Unique growth and morphology properties of Clade 5 *Clostridioides difficile* strains revealed by single-cell time-lapse microscopy

John W. Ribis[1,2], César A. Nieto-Acuña[3], Qiwen Dong[1], Nicholas V. DiBenedetto[1,2], Anchal Mehra,[4] Irene Nagawa,[1] Imane El Meouche,[5] Bree B. Aldridge,[1] Mary J. Dunlop,[6] Rita Tamayo,[4] Abyudhai Singh,[3] Aimee Shen[1*]

Affiliations

[1]Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts, USA, [2] Graduate School of Biomedical Sciences, Tufts University School of Medicine, Boston, Massachusetts, USA. [3]Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA. [4]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA, [5]INSERM, Université de Paris Cité, Université Sorbonne Paris, Nord, IAME, F-75018, Paris, France, [6]Biomedical Engineering, Boston University, Boston, MA 02215.

[*]Address correspondence to
Aimee Shen, aimee.shen@tufts.edu
Phone number: (617)636-3792

Abstract

Clostridioides difficile is a major One Health threat as an important gastrointestinal pathogen of both humans and agricultural animals. The *C. difficile* species can be subdivided into 5 main clades, with Clade 5 currently undergoing speciation from Clades 1-4. Since Clade 5 strains are found more frequently in agricultural animals and can cause zoonotic infections, Clade 5 strains likely have evolved phenotypes that distinguish them from Clade 1-4 strains. Here, we compare the growth properties of Clade 5 strains to Clade 1, 2, and 4 strains using an anaerobic time-lapse microscopy system coupled with automated image analysis. These analyses revealed that Clade 5 strains grow faster than Clade 1, 2, and 4 strains and are more likely to form long chains of cells. Notably, the chaining phenotype was not shared among all Clade 5 strains examined, since 1 of the 9 strains did not form chains. Genomic analyses of the Clade 5 strains revealed that the orientation of the *cmr* switch, an invertible DNA element controlling the expression of a signal transduction system that regulates chaining, correlates with the propensity of a given Clade 5 strain to form chains. Taken together, Clade 5 strains appear to have distinct growth properties that allow them to inhabit more diverse ecological niches.

Introduction

Clostridioides difficile is a leading cause of nosocomial infection in the United States, with approximately 500,000 new infections and 14,000 deaths attributed to this organism annually (1, 2). As an obligate anaerobe, *C. difficile* relies on its ability to form hardy, metabolically dormant spores to survive outside the gut environment and transmit disease (3). When *C. difficile* spores are ingested by susceptible hosts (4, 5), they encounter bile salt signals in the small intestine that trigger their germination. These stimuli allow them to exit dormancy and outgrow into vegetative cells, which colonize the colon and secrete toxins that damage gut epithelial tissue (6). This damage triggers a significant inflammatory response that can cause disease pathologies ranging from mild diarrhea to pseudomembranous colitis, and even death (2, 4). Beyond these acute complications, *C. difficile* frequently causes recurrent infections (~20% of infections), which can lead to more severe disease symptoms (14–16) (2, 7, 8).

Notably, as a species, *C. difficile* exhibits tremendous genetic diversity (9-11). Genomic analyses indicate that *C. difficile*'s core genome represents only ~10-20% of its pan-genome (9, 10). The extreme plasticity of the "open" pan-genome of *C. difficile* likely helps the bacterium inhabit the gastrointestinal tract of diverse animals, from mammals to invertebrates, and environmental reservoirs like sewage and compost (12). Multi-locus sequence typing (MLST) analyses have revealed that the *C. difficile* species comprises 5 distinct phylogenetic clades that can be further subdivided into different ribotypes (RTs) or sequence types (STs) (10, 13). While whole-genome sequencing analyses identified three additional cryptic clades, subsequent analyses have revealed that the genomic divergence between the 5 main clades and three additional cryptic clades is so significant that the cryptic clade species fall below the threshold determined by average nucleotide identity (ANI) to qualify as part of the *C. difficile* species (10, 14).

Between the five remaining phylogenetic clades, there are notable differences in geographic and host distributions. Clade 1 is the largest and most heterogeneous (and likely ancient) clade, containing over 200 STs that have the broadest geographic distribution (15, 16). It includes both toxin-producing and non-toxigenic strains, including the well-characterized, genetically tractable toxigenic strain 630 (9). Clade 2 harbors epidemic-associated strains found within the ribotype 027 (RT027) lineage. Strains from this ribotype have been associated with

74 outbreaks in hospitals, particularly in North America, due to their frequent resistance to
75 fluoroquinolones (17-19). These epidemic strains can also cause severe disease symptoms in
76 part due to their production of three toxins: TcdA, TcdB, and CDT (binary toxin) (6). Although
77 RT027 strains have frequently been associated with "hypervirulence," there is considerable
78 phenotypic diversity within this lineage with respect to virulence, toxin production levels, and
79 sporulation (20-22).
80 Clade 3 strains are relatively uncommon and harbor a unique cell surface due to their
81 lack of CwpV production (23), but phenotypic analyses of biofilm formation and motility
82 suggest that they share similarities with Clade 2 strains like R20291 (24). Clade 4 contains
83 RT017 (ST37) strains that only encode a single toxin, TcdB (i.e., TcdA–CDT–), and are often
84 clindamycin- and fluoroquinolone-resistant. RT017 strains have been associated with outbreaks
85 in Europe and North America and are the most common strains found in Asia (25). Clade 5 is
86 the most genetically distant from the other 4 main *C. difficile* clades (94) and is thought to have
87 emerged before Clades 1-4 (26). While the ANI for Clades 1-4 ranges between 97.1 - 99.8%,
88 Clade 5 strains exhibit ANI values around 96%, which is close to the ANI value demarcation
89 used by NCBI to define organisms of the same species (10, 14). Thus, Clade 5 strains appear to
90 be actively diverging from Clades 1-4 (10).
91 Clade 5 strains are an increasing problem in healthcare and agricultural settings because
92 they can cause severe disease in humans and are commonly found in livestock, particularly pigs
93 (12, 27). While other *C. difficile* strains have been known to infect both humans and animals,
94 only Clade 5 strains have been associated with zoonotic transmission from both animal-to-
95 human and human-to-animal. The mechanisms underlying this bidirectional zoonotic
96 transmission are poorly understood, but the increased carriage of antimicrobial resistance genes
97 by Clade 5 strains may contribute to their ability to persist in livestock and humans. The
98 prevalence of antibiotic resistance genes in Clade 5 strains has major implications for One
99 Health, particularly with respect to antibiotic use in agriculture and humans (117, 119, 121) (12,
100 27, 28). Furthermore, Clade 5 strains frequently cause community-acquired infections, which
101 likely reflects wide environmental distribution and highlights the importance of understanding
102 the unique properties of this group of strains. Indeed, a recent genomic analysis suggests that
103 RT078/ST11 strains within Clade 5 frequently carry zinc acquisition and homeostasis genes
104 (11).

Despite numerous genomic analyses revealing the remarkable genetic diversity of *C. difficile* strains, relatively few studies have investigated the phenotypic diversity between strains from different clades. Clade-specific differences in colony morphology between Clade 5 strains relative to Clade 1-4 strains have been described in a limited set of analyses (26, 29), suggesting that differences in growth and/or cellular morphology may exist within clades. While differences in bulk growth rates between *C. difficile* strains have been reported (30), most phenotypic analyses have been conducted on a limited subset of strains within a given clade. Indeed, no systematic comparisons of the growth properties of the different clades or comparisons of cell morphology have been performed to date.

Here, we compare the growth properties of multiple strains derived from the four predominant main phylogenetic clades of *C. difficile* (Clade 1, 2, 4, and 5) using anaerobic time-lapse microscopy. These analyses unexpectedly revealed striking differences in the growth and cell morphology of the Clade 5 lineage relative to strains from Clades 1, 2, and 4. These morphological differences correlate with the orientation of the *cmr* switch, which has previously been implicated in modulating cell and colony morphology and cell chaining (31). Thus, the *cmr* switch appears to be an important contributor to the chaining phenotype frequently observed in Clade 5 strains.

Results

Development of a simple method for time-lapse imaging under anaerobic conditions

Time-lapse imaging of single cells has been widely used in bacteria due to the growing appreciation that clonal bacterial populations can exhibit considerable phenotypic heterogeneity, with implications for antibiotic resistance and virulence (31-35). However, live single-cell analyses in *C. difficile* have been complicated by its inability to grow in the presence of atmospheric oxygen (36). While there have been limited reports of live cell time-lapse microscopy of *C. difficile*, these analyses have required custom growth chambers to maintain anaerobic conditions (37), which may limit the accessibility of these experimental systems to most investigators.

To overcome these limitations, we established a simple system that relies solely upon commercially available reagents and materials to grow *C. difficile* cells under anaerobic conditions. This system uses gas-tight, adhesive Gene Frames™, which have been used extensively in imaging applications for bacteria (38). Notably, the gas-impermeability of these commercial seals allows maintenance of anaerobic conditions when the agarose pads, which contain growth media, are prepared in the anaerobic chamber (**Figure 1**). Gene Frames also allow highly uniform, thick agarose pads to be generated, which are properties critical for conducting automated time-lapse microscopy analyses and enabling *C. difficile* to grow in a sealed system under ambient conditions.

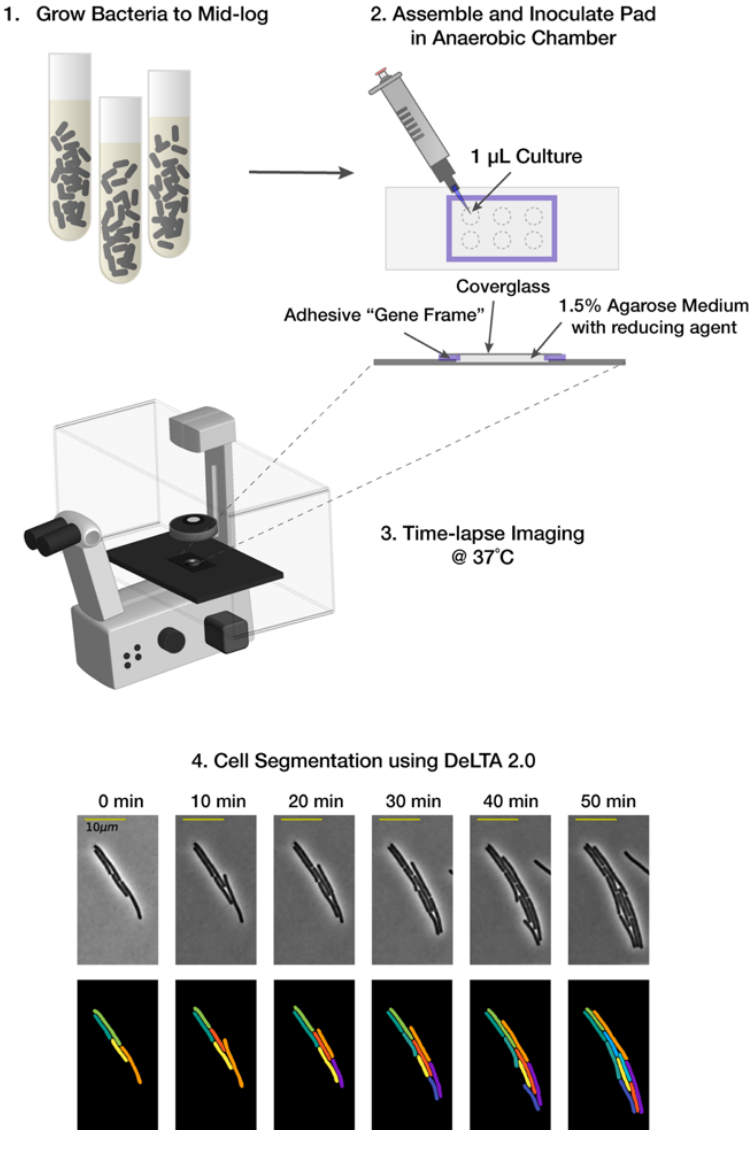


Figure 1. Schematic of the anaerobic single-cell imaging set-up. Exponentially growing *C. difficile* cells in TY medium supplemented with cysteine (TYC) are spotted onto a 1.5% agarose pads formed

148 within gas-tight adhesive Gene Frames inside the anaerobic chamber. Up to 6 strains can be spotted
149 onto a pad. The pad is sealed with a coverslip, and the imaging chamber is removed from the anaerobic
150 chamber and transferred to a heated (37˚C) microscope stage. Time-lapse microscopy is used to
151 visualize the growth of individual bacterial cells for 2-6 hrs. The output data is segmented and tracked
152 with the DeLTA 2.0 Python package (191). An example filmstrip of strain 630 grown on TYC medium
153 over time (Bottom).
154
155
156 **Time-lapse microscopy reveals clade-specific differences in elongation rate and cell length**
157
158 Following the establishment of our time-lapse imaging setup, we compared the single-
159 cell growth properties of representative *C. difficile* strains from Clade 1 (630, ribotype (RT)
160 012), Clade 2 (R20291, RT027), Clade 4 (M68, RT017), and Clade 5 (M120, RT078) (**Figure**
161 **2, Table 1**). We did not have access to Clade 3 strains, so these were not included in this study.
162 The four "representative" strains were all isolated from patients with *C. difficile*-associated
163 disease and are frequently used as reference genomes for their clades and ribotype groups.
164 Notably, RT027 (ST1), RT017 (ST45), and RT078 (ST11) strains are from ribotypes/
165 multilocus sequencing types that are frequently isolated from patients with *C. difficile* infection
166 (CDI) (12, 19, 22, 25).
167 To quantify the growth properties of single cells visualized by time-lapse microscopy,
168 we used DeLTA 2.0, a Deep Learning for Time-lapse Analysis (DeLTA) software that allows
169 rapid and accurate segmentation and tracking of bacteria growing in two dimensions on agarose
170 pads (39, 40). This software uses deep convolutional neural networks to obtain information on
171 the growth properties of individual cells growing in microcolonies on agarose pads. The
172 segmentation and tracking of *C. difficile* cells was highly accurate (**Figure 1**), and minimal user
173 input or post-image processing was needed to obtain growth property measurements.
174 Robust growth was observed for all strains using our system, with Clade 5 strain M120
175 elongating the fastest (2.1 doublings/hr), followed by Clade 2 strain R20291 (1.8 doublings/hr),
176 and then Clade 1 strain 630 and Clade 4 strain M68 (1.6-1.7 doublings/hr) (**Fig 2A**). This
177 elongation rate was defined as doublings/hr to indicate the number of times that a cell's size
178 doubles in one hour. For example, an elongation rate of 2 doublings/hr means that the cell size
179 doubled twice in 1 hr, which corresponds to an exponent of 2ln(2) 1/hr. Notably, the elongation
180 rate (doublings/hr) should not be confused with doubling time, or generation time, which are

defined as the length of *time* that it takes before a bacterium divides. Instead, the elongation rate reflects how fast the cell is increasing in length over time. Nevertheless, elongation rate and doubling time are related because most cells undergo a division event when they reach twice their original size.

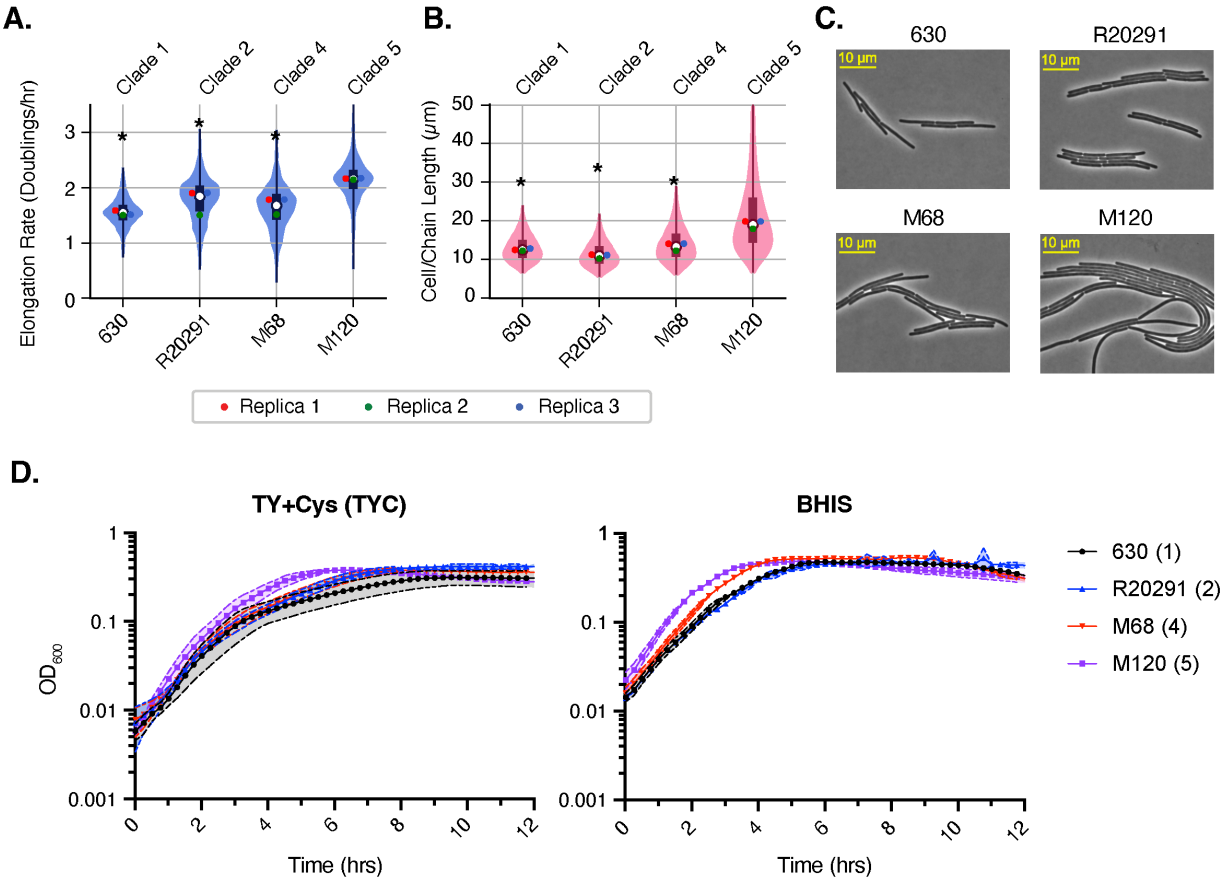


Figure 2. Clade 5 strain M120 elongates more quickly and exhibits cell chaining. A. Violin plot of the elongation rates measured during time-lapse microscopy analyses of strains 630 (Clade 1), R20291 (Clade 2), M68 (Clade 4), M120 (Clade 5) grown on TY supplemented with cysteine (TYC) agar. Data are from three biological replicates, with the mean of each replicate shown as a point on the violin. B. Violin plot of the cell length measured during time-lapse microscopy for strains 630 (Clade 1), R20291 (Clade 2), M68 (Clade 4), M120 (Clade 5). Each replicate mean is shown as a point on the violin. Statistical significance for A and B was determined by comparing the mean of the three replicates of Clade 1, 2, and 4 strains to the Clade 5 M120 strain using a Kruskal-Wallis test * $p < 0.05$. C. Phase-contrast image from time-lapse microscopy movies. Scale bar is 10 μm. D. Optical density-based analyses of bulk population growth of the indicated strains in TYC or BHIS medium. The number in brackets indicates the clade to which a given strain belongs.

Importantly, the differences in single-cell elongation rates measured for the four strains using DeLTA were also observed in bulk population analyses of their growth using optical

density in TYC and BHIS media (**Figure 2D**). These analyses confirmed that the Clade 5 strain M120 grows faster (using optical density as a proxy for bacterial growth) than the Clade 1, 2, and 4 strains analyzed. Negligible differences were observed in the elongation rates for Clade 1, 2, and 4 strains in these bulk population analyses (**Figure 2A**), consistent with the relatively subtle difference observed in the single-cell elongation rates measured for the Clade 2 R20291 strain relative to the Clades 1 and 4 strains.

The Clade 5 strain M120 exhibited another distinct growth property from the strains from Clades 1, 2, and 4. While strains from Clades 1, 2, and 4 produced cells of similar length prior to division, with an average cell size of ~13 µm prior to division, cells of Clade 5 strain M120 were significantly longer, appearing as long filaments (or chains) with an average length of 22 µm but chains close to 50 µm readily being measured (**Figure 2B**). These filaments or chains also appeared to bend readily (**Figure 2B**). Notably, to visualize an entire filament (or chain), it was necessary to stitch together several fields of view, revealing that the bacterial filaments (or chains) approached several hundred microns and even up to ~1 mm in length (**Figure 3**). To determine whether strain M120 forms filaments, i.e. long cells that lack division septa, or chains of cells, i.e. cells that remain attached despite forming division septa, we incorporated the lipophilic fluorescent dye FM4-64 which visualizes membranes (Pogliano 1999), directly into the agarose pads. This stain revealed division septa throughout the length of the cell, indicating that the extremely long cells were not filaments but instead extended chains of cells (**Figure 3, inset**). Since the spacing between division septa generated cells that were relatively similar in length to the other *C. difficile* strains analyzed, Clade 5 strain M120 is distinguished by a decreased efficiency in cell separation.

Clade 5 clinical isolates typically form long chains and grow more quickly than strains from other clades

Since prior work indicated that Clade 5 strains produce colony morphologies distinct from Clade 1, 2, and 4 strains (26), we sought to determine whether the (i) striking cell chaining phenotype and (ii) faster growth rate observed in Clade 5 strain M120 is observed in other Clade 5 strains. To address these questions, we compared the single-cell growth properties on TYC agarose of 5 additional Clade 5 clinical isolates obtained from three separate hospitals

(**Table 1**) using time-lapse microscopy analyses (**Figure 2**). Clade 1 strain 630 was included as a control since it does not form chains under these conditions.

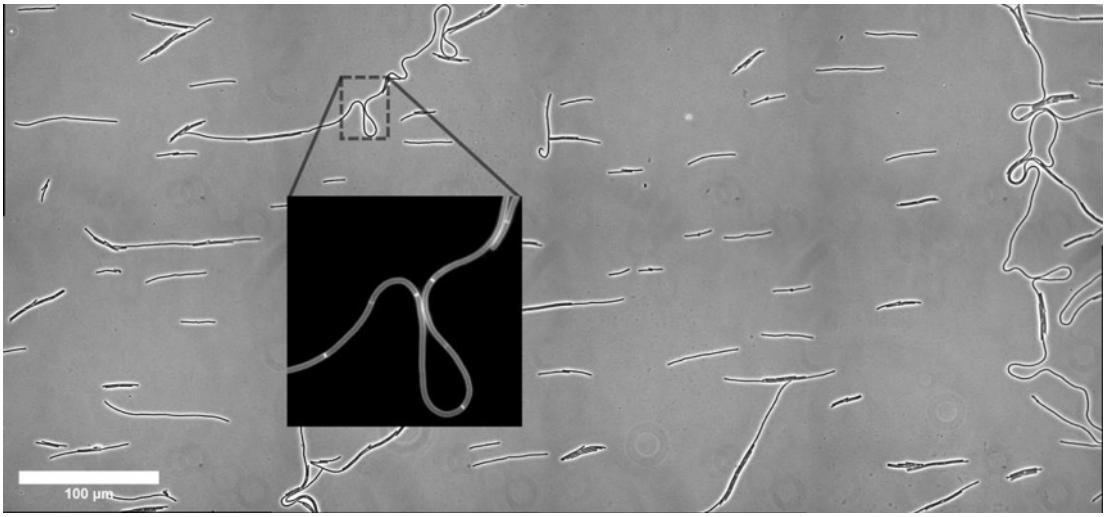


Figure 3. *C. difficile* Clade 5 can form large heterogenous chains. Large mosaic phase-contrast image of the Clade 5 *C. difficile* strain M120. Zoomed image shows chains revealed by staining with the membrane dye FM4-64. Image was stitched from 8 individual fields of view at 63x magnification. Scale bar is 100 µm.

These analyses revealed that all but one of the Clade 5 strains tested formed long chains, with TAL29600 forming the longest chains (29 µm on average, **Figure 4A-C**). In contrast, strain TAL29996 formed shorter chains that were comparable in length to those observed for strains from Clades 1, 2, and 4 (12-13 µm, **Figures 2B, 4C**). Notably, DeLTA over-estimated cell length in Clade 5 strains that form chains because it could not identify cells that had formed division septa (based on FM4-64 staining) but had not yet initiated cell separation.

To overcome this limitation and accurately quantify cell length within long chains, namely the distance between division septa, we modified our image processing pipeline to use a thresholding method to detect division septa. After generating masks in DeLTA to segment the chains, we modified the mask so that only the interior of the contour was analyzed. We then applied an adaptive thresholding method to identify division septa based on their elevated fluorescence relative to the long axis of the cell; the thresholds were defined using a Gaussian-weighted method (**Figure 4A**, red dots). These analyses revealed that Clade 5 strains that form

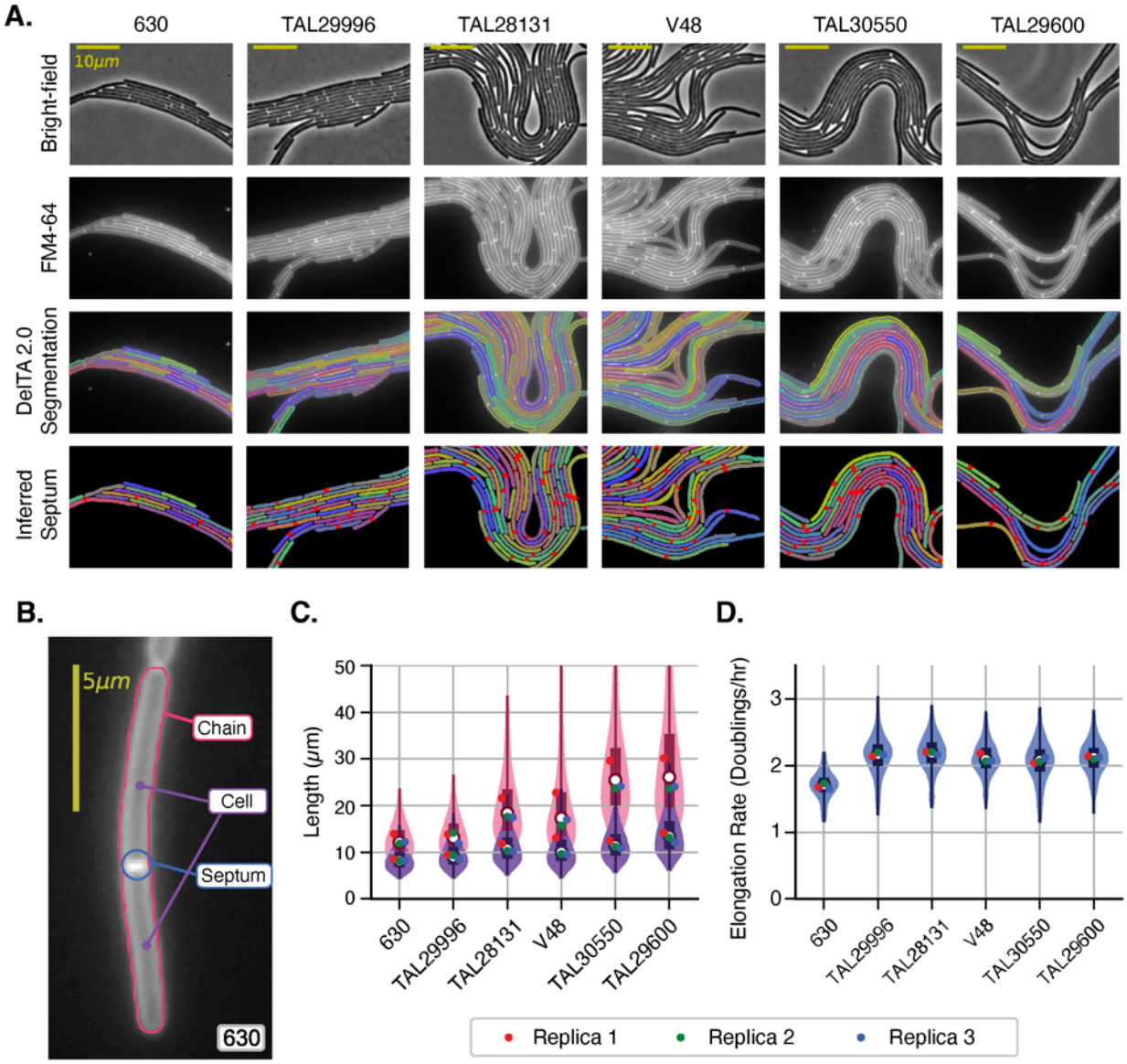


Figure 4. Faster growth is a common feature of Clade 5 strains, but the chaining phenotype is not fully penetrant in Clade 5 strains. (A) Phase-contrast microscopy images from time-lapse microscopy studies of Clade 5 strains; Clade 1 strain 630 is included for comparison. Fluorescence microscopy was used to visualize the FM4-64 stain incorporated into the agarose pads. Masks generated using DeLTA 2.0 are shown, with the bottom panel showing division septa identified with our adaptive thresholding approach. (B) Example image showing the parameters identified using DeLTA 2.0 combined with our adaptive threshold method for detecting division septa within a chain of cells. (C) Cell length measured based on automated DeLTA 2.0 analyses (pink violin plot) or DeLTA 2.0 combined with adaptive threshold analyses (purple violin plot). The former method is more likely to measure chain length, while the latter method more accurately measures cell length. (D) Violin plot of the elongation rates measured for the indicated strains based on three biological replicates. For all the violin plots shown, each replicate mean is shown as a point on the violin, and statistical significance was determined by comparing the mean of the three replicates using a Kruskal-Wallis test * $p < 0.05$.

long chains do not produce cells that are dramatically longer than strains that do not form chains; cells within the Clade 5 strains that produced the longest chains had average cell lengths (TAL29600 and TAL30550, 14 and 12 µm, respectively) that were 30-50% longer than strains that do not form chains like strains 630 (Clade 1, 8.5 µm) and TAL29996 (Clade 5, 9.4 µm). Notably, even though TAL29996 did not form long chains like the other Clade 5 strains analyzed, it elongated at rates similar to those of other Clade 5 strains (2.1 doublings/hr vs. 1.7 doublings/hr for Clade 1 strain 630**, Figure 4D**). Taken together, these analyses strongly suggest that Clade 5 strains grow faster than Clade 1, 2, and 4 strains and are more likely to form chains presumably because their cell separation mechanisms are less efficient.

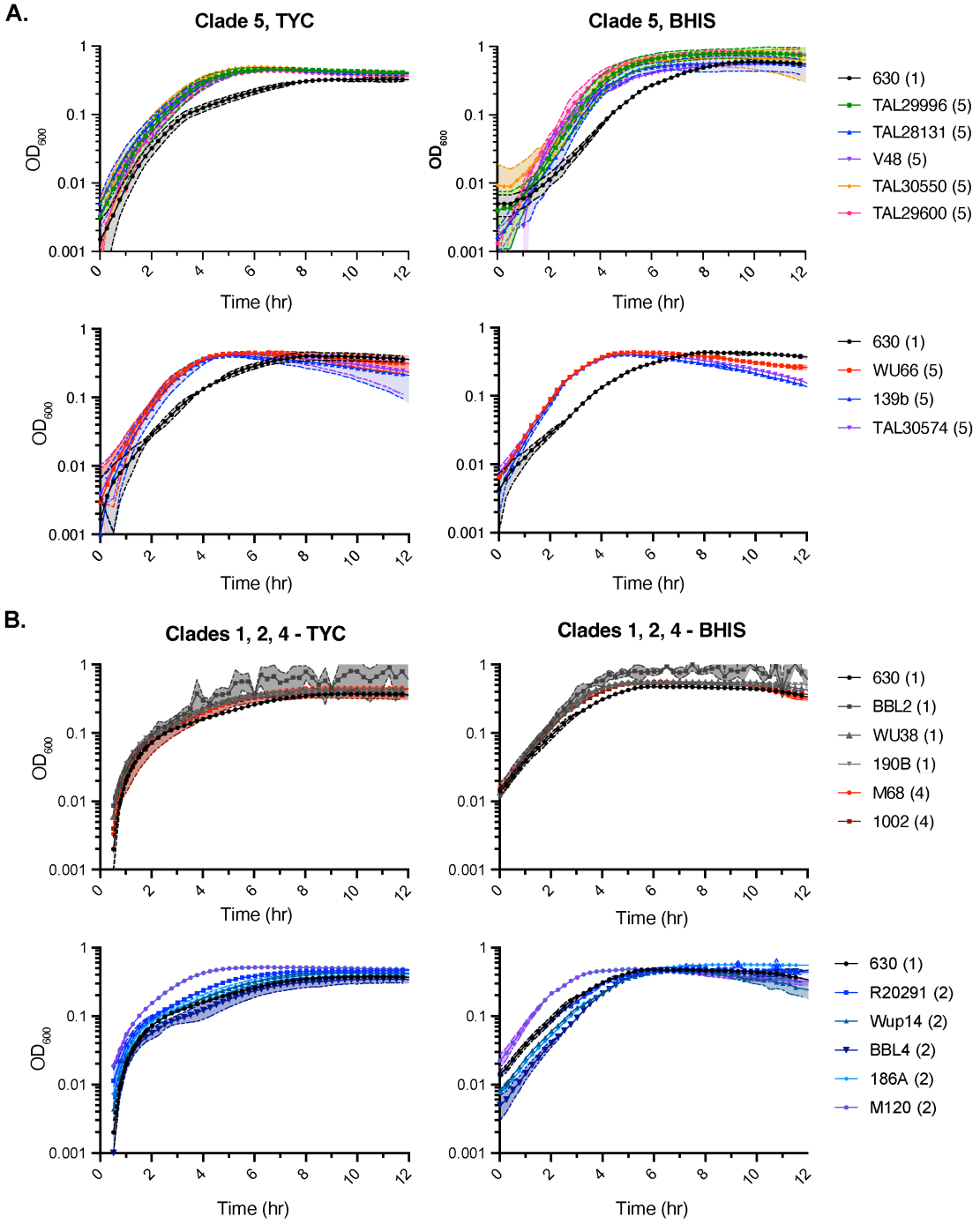
While the long chains formed by Clade 5 strains limited the number of divisions that could be analyzed within a given field using time-lapse microscopy, the faster elongation rates measured on the agarose pads were also observed in bulk analyses of growth in broth culture. These optical density-based analyses revealed that Clade 5 strains all grow faster than Clade 1, 2, and 4 strains irrespective of their ability to form chains or the growth media used (**Figure S1**).

To assess whether these conclusions could be generalized to additional Clade 5 strains vs. Clade 1, 2, and 4 strains, we analyzed the growth of additional strains from all four clades in broth culture. These analyses confirmed that Clade 1, 2, and 4 strains grow at similar rates, which are slower than those measured for the nine Clade 5 strains analyzed in both TYC and BHIS media (**Figure S1**).

Cell chaining in Clade 5 strains is not dependent on growth on a solid medium.

C. difficile has previously been shown to promote cell elongation and chain formation upon induction of the *cmrRST* locus (31), which encodes a signal transduction system. Expression of this locus is also responsive to c-di-GMP levels (29), which increases in cells grown on solid surfaces such as in a biofilm or on an agar plate (29). To determine whether the chaining phenotype observed in most Clade 5 strains is induced by growth on a surface, we assessed whether cell chaining would be observed in cells grown in broth culture. The Clade 5 strains along with strains from other clades were grown to log-phase in rich media, and their ability to form chains was assessed by visualizing division septa with the fluorescent D-amino

302 acid label, HADA. Chains of cells were still observed in the Clade 5 strains by time-lapse
303 microscopy analyses, although the chains formed in broth culture were not as long as those
304



305
306
307 **Figure S1.** Optical density-based analyses of bulk population growth of the indicated strains during
308 growth in TYC or BHIS medium. The number in brackets indicates the clade to which a given strain
309 belongs.

observed during growth on the agarose pads (**Figure 5, S2**). Chains were also observed in 4 additional Clade 5 strains analyzed during exponential-phase growth in broth culture. Interestingly, two of the Clade 5 strains, V48 and WU66, did not appear to incorporate the HADA label efficiently, and the cells appeared lighter in phase-contrast images despite undergoing the fixation process shortly after HADA addition. It is unclear why the HADA was poorly incorporated by these strains across multiple independent experiments. Regardless, these results indicate that many Clade 5 strains undergo cell separation less efficiently than strains from other clades irrespective of whether they are grown on a solid surface or in broth culture.

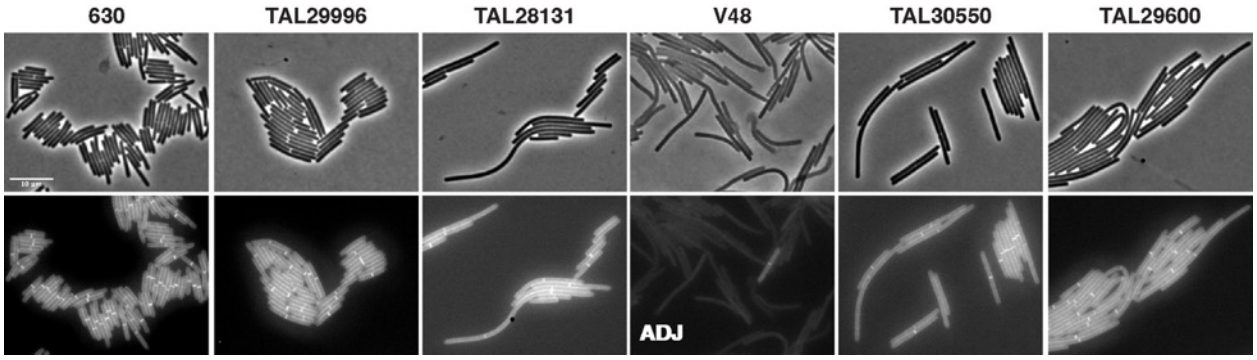


Figure 5. Clade 5 strains form chains during logarithmic growth in broth culture. Representative micrographs showing phase-contrast (top) and peptidoglycan labeling with the fluorescent D-amino acid, HADA, (bottom) following growth in rich broth (BHIS) to logarithmic phase. Scale bar, 10 µm. Data are representative of multiple independent experiments. **ADJ** indicates that the brightness of the image was enhanced to detect HADA labeling in V48.

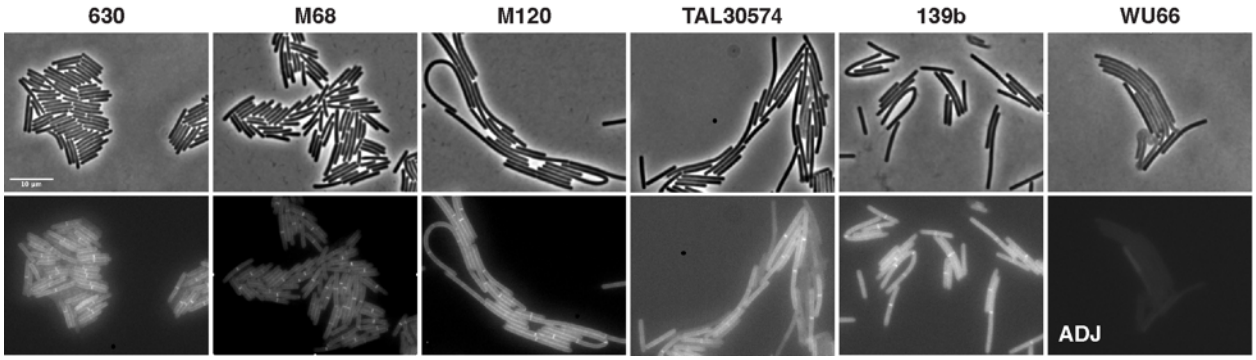


Figure S2. Clade 5, but not Clade 1 or 3, strains form chains during logarithmic growth in broth culture. Representative micrographs showing phase-contrast (top) and peptidoglycan labeling with the fluorescent D-amino acid, HADA (bottom) following growth in rich broth (BHIS) to logarithmic phase. Scale bar, 10 µm. Data are representative of multiple independent experiments. **ADJ** indicates that the brightness of the image was enhanced to detect HADA labeling in WU66.

Cell length does not correlate with the propensity to sporulate.

We next wondered whether the propensity to form chains impacts the ability of Clade 5 strains to sporulate because analyses in *Bacillus subtilis* suggest that smaller cells, such as those formed during stationary phase growth (41), may concentrate proteins involved in the phosphorelay that induces sporulation. In *B. subtilis*, the decrease in cell length helps the kinase KinA reach the threshold concentration needed to trigger sporulation initiation (42, 43). Although *C. difficile* lacks homologs of KinA and other components of the phosphorelay system, it is possible that the longer cell length observed in Clade 5 strains may similarly reduce their ability to induce sporulation by diluting the levels of a currently unknown sporulation regulator. To test this hypothesis, we analyzed the propensity of Clade 5 strains to form spores when plated on 70:30 sporulation medium using phase-contrast microscopy and heat resistance assays.

Striking differences in the frequency of spore formation were observed across the Clade 5 strains, with several strains exhibiting close to 100% sporulation levels (**Figure 6A, S3**). Although TAL29600, which forms the longest chains of the strains tested, exhibited extremely low levels of sporulation (0.003%), we did not observe a correlation between sporulation and cell length among the strains (**Figure 6B, S3**). For example, M120 forms long chains yet readily forms spores (~100% sporulation frequency). The low sporulation levels of TAL29600 were rather surprising since this strain was isolated from a clinical setting, so it likely forms spores more efficiently in conditions encountered during infection. Notably, when grown on 70:30 medium, Clade 5 strains did not appear to form the long chains observed during broth culture and grown on agarose pads. In addition, the spores produced by all *C. difficile* isolates analyzed exhibited similar lengths and proportions despite arising from vegetative cells that varied widely in their cell length (**Figure S4**). Thus, our data suggest that Clade 5 strains alter their cell length and propensity to form chains depending on the media encountered.

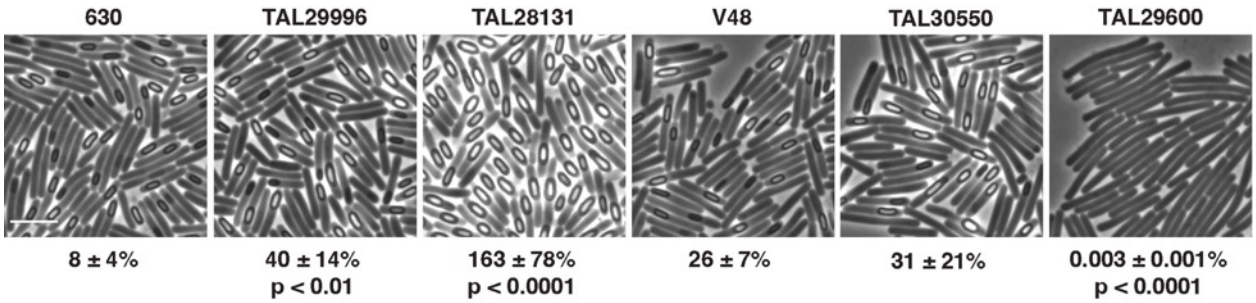


Figure 6. Sporulation levels in Clade 5 isolates grown on 70:30 medium. Phase-contrast microscopy of the indicated strains ~24 hrs after sporulation induction. Heat resistance efficiencies, indicated below

the respective images, are based on 20-24 hr sporulating cultures and represent the mean and standard deviation for a given strain relative to wild type based on a minimum of three biological replicates. Statistical significance relative to strain 630 was determined using a one-way ANOVA and Tukey's test. The scale bar represents 5 µm.

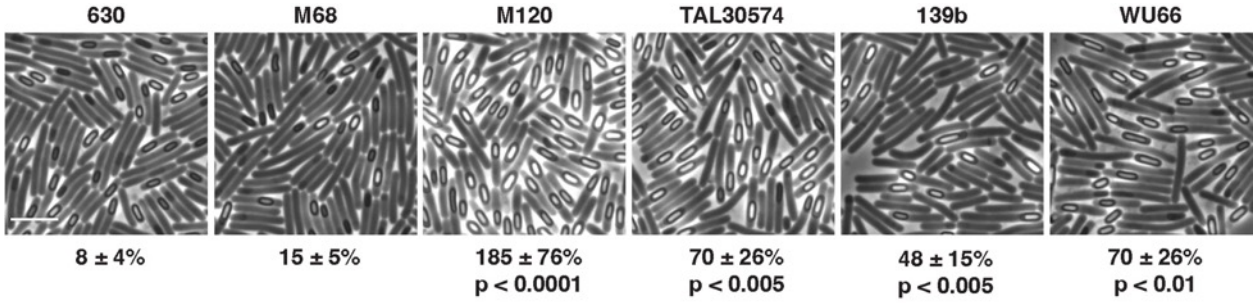


Figure S3. Sporulation levels in clinical isolates grown on 70:30 medium. Phase-contrast microscopy of the indicated strains ~24 hrs after sporulation induction. Heat resistance efficiencies are based on 20-24 hr sporulating cultures and represent the mean and standard deviation for a given strain relative to wild type based on a minimum of three biological replicates. Statistical significance relative to strain 630 was determined using a one-way ANOVA and Tukey's test. The scale bar represents 5 µm.

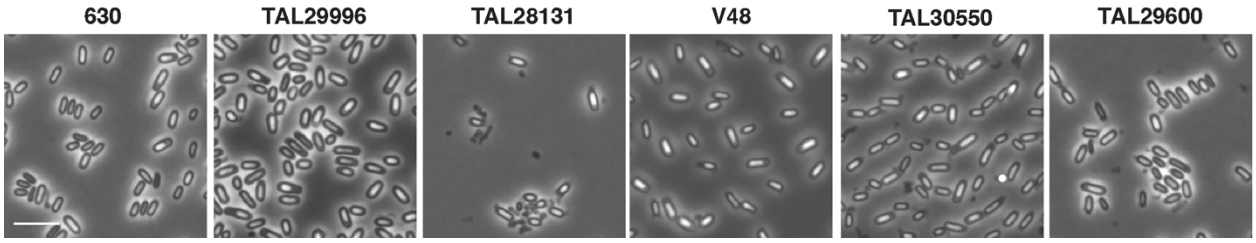


Figure S4. Spores purified from Clade 5 strains. Spores were purified using a Histodenz gradient. Strain 630 is provided as a reference. The scale bar represents 5 µm.

Genomic comparisons of Clade 5 strains

Since our phenotypic analyses revealed that TAL29996 differs from the 8 additional Clade 5 strains analyzed, we sought to gain insight into the molecular basis for this difference by comparing their genomes. To this end, we sequenced 5 of the Clade 5 strains, including TAL29996, and determined that the average nucleotide identity (ANI) for orthologous genes between the strains ranged from 99.83-99.99% (**Table S1**). Since the pan-genome between the 5 strains is 12%, they are quite closely related. A comparison of their pan-genomes revealed that TAL29996 is missing one duplication of *blaR1*, which encodes an integral membrane protein that senses beta-lactams, and a gene region predicted to be involved in nicotinate metabolism compared to the other 4 strains. We also compared the genomes using breseq (44) to identify SNPs that might distinguish TAL29996 from the other strains using strain M120 as the

reference, since this strain has been the Clade 5 strain traditionally characterized (30, 45). Around 10 SNPs were unique to TAL29996, but none were clearly implicated in regulating cell separation or peptidoglycan synthesis (**Tables S2, S3**).

Since testing whether one or some of these SNPs or absent genes help TAL29996 undergo cell separation more efficiently would require genetic manipulation of TAL29996, we instead analyzed the orientation of the *cmr* switch, also known as the Cdi6 DNA invertible element (31, 46). The orientation of the *cmr* switch affects expression of the adjacent *cmrRST* operon, which encodes a non-canonical two-component system, resulting in phase variation of CmrRST. Because CmrRST promotes cell chaining, the orientation of the *cmr* switch determines whether *C. difficile* produces rough or smooth colony morphologies (29, 31, 46). Specifically, cells from rough colonies are highly biased to the ON orientation of the *cmr* switch and exhibit a chaining phenotype, whereas cells from smooth colonies are highly biased to the OFF orientation and do not form chains (31). Consistent with the breseq analyses, manual inspection of the *cmrRST* upstream region revealed that this region is highly similar between all the Clade 5 strains analyzed (98% identity), including strain TAL29996. In the Illumina whole genome sequencing performed on the subset of 5 Clade 5 strains, the *cmr* switch is in the OFF orientation; these results are consistent with the genomic DNA from these samples being harvested from liquid culture growth, which favors the *cmr*-OFF state at least for Clade 2 strain R20291.

However, in contrast with the Illumina analyses, qPCR analyses of the *cmr* switch region in the Clade 5 isolates revealed that TAL29996 exhibits a *cmr*-ON frequency of ~20%, whereas the four remaining Clade 5 strains exhibit a *cmr*-ON frequency of between 70-96% on average **(Fig 7**). Furthermore, when we extended these analyses to the four additional Clade 5 strains characterized that also produce chains, we found that they exhibit a *cmr*-ON frequency of between 80-93%. Collectively, these results suggest that the orientation of the *cmr* switch orientation correlates with the propensity to form chains in Clade 5 strains and that most Clade 5 strains are biased to the *cmr*-ON orientation even during growth in broth culture.

Since the *cmr*-ON state has also been correlated with increased surface motility, we next analyzed the surface motility of Clade 5 strains. After incubation for 5 days, *cmr*-ON strains spread further on agar plates than the TAL29996 *cmr*-OFF strain (**Fig 7B**). Consistent with prior reports, Clade 5 strains whose *cmr* switch is predominantly in the ON orientation exhibited

greater and more uniform surface motility. In contrast, TAL29996 exhibited less and more asymmetric surface motility, with fractal-like extensions extending from a couple of sites. The surface motility of TAL29996 resembles the phenotype previously reported for Clade 2 strain R20291, whose *cmr* switch is also predominantly in the OFF position in liquid cultures. However, growth of R20291 on agar medium selects for bacteria with the *cmr* switch in the ON orientation, which leads to asymmetric surface motility on plates (31). These observations suggest that even though TAL29996 is biased to the *cmr*-OFF orientation during broth culture growth, during growth on BHIS agar, a subset of cells switch to the *cmr*-ON orientation, leading to the asymmetric spreading phenotype.

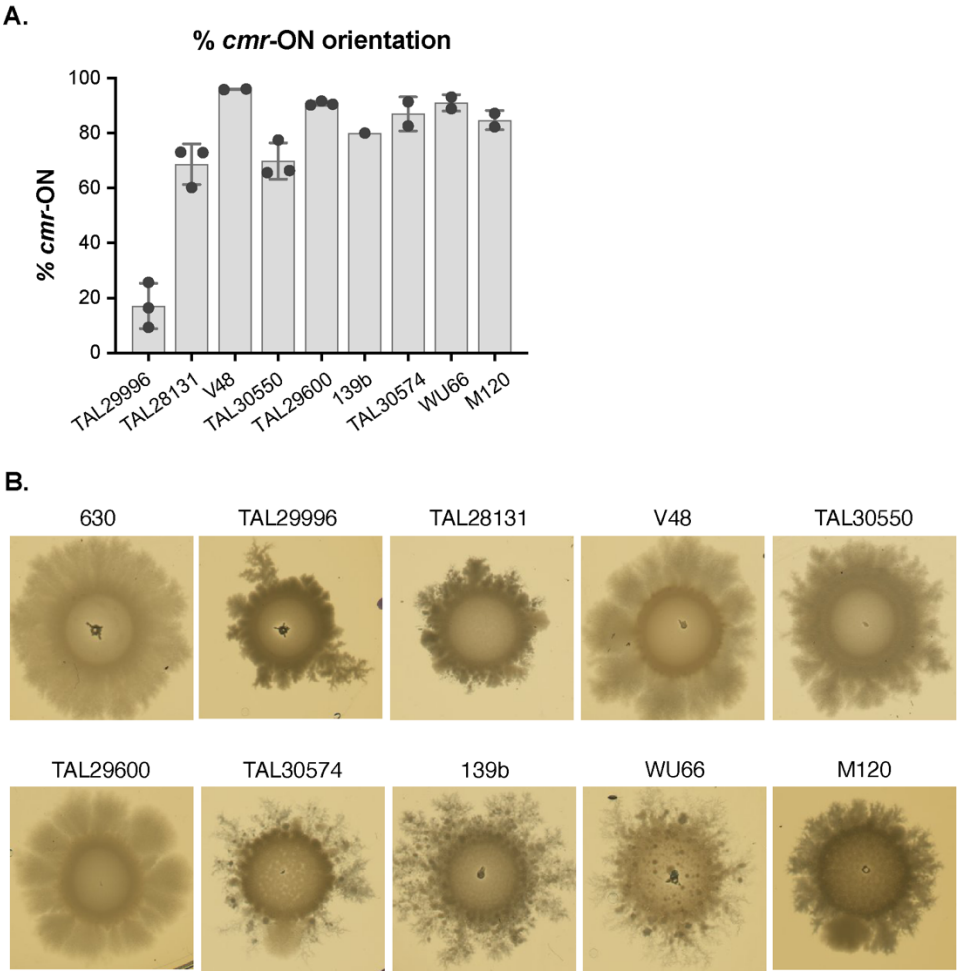


Figure 7. Orientation of the *cmr* switch correlates with chaining phenotype and surface motility in Clade 5 strains. A. Orientation-specific PCR for detecting the orientation of the *cmr* switch in the indicated strains. B. Representative images of surface motility 5 days after exponentially growing liquid cultures were spotted onto BHIS agar plates.

Analyses of the surface motility of strains from Clade 1, 2, and 4 revealed that asymmetric motility was more frequently observed in Clade 2 strains. Clade 1 strains exhibited a range of surface motility, from high surface motility with strain 630 to lower surface motility with strain WU38 (**Figure S5**). While analyzing a larger number of Clade 1, 2, and 4 strains would reveal whether there are clade-specific differences in surface motility, the data suggest that Clade 2 strains may exhibit more variability in *cmr* switching within a population during growth on plates, with strain Wup14 exhibiting the least surface motility of the Clade 2 strains tested. Collectively, the data imply that the ON orientation of the *cmr* switch during broth culture growth contributes to the chaining phenotype observed in most Clade 5 strains.

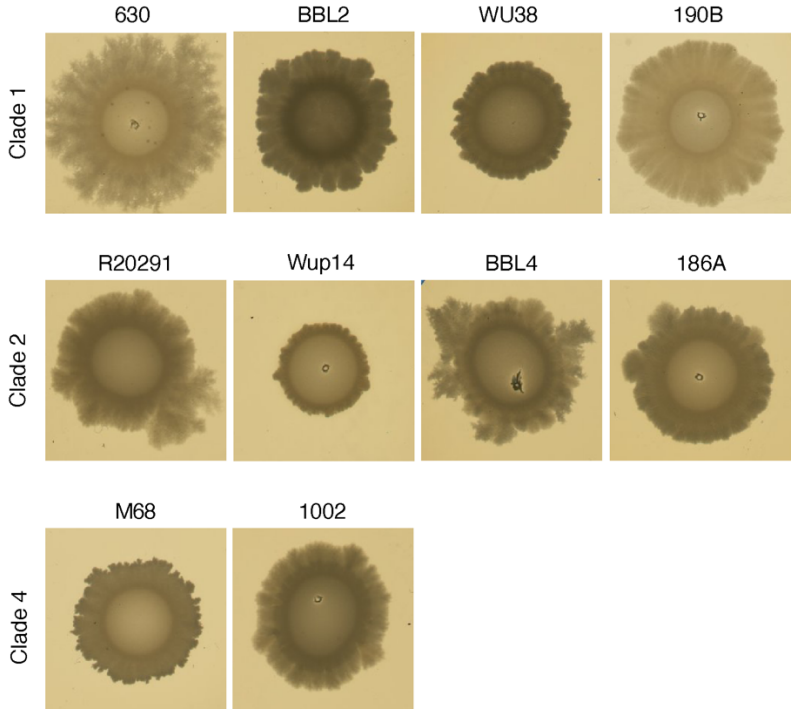


Figure S5. Surface motility in Clade 1, 2, and 4 strains. Representative images of surface motility 5 days after exponentially growing liquid cultures were spotted onto BHIS agar plates.

Discussion

While the genetically distinct nature of Clade 5 strains from Clade 1-4 strains is well-established (10, 26), the phenotypic impacts of this genetic diversity are poorly understood, with only one study to our knowledge systematically comparing Clade 5 strains to Clade 1-4 strains (10, 26). By phenotypically characterizing *C. difficile* strains from multiple clades using

462 time-lapse microscopy, we discovered that Clade 5 strains have distinct growth properties from
463 the Clade 1, 2, and 4 strains tested. Our analyses of multiple strains from each of Clades 1, 2, 4,
464 and 5 revealed that Clade 5 strains grow more quickly (**Figures 2, 4, S1**) and are more likely to
465 form chains than strains from Clades 1, 2, and 4, irrespective of media type or growth on a
466 surface (**Figures 2, 4, S2**). Notably, these analyses represent the first report of the single-cell
467 growth properties of *C. difficile* to our knowledge and were enabled by our development of a
468 simple method for conducting anaerobic time-lapse microscopy analyses (**Figure 1**) and recent
469 developments in automated image analyses of time-lapse microscopy data (**rf**).
470 Using this time-lapse microscopy method, we discovered that many Clade 5 strains
471 produce extremely long chains of cells (>100 μm in length, **Figures 3, 4**). This chaining
472 phenotype was observed in 8 of 9 Clade 5 strains versus none of the Clade 1, 2, and 4 strains,
473 regardless of whether the cells were grown on agarose pads or in broth culture. Thus, Clade 5
474 strains appear to undergo cell separation less frequently than strains from other Clades. The
475 exception to this conclusion is Clade 5 strain TAL29996, which did not form chains but
476 nevertheless grew at similarly fast rates as the other Clade 5 strains (**Figure 4**). Taken together,
477 these data suggest that faster growth is likely a characteristic that distinguishes Clade 5 strains
478 from Clade 1, 2, and 4 strains.
479 Our genomic analyses suggest that a driving factor of the chaining phenotype of Clade 5
480 strains is their propensity to have the *cmr* switch in the ON orientation (**Figure 7**), which
481 increases the expression of an operon encoding a signal transduction system that induces
482 chaining (31). This conclusion is based on our finding that the non-chaining TAL29996 strain
483 exhibits a *cmr*-ON orientation frequency of 20%, whereas the 8 additional Clade 5 strains tested
484 exhibited *cmr*-ON orientation frequencies of between ~70-95% (**Figure 7**). In addition, strain
485 TAL29996 exhibited less surface motility during growth on agar plates, with only a fraction of
486 the population spreading on the surface relative to the more uniform spreading observed for the
487 other 8 Clade 5 strains (**Figure 7**).
488 Previous work in Clade 2 strain R20291 has shown that growth on agar plates promotes
489 selection of *cmr*-ON bacteria, and increased c-di-GMP during surface growth may also promote
490 *cmrRST* expression (29), so Clade 5 strain TAL29996 may exhibit a similar surface-inducible
491 expression of the *cmrRST* operon. Indeed, the chaining phenotype of Clade 5 strains was more
492 frequent and extensive during time-lapse microscopy analyses on agarose pads compared to

broth culture-grown cells (**Figures 4, 5**). Thus, the chaining phenotype of Clade 5 strains may be similarly responsive to c-di-GMP levels during growth on a surface (29). While testing this hypothesis will depend on the ability to genetically manipulate Clade 5 strains, it is tempting to speculate that *cmrRST* expression is an important determinant of the growth properties of Clade 5 strains. Expression of this operon in Clade 2 strain R20291 not only promotes chaining and surface motility in a Type IV pilus-independent manner, it also negatively regulates biofilm formation and flagellar motility (31). In addition, growth conditions favoring flagellar motility select against *cmr*-ON bacteria. Since Clade 5 strains lack flagellar motility (12), they may be "primed" to exhibit surface motility in the absence of selection against the *cmr*-ON state. Regardless, analyzing the regulation of c-di-GMP in different growth conditions in Clade 5 strains, particularly for TAL29996 relative to the other strains, will likely provide insight into the mechanisms that drive the higher incidence of *cmr*-ON cells in most Clade 5 strains.

Even though Clade 5 strains form long chains and produce longer cells than strains from the other clades, neither cell nor chain length was predictive of the propensity to sporulate (**Figure 6**). Clade 5 strains exhibited wildly different sporulation frequencies, with TAL29600 forming spores inefficiently (~5,000-fold less efficiently than most Clade 5 strains). While little is known about the mechanisms regulating sporulation initiation outside of strains 630 and R20291, future analyses of strain TAL29600, which sporulates poorly under the laboratory media tested (**Figure 6**), could provide insight into the molecular determinants of sporulation in Clade 5 strains. For example, differences in gene presence or polymorphisms in several c-di-GMP-related genes were observed in TAL29600 relative to the other Clade 5 strains (**Table S2**), and c-di-GMP has also been implicated in regulating sporulation initiation events through an unknown mechanism (47, 48). Thus, analyzing the contribution of these genes in future work could provide important insight into c-di-GMP-mediated regulation of sporulation.

The variable sporulation levels observed for the nine Clade 5 strains analyzed in this manuscript contrast with the findings of a previous study, which reported that Clade 5 strains sporulate at lower rates in the presence of glucose and fructose (26). Kumar *et al*. previously correlated phenotypic differences in sporulation and mouse colonization between Clade 5 strains vs. Clade 1-4 strains to their finding that genes involved in sporulation and the metabolism of simple sugars are under positive selection in Clade 1-4 strains (26). However, the five Clade 5 strains analyzed in this study all exhibited very low sporulation frequencies under

the conditions tested. While the conditions used in this study to induce sporulation differ from those used by Kumar *et al*., our analyses of nine Clade 5 strains indicate that there are marked phenotypic differences in sporulation frequency within this clade. Thus, analyzing more strains within a clade may reveal fewer differences between clades. A caveat to our analyses is that *C. difficile* strains exhibit considerable genetic diversity even within a ribotype (20, 22), so our conclusions may only reflect the behavior of RT078 Clade 5 strains. It is also possible that analyzing more strains within this ribotype will reveal a broader spectrum of phenotypes than the somewhat bimodal ones we have reported with respect to cell length and chaining propensity (**Figures 2, 3**). Indeed, a recent analysis of a large number of RT027 strains indicates that strains within this ribotype can exhibit marked differences in virulence in mice (22) despite these strains traditionally being associated with "epidemic" outbreaks in hospital setting.

While addressing these questions will depend on analyzing a larger number of strains from each clade, an important outcome of our study is its development of a facile method for conducting time-lapse microscopy under anaerobic conditions. Although this method is a closed system, i.e. no additional reagents can be introduced once the bacteria are inoculated onto the pad in the anaerobic chamber, it could be coupled to recently developed anaerobic imaging tags based on fluorogen-activated proteins (49). Thus, there are many potential applications for this simple method for studying the growth of anaerobes over time at the single-cell level.

Materials and Methods

Bacterial strains and growth conditions

All *C. difficile* strains were grown on brain heart infusion (BHIS) medium supplemented with 0.5% w/v yeast extract and 0.1% w/v L-cysteine with taurocholate (TCA; 0.1% w/v; 1.9 mM). Strains were sub-cultured into tryptone yeast extract (TY) broth supplemented with 0.1% w/v L-cysteine (TYC medium) prior to inoculation onto the time-lapse microscopy agarose pads. All strains were grown at 37°C under anaerobic conditions using a gas mixture of 85% hydrogen, 5% CO2, and 10% H2. For time-lapse experiments, 1.5% agarose pads were used supplemented with TYC medium as described above. Sporulation analyses were carried out on 70:30 medium (70% BHIS and 30% SMC) for 24h as described previously (50).

Anaerobic time-lapse imaging of C. difficile growth

All imaging was carried out on a Leica DMi8 inverted microscope with a HC plan apochromat 63x 1.4 NA oil immersion phase contrast objective. Fluorescent membrane staining experiments were done with a Lumencor Spectra X light source, coupled with an XLED-QP quadruple-band dichroic beam-splitter (Leica) (transmission: 415, 470, 570, and 660 nm) along

with an external emission filter wheel (Leica). FM4-64 was excited using a 470nm LED through a 470/20nm excitation filter and emitted light was filtered through 590/50nm emission filter and captured with a Leica DFC9000GTC sCMOS camera. All experiments were carried out at 37°C using a microscope incubation system (Pecon), Leica Adaptive Focus Control hardware autofocus, and a high precision stage (Pecon) were used for all imaging experiments.

For time-lapse imaging of *C. difficile* growth all bacterial strains were grown in 2 mL liquid TY medium to a turbid OD600 > 2-3; after 2 hours of growth, bacteria were diluted 1:100 for Clade 5 strains and all other strains were diluted 1:50 in fresh media and grown to mid-log stage (OD600 0.4-0.7).

An imaging chamber with a gas-tight seal was constructed using a 125 µL Gene Frame (Thermo Fisher) adhered to a glass slide generating a well for growth medium. The slide was then transferred to the anaerobic chamber. In the anaerobic chamber, the gene frame was filled with 500µl 1.5% Top vision low melting point agarose and tryptone yeast extract media containing 0.1% w/v L-cysteine to scavenge oxygen and maintain anaerobic conditions. While the agarose was molten, a second clean slide was placed over the top and the agar pad was placed on a frozen small freezer block (for holding PCR strip tubes) for 10-30 minutes until the agarose-media mixture was solid. For experiments where FM4-64 was used, agarose pads were made the same way, with the addition of FM4-64 to a final concentration of 1 µg/mL directly to the agarose/media solution prior to making the agar pad.

The agar pad was dried for 5-10 minutes until visible liquid on the surface of the pad was evaporated. 1 µL of mid-log cells were spotted on the pad, dried, and a #1.5 coverslip (VWR) was adhered to the Gene Frame. The cells were imaged at 37°C until they reached confluency in the field of view. This was anywhere from 2.5 hours for Clade 5 strains to 6 hours for Clades 1-4 for all experiments with images taken at 5-minute intervals.

Image analysis, computing hardware, and statistical analysis.

All movie frames were trimmed to the point when cells were not overlapping and out of focus regions were cropped. The resulting images were analyzed using the Python library DeLTA 2.0 (39, 40). All image and data analyses were done on a PC running Windows 10 equipped with an AMD Ryzen 5900HX 8-core CPU, 32GB DDR4 RAM, 2 1TB NVME SSDs, and an NVIDIA RTX3080 GPU with 16GB VRAM. Analysis of the output data and data visualizations were done in Python using Matplotlib/Seaborn, Pandas, Numpy, Scipy, and the Statannotations library.

Septum detection

The image processing starts with the masks generated using DeLTA 2.0. First, we performed erosion on the mask with a disk of radius of 1 pixel to avoid effects of the membrane fluorescence. The photo was cropped following the eroded mask contour. Hence, only the interior of the contour was considered. Pixel values were rescaled such as the minimum pixel value of inner pixels was mapped to 0 and the maximum is mapped to 255. We set a threshold intensity as the pixel value of the 95% quantile of pixel intensities. The image was slightly blurred by convolution with a Gaussian filter with standard deviation of 2 pixels. The septum corresponds to the contours found after performing an adaptive thresholding based on thresholds defined using a Gaussian-weighted method. This thresholding was performed using the open cv2 library in Python. The resultant contours were manually verified from videos generated using the tracking from the outputs from DeLTA. The manual validation mainly involved filling in the time points where the algorithm

missed a ring that was visible in the video. We did not assume the existence of a ring if the algorithm did not detect it in a previous ,m,frame.

Cell length estimation

We measured the cell projected area from the DeLTA contours of the images as the pixel amount of the contours. However, estimating the cell length was challenging because some cells were very long and bent. To overcome this problem, we selected 30 images of three different cells (from strains 630, TAL3050 and TAL28131) that were straight and had different lengths. For these cells, we calculated the cell length as the longest side of the minimum bounding rectangle of the contour. From these lengths of straight cells, we also estimated the best cell width as the average of the projected area divided by the length. Considering the extreme cell length, the effects of the rounded tips were negligible and the rectangle shape adequately approximated length. We verifed that this mean value showed low variability for the three strains. Then, we used this width value to estimate the length of all the cells, including the bent ones, by dividing their projected area by the width. This way, we obtained a consistent measure of cell length that was independent of bending.

Elongation rate estimation

We tracked the cell size over time and identified the division points as the ones where the cell size (projected area) dropped by more than 30% compared to the current cell size value. We fitted an exponential function (with base e) of time to the data points between two divisions and estimated the elongation rate from the exponent of the best fit. We expressed the elongation rate in doublings/hr, which means how many times the cell size doubles in one hour. For example, an elongation rate of 2 doubling/hr means that the cell size doubles two times in 1 hr, which corresponds to an exponent of 2ln(2) 1/hr. For the statistics, we only included the elongation rates that had a high quality of fit, with an R2 coefficient greater than 0.9.

Bulk growth measurements. Starter cultures were grown until early stationary phase in BHIS (or TYC medium as indicated) then back-diluted 1:50 into BHIS (or TYC medium). When the cultures reached an OD_{600} of 0.5, they were back-diluted 1:25 into 200 µL of either BHIS or TYC in a flat 96 well polystyrene plate (CellTreat). The OD_{600} was analyzed every 15 min for 24 hrs in a BioTek Epoch plate reader with shaking. Bulk growth measurements are based on a minimum of three independent replicates across a minimum of 2 experiments.

Sporulation assays. Starter cultures were grown until early stationary phase in BHIS then back-diluted 1:50 into BHIS. When the cultures reached an OD_{600} between 0.35 and 0.75, 120 µL of the culture was spread onto 70:30 (70% SMC media and 30% BHIS media) agar plates (40 ml media per plate) and then incubated for 20-24 hrs before the sporulating cells were scraped from the plate into phosphate-buffered saline (PBS). Sporulation levels were visualized by phase-contrast microscopy as previously described (51).

Heat resistance assay. Heat-resistant spore formation was measured 20-24 hrs after sporulation was induced on 70:30 agar plates as previously described (50). The percent sporulation of given culture represents the ratio of heat-resistant colony-forming units (CFUs) to total CFUs. Percent sporulation was determined from a minimum of 3 biological replicates.

Spore purification. Spores were purified as previously described (52) by scraping up sporulating cells incubated on 70:30 medium for 3 days into ice-cold H_2O. The cells were

washed several times in ice-water over the course of a day and incubated on ice overnight. The following morning, the sample was pelleted, and cells were resuspended in 1 X DNAse buffer (New England Biolabs) and then treated with DNAse (New England Biolabs) for 30 min at 37˚C. The samples was washed one more time before being resuspended in 20% Histodenz and then layered onto a 50% Histodenz layer. The resulting mixture was pelleted and the supernatant was aspirated off using a vacuum aspirator. The pelleted spores were washed in ice-cold water 2-3 times and the optical density of the purified spores was measured.

Genomic DNA preparation. Starter cultures were grown until early stationary phase in BHIS then back-diluted 1:50 into BHIS and grown until an OD_{600} of around 0.7-0.8 was reached. 10 mL of the culture was pelleted and then frozen at -80˚C. After thawing the sample, it was resuspended in a 25% sucrose TE buffer (10 mM Tris, 1mM EDTA), incubated with 100 mg/mL lysozyme for 37˚C for 1 hr. After the cultures tarted to lyse, proteinase K, RNAse A, EDTA, Sarkosyl, and NaCl was added. Phenol:Chloroform:IAA (25:24:1) was added to extract proteins, gently mixed, and then the sample was pelleted to separate the phenol and aqueous layer. The aqueous layer was then added to Chloroform:IAA (24:1), mixed gently, then centrifuged. The aqueous layer was then precipitated using isopropanol and incubated at -20˚C for a minimum of 15 min. The precipitated DNA was pelleted and then washed with 70% ethanol. The pellet was air dried and then gently resuspended in 10 mM Tris pH 8.0 elution buffer.

Genomic analyses. Genomic DNA was sequenced by MiGS at the University of Pittsburgh (now SeqCenter) according to their standard protocol. Libraries were sequenced on an Illumina NextSeq 500 platform to generate paired-end 150 bp reads. Illumina reads of RT078 genomes were assembled into contigs using SPAdes (v3.13.0),(53) and genes were called and annotated using Prokka (v1.11).(54) Assembled and annotated contigs of five RT078 strains (TAL28131, TAL29600, TAL29996, TAL30550, TAL30574) were applied for pangenomic analysis. Default settings were used based on the Anvi'o workflow for microbial pangenomcis with adjustments for minbit as 0.5 and mcl-inflation as 10.(55-57) For SNPs analyses, reads of five RT078 genomes were aligned to the reference M120 and variants were called by breseq (v. 0.38.1) by default settings.(58)

qPCR analyses. Each genomic DNA sample was analyzed by qPCR with primers that amplify the *cmr*-ON sequence orientation, the *cmr*-OFF orientation, or the reference gene *rpoA* (31) . Each 20-µL qPCR reaction consisted of 100 ng genomic DNA, 100 nM primers, and SensiMix™ SYBR® reagents (Bioline). The reactions were run on a LightCycler® 96 (Roche Diagnostics), and *cmr* switch orientation frequencies were calculated as described previously (29).

Surface motility assays

Starter cultures were grown until early stationary phase in BHIS then back-diluted 1:50 into BHIS and grown until an OD_{600} of 0.5 was reached. 10 µL of the exponential-phase cultures were then spotted onto BHIS plates and incubated at 37˚C for 5 days after which the plates were scanned using a flatbed scanner.

Author contribution statement

J.W.R. conducted all live-cell microscopy experiments and analyzed the single-cell data. C.A.N.A. generated data visualizations from the time-lapse microscopy data, developed the

method for identifying septa within chains that had not initiated cell separation, and calculated elongation rates and cell lengths. I.N. and A.S. conducted sporulation analyses and liquid culture experiments. J.W.R., R.T., and A.S conceptualized all experiments with input from . Q.D. and A.S extracted genomic DNA and visualized cells at different growth phases. Q.D. conducted all the genomic analyses and genome comparisons. N.D. and I.N. conducted spore purification analyses and imaging. B.A. and IE developed and optimized methods for constructing agarose pads that could support anaerobic growth in an ambient atmosphere. M.J.D. developed the DeLTA 2.0 analysis pipeline. A.M. conducted qPCRs to determine *cmr* switch orientation in Clade 5 isolates. Writing was by J.W.R and A.S with input from all the authors.

Table 1. *Clostridioides difficile* clinical isolates used in this study

Strain Name	Clade	Ribotype	ST group	Source	Reference
630	1	012	2	Zurich, 1982 (Sanger Institute)	(59)
R20291	2	027	1	London, 2006 (Sanger Institute)	(60)
M68	4	017	81	Dublin, 2006 (Sanger Institute)	(45)
M120	5	078	11	UK, 2007 (Sanger Institute)	(45)
TAL28131	5	078	11	Tufts Medical Center	This paper
TAL29600	5	078	11	Tufts Medical Center	This paper
TAL29996	5	078	11	Tufts Medical Center	This paper
TAL30550	5	078	11	Tufts Medical Center	This paper
TAL30574	5	078	11	Tufts Medical Center	This paper
V48	5	078	11	Brigham & Women's Hospital	This paper
139b	5	078	11	Memorial Sloan Kettering	(20)
WU66	5	078	11	Barnes-Jewish Hospital	(20)
BBL2	1	012	2	Memorial Sloan Kettering	(20)
WU38	1	012	2	Barnes-Jewish Hospital	(20)
190B	1	087	46	Memorial Sloan Kettering	(20)
Wup14	2	027	1	Barnes-Jewish Hospital	(20)
BBL4	2	027	1	Memorial Sloan Kettering	(20)
186A	2	027	1	Memorial Sloan Kettering	(20)

Memorial Sloan Kettering Cancer Center (MSK), Barnes-Jewish Hospital (BJH)

References

1. Desai K, Gupta SB, Dubberke ER, Prabhu VS, Browne C, Mast TC. 2016. Epidemiological and economic burden of Clostridium difficile in the United States: estimates from a modeling approach. BMC Infect Dis 16:303.
2. Guh AY, Mu Y, Winston LG, Johnston H, Olson D, Farley MM, Wilson LE, Holzbauer SM, Phipps EC, Dumyati GK, Beldavs ZG, Kainer MA, Karlsson M, Gerding DN, McDonald LC, Emerging Infections Program Clostridioides difficile Infection Working G. 2020. Trends in U.S. Burden of Clostridioides difficile Infection and Outcomes. N Engl J Med 382:1320-1330.

3. Shen A. 2020. Clostridioides difficile Spore Formation and Germination: New Insights and Opportunities for Intervention. Annu Rev Microbiol 74:545-566.
4. Abt MC, McKenney PT, Pamer EG. 2016. Clostridium difficile colitis: pathogenesis and host defence. Nat Rev Microbiol doi:10.1038/nrmicro.2016.108.
5. Theriot CM, Young VB. 2015. Interactions Between the Gastrointestinal Microbiome and Clostridium difficile. Annu Rev Microbiol 69:445-61.
6. Kordus SL, Thomas AK, Lacy DB. 2022. Clostridioides difficile toxins: mechanisms of action and antitoxin therapeutics. Nat Rev Microbiol 20:285-298.
7. Maroo S, Lamont J. 2006. Recurrent Clostridium difficile. Gastroenterology 130:1311-1316.
8. Zanella Terrier MC, Simonet ML, Bichard P, Frossard JL. 2014. Recurrent Clostridium difficile infections: the importance of the intestinal microbiota. World J Gastroenterol 20:7416-23.
9. Knight DR, Elliott B, Chang BJ, Perkins TT, Riley TV. 2015. Diversity and Evolution in the Genome of Clostridium difficile. Clin Microbiol Rev 28:721-41.
10. Knight DR, Imwattana K, Kullin B, Guerrero-Araya E, Paredes-Sabja D, Didelot X, Dingle KE, Eyre DW, Rodriguez C, Riley TV. 2021. Major genetic discontinuity and novel toxigenic species in Clostridioides difficile taxonomy. Elife 10.
11. Norsigian CJ, Danhof HA, Brand CK, Midani FS, Broddrick JT, Savidge TC, Britton RA, Palsson BO, Spinler JK, Monk JM. 2022. Systems biology approach to functionally assess the Clostridioides difficile pangenome reveals genetic diversity with discriminatory power. Proc Natl Acad Sci U S A 119:e2119396119.
12. Knight DR, Kullin B, Androga GO, Barbut F, Eckert C, Johnson S, Spigaglia P, Tateda K, Tsai PJ, Riley TV. 2019. Evolutionary and Genomic Insights into Clostridioides difficile Sequence Type 11: a Diverse Zoonotic and Antimicrobial-Resistant Lineage of Global One Health Importance. mBio 10.
13. Seth-Smith HMB, Biggel M, Roloff T, Hinic V, Bodmer T, Risch M, Casanova C, Widmer A, Sommerstein R, Marschall J, Tschudin-Sutter S, Egli A. 2021. Transition From PCR-Ribotyping to Whole Genome Sequencing Based Typing of Clostridioides difficile. Front Cell Infect Microbiol 11:681518.
14. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114.
15. Dingle KE, Elliott B, Robinson E, Griffiths D, Eyre DW, Stoesser N, Vaughan A, Golubchik T, Fawley WN, Wilcox MH, Peto TE, Walker AS, Riley TV, Crook DW, Didelot X. 2014. Evolutionary history of the Clostridium difficile pathogenicity locus. Genome Biol Evol 6:36-52.
16. Mengoli M, Barone M, Fabbrini M, D'Amico F, Brigidi P, Turroni S. 2022. Make It Less difficile: Understanding Genetic Evolution and Global Spread of Clostridioides difficile. Genes (Basel) 13.
17. Endres BT, Begum K, Sun H, Walk ST, Memariani A, Lancaster C, Gonzales-Luna AJ, Dotson KM, Basseres E, Offiong C, Tupy S, Kuper K, Septimus E, Arafat R, Alam MJ, Zhao Z, Hurdle JG, Savidge TC, Garey KW. 2019. Epidemic Clostridioides difficile Ribotype 027 Lineages: Comparisons of Texas Versus Worldwide Strains. Open Forum Infect Dis 6:ofz013.

18. McDonald LC, Killgore GE, Thompson A, Owens RC, Jr., Kazakova SV, Sambol SP, Johnson S, Gerding DN. 2005. An epidemic, toxin gene-variant strain of Clostridium difficile. N Engl J Med 353:2433-41.
19. Clements AC, Magalhaes RJ, Tatem AJ, Paterson DL, Riley TV. 2010. Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. Lancet Infect Dis 10:395-404.
20. Lewis BB, Carter RA, Ling L, Leiner I, Taur Y, Kamboj M, Dubberke ER, Xavier J, Pamer EG. 2017. Pathogenicity Locus, Core Genome, and Accessory Gene Contributions to Clostridium difficile Virulence. mBio 8.
21. Merrigan M, Venugopal A, Mallozzi M, Roxas B, Viswanathan VK, Johnson S, Gerding DN, Vedantam G. 2010. Human hypervirulent Clostridium difficile strains exhibit increased sporulation as well as robust toxin production. J Bacteriol 192:4904-11.
22. Dong Q, Lin H, Allen MM, Garneau JR, Sia JK, Smith RC, Haro F, McMillen T, Pope RL, Metcalfe C, Burgo V, Woodson C, Dylla N, Kohout C, Sundararajan A, Snitkin ES, Young VB, Fortier LC, Kamboj M, Pamer EG. 2023. Virulence and genomic diversity among clinical isolates of ST1 (BI/NAP1/027) Clostridioides difficile. Cell Rep 42:112861.
23. Shaw HA, Preston MD, Vendrik KEW, Cairns MD, Browne HP, Stabler RA, Crobach MJT, Corver J, Pituch H, Ingebretsen A, Pirmohamed M, Faulds-Pain A, Valiente E, Lawley TD, Fairweather NF, Kuijper EJ, Wren BW. 2020. The recent emergence of a highly related virulent Clostridium difficile clade with unique characteristics. Clin Microbiol Infect 26:492-498.
24. Li C, Harmanus C, Zhu D, Meng X, Wang S, Duan J, Liu S, Fu C, Zhou P, Liu R, Wu A, Kuijper EJ, Smits WK, Fu L, Sun X. 2018. Characterization of the virulence of a non-RT027, non-RT078 and binary toxin-positive Clostridium difficile strain associated with severe diarrhea. Emerg Microbes Infect 7:211.
25. Imwattana K, Rodriguez C, Riley TV, Knight DR. 2021. A species-wide genetic atlas of antimicrobial resistance in Clostridioides difficile. Microb Genom 7.
26. Kumar N, Browne HP, Viciani E, Forster SC, Clare S, Harcourt K, Stares MD, Dougan G, Fairley DJ, Roberts P, Pirmohamed M, Clokie MRJ, Jensen MBF, Hargreaves KR, Ip M, Wieler LH, Seyboldt C, Noren T, Riley TV, Kuijper EJ, Wren BW, Lawley TD. 2019. Adaptation of host transmission cycle during Clostridium difficile speciation. Nat Genet 51:1315-1320.
27. Knetsch CW, Kumar N, Forster SC, Connor TR, Browne HP, Harmanus C, Sanders IM, Harris SR, Turner L, Morris T, Perry M, Miyajima F, Roberts P, Pirmohamed M, Songer JG, Weese JS, Indra A, Corver J, Rupnik M, Wren BW, Riley TV, Kuijper EJ, Lawley TD. 2018. Zoonotic Transfer of Clostridium difficile Harboring Antimicrobial Resistance between Farm Animals and Humans. J Clin Microbiol 56.
28. Lim SC, Knight DR, Riley TV. 2020. Clostridium difficile and One Health. Clin Microbiol Infect 26:857-863.
29. Garrett EM, Mehra A, Sekulovic O, Tamayo R. 2021. Multiple Regulatory Mechanisms Control the Production of CmrRST, an Atypical Signal Transduction System in Clostridioides difficile. mBio 13:e0296921.
30. Thanissery R, Winston JA, Theriot CM. 2017. Inhibition of spore germination, growth, and toxin activity of clinically relevant C. difficile strains by gut microbiota derived secondary bile acids. Anaerobe 45:86-100.

31. Garrett EM, Sekulovic O, Wetzel D, Jones JB, Edwards AN, Vargas-Cuebas G, McBride SM, Tamayo R. 2019. Phase variation of a signal transduction system controls Clostridioides difficile colony morphology, motility, and virulence. PLoS Biol 17:e3000379.
32. Ackermann M. 2015. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol 13:497-508.
33. Aldridge BB, Fernandez-Suarez M, Heller D, Ambravaneswaran V, Irimia D, Toner M, Fortune SM. 2012. Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility. Science 335:100-4.
34. Anjuwon-Foster BR, Tamayo R. 2017. A genetic switch controls the production of flagella and toxins in Clostridium difficile. PLoS Genet 13:e1006701.
35. Anjuwon-Foster BR, Tamayo R. 2018. Phase variation of Clostridium difficile virulence factors. Gut Microbes 9:76-83.
36. Kint N, Janoir C, Monot M, Hoys S, Soutourina O, Dupuy B, Martin-Verstraete I. 2017. The alternative sigma factor sigma(B) plays a crucial role in adaptive strategies of Clostridium difficile during gut infection. Environ Microbiol 19:1933-1958.
37. Courson DS, Pokhrel A, Scott C, Madrill M, Rinehold AJ, Tamayo R, Cheney RE, Purcell EB. 2019. Single cell analysis of nutrient regulation of Clostridioides (Clostridium) difficile motility. Anaerobe 59:205-211.
38. de Jong IG, Beilharz K, Kuipers OP, Veening JW. 2011. Live Cell Imaging of Bacillus subtilis and Streptococcus pneumoniae using Automated Time-lapse Microscopy. J Vis Exp doi:10.3791/3145.
39. Lugagne JB, Lin H, Dunlop MJ. 2020. DeLTA: Automated cell segmentation, tracking, and lineage reconstruction using deep learning. PLoS Comput Biol 16:e1007673.
40. O'Connor OM, Alnahhas RN, Lugagne JB, Dunlop MJ. 2022. DeLTA 2.0: A deep learning pipeline for quantifying single-cell spatial and temporal dynamics. PLoS Comput Biol 18:e1009797.
41. Sauls JT, Cox SE, Do Q, Castillo V, Ghulam-Jelani Z, Jun S. 2019. Control of Bacillus subtilis Replication Initiation during Physiological Transitions and Perturbations. mBio 10.
42. Narula J, Kuchina A, Zhang F, Fujita M, Suel GM, Igoshin OA. 2016. Slowdown of growth controls cellular differentiation. Mol Syst Biol 12:871.
43. Eswaramoorthy P, Duan D, Dinh J, Dravis A, Devi SN, Fujita M. 2010. The threshold level of the sensor histidine kinase KinA governs entry into sporulation in Bacillus subtilis. J Bacteriol 192:3870-82.
44. Deatherage DE, Traverse CC, Wolf LN, Barrick JE. 2014. Detecting rare structural variation in evolving microbial populations from new sequence junctions using breseq. Front Genet 5:468.
45. He M, Sebaihia M, Lawley T, Stabler R, Dawson L, Martin M, Holt K, Seth-Smith H, Quail M, Rance R, Brooks K, Churcher C, Harris D, Bentley S, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren B, Dougan G, Parkhill J. 2010. Evolutionary dynamics of Clostridium difficile over short and long time scales. Proceedings of the National Academy of Sciences of the United States of America 107:7527-7532.

46. Sekulovic O, Mathias Garrett E, Bourgeois J, Tamayo R, Shen A, Camilli A. 2018. Genome-wide detection of conservative site-specific recombination in bacteria. PLoS Genet 14:e1007332.
47. Dhungel BA, Govind R. 2021. Phase-variable expression of pdcB, a phosphodiesterase, influences sporulation in Clostridioides difficile. Mol Microbiol 116:1347-1360.
48. Edwards AN, Willams CL, Pareek N, McBride SM, Tamayo R. 2021. c-di-GMP Inhibits Early Sporulation in Clostridioides difficile. mSphere 6:e0091921.
49. Streett H, Charubin K, Papoutsakis ET. 2021. Anaerobic fluorescent reporters for cell identification, microbial cell biology and high-throughput screening of microbiota and genomic libraries. Curr Opin Biotechnol 71:151-163.
50. Shen A, Fimlaid KA, Pishdadian K. 2016. Inducing and Quantifying Clostridium difficile Spore Formation. Methods Mol Biol 1476:129-42.
51. Pishdadian K, Fimlaid KA, Shen A. 2015. SpoIIID-mediated regulation of sigma(K) function during Clostridium difficile sporulation. Mol Microbiol 95:189-208.
52. Fimlaid KA, Jensen O, Donnelly ML, Francis MB, Sorg JA, Shen A. 2015. Identification of a Novel Lipoprotein Regulator of Clostridium difficile Spore Germination. PLoS Pathogens doi:10.1371/journal.ppat.1005239.
53. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes De Novo Assembler. Curr Protoc Bioinformatics 70:e102.
54. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-9.
55. Benedict MN, Henriksen JR, Metcalf WW, Whitaker RJ, Price ND. 2014. ITEP: an integrated toolkit for exploration of microbial pan-genomes. BMC Genomics 15:8.
56. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen OC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjodin A, Scott JJ, Vazquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 6:3-6.
57. van Dongen S, Abreu-Goodger C. 2012. Using MCL to extract clusters from networks. Methods Mol Biol 804:281-95.
58. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol 1151:165-88.
59. Sebaihia M, Wren B, Mullany P, Fairweather N, Minton N, Stabler R, Thomson N, Roberts A, Cerdeño-Târraga A, Wang H, Holden M, Wright A, Churcher C, Quail M, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B, Parkhill J. 2006. The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. Nature genetics 38:779-786.
60. Stabler R, He M, Dawson L, Martin M, Valiente E, Corton C, Lawley T, Sebaihia M, Quail M, Rose G, Gerding D, Gibert M, Popoff M, Parkhill J, Dougan G, Wren B. 2009. Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome biology 10.