HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2022 December 25.

Published in final edited form as:

Cell Rep. 2022 November 29; 41(9): 111725. doi:10.1016/j.celrep.2022.111725.

Clostridia isolated from helminth-colonized humans promote the life cycle of *Trichuris* species

Shushan Sargsian[1,2], **Ze Chen**[2], **Soo Ching Lee**[3], **Amicha Robertson**[1,2], **Rafaela Saes Thur**[4], **Julia Sproch**[2], **Joseph C. Devlin**[2], **Mian Zi Tee**[5], **Yi Xian Er**[6], **Richard Copin**[2], **Adriana Heguy**[7,8], **Alejandro Pironti**[2,9,10], **Victor J. Torres**[2,9], **Kelly V. Ruggles**[11,12], **Yvonne A.L. Lim**[6], **Jeffrey Bethony**[4], **P'ng Loke**[3,*], **Ken Cadwell**[1,2,13,14,*]

[1]Kimmel Center for Biology and Medicine at the Skirball Institute, New York University Grossman School of Medicine, New York, NY 10016, USA

[2]Department of Microbiology, New York University Grossman School of Medicine, New York, NY 10016, USA

[3]Type 2 Immunity Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[4]Department of Microbiology, Immunology and Tropical Medicine, George Washington University, Washington, DC 20052, USA

[5]Department of Biomedical Science, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

[6]Department of Parasitology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

[7]Genome Technology Center, Office of Science and Research, New York University Langone Health, New York, NY 10016, USA

[8]Department of Pathology, New York University Grossman School of Medicine, New York, NY 10016, USA

[9]Antimicrobial-Resistant Pathogens Program, New York University Grossman School of Medicine, New York, NY 10016, USA

*Correspondence: png.loke@nih.gov (P.L.), ken.cadwell@nyulangone.org (K.C.).

AUTHOR CONTRIBUTIONS

S.S., P.L., and K.C. conceived and designed the study. M.Z.T., Y.X.E., and Y.A.L.L. carried out and oversaw the collection of samples from Malaysia. S.S. performed the experiments and analyzed and interpreted the data. A.R. and J.S. helped perform *T. muris* hatching experiments, and A.R. took confocal images. R.S.T. and J.B. provided guidance with *T. trichiura* experiments. A.H. oversaw whole-genome sequencing. A.P. and R.C. assembled and analyzed bacterial genomes. Z.C. and J.C.D. ran the metagenomic mapping analysis with guidance from K.V.R. and V.J.T. S.C.L. provided metadata and help with Malaysian metagenome datasets. K.C., P.L., and S.C.L. oversaw analysis and interpretation of all experiments described. S.S. and K.C. wrote the paper with input from all authors.

DECLARATION OF INTERESTS

K.C. has received research support from Pfizer, Takeda, Pacific Biosciences, Genentech, and Abbvie, consulted for or received honoraria from Vedanta, Genentech, and Abbvie, and is an inventor on US patent 10,722,600 and provisional patents 62/935,035 and 63/157,225.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111725.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

[10]Microbial Computational Genomic Core Lab, New York University Grossman School of Medicine, New York, NY 10016, USA

[11]Institute for System Genetics, New York University Langone Health, New York, NY 10016, USA

[12]Division of Precision Medicine, Department of Medicine, New York University Langone Health, New York, NY 10016, USA

[13]Division of Gastroenterology and Hepatology, Department of Medicine, New York University Langone Health, New York, NY 10016, USA

[14]Lead contact

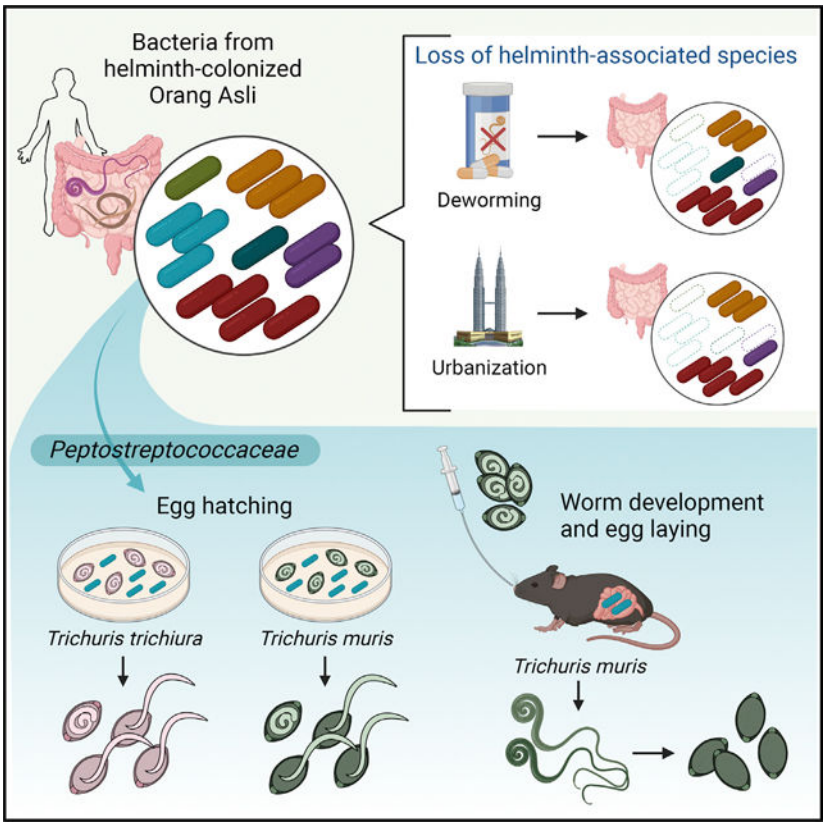
SUMMARY

Soil-transmitted intestinal worms known as helminths colonize over 1.5 billion people worldwide. Although helminth colonization has been associated with altered composition of the gut microbiota, such as increases in Clostridia, individual species have not been isolated and characterized. Here, we isolate and sequence the genome of 13 Clostridia from the Orang Asli, an indigenous population in Malaysia with a high prevalence of helminth infections. Metagenomic analysis of 650 fecal samples from urban and rural Malaysians confirm the prevalence of species corresponding to these isolates and reveal a specific association between *Peptostreptococcaceae* family members and helminth colonization. Remarkably, *Peptostreptococcaceae* isolated from the Orang Asli display superior capacity to promote the life cycle of whipworm species, including hatching of eggs from *Trichuris muris* and *Trichuris trichiura*. These findings support a model in which helminths select for gut colonization of microbes that support their life cycle.

In brief

Helminth-endemic regions are underrepresented in microbiome studies. Sargsian et al. isolate bacteria from individuals in rural Malaysia consisting of both well-studied and uncharacterized taxa. Clostridia associated with helminths in the microbiomes of these individuals display enhanced capacity to support the life cycle of the helminths *Trichuris muris* and *Trichuris trichiura*.

Graphical Abstract



INTRODUCTION

Intestinal parasitic worms known as helminths affect roughly 24% of the world population.[1] The most common species that infect humans are the whipworm *Trichuris trichiura*, roundworm *Ascaris lumbricoides*, and hookworms *Necator americanus* and *Ancylostoma duodenale*, which are associated with maladies ranging from diarrhea to stunted growth and malnutrition.[1,2] Although helminths cohabitate the mammalian gastrointestinal tract alongside trillions of microbes and can influence the diversity and composition of the microbiome,[3] species-level resolution and functional characterization of helminth-associated bacteria have lagged behind. Microbiome studies have been dominated by those examining individuals in industrialized nations,[4] where the prevalence of helminths is low. Neglecting to study helminth-associated microbes may result in a lack of knowledge of "missing microbes" not present in the microbiomes of better-characterized populations.[5,6] Furthermore, medical advances gained from microbiome studies may be biased toward individuals in high-income countries and insufficient in addressing the healthcare needs of marginalized groups in low- and middle-income countries.

The microbiomes of helminth-colonized individuals display cohort-specific compositions.[7–13] Inconsistent associations between microbes and helminth colonization likely reflect the unique ethnic populations, environment, and parasites associated with different geographic regions, highlighting the importance of deeper characterization of population-specific microbiomes. We previously reported higher gut microbial diversity in

the Orang Asli, an indigenous population in Malaysia with a high prevalence of helminth infections, especially *T. trichiura*, compared with nearby urbanized Kuala Lumpur.[14] Following deworming treatment with the anthelmintic drug albendazole, the microbiomes of the Orang Asli showed decreased alpha diversity accompanied by increased Bacteroidales and decreased Clostridiales. Colonization of mice with *Trichuris muris*, the *Trichuris* species that infects mice, or the roundworm *Heligmosomoides polygyrus bakeri* reveal similar shifts in the microbiome that support a positive relationship between helminth colonization and species belonging to the Clostridiales order.[14,15]

Clostridia is a polyphyletic class of Firmicutes that includes pathogenic and commensal Clostridiales species with ill-defined taxonomic relationships sharing the key feature of being sporeforming obligate anaerobes.[16,17] Clostridia species have received considerable attention as immunomodulatory members of the microbiota due to their metabolic activity and capacity to confer colonization resistance against pathogens.[18–23] Expansion of Clostridia during helminth colonization mediates amelioration of disease in mouse models of inflammatory bowel disease and asthma,[14,15] supporting the importance of genomic and functional characterization of helminth-associated bacteria in humans.

Furthermore, intestinal bacteria promote gastrointestinal colonization by helminths, indicating bidirectional communication between parasites and the microbiome.[24,25] For example, bacteria promote hatching of *T. muris eggs*, and *T. muris* cannot complete its life cycle in germ-free mice that lack a microbiome.[26–28] It is unclear if specific bacterial species display differential effects on helminth reproduction. Here, we report the genomic characterization of bacteria isolated from helminth-colonized Orang Asli villagers. Our findings identify specific taxa, especially those belonging to the *Peptostreptococcaceae* family, that are uniquely associated with helminth colonization and promote *T. muris* egg hatching and egg laying. Finally, we establish an egg-hatching assay for *T. trichiura* to demonstrate a similar effect of a helminth-associated *Peptostreptococcaceae* species on promoting the human parasite, providing evidence of parasite-microbiome coadaptation.

RESULTS

Isolation and identification of spore-forming Firmicutes from helminth-colonized individuals

Our previous analyses of stool specimens from the Orang Asli using 16S sequencing showed a positive correlation between *T. trichiura* burden and relative abundance of Clostridiales;[14] however, we lacked species-level resolution of helminth-associated bacteria. We screened stool from Orang Asli in Pangsun village (*Trichuris* prevalence = 52.63%) to identify individuals with high *T. trichiura* burden to increase the probability of isolating Clostridiales (Table S1). We applied a previously described procedure, in which fecal samples are enriched for spore-forming bacteria,[23] to specimens from Orang Asli individuals with high *Trichuris* burdens (Table S1; Figure S1A). From 75 initial colonies, we used 16S sequencing to identify 14 isolates with unique sequences and/or morphologies, designated OA1–14. Bacterial colonies displayed distinct visual properties including mucoid and filamentous morphologies and hemolytic activity, providing evidence that they represent non-redundant isolates (Figure 1A).

Taxonomic lineage was identified through whole-genome sequencing and the GTDB-Tk approach for genomic classification.[29] Thirteen isolates belonged to Clostridiales, representing five families: *Lachnospiraceae* (OA1, OA3, OA11, and OA13), *Peptostreptococcaceae* (OA2, OA6, and OA8), *Erysipelotrichaceae* (OA4, OA5, and OA7), *Oscillospiraceae* (OA9 and OA10), and *Clostridiaceae* (OA14) (Figure 1B). OA12 was identified as *Enterococcus hirae*. Although *Enterococci* are not spore formers, they can be partially resistant to chemical agents, including chloroform.[30,31] All isolates except OA3 had an average nucleotide identity (ANI) of at least 95% to previously sequenced bacteria in public databases, allowing us to assign a species identifier. With an ANI of 88.32% to the closest known relative, OA3 likely represents a new species belonging to the *Ruminococcus* genus, which we have provisionally named *Ruminococcus pangsunibacterium*.

Genome sizes and GC content for OA isolates were generally typical for their taxa and range from 2.85 to 6.95 Mbp and 27.91% to 60.26%, respectively. Phylogenetic comparison with related Firmicutes (Figure 1C; Table S2) confirmed the taxonomic relation of OA isolates to previously sequenced bacteria. Although some isolates such as OA12 (*E. hirae*) and OA14 (*Clostridium perfringens*) are extensively studied species, others like OA2 and OA6 (*Romboutsia hominis*) and OA9 and OA10 (genus *Lawsonibacter*) have fewer than 8 other deposited genomes and have yet to be functionally characterized (Figure 1B; Table S2).[32,33]

We created uniform manifold approximation and projection (UMAP) plots of orthogroups of genes for each species to compare the total coding potential of the OA isolates with their relatives.[34] We restricted this analysis to species for which a substantial number of sequenced genomes are available: *Clostridium innocuum* (OA4 and −5), *Paraclostridium sordellii* (OA8), *E. hirae* (OA12), and *C. perfringens* (OA14). OA4 and OA5 fell into a distinct cluster of *C. innocuum* genomes, while the other OA isolates aligned with the dominant cluster of their respective species (Figures 1D, 1E, S1B, and S1C). The availability of 622 *C. perfringens* genomes facilitated additional analysis of OA14. A UMAP plot of KEGG orthologs in the metabolism category revealed that OA14 clusters away from most other *C. perfringens* isolates (Figure 1F). Further analysis revealed that the OA14 cluster, consisting of 28 *C. perfringens* genomes (4.5% of all available genomes) (Figure S1D; Table S2), is characterized by a significantly higher representation of metabolic pathways involved in glycosphingolipid and N-glycan biosynthesis and metabolism of various compounds, as well as several orthologs involved in lantibiotic biosynthesis and transport. In addition, OA14 lacks a gene ortholog for an enterotoxin present in over 94% of previously identified *C. perfringens*, suggesting lower virulence of our isolate (Table S3).[35]

Quantification of how the pangenome, the whole-genomic repertoire of each microorganism,[36] increases with each added genome revealed that while there have been many genomes sequenced for these species, there is great intra-species diversity, and information about these taxa has not yet been saturated (Figure 1G and S1E–S1G). Thus, OA isolates include both uncharacterized taxa and well-investigated species, which may include those with unique properties.

OA isolates belonging to the Peptostreptococcaceae family are associated with helminth colonization

Given that the OA isolates were isolated from two Orang Asli individuals, we tested their broader presence in human microbiomes. We collected stool from a large cohort of 351 Orang Asli from 6 villages and 56 individuals living in Kuala Lumpur.[37] In addition to comprehensive analyses exploring how helminth infection is associated with microbiome composition and diversity, which is reported in Tee et al.,[37] we used the metagenomics classification tool krakenuniq[38] to map 650 metagenomes generated from the above Malaysian populations to unique sequences from the whole genomes of OA1–14. We also compared these results with 544 stool metagenomes from the Human Microbiome Project (HMP), which sampled 242 healthy individuals in the United States.[39,40] As expected, we detected higher mapping frequency in the Malaysian metagenomes than HMP metagenomes to sequences specific to OA1–14, as well as to the OA isolates and their close relatives combined (Figures 2A and 2B; Table S4). This was also the case when we examined mapping to individual isolates, representing the abundance of taxa closely related to each isolate (Figure 2C). Because our computational approach discriminates bacterial genomes belonging to the same species, these results do not exclude the potential presence of closely related bacteria in HMP metagenomes.

Lachnospiraceae isolates OA11 and OA13 were not detected due to low representation across all metagenomes. As krakenuniq functions by mapping metagenomes to unique sequences in each genome that are not found in any other genomes, we hypothesized that OA5 (*C. innocuum*) was not detected because its genome is highly similar to OA4 with an ANI of 99.99%. Indeed, when OA4 was removed from the mapping analysis, the mapping percentage of OA5 increased (Figure S1H), indicating that krakenuniq could not differentiate between OA4 and OA5. In contrast, our computational approach was able to distinguish OA2 and OA6 (both *R. hominis*), which have an ANI of 99.96%, indicating that closely related members of the same species can be differentiated. It is also important to note that although krakenuniq can differentiate between two highly similar strains, it cannot accurately determine the presence of a specific strain or isolate within a metagenome. Therefore, the outcome of our analysis more likely reflects the abundance of species or closely related taxa that represent the OA isolates.

We detected several positive correlations between pairs of OA isolates within the Malaysian metagenomes, with a particularly strong association between OA2 and OA6, both *R. hominis* (Figure S1I). The lack of strong negative correlations between isolates suggests that these bacteria are not in an antagonistic relationship with each other in this population.

The above Malaysian dataset includes individuals living in rural Orang Asli villages and urbanized Kuala Lumpur (KL). The Orang Asli samples can be further segregated into those taken prior to and after treatment with the anthelmintic drug albendazole (ABZ). The helminth prevalence in the Orang Asli population decreased from 67.2% at baseline to 22.5% 21 days after ABZ. The helminth prevalence in the urban control population was 0%.[37] To identify helminth-associated bacteria, we determined which taxa corresponding to the OA isolates were enriched in pre-ABZ samples compared with post-ABZ and KL. A heatmap showing the normalized percentage of k-mers in each metagenome that specifically

map to each OA genome revealed that taxa corresponding to OA2, OA6, OA8, OA12, and OA14 were generally rare but exhibited high mapping in a small proportion of individuals, whereas taxa corresponding to OA1, OA3, and OA10 had a wider distribution of mapping across individuals (Figure S2A). While OA2, OA6, OA8, OA9, OA12, and OA14 exhibited higher mapping in pre-ABZ Orang Asli villagers than KL, OA1, OA3, OA4, OA7, and OA10 displayed similar or lower mapping in the Orang Asli compared with KL, indicating that these taxa are associated with the geographical region rather than helminths (Figures 3A–3H and S2C–S2E).

Strikingly, all three isolates belonging to the *Peptostreptococcaceae* family, OA2, OA6, and OA8, had significantly decreased mapping in post-ABZ Orang Asli metagenomes, with these mapping percentages resembling those for KL (Figures 3A–3C and S2B). This finding is reminiscent of observations in mice showing that a *Peptostreptococcaceae* species was among the most expanded following *H. polygyrus* inoculation.[41] *Lawsonibacter sp900066645* (OA9) and *E. hirae* (OA12) were also reduced in the microbiomes of ABZ-treated individuals (Figures 3D, 3E, and S2B). The abundance of taxa representing the remaining OA isolates had no significant change (Figures 3F–3H and S2C–S2E). For 129 individuals, matched pre- and post-ABZ longitudinal samples were available. Pairwise analyses for these samples confirmed that OA2, OA6, OA8, OA9, and OA12 were reduced within individuals following ABZ treatment (Figures 3I–3M). We observed decreased abundance of *Lachnospiraceae* species corresponding to OA1 and OA3 within individuals post-ABZ (Figures S2B and S2F–S2K), even though mapping to these isolates was not significantly different between pre- and post-ABZ at the population level (Figures S2C and S2D). Considering the relatively high mapping of metagenomes to OA1 and OA3 across the entire Malaysian cohort (Figure 2C), the data suggest that these taxa are present at high enough levels to persist in microbiomes in the region despite decreasing as a result of deworming within an individual.

To test if the relationship between OA2, OA6, and OA8 and helminth infection was species specific or indicative of a relationship between helminths and all *Peptostreptococcaceae*, we mapped metagenomes to all 113 available complete genomes in this family (Table S4) and found that the *Peptostreptococcaceae* family is indeed enriched in Malaysian microbiomes compared with the HMP (Figure S2L). Additionally, there was a significant decrease in the abundance of *Peptostreptococcaceae* within individuals post-ABZ, suggesting a specific interaction of this family with helminths (Figure S2M).

The reduced mapping to OA2, OA6, OA8, OA9, and OA12 in KL microbiomes increased our confidence that these taxa are helminth associated. Although we cannot rule out helminth-independent effects of ABZ, we found that even the highest soluble concentration of ABZ did not impair *in vitro* growth of OA isolates (Figure S2N). We found that the individuals who were fully cured from *T. trichiura* post-ABZ had a lower initial worm burden than those who merely decreased in worm burden but were not cured (Figure S3A). In addition, individuals who were fully cured of helminths displayed decreased abundance of the *Peptostreptococcaceae* species represented by OA2, OA6, and OA8, whereas those who remained colonized by helminths had stable levels of these taxa (Figure S3B). Thus, the reduced presence of species corresponding to the above OA isolates in the microbiomes

of ABZ-treated individuals is unlikely due to a direct effect of the drug on these bacteria and instead is a consequence of deworming, as the continued presence of helminth infection within an individual maintains colonization of the OA isolate despite deworming treatment.

Helminth-associated Clostridia promote the Trichuris life cycle

The host acquires *Trichuris* species following ingestion of embryonated eggs in contaminated food or water, which then hatch to larvae in the microbiota-rich cecum and mature to adult worms. Adult females lay thousands of unembryonated eggs per day that are released through stool to the environment, where they embryonate and can transmit to a new host or reinfect the same individual.[42,43] To test the hypothesis that *Trichuris* colonization favors the presence of microbiota members that facilitate its life cycle, we examined whether helminth-associated OA isolates can promote hatching of *T. muris* eggs, the helminth species for which a bacteria-mediated hatching assay under aerobic conditions has been established.[27,44] Because Clostridia are obligate anaerobes, we first determined whether *T. muris* eggs (Figure 4A) can hatch in the presence of bacteria under anaerobic conditions. As a positive control, we chose the facultative anaerobe *Pseudomonas aeruginosa*, which was previously shown to be a potent inducer of hatching in aerobic conditions.[27]

Incubation of embryonated *T. muris* eggs with *P. aeruginosa* resulted in nearly 100% hatching over 3 h compared with the slower inefficient hatching of <30% for eggs incubated with media alone (Figure 4B). Thus, bacteria can enhance *T. muris* egg hatching under anaerobic conditions. Using this assay, we tested four OA isolates associated with helminths (OA2, OA6, OA8, and OA12) and compared them with two isolates not associated with helminths (OA4 and OA14). We also included two Bacteroidales species, *Bacteroides thetaiotaomicron* and *Phocaeicola vulgatus* (previously *Bacteroides vulgatus*),[45] which are ubiquitous anaerobes in the human microbiome. In addition, *P. vulgatus*'s growth rate was negatively correlated with *Trichuris* egg burden in the Orang Asli.[37] The *Peptostreptococcaceae* strains OA2, OA6, and OA8 were rapid inducers, with around 20% of eggs hatched after 1 h of incubation and >60% hatched by 3 h. Eggs incubated with OA4 and *B. thetaiotaomicron* were slower to hatch, while OA12, OA14, and *P. vulgatus* were poor inducers that potentially inhibit hatching (Figures 4B–4E). The degree of hatching did not correlate with the rate of bacterial growth (Figures S4A–S4C).

Bacteria-mediated hatching of eggs from *T. trichiura* has not previously been reported. We reasoned that using bacteria isolated from individuals colonized by *T. trichiura* may overcome this obstacle. Compared with a mean hatching rate of 8.8% for *T. trichiura* eggs incubated in media alone for 6 days, incubation with OA8, *P. sordellii*, increased the mean hatching rate to over 26% (Figures 4F and 4G). This effect was species specific, as other *Peptostreptococcaceae* family members and non-helminth-associated controls did not increase hatching rates compared with media alone. In addition, OA12, *E. hirae*, inhibited hatching of *T. trichiura* eggs, mirroring what we observed with *T. muris.* Although hatching occurred more slowly than for *T. muris* eggs, viability was not impacted by the long incubation, as larvae were still motile as observed by light microscopy.

To validate these findings within the mammalian host, we applied a monocolonization mouse model we recently established.[26] Germ-free mice are not permissive to *T. muris*

colonization; however, reproductively viable adult *T. muris* can be recovered from germ-free mice inoculated with *E. coli* alone, albeit at lower levels than conventional mice with a complete microbiota.[26] The number of adult worms from mice monocolonized with OA8 (*P. sordellii*) was comparable to those recovered from mice monocolonized with *E. coli*, despite lower levels of OA8 colonization (Figures 4H and S4D). Worms from OA8-monocolonized mice were developmentally mature based on their ability to lay eggs after overnight incubation. In fact, the number of eggs laid by *T. muris* recovered from mice monocolonized with OA8 was greater than *E. coli* (Figure 4I). Thus, OA8 is sufficient to support *T. muris* in the murine host and may play a role in the *Trichuris* life cycle at stages beyond egg hatching.

Because we previously showed that the increase in Clostridiales in mice and humans colonized with *Trichuris* is associated with a corresponding decrease in Bacteroidales,[14] we tested whether the OA isolates directly impact Bacteroidales growth. We found that OA12 and OA14, two isolates that did not induce *T. muris* egg hatching, were potent inhibitors of *P. vulgatus* and *B. theta* growth when cultured together in an *in vitro* bacterial competition assay (Figures S4E and S4F). Altogether, these results indicate that helminth-associated *Peptostreptococcaceae* species promote the *Trichuris* life cycle, while competition with *Bacteroides* species may not be a unique property of helminth-associated bacteria.

DISCUSSION

Whole-genome sequencing of the OA isolates and the availability of the largest Malaysian microbiome dataset to date[37] allowed us to uncover differential and high-resolution relationships between specific taxa and helminth colonization. Taxa corresponding to only a subset of OA isolates decreased in the microbiomes of individuals treated with deworming medication. The *Peptostreptococcaceae* isolates were particularly notable because they also mediated superior *T. muris* and *T. trichiura* egg hatching. The finding that OA8, *P. sordellii*, conferred greater fecundity of *T. muris* worms that develop in mice suggests additional roles for the microbiome post-hatching. Clostridia species are producers of immunogenic metabolites that impact T cell differentiation.[23,46–50] Given the known role of T helper lineages, including regulatory T cells on *T. muris* burden,[42,51,52] it will be important to analyze how the mammalian host immune system is impacted by various OA isolates and if this contributes to helminth susceptibility.

These results demonstrating that OA isolates influence hatching rates and adult worm fecundity raise the possibility that helminth colonization “primes” the host microbiome to be a more hospitable environment for either subsequent infections in the same host or infections in other hosts within the same community. Individuals living in the same household or environment share microbiome compositions.[53,54] Thus, naive hosts may acquire a pro-helminth microbiome by living in proximity with colonized individuals, thereby creating a positive feedback loop that favors propagation of helminths at the community level. A non-mutually exclusive possibility is that eggs shed into the environment are coated by bacteria that promote hatching when ingested together. These hypothetical mechanisms will be important to test in the future as factors that determine helminth burden and susceptibility to reinfection are poorly understood.

A previous study found that alterations to the microbiome following *T. muris* colonization of mice increases resistance to subsequent parasite colonization.[28] One reason for this potential discrepancy with our findings is that, depending on the mouse facility, the gut microbiome of inbred laboratory mice may not harbor a sufficient number of *Peptostreptococcaceae* or the proper strains necessary to observe enhancement. Similarly, a modest outgrowth of pro-helminth bacteria could be insufficient to benefit the parasite if the initial microbiome contains a high proportion of bacterial taxa that inhibit hatching. In fact, OA12, *E. hirae*, was found to be helminth associated in our metagenomic analyses while also inhibiting *Trichuris* egg hatching, suggesting that helminth colonization could result in increased inhibitory strains as well. However, it is also possible there are fundamental differences between the mouse and human parasite. In this context, it is notable that we successfully established a bacteria-induced hatching assay for *T. trichiura*. The criteria for efficient hatching may be more stringent for the human *Trichuris* species because only OA8 enhanced hatching. Serial passaging of *T. muris* in mice potentially selected for parasites that are optimally adapted to laboratory conditions. For example, eggs harvested under sterile laboratory conditions would not be in danger of pre-mature hatching outside the host upon contact with microbes in the soil. We hope that our newly established assay to study *T. trichiura* outside the human host can be applied to better understand the mechanisms involved in host adaptation and the dependencies of helminths on specific bacterial taxa.

Finally, regions with the highest incidence of immune diseases are associated with the lowest prevalence of helminth infections,[55,56] which may be explained by either direct effects of parasites on the host immune system or indirect effects through the microbiome.[14,24,57–59] If the difference in the abundance of species representing the OA isolates upon deworming or between rural and urban Malaysian populations reflects a larger shift in microbiome compositions, then it is possible that transition from a rural lifestyle in a helminth-endemic region to an industrialized setting may be leading to a loss of vulnerable low-abundance taxa. Whether the consequence of losing certain microbes from our microbiome is harmful or reflects an adaptation to modernization remains to be determined.[60] Considering the diversity of physiological processes associated with intestinal bacteria, we suggest additional investigation of understudied populations, including those in helminth-endemic regions, will yield valuable knowledge regarding the consequences of a shifting microbiota.

Limitations of the study

Although we validated the presence of taxa representing the OA isolates in a large metagenome dataset, our focus was limited to a modest number of species. Our ability to compare their abundance between samples was dependent on computational approaches that discriminate based on sequence identity and the availability of existing databases. We can confidently compare microbiomes at the species level because we are able to distinguish taxa with >99% sequence identity, but strains belonging to the same species can have important functional differences. A recent phylogenetic study of *Lachnospiraceae* (family within Clostridiales) isolated from individuals in the USA provides compelling evidence of substantial inter- and intra-species diversity in metabolic pathways linked with human health.[61] Given that our OA isolates include understudied taxa and a new species, similar

large-scale efforts to isolate and sequence bacteria from the microbiomes of individuals living in low- and middle-income countries are warranted. A second limitation of our study is the use of *T. muris* as a surrogate for the human parasite to study colonization of the mammalian host. Nevertheless, our observation that bacteria-mediated hatching of *T. muris* and *T. trichiura* are at least partially translatable is promising for further studies exploring the role of *Peptostreptococcaceae* species, especially *P. sordellii*, in the helminth life cycle.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by lead contact, Dr. Ken Cadwell (ken.cadwell@nyulangone.org).

Materials availability—All unique and stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- PacBio raw sequences and assembled whole genome sequences reported in this study are available under NCBI BioProject: PRJNA800461.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects—Prior to collection of samples from Kuala Pangsun village (Table S1), the human study was approved by the Research and Ethics Committee of Universiti Malaya Medical Center (UMMC) (i.e., MEC Ref. No. 824.11, No. 943.14 and No. 2017925-5593), National Medical Research Register (NMRR), Ministry of Health, Malaysia (Reference No.: NMRR-17-3055-37252), the Department of Orang Asli Development (JAKOA) [Ref. No.: JAKOA/pp.30.052Jld13 (12) & JAKOA/pp.30.052Jld14 (47)], and NYU IRB# i17-01068. Permission was also obtained from the Tok Batin, chieftain of the village. Kuala Pangsun village (101.88°E longitude, 3.21°N latitude) is situated in Hulu Langat district, the fifth largest district in Selangor state, Malaysia. The selection of this village was favorable in terms of logistics and feasibility, coupled with good cooperation from the villagers. The purpose and the procedure of this study was explained orally to all the participants by the investigator. Written consent was attained from all adult participants aged 18 and above. For children under 18 years old, written parental consent was obtained from their respective parents or guardian. Study exclusion criteria consisted of pregnant women, breastfeeding mothers and presence or perceived presence of any clinically significant disease. Information about human subjects from the large Orang Asli and Kuala Lumpur cohort which provided metagenomes for our analyses is described in.[37]

Germfree mice—Germfree (GF) C57BL/6J were bred and maintained in flexible-film isolators at the New York University Grossman School of Medicine Gnotobiotics Animal Facility. Absence of fecal bacteria was confirmed monthly by evaluating the presence of 16S DNA in stool samples by qPCR as previously described.[79] For inoculation with spore-forming stool fractions, bacteria and/or *Trichuris muris* eggs, GF mice were housed in Bioexclusion cages (Tecniplast) with access to sterile food and water. An equal amount of male and female mice 6–8 weeks of age were used for all experiments. All animal studies were performed according to protocols approved by the NYU Grossman School of Medicine Institutional Animal Care and Use Committee (protocol # IA16-00087).

Bacterial strains—In addition to the OA isolates from this study, *Phocaeicola* (previously *Bacteroides*) *vulgatus* was isolated by our lab previously,[62] *Bacteroides thetaiotaomicron* VPI-5482 was kindly provided by E. Martens (University of Michigan Medical School), *Pseudomonas aeruginosa* was kindly provided by A. Darwin (NYU Grossman School of Medicine)[63] and *Escherichia coli* strain BW25113 is from the National BioResource Project at the National Institute of Genetics, Japan.[64] All bacteria except *P. aeruginosa* and *E. coli* were cultured under anaerobic conditions in an anaerobic chamber (Coy Labs). Frozen glycerol stocks (30% glycerol) of all bacteria were prepared. Glycerol stocks of the OA isolates and *Bacteroides* species were streaked onto BRU plates (Anaerobe Systems) and incubated anaerobically for 48 hours at 37°C. PYG broth (Anaerobe Systems) inoculated with single colonies was grown at 37°C. OA1, 2, 4, 5, 6, 8, 11, 12, 13, and 14 were grown for 24 hours. OA3, 7, 9, and 10 required 3 days to reach similar turbidity. *P. aeruginosa* and *E. coli* glycerol stocks were streaked onto TSA or LB agar respectively and incubated aerobically for 24 hours at 37°C, after which single colonies were spiked into TSA or LB broth and grown overnight at 37°C with shaking at 225 rpm. To quantify colony forming units, we performed serial dilutions of liquid culture in sterile PBS and plated on BRU agar for the OA isolates, BBE agar (BD) for *Bacteroides* species, TSA agar for *P. aeruginosa* and LB agar for *E. coli.*

Parasite maintenance—Stock eggs of *Trichuris muris* E strain[14] were propagated and maintained as previously described.[80] Each egg batch was confirmed to hatch at ≥80% *in vitro* using the method described in Venzon et al.[26] before use in subsequent experiments.

Stock eggs of *Trichuris trichiura* were provided by the *Trichuris trichiura* egg Production Unit (TTPU) located at the Clinical Immunology Laboratory at the George Washington University. Whipworm eggs were isolated from the feces of a chronically infected human volunteer following a qualified standard procedure that includes a modified Simulated Gastric Fluid (SGF) method. After isolation, the eggs were stored for two months in flasks containing sulfuric acid (H_2SO_4) maintained at 25–30°C in a monitored incubator. Once embryonation was achieved, the eggs were transferred to a locked and monitored refrigerator at 2–8°C until further use. Controls for the manufacturing process involved: (i) tests for viability (hatching), which confirmed that more than 80% of the eggs were viable; (ii) species confirmation by polymerase chain reaction (PCR); and (iii) evaluation of the microbiological burden determined by bioburden testing by an outside-certified laboratory.

METHOD DETAILS

Sample collection from Orang Asli—A total number of 19 fecal samples were collected from 19 individuals (9 female, 10 male, ages 5–51) (Table S1). Participants were given screw-capped containers labeled with names. Containers with participants' samples were collected the following day, immediately frozen on dry ice, and transported on dry ice to the Department of Parasitology, Faculty of Medicine, Universiti Malaya. Fecal samples were then separated into three portions: (i) preserved in 2.5% potassium dichromate and stored at 4°C for intestinal helminth infection screening, (ii) aliquoted into anaerobic glass hungate tubes for isolation of bacteria, and (iii) aliquoted in 1.5mL cryovial tubes for long-term storage of samples. Aliquots (ii) and (iii) were transported on dry ice to NYU Grossman School of Medicine while aliquot (i) was kept at the University of Malaya to quantify helminth burden using Kato-Katz and formalin ether as described in[14]

Sample collection and processing for the large Orang Asli and Kuala Lumpur cohort which provided metagenomes for our analyses is described in.[37] In brief, fecal samples were collected from Orang Asli villagers right before administration of albendazole (ABZ) at 400mg/day for 3 consecutive days and follow-up samples were collected after 21 days. Individuals in urban Kuala Lumpur were not given ABZ and samples were only collected at one timepoint.

Isolation and identification of OA isolates from human stool—Spores were enriched from stool using a previously published protocol.[23] The following procedure was applied to specimens from the Orang Asli individuals with highest *Trichuris* burdens (Table S1 and Figure S1A). However, isolates from samples #88 and #135 were discarded due to contamination during the procedure and we proceeded with isolating bacteria from the remaining two samples (#20110 and #762) with Kato Katz intensities of 4,375 and 420 eggs.

Fecal samples frozen in hungate tubes were moved into an anaerobic chamber (Coy Labs). ~50mg of the sample was suspended in 10 times volume (w/v) of sterile pre-reduced PBS and passed through a 70 mm cell strainer. Suspensions were mixed with chloroform at a final concentration of 3%, shaken vigorously for 30 seconds, then incubated at 37°C for 1 hour. Chloroform was then removed through evaporation by bubbling with N_2 gas for 30 min. Suspensions from each donor were inoculated into 4–5 germ-free C57BL/6J mice by intra-gastric gavage (250ul per mouse). Stools on day 7 post-inoculation were collected and suspended in sterile PBS, homogenized in a bead beater in safe-lock tubes with 1.0 mm silica beads, streaked onto BRU agar in an anaerobic chamber, and incubated for 2–4 days at 37°C. 75 colonies were selected, with the aim of picking as many different morphologies as possible, spiked into PYG broth, incubated for 1–3 days at 37°C, and used to prepare sterile glycerol stocks. The inoculation loop used to pick each isolate was dipped into PCR grade water to perform a 16S colony PCR on each isolate using primers 16S-Fwd and 16S-Rev[65] to remove duplicates from the pool. After identifying 14 isolates with unique 16S sequences and/or morphologies, we extracted DNA from liquid cultures of these isolates using phenol-chloroform, then subcloned and sequenced the full 16S gene using the primers 16S-Fwd and 16S-Rev and the TOPO™ TA Cloning™ Kit with pCR™2.1-TOPO™ plasmid

(ThermoFisher). Sequences were aligned to the EzBioCloud 16S rRNA database[66] for initial taxonomic identification.

Whole-genome sequencing and genome assembly—DNA was isolated from liquid culture of all isolates and lysed overnight in Buffer B1 (QIAGEN) containing 200ug/mL RNase A (ThermoScientific), 1250U/ul lysozyme (Epicentre) and 100ul Proteinase K (QIAGEN), at 37°C. High molecular weight DNA was isolated using the Genomic-tip 100/G kit (QIAGEN) and run on the Agilent 2200 TapeStation system with Genomic ScreenTape (cat# 5067-5365) to verify large fragment DNA. Once verified, 1.5ug of DNA was sheared using the Covaris G-TUBE (cat# 520079) and spun for 2min in each direction at 5.4rpm in the Eppendorf 5415 D centrifuge to shear the DNA to an average size of 13,000bp length. Each sheared sample was barcoded with the PacBio Barcoded Overhang Adapter Kit 8A (cat# 101-628-400) during library prep using the PacBio SMRTbell Express Template Prep Kit 2.0 (cat# 100-938-900). After barcoding, each sample was verified on the Invitrogen Qubit system with the dsDNA HS kit (cat# Q32854) for concentration and on the Agilent 2200 TapeStation system for size. Samples were combined to an equal concentration using the PacBio Express microbial multiplexing calculator to give an estimated size of about 30mb total. The PacBio Sequencing Primer v4 was conditioned and annealed to the pooled sample based on the Diffusion loading instructions found on the PacBio SMRT Link software version 8.0.0805529. Following primer binding, the diluted PacBio Sequel Polymerase 3.0 (cat# 100-500-400) was bound to the sample. After binding, the sample was cleaned with PacBio Ampure Beads (cat# 100-265-900) and eluted in PacBio Elution Buffer (cat# 101633–500), and the final concentration of the pool was verified on the Invitrogen Qubit system as before. The pool was combined with final loading reagents (part of the Express kit) and run on CLR mode (continuous long reads) on the PacBio Sequel using a 20hr run time with the SMRT Cell 1M v3 LR tray (cat# 101–531-001) and the Sequel Sequencing Kit 3.0 reagent plate (cat# 101-597-900). Following the sequencing, all genomes were assembled through SMRT Link software except OA2. The OA2 genome was assembled using Flye 2.8.1 in pacbio-raw mode with asm-coverage 150, sub-sampling the reads for the initial assembly to a coverage of 150x.[67]

Species identification, phylogenetic analysis, orthogroup and KEGG UMAP plots, and pangenome plots—To obtain a phylogenetic context for the OA isolates, a total of 190,173 genome assemblies that had been assigned to the phylum Firmicutes (NCBI taxonomy ID: 1239) were downloaded from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) on November 24, 2021. Of those, only 1,009 were retained after comparing them to all assemblies of the OA isolates with fastANI v.1.32[68] and filtering for estimated average nucleotide identity (ANI) values R0.95. The workflow classify_wf from GTDB-Tk v.1.5.0[29] was applied to 1,023 retained GenBank and OA assemblies to obtain a taxonomic classification, using default parameters. We adopted GTDB-Tk taxonomic designations. Using fasttree v.1.11.2,[69] a phylogenetic tree was calculated with the alignment of concatenated marker genes that was produced by GTDB-Tk from all 1,023 assemblies, using default parameters. Phylogenies were plotted with iToL.[70]

The 1,023 assemblies (above) were annotated with dfast v.1.2.13[71] and grouped by GTDB-Tk species designation. Orthofinder v.2.5.472 was used with default parameters for computing orthogroups (also called clusters of orthologous genes) for each individual species group and for all 14 OA assemblies as a group. The phylogenetic tree in Figure 1B was produced by orthofinder using the group of OA assemblies only as an input. However, the largest group of 623 assemblies classified as *Clostridium_P perfringens* could not be processed with orthofinder due to its size. For this reason, a subset of 221 assemblies was chosen from the 623 such that the ANI between the assemblies in the subset was maximized. KEGG orthologs were mapped to the predicted proteomes of each assembly (dfast, above) using KofamScan v.1.3.0 and HMM profiles downloaded from KEGG on June 24, 2021.[73,74] The resulting sets of KEGG ortholog identifiers were mapped and annotated with all KEGG BRITE hierarchies released on March 30, 2020.

The R programming language v.4.1.1[75] was used for plotting, using ggplot2 v.3.3.5[78]. UMAP dimensionality reduction was carried out using UMAP v.0.2.7 within R (https://github.com/tkonopka/umap). For the orthogroup plots of each species, the input to UMAP was a binary matrix indicating presence or absence of an orthogroup in each strain. For the plot of KEGG metabolic genes in *Clostridium_P perfringens*, the numbers of the KEGG orthologs in the KEGG BRITE hierarchy *09100 Metabolism* were tabulated according to the categories of the most specific level of this hierarchy, *D*, and a matrix indicating the numbers of genes in each category and strain was input to UMAP.

On the *Clostridium_P Perfringens* UMAP plot of metabolic KEGG genes, OA14 appeared to form a cluster with 28 other assemblies (Figure S1D and Table S2), henceforth called “OA14 cluster”. To learn more about what metabolic functions could set the OA14 cluster apart from the rest of the assemblies, we tabulated the occurrence of KEGG orthologs and of terms from the KEGG BRITE hierarchy *09100 Metabolism* at level per *D* in each *Clostridium_P perfringens* assembly. We then used the Wilcoxon rank sum test to quantify the statistical significance of differences in the prevalence of KEGG orthologs or of BRITE terms between the OA14 cluster and the rest of the assemblies.[81] The Bonferroni method was used to correct p-values for multiple testing.[82]

To produce pangenome plots for each species, the numbers of orthogroups in the core and pan-genome were successively computed for up to 100 size-k subsets of the available assemblies. Specifically, we started with $k = 2$ and successively increased k until it reached the number of available assemblies for a species, n. For each k, the number of size-k combinations of n genomes, $n\mathrm{C}k$ was computed. If $n\mathrm{C}k$ % 100, all possible combinations of the assemblies were considered when computing core and pangenome sizes. If $n\mathrm{C}k > 100$, long computation times and combinatorial explosion were prevented by only considering 100 size-k random samples of the assemblies that were drawn without replacement. Using linear regression, two Heap’s law curves $N = \beta k^{\alpha}$ were fit for each species: one for the median core genome sizes and one for the median pangenome sizes. Only 50% highest values of k were used for fitting Heap’s law curves.[83]

Metagenome classification using Krakenuniq—Malaysian metagenomes were obtained from NCBI BioProject: PRJNA797994[37] and metagenomes from the Human

Microbiome Project (HMP) were obtained from https://www.hmpdacc.org/hmp/.[39,40] The file type used in the HMP data portal was wgs_raw_seq_set and we only used fecal samples from healthy human subjects. Quality control, trimming and human host genome removal were performed by KneadData v.0.7.4 with default settings. In KneadData, sequence reads were mapped against human reference genome (hg37) using Bowtie2 with default parameters (very-sensitive end-to-end alignment) to remove the human host genome.[77,84] Human host genome removal was not performed on HMP metagenomes because the human reference genome had already been filtered in the HMP data. Next, FastQC was used to verify quality control.[78] Krakenuniq was used for taxonomic classification.[38] The filtered reads from the Malaysia dataset and HMP were mapped against the OA isolate genomes. Whole genomes from the 47 closest relatives of the OA isolates were used to ensure that mapping reflected matches to sequences found exclusively in our isolates (Table S4). In Figure S1H, classifications with lower than 50 k-mer counts were removed due to low confidence. To determine the abundance of *Peptostreptococcaceae* family members, filtered reads from metagenomes were mapped against all 113 complete genomes within the *Peptostreptococcaceae* family available on NCBI as of July 2022 (Table S4). The plots in Figures S2L and S2M represent mapping percentages at the family level.

Albendazole assay—Albendazole (ABZ) (Sigma) was dissolved in DMSO to make 2.5mM and 500mM stocks that were diluted with sterile PYG to a final concentration of 10uM and 200uM. Concentrations used in *in vitro* experiments were chosen based on max solubility in DMSO and max amount of DMSO added to culture (0.5%). Single colonies of bacteria were spiked into 5mL PYG containing 10um ABZ, 200uM ABZ or 0.5% DMSO as a vehicle control, in duplicates in sterile polypropylene tubes. After incubation for 24 hours at 37°C in anaerobic conditions, serial dilutions of the bacterial culture in sterile PBS were plated on BRU agar for OA isolates and BBE agar for *Bacteroides* species and incubated for 48 hours to count colonies and calculate CFUs/mL of culture.

T. muris egg hatching assay—25ul embryonated *T. muris* eggs at a concentration of 1 egg/1ul in sterile water were mixed with 15ul sterile PYG media and 10ul overnight bacterial culture in triplicates in a 48 well plate. PYG-control wells contained an additional 10ul sterile PYG media instead of bacterial culture. Plates were incubated at 37°C in an anaerobic chamber and hatched eggs were quantified every hour over the course of three hours on the Zeiss Primovert microscope, by counting hatched and embryonated unhatched eggs in each well. Unembryonated eggs, which lack visible larvae and have disordered contents, were excluded due to their inability to hatch. A separate 48-well plate was used per timepoint as the plate needed to be removed from the anaerobic chamber to count colonies at each timepoint.

T. muris in vivo infection in germ-free mice—Male and female germ-free C57BL/6J mice were monocolonized at 6–8 weeks of age by oral gavage with ~1×10^7 colony forming units (CFU) of indicated bacteria. 7 and 28 days later, mice were infected by oral gavage with ~100 embryonated *T. muris* eggs. 14 days after the second gavage of *T. muris* eggs, mice were euthanized, and individual worms were collected from the cecal contents of all infected mice and washed in RPMI 1640 (Gibco) supplemented with penicillin (100IU/mL)

and streptomycin (100ug/mL; Corning). To evaluate egg laying, each worm was placed into individual wells of a 48 well plate with 200mL supplemented RPMI. Plates were then incubated overnight at 37°C in a closed tupperware (Sistema) lined with damp paper towels. The following day, eggs laid were enumerated using a Zeiss Primovert light microscope at 10X.

T. trichiura egg hatching assay—50ul embryonated *T. muris* eggs at a concentration of 1 egg/1ul in sterile water were mixed with 240ul sterile PYG media and 10ul overnight bacterial culture in triplicates in a 96-well plate. PYG-control wells contained an additional 10ul sterile PYG media instead of bacterial culture. Plates were sealed with parafilm to prevent evaporation and incubated at 37°C in an anaerobic chamber for 6 days. Hatched eggs were quantified on the Zeiss Primovert microscope, by counting hatched and embryonated unhatched eggs in each well, as well as larvae. Unembryonated eggs, which lack visible larvae and have disordered contents, were excluded due to their inability to hatch. For conditions in which bacterial overgrowth prohibited clearly seeing eggs and larvae, supernatant from the wells were transferred to a different well and the remaining contents resuspended in 300ul PBS. If needed for better visibility. this was repeated a second time. Confocal images were acquired by transferring the contents of one well containing eggs and larvae to a glass-bottom dish (MatTek) and imaging on the Nikon Eclipse Ti2-E inverted microscope at 60X oil. Images were processed using NIS-Elements (Nikon).

Bacterial competition assay—Single colonies of *P. vulgatus* or *B. thetaiotaomicron* were inoculated into duplicates of 7ul PYG media along with single colonies of OA isolates. Cultures were grown for 24 hours at 37°C under anaerobic conditions after which serial dilutions of the bacterial culture in sterile PBS were plated on BRU agar for Clostridia and BBE agar for Bacteroides and incubated for 48 hours to count colonies and calculate CFUs/mL of culture.

QUANTIFICATION AND STATISTICAL ANALYSIS

For *in vitro* experiments, the number of repeats per group is annotated in corresponding figure legends. Significance for all experiments was assessed using GraphPad Prism software (GraphPad). Specific tests are annotated in corresponding figure legends. p values correlate with symbols: ns or no symbol = not significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We would like to thank Margie Alva, Juan Carrasquillo, and David Basnight for their help in the NYU Gnotobiotic Facility, Paul Zappile, Dacia Dimartino, Kim Castelli, and Christian Marier at the NYU Langone Genome Technology Center for their help with whole-genome sequencing and assemblies, and the NYU Reagent Preparation service for providing bacterial media. We thank Gregory Putzel for help with genome analyses, Menghan Liu for help with krakenuniq, and Mericien Venzon for providing *T. muris* eggs. We would also like to thank Juan Lafaille and members of the Cadwell and Loke Labs for their constructive comments. Figure S1A was created using BioRender.com This work was in part funded by NIH grants DK093668 (K.C.), HL123340 (K.C.),

AI130945 (K.C., P.L., and Y.A.L.L.), AI140754 (K.C.), DK124336 (K.C.), AI121244 (K.C.), and AI133977 (P.L. and V.J.T.). Further funding was provided by Faculty Scholar grant from the Howard Hughes Medical Institute (K.C.), Crohn's and Colitis Foundation (K.C.), Kenneth Rainin Foundation (K.C.), Judith & Stewart Colton Center of Autoimmunity (K.C.), NIH grant 2T32AI007180 (S.S.), and the NYU Langone Health Antimicrobial-Resistant Pathogens Program (A.P.). PacBio sequencing reagents were partially paid for through the PacBio Local SMRT Grant (K.C.). NYU Langone's Genome Technology Center and Microscopy Laboratory are supported by NIH Cancer Center Support Grant P30CA016087 and further support for the PacBio Sequel was provided by NIH Shared Instrumentation Grant 1S10OD023423-01. The research was supported in part by the Intramural Research Program of NIAID.

REFERENCES

1. WHO (2020). Soil-Transmitted Helminth Infections. WHO Fact Sheets (World Health Organization).
2. Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, and Hotez PJ (2006). Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet 367, 1521–1532. 10.1016/s0140-6736(06)68653-4. [PubMed: 16679166]
3. Kupritz J, Angelova A, Nutman TB, and Gazzinelli-Guimaraes PH (2021). Helminth-induced human gastrointestinal dysbiosis: a systematic review and meta-analysis reveals insights into altered taxon diversity and microbial gradient collapse. mBio 12, e0289021. 10.1128/mBio.02890-21. [PubMed: 34933444]
4. Abdill RJ, Adamowicz EM, and Blekhman R (2022). Public human microbiome data are dominated by highly developed countries. PLoS Biol. 20, e3001536. 10.1371/journal.pbio.3001536. [PubMed: 35167588]
5. Blaser MJ (2016). Antibiotic use and its consequences for the normal microbiome. Science 352, 544–545. 10.1126/science.aad9358. [PubMed: 27126037]
6. Blaser MJ, and Falkow S (2009). What are the consequences of the disappearing human microbiota? Nat. Rev. Microbiol. 7, 887–894. 10.1038/nrmicro2245. [PubMed: 19898491]
7. Easton AV, Quiñones M, Vujkovic-Cvijin I, Oliveira RG, Kepha S, Odiere MR, Anderson RM, Belkaid Y, and Nutman TB (2019). The impact of anthelmintic treatment on human gut microbiota based on cross-sectional and pre- and postdeworming comparisons in Western Kenya. mBio 10, e00519–19. 10.1128/mBio.00519-19. [PubMed: 31015324]
8. Lee SC, Tang MS, Lim YAL, Choy SH, Kurtz ZD, Cox LM, Gundra UM, Cho I, Bonneau R, Blaser MJ, et al. (2014). Helminth colonization is associated with increased diversity of the gut microbiota. PLoS Negl. Trop. Dis. 8, e2880. 10.1371/journal.pntd.0002880. [PubMed: 24851867]
9. Kay GL, Millard A, Sergeant MJ, Midzi N, Gwisai R, Mduluza T, Ivens A, Nausch N, Mutapi F, and Pallen M (2015). Differences in the faecal microbiome in schistosoma haematobium infected children vs. Uninfected children. PLoS Negl. Trop. Dis. 9, e0003861. 10.1371/journal.pntd.0003861. [PubMed: 26114287]
10. Jenkins TP, Rathnayaka Y, Perera PK, Peachey LE, Nolan MJ, Krause L, Rajakaruna RS, and Cantacessi C (2017). Infections by human gastrointestinal helminths are associated with changes in faecal microbiota diversity and composition. PLoS One 12, e0184719. 10.1371/journal.pone.0184719. [PubMed: 28892494]
11. Rosa BA, Supali T, Gankpala L, Djuardi Y, Sartono E, Zhou Y, Fischer K, Martin J, Tyagi R, Bolay FK, et al. (2018). Differential human gut microbiome assemblages during soil-transmitted helminth infections in Indonesia and Liberia. Microbiome 6, 33. 10.1186/s40168-018-0416-5. [PubMed: 29486796]
12. Cooper P, Walker AW, Reyes J, Chico M, Salter SJ, Vaca M, and Parkhill J (2013). Patent human infections with the whipworm, Trichuris trichiura, are not associated with alterations in the faecal microbiota. PLoS One 8, e76573. 10.1371/journal.pone.0076573. [PubMed: 24124574]
13. Martin I, Djuardi Y, Sartono E, Rosa BA, Supali T, Mitreva M, Houwing-Duistermaat JJ, and Yazdanbakhsh M (2018). Dynamic changes in human-gut microbiome in relation to a placebo-controlled anthelminthic trial in Indonesia. PLoS Negl. Trop. Dis. 12, e0006620. 10.1371/journal.pntd.0006620. [PubMed: 30091979]

14. Ramanan D, Bowcutt R, Lee SC, Tang MS, Kurtz ZD, Ding Y, Honda K, Gause WC, Blaser MJ, Bonneau RA, et al. (2016). Helminth infection promotes colonization resistance via type 2 immunity. Science 352, 608–612. 10.1126/science.aaf3229. [PubMed: 27080105]
15. Zaiss MM, Rapin A, Lebon L, Dubey LK, Mosconi I, Sarter K, Piersigilli A, Menin L, Walker AW, Rougemont J, et al. (2015). The intestinal microbiota contributes to the ability of helminths to modulate allergic inflammation. Immunity 43, 998–1010. 10.1016/j.immuni.2015.09.012. [PubMed: 26522986]
16. Dürre P (2014). Physiology and sporulation in Clostridium. Microbiol. Spectr. 2, TBS-0010–2012. 10.1128/microbiolspec.TBS-0010-2012.
17. Nagano Y, Itoh K, and Honda K (2012). The induction of Treg cells by gut-indigenous Clostridium. Curr. Opin. Immunol. 24, 392–397. 10.1016/j.coi.2012.05.007. [PubMed: 22673877]
18. Stefka AT, Feehley T, Tripathi P, Qiu J, McCoy K, Mazmanian SK, Tjota MY, Seo GY, Cao S, Theriault BR, et al. (2014). Commensal bacteria protect against food allergen sensitization. Proc. Natl. Acad. Sci. USA 111, 13145–13150. 10.1073/pnas.1412008111. [PubMed: 25157157]
19. Becattini S, Littmann ER, Carter RA, Kim SG, Morjaria SM, Ling L, Gyaltshen Y, Fontana E, Taur Y, Leiner IM, and Pamer EG (2017). Commensal microbes provide first line defense against Listeria monocytogenes infection. J. Exp. Med. 214, 1973–1989. 10.1084/jem.20170495. [PubMed: 28588016]
20. Kim YG, Sakamoto K, Seo SU, Pickard JM, Gillilland MG 3rd, Pudlo NA, Hoostal M, Li X, Wang TD, Feehley T, et al. (2017). Neonatal acquisition of Clostridia species protects against colonization by bacterial pathogens. Science 356, 315–319. 10.1126/science.aag2029. [PubMed: 28428425]
21. Caballero S, Kim S, Carter RA, Leiner IM, Sušac B, Miller L, Kim GJ, Ling L, and Pamer EG (2017). Cooperating commensals restore colonization resistance to vancomycin-resistant Enterococcus faecium. Cell Host Microbe 21, 592–602.e4. 10.1016/j.chom.2017.04.002. [PubMed: 28494240]
22. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G, et al. (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. USA 105, 16731–16736. 10.1073/pnas.0804812105. [PubMed: 18936492]
23. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, Fukuda S, Saito T, Narushima S, Hase K, et al. (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature 500, 232–236. 10.1038/nature12331. [PubMed: 23842501]
24. Lawson MAE, Roberts IS, and Grencis RK (2021). The interplay between Trichuris and the microbiota. Parasitology 148, 1806–1813. 10.1017/s0031182021000834.
25. Reynolds LA, Smith KA, Filbey KJ, Harcus Y, Hewitson JP, Redpath SA, Valdez Y, Yebra MJ, Finlay BB, and Maizels RM (2014). Commensal-pathogen interactions in the intestinal tract: lactobacilli promote infection with, and are promoted by, helminth parasites. Gut Microbes 5, 522–532. 10.4161/gmic.32155. [PubMed: 25144609]
26. Venzon M, Das R, Luciano DJ, Burnett J, Park HS, Devlin JC, Kool ET, Belasco JG, Hubbard EJA, and Cadwell K (2022). Microbial byproducts determine reproductive fitness of free-living and parasitic nematodes. Cell Host Microbe 30, 786–797.e8. 10.1016/j.chom.2022.03.015. [PubMed: 35413267]
27. Hayes KS, Bancroft AJ, Goldrick M, Portsmouth C, Roberts IS, and Grencis RK (2010). Exploitation of the intestinal microflora by the parasitic nematode Trichuris muris. Science 328, 1391–1394. 10.1126/science.1187703. [PubMed: 20538949]
28. White EC, Houlden A, Bancroft AJ, Hayes KS, Goldrick M, Grencis RK, and Roberts IS (2018). Manipulation of host and parasite microbiotas: survival strategies during chronic nematode infection. Sci. Adv. 4, eaap7399. 10.1126/sciadv.aap7399. [PubMed: 29546242]
29. Chaumeil P-A, Mussig AJ, Hugenholtz P, and Parks DH (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36, 1925–1927. 10.1093/bioinformatics/btz848. [PubMed: 31730192]

30. Maria R, Dutta SD, Thete SG, and AlAttas MH (2021). Evaluation of antibacterial properties of organic gutta-percha solvents and synthetic solvents against Enterococcus faecalis. J. Int. Soc. Prev. Community Dent. 11, 179–183. 10.4103/jispcd.JISPCD_422_20. [PubMed: 34036080]

31. Suchomel M, Lenhardt A, Kampf G, and Grisold A (2019). Enterococcus hirae, Enterococcus faecium and Enterococcus faecalis show different sensitivities to typical biocidal agents used for disinfection. J. Hosp. Infect. 103, 435–440. 10.1016/j.jhin.2019.08.014. [PubMed: 31449920]

32. Gerritsen J, Umanets A, Staneva I, Hornung B, Ritari J, Paulin L, Rijkers GT, de Vos WM, and Smidt H (2018). Romboutsia hominis sp. nov., the first human gut-derived representative of the genus Romboutsia, isolated from ileostoma effluent. Int. J. Syst. Evol. Microbiol. 68, 3479–3486. 10.1099/ijsem.0.003012. [PubMed: 30226461]

33. Sakamoto M, Iino T, Yuki M, and Ohkuma M (2018). Lawsonibacter asaccharolyticus gen. nov., sp. nov., a butyrate-producing bacterium isolated from human faeces. Int. J. Syst. Evol. Microbiol. 68, 2074–2081. 10.1099/ijsem.0.002800. [PubMed: 29745868]

34. McInnes L, Healy J, and Melville J (2018). Umap: Uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. 10.48550/arXiv.1802.03426.

35. Shimizu T, Ohtani K, Hirakawa H, Ohshima K, Yamashita A, Shiba T, Ogasawara N, Hattori M, Kuhara S, and Hayashi H (2002). Complete genome sequence of Clostridium perfringens, an anaerobic flesheater. Proc. Natl. Acad. Sci. USA 99, 996–1001. 10.1073/pnas.022493799. [PubMed: 11792842]

36. Vernikos G, Medini D, Riley DR, and Tettelin H (2015). Ten years of pan-genome analyses. Curr. Opin. Microbiol. 23, 148–154. 10.1016/j.mib.2014.11.016. [PubMed: 25483351]

37. Tee MZ, Er YX, Easton AV, Yap NJ, Lee IL, Devlin J, Chen Z, Ng KS, Subramanian P, Angelova A, et al. (2022). Gut microbiome of helminth infected indigenous Malaysians is context dependent. Preprint at bioRxiv. 10.1101/2022.01.21.477162.

38. Breitwieser FP, Baker DN, and Salzberg SL (2018). KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. Genome Biol. 19, 198. 10.1186/s13059-018-1568-0.. [PubMed: 30445993]

39. Human Microbiome Project Consortium (2012). A framework for human microbiome research. Nature 486, 215–221. 10.1038/nature11209. [PubMed: 22699610]

40. Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. 10.1038/nature11234. [PubMed: 22699609]

41. Rapin A, Chuat A, Lebon L, Zaiss MM, Marsland BJ, and Harris NL (2020). Infection with a small intestinal helminth, Heligmosomoides polygyrus bakeri, consistently alters microbial communities throughout the murine small and large intestine. Int. J. Parasitol. 50, 35–46. 10.1016/j.ijpara.2019.09.005. [PubMed: 31759944]

42. Klementowicz JE, Travis MA, and Grencis RK (2012). Trichuris muris: a model of gastrointestinal parasite infection. Semin. Immunopathol. 34, 815–828. 10.1007/s00281-012-0348-2. [PubMed: 23053395]

43. Bogitsh BJ, Carter CE, and Oeltmann TN (2013). Chapter 16 - intestinal nematodes. In Human Parasitology, Fourth Edition, Bogitsh BJ, Carter CE, and Oeltmann TN, eds. (Academic Press), pp. 291–327. 10.1016/B978-0-12-415915-0.00016-9.

44. Koyama K (2016). Bacteria-induced hatching of Trichuris muris eggs occurs without direct contact between eggs and bacteria. Parasitol. Res. 115, 437–440. 10.1007/s00436-015-4795-2. [PubMed: 26481492]

45. Lück R, and Deppenmeier U (2022). Genetic tools for the redirection of the central carbon flow towards the production of lactate in the human gut bacterium Phocaeicola (Bacteroides) vulgatus. Appl. Microbiol. Biotechnol. 106, 1211–1225. 10.1007/s00253-022-11777-6. [PubMed: 35080666]

46. Kanamori M, Nakatsukasa H, Okada M, Lu Q, and Yoshimura A (2016). Induced regulatory T cells: their development, stability, and applications. Trends Immunol. 37, 803–811. 10.1016/j.it.2016.08.012. [PubMed: 27623114]

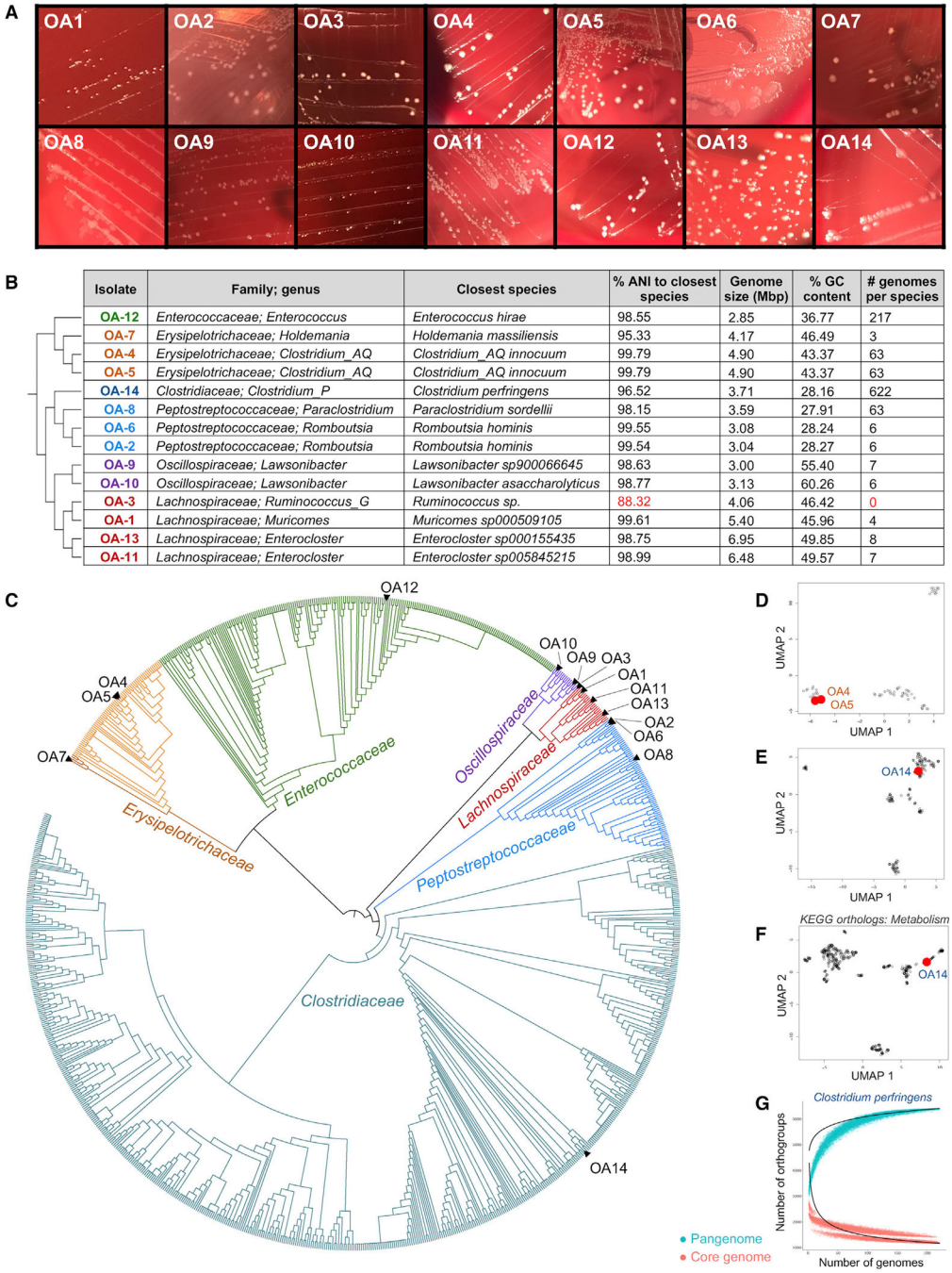
47. Li YN, Huang F, Cheng HJ, Li SY, Liu L, and Wang LY (2014). Intestine-derived Clostridium leptum induces murine tolerogenic dendritic cells and regulatory T cells in vitro. Hum. Immunol. 75, 1232–1238. 10.1016/j.humimm.2014.09.017. [PubMed: 25300998]

48. Cantorna MT, Lin YD, Arora J, Bora S, Tian Y, Nichols RG, and Patterson AD (2019). Vitamin D regulates the microbiota to control the numbers of RORγt/FoxP3+ regulatory T cells in the colon. Front. Immunol. 10, 1772. 10.3389/fimmu.2019.01772. [PubMed: 31417552]
49. Narushima S, Sugiura Y, Oshima K, Atarashi K, Hattori M, Suematsu M, and Honda K (2014). Characterization of the 17 strains of regulatory T cell-inducing human-derived Clostridia. Gut Microbes 5, 333–339. 10.4161/gmic.28572. [PubMed: 24642476]
50. Cai J, Sun L, and Gonzalez FJ (2022). Gut microbiota-derived bile acids in intestinal immunity, inflammation, and tumorigenesis. Cell Host Microbe 30, 289–300. 10.1016/j.chom.2022.02.004. [PubMed: 35271802]
51. D'Elia R, Behnke JM, Bradley JE, and Else KJ (2009). Regulatory T cells: a role in the control of helminth-driven intestinal pathology and worm survival. J. Immunol. 182, 2340–2348. 10.4049/jimmunol.0802767. [PubMed: 19201888]
52. McSorley HJ, and Maizels RM (2012). Helminth infections and host immune regulation. Clin. Microbiol. Rev. 25, 585–608. 10.1128/CMR.05040-11. [PubMed: 23034321]
53. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, Caporaso JG, Knights D, Clemente JC, Nakielny S, et al. (2013). Cohabiting family members share microbiota with one another and with their dogs. Elife 2, e00458. 10.7554/eLife.00458. [PubMed: 23599893]
54. Gacesa R, Kurilshikov A, Vich Vila A, Sinha T, Klaassen MAY, Bolte LA, Andreu-Sánchez S, Chen L, Collij V, Hu S, et al. (2022). Environmental factors shaping the gut microbiome in a Dutch population. Nature 604, 732–739. 10.1038/s41586-022-04567-7. [PubMed: 35418674]
55. Elliott DE, and Weinstock JV (2012). Helminth-host immunological interactions: prevention and control of immune-mediated diseases. Ann. N. Y. Acad. Sci. 1247, 83–96. 10.1111/j.1749-6632.2011.06292.x. [PubMed: 22239614]
56. Bach JF (2018). The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. Nat. Rev. Immunol. 18, 105–120. 10.1038/nri.2017.111. [PubMed: 29034905]
57. Maizels RM (2020). Regulation of immunity and allergy by helminth parasites. Allergy 75, 524–534. 10.1111/all.13944. [PubMed: 31187881]
58. Maizels RM, Smits HH, and McSorley HJ (2018). Modulation of host immunity by helminths: the expanding repertoire of parasite effector molecules. Immunity 49, 801–818. 10.1016/j.immuni.2018.10.016. [PubMed: 30462997]
59. Loke P, and Lim YAL (2015). Helminths and the microbiota: parts of the hygiene hypothesis. Parasite Immunol. 37, 314–323. 10.1111/pim.12193. [PubMed: 25869420]
60. Carmody RN, Sarkar A, and Reese AT (2021). Gut microbiota through an evolutionary lens. Science 372, 462–463. 10.1126/science.abf0590. [PubMed: 33926939]
61. Sorbara MT, Littmann ER, Fontana E, Moody TU, Kohout CE, Gjonbalaj M, Eaton V, Seok R, Leiner IM, and Pamer EG (2020). Functional and genomic variation between human-derived isolates of Lachnospiraceae reveals inter- and intra-species diversity. Cell Host Microbe 28, 134–146.e4. 10.1016/j.chom.2020.05.005. [PubMed: 32492369]
62. Ramanan D, Tang MS, Bowcutt R, Loke P, and Cadwell K (2014). Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. Immunity 41, 311–324. 10.1016/j.immuni.2014.06.015. [PubMed: 25088769]
63. Srivastava D, Seo J, Rimal B, Kim SJ, Zhen S, and Darwin AJ (2018). A proteolytic complex targets multiple cell wall hydrolases in Pseudomonas aeruginosa. mBio 9, 009722–18. 10.1128/mBio.00972-18.
64. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, and Mori H (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006.0008. 10.1038/msb4100050.
65. Brugiroux S, Beutler M, Pfann C, Garzetti D, Ruscheweyh HJ, Ring D, Diehl M, Herp S, Lötscher Y, Hussain S, et al. (2016). Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium. Nat. Microbiol. 2, 16215. 10.1038/nmicrobiol.2016.215. [PubMed: 27869789]
66. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, and Chun J (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int. J. Syst. Evol. Microbiol. 67, 1613–1617. 10.1099/ijsem.0.001755. [PubMed: 28005526]

67. Kolmogorov M, Yuan J, Lin Y, and Pevzner PA (2019). Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37, 540–546. 10.1038/s41587-019-0072-8. [PubMed: 30936562]
68. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, and Aluru S (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9, 5114. 10.1038/s41467-018-07641-9. [PubMed: 30504855]
69. Price MN, Dehal PS, and Arkin AP (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26, 1641–1650. 10.1093/molbev/msp077. [PubMed: 19377059]
70. Letunic I, and Bork P (2021). Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. 10.1093/nar/gkab301. [PubMed: 33885785]
71. Tanizawa Y, Fujisawa T, and Nakamura Y (2018). DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34, 1037–1039. 10.1093/bioinformatics/btx713. [PubMed: 29106469]
72. Emms DM, and Kelly S (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20, 238. 10.1186/s13059-019-1832-y. [PubMed: 31727128]
73. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, and Ogata H (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36, 2251–2252. 10.1093/bioinformatics/btz859. [PubMed: 31742321]
74. Kanehisa M, and Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30. 10.1093/nar/28.1.27. [PubMed: 10592173]
75. R Core Team (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.
76. Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).
77. Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]
78. Andrews S (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data.
79. Kernbauer E, Ding Y, and Cadwell K (2014). An enteric virus can replace the beneficial function of commensal bacteria. Nature 516, 94–98. 10.1038/nature13960. [PubMed: 25409145]
80. Antignano F, Mullaly SC, Burrows K, and Zaph C (2011). Trichuris muris infection: a model of type 2 immunity and inflammation in the gut. J. Vis. Exp. 10.3791/2774.
81. Wilcoxon F (1992). Individual comparisons by ranking methods. In Break-throughs in Statistics: Methodology and Distribution, Kotz S and Johnson NL, eds. (Springer), pp. 196–202. 10.1007/978-1-4612-4380-9_16.
82. Noble WS (2009). How does multiple testing correction work? Nat. Biotechnol. 27, 1135–1137. 10.1038/nbt1209-1135. [PubMed: 20010596]
83. Tettelin H, Riley D, Cattuto C, and Medini D (2008). Comparative genomics: the bacterial pan-genome. Curr. Opin. Microbiol. 11, 472–477. 10.1016/j.mib.2008.09.006. [PubMed: 19086349]
84. Bolger AM, Lohse M, and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. 10.1093/bioinformatics/btu170 [PubMed: 24695404]

Highlights

- Isolation of Clostridia from the gut microbiome of helminth-colonized human
- *Peptostreptococcaceae* family of Clostridia are associated with helminths in humans
- *Peptostreptococcaceae* induce *Trichuris muris* and *Trichuris trichiura* egg hatching



Isolate	Family; genus	Closest species	% ANI to closest species	Genome size (Mbp)	% GC content	# genomes per species
OA-12	*Enterococcaceae; Enterococcus*	*Enterococcus hirae*	98.55	2.85	36.77	217
OA-7	*Erysipelotrichaceae; Holdemania*	*Holdemania massiliensis*	95.33	4.17	46.49	3
OA-4	*Erysipelotrichaceae; Clostridium_AQ*	*Clostridium_AQ innocuum*	99.79	4.90	43.37	63
OA-5	*Erysipelotrichaceae; Clostridium_AQ*	*Clostridium_AQ innocuum*	99.79	4.90	43.37	63
OA-14	*Clostridiaceae; Clostridium_P*	*Clostridium perfringens*	96.52	3.71	28.16	622
OA-8	*Peptostreptococcaceae; Paraclostridium*	*Paraclostridium sordellii*	98.15	3.59	27.91	63
OA-6	*Peptostreptococcaceae; Romboutsia*	*Romboutsia hominis*	99.55	3.08	28.24	6
OA-2	*Peptostreptococcaceae; Romboutsia*	*Romboutsia hominis*	99.54	3.04	28.27	6
OA-9	*Oscillospiraceae; Lawsonibacter*	*Lawsonibacter sp900066645*	98.63	3.00	55.40	7
OA-10	*Oscillospiraceae; Lawsonibacter*	*Lawsonibacter asaccharolyticus*	98.77	3.13	60.26	6
OA-3	*Lachnospiraceae; Ruminococcus_G*	*Ruminococcus sp.*	88.32	4.06	46.42	0
OA-1	*Lachnospiraceae; Muricomes*	*Muricomes sp000509105*	99.61	5.40	45.96	4
OA-13	*Lachnospiraceae; Enterocloster*	*Enterocloster sp000155435*	98.75	6.95	49.85	8
OA-11	*Lachnospiraceae; Enterocloster*	*Enterocloster sp005845215*	98.99	6.48	49.57	7

Figure 1. Isolation and identification of spore-forming Firmicutes from helminth-colonized individuals

(A) Representative images of OA isolates 48 h after streaking on brucella blood agar.

(B) Taxonomic identities and genomic features of OA isolates based on full genome sequences. “Closest species” refers to the species with the highest average nucleotide identity (ANI). Red: denotes new species with no matches ≥95% ANI. The phylogenetic tree to the left of the table depicts the relationships between OA isolates.

(C) Phylogeny of OA isolates and all Firmicutes assemblies from NCBI GenBank with an ANI ≥95% to at least one OA isolate.

(D and E) UMAP plots of all gene clusters from *Clostridium_AQ innocuum* (D) and *Clostridium perfringens* (E) according to orthogroup presence or absence. OA isolates corresponding to each species are highlighted in red.

(F) UMAP plot of KEGG orthologs involved in metabolism for OA14 (in red) and all related *C. perfringens* with an ANI ≥95%.

(G) Number of orthogroups in the core genome (red) and pangenome (teal) of *C. perfringens* as a function of number of genomes analyzed. Each dot represents the core or pangenome size N at a given number of genomes k, which was computed with up to 100 different combinations of genomes. Heap's law (black lines) was fit to the median core genome ($N = 607{,}029\ k^{-0.31}$) and pan-genome ($N = 429{,}650\ k^{0.06}$) sizes.

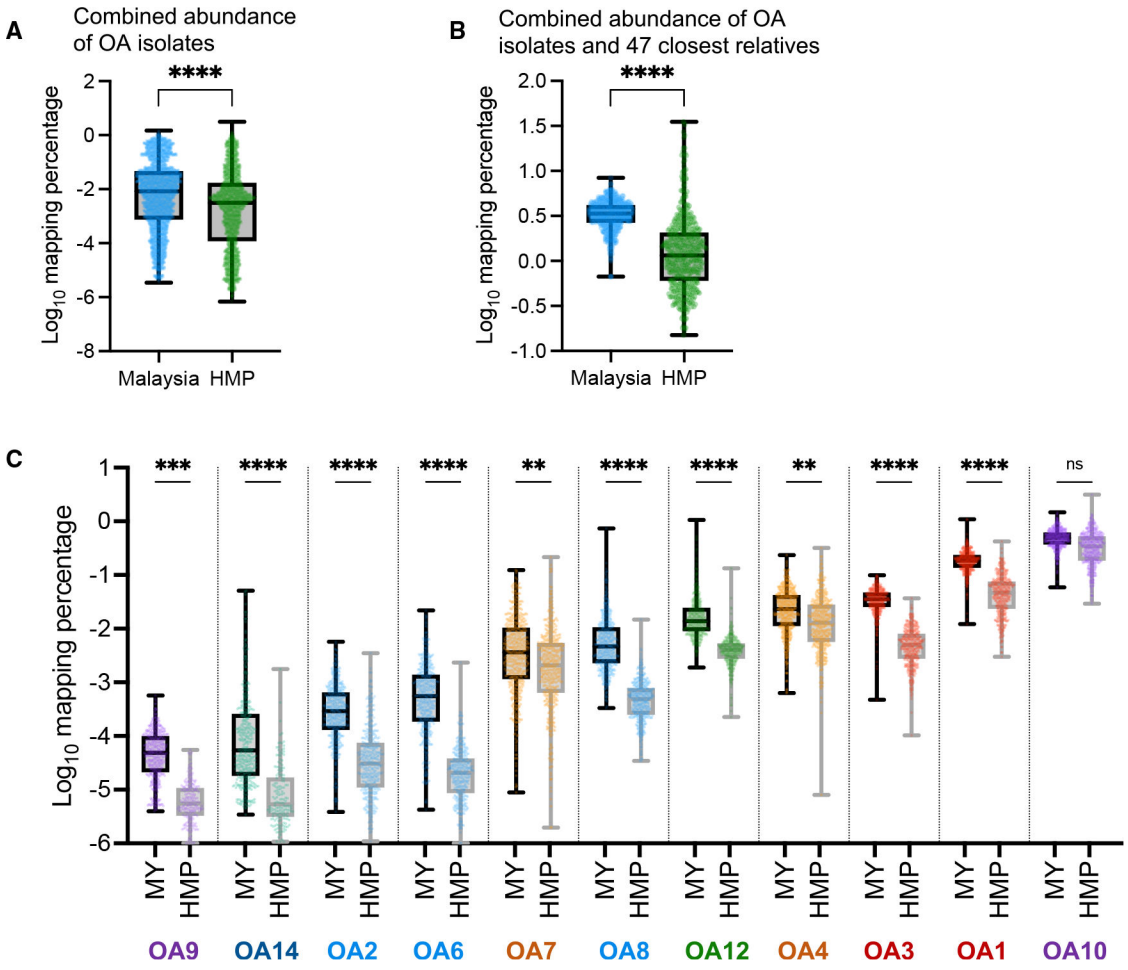


Figure 2. OA genomes are enriched in Malaysian microbiomes

(A) Mapping percentage of Malaysian and HMP metagenomes to genomes of the 14 OA isolates.

(B) Mapping percentage of Malaysian and HMP metagenomes to genomes of the 14 OA isolates and their 47 closest relatives (see Table S4).

(C) Mapping percentage of Malaysian and HMP metagenomes to individual OA isolate genomes. OA5, OA11, and OA13 were not detected by krakenuniq due to low representation.

(A–C) Each dot represents one metagenome from one individual. Whiskers represent minimum (Min) and maximum (Max). Mapping percentage denotes the percentage of k-mers from each metagenome that specifically map to one of the genomes in the group being measured. Mann-Whitney test in (A) and (B), ****$p < 0.0001$. Kruskal-Wallis with Dunn's multiple comparisons test in (C). **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

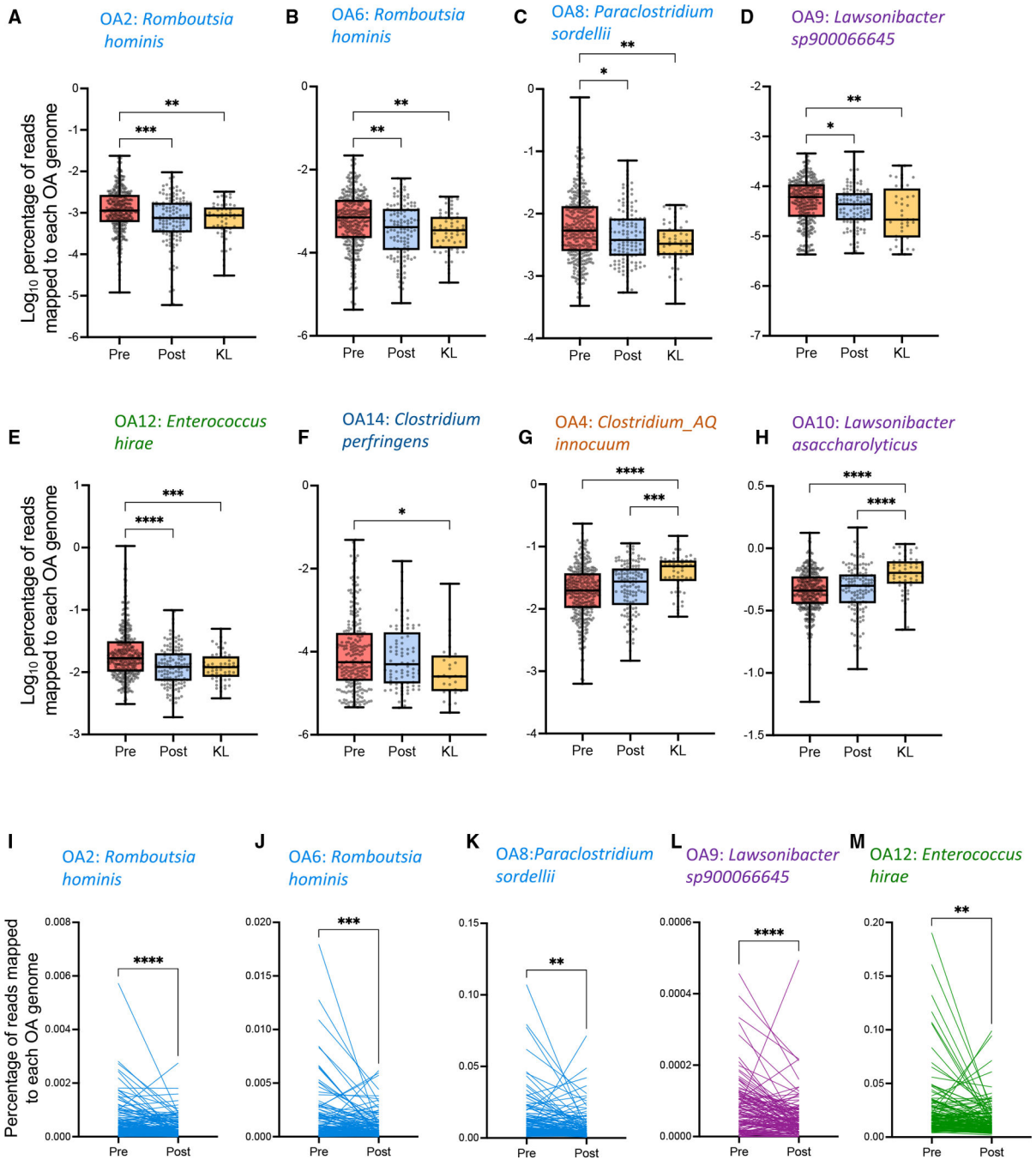


Figure 3. Identification of OA isolates associated with helminth colonization

(A–H) Mapping percentages of pre-albendazole (ABZ), post-ABZ, and urban control Kuala Lumpur (KL) metagenomes to indicated OA isolate genomes. Each dot represents one metagenome from one individual. Whiskers represent Min and Max.

(I–M) Pairwise analysis of the mapping percentage of metagenomes derived from matched pre- and post-ABZ longitudinal sampling to indicated OA isolate genomes.

Each line represents paired metagenomes from one individual. Kruskal-Wallis with Dunn's multiple comparisons test was used in (A)–(H). Paired t test was used in (I)–(M). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

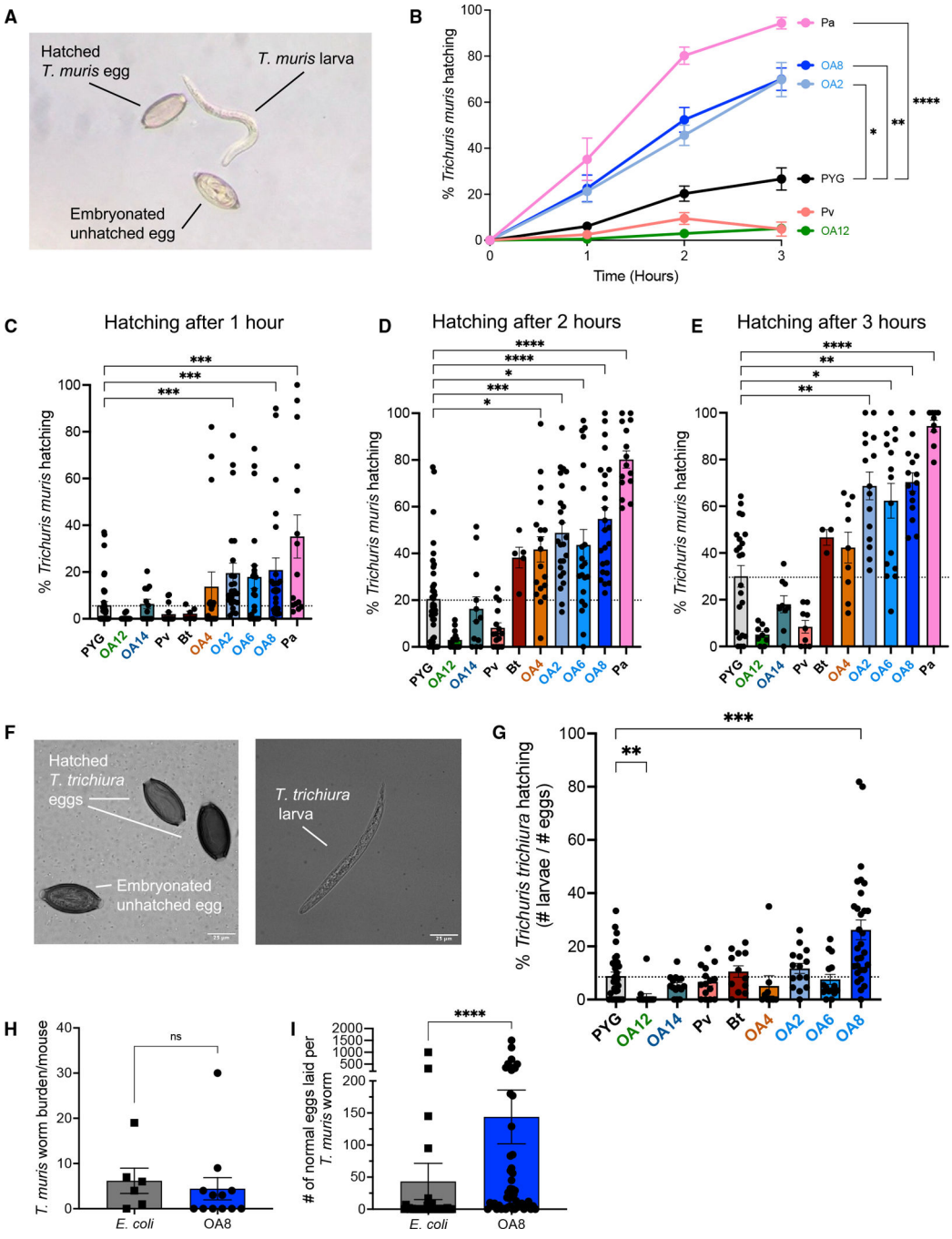


Figure 4. Helminth-associated Clostridia promote the Trichuris life cycle

(A) Representative light microscopy image of a *T. muris* larva and hatched and unhatched eggs.

(B) Time course of *T. muris* egg hatching in the presence of indicated bacteria. Data are combined from 11 independent experiments, with each condition tested in at least 3 separate experiments, and plotted as mean ± SEM.

(C–E) Percentage of hatched *T. muris* eggs after 1 (C), 2 (D), and 3 h (E) of incubation with indicated bacteria corresponding to the time series graph in (B) to allow visualization of all

bacterial strains at each time point. Experiments were performed in at least 3 independent repeats for each strain, and each dot represents one well of eggs. Bars indicate mean ± SEM.
(F) Representative confocal images of a *T. trichiura* larva and hatched and unhatched eggs. Scale bar represents 25μm.
(G) Percentage of hatched *T. trichiura* eggs after 6 days of incubation with indicated bacteria. Experiments were performed in at least 3 independent repeats for each strain, and each dot represents one well of eggs. Bars indicate mean ± SEM.
(H) Adult *T. muris* worms harvested from mice monocolonized with *E. coli* or OA8 (*P. sordellii*). Each dot represents one mouse, and bars indicate mean ± SEM.
(I) Number of eggs laid with normal morphology per worms harvested from mice monocolonized with *E. coli* or OA8 (*P. sordellii*). Each dot represents eggs laid by one individual worm, and bars indicate mean ± SEM.
PYG, Peptone Yeast Glucose media; Pa, *Pseudomonas aeruginosa;* Pv, *Phocaeicola vulgatus*; Bt, *Bacteroides thetaiotaomicron*. Welch's t test was used to compare between area under curve of each condition for (B). Kruskal-Wallis with Dunn's multiple comparisons test was used to compare between each OA isolate and PYG in (C)–(E) and (G). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. Mann-Whitney test in (H) and (I). ****$p < 0.0001$.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
OA isolates OA1-OA14	This study	N/A
Phocaeicola (previously *Bacteroides*) *vulgatus*	Ramanan et al.[62]	N/A
Bacteroides thetaiotaomicron VPI-5482	ATCC	Cat# 29148
Pseudomonas aeruginosa strain PAO1	Srivastava et al.[63]	N/A
Escherichia coli strain BW25113	Baba et al.[64]	N/A
Biological samples		
Fecal samples from Orang Asli villagers from Kuala Pangsun village	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Chloroform	Fisher Scientific	Cat# C298-500
Peptone Yeast Broth with Glucose - PYG	Anaerobe Systems	Cat# AS-822
Brucella Blood Agar - BRU	Anaerobe Systems	Cat# AS-141
BD BBL Prepared Plated Media: Bacteroides Bile Esculin Agar - BBE	Fisher Scientific	Cat# L21836
Tryptic Soy Broth - TSB	NYU Reagent Preparation	N/A
Tryptic Soy Agar - TSA	NYU Reagent Preparation	N/A
Luria Bertani (LB) Broth	NYU Reagent Preparation	N/A
Luria Bertani (LB) Agar	NYU Reagent Preparation	N/A
Genomic DNA Buffer Set	QIAGEN	Cat# 19060
Proteinase K	QIAGEN	Cat# 19131
Ready-Lyse Lysozyme Solution	Epicentre/Lucigen	Cat# R1810M
RNase A, DNase and protease free	Thermo Scientific	Cat# EN0531
AMPure Beads	Pacific Biosciences	Cat# 100-265-900
Elution Buffer	Pacific Biosciences	Cat# 101-633-500
Albendazole	Sigma-Aldrich	Cat# A4673
RPMI 1640 Medium with GlutaMAX	Gibco	Cat# 61870036
Penicillin Streptomycin Solution, 100X	Corning	Cat# 30-002-CI
Critical commercial assays		
TOPO TA Cloning Kit, with pCR2.1-TOPO, One Shot TOP10 Chemically Competent *E. coli,* and PureLink Quick Plasmid Miniprep Kit	ThermoFisher Scientific	Cat# K450002
QIAGEN Genomic-tip 100/G	QIAGEN	Cat# 10243
Qubit™ dsDNA HS and BR Assay Kits	Invitrogen	Cat# Q32851
Barcoded Overhang Adapter Kit 8A	Pacific Biosciences	Cat# 101-628-400
SMRTbell Express Template Prep Kit 2.0	Pacific Biosciences	Cat# 101-938-900
Sequel DNA Polymerase Binding Kit 3.0	Pacific Biosciences	Cat# 100-500-400
SMRT Cell 1M v3 LR Tray	Pacific Biosciences	Cat# 101-531-001

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Sequel Sequencing Kit 3.0	Pacific Biosciences	Cat# 101-597-900
Deposited data		
OA isolate genomes	This paper	NCBI BioProject: PRJNA800461
Experimental models: Organisms/strains		
Germfree C57BL/6J mice	NYU gnotobiotic facility	N/A
Trichuris muris E strain	Ramanan et al.[14]	N/A
Trichuris trichiura eggs	The *Trichuris trichiura* egg Production Unit at the Clinical Immunology Laboratory at George Washington University	N/A
Oligonucleotides		
16S-Fwd: CCGATATCTCTAGAAG AGTTTGATCCTGGCTCAG	Brugiroux et al.[65]	N/A
16S-Rev: CCGATATCGGATCCACGGTTAC CTTGTTACGACTT	Brugiroux et al.[65]	N/A
Software and algorithms		
EZBioCloud	Yoon et al.[66]	https://www.ezbiocloud.net/
SMRT Link	Pacific Biosciences	https://www.pacb.com/support/software-downloads/
Flye v.2.8.1	Kolmogorov et al.[67]	https://github.com/fenderglass/Flye
fastANI v.1.32	Jain et al.[68]	https://github.com/ParBLiSS/FastANI
GTDB-Tk v.1.5.0	Chaumeil et al.[29]	https://github.com/Ecogenomics/GTDBTk
fasttree v.1.11.2	Price et al.[69]	http://microbesonline.org/fasttree
iTOL v5	Letunic and Bork,[70]	https://itol.embl.de
dfast v.1.2.13	Tanizawa et al.[71]	https://dfast.nig.ac.jp/
OrthoFinder v.2.5.4	Emms and Kelly,[72]	https://github.com/davidemms/OrthoFinder
KofamScan v.1.3.0	Aramaki et al.,[73]	https://www.genome.jp/tools/kofamkoala/
UMAP v.0.2.7.0	McInnes et al.[34]	https://arxiv.org/abs/1802.03426 and https://github.com/tkonopka/umap
KEGG	Kanehisa and Goto,[74]	https://www.genome.jp/kegg/
R programming language v.4.1.1	R Core Team,[75]	https://www.r-project.org/
ggplot2 v.3.3.5	Wickham,[76]	https://ggplot2.tidyverse.org
KneadData v.0.7.4	Huttenhower lab	http://huttenhower.sph.harvard.edu/kneaddata
Bowtie2	Langmead and Salzberg,[77]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
FastQC	Andrews,[78]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
KrakenUniq	Breitwieser et al.[38]	https://github.com/fbreitwieser/krakenuniq
NIS-Elements	Nikom	https://www.microscope.healthcare.nikon.com/products/software
GraphPad Prism	GraphPad Software	http://www.graphpad.com/scientific-software/prism/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Malaysian metagenomes	Tee et al.[37]	NCBI BioProject: PRJNA797994
HMP metagenomes	Human Microbiome Project Consortium,[39,40]	https://www.hmpdacc.org/hmp/