github.com/mtorchiano/effsize
car	Fox et al., 2023[62]	https://CRAN.R-project.org/package=car
openxlsx	Schauberger et al., 2023[63]	https://github.com/ycphs/openxlsx
R	R Core Team, 2021	https://www.R-project.org/
R studio	R Studio Team, 2021	https://www.rstudio.com/
MZmine 3	Schmid et al., 2023[64]	https://mzmine.github.io/
GNPS	Wang et al., 2016[65]	https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp
CAZy	Lombard et al., 2014[66]	http://www.cazy.org
PUL DB	Terrapon et al., 2018[22]	http://www.cazy.org/PULDB/
InterPro	Mitchell et al., 2019[67]	https://www.ebi.ac.uk/interpro/search/sequence/
Clustal Omega	Sievers and Higgins et al., 2018[68]	https://www.ebi.ac.uk/Tools/msa/clustalo/
Signal P 5.0	Almagro Amenteros et al., 2019[69]	https://services.healthtech.dtu.dk/services/SignalP-5.0/
NEBuilder assembly tool	New England Biolabs	https://nebuilder.neb.com
ProteomeDiscoverer 2.5	Thermo Scientific	ProteomeDiscoverer 2.5
GraphPad Prism	GraphPad Version 9.0 or higher	https://www.graphpad.com
Other		
Stainless steel beads 3.2 mm RNase-free	Next Advance	Cat# SSB32-RNA
TSKgel Amide-80 column	Tosh Bioscience	Cat# 0021865
Bio-Gel P2 Media	Bio-Rad	Cat# 150-4114
HPLC column	Thermo Scientific	Cat# 720105-254630
Fused Silica Capillary Tubing	Polymicro Technologies	Cat# 106815-0023
Orbitrap Fusion Tribird Mass Spectrometer	Thermo Scientific	IQLAAEGAAPFADBMBCX
QuantStudio 5.0	Thermo Scientific	Cat# A28140
Bullet blender	Next Advance	Storm 24