Endospores and other lysis-resistant bacteria comprise a widely shared core community within the human microbiota

Sean M. Kearney^{1,2,3} · Sean M. Gibbons^{1,2,3} · Mathilde Poyet^{1,2,3} · Thomas Gurry^{1,2,3} · Kevin Bullock² · Jessica R. Allegretti^{4,5} · Clary B. Clish² · Eric J. Alm^{1,2,3}

Received: 5 December 2017 / Revised: 20 March 2018 / Accepted: 27 March 2018 / Published online: 13 June 2018
© International Society for Microbial Ecology 2018

Abstract

Endospore-formers in the human microbiota are well adapted for host-to-host transmission, and an emerging consensus points to their role in determining health and disease states in the gut. The human gut, more than any other environment, encourages the maintenance of endospore formation, with recent culture-based work suggesting that over 50% of genera in the microbiome carry genes attributed to this trait. However, there has been limited work on the ecological role of endospores and other stress-resistant cellular states in the human gut. In fact, there is no data to indicate whether organisms with the genetic potential to form endospores actually form endospores in situ and how sporulation varies across individuals and over time. Here we applied a culture-independent protocol to enrich for endospores and other stress-resistant cells in human feces to identify variation in these states across people and within an individual over time. We see that cells with resistant states are more likely than those without to be shared among multiple individuals, which suggests that these resistant states are particularly adapted for cross-host dissemination. Furthermore, we use untargeted fecal metabolomics in 24 individuals and within a person over time to show that these organisms respond to shared environmental signals, and in particular, dietary fatty acids, that likely mediate colonization of recently disturbed human guts.

Introduction

To date, there is limited work investigating the relevance of stress-resistant cellular states in the propagation, survival, and function of organisms in the mammalian gastrointestinal (GI) tract. The gut is the only known environment with such a considerable abundance of organisms that form

endospores, considered the most stress-resistant of all cell-types [1]. Anaerobic, endospore-forming Firmicutes are numerically dominant members of the GI tract of most animal species [2, 3]. Within this group of organisms, the presence of genes for endospore formation suggests that growth in the GI tract favors the maintenance of this large gene repertoire [3]. The apparent utility of these genes is to allow organisms to enter metabolically dormant states that aid in survival and transmission to new hosts. Passage through the GI tract is likely to trigger sporulation [4, 5], but the mechanisms by which this process occurs and the signals that induce sporulation here are mostly unknown, even for well-studied pathogens like *Clostridioides difficile*.

Many endospore-forming organisms in the human gut are in the class Clostridia, the most well-studied of which are the pathogens *C. difficile* and *Clostridium perfringens* [6, 7]. However, Clostridia also includes abundant organisms not known to form endospores, like *Faecalibacterium prausnitzii* [8] and *Roseburia intestinalis* [9]. For *C. difficile*, the role of sporulation is central to disease etiology [10], particularly in patients who experience recurrence. Sporulation and rising levels of antibiotic resistance allow *C. difficile* to persist in the face of antibiotic assault,

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41396-018-0192-z>) contains supplementary material, which is available to authorized users.

✉ Eric J. Alm
ejalm@mit.edu

¹ Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

² The Broad Institute, Cambridge, MA, USA

³ The Center for Microbiome Informatics and Therapeutics, Cambridge, MA, USA

⁴ Division of Gastroenterology, Brigham and Women's Hospital, Boston, MA, USA

⁵ Harvard Medical School, Boston, MA, USA

Fig. 1 Resistant fraction sequencing of human fecal bacteria. **a** Overview of resistant cell enrichment and 16S rDNA sequencing protocol. Resistant fraction samples are treated with a series of physical, enzymatic, and chemical lysis steps to deplete vegetative cells. DNA from bulk community and resistant fraction samples are extracted via a mechanical lysis protocol, and 16S rDNA libraries prepared. Communities are analyzed to determine the change in abundance of each OTU in the resistant fraction relative to the bulk community. (right) Phase contrast images of bulk community and resistant fraction—phase-bright cells are endospores. Endospores stain green when heat fixed with malachite green, vegetative cells appear red from safranin counter stain. **b** Representative results of 16S rDNA

profile for bulk community and endospore-enriched samples. Reads from each OTU are summed across 24 individuals to give a meta-bulk and meta-endospore community. Phylogenetic classes in black text increase with resistant fraction; gray text classes decrease with resistant fraction. **c** Distribution of resistant fraction proportion across phyla aggregated across individuals filtered to remove OTUs with single counts in a sample (for visualization purposes). Colors represent phyla. OTUs with a resistant fraction proportion of 0 are absent from the resistant fraction; OTUs with a resistant fraction proportion of 1 are absent from the bulk community and only found in the resistant fraction

ensuring that it remains in the environment to rapidly re-colonize its host.

Among Clostridia that do not cause disease, multiple strains of endospore-forming organisms have the capacity to induce T regulatory cells and associated anti-inflammatory cytokines [11, 12] involved in sensitivity to, for example, peanut antigen [13]. These organisms have recently been shown to provide pathogen resistance in neonatal mice [14]. Similarly, endospore-forming commensals of the murine GI

tract have a central role in mediating the induction of a Th17-type T helper cell response [15–17]. Many Clostridia also produce butyrate as an end-product of metabolism, which can regulate how immune cells interact with gut commensal bacteria [18–22]. This group of organisms also boosts production of serotonin by enterochromaffin cells in the intestine, crucial for motility in the gut [23]. An ecological understanding of sporulation and induction of other resistant states could be informative for how these phenotypes interact with

host immunity. For example, such an understanding could inform whether inflammation acts positively or negatively on endospore formation, or whether endospores themselves have immunomodulatory effects.

Resistant cellular states like endospores appear to be adaptive in the mammalian gut environment. It is likely that other non-endospore-forming taxa have evolved analogous resistance strategies for passing between hosts. Persister states may allow non-endospore-forming organisms to enter a metabolically inert state upon exit from the gastrointestinal tract. Toxin-antitoxin systems, associated with persistence in *E. coli*, are well-represented in Bacteroidetes, Alpha- and Gammaproteobacteria [24], and Bacteroidetes are among the most metabolically inactive cells in human fecal samples [25]. Further, viable but non-culturable cell (VBNC) states may enable passage through the environment by reducing metabolic needs and affecting cell wall and membrane composition and morphology [26]. These strategies and others not yet studied may play an important role in mediating cross-host bacterial transmission.

Environmental stress resistance protects cells faced with unfavorable conditions. The signals triggering resistance are likely quite varied. Even for well-studied endospore-forming bacteria, inducing sporulation in vitro can be difficult, and across strains of one species, signals that induce sporulation in one strain may be insufficient to induce sporulation in others [27]. Further, even organisms that abundantly form endospores in their native environment may not do so under conditions permitting vegetative growth. For instance, *Paenibacillus larvae* in honeybees will only form endospores in vitro under idiosyncratic conditions designed to mimic the host environment [28]. Similarly, certain strains of *Clostridium perfringens* rarely form endospores in vitro unless exposed to a specific set of environmental stressors [27]. The discrepancy in phenotype of organisms in their native environments compared to in vitro argues for culture-free approaches to investigate such phenotypes in situ. Enriching for stress-resistant cells in environmental samples provides a means to uncover the actual context in which these states form.

Here we investigate which organisms are present as endospores or as other resistant cell types in the human gastrointestinal tract. We modified previously described methods to enrich fecal samples for endospores and obtain paired bulk community and resistant fraction 16S rDNA sequence data for 24 healthy individuals and one individual across 24 days. We consistently enriched for putatively endospore-forming taxa in all samples, as well as other taxa, predominantly from the Actinobacteria phylum, that show high levels of lysis resistance. We compared resistant OTUs (rOTUs) and non-resistant OTUs (nOTUs) to identify ecological characteristics differing between these groups. Using a database of rOTUs identified in this study, we find

consistent signals for these organisms in their responses to a variety of successional states across multiple independent datasets from prior published studies (Supplementary Table 7). Overall, we show a tight association between the ecological role of these resistant organisms and their distribution within and across human hosts.

Results and discussion

Sequencing resistant fraction reveals resistant taxa present in human feces

We modified a culture-independent method [29] to generate resistant fraction 16S rDNA amplicon data from human feces (Fig. 1a). This method entails a series of lysis treatments, including heating lysozyme, and proteinase treatment, alkaline and SDS treatment, and hypotonic wash steps followed by DNase treatment. We extract the resultant resistant fraction pellet alongside the bulk community sample using a bead-beating protocol (Methods section). We validated this method on pure bacterial cultures and endospore cultures prior to treatment of human fecal samples (Supplementary Fig. 2) and conducted a small study to validate the reproducibility of the method across samples (Supplementary Table 5). We then proceeded to test the method on feces from a healthy human cohort and from a time series of a single individual.

When aggregating the fecal data across our cohort, we see expansion of classes with known endospore-formers in the resistant fraction: Clostridia, Erysipelotrichia, and Bacilli (Fig. 1b). We also see depletion of classes lacking endospore-formers (Bacteroidia, Betaproteobacteria, Verrucomicrobia, Gammaproteobacteria). Organisms in the class Actinobacteria were enriched in the resistant fraction, but lack genes considered essential for endospore formation. Although exospore formation is well documented in some families of Actinobacteria (e.g., Actinomycetaceae and Streptomycetaceae), these families have only modest representation in our data. We see high-level resistance primarily from *Bifidobacterium* and *Collinsella*, whose representative genomes lack orthologs for genes thought to be essential to exospore formation.

We suspect that high-level resistance in the Actinobacteria is mediated primarily by resistance to lysozyme conferred by cell wall structures common to Actinobacteria [30]. Lysozyme is one of the most common and important defense mechanisms used by neutrophils, monocytes, macrophages, and epithelial cells [31, 32]. It is abundant in human milk, a source of *Bifidobacterium* species transferred to breast-feeding infants, and in saliva and mucus, where it serves an antibacterial role [33]. Attempts to deplete Actinobacteria with achromopeptidase, which has previously

Fig. 2 Resistant fraction OTUs are more shared across individuals than bulk community OTUs. **a** Alpha-diversity metrics measured for the bulk community (x-axis) and resistant fraction enrichments (y-axis). *P*-values are for the test of differences between alpha-diversity metric distributions using paired Wilcoxon rank-sum test. **b** Distribution of Jaccard distance between resistant fraction and bulk communities, within the resistant fraction, and within the bulk communities. **c**

Multidimensional scaling on the Jensen-Shannon divergence of all resistant fraction and bulk community samples. Black dots represent resistant fraction communities, gray dots represent bulk community samples. **d** Comparison of the number of rOTU sequence variants in only one person and in multiple people to nOTU sequence variants in only one person and in multiple people

been shown to break down Actinobacterial peptidoglycan, had variable efficacy across samples (data not shown). Thus, factors other than cell wall structure may contribute to Actinobacteria resistance.

To quantify the extent of lysis resistance, we calculated the proportion of normalized reads for each sequence variant in the resistant fraction to the sum of its reads in the bulk community and the resistant fraction. We then obtain a finite quantity even for organisms not observed in one of the paired samples. When the proportion exceeds 0.5, we call an OTU enriched in the resistant fraction (Fig. 1a). An OTU that is enriched in at least one of the samples in which it is present is considered a resistant OTU (rOTU), and non-resistant (nOTU) otherwise. Using the above definitions, all of the rOTUs are either Firmicutes or Actinobacteria (Fig. 1c). In fact, when grouping OTUs at the genus level, the top two most enriched genera (*Bifidobacterium*, *Colinsella*) are both Actinobacteria.

Resistant fractions consist of few dominant and many rare OTUs

In order to investigate ecological properties of the resistant cell fraction, we first examined the community

structure of resistant fractions and compared these to their bulk community counterparts. After rarefying to the minimum read depth (28,639 reads, Supplementary Table 6), we find that resistant fractions are significantly less diverse both in species richness and evenness than their bulk community counterparts (Fig. 2a). As we are sampling a subset of the community, this result is not necessarily surprising. However, reduced evenness of the resistant fraction compared to the bulk community suggests dominance of a few organisms coupled with many low abundance organisms. This difference in community structure could entail that resistant fractions are more dissimilar from each other than their bulk community counterparts. Instead, the compositions (estimated by Jaccard Distance) of resistant fractions tend to be more similar to each other than bulk communities to each other across different people (PERMANOVA *p*-value <0.001, Fig. 2b). Similarly, using Jensen-Shannon divergence, resistant fractions cluster together separately from bulk communities implying that resistant fractions may assemble in similar ways across people (Fig. 2c). We hypothesized that the coherence between resistant fractions may lead to increased prevalence of organisms found in the resistant fraction across people.

rOTUs are more shared than nOTUs among individuals

To test the hypothesis that resistant states contribute to prevalence, we examined the frequency with which rOTUs were found among the bulk communities across individuals compared to nOTUs (Fig. 2d). First, nOTUs, which are never enriched in the resistant fraction, are significantly less likely to be shared among multiple individuals than rOTUs (Mann–Whitney *U*-test comparing the distribution of the number of individuals sharing each rOTU to the number of individuals sharing each nOTU, p -value = $1e-36$). We again see this result by calculating the correlation between the frequency of resistance (the number of times an organism is enriched in the resistant fraction divided by the number of times it is observed) and sharedness (number of individuals an OTU is observed in divided by the total number of individuals), giving a weak, but positive and highly significant correlation (Spearman rho, correlation = 0.23, p -value = $1e-17$; Kendall tau, correlation = 0.19, p -value = $1e-17$). Finally, when we compare the proportion of rOTUs found in only one person compared to multiple people with that proportion for nOTUs in bulk communities, we find rOTUs are about four times as likely to be found in multiple individuals (Fisher exact test, p -value = $1e-33$, odds ratio = 4.0).

These results suggest that organisms that do not form resistant states are less likely to be found across multiple individuals than those that do. Yet, rOTUs tend to be less dominant members of the community (median rOTU abundance = 13.5 counts, median nOTU abundance = 18 counts, Mann–Whitney *U*-test, p -value = 0.004). Even though rOTUs are not as dominant as nOTUs, they are more widespread within this cohort.

Representation of organisms in resistant fractions is heterogeneous across and within individuals

We suspected that the increased prevalence of rOTUs might indicate positive selection on resistance capabilities in this environment. Variation in this trait among related organisms could be indicative of selection. In order to visualize how much of a population is present in a resistant state within a given sample, we scaled 16S rDNA abundance data using V4 16S rDNA qPCR-based estimates of community size (Supplementary Fig. 1, Supplementary Table 6, and Supplementary Text) and defined the resistant fraction as the ratio of these scaled reads for each OTU. We plot this quantity on a phylogeny representing 99% OTUs (clustered at 99% nucleotide ID using usearch [34]) present in at least 8 individuals and up to 24 individuals (Fig. 3). First, we note the high variability in the resistant fraction within and

Fig. 3 Taxa show heterogeneous patterns of resistant cell fractions across individuals. Phylogenetic placement of the fraction of resistant organisms for taxa present within at least eight individuals estimated by the ratio of counts scaled by qPCR-estimates of biomass in the resistant fractions and bulk communities. Tree branch colors represent the degree to which a taxonomic group was enriched (higher relative

abundance) in the resistant fraction with pink branches never enriched and blue and green branches enriched at least once across samples. Classes within each phylum are shown with a colored bar. Arrows indicate OTUs showing the maximum (black) and minimum (gray) within-OTU variability in enrichment scores. (Color figure online)

Fig. 4 Common signals govern resistant state exit and growth in the GI tract. **a** Boxplot of the distribution of correlation distances (pairwise Euclidean distance between the Spearman correlation vectors for two OTUs) between rOTUs and nOTUs, within rOTUs (enriched), and

across taxa (the average variation is over 50-fold within each taxon). For one *Roseburia* 99% OTU in particular, this quantity varies over 3 orders of magnitude, suggesting this OTU contains organisms present in a resistant state in some individuals, but not in others.

Furthermore, within a person, OTUs with the same genus classification can be discordant in their degree of resistance. In the individual time series, for example, one *Ruminococcus* 100% OTU is almost always enriched, and another is never enriched (Supplementary Fig. 3). The closest matching genomes to these two organisms show differences in sporulation gene content, with the resistant *Ruminococcus* sharing 48/58 core sporulation genes [35], and the non-resistant only 41/58 (Supplementary Fig. 4 and Supplementary Tables 1 and 4). We also see that spore maturation proteins *spmA* and *spmB* vary in their presence in genomes of genera with variable enrichment phenotypes. These genes are involved in spore cortex dehydration and heat resistance in *B. subtilis* and *C. perfringens*, so their loss might contribute to differences in the recovery of resistant cells in this work.

The process of entering a resistant state itself might be selected on in this system. There is evidence that the sporulation phenotype is evolving in mammalian guts, as several gut isolates of *B. subtilis* lack genes that negatively regulate sporulation compared to their laboratory counterparts [36]. Knowing which organisms can form resistant cells in a community does not provide complete information about which organisms do (Supplementary results). Formation of resistant states in vivo seems to be highly context dependent. We also note that loss of a single gene (i.e., *spo0A*) in *C. difficile* is sufficient for loss-of-sporulation,

within nOTUs (unenriched). **b** Abundance of OTUs in the resistant fraction as a function of bile acid exposure for nine phylogenetically distant OTUs

such that retaining endospore formation requires strong purifying selection.

rOTUs share signals for growth within an individual

Previous evidence has shown that bile acids contribute to outgrowth of *C. difficile* endospores in vivo [37]. As taurocholate is a known germinant for several endospore-forming species [3], we wondered whether endospore-formers and resistant organisms more broadly would share dynamic behavior over time, suggesting coherent responses to environmental signals. We compared the Euclidean distance of correlation profiles between organisms to determine whether there were differences between the correlation profiles of rOTUs and nOTUs (Fig. 4a). We find that rOTUs are more similar in their correlation profiles than nOTUs (PERMANOVA, $p < 0.001$). The average correlation between rOTUs in the time series to each other is 0.211 compared to 0.162 for nOTUs to each other (Wilcoxon rank-sum test, p -value = $5e-08$): strong correlated behavior in this person associated primarily with rOTUs (Fig. 4a). We interpret this result to mean that the dynamic behavior of rOTUs is more strongly coupled: these OTUs respond coherently to environmental signals.

rOTUs link growth to fatty acid metabolism

To address whether bile-related signals relate to the dynamics of rOTUs in the time series, we conducted untargeted metabolomics with standards for fatty acid

Fig. 5 Resistant OTUs show disproportionate turnover in diverse contexts. **a** Overview of approach for identifying resistant-cell forming OTUs in 16S rDNA sequencing datasets. rOTU database sequences are matched to sequences in other datasets, and then patterns within those datasets among the identified rOTUs are determined. **b** Fraction of rOTUs present during microbial colonization of an infant gut annotated with major diet and health perturbations. rOTUs encompass both putative endospore-forming organisms and those not known to form endospores, but which possess a resistant state (Actinobacteria

and non-endospore-forming Firmicutes). **c** Fraction of rOTUs present as a function of *C. difficile* infection status (fCDI = first time *C. difficile* diagnosis, rCDI = at least 3 episodes of *C. difficile* infection following initial treatment). **d** Fraction of rOTUs and all other OTUs (non-resistant OTUs) transferred from donors to recipients by fecal microbiota transplant. **e** Time series of rOTUs (top) and all other (non-resistant) OTUs (bottom) from a human male infected with *Salmonella*, with OTUs significantly more abundant pre-infection (dark gray) and significantly more abundant post-infection (light gray)

metabolism. We then calculated the Spearman correlations between the median abundance profile of all OTUs clustering with the highly correlated rOTUs and metabolites for which we had standard markers. This cluster tends to correlate positively with long-chain saturated fatty acids, and negatively with long-chain polyunsaturated fatty acids and,

notably, taurocholate (Supplementary Table 2). An OTU in the genus *Bilophila*, known to use taurocholate for sulfite reduction [38] also clusters with these organisms, and shows a strong relationship to markers of milkfat consumption (Supplementary Text). We suspect that taurocholate metabolism by members of this group drives down

the concentration of taurocholate in stool. Additionally, saturated fatty acid concentration in the stool measures fatty acids escaping absorption in the small intestine. This process would be negatively impacted by microbial metabolism of taurocholate, as it thought to more efficiently emulsify saturated fats than glycine-conjugated primary bile acids [38]. Fecal concentrations of taurocholate reflect secretion of unmetabolized taurocholate, which should increase if taurocholate metabolism by the gut microbiota decreases.

Resistant cells lose resistance in response to physiological bile acid concentrations

As a more direct test of the coupling of rOTUs to bile acid concentration, we dosed ethanol-treated feces (to kill vegetative cells without the additional harshness of the resistant fraction DNA enrichment protocol) with increasing concentrations of bovine bile in aqueous solution. We then measured the depletion of OTUs from the endospore-enrichment using 16S rDNA sequencing (Fig. 4b). When correcting for biomass via qPCR, nearly 20% of OTUs identified in the resistant fraction apparently germinated in response to bile acids (log-link quasi-poisson generalized linear model, p -value < 0.05 , Supplementary Table 3). The true fraction of resistant cells that lose resistance in response to bile acids is likely higher, as many endospores require an activation step (i.e., heating at 80 °C or treatment with lysozyme as for *C. difficile* [39]) before they will respond to germinants.

Notably, most ethanol-resistant OTUs began to show a germination-like response at 0.5% bile (Fig. 4b), which is near the concentrations found in the human small intestine [40]. Although Clostridia and other putative endospore-formers make up the majority of organisms that lose resistance in response to bile acids, genera in the Actinobacteria and other resistant cells also show this response when approaching physiological concentrations. These conserved responses suggest that the same cues can mediate loss-of-resistance in distantly related organisms, similar to the conserved resuscitation response of dormant bacteria to peptidoglycan [41].

rOTUs exhibit shared dynamics in diverse contexts

Correlated behavior, increased prevalence, and shared signals for growth among rOTUs indicated that these organisms might exhibit a global response during disturbances of various kinds. To test this hypothesis, we made a sequence database of rOTUs within our cohort, and used this database to identify putative rOTUs in other datasets (Fig. 5a). While certainly not all rOTUs in our dataset will map to organisms forming resistant states in other datasets, we assume that some strains or species within an rOTU are capable of forming a resistant state at some time.

We expected that increased prevalence and shared signals for growth would lead to enhanced colonization of the developing infant gut microbiota [42]. The lysozyme-resistant members of the Actinobacteria and Bacillales dominate the infant gut microbiota for most of the first 80 days of life and do not equilibrate until the infant starts a full adult diet (Fig. 5b). Early colonization by these rOTUs connects a resistant state to development of the infant gut microbiome. Here, lysozyme resistance might be essential for semi-selective transmission of *Bifidobacterium*, as human breast milk is rich in lysozyme [43], potentially lysing non-resistant cells. Others have shown endospore-formers negatively associate with vertical transmission from mother to infant [44], but other environmentally resistant states as in the Actinobacteria may be important for vertical transmission as well.

Depletion of endospore-forming clades is common during infection with *C. difficile*. We predicted a strong signal for rOTUs in individuals infected with *C. difficile*, due to its sporulation requirement for transmission [10]. We find a significant depletion of rOTUs dependent on *C. difficile* infection status (Fig. 5c), with a serial depletion of rOTUs from healthy to first time diagnosis to recurrent patients [45]. Because of this depletion in rOTUs, we expected that fecal microbiota transplant (FMT) might transfer relatively more rOTUs than other OTUs [46]. Indeed, among OTUs shared with donors, 90% of rOTUs increase in abundance following FMT, compared to 77% for the rest of the community (Fisher exact test, p -value = 0.008) (Fig. 5d).

We suspected that rOTUs are a particularly malleable component of the microbiota. To test this hypothesis, we measured the turnover of rOTUs in the time series of an otherwise healthy male individual who was infected by Salmonella [47]. New rOTUs almost completely replaced old rOTUs following this perturbation. By contrast, fewer OTUs from the rest of the community were lost and gained. This result holds both when examining the number of OTUs replaced (Fisher exact test, p -value = $6e-12$), as well as the change in abundance of these OTUs (Fig. 5e). We see again that rOTUs exhibit coherent responses to changes in the gut environment, most pronounced in systems with dramatic perturbations. Colonization of newly vacant niches favors rOTUs, likely transmitted in an endospore or other resistant state to germinate in an environment replete with nutrients (including untransformed bile acids). In the absence of a fully functioning microbiota, rOTUs appear to fill open niches more readily than nOTUs.

Conclusion

Gut bacteria in the resistant fraction were more shared across individuals and showed more correlated dynamics

compared to non-resistant organisms. Resistant taxa show greater turnover following large-scale disturbance events, as in the case of *C. difficile* and *Salmonella* infection, which suggests that many of these organisms are sensitive to environmental fluctuations and respond to stress by entering into a dormant, seed-like state. Environmental sensitivity and high-turnover rates of resistant taxa provide an opportunity to manipulate the composition of the human gut microbiota through targeted perturbations and replacements. Because of the therapeutic relevance of Clostridia endospores [11–14, 23], determining the exact conditions that permit their replacement may be of high value for future microbiota-based therapeutics. Here we found that the growth of many resistant organisms was associated with dietary fatty acids. If this result extends to more individuals, one can imagine a therapeutic strategy coupling dietary changes with introduced resistant cells to enable robust colonization and engraftment.

Materials and methods

Contact for reagent and resource sharing

Further information may be obtained from the Lead Contact Eric J. Alm (Email: ejalm@mit.edu; address: Massachusetts Institute of Technology Cambridge, MA, 02139, USA)

Experimental model and subject details

Human subjects

Human subject enrollment and sample collection was approved by the Institutional Review Board of the Massachusetts Institute of Technology (IRB Approval Number: 1510271631). Informed consent was obtained from all subjects. A total of 12 male and 12 female healthy human subjects (age range 21–65) with no history of antibiotic use in the last 6 months were enrolled in the study. In total, 24 fecal samples were collected from these individuals and an additional 24 fecal samples were collected from one male individual (age 24) over 24 days for culturing and DNA isolation.

Method details

Fecal sample processing and storage

Fecal samples were collected and processed in a biosafety cabinet within 30 min of defecation. Samples (5 g) were suspended in 20 mL of 1% sodium hexametaphosphate solution (a flocculant) in order to bring biomass into solution as described previously [29]. Fecal samples were bump vortexed with glass beads to homogenize, and centrifuged at

50×g for 5 min at room temperature to sediment particulate matter and beads. Triplicate aliquots of 1 mL of the supernatant liquid were transferred into cryovials and stored at −80 °C until processing. For the time series, samples were collected at ~24-h intervals.

Resistant fraction enrichment from fecal samples

We modified a previously published method [29] for endospore sequencing to increase throughput and decrease signal from contaminating, non-endospore-forming organisms. Fecal samples previously frozen at −80 °C were thawed at 4 °C prior to use, and 500 µL was aliquoted for resistant fraction, while the remaining 500 µL was saved for bulk community DNA extraction. Samples were centrifuged at 4 °C and 10,000×g for 5 min, washed and then resuspended in 1 mL Tris-EDTA pH 7.6. Samples were heated at 65 °C for 30 min with shaking at 100 r.p.m. and then cooled on ice for 5 min. Lysozyme (10 mg/mL) was added to a final concentration of 2 mg/mL and the samples were incubated at 37 °C for 30 min with shaking at 100 r.p.m. At 30 min, 50 µL Proteinase K (>600 mAU/ml) (Qiagen) was added and the samples incubated for an additional 30 min at 37 °C. Next, 200 µL 6% SDS, 0.3 N NaOH solution was added and the samples incubated for 1 h at room temperature with shaking at 100 r.p.m. Samples were then centrifuged at 10,000 r.p.m. for 30 min. At this step, a pellet containing resistant endospores should be visible or slightly visible in the sample, and the pellet is washed three times at 10,000×g with 1 mL chilled sterile ddH₂O. The pellet is then resuspended in 100 µL ddH₂O, and treated with 2 µL DNase I (Ambion) to remove residual contaminating DNA with incubation at 37 °C for 30 min. The DNase is killed by addition of 10 µL Proteinase K (Qiagen) and incubation at 50 °C for 15 min, followed by incubation at 70 °C for 10 min to inactivate Proteinase K. At this step, microscopic examination of samples is used to confirm the presence of phase-bright spores. The sample is then ready for downstream extraction and sequencing.

Extraction of nucleic acids

We extracted DNA from both the original sample suspended in sodium hexametaphosphate and the output of the resistant fraction. Both the original sample and the resistant fraction were extracted with MoBio PowerSoil Isolation Kit (MoBio Laboratories, Inc.) with three 10 min bead-beating steps followed by sequential collection of 1/3 of the solution to enhance recovery of endospore DNA as shown previously [48]. DNA was extracted from bacterial pure cultures, fecal enrichment cultures, and endospores using the same protocol as for fecal samples. DNA from bacterial colonies for 16S rDNA Sanger sequencing confirmation or

qPCR was obtained by homogenizing colonies in alkaline polyethylene glycol buffer as described previously [49].

16S rDNA library preparation and sequencing

Libraries for paired-end Illumina sequencing were constructed using a two-step 16S rRNA PCR amplicon approach as described previously with minor modifications [50]. In order to account for cross-sample and buffer contamination, triplicate negative controls comprising resistant fraction extraction blanks, nucleic acid extraction blanks, and PCR negatives were included during library preparation and samples were randomized across the plate. The first-step primers (PE16S_V4_U515_F, 5'-ACACGACGCTCTCCGATCTYRYRGTGCCAGCMGCCGCGGTAA-3'; PE16S_V4_E786_R, 5'-CGGCATTCCTGCTGAACCGCTCTTCCGATCTGGACTACHVGGGTWTCTAAT-3') contain primers U515F and E786R targeting the V4 region of the 16S rRNA gene, as described previously [50]. Additionally, a complexity region in the forward primer (5'-YRYR-3') was added to help the image-processing software used to detect distinct clusters during Illumina next-generation sequencing. A second-step priming site is also present in both the forward (5'-ACACGACGCTCTCCGATCT-3') and reverse (5'-CGGCATTCCTGCTGAACCGCTCTTCCGATCT-3') first-step primers. The second-step primers incorporate the Illumina adapter sequences and a 9-bp barcode for library recognition (PE-III-PCR-F, 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'; PE-III-PCR-001-096, 5'-CAAGCAGAAGACGGCATACGAGATNNNNNNNNCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-3', where N indicates the presence of a unique barcode).

Real-time qPCR before the first-step PCR was done to ensure uniform amplification and avoid overcycling all templates. Both real-time and first-step PCRs were done similarly to the manufacturer's protocol for Phusion polymerase (New England BioLabs, Ipswich, MA). For qPCR, reactions were assembled into 20 µL reaction volumes containing the following: DNA-free H₂O, 8.9 µL, HF buffer, 4 µL, dNTPs 0.4 µL, PE16S_V4_U515_F (3 µM), 2 µL, PE16S_V4_E786_R (3 µM) 2 µL, BSA (20 mg/mL), 0.5 µL, EvaGreen (20×), 1 µL, Phusion, 0.2 µL, and template DNA, 1 µL. Reactions were cycled for 40 cycles with the following conditions: 98 °C for 2 min (initial denaturation), 40 cycles of 98 °C for 30 s (denaturation), 52 °C for 30 s (annealing), and 72 °C for 30 s (extension). Samples were diluted based on qPCR amplification to the level of the most dilute sample, and amplified to the maximum number of cycles needed for PCR amplification of the most dilute sample (18 cycles, maximally, with no more than 8 cycles of second-step PCR). For first-step PCR, reactions were

scaled (EvaGreen dye excluded, water increased) and divided into three 25-µL replicate reactions during both first- and second-step cycling reactions and cleaned after the first- and second-step using Agencourt AMPure XP-PCR purification (Beckman Coulter, Brea, CA) according to manufacturer instructions. Second-step PCR contained the following: DNA-free H₂O, 10.65 µL, HF buffer, 5 µL, dNTPs 0.5 µL, PE-III-PCR-F (3 µM), 3.3 µL, PE-III-PCR-XXX (3 µM) 3.3 µL, Phusion, 0.25 µL, and first-step PCR DNA, 2 µL. Reactions were cycled for 10 cycles with the following conditions: 98 °C for 30 s (initial denaturation), 10 cycles of 98 °C for 30 s (denaturation), 83 °C for 30 s (annealing), and 72 °C for 30 s (extension). Following second-step clean-up, product quality was verified by DNA gel electrophoresis and sample DNA concentrations determined using Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific). The libraries were multiplexed together and sequenced using the paired-end with 250-bp paired-end reads approach on the MiSeq Illumina sequencing machine at the BioMicro Center (Massachusetts Institute of Technology, Cambridge, MA).

qPCR

For testing of the resistant fraction protocol, qPCR was carried out as described in the 16S rDNA Library Preparation and Sequencing section. Total bacterial abundance was quantified using the same primers. For quantification of Firmicutes and Actinobacteria, primer sequences were obtained from [51]. Primers were used at the same concentrations as 16S primers, and annealing temperatures were adjusted to the appropriate temperature for the corresponding primer pairs.

16S rDNA sequence data processing and quality control

Paired-end reads were joined with PEAR [52] using default settings. After read joining, the complexity region between the adapters and the primer along with the primer sequence and adapters were removed. Except where specified otherwise, sequences were processed using the DADA2 [53] pipeline in R, trimming sequences to 240 bp long after quality filtering (quality trim Q10) with maximum expected errors set to 1. A sequence variant table was generated using DADA2. Sequence variants were classified using RDP [54, 55]. The resulting count tables were used as input for analysis within R.

Identifying high-confidence endospore-forming and resistant OTUs

We developed a workflow for identifying organisms showing increased abundance in the resistant fraction

relative to the bulk community. First sequences present at >1% in negative control samples were removed from the DADA2 sequence variant table. The resultant pruned sequence variant table was down-sampled to the minimum read depth (25808) and then used to calculate a resistance score for each sequence variant in each sample as $\text{Resistance Score} = (\# \text{ of reads in resistant fraction}) / (\# \text{ of reads in resistant fraction} + \# \text{ of reads in bulk community})$. We then identified sequence variants that had a resistance score >0.5 (more reads in the resistant fraction than in the bulk) at least once across samples, denoting these sequence variants rOTUs. All other OTUs were considered nOTUs. Next, because there were several sequence variants found in the resistant fraction that were absent from all bulk communities (291 with 0 prevalence of 795 total rOTUs), we excluded calculations (as in Fig. 4d) involving these OTUs, which would have apparently deflated prevalence estimate from the bulk community samples.

To compile a list of high-confidence resistant fraction-enriched organisms (as the reference database for use with other datasets), we took a similar strategy as before, but also included OTUs which had 0 counts in the bulk community but non-zero counts in the resistant fraction. The OTUs increased in abundance in the resistant fraction compared to the bulk community in more than half of the samples present (excluding singletons) were included in this list.

Genomic spore gene content

Protein sequences in *Bacillus subtilis subtilis* 168 from genes identified as shared among all spore-forming Bacilli and Clostridia [35] were downloaded from UniProt (<http://www.uniprot.org/>) to make a spore gene database. All genomes as of August 2016 from 9 genera of the Clostridia in containing OTUs that were both significantly enriched at times in the resistant fraction and significantly unenriched (based on Mann-Whitney U test) were downloaded from NCBI. A standard tblastn approach was used to identify homologs in the downloaded genomes with the corresponding genes in the spore gene database. After identifying presence/absence of spore genes, genome spore gene profiles were hierarchically clustered using UPGMA on the binary distance (Jaccard) matrix.

Metabolite profiling

Metabolites were measured using liquid chromatography tandem-mass spectrometry (LC-MS) method operated on a Nexera X2 U-HPLC (Shimadzu Scientific Instruments; Marlborough, MA) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA) methods. Stool samples (200 mg/mL in 1% sodium hexametaphosphate) were homogenized using a

TissueLyser II (Qiagen). Stool homogenates (30 µL) were extracted using 90 µL of methanol containing PGE2-d4 as an internal standard (Cayman Chemical Co.; Ann Arbor, MI) and centrifuged (10 min, 10,000×g, 4 °C). The supernatants (2 µL) were injected onto a 150 × 2.1 mm ACQUITY UPLC BEH C18 column (Waters; Milford, MA). The column was eluted isocratically at a flow rate: 450 µL/min with 20% mobile phase A (0.1% formic acid in water) for 3 min followed by a linear gradient to 100% mobile phase B (acetonitrile with 0.1% formic acid) over 12 min. MS analyses were carried out using electrospray ionization in the negative ion mode using full scan analysis over m/z 70–850 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, −3.5 kV; capillary temperature, 320 °C; probe heater temperature, 300 °C; sheath gas, 45; auxiliary gas, 10; and S-lens RF level 60. Raw data were processed using TraceFinder 3.3 (Thermo Fisher Scientific; Waltham, MA) and Progenesis QI (Nonlinear Dynamics; Newcastle upon Tyne, UK) software for detection and integration of LC-MS peaks.

Bile germination tests

Treatment of fecal samples with ethanol has previously been shown to allow culture-based recovery of endospore-forming organisms [3]. To this end, fresh fecal samples were homogenized in 50% ethanol (250 mg/mL), incubated for 1 h under aerobic conditions with shaking at 100 r.p.m., and washed three times (5 min, 10,000×g) with sterile water to remove residual ethanol. Serial dilutions from 1e−4–10% (w/v) bile bovine oxgall (Sigma) were prepared in sterile water and 2.5 mL ethanol-treated fecal suspension mixed in triplicate with 2.5 mL each of these bile solutions. Samples were incubated under aerobic conditions for 2 h at 37 °C with 200 r.p.m. shaking, and then transferred to −80 °C prior to resistant fraction extraction and 16S rDNA library preparation.

Bile germination analysis

We transformed 16S rDNA sequencing counts generated by the bile germination tests using the cumulative sum-scaling transformation [56]. Under the assumption that cells in the resistant fraction can only decrease or remain the same during treatment, we searched for negative relationships between bile acid concentration and abundance that would indicate an OTU had germinated. To identify significant negative relationships, we first fit a generalized linear model (GLM) with a log-link quasi-poisson distribution to the normalized counts of OTUs present in the control sample with bile acid concentration as the predictor variable. We then identified the OTU with the strongest positive trend in

the data (that with the highest positive slope and lowest p -value). We assume that OTUs increase due only to compositional effects (that is, this OTU has not germinated but its abundance apparently increases due to loss of other OTUs), and we use the slope estimated from the fit of this model to detrend the other dose-response data so as to constrain the abundance of this apparently increasing OTU to be constant. We do so by dividing counts of all OTUs by $\exp(\text{slope} \times \text{bile acid concentration})$, which is also a measure of the depletion of the endospore-enrichment biomass. From this detrended dose-response data, we again fit a quasi-poisson GLM and identify putatively germinating OTUs as those having a significant ($p < 0.05$) negative slope.

Analysis of infant gut time series

SRA files containing 16S rDNA Sequences were downloaded from Genbank under accession no. [SRA012472](#) [42]. Sequences were generated using a Roche 454 pyrosequencer. In order to simplify analysis of the dataset, these sequences were again processed using the protocol outlined for processing of the original dataset in this paper. However, sequences were quality trimmed using Q20 to 230 base pairs, and the retained sequences were used to call 100% OTUs. OTUs were assigned taxonomies using RDP and 100% OTUs were collapsed into taxonomic names. As very few sequences matched between datasets when using uclust, these taxonomic names were instead used to identify organisms as potential resistant cell-formers based on the correspondence to the RDP-assigned taxonomic names of high-confidence resistant cell-formers identified previously. While this approach loses information given the noted heterogeneity in resistance phenotypes even among closely related strains, the original sequences themselves are still proxies for having this phenotype, and so the results of such analysis must be interpreted keeping this observation in mind.

The relative abundance of organisms identified in the infant gut time series as putative resistant-cell formers were summed, and the dynamics of this resistant cell-forming population in the infant gut was visualized over time.

Analysis of 16S rDNA sequence files from first time and recurrent *C. difficile* infection

The open reference 97% OTU table including RDP taxonomic annotations from [45] was used for this analysis [45]. OTU IDs were mapped using uclust to the corresponding genus level OTUs identified as rOTUs from this study. Patients were grouped either as healthy, first-time *C. difficile* infection (fCDI), or recurrent *C. difficile* infection (rCDI), and the fraction of rOTUs was calculated by

summing their relative abundances within each patient. A Mann–Whitney U -test was used to determine whether there were significant differences in the total relative abundance of rOTUs across groups with a Bonferroni multiple hypothesis test correction.

Analysis of 16S rDNA sequence files from fecal microbiota transplant in relapsing *C. difficile* infection

This dataset was obtained from [46]. To simplify analysis, an existing closed-reference GreenGenes 97% OTU table generated by the original authors was used. Closed-reference OTU IDs were mapped back to GreenGenes reference sequences, and sequences were assigned to the resistant cell-former database sequences again using uclust as for the adult time series.

Unique pre-FMT, post-FMT, and donor samples were separated in the dataset. We again identified organisms that had significantly different relative abundance (Benjamini–Hochberg adjusted Mann–Whitney U -test $p < 0.05$) across the groups for our analysis. We again obtained four categories of OTUs: non-resistant and resistant cell-formers that were elevated in the donor and the post-FMT samples relative to the pre-FMT samples. We used the Fisher exact test on the contingency table containing the number of OTUs in each of the previously mentioned categories to identify whether OTU engraftment from the donor was different across the groupings.

Analysis of 16S rDNA sequence files in adult time series pre- and post-Salmonella Infection

Illumina HiSeq sequencing files containing 16S rDNA sequences from the stool of a healthy adult male [47] were downloaded and processed as described for the original dataset in this paper, except that sequences were trimmed to 101 base pairs as described previously before calling 100% OTUs due to the use of shorter read sequencing technology. Sequences were assigned to the resistant cell-former database sequences using uclust constrained with the parameters: `--id 99 --usersort --libonly`, in order that sequences from this dataset would be assigned only to resistant cell-formers.

In order to assess the presence of differential turnover between resistant and non-resistant cell formers in this dataset, we identified organisms that had significantly different relative abundance (Benjamini–Hochberg adjusted Mann–Whitney U -test $p < 0.05$) before Salmonella infection starting at day 151 (days 0–150) and after the end of infection at day 159 (days 160–252). We partitioned these OTUs into four sets for our analysis: non-resistant and resistant cell formers whose median abundance was higher post-infection and those whose median abundance was

lower post-infection. We used the Fisher exact test on the contingency table containing the number of OTUs in each of the previously mentioned categories to identify whether the OTU turnover was different across the groupings.

Data availability

All amplicon sequencing data generated in this study have been can be accessed on the US National Center for Biotechnology Information SRA database under BioProject PRJNA389431. Metabolomic data and DADA2 sequence variant tables are available online through Github (<https://github.com/microbetraîner/Spores>).

Code availability

All custom scripts generated in R to analyze the data in this paper will be made available through GitHub (<https://github.com/microbetraîner/Spores>).

Acknowledgements We thank Fatima Hussain and Mathieu Groussin for extensive discussion and experimental advice. We thank the MIT BioMicro Center for sequencing service. We thank the members of the Losick lab for helpful advice and feedback on assaying sporulation phenotypes. Special thanks to Claire Duvallet for feedback on figure design and editing and manuscript clarity as well as the MIT Biological Engineering Communication Lab. We thank the members of the Alm Lab in general for intellectual support, discussion of data, and design of experiments. We thank Hilary Browne for discussions involving the cultivation and revival of endospore-forming organisms in the gut. We thank the MIT BioMicro Center for sequencing service. We thank the Chisholm lab and the Weiss Lab at MIT for allowing us to use their microscopes and qPCR machines.

Funding The funding was provided by the Broad Institute BN10 Training Grants. S.M.K. was funded by an NSF Graduate Research Fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Filippidou S, Junier T, Wunderlin T, Lo C-C, Li P-E, Chain PS, et al. Under-detection of endospore-forming Firmicutes in metagenomic data. *Comput Struct Biotechnol J*. 2015;13:299–306.
- Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. *Science*. 2008;320:1647–51.
- Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, et al. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporulation. *Nature*. 2016;533:543–6.
- Flint JF, Drzymalski D, Montgomery WL, Southam G, Angert ER. Nocturnal production of endospores in natural populations of epulopiscium-like surgeonfish symbionts. *J Bacteriol*. 2005;187:7460–70.
- Angert ER, Losick RM. Propagation by sporulation in the guinea pig symbiont *Metabacterium polyspora*. *Proc Natl Acad Sci USA*. 1998;95:10218–23.
- Alexander CJ, Citron DM, Brazier JS, Goldstein EJ. Identification and antimicrobial resistance patterns of clinical isolates of *Clostridium clostridioforme*, *Clostridium innocuum*, and *Clostridium ramosum* compared with those of clinical isolates of *Clostridium perfringens*. *J Clin Microbiol*. 1995;33:3209–15.
- Paredes-Sabja D, Torres JA, Setlow P, Sarker MR. *Clostridium perfringens* Spore Germination: characterization of germinants and their receptors. *J Bacteriol*. 2008;190:1190–201.
- Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA*. 2008;105:16731–6.
- Png CW, Lindén SK, Gilshenan KS, Zoetendal EG, McSweeney CS, Sly LI, et al. Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am J Gastroenterol*. 2010;105:2420–8.
- Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, et al. The *Clostridium difficile* *spo0A* gene is a persistence and transmission factor. *Infect Immun*. 2012;80:2704–11.
- Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, et al. Treg induction by a rationally selected mixture of *Clostridia* strains from the human microbiota. *Nature*. 2013;500:232–6.
- Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, et al. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science*. 2011;331:337–41.
- Stefka AT, Feehley T, Tripathi P, Qiu J, McCoy K, Mazmanian SK, et al. Commensal bacteria protect against food allergen sensitization. *Proc Natl Acad Sci USA*. 2014;111:13145–50.
- Kim Y-G, Sakamoto K, Seo S-U, Pickard JM, Gilliland MG, Pudlo NA, et al. Neonatal acquisition of *em* *Clostridia/em* species protects against colonization by bacterial pathogens. *Science*. 2017;356:315.
- Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. 2009;139:485–98.
- Kuwahara T, Ogura Y, Oshima K, Kurokawa K, Ooka T, Hirakawa H, et al. The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. *DNA Res*. 2011;18:291–303.
- Szczesnak A, Segata N, Qin X, Gevers D, Petrosino JF, Huttenhower C, et al. The genome of Th17 cell-inducing segmented filamentous bacteria reveals extensive auxotrophy and adaptations to the intestinal environment. *Cell Host Microbe*. 2011;10:260–72.
- Eeckhaut V, Van Immerseel F, Croubels S, De Baere S, Haesebrouck F, Ducatelle R, et al. Butyrate production in phylogenetically diverse Firmicutes isolated from the chicken caecum: butyrate-producing bacteria from the chicken caecum. *Microb Biotechnol*. 2011;4:503–12.
- Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature*. 2013;504:446–50.
- Louis P, Young P, Holtrop G, Flint HJ. Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. *Environ Microbiol*. 2010;12:304–14.
- Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly-Y M, et al. The microbial metabolites, short-chain fatty

- acids, regulate colonic treg cell homeostasis. *Science*. 2013;341:569–73.
22. Van den Abbeele P, Belzer C, Goossens M, Kleerebezem M, De Vos WM, Thas O, et al. Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model. *ISME J*. 2013;7:949–61.
 23. Yano JM, Yu K, Donaldson GP, Shastri GG, Ann P, Ma L, et al. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell*. 2015;161:264–76.
 24. Makarova KS, Wolf YI, Koonin EV. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct*. 2009;4:19.
 25. Maurice CF, Haider HJ, Turnbaugh PJ. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell*. 2013;152:39–50.
 26. Li L, Mendis N, Trigui H, Oliver JD, Faucher SP. The importance of the viable but non-culturable state in human bacterial pathogens. *Front Microbiol*. 2014;5:258.
 27. Kaplan I, Williams JW. Spore formation among the anaerobic bacteria: I. The formation of spores by *Clostridium sporogenes* in nutrient agar media. *J Bacteriol*. 1941;42:265.
 28. Dingman DW, Stahly DP. Medium promoting sporulation of *Bacillus* larvae and metabolism of medium components. *Appl Environ Microbiol*. 1983;46:860–9.
 29. Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Endospore-enriched sequencing approach reveals unprecedented diversity of *Firmicutes* in sediments: endospore-forming enrichment. *Environ Microbiol Rep*. 2014;6:631–9.
 30. Sekar R, Pernthaler A, Pernthaler J, Warnecke F, Posch T, Amann R. An improved protocol for quantification of freshwater actinobacteria by fluorescence in situ hybridization. *Appl Environ Microbiol*. 2003;69:2928–35.
 31. Fahlgren A, Hammarström S, Danielsson Å, HAMMARSTRÖM M-L. Increased expression of antimicrobial peptides and lysozyme in colonic epithelial cells of patients with ulcerative colitis. *Clin Exp Immunol*. 2003;131:90–101.
 32. Keshav S, Chung P, Milon G, Gordon S. Lysozyme is an inducible marker of macrophage activation in murine tissues as demonstrated by in situ hybridization. *J Exp Med*. 1991;174:1049.
 33. Gueimonde M, Laitinen K, Salminen S, Isolauri E. Breast milk: a source of bifidobacteria for infant gut development and maturation? *Neonatology*. 2007;92:64–6.
 34. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26:2460–1.
 35. Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Ríden DJ. Genomic determinants of sporulation in *Bacilli* and *Clostridia*: towards the minimal set of sporulation-specific genes: distribution of sporulation genes in *Bacilli* and *Clostridia*. *Environ Microbiol*. 2012;14:2870–90.
 36. Serra CR, Earl AM, Barbosa TM, Kolter R, Henriques AO. Sporulation during growth in a gut isolate of *Bacillus subtilis*. *J Bacteriol*. 2014;196:4184–96.
 37. Francis MB, Allen CA, Shrestha R, Sorg JA. Bile acid recognition by the *clostridium difficile* germinant receptor, CspC, is important for establishing infection. *PLoS Pathog*. 2013;9:e1003356.
 38. Devkota S, Wang Y, Musch MW, Leone V, Fehlner-Peach H, Nadimpalli A, et al. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in *IL10*^{−/−} mice. *Nature*. (2012). e-pub ahead of print, <https://doi.org/10.1038/nature11225>.
 39. Sorg JA, Sonenshein AL. Inhibiting the initiation of *clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. *J Bacteriol*. 2010;192:4983–90.
 40. Ceuppens S, Uyttendaele M, Drieskens K, Heyndrickx M, Rajkovic A, Boon N, et al. Survival and germination of *Bacillus cereus* spores without outgrowth or enterotoxin production during in vitro simulation of gastrointestinal transit. *Appl Environ Microbiol*. 2012;78:7698–705.
 41. Shah IM, Laaberki M-H, Popham DL, Dworkin J. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell*. 2008;135:486–96.
 42. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci USA*. 2011;108:4578–85.
 43. Chandan RC, Shahani KM, Holly RG. Lysozyme content of human milk. *Nature*. 1964;204:76.
 44. Nayfach S, Rodriguez-Mueller B, Garud N, Pollard KS. An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res*. 2016;26:1612–25.
 45. Allegretti JR, Kearney S, Li N, Bogart E, Bullock K, Gerber GK, et al. Recurrent *Clostridium difficile* infection associates with distinct bile acid and microbiome profiles. *Aliment Pharmacol Ther*. 2016;43:1142–53.
 46. Youngster I, Sauk J, Pindar C, Wilson RG, Kaplan JL, Smith MB, et al. Fecal microbiota transplant for relapsing *Clostridium difficile* infection using a frozen inoculum from unrelated donors: a randomized, open-label, controlled pilot study. *Clin Infect Dis*. 2014;58:1515–22.
 47. David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, et al. Host lifestyle affects human microbiota on daily timescales. *Genome Biol*. 2014;15:R89.
 48. Bueche M, Wunderlin T, Roussel-Delif L, Junier T, Sauvain L, Jeanneret N, et al. Quantification of endospore-forming firmicutes by quantitative PCR with the functional gene *spo0A*. *Appl Environ Microbiol*. 2013;79:5302–12.
 49. Chomczynski P, Rymaszewski M. Alkaline polyethylene glycol-based method for direct PCR from bacteria, eukaryotic tissue samples, and whole blood. *Biotechniques*. 2006;40:454.
 50. Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ. Distribution-based clustering: using ecology to refine the operational taxonomic unit. *Appl Environ Microbiol*. 2013;79:6593–603.
 51. Fierer N, Jackson JA, Vilgalys R, Jackson RB. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Environ Microbiol*. 2005;71:4117–20.
 52. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate illumina paired-end read merger. *Bioinformatics*. 2014;30:614–20.
 53. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from illumina amplicon data. *Nat Methods*. 2016;13:581.
 54. Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR. The ribosomal database project (RDP). *Nucleic Acids Res*. 1996;24:82–85.
 55. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. 2007;73:5261–7.
 56. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10:1200–2.