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.

43 **Introduction**

44

45 *Clostridioides difficile* is a leading cause of nosocomial infection in the United
46 States, with approximately 500,000 new infections and 14,000 deaths attributed to this
47 organism annually (1, 2). As an obligate anaerobe, *C. difficile* relies on its ability to form hardy,
48 metabolically dormant spores to survive outside the gut environment and transmit disease (3).
49 When *C. difficile* spores are ingested by susceptible hosts (4, 5), they encounter bile salt signals
50 in the small intestine that trigger their germination. These stimuli allow them to exit dormancy
51 and outgrow into vegetative cells, which colonize the colon and secrete toxins that damage gut
52 epithelial tissue (6). This damage triggers a significant inflammatory response that can cause
53 disease pathologies ranging from mild diarrhea to pseudomembranous colitis, and even death
54 (2, 4). Beyond these acute complications, *C. difficile* frequently causes recurrent infections
55 (~20% of infections), which can lead to more severe disease symptoms (14–16) (2, 7, 8).
56 Notably, as a species, *C. difficile* exhibits tremendous genetic diversity (9-11). Genomic
57 analyses indicate that *C. difficile*'s core genome represents only ~10-20% of its pan-genome (9,
58 10). The extreme plasticity of the "open" pan-genome of *C. difficile* likely helps the bacterium
59 inhabit the gastrointestinal tract of diverse animals, from mammals to invertebrates, and
60 environmental reservoirs like sewage and compost (12). Multi-locus sequence typing (MLST)
61 analyses have revealed that the *C. difficile* species comprises 5 distinct phylogenetic clades that
62 can be further subdivided into different ribotypes (RTs) or sequence types (STs) (10, 13). While
63 whole-genome sequencing analyses identified three additional cryptic clades, subsequent
64 analyses have revealed that the genomic divergence between the 5 main clades and three
65 additional cryptic clades is so significant that the cryptic clade species fall below the threshold
66 determined by average nucleotide identity (ANI) to qualify as part of the *C. difficile* species (10,
67 14).
68 Between the five remaining phylogenetic clades, there are notable differences in
69 geographic and host distributions. Clade 1 is the largest and most heterogeneous (and likely
70 ancient) clade, containing over 200 STs that have the broadest geographic distribution (15, 16).
71 It includes both toxin-producing and non-toxigenic strains, including the well-characterized,
72 genetically tractable toxigenic strain 630 (9). Clade 2 harbors epidemic-associated strains found
73 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).

105 Despite numerous genomic analyses revealing the remarkable genetic diversity of *C.*
106 *difficile* strains, relatively few studies have investigated the phenotypic diversity between strains
107 from different clades. Clade-specific differences in colony morphology between Clade 5 strains
108 relative to Clade 1-4 strains have been described in a limited set of analyses (26, 29), suggesting
109 that differences in growth and/or cellular morphology may exist within clades. While
110 differences in bulk growth rates between *C. difficile* strains have been reported (30), most
111 phenotypic analyses have been conducted on a limited subset of strains within a given clade.
112 Indeed, no systematic comparisons of the growth properties of the different clades or
113 comparisons of cell morphology have been performed to date.
114 Here, we compare the growth properties of multiple strains derived from the four
115 predominant main phylogenetic clades of *C. difficile* (Clade 1, 2, 4, and 5) using anaerobic
116 time-lapse microscopy. These analyses unexpectedly revealed striking differences in the growth
117 and cell morphology of the Clade 5 lineage relative to strains from Clades 1, 2, and 4. These
118 morphological differences correlate with the orientation of the *cmr* switch, which has previously
119 been implicated in modulating cell and colony morphology and cell chaining (31). Thus, the
120 *cmr* switch appears to be an important contributor to the chaining phenotype frequently
121 observed in Clade 5 strains.
122
123 **Results**
124
125 **Development of a simple method for time-lapse imaging under anaerobic conditions**
126
127 Time-lapse imaging of single cells has been widely used in bacteria due to the growing
128 appreciation that clonal bacterial populations can exhibit considerable phenotypic heterogeneity,
129 with implications for antibiotic resistance and virulence (31-35). However, live single-cell
130 analyses in *C. difficile* have been complicated by its inability to grow in the presence of
131 atmospheric oxygen (36). While there have been limited reports of live cell time-lapse
132 microscopy of *C. difficile*, these analyses have required custom growth chambers to maintain
133 anaerobic conditions (37), which may limit the accessibility of these experimental systems to
134 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.

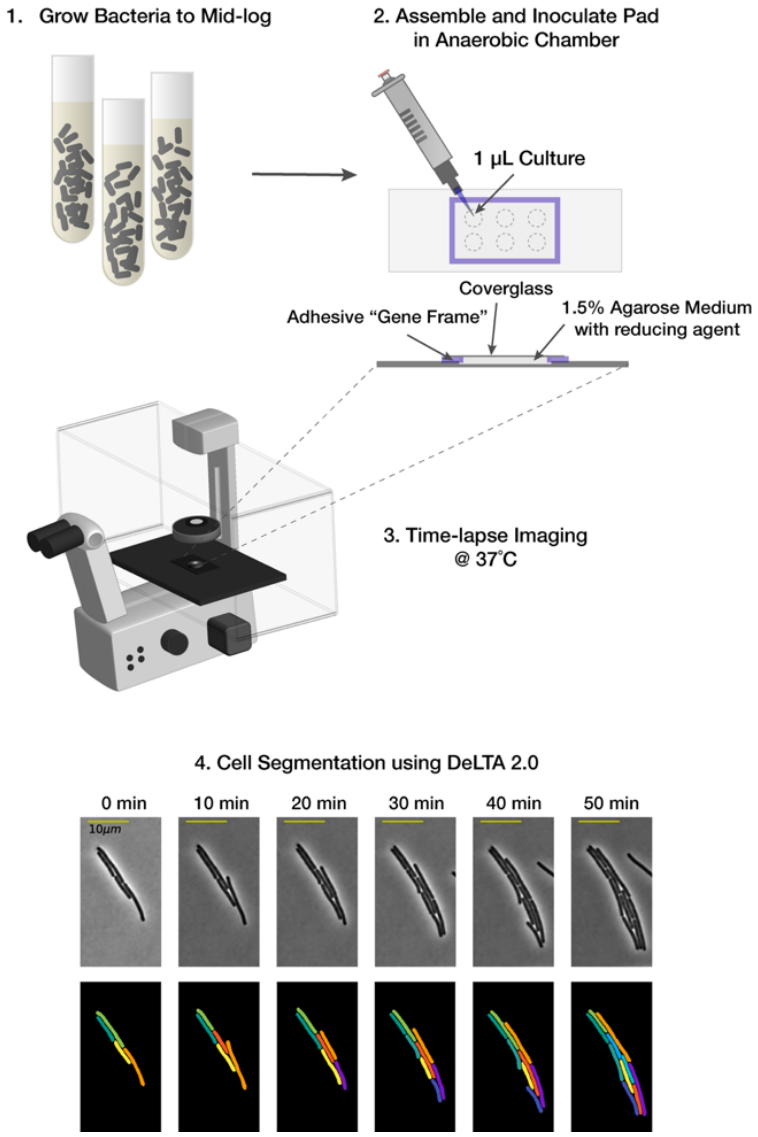
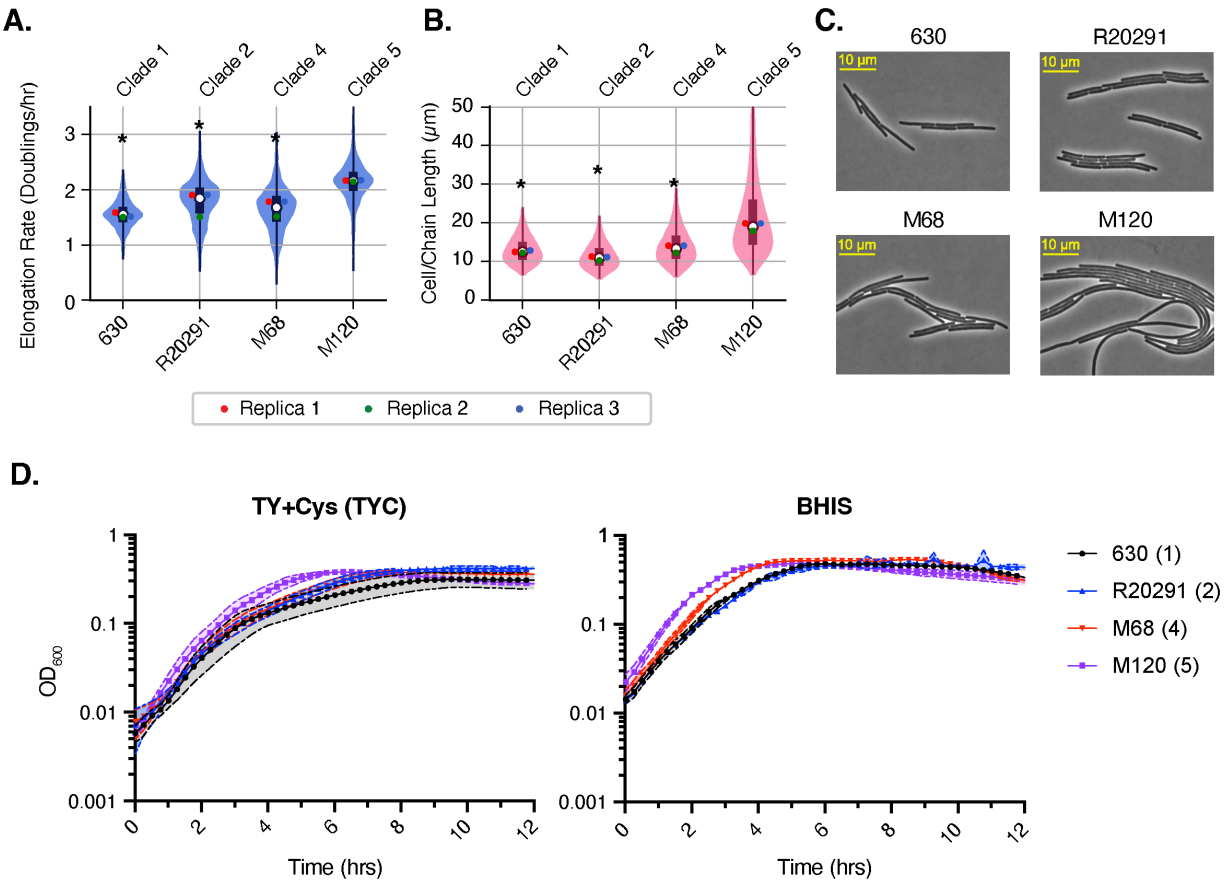


Figure 1. Schematic of the anaerobic single-cell imaging set-up. Exponentially growing *C. difficile* cells in TY medium supplemented with cysteine (TYC) are spotted onto a 1.5% agarose pads formed

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

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



185
186
187 **Figure 2. Clade 5 strain M120 elongates more quickly and exhibits cell chaining.** A. Violin plot of
188 the elongation rates measured during time-lapse microscopy analyses of strains 630 (Clade 1), R20291
189 (Clade 2), M68 (Clade 4), M120 (Clade 5) grown on TY supplemented with cysteine (TYC) agar. Data
190 are from three biological replicates, with the mean of each replicate shown as a point on the violin. B.
191 Violin plot of the cell length measured during time-lapse microscopy for strains 630 (Clade 1), R20291
192 (Clade 2), M68 (Clade 4), M120 (Clade 5). Each replicate mean is shown as a point on the violin.
193 Statistical significance for A and B was determined by comparing the mean of the three replicates of
194 Clade 1, 2, and 4 strains to the Clade 5 M120 strain using a Kruskal-Wallis test * $p < 0.05$. C. Phase-
195 contrast image from time-lapse microscopy movies. Scale bar is 10 μm. D. Optical density-based
196 analyses of bulk population growth of the indicated strains in TYC or BHIS medium. The number in
197 brackets indicates the clade to which a given strain belongs.
198

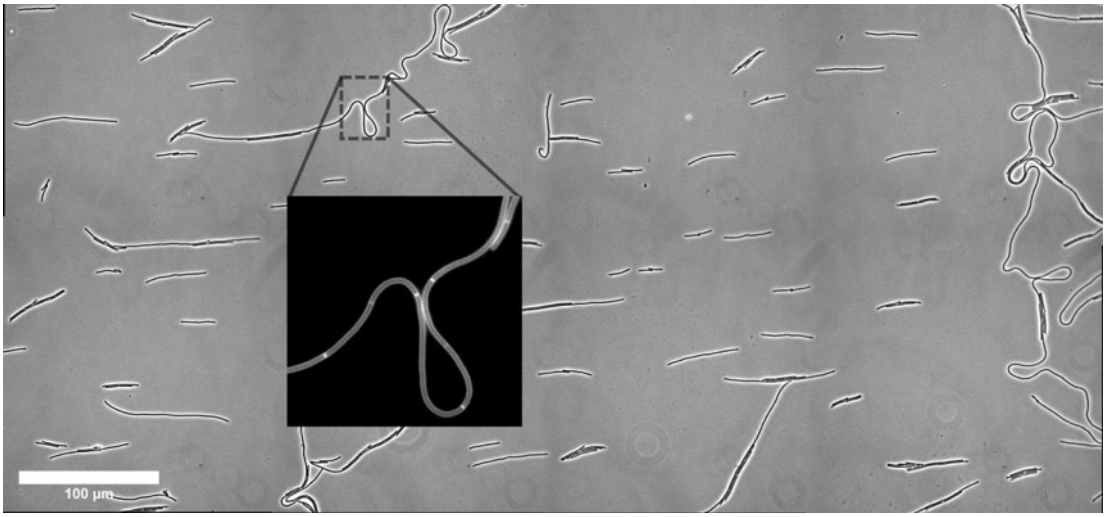
199 Importantly, the differences in single-cell elongation rates measured for the four strains
200 using DeLTA were also observed in bulk population analyses of their growth using optical

201 density in TYC and BHIS media (**Figure 2D**). These analyses confirmed that the Clade 5 strain
202 M120 grows faster (using optical density as a proxy for bacterial growth) than the Clade 1, 2,
203 and 4 strains analyzed. Negligible differences were observed in the elongation rates for Clade 1,
204 2, and 4 strains in these bulk population analyses (**Figure 2A**), consistent with the relatively
205 subtle difference observed in the single-cell elongation rates measured for the Clade 2 R20291
206 strain relative to the Clades 1 and 4 strains.
207 The Clade 5 strain M120 exhibited another distinct growth property from the strains
208 from Clades 1, 2, and 4. While strains from Clades 1, 2, and 4 produced cells of similar length
209 prior to division, with an average cell size of ~13 µm prior to division, cells of Clade 5 strain
210 M120 were significantly longer, appearing as long filaments (or chains) with an average length
211 of 22 µm but chains close to 50 µm readily being measured (**Figure 2B**). These filaments or
212 chains also appeared to bend readily (**Figure 2B**). Notably, to visualize an entire filament (or
213 chain), it was necessary to stitch together several fields of view, revealing that the bacterial
214 filaments (or chains) approached several hundred microns and even up to ~1 mm in length
215 (**Figure 3**). To determine whether strain M120 forms filaments, i.e. long cells that lack division
216 septa, or chains of cells, i.e. cells that remain attached despite forming division septa, we
217 incorporated the lipophilic fluorescent dye FM4-64 which visualizes membranes (Pogliano
218 1999), directly into the agarose pads. This stain revealed division septa throughout the length of
219 the cell, indicating that the extremely long cells were not filaments but instead extended chains
220 of cells (**Figure 3, inset**). Since the spacing between division septa generated cells that were
221 relatively similar in length to the other *C. difficile* strains analyzed, Clade 5 strain M120 is
222 distinguished by a decreased efficiency in cell separation.
223

224 **Clade 5 clinical isolates typically form long chains and grow more quickly than strains**
225 **from other clades**
226

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

232 (**Table 1**) using time-lapse microscopy analyses (**Figure 2**). Clade 1 strain 630 was included as
233 a control since it does not form chains under these conditions.
234



235
236
237 **Figure 3. *C. difficile* Clade 5 can form large heterogenous chains.** Large mosaic phase-contrast
238 image of the Clade 5 *C. difficile* strain M120. Zoomed image shows chains revealed by staining with
239 the membrane dye FM4-64. Image was stitched from 8 individual fields of view at 63x magnification.
240 Scale bar is 100 µm.
241

242 These analyses revealed that all but one of the Clade 5 strains tested formed long chains,
243 with TAL29600 forming the longest chains (29 µm on average, **Figure 4A-C**). In contrast,
244 strain TAL29996 formed shorter chains that were comparable in length to those observed for
245 strains from Clades 1, 2, and 4 (12-13 µm, **Figures 2B, 4C**). Notably, DeLTA over-estimated
246 cell length in Clade 5 strains that form chains because it could not identify cells that had formed
247 division septa (based on FM4-64 staining) but had not yet initiated cell separation.
248 To overcome this limitation and accurately quantify cell length within long chains,
249 namely the distance between division septa, we modified our image processing pipeline to use a
250 thresholding method to detect division septa. After generating masks in DeLTA to segment the
251 chains, we modified the mask so that only the interior of the contour was analyzed. We then
252 applied an adaptive thresholding method to identify division septa based on their elevated
253 fluorescence relative to the long axis of the cell; the thresholds were defined using a Gaussian-
254 weighted method (**Figure 4A**, red dots). These analyses revealed that Clade 5 strains that form

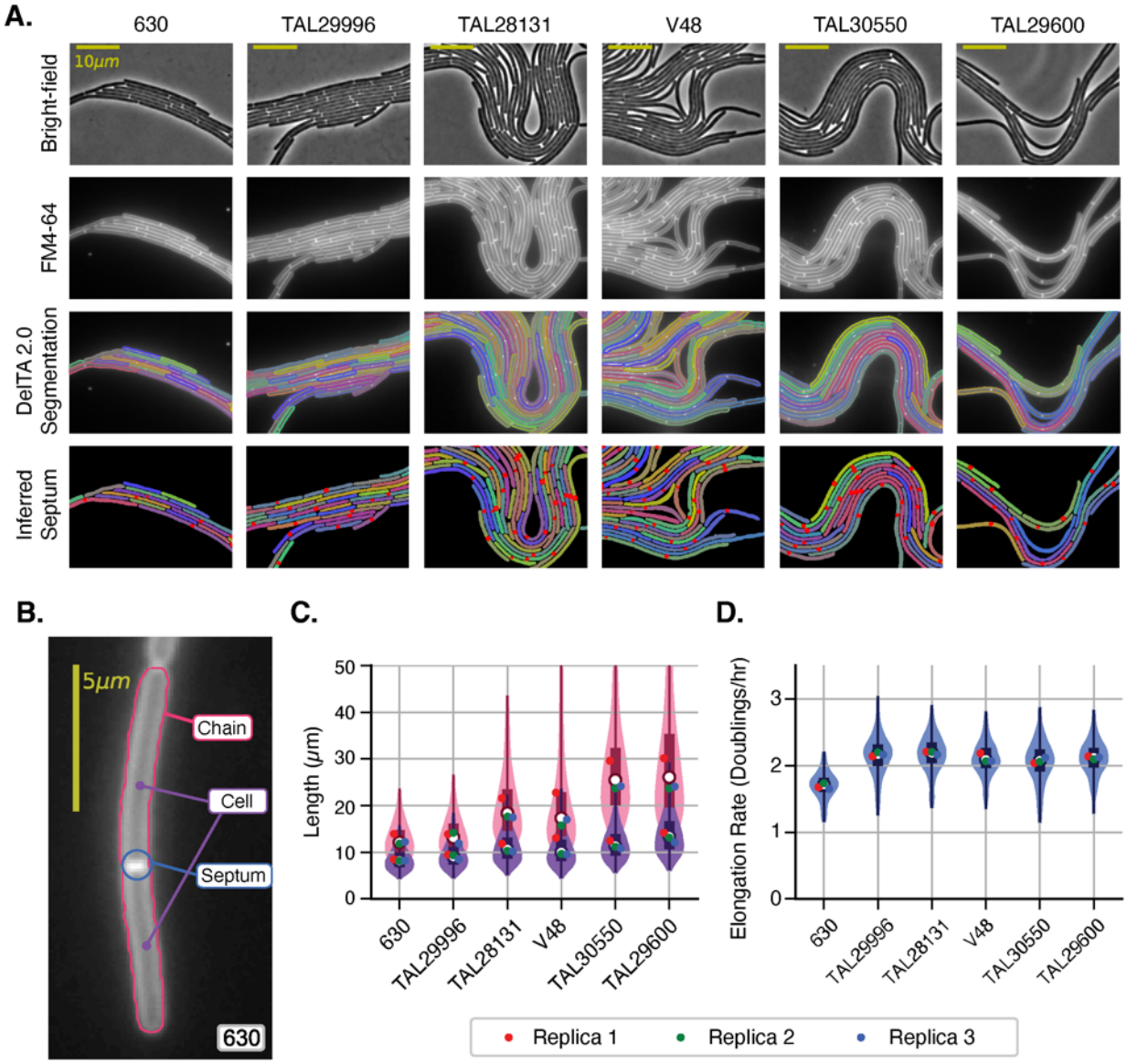


Figure 4. Faster growth is a common feature of Clade 5 strains, but the chaining phenotype is not fully penetrant in Clade 5 strains. (A) Phase-contrast microscopy images from time-lapse microscopy studies of Clade 5 strains; Clade 1 strain 630 is included for comparison. Fluorescence microscopy was 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 * $p < 0.05$.

271 long chains do not produce cells that are dramatically longer than strains that do not
272 form chains; cells within the Clade 5 strains that produced the longest chains had average cell
273 lengths (TAL29600 and TAL30550, 14 and 12 µm, respectively) that were 30-50% longer than
274 strains that do not form chains like strains 630 (Clade 1, 8.5 µm) and TAL29996 (Clade 5, 9.4
275 µm). Notably, even though TAL29996 did not form long chains like the other Clade 5 strains
276 analyzed, it elongated at rates similar to those of other Clade 5 strains (2.1 doublings/hr vs. 1.7
277 doublings/hr for Clade 1 strain 630**, Figure 4D**). Taken together, these analyses strongly
278 suggest that Clade 5 strains grow faster than Clade 1, 2, and 4 strains and are more likely to
279 form chains presumably because their cell separation mechanisms are less efficient.
280 While the long chains formed by Clade 5 strains limited the number of divisions that
281 could be analyzed within a given field using time-lapse microscopy, the faster elongation rates
282 measured on the agarose pads were also observed in bulk analyses of growth in broth culture.
283 These optical density-based analyses revealed that Clade 5 strains all grow faster than Clade 1,
284 2, and 4 strains irrespective of their ability to form chains or the growth media used (**Figure**
285 **S1**).
286 To assess whether these conclusions could be generalized to additional Clade 5 strains
287 vs. Clade 1, 2, and 4 strains, we analyzed the growth of additional strains from all four clades in
288 broth culture. These analyses confirmed that Clade 1, 2, and 4 strains grow at similar rates,
289 which are slower than those measured for the nine Clade 5 strains analyzed in both TYC and
290 BHIS media (**Figure S1**).
291
292 **Cell chaining in Clade 5 strains is not dependent on growth on a solid medium.**
293
294 *C. difficile* has previously been shown to promote cell elongation and chain formation
295 upon induction of the *cmrRST* locus (31), which encodes a signal transduction system.
296 Expression of this locus is also responsive to c-di-GMP levels (29), which increases in cells
297 grown on solid surfaces such as in a biofilm or on an agar plate (29). To determine whether the
298 chaining phenotype observed in most Clade 5 strains is induced by growth on a surface, we
299 assessed whether cell chaining would be observed in cells grown in broth culture. The Clade 5
300 strains along with strains from other clades were grown to log-phase in rich media, and their
301 ability to form chains was assessed by visualizing division septa with the fluorescent D-amino

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

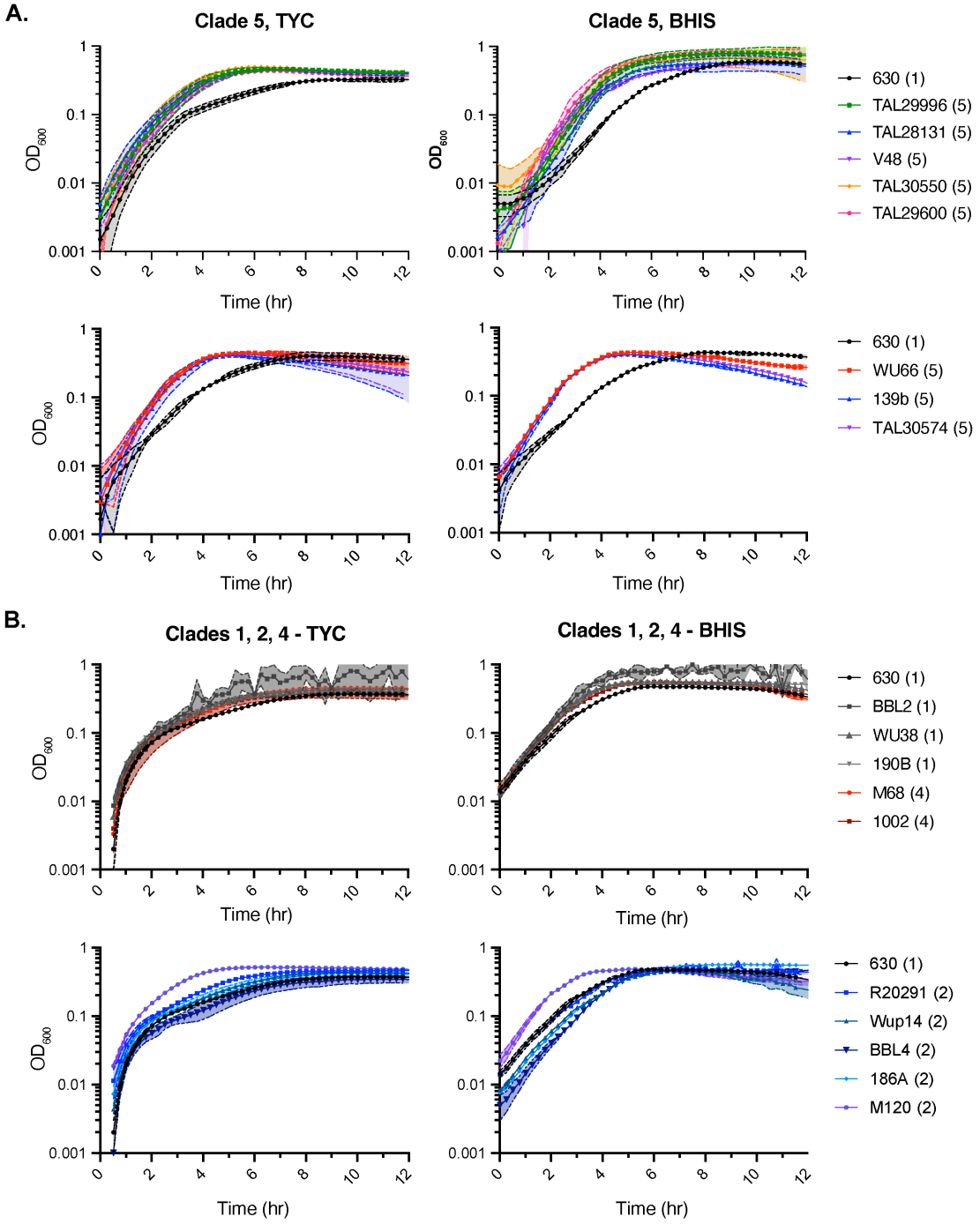
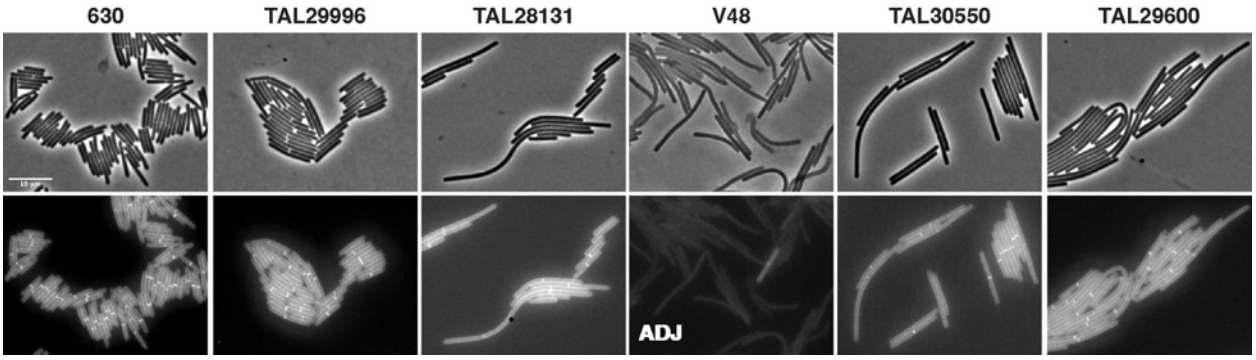
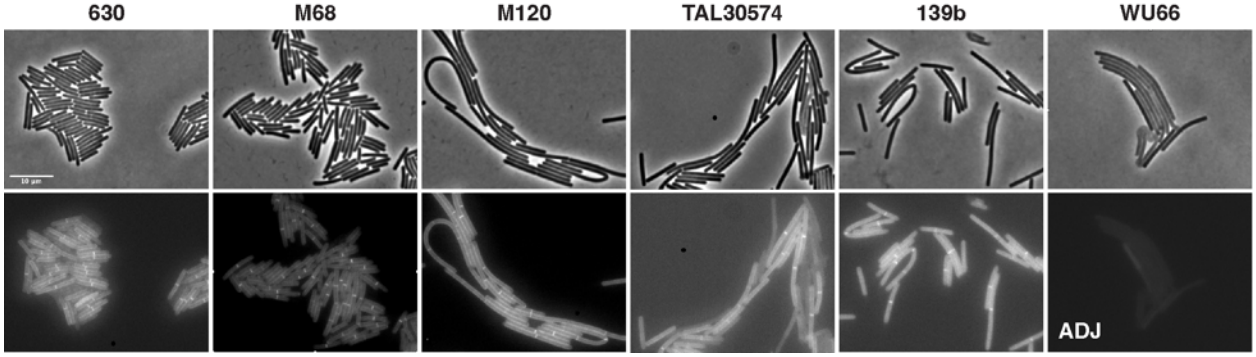


Figure S1. Optical density-based analyses of bulk population growth of the indicated strains during growth in TYC or BHIS medium. The number in brackets indicates the clade to which a given strain belongs.

310 observed during growth on the agarose pads (**Figure 5, S2**). Chains were also observed in 4
311 additional Clade 5 strains analyzed during exponential-phase growth in broth culture.
312 Interestingly, two of the Clade 5 strains, V48 and WU66, did not appear to incorporate the
313 HADA label efficiently, and the cells appeared lighter in phase-contrast images despite
314 undergoing the fixation process shortly after HADA addition. It is unclear why the HADA was
315 poorly incorporated by these strains across multiple independent experiments. Regardless, these
316 results indicate that many Clade 5 strains undergo cell separation less efficiently than strains
317 from other clades irrespective of whether they are grown on a solid surface or in broth culture.
318

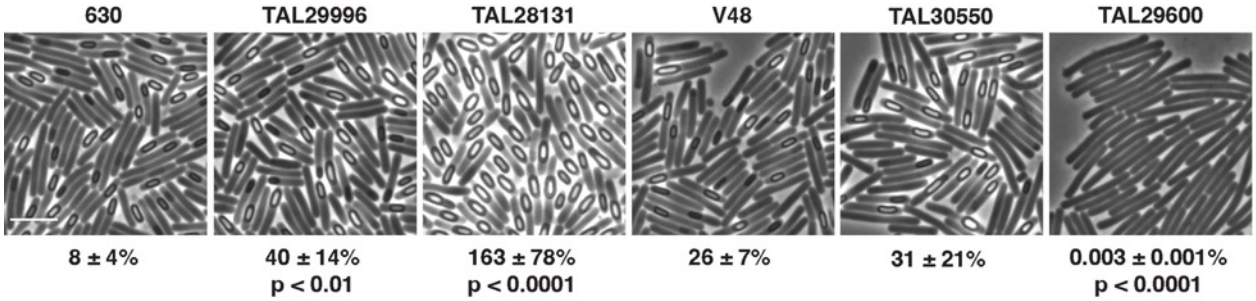


319
320 **Figure 5. Clade 5 strains form chains during logarithmic growth in broth culture**. Representative
321 micrographs showing phase-contrast (top) and peptidoglycan labeling with the fluorescent D-amino
322 acid, HADA, (bottom) following growth in rich broth (BHIS) to logarithmic phase. Scale bar, 10 µm.
323 Data are representative of multiple independent experiments. **ADJ** indicates that the brightness of the
324 image was enhanced to detect HADA labeling in V48.
325



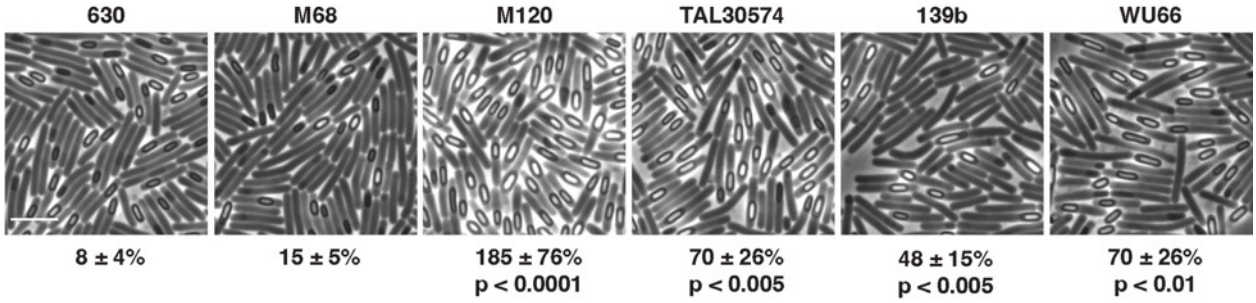
326
327
328 **Figure S2. Clade 5, but not Clade 1 or 3, strains form chains during logarithmic growth in broth**
329 **culture**. Representative micrographs showing phase-contrast (top) and peptidoglycan labeling with the
330 fluorescent D-amino acid, HADA (bottom) following growth in rich broth (BHIS) to logarithmic
331 phase. Scale bar, 10 µm. Data are representative of multiple independent experiments. **ADJ** indicates
332 that the brightness of the image was enhanced to detect HADA labeling in WU66.
333
334 **Cell length does not correlate with the propensity to sporulate.**
335

336 We next wondered whether the propensity to form chains impacts the ability of Clade 5
337 strains to sporulate because analyses in *Bacillus subtilis* suggest that smaller cells, such as those
338 formed during stationary phase growth (41), may concentrate proteins involved in the
339 phosphorelay that induces sporulation. In *B. subtilis*, the decrease in cell length helps the kinase
340 KinA reach the threshold concentration needed to trigger sporulation initiation (42, 43).
341 Although *C. difficile* lacks homologs of KinA and other components of the phosphorelay
342 system, it is possible that the longer cell length observed in Clade 5 strains may similarly reduce
343 their ability to induce sporulation by diluting the levels of a currently unknown sporulation
344 regulator. To test this hypothesis, we analyzed the propensity of Clade 5 strains to form spores
345 when plated on 70:30 sporulation medium using phase-contrast microscopy and heat resistance
346 assays.
347 Striking differences in the frequency of spore formation were observed across the Clade
348 5 strains, with several strains exhibiting close to 100% sporulation levels (**Figure 6A, S3**).
349 Although TAL29600, which forms the longest chains of the strains tested, exhibited extremely
350 low levels of sporulation (0.003%), we did not observe a correlation between sporulation and
351 cell length among the strains (**Figure 6B, S3**). For example, M120 forms long chains yet readily
352 forms spores (~100% sporulation frequency). The low sporulation levels of TAL29600 were
353 rather surprising since this strain was isolated from a clinical setting, so it likely forms spores
354 more efficiently in conditions encountered during infection. Notably, when grown on 70:30
355 medium, Clade 5 strains did not appear to form the long chains observed during broth culture
356 and grown on agarose pads. In addition, the spores produced by all *C. difficile* isolates analyzed
357 exhibited similar lengths and proportions despite arising from vegetative cells that varied
358 widely in their cell length (**Figure S4**). Thus, our data suggest that Clade 5 strains alter their cell
359 length and propensity to form chains depending on the media encountered.

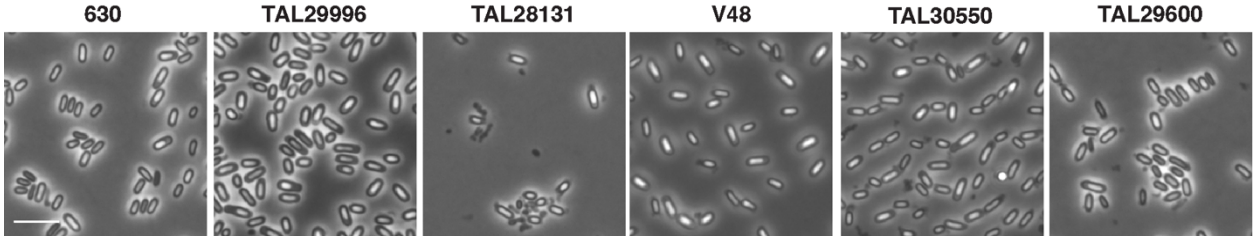


360
361
362 **Figure 6. Sporulation levels in Clade 5 isolates grown on 70:30 medium**. Phase-contrast microscopy
363 of the indicated strains ~24 hrs after sporulation induction. Heat resistance efficiencies, indicated below

364 the respective images, are based on 20-24 hr sporulating cultures and represent the mean and standard
365 deviation for a given strain relative to wild type based on a minimum of three biological replicates.
366 Statistical significance relative to strain 630 was determined using a one-way ANOVA and Tukey's test.
367 The scale bar represents 5 µm.
368



369
370
371 **Figure S3. Sporulation levels in clinical isolates grown on 70:30 medium**. Phase-contrast microscopy
372 of the indicated strains ~24 hrs after sporulation induction. Heat resistance efficiencies are based on 20-
373 24 hr sporulating cultures and represent the mean and standard deviation for a given strain relative to
374 wild type based on a minimum of three biological replicates. Statistical significance relative to strain 630
375 was determined using a one-way ANOVA and Tukey's test. The scale bar represents 5 µm.
376



377
378 **Figure S4. Spores purified from Clade 5 strains.** Spores were purified using a Histodenz gradient. Strain
379 630 is provided as a reference. The scale bar represents 5 µm.
380

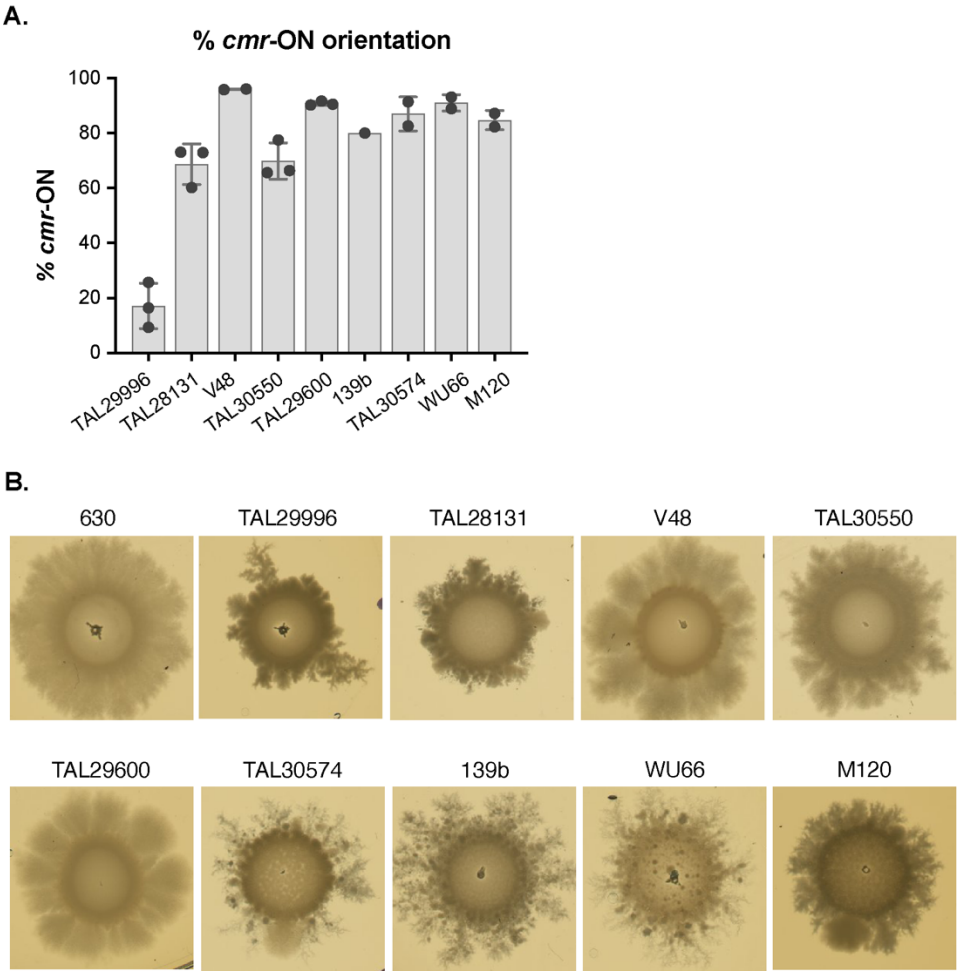
381 **Genomic comparisons of Clade 5 strains**

382

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

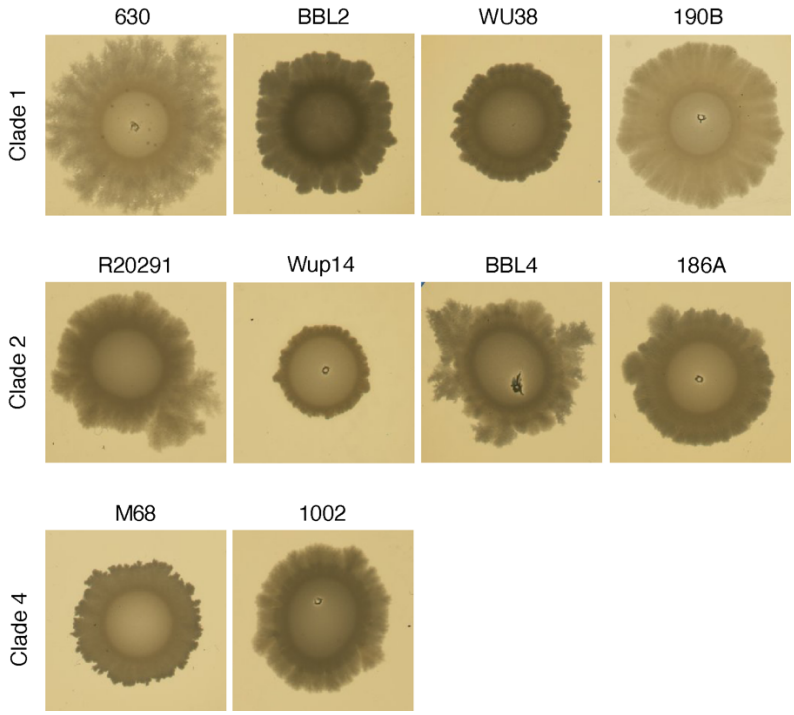
393 reference, since this strain has been the Clade 5 strain traditionally characterized (30, 45).
394 Around 10 SNPs were unique to TAL29996, but none were clearly implicated in regulating cell
395 separation or peptidoglycan synthesis (**Tables S2, S3**).
396 Since testing whether one or some of these SNPs or absent genes help TAL29996
397 undergo cell separation more efficiently would require genetic manipulation of TAL29996, we
398 instead analyzed the orientation of the *cmr* switch, also known as the Cdi6 DNA invertible
399 element (31, 46). The orientation of the *cmr* switch affects expression of the adjacent *cmrRST*
400 operon, which encodes a non-canonical two-component system, resulting in phase variation of
401 CmrRST. Because CmrRST promotes cell chaining, the orientation of the *cmr* switch determines
402 whether *C. difficile* produces rough or smooth colony morphologies (29, 31, 46). Specifically,
403 cells from rough colonies are highly biased to the ON orientation of the *cmr* switch and exhibit a
404 chaining phenotype, whereas cells from smooth colonies are highly biased to the OFF orientation
405 and do not form chains (31). Consistent with the breseq analyses, manual inspection of the
406 *cmrRST* upstream region revealed that this region is highly similar between all the Clade 5
407 strains analyzed (98% identity), including strain TAL29996. In the Illumina whole genome
408 sequencing performed on the subset of 5 Clade 5 strains, the *cmr* switch is in the OFF
409 orientation; these results are consistent with the genomic DNA from these samples being
410 harvested from liquid culture growth, which favors the *cmr*-OFF state at least for Clade 2 strain
411 R20291.
412 However, in contrast with the Illumina analyses, qPCR analyses of the *cmr* switch region
413 in the Clade 5 isolates revealed that TAL29996 exhibits a *cmr*-ON frequency of ~20%, whereas
414 the four remaining Clade 5 strains exhibit a *cmr*-ON frequency of between 70-96% on average
415 **(Fig 7**). Furthermore, when we extended these analyses to the four additional Clade 5 strains
416 characterized that also produce chains, we found that they exhibit a *cmr*-ON frequency of
417 between 80-93%. Collectively, these results suggest that the orientation of the *cmr* switch
418 orientation correlates with the propensity to form chains in Clade 5 strains and that most Clade 5
419 strains are biased to the *cmr*-ON orientation even during growth in broth culture.
420 Since the *cmr*-ON state has also been correlated with increased surface motility, we next
421 analyzed the surface motility of Clade 5 strains. After incubation for 5 days, *cmr*-ON strains
422 spread further on agar plates than the TAL29996 *cmr*-OFF strain (**Fig 7B**). Consistent with prior
423 reports, Clade 5 strains whose *cmr* switch is predominantly in the ON orientation exhibited

424 greater and more uniform surface motility. In contrast, TAL29996 exhibited less and more
425 asymmetric surface motility, with fractal-like extensions extending from a couple of sites. The
426 surface motility of TAL29996 resembles the phenotype previously reported for Clade 2 strain
427 R20291, whose *cmr* switch is also predominantly in the OFF position in liquid cultures.
428 However, growth of R20291 on agar medium selects for bacteria with the *cmr* switch in the ON
429 orientation, which leads to asymmetric surface motility on plates (31). These observations
430 suggest that even though TAL29996 is biased to the *cmr*-OFF orientation during broth culture
431 growth, during growth on BHIS agar, a subset of cells switch to the *cmr*-ON orientation, leading
432 to the asymmetric spreading phenotype.
433



434
435
436 **Figure 7. Orientation of the *cmr* switch correlates with chaining phenotype and surface motility in**
437 **Clade 5 strains.** A. Orientation-specific PCR for detecting the orientation of the *cmr* switch in the
438 indicated strains. B. Representative images of surface motility 5 days after exponentially growing liquid
439 cultures were spotted onto BHIS agar plates.
440

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



450
451
452 **Figure S5. Surface motility in Clade 1, 2, and 4 strains.** Representative images of surface motility 5
453 days after exponentially growing liquid cultures were spotted onto BHIS agar plates.
454
455
456 **Discussion**
457
458 While the genetically distinct nature of Clade 5 strains from Clade 1-4 strains is well-
459 established (10, 26), the phenotypic impacts of this genetic diversity are poorly understood,
460 with only one study to our knowledge systematically comparing Clade 5 strains to Clade 1-4
461 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

493 broth culture-grown cells (**Figures 4, 5**). Thus, the chaining phenotype of Clade 5 strains may
494 be similarly responsive to c-di-GMP levels during growth on a surface (29). While testing this
495 hypothesis will depend on the ability to genetically manipulate Clade 5 strains, it is tempting to
496 speculate that *cmrRST* expression is an important determinant of the growth properties of Clade
497 5 strains. Expression of this operon in Clade 2 strain R20291 not only promotes chaining and
498 surface motility in a Type IV pilus-independent manner, it also negatively regulates biofilm
499 formation and flagellar motility (31). In addition, growth conditions favoring flagellar motility
500 select against *cmr*-ON bacteria. Since Clade 5 strains lack flagellar motility (12), they may be
501 "primed" to exhibit surface motility in the absence of selection against the *cmr*-ON state.
502 Regardless, analyzing the regulation of c-di-GMP in different growth conditions in Clade 5
503 strains, particularly for TAL29996 relative to the other strains, will likely provide insight into
504 the mechanisms that drive the higher incidence of *cmr*-ON cells in most Clade 5 strains.
505 Even though Clade 5 strains form long chains and produce longer cells than strains from
506 the other clades, neither cell nor chain length was predictive of the propensity to sporulate
507 (**Figure 6**). Clade 5 strains exhibited wildly different sporulation frequencies, with TAL29600
508 forming spores inefficiently (~5,000-fold less efficiently than most Clade 5 strains). While little
509 is known about the mechanisms regulating sporulation initiation outside of strains 630 and
510 R20291, future analyses of strain TAL29600, which sporulates poorly under the laboratory
511 media tested (**Figure 6**), could provide insight into the molecular determinants of sporulation in
512 Clade 5 strains. For example, differences in gene presence or polymorphisms in several c-di-
513 GMP-related genes were observed in TAL29600 relative to the other Clade 5 strains (**Table**
514 **S2**), and c-di-GMP has also been implicated in regulating sporulation initiation events through
515 an unknown mechanism (47, 48). Thus, analyzing the contribution of these genes in future work
516 could provide important insight into c-di-GMP-mediated regulation of sporulation.
517 The variable sporulation levels observed for the nine Clade 5 strains analyzed in this
518 manuscript contrast with the findings of a previous study, which reported that Clade 5 strains
519 sporulate at lower rates in the presence of glucose and fructose (26). Kumar *et al*. previously
520 correlated phenotypic differences in sporulation and mouse colonization between Clade 5
521 strains vs. Clade 1-4 strains to their finding that genes involved in sporulation and the
522 metabolism of simple sugars are under positive selection in Clade 1-4 strains (26). However, the
523 five Clade 5 strains analyzed in this study all exhibited very low sporulation frequencies under

524 the conditions tested. While the conditions used in this study to induce sporulation differ from
525 those used by Kumar *et al*., our analyses of nine Clade 5 strains indicate that there are marked
526 phenotypic differences in sporulation frequency within this clade. Thus, analyzing more strains
527 within a clade may reveal fewer differences between clades. A caveat to our analyses is that *C.*
528 *difficile* strains exhibit considerable genetic diversity even within a ribotype (20, 22), so our
529 conclusions may only reflect the behavior of RT078 Clade 5 strains. It is also possible that
530 analyzing more strains within this ribotype will reveal a broader spectrum of phenotypes than
531 the somewhat bimodal ones we have reported with respect to cell length and chaining
532 propensity (**Figures 2, 3**). Indeed, a recent analysis of a large number of RT027 strains indicates
533 that strains within this ribotype can exhibit marked differences in virulence in mice (22) despite
534 these strains traditionally being associated with "epidemic" outbreaks in hospital setting.
535 While addressing these questions will depend on analyzing a larger number of strains
536 from each clade, an important outcome of our study is its development of a facile method for
537 conducting time-lapse microscopy under anaerobic conditions. Although this method is a closed
538 system, i.e. no additional reagents can be introduced once the bacteria are inoculated onto the
539 pad in the anaerobic chamber, it could be coupled to recently developed anaerobic imaging tags
540 based on fluorogen-activated proteins (49). Thus, there are many potential applications for this
541 simple method for studying the growth of anaerobes over time at the single-cell level.
542
543
544 **Materials and Methods**
545
546 *Bacterial strains and growth conditions*
547 All *C. difficile* strains were grown on brain heart infusion (BHIS) medium supplemented
548 with 0.5% w/v yeast extract and 0.1% w/v L-cysteine with taurocholate (TCA; 0.1% w/v; 1.9
549 mM). Strains were sub-cultured into tryptone yeast extract (TY) broth supplemented with 0.1%
550 w/v L-cysteine (TYC medium) prior to inoculation onto the time-lapse microscopy agarose
551 pads. All strains were grown at 37°C under anaerobic conditions using a gas mixture of 85%
552 hydrogen, 5% CO2, and 10% H2. For time-lapse experiments, 1.5% agarose pads were used
553 supplemented with TYC medium as described above. Sporulation analyses were carried out on
554 70:30 medium (70% BHIS and 30% SMC) for 24h as described previously (50).
555
556 *Anaerobic time-lapse imaging of C. difficile growth*
557 All imaging was carried out on a Leica DMi8 inverted microscope with a HC plan
558 apochromat 63x 1.4 NA oil immersion phase contrast objective. Fluorescent membrane staining
559 experiments were done with a Lumencor Spectra X light source, coupled with an XLED-QP
560 quadruple-band dichroic beam-splitter (Leica) (transmission: 415, 470, 570, and 660 nm) along

561 with an external emission filter wheel (Leica). FM4-64 was excited using a 470nm LED
562 through a 470/20nm excitation filter and emitted light was filtered through 590/50nm emission
563 filter and captured with a Leica DFC9000GTC sCMOS camera. All experiments were carried
564 out at 37°C using a microscope incubation system (Pecon), Leica Adaptive Focus Control
565 hardware autofocus, and a high precision stage (Pecon) were used for all imaging experiments.
566 For time-lapse imaging of *C. difficile* growth all bacterial strains were grown in 2 mL
567 liquid TY medium to a turbid OD600 > 2-3; after 2 hours of growth, bacteria were diluted 1:100
568 for Clade 5 strains and all other strains were diluted 1:50 in fresh media and grown to mid-log
569 stage (OD600 0.4-0.7).
570 An imaging chamber with a gas-tight seal was constructed using a 125 µL Gene
571 Frame (Thermo Fisher) adhered to a glass slide generating a well for growth medium. The slide
572 was then transferred to the anaerobic chamber. In the anaerobic chamber, the gene frame was
573 filled with 500µl 1.5% Top vision low melting point agarose and tryptone yeast extract media
574 containing 0.1% w/v L-cysteine to scavenge oxygen and maintain anaerobic conditions. While
575 the agarose was molten, a second clean slide was placed over the top and the agar pad was
576 placed on a frozen small freezer block (for holding PCR strip tubes) for 10-30 minutes until the
577 agarose-media mixture was solid. For experiments where FM4-64 was used, agarose pads were
578 made the same way, with the addition of FM4-64 to a final concentration of 1 µg/mL directly to
579 the agarose/media solution prior to making the agar pad.
580 The agar pad was dried for 5-10 minutes until visible liquid on the surface of the pad
581 was evaporated. 1 µL of mid-log cells were spotted on the pad, dried, and a #1.5 coverslip
582 (VWR) was adhered to the Gene Frame. The cells were imaged at 37°C until they reached
583 confluency in the field of view. This was anywhere from 2.5 hours for Clade 5 strains to 6 hours
584 for Clades 1-4 for all experiments with images taken at 5-minute intervals.
585
586 *Image analysis, computing hardware, and statistical analysis.*
587 All movie frames were trimmed to the point when cells were not overlapping and out of
588 focus regions were cropped. The resulting images were analyzed using the Python library
589 DeLTA 2.0 (39, 40). All image and data analyses were done on a PC running
590 Windows 10 equipped with an AMD Ryzen 5900HX 8-core CPU, 32GB DDR4 RAM, 2 1TB
591 NVME SSDs, and an NVIDIA RTX3080 GPU with 16GB VRAM. Analysis of the output data
592 and data visualizations were done in Python using Matplotlib/Seaborn,
593 Pandas, Numpy, Scipy, and the Statannotations library.
594
595 **Septum detection**
596 The image processing starts with the masks generated using DeLTA 2.0. First, we performed
597 erosion on the mask with a disk of radius of 1 pixel to avoid effects of the membrane fluorescence.
598 The photo was cropped following the eroded mask contour. Hence, only the interior of the contour
599 was considered. Pixel values were rescaled such as the minimum pixel value of inner pixels was
600 mapped to 0 and the maximum is mapped to 255. We set a threshold intensity as the pixel value
601 of the 95% quantile of pixel intensities. The image was slightly blurred by convolution with a
602 Gaussian filter with standard deviation of 2 pixels. The septum corresponds to the contours found
603 after performing an adaptive thresholding based on thresholds defined using a Gaussian-weighted
604 method. This thresholding was performed using the open cv2 library in Python. The resultant
605 contours were manually verified from videos generated using the tracking from the outputs from
606 DeLTA. The manual validation mainly involved filling in the time points where the algorithm

607 missed a ring that was visible in the video. We did not assume the existence of a ring if the
608 algorithm did not detect it in a previous ,m,frame.
609
610 **Cell length estimation**
611 We measured the cell projected area from the DeLTA contours of the images as the pixel amount
612 of the contours. However, estimating the cell length was challenging because some cells were very
613 long and bent. To overcome this problem, we selected 30 images of three different cells (from
614 strains 630, TAL3050 and TAL28131) that were straight and had different lengths. For these cells,
615 we calculated the cell length as the longest side of the minimum bounding rectangle of the contour.
616 From these lengths of straight cells, we also estimated the best cell width as the average of the
617 projected area divided by the length. Considering the extreme cell length, the effects of the rounded
618 tips were negligible and the rectangle shape adequately approximated length. We verifed that this
619 mean value showed low variability for the three strains. Then, we used this width value to estimate
620 the length of all the cells, including the bent ones, by dividing their projected area by the width.
621 This way, we obtained a consistent measure of cell length that was independent of bending.
622
623 **Elongation rate estimation**
624 We tracked the cell size over time and identified the division points as the ones where the cell size
625 (projected area) dropped by more than 30% compared to the current cell size value. We fitted an
626 exponential function (with base e) of time to the data points between two divisions and estimated
627 the elongation rate from the exponent of the best fit. We expressed the elongation rate in
628 doublings/hr, which means how many times the cell size doubles in one hour. For example, an
629 elongation rate of 2 doubling/hr means that the cell size doubles two times in 1 hr, which
630 corresponds to an exponent of 2ln(2) 1/hr. For the statistics, we only included the elongation rates
631 that had a high quality of fit, with an R2 coefficient greater than 0.9.
632
633 **Bulk growth measurements**. Starter cultures were grown until early stationary phase in BHIS
634 (or TYC medium as indicated) then back-diluted 1:50 into BHIS (or TYC medium). When the
635 cultures reached an OD_{600} of 0.5, they were back-diluted 1:25 into 200 µL of either BHIS or
636 TYC in a flat 96 well polystyrene plate (CellTreat). The OD_{600} was analyzed every 15 min for 24
637 hrs in a BioTek Epoch plate reader with shaking. Bulk growth measurements are based on a
638 minimum of three independent replicates across a minimum of 2 experiments.
639
640 **Sporulation assays**. Starter cultures were grown until early stationary phase in BHIS then back-
641 diluted 1:50 into BHIS. When the cultures reached an OD_{600} between 0.35 and 0.75, 120 µL of
642 the culture was spread onto 70:30 (70% SMC media and 30% BHIS media) agar plates (40 ml
643 media per plate) and then incubated for 20-24 hrs before the sporulating cells were scraped from
644 the plate into phosphate-buffered saline (PBS). Sporulation levels were visualized by phase-
645 contrast microscopy as previously described (51).
646
647 **Heat resistance assay**. Heat-resistant spore formation was measured 20-24 hrs after sporulation
648 was induced on 70:30 agar plates as previously described (50). The percent sporulation of given
649 culture represents the ratio of heat-resistant colony-forming units (CFUs) to total CFUs. Percent
650 sporulation was determined from a minimum of 3 biological replicates.
651
652 **Spore purification**. Spores were purified as previously described (52) by scraping up
653 sporulating cells incubated on 70:30 medium for 3 days into ice-cold H_2O. The cells were

654 washed several times in ice-water over the course of a day and incubated on ice overnight. The
655 following morning, the sample was pelleted, and cells were resuspended in 1 X DNAse buffer
656 (New England Biolabs) and then treated with DNAse (New England Biolabs) for 30 min at 37˚C.
657 The samples was washed one more time before being resuspended in 20% Histodenz and then
658 layered onto a 50% Histodenz layer. The resulting mixture was pelleted and the supernatant was
659 aspirated off using a vacuum aspirator. The pelleted spores were washed in ice-cold water 2-3
660 times and the optical density of the purified spores was measured.
661
662 **Genomic DNA preparation**. Starter cultures were grown until early stationary phase in BHIS
663 then back-diluted 1:50 into BHIS and grown until an OD_{600} of around 0.7-0.8 was reached. 10
664 mL of the culture was pelleted and then frozen at -80˚C. After thawing the sample, it was
665 resuspended in a 25% sucrose TE buffer (10 mM Tris, 1mM EDTA), incubated with 100 mg/mL
666 lysozyme for 37˚C for 1 hr. After the cultures tarted to lyse, proteinase K, RNAse A, EDTA,
667 Sarkosyl, and NaCl was added. Phenol:Chloroform:IAA (25:24:1) was added to extract proteins,
668 gently mixed, and then the sample was pelleted to separate the phenol and aqueous layer. The
669 aqueous layer was then added to Chloroform:IAA (24:1), mixed gently, then centrifuged. The
670 aqueous layer was then precipitated using isopropanol and incubated at -20˚C for a minimum of
671 15 min. The precipitated DNA was pelleted and then washed with 70% ethanol. The pellet was
672 air dried and then gently resuspended in 10 mM Tris pH 8.0 elution buffer.
673
674 **Genomic analyses**. Genomic DNA was sequenced by MiGS at the University of Pittsburgh
675 (now SeqCenter) according to their standard protocol. Libraries were sequenced on an Illumina
676 NextSeq 500 platform to generate paired-end 150 bp reads. Illumina reads of RT078 genomes
677 were assembled into contigs using SPAdes (v3.13.0),(53) and genes were called and annotated
678 using Prokka (v1.11).(54) Assembled and annotated contigs of five RT078 strains (TAL28131,
679 TAL29600, TAL29996, TAL30550, TAL30574) were applied for pangenomic analysis. Default
680 settings were used based on the Anvi'o workflow for microbial pangenomcis with adjustments
681 for minbit as 0.5 and mcl-inflation as 10.(55-57) For SNPs analyses, reads of five RT078
682 genomes were aligned to the reference M120 and variants were called by breseq (v. 0.38.1) by
683 default settings.(58)
684
685 **qPCR analyses**. Each genomic DNA sample was analyzed by qPCR with primers that amplify
686 the *cmr*-ON sequence orientation, the *cmr*-OFF orientation, or the reference gene *rpoA* (31) .
687 Each 20-µL qPCR reaction consisted of 100 ng genomic DNA, 100 nM primers, and SensiMix™
688 SYBR® reagents (Bioline). The reactions were run on a LightCycler® 96 (Roche Diagnostics),
689 and *cmr* switch orientation frequencies were calculated as described previously (29).
690
691 **Surface motility assays**
692 Starter cultures were grown until early stationary phase in BHIS then back-diluted 1:50 into
693 BHIS and grown until an OD_{600} of 0.5 was reached. 10 µL of the exponential-phase cultures
694 were then spotted onto BHIS plates and incubated at 37˚C for 5 days after which the plates were
695 scanned using a flatbed scanner.
696
697 *Author contribution statement*
698 J.W.R. conducted all live-cell microscopy experiments and analyzed the single-cell data.
699 C.A.N.A. generated data visualizations from the time-lapse microscopy data, developed the

700 method for identifying septa within chains that had not initiated cell separation, and calculated
701 elongation rates and cell lengths. I.N. and A.S. conducted sporulation analyses and liquid
702 culture experiments. J.W.R., R.T., and A.S conceptualized all experiments with input from .
703 Q.D. and A.S extracted genomic DNA and visualized cells at different growth phases. Q.D.
704 conducted all the genomic analyses and genome comparisons. N.D. and I.N. conducted spore
705 purification analyses and imaging. B.A. and IE developed and optimized methods for
706 constructing agarose pads that could support anaerobic growth in an ambient atmosphere.
707 M.J.D. developed the DeLTA 2.0 analysis pipeline. A.M. conducted qPCRs to determine *cmr*
708 switch orientation in Clade 5 isolates. Writing was by J.W.R and A.S with input from all the
709 authors.
710
711
712 **Table 1. *Clostridioides difficile* clinical isolates used in this study**
713

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)

714
715 Memorial Sloan Kettering Cancer Center (MSK), Barnes-Jewish Hospital (BJH)
716
717 **References**
718

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

727 3. Shen A. 2020. Clostridioides difficile Spore Formation and Germination: New Insights
728 and Opportunities for Intervention. Annu Rev Microbiol 74:545-566.
729 4. Abt MC, McKenney PT, Pamer EG. 2016. Clostridium difficile colitis: pathogenesis and
730 host defence. Nat Rev Microbiol doi:10.1038/nrmicro.2016.108.
731 5. Theriot CM, Young VB. 2015. Interactions Between the Gastrointestinal Microbiome
732 and Clostridium difficile. Annu Rev Microbiol 69:445-61.
733 6. Kordus SL, Thomas AK, Lacy DB. 2022. Clostridioides difficile toxins: mechanisms of
734 action and antitoxin therapeutics. Nat Rev Microbiol 20:285-298.
735 7. Maroo S, Lamont J. 2006. Recurrent Clostridium difficile. Gastroenterology 130:1311-
736 1316.
737 8. Zanella Terrier MC, Simonet ML, Bichard P, Frossard JL. 2014. Recurrent Clostridium
738 difficile infections: the importance of the intestinal microbiota. World J Gastroenterol
739 20:7416-23.
740 9. Knight DR, Elliott B, Chang BJ, Perkins TT, Riley TV. 2015. Diversity and Evolution in
741 the Genome of Clostridium difficile. Clin Microbiol Rev 28:721-41.
742 10. Knight DR, Imwattana K, Kullin B, Guerrero-Araya E, Paredes-Sabja D, Didelot X,
743 Dingle KE, Eyre DW, Rodriguez C, Riley TV. 2021. Major genetic discontinuity and
744 novel toxigenic species in Clostridioides difficile taxonomy. Elife 10.
745 11. Norsigian CJ, Danhof HA, Brand CK, Midani FS, Broddrick JT, Savidge TC, Britton
746 RA, Palsson BO, Spinler JK, Monk JM. 2022. Systems biology approach to functionally
747 assess the Clostridioides difficile pangenome reveals genetic diversity with
748 discriminatory power. Proc Natl Acad Sci U S A 119:e2119396119.
749 12. Knight DR, Kullin B, Androga GO, Barbut F, Eckert C, Johnson S, Spigaglia P, Tateda
750 K, Tsai PJ, Riley TV. 2019. Evolutionary and Genomic Insights into Clostridioides
751 difficile Sequence Type 11: a Diverse Zoonotic and Antimicrobial-Resistant Lineage of
752 Global One Health Importance. mBio 10.
753 13. Seth-Smith HMB, Biggel M, Roloff T, Hinic V, Bodmer T, Risch M, Casanova C,
754 Widmer A, Sommerstein R, Marschall J, Tschudin-Sutter S, Egli A. 2021. Transition
755 From PCR-Ribotyping to Whole Genome Sequencing Based Typing of Clostridioides
756 difficile. Front Cell Infect Microbiol 11:681518.
757 14. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput
758 ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun
759 9:5114.
760 15. Dingle KE, Elliott B, Robinson E, Griffiths D, Eyre DW, Stoesser N, Vaughan A,
761 Golubchik T, Fawley WN, Wilcox MH, Peto TE, Walker AS, Riley TV, Crook DW,
762 Didelot X. 2014. Evolutionary history of the Clostridium difficile pathogenicity locus.
763 Genome Biol Evol 6:36-52.
764 16. Mengoli M, Barone M, Fabbrini M, D'Amico F, Brigidi P, Turroni S. 2022. Make It Less
765 difficile: Understanding Genetic Evolution and Global Spread of Clostridioides difficile.
766 Genes (Basel) 13.
767 17. Endres BT, Begum K, Sun H, Walk ST, Memariani A, Lancaster C, Gonzales-Luna AJ,
768 Dotson KM, Basseres E, Offiong C, Tupy S, Kuper K, Septimus E, Arafat R, Alam MJ,
769 Zhao Z, Hurdle JG, Savidge TC, Garey KW. 2019. Epidemic Clostridioides difficile
770 Ribotype 027 Lineages: Comparisons of Texas Versus Worldwide Strains. Open Forum
771 Infect Dis 6:ofz013.

772 18. McDonald LC, Killgore GE, Thompson A, Owens RC, Jr., Kazakova SV, Sambol SP,
773 Johnson S, Gerding DN. 2005. An epidemic, toxin gene-variant strain of Clostridium
774 difficile. N Engl J Med 353:2433-41.
775 19. Clements AC, Magalhaes RJ, Tatem AJ, Paterson DL, Riley TV. 2010. Clostridium
776 difficile PCR ribotype 027: assessing the risks of further worldwide spread. Lancet Infect
777 Dis 10:395-404.
778 20. Lewis BB, Carter RA, Ling L, Leiner I, Taur Y, Kamboj M, Dubberke ER, Xavier J,
779 Pamer EG. 2017. Pathogenicity Locus, Core Genome, and Accessory Gene Contributions
780 to Clostridium difficile Virulence. mBio 8.
781 21. Merrigan M, Venugopal A, Mallozzi M, Roxas B, Viswanathan VK, Johnson S, Gerding
782 DN, Vedantam G. 2010. Human hypervirulent Clostridium difficile strains exhibit
783 increased sporulation as well as robust toxin production. J Bacteriol 192:4904-11.
784 22. Dong Q, Lin H, Allen MM, Garneau JR, Sia JK, Smith RC, Haro F, McMillen T, Pope
785 RL, Metcalfe C, Burgo V, Woodson C, Dylla N, Kohout C, Sundararajan A, Snitkin ES,
786 Young VB, Fortier LC, Kamboj M, Pamer EG. 2023. Virulence and genomic diversity
787 among clinical isolates of ST1 (BI/NAP1/027) Clostridioides difficile. Cell Rep
788 42:112861.
789 23. Shaw HA, Preston MD, Vendrik KEW, Cairns MD, Browne HP, Stabler RA, Crobach
790 MJT, Corver J, Pituch H, Ingebretsen A, Pirmohamed M, Faulds-Pain A, Valiente E,
791 Lawley TD, Fairweather NF, Kuijper EJ, Wren BW. 2020. The recent emergence of a
792 highly related virulent Clostridium difficile clade with unique characteristics. Clin
793 Microbiol Infect 26:492-498.
794 24. Li C, Harmanus C, Zhu D, Meng X, Wang S, Duan J, Liu S, Fu C, Zhou P, Liu R, Wu A,
795 Kuijper EJ, Smits WK, Fu L, Sun X. 2018. Characterization of the virulence of a non-
796 RT027, non-RT078 and binary toxin-positive Clostridium difficile strain associated with
797 severe diarrhea. Emerg Microbes Infect 7:211.
798 25. Imwattana K, Rodriguez C, Riley TV, Knight DR. 2021. A species-wide genetic atlas of
799 antimicrobial resistance in Clostridioides difficile. Microb Genom 7.
800 26. Kumar N, Browne HP, Viciani E, Forster SC, Clare S, Harcourt K, Stares MD, Dougan
801 G, Fairley DJ, Roberts P, Pirmohamed M, Clokie MRJ, Jensen MBF, Hargreaves KR, Ip
802 M, Wieler LH, Seyboldt C, Noren T, Riley TV, Kuijper EJ, Wren BW, Lawley TD. 2019.
803 Adaptation of host transmission cycle during Clostridium difficile speciation. Nat Genet
804 51:1315-1320.
805 27. Knetsch CW, Kumar N, Forster SC, Connor TR, Browne HP, Harmanus C, Sanders IM,
806 Harris SR, Turner L, Morris T, Perry M, Miyajima F, Roberts P, Pirmohamed M, Songer
807 JG, Weese JS, Indra A, Corver J, Rupnik M, Wren BW, Riley TV, Kuijper EJ, Lawley
808 TD. 2018. Zoonotic Transfer of Clostridium difficile Harboring Antimicrobial Resistance
809 between Farm Animals and Humans. J Clin Microbiol 56.
810 28. Lim SC, Knight DR, Riley TV. 2020. Clostridium difficile and One Health. Clin
811 Microbiol Infect 26:857-863.
812 29. Garrett EM, Mehra A, Sekulovic O, Tamayo R. 2021. Multiple Regulatory Mechanisms
813 Control the Production of CmrRST, an Atypical Signal Transduction System in
814 Clostridioides difficile. mBio 13:e0296921.
815 30. Thanissery R, Winston JA, Theriot CM. 2017. Inhibition of spore germination, growth,
816 and toxin activity of clinically relevant C. difficile strains by gut microbiota derived
817 secondary bile acids. Anaerobe 45:86-100.

818 31. Garrett EM, Sekulovic O, Wetzel D, Jones JB, Edwards AN, Vargas-Cuebas G, McBride
819 SM, Tamayo R. 2019. Phase variation of a signal transduction system controls
820 Clostridioides difficile colony morphology, motility, and virulence. PLoS Biol
821 17:e3000379.
822 32. Ackermann M. 2015. A functional perspective on phenotypic heterogeneity in
823 microorganisms. Nat Rev Microbiol 13:497-508.
824 33. Aldridge BB, Fernandez-Suarez M, Heller D, Ambravaneswaran V, Irimia D, Toner M,
825 Fortune SM. 2012. Asymmetry and aging of mycobacterial cells lead to variable growth
826 and antibiotic susceptibility. Science 335:100-4.
827 34. Anjuwon-Foster BR, Tamayo R. 2017. A genetic switch controls the production of
828 flagella and toxins in Clostridium difficile. PLoS Genet 13:e1006701.
829 35. Anjuwon-Foster BR, Tamayo R. 2018. Phase variation of Clostridium difficile virulence
830 factors. Gut Microbes 9:76-83.
831 36. Kint N, Janoir C, Monot M, Hoys S, Soutourina O, Dupuy B, Martin-Verstraete I. 2017.
832 The alternative sigma factor sigma(B) plays a crucial role in adaptive strategies of
833 Clostridium difficile during gut infection. Environ Microbiol 19:1933-1958.
834 37. Courson DS, Pokhrel A, Scott C, Madrill M, Rinehold AJ, Tamayo R, Cheney RE,
835 Purcell EB. 2019. Single cell analysis of nutrient regulation of Clostridioides
836 (Clostridium) difficile motility. Anaerobe 59:205-211.
837 38. de Jong IG, Beilharz K, Kuipers OP, Veening JW. 2011. Live Cell Imaging of Bacillus
838 subtilis and Streptococcus pneumoniae using Automated Time-lapse Microscopy. J Vis
839 Exp doi:10.3791/3145.
840 39. Lugagne JB, Lin H, Dunlop MJ. 2020. DeLTA: Automated cell segmentation, tracking,
841 and lineage reconstruction using deep learning. PLoS Comput Biol 16:e1007673.
842 40. O'Connor OM, Alnahhas RN, Lugagne JB, Dunlop MJ. 2022. DeLTA 2.0: A deep
843 learning pipeline for quantifying single-cell spatial and temporal dynamics. PLoS
844 Comput Biol 18:e1009797.
845 41. Sauls JT, Cox SE, Do Q, Castillo V, Ghulam-Jelani Z, Jun S. 2019. Control of Bacillus
846 subtilis Replication Initiation during Physiological Transitions and Perturbations. mBio
847 10.
848 42. Narula J, Kuchina A, Zhang F, Fujita M, Suel GM, Igoshin OA. 2016. Slowdown of
849 growth controls cellular differentiation. Mol Syst Biol 12:871.
850 43. Eswaramoorthy P, Duan D, Dinh J, Dravis A, Devi SN, Fujita M. 2010. The threshold
851 level of the sensor histidine kinase KinA governs entry into sporulation in Bacillus
852 subtilis. J Bacteriol 192:3870-82.
853 44. Deatherage DE, Traverse CC, Wolf LN, Barrick JE. 2014. Detecting rare structural
854 variation in evolving microbial populations from new sequence junctions using breseq.
855 Front Genet 5:468.
856 45. He M, Sebaihia M, Lawley T, Stabler R, Dawson L, Martin M, Holt K, Seth-Smith H,
857 Quail M, Rance R, Brooks K, Churcher C, Harris D, Bentley S, Burrows C, Clark L,
858 Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren B, Dougan G,
859 Parkhill J. 2010. Evolutionary dynamics of Clostridium difficile over short and long time
860 scales. Proceedings of the National Academy of Sciences of the United States of America
861 107:7527-7532.

862 46. Sekulovic O, Mathias Garrett E, Bourgeois J, Tamayo R, Shen A, Camilli A. 2018.
863 Genome-wide detection of conservative site-specific recombination in bacteria. PLoS
864 Genet 14:e1007332.
865 47. Dhungel BA, Govind R. 2021. Phase-variable expression of pdcB, a phosphodiesterase,
866 influences sporulation in Clostridioides difficile. Mol Microbiol 116:1347-1360.
867 48. Edwards AN, Willams CL, Pareek N, McBride SM, Tamayo R. 2021. c-di-GMP Inhibits
868 Early Sporulation in Clostridioides difficile. mSphere 6:e0091921.
869 49. Streett H, Charubin K, Papoutsakis ET. 2021. Anaerobic fluorescent reporters for cell
870 identification, microbial cell biology and high-throughput screening of microbiota and
871 genomic libraries. Curr Opin Biotechnol 71:151-163.
872 50. Shen A, Fimlaid KA, Pishdadian K. 2016. Inducing and Quantifying Clostridium difficile
873 Spore Formation. Methods Mol Biol 1476:129-42.
874 51. Pishdadian K, Fimlaid KA, Shen A. 2015. SpoIIID-mediated regulation of sigma(K)
875 function during Clostridium difficile sporulation. Mol Microbiol 95:189-208.
876 52. Fimlaid KA, Jensen O, Donnelly ML, Francis MB, Sorg JA, Shen A. 2015. Identification
877 of a Novel Lipoprotein Regulator of Clostridium difficile Spore Germination. PLoS
878 Pathogens doi:10.1371/journal.ppat.1005239.
879 53. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes
880 De Novo Assembler. Curr Protoc Bioinformatics 70:e102.
881 54. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics
882 30:2068-9.
883 55. Benedict MN, Henriksen JR, Metcalf WW, Whitaker RJ, Price ND. 2014. ITEP: an
884 integrated toolkit for exploration of microbial pan-genomes. BMC Genomics 15:8.
885 56. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef
886 M, Fogarty EC, Trigodet F, Watson AR, Esen OC, Moore RM, Clayssen Q, Lee MD,
887 Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjodin
888 A, Scott JJ, Vazquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE,
889 Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-
890 led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 6:3-6.
891 57. van Dongen S, Abreu-Goodger C. 2012. Using MCL to extract clusters from networks.
892 Methods Mol Biol 804:281-95.
893 58. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved
894 microbes from next-generation sequencing data using breseq. Methods Mol Biol
895 1151:165-88.
896 59. Sebaihia M, Wren B, Mullany P, Fairweather N, Minton N, Stabler R, Thomson N,
897 Roberts A, Cerdeño-Târraga A, Wang H, Holden M, Wright A, Churcher C, Quail M,
898 Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A,
899 Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch
900 E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G,
901 Barrell B, Parkhill J. 2006. The multidrug-resistant human pathogen Clostridium difficile
902 has a highly mobile, mosaic genome. Nature genetics 38:779-786.
903 60. Stabler R, He M, Dawson L, Martin M, Valiente E, Corton C, Lawley T, Sebaihia M,
904 Quail M, Rose G, Gerding D, Gibert M, Popoff M, Parkhill J, Dougan G, Wren B. 2009.
905 Comparative genome and phenotypic analysis of Clostridium difficile 027 strains
906 provides insight into the evolution of a hypervirulent bacterium. Genome biology 10.
907