A *Caenorhabditis elegans* based system for high-throughput functional phenotyping of human gut microbiota

Achuthan Ambat[1,2], Vishnu Thayyil Valappil[1,2], Sudeep Ghimire[1], Phurt Harnvoravongchai[2], Shalabh Mishra[1,2], Viju Vijayan Pillai[4], Purna Kashyap[3], and Joy Scaria[1,2]*

[1]Department of Veterinary and Biomedical Sciences, South Dakota State University, Brookings, SD, USA
[2]Department of Veterinary Pathobiology, Oklahoma State University, Stillwater, OK, USA
[3]Enteric Neuroscience Program, Department of Medicine and Physiology, Mayo Clinic, Rochester, MN, USA
[4]Department of Comparative Pathobiology, Animal Disease Diagnostic Laboratory, Purdue University College of Veterinary Medicine, West Lafayette, IN, USA

* Correspondence:
Corresponding Author
Joy.scaria@okstate.edu

Summary

Bottoms-up approach of mono or poly colonizing microbes in germfree model is an important tool for mechanistic understanding of human gut microbiota. However, doing this in models such as germfree mouse is expensive and time consuming. To address this problem, we developed a *Caenorhabditis elegans* based screening system. We used a gut microbiota culture collection that represents more than 70% functional capacity of the human gut microbiome to anaerobically colonize *C. elegans*. We chose colonization resistance as the phenotype of the microbiome for further screening and found that half of the strains, previously identified *in vitro* as inhibiting *Clostridioides difficile*, also did so in *C. elegans*. When validated using germ-free mouse model, results were in concordance with that obtained from *C. elegans* model. Our system therefore could be used for screening large number of bacterial species to better understand host-microbiome interaction.

Introduction

The human gut microbiota (GM) is an intricate and dynamic microbial ecosystem, constantly interacting with its host[1-4]. This ecosystem, home to thousands of symbiotic species, significantly influences the host's health[5,6]. The immune system, both innate and adaptive, faces a substantial challenge in maintaining the diversity and balance of these species[7]. While association studies correlating gut microbiome species with immunological markers have been common approaches to understanding these interactions, such generalizations are challenging due to gut microbes' diverse and dynamic nature[8-10]. A more focused strategy involves mono-colonization studies to observe physiological changes, moving away from broader screening methods that struggle to encompass the sheer diversity within the gut. Previous studies employing similar strategies on germ-free rodents were limited to a maximum of 54 monocolonized species[11].

Notably, the gut microbiota serves as a vital defense against enteric infections, including *Clostridioides difficile* infection (CDI)[12]. CDI has been established as the most common infection in healthcare-associated individuals in the USA[13]. Disturbances and loss of diversity within the gut microbiome have been linked to the colonization and pathogenesis of CDI[14,15].

Addressing this, Fecal Microbiota Transplant (FMT) has been recognized as a safe and effective treatment for recurrent CDI, boasting an efficacy rate of 80-90%[16,17]. Despite the success of FMT, the precise mechanisms of how it assists in recovery from dysbiosis remain unclear. Moreover, there is a concern about the potential harm of exposing patients with various immune dysregulations to the allogenic strains used in FMT. This gap in understanding highlights the need for a defined bacterial therapy that can provide a clear understanding of biotherapeutic mechanisms, enabling precise dosing. The random selection of bacteria in clinical trials has not always yielded the expected therapeutic efficacy, raising questions about the best approach to biotherapeutic interventions[18,19]. Even though multiple studies address colonization *in vitro*, one of the significant challenges in this field is replicating in-vitro inhibition data in in-vivo environments. Considering the vast number of bacteria combinations possible from a moderately sized library, finding a reliable therapeutic option in in-vivo models like mice is impractical due to the sheer volume of combinations and the associated time and cost.

In light of these challenges, this study pivots to the potential of using *Caenorhabditis elegans* as a high-throughput model organism for such studies. *C. elegans*, first isolated from soil in Algeria in 1900 and later popularized by Sydney Brenner in 1963 for genetic manipulation studies, presents a simpler biological system[20-22]. Its intestine, comprising 20 cells, serves the primary functions of digestion and nutrient absorption, mirrors liver functions, and plays a central role in pathogen infection, immunity, longevity, and detoxification[23-27]. Recent studies have demonstrated the presence of a microbiome in the *C. elegans* intestine, with around 300 different bacterial genera identified, showing variable microbiome compositions depending on different substrates[28-30].

We have developed and validated an innovative protocol for the colonization of *C. elegans* with a spectrum of anaerobic bacteria isolated from the human gut, thus circumventing the constraints imposed by aerobic conditions and the limitations of continuous bacterial feeding. Our confirmation of successful bacterial colonization, as evidenced by fluorescence microscopy and quantitative CFU analysis of *B. longum* and *C. difficile*, underscores the model's reliability. This work also sheds light on the pivotal role of mucin-associated sugar metabolism in the colonization process and the establishment of a beneficial symbiosis. Notably, *B. longum* not only successfully colonizes but also demonstrates the capacity to impart resistance to *C. difficile* infection, an effect substantiated in germ-free mouse models, thereby enhancing the model's translational potential for high-throughput investigations. Further, the dissection of key immune pathways, with an emphasis on the p38 MAP kinase cascade, provides insights into the nematode's immune response mechanisms paralleling the response observed in germ-free mouse models and in humans. These advances offer a more nuanced understanding of colonization resistance, positioning *C. elegans* as a valuable model for probing host-microbe interactions and their implications for human health. Our findings set the stage for future investigations into the mechanistic underpinnings of microbiota-mediated host defenses, potentially guiding the development of novel microbial-based therapeutic strategies.

Results

Anaerobic feeding of *C. elegans* L1 larvae with human gut bacteria showed successful colonization.

In *C. elegans* host-microbe interaction research, the focus has traditionally been on the colonization of aerobic bacteria through continuous feeding, involving the cultivation of worms on growth medium plates with the bacteria of interest. This method encounters difficulties with obligate anaerobes due to *C. elegans'* aerobic growth conditions. Also, the method is different from the concept of colonization as there is a continuous feeding regime rather than dosage feeding as used in other animal models. To address this issue, we devised a new protocol for introducing and establishing anaerobic human gut bacteria in the *C. elegans* intestine, "short term feeding" rather than continuous. Briefly, germ-free L1 *C. elegans* larvae were subjected to anaerobic bacteria within an anaerobic chamber for a duration of three hours. Post-exposure, these larvae were extracted, thoroughly cleansed, and then relocated to NGM plates already seeded with heat-killed *E. coli OP50* (Figure 1G). To ensure high throughput in our method, we used a 96-well-based liquid growing and colonization (Figure S1A) of the human gut bacterial isolates previously cultured[31]. We reasoned that since the defecation rate of *C. elegans* is an average of 45 seconds[32,33], most of the very slow-growing bacteria would not survive in the intestine of *C. elegans*. We selected 83 bacteria from the library of 102 bacteria, eliminating 19 very slow-growing species as previously described in the colonization resistance screen invitro[31] against *C. difficile*. The colonization trial resulted in the successful mono-colonization of 39 out of the 83 bacteria (Figure 1A). 10 of the colonized bacteria showed a *C. difficile* inhibition of more than 60 % *in vitro*. To understand the factors determining the colonization potential of the strain, we checked multiple genomic and physiological factors. Phylogenetic distance-based analysis showed no relationship between the colonized and non-colonized fractions against the natural microbiome member of *C. elegans*[34] (Figure 1B) and between themselves (Figure S1C). There was also no relation between the percentage GC content, which has otherwise been proven to exist in bacterial genomes in extremities where carbon availabilities are scarce[35](Figure S1B). We also considered pH, temperature dependence, and oxygen tolerance as potential determinants of colonization since the *C. elegans* intestine is slightly acidic[36], grown at 25°C, and weakly anaerobicity because of smaller GI tract. To explore this, we assessed the general pH tolerance, temperature sensitivity, and oxygen tolerance of the selected bacteria using the data for the type strain[37] or the representative species[38] and found no direct relation (Figure 1C and S1D). We examined carbon utilization ability, as bacteria colonizing *C. elegans* require a carbon source for growth. There were no separation between the clustering for the carbon utilization between colonized and uncolonized bacteria (Figure S2A). While gut bacteria in larger model organisms like mice often rely on host diets or mucin, *C. elegans* primarily consume bacteria. Therefore, we hypothesized that colonizing bacteria might utilize sugars associated with host glycoproteins in the nematode's intestine. Hence, we further delved into mucin-associated sugars (N-Acetyl-D-Galactosamine, N-Acetyl-D-Glucosamine, D-Galactose, N-acetylneuraminate, and D-Glucose utilization). We observed a significant increase in the median utilization of N-Acetyl-D-Galactosamine by colonized worms (Figure 1C). None of the other mucin-related sugars showed significant differences (Figure S2C-F). Additionally, analysis of 91 other sugars revealed a

significant increase in the utilization of Glycyl-L-Methionine and B-Methyl-D-Glucoside for the uncolonized fraction (Figure S2B). To check the dependence of SCFA production on colonization, we studied the SCFA profile for all the colonized and not colonized bacteria invitro. The results showed no significant difference between all acetate, butyrate, isobutyrate, valerate, isovalerate, and propionate in colonization (Figure S3A-H). The Random Forest analysis on the different bacterial traits predicted from the bacterial genome using traitar revealed nitrate reduction, oxidase activity, and growth on ordinary media as the most influential factors in colonization (Figure S4A). Principal Component Analysis (PCA) corroborated these findings, demonstrating a discernible but not exclusive clustering of colonizing and non-colonizing bacteria, indicating that while certain traits are influential, they do not entirely segregate colonizing from non-colonizing bacterial phenotypes (Figure S4B).

To gain insights into the genomic potential for mucin utilization, we analyzed genes involved in various degradation pathways predicted by Biocyc[39] (Table-1) in both colonized and non-colonized bacteria, including genomes of known natural microbiome members of the *C. elegans* intestine for comparison. The analysis revealed an increase in frequency of hits for the greatest Bitscore bins only for colonized bacteria for the 3 degradation pathways except Fucose (Figure 1E and S1E). Additionally, we also checked for cell attachment related genes (Table-2) in the bacterial genome and saw a similar increase in frequency of hits (Figure 1F).

To study the limitations and extent to which we can use *C. elegans* and the standardized protocol to study colonization and colonization resistance, we selected *B. longum* which is previously demonstrated as a symbiont in multiple studies[40] which also showed a 95% inhibition in our invitro colonization resistance screen. We utilized *B. longum* and *C. difficile* strains, to assess colonization patterns and examine physiological differences using our colonization method. Successful colonization was verified using two different approaches; CFU enumeration, and bacterial tagging and imaging (Figure 1G). CFU enumeration showed successful colonization of both *B. longum* and *C. difficile*. Notably, we observed bacterial colonies with counts exceeding 1000 CFU per worm during the five days post-colonization assessment (Figure 1H). Furthermore, we labeled all three bacteria with Carboxyfluorescein diacetate succinimidyl ester (CFDA-SE) before introducing them to *C. elegans* for colonization. Subsequently, we conducted imaging of the *C. elegans* gut immediately after colonization of *B. longum*, *E. coli OP50* and *C. difficile* and successfully visualized within the intestine (Figure 1I). These findings affirm that our standardized protocol effectively facilitates the colonization of human anaerobic gut bacteria within the *C. elegans* intestine. These results also emphasize the importance of carbon utilization, particularly sugars associated with host glycoproteins, in governing bacterial colonization within the *C. elegans* intestine.

B. longum pre-colonized worms showed improved gut integrity, rescued lifespan, and colonization resistance against _C. difficile_.

Despite having a standardized protocol for colonizing anaerobic human gut isolates in *C. elegans*, the optimal utilization of this model lies in its potential for *in vivo* screening of colonization resistance. We made slight modifications to the existing protocol by introducing an additional step where *C. difficile* colonization was initiated a day after the colonization of

selected top 10 bacteria which showed more than 60% inhibition invitro (Figure 1A). Subsequently, we quantified the amount of *C. difficile* using selective media (Figure 2A). The enumeration results revealed a significant decrease in the total CFU load for *B. thetaiotaomicron, E. tayi, B. adolescentis, B. eggerthi,* and *B. longum* compared to the control *C. difficile* (Figure 2B). To dig deeper into the mechanism of colonization resistance, we selected *B. longum* as the representative bacteria and check for the CFU at different timepoint compared to the control *C. difficile*.

To understand the mechanisms behind the colonization resistance mediated by *B. longum*, we investigated whether it was associated with the activation of innate immune genes. Gene expression analysis was done for the genes involved in the major innate immune pathways (Figure S5A) in *C. elegans* and showed a significant increase in Scav-6 and FSHR-1 gene expression when compared to *C. difficile*-only treated controls (Figure 2D). However, aside from these genes, there was no significant increase in the expression of other genes, suggesting that the inhibition of *C. difficile* by *B. longum* is not primarily due to the activation of innate immune genes. We also didn't notice significant increase of gene expression for *C. difficile* only control worms (Figure S6A).

Furthermore, we examined whether pre-colonization with *B. longum* in *C. elegans* could protect against *C. difficile*-induced disease pathology. One of the hallmark features of *C. difficile* infection (CDI), observed in various hosts, including humans, is a phenomenon called leaky gut[41]. We assessed leaky gut behavior in *C. elegans* using the Smurf assay. We observed a notable decrease in the percentage of smurf worms when *C. elegans* were pre-colonized with *B. longum* (Figure 2E, F). Lastly, we investigated the potential extension of lifespan or protection against mortality upon pre-colonization. Our findings indicated a reduced rate of death in pre-colonized worms (Figure 2G), clearly demonstrating that *C. elegans* can effectively recapitulate *B. longum*-mediated colonization resistance *in vivo*. Altogether, our results show the ability of *C. elegans* to be used as a model to study colonization resistance against *C. difficile* by human gut isolated anaerobic bacteria.

Continuous feeding of *C. elegans* with *B. longum* enhances, while *C. difficile* deteriorates, gut integrity, immune outcomes, and shows avoidance behavior.

Feeding of *C. elegans* by bacteria to understand the host-microbiome interactions is not a novel concept in *C. elegans* biology but as mentioned earlier most of these studies follow a continuous feeding regime. Although there may be variations in the larval stages at which feeding occurs, in terms of experimental methodology, most studies follow a similar approach. In essence, *C. elegans* are introduced onto nematode growth medium plates that have been seeded with the chosen bacteria, and they are allowed to grow under these conditions until they are ready for subsequent experiments (Figure 3A). In the case of *B. longum* and *C. difficile*, we employed a similar approach with a minor adjustment. We seeded both *B. longum* and *C. difficile* onto a NGM plate under anaerobic conditions and subsequently introduced *C. elegans* aerobically to study the host behavioral response post exposure (Figure 3A). Even though this is not a perfect solution to feed anaerobic bacteria, we assume to maintain some viability in the cells fed and seek to maintain the freshness of the metabolites produced by the bacteria. First, we looked for

the median survival time for *C. difficile* exposed *C. elegans*. We discovered that 50% of the *C. elegans* died by 57 hours significantly for *C. difficile R20291* exposure only, not for *B. longum* or control *E. coli OP50* (Figure 3B). This demonstrates the effective pathogenic behavior of *C. difficile* towards *C. elegans*. It has been previously established that the pathogenicity of *C. difficile* is attributed to the production of toxins. We also utilize a distinct strain of *C. difficile*, specifically strain 43598, a non-pathogenic strain[42] that exclusively produces only toxin B (*tcdB*). This choice allows us to complement the findings from the pathogenic R20291 strain of *C. difficile*, which possesses the capability to produce toxins A (*tcdA*), B (*tcdB*), and binary toxin (*CDT*)[43]. Our survival assay shows no changes in the viability of *C. elegans* by 57 hours upon feeding the non-pathogenic *C. difficile 43598*. The avoidance assay yielded noteworthy findings, with a considerably higher avoidance percentage observed for the pathogenic strain of *C. difficile R20291*. Surprisingly, we also observed a significant avoidance response to *C. difficile 43598* (Figure 3C, S8A). These results indicate that *C. elegans* identifies both toxigenic and non-toxigenic strains of *C. difficile* as potential pathogens, although lethality is significantly noted only with toxigenic strains. Additionally, further avoidance and lethality testing using culture supernatant and synthetic toxin A we found out that *C. elegans* avoid even the culture supernatant of *C. difficile* but showed a higher reduction in the viability on the presence of toxin A indicating that avoidance behavior might be the result of some metabolite produced by *C. difficile* (Figure S9A-C). Finally, it is worth noting that feeding *B. longum* had no discernible impact on the viability or avoidance behavior of *C. elegans*, which mirrors the results seen with *E. coli OP50* feeding.

Subsequently, we investigated whether the continuous feeding of *C. difficile*, as observed in this study, leads to any physiological variations like what has been observed in other model systems. We also observed a significant rise in leaky gut behavior among worms that were fed with the pathogenic strain of *C. difficile R20291* compared to those fed non-pathogenic *C. difficile 43598*, *B. longum* or *E. coli OP50* (Figure 3D).

Furthermore, we explored whether the feeding of *C. difficile* in *C. elegans* prompted any orthologous patterns in the nematode's immune response. We quantified gene expression levels at various time points following feeding *B. longum*, *C. difficile R20291*, and *C. difficile 43598*, compared to the control group fed with *E. coli OP50* (Figure 3E-G). We examined a total of 60 different genes related to major innate immune pathways in *C. elegans* (Figure S5A). Notably, in the case of pattern recognition, we observed an intriguing trend when *C. elegans* was exposed to anaerobic bacteria. Specifically, there was a significant increase in the *Scav-2* gene expression for all anaerobic bacteria compared to *E. coli OP50* (Figure 3E-G). Previously it has documented that the role of *Scav-5* in the pathogenicity of both *Pseudomonas aeruginosa* and *Salmonella typhimurium*[44]. We hypothesize that *C. elegans* may detect human anaerobic bacteria through the involvement of the Scav-2 receptor, although further experiments are needed to confirm this assumption. Furthermore, our analysis of gene expression data revealed a significant clustering pattern where we observed a distinct clustering for *B. longum* and *C. difficile R20291* indicating the difference in immune response as expected (Figure S8B).

Notably, in *C. elegans* fed with *B. longum*, there was a substantial decrease in Scav-5, and an increase in the *DAF-16* and *cnc-2* gene expression at the 48-hour mark. While the precise role of cnc-2 upregulation remains unclear, it's important to note that there were no significant increases in other innate immune responses within the same timeframe (Figure 3E). Increase in *DAF-16* expression was previously noted when *C. elegans* were fed with heat killed *B. longum* and was associated with the observed increase in longevity[45,46]. These findings underscore the symbiotic relationship between *B. longum* and *C. elegans*. In contrast, when *C. elegans* were fed with the toxigenic strain of *C. difficile*, there was a significant upregulation of almost all the major innate immune gene expression by the 48-hour point after infection (Figure S8C). To highlight there was a noteworthy increase in expression of genes related to the p38 MAPK pathway (Figure 3F, S8C). These observations were consistent with changes in gene expression previously observed during *C. difficile* colonization in mice and humans, highlighting the relevance of these findings to broader host-pathogen interactions[47,48]. We also saw an increase in expression for autophagy related genes in *C. elegans* fed with pathogenic *C. difficile* (Figure 3E). Recent articles have also highlighted a comparable occurrence in different model systems where the host responds to toxins produced by *C. difficile*[49,50]. Apart from this we also saw an increase in expression for *NLP-29, HLH-30, EGL-5, SCAV-3, CED-1, XBP-1, TOL-1, ELT-3, FSHR-1,* and *SKN-1* (Figure S8C). In the case of the non-toxigenic strain of *C. difficile*, we did observe some immune response at 12 hours post-infection (PI), but notably, there was no significant upregulation of innate immune responses at later time points (Figure 3G). This finding strongly suggests that the immune response observed when feeding the toxigenic strain of *C. difficile* is likely a result of the action of toxin-A or Binary toxin.

In summary, these results point to the activation of orthologous innate immune genes akin to those seen in other model systems, such as mice. Also, it opens a wide array of investigatory immune reactions that are not otherwise observed in other model systems.

Short-term anaerobic bacterial feeding induces differential but temperate responses in *C. elegans* gut integrity, toxicity, and immune dynamics.

Since we have a different method for colonization (short-term feeding) in *C. elegans* rather than continuous feeding, we next aimed to determine whether short-term feeding had a comparable physiological impact as observed in the case of continuous feeding. Initially, we assessed gut integrity following short-term feeding using the Smurf assay. We observed a similar trend, wherein *C. elegans* short-term fed with the toxigenic strain of *C. difficile* exhibited a higher percentage of leaky gut worms compared to control *E. coli OP50* fed worms (Figure 4A). As expected, we noticed a reduction in the percentage of leaky gut worm in short-term feeding when compared to continuous feeding. The phenomenon remained true for not only *C. difficile R20291* but for all the bacteria tested (Figure S10A). To further investigate the toxicity of *C. difficile* towards *C. elegans*, we conducted a lifespan analysis. Notably, we observed a much lower mortality rate for short term feeding compared to continuous feeding. Although the mortality appeared to be slow for *C. difficile R20291* short-term fed worms, they exhibited a higher death rate than those colonized with non-toxigenic *C. difficile*, *B. longum*, or *E. coli OP50*. Short-term feeding of *C. elegans* with *B. longum* showed a slight increase in the nematode's longevity

(Figure 4B). These findings suggest that colonization through partial exposure can recapitulate most physiological responses while reducing the overall toxicity in *C. elegans* leaving the room for experimental interventions.

Furthermore, we examined the expression of innate immune genes during short-term feeding. As anticipated, the changes in gene expression were not as pronounced or distinct compared to continuous colonization. The expression patten for all the 60 innate immune genes were compared against the expression of continues fed *C. elegans*. Our results indicate a significant negative to near zero correlation for all *B. longum*, *C. difficile R20291* and *C. difficile 43598* indicating a differential response. However, we did observe differences in the clustering of *B. longum* and *C. difficile* based on gene expression changes (Figure S10B). We also saw a unanimous decrease in expression for all the genes for short-term feeding compared to continues feeding (Figure S10F-G)

In the case of *C. elegans* colonized using short-term feeding with *B. longum*, there was an unexpected increase in some innate immune genes at 12 hours (Figure S10C). One potential explanation for this could be fungal contamination in the agar plates used for growth. To rule out this possibility, we repeated the experiment in biological triplicates and observed a similar trend of increased gene expression (Figure S10H). Notably, the Insulin-like signaling pathway showed upregulation during this period, with an increase in the expression of *DAF-16*, but not its inhibitor *DAF-2*. This compliments the earlier longevity results (Figure S10C) and suggests that the increase in *DAF-16* expression might be responsible for the observed increase in life span. However, it's worth noting that this expression change seemed to diminish at later time points. One hypothesis is that the data normalization with ΔΔCT values of *E. coli OP50* masked physiological differences and hence we need to have a better non-bacterial feed control for the study. Finally, we assessed gene expression during colonization by non-toxigenic *C. difficile 43598*, and we observed only a nearly significant difference in the *DBL-1* gene expression shortly after colonization (Figure S10D). Finally, we also investigated the variations in colonization and associated physiology when using different larval stages for *C. difficile* colonization. Notably, we observed that colonization with Larval 1 (L1) led to a significant increase in CFU loads within the intestine, while colonization with Larval 3 (L3) initially exhibited a reduction in colonization, eventually getting stabilized (Figure S7A). Interestingly, there were no discernible differences in the percentage of leaky gut observed during colonization at different life stages (Figure S7B). Furthermore, we noted an elevated death rate in *C. elegans* colonized at L1 compared to those colonized at L3, likely attributed to the higher colonization levels observed at the L1 stage (Figure S7C). These findings highlight the significance of the timing of colonization, as it influences both the extent of colonization and disease severity.

To further investigate this expressional change and its associated physiological response, we utilized a transgene mutant colonization approach. We colonized CF1139 ((pKL78) *daf16*::GFP + *rol-6*(su1006)), a strain with *DAF-16* tagged with a GFP fluorophore, with all four bacteria, and observed an increase in fluorescence intensity at 12, 24 and 48 hours post-colonization, specifically for *B. longum* short-term fed *C. elegans* (Figure 5A&D). These findings suggest the possibility of strong RNA-level transcriptional/translational control of *DAF-16* expression. For

the toxigenic strain of *C. difficile R20291*, we observed an increase in the expression of only two essential genes involved in antimicrobial defense, *LGG-1* and *GPA-12* (Figure S10D). *LGG-1* is part of an autophagy-related pathway[51], and *GPA-12* is a G protein-coupled receptor in the p38 MAPK pathway[52]. Both genes were significantly upregulated during continuous feeding (Figure 3F, S8C). This led us to hypothesize a similar level of RNA-level control in their expression. This hypothesis was further confirmed using a transgenic mutant of *LGG-1*. We colonized all four strains in *C. elegans* MAH236 (*lgg-1*p::gfp::lgg-1p + *odr-1*p::rfp), where *LGG-1* was tagged with GFP. Upon colonization, we observed an increase in *LGG-1*expression only when the worms were short-term fed with the toxigenic strain of *C. difficile R20291* and *C. difficile 43598* at 12- and 48-hours post-colonization (Figure 5B&E). Building on the previous result that demonstrated robust mRNA-level expression control for *DAF-16*, we investigated whether a similar control existed for *DAF-2* gene expression. We used the transgenic mutant QQ202 (cv20 (*daf-2* :: GFP)), where the *DAF-2* gene was tagged with GFP, and short-term feeding it with all four strains. The results confirmed our hypothesis, showing an increase in DAF-2 expression only for the toxigenic strain of *C. difficile R20291* at 12-, 24- and 48-hours post-colonization (Figure 5C&F).

These results clearly demonstrate distinct patterns of innate immune gene expression during colonization compared to continuous feeding. Furthermore, they provide strong evidence for mRNA-level control of expression in the Insulin-like signaling pathways in *C. elegans*, which becomes apparent only when the bacteria are short-term fed, not when they are continuously fed.

C. difficile colonization induces differential lipid accumulation in _C. elegans_.

Histopathology serves a pivotal role in infectious disease investigations, providing valuable insights into tissue-level host-pathogen interactions and allowing researchers to visualize and comprehend the structural and cellular alterations induced by infectious agents. This approach prompted us to explore whether a similar methodology could be applied to *C. elegans*, facilitating an in-depth examination of cellular changes through basic histology staining. After short-term feeding *C. elegans* with *E. coli* and *C. difficile*, we utilized hematoxylin and eosin staining on 10µm-thick sections and observed an intriguing disparity in the distribution of total lipids within *C. elegans* (Figure 6A). Specifically, short-term feeding by the toxigenic strain of *C. difficile R20291* led to a reduction in total lipids, suggesting a potential adaptive strategy employed by the nematodes in response to the reduced lifespan observed upon *C. difficile* colonization, as documented previously in *C. elegans*. To validate this observed lipid accumulation difference, we conducted Nile Red staining to visualize lipids and quantified staining intensity through microscopy. The result confirmed the observations from the histopathology section with higher and significant fluorescence intensity shown for *C. elegans* short-term fed with *C. difficile R20291* (Figure 6B). Altogether our histopathological sectioning approach open doors to the more pathological exploration in *C. elegans* has in microbiome studies.

B. longum-associated protective effect is recapitulated in a germ-free mouse model of _C. difficile_ pathogenesis.

While previous studies have demonstrated colonization resistance *in vivo*, they primarily employed conventional mouse models or different strains of *C. difficile*[53-55]. To bridge this gap in our studies, we sought to replicate our findings from *C. elegans* in a well-established model system of germ-free mice. We colonized these mice with *B. longum* and allowed stable colonization before gavaging them with *C. difficile* (Figure 7A). Stable colonization was noted by daily CFU counting of the feces (Figure S11A). Post *C. difficile* infection subsequent examination of cecal CFU loads revealed a significant but modest reduction in *C. difficile* levels following *B. longum* pre-colonization, compared to the control group without treatment (Figure 7B).

Upon euthanasia, we observed visible differences in cecal and colon architecture such as cecum and spleen size and necrosis features in the control group exposed to *C. difficile*, with a decrease in the *B. longum* pretreatment group (Figure 7C). We also measured colon length to investigate whether *B. longum* protects mice from *C. difficile*-induced colon shortening. Unfortunately, we did not observe a significant reduction in colon length compared to germ-free control, but intriguingly, *B. longum*-only colonized germ-free mice showed increased colon length (Figure 7D), possibly due to its ability to produce essential SCFA from dietary fibers as detected (Figure S12D-G).

Next, we assessed immune gene expression changes in response to *C. difficile* pathogenesis and *B. longum* pretreatment. We selected 14 different genes (Table 3) previously known to be activated during *C. difficile* pathogenesis[56]. We observed a significant increase in the expression of MAPK signaling genes, ultimately leading to the upregulation of both *IL-6* and *IL1r1*. However, this elevated expression was reduced or lost significance in *B. longum* pre-colonization (Figure 7E). This implies that pretreatment with *B. longum* significantly diminishes the inflammatory response linked to *C. difficile* pathogenesis and does not trigger gene activation in response to the pathogen. These observations are consistent with the orthologous gene expression patterns observed in *C. elegans*. More importantly we checked for improvement in cell architecture and inflammation from the histopathological slides of the whole colon from all the different treatments mentioned above and scored using the scheme adapted from[57]. We found a higher mucosal, submucosal and transmural inflammation with increased polymorphonuclear infiltrate and moderate increase in mononuclear infiltrate for *C. difficile* control mice. This morphology was very significantly diminished or not seen in the case of *B. longum* pretreated mice (Figure 7F&S11B). Furthermore, to elucidate the importance of *B. longum* in resisting *C. difficile* colonization, we analyzed human fecal metagenomic data from *C. difficile* infection (CDI) patient's databases. The results were intriguing and supported our hypothesis that *B. longum* plays a crucial role in colonization resistance against *C. difficile* in humans. *B. longum* abundance was significantly lower in CDI patients compared to healthy fecal controls, and this was further supported by an increase in abundance following fecal microbiota transplantation (FMT) treatment (Figure 7G).

Finally, we aimed to understand the mechanism of inhibition. We hypothesized that either *B. longum* produces antimicrobial peptides that inhibit *C. difficile* or it reduces the environmental pH, thereby inhibiting *C. difficile*.

To test the first hypothesis, we performed a coculture assay using cell-free culture supernatants of *B. longum*. This assay demonstrated successful inhibition of *C. difficile* for different fractions tested (Supplementary Figure 8C). Interestingly, the complete inhibition phenotype was reversed both for the coculture and treatment with cell free supernatant, when the media was supplemented with a buffering reagent maintaining a neutral pH (Figure S12B&C). We also measured the SCFA produced by *B. longum in vivo* in GF mice. Our results show that *B. longum* colonization both in control and *C. difficile* infection produced significantly higher amount of formic, propionic and valeric acid (Figure S12D-G). These results confirm that the inhibition is primarily due to *B. longum's* ability to acidify the environment. In conclusion, our findings indicate that *B. longum* can inhibit *C. difficile* in mice, corroborating the results obtained from the *C. elegans* trials.

Discussion

The quest to harness gut microbiota for the treatment of enteric infections has intensified with the expansion of the culturomics library[58,59], which presents a challenge in selecting appropriate bacterial candidates due to the increased diversity of potential species. To address the limitations of *in vitro* approaches, which may not replicate *in vivo* conditions, and the costly, time-intensive nature of widely used rodent-based *in vivo* models[60-62], our study proposes the nematode *C. elegans* as an alternative high-throughput model to investigate colonization resistance.

C. elegans are a model organism that have been explored from the late 1940s to understand host biology. Due to its microscopic nature *C. elegans* intestine can't be colonized with anaerobic bacteria by simple gavaging, unlike other models. The only way one could introduce bacteria into the intestine is by exposing the worms with the bacteria of interest and hoping *C. elegans* to feed on them. Since *C. elegans* is an aerobic organism, this method won't work efficiently with obligate anaerobic bacteria. Even though previous studies tried feeding *C. elegans* with anaerobic bacteria[45,46,63,64] no proper method has been established to colonize anaerobic bacterial isolates. Our avoidance assay findings with continuous exposure indicate that *C. elegans* can distinguish between pathogenic and non-pathogenic strains, responding specifically to pathogen-associated molecular patterns and associated toxin production (Figure 3). This is further evidenced by the differential survival outcomes and the induction of the smurf phenotype in *C. elegans* when exposed to toxigenic strains of *C. difficile*. Further analysis revealed a pronounced activation of the MAPK pathway in response to exposure by toxigenic strains, mirroring immune responses observed in human gut physiology[56]. This underscores the translational potential of *C. elegans*, despite the absence of an adaptive immune system in *C. elegans.*

For the first time, we have developed a pioneering protocol to facilitate the colonization of anaerobic human gut bacteria in *C. elegans* without continuous feeding (Figure 1). Our colonization protocol, involving short-term anaerobic exposure to mimic gavage techniques used in mammalian studies, successfully established colonization in *C. elegans* for a subset of gut microbiota. Genotypic and phenotypic exploration in understanding the differential colonization potential unveiled the ability of a bacteria to utilize mucin associated sugars such as N-acetyl galactosamine utilization and cell attachment factors as a critical determinant for colonization,

emphasizing the role of mucin as a nutritional and supportive factor for bacterial growth within the host. This was also proven to be true for some of the pathogen-based studies in *C. elegans*[65-68]. Similar Physiological effect on colonization showed a much milder and differential response in smurf behavior and life span analysis (Figure 4). The differential expression of innate immune genes and the attenuated decline in viability suggest that colonization elicits a more gradual physiological response compared to continuous bacterial feeding. But this attenuation in response in gene expression but a significant higher physiological response behavior forced us to ask whether there is any post translational control existing which is significantly effecting the change. Our transgene mutant-based colonization studies showed differential protein production for *DAF-2, DAF-16* and *LGG-1,* even though not seen in RT-PCR data. Clearly indication of a higher-level post transcriptional control existing otherwise was masked due to continues exposure-based overexpression of the gene transcript. This opens windows to explore new post transcriptional control mechanism or new control for the gene itself similar to what is previously known from *C. elegans*[69,70].

Our *in vivo* colonization resistance screen of to 10 C. difficile inhibiting bacteria showed a significant inhibition compared to the control for 5 bacteria. We also demonstrated in detail that *B. longum* can confer colonization resistance against *C. difficile in vivo*, as evidenced by reduced pathogen load and attenuation of adverse health outcomes in *C. elegans* in the form of gut integrity and immune expression. This resistance was further validated in a germ-free mouse model, reinforcing the translational relevance of our findings and the utility of *C. elegans* as a model for preclinical screening of bacterial therapeutics. Our results further corroborate the mechanism of inhibition of *B. longum* against vancomycin-resistant *Enterococcus faecium* by acidifying the lumen[71]. The inhibition was reversed when there was a buffering reagent in the media. This acidification is also shown to be very important as the abundance of *B. longum* is significantly reduced in CDI patients and increased when microbiome is replenished with FMT. This research not only advances our understanding of host-microbe interactions but also establishes a novel, efficient methodology for exploring colonization dynamics and resistance mechanisms in the context of gut microbiota. The limitation of the study is that in our colonization screening only half of them showed successful colonization, also we found that sugar metabolism and cell attachments are some of the factors that determined the colonization. We assume this factor dependencies change when introduced as a community or even the media in which the *C. elegans* are changed into a much more complex one, something that mimic the diet of any other models. Future studies could includes testing for the combination of colonized and non-colonized bacteria to study the colonization profile in a feeder media similar to what have been previously reported[72].

Methodology

Experimental model and growth

Caenorhabditis elegans Strains

Mutant transgenic *C. elegans* strains used in this study are presented in the Key Resources Table. The Bristol strain N2 was used as the wild-type *C. elegans* strain (Brenner, 1974). All strains were maintained with an *E. coli* OP50 diet on Nematode Growth Media (NGM). All animals were maintained at 20°C unless otherwise indicated.

Mice Strains

We purchased six-week-old C57BL/6 Germ-free (GF) mice from Taconic Biosciences. Inc (New York, USA). To avoid any bias related to gender differences, we included both male and female mice in our experiments wherever it was possible. All animals were maintained in a 12-hour light and dark cycle with ad-libitum autoclaved drinking water and a standard chow diet. All animal experiments were performed with the prior approval of the protocols (19-014A, 19-044A) by the South Dakota State University (SDSU) animal studies committee and under the regulations of the Institutional Animal Care and Use Committee (IACUC) and SDSU policies.

Bacterial Strains

All the bacterial strains used in this study are presented in the Key Resources Table. All the anaerobic strains were grown in a Coy anaerobic chamber purged with 5% Hydrogen. All the media for anaerobic strain growth was purged with Nitrogen until the indicator (Resazurin) turned colorless.

Method details

***C. elegans* growth conditions**

Nematodes were handled according to standard practices (Brenner, 1974). Worm strains were grown on NGM plates for experiments with vegetative cells. All strains were grown at 20°C unless otherwise indicated. For regular strain maintenance, OP50 at a concentration of 10x was used. For anaerobic bacterial colonization experiments, the same 10X concentration was seeded 0.5mL for 100mm plates and 0.05mL for 60mm plates and spread using an L-rod for forced feeding. For dead *E. coli* OP50, the pre-frozen OP50 pellet was heated at a temperature of 90 C in a water bath for 10 min, resuspended in M9 to a 10x concentration, and seeded 0.5mL for 100 and 0.05mL for 60mm plates, for more than three days of colonization or the life span assay NGM containing 50mM Fluroxydiuridine (FUDR) was used. For high through colonization, the worms were fed in anaerobic axenic media[73] containing specific bacteria of interest, as mentioned, then transferred to aerobic axenic media containing 10x dead *E. coli* OP50.

Bacterial growth conditions

All the anaerobic bacteria for the study were grown in modified Brain Heart Infusion (mBHI) broth, as mentioned previously in [74]. Briefly, the media was prepared by adding yeast extract (5.0 g/L), L-cysteine (0.3 g/L), and 1 ml/l Resazurine (0.25 mg/ml) to standard BHI base ingredients. The media was purged until anaerobic and was autoclaved at 121°C and 15 lbs pressure for 30 minutes. The media was then transferred to the anaerobic chamber and further supplemented with 1 ml/L of hemin solution (0.5mg/ml), 0.1 ml/L Menadione (58 mM), 10 ml of ATCC vitamin mixture (ATCC, USA), and 10 ml of ATCC mineral mixture (ATCC, USA). For making *Clostridium difficile* selective agar, mBHI was supplemented with 0.3g/L of D-cycloserine and 0.016g/L of Cefoxitin. *Bifidobacterium* selective Medium was made as per the manufacturer's

instruction. Finally, *E. coli* OP50 was grown in Luria Bertani media aerobically unless mentioned otherwise.

C. elegans gut bacterial colonization and enumeration

Colonization of anaerobic and aerobic bacteria was done similarly. The axenic L1 worms were generated as previously mentioned [75]. First, 8-10 of 3-4 days-old NGM plates were washed with 5mL sterilized M9 buffer. Next, the worms were washed twice with 10mL of M9 buffer. Worms were then treated with 5mL bleach solution (Bleach:5N NaoH:Water at a ratio 2:1:2) and vigorously shaken for 3 minutes. Finally, bleached eggs were washed twice with 10mL of M9 buffer, redissolved in 2mL of M9, and incubated in a rotor shaker at room temperature for ~10 hours.

Axenic worms were then anaerobically exposed to previously anaerobically seeded NGM plates with 500uL of 10x *B. longum*, *C. difficile* R20291, *C. difficile* 43598, or *E. coli* OP50 for 3 hours. After the exposure, plates were placed in aerobic condition for another 1 hour. The worms were transferred to a microcentrifuge tube with 1mL of M9 buffer, washed 3-4 times, and resuspended in ~100uL of M9. The worms were then transferred to previously seeded NGM plates with dead OP50. For the colonization experiment of more than 3 days, worms were picked and transferred onto NGM plates containing FUDR.

For enumeration, worms were transferred from the plates using 5mL of M9 and washed at least 8-10 times with 10mL of M9. The worms were then exposed to 5mL of 5mg/mL of vancomycin in M9 for 40-50 minutes. Worms were then washed thrice with M9 and redissolved in 200uL of M9. Next, pulverization was done, as previously mentioned, using a mortar and pestle [76,77]. Pulverized worms were then diluted in PBS and plated on Bifidobacterium selective medium (BSM) for *B. longum*, *Clostridium difficile* selective agar (CDSA) for *C. difficile* and Luria Bertani agar for *E. coli OP50*.

High throughput colonization: 10 mL of previously grown anaerobic bacteria, as mentioned in (Key resource Table) was pelleted down and resuspended in 1 ml of anaerobic axenic media. In a 96-well 1mL block containing 800 µl of axenic media,100 µl of bacterial suspension and 100 µl of worms (~1worm/uL) in anaerobic axenic media were added. The plate was sealed with a breathable membrane (Sigma) and incubated anaerobically for 3 hours at 25°C. Following this, the plate was aerated for 30-45 minutes in a laminar hood. Afterward, the worms were washed with 500 µl of aerobic axenic media 3-4 times and then resuspended in 500 µl. The worms were then treated with 500 µl of heat-killed OP50 in axenic media. The plate was resealed with a breathable membrane and incubated at 20°C for 24 hours. For bacterial isolation, the worms underwent multiple washings using 0.1% Triton X in M9 buffer, including treatment with 0.2% bleach in M9 for 6 minutes at 4°C. Finally, worms were disrupted with silicon grits, and the bacterial load was quantified by serial dilution and plating on mBHI agar, followed by incubation at 37°C overnight. For each batch of the experiments, *B. longum* and *E. coli* OP50 served as a positive control, and no bacteria treatment served as a negative control.

Colonization resistance in *C. elegans*

For colonization resistance, initial colonization of the top 10 *in vitro C. difficile* inhibitors/*B. longum* or dead *E. coli OP50* was done per the previously mentioned colonization protocol. After 24 hours of incubation in NGM plates with dead *E. coli* OP50, the worms were washed and re-

inoculated anaerobically with previously seeded NGM plates with *C. difficile* and fed for 3 hours anaerobically. After the plates were placed in an aerobic condition for another 1 hour, the worms were transferred into microfuge tubes washed 3-4 times with M9 and resuspended in ~100uL of M9. These worms were then transferred to NGM plates with dead OP50. Enumeration of *C. difficile* was done in CDSA, as previously mentioned.

In-vitro coculture inhibition and antimicrobial inhibition

Coculture inhibition was performed as previously mentioned in [74]. Briefly, for the antimicrobial inhibition assay, *B. longum* strains were grown for 24 hours in mBHI medium and then centrifuged at 10,000 g for 5 min. The supernatant was filter-sterilized using a 0.22 μ filter in the anaerobic chamber and diluted in a ratio 1:1 with the 1X mBHI medium. pH was adjusted to 6.8. An overnight culture of *C. difficile*R20291 was adjusted to OD600= 0.5. Twenty microliters of OD600 adjusted suspension was added to 1 ml of the 1:1 diluted cell-free supernatant and incubated for 24 hours in triplicate. *C. difficile*R20291 was grown in 50% mBHI diluted with anaerobic PBS as a positive control. After 24 hours, the cultures were serial diluted with anaerobic PBS, plated on CDSA, and incubated anaerobically for 24 hours for enumeration. For Heat treatment, the centrifuged supernatant was transferred to sterile tubes and pasteurized at 90 °C for 45 min (1 mL portions) suspended in a water-filled heating block. Pasteurized supernatant preparations were stored at 4 °C for near-term experiments or frozen at 20 °C. For Proteinase K, portions of supernatants were treated with proteinase K) at a concentration of 1mg mL-1 for 1 hour at 37°C. Following treatment, the enzyme was inactivated by the addition of phenylmethylsulfonyl fluoride.

Avoidance Behavior and death assay

Avoidance behavior was measured as mentioned previously [78] with slight modification. Briefly, NGM plates were anaerobically seeded with 5x concentration of *B. longum*, *C. difficile* R20291, *C. difficile* 43598, or *E. coli* OP50. The plates were then inoculated with ~30 worms onto the center of the previously seeded NGM plates. The number of worms inside and outside the bacterial lawn was counted at different life stages, and percentage avoidance was calculated. Dead worms / Crawled out worms were not included in the analysis. To evaluate the toxicity of strains, the number of dead/alive worms was counted at different life stages. Life stages for worms were different for different bacteria.

Life span Analysis

Life Span analysis was performed as mentioned previously [79] with slight modifications. Briefly, colonization of *B. longum*, *C. difficile* R20291, *C. difficile* 43598, or *E. coli* OP50 was done as discussed earlier and transferred to NGM plates containing heat-killed *E. coli* OP50. During the L4 stage of their life cycle, the 20-30 worms were transferred to NGM plates supplemented with 50mM FUDR and seeded with heat-killed *E. coli* OP50. The worms were then observed using the dissecting scope daily for their viability, and data were recorded. Worms that crawled out of the plate were omitted from the analysis, and dead worms were removed. Worms are picked up and transferred to new NGM with FUDR plates weekly to replenish the food.

Mice experiments

For simple consortium gnotobiotic mice, *B. longum/C. difficile* was used for inoculation. For *B. longum* pre-treatment, mice were gavaged with a volume of 200uL of 20x concentrated *B.*

longum / sterile saline freshly prepared for 2 days[80]. The mice were undisturbed for 7 days until stable colonization was observed. For *C. difficile* pathogenesis, we used 200uL of 10^4 CFU of *C. difficile R20291* for consecutive 2 days, 7 days post *B. longum*/saline gavage. A volume of 200uL of the bacterial mix was orally administered to the treatment group. Mice were sacrificed 5 days post *C. difficile* infection. We used a triplicate treatment group with 4-6 mice in each group, while a group of uninoculated mice was used as GF control.

Bacterial enumeration

Total bacterial count in the fecal samples collected at different time points was assessed to evaluate the bacterial load, which indicates the bacteria's colonization. Anaerobic mBHI/CDSA was used to enumerate the bacterial load used in this study.

RNA Isolation and RT-PCR

***C. elegans*:** RNA isolation was performed using Trizol reagent, as mentioned in [81], with slight modifications. First, continuous colonized worms were harvested at 24-hour intervals, washed off from the plate using 5 mL of M9, and transferred into a centrifuge tube. The worms were washed at least five times at 1500 rpm for 1 minute to remove any existing bacteria. Next, the worms were resuspended in 500uL of Trizol reagent and transferred to a microfuge tube. The tubes containing the worms were then snap-frozen in liquid nitrogen and thawed at 37°C at least ten times for the worms to lyse properly. The Trizol-containing worms were pipetted up and down until all the worms were sufficiently dissolved.

Next, 100uL of chloroform was added to the lysed Trizol reagent, vortexed properly, and incubated for 3 min at 25°C. The Trizol/chloroform mixture is centrifugated at 12500rpm for 15 min at 4°C. The aqueous phase was separated into a fresh tube, and 250uL of cold isopropanol was added. The mixture was vortexed and incubated for 10 min at 25°C. Centrifugation was done at 12500 rpm for 10 min at 4°C. The supernatant was discarded slowly by not disturbing the pellet at the bottom of the tube. The pellet was further washed with 500uL of 80% ethanol by centrifugation at 7500 rpm for 5 minutes at 4°C. The pellet was dried in air for 3 min to remove the residual ethanol. The pellet was then resuspended in 20uL of nuclease-free water, and the concentration was checked using an RNA HS assay kit in a Qubit 4 fluorimeter. RNA was stored at -80°C until used.

50 ng of total RNA (see above) was reverse transcribed according to the manufacturer's protocol (Protoscript II, NEB). The reaction was done in technical triplicates, and the product was pooled together and diluted with 80uL of NFW. Reverse transcribed cDNA was frozen at -20°C until further use. Q-PCR was done using Syber green PCR Master mix (Invitrogen) with a primer concentration of 10uM with 4uL of cDNA product. Q-PCR was done in Quantstudio Real-time system following the program: 50°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C for 15 sec, 55°C for 30 sec. We used two housekeeping genes, Actin and Peroxisomal membrane protein-related genes, for calculating the fold change of gene expression relative to worms fed with *E. coli OP50* using the $2^{-\Delta\Delta Ct}$ method.

Mice colon: To assess the localized response of host-microbe interaction resulting from the colonization of *C. difficile* and *B. longum*, 100mg of tissue from the colon was sampled and snap-frozen in liquid nitrogen. The samples were then saved at -80ºC until further processing.

For RNA extraction, we used a TRIzol® (Ambion, life technologies)- chloroform (Sigma-Aldrich) method. According to the manufacturer's protocol, a DNase treatment was performed using an RNase-Free DNase kit (Qiagen, Maryland) to remove any contaminant genomic DNA. The RNA quality and quantity were evaluated using NanoDrop One (Thermo Scientific) and saved at -80ºC until further use.

Complimentary DNA (cDNA) was prepared from 250 ng of total RNA using the First strand cDNA synthesis kit (New England BioLabs Inc.) as per the manufacturer's protocol. Host gene expressions were assessed by qRT-PCR using the 25 define immune-related genes. Reactions were prepared using Power SYBR® Green PCR Master mix (Applied Biosystems). PCR run was performed in an ABI7500 standard (Applied Biosystems) RT-PCR machine using the following cycling conditions: 95ºC for 10 mins, 40 cycles of 95ºC for 15 sec and 60ºC for 1 min.

Raw cycle threshold (C_T) values at a threshold of 0.15 were exported and then uploaded to the Qiagen data analysis center for further analysis. A C_T cut-off value of 30 was set for analysis. An average geometric mean of C_T values of two housekeeping genes - mouse beta-actin (Actb) was used for data normalization and ΔC_T calculation. Fold change and fold regulation in the gene expression were calculated using the $\Delta\Delta C_T$ method.

Imaging

CFDA-SE staining

CFDA-SE (5,6-carboxyfluorescein diacetate succinimidyl ester) staining was done as per the manufacturer's protocol. For colonization, the NGM plates were seeded with stained *B. longum*, *C. difficile* R20291, *C. difficile* 43598, or *E. coli* OP50 3 hours before worm inoculation anaerobically. Worms were then inoculated anaerobically, and after 3 hours, the plates were kept in aerobic condition for another 1 hour. Visualization of bacteria was done in the 4th hour. Agar pads were made on the microscopic slide using 2% agarose for fixing the worms. Then, 10uL of stained bacteria-treated worms were placed on the slides, and 5uL of 1mM levamisole was added to anesthetize the worms. The worms were covered with a coverslip and visualized under a fluorescent microscope with FITC filter illumination (419/516 Ex/Em).

Transgenic worms

Colonization and colonization resistance were done for the Transgenic worm, same as the wildtype mentioned before for MAH236, CF1139, and QQ202 *C. elegans* strains tagged with GFP. The worms were visualized 3, 24, and 48 hours after the colonization. Agar pads were made on microscopic slides using 2% agarose to fix the worms. 10uL of worms at 3 and 24 hours were placed on the slides, and for 48 hours, individual worms were picked. 5uL of 1mM levamisole was added to the pad to anesthetize the worms. The worms were covered with a coverslip and visualized under a fluorescent microscope with GFP filter illumination (395/509 Ex/Em).

Histopathology sectioning and analysis for *C. elegans*

Colonization and colonization resistance for the worms was done, as previously mentioned. 4 days post colonized worms were then fixed using buffered formalin and stained with Eosin. The worms were then washed thoroughly with M9 to remove excess stains. Finally, the worms were mixed with 5% agarose and poured onto a disposable mold. The mold was further filled with 5% agarose. The blocks were then cut into 10um sections and placed on a lysine-coated slide. This slide was imaged using a slide scanner (Motic).

Histopathology sectioning and analysis of mice colon

Post euthanasia, the complete gut was exteriorized, and the colon (from cecum to distal colon, without anus) was opened longitudinally. The surface was cleaned and flushed with cold PBS, and the tissue was rolled transversally with the luminal surface inside. The rolls were fixed using a wooden toothpick, preserved in formaldehyde for 24 hours, and placed in 100% ethanol until paraffin embedding. Tissues were embedded in paraffin and sectioned into 4 µm thick slices before staining with hematoxylin and eosin. A board-certified veterinary pathologist microscopically examined the slides, and the interpretations were based on standard histopathological morphologies. The pathologist, blinded to the treatment, compared the treatment sections to the controls. Histopathology was scored on a histomorphological scale from 0 to 3 (normal to severe abnormality) for each of the following: Extent of inflammation, mucosal hyperplasia, crypt architectural distortion, erosion, mononuclear infiltrate, polymorphonuclear and leukocyte infiltrate [82].

Nile Red staining

We performed Nile Red fat staining of post-colonized animals as mentioned previously [83]. Briefly, the worms were washed off from NGM Petri dishes with PBS + 0.01% Triton X-100 (PBST) into a 1.5 mL tube, targeting approximately 100 worms per strain. The samples were centrifuged at 25 × g for 1 minute, the supernatant was reduced to 0.1 mL, and the washing process was repeated until the supernatant appeared clear. The worms were then fixed with 600 µL of 40% isopropanol for 3 minutes at room temperature. A Nile Red staining solution was prepared by diluting a 0.5 mg/mL stock solution in 40% isopropanol. Following the fixation, the samples were centrifuged again, the supernatant reduced, and 600 µL of the staining solution was added, with an incubation period of 2 hours in the dark. Subsequently, the samples were centrifuged, the supernatant replaced with PBST and incubated for 30 minutes in the dark. Approximately 14 µL of the worm suspension was then mounted on a microscope slide, covered with a coverslip. The samples were imaged using a Zeiss LSM 980 with Airyscan 2 with a 20× objective and a red (514) filter to visualize the worms.

Bioinformatics Analysis

Phylogenetic tree and distance analysis

The whole genome for the bacteria used in the colonization trial, as mentioned previously [74], and the natural microbiome members [34] were used. Raw FASTQ files from the two studies were downloaded from the NCBI database, assembled using SPAdes (v3.15.5) [84], and annotated with Prokka (1.14.6) [85]. The 16S rRNA sequences were then extracted using Barrnap (0.9) [86]. For phylogenetic analysis, essential R libraries such as 'ape' [87]and 'vegan' [88]were utilized. The pre-

constructed phylogenetic tree was read using ‘read.tree’ from the 'ape' library, and cophenetic distances were computed to assess the phylogenetic relationships.

GC content analysis

For GC content analysis, a custom R function, calculate-GC-Content, was developed using the 'Biostrings' package [89]. This function read the sequences from FASTA files and calculated the GC content by aggregating guanine and cytosine nucleotide counts across all sequences, measuring the genomic GC proportion of the bacteria.

tBLASTn Analysis

The whole genome for the bacteria used in the colonization trial, as mentioned previously [74], and the natural microbiome members [34] were used. Raw FASTQ files from the two studies were downloaded from the NCBI database and assembled using SPAdes (v3.15.5) [84]. The FASTA file generated was converted into the BLAST database using makeblastdb (2.14.1) tool. Query gene sequences downloaded from the NCBI database (Table S1) were made onto a FASTA file and used as query sequences. Finally, we ran a tBLASTn (2.14.1) search using the specified query file, saving the results in a TSV format with detailed alignment information. The results were plotted as a frequency histogram without a cutoff using R.

Random forest analysis of Traitar data

Previously analyzed Traitar data from [74] was used for the analysis. For the Random Forest analysis, the 'randomForest' [90] package in R was employed, utilizing a dataset preprocessed to select relevant features. The colonization variable was transformed for binary classification (1 for colonized, 0 for non-colonized). A seed value of 123 was set for reproducibility before applying the Random Forest model, which consisted of 500 trees, ensuring robustness in the analysis. The algorithm was tasked with identifying feature importance, ranking variables based on their contribution to predictive accuracy, and helping understand their influence on bacterial colonization across the dataset.

Public CDI metagenome dataset analysis

To assess the association of *B. longum* in CDI conditions, an abundance mapping at the strain level was performed using publicly available gut metagenome sequencing data sets collected from research studies, as mentioned (Table S2). For all the metagenome reads, quality trimming and adapter clipping were done with Trimmomatic [91] . Further, the reads were aligned against the human genome to filter out human reads and were assembled with Bowtie2 v2.3.2 [92] and SAMtools [93]. The resulting contigs were classified taxonomically by k-mer analysis using Kraken2 [94], with Kraken 2 standard database. Subsequent species abundance estimation was performed using Kraken-tools [95] with a threshold set to ignore species with fewer than 10 classified reads.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA) or R. Comparisons between groups were analyzed using either the Mann-Whitney t-test or the t-test with Welch's correction, as mentioned in the figure legends. A p-value of < 0.05 was considered statistically significant. All illustrations are created with Biorender.com.

References

1. Nicholson, J.K., Holmes, E., and Wilson, I.D. (2005). Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol *3*, 431-438. 10.1038/nrmicro1152.
2. Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., and Gordon, J.I. (2008). Evolution of mammals and their gut microbes. Science *320*, 1647-1651. 10.1126/science.1155725.
3. Cani, P.D., Possemiers, S., Van de Wiele, T., Guiot, Y., Everard, A., Rottier, O., Geurts, L., Naslain, D., Neyrinck, A., Lambert, D.M., et al. (2009). Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. Gut *58*, 1091-1103. 10.1136/gut.2008.165886.
4. Hill, M.J. (1997). Intestinal flora and endogenous vitamin synthesis. Eur J Cancer Prev *6 Suppl 1*, S43-45. 10.1097/00008469-199703001-00009.
5. Kamada, N., Seo, S.U., Chen, G.Y., and Núñez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol *13*, 321-335. 10.1038/nri3430.
6. Kau, A.L., Ahern, P.P., Griffin, N.W., Goodman, A.L., and Gordon, J.I. (2011). Human nutrition, the gut microbiome and the immune system. Nature *474*, 327-336. 10.1038/nature10213.
7. Hooper, L.V., and Macpherson, A.J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. Nat Rev Immunol *10*, 159-169. 10.1038/nri2710.
8. Kostic, A.D., Gevers, D., Siljander, H., Vatanen, T., Hyötyläinen, T., Hämäläinen, A.M., Peet, A., Tillmann, V., Pöhö, P., Mattila, I., et al. (2015). The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. Cell Host Microbe *17*, 260-273. 10.1016/j.chom.2015.01.001.
9. Gevers, D., Kugathasan, S., Denson, L.A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M., et al. (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe *15*, 382-392. 10.1016/j.chom.2014.02.005.
10. Smith, P.A. (2015). The tantalizing links between gut microbes and the brain. Nature *526*, 312-314. 10.1038/526312a.
11. Geva-Zatorsky, N., Sefik, E., Kua, L., Pasman, L., Tan, T.G., Ortiz-Lopez, A., Yanortsang, T.B., Yang, L., Jupp, R., Mathis, D., et al. (2017). Mining the Human Gut Microbiota for Immunomodulatory Organisms. Cell *168*, 928-943.e911. 10.1016/j.cell.2017.01.022.
12. Britton, R.A., and Young, V.B. (2014). Role of the intestinal microbiota in resistance to colonization by Clostridium difficile. Gastroenterology *146*, 1547-1553. 10.1053/j.gastro.2014.01.059.
13. Lessa, F.C., Mu, Y., Bamberg, W.M., Beldavs, Z.G., Dumyati, G.K., Dunn, J.R., Farley, M.M., Holzbauer, S.M., Meek, J.I., Phipps, E.C., et al. (2015). Burden of Clostridium difficile infection in the United States. N Engl J Med *372*, 825-834. 10.1056/NEJMoa1408913.
14. Slimings, C., and Riley, T.V. (2014). Antibiotics and hospital-acquired Clostridium difficile infection: update of systematic review and meta-analysis. J Antimicrob Chemother *69*, 881-891. 10.1093/jac/dkt477.
15. George, W.L., Sutter, V.L., Goldstein, E.J., Ludwig, S.L., and Finegold, S.M. (1978). Aetiology of antimicrobial-agent-associated colitis. Lancet *1*, 802-803. 10.1016/s0140-6736(78)93001-5.
16. Quraishi, M.N., Widlak, M., Bhala, N., Moore, D., Price, M., Sharma, N., and Iqbal, T.H. (2017). Systematic review with meta-analysis: the efficacy of faecal microbiota transplantation for the treatment of recurrent and refractory Clostridium difficile infection. Aliment Pharmacol Ther *46*, 479-493. 10.1111/apt.14201.

17. Ianiro, G., Maida, M., Burisch, J., Simonelli, C., Hold, G., Ventimiglia, M., Gasbarrini, A., and Cammarota, G. (2018). Efficacy of different faecal microbiota transplantation protocols for Clostridium difficile infection: A systematic review and meta-analysis. United European Gastroenterol J *6*, 1232-1244. 10.1177/2050640618780762.
18. Rode, A.A., Chehri, M., Krogsgaard, L.R., Heno, K.K., Svendsen, A.T., Ribberholt, I., Helms, M., Engberg, J., Schønning, K., Tvede, M., et al. (2021). Randomised clinical trial: a 12-strain bacterial mixture versus faecal microbiota transplantation versus vancomycin for recurrent Clostridioides difficile infections. Aliment Pharmacol Ther *53*, 999-1009. 10.1111/apt.16309.
19. Tvede, M., Tinggaard, M., and Helms, M. (2015). Rectal bacteriotherapy for recurrent Clostridium difficile-associated diarrhoea: results from a case series of 55 patients in Denmark 2000-2012. Clin Microbiol Infect *21*, 48-53. 10.1016/j.cmi.2014.07.003.
20. Nigon, V., and Dougherty, E.C. (1949). Reproductive patterns and attempts at reciprocal crossing of Rhabditis elegans Maupas, 1900, and Rhabditis briggsae Dougherty and Nigon, 1949 (Nematoda: Rhabditidae). J Exp Zool *112*, 485-503. 10.1002/jez.1401120307.
21. Ferris, H., and Hieb, W.F. (2015). Ellsworth C. Dougherty: A Pioneer in the Selection of Caenorhabditis elegans as a Model Organism. Genetics *200*, 991-1002. 10.1534/genetics.115.178913.
22. Dougherty, E.C., and Hansen, E.L. (1956). Axenic cultivation of Caenorhabditis briggsae (Nematoda: Rhabditidae). V. Maturation on synthetic media. Proc Soc Exp Biol Med *93*, 223-227. 10.3181/00379727-93-22714.
23. Dimov, I., and Maduro, M.F. (2019). The C. elegans intestine: organogenesis, digestion, and physiology. Cell Tissue Res *377*, 383-396. 10.1007/s00441-019-03036-4.
24. Block, D.H., Twumasi-Boateng, K., Kang, H.S., Carlisle, J.A., Hanganu, A., Lai, T.Y., and Shapira, M. (2015). The Developmental Intestinal Regulator ELT-2 Controls p38-Dependent Immune Responses in Adult C. elegans. PLoS Genet *11*, e1005265. 10.1371/journal.pgen.1005265.
25. Ezcurra, M., Benedetto, A., Sornda, T., Gilliat, A.F., Au, C., Zhang, Q., van Schelt, S., Petrache, A.L., Wang, H., de la Guardia, Y., et al. (2018). C. elegans Eats Its Own Intestine to Make Yolk Leading to Multiple Senescent Pathologies. Curr Biol *28*, 2544-2556.e2545. 10.1016/j.cub.2018.06.035.
26. Peres, T.V., Arantes, L.P., Miah, M.R., Bornhorst, J., Schwerdtle, T., Bowman, A.B., Leal, R.B., and Aschner, M. (2018). Role of Caenorhabditis elegans AKT-1/2 and SGK-1 in Manganese Toxicity. Neurotox Res *34*, 584-596. 10.1007/s12640-018-9915-1.
27. Pukkila-Worley, R., and Ausubel, F.M. (2012). Immune defense mechanisms in the Caenorhabditis elegans intestinal epithelium. Curr Opin Immunol *24*, 3-9. 10.1016/j.coi.2011.10.004.
28. Berg, M., Stenuit, B., Ho, J., Wang, A., Parke, C., Knight, M., Alvarez-Cohen, L., and Shapira, M. (2016). Assembly of the Caenorhabditis elegans gut microbiota from diverse soil microbial environments. Isme j *10*, 1998-2009. 10.1038/ismej.2015.253.
29. Dirksen, P., Marsh, S.A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Petersen, C., Kowallik, V., Rosenstiel, P., et al. (2016). The native microbiome of the nematode Caenorhabditis elegans: gateway to a new host-microbiome model. BMC Biology *14*, 38. 10.1186/s12915-016-0258-1.
30. Samuel, B.S., Rowedder, H., Braendle, C., Félix, M.A., and Ruvkun, G. (2016). Caenorhabditis elegans responses to bacteria from its natural habitats. Proc Natl Acad Sci U S A *113*, E3941-3949. 10.1073/pnas.1607183113.
31. Ghimire, S., Roy, C., Wongkuna, S., Antony, L., Maji, A., Keena, M.C., Foley, A., and Scaria, J. (2020). Identification of Clostridioides difficile-Inhibiting Gut Commensals Using Culturomics,

Phenotyping, and Combinatorial Community Assembly. mSystems *5*. 10.1128/mSystems.00620-19.

32. Branicky, R., and Hekimi, S. (2006). What keeps C. elegans regular: the genetics of defecation. Trends Genet *22*, 571-579. 10.1016/j.tig.2006.08.006.

33. Iwasaki, S., Tsuruoka, N., Hattori, A., Sato, M., Tsujimoto, M., and Kohno, M. (1995). Distribution and characterization of specific cellular binding proteins for bone morphogenetic protein-2. J Biol Chem *270*, 5476-5482. 10.1074/jbc.270.10.5476.

34. Dirksen, P., Assie, A., Zimmermann, J., Zhang, F., Tietje, A.M., Marsh, S.A., Felix, M.A., Shapira, M., Kaleta, C., Schulenburg, H., and Samuel, B.S. (2020). CeMbio - The Caenorhabditis elegans Microbiome Resource. G3 (Bethesda) *10*, 3025-3039. 10.1534/g3.120.401309.

35. Hellweger, F.L., Huang, Y., and Luo, H. (2018). Carbon limitation drives GC content evolution of a marine bacterium in an individual-based genome-scale model. The ISME Journal *12*, 1180-1187. 10.1038/s41396-017-0023-7.

36. Allman, E., Johnson, D., and Nehrke, K. (2009). Loss of the apical V-ATPase a-subunit VHA-6 prevents acidification of the intestinal lumen during a rhythmic behavior in C. elegans. Am J Physiol Cell Physiol *297*, C1071-1081. 10.1152/ajpcell.00284.2009.

37. Reimer, L.C., Sardà Carbasse, J., Koblitz, J., Ebeling, C., Podstawka, A., and Overmann, J. (2021). BacDive in 2022: the knowledge base for standardized bacterial and archaeal data. Nucleic Acids Research *50*, D741-D746. 10.1093/nar/gkab961.

38. Ng, K.M., Pannu, S., Liu, S., Burckhardt, J.C., Hughes, T., Treuren, W.V., Nguyen, J., Naqvi, K., Nguyen, B., Clayton, C.A., et al. (2023). Single-strain behavior predicts responses to environmental pH and osmolality in the gut microbiota. mBio *14*, e00753-00723. doi:10.1128/mbio.00753-23.

39. Caspi, R., Billington, R., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P.E., Ong, Q., Ong, W.K., et al. (2017). The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Research *46*, D633-D639. 10.1093/nar/gkx935.

40. Floch, M.H., Ringel, Y., and Walker, W.A. (2016). The microbiota in gastrointestinal pathophysiology: implications for human health, prebiotics, probiotics, and dysbiosis (Academic Press).

41. Koenigsknecht, M.J., Theriot, C.M., Bergin, I.L., Schumacher, C.A., Schloss, P.D., and Young, V.B. (2015). Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. Infect Immun *83*, 934-941. 10.1128/IAI.02768-14.

42. Depitre, C., Delmee, M., Avesani, V., L'Haridon, R., Roels, A., Popoff, M., and Corthier, G. (1993). Serogroup F strains of Clostridium difficile produce toxin B but not toxin A. J Med Microbiol *38*, 434-441. 10.1099/00222615-38-6-434.

43. Rupnik, M., Avesani, V., Janc, M., von Eichel-Streiber, C., and Delmée, M. (1998). A novel toxinotyping scheme and correlation of toxinotypes with serogroups of Clostridium difficile isolates. J Clin Microbiol *36*, 2240-2247. 10.1128/jcm.36.8.2240-2247.1998.

44. Luo, A., Jing, H., Yuan, L., Wang, Y., Xiao, H., and Zheng, Q. (2021). Loss of Function of Scavenger Receptor SCAV-5 Protects C. elegans Against Pathogenic Bacteria. Front Cell Infect Microbiol *11*, 593745. 10.3389/fcimb.2021.593745.

45. Zhao, L., Zhao, Y., Liu, R., Zheng, X., Zhang, M., Guo, H., Zhang, H., and Ren, F. (2017). The Transcription Factor DAF-16 is Essential for Increased Longevity in C. elegans Exposed to Bifidobacterium longum BB68. Sci Rep *7*, 7408. 10.1038/s41598-017-07974-3.

46. Sugawara, T., and Sakamoto, K. (2018). Killed Bifidobacterium longum enhanced stress tolerance and prolonged life span of Caenorhabditis elegans via DAF-16. Br J Nutr *120*, 872-880. 10.1017/S0007114518001563.

47. Meyer, G.K., Neetz, A., Brandes, G., Tsikas, D., Butterfield, J.H., Just, I., and Gerhard, R. (2007). Clostridium difficile toxins A and B directly stimulate human mast cells. Infect Immun *75*, 3868-3876. 10.1128/iai.00195-07.
48. Warny, M., Keates, A.C., Keates, S., Castagliuolo, I., Zacks, J.K., Aboudola, S., Qamar, A., Pothoulakis, C., LaMont, J.T., and Kelly, C.P. (2000). p38 MAP kinase activation by Clostridium difficile toxin A mediates monocyte necrosis, IL-8 production, and enteritis. J Clin Invest *105*, 1147-1156. 10.1172/JCI7545.
49. He, R., Peng, J., Yuan, P., Yang, J., Wu, X., Wang, Y., and Wei, W. (2017). Glucosyltransferase Activity of Clostridium difficile Toxin B Triggers Autophagy-mediated Cell Growth Arrest. Scientific Reports *7*, 10532. 10.1038/s41598-017-11336-4.
50. Amirkamali, S., Azimirad, M., Nasiri, G., Goudarzi, H., Noori, M., Yadegar, A., Ghalavand, Z., and Zali, M.R. (2022). Surface layer protein A from hypervirulent Clostridioides difficile ribotype 001 can induce autophagy process in human intestinal epithelial cells. Microb Pathog *169*, 105681. 10.1016/j.micpath.2022.105681.
51. Meléndez, A., Tallóczy, Z., Seaman, M., Eskelinen, E.L., Hall, D.H., and Levine, B. (2003). Autophagy genes are essential for dauer development and life-span extension in C. elegans. Science *301*, 1387-1391. 10.1126/science.1087782.
52. Ziegler, K., Kurz, C.L., Cypowyj, S., Couillault, C., Pophillat, M., Pujol, N., and Ewbank, J.J. (2009). Antifungal innate immunity in C. elegans: PKCdelta links G protein signaling and a conserved p38 MAPK cascade. Cell Host Microbe *5*, 341-352. 10.1016/j.chom.2009.03.006.
53. Yun, B., Song, M., Park, D.J., and Oh, S. (2017). Beneficial Effect of Bifidobacterium longum ATCC 15707 on Survival Rate of Clostridium difficile Infection in Mice. Korean J Food Sci Anim Resour *37*, 368-375. 10.5851/kosfa.2017.37.3.368.
54. Valdes-Varela, L., Hernandez-Barranco, A.M., Ruas-Madiedo, P., and Gueimonde, M. (2016). Effect of Bifidobacterium upon Clostridium difficile Growth and Toxicity When Co-cultured in Different Prebiotic Substrates. Front Microbiol *7*, 738. 10.3389/fmicb.2016.00738.
55. Wei, Y., Yang, F., Wu, Q., Gao, J., Liu, W., Liu, C., Guo, X., Suwal, S., Kou, Y., Zhang, B., et al. (2018). Protective Effects of Bifidobacterial Strains Against Toxigenic Clostridium difficile. Front Microbiol *9*, 888. 10.3389/fmicb.2018.00888.
56. Pechine, S., and Collignon, A. (2016). Immune responses induced by Clostridium difficile. Anaerobe *41*, 68-78. 10.1016/j.anaerobe.2016.04.014.
57. Drolia, R., Tenguria, S., Durkes, A.C., Turner, J.R., and Bhunia, A.K. (2018). Listeria Adhesion Protein Induces Intestinal Epithelial Barrier Dysfunction for Bacterial Translocation. Cell Host Microbe *23*, 470-484.e477. 10.1016/j.chom.2018.03.004.
58. Vanstokstraeten, R., Demuyser, T., Ellen, C., Blockeel, C., Tournaye, H., De Geyter, D., Muyldermans, A., Wybo, I., Piérard, D., and Mackens, S. (2022). P-321 Culturomics-generated vaginal and endometrial microbiome profiles in subfertile patients. Human Reproduction *37*, deac107. 306.
59. Huang, Y., Sheth, R.U., Zhao, S., Cohen, L.A., Dabaghi, K., Moody, T., Sun, Y., Ricaurte, D., Richardson, M., and Velez-Cortes, F. (2023). High-throughput microbial culturomics using automation and machine learning. Nature Biotechnology, 1-10.
60. Nguyen, T.L., Vieira-Silva, S., Liston, A., and Raes, J. (2015). How informative is the mouse for human gut microbiota research? Dis Model Mech *8*, 1-16. 10.1242/dmm.017400.
61. Hugenholtz, F., and de Vos, W.M. (2018). Mouse models for human intestinal microbiota research: a critical evaluation. Cell Mol Life Sci *75*, 149-160. 10.1007/s00018-017-2693-8.
62. Kennedy, E.A., King, K.Y., and Baldridge, M.T. (2018). Mouse microbiota models: comparing germ-free mice and antibiotics treatment as tools for modifying gut bacteria. Frontiers in physiology *9*, 1534.

63. Yun, B., Oh, S., Song, M., Hong, Y.S., Park, S., Park, D.J., Griffiths, M.W., and Oh, S. (2015). Inhibitory Effect of Epigallocatechin Gallate on the Virulence of Clostridium difficile PCR Ribotype 027. J Food Sci *80*, M2925-2931. 10.1111/1750-3841.13145.
64. Walker, A.C., Bhargava, R., Vaziriyan-Sani, A.S., Pourciau, C., Donahue, E.T., Dove, A.S., Gebhardt, M.J., Ellward, G.L., Romeo, T., and Czyż, D.M. (2021). Colonization of the Caenorhabditis elegans gut with human enteric bacterial pathogens leads to proteostasis disruption that is rescued by butyrate. PLOS Pathogens *17*, e1009510. 10.1371/journal.ppat.1009510.
65. Hoffman, C.L., Lalsiamthara, J., and Aballay, A. (2020). Host Mucin Is Exploited by Pseudomonas aeruginosa To Provide Monosaccharides Required for a Successful Infection. mBio *11*. 10.1128/mBio.00060-20.
66. He, Y., Hao, F., Fu, H., Tian, G., Zhang, Y., Fu, K., and Qi, B. (2023). N-glycosylated intestinal protein BCF-1 shapes microbial colonization by binding bacteria via its fimbrial protein. Cell Rep *42*, 111993. 10.1016/j.celrep.2023.111993.
67. Head, B.P., Olaitan, A.O., and Aballay, A. (2017). Role of GATA transcription factor ELT-2 and p38 MAPK PMK-1 in recovery from acute P. aeruginosa infection in C. elegans. Virulence *8*, 261-274. 10.1080/21505594.2016.1222334.
68. Troemel, E.R., Chu, S.W., Reinke, V., Lee, S.S., Ausubel, F.M., and Kim, D.H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity in C. elegans. PLoS Genet *2*, e183. 10.1371/journal.pgen.0020183.
69. Malin, J.A., Kinet, M.J., Abraham, M.C., Blum, E.S., and Shaham, S. (2016). Transcriptional control of non-apoptotic developmental cell death in C. elegans. Cell Death & Differentiation *23*, 1985-1994. 10.1038/cdd.2016.77.
70. Maheshwari, R., Pushpa, K., and Subramaniam, K. (2016). A role for post-transcriptional control of endoplasmic reticulum dynamics and function in C. elegans germline stem cell maintenance. Development *143*, 3097-3108. 10.1242/dev.134056.
71. Odenwald, M.A., Lin, H., Lehmann, C., Dylla, N.P., Cole, C.G., Mostad, J.D., Pappas, T.E., Ramaswamy, R., Moran, A., Hutchison, A.L., et al. (2023). Bifidobacteria metabolize lactulose to optimize gut metabolites and prevent systemic infection in patients with liver disease. Nature Microbiology *8*, 2033-2049. 10.1038/s41564-023-01493-w.
72. Samuel, T.K., Sinclair, J.W., Pinter, K.L., and Hamza, I. (2014). Culturing Caenorhabditis elegans in axenic liquid media and creation of transgenic worms by microparticle bombardment. J Vis Exp, e51796. 10.3791/51796.
73. Ortiz, A., Vega, N.M., Ratzke, C., and Gore, J. (2021). Interspecies bacterial competition regulates community assembly in the C. elegans intestine. ISME J *15*, 2131-2145. 10.1038/s41396-021-00910-4.
74. Ghimire, S., Roy, C., Wongkuna, S., Antony, L., Maji, A., Keena, M.C., Foley, A., and Scaria, J. (2020). Identification of Clostridioides difficile-inhibiting gut commensals using culturomics, phenotyping, and combinatorial community assembly. Msystems *5*, e00620-00619.
75. Stiernagle, T. (2006). Maintenance of C. elegans. WormBook, 1-11. 10.1895/wormbook.1.101.1.
76. Taylor, M.N., Spandana Boddu, S., and Vega, N.M. (2022). Using Single-Worm Data to Quantify Heterogeneity in Caenorhabditis elegans-Bacterial Interactions. J Vis Exp. 10.3791/64027.
77. Taylor, M., and Vega, N.M. (2021). Host Immunity Alters Community Ecology and Stability of the Microbiome in a Caenorhabditis elegans Model. mSystems *6*. 10.1128/mSystems.00608-20.
78. Meisel, J.D., and Kim, D.H. (2014). Behavioral avoidance of pathogenic bacteria by Caenorhabditis elegans. Trends Immunol *35*, 465-470. 10.1016/j.it.2014.08.008.
79. Cornwell, A.B., and Samuelson, A.V. (2020). Analysis of Lifespan in C. elegans: Low- and High-Throughput Approaches. Methods Mol Biol *2144*, 7-27. 10.1007/978-1-0716-0592-9_2.

80. Ambat, A., Antony, L., Maji, A., Ghimire, S., Mattiello, S., Kashyap, P., Sebastian, V., and Scaria, J. (2023). Enhancing Recovery from Gut Microbiome Dysbiosis and Alleviating DSS-Induced Colitis in Mice with a Consortium of Rare Short-Chain Fatty Acid-Producing Bacteria. bioRxiv, 2023.2009. 2011.556543.
81. Chomczynski, P., and Mackey, K. (1995). Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. Biotechniques *19*, 942-945.
82. Drolia, R., Tenguria, S., Durkes, A.C., Turner, J.R., and Bhunia, A.K. (2018). Listeria Adhesion Protein Induces Intestinal Epithelial Barrier Dysfunction for Bacterial Translocation. Cell Host Microbe *23*, 470-484 e477. 10.1016/j.chom.2018.03.004.
83. Stuhr, N.L., Nhan, J.D., Hammerquist, A.M., Van Camp, B., Reoyo, D., and Curran, S.P. (2022). Rapid Lipid Quantification in Caenorhabditis elegans by Oil Red O and Nile Red Staining. Bio Protoc *12*, e4340. 10.21769/BioProtoc.4340.
84. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol *19*, 455-477. 10.1089/cmb.2012.0021.
85. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068-2069. 10.1093/bioinformatics/btu153.
86. Seemann, T. (2015). Barrnap 0.9 – Bacterial ribosomal RNA predictor. https://mybiosoftware.com/barrnap-bacterial-ribosomal-rna-predictor.html.
87. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics *20*, 289-290. 10.1093/bioinformatics/btg412.
88. Oksanen J, S.G., Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R,, Solymos P, S.M., Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B,, Borcard D, C.G., Chirico M, De Caceres M, Durand S, Evangelista H,, FitzJohn R, F.M., Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D,, and Ouellette M, R.C.E., Smith T, Stier A, Ter Braak C, Weedon J (2022). vegan: Community Ecology Package_. R package version 2.6-4,.
89. Lifschitz, S., Haeusler, E.H., Catanho, M., de Miranda, A.B., Molina de Armas, E., Heine, A., Moreira, S.G., and Tristão, C. (2022). Bio-Strings: A Relational Database Data-Type for Dealing with Large Biosequences. BioTech *11*, 31.
90. Breiman, L. (2001). Random forests. Machine learning *45*, 5-32.
91. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.
92. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.
93. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The sequence alignment/map format and SAMtools. bioinformatics *25*, 2078-2079.
94. Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. Genome Biol *20*, 257. 10.1186/s13059-019-1891-0.
95. Davis, M.P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N., and Enright, A.J. (2013). Kraken: a set of tools for quality control and analysis of high-throughput sequence data. Methods *63*, 41-49.

Figures

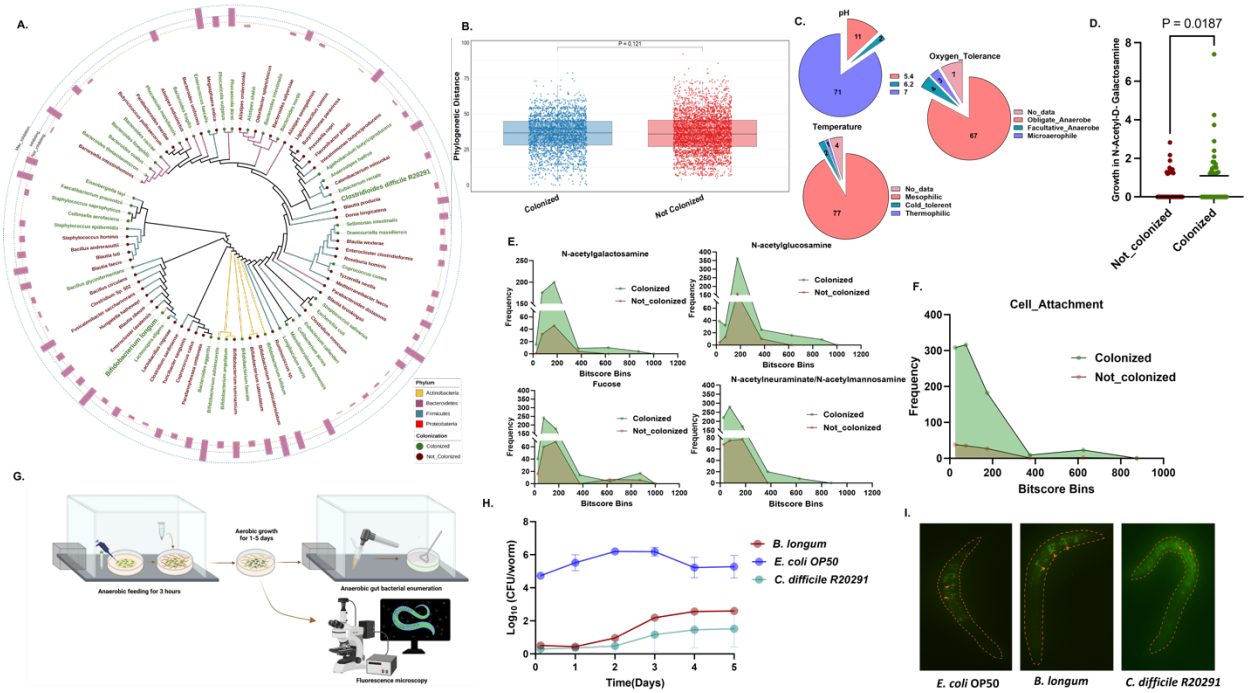


Figure 1. Carbohydrate degradation capability correlates with successful anaerobic bacterial colonization in *C. elegans*.

A. Displaying a phylogenetic tree derived from 16s rRNA sequence alignment of 82 bacterial species from the human gut and *C. difficile R20291*. Branch colors indicate phylum classification, while circle/text colors represent colonization ability. Surrounding bar graph depicts *C. difficile* inhibition percentages from previous study[1], with dotted lines at 100%, 50%, and 20% inhibition thresholds.

B. Median phylogenetic distances between colonized and non-colonized bacteria and natural microbiome members of *C. elegans*[2], visualized as a scatter plot. Distance calculations are based on pairwise comparisons, with statistical analyses (two-tailed unpaired t-tests) indicated.

C. Pie chart showing the distribution of bacteria based on tolerance to varying pH levels, oxygen, and temperature conditions, informed by predictions for type strains[3] and supplementary pH data sources[4].

D. Comparison of growth between colonized and non-colonized bacteria in minimal media containing N-Acetyl-D-Galactosamine, measured via OD (600nm) and normalized against a water blank. The median line and statistical significance (two-tailed unpaired t-tests) are highlighted.

E. Histogram representing the frequency distribution of bit scores from tblastn for genes involved in the metabolism of N-acetylgalactosamine, N-acetylglucosamine, Fucose, N-acetylneuraminate/N-acetylmannoseamine, and

F. Cell attachment in the colonized and non-colonized bacterial genome sets, including genome data from natural *C. elegans* microbiome isolates[2].

G. Graphical representation of the methodology for colonizing *C. elegans* with *B. longum* and *C. difficile*, including details of feeding protocols, worm transfer, and colonization assessment techniques. Created with BioRender.com.
H. Log_{10} CFU/worm graph showing the bacterial counts for *B. longum*, *C. difficile R20291*, and *E. coli OP50* colonized *C. elegans* at various time intervals, with mean ± SEM values from biological triplicates.
I. Fluorescence microscopy image displaying CFDA-SE stained bacteria in the *C. elegans* intestine post-anaerobic feeding, with arrows indicating colonized bacteria.

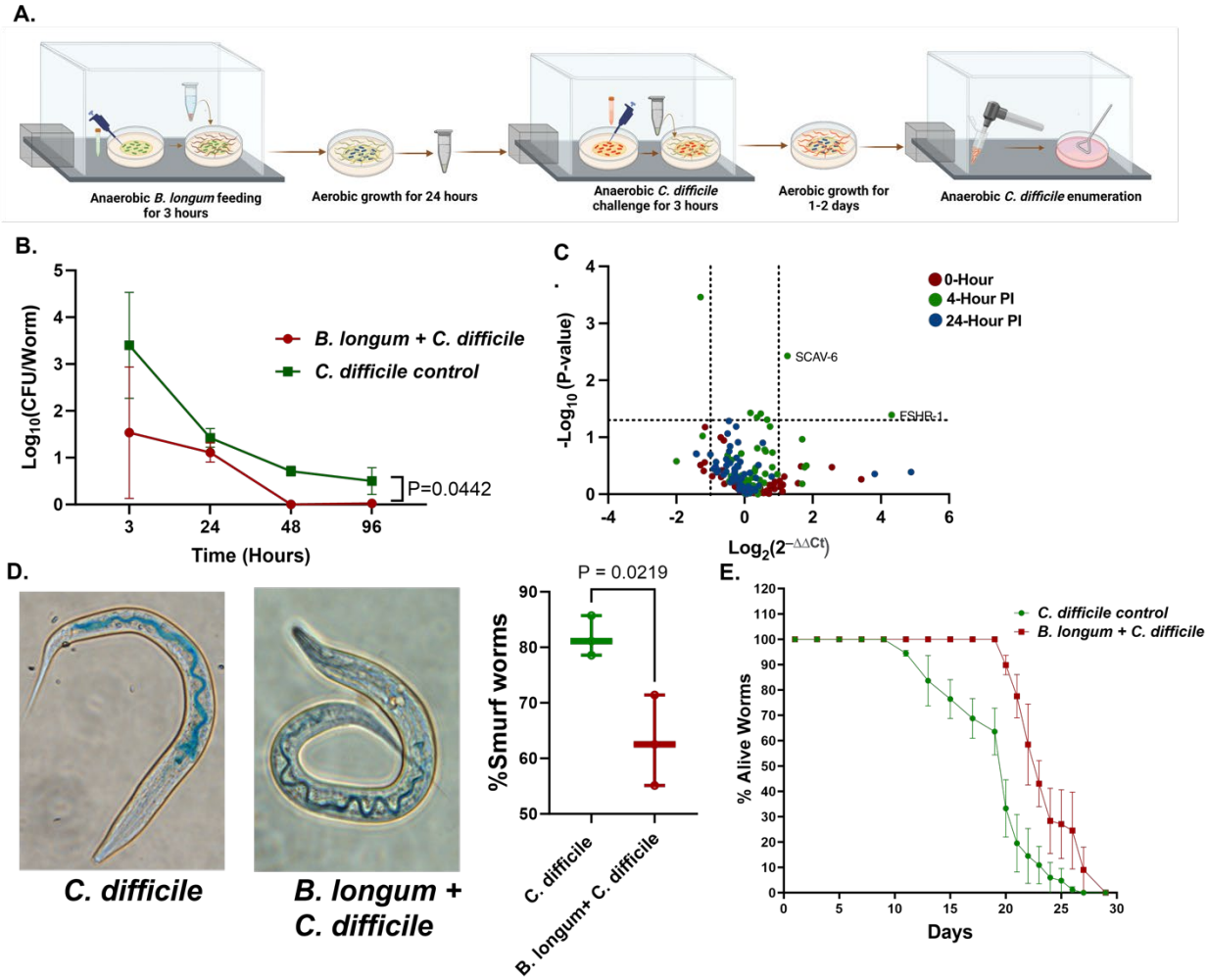


Figure 2. *B. longum* protects *C. elegans* from *C. difficile*-related gut damage and reduced lifespan.
A. Graphical Representation of the methodology used for assessing colonization resistance in *C. elegans* against *B. longum* and *C. difficile*. It includes detailed feeding protocols, procedures for worm transfer, and various techniques employed to evaluate colonization. Created with BioRender.com.

B. Log_{10} CFU/Worm graph showing the counts of *C. difficile* in *C. elegans* that were pre-colonized with *B. longum* or fed control dead *E. coli OP50*, followed by *C. difficile R20291* infection. The data is represented at various time intervals, with mean ± SEM values from biological triplicates. Statistical significance is determined using unpaired students t-tests.
C. Volcano Plot displays gene expression data quantified using 62 innate immune-associated genes in *C. elegans* post *C. difficile* feeding for *B. longum* pre colonized worms at 0, 4, and 24 hours. Fold change is calculated using the $2^{\Delta\Delta}Ct$ method with *E. coli OP50* fed and *C. difficile* infected expression data as control, and P-Value is calculated using unpaired students t-tests.
D. Representative images of *C. elegans* intestine, focusing on the display of smurf behavior, and Mean ± SEM values for the percentage of smurf worms in *C. elegans* colonized with either *B. longum* or fed with *dead E. coli OP50* and subsequently infected with *C. difficile R20291*, 24 hours post-infection. The P-Value is calculated using unpaired students t-tests.
E. Mean ± SEM percentage of live worms over time for both control groups – those pre-colonized with *B. longum* and those fed with *E. coli OP50*, following infection with *C. difficile R20291*.

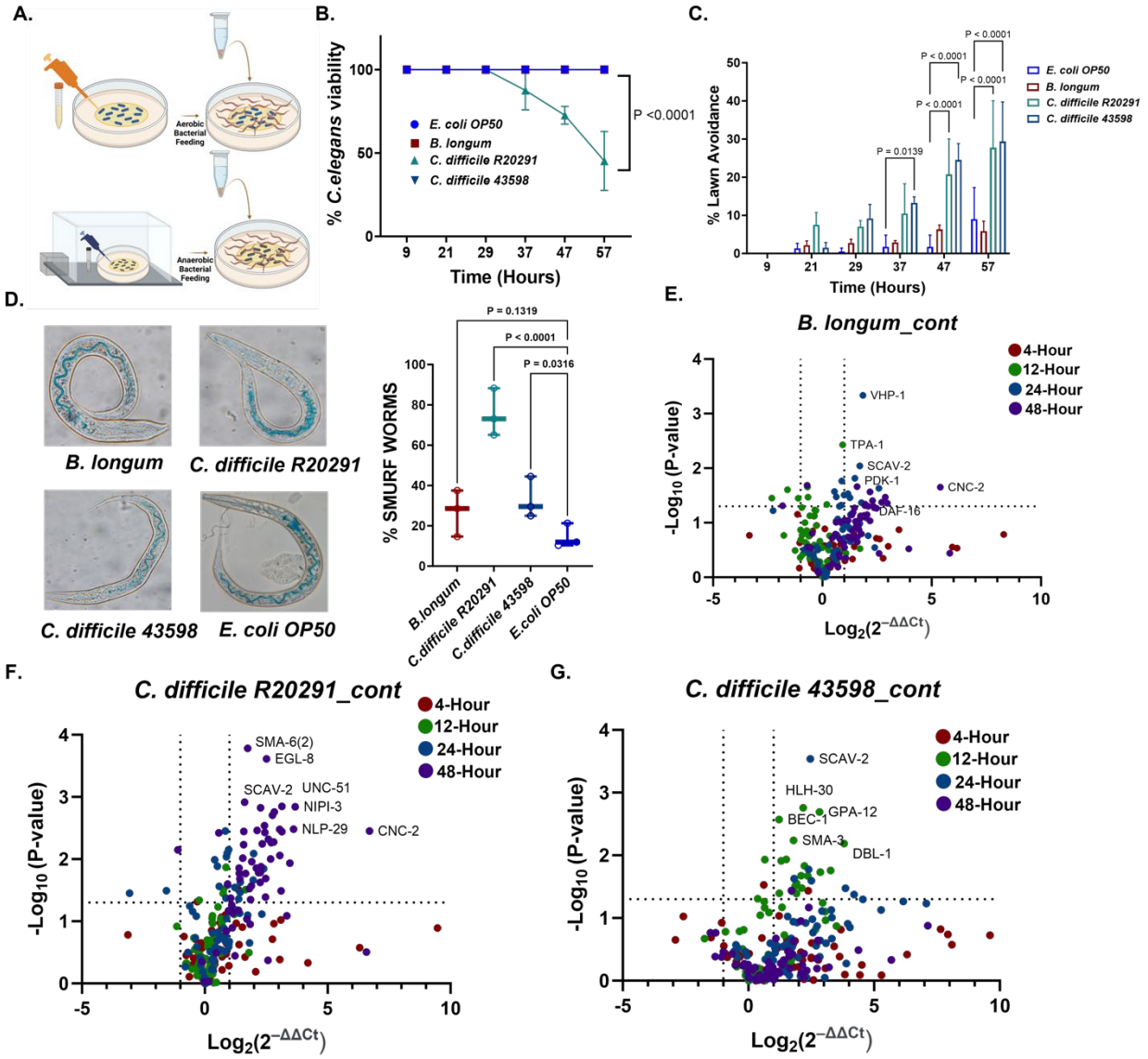


Figure 3. Continuous intake of *B. longum* and *C. difficile* leads to varied avoidance, toxicity, alterations in gut integrity, and immune response in *C. elegans*.

A. Graphical representation of the methods used for assessing *C. elegans* physiology upon aerobic feeding and a modified version for assessing physiology with anaerobic bacteria. The process involves seeding anaerobic bacteria onto NGM plates within an anaerobic chamber and, after drying, introducing *C. elegans* to these bacteria outside the chamber. Created with BioRender.com.

B. *C. elegans* viability percentage represented as mean ± SEM percentage of *C. elegans* viability after continuous exposure to *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598*. Measurements are taken at different time points with more than 30 *C. elegans* per replicate in biological triplicates. Statistical significance is determined using two-way ANOVA with Dunnett's correction.

C. Avoidance behavior percentage as mean ± SEM percentage of *C. elegans* avoiding the bacterial lawn of *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598*, measured at the same time points as in **(B).** ≥30 C. elegans per replicate are used in biological triplicates, with statistical significance calculated by two-way ANOVA with Dunnett's correction.
D. Representative images of the *C. elegans* intestine, focusing on smurf behavior. Also presented are the mean ± SEM values for the percentage of smurf worms when *C. elegans* are continuously fed with *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598*, using ≥50 *C. elegans* per replicate in biological triplicates. Statistical significance is assessed using two-way ANOVA with Dunnett's correction.
E-G. Volcano plots for gene expression data quantified for 62 innate immune-associated genes in C. elegans post continuous feeding with *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598* at 4, 12, 24, and 48 hours. The fold change is calculated using the $2^{\Delta\Delta Ct}$ method with *E. coli OP50* continuously fed expression data as the control at the respective time point. P-Values are calculated using unpaired students t-tests.

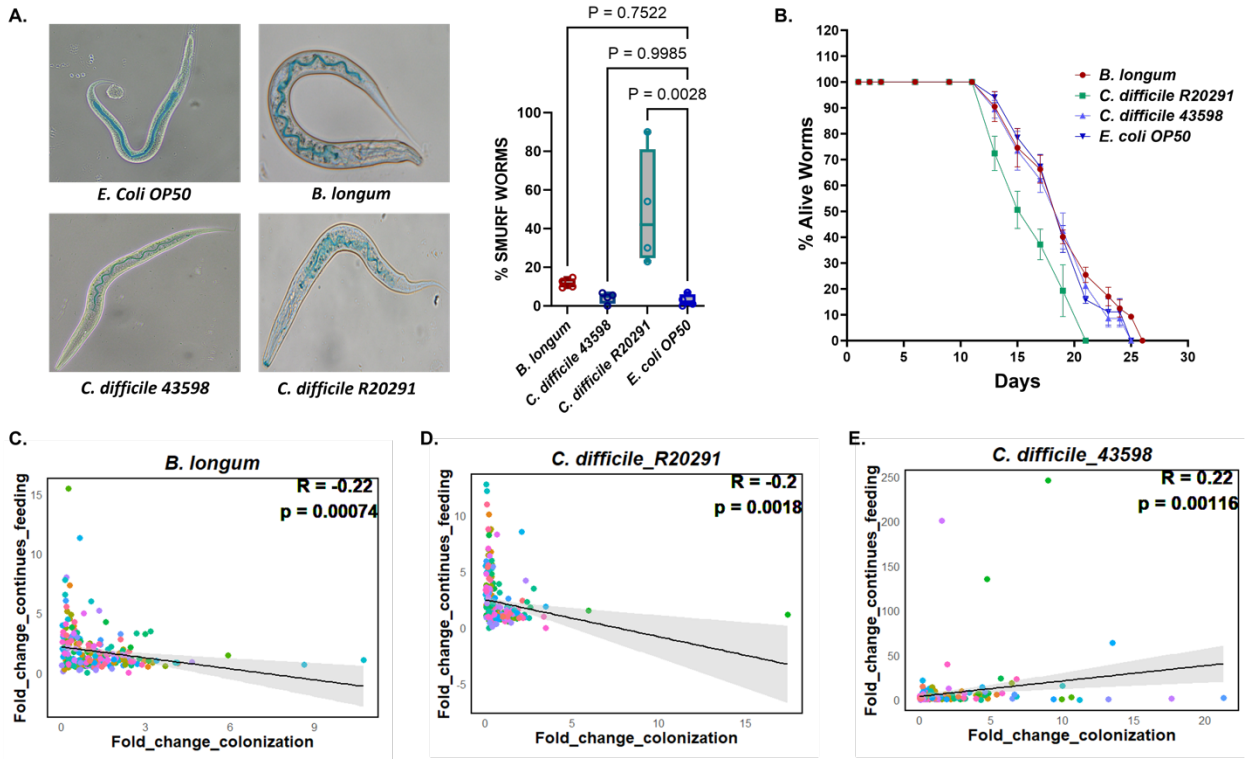


Figure 4. Anaerobic bacterial colonization of *B. longum* and *C. difficile R20291* results in nuanced differences in Avoidance, Toxicity, Alterations in Gut Integrity, and Immune Response in *C. elegans*.
A. Representative images of the *C. elegans* intestine, with a focus on illustrating smurf behavior. Additionally, mean ± SEM values for the percentage of smurf worms are presented for *C. elegans* colonized following the protocol outlined in Figure 1G for *E. coli OP50, B. longum, C.*

difficile R20291, and *C. difficile 43598*, with each replicate involving ≥50 C. elegans in biological triplicates. Statistical significance of the observed differences in smurf behavior across the groups is assessed using two-way ANOVA with Dunnett's correction.

B. Viability over time shows the mean ± SEM percentage of live worms over time for *C. elegans* colonized using the protocol mentioned in Figure 1G. The worms were colonized with *E. coli OP50, B. longum, C. difficile R20291,* and *C. difficile 43598*. The assay includes is done with ≥50 *C. elegans* per replicate, conducted in biological triplicates.

C-E. Correlation plots with trendlines indicating the linear relationship between $2^{\Delta\Delta Ct}$ values for continuous feeding, as detailed in Figures 3E-G, and colonization levels as mentioned in Figure S9C-E for include *B. longum, C. difficile R20291,* and *C. difficile 43598*. Each graph also includes the correlation coefficient (R) and the statistical significance of the observed correlations, represented as a p-value.

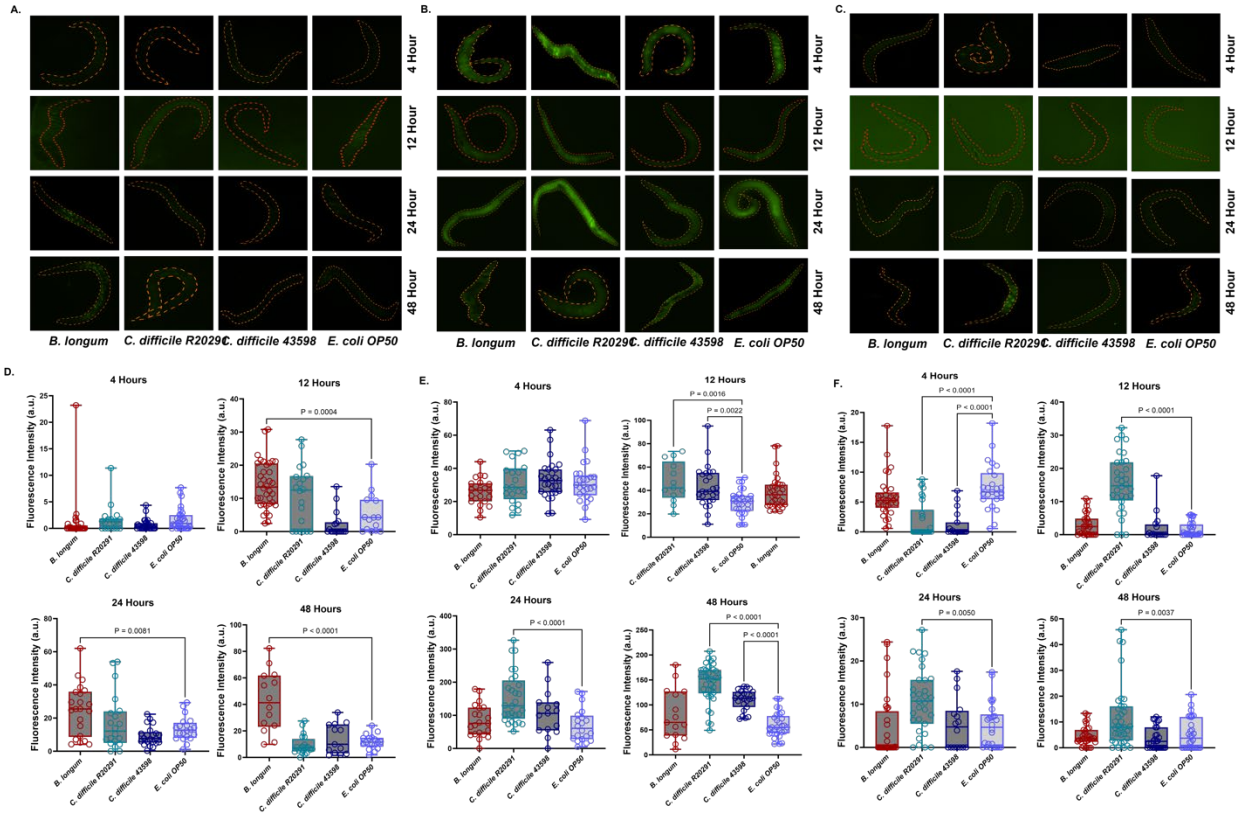


Figure 5. Colonization of transgenic mutants validates toxicity and lifespan responses in *C. elegans*.

A-C. Microscopic fluorescence images of three different GFP transgene mutants of C. elegans. Each mutant is tagged with GFP for distinct gene markers: **A.** *Daf-16*, represented by CF1139 ((pKL78) *daf16*::GFP + *rol-6*(su1006)); **B.** *lgg-1*, represented by MAH236 (*lgg-1p*::gfp::*lgg-1p* + *odr-1p*::rfp); and **C.** *Daf-2*, represented by QQ202 (cv20 (*daf-2*::gfp)). These mutants were colonized with *B. longum, C. difficile R20291, C. difficile 43598,* and *E. coli OP50* for 4, 12, 24, and 48 hours post colonization, following the protocol outlined in Figure 1G. These images

allow for the visualization of GFP expression, providing insights into the localization and regulation of the tagged genes under different colonization conditions.
D-F. Quantification of the fluorescence intensity observed in the GFP transgene mutants under the various conditions described in **A-C.** The mean fluorescence intensity for each group is expressed in arbitrary units (a.u) and includes the standard error of the mean (SEM). This quantification is conducted for at least 25 worms for every condition, ensuring a robust analysis. Statistical significance of the fluorescence intensity across different conditions and time points is determined using unpaired students t-tests.

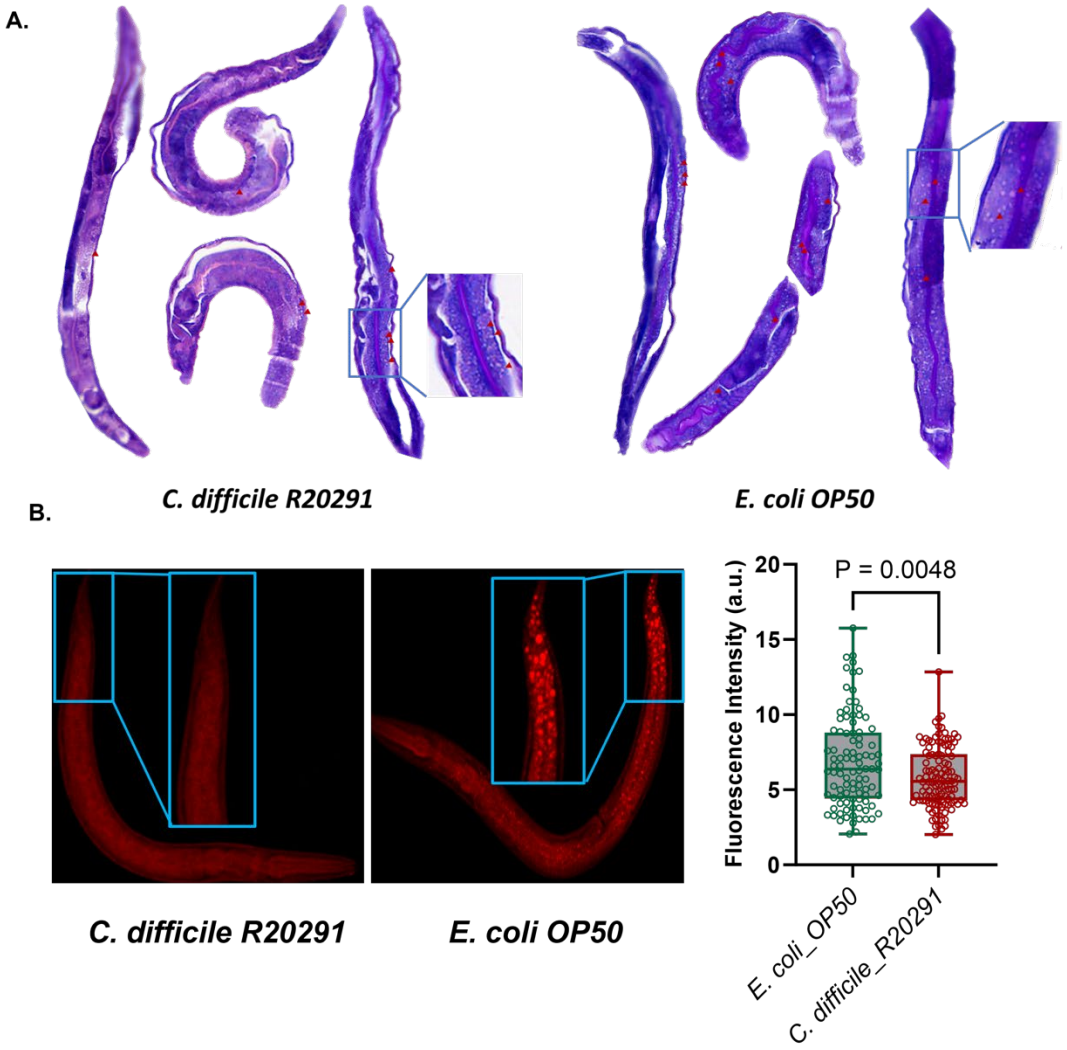


Figure 6. *C. difficile* colonization shows differential lipid accumulation in *C. elegans.*
A. Microscopic image of eosin stained 10µm thick histopathology section and
B. Fluorescence image of Nile red stained and intensity quantified for *C. elegans* colonized with *C. difficile R20291* and *E. coli OP50* following the protocol outlined in Figure 1G. Mean fluorescence intensity ± SEM is represented as arbitrary units (a.u), with p-value calculated using unpaired students t-tests. Red arrows and zoomed in images highlight the lipid granules.

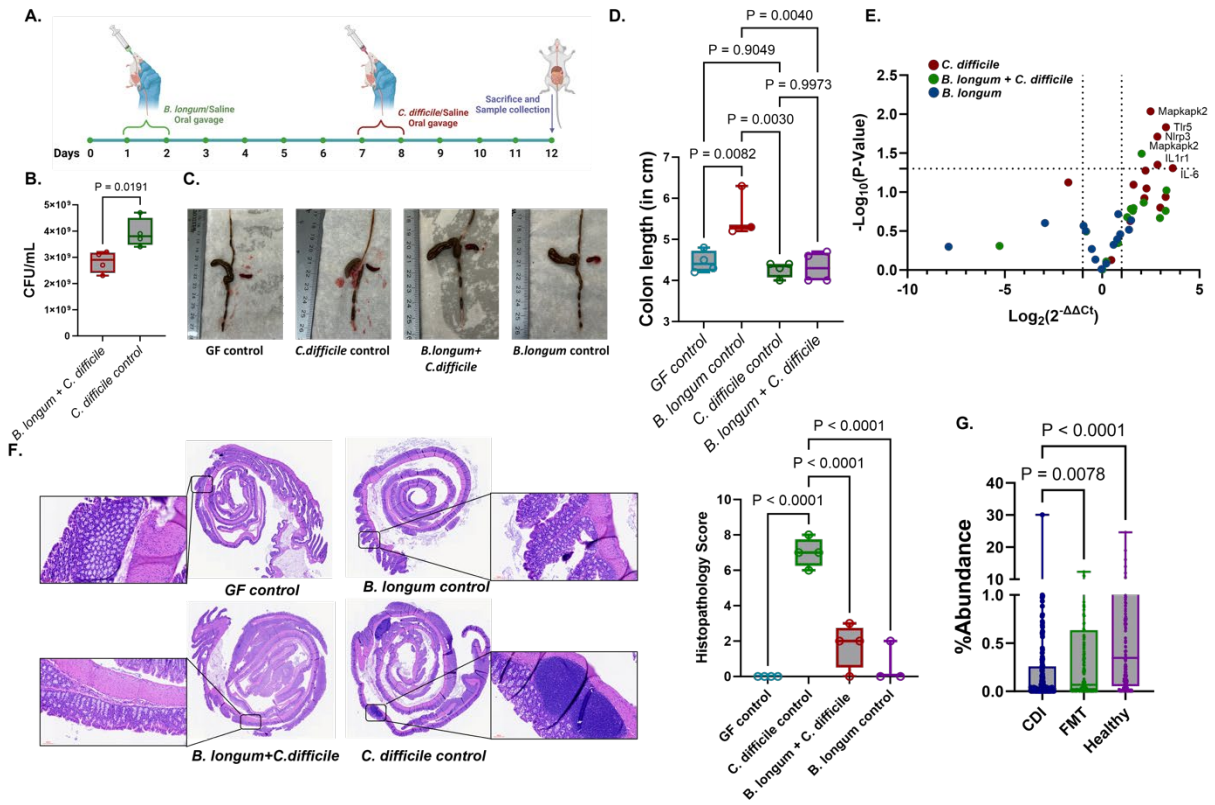


Figure 7. *B. longum* demonstrated effective colonization resistance to *C. difficile* in mice, resulting in a parallel enhancement of intestinal integrity.

A. Schematic representation of the experimental workflow used to assess colonization resistance in mice treated with *B. longum*. The schematic provides a clear overview of the experimental design, including the treatment and observation timelines. Created with BioRender.com.

B. CFU/mL of *C. difficile R20291* in fecal samples collected from the cecum of mice 4 days post-infection (DPI) after euthanasia. The data compares control mice treated with *C. difficile* only and mice pre-treated with *B. longum*.

C. Images of the cecum, colon, and spleen from the experimental mice and

D. Colon length measurement in centimeters for various groups: Germfree (GF) control (n=4), *C. difficile* control (n=4), mice pre-treated with *B. longum* and then infected with *C. difficile* (n=4), and *B. longum* control (n=3). Statistical significance of the differences in colon length across these groups is assessed using two-way ANOVA with Dunnett's correction.

E. Volcano plot for RT-PCR expression data of 14 immune genes associated with *C. difficile* infection in the mice colon. The groups compared include mice treated with *C. difficile* only, mice pre-treated with *B. longum* and then challenged with *C. difficile*, and mice treated with *B. longum* only. Fold changes in gene expression are calculated using the $2^{\Delta\Delta Ct}$ method with expression data from Germfree mice as the control. P-values are calculated using unpaired students t-tests.

F. H&E stained histopathological sections of the mice colon for the different groups: Germfree (GF) control, *C. difficile* control, *B. longum* pre-treated and *C. difficile* infection, and *B. longum* control and histopathological score is presented as mean ± SEM values for the sum of scores for different categories, as mentioned in Figure S11B. The statistical significance of observed differences in histological scores across groups is determined using two-way ANOVA with Dunnett's correction.

G. *B. longum* abundance analysis illustrating the percentage abundance of *B. longum* during *Clostridioides difficile* infection (CDI) and consecutive fecal microbiota transplantation (FMT) treatment, compared to healthy control individuals. Data is extracted from patient samples, with methodological details provided in the study methods section.

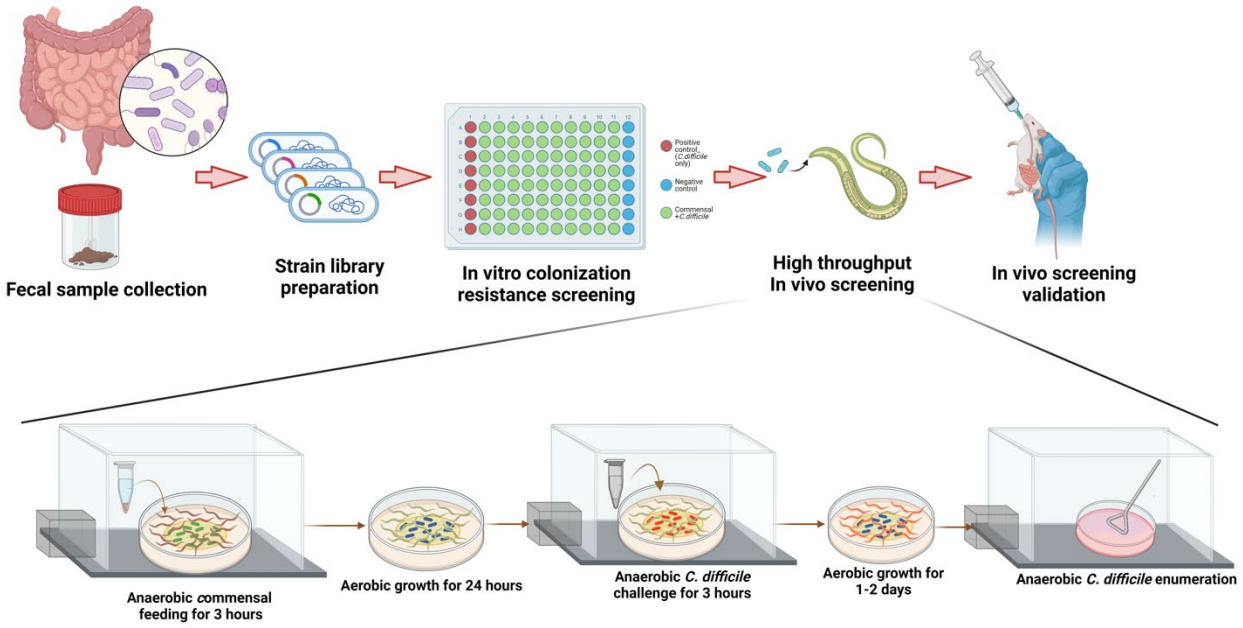


Figure 8. Summary schematic for the high throughput screening of the colonization resistance against *C. difficile* in *C. elegans*.

The figure provides a summary of the entire process, starting with the collection of cultures and preparation of a strain library, followed by an in vitro co-culture-based screening for colonization resistance. The process then advances to the confirmation of the efficacy of individual strains or mixtures against *C. difficile* in vivo using a germfree mouse model. Additionally, this study includes the development of an intermediate system for high-throughput screening based on *C. elegans*, with a brief overview of the foundational method included. Created with BioRender.com.

Supplementary Figures

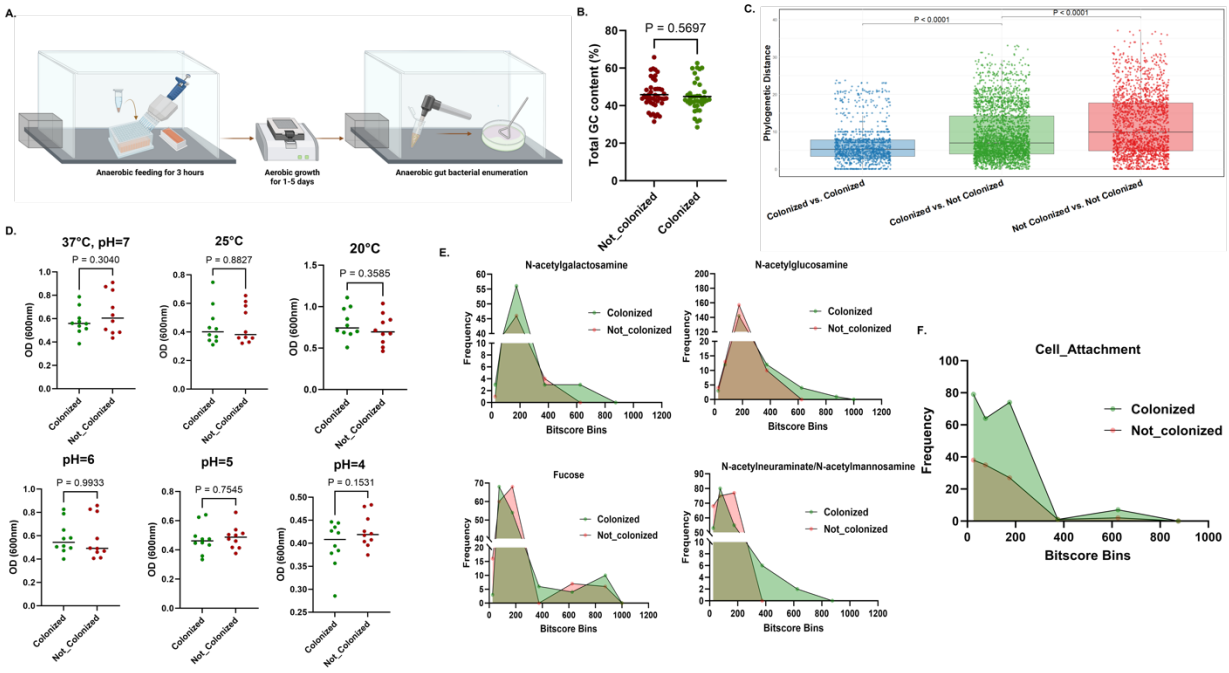


Figure S1. Factors influencing colonization of Human gut anaerobic gut isolates in *C. elegans*, related to Figure 1.

A. Graphical representation of the methodology used for the colonization of anaerobic human gut isolates in a high-throughput manner. Created with BioRender.com.

B. Presented here are the mean ± SEM values for the percentage of GC content identified from the whole genome sequencing of both colonized and non-colonized fractions of bacteria from the human gut microbiome collection. The statistical significance of differences in GC content is calculated using unpaired students t-tests.

C. The figure shows the mean ± SEM values of the phylogenetic distances calculated between and within colonized and non-colonized species. Statistical significance is determined using two-way ANOVA with Dunnett's correction.

D. Displayed here are the mean ± SEM OD (600nm) values of 10 blindly selected bacteria from both the colonized and non-colonized fractions of isolates. The growth of these bacteria is assessed post 24 hours at different pH levels (7, 6, 5, and 4) and temperatures (37°C, 25°C, and 20°C). The P-values, calculated using unpaired students t-tests.

E. The frequency histogram represents the frequency distribution of bit scores from tblastn for genes involved in the metabolism of N-acetylgalactosamine, N-acetylglucosamine, Fucose, and N-acetylneuraminate/N-acetylmannoseamine and

F. the cell attachment related genes in the colonized and non-colonized bacterial genome sets from the human gut bacterial culture collection used in the colonization trial.

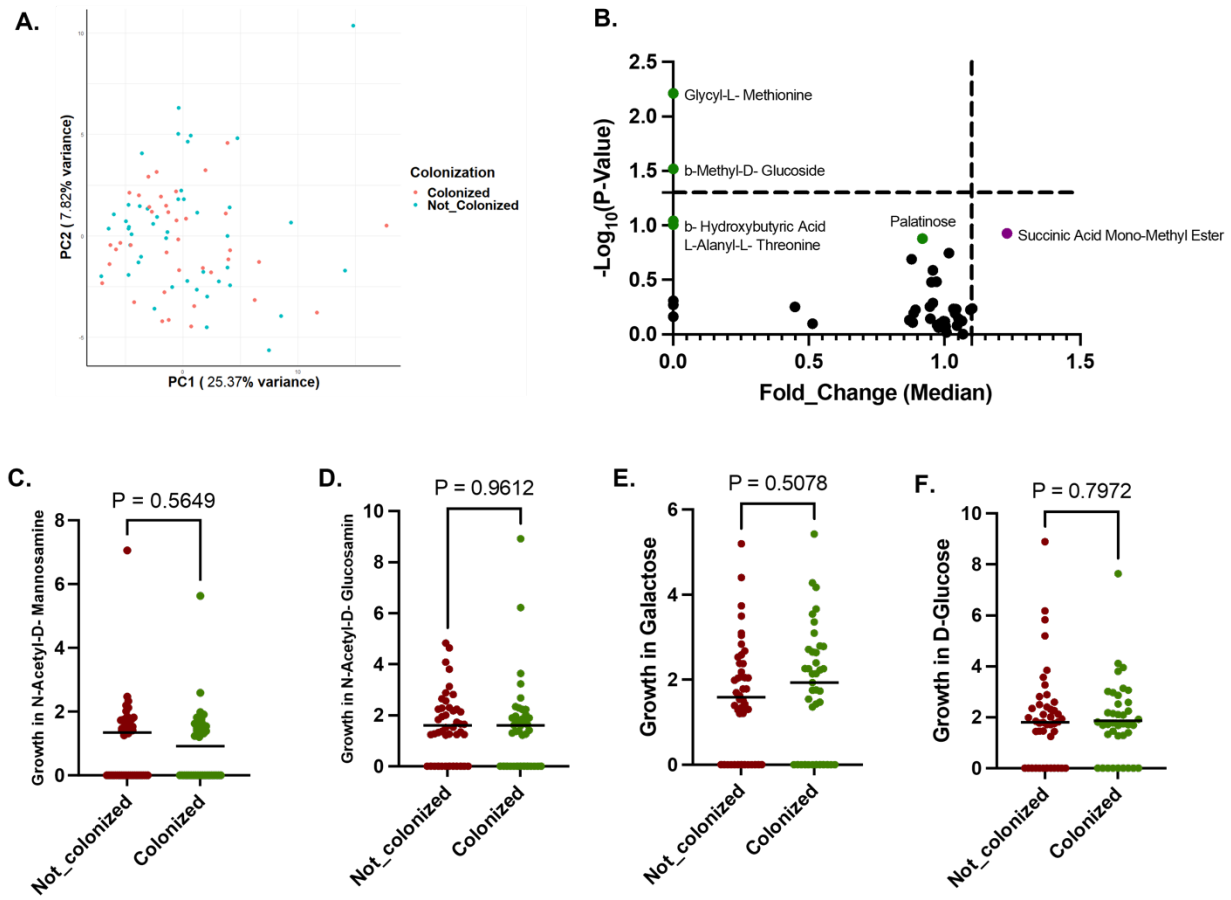


Figure S2. Carbohydrate degradation determination of colonization status, related to Figure 1.

A. The scatter plot and **B.** Volcano plot visualizing the results of a Principal Component Analysis (PCA) on growth between colonized and non-colonized bacteria in minimal media containing 95 different individual sugars, measured via OD (600nm) and normalized against a water blank. The median line and statistical significance (two-tailed unpaired t-tests) are highlighted.

C-F. Comparison of growth between colonized and non-colonized bacteria in minimal media containing selected mucin associated sugars **C.** N-Acetyl-D-Mannosamine, **D.** N-Acetyl-D-Glucosamine, **E.** Galactose and **F.** Glucose, measured via OD (600nm) and normalized against a water blank. The median line and statistical significance (two-tailed unpaired t-tests) are highlighted.

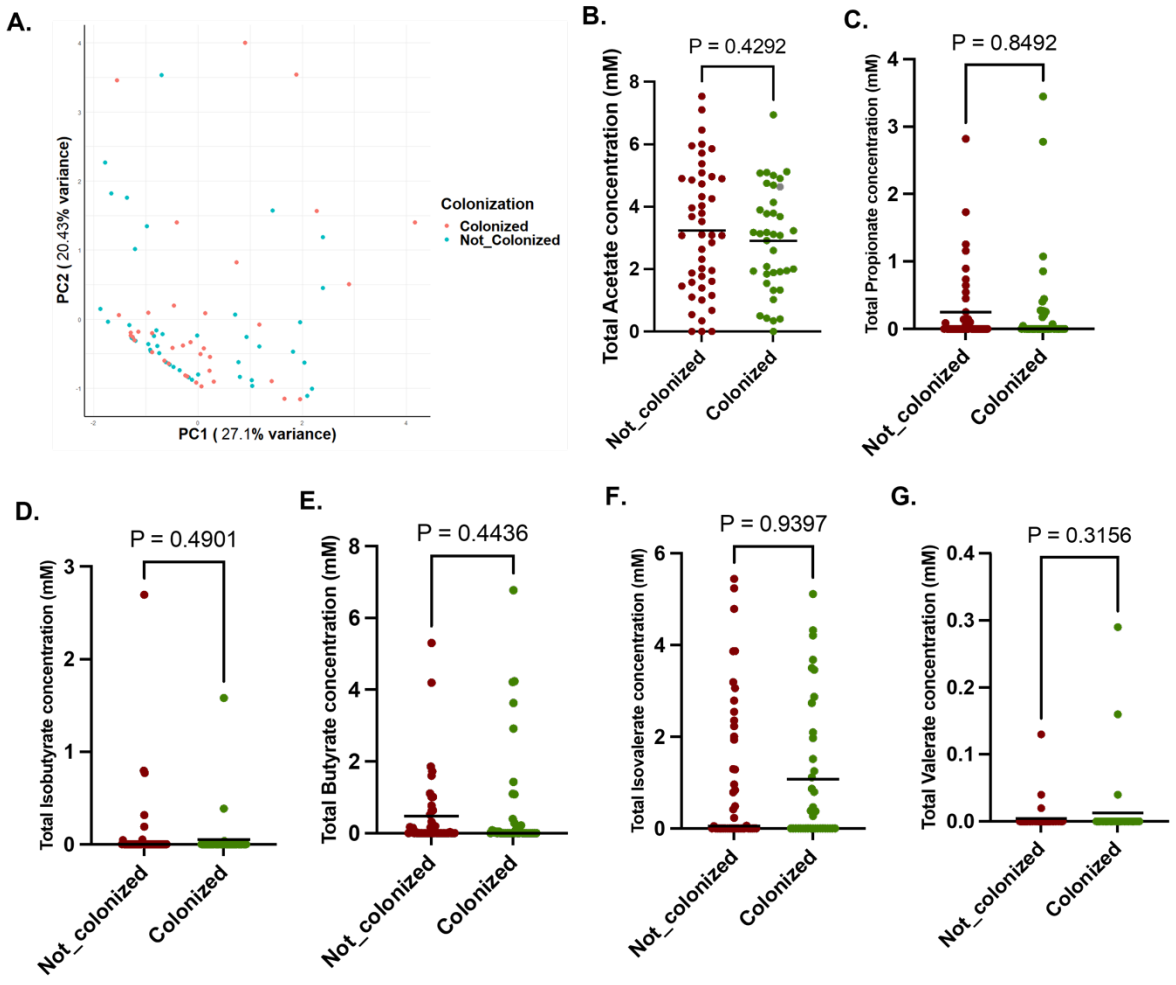


Figure S3. SCFA concentration doesn't correlates with the colonization status, related to Figure 1.

A. The scatter plot visualizing the results of a Principal Component Analysis (PCA) on SCFA produced by Individual human gut isolated bacteria in an mBHI medium previously reported in [1]. The plot displays data points on the first two principal components (PC1 and PC2), which capture the most significant variance in the dataset. Each point represents a bacterial sample, color-coded based on its colonization status.

B-G. Comparison of **B.** Acetate **C.** Propionate **D.** Isobutyrate **E.** Butyrate **F.** Isovalerate **G.** Valerate concentration between colonized and non-colonized bacteria in mBHI media. The median line and statistical significance (two-tailed unpaired t-tests) are highlighted.

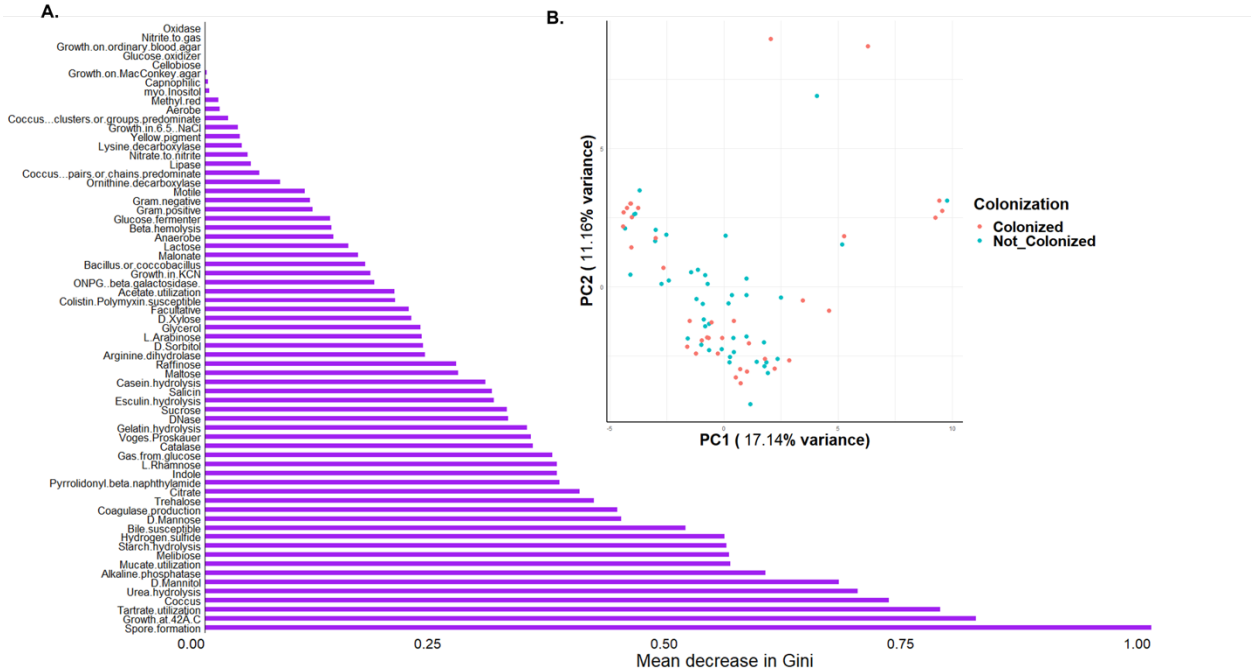


Figure S4. Traitar phenotype measures from colonized and non-colonized, related to Figure 1.

A. The bar chart displays the results of a Random Forest analysis identifying the importance of different bacterial strains in relation to colonization status. Each bar represents Mean Decrease in Gini Impurity. The chart is designed for ease of interpretation, with strains reordered for clarity and visual impact, highlighting those with the greatest influence on the model's predictions.

B. The scatter plot visualizing the results of a Principal Component Analysis (PCA) on Traitar data. The plot displays data points on the first two principal components (PC1 and PC2), which capture the most significant variance in the dataset. Each point represents a bacterial sample, color-coded based on its colonization status.

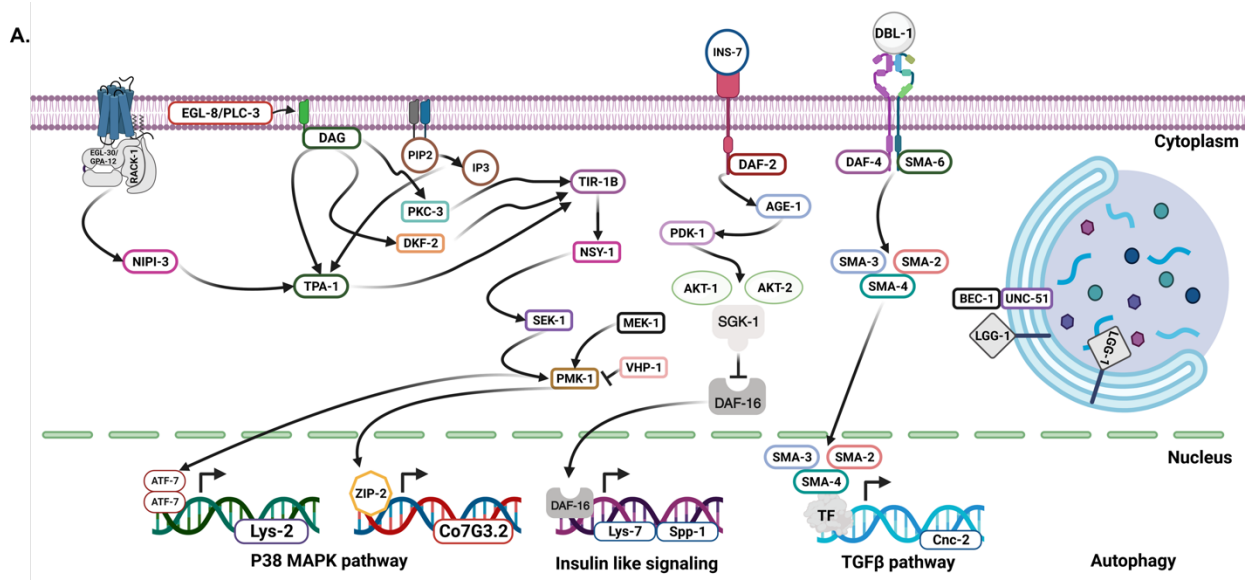


Figure S5. Signaling summary of the selected innate immune genes used for RT-PCR in *C. elegans* colonization experiment. Related to Figure 2. Created with BioRender.com.

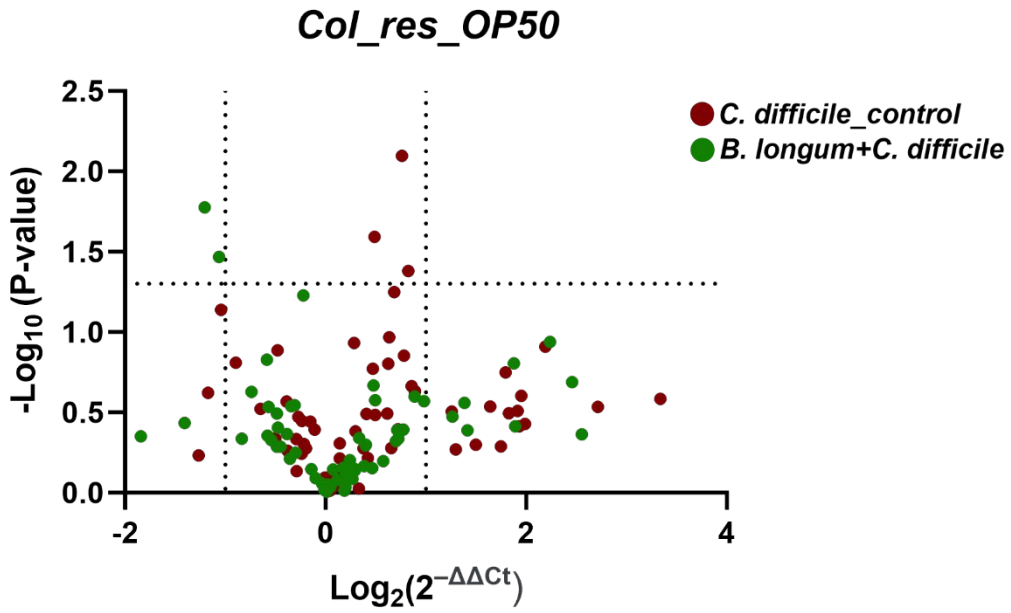
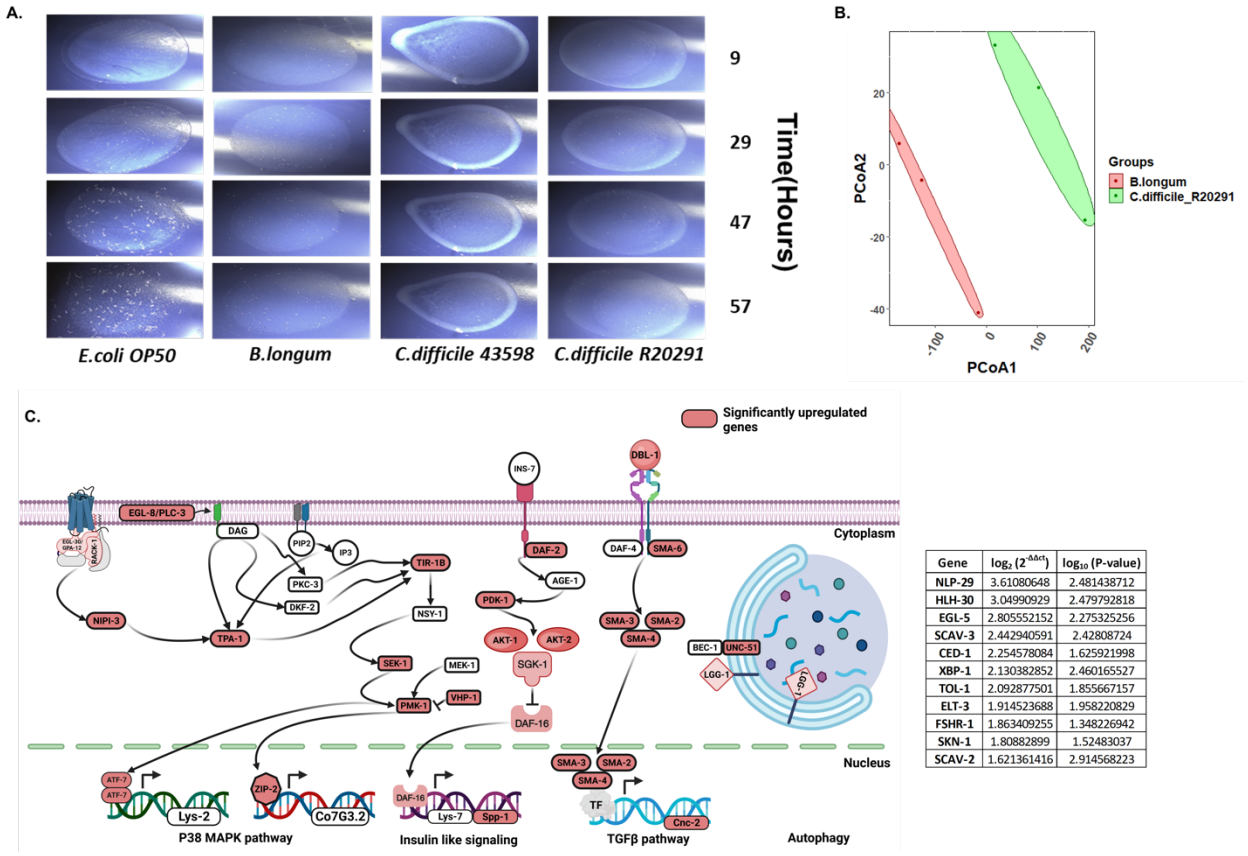


Figure S6. Gene expression of innate immune genes for colonization resistance experiments in C. elegans. Related to Figure 2.

A. Volcano Plot displays gene expression data quantified using 62 innate immune-associated genes in *C. elegans* post *C. difficile* feeding for *B. longum* pre colonized and *C. difficile* control worms at 24 hours normalized with OP50 colonized worms post 48 hours. Fold change is calculated using the $2^{\Delta\Delta}$Ct method with *E. coli OP50* fed expression data as control, and P-value is calculated using unpaired student t-tests.



Gene	$log_2 (2^{-\Delta\Delta ct})$	log_{10} (P-value)
NLP-29	3.61080648	2.481438712
HLH-30	3.04990929	2.479792818
EGL-5	2.805552152	2.275325256
SCAV-3	2.442940591	2.42808724
CED-1	2.254578084	1.625921998
XBP-1	2.130382852	2.460165527
TOL-1	2.092877501	1.855667157
ELT-3	1.914523688	1.958220829
FSHR-1	1.863409255	1.348226942
SKN-1	1.80882899	1.52483037
SCAV-2	1.621361416	2.914568223

Figure S7. *C. difficile* continues feeding shows avoidance behavior and an increase in innate immune gene expression in *C. elegans*, related to Figure 3.

A. Representative Plate image for the avoidance behavior result as mentioned in Figure 3C.

B. The scatter plot visualizing the results of a Principal Component Analysis (PCA) on gene expression data quantified for 62 innate immune-associated genes in *C. elegans* post continuous feeding with *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598* at 4, 12, 24, and 48 hours as mentioned in Figure 3E-G.

C. Signaling summary highlighting the genes with red color which are significantly upregulated more than 2-fold and P-value less than 0.05 for only at 48 hours post feeding. The adjacent table indicates the genes present in the pool of 62 innate immune genes which are not part of the signaling pathway mentioned but is significantly upregulated. Created with BioRender.com.

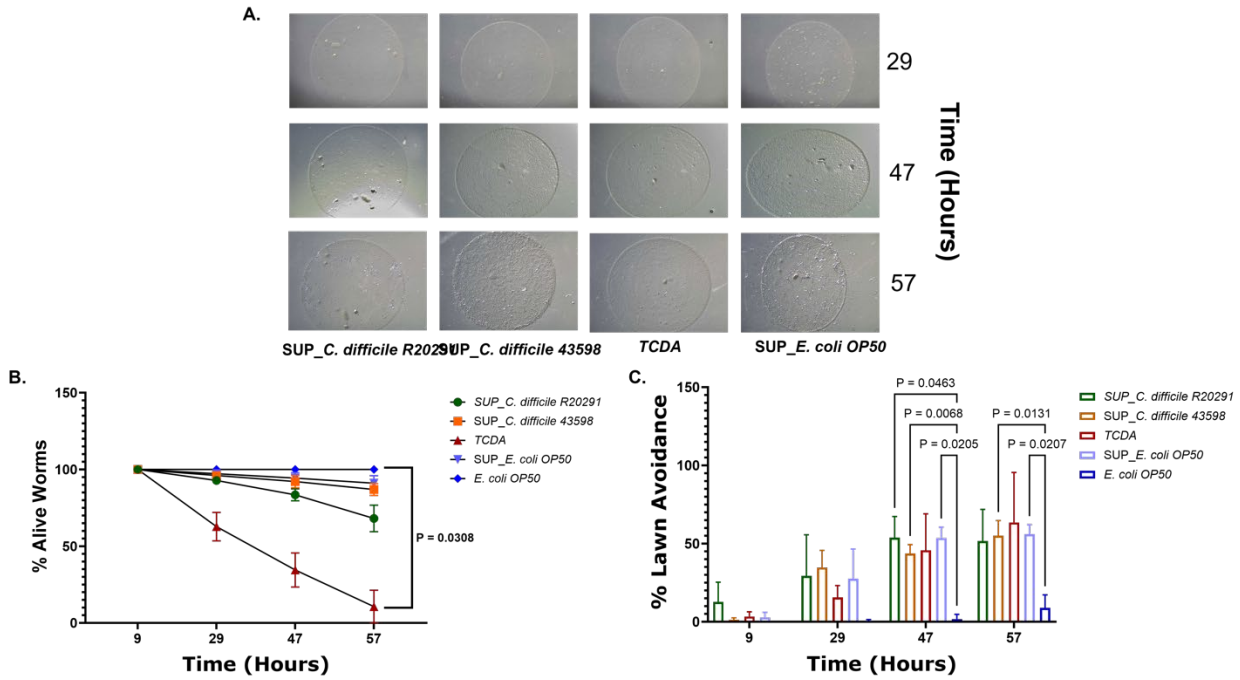


Figure S8. *C. difficile tcdA* toxins decrease viability and avoidance in *C. elegans*, related to Figure 3.

A. Representative plate images for the bacterial lawn for *C. elegans* fed with dead *E. coli OP50* pellet dissolved in Supernatant from *C. difficile R20291, C. difficile 43598, E. coli OP50,* and 1nM *tcdA* at different time points.

B. *C. elegans* viability percentage represented as mean ± SEM percentage of *C. elegans* viability after continuous exposure *E. coli OP50* pellet dissolved in Supernatant from *C. difficile R20291, C. difficile 43598, E. coli OP50,* and 1nM *tcdA* at different time points with more than 30 *C. elegans* per replicate in biological triplicates. Statistical significance is determined using two-way ANOVA with Dunnett's correction.

C. Avoidance behavior percentage as mean ± SEM percentage of *C. elegans* avoiding the bacterial lawn *E. coli OP50* pellet dissolved in Supernatant from *C. difficile R20291, C. difficile 43598, E. coli OP50,* and 1nM *tcdA* at the same time points as in **(B).** ≥30 C. elegans per replicate are used in biological triplicates, with statistical significance calculated by two-way ANOVA with Dunnett's correction.

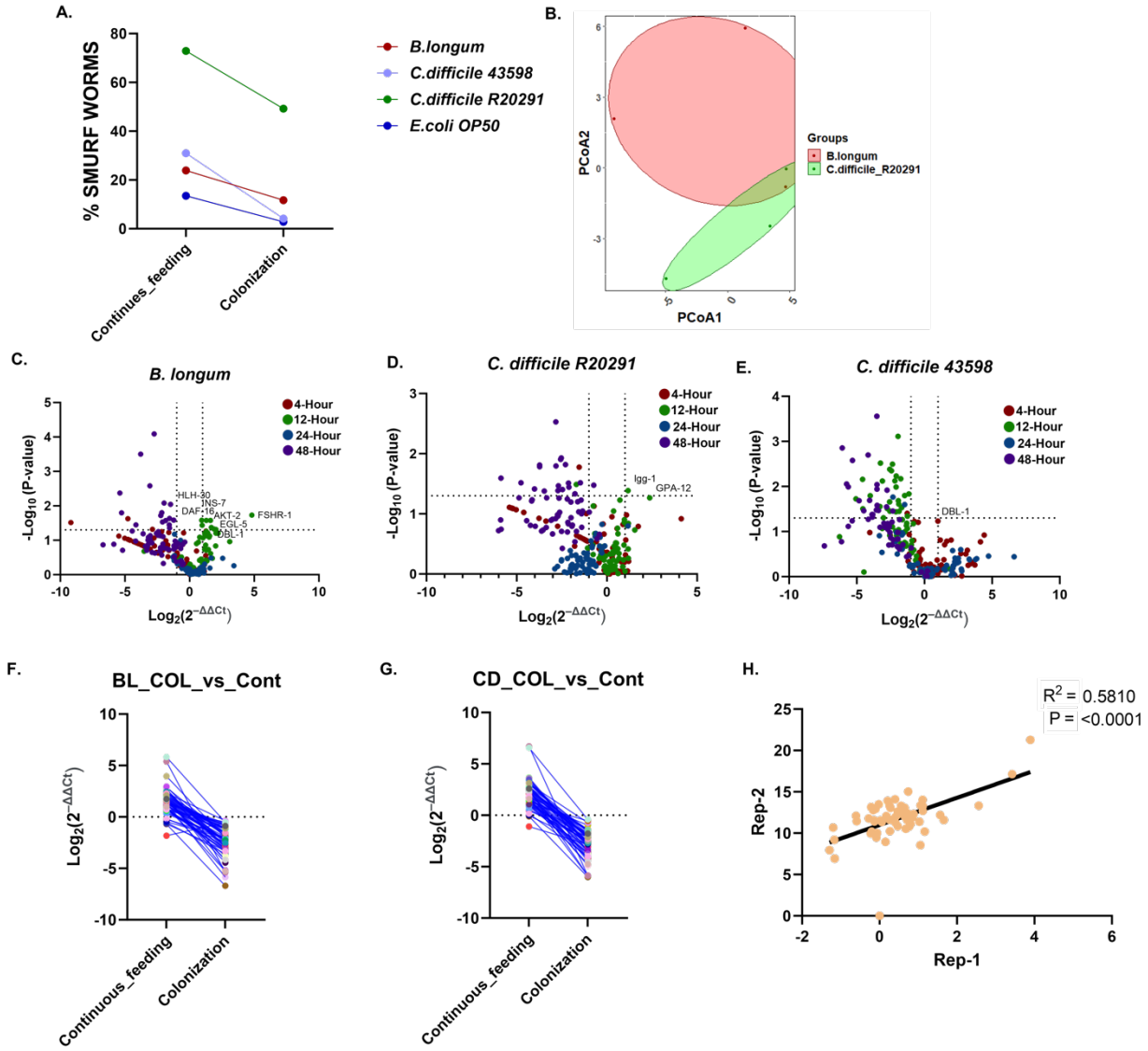


Figure S9. Differential physiology and immune response in *C. elegans* against colonization and continued feeding related to Figure 4.

A. Mean values for the percentage of smurf worms shown for *C. elegans* colonized and continuously fed with *E. coli OP50, B. longum, C. difficile R20291,* and *C. difficile 43598*, as mentioned in Figures 3D and 4A showing the difference.

B. The scatter plot visualizing the results of a Principal Component Analysis (PCA) on gene expression data quantified for 62 innate immune-associated genes in *C. elegans* post continuous feeding with *E. coli OP50, B. longum, C. difficile R20291*, and *C. difficile 43598* at 4, 12, 24, and 48 hours altogether as mentioned in Figure 4C-E.

C-E. Volcano plots for gene expression data quantified for 62 innate immune-associated genes in C. elegans colonized with *E. coli OP50,* **C.** *B. longum,* **D.** *C. difficile R20291*, and **E.** *C. difficile 43598* at 4, 12, 24, and 48 hours. The fold change is calculated using the $2^{\Delta\Delta Ct}$ method with *E.*

coli OP50 colonized expression data as the control at the respective time point. P-Values are calculated using unpaired students t-tests.

F-G. Mean values of gene expression data shown for *C. elegans* colonized and continuously fed with **F.** *B. longum and* **G.** *C. difficile R20291,* as shown in Figure 3E-F and S9C-D.

H. Replicate gene expression data, represented a correlation plot for the same as mentioned in **(C).**

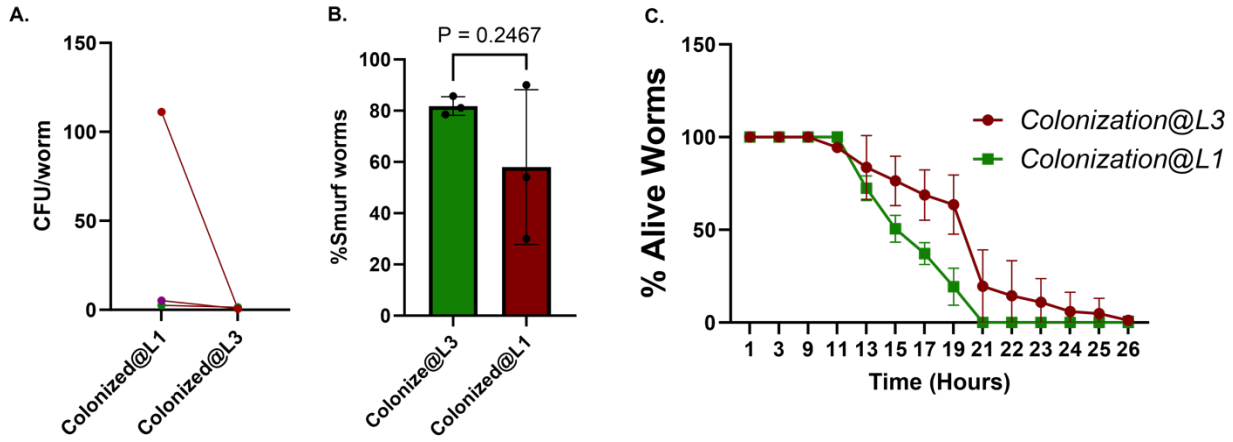
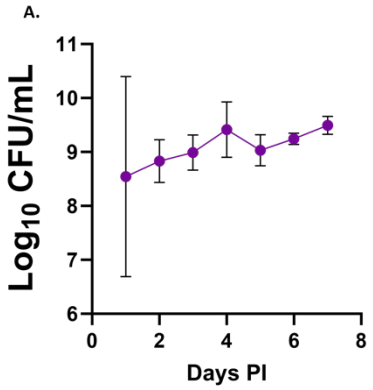


Figure S10. *C. elegans* performs differentially in colonization and disease physiology when colonized at different life stages.

A. CFU/worm graph showing the bacterial counts for *C. difficile R20291* colonized *C. elegans* at various time intervals for colonized in L1 worms as mentioned in Figure 1H and L3 worm as mentioned 2B.

B. Mean ± SEM values for the percentage of smurf worms shown for *C. elegans* colonized with *C. difficile R20291* at L3 as mentioned in Figures 2D and L1 mentioned in 4A showing the difference.

C. Mean ± SEM percentage of alive worms for *C. difficile R20291* colonized *C. elegans* at L1 worms as mentioned in Figure 4B and L3 worm as mentioned 2E.



B.

Sample	Erosion	Inflammation	PMN	MN	Mucosal hyperplasia	Crypt	Total
Germfree_M1	0	0	0	0	0	0	0
Germfree_M2	0	0	0	0	0	0	0
Germfree_M3	0	0	0	0	0	0	0
Germfree_M4	0	0	0	0	0	0	0
*C.difficile*_control_M1	0	3	3	1	0	0	7
*C.difficile*_control_M2	0	3	3	1	0	0	7
*C.difficile*_control_M3	0	3	3	1	0	1	8
*C.difficile*_control_M4	0	3	2	1	0	0	6
*B.longum+C.difficile*_M1	0	2	1	0	0	0	3
*B.longum+C.difficile*_M2	0	1	1	0	0	0	2
*B.longum+C.difficile*_M3	0	0	0	0	0	0	0
*B.longum+C.difficile*_F1	0	1	1	0	0	0	2
*B.longum*_control_M1	0	0	0	0	0	0	0
*B.longum*_control_M2	0	0	0	0	0	0	0
*B.longum*_control_M3	0	1	1	0	0	0	2

Figure S11. *B. longum* mediated colonization resistance in mice against *C. difficile*.

A. Mean ± SEM Log_{10} CFU/mL of *B. longum* isolated from the feces of mice colonized with *B. longum* and collected at different days post-gavaging.

B. Detailed histopathology scores for individual mice in all the treatments used in the colonization resistance studies in mice, as mentioned in Figure 7F.

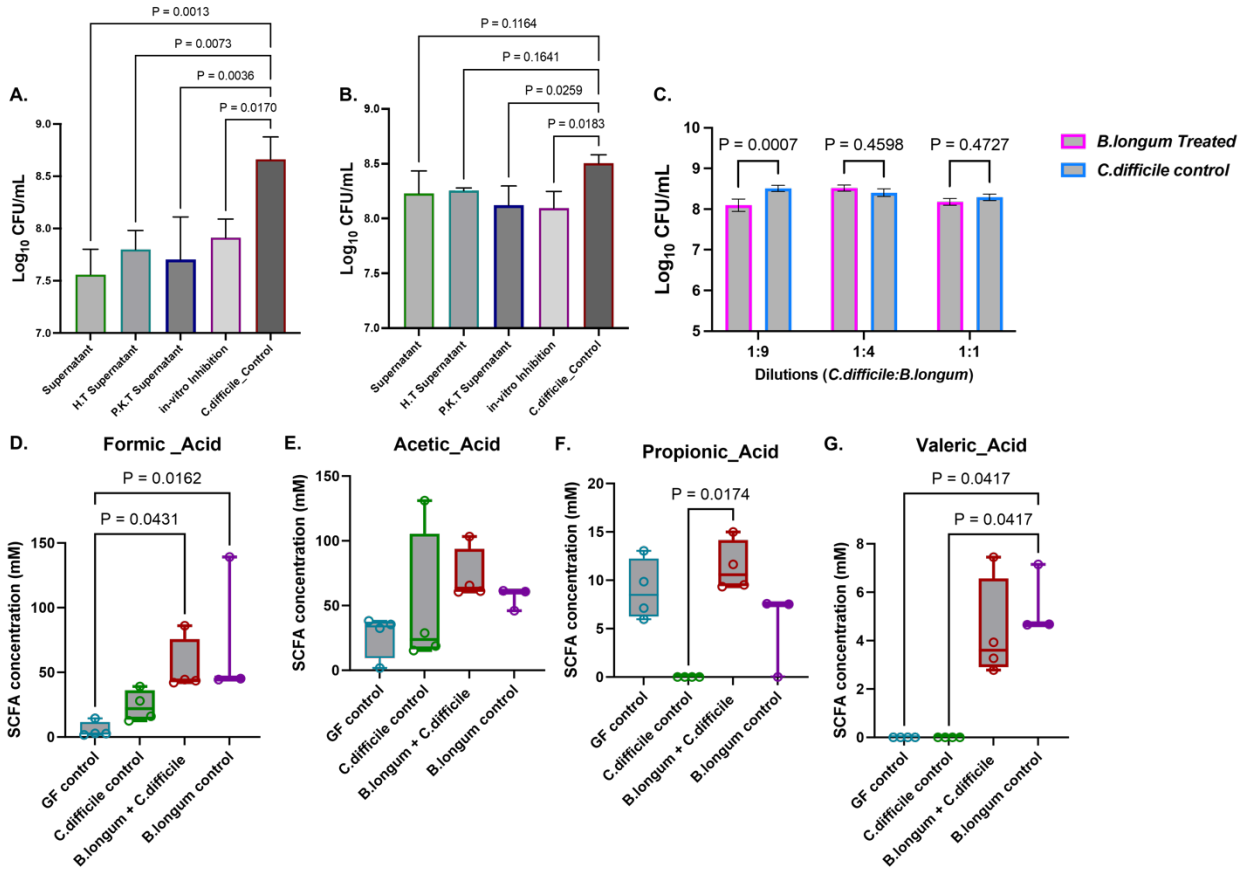


Figure S12. Intestinal PH and SCFA as the determinant of colonization resistance against *C. difficile*.

A. Mean ± SEM Log_{10} CFU/mL of *C. difficile R20291* when treated with a 1:1 ratio of *B. longum* grown mBHI culture supernatant with mBHI for 24 hours with statistical significance calculated by two-way ANOVA with Dunnett's correction.

B. Mean ± SEM Log_{10} CFU/mL of *C. difficile R20291* when treated with a 1:1 ratio of *B. longum* grown mBHI culture supernatant but adjusted the pH to 7 and mBHI for 24 hours with statistical significance calculated by two-way ANOVA with Dunnett's correction.

C. Mean ± SEM Log_{10} CFU/mL of *C. difficile R20291* cocultred with different ratios (1:1, 1:4, 1:9) of *C. difficile* and *B. longum* grown in mBHI for 24 hours with statistical significance calculated by two-way ANOVA with Dunnett's correction.

D-G. Concentration of **D.** Formic acid, **E.** Acetic acid, **F.** Propionic acid and **G.** Valeric acid identified from the cecal sample of mice colonized with Germfree (GF) control (n=4), *C. difficile* control (n=4), mice pre-treated with *B. longum* and then infected with *C. difficile* (n=4), and *B. longum* control (n=3). Statistical significance of the differences in colon length across these groups is assessed using two-way ANOVA with Dunnett's correction.