Flagellar switch inverted repeat impacts flagellar invertibility and varies *Clostridioides difficile* RT027/MLST1 virulence.

Nguyen T. Q. Nhu[a,b], Huaiying Lin[b], Ying Pigli[c], Jonathan K. Sia[d], Pola Kuhn[e], Evan S. Snitkin[f], Vincent Young[f,g], Mini Kamboj[h], Eric G. Pamer[a,b], Phoebe A. Rice[c], Aimee Shen[e], Qiwen Dong[a,b,e,#].

[a]Department of Medicine, University of Chicago, Chicago, Illinois, USA

[b]Duchossois Family Institute, University of Chicago, Chicago, Illinois, USA

[c]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL, USA

[d]Immunology Program, Memorial Sloan Kettering Cancer Center, New York City, New York, USA

[e]Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts, USA

[f]Division of Infectious Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA

[g]Department of Microbiology & Immunology, University of Michigan, Ann Arbor, MI, USA

[h]Infection Control, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA

Running Head: Flagellar switch inverted repeat varies *C. difficile* RT027 virulence.

[#]Address correspondence to Qiwen Dong, qiwendong0721@gmail.com.

*Present address: Nguyen T. Q. Nhu, Division of Infectious Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA; Qiwen Dong, Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts, USA

SUMMARY

Clostridioides difficile RT027 strains cause infections that vary in severity from asymptomatic to lethal, but the molecular basis for this variability is poorly understood. Through comparative analyses of RT027 clinical isolates, we determined that isolates that exhibit greater variability in their flagellar gene expression exhibit greater virulence *in vivo*. *C. difficile* flagellar genes are phase-variably expressed due to the site-specific inversion of the *flgB* 5'UTR region, which reversibly generates ON vs. OFF orientations for the flagellar switch. We found that longer inverted repeat (IR) sequences in this switch region correlate with greater disease severity, with RT027 strains carrying 6A/6T IR sequences exhibiting greater phenotypic heterogeneity in flagellar gene expression (60%-75% ON) and causing more severe disease than those with shorter IRs (> 99% ON or OFF). Taken together, our results reveal that phenotypic heterogeneity in flagellar gene expression may contribute to the variable disease severity observed in *C. difficile* patients.

KEYWORDS *Clostridioides difficile*, RT027, virulence, flagellar switch, inverted repeats, recombinase

40 **INTRODUCTION**

41 *Clostridioides difficile* (*C. difficile*) is the leading cause of hospital-acquired infection in
42 the U.S., with an estimated incidence of approximately 224,000 cases per year.[1] The incidence of
43 community-acquired *C. difficile* infection (CDI) is also increasing, with up to 51.2 cases /100,000
44 population.[2] Major risk factors for developing CDI include broad-spectrum antibiotic usage, as
45 antibiotics deplete the commensal bacterial species that provide colonization resistance against
46 CDI. However, CDI disease severity in humans ranges from asymptomatic colonization to diarrhea
47 to severe pseudomembranous colitis and even death. Recent studies have revealed that the gut
48 microbiota and host immunity impact CDI pathogenicity,[3] but genetic features of *C. difficile* likely
49 also regulate disease severity in humans.

50 *C. difficile* virulence requires the production of at least one of its glucosylating toxins,
51 TcdA and/or TcdB.[4] These toxins are internalized by epithelial cells via endocytosis[5–7] and
52 released into the cytosol, where they glucosylate and inactivate intracellular GTPases.[8–10] These
53 activities disrupt cell signaling pathways and cytoskeletal structure, leading to cell death.[11] The
54 impact of TcdA and TcdB on disease progression and severity, however, can vary, depending on
55 the toxin titer, the strain of *C. difficile*, and the host's innate and adaptive immune defenses.[12–14]

56 In addition to toxin expression, flagellar-mediated motility has been suggested to modulate
57 *C. difficile* virulence. Flagellar motility contributes to the virulence of *S. enterica* and *E. coli* by
58 enhancing host cell invasion, adhesion to epithelial cells, and systemic inflammation.[15–17] However,
59 the impact of flagellar motility on *C. difficile* infection *in vivo* remains controversial and is likely
60 strain- and host-dependent.[18–20] Deletion of *fliC*, which encodes flagellin, did not impact the
61 virulence of *C. difficile* R20291 in a germ-free mouse model, yet it reduced its virulence in
62 antibiotic-treated SPF mice.[18,19] In contrast, the deletion of *fliC* in *C. difficile* strain CD630

63 increased its virulence in a hamster model of CDI.[20] Notably, flagellar genes are heterogeneously
64 expressed in *C. difficile* due to the inversion of a flagellar switch sequence by the tyrosine
65 recombinase RecV.[21] Inversion of this region in the 5'UTR of *flgB* to the ON orientation allows
66 for expression of a flagellar gene operon that includes the gene encoding the sigma factor, SigD,
67 which directly induces the expression of flagellar gene operons as well as *tcdR*, which encodes
68 another sigma factor that activates toxin gene expression.[22] Thus, flagellar gene expression is
69 coupled with toxin gene expression.[21,23] Despite these insights, the impact of *C. difficile* flagella
70 and the heterogeneity of the flagellar gene expression on CDI virulence remains unclear.

71 In this study, we investigated the relationship between flagellar gene expression and
72 virulence in *C. difficile* ST1 strains isolated from patients with CDI. Strains of the multi-locus
73 sequence type 1 (ST1), also known as the NAP1/B1/ribotype 027, are highly transmissible, have
74 increased multidrug resistance,[24,25] and were initially reported to be hyper-virulent.[26,27] However,
75 heterogeneity in the virulence[28,29] and toxin production of ST1 strains[30,31] have been reported,
76 revealing that there is considerable variation between individual ST1 trains. By examining 22 ST1
77 clinical isolates in a mouse model of infection, we observed marked differences in virulence. We
78 found that sequence differences in the flagellar switch region, specifically in the inverted repeat
79 (IR) region, of these ST1 strains correlated with disease severity. Our data reveal that these
80 sequence differences alter the invertibility of the flagellar switch, with longer IR sequences causing
81 greater heterogeneity within the population, i.e., a mixture of flagellar gene ON and OFF, and
82 shortened IR sequences being associated with more fixed populations of either flagellar gene ON
83 or OFF. Furthermore, exchanging a longer IR sequence with a shorter IR sequence in *C. difficile*
84 R20291 significantly reduced its virulence, particularly when combined with the OFF flagellar
85 switch orientation. Our results argue that flagellar switch IR sequences alter the invertibility of the

86 flagellar switch, which contributes to the considerable heterogeneity in virulence observed
87 between *C. difficile* ST1 strains.
88
89 **RESULTS**
90 **The flagellar expression associates with the virulence of clinical *C. difficile* RT027 isolates**
91 We previously showed that the *in vivo* virulence of *C. difficile* ST1 isolates in C57BL/6
92 mice varies widely, ranging from mortality to the absence of any weight loss or diarrhea.[31] While
93 we showed that the avirulence of 2 ST1 strains was due to a small internal deletion in the *cdtR*
94 gene, the remaining ST1 strains did not exhibit any genetic variability in their pathogenicity and
95 CDT loci, which encode the TcdA and TcdB glucosylating toxins and the binary toxin CDT,
96 respectively[31] (**Figure S1A**). In accordance, the virulence differences (% weight loss) observed
97 between these strains did not correlate with the fecal toxin levels measured using a cell-based
98 toxicity assay (**Figures S1B-C**). In addition, no correlation in the colonization (CFU) levels on
99 day 1 post-infection and virulence of the strains was observed (**Figure S1B-S1D**). To identify
100 additional mechanisms for variable virulence among ST1 strains, we selected two isolates that
101 induced >10% weight loss (ST1-12 and ST1-53) and two isolates that resulted in < 10% weight
102 loss (ST1-6 and ST1-27) for further study (**Fig. S1**). Here, we refer them as high-virulence and
103 low-virulence isolates.
104 We first compared the transcriptional profile of the four isolates during infection of
105 antibiotic-treated mice by harvesting cecal contents one-day post-infection and conducting
106 RNAseq profiling of the isolates (**Figure 1A**). The transcriptomic profile revealed that flagellar-
107 related genes are over-expressed in high-virulence isolates relative to low-virulence isolates
108 (**Figure 1B**). These differences in flagellar gene expression were also observed when RT-qPCR

109 was used to compare the expression of two flagellar genes, *fliE* and *fliS1*, between the strains using
110 the same cecal content samples (**Figure S2A**). High-virulence isolates also exhibited greater
111 flagellation and spread further on swim plates 24-hours after inoculation (**Fig. 1C-1D)**. In contrast,
112 the low-virulence isolates were aflagellate and exhibited delayed spreading on the swim plates
113 (**Fig. 1C-1D)**. Notably, the RNA-Seq analyses confirmed that toxin genes, including *tcd* genes and
114 *cdt* genes, were expressed at similar levels between the four strains during murine infection (**Fig.**
115 **1B**), consistent with the similar levels of fecal toxins detected during infection in mice. While ST1-
116 27 appeared to colonize mice at higher levels compared to the other three isolates (**Fig. S2B-S2C**),
117 the growth of these 4 isolates did not differ from each other in broth culture (**Fig. S2D**).
118 Interestingly, according to the RT-qPCR results performed with the same cecal content samples,
119 the high-virulence strain ST1-12 expressed significantly higher *tcdA* and *tcdB* relative to the other
120 three strains, including the other high-virulence strain ST1-53. Though it did not fully reproduce
121 the RNAseq observation, it suggests again that toxin production differences cannot fully explain
122 the variation in virulence observed among ST1 isolates (**Figure S2A**).
123

124 **Variation in the flagellar switch inverted repeats of RT027/ST1 clinical isolates.**

125 The difference in flagellar gene expression between high-virulence and low-virulence
126 isolates led us to compare their flagellar genomic regions. While the flagellar genes were identical
127 between the isolates, we identified variation in the flagellar switch region, which plays an
128 important role in regulating the expression of the flagellar operons.[21] The flagellar switch region
129 comprises a central region flanked by a left-inverted repeat (LIR) and a right-inverted repeat (RIR)
130 (**Figure 2A**). The inverted repeats are recognized by the tyrosine recombinase RecV, which inverts
131 the central region. This leads to the switch region exhibiting either an ON or OFF orientation,

132 which allows or inhibits flagellar gene expression, respectively [21]. When we expanded the analysis
133 of the flagellar switch region to all 68 ST1 isolates in our collection, we found that 65 of the isolates
134 have a flagellar switch central sequence identical to the reference RT027/ST1 strain, R20291, and
135 thus carried for further analyses. Most notably, we found variability in the sequence of the left and
136 right IRs that flank the central flagellar switch region. Specifically, while most isolates have IRs
137 identical to the reference strain *C. difficile* R20291, 40% of isolates have at least one less A or T
138 in either the left or right IRs (**Figure 2A and Table S1**). In total, four types of IR flanking regions
139 were identified in our strain collection; we named them: Common1 (C1-59.38% 6A/6T),
140 Common2 (C2-32.81%, 6A/5T-ON or 5A/6T-OFF), Rare1 (R1-6.25%, 5A/6T-ON or 6A/5T-
141 OFF), and Rare2 (R2-1.56%, 5A/5T). The C1 switch region, the most common sequence, is
142 identical to that observed in *C. difficile* R20291, with 6 A's on the left IR (LIR) and 6 T's on the
143 right IR (RIR). When C1 inverts between the OFF vs. ON orientation, the number of A's and T's
144 in both IRs remains the same. In contrast, the composition of the Common 2 (C2) region differs
145 depending on the orientation of the switch region: in the OFF orientation, the LIR consists of 5
146 A's and the RIR consists of 6 T's (C2-OFF); in the ON orientation, the LIR consists of 6 A's and
147 the RIR consists of 5 T's (C2-ON). The Rare 1 region (R1) also has an asymmetric distribution
148 depending on the switch orientation except that in the OFF orientation, the LIR consists of 6 A's
149 and the RIR consists of 5 T's (R1-OFF). The Rare 2 region (R2) IRs remain the same regardless
150 of the switch orientation, with the LIR consisting of 5 A's and the RIR consisting of 5 T's.
151 However, the R2 is only observed in one strain, ST1-67, where it is in the OFF orientation.
152 Notably, Sanger sequencing of the flagellar switch region confirmed the HiSeq whole genome
153 sequencing analyses (**Table S1**).

154 To assess the conservation of the IR types observed, we screened 1,359 RT027/ST1 isolates
155 whose whole-genome sequences were available from NCBI BioSamples. These isolates derive
156 from the CDC HAI-Seq *C. difficile* collection (PRJNA629351)[32] and Texas and Michigan medical
157 centers (PRJNA595724, PRJNA561087).[33] While isolates from more recent publications were also
158 included,[34–37] all *C. difficile* isolates were collected from 1988 to 2020. Most isolates (1318/1359)
159 derive from Continental Europe, UK, Ireland, and the U.S., with a small number (20/1359) of
160 isolates from East Asia and Australia. Most *C. difficile* ST1 isolates carry the flagellar switch IR
161 type C1, representing 96.5% (1311/1359) of isolates analyzed. Type C2 was observed in only 2.4%
162 (33/1359) of isolates, of which most (27/33) were obtained in the U.S between 2007 and 2020.
163 The other IR types were detected at even lower frequencies, representing only 0.66% of isolates
164 (**Figure 2B**).
165 When more recent CDC HAI-Seq collection strains were analyzed, which were
166 predominantly submitted between 2009 and 2020, the IR type C1 remained dominant, but the C2
167 represented 30% (19/63) isolates, similar to our collection (32.8%). Interestingly, the proportion
168 of C2 increased in isolates collected in 2020, accounting for 53% (9/17) of isolates collected in
169 2020 (**Figure 2C**). Thus, our data suggest that variability in *C. difficile* flagellar IR regions has
170 increased over time, although their impact on *C. difficile* virulence remains unclear.
171
172 **Flagellar invertibility has a strong association with virulence of RT027/ST1**
173 To analyze the impact of inverted repeat types on RT027/ST1 virulence, we correlated the
174 weight loss in mice caused by *C. difficile* infection with different IR types. We found a strong
175 correlation between IR type C1 and more severe weight loss in infected mice ($p = 3.2e-05$) (**Figure**
176 **2A and 3A**). Since a small deletion in the RIR and surrounding region greatly reduces the

177 invertibility of *C. difficile* flagellar switch, essentially locking it in one orientation,[23] we tested
178 whether the different IR types impact the invertibility of the flagellar switch region. Using
179 orientation-specific qPCR analyses to quantify the fraction of ON and OFF cells in *C. difficile*
180 population from liquid culture (**Figure 3B**),[21] we found a weak correlation between greater weight
181 loss and strains with a higher proportion of ON cells in a population by linear regression (R^2 =
182 0.39). However, the most virulent isolates had ~70 % of cells in the ON orientation and consisted
183 of the C1 IR type (**Figure 3C**, green circles). Furthermore, isolates with the other IR types
184 exhibited lower virulence than the C1 isolates and were either nearly 100% ON or 100% OFF
185 (**Figure 3C**). Thus, our results suggest that IR type governs the flexibility of the flagellar switch
186 region and that IR types that lead to more heterogeneous flagellar gene expression, i.e. the C1 IR
187 type, enhance the virulence of *C. difficile* strains.
188
189 **The Inverted repeat type correlates with the invertibility of the flagellar switch.**
190 To further investigate whether different IR types impact the flagellar switch invertibility,
191 analyzed the flagellar orientation of our 64 ST1 isolates across 3-4 biological replicates per isolate.
192 We cultured individual isolates in broth to exponential growth and prepared the genomic DNA for
193 orientation-specific qPCR analyses. When calculating the proportion of C1 cultures in the ON
194 orientation, we found that the C1 IR type is slightly biased toward the ON orientation, with
195 approximately 73% (55-99%) in the ON orientation (**Figure 4A**). In contrast, the biased-OFF
196 isolates, including C2-OFF, R1-OFF and R2-OFF, have 95.6% (88.7-100.0) of their population in
197 the OFF orientation, whereas the biased-ON isolates, including C2-ON and R2-ON, have 96.7%
198 (92.6-100.0) of their population in the ON orientation (**Figure 4A**). Such a strong correlation

199 between IR types and biased-flagellar orientations indicates that mutations in the IR regions greatly
200 impact the switching invertibility.

201 To test if the biased IR types also restrict the invertibility of *C. difficile* isolates *in vivo*, we
202 infected mice with *C. difficile* R20291 and clinical isolates with different IR types and monitored
203 the ON and OFF cell fractions in fecal or cecal contents over the course of eight days. On day 1
204 post-infection, we found that *C. difficile* isolates with the C1 IR type had about ~75% of their
205 population with the ON orientation (**Figure 4B**), similar to our broth culture analyses (**Figure 4A**).
206 Interestingly, the proportion of C1 IR type strains in the ON vs. OFF cells exhibited greater
207 variability over the course of the infection compared to the isolates with IR types C2, R1, and R2.
208 These latter three IR type strains maintained their flagellar switch region in a biased orientation
209 over the eight-day infection course (**Figure 4B and S3**). These data further highlight the strong
210 correlation between IR type and flagellar switch invertibility. Specifically, the C1 IR type (6A/6T)
211 exhibits greater invertibility in its flagellar switch region, while the other IR types appear to "lock"
212 the flagellar switch in either the ON or OFF orientation.

213

214 ***C. difficile* flagellar IR type impacts RecV-mediated DNA inversion.**

215 Since IR types correlate with the invertibility of ST1 isolates (**Figure 4**), we wondered if
216 the IR types specifically affect RecV-mediated DNA inversion. To test this possibility, we assessed
217 the invertibility of IR types in a heterologous host using an *E. coli*-based colorimetric assay.[38] *E.*
218 *coli* were co-transformed with plasmids encoding *C. difficile recV* and a second plasmid containing
219 a promoter flanked by either the C1, C2 or R2 IR type. Depending on the orientation of the
220 promoter within the switch region flanked by the IR variants, it will drive the expression of either
221 *gfp* or *rfp*. Thus, two variations of the reporter plasmid were transformed into *E. coli* for each IR

222 type variant: One with the promoter initially driving *gfp* (left panel), and the second with the
223 promoter driving *rfp* (right panel). RecV-mediated DNA inversion of the reporter plasmid results
224 in a switch from red to green (or green to red) (**Figure 5A)**. When the reporter plasmid carried the
225 C1 IR type, a high percentage of dual plasmid-transformed colonies exhibited a color switch
226 (**Figure 5B).** In contrast, when the reporter plasmid carried the C2 IR type, only 1-2 colonies
227 exhibited a color switch, whereas no colonies were observed to have color-switched when the
228 reporter plasmid carried the R2 IR type (**Figure 5B**). These results align with our qPCR analyses
229 of the switch region, strongly suggesting that the IR sequence controls the invertibility of the
230 flagellar switch region.
231
232 **IR determines the flexibility of flagellar switch in *C. difficile* and impacts *in vivo* virulence.**
233 To directly test the ability of the flagellar IR type to alter the virulence of *C. difficile* in
234 mice, we determined the impact of mutating the IR switch region of the reference RT027/ST1
235 strain, R20291, on its virulence using CRISPR.[39] To this end, we deleted ~200 bp of the region
236 upstream of the R20291 *flgB* operon first and then knocked in the wild-type C1 IR (6A/6T), as
237 well as the variants including R2 variants 5A/5T-ON (R2-ON) and 5A/5T-OFF (R2-OFF). It was
238 necessary to generate the R2-ON and R2-OFF variants because our data suggested that the R2 IR
239 sequence greatly reduces the inversion of the flagellar switch region (**Figures 4, 5**). Consistent
240 with our prior data, we observed that the R2-ON resulted in almost all cells in the population
241 carrying the ON orientation (99.9%), while almost all cells in the R2-OFF population carrying the
242 OFF orientation (0.02%) (**Figure S4A**). Furthermore, the R20291_R2-ON (5A/5T-ON) variant
243 exhibited spread using flagellar motility to similar levels as the parental R20291 strain, while the
244 R20291_R2-OFF (5A/5T-OFF) was largely non-motile (**Figure S4B**).

245 To explore if these R20291 flagellar IR variants have differential virulence in a mouse
246 model, we infected antibiotic-treated mice with the mutant strains and monitored their weight loss
247 (**Figure 6A**). All mutant strains colonized mice to similar levels as the parental R20291 strain,
248 although more variability in colonization levels was observed in strains carrying the flagellar
249 switch in the ON orientation (R20291, **Figure 6B**). Interestingly, fecal CFU levels were below the
250 limit of detection for 4 out of 10 mice infected with R20291_R2-OFF (5A/5T-OFF) and 3 out of
251 20 mice infected with R20291_R2-ON (5A/5T-ON), implying that the ability to invert the flagellar
252 switch region enhances the initial colonization of *C. difficile* (**Figure 6B**). To focus on the impact
253 of the IR variants on the infection dynamics of *C. difficile*, we excluded these 7 mice from the
254 downstream analyses in mice. All colonized strains maintained high levels of *C. difficile*
255 colonization and shedding up to 14 days post-infection (**Figure 6C and S4C**). Mice infected with
256 the parental R20291 strain exhibited severe weight loss on day 2 post-infection, while the clinical
257 isolate ST1-67, which carries the R2-OFF variant (5A/5T-OFF) did not induce weight loss,
258 consistent with our previous observations(**Figure 6D**).[31] Mice infected with the "restored" C1
259 variant, R20291_C1 (6A/6T) phenocopied the weight loss of the parental R20291 strain (**Figure**
260 **6D and 6E**), while the R20291 variant carrying R2-OFF, R20291_R2-OFF (5A/5T-OFF), caused
261 minimal weight-loss, essentially phenocopying the avirulence of the ST1-67 strain, which carries
262 the flagellar switch in the R2-OFF orientation (5A/5T-OFF) (**Figure 6D and 6E**).

263 Since orientation-specific qPCR revealed that *C. difficile* strains carrying the R2 IR
264 (5A/5T) IRs remained skewed to either ~100% ON or ~100% OFF throughout the infection time
265 course (**Figure 6F**), these data confirm that the R2 IR type largely prevents the flagellar switch
266 from inverting. Notably, since the R20291_R2-OFF variant caused only mild disease, and the only
267 change between the R20291_R2-OFF variant and the parental R20291 strain is a single nucleotide

268 in each of its flagellar switch IR sequences, these data demonstrate that the loss of flagellar gene
269 expression at the population level reduces *C. difficile*'s virulence. In contrast, the R20291_R2-ON
270 (5A/5T-ON) variant exhibited a virulence phenotype in comparison to R20291_R2-OFF (5A/5T-
271 OFF) (**Figure 6D and 6E**), strongly suggesting that flagellar gene expression during infection
272 enhances *C. difficile*'s virulence. However, it should be noted that a majority of mice (11/16, 69%)
273 infected with R20291_R2-ON showed a low-virulence phenotype, causing less than 10% weight
274 loss (which we defined as "low-virulence" for the clinical isolates, **Figure S1**), whereas only 6/17
275 (35%) of mice infected with the parental R20291 or R20291_C1 variant exhibited less than 10%
276 weight loss. Additionally, pooling the R20291 strains tested by their IR types revealed that *C.*
277 *difficile* strains carrying the C1 IR type (6A/6T) are significantly more virulent than their
278 counterparts carrying the R2 IR type (5A/5T), regardless of the orientation of the flagellar switch
279 (**Figure 6G**). This is consistent with our finding that clinical isolates carrying the C1 IR type are
280 associated with greater virulence (**Figure 3A**). Taken together, these data demonstrate that the IR
281 types impact the invertibility of *C. difficile* flagellar switch during mouse infection, contributing
282 to heterogeneity in flagellar gene expression at the population level. This heterogeneity appears to
283 promote more severe disease because greatly reducing the invertibility of the flagellar switch
284 region by introducing the R2 IR type variant reduces disease severity in the R20291 strain
285 background, particularly when the R2 IR variant is fixed in the OFF orientation (**Figure 6**).
286
287 **DISCUSSION**
288
289 *C. difficile* infection induces a wide range of disease severity in patients, yet the molecular
290 basis for the variation is poorly understood. Here, we applied a mouse model to study how genetic

291 variation between *C. difficile* isolates impacts their virulence. Using a collection of more than 60
292 ST1/RT027 clinical isolates, we discovered that variation in the inverted repeat (IR) regions
293 flanking *C. difficile*'s flagellar switch region regulates the invertibility of this switch by the RecV
294 recombinase. Our data indicate that the most common variant found in ST1 strains, C1 (6A/6T),
295 results in greater heterogeneity in flagellar gene expression within a population, with ~70% of the
296 population carrying the flagellar switch in the ON orientation (**Figures 4, 6**). We further showed
297 that the C1 variant leads to greater disease severity relative to strains with shorter IR variants
298 (**Figure 3, 6G**). The shorter IR variant types - C2 (6A/5T), R1 (5A/6T), and R2 (5A/5T) - restrict
299 the RecV-mediated invertibility of the flagellar switch, effectively "locking" the flagellar switch
300 region into either the ON or OFF orientation (**Figure 4 and 5**). Notably, the decreased invertibility
301 of the flagellar switch region was associated with decreased virulence in ST1 clinical isolates
302 (**Figure 3**) as well as in CRISPR-engineered R20291 strains (**Figure 6**), especially when the
303 engineered R2 switch was in the OFF orientation. Since the IR sequences are sufficient to regulate
304 the inversion of promoter regions in *E. coli* (**Figure 5**) and *C. difficile* R20291 (**Figure 6**), our data
305 reveal that flagellar switch regions IR types control the inversion frequency of this locus (i.e.,
306 phase variation), which impacts the virulence of *C. difficile* ST1 strains. In demonstrating that
307 "fixing" the flagellar switch region in the OFF orientation can render ST1 strains avirulent (**Figure**
308 **6**), our study reveals another mechanism by which RT027 strains can cause disease of varying
309 severity, which may help explain the wide range of disease outcomes observed in CDI patients.

310 The invertibility of the flagellar switch likely confers an evolutionary advantage, as more
311 than 95% of isolates worldwide harbor C1 (6A/6T) IR type. Interestingly, in both our strain
312 collection and in the CDC HAI-Seq *C. difficile* strain collection, we found that the C2 (6A/5T)-
313 OFF IR variant accounts for approximately 30% of isolates. This could be due to rigorous testing

314 and *C. difficile* cultivation in the U.S.,[40,41] as C2-OFF isolates are mostly non-flagellated and are
315 associated with reduced virulence. Since flagellar synthesis and motility is costly[42] and flagella are
316 targeted by mucosal innate immune response,[19,43] reducing flagellar motility could be energetically
317 favorable and allow *C. difficile* to evade the immune system as a strategy to increase the persistence
318 of *C. difficile*. However, the ability to switch between flagellar ON and OFF (and vice versa) could
319 enhance *C. difficile*'s fitness during infection by allowing sub-populations of *C. difficile* expressing
320 flagellar genes to inhabit different locations within the gut or function at different stages during
321 the infection process. For example, flagellar motility may help bring *C. difficile* cells closer to the
322 host epithelium to improve toxin binding to its target cells, which could explain why C1 (6A/6T)
323 isolates, which have the most flexible flagellar switch, exhibit the highest virulence among the
324 ST1 strains tested. Future investigations could focus on examining the relative locations of
325 flagellar-ON vs. flagellar-OFF cells for *C. difficile* with various IRs to gain mechanistic insight
326 into the relationship between flagellar heterogeneity and *in vivo* virulence.
327
328 Bacteria apply phase variation to reach population phenotypical heterogeneity to promote
329 adaptation to various environmental conditions. Such heterogeneity may increase motility,
330 antibiotic resistance and pathogenic potentials.[44,45] Notably in *C. difficile*, RecV regulates multiple
331 phase variable regions including flagella as well as a cell wall protein CwpV.[46,47] CwpV deposited
332 on a proteinaceous layer on the cell surface of *C. difficile* promotes bacterial aggregation *in*
333 *vitro* and potentially promotes intestinal colonization.[48] Moreover, *C. difficile* regulates it surface
334 motility via regulation by CmrRST system, which induces cell chaining phenotype on surfaces
335 and contributes to disease development in hamsters.[49,50] Here, we observed the heterogeneity of
336 *C. difficile* flagella upon infection in mice and the flagellar heterogeneity is important for *C.*

337 *difficile* virulence. Together, we and others underscore the importance of phenotypical
338 heterogeneity in bacterial survival and virulence and encourage more studies on such phase
339 variable traits of otherwise genetically identical strains.
340
341 *C. difficile* flagellar expression has been also linked to toxin production. Turning on the
342 expression of flagellar operon increases the expression of SigD, which is a sigma factor that not
343 only further regulates flagellar synthesis and motility but also positively regulates *tcdR* for toxin
344 production.[22] Thus, whether flagellar expression contributes to virulence beyond regulating toxin
345 levels is unclear. Here, we demonstrated that the engineered R20291 mutant with a biased ON
346 flagellar orientation is significantly more virulent than its counterpart with a biased OFF flagellar
347 orientation, supporting the role of flagella in *C. difficile* virulence. However, the increase of
348 virulence by tuning ON flagella did not surpass the virulence of *C. difficile* with heterogeneous
349 flagellar orientations, as for both clinical isolates and CRISPR-engineered R20291 mutants.
350 Moreover, we did not find an association between toxin levels in fecal samples and the degree of
351 virulence on a panel of 22 isolates. These suggest that the expression of flagellar (and with coupled
352 toxin expression) provides a relatively minor contribution to virulence and is likely to be strain-
353 dependent, while the invertibility of the flagellar switch may provide major fitness and impact on
354 *C. difficile* virulence (i.e. via impacting the locations of toxin production).
355
356 The flagellar phase variation is regulated by DNA recombinase RecV.[21,51] RecV reversibly
357 inverts the flagellar switch between ON and OFF orientations, and the proportion of ON vs. OFF
358 in the resulting population is likely selected by the residing environments.[52] Thus, the flexible
359 switch leads to more variable ON and OFF compositions spanning different conditions (from

360 vegetative broth culture and prepared spores to mouse infection), while the less flexible switches
361 result in their highly biased orientation all the time, regardless of the conditions. Using a
362 heterogenous *E. coli* strain, we demonstrated that the flagellar IR type variants reduce RecV-
363 mediated DNA inversion. This is further confirmed with CRIPSR-engineered mutants as R20291
364 with R2-5A/5T IRs demonstrated highly biased flagellar orientations. RecV targets IRs for
365 inversion, and the potential recombination sites were speculated to be at the borders of these IRs.[51]
366 Notably, deleting the 1st T and its adjacent nucleotides (Δ3) also "locked" the flagellar orientation
367 into ON or OFF.[23] Together with our observation, lacking one T in the 6T track in Right IR is
368 likely impacting the efficiency of RecV function. Further experiments should dissect out at what
369 steps that the IR variants impact RecV, such as sequence-specific DNA binding, DNA cleavage,
370 strand exchange, or religation.

371 In summary, our study identifies the flagellar switch IR as a determinant of the
372 heterogeneity *C. difficile* flagellar phase variation, which also introduces variable virulence
373 outcomes within a single RT027 strain type. Since the invertibility of the flagellar switch is highly
374 associated with the virulence of clinical isolates, we highlight the potential of using flagellar switch
375 inverted repeats as an easily accessible genetic trait to predict pathogen virulence. Further research
376 on the flagellar switch regions of clinical non-ST1 strains may provide additional insights into the
377 wide range of disease severities in patients infected with *C. difficile*.

378

379 **Acknowledgement**

380 We would like to thank Dr. Rita Tamayo for providing the primers and reference strains for the
381 flagellar orientation-specific qPCR. We thank Dr. Craig D. Ellermeier for generously providing
382 us the CRISPR editing system of *C. difficile* and Dr. Louis-Charles Fortier for providing *C. difficile*

R20291 strain. We thank Dr. Femi Olorunniji for providing the recombinase-switch plasmid. We thank the animal facilities of the University of Chicago and Tufts University for their help with mouse experiments. We thank the Pamer lab members and Shen lab members for helpful discussions. This work was supported by National Institutes of Health R01 AI095706 (to E.G.P), R21 AI168849 (to A.S), R35 GM149586 (to P.A.R), the Duchossois Family Institute of the University of Chicago, and a Burroughs Wellcome Fund Investigators in the Pathogenesis of Disease Award to A.S. The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication. The graphical schematics were created with BioRender. com.

Author contributions

N.T.Q.N, Q.D., A.S. and E.G.P. conceived the project. N.T.Q.N, Q.D., H.L. and P.A.R. analyzed the data. N.T.Q.N., Q.D., Y.P., J.K.S., and P.K., performed experiments. M.K. isolated *C. difficile* isolates. V.B.Y. and E.S.S. sequenced clinical isolates. N.T.Q.N., Q.D., A.S., P.A.R. and E.G.P. interpreted the results and wrote the manuscript.

Declaration of interests

None.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used Grammarly in order to proofread the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

405 **Supplemental information titles and legends**

406 Document S1: Supplemental figures and legends

407 Table S1: Sanger sequencing on flagellar inverted repeats

408 Table S2: Oligos used for CRISPR

409 Table S3: List of plasmids used in the study

410

411 **Figure titles and legends**

412 **Figure 1**: **High virulence isolates make more flagella**.

413 (A) The schematic of the mouse infection. (B) Heat map of RNA sequencing results from the
414 cecal contents of mice infected with the four ST1 isolates shown. The heat map is generated
415 based on the number of transcripts per sample after normalization. RNA was extracted from the
416 cecal contents of mice one-day post-infection (n= 5 mice per isolate). (C) Transmission electron
417 micrographs of the four isolates. (D) Swim plates results of four isolates after 24 hours
418 incubation (n = 3 replicates per isolate). High-virulence isolates: pink. Low-virulence isolates:
419 black. Statistical significance was calculated by One-way ANOVA, * $p < 0.05$.

420

421 **Figure 2: Variations in flagellar inverted repeat types observed among ST1 isolates.**

422 (A) Schematic of the flagellar switch and the alignment of the flagellar switch regions from
423 representative isolates. In Illumina sequencing the flagellar switch region of the Common 1 IR
424 type (C1, 6A/6T) is observed in the ON orientation. The Common 2 IR type (C2, 6A/5T) is
425 observed in isolates with the flagellar switch region being both in the ON and OFF orientations.

426 Rare 1 (R1-5A/6T) was observed in isolates with the flagellar switch being ON and OFF. The
427 Rare 2 IR type (R2, 5A/5T) was observed only in one isolate with the flagellar switch region
428 being in the OFF orientation. Full list is in Table S1. (B) IR types of isolates collected from
429 around the world. (C) IR types of isolates collected in the US and by the CDC.

430

431 **Figure 3: High virulence isolates exhibit more heterogeneous flagellar gene expression.**
432 (A) %maximum weight loss graphed by IR types. (B) Schematic of the qPCR strategy to measure
433 the proportion of flagella - ON vs. OFF cells in the population. (C) Correlation between %
434 maximum weight loss and % flagella – ON cells. DNA was extracted from bacterial cultures during
435 logarithmic growth.

436

437 **Figure 4: Flagellar inverted repeat types are associated with flagellar switch invertibility in**
438 **culture and in mice.**
439 (A) Proportion of the population with the flagellar switch region in the ON orientation during
440 logarithmic growth in broth culture and (B) during infection over time as determined by qPCR.
441 The results reflect samples obtained from the feces (day1-7) and cecum (day 8) of *C. difficile*-
442 infected mice. Statistical significance was calculated by One-way ANOVA, ** $p < 0.01$, , *** $p <$
443 0.001, **** $p < 0.0001$.

444

445 **Figure 5: Flagellar inverted repeat (IR) types determine the flagellar switch invertibility.**
446 (A) Schematic of the color switch assay in *E. coli*. Two plasmids were transformed into *E. coli.*
447 One plasmid encodes *C. difficile* RecV. The other plasmid contains an invertible sequence where
448 the promoter is flanked by the inverted repeat (IR) variants. Depending on the orientation of the

449 promoter, either *gfp or rfp* will be expressed, and color switch from red to green (or green to red)
450 is expected upon RecV acting on flexible switches. (B) Color switching assay when different types
451 of flagellar inverted repeats flank the invertible region. The switch frequency of the Common 1
452 (C1-6A/6T), Common 2 (C2-5A/6T), Rare 2 (R2-5A/5T) IR variants was visualized. White arrows
453 point to colonies that underwent a promoter inversion (i.e. switched colors). Two dilutions of the
454 cultures were plated (left = lower dilution).
455
456 **Figure 6. The flagellar repeat variants and flagellar switch orientation impact the virulence**
457 **of *C. difficile* R20291.**
458 (A) Schematic of the experimental procedure. Wild-type C57BL/6 mice (n = 8-20 per group) were
459 treated with Metronidazole, Neomycin and Vancomycin (MNV, 0.25 g/L for each) in drinking
460 water for 3 days, followed by one intraperitoneal injection of Clindamycin (200 mg/mouse),
461 indicated as C in the schematic, 2 days after antibiotic recess. Then, mice were inoculated with
462 200 indicated *C. difficile* spores via oral gavage. Daily body weight was monitored for 7 days post-
463 infection. (B-C) Fecal colony-forming units were measured by plating on selective agar on 1 day
464 (B) and 14 days (C) post-infection. (D) %Weight loss relative to the baseline of mice infected with
465 the indicated strains up to 7 days post-infection. (E) % Maximum weight loss relative to the
466 baseline of mice infected with indicated strains up to 7 days post-infection. (F) % Flagellar ON vs
467 OFF cells in spore inoculum and in fecal pellets by qPCR. (G) % Maximum weight loss relative
468 to the baseline of mice infected with strains (R20291 background only) grouped by IR types.
469 Statistical significance was calculated by unpaired t-test and One-way ANOVA, * $p < 0.05$, ** p
470 < 0.01, *** $p < 0.001$, **** $p < 0.0001$.
471

472 **Methods**

473 **Clinical *C. difficile* isolates collection**

474 Clinical *C. difficile* isolates were collected from 2013 to 2017 at Memorial Sloan Kettering
475 Cancer Center (MSKCC) from patients receiving bone marrow transplants and cancer
476 chemotherapy.[40] The isolates were sequenced whole genome in an Illumina Hiseq platform[33] and
477 circularized at Duchossois Family Institute by MinION Nanopore sequencing (Oxford Nanopore
478 Technologies). Whole genome sequences are available at National Center for Biotechnology
479 Information, BioProject PRJNA595724.

480

481 **Bacterial growth and spore collection**

482 Frozen stock of *C. difficile* was struck onto Brain heart infusion (BHI) agar plus 0.1% (w/v) sodium
483 taurocholate hydrate (Cat. 86339, Sigma-Aldrich). A single colony was picked and sub-cultured
484 in BHI medium (BD, 237500) with yeast extract and 0.1% L-Cysteine (BHIS) at 37ºC in anaerobic
485 chamber (Coy lab). To prepare spores for the mouse infections, *C. difficile* was either incubated in
486 BHIS broth for approximately two months or on 70:30 agar for 4-5 days to encourage
487 sporulation.[53] Next, spores were separated from cell debris by gradient centrifugation using a
488 20%/50% (wt/vol) HistoDenz (D2158, Sigma-Aldrich) or 50% (wt/vol) sucrose gradient, then
489 washed five times in sterile water at 14,000 x g for 5 minutes.[53] The spore solution was further
490 incubated at 65ºC for 20 minutes to kill vegetative cells. Spore purity was confirmed by the
491 absence of vegetative cell growth on BHIS plates or by microscopy.

492

493 **Mouse experiment**

494 C57BL/6 female mice from six- to eight-week-old were purchased from Jackson laboratory and
495 housed in the specific-pathogen-free (SPF) facility at University of Chicago or the Tufts University
496 School of Medicine. Mice were randomized and administered *ad libitum* with a cocktail of
497 metronidazole (0.25g/l), neomycin (0.25g/l), and vancomycin (0.25g/l) for three days. Two days
498 after antibiotics removal, mice were intraperitoneally injected with 200µg/mouse clindamycin.
499 After 24 hours, mice were oral gavage with approximately 200 spores of *C. difficile*. All mice
500 infected with *C. difficile* were single-housed then monitored for weight loss and clinical sign of
501 disease for seven days. CFU/g feces counting, and toxin titer measurement was done from fecal
502 pellet collected one day after infection as described elsewhere.[31]
503
504 **RNA sequencing**
505 Mice were infected with spores from the four selected isolates (ST1-6, ST1-12, ST1-27, and ST1-
506 53). After 24 hours, mice were sacrificed to collect cecal content for RNA extraction (Qiagen
507 RNeasy PowerMicrobiome kit). Total RNA was sent to Genewiz for rRNA removal and library
508 preparation before sequencing. Adapters were trimmed off from the raw reads, and their quality
509 was assessed and controlled using Trimmomatic (v.0.39),[54] then human genome was identified
510 and removed by kneaddata (v0.7.10, https://github.com/biobakery/kneaddata), while ribosomal
511 RNA was removed by aligned the clean non-host reads to silva database (138.1 SSURef).[55–57] The
512 remaining reads from each sample were mapped to their corresponding circularized genome using
513 bowtie2,[58] and reads counts of each gene were obtained by running featureCounts from Subread
514 (v2.0.1),[59] and the core gene counts were normalized by DESeq2.[60]
515
516 **Sequencing the inverted-repeat sequences**

517 64 isolates with the flagellar switch sequence resembling that of *C. difficile* R20291 were
518 submitted for Sanger sequencing. Firstly, the ON and OFF sequence of each isolate was amplified
519 using the primer pairs as designed by.[21] The PCR product was purified using QIAquick PCR
520 purification kit (Cat. 28104, Qiagen). Purified products were sent to University of Chicago DNA
521 Sequencing facility and sequenced using the same primer pairs for PCR reaction.
522
523 **Swim plates**
524 *C. difficile* was grown in BHIS broth to late-log phase then diluted to OD_{600nm} ~0.5. To ensure a
525 small number of *C. difficile* cells was added for testing, we submerged a 10 µL pipette tip into
526 diluted *C. difficile* culture without drawing any liquid. The pipette tip carrying *C. difficile* cells on
527 it was then stabbed into 0.3% agarose buffered with BHIS, pH 7.0. The swim plates were then
528 incubated at 37˚C. After 24 hours, the swim zone size was measured and compared between
529 isolates.
530
531 **Growth curve**
532 To ensure that only live cells were used for growth curve, we prepared a fresh culture from
533 overnight culture at the ratio 1:100. When the back-diluted culture reached log phase, we
534 subculture it into fresh BHI broth at a 1:100 dilution. The culture was then loaded into a 96-well
535 plate for OD_{600} measurements. The incubation time is 24 hours with OD measurements every 10
536 minutes at 37˚C with shaking.
537
538 **Transmission electron microscope (TEM)**

539 *C. difficile* was inoculated from frozen stocks onto a BHIS agar plate and incubated overnight. To
540 collect cells for TEM, 10µl of distilled water was dropped onto a colony. After two minutes, a
541 copper grid (CF400-Cu, EMS) was placed on top of the soaked colony so that vegetative cells
542 were passively transferred to the grid. To stain the cells, one drop of uranyl acetate 1% was added
543 to the grid and incubate for 30 seconds. TEM pictures of the stained cells were then taken at
544 2900X1.4 or 5900X1.4 using FEI Tecnai F30 microscope at University of Chicago Advanced
545 Electron Microscopy Facility.
546
547 **Quantitative PCR of the flagellar switch**
548 *C. difficile* was grown in BHIS broth to late log phase (OD_{600nm} ~ 0.5) with three to four
549 replicates per isolate. *C. difficile* DNA was extracted by Qiamp PowerFecal Pro DNA kit (Cat.
550 51804). Fecal samples were harvested from infected mice and fecal DNA was extracted either
551 using QIAamp PowerFecal Pro DNA Kit or as previously described.[23] Quantitative real-time
552 PCR was done using 20-30 ng template and flagellar switch primers.[21] *C. difficile* R20291
553 mutated strains with flagellar switch locked ON and OFF were used as the control.[23] Adenosine
554 kinase (*adk)* gene is used as a reference gene. qPCR was set up using PowerUp™ SYBR™
555 Green Master Mix (Cat. A25742) and run by QuantStudio Real-Time PCR Systems or NEB
556 Luna qPCR Master Mix and run by Applied Biosystems StepOnePlus system. The switch
557 direction percentage was calculated using the ΔΔCt method as previously reported,[23] while
558 integrating the primers' amplification efficiencies.
559
560 **Generation of *C. difficile* mutants using CRISPR**

561 CRISPR editing on *C. difficile* strains R20291 was performed as described previously.[39] Briefly,
562 donor regions for homology were generated by separately amplifying regions ~500 bp upstream
563 and ~500 bp downstream of the target of interest. The resulting regions were cloned into pCE677
564 between NotI and XhoI sites by Gibson Assembly. Geneious Prime (v11) was used to design
565 sgRNAs targeting each deleted target. sgRNA fragments were then amplified by PCR from
566 pCE677, using an upstream primer that introduces the altered guide and inserted at the MscI and
567 MluI sites of the pCE677-derivative with the appropriate homology region. The regions of
568 plasmids constructed using PCR were verified by Sanger sequencing. Plasmids were then passaged
569 through NEBturbo *E. coli* strain before transformation into *Bacillus subtilis* strain BS49. The
570 CRISPR-Cas9 deletion plasmids which harbor the *oriT* (*Tn916*) origin of transfer, were then
571 introduced into *C. difficile* strains by conjugation.[61] *C. difficile* colonies were then screened for
572 proper mutations in the genomes by PCR and Sanger sequencing. To generate *C. difficile* IR
573 repeats mutants, two rounds of CRISPR editing were conducted. The first round was to delete
574 ~200 bp region containing the flagellar switch and IRs while introducing an *RFP* landing pad
575 (GGCGCCCAGACCGCTAAACTGAAAGTT) into the place. The second round gRNA targeted
576 the *RFP* landing pad, with the mutant IR variant flagellar switch region (either in the ON or OFF
577 orientation) template supplemented for repair. Primers used for CRISPR editing were included in
578 Table S2.
579
580 **Flagella switch alignment and IR type survey in public databases**
581 *C. difficile* isolates (N=131) from BioProject PRJNA595724 and PRJEB2318 (N=1198) were
582 downloaded from NCBI and assembled into contigs using SPAdes.[62] Eleven of those isolates did
583 not pass the assembling process. In addition, a collection of 66 *C. difficile* genomes from Patric
584 (date: Mar. 15 2022, https://www.bv-brc.org/ were also downloaded. MLST was determined on

585 those contigs by mlst (Seemann T, mlst Github https://github.com/tseemann/mlst).[63] Flagellar
586 switch region and 50 bp upstream and downstream of 68 isolates in our collection of both the ON
587 and OFF sequences were used as query to BLAST[64] against the assembled contigs, and hits with
588 at least 85% identity and 85% coverage of the query are considered a valid match.
589

590 ***E. coli* colorimetric assay**

591 Plasmids for the E. coli experiments (see Table S3) were ordered from Twist Bioscience and were
592 based on backbone vectors kindly onboarded with Twist by Dr. Femi Olorunniji (Liverpool John
593 Moores University). Plasmids were checked by full-plasmid sequencing (Plasmidsaurus). pQD1
594 is a pBAD derivative for tightly controlled arabinose-inducible expression of RecV. Test plasmids
595 pQD2-7 are based on pϕC31-invPB described previously,[38] with IRs from the flagellar switch
596 replacing the att site for ϕC31 integrase.

597 Assays were performed as described previously with minor variations.[65] *E. coli* DS941[66]
598 was co-transformed with pQD1 plus one of the test plasmids, then after recovery grown overnight
599 at 37C in 10ml LB supplemented with 0.2% glucose (to enhance repression of RecV), kanamycin
600 (50 µg/ml; to maintain the test plasmid) and chloramphenicol (30 µg/ml; to maintain pQD1). In
601 the morning, the OD_{600} was ~2. 200 µl of that culture was diluted into 10ml LB plus 0.2% glucose,
602 kanamycin (50 µg/ml) and chloramphenicol (30 µg/ml) and grown at 37C until the OD_{600} reached
603 ~0.5. They were then switched to arabinose to induce RecV expression by pelleting the cells,
604 removing the supernatant, and resuspending in 10ml LB plus 0.2% arabinose, kanamycin (50
605 µg/ml) and chloramphenicol (30 µg/ml), and growing at 37C for 4 hours. When cultures were
606 plated immediately after the RecV induction period the colonies for experiments with pQD1 plus
607 test plasmids containing the 6A/6T IRs (pQD2 or pQD3) were mostly yellow due to mixed

608 populations of substrate and product plasmid within each founder cell (the test plasmid replicates
609 to high copy number within each host cell).

610 To be separated from one another the test plasmids needed to be recovered then
611 retransformed. After the 4 hours of RecV expression in arabinose, the cultures were switched back
612 to glucose for overnight growth: 1 ml was removed, pelleted, and resuspended in 1 ml LB
613 supplemented with 0.2% glucose and kanamycin (50 µg/ml), then 50ul of that was used to
614 inoculate 5 ml of LB supplemented with 0.2% glucose and kanamycin (50 µg/ml), which was
615 grown overnight. Plasmids were recovered by miniprep. 1ul of each plasmid was used to transform
616 competent DS941 E. coli, then two different volumes were plated on LB plus kanamycin. Colony
617 color was visualized using a ChemiDoc imager (BioRad). The red and green channel images are
618 overlaid in Figure 5.

619 Full-plasmid sequencing (Plasmidsaurus) was used to verify the recombination products.
620 Plasmids recovered from a green colony resulting from the pQD1 + pQD2 (6A/6T – OFF /red)
621 experiment were identical to pQD3 (6A/6T – ON /green), confirming the expected inversion. To
622 isolate products from the experiments using pQD4 (5A/6T – OFF /red) and pQD5 (5A/6T – ON
623 /green), the 1 or 2 product-color colonies seen in Figure 5 were picked, then restreaked to ensure
624 separation from their substrate-containing neighbors. Sequencing confirmed that product plasmids
625 from the pQD4 experiment matched the sequence of pQD5, and vice versa.

626 **RNA extraction, reverse transcription and RT-qPCR**

627 Cecal RNA was extracted using Rneasy PowerMicrobiome Kit (Qiagen) according to the
628 manufacturer's instructions. Complementary DNA was generated using the QuantiTect reverse
629 transcriptase kit (Qiagen) according to the manufacturer's instructions. Quantitative PCR was
630 performed on complementary DNA using primers with PowerTrack SYBR Green Master Mix

(Thermo Fisher). Reactions were run on a QuantStudio 6 pro (Thermo Fisher). Relative abundance was normalized by ΔΔCt. TcdA_qFor 5'-GTATGGATAGGTGGAGAAGTCA-3'; TcdA_qRev 5'-CTCTTCCTCTAGTAGCTGTAATGC-3'[67]; TcdB_qFor 5'-AGCAGTTGAATATAGTGGTTTAGTTAGAGTTG-3'; TcdB_qRev 5'-CATGCTTTTTTAGTTTCTGGATTGAA-3'[68]; FliS1_qFor 5'-TGCAGGACAATGGGCAAAGG-3'; FliS1_qRev 5'-CAGGCAACACATTATCTATTACCTGG-3'; FliE_qFor 5'-AGGCGAAGATGTTTCTATGCA-3'; FliE_qRev 5'-ACCTTATTCATTTCTTGATATGCATCA-3'

Statistical analysis

Kruskal-Wallis, T-tests and One-way ANOVA tests was performed to test the difference in maximum weight loss, swim plate, and RT-qPCR results. Statistical significance was determined by using a P value of <0.05. Linear regression analysis was used to estimate the correlation between IR type and weight loss.

648 **REFERENCES**

649 1. CDC, A. (2019). Antibiotic resistance threats in the United States. US Department of Health
650 and Human Services: Washington, DC, USA.

651 2. CDC (2020). 2020 Annual Report.

652 3. Abt, M.C., McKenney, P.T., and Pamer, E.G. (2016). *Clostridium difficile* colitis:
653 pathogenesis and host defence. Nat Rev Microbiol *14*, 609–620.
654 https://doi.org/10.1038/nrmicro.2016.108.

655 4. Lyerly, D.M., Krivan, H.C., and Wilkins, T.D. (1988). *Clostridium difficile*: its disease and
656 toxins. Clin Microbiol Rev *1*, 1–18.

657 5. Ciesielski-Treska, J., Ulrich, G., Rihn, B., and Aunis, D. (1989). Mechanism of action of
658 *Clostridium difficile* toxin B: role of external medium and cytoskeletal organization in
659 intoxicated cells. Eur J Cell Biol *48*, 191–202.

660 6. Florin, I., and Thelestam, M. (1983). Internalization of *Clostridium difficile* cytotoxin into
661 cultured human lung fibroblasts. Biochim Biophys Acta *763*, 383–392.
662 https://doi.org/10.1016/0167-4889(83)90100-3.

663 7. Henriques, B., Florin, I., and Thelestam, M. (1987). Cellular internalisation of *Clostridium*
664 *difficile* toxin A. Microb Pathog *2*, 455–463. https://doi.org/10.1016/0882-4010(87)90052-0.

665 8. Just, I., Selzer, J., Hofmann, F., Green, G.A., and Aktories, K. (1996). Inactivation of Ras by
666 *Clostridium sordellii* lethal toxin-catalyzed glucosylation. J Biol Chem *271*, 10149–10153.
667 https://doi.org/10.1074/jbc.271.17.10149.

668 9. Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M., and Aktories, K. (1995).
669 Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature *375*, 500–503.
670 https://doi.org/10.1038/375500a0.

671 10. Just, I., Wilm, M., Selzer, J., Rex, G., von Eichel-Streiber, C., Mann, M., and Aktories, K.
672 (1995). The enterotoxin from *Clostridium difficile* (ToxA) monoglucosylates the Rho
673 proteins. J Biol Chem *270*, 13932–13936. https://doi.org/10.1074/jbc.270.23.13932.

674 11. Just, I., Fritz, G., Aktories, K., Giry, M., Popoff, M.R., Boquet, P., Hegenbarth, S., and von
675 Eichel-Streiber, C. (1994). *Clostridium difficile* toxin B acts on the GTP-binding protein
676 Rho. J Biol Chem *269*, 10706–10712.

677 12. Schnizlein, M.K., and Young, V.B. (2022). Capturing the environment of the *Clostridioides*
678 *difficile* infection cycle. Nat Rev Gastroenterol Hepatol *19*, 508–520.
679 https://doi.org/10.1038/s41575-022-00610-0.

680 13. Vedantam, G., Clark, A., Chu, M., McQuade, R., Mallozzi, M., and Viswanathan, V.K.
681 (2012). *Clostridium difficile* infection: toxins and non-toxin virulence factors, and their

682 contributions to disease establishment and host response. Gut Microbes *3*, 121–134.
683 https://doi.org/10.4161/gmic.19399.

684 14. Burdon, D.W., George, R.H., Mogg, G.A., Arabi, Y., Thompson, H., Johnson, M.,
685 Alexander-Williams, J., and Keighley, M.R. (1981). Faecal toxin and severity of antibiotic-
686 associated pseudomembranous colitis. J Clin Pathol *34*, 548–551.
687 https://doi.org/10.1136/jcp.34.5.548.

688 15. La Ragione, R.M., Sayers, A.R., and Woodward, M.J. (2000). The role of fimbriae and
689 flagella in the colonization, invasion and persistence of *Escherichia coli* O78:K80 in the day-
690 old-chick model. Epidemiol Infect *124*, 351–363.
691 https://doi.org/10.1017/s0950268899004045.

692 16. Eaves-Pyles, T., Murthy, K., Liaudet, L., Virag, L., Ross, G., Soriano, F.G., Szabo, C., and
693 Salzman, A.L. (2001). Flagellin, a novel mediator of *Salmonella*-induced epithelial
694 activation and systemic inflammation: I kappa B alpha degradation, induction of nitric oxide
695 synthase, induction of proinflammatory mediators, and cardiovascular dysfunction. J
696 Immunol *166*, 1248–1260. https://doi.org/10.4049/jimmunol.166.2.1248.

697 17. Horstmann, J.A., Lunelli, M., Cazzola, H., Heidemann, J., Kuhne, C., Steffen, P., Szefs, S.,
698 Rossi, C., Lokareddy, R.K., Wang, C., et al. (2020). Methylation of *Salmonella*
699 Typhimurium flagella promotes bacterial adhesion and host cell invasion. Nat Commun *11*,
700 2013. https://doi.org/10.1038/s41467-020-15738-3.

701 18. Baban, S.T., Kuehne, S.A., Barketi-Klai, A., Cartman, S.T., Kelly, M.L., Hardie, K.R.,
702 Kansau, I., Collignon, A., and Minton, N.P. (2013). The Role of Flagella in *Clostridium*
703 *difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. PLoS
704 One *8*, e73026. https://doi.org/10.1371/journal.pone.0073026.

705 19. Batah, J., Kobeissy, H., Bui Pham, P.T., Denève-Larrazet, C., Kuehne, S., Collignon, A.,
706 Janoir-Jouveshomme, C., Marvaud, J.-C., and Kansau, I. (2017). *Clostridium difficile*
707 flagella induce a pro-inflammatory response in intestinal epithelium of mice in cooperation
708 with toxins. Sci Rep *7*, 3256. https://doi.org/10.1038/s41598-017-03621-z.

709 20. Dingle, T.C., Mulvey, G.L., and Armstrong, G.D. (2011). Mutagenic Analysis of the
710 *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence
711 in Hamsters ▿. Infect Immun *79*, 4061–4067. https://doi.org/10.1128/IAI.05305-11.

712 21. Anjuwon-Foster, B.R., and Tamayo, R. (2017). A genetic switch controls the production of
713 flagella and toxins in *Clostridium difficile*. PLOS Genetics *13*, e1006701.
714 https://doi.org/10.1371/journal.pgen.1006701.

715 22. El Meouche, I., Peltier, J., Monot, M., Soutourina, O., Pestel-Caron, M., Dupuy, B., and
716 Pons, J.-L. (2013). Characterization of the SigD Regulon of *C. difficile* and Its Positive
717 Control of Toxin Production through the Regulation of tcdR. PLoS One *8*, e83748.
718 https://doi.org/10.1371/journal.pone.0083748.

719 23. Trzilova, D., Warren, M.A.H., Gadda, N.C., Williams, C.L., and Tamayo, R. Flagellum and
720 toxin phase variation impacts intestinal colonization and disease development in a mouse
721 model of *Clostridioides difficile* infection. Gut Microbes *14*, 2038854.
722 https://doi.org/10.1080/19490976.2022.2038854.

723 24. Tenover, F.C., Tickler, I.A., and Persing, D.H. (2012). Antimicrobial-resistant strains of
724 *Clostridium difficile* from North America. Antimicrob Agents Chemother *56*, 2929–2932.
725 https://doi.org/10.1128/AAC.00220-12.

726 25. Wieczorkiewicz, J.T., Lopansri, B.K., Cheknis, A., Osmolski, J.R., Hecht, D.W., Gerding,
727 D.N., and Johnson, S. (2016). Fluoroquinolone and Macrolide Exposure Predict *Clostridium*
728 *difficile* Infection with the Highly Fluoroquinolone- and Macrolide-Resistant Epidemic *C.*
729 *difficile* Strain BI/NAP1/027. Antimicrob Agents Chemother *60*, 418–423.
730 https://doi.org/10.1128/AAC.01820-15.

731 26. Hubert, B., Loo, V.G., Bourgault, A.M., Poirier, L., Dascal, A., Fortin, E., Dionne, M., and
732 Lorange, M. (2007). A portrait of the geographic dissemination of the *Clostridium difficile*
733 North American pulsed-field type 1 strain and the epidemiology of *C. difficile*-associated
734 disease in Quebec. Clin Infect Dis *44*, 238–244. https://doi.org/10.1086/510391.

735 27. Labbe, A.C., Poirier, L., Maccannell, D., Louie, T., Savoie, M., Beliveau, C., Laverdiere, M.,
736 and Pepin, J. (2008). *Clostridium difficile* infections in a Canadian tertiary care hospital
737 before and during a regional epidemic associated with the BI/NAP1/027 strain. Antimicrob
738 Agents Chemother *52*, 3180–3187. https://doi.org/10.1128/AAC.00146-08.

739 28. Morgan, O.W., Rodrigues, B., Elston, T., Verlander, N.Q., Brown, D.F., Brazier, J., and
740 Reacher, M. (2008). Clinical severity of *Clostridium difficil*e PCR ribotype 027: a case-case
741 study. PLoS One *3*, e1812. https://doi.org/10.1371/journal.pone.0001812.

742 29. Sirard, S., Valiquette, L., and Fortier, L.C. (2011). Lack of association between clinical
743 outcome of *Clostridium difficile* infections, strain type, and virulence-associated phenotypes.
744 J Clin Microbiol *49*, 4040–4046. https://doi.org/10.1128/JCM.05053-11.

745 30. Carlson, P.E., Walk, S.T., Bourgis, A.E., Liu, M.W., Kopliku, F., Lo, E., Young, V.B.,
746 Aronoff, D.M., and Hanna, P.C. (2013). The relationship between phenotype, ribotype, and
747 clinical disease in human *Clostridium difficile* isolates. Anaerobe *24*, 109–116.
748 https://doi.org/10.1016/j.anaerobe.2013.04.003.

749 31. Dong, Q., Lin, H., Allen, M.-M., Garneau, J.R., Sia, J.K., Smith, R.C., Haro, F., McMillen,
750 T., Pope, R.L., Metcalfe, C., et al. (2023). Virulence and genomic diversity among clinical
751 isolates of ST1 (BI/NAP1/027) *Clostridioides difficile*. Cell Rep *42*, 112861.
752 https://doi.org/10.1016/j.celrep.2023.112861.

753 32. Paulick, A., Adamczyk, M., Anderson, K., Vlachos, N., Machado, M.J., McAllister, G.,
754 Korhonen, L., Guh, A.Y., Halpin, A.L., Rasheed, J.K., et al. (2021). Characterization of
755 *Clostridioides difficile* Isolates Available through the CDC & FDA Antibiotic Resistance
756 Isolate Bank. Microbiol Resour Announc *10*. https://doi.org/10.1128/MRA.01011-20.

757 33. Miles-Jay, A., Young, V.B., Pamer, E.G., Savidge, T.C., Kamboj, M., Garey, K.W., and
758 Snitkin, E.S. (2021). A multisite genomic epidemiology study of *Clostridioides difficile*
759 infections in the USA supports differential roles of healthcare versus community spread for
760 two common strains. Microb Genom *7*, 000590. https://doi.org/10.1099/mgen.0.000590.

761 34. Dingle, K.E., Didelot, X., Quan, T.P., Eyre, D.W., Stoesser, N., Golubchik, T., Harding,
762 R.M., Wilson, D.J., Griffiths, D., Vaughan, A., et al. (2017). Effects of control interventions
763 on *Clostridium difficile* infection in England: an observational study. Lancet Infect Dis *17*,
764 411–421. https://doi.org/10.1016/S1473-3099(16)30514-X.

765 35. Eyre, D.W., Davies, K.A., Davis, G., Fawley, W.N., Dingle, K.E., De Maio, N., Karas, A.,
766 Crook, D.W., Peto, T.E.A., Walker, A.S., et al. (2018). Two Distinct Patterns of *Clostridium*
767 *difficile* Diversity Across Europe Indicating Contrasting Routes of Spread. Clin Infect Dis
768 *67*, 1035–1044. https://doi.org/10.1093/cid/ciy252.

769 36. Eyre, D.W., Didelot, X., Buckley, A.M., Freeman, J., Moura, I.B., Crook, D.W., Peto,
770 T.E.A., Walker, A.S., Wilcox, M.H., and Dingle, K.E. (2019). *Clostridium difficile* trehalose
771 metabolism variants are common and not associated with adverse patient outcomes when
772 variably present in the same lineage. EBioMedicine *43*, 347–355.
773 https://doi.org/10.1016/j.ebiom.2019.04.038.

774 37. Stoesser, N., Crook, D.W., Fung, R., Griffiths, D., Harding, R.M., Kachrimanidou, M.,
775 Keshav, S., Peto, T.E., Vaughan, A., Walker, A.S., et al. (2011). Molecular epidemiology of
776 *Clostridium difficile* strains in children compared with that of strains circulating in adults
777 with *Clostridium difficile*-associated infection. J Clin Microbiol *49*, 3994–3996.
778 https://doi.org/10.1128/JCM.05349-11.

779 38. Olorunniji, F.J., Lawson-Williams, M., McPherson, A.L., Paget, J.E., Stark, W.M., and
780 Rosser, S.J. (2019). Control of ϕC31 integrase-mediated site-specific recombination by
781 protein trans-splicing. Nucleic Acids Res *47*, 11452–11460.
782 https://doi.org/10.1093/nar/gkz936.

783 39. Kaus, G.M., Snyder, L.F., Müh, U., Flores, M.J., Popham, D.L., and Ellermeier, C.D. (2020).
784 Lysozyme Resistance in *Clostridioides difficile* Is Dependent on Two Peptidoglycan
785 Deacetylases. J Bacteriol *202*, e00421-20. https://doi.org/10.1128/JB.00421-20.

786 40. Kamboj, M., McMillen, T., Syed, M., Chow, H.Y., Jani, K., Aslam, A., Brite, J., Fanelli, B.,
787 Hasan, N.A., Dadlani, M., et al. (2021). Evaluation of a Combined Multilocus Sequence
788 Typing and Whole-Genome Sequencing Two-Step Algorithm for Routine Typing of
789 *Clostridioides difficile*. J Clin Microbiol *59*. https://doi.org/10.1128/JCM.01955-20.

790 41. Tariq, R., Weatherly, R.M., Kammer, P.P., Pardi, D.S., and Khanna, S. (2017). Experience
791 and Outcomes at a Specialized *Clostridium difficile* Clinical Practice. Mayo Clin Proc Innov
792 Qual Outcomes *1*, 49–56. https://doi.org/10.1016/j.mayocpiqo.2017.05.002.

793 42. Schavemaker, P.E., and Lynch, M. (2022). Flagellar energy costs across the tree of life. Elife
794 *11*. https://doi.org/10.7554/eLife.77266.

43. Batah, J., Denève-Larrazet, C., Jolivot, P.-A., Kuehne, S., Collignon, A., Marvaud, J.-C., and Kansau, I. (2016). *Clostridium difficile* flagella predominantly activate TLR5-linked NF-κB pathway in epithelial cells. Anaerobe *38*, 116–124. https://doi.org/10.1016/j.anaerobe.2016.01.002.

44. Jiang, X., Hall, A.B., Arthur, T.D., Plichta, D.R., Covington, C.T., Poyet, M., Crothers, J., Moses, P.L., Tolonen, A.C., Vlamakis, H., et al. (2019). Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. Science *363*, 181–187. https://doi.org/10.1126/science.aau5238.

45. Krinos, C.M., Coyne, M.J., Weinacht, K.G., Tzianabos, A.O., Kasper, D.L., and Comstock, L.E. (2001). Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature *414*, 555–558. https://doi.org/10.1038/35107092.

46. Emerson, J.E., Reynolds, C.B., Fagan, R.P., Shaw, H.A., Goulding, D., and Fairweather, N.F. (2009). A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol *74*, 541–556. https://doi.org/10.1111/j.1365-2958.2009.06812.x.

47. Sekulovic, O., Ospina Bedoya, M., Fivian-Hughes, A.S., Fairweather, N.F., and Fortier, L.-C. (2015). The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. Mol Microbiol *98*, 329–342. https://doi.org/10.1111/mmi.13121.

48. Reynolds, C.B., Emerson, J.E., de la Riva, L., Fagan, R.P., and Fairweather, N.F. (2011). The *Clostridium difficil*e cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog *7*, e1002024. https://doi.org/10.1371/journal.ppat.1002024.

49. Garrett, E.M., Mehra, A., Sekulovic, O., and Tamayo, R. (2021). Multiple Regulatory Mechanisms Control the Production of CmrRST, an Atypical Signal Transduction System in *Clostridioides difficile*. mBio *13*, e0296921. https://doi.org/10.1128/mbio.02969-21.

50. Ribis, J.W., Nieto-Acuna, C.A., DiBenedetto, N., Mehra, A., Dong, Q., Nagawa, I.L., Meouche, I.E., Aldridge, B.B., Dunlop, M.J., Tamayo, R., et al. (2024). Unique growth and morphology properties of Clade 5 *Clostridioides difficile* strains revealed by single-cell time-lapse microscopy. Preprint at bioRxiv, https://doi.org/10.1101/2024.02.13.580212 https://doi.org/10.1101/2024.02.13.580212.

51. Sekulovic, O., Garrett, E.M., Bourgeois, J., Tamayo, R., Shen, A., and Camilli, A. (2018). Genome-wide detection of conservative site-specific recombination in bacteria. PLOS Genetics *14*, e1007332. https://doi.org/10.1371/journal.pgen.1007332.

52. Trzilova, D., and Tamayo, R. (2021). Site-specific recombination – how simple DNA inversions produce complex phenotypic heterogeneity in bacterial populations. Trends Genet *37*, 59–72. https://doi.org/10.1016/j.tig.2020.09.004.

831 53. Weldy, M., Evert, C., Dosa, P.I., Khoruts, A., and Sadowsky, M.J. (2020). Convenient
832 Protocol for Production and Purification of *Clostridioides difficile* Spores for Germination
833 Studies. STAR Protoc *1*, 100071. https://doi.org/10.1016/j.xpro.2020.100071.

834 54. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for
835 Illumina sequence data. Bioinformatics *30*, 2114–2120.
836 https://doi.org/10.1093/bioinformatics/btu170.

837 55. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and
838 Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data
839 processing and web-based tools. Nucleic Acids Research *41*, D590–D596.
840 https://doi.org/10.1093/nar/gks1219.

841 56. Glöckner, F.O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G.,
842 Yarza, P., Peplies, J., Westram, R., et al. (2017). 25 years of serving the community with
843 ribosomal RNA gene reference databases and tools. Journal of Biotechnology *261*, 169–176.
844 https://doi.org/10.1016/j.jbiotec.2017.06.1198.

845 57. Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies,
846 J., Ludwig, W., and Glöckner, F.O. (2014). The SILVA and "All-species Living Tree Project
847 (LTP)" taxonomic frameworks. Nucleic Acids Research *42*, D643–D648.
848 https://doi.org/10.1093/nar/gkt1209.

849 58. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat
850 Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

851 59. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose
852 program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.
853 https://doi.org/10.1093/bioinformatics/btt656.

854 60. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
855 dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
856 https://doi.org/10.1186/s13059-014-0550-8.

857 61. McAllister, K.N., Bouillaut, L., Kahn, J.N., Self, W.T., and Sorg, J.A. (2017). Using
858 CRISPR-Cas9-mediated genome editing to generate *C. difficile* mutants defective in
859 selenoproteins synthesis. Sci Rep *7*, 14672. https://doi.org/10.1038/s41598-017-15236-5.

860 62. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using
861 SPAdes De Novo Assembler. Curr Protoc Bioinformatics *70*, e102.
862 https://doi.org/10.1002/cpbi.102.

863 63. Griffiths, D., Fawley, W., Kachrimanidou, M., Bowden, R., Crook, D.W., Fung, R.,
864 Golubchik, T., Harding, R.M., Jeffery, K.J.M., Jolley, K.A., et al. (2010). Multilocus
865 sequence typing of *Clostridium difficile*. J Clin Microbiol *48*, 770–778.
866 https://doi.org/10.1128/JCM.01796-09.

867 64. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and
868 Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.
869 https://doi.org/10.1186/1471-2105-10-421.

870 65. Shin, H., Holland, A., Alsaleh, A., Retiz, A.D., Pigli, Y.Z., Taiwo-Aiyerin, O.T., Reyes,
871 T.P., Bello, A.J., Olorunniji, F.J., and Rice, P.A. (2024). Identification of cognate
872 recombination directionality factors for large serine recombinases by virtual pulldown.
873 Preprint at bioRxiv, https://doi.org/10.1101/2024.06.11.598349
874 https://doi.org/10.1101/2024.06.11.598349.

875 66. Summers, D.K., and Sherratt, D.J. (1988). Resolution of ColE1 dimers requires a DNA
876 sequence implicated in the three-dimensional organization of the cer site. EMBO J *7*, 851–
877 858. https://doi.org/10.1002/j.1460-2075.1988.tb02884.x.

878 67. Babakhani, F., Bouillaut, L., Sears, P., Sims, C., Gomez, A., and Sonenshein, A.L. (2013).
879 Fidaxomicin inhibits toxin production in *Clostridium difficile*. Journal of Antimicrobial
880 Chemotherapy *68*, 515–522. https://doi.org/10.1093/jac/dks450.

881 68. Wroblewski, D., Hannett, G.E., Bopp, D.J., Dumyati, G.K., Halse, T.A., Dumas, N.B., and
882 Musser, K.A. (2009). Rapid molecular characterization of *Clostridium difficile* and
883 assessment of populations of *C. difficile* in stool specimens. J Clin Microbiol *47*, 2142–2148.
884 https://doi.org/10.1128/JCM.02498-08.

885

Fig 1

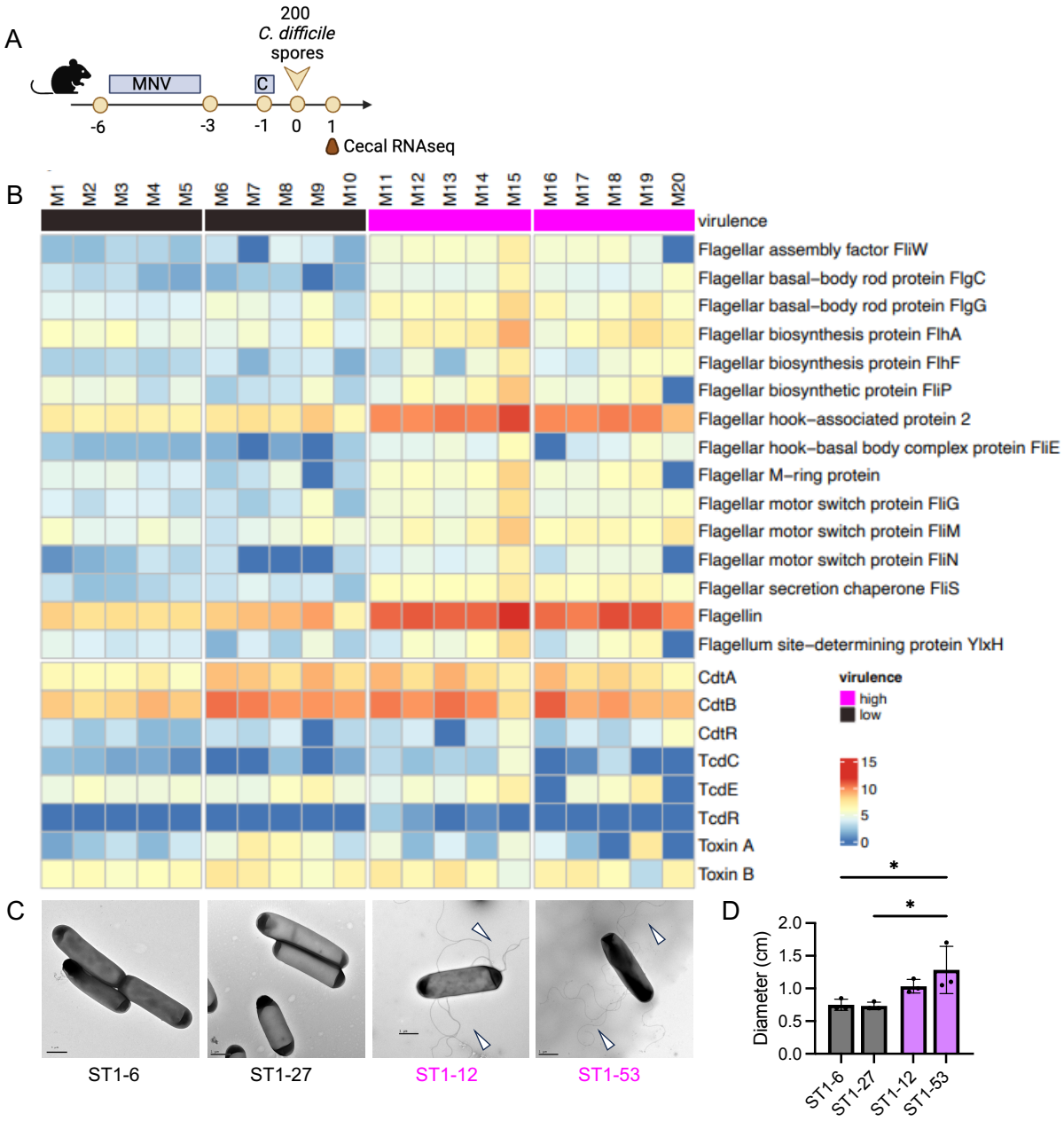
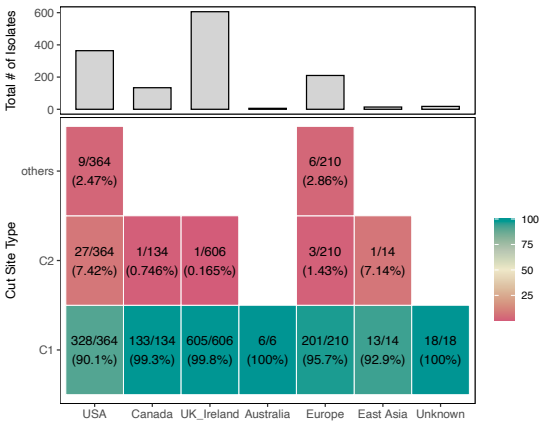

A
200
C. difficile
spores
MNV
C
-6
-3
-1
0
1
Cecal RNAseq
B
M1
M2
M3
M4
M5
M6
M7
M8
M9
M10
M11
M12
M13
M14
M15
M16
M17
M18
M19
M20
virulence
Flagellar assembly factor FliW
Flagellar basal–body rod protein FlgC
Flagellar basal–body rod protein FlgG
Flagellar biosynthesis protein FlhA
Flagellar biosynthesis protein FlhF
Flagellar biosynthetic protein FliP
Flagellar hook–associated protein 2
Flagellar hook–basal body complex protein FliE
Flagellar M–ring protein
Flagellar motor switch protein FliG
Flagellar motor switch protein FliM
Flagellar motor switch protein FliN
Flagellar secretion chaperone FliS
Flagellin
Flagellum site–determining protein YlxH
CdtA
CdtB
CdtR
TcdC
TcdE
TcdR
Toxin A
Toxin B
virulence
high
low
15
10
5
0
*
C
ST1-6
ST1-27
ST1-12
ST1-53
D
*
Diameter (cm)
2.0
1.5
1.0
0.5
0.0
ST1-6
ST1-27
ST1-12
ST1-53

Fig 2

A

Isolate	Type	Left inverted repeat	Central region	Right inverted repeat	Virulence
R20291	C1	AAGTTACTATTTTACAAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	High
ST1-12	C1	AAGTTACTATTTTACAAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	High
ST1-53	C1	AAGTTACTATTTTACAAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	High
ST1-19	C2-ON	AAGTTACTATTTTACAAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	Intermediate
ST1_65	C2-ON	AAGTTACTATTTTACAAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	Low
ST1-6	C2-OFF	AAGTTACTATTTTACAAAAA	TGACTT..... TTGCCT	TTTTTTGTAATATAGCAACTT	Low
ST1-27	C2-OFF	AAGTTACTATTTTACAAAAA	TGACTT..... TTGCCT	TTTTTTGTAATATAGCAACTT	Low
ST1-51	R1-ON	AAGTTACTATTTTACAAAAA	AGGCAA.....AAGTCA	TTTTTTGTAATATAGCAACTT	Low
ST1-68	R1-OFF	AAGTTACTATTTTACAAAAAA	TGACTT..... TTGCCT	TTTTTTGTAATATAGCAACTT	Low
ST1-67	R2-OFF	AAGTTACTATTTTACAAAAA	TGACTT..... TTGCCT	TTTTTTGTAATATAGCAACTT	Low

B



C

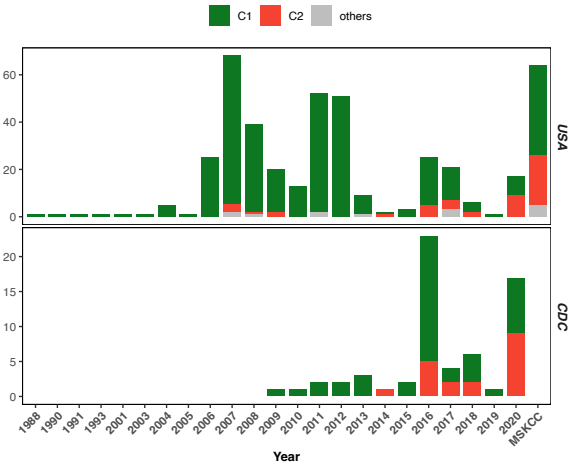


Fig 3

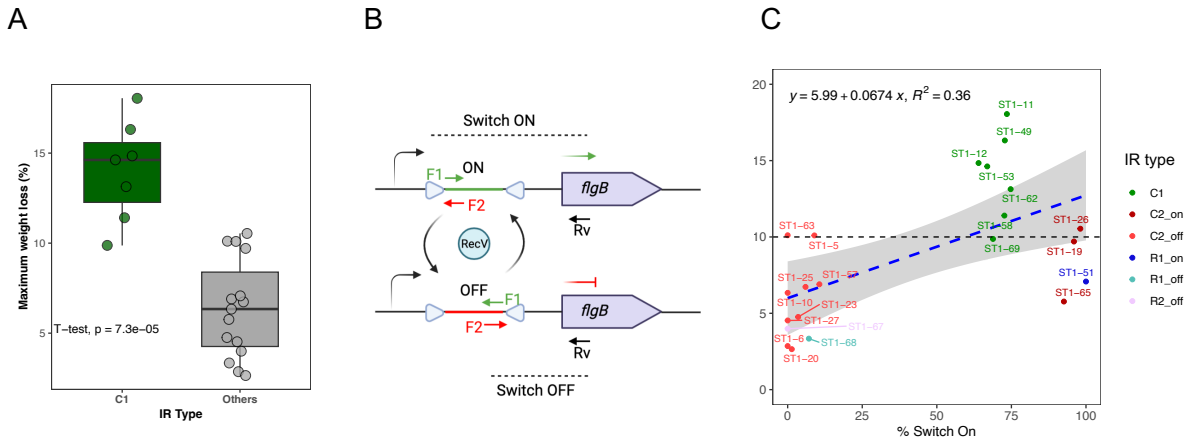

A
B
C
Maximum weight loss (%)
T−test, p = 7.3e−05
C1
Others
IR Type
Switch ON
F1
ON
F2
flgB
RecV
Rv
OFF
F1
F2
flgB
Rv
Switch OFF
y = 5.99 + 0.0674 x, R2 = 0.36
% Switch On
IR type
C1
C2_on
C2_off
R1_on
R1_off
R2_off

Fig 4

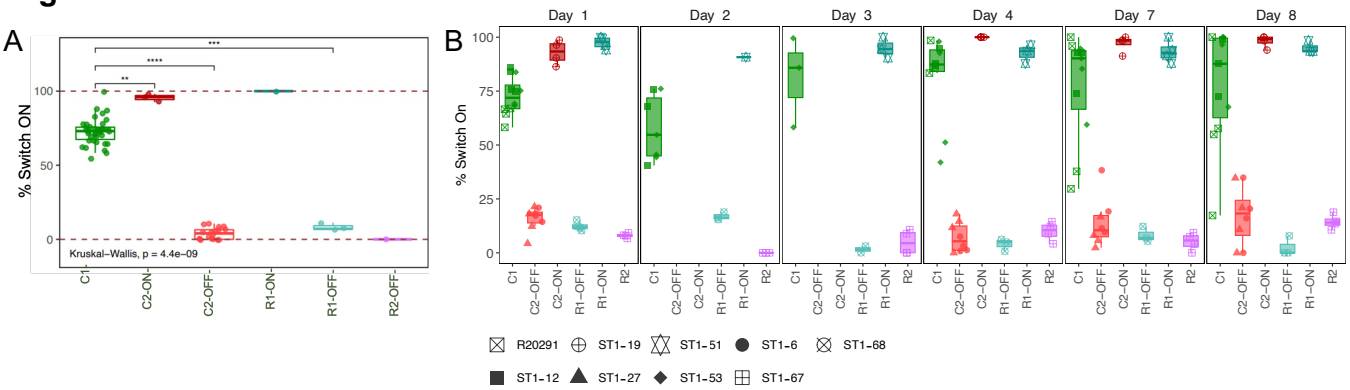

A
% Switch ON
Kruskal–Wallis, p = 4.4e−09
C1
C2-ON-
C2-OFF-
R1-ON-
R1-OFF-
R2-OFF-
B
% Switch On
Day 1
Day 2
Day 3
Day 4
Day 7
Day 8
C1
C2−OFF
C2−ON
R1−OFF
R1−ON
R2
R20291
ST1−19
ST1−51
ST1−6
ST1−68
ST1−12
ST1−27
ST1−53
ST1−67

Fig 5

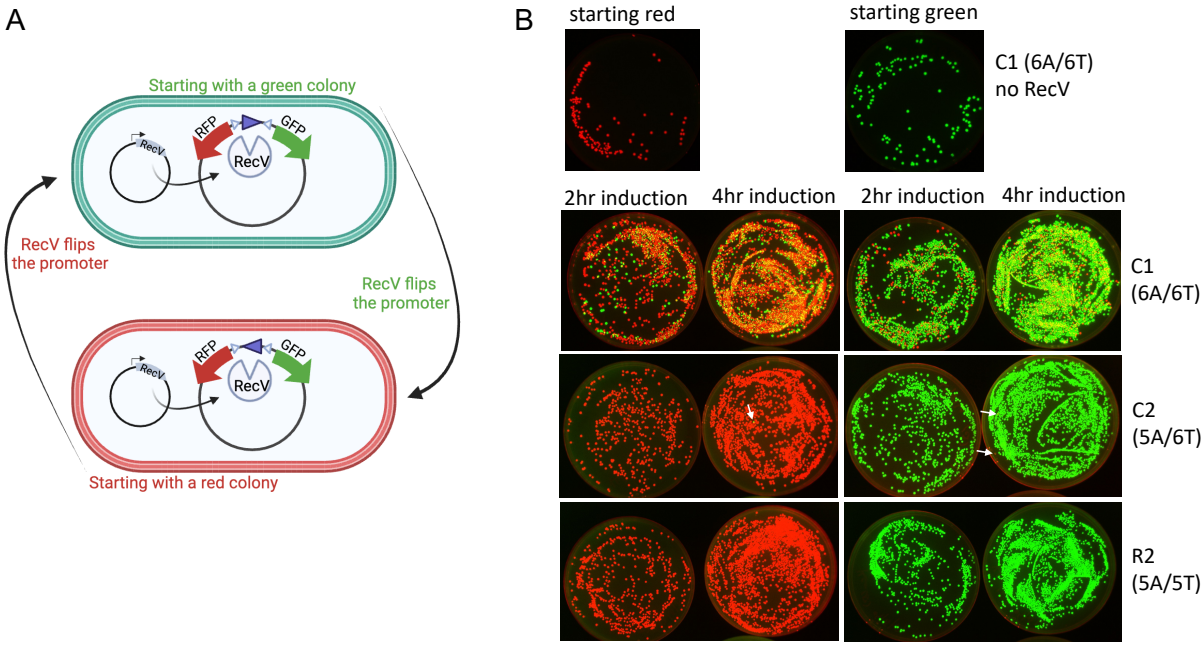
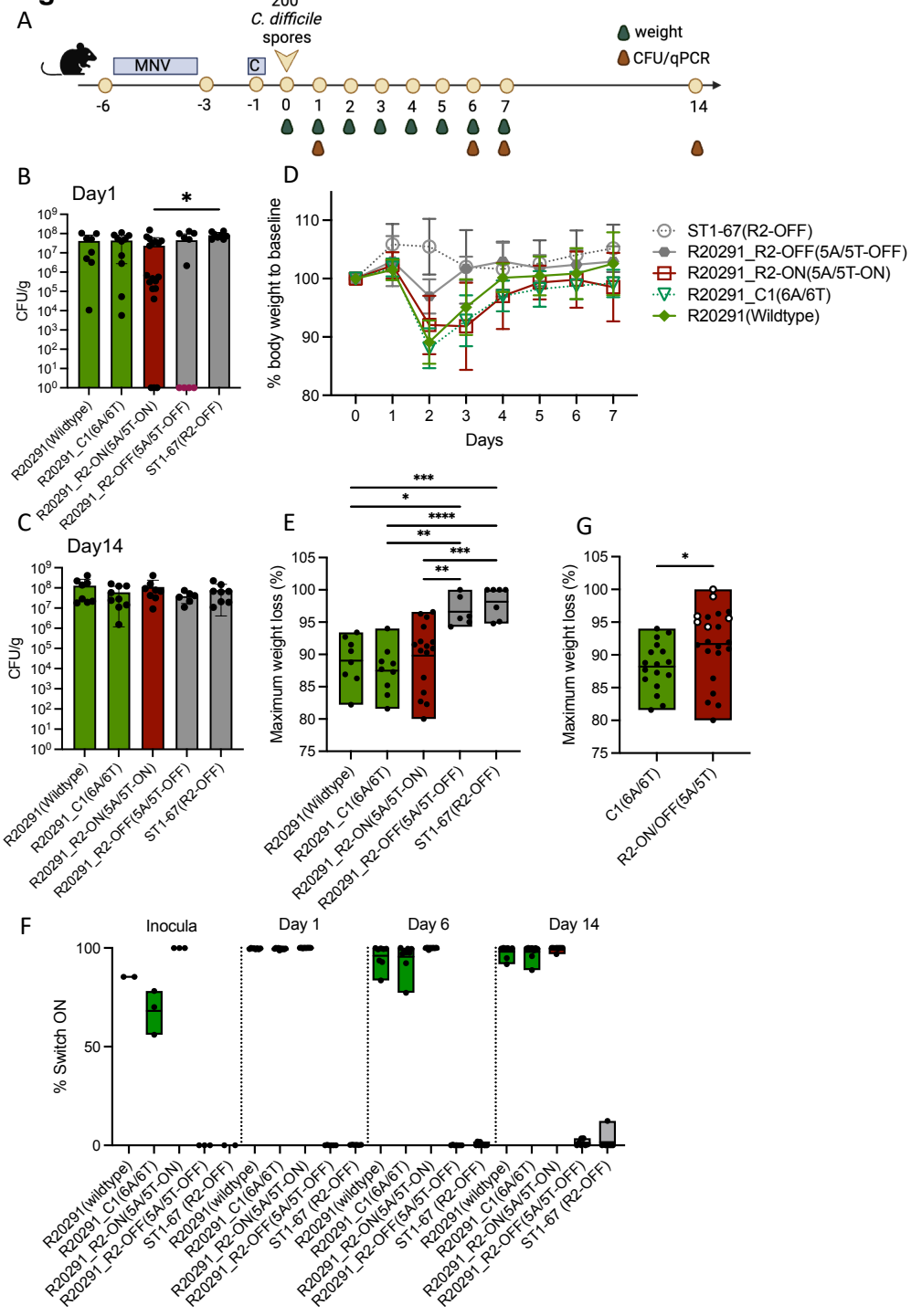

A
Starting with a green colony
RFP
GFP
RecV
RecV
RecV flips
the promoter
RecV flips
the promoter
RFP
GFP
RecV
RecV
Starting with a red colony
B
starting red
starting green
C1 (6A/6T)
no RecV
2hr induction
4hr induction
2hr induction
4hr induction
C1
(6A/6T)
C2
(5A/6T)
R2
(5A/5T)

Fig 6


A
200
C. difficile
spores
MNV
C
weight
CFU/qPCR
-6
-3
-1
0
1
2
3
4
5
6
7
14
B
Day1
CFU/g
R20291(Wildtype)
R20291_C1(6A/6T)
R20291_R2-ON(5A/5T-ON)
R20291_R2-OFF(5A/5T-OFF)
ST1-67(R2-OFF)
D
% body weight to baseline
Days
ST1-67(R2-OFF)
R20291_R2-OFF(5A/5T-OFF)
R20291_R2-ON(5A/5T-ON)
R20291_C1(6A/6T)
R20291(Wildtype)
C
Day14
CFU/g
E
Maximum weight loss (%)
G
Maximum weight loss (%)
C1(6A/6T)
R2-ON/OFF(5A/5T)
F
Inocula
Day 1
Day 6
Day 14
% Switch ON