View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository@Nottingham

1 **What's a SNP between friends: the lineage of *Clostridioides difficile* R20291 can**
2 **effect research outcomes.**

3 Jorge Monteford[1]*, Terry W. Bilverstone[1,2]*, Patrick Ingle[1,2], Sheryl Philip[1,3], Sarah A.
4 Kuehne[1,4], Nigel P. Minton[1,2].

5 * These authors contributed equally.

6

7 [1] Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre (SBRC), School
8 of Life Sciences, Centre for Biomolecular Sciences, The University of Nottingham, Nottingham,
9 NG7 2RD, UK.

10 [2] NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust
11 and the University of Nottingham, Nottingham, NG7 2RD, UK.

12 *Current Address:*

13 [3]Biocatalysts Limited, Unit 1, Cefn Coed, Parc Nantgarw, Cardiff, CF15 7QQ, UK,

14 [4]Oral Microbiology Group, School of Dentistry and Institute of Microbiology and Infection,
15 College of Medical and Dental Sciences, The University of Birmingham, Birmingham, B5 7EG,
16 UK.
17
18 Corresponding author: nigel.minton@nottingham.ac.uk

19
20 **Key Words**

21 *Clostridioides difficile* R20291; motility; biofilm; toxin production; conjugation; genomic
22 variation

23

24 **Abstract:**

25 *Clostridioides difficile* R20291 is the most studied PCR-Ribotype 027 isolate. The two
26 predominant lineages of this hypervirulent strain, however, exhibit substantive phenotypic
27 differences and possess genomes that differ by a small number of nucleotide changes. It
28 is important that the source of R20291 is taken into account in research outcomes.

29

30 *Clostridioides difficile* (formerly *Clostridium difficile* [1]) is the leading cause of hospital-
31 associated diarrhoea in the developed world. Its prevalence in recent years has been
32 attributed to the emergence of hypervirulent strains, and in particular those belonging to
33 BI/NAP1/PCR ribotype 027 (RT 027) which elaborate high titres of Toxin A/B, produce
34 binary toxin and exhibit an increased propensity to form spores [2]. The first RT 027 strain
35 to have its genome sequenced was strain R20291 [3] responsible for a major outbreak in
36 2006 at Stoke Mandeville Hospital, UK. Consequently, R20291 has become one of the most
37 studied laboratory strains of *C. difficile.*

38 Full exploitation of clostridial genome sequence data has relied on the application of
39 forward and reverse genetics tools [4], most notably ClosTron technology based on intron
40 re-targeting [5]. Initial attempts to generate mutants in R20291, however, found that the
41 effective transfer of the ClosTron plasmid was dependent on the R20291 stock used.
42 Transfer was reproducibly possible using CRG0825, a stock of R20291 obtained by
43 Nottingham's Clostridia Research Group (CRG) from the UK Anaerobe reference unit
44 (ARU), Cardiff, UK. In comparison, transfer to a stock of R20291 (CRG2021) originating
45 from the Brendon Wren laboratory at The London School of Hygiene and Tropical Medicine
46 (LSHTM), was extremely ineffective. Consequently, the CRG0825 was taken forward in
47 reverse genetic studies using the ClosTron and as the basis for the development of allelic-
48 exchange (AE) technology based on *pyrE* alleles [6]. As a result, CRG0825 and its Δ*pyrE*
49 derivative have been widely distributed to research laboratories wishing to study R20291.

50 The inefficient nature of CRG2021 as a conjugative recipient is not confined to the ClosTron
51 plasmid but affects a range of different vectors which are transferred to CRG0825 at rates
52 that are an order of magnitude higher (Fig.S1). To shed light on this phenomenon the
53 genome sequences and the phenotypes of the two strains were compared. A third R20291
54 strain used by Novartis (CRG3661) was included for comparative purposes.

55 Genomic DNA from all three strains was subjected to Illumina paired-end sequencing and
56 the reads assembled and aligned with the reference genome sequence (Accession number:
57 FN545816). This analysis identified six single nucleotide polymorphisms (SNPs) across all
58 three strains that deviate from the reference sequence, alongside thirteen insertions and
59 eleven deletions (Table 1). In addition to the mutations that were conserved across all
60 three strains, CRG0825 possessed three deletions and two SNPs that were not present in
61 the reference or CRG2021 sequence, whilst the CRG3661 possessed three unique SNPs
62 (Table 1). CRG2091 did not possess any unique mutations compared with the reference
63 genome sequence.

64 Flagella likely play an important role in the conjugation process. We had previously noted
65 that CRG0825 carried a single, polar flagella [7]. A separate study suggested that
66 CRG2091 was peritrichously flagellated [8]. These differences were confirmed here using
67 Transmission Electron Microscopy (TEM) and extended to establish that CRG3661 was also
68 peritrichously flagellated (Fig. 1c-e). Further analysis demonstrated that CRG0825
69 exhibited an approximate 50% reduction in swimming motility relative to the other two
70 strains (Fig. 1a). Moreover, consistent with its reduced motility, strain CRG0825 was also
71 found to show a greater propensity to form biofilm, as measured by a biomass formation
72 using crystal violet [9], than strains CRG2021 and CRG3661 (Fig. 1b).

73 Other studies have linked flagella-mediated motility with toxigenesis in *C. difficile* [10].
74 Therein, inactivation of early-stage flagella genes led to increased toxin production
75 corresponding with enhanced *in vivo* virulence, whilst inactivation of late-stage flagella
76 genes had the opposite effect [11, 12]. Accordingly, we assessed the levels of toxin
77 production in the three strains using a commercial ELISA kit on 72h filter-sterilised

78 supernatants as described previously [13]. An approximately 3.5-fold increase in toxin
79 production was observed for the CRG0825 compared with the CRG2021 strain which
80 produced around 22% less combined Toxin A/B than CRG3661 (Fig. 2a).

81 Having established that genetic differences between the strains had affected the important
82 virulence determinants of motility and toxin production, we tested to see whether the
83 capacity to form endospores had been altered as spores represent a critical attribute of
84 disease transmission. Under the conditions tested it was established that the final titre of
85 spores obtained from 96h onwards was the same regardless of the strain (Fig. 2b). The
86 first appearance of spores in cultures of CRG0825, however, was significantly delayed by
87 some 24h compared to the other two strains (Fig. 2b).

88 Finally, the growth performance of each strain was compared. On complex medium,
89 CRG0825 grew to a lower optical density (OD) during the exponential and stationary
90 growth phases than the CRG3661 or CRG2021 strains, where the greatest disparity was
91 observed between the CRG0825 and CRG3661 (Fig S2a). Intriguingly, the observed
92 difference was reversed when the strains were cultured on minimal medium containing
93 either glucose, fructose or mannitol (1% w/v) as the primary carbohydrate source (Fig.
94 S2b-d).

95 The net result of our analysis was that the two R2091 strain CRG0825 and CRG2021 exhibit
96 significant phenotypic differences. Aside from its greater efficiency as a recipient in
97 conjugations with *E. coli* donor strains, CRG0825 was less motile and exhibited a greater
98 propensity to form biofilm, as measured by a standard crystal violet assay. These
99 differences may represent a consequence of its apparent possession of a single, polar
100 flagella as opposed to the peritrichous flagella of CRG2021, as visualised under TEM.
101 CRG0825 was also shown to produce higher levels of toxins, delayed sporulation and
102 different growth characteristics to CRG2021 on rich and minimal media. The cause of
103 these phenotypic differences are undoubtedly the SNPs and Indels present in CRG0825.
104 A number of pivotal questions emerge from these observations.

105 **What are the specific causes of the observed changes in phenotype?** Many of the
106 changes appear linked to flagella and motility, yet none of the five mutations in CRG0825
107 reside directly within, or flank any known flagella genes, and are most likely impinging on
108 the regulation of these processes. Moreover, regulation of flagella, toxin production and
109 virulence are known to be linked in *C. difficile* [10-12]. Three of the four non-synonymous
110 SNPs present in CRG0825 do indeed affect apparent regulatory genes, namely *vncR,* TCS
111 HK and the anti-sigma factor *rsbW.* Two of the three non-synonymous changes in
112 CRG3661 are also in regulatory genes, *codY* and a *gntR* family regulator. However, to pin
113 down exactly which SNP(s) or Indels, are responsible for the observed phenotypic
114 differences between CRG0825 and CRG2021, for instance, would require a considerable
115 effort in which all combinations of mutation would need to be generated in allele
116 replacement experiments during which the generation of additional changes would need
117 to be excluded.

118 **How did these changes arise?** Following their discovery, correspondence with Val Hall
119 at the ARU revealed that at the time R20291 was sent to Nottingham, it was routine ARU
120 practice to "*keep a small number of isolates that are used as internal lab controls on agar*
121 *plates, sub-culturing weekly plate-to-plate and retrieving fresh cultures from the original*
122 *vial periodically".* The sequence presented here is from Nottingham's -80° C, Master seed
123 bank (red tube) prepared immediately on receipt of the strain in 2006. We can conclude,
124 that during the repeated subculture of the R20291 stock at the ARU in 2006, the 5
125 described mutations arose. This practice of maintaining a stock plate no longer takes place
126 at the ARU. The consequences of subculturing have previously been noted in the case of
127 the *C. difficile* strain 630, where deliberate, repeated subculture led to the emergence of
128 two very different cell lines (630Δ*erm* and 630E) carrying distinct SNPs, inversions and
129 deletions and which exhibited differences in motility, spore formation and toxin production,
130 as well as overall virulence in the hamster model of infection [14].

131 **What are the lessons to be learnt?** The take home message of this investigation is
132 that stock cultures need to be appropriately maintained. At Nottingham, a traffic light

133 system is used to store bacterial cultures. Upon receipt of cultures, aliquots of cells (never
134 single colonies) are used to inoculate an overnight which following addition of 10% glycerol
135 is allocated to three 2 ml screwed capped tubes with a red, amber and green coloured cap
136 insert and stored at -80° C. Red tubes remain untouched and are stored in a separate
137 freezer. Green tubes represent the working stock which may be restocked from the amber
138 tube where necessary.

139 The R20291 strain maintained at LSHTM has a genome sequence consistent with the
140 sequence held at GenBank (Accession number: FN545816). The differences listed in Table
141 1 are common to all strains, and therefore likely represent errors in the original sequence.
142 The strain CRG3661 began its journey at LSHTM and found its way to Novartis via the
143 Trevor Lawley laboratory at the Sanger Institute, and thence to Nottingham. It is not clear
144 when its three mutations arose. The Nottingham CRG0825 apparently arose as a result of
145 repeated subculture at the ARU.

146 **What is the way forward?** It is clearly advisable that the genome sequences of any
147 stock culture of any bacterial strain stored in a laboratory should be confirmed, regardless
148 of source, prior to use. This principal should equally apply to any mutant derivative of a
149 strain made by whatever means, to ensure that additional SNPs/ Indels have not arisen.

150 On the specific subject of studies dealing with R20291, it is important that experimentalists
151 are aware of the differences between the strain lineages discussed here, and that the
152 strain used is made clear in any meeting presentation or published article. CRG0825 is a
153 widely distributed strain, owing to its superior conjugative efficiencies and its usage in the
154 development of AE mutagenesis technologies [6]. The absence of polymorphisms specific
155 to CRG2021, however, suggests that this strain is the closest ancestor of the original
156 R20291 clinical isolate. Although the lack of a characterised uracil auxotroph, in addition
157 to difficulties concerning conjugal transfer, formally reduced its attractiveness compared
158 to CRG0825, recent advances that improve gene transfer frequencies [15, 16] and the
159 advent of multiple CRISPR-Cas methodologies for use in *C. difficile* research [17-20], have
160 improved the tractability of CRG2021 to genetic studies. For those researchers who wish

161 to use AE technologies based on *pyrE* [6], the requisite auxotrophic uracil mutant of
162 CRG2021 is available from www.plasmidvectors.com.

163

164 **References**

165 1. Lawson P.A., Citron D.M., Tyrrell K.L., Finegold S.M. Reclassification of *Clostridium*
166 *difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prevot 1938. *Anaerobe.*
167 2016;**40:** 95-99.

168 2. Cartman S.T., Heap J.T., Kuehne S.A., Cockayne A., Minton N.P. The emergence of
169 'hypervirulence' in *Clostridium difficile*. *Int. J. Med. Microbiol*. 2010;**300(6):**387-95.

170 3. Stabler R.A., He M., Dawson L. et al. Comparative genome and phenotypic analysis
171 of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent
172 bacterium. *Genome Biol*. 2009;**10(9):**R102

173 4. Minton N.P., Ehsaan M., Humphreys C.M., Little G.T., Baker J., Henstra A.M., Liew F.,
174 Kelly M.L., Sheng L., Schwarz K., Zhang Y. A roadmap for gene system development
175 in *Clostridium. Anaerobe*. 2016;**41:**104-112.

176 5. Heap J.T., Kuehne S.A., Ehsaan M., Cartman S.T., Cooksley C.M., Scott J.C., Minton
177 N.P. *J. Microbiol. Methods*. The ClosTron: Mutagenesis in *Clostridium* refined and
178 streamlined. 2010;**80(1):**49-55.

179 6. Ng Y.K., Ehsaan M., Philip S., Collery M.M., Janoir C., Collignon A., Cartman S.T.,
180 Minton N.P. Expanding the repertoire of gene tools for precise manipulation of the
181 *Clostridium difficile* genome: Allelic Exchange using *pyrE* alleles. *PLoS ONE.*
182 2013;**8(2):**e56051.

183 7. Baban S.T., Kuehne S.A., Barketi-Klai A., Cartman S.T., Kelly M.L., Hardie K.R.,
184 Kansau I., Collignon A., Minton N.P. The role of flagella in *Clostridium difficile*
185 pathogenesis: comparison between a non-epidemic and an epidemic strain. *PLoS One*.
186 2013;**8(9):**e73026.

187 8. Martin M.J., Clare S., Goulding D., Faulds-Pain A., Barquist L., Browne H.P., Pettit L.,
188 Dougan G., Lawley T.D., Wren B.W. The *agr* locus regulates virulence and colonization
189 genes in *Clostridium difficile* 027. *J. Bacteriol*. 2013;**195(16):**3672-81.

190 9. Đapa T., Leuzzi R., Ng Y.K., Baban S.T., Adamo R., Kuehne S.A., Scarselli M., Minton
191 N.P., Serruto D., Unnikrishnan M. Multiple factors modulate biofilm formation by the
192 anaerobic pathogen *Clostridium difficile*. *J Bacteriol*. 2013;**195(3):**545-55.

193 10. Aubry A., Hussack G., Chen W., KuoLee R., Twine S.M., Fulton K.M., Foote S., Carrillo
194 C.D., Tanha J., Logan S.M. Modulation of toxin production by the flagellar regulon in
195 *Clostridium difficile*. *Infect. Immun.* 2012;**80(10):**3521-3532.

196 11. Dingle T.C., Mulvey G.L., Armstrong G.D. Mutagenic analysis of the *Clostridium difficile*
197 flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. *Infect*
198 *Immun*. 2011;**79(10):**4061-4067.

199 12. Anjuwon-Foster B.R., Tamayo R. A genetic switch controls the production of flagella
200 and toxins in *Clostridium difficile*. *PLoS Genet.* 2017;**13(3):**e1006701-e1006701.

201 13. Bilverstone T.W., Kinsmore N.L., Minton N.P., Kuehne S.A. Development of
202 *Clostridium difficile* R20291ΔPaLoc model strains and in vitro methodologies reveals
203 CdtR is required for the production of CDT to cytotoxic levels. *Anaerobe.* 2017;**44:**51-
204 54.

205 14. Collery M.M., Kuehne S.A., McBride S.M., Kelly M.L., Monot M., Cockayne A., Dupuy
206 B., Minton, N.P. What's a SNP between friends: The influence of single nucleotide
207 polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and
208 derivatives. *Virulence* 2017;**8(6):**767-781.

209 15. Kirk J.A., Fagan R.P. Heat shock increases conjugation efficiency in Clostridium
210 difficile. *Anaerobe*, 2016;**42:**1-5.

211 16. Woods C., Humphreys C.M., Rodrigues R.M., Ingle P., Rowe P., Henstra A.M., Köpke
212 M., Simpson S.D., Winzer K., Minton N.P. A novel conjugal donor strain for improved
213 DNA transfer into *Clostridium* spp. *Anaerobe*. 2019;**59:**184-191.

17. McAllister K.N., Bouillaut L., Kahn J.N., Self W.T., Sorg J.A. Using CRISPR-Cas9-mediated genome editing to generate *C. difficile* mutants defective in selenoproteins synthesis. *Sci. Rep.* 2017;**7(1)**:14672.

18. Hong W., Zhang J., Cui G., Wang L., Wang Y. Multiplexed CRISPR-Cpf1-Mediated Genome Editing in *Clostridium difficile* toward the Understanding of Pathogenesis of *C. difficile* Infection. *ACS Synth. Biol.* 2018;**7(6):**1588-1600.

19. Maikova A., Kreis V., Boutserin A., Severinov K., Soutourina O. Using an Endogenous CRISPR-Cas System for Genome Editing in the Human Pathogen *Clostridium difficile. Appl. Environ. Microbiol*. 2019;**85(20):**e01416-19.

20. Cañadas IC, Groothuis D, Zygouropoulou M, Rodrigues R, Minton NP. RiboCas: A Universal CRISPR-Based Editing Tool for *Clostridium*. *ACS Synth. Biol.* 2019;**8(6):**1379-1390.

ACKNOWLEDGEMENT

This work was supported by the award of a PhD scholarship to JM by CONACYT (Consejo nacional de ciencia y tecnologia, Mexico), through the UK Biotechnology and Biological Sciences Research Council [grant number BB/J014508/1] and Medical Research Council [grant number G090085] and by the Swiss National Science Foundation [Sinergia grant CRSII5_180317]. The views expressed are those of the authors and not necessarily those of the funders.

Author contributions

Conceived the experiments: NPM. Performed the experiments: JM, TWB and PI. Undertook genome sequence determination and analysis: SP, TWB and NPM. Analysed the data: JM, TWB, PI, SAK and NPM. Wrote the paper: TWB and NPM. All authors read and commented on the final manuscript.

240 **Conflicts of interest**

241 The authors declare no conflicts of interest.

242

243 **Table 1: Genomic mutations compared with the reference genome sequence for**
244 **R02921.**

245 The region encompassing the mutation was aligned with multiple *C. difficile* genome
246 sequences using NCBI Blastn. *Insertion here results in a frameshift mutation to a full-
247 length pseudogene encoding an 87 AA protein. This gene without mutation encodes only
248 8 AAs. TCS-HC: Two-component system histidine kinase.

249

250 **Figure Legends**

251 **Fig. 1: Motility, biofilm formation and Transmission electron microscopy of**
252 **R20291 derivatives. a)** The three derivatives of *C. difficile* R20291 were assessed for
253 their motility characteristics using a swarming motility assay. Motility is represented by
254 the distance travelled from the initial inoculum to the outermost edge of the ensuing halo
255 following 48h incubation. **b**) The three R20291 derivatives were assessed for their
256 propensity to form biofilms by means of a biofilm assay. Biofilm production is represented
257 by the enumeration of crystal violet dye extracted from 120h broth cultures. Data
258 represent the mean ±SD of three independent experiments. Statistical significance
259 according to One-way ANOVA followed by the Dunnet's multiple comparison test.
260 P=**<0.01; P-***<0.001. Transmission electron microscopy analysis of **c)** CRG0825; **d)**
261 CRG2091; **e)** CRG3661.

262

263 **Fig. 2: Toxin and sporulation profiles of R20291 derivatives.** The three derivatives
264 of R20291 were assessed for **a)** Their ability to produce and secrete toxin through a
265 combined ELISA for TcdA and TcdB on sterile-filtered 72h supernatants **b)** Their ability to

266 form heat-resistant endospores (heat-resistant colony forming units HR-CFU/mL) across

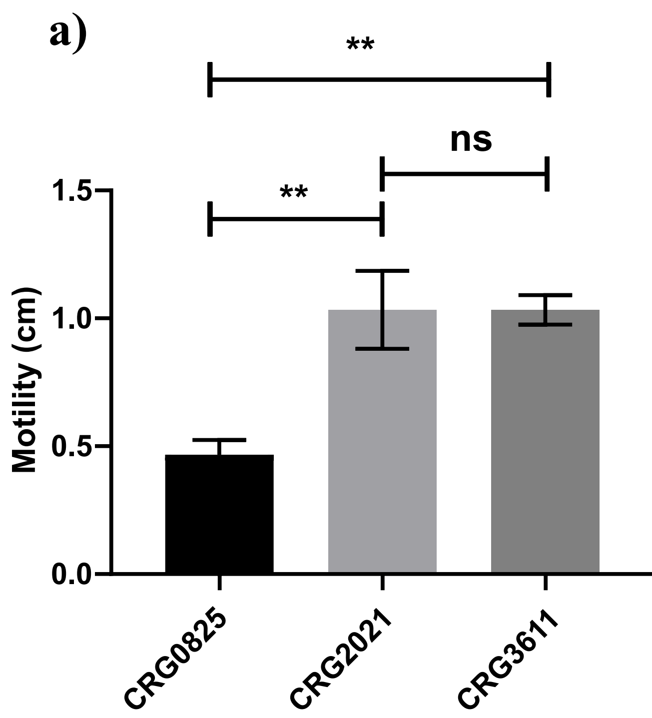
267 six time-points between 0 and 120h. Data represent the mean ±SD of three independent

268 experiments. Statistical significance according to One-way ANOVA followed by the

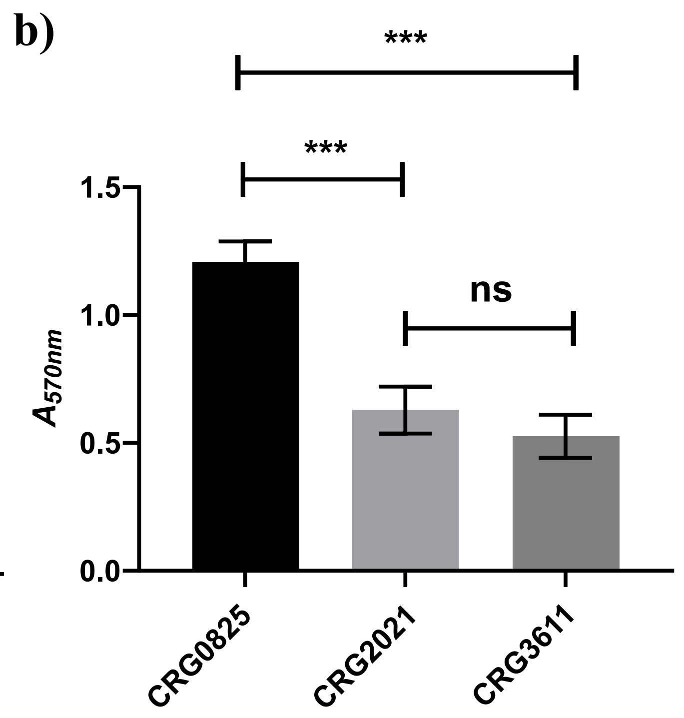
269 Dunnet's multiple comparison test (P=*<0.05; ****<0.0001).

270

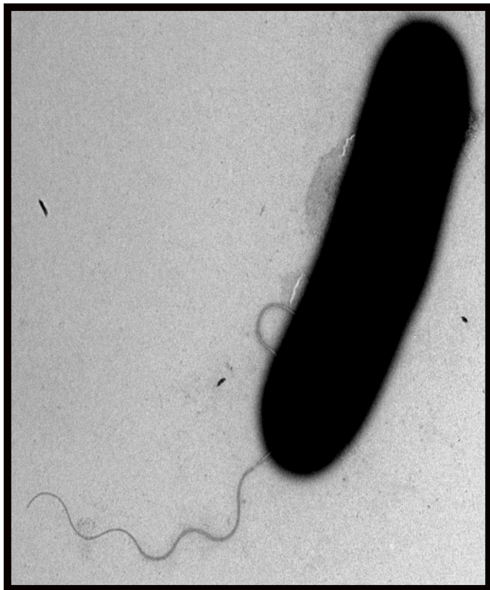
a)



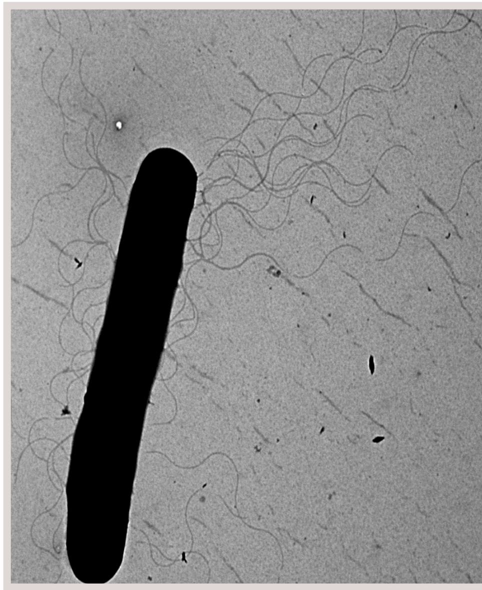
b)



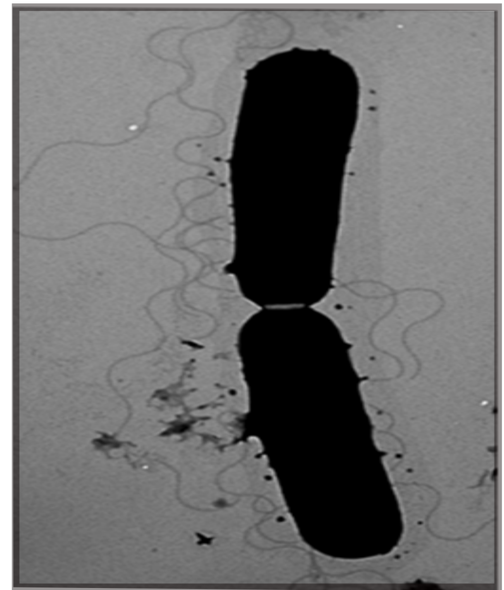
c)

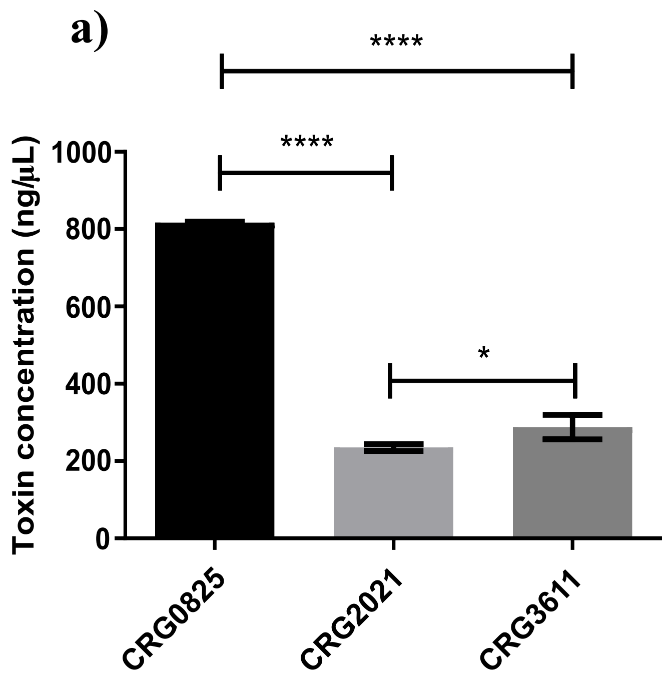


d)



e)




a)
****
****
*
Toxin concentration (ng/µL)
1000
800
600
400
200
0
CRG0825
CRG2021
CRG3611

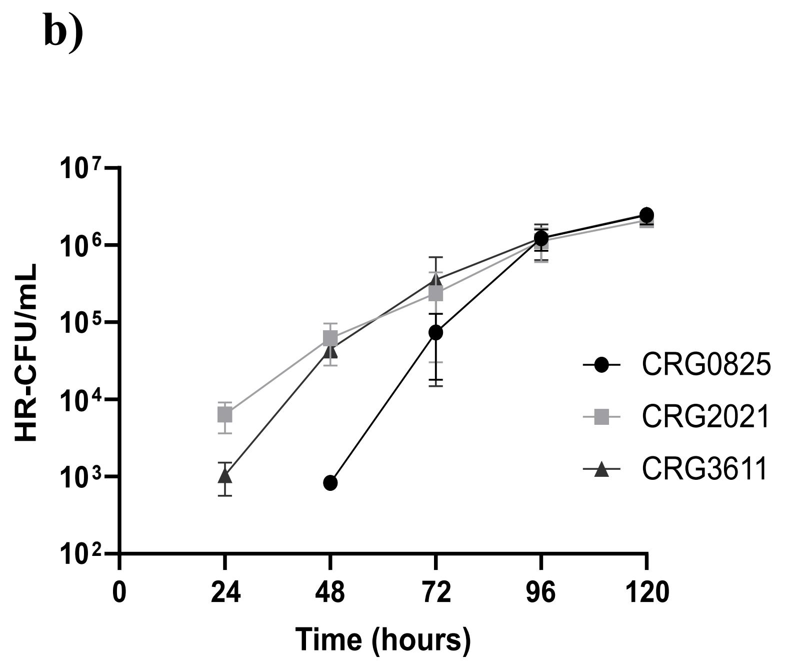

b)
HR-CFU/mL
10^7
10^6
10^5
10^4
10^3
10^2
0
24
48
72
96
120
Time (hours)
CRG0825
CRG2021
CRG3611

Table 1: Genomic mutations of the three R20291 stocks compared with the reference genome sequence.

Position	Gene	Locus Tag	Type	Reference ID	Mutated ID	AA substitution
[1] Common to all three strains						
132924	*met-tRNA*		Insertion	-	A	
132939	*met-tRNA*		SNP	G	T	
132955	*met-tRNA*		SNP	C	A	
132958-59	*met-tRNA*		SNP	TT	CG	
143463	Intergenic		Insertion	-	A	
206399	Intergenic		Insertion	-	A	
581481	Intergenic		Insertion	-	A	
581488	Intergenic		Insertion	-	A	
581495	Intergenic		Insertion	-	A	
1564432	Intergenic		Deletion	A	-	
1568676	Ruberythrin	CDR20291_1323	SNP	C	A	Gln138Lys
1578167	Intergenic		Deletion	T	-	
1578203	Intergenic		Insertion	-	A	
1592813	Intergenic		SNP	A	T	
1864417	Pseudogene	CDR20291_1576	Insertion	-	T	Ser7Frameshift*
1899596	Intergenic		Deletion	A	-	
2235738	Membrane protein	CDR20291_1913	Deletion	T	-	Val83Frameshift
2262060	Intergenic		Insertion	-	A	
2264190	Intergenic		Deletion	T	-	
2298111	Intergenic		Insertion	-	T	
2361948	Intergenic		SNP	C	A	
2361957	Intergenic		Insertion	-	A	
2367942	Intergenic		Insertion	-	T	
2578157	Intergenic		Deletion	T	-	
2674744	Intergenic		Deletion	T	-	
2680787	Intergenic		Insertion	-	T	
2772179	Pseudogene**	CDR20291_2368	Deletion	T	-	
3077986	Intergenic		Deletion	A	-	
3162098	Intergenic		Deletion	T	-	
3361915	Intergenic		Deletion	A	-	
[2] Specific to CRG0825						
9694	*rsbW*	CDR20291_3551	SNP	G	T	Gly82Val
358260	*rbsK*	CDR20291_0302	Deletion	A	-	Met57Stop
2077305	Intergenic	(CDR20291_1777 to CDR20291_1778)	Deletion	C	-	
2120669	*vncR*	CDR20291_1806	SNP	A	G	Asp202Gly
2881467	TCS-HK***	CDR20291_2456	Deletion	T	-	Leu434Stop
[3] Specific to CRG3661						
1340128	*codY*	CDR20291_1115	SNP	T	A	Try146Asn
3292465	*gntR* regulator	CDR20291_2781	SNP	T	C	Ile58Met
3472928	Intergenic	(CDR20291_2929 to CDR20291_2928)	SNP	G	A	

The region encompassing the mutation was aligned with multiple *C. difficile* genome sequences using NCBI Blastn. *Insertion here results in a frameshift mutation to a full-length pseudogene encoding an 87 AA protein. This gene without mutation encodes only 6 AAs. **Putative competence membrane protein ***TCS-HC: Two-component system histidine kinase. The gene resides immediately downstream of an adjacent gene (CDR20291_2457) encoding a putative response regulator.

1 **What's a SNP between friends: the lineage of *Clostridioides difficile* R20291 can**
2 **effect research outcomes.**

3

4 Jorge Monteford[1]*, Terry W. Bilverstone[1,2]*, Patrick Ingle[1,2], Sheryl Philip[1,3], Sarah A.
5 Kuehne[1,4], Nigel P. Minton[1,2].

6 * These authors contributed equally.

7

8 [1] Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre (SBRC), School
9 of Life Sciences, Centre for Biomolecular Sciences, The University of Nottingham, Nottingham,
10 NG7 2RD, UK.

11 [2] NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust
12 and the University of Nottingham, Nottingham, NG7 2RD, UK.

13

14 Corresponding author: nigel.minton@nottingham.ac.uk

15

16 *Current Address:*

17 [3]Biocatalysts Limited, Unit 1, Cefn Coed, Parc Nantgarw, Cardiff, CF15 7QQ, UK,

18 [4]Oral Microbiology Group, School of Dentistry and Institute of Microbiology and Infection,
19 College of Medical and Dental Sciences, The University of Birmingham, Birmingham, B5 7EG,
20 UK.

21 **Supplementary material**

22 **Table S1: *C. difficile* strains used in this study**

Strain	Description
CRG0825	R20291 sent to Nottingham by Val Hall in 2006 from the Anaerobe Reference Unit (ARU), Cardiff, UK.
CRG2021	R20291 sent to Nottingham by Lisa Dawson in 2010 from the laboratory of Brendan Wren at the London School of Hygiene and Tropical Medicine (LSHTM), UK. Originally obtained from the ARU, Cardiff, UK.
CRG3661	R20291 sent to Nottingham by Meera Unnikrishnan in 2013 from Novartis, Sienna, Italy. Sent to Novartis by Trevor Lawley at the Sanger Institute, Cambridge, UK. Originally obtained from the Brendan Wren laboratory at LSHTM, London UK.
CRG1375	R20291 (CRG0825) *spo0A* ClosTron mutant [1].

23

24 **Experimental**

25 **Bacterial Strains and Growth Conditions**

26 Strains were routinely cultured anaerobically at 37 °C in an anaerobic MACS1000
27 workstation (Don Whitely, Yorkshire, UK) in BHIS (Brain Heart Infusion supplemented with
28 yeast extract [5 mg.ml−1] and L-cysteine [0.1% w/v]) medium supplemented with d-

29 cycloserine (250 μg.ml−1), cefoxitin (8 μg.ml−1) and thiamphenicol (15 μg.ml−1) or Em
30 (10 μg.ml−1) where appropriate.

31

32 **Comparative conjugations**

33 Conjugations of shuttle vectors into *C. difficile* R20291 were performed as described in
34 [2]. Briefly, *E. coli* CA434 donor strains harbouring either pMTL82151, pMTL83151 or
35 pMTL84151 were grown overnight in LB supplemented with chloramphenicol and
36 kanamycin. From which, aliquots (1ml) were pelleted, washed in PBS and resuspended in
37 200 µl of *C. difficile* R20291 cultures grown anaerobically overnight in BHIS broth. The
38 resulting conjugal mixtures were spotted onto BHIS plates lacking antibiotics and
39 incubated anaerobically for 24h. Subsequent growth was resuspended in PBS (500µl) and
40 spread onto BHIS plates supplemented with d-cycloserine and cefoxitin, both with and
41 without thiamphenicol. After 72h, Thiamphenicol (Tm) resistant (^{R}) CFU.ml^{-1} and total *C.*
42 *difficile* CFU.ml^{-1} values were determined from the subsequent growth of *C. difficile* R20291
43 strains in the presence (Tm^{R}) or absence (total) of thiamphenicol. Conjugation efficiency
44 was calculated as the Tm^{R} CFU.ml^{-1} divided by total *C. difficile* CFU.ml^{-1}.

45 **24h growth curve**

46 The growth characteristics of *C. difficile* R02921 was assessed by manual growth curve.
47 Therein, colonies of *C. difficile* R20291 were subcultured into fresh BHIS broth in an
48 Anaerobic Work Station (Don Whitley, UK), at 37°C with an anaerobic atmosphere
49 comprising 80% N_2, 10% H_2 and 10% CO_2. The resultant cultures were then diluted 1/100
50 in fresh BHIS and grown to an optical density value ($OD_{600nm)}$ of 0.2-0.5. This generated
51 replicates of each strain with similar starting OD values for downstream growth
52 assessment. Once target OD values had been reached, the cultures were diluted 1/100 in
53 fresh medium incubated for 24h. 1ml of sample was taken for each replicate at hourly
54 intervals which was then assessed for its optical density using a Novaspec II
55 spectrophotometer (Geminibv, Netherlands).

56

57 **Motility assay**

58 The motility of *C. difficile* R20291 derivatives was assessed by swimming motility assays
59 as previously described [3]. Therein, single colonies of R20291 were isolated using a
60 toothpick and stabbed onto the centre of semi-solid BHIS plate containing 0.3% (w/v)
61 agar. Following 48h incubation as described above, the diameter of the ensuing halo was
62 measured. Motility is represented as the distance between the centremost and outermost
63 points of detected colonisation (cm).

64

65 **Crystal Violet, Biofilm Assay**

66 The assay was undertaken essentially as described by Dapa and co-workers [4]. A starter
67 culture of *C. difficile* was prepared by inoculating fresh BHIs broth containing 0.1M glucose
68 with an overnight culture of the desired strain in a 1:100 dilution. A 1ml aliquot of this
69 culture added to each well of a 24-well Tissue culture plate (Costar,USA) and incubated
70 anaerobically for 120h. Plates were pre-reduced in the anaerobic cabinet for 48h prior to
71 use. To avoid liquid evaporation, each plate was wrapped in parafilm. Following
72 incubation, wells were washed with PBS and the plate allowed to dry for 10m. The biofilm
73 was stained with 1ml of filter-sterilised 0.2% (w/v) crystal violet solution per well and
74 incubated for 30m at 37°C under anaerobic conditions. The staining solution was removed
75 and the wells were washed twice with PBS. The plate was removed from the anaerobic
76 cabinet and 1ml methanol was added to the wells for the removal of the dye from the
77 biofilm and it was incubated for 30m at room temperature. The methanol extracted dye
78 was diluted 1:1, 1:10 and 1:100 and the absorbance A570 was measured with Ultrospec
79 500 pro spectrophotometer.

80

81 **Sporulation assay**

82 Cultures were generated for each strain with similar staring OD values as described for
83 the 24h growth curve. For the sporulation assay, cultures were incubated for a 120h
84 period. Samples were taken at 24h intervals which were heated at 65°C for 30m in order
85 to eradicate vegetative cells and diluted $1x10^{-1}$-$1x10^{-8}$ before plating onto BHIS
86 supplemented with 0.1% taurocholate germinant. Spores were then enumerated for each
87 R20291 derivative alongside a ClosTron insertional mutant for the master regulator of
88 sporulation *spo0A* [1].

89

90 **Detection of combined TcdA and TcdB**

91 Combined TcdA and TcdB was detected as previously described [5]. Cultures of each strain
92 were collected after 72h, the OD measured, and normalised to the lowest OD value.
93 Normalised samples were centrifuged and the supernatant passed through a 0.22µm filter
94 and frozen at -20°C until required (<1 week). Thawed samples were diluted $1X10^{1}$ – $1X10^{8}$
95 in sterile PBS and processed for ELISA quantification of total TcdA and TcdB using a C.
96 DIFFICILE Tox A/B II detections kit (TechLab, USA) according to the manufacturer's
97 instructions. Absorbance values were converted into toxin concentration by determining
98 the R2 value of the assay's standard curve by running a range of defined combined TcdA
99 and TcdB toxin standards from 0-125ng/ml (The Native Antigen Company).

100

101 **Genome Sequencing**

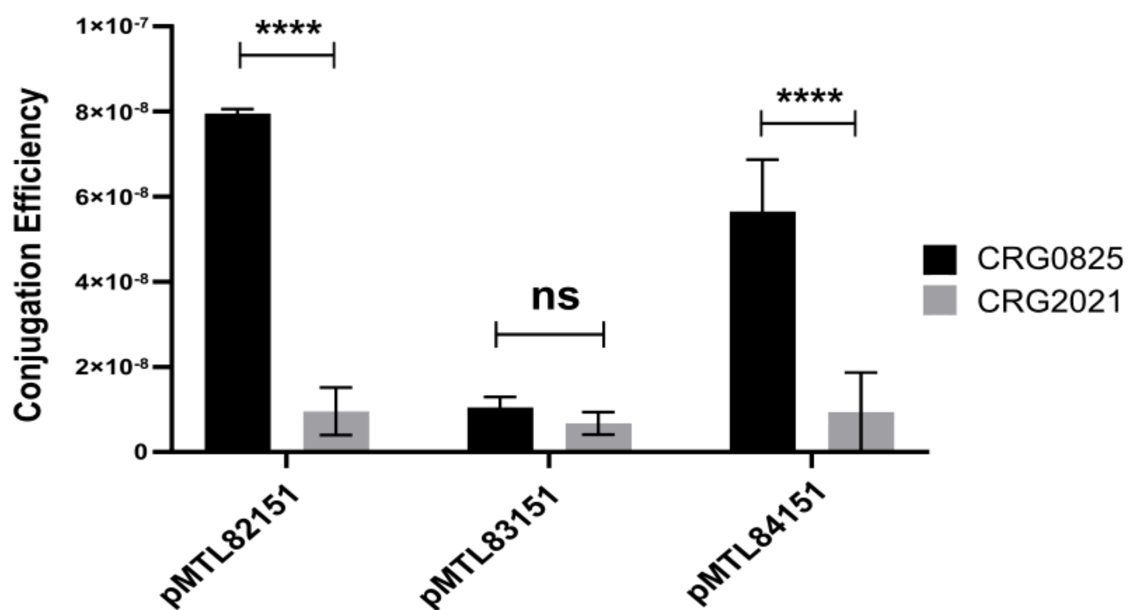
102 Chromosomal DNA of each strain was prepared and subjected to Illumina paired-end
103 sequencing by DeepSeq (University of Nottingham) using the MiSeq v3 600 platform.
104 Paired reads were trimmed, before mapping the trimmed reads to the reference genome
105 sequence for R20291 (Accession number: FN545816) using the quality-based variant
106 detection workflow from CLC Genomics Workbench (Qiagen, Germantown, USA). The
107 software was then used to identify single nucleotide variations (SNVs), insertions and
108 deletions compared with the reference genome sequence. Sequencing reads were

109 deposited to the NCBI Sequencing Reads Archive with the Bioproject accession
110 PRJNA689976 and the following individual accession numbers: CRG2021 (SRR13366486);
111 CRG0825 (SRR13366485); CRG03661 (SRR13366484).

112

113 **Supplementary Figures**

114



115

116 **Fig S1: Comparative conjugation efficiencies of plasmid transfer from *E. coli***
117 **CA434 into *C. difficile* R20291 stocks CRG0825 and CRG2021.** Conjugations from
118 *E. coli* CA434 strains harbouring the indicated shuttle vectors, differing only in the Gram-
119 positive replicon present, into *C. difficile* R20291 CRG2021 (grey bars) and CRG0825
120 (black bars) were performed as indicated in Materials and Methods. Conjugation efficiency
121 was calculated as thiamphenicol resistant CFU.ml^{-1} divided by the total recipient *C. difficile*
122 R20291 CFU.ml^{-1}. Data represent the mean ±SD of three independent experiments.
123 Statistical significance was determined using multiple unpaired t-tests. P=****<0.0001;
124 ns= not significant.

125

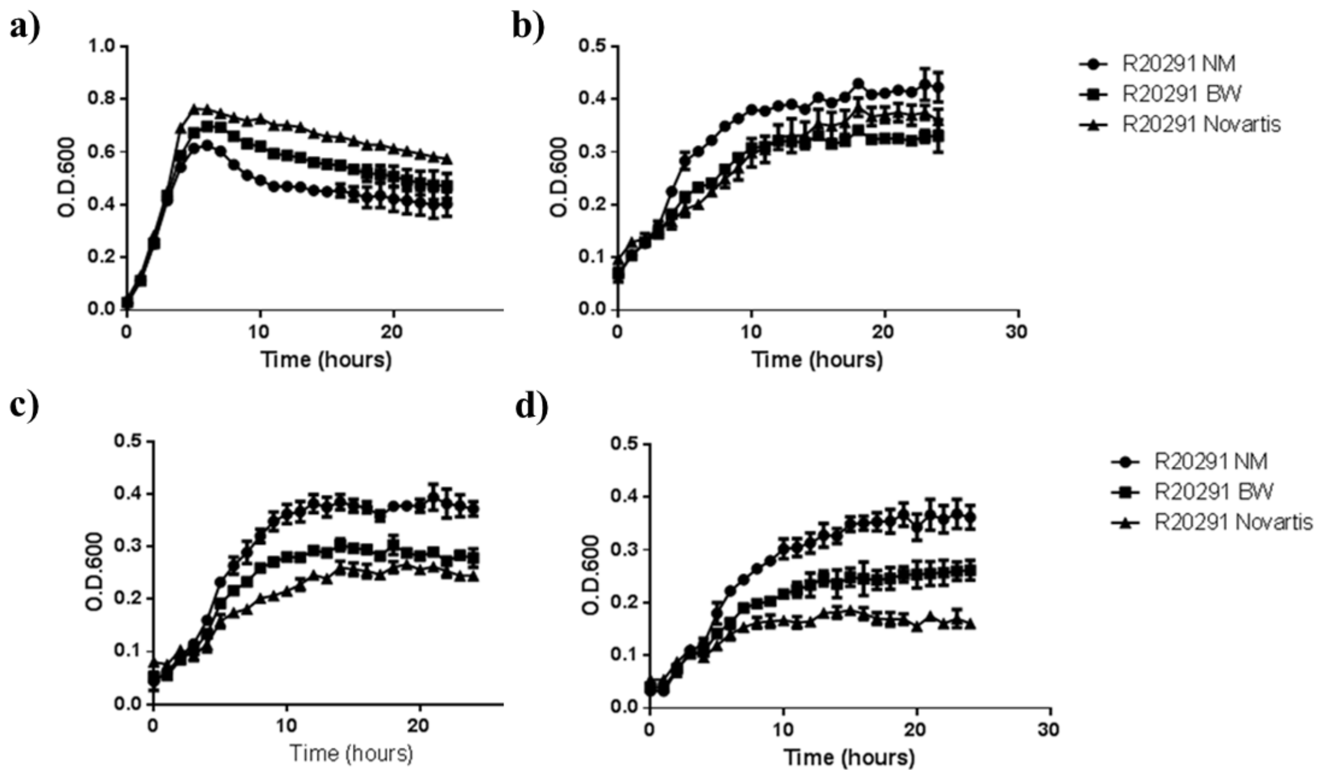


Figure S2: Comparison of growth characteristics between the three derivatives of R20291. R02921 strains were grown for 24h in a) BHIS broth; b) CDMM 1% (w/v) glucose; c) CDMM 1% fructose; d) CDMM 1% mannitol. Data points indicate the mean ±SD of three independent experiments.

References

1. Heap J.T., Kuehne S.A., Ehsaan M., Cartman S.T., Cooksley C.M., Scott J.C., Minton N.P. *J Microbiol Methods*. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. 2010;**80(1):**49-55.

141 2. Woods C., Humphreys C.M., Rodrigues R.M., Ingle P., Rowe P., Henstra A.M., Köpke
142 M., Simpson S.D., Winzer K., Minton N.P. A novel conjugal donor strain for improved
143 DNA transfer into *Clostridium* spp. *Anaerobe*. 2019;**59:**184-191.
144 3. Baban S.T., Kuehne S.A., Barketi-Klai A., Cartman S.T., Kelly M.L., Hardie K.R.,
145 Kansau I., Collignon A., Minton N.P. The role of flagella in *Clostridium difficile*
146 pathogenesis: comparison between a non-epidemic and an epidemic strain. *PLoS One*.
147 2013;**8(9):**e73026.
148 4. Đapa T., Leuzzi R., Ng Y.K., Baban S.T., Adamo R., Kuehne S.A., Scarselli M., Minton
149 N.P., Serruto D., Unnikrishnan M. Multiple factors modulate biofilm formation by the
150 anaerobic pathogen *Clostridium difficile*. *J Bacteriol*. 2013;**195(3):**545-55.
151 5. Bilverstone T.W., Kinsmore N.L., Minton N.P., Kuehne S.A. Development of
152 *Clostridium difficile* R20291ΔPaLoc model strains and in vitro methodologies reveals
153 CdtR is required for the production of CDT to cytotoxic levels. *Anaerobe.* 2017;**44:**51-
154 54.