Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

A Catalytic DNA Activated by a Specific Strain of Bacterial Pathogen

*Zhifa Shen, Zaisheng Wu, Dingran Chang, Wenqing Zhang, Kha Tram, Christine Lee, Peter Kim, Bruno J. Salena, and Yingfu Li**

anie_201510125_sm_miscellaneous_information.pdf

MATERIALS AND METHODS

Synthesis and purification of oligonucleotides. The sequences of oligonucleotides are listed in Figure S2. The DNA library CDL1, the special fluorogenic substrate CDFS1, the forward PCR primer CDFP1, the two reverse PCR primers CDRP1 and CDRP2, the template CDT1 for ligating CDFS1 to CDL1 were purchased as synthetic oligonucleotides either from Yale University Keck Facilities or from Integrated DNA Technologies (IDT). All oligonucleotides were purified by 10% denaturing polyacrylamide gel electrophoresis (dPAGE) before use. Each random position in CDL1 represents a 25% probability of A, C, G or T nucleotide. CDFS1 contains an adenosine ribonucleotide (R in dark blue), flanked by a fluorescein-dT (F in green) and a DABCYL-dT (Q in red). The reverse primer CDRP2 contains a triethylene glycol spacer (L in grey) and A20 tail at the 5' end. The spacer prevents the poly(dA) tail from being amplified, making the non-DNAzyme-coding strand 20 nucleotides longer than the coding strand. This allows for the separation of the two strands by 10% dPAGE. The RNA-containing substrate CDFS1 was deprotected and purified by 10% dPAGE following a previously reported protocol.[1]

Bacterial strains and culture conditions. *C. difficile* BI/027-H strain was cultured from a patient with CDI at St. Joseph's Healthcare Hamilton (Ontario, Canada). Other *C. difficile* strains in this study were obtained from the American Type Culture Collection (ATCC; Manassas, Va.). *E. coli* used was *E. coli* K12 BW25113Δrna. Gram-negative bacteria *Serratia fonticola, Acinetobacter lwoffi, Moraxella osloensis, Achromobacter xylosoxidans, Ochrobactrum grignonese, Yersinia ruckeri, Hafnia alvei, Brevundimonas diminuta, Pseudomonas peli* and gram-positive bacteria *Leuconostoc mesenteroides, Pediococcus acidilactici, Bacillus subtilis* are routinely cultured and maintained in our laboratory. For aerobic bacteria, cells were cultured (37°C; shaking at 200 rpm) in LB medium. *C. difficile* cultures were grown in chopped meat glucose medium (ATCC medium 593), anaerobically, at 37 °C in an anaerobic workstation (D. Whitley).

Preparation of crude extracellular mixture (CEM) from bacterial strains. Both aerobic and anaerobic bacteria were grown in 5 mL of appropriate medium (LB for aerobic bacteria, chopped meat glucose medium for anaerobic bacteria) until OD_{600} reached ~1. Each culture was then heated at 90ºC for 5 min, the pellets were removed by centrifugation at 11,000 g at 4ºC for 5 min. The crude supernatant was collected and supplemented with complete protease inhibitor cocktail (Roche), passed through 0.22 μm filter, and then concentrated to 1 mL, aliquoted into microcentrifuge tubes, and stored at -20ºC.

***In vitro* selection.** Procedures were performed similarly as previously described.[2] Briefly, 200 pmol of CDFS1 was phosphorylated (reaction volume: 100 μL) with 10 μCi [γ-^{32}P]ATP and 20 U of T4 polynucleotide kinase (PNK) in 1× PNK buffer A (Thermo Scientific) at 37°C for 15 min. This was followed by the addition of non-radioactive ATP to the final concentration of 1 mM and a further incubation at 37°C for 20 min. The reaction was quenched by heating the mixture at 90°C for 5 min. Equimolar CDL1 and CDT1 were then added to this solution and the mixture was heated at 90°C for 1 min and cooled to room temperature. Then, 15 μL of 10× T4 DNA ligase buffer (Thermo Scientific) was added and the volume was adjusted to 150 μL with ddH_2O. T4 DNA ligase (20 U) was added, followed by incubation at room temperature for 2 h. The DNA molecules in the mixture were concentrated by ethanol precipitation and the ligated CDFS1-CDL1 molecules were purified by 10% dPAGE. The purified CDFS1-CDL1 was dissolved in 50 μL of 1× selection buffer (1× SB) (50 mM HEPES, pH 7.5, 150 mM NaCl, 15 mM $MgCl_2$, and 0.01% Tween 20). This mixture was incubated

with 50 μL of mixed CEMs prepared from *E. coli*, *B. subtitlis* and CD630 (equal amounts) in 1× SB at room temperature for 5 h. The reaction was quenched by the addition of 100 μL of stop solution containing 200 mM EDTA and 16 M urea. After ethanol precipitation, the uncleaved CDFS1-RFL1 molecules were purified by 10% dPAGE, dissolved in 10 μL of ddH_2O. The purified CDFS1-CDL1 was again dissolved in 50 μL of 1× SB. This mixture was incubated with 50 μL of CEM-CD in 1× SB at room temperature for 30 min. The reaction was quenched by the addition of 150 μL of stop solution containing 100 mM EDTA and 8 M urea. After ethanol precipitation, the cleaved CDFS1-RFL1 molecules were purified by 10% dPAGE, dissolved in 10 μL of ddH_2O, and stored at -20°C. The PCR1 mixture (50 μL) contained 4 μL of the template prepared above, 0.5 μM each of CDFP1 and CDRP1, 200 μM each of dNTPs (dATP, dCTP, dGTP and dTTP), 1× PCR buffer (75 mM Tris-HCl, pH 9.0, 2 mM $MgCl_2$, 50 mM KCl, 20 mM $(NH_4)_2SO_4$) and 2.5 U of *Thermus thermophilus* (Tth) DNA polymerase. The DNA was amplified using the following thermocycling steps: 94°C for 1 min; 11-13 (dependent on the amount of cleavage of the DNA pool) cycles of 94°C for 30 s, 50°C for 45 s and 72°C for 45 s; 72°C for 1 min. For the PCR2 reaction, 1 μL of the PCR1 product was diluted with ddH_2O to 20 μL, 2-5 μL of which was used as the template for this additional PCR step (a total of 8× 50-μL reactions were conducted to generate enough DNA) using primers CDFP1 and CDRP2 and the same protocol for PCR1. The CDL1 strand was purified by 10% dPAGE (approximately 100 pmol) and used for next selection round. A total of 19 cycles of selection were conducted. The DNA population from round 19 was cloned and sequenced.

PCR Ribotyping of *C. difficile* strains. PCR ribotyping was carried out using a method previously described.[3] Briefly, primers located partially within the *C. difficile* 16S–23S rRNA intergenic spacer region and partially within the 16S (forward primer 5′-GCTGG ATCAC CTCCT TTCTA AG-3′) and 23S (reverse primer 5′-TGACC AGTTA AAAAG GTTTG ATAGA TT-3′) rRNA genes. The PCR mixture (50 μl) contained 80 ng genomic DNA, 50 pmol of each primer, 1.5 mM $MgCl_2$, 0.1 mg/ml bovine serum albumin (BSA), and 2 U of Tth DNA polymerase (Biotools). Amplified products were concentrated to 25 μl by heating at 75°C for 45 min before electrophoresis in 3% agarose gel for 5 h at 2.5 V/cm. PCR ribotypes from ATCC strains were used as the references for the clinical strain.

Restriction endonuclease digestion of *C. difficile* strains. Genomic DNA was prepared as previously described.[4] Briefly, ATCC medium 593 was inoculated with a single colony from an anaerobic blood agar plate and then incubated overnight. Cells were washed with 1 mL of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), resuspended in 200 μL of genomic DNA solution (10 mM Tris-HCl, 1 mM EDTA [pH 8.0], 1 M sucrose) with lysozyme (50 mg/mL; Sigma), incubated at 37°C for 2 h. This was followed by the addition of 100 μL of 20% Sarkosyl (Sigma) and 15 μL of RNase A (10 mg/mL, Thermo Scientific), and further incubation at 37°C for 30 min. Next, 15 μL of proteinase K (10 mg/mL) was added and the mixture was incubated at 37°C for 30 min. The solution was then mixed with 600 μL of phenol-chloroform-isoamyl alcohol (25:24:1). The DNA material in the aqueous phase was precipitated with cold ethanol. For restriction digestion, 5 μg DNA was incubated with 2 μL of *Hind*III (2.5 U/μL, Thermo Scientific) at 37°C for 40 min. Restriction fragments were separated using 0.7% agarose gel electrophoresis (50 volts, 16 h). DNA banding patterns were compared to a published library of known REA groups.[5] Finally, the BI type was experimentally confirmed by analyzing a known BI strain, BAA-1803, from ATCC.

Specificity test with different bacterial cells. The CEM from each bacterium listed in Figure 2

was prepared according to the procedures described in the section "Bacterial Strains and Culture Conditions". The dPAGE based specificity test for each CEM was carried out as follows: 50 μL of 1× SB containing 2 pmol RFD-CD1 was incubated with 50 μL of CEM-CD (25 μL of CEM-CD was mixed with 25 μL of 2× SB) at room temperature for 30 min. The reaction was quenched by addition of 100 μL of 2× stop solution. After ethanol precipitation, the cleavage was analyzed by 10% dPAGE.

Cleavage test using CEM-CD treated with proteases. To treat CEM-CD with proteinase K (PK; New England Biolabs), 1 μL of 5 U/μL PK stock was mixed with 23 μL of CEM-CD and 25 μL of 2× SB, and the resulted solution was incubated at 50ºC for 1 h. Treatment of CEM-CD with 5 U of trypsin (New England Biolabs) or subtilisin (Sigma-Aldrich) was carried out in a similar way except that the digestion was done at room temperature. Following this step, 1 μL of 5 μM RFD-CD1 stock was added to each mixture, which was further incubated at 37ºC for 60 min. Each reaction was then quenched by 50 μL of 2× stop solution and the reaction mixture was analyzed using 10% dPAGE.

Estimation of molecular weight of the target. Two milliliters of CEM-CD was prepared as described above. Five 200-μL aliquots were taken and individually passed through a membrane based molecular sizing centrifugal column with a molecular weight cut-off of 3K (3,000 Daltons), 10K, 30K, 50K and 100K, respectively (NANOSEP OMEGA, Pall Incorporation). The filtrate from each column was then used to induce the cleavage of RFD-CD1 using the dPAGE-based assay described above.

Bioinformatics analysis. All 3798 protein sequences of CD630 from NCBI were collected, and filtered for the proteins with length between 230-340 amino acids (which have estimated molecular weight of 25,000-35,000 Daltons according to method from www.Lenntech.com) using Procom program (at http://ural.wustl.edu/~billy/Procom). All the proteins from this screening were further filtered using Procom so that transcription factors or transcriptional regulators can be identified. In addition, the sequence identity of each protein in CD630 was compared with that in CD196 (a known BI/027 strain) using BLAST.[6]

Molecular cloning. Candidate genes were cloned into pBADMyc-HisB vector and expressed in *E. coli* K12 BW25113Δrna. PCR for each candidate gene (*cdtR, tcdC, merR, rpiR* and *trmB*) was performed with Tth DNA polymerase using the genomic DNA from BI/027-H as the template (the primers used listed in Table S1). The amplified products, which contained PstI and XbaI restriction sites, were purified with the Qiagen PCR purification kit. The purified PCR products and the vector were digested with PstI and XbaI, and digested products were gel-extracted using the Qiagen QIAquick Gel Extraction kit. A ligation reaction (30 μL) was conducted with 100 ng digested PCR product, 100 ng vector, 3 μL of 10× T4 DNA ligase buffer, T4 DNA ligase (2.5 U) at room temperature for 1 h. The ligated DNA was transformed into *E. coli* competent cells by heat-shocking the cells at 42°C for 40 s. Potential clones were screened by restriction digestion and positive clones were sent to sequencing and confirmation of the desired insert (MOBIX Lab, McMaster University). For protein expression, *E. coli* cells containing the desired construct were cultured at 37°C for 16 h in LB containing 100 μg/mL ampicillin and 50 μg/mL kanamycin. Following this incubation a 1/50 subculture into 5 mL of LB (to have OD_{600} of ~0.05) was performed. Cells were grown at 37°C until an OD_{600} of 0.2 was reached. At this point arabinose was added at a final concentration of 0.1% to induce expression. This was followed by incubation at 37°C for another 4 h. CEM was then made for the cleavage assay with RFD-CD1 as described above.

TcdC purification. The *tcdC* gene amplified from both CD630 and BI/027-H using PCR with the use of the following two primers: 5′-CCCGG GCATA TGATG TTTTC TAAAA AAAAT G-3′

(forward primer) and 5′-CCCGG GGGAT CCTTA ATTAA TTTTC TCTAC AG-3′ (reverse primer). The amplified genes were cloned into the NdeI and BamHI sites of pET15b (Novagen) using *E. coli* BL21 (DE3). Overexpression of 6-N-His-TcdC was achieved by growing cells (1 L) at 37°C to mid-exponential phase, followed by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and further incubation for 5 h at 30°C. Cells were resuspended in 30 mL of binding buffer (50 mM NaH_2PO_4, 300 mM NaCl and 10 mM imidazole, pH 8.0) containing 1 mg/ml lysozyme and one mini EDTA-free protease inhibitor tablet (Roche). The cells were incubated on ice for 30 min before being lysed by sonication. Each protein was purified by passing the soluble protein extract over a Ni-NTA column (BioRad). The column was washed with 10 mL of binding buffer supplemented with increasing concentrations of imidazole and the protein was eluted with 150 mM imidazole. The purified protein was dialyzed into storage buffer [50 mM Tris, pH 7.5, 150 mM NaCl and 1 mM dithiothreitol (DTT)] and then concentrated to 3.5 mg/mL and stored at -80°C in 20 mM Tris-Cl (pH 7.5), 1.4 mM β-mercaptoethanol, 100 mM KCl and 25% glycerol.

Total DNA extraction. To extract genomic DNA from *C. difficile* strains, 48-h cultures were suspended in 5% Chelex 100 (Bio-Rad), vortexed and boiled for 10 min. After centrifugation for 10 min at 10,000 g, the supernatant was transferred into a fresh tube and stored at 4°C until use.[7]

Sequencing of *tcdC* genes. Each *tcdC* gene was amplified by PCR as previously described.[8] The reaction mixture contained 1× PCR buffer, 200 pmol of each deoxynucleoside triphosphate, 25 pmol of the forward and reverse primers (Table S2), and 2.5 U of Tth DNA polymerase. The template was denatured for 5 min at 94°C, and DNA was amplified for 30 cycles consisting of 1 min at 94°C, 1 min at 50°C, and 1 min at 72°C. The PCR products were sequenced at the Mobix Lab (McMaster University). The nucleotide sequences were compared with database entries by using the BLAST program.

SUPPLEMENTARY REFERENCES

1. S. H. Mei, Z. Liu, J. D. Brennan, Y. Li. *J. Am. Chem. Soc.* **2003**, *125*, 412-420.
2. M. M. Ali, S. D. Aguirre, H. Lazim, Y. Li. *Angew. Chem. Int. Ed.* **2011**, *50*, 3751-3754.
3. C. P. Cartwright, F. Stock, S. E. Beekmann, E. C. Williams, V. J. Gill. *J. Clin. Microbiol.* **1995**, *33*, 184-187.
4. L. Bouillaut, S. M. McBride, J. A. Sorg. *Curr Protoc Microbiol.* **2011**, *Chapter 9*: Unit 9A.2.
5. C. R. Clabots, S. Johnson, K. M. Bettin, P. A. Mathie, M. E. Mulligan, D. R. Schaberg, L. R. Peterson, D. N. Gerding. *J. Clin. Microbiol.* **1993**, *31*, 1870-1875.
6. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman. *J. Mol. Biol.* **1990**, *215*, 403-410.
7. B. G. Stare, M. Delmee, M. Rupnik. *J. Med. Microbiol.* **2007**, *56*, 329-335.
8. P. Spigaglia, P. Mastrantonio. *J. Clin. Microbiol.* **2002**, *40*, 3470-3475.

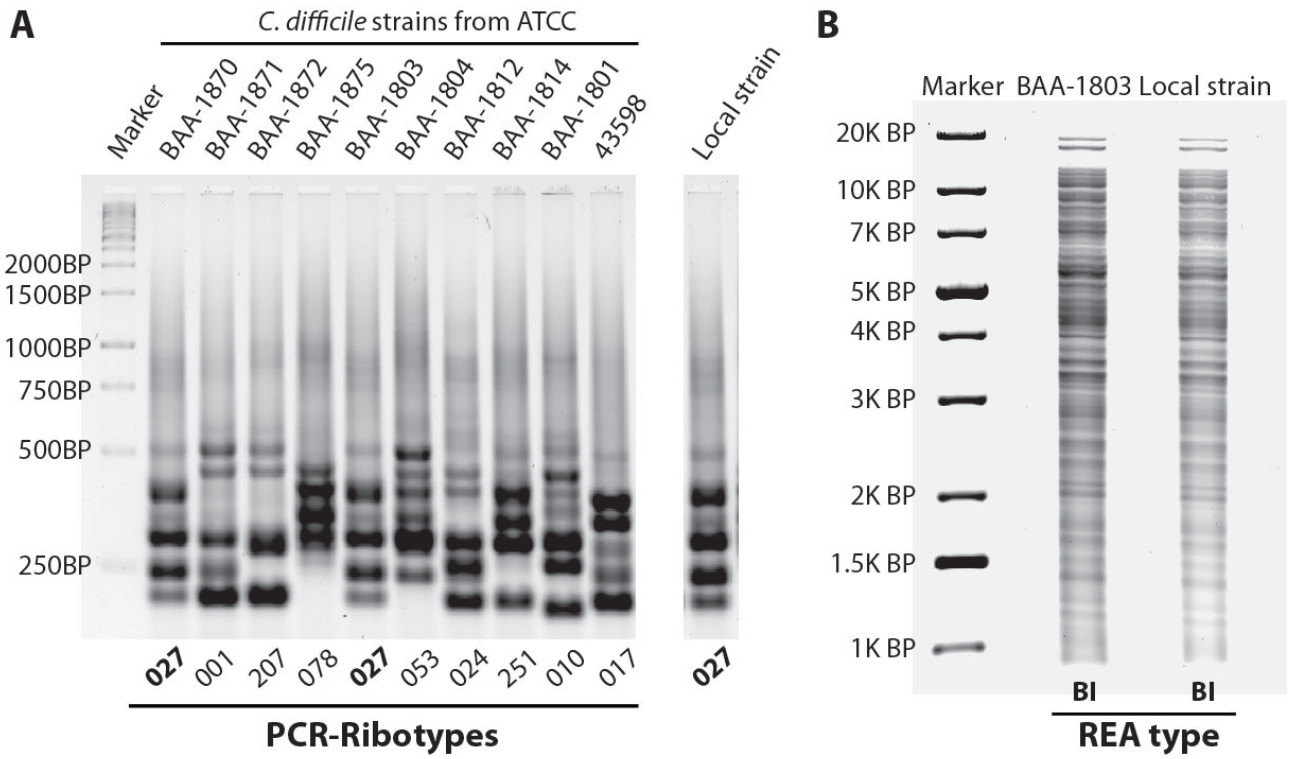

Figure S1. Genetic typing of *C. difficile* strains. (A) PCR-ribotyping of the local isolate as well as 10 *C. difficle* strains obtained from ATCC. (B) Results from restriction endonuclease analysis (REA) of the local isolate and ATCC BAA-1803 strain, which is known to have a BI type. Based on the typing results, we named our local strain as BI/027-H (H stands for Hamilton). BP: Base-pair.

A

CDL1: 5′-CACGGATCCTGACAAG-N_{40}-CAGCTCCGTCCG

CDFS1: 5′-ACTCTTCCTAGC-F-R-Q-GGTTCGATCAAGA

CDFP1: 5′-CACGGATCCTGACAA

CDRP1: 5′-CGGACGGAGCTG

CDRP2: 5′-AAAAAAAAAAAAAAAAAAAA-L-CGGACGGAGCTG

CDT1: 5′-CTAGGAAGAGTCGGACGGAGCTG

B

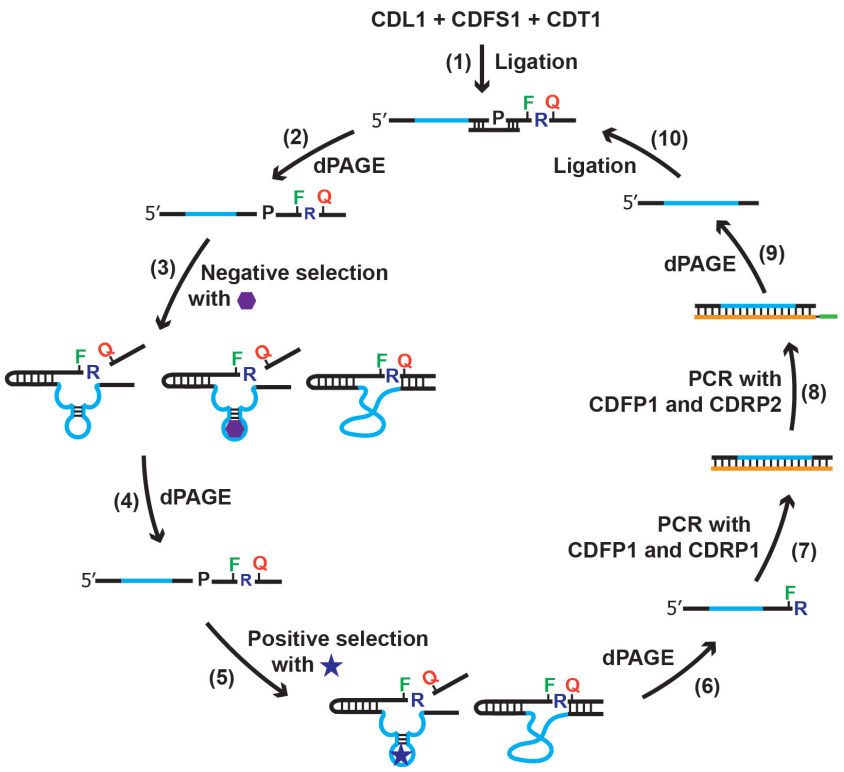

Figure S2. In vitro selection of RFD-CD1. (A) The sequences of the DNA molecules used for in vitro selection. CDL1: DNA library with 40 random nucleotides (N_{40}); CDFS1: fluorogenic substrate; CDFP1: forward PCR primer; CDRP1 and CDRP2: two reverse PCR primers; CDT1: the template for ligating CDFS1 to CDL1. (B) In vitro selection scheme – (1) Ligation of CDFS1 to CDL1; (2) purification of ligated CDFS1-CDL1; (3) negative selection with CEM-EC, CEM-BS and CEM-CD630; (4) purification of uncleaved CDFS1-CDL1 by dPAGE; (5) positive selection with CEM-CD; (6) purification of cleaved products by dPAGE; (7) PCR using CDFP1 and CDRP1 as primers; (8) PCR with CDFP1 and CDRP2 as primers (note: CDRP2 contains a triethylene glycol spacer and A20 tail at the 5' end – the triethylene glycol spacer prevents the poly-A tail from being amplified, making the non-DNAzyme-coding strand 20 nucleotides longer than the coding strand); (9) purification of CDL1 strand by dPAGE; (10) ligation of CDL1 to CDFS1 The cycle of steps 2-10 was repeated for 19 times in this study. It should be noted that steps 3 and 4 serve as a counter selection strategy to eliminate cross-reactive RFD molecules.

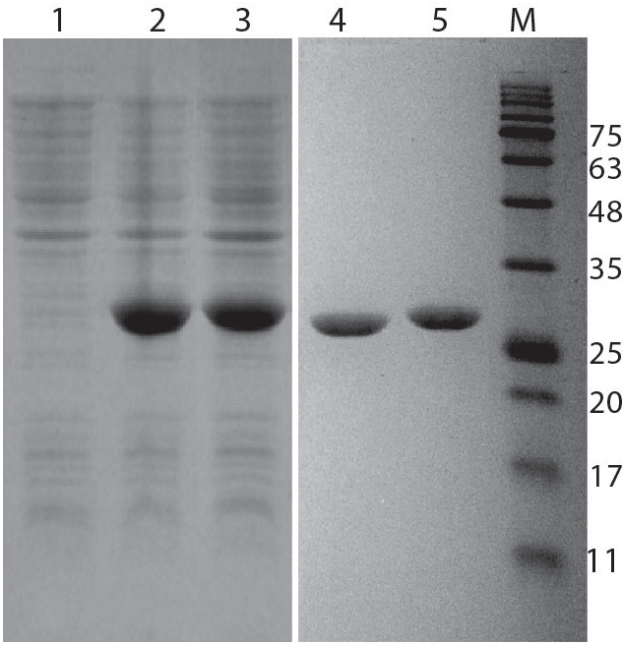

Figure S3. SDS-PAGE analysis of crude and purified TcdC-WT and TcdC-24. Lanes 1-3, crude cell extracts from *E. coli* cells carrying the empty vector pET15b, pET15b with *tcdC-24* and pET15b with *tcdC-WT*, respectively. Lanes 4 and 5: purified His6-Tcd-24 (2.5 μg) and His6-TcdC-WT (2.5 μg), respectively. Proteins were stained by Coomassie brilliant blue. M: protein markers (sizes in KDa).

Gene	Primer name	Primer sequence
cdtR	FP-cdtR	5′CCCGGGATGGATATATTAATCTTTGATAATGATGTTTG3′
	RP-cdtR	5′CCCGGGTCTAGATTATGTTTTAATAATATTCTTTAAAATA3′
trmB	FP-trmB	5′CCCGGGCTGCAGATGGATAATATAATTAATGAATTACAA3′
	RP-trmR	5′CCCGGGTCTAGATTAATTTTGTTCATTTTTATAACCTCC3′
rpiR	FP-rpiR	5′CCCGGGCTGCAGATGATAATTGAGCTAAACAAAAGTAATTC3′
	RP-rpiR	5′CCCGGGTCTAGATTAATTTTCAGAATTATTATATCTGACC3′
tcdC	FP-tcdC	5′CCCGGGCTGCAGATGTTTTCTAAAAAAAATGATGGTAAC3′
	RP-tcdC	5′CCCGGGTCTAGATTAATTAATTTTCTCTACAGCTATCCCTG3′
merR	FP-merR	5′CCCGGGCTGCAGTTGGAAAAACAGTATTTTACAACAGG3′
	RP-merR	5′CCCGGGTCTAGATTATCTGGATACTTGGATTGAAATTTC3′

Table S1. Primers used in the experiment of cloning five candidate genes.

Gene	Primer name	Primer sequence
tcdC	FP-tcdC1	5′TTAATTAATTTTCTCTACAGCTATCC3′
	RP-tcdC1	5′TCTAATAAAAGGGAGATTGTATTATG3′

Table S2. Primers used for the amplification of *tcdC* gene in PaLoc.

References in main manuscript with full authorship:

[10b] V. G. Loo, L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. René, Y. Monczak, A. Dascal, *N. Engl. J. Med.* **2005**, *353*, 2242-2249.

[10d] G. Killgore, A. Thompson, S. Johnson, J. Brazier, E. Kuijper, J. Pepin, E. H. Frost, P. Savelkoul, B. Nicholson, R. J. van den Berg, H. Kato, S. P. Sambol, W. Zukowski, C. Woods, B. Limbago, D. N. Gerding, L. C. McDonald, *J. Clin. Microbiol.* **2008**, *46*, 431-437.

[12] M. Sebaihia, B. W. Wren, P. Mullany, N. F. Fairweather, N. Minton, R. Stabler, N. R. Thomson, A. P. Roberts, A. M. Cerdeño-Tárraga, H. Wang, M. T. Holden, A. Wright , C. Churcher, M. A. Quail, S. Baker, N. Bason, K. Brooks, T. Chillingworth, A. Cronin, P. Davis, L. Dowd, A. Fraser, T. Feltwell, Z. Hance, S. Holroyd, K. Jagels, S. Moule, K. Mungall, C. Price, E. Rabbinowitsch, S. Sharp, M. Simmonds, K. Stevens, L. Unwin, S. Whithead, B. Dupuy, G. Dougan, B. Barrell, J. Parkhill, *Nat. Genet.* **2006**, *38*, 779-786.

[13] R. A. Stabler, M. He, L. Dawson, M. Martin, E. Valiente, C. Corton, T. D. Lawley, M. Sebaihia, M. A. Quail, G. Rose, D. N. Gerding, M. Gibert, M. R. Popoff, J. Parkhill, G. Dougan, B. W. Wren, *Genome Biol*. **2009**, *10*, R102.